



# **Identifying Transcription Factor Dysregulation in RUNX1-ETO Leukaemia**

---

A thesis submitted to Cardiff University in candidature for  
the degree of Doctor of Philosophy

**Aleksandra Mazurek Azevedo**

Division of Cancer and Genetics,  
School of Medicine  
Cardiff University

**February 2022**

## Abstract

Acute myeloid leukaemia (AML) results from clonal expansion of primitive myeloid cells incapable of terminal differentiation, giving rise to an accumulation of ‘blast’ cells at various stages of maturation within the bone marrow niche. AML is a heterogenous disease with multiple morphological, immunophenotypic and genetic features. This includes the t(8;21) which results in the expression of RUNX1-ETO, and occurs in 12% of AML cases.

To understand the role of RUNX1-ETO in the pathogenesis of AML, our group previously ectopically expressed RUNX1-ETO in normal human haematopoietic stem progenitor cells (HSPC). This resulted in a block in granulocytic differentiation and was associated with increased self-renewal - hallmarks of leukaemia. A subsequent study analysed the transcriptome of these cells and identified 380 differentially expressed genes using an unsupervised approach. This current study has now refined this analysis to determine the most significant changing transcription factors (TFs). Using Pathway Analysis programme (Metacore™), this study identified ZNF217 to be significantly overexpressed compared to control (1.5-fold;  $p=0.003$ ). ZNF217 is a TF responsible for binding to the promoters of several target genes, such as E-cadherin, as well as cooperating in transcriptional silencing programs by recruiting chromatin modifiers. This study determined that ZNF217 overexpression, as single abnormality, induced myeloid differentiation of HSPC, particularly within the monocytic population, suggesting that it is unlikely that this TF possesses a role in leukemogenesis on its own. Additionally, ZNF217 was found to be dispensable for myeloid differentiation, as knockdown (KD) of this TF failed to inhibit this process.

Whilst studies have determined the transcriptomic changes observed in cells expressing RUNX1-ETO, there is a paucity of studies quantitating proteomic changes. Therefore, this study also aimed at analysing the proteomic profile of RUNX1-ETO expressing HSPC using quantitative proteomics by SWATH-MS (on different subcellular structures including cytosolic or nuclear fractions). 4,635 proteins were quantified, of which 2,787 were detected in the cytoplasm, and 1,848 in the nucleus. Statistical analysis identified 257 significantly differentially expressed proteins in RUNX1-ETO compared to controls; of which 71% were detected in cytoplasm and 29% in the nucleus. RUNX1-ETO significantly downregulated the expression of C/EBP $\beta$  protein and mRNA vs control suggesting transcriptional suppression by RUNX1-ETO. Knocking-down C/EBP $\beta$  expression in HSPC, however, failed to induce significant changes in both monocytic and granulocytic development. Interestingly, KD of C/EBP $\beta$  in the RUNX1-ETO-expressing cell line, SKNO-1, completely suppressed myeloid cell surface marker expression and gave a concomitant increase in cell proliferation. In non-t(8;21) cells lines (HEL and U937), on the other hand, KD of C/EBP $\beta$  ablated cell growth and increased apoptotic frequency, suggesting that the effects of C/EBP $\beta$  KD are context dependent.

In conclusion, both transcriptomic and proteomic analysis proved to be useful tools for the identification of potential mediators of the block in terminal differentiation observed in RUNX1-ETO-expressing cells. Subsequently, ZNF217 and C/EBP $\beta$  were identified as targets of interest in the context of t(8;21). Whilst it is unlikely that ZNF217 overexpression contributes to leukaemogenic development, additional studies would be necessary to fully determine the role of C/EBP $\beta$  in this process.

## **Acknowledgements**

I would like to express my deepest gratitude to my supervisors Prof. Alex Tonks and Prof. Richard Darley for their continuous support and guidance throughout my PhD. Thanks to your encouraging words, your patience, and your knowledge, you have allowed me to grow as a scientist over the last three years, and provided me with invaluable skills for the future. I would also like to thank Dr. Steven Knapper, my mentor, for his advice and for providing me insights into the clinical aspects of AML. I would also like to extend my appreciation to Prof. Tony Whetton, Dr. Andrew Pierce, and Dr. Bethany Geary (University of Manchester), for their contribution to the project, and help and advice on mass spectrometry techniques and analysis.

Within the Haematology Department, I would like to particularly thank Amanda Gilkes for all her hard work during these unprecedented times, which allowed me to safely return to the laboratory to conclude my studies. A special thank you to Dr. (!) Rachael Nicholson, for always being there for me, to help and to listen, for both the good and the bad days that inevitably come with the PhD. I would also like to thank Sara Davies for her training and support during the beginning stages of my PhD. Additionally, I would like to thank Dr. Amanda Tonks, director of PGR, for not only helping me in securing funding for the last months of my PhD, but also for your continuous advice and support throughout this time. Thank you to Dr. Ann Kift-Morgan and Dr. Catherine Naseriyan from the Central Biotechnology Service at Cardiff University for your support and guidance.

I would particularly like to thank my family. Words cannot express how grateful I am for having you in my life, whether you're near or far. I am endlessly grateful for your constant support and words of encouragement, for advising me when stress got in the way and for believing in me, always. Adam, I want to especially thank you for everything you've done for me, for your incredible patience, for your constant support, and for never leaving my side. I couldn't have done this without you! During my PhD studies, I was lucky enough to have found a Portuguese family in the middle of Cardiff. To Ana, Sara and Tatiana, I would specially like to thank you for your friendship and support during these years, and for constantly encouraging me to keep going. Thank you for always being there, never complaining about hearing my numerous problems.

Lastly, I would like to thank School of Medicine in Cardiff university for providing the funding for a scholarship to undertake this research.

**Thank you! Obrigado! Dziękuję!**

## **Publications and Presentations**

### **Publications**

**Abstract:** SWATH-MS Proteomic Analysis Identifies C/EBP $\beta$  Suppression by RUNX1-ETO as a Potential Mediator of Block in Differentiation; A. Azevedo, A. Pierce, B. Geary, A. Whetton, R. L. Darley, A. Tonks; EHA2021 Virtual Congress Abstract Book, HemaSphere: June 2021; Volume 5; Issue p e566; doi: <https://doi.org/10.1097/HS9.0000000000000566>.

### **Oral Presentations**

2018-2021 - Haematology Department Seminar Series, Cardiff University, UK

2019 - Division of Cancer and Genetics PGR Symposium, Cardiff University, UK

2021 - Schools of Medicine and Dentistry PGR symposium, Cardiff University, UK

### **Poster Presentations**

2021 – European Hematology Association; Virtual

## Table of Contents

<i>Abstract</i> .....	<i>i</i>
<i>Acknowledgements</i> .....	<i>ii</i>
<i>Publications and Presentations</i> .....	<i>iii</i>
<i>Table of Contents</i> .....	<i>iv</i>
<i>Summary of Figures</i> .....	<i>xii</i>
<i>Summary of Tables</i> .....	<i>xix</i>
<i>List of Abbreviations</i> .....	<i>xx</i>

## **Chapter 1 ..... 1**

<b>1.1 Haematopoiesis</b> .....	<b>2</b>
1.1.1 Overview .....	2
1.1.2 Haematopoietic stem cells .....	3
1.1.2.1 Properties of haematopoietic stem cells .....	5
1.1.2.2 Haematopoietic stem cell niche.....	5
1.1.3 Hierarchical organisation of haematopoiesis.....	9
1.1.4 Regulation of normal human haematopoiesis .....	11
1.1.4.1 Transcription factor regulation of haematopoietic development.....	11
1.1.4.2 Cytokines and growth factors.....	14
<b>1.2 Acute Myeloid Leukaemia</b> .....	<b>16</b>
1.2.1 Overview .....	16
1.2.2 Pathophysiology of AML.....	16
1.2.2.1 Leukaemic Stem Cells.....	19
1.2.3 Classification of AML.....	21
1.2.4 Epidemiology, diagnosis, and prognosis of AML.....	24
1.2.5 Treatment of AML .....	26
1.2.5.1 Targeted therapies .....	27
<b>1.3 Pathophysiology of RUNX1-ETO</b> .....	<b>29</b>
1.3.1 RUNX1.....	30

1.3.1.1	The RUNX family of transcription factors.....	30
1.3.1.2	RUNX1 structure and function.....	30
1.3.1.3	Transcriptional regulation by RUNX1 .....	33
1.3.1.4	Post-translational modifications .....	35
1.3.1.5	The role of RUNX1 in normal haematopoiesis .....	36
1.3.1.6	The role of RUNX1 in AML.....	37
1.3.2	ETO .....	38
1.3.3	RUNX1-ETO.....	41
1.3.3.1	RUNX1-ETO structure.....	41
1.3.3.2	Cooperation between RUNX1 and RUNX1-ETO.....	43
1.3.3.3	Identification of RUNX1-ETO-mediated transcriptional de-regulation.....	46
1.3.3.4	Molecular mechanisms of RUNX1-ETO-induced leukaemia.....	48
<b>1.4</b>	<b>ZNF217.....</b>	<b>50</b>
1.4.1	ZNF217 in haematopoietic development .....	53
1.4.2	ZNF217 as an oncogene .....	54
1.4.2.1	Proliferation.....	54
1.4.2.2	Replicative immortality .....	55
1.4.2.3	Drug resistance .....	56
1.4.2.4	Cell differentiation.....	56
<b>1.5</b>	<b>C/EBP<math>\beta</math>.....</b>	<b>57</b>
1.5.1	Structure and function .....	59
1.5.2	C/EBP $\beta$ in haematopoietic development.....	62
1.5.3	C/EBP $\beta$ as an oncogene.....	64
1.5.3.1	Cell survival .....	64
1.5.3.2	Tumour aggressiveness .....	65
1.5.3.3	Drug resistance .....	66
1.5.4	C/EBP $\beta$ in leukaemogenesis.....	66
<b>1.6</b>	<b>Aims and objectives.....</b>	<b>67</b>

<b>Chapter 2 .....</b>	<b>71</b>
<b>2.1 Molecular Biology .....</b>	<b>72</b>
2.1.1 Plasmids used in the study.....	72
2.1.2 Bacterial growth medium and selective agar plates .....	72
2.1.3 Isolation and quantification of recombinant plasmid DNA from bacteria .....	76
2.1.4 DNA Sequencing.....	76
<b>2.2 Cell culture.....</b>	<b>77</b>
2.2.1 Culture of cell lines .....	77
2.2.1.1 Cell culture of non-adherent cell lines.....	77
2.2.1.2 Cell culture of adherent cell lines .....	77
2.2.2 Determination of cell density .....	80
2.2.3 Determination of cell viability.....	81
2.2.4 Cryopreservation and thawing of cell lines .....	81
<b>2.3 Primary cell culture .....</b>	<b>82</b>
2.3.1 Isolation of mononuclear cells from human neonatal cord blood .....	82
2.3.2 Isolation of CD34 <sup>+</sup> HSPC from mononuclear cells .....	82
2.3.3 Culture of CD34 <sup>+</sup> HSPC.....	83
2.3.4 Colony forming efficiency .....	85
2.3.5 Assessment of cell morphology.....	85
<b>2.4 Virus transfection.....</b>	<b>85</b>
2.4.1 Transfection using Lipofection.....	87
2.4.2 Transfection using calcium-phosphate precipitation .....	88
2.4.3 Virus titration.....	88
<b>2.5 Transduction of cell lines and human CD34<sup>+</sup> HSPC.....</b>	<b>89</b>
2.5.1 Transduction of cell lines .....	89
2.5.2 Transduction of human CD34 <sup>+</sup> HSPCs .....	89
<b>2.6 Determination of protein expression by western blot .....</b>	<b>90</b>
2.6.1 Total protein extraction .....	90

2.6.2	Fractionated protein extraction.....	90
2.6.3	Protein quantification using Bradford assay.....	91
2.6.4	Protein electrophoresis and electroblotting .....	92
2.6.5	Immunoblotting and protein detection .....	93
2.6.6	Membrane and gel protein staining .....	93
2.6.7	Densitometry analysis .....	95
<b>2.7</b>	<b>Flow cytometry .....</b>	<b>95</b>
2.7.1	Equipment and Analysis.....	95
2.7.2	Estimation of CD34 <sup>+</sup> cell count .....	95
2.7.3	Fluorescence Activated Cell Sorting .....	98
2.7.4	Immunophenotypic analysis of transduced CD34 <sup>+</sup> HSPCs.....	98
2.7.5	Cell cycle analysis .....	102
2.7.6	Apoptosis.....	102
2.7.7	Growth and proliferation assay.....	105
<b>2.8</b>	<b>Proteomic analysis using SWATH®.....</b>	<b>105</b>
2.8.1	Sample preparation.....	105
2.8.2	Tryptic digestion.....	105
2.8.3	Strong Cation Exchange Clean-up .....	106
2.8.4	Relative quantification by SWATH acquisition .....	106
2.8.5	SWATH-MS statistical analysis.....	107
<b>2.9</b>	<b>Analysis of existing mRNA expression data .....</b>	<b>107</b>
2.9.1	Analysis of microarray data of RUNX1-ETO induced changes in mRNA expression in CD34 <sup>+</sup> HSPC.....	107
2.9.2	Publicly available databases .....	108
2.9.2.1	Bloodspot.....	108
2.9.2.2	cBioPortal.....	108
<b>2.10</b>	<b>Meta-analysis of mRNA and proteomic data .....</b>	<b>108</b>
2.10.1	Pathway analysis using Metacore.....	109

2.10.2	Interactome analysis using Metacore .....	111
2.10.3	Pathway analysis and identification of transcription factors using Ingenuity Pathway Analysis	112
<b>2.11</b>	<b>Statistical analysis .....</b>	<b>112</b>
<b>Chapter 3 .....</b>		<b>113</b>
<b>3.1</b>	<b>Introduction .....</b>	<b>114</b>
<b>3.2</b>	<b>Hypothesis and Aims.....</b>	<b>116</b>
<b>3.3</b>	<b>Results .....</b>	<b>119</b>
3.3.1	Transcriptional dysregulation mediated by RUNX1-ETO .....	119
3.3.1.1	RUNX1-ETO expression in HSPC alters TF gene expression and is associated with GO processes that regulate transcription .....	119
3.3.1.2	Identification of aberrantly expressed TFs driving transcriptional change in AML	127
3.3.2	Quantitative SWATH-MS analysis reveals proteomic changes arising from RUNX1-ETO expression, in CD34 <sup>+</sup> HSPC.....	136
3.3.2.1	Generation of human CD34 <sup>+</sup> HSPC expressing RUNX1-ETO for SWATH-MS analysis	136
3.3.2.2	Assessment of CD34 <sup>+</sup> HSPC nuclear and cytosolic extracts shows sample purity and integrity .....	136
3.3.2.3	SWATH-MS analysis detected over 4000 proteins in CD34 <sup>+</sup> HSPC.....	137
3.3.2.4	Exploratory data analysis .....	142
3.3.2.5	Statistical model used for data analysis.....	143
3.3.2.6	Expression of RUNX1-ETO disrupts normal protein expression in CD34 <sup>+</sup> cord blood derived HSPC .....	147
3.3.2.7	Identification of aberrantly expressed TF in CD34 <sup>+</sup> HSPC expressing RUNX1-ETO	150
3.3.2.8	RUNX1-ETO leads to the aberrant expression of 53 proteins .....	153
3.3.2.9	RUNX1-ETO leads to a shift in protein sub-cellular localisation in CD34 <sup>+</sup> HSPC.	157
<b>3.4</b>	<b>Discussion.....</b>	<b>162</b>

3.4.1	Transcriptional dysregulation imposed by RUNX1-ETO expression .....	162
3.4.2	Dysregulation of the proteome arising from RUNX1-ETO expression .....	167
3.4.2.1	Alternative analysis of RUNX1-ETO-induced proteomic changes in CD34 <sup>+</sup> HSPC	171

## **Chapter 4 .....174**

<b>4.1</b>	<b>Introduction.....</b>	<b>175</b>
<b>4.2</b>	<b>Hypothesis and Aims.....</b>	<b>176</b>
<b>4.3</b>	<b>Results .....</b>	<b>177</b>
4.3.1	Expression of <i>ZNF217</i> mRNA increases throughout myeloid cell development. ....	177
4.3.2	<i>ZNF217</i> is variably expressed across AML subtypes .....	177
4.3.3	Overexpression of <i>ZNF217</i> promotes human myeloid development.....	184
4.3.3.1	Generation of <i>ZNF217</i> overexpressing CD34 <sup>+</sup> HSPC .....	184
4.3.3.2	<i>ZNF217</i> overexpression suppresses myeloid colony formation and self-renewal	186
4.3.3.3	<i>ZNF217</i> overexpression suppresses the growth of myeloid cells .....	186
4.3.3.4	<i>ZNF217</i> overexpression promotes myeloid differentiation.....	190
4.3.3.5	<i>ZNF217</i> overexpression in HSPC induces changes in cell cycle .....	194
4.3.4	Knockdown of <i>ZNF217</i> does not impact myeloid differentiation.....	199
4.3.4.1	Selection and validation of <i>ZNF217</i> shRNA constructs.....	199
4.3.4.2	Knockdown of <i>ZNF217</i> leads to a decrease in CD34 <sup>+</sup> HSPC colony forming ability	199
4.3.4.3	Knockdown of <i>ZNF217</i> disrupts myeloid cell growth and development.....	203
4.3.4.4	Knockdown of <i>ZNF217</i> does not affect cell cycle and apoptosis .....	206
4.3.5	The role of <i>ZNF217</i> in AML cell lines.....	210
4.3.5.1	<i>ZNF217</i> expression is variable across AML cell lines .....	210
4.3.5.2	Generation of <i>ZNF217</i> -knockdown AML cell lines .....	210
4.3.5.3	<i>ZNF217</i> knockdown in AML cell lines reduces cell growth by promoting changes	214
	in cell cycle and apoptosis .....	
<b>4.4</b>	<b>Discussion.....</b>	<b>217</b>

4.4.1	High <i>ZNF217</i> in AML is associated with increased risk of relapse .....	218
4.4.2	<i>ZNF217</i> overexpression of promotes myeloid development in HSPC .....	219
4.4.3	<i>ZNF217</i> knockdown in HSPC does not impact myeloid development .....	222
<b>Chapter 5 .....</b>		<b>223</b>
<b>5.1</b>	<b>Introduction .....</b>	<b>224</b>
<b>5.2</b>	<b>Hypothesis and Aims.....</b>	<b>226</b>
<b>5.3</b>	<b>Results .....</b>	<b>227</b>
5.3.1	Expression of <i>CEBPB</i> mRNA increases throughout myeloid cell development..	227
5.3.2	<i>CEBPB</i> is variably expressed across AML subtypes .....	228
5.3.3	Knockdown of <i>C/EBPβ</i> influences normal myeloid development.....	237
5.3.3.1	Generation of <i>C/EBPβ</i> -knockdown CD34 <sup>+</sup> HSPC .....	237
5.3.3.2	Knockdown of <i>C/EBPβ</i> leads to a decrease in HSPC colony forming ability ..	237
5.3.3.3	Knockdown of <i>C/EBPβ</i> supresses myeloid cell growth.....	240
5.3.3.4	Knockdown of <i>C/EBPβ</i> influences myeloid differentiation.....	240
5.3.3.5	Knockdown of <i>C/EBPβ</i> does not result in significant cell cycle or apoptotic changes	247
5.3.4	Overexpression of <i>C/EBPβ</i> influences normal myeloid development .....	247
5.3.4.1	Generation of <i>C/EBPβ</i> -overexpressing HSPC .....	247
5.3.4.2	<i>C/EBPβ</i> overexpression promotes myeloid colony formation and self-renewal ability	250
5.3.4.3	<i>C/EBPβ</i> promotes myeloid cell growth.....	250
5.3.4.4	<i>C/EBPβ</i> expression promotes myeloid differentiation .....	253
5.3.4.5	Overexpression of <i>C/EBPβ</i> induces change in the cell cycle of HSPC.....	255
5.3.5	The effect of <i>C/EBPβ</i> mis-regulation in AML cell lines.....	260
5.3.5.1	Expression of <i>C/EBPβ</i> is variable across AML cell lines .....	260
5.3.5.2	Generation of <i>C/EBPβ</i> -knockdown cell lines.....	260
5.3.5.3	The impact of <i>C/EBPβ</i> knockdown on the AML t(8;21) cell line SKNO-1 .....	265
5.3.5.4	The impact of <i>C/EBPβ</i> knockdown on AML non-t(8;21) cell lines.....	267

<b>5.4</b>	<b>Discussion.....</b>	<b>277</b>
5.4.1	<i>CEBPB</i> expression increases during haematopoietic development.....	278
5.4.2	Increased <i>CEBPB</i> expression is associated with poor prognosis in AML .....	279
5.4.3	Knockdown of C/EBP $\beta$ in HSPC does not influence normal myeloid development 279	
5.4.4	C/EBP $\beta$ overexpression in HSPC promotes myeloid cell growth and development 281	
5.4.5	The consequences of C/EBP $\beta$ KD in AML cell lines is context-dependant .....	283
<b>Chapter 6</b>	<b>.....</b>	<b>285</b>
<b>6.1</b>	<b>Conclusions.....</b>	<b>286</b>
6.1.1	RUNX1-ETO-induced transcriptomic and proteomic changes in HSPC .....	287
6.1.2	ZNF217 does not contribute to the pathogenesis of AML .....	288
6.1.3	KD of C/EBP $\beta$ does not contribute to the pathogenesis of AML.....	289
6.1.4	Limitations of the study.....	291
6.1.5	Future directions.....	293
	<i>Supplementary Tables</i> .....	297
	<i>Supplementary Figures</i> .....	306
	<i>References</i> .....	366

## Summary of Figures

Figure 1.1 – Hierarchical representation of human haematopoietic development.....	4
Figure 1.2 – HSC regulation by the BM microenvironment.....	7
Figure 1.3 – Transcriptional regulation of myeloid development .....	12
Figure 1.4 – Structural representation of RUNX1 isoforms genes and proteins .....	32
Figure 1.5 – Transcriptional regulation mediated by RUNX1.....	34
Figure 1.6 – Structural representation of the ETO gene and protein .....	40
Figure 1.7 – RUNX1-ETO structure and isoforms.....	42
Figure 1.8 – Mechanisms of transcriptional dysregulation induced by RUNX1-ETO.....	45
Figure 1.9 – Structural representation of the ZNF217 gene and protein .....	51
Figure 1.10 – ZNF217 mediated transcriptional repression .....	52
Figure 1.11 – Structural representation of the C/EBP $\beta$ gene and protein isoforms.....	60
Figure 1.12 – Hypothesis of the study .....	70
Figure 2.1 – Plasmid DNA constructs used in this study.....	75
Figure 2.2 – Morphological analysis of HSPC .....	86
Figure 2.3 – Estimation of CD34 <sup>+</sup> HSPC in cord blood by flow cytometry.....	97
Figure 2.4 – Gating strategy used to sort GFP-expressing CD34 <sup>+</sup> HSPC.....	99
Figure 2.5 – Immunophenotypic analysis of CD34 <sup>+</sup> HSPC.....	101
Figure 2.6 – Gating strategy used to perform cell cycle analysis .....	103
Figure 2.7 –Gating strategy used to analyse apoptosis .....	104
Figure 2.8 - Strategy for the analysis of differentially expressed genes/protein to study the effects of RUNX1-ETO overexpression in human CD34 <sup>+</sup> HSPC.....	110
Figure 3.1 – Strategy for the identification of changes in gene/protein expression arising from the expression of the fusion protein RUNX1-ETO .....	118
Figure 3.2 – Functional Enrichment Analysis of mRNA changes observed in human CD34 <sup>+</sup> HSPC expressing RUNX1-ETO.....	120
Figure 3.3 – TF were identified as the most significantly represented functional class due to RUNX1-ETO expression in CD34 <sup>+</sup> HSPC.....	122
Figure 3.4 – RUNX1-ETO does not influence the expression of IRF7 in CD34 <sup>+</sup> HSPC.....	130
Figure 3.5 – RUNX1-ETO leads to the upregulation of IRF9 in CD34 <sup>+</sup> HSPC.....	131
Figure 3.6 – RUNX1-ETO leads to the upregulation of ARID5B in CD34 <sup>+</sup> HSPC.....	132
Figure 3.7 – RUNX1-ETO leads to the upregulation of ZNF217 in CD34 <sup>+</sup> HSPC .....	133

Figure 3.8 – RUNX1-ETO leads to the upregulation of protein targets, in CD34 <sup>+</sup> HSPC .....	134
Figure 3.9 – Viral infection of CD34 <sup>+</sup> cells. CD34 <sup>+</sup> HSPC cells were analysed for GFP expression on day 3 of cell culture .....	138
Figure 3.10 – Quality control and characterisation of CD34 <sup>+</sup> HSPC overexpressing RUNX1-ETO .....	139
Figure 3.11 – GAPDH and Histone H1 protein quantitation by SWATH-MS supports acceptable subcellular fraction efficiencies for cytosol and nuclear protein isolation .....	140
Figure 3.12 – RUNX1 is significantly overexpressed in RUNX1-ETO transfected CD34 <sup>+</sup> HSPC .....	141
Figure 3.13 – Principal Component Analysis (PCA) of the protein expression data .....	144
Figure 3.14 – Hierarchical clustering of proteins expressed in the cytoplasmic and nuclear compartment of CD34 <sup>+</sup> HSPC .....	145
Figure 3.15 – Differential protein expression analysis between CD34 <sup>+</sup> HSPC expressing RUNX1-ETO and control.....	146
Figure 3.16 – Functional Enrichment Analysis of protein changes observed in normal haematopoietic development vs. RUNX1-ETO expressing CD34 <sup>+</sup> HSPC.....	148
Figure 3.17 – Classification of differentially expressed proteins due to the expression of RUNX1-ETO in CD34 <sup>+</sup> HSPC.....	149
Figure 3.18 – C/EBPβ expression profile in CD34 <sup>+</sup> HSPC expressing RUNX1-ETO.....	154
Figure 3.19 – Target protein and mRNA expression profiles of TF found exclusively in RUNX1-ETO CD34 <sup>+</sup> HSPC.....	156
Figure 3.20 – Proteins are mis-localised to the nucleus or cytoplasm as a consequence of RUNX1-ETO expression, in CD34 <sup>+</sup> HSPC.....	159
Figure 3.21 – RUNX1-ETO leads to a shift in localisation of 9 proteins .....	160
Figure 4.1 – <i>ZNF217</i> mRNA expression in normal haematopoiesis .....	178
Figure 4.2 – <i>ZNF217</i> mRNA expression levels are variable across AML subtypes .....	179
Figure 4.3 – High levels of <i>ZNF217</i> is associated with poorer disease-free survival of AML patients .....	181
Figure 4.4 – High <i>ZNF217</i> expression is associated with decreased BM blasts and lower mutation count .....	182
Figure 4.5 – Dysregulated <i>ZNF217</i> expression is not associated with patient characteristics.....	183
Figure 4.6 – Transduction of <i>ZNF217</i> in HEK 293T cells and human HSPC.....	185
Figure 4.7 – Expression and enrichment of GFP <sup>+</sup> populations in human CD34 <sup>+</sup> HSPC .....	187

Figure 4.8 – ZNF217 overexpression inhibits myeloid colony formation and decreases self-renewal capacity .....	188
Figure 4.9 – ZNF217 overexpression inhibits monocytic and granulocytic growth during myeloid development.....	189
Figure 4.10 – ZNF217 overexpression disrupts the balance between granulocytic and erythroid populations during myeloid cell development.....	191
Figure 4.11 – ZNF217 overexpression promotes monocytic differentiation.....	193
Figure 4.12 – ZNF217 overexpression promotes granulocytic differentiation by upregulating monocytic markers.....	195
Figure 4.13 – ZNF217 overexpression alters cell morphology during myeloid development .....	196
Figure 4.14 – ZNF217 overexpression in CD34 <sup>+</sup> HSPC significantly disrupts normal cell cycle.....	197
Figure 4.15 – ZNF217 shRNA constructs successfully reduced the levels of endogenous ZNF217 in K562 cells .....	200
Figure 4.16 – Infection efficiency of transduced CD34 <sup>+</sup> HSPC using shRNA lentiviral constructs .....	201
Figure 4.17 – ZNF217 knockdown disrupts myeloid colony formation of HSPC .....	202
Figure 4.18 – Knockdown of ZNF217 promotes monocytic cell growth in myeloid development .....	204
Figure 4.19 – ZNF217 knockdown disrupts the balance between the monocytic and granulocytic populations in CD34 <sup>+</sup> HSPC during myeloid cell development.....	205
Figure 4.20 – Knockdown of ZNF217 does not significantly alter monocytic development.....	207
Figure 4.21 – Knockdown of ZNF217 does not significantly alter granulocytic development.....	208
Figure 4.22 – Knockdown of ZNF217 has no effect on normal cell cycle or apoptosis .....	209
Figure 4.23 – ZNF217 expression in leukaemic cell lines.....	211
Figure 4.24 – ZNF217 is knocked-down in the erythroleukaemia cell line HEL.....	212
Figure 4.25 – ZNF217 is knocked-down in the t(8;21) cell line SKNO-1 .....	213
Figure 4.26 – ZNF217 knockdown significantly decreases HEL cell proliferation .....	215
Figure 4.27 – ZNF217 knockdown significantly decreases SKNO-1 cell proliferation.....	216
Figure 5.1 – <i>CEBPB</i> mRNA expression in normal haematopoiesis .....	229
Figure 5.2 – <i>C/EBPβ</i> endogenous expression throughout haematopoiesis.....	230
Figure 5.3 – <i>CEBPB</i> mRNA expression levels are variable across AML subtypes .....	232
Figure 5.4 – High levels of <i>CEBPB</i> is associated with poorer disease-free survival of AML patients .....	233

Figure 5.5 – High <i>CEBPB</i> expression is associated with increased WBC and lower mutation count .....	234
Figure 5.6 – Dysregulated <i>CEBPB</i> expression is associated with patient characteristics .....	235
Figure 5.7 – <i>CEBPB</i> expression correlates with cytogenetic and molecular abnormalities .....	236
Figure 5.8 – C/EBP $\beta$ is knocked-down in human CD34 <sup>+</sup> HSPC.....	238
Figure 5.9 – C/EBP $\beta$ KD disrupts myeloid colony formation of HSPC.....	239
Figure 5.11 – C/EBP $\beta$ KD disrupts the balance between the monocytic and granulocytic populations in CD34 <sup>+</sup> HSPC during myeloid cell development .....	242
Figure 5.12 – KD of C/EBP $\beta$ has little impact on normal monocytic differentiation .....	243
Figure 5.13 – KD of C/EBP $\beta$ promotes granulocytic development .....	245
Figure 5.14 – C/EBP $\beta$ KD affects cell morphology during myeloid development .....	246
Figure 5.15 – KD of C/EBP $\beta$ has no significant effect on normal cell cycle or apoptosis.....	248
Figure 5.16 – C/EBP $\beta$ overexpression in human HSPC.....	249
Figure 5.17 – C/EBP $\beta$ overexpression promotes myeloid colony formation of HSPC .....	251
Figure 5.18 – Overexpression of C/EBP $\beta$ promotes monocytic and granulocytic growth in myeloid development.....	252
Figure 5.19 – C/EBP $\beta$ overexpression disrupts the balance between myeloid population during haematopoietic development .....	254
Figure 5.20 – C/EBP $\beta$ overexpression promotes the upregulation of monocytic markers in monocytic progenitors .....	256
Figure 5.21 – C/EBP $\beta$ overexpression promotes the upregulation of granulocytic markers in granulocytic progenitors .....	257
Figure 5.22 – C/EBP $\beta$ overexpression alters cell morphology during myeloid development.....	258
Figure 5.23 – Overexpression of C/EBP $\beta$ influences normal cell cycle.....	259
Figure 5.24 – C/EBP $\beta$ expression in leukaemic cell lines .....	261
Figure 5.25 – C/EBP $\beta$ KD in the t(8;21) cell line SKNO-1 .....	262
Figure 5.26 – C/EBP $\beta$ KD in the erythroleukaemia cell line HEL.....	263
Figure 5.27 – C/EBP $\beta$ KD in the leukaemic cell line U937 .....	264
Figure 5.28 – C/EBP $\beta$ KD promotes cell growth in SKNO-1 cells, associated with changes in normal cell cycle and apoptotic rate.....	266
Figure 5.29 – The effect of C/EBP $\beta$ KD on cell surface marker expression in SKNO-1 cells .....	268
Figure 5.30 – Morphological effect of C/EBP $\beta$ KD in SKNO-1 cells .....	269

Figure 5.31 – C/EBP $\beta$ KD significantly reduces cell growth in HEL cells, associated with changes in apoptotic frequency .....	271
Figure 5.32 – The effect of C/EBP $\beta$ KD on cell surface marker expression in HEL cells.....	272
Figure 5.33 – Morphological effect of C/EBP $\beta$ KD in HEL cells.....	273
Figure 5.34 – C/EBP $\beta$ KD significantly reduces cell growth in U937 cells, associated with changes in normal cell cycle and apoptotic rate .....	274
Figure 5.35 – The effect of C/EBP $\beta$ KD on cell surface marker expression in U937 cells.....	275
Figure 5.36 – Morphological effect of C/EBP $\beta$ KD in U937 cells.....	276
Figure 6.1 – Summary of the main findings described on Chapters 3 and 4 .....	295
Figure 6.2 – Summary of the main findings described on Chapter 5 .....	296
Supplementary Figure 1 - Representative gating strategy used to immunophenotype HSPC	306
Supplementary Figure 2 – Expression of control GFP and ZNF217 in human HSPC .....	307
Supplementary Figure 3 – Expression of the stem cell marker CD34 in monocytes during differentiation, following ZNF217 overexpression .....	308
Supplementary Figure 4 – Expression of the myeloid marker CD11b in monocytes during differentiation, following ZNF217 overexpression .....	309
Supplementary Figure 5 – Expression of the monocytic marker CD14 in monocytes during differentiation, following ZNF217 overexpression .....	310
Supplementary Figure 6 – Expression of the stem cell marker CD34 in granulocytes during differentiation, following ZNF217 overexpression .....	311
Supplementary Figure 7 – Expression of the myeloid marker CD11b in granulocytes during differentiation, following ZNF217 overexpression .....	312
Supplementary Figure 8 – Expression of the granulocytic marker CD15 in granulocytes during differentiation, following ZNF217 overexpression .....	313
Supplementary Figure 9 – Expression of control GFP and shZNF217 KD in human HSPC.....	315
Supplementary Figure 10 – Expression of the stem cell marker CD34 in monocytes during differentiation, following ZNF217 KD.....	317
Supplementary Figure 11 – Expression of the myeloid marker CD11b in monocytes during differentiation, following ZNF217 KD.....	319
Supplementary Figure 12 – Expression of the monocytic marker CD14 in monocytes during differentiation, following ZNF217 KD.....	321

Supplementary Figure 13 – Expression of the stem cell marker CD34 in granulocytes during differentiation, following ZNF217 KD.....	323
Supplementary Figure 14 – Expression of the myeloid marker CD11b in granulocytes during differentiation, following ZNF217 KD.....	325
Supplementary Figure 15 – Expression of the granulocytic marker CD15 in granulocytes during differentiation, following ZNF217 KD.....	327
Supplementary Figure 16 – Expression of control GFP and shC/EBP $\beta$ KD in human HSPC .....	329
Supplementary Figure 17 – Expression of the stem cell marker CD34 in monocytes during differentiation, following C/EBP $\beta$ KD .....	331
Supplementary Figure 18 – Expression of the myeloid marker CD11b in monocytes during differentiation, following C/EBP $\beta$ KD .....	333
Supplementary Figure 18 – Expression of the monocytic marker CD14 in monocytes during differentiation, following C/EBP $\beta$ KD .....	335
Supplementary Figure 20 – Expression of the stem cell marker CD34 in granulocytes during differentiation, following C/EBP $\beta$ KD .....	337
Supplementary Figure 21 – Expression of the myeloid marker CD11b in granulocytes during differentiation, following C/EBP $\beta$ KD .....	339
Supplementary Figure 22 – Expression of the granulocytic marker CD15 in granulocytes during differentiation, following C/EBP $\beta$ KD .....	341
Supplementary Figure 23 – Expression of control GFP and C/EBP $\beta$ in human HSPC .....	342
Supplementary Figure 24 – Expression of the stem cell marker CD34 in monocytes during differentiation, following C/EBP $\beta$ overexpression.....	343
Supplementary Figure 25 – Expression of the myeloid marker CD11b in monocytes during differentiation, following C/EBP $\beta$ overexpression.....	344
Supplementary Figure 26 – Expression of the myeloid marker CD14 in monocytes during differentiation, following C/EBP $\beta$ overexpression.....	345
Supplementary Figure 27 – Expression of the stem cell marker CD34 in granulocytes during differentiation, following C/EBP $\beta$ overexpression.....	346
Supplementary Figure 28 – Expression of the myeloid marker CD11b in granulocytes during differentiation, following C/EBP $\beta$ overexpression.....	347
Supplementary Figure 29 – Expression of the granulocytic marker CD15 in granulocytes during differentiation, following C/EBP $\beta$ overexpression.....	348
Supplementary Figure 30 – ZNF217 overexpression inhibits monocytic and granulocytic growth during myeloid development .....	349

Supplementary Figure 31– ZNF217 overexpression disrupts the balance between granulocytic and erythroid populations during myeloid cell development .....	350
Supplementary Figure 32 – ZNF217 overexpression promotes monocytic differentiation.....	351
Supplementary Figure 33 – ZNF217 overexpression promotes granulocytic differentiation by upregulating monocytic markers.....	352
Supplementary Figure 34 – Knockdown of ZNF217 promotes monocytic cell growth in myeloid development.....	353
Supplementary Figure 35 – ZNF217 knockdown disrupts the balance between the monocytic and granulocytic populations in CD34 <sup>+</sup> HSPC during myeloid cell development .....	354
Supplementary Figure 36 – Knockdown of ZNF217 does not significantly alter monocytic development.....	355
Supplementary Figure 37 – Knockdown of ZNF217 does not significantly alter granulocytic development.....	356
Supplementary Figure 38 – KD of C/EBP $\beta$ decreases granulocytic growth in myeloid development .....	357
Supplementary Figure 39 – C/EBP $\beta$ KD disrupts the balance between the monocytic and granulocytic populations in CD34 <sup>+</sup> HSPC during myeloid cell development.....	358
Supplementary Figure 40 – KD of C/EBP $\beta$ has little impact on normal monocytic differentiation .....	359
Supplementary Figure 41 – KD of C/EBP $\beta$ promotes granulocytic development .....	360
Supplementary Figure 42 – Overexpression of C/EBP $\beta$ promotes monocytic and granulocytic growth in myeloid development .....	361
Supplementary Figure 43 – C/EBP $\beta$ overexpression disrupts the balance between myeloid population during haematopoietic development.....	362
Supplementary Figure 44 – C/EBP $\beta$ overexpression promotes the upregulation of monocytic markers in monocytic progenitors .....	363
Supplementary Figure 45 – C/EBP $\beta$ overexpression promotes the upregulation of granulocytic markers in granulocytic progenitors .....	364
Supplementary Figure 46 – C/EBP $\beta$ KD in the Kasumi-1 cell line.....	365

## Summary of Tables

Table 1.1 – Recurrent mutations in AML.....	18
Table 1.2 – FAB classification of AML .....	22
Table 1.3 – WHO (2016) classification of AML.....	23
Table 1.4 – ELN 2017 risk stratification of AML .....	25
Table 2.1 – Plasmid vectors used in this study .....	73
Table 2.2 – Primers used for direct sequencing .....	77
Table 2.3 – Culture conditions for non-adherent cells.....	78
Table 2.4 – Culture conditions for adherent cells .....	80
Table 2.5 – Cell culture media used for the growth and development of CD34 <sup>+</sup> HSPC .....	84
Table 2.6 – List of primary antibodies used in this study .....	94
Table 2.7 – Technical specifications of the Accuri™ C6 Plus used in this study .....	96
Table 2.8 – List of antibodies used for cord blood immunophenotyping .....	96
Table 2.9 – Antibodies used for immunophenotypic analysis of transduced CD34 <sup>+</sup> HSPC.....	100
Table 3.1 – Identification of RUNX1-ETO induced TF dysregulation in CD34 <sup>+</sup> HSPC .....	123
Table 3.2 – Summary showing the effects of RUNX1-ETO expression in CD34 <sup>+</sup> HSPC on gene and protein expression in prioritised targets.....	135
Table 3.3 – Identification of RUNX1-ETO induced TF protein dysregulation in CD34 <sup>+</sup> HSPC..	151
Table 3.4 – TF detected exclusively in RUNX1-ETO expressing CD34 <sup>+</sup> HSPC.....	155
Table 3.5 – Identification of ‘shifted’ proteins as a result of RUNX1-ETO expression .....	161
Table 4.1 - Summary of the main findings regarding the role of ZNF217 in normal human haematopoiesis.....	221
Table 5.1 – Summary of the main findings regarding the role of C/EBPβ in normal human haematopoiesis.....	282
Supplementary Table 1 – Compensation matrix.....	297
Supplementary Table 2 – Identification of significantly dysregulated proteins, as a result of RUNX1-ETO expression.....	298

## List of Abbreviations

<b>1,25(OH)<sub>2</sub>D<sub>3</sub></b>	1,25-Dihydroxyvitamin D <sub>3</sub>
<b>ALL</b>	Acute Lymphoblastic Leukaemia
<b>AmBic</b>	Ammonium Bicarbonate
<b>αMEM</b>	Minimum Essential Medium Eagle - Alpha Modification
<b>AML</b>	Acute Myeloid Leukaemia
<b>ANG1</b>	Angiopoietin 1
<b>ANXA2</b>	Annexin 2
<b>APC</b>	Allophycocyanin
<b>APL</b>	Acute Promyelocytic Leukaemia
<b>ATO</b>	Arsenic Trioxide
<b>ATRA</b>	All-Trans Retinoic Acid
<b>BM</b>	Bone Marrow
<b>β-ME</b>	β-Mercaptoethanol
<b>BSA</b>	Bovine Serum Albumin
<b>bZIP</b>	C-terminal Leucine-Zipper Dimerization Domain
<b>C/EBP</b>	CCAAT/Enhancer Binding Protein
<b>CAR</b>	CXCL12-abundant reticular cell
<b>CB</b>	Cord Blood
<b>CBF</b>	Core Binding Factor
<b>CBW</b>	Cord Blood Wash
<b>CD</b>	Cluster of Differentiation
<b>CDK</b>	Cyclin Dependent Kinase
<b>CDS</b>	Coding Sequence
<b>CFU</b>	Colony-Forming Unit
<b>ChIP-Seq</b>	Chromatin Immunoprecipitation Sequencing
<b>CHOP</b>	C/EBP-Homologous Protein
<b>c-KIT</b>	Tyrosine-Protein Kinase KIT
<b>CLP</b>	Common Lymphoid Progenitor Cell
<b>CML</b>	Chronic Myeloid Leukaemia
<b>CMP</b>	Common Myeloid Progenitor Cell
<b>CR</b>	Complete remission
<b>CSC</b>	Cancer Stem Cell
<b>CSF</b>	Colony Stimulating Factor
<b>CtBP</b>	C-Terminal Binding Protein
<b>CXCL12</b>	CXC-Chemokine ligand 12
<b>DC</b>	Dendritic Cell
<b>DEG</b>	Differentially Expressed Genes
<b>DEP</b>	Differentially Expressed Proteins
<b>dH<sub>2</sub>O</b>	Deionised Water
<b>DIA</b>	Data-Independent Acquisition
<b>DMEM</b>	Dulbecco's Modified Eagle Medium
<b>DMNT</b>	DNA Methyltransferase

<b>DMSO</b>	Dimethyl Sulfoxide
<b>EGR</b>	Early Growth Response
<b>ELN</b>	European LeukemiaNet
<b>EMP</b>	Erythro-myeloid Progenitor Cell
<b>EMT</b>	Epithelial-Mesenchymal Transition
<b>EPO</b>	Erythropoietin
<b>ERK</b>	Extracellular Signal-Regulated Kinases
<b>ETO</b>	Eight-Twenty-One
<b>ETS</b>	E26 Transformation-Specific
<b>FAB</b>	French-American-British
<b>FACS</b>	Fluorescently Activated Cell Sorting
<b>FBS</b>	Foetal Bovine Serum
<b>FC</b>	Fold Change
<b>FDA</b>	US Food and Drug Administration
<b>FDR</b>	False Discovery Rate
<b>FGF</b>	Fibroblast Growth Factor
<b>FITC</b>	Fluorescein isothiocyanate
<b>FLI1</b>	Friend Leukaemia Integration 1
<b>FLT3</b>	Fms-like Tyrosine Kinase 3
<b>FSC</b>	Forward Scatter
<b>GAPDH</b>	Glyceraldehyde 3-Phosphate Dehydrogenase
<b>GATA</b>	Globin Transcription Factor
<b>G-CSF</b>	Granulocyte Colony Stimulating Factor
<b>GFI1</b>	Growth-Factor Independent 1
<b>GFP</b>	Green Fluorescent Protein
<b>GlyA</b>	Glycophorin A (CD235a)
<b>GM-CSF</b>	Granulocyte/Macrophage Colony Stimulating Factor
<b>GMP</b>	Granulocyte-Monocyte Progenitor Cell
<b>GO</b>	Gene Ontology
<b>GSC</b>	Glioma Stem Cell
<b>HBSS</b>	Hank's Balanced Salt Solution
<b>HDAC</b>	Histone Deacetylase
<b>HMA</b>	Hypomethylating Agents
<b>HMEC</b>	Human Mammary Epithelial Cells
<b>HR</b>	Hazard Ratio
<b>HRP</b>	Horseradish Peroxidase
<b>HSC</b>	Haematopoietic Stem Cell
<b>HSPC</b>	Haematopoietic Stem and Progenitor Cell
<b>IDH</b>	Isocitrate Dehydrogenase (NADP(+))
<b>IL</b>	Interleukin
<b>IMDM</b>	Iscove Modified Dulbecco's Medium
<b>Inv</b>	Inversion
<b>IOSE</b>	Human Ovarian Surface Epithelial Cell Line
<b>IPA</b>	Ingenuity Pathway Analysis

<b>IRF</b>	Interferon-Regulatory Factor
<b>IT-HSC</b>	Intermediate-Term Haematopoietic Stem Cell
<b>ITD</b>	Internal Tandem Duplication
<b>KD</b>	Knockdown
<b>KO</b>	Knockout
<b>KRAS</b>	Kirsten Rat Sarcoma Virus
<b>LAP</b>	Liver-Enriched Activating Protein
<b>LAP*</b>	Liver-Enriched Activating Protein*
<b>LB</b>	Luria-Bertani
<b>LDS</b>	Lithium Dodecyl Sulphate Sample Buffer
<b>LepR</b>	Leptin Receptor Perivascular-Expressing Cell
<b>Lin</b>	Lineage
<b>LIP</b>	Liver-Enriched Inhibitory Protein
<b>LMPP</b>	Lymphoid-Primed Multipotent Progenitor Cell
<b>LSC</b>	Leukaemic Stem Cell
<b>LSD</b>	Lysine-Specific Histone Demethylase
<b>LT-HSC</b>	Long-Term Haematopoietic Stem Cell
<b>MACS</b>	Magnetic-Activated Cell Sorting
<b>M-CSF</b>	Macrophage Colony Stimulating Factor
<b>MDP</b>	Monocyte-Dendritic Progenitor Cell
<b>MDS</b>	Myelodysplastic syndrome
<b>MEP</b>	Megakaryocytic-Erythroid Progenitor Cell
<b>MFI</b>	Mean Fluorescent Intensity
<b>MgCl<sub>2</sub></b>	Magnesium Chloride
<b>miRNA</b>	microRNAs
<b>MLL</b>	Mixed Lineage Leukaemia
<b>MPP</b>	Multipotent Progenitor Cell
<b>MS</b>	Mass Spectrometry
<b>MSC</b>	Mesenchymal Stem Cell
<b>mSin3a</b>	Mammalian Sin3a
<b>N-CoR</b>	Human Nuclear Receptor Co-Repressor
<b>Nes</b>	Nestin-positive Cell
<b>NF-κB</b>	Nuclear Factor-κB
<b>NHR</b>	Nervy Homology Region
<b>NK</b>	Natural Killer Cell
<b>NLS</b>	Nuclear Localisation Signal
<b>NPM1</b>	Nucleophosmin-1
<b>NRAS</b>	Neuroblastoma Rat Sarcoma Virus
<b>OE</b>	Overexpression
<b>OIS</b>	Oncogene-Induced Senescence
<b>OPN</b>	Osteopontin
<b>OS</b>	Overall Survival
<b>P300</b>	Histone Acetyltransferase p300
<b>PB</b>	Peripheral Blood

<b>PBS</b>	Phosphate Buffer Saline
<b>PCA</b>	Principal Component Analysis
<b>PE</b>	Phycoerythrin
<b>PerCP-Cy5.5</b>	Peridinin chlorophyll protein – Cyanine5.5
<b>PI</b>	Propidium iodide
<b>PLZF</b>	Promyelocytic Leukaemia Zinc-Finger
<b>PRMT</b>	Protein Arginine N-Methyltransferase
<b>PVDF</b>	Polyvinylidene Difluoride
<b>RD</b>	Regulatory Domain
<b>RHD</b>	Runt Homology Domain
<b>RNASeq</b>	RNA Sequencing
<b>ROI</b>	Region of Interest
<b>RPMI</b>	Roswell Park Memorial Institute-1640 medium
<b>RUNX</b>	Runt-Related Transcription Factor
<b>SA-PerCP</b>	Streptavidin-Peridinin Chlorophyll Protein
<b>SB</b>	Staining Buffer
<b>SCF</b>	Stem Cell Factor
<b>SCL</b>	Stem-Cell Leukaemia Factor
<b>SCX</b>	Strong Cation Exchange
<b>shRNA</b>	Short Hairpin RNA
<b>SMRT</b>	Silencing Mediator of Retinoic Acid and Thyroid Hormone Receptors
<b>SSC</b>	Side Scatter
<b>ST-HSC</b>	Short-Term Haematopoietic Stem Cell
<b>STAT3</b>	Signal Transducer and Activator of Transcription 3
<b>SWATH-MS</b>	Sequential Window Acquisition of all Theoretical Mass Spectra
<b>t</b>	Translocation
<b>TAD</b>	Transactivation Domain
<b>TBS</b>	Tris-Buffered Saline
<b>TBS-T</b>	Tris-Buffered Saline with Tween-20
<b>TCGA</b>	The Cancer Genome Atlas
<b>TET2</b>	Tet Methylcytosine Dioxygenase 2
<b>TF</b>	Transcription Factor
<b>TGF-<math>\beta</math></b>	Transforming Growth Factor Beta
<b>TKD</b>	Tyrosine Kinase Domain Mutations
<b>TP53</b>	Tumour Protein p53
<b>TPO</b>	Thrombopoietin
<b>TRCN</b>	RNAi Consortium Number
<b>TSP1</b>	Thrombospondin-1
<b>UC</b>	Universal Container
<b>WBC</b>	White Blood Count
<b>WHO</b>	World Health Organization
<b>WT1</b>	Wilms Tumour 1
<b>ZNF</b>	Zinc Finger Protein

# **Chapter 1**

---

## **Introduction**

## 1.1 Haematopoiesis

### 1.1.1 Overview

Haematopoiesis is the process through which cellular constituents of the blood are continually replenished throughout an individual's lifetime. It is composed of various highly specialised cells with numerous functions, including oxygen transport and immune defence (Jagannathan-Bogdan and Zon, 2013). This process is initiated by a common precursor, a haematopoietic stem cell (HSC), as found in the early 1900s. These cells are characterised by their ability to self-renew or give rise to different progenitor cells that can proliferate and differentiate into mature cells (Laurenti and Göttgens, 2018). Early studies showed that bone marrow (BM) failure in radiation exposed recipients could be rescued by injecting spleen or BM cells from non-exposed donors (Lorenz *et al.*, 1951). Research on this topic intensified when Till and McCulloch demonstrated that the regenerative potential of HSCs could be assayed by performing *in vivo* repopulation assays, thus supporting the existence of multipotential HSCs (Till and McCulloch, 1961). These studies contributed to an understanding of a development hierarchy, in which multipotent HSCs are found at the top, whilst terminally differentiated cells sit on the bottom (1.1.3).

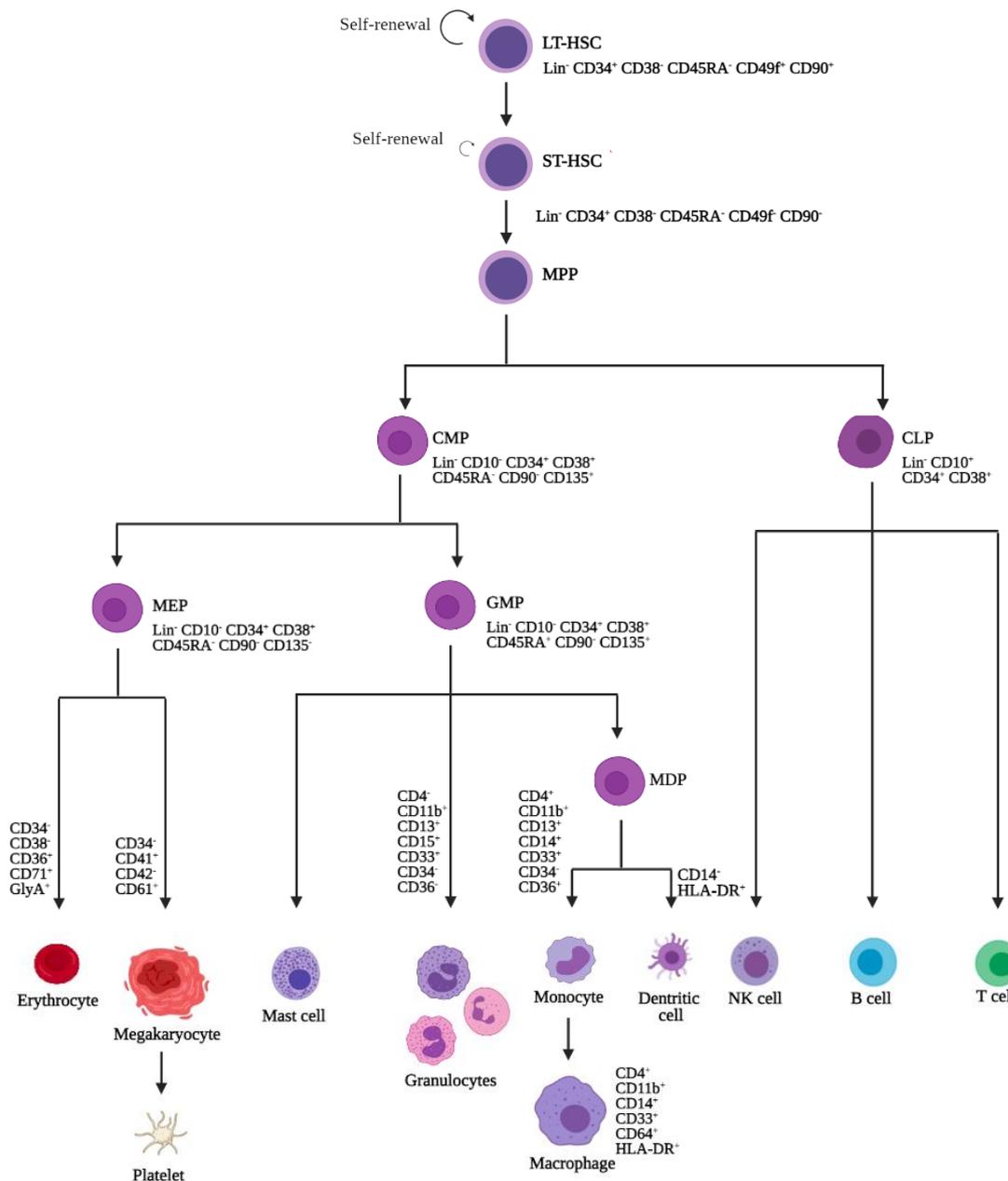
Throughout embryonic development, haematopoiesis occurs in successive waves, each one of them temporarily and spatially restricted, resulting in the development of haematopoietic progenitor cells. This process can be further divided into two main stages. In early embryonic development, blood cells are produced in the yolk sac, in a process termed 'primitive' haematopoiesis. During this time, there is an increased production of red blood cells, to promote tissue oxygenation as the embryo undergoes rapid growth. As well as primitive nucleated erythrocytes, there is also a low frequency of primitive macrophages and megakaryocytes (Fukuda, 1973; Lockett, 1978). The next stage of blood cell development is termed 'definitive' haematopoiesis, with the development of erythro-myeloid progenitors (EMP), which will give rise to definitive erythrocytes and most myeloid lineages, as well as early B and T progenitors. This process occurs initially in the aorta-gonad-mesonephros (AGM), later occurring in the placenta, foetal liver, spleen and BM (Ivanovs *et al.*, 2017). Postnatally, haematopoiesis occurs primarily in the BM; however, in times of haematopoietic stress or injury, this process can occur in the liver or spleen (Butler *et al.*, 2010).

### 1.1.2 Haematopoietic stem cells

HSCs possess two main characteristics that make them unique in the haematopoietic system. Firstly, these cells are multipotential and are therefore able to differentiate into all functional blood cells. Secondly, they have the ability to self-renew, giving rise to identical daughter cells without differentiating. Loss and/or gain of function studies in mice have identified basic developmental principles that control the emergence of haemogenic tissues during ontogeny, haematopoiesis in the adult. However, due to differences observed in basic biology and haematology, as well as therapy development, there was a constant need to complement these mouse studies with human primary cells. Subsequently, based on Till and McCulloch experiments, several groups focused on investigating human haematopoiesis using human colony-forming progenitors scored in *in vitro* colony-forming unit (CFU) assays (Moore *et al.*, 1973; Pike and Robinson, 1970). Subsequently, the ability to engraft human haematopoietic cells into immune-deficient mice, allowed the development of several humanised mouse models to study HSC development *in vivo* (Bosma *et al.*, 1983; McCune *et al.*, 1988; Shultz *et al.*, 2005; Shultz *et al.*, 1995; Rongvaux *et al.*, 2011).

A major obstacle in studying HSCs relies in the rarity of these cells. In general, only 1 in  $10^6$  cells in the human BM is defined as a transplantable HSC (Wang *et al.*, 1997) based on the simultaneous detection of several independent cell surface markers. The first marker to be identified in these cells was CD34, expressed in less than 5% of all blood cells, and in > 99% of human HSC (Civin *et al.*, 1984). However, as CD34 is expressed both in HSCs, and in progenitor cells, there was a need to search for additional markers of stemness. Subsequent studies identified CD90 as a stem cell marker (Baum *et al.*, 1992), whilst CD45RA and CD38 were identified as markers of progenitor cells (Bhatia *et al.*, 1997; Conneally *et al.*, 1997; Lansdorp *et al.*, 1990). Furthermore, human HSC were characterised by the absence of lineage markers (Lin<sup>-</sup>), resulting in their classification as Lin<sup>-</sup>CD34<sup>+</sup>CD38<sup>-</sup>CD90<sup>+</sup>CD45RA<sup>-</sup> (**Figure 1.1**).

At the cellular level, the inhibition of self-renewal occurs as lineage programming is initiated. Hence, it would be anticipated that this would also happen at the molecular level, resulting in the concept of multilineage priming, proposed as a mechanism through which HSCs are able to maintain their multipotency potential (Miyamoto *et al.*, 2002).



**Figure 1.1 – Hierarchical representation of human haematopoietic development**

HSC possess an increased self-renewal capacity (curved arrow), gradually giving rise to mature progeny with reduced capacity for self-renewal and increased differentiation potential. The phenotypic cell surface marker of each population is shown. Adapted from (Weiskopf *et al.*, 2016; Guilliams *et al.*, 2018).

**HSC** – Haematopoietic Stem Cell; **MPP** – Multipotent Progenitor Cell; **CMP** – Common Myeloid Progenitor Cell; **MEP** – Megakaryocytic-Erythroid Progenitor Cell; **GMP** – Granulocyte-Monocyte Progenitor Cell; **MDP** – Monocyte-Dendritic cell Progenitor Cell; **CLP** – Common Lymphoid Progenitor Cell; **NK cell** – Natural Killer Cell; **GlyA** – Glycophorin A (CD235a); **CD** – Cluster of Differentiation; **Lin** – Lineage markers.

However, the HSC transcriptional programme is characterised by several unique metabolic and cellular properties, not intuitively linked with multipotency (Laurenti and Göttgens, 2018). In fact, approximately 70% of all transcriptional changes observed between HSCs and early progenitor cells occur independently of lineage fate (Laurenti *et al.*, 2013). HSCs typically reside in a quiescent (Wilson *et al.*, 2008; Foudi *et al.*, 2009), autophagy-dependent (Warr *et al.*, 2013; Ho *et al.*, 2017) and glycolytic state (Simsek *et al.*, 2010; Takubo *et al.*, 2013), with low mitochondrial activity (Vannini *et al.*, 2016; Ito *et al.*, 2016) and tight regulation of protein synthesis (Signer *et al.*, 2014). Progenitor cells, on the other hand, are highly proliferative and metabolically active cells (Laurenti and Göttgens, 2018). However, it is important to acknowledge that these characteristics are not absolute. Whilst the initiation of transcriptional programs specific to lineage determination may occur regardless of loss of stem cell characteristics, several regulators can be involved in both, like runt-related transcription factor 1 (*RUNX1*, aka *AML1*) (Cai *et al.*, 2015). These observations suggest that the processes through which lineage determination coordinates with changes in the cellular state remains to be fully understood.

#### 1.1.2.1 Properties of haematopoietic stem cells

It is currently recognised that the HSC population is heterogeneous, comprising, at least, two subsets, differing in their repopulation abilities and cycling properties (Foudi *et al.*, 2009; Qiu *et al.*, 2014; Takizawa *et al.*, 2011). These are long-term HSCs (LT-HSC), capable of long-term engraftment, and short-term HSCs (ST-HSC), derived from LT-HSCs. Whilst LT-HSCs have a lifelong self-renewing potential (months-years), the ST-HSCs have limited ability (days-weeks), as they firstly give rise to multipotent progenitors (MPPs), further branching into common-myeloid progenitors (CMPs) and common-lymphoid progenitors (CLPs), with the aim of replenishing the haematopoietic system (Morrison and Weissman, 1994; Guenechea *et al.*, 2001; Christensen and Weissman, 2001). Furthermore, phenotypically, LT-HSC are defined as  $\text{Lin}^- \text{CD34}^+ \text{CD45RA}^- \text{CD49f}^+ \text{CD90}^+ \text{CD38}^-$ , lacking CD38 or any lineage-restricted antigen. Upon loss of CD49f and CD90 expression, these give rise to transiently engrafting multipotent progenitors, the ST-HSCs, characterised as  $\text{Lin}^- \text{CD34}^+ \text{CD45RA}^- \text{CD49f}^- \text{CD90}^- \text{CD38}^-$  cells (Notta *et al.*, 2011).

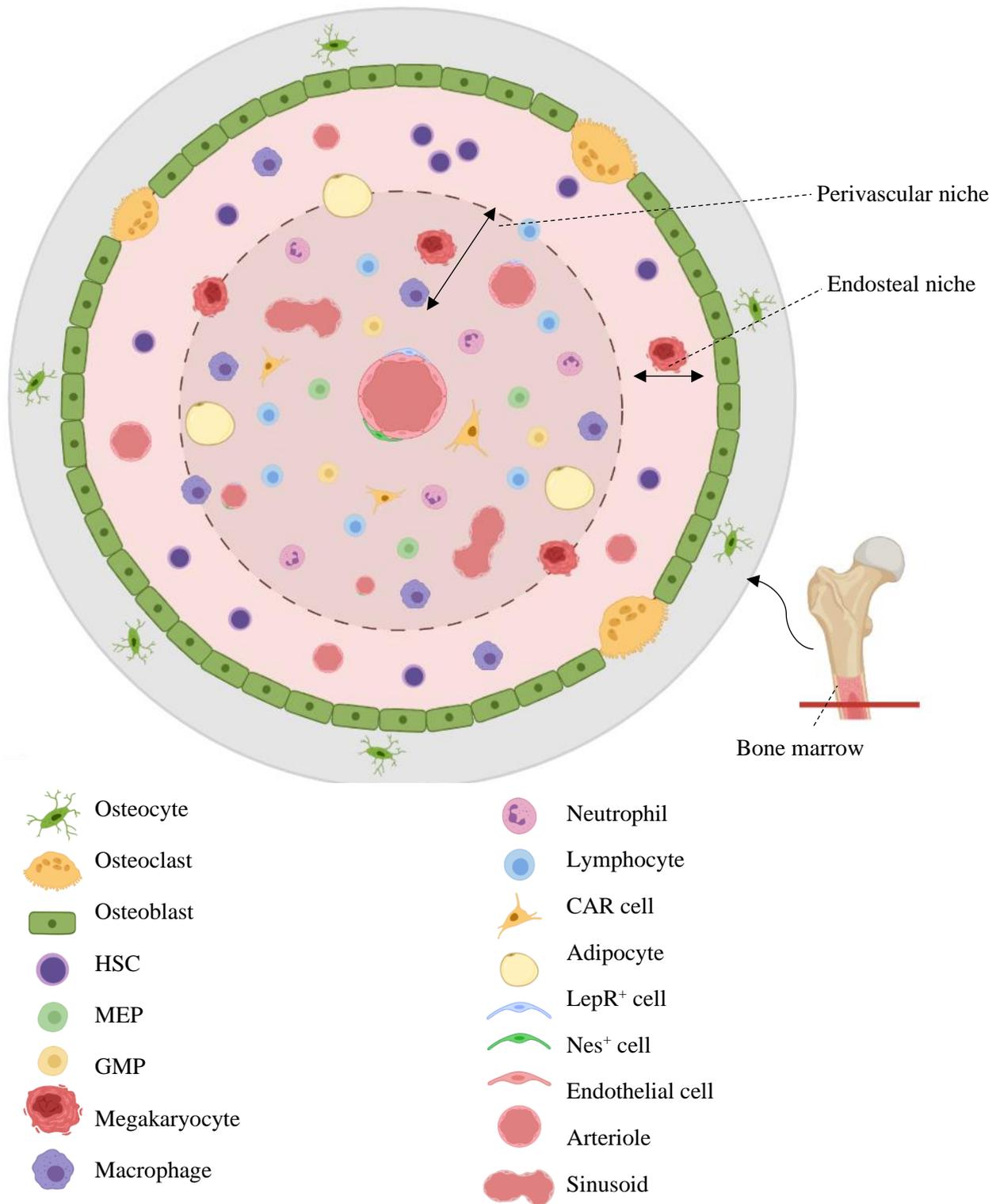
#### 1.1.2.2 Haematopoietic stem cell niche

The concept of an HSC niche was first suggested by Schofield in 1978, who proposed that a physical niche of cells resides within the BM (Schofield, 1978). HSCs largely rely on their

microenvironment, made up of a complex network of cells and secreted factors, for the regulation of quiescence, proliferation, self-renewal and differentiation (Morrison and Scadden, 2014). Normally, most of the HSCs are in a quiescent state, but can become active as a response to infectious stress, such as interferon-mediated signalling, and increase their proliferative rate or promote differentiation (Essers *et al.*, 2009; Wilson *et al.*, 2008; Baldrige *et al.*, 2010). Currently, it is considered that the BM microenvironment is composed of two niches, able to maintain and regulate HSCs: the endosteal and the vascular niches (**Figure 1.2**). In the outer BM, HSCs reside closely to the endosteal bone surface, as several studies showed that HSCs isolated from this region presented a higher degree of proliferation and an increased long-term hematopoietic reconstitution potential (Haylock *et al.*, 2007; Grassinger *et al.*, 2010). More differentiated cells, on the other hand, are generally found in the central BM region, within the perivascular niche.

The endosteum comprises the region between bone and BM, and it's lined with a heterogeneous group of osteoblastic cells at various stages of differentiation, from which a fraction of are fully mature osteoblasts able to synthesise bone. Moreover, osteoclasts, bone-absorbing cells, also line the endosteum and together regulate bone formation. Osteoblasts synthesize several cytokines that are suggested to contribute to the maintenance and regulation of HSCs, such as such as CXC-chemokine ligand 12 (CXCL12 aka stromal-derived factor 1 [SDF-1]), stem-cell factor (SCF), osteopontin (OPN), granulocyte colony-stimulating factor (G-CSF), annexin 2 (ANXA2), angiopoietin 1 (ANG1), and thrombopoietin (TPO) (Taichman and Emerson, 1994; Ponomaryov *et al.*, 2000; Calvi *et al.*, 2003; Arai *et al.*, 2004; Stier *et al.*, 2005; Jung *et al.*, 2007; Yoshihara *et al.*, 2007).

CXCL12 is mainly produced by immature osteoblasts and endothelial cells, and controls HSC homing, retention, and repopulation (Ponomaryov *et al.*, 2000). Even though SCF is mainly produced by perivascular cells (Ding *et al.*, 2012), this cytokine is also produced by osteoblasts, and plays an important role in HSC maintenance. Equally, OPN, secreted by osteoblasts and other cells, plays a critical role in the retention, migration, and control of HSC proliferation and differentiation in the endosteal surface (Nilsson *et al.*, 2005; Stier *et al.*, 2005; Grassinger *et al.*, 2009). G-CSF is necessary to support myelopoiesis (Lord *et al.*, 1975; Morad *et al.*, 2008), whilst ANXA2 regulates HSC homing and engraftment (Jung *et al.*, 2007).



**Figure 1.2 – HSC regulation by the BM microenvironment**

HSC reside in a specialised environment within the BM, divided into two main niches: endosteal and perivascular. Adapted from (Galán-Díez *et al.*, 2018).

**HSC** – Haematopoietic Stem Cell; **MEP** – Megakaryocytic-Erythroid Progenitor Cell; **GMP** – Granulocyte-Monocyte Progenitor Cell; **CAR cell** - CXCL12-abundant reticular cell.

In addition to secreting cytokines, osteoblastic cells have also been shown to express several membrane-bound ligands and adhesion receptors which contribute to HSC maintenance, including Jagged, a ligand for Notch receptors expressed in HSCs, and N-cadherin. However, both of these factors have proven to be controversial, with studies conflicting in regard to their specific role in HSC maintenance (Varnum-Finney *et al.*, 2000; Butler *et al.*, 2010; Maillard *et al.*, 2008; Mancini *et al.*, 2005; Li and Zon, 2010; Hosokawa *et al.*, 2010). Several other factors are thought to play a role in HSC maintenance, including bone-degrading osteoclasts (Batard *et al.*, 2000; Bhatia *et al.*, 1999; Silver *et al.*, 1988) and macrophages (Chang *et al.*, 2008; Winkler *et al.*, 2010).

In addition to the endosteal region, several studies have suggested that vascular environments are also involved in the maintenance of HSC, comprising the perivascular niche, essential for gas exchange, delivery of nutrients to cells and waste removal in the BM (Kiel *et al.*, 2005) (**Figure 1.2**). The perivascular niche is highly heterogenous and contains distinct cell types, including endothelial cells, blood vessels that deliver both oxygen and nutrients, and can be found in the lining of sinusoids, specialised blood vessels that form an extensive network throughout the BM (Kiel *et al.*, 2005). These cells regulate HSC fate both by direct cell-to-cell contact, as well as by the secretion of angiocrine factors (Butler *et al.*, 2010; Kobayashi *et al.*, 2010). These include CXCL12, vascular endothelial growth factor A (VEGF-A), fibroblast growth factor 2 (FGF2), ANG1, thrombospondin-1 (TSP1), and Notch ligands, which have all been described as playing important roles in maintaining the stem-cell pool and regulate HSC self-renewal capacity (Butler *et al.*, 2010; Kobayashi *et al.*, 2010; Poulos *et al.*, 2013; Rafii *et al.*, 2016). Other cell constituents present in the perivascular niche include Leptin Receptor Perivascular-Expressing Cells (LepR<sup>+</sup>), a major source of several growth factors, including CXCL12 and SCF (Ding *et al.*, 2012; Ding and Morrison, 2013). Moreover, these cells are responsible for the differentiation of mesenchymal stem cells (MSCs) into adipocytes, osteoblasts, and chondrocytes (Zhou *et al.*, 2014). Equally, Nestin-positive (Nes<sup>+</sup>) cells produce soluble factors involved in HSC maintenance, including CXCL12 and SCF, suggesting that MSCs are directly involved with HSCs (Méndez-Ferrer *et al.*, 2010). CXCL12 abundant reticular (CAR) cells have been shown to be important in the homing and localisation of HSC within the BM. As the name suggests, these cells secrete extremely high levels of CXCL12, as well as SCF. Other molecules found in the perivascular niche include megakaryocytes and platelets (Banu and Williams, 1995; Alexander *et al.*, 1996).

### 1.1.3 Hierarchical organisation of haematopoiesis

In order to define the relationship between HSCs and their progenies, a model for the haematopoietic hierarchy was initially proposed as a tree-like roadmap detailing the differentiation process starting from an HSC, into cells with declining multilineage potential leading to unilineage commitment (Kondo *et al.*, 1997; Akashi *et al.*, 2000; Adolfsson *et al.*, 2005; Wilson *et al.*, 2008) (**Figure 1.1**). Such distinction was made possible due to the use of phenotypical analysis of the different lineages, based on flow cytometry-based cell sorting (Rieger and Schroeder, 2012). LT-HSC sit at the apex of the hierarchical tree and, upon certain stimuli, can enter cell cycle (Schoedel *et al.*, 2016), and progress into ST-HSC, which in turn differentiate into MPP. These cells present a higher frequency of cell-cycle progression and differentiation activity, but are incapable of self-renewing (Morrison *et al.*, 1997). MPPs, in turn, give rise to CLPs, that only possess lymphoid-restricted differentiation abilities (Kondo *et al.*, 1997), and CMPs, that can differentiate into megakaryocytic/erythroid progenitors (MEPs) and granulocyte/macrophage progenitors (GMPs) (Akashi *et al.*, 2000; Na Nakorn *et al.*, 2002). GMPs ultimately give rise to monocytes, macrophages and granulocytes, including basophils, neutrophils and eosinophils, whilst MEP cells differentiate into erythrocytes, megakaryocytes and platelets (Rieger and Schroeder, 2012). Cells derived from the lymphoid lineage are responsible for mediating immune response, whilst cells from the myeloid lineage are involved in several physiological processes, including blood clotting, oxygen transport and innate immunity.

Additional studies have shown that this classical model of haematopoiesis is more complex than initially thought. For instance, intermediate-term HSCs (IT-HSC) represent a transitory population between LT-HSC and ST-HSC (Yamamoto *et al.*, 2013; Benveniste *et al.*, 2010; Zhang *et al.*, 2018). Furthermore, the MPP population can be further characterised into MPP1, MPP2, MPP3 and MPP4 (Wilson *et al.*, 2008; Pietras *et al.*, 2015). In the revised model, MPP1 resembles ST-HSC, whilst MPP2-4 are parallel subpopulations at a downstream level, responsible for sustaining blood production at steady-state (Pietras *et al.*, 2015). In addition, another subpopulation has been identified, expressing high levels of Fms-like tyrosine kinase 3 (FLT3), defined as lymphoid-primed multipotent progenitors (LMPPs). Functionally, these generate cells within the lymphoid lineages (Adolfsson *et al.*, 2001; Adolfsson *et al.*, 2005;

Forsberg *et al.*, 2006; Boyer *et al.*, 2011), and phenotypically resemble MPP4 (Pietras *et al.*, 2015).

The traditional hierarchical model suggests that HSC differentiation is a stepwise process, however, several studies have shown that megakaryocyte differentiation is able to bypass most of these steps, and its lineage outcome is mainly derived from HSCs (Woolthuis and Park, 2016). In mice, HSCs were found to highly express von Willebrand factor (*Vwf*), a megakaryocyte-specific gene (Månsson *et al.*, 2007). These cells were shown to possess robust short- and long-term reconstitution ability for megakaryocytes post-transplantation, and their maintenance was modulated by TPO, essential for megakaryocyte development and platelet production (Sanjuan-Pla *et al.*, 2013; de Sauvage *et al.*, 1996). Subsequent studies equally showed that megakaryocytes are predominately and directly derived from HSCs, further contributing to the hypothesis that these cells can bypass the differentiation stages between MPPs and MEPs (Notta *et al.*, 2016).

More recent studies using single-cell technology, combined with computational analysis, have shown that haematopoiesis is a continuous process, as opposed to a stepwise progress to lineage commitment (Velten *et al.*, 2017). Moreover, unilineage-restricted cells have been suggested to originate directly from a ‘continuum of low-primed undifferentiated HSPCs’ (CLOUD-HSPC), without any major transition through the multi- and bi-potent stages (Velten *et al.*, 2017). However, even though these new techniques have further allowed the understating of normal human haematopoiesis, it has led to the generation of conflicting observations. Instead, some authors defend the existence of a structured hierarchy with a heterogenous haematopoietic landscape (Buenrostro *et al.*, 2018; Karamitros *et al.*, 2018; Pellin *et al.*, 2019).

Altogether, with the use of high-throughput and single-cell methodologies, such as RNA sequencing, mass cytometry and immunophenotypic analysis, it was possible to not only identify new populations, but also enlighten the specific processes cell undergo, until reaching a terminal differentiated state, suggesting a need to revise the classical hematopoietic hierarchy roadmap.

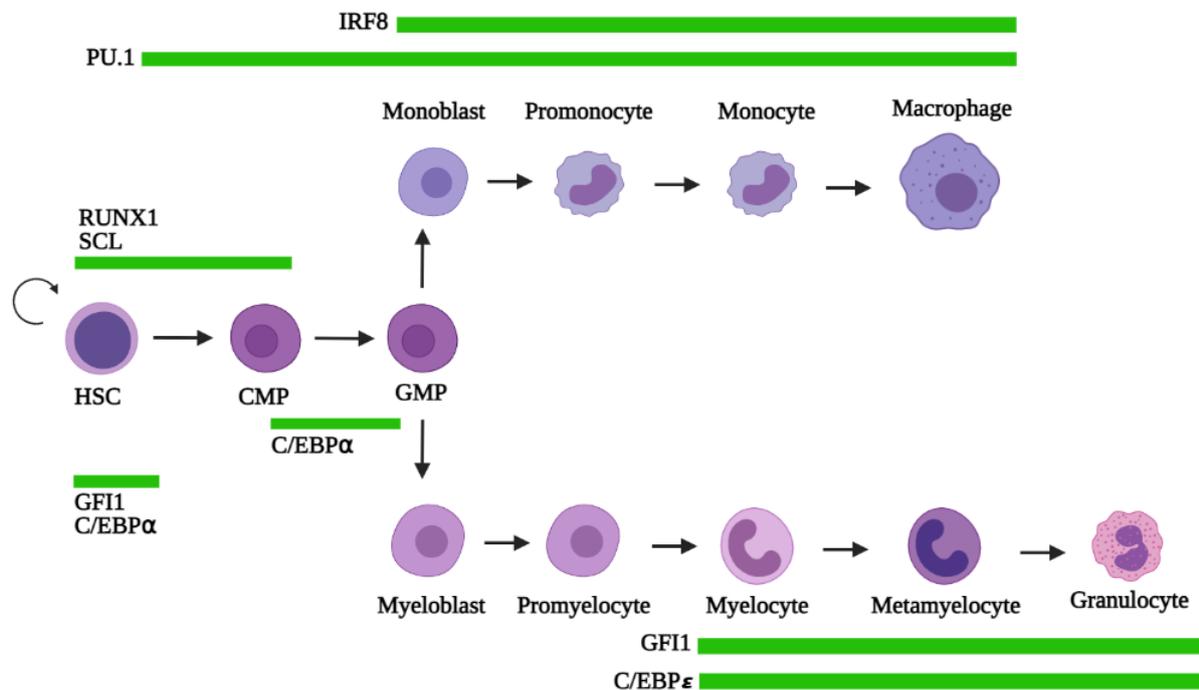
## 1.1.4 Regulation of normal human haematopoiesis

### 1.1.4.1 Transcription factor regulation of haematopoietic development

Transcription factors (TFs) play an essential role during haematopoiesis, being responsible for several processes, such as stem cell maintenance, lineage commitment and differentiation. These are often cell-type restricted, which means they can drive the expression of characteristic lineage-specific target genes, leading to the development of a specific cell subtype (**Figure 1.3**). The differentiation of haematopoietic precursors is associated with two main processes: the reduction on the cells' self-renewal potential and the stepwise acquisition of specific lineage identity. However, the interpretation of the roles of these TFs was framed on the classical model, and the current understanding shows this to be oversimplistic (**1.1.3**).

Myeloid cell differentiation is orchestrated by a small number of TFs, which include *PU.1* (aka *SPI1*) (Klemsz *et al.*, 1990), CCAAT/enhancer binding proteins (including *CEBPA*, *CEBPB* and *CEBPE*) (Zhang *et al.*, 1997; Tanaka *et al.*, 1995a; Yamanaka *et al.*, 1997), growth-factor independent 1 (*GF11*) (Hock *et al.*, 2003) and interferon-regulatory factor 8 (*IRF8*) (Holtschke *et al.*, 1996). These also include, at the stem-cell level, *RUNX1* (Okuda *et al.*, 1996) and stem-cell leukaemia factor (*SCL* or *TALI*) (Shivdasani *et al.*, 1995). These TFs are often classified as master regulators of haematopoiesis due to their ability to promote the expression of several myeloid genes, such as those encoding receptors for macrophage-CSF (M-CSF), G-CSF or granulocyte/macrophage-CSF (GM-CSF).

The first TF to be expressed are those that orchestrate the formation of the HSC pool, which include *SCL* and *RUNX1*. Mice lacking either one of these factors are embryonically lethal and have no detectable haematopoiesis, indicating that these are essential in the generation of foetal HSCs (Goode *et al.*, 2016; Suter *et al.*, 2011). In adult haematopoietic cells, conditional deletion of these genes showed that they are not required for the maintenance of HSCs in the BM; instead, the effects observed are more lineage-specific. Regarding *RUNX1*, its conditional ablation inhibited CLP production, blocked B- and T- cell maturation and decreased the formation of platelets (Ichikawa *et al.*, 2004). A subsequent study showed that this ablation contributed to the development of mild myeloproliferative syndrome (Growney *et al.*, 2005).



**Figure 1.3 – Transcriptional regulation of myeloid development**

Developmental stages and transcriptional regulation of monocytic and granulocytic differentiation. RUNX1 and stem-cell leukaemia factor (SCL) are required for the generation HSC, whilst growth-factor independent 1 (GFI1) and CCAAT/enhancer binding protein- $\alpha$  (C/EBP $\alpha$ ) regulate HSC self-renewal potential. Moreover, C/EBP $\alpha$  is essential for the differentiation of CMP into GMPs. For granulocytic development, both GFI1 and C/EBP $\epsilon$ , are crucial mediators. Monocytic differentiation relies on the expression of PU.1 and IRF8. Bars represent the controlled expression of different TF during myeloid cell development. Adapted from (Rosenbauer and Tenen, 2007).

Whereas expression of PU.1 is essential for the differentiation of HSCs into CMPs, the next step of development into GMPs requires the expression of C/EBP $\alpha$ . C/EBP $\alpha$  is a basic-region leucine zipper TF expressed by HSCs, myeloid progenitors and granulocytes, but absent in macrophages (Akashi *et al.*, 2000; Radomska *et al.*, 1998). In mice, C/EBP $\alpha$  deficiency led to the absence of GMPs and granulocytic-committed cells, but normal numbers of CMPs (Zhang *et al.*, 1997; Zhang *et al.*, 2004b). Interestingly, C/EBP $\alpha$  is not required for granulocytic differentiation following GMP commitment (Zhang *et al.*, 2004b). In addition to having a role in myeloid differentiation, C/EBP $\alpha$  is also responsible for controlling the cells' self-renewal potential, as ablation of C/EBP $\alpha$  in HSCs results in an increased repopulation activity in mouse transplant models (Zhang *et al.*, 2004b). Additionally, this TF has also been shown to be a strong promoter of cell-growth arrest by regulating cell cycle exit (Johnson, 2005).

Following development into GMPs, cells undergo further differentiation into macrophages or granulocytes, following the expression of both PU.1 and IRF8. IRF8 is expressed in several haematopoietic lineages, including HSCs, B cells, dendritic cells and in resting T cells (Driggers *et al.*, 1990). Furthermore, expression of IRF8 is higher in macrophages, but not in granulocytes (Tamura *et al.*, 2000). In mice, loss of IRF8 led to the development of myeloproliferative syndrome resembling CML (Holtschke *et al.*, 1996), characterised by an increased number of granulocytic-derived cells at the expense of macrophages (Scheller *et al.*, 1999).

Granulocytic differentiation requires the expression of two additional TF, *GFI1* and *C/EBP $\epsilon$* . *GFI1* is a transcriptional repressor detected in HSCs, neutrophils and B- and T-precursor cells (Hock *et al.*, 2003; Karsunky *et al.*, 2002a; Hock and Orkin, 2006). Mice deficient for *GFI1B* were seen to lack lymphocytic progenitor differentiation, as well as neutrophilic granulocytes (Hock *et al.*, 2003; Karsunky *et al.*, 2002b), even though the development of early myeloid progenitors remained unchanged. Similarly, ablation of *C/EBP $\epsilon$*  led to abnormal granulocytic development beyond the promyelocyte stage (Yamanaka *et al.*, 1997). In addition to having a role in granulocytic differentiation, deletion of *GFI1* also impacts normal HSC development. Even though, phenotypically, *GFI1*-deficient cells resembled normal HSCs, these were functionally impaired, as they displayed reduced self-renewal capacity due to cell cycle changes (Hock *et al.*, 2004).

Taken together, haematopoietic development is a complex process, relying on the hierarchical expression of key TFs, responsible for promoting HSC differentiation into mature

myeloid cells. The dysregulation of these TFs has been previously shown to play an important role in the leukaemogenic process, and often present themselves as therapeutic targets (Tenen, 2003). Several studies have shown that absence of a lineage specific TF results in a perturbed differentiation process or complete absence of the respective lineage. For instance, knocking-out of GATA-1, an erythroid TF, results in a complete absence of erythroid and megakaryocytic differentiation in mice, whilst other lineages remained unperturbed (Pevny *et al.*, 1991). Moreover, whilst removal of RUNX1 or TAL1 influences HSC formation and the development of the entire haematopoietic system, absence of PU.1 largely affects the myeloid and B-cell lineages.

#### 1.1.4.2 Cytokines and growth factors

The production of haematopoietic-derived cells is under the control of several haematopoietic cytokines (Kondo *et al.*, 2000; Mossadegh-Keller *et al.*, 2013; Brown *et al.*, 2018). Cytokines are extracellular ligands that can stimulate biological responses through the binding and activation of cytokine receptors. Important cytokines within the haematopoietic system include interleukins (ILs), CSFs, interferons, erythropoietin (EPO) and TPO. Similarly to haematopoietic development, cytokines have been arranged in a hierarchical system, as sub-populations of HSC express certain cytokine receptors associated with specific cell lineages (**Figure 1.1**). Moreover, certain cell types require the action of multiple cytokines simultaneously, observed in HSCs and megakaryocyte progenitors. For instance, whilst IL-3 stimulates the growth of most lineages, EPO exclusively regulates erythroid development (Ihle, 1992).

EPO supports erythroid progenitor cells by promoting the expression of the erythroid-affiliated TF GATA-1 (Koury and Bondurant, 1990; Grover *et al.*, 2014). Ablation of EPO and its respective receptor results in *in vivo* death mid-gestation due to severe anaemia, even though these mice displayed erythroid progenitor cells within the BM (Wu *et al.*, 1995; Kieran *et al.*, 1996; Lin *et al.*, 1996). Moreover, these subjects showed signs of early erythropoiesis, suggesting that another cytokine, possibly TPO, is responsible for supporting early erythropoiesis (Kieran *et al.*, 1996). Indeed, absence of TPO or its receptor resulted not only in loss of normal number of platelets and megakaryocytes, but also reduced numbers of HSCs, associated with a significant reduction in stem cell expansion following BM transplantation,

suggesting a non-redundant role in stem cell activity (Gurney *et al.*, 1994; Alexander *et al.*, 1996; Carver-Moore *et al.*, 1996; Solar *et al.*, 1998; Fox *et al.*, 2002).

In high concentrations, FLT3 has been shown to drive HSC development towards a myeloid-lymphoid fate, whilst suppressing megakaryocyte and erythroid development, by promoting to the upregulation of PU.1 (Tsapogas *et al.*, 2014; Onai *et al.*, 2006). Targeted disruption of FLT3 resulted in a significant reduction in the BM haematopoietic progenitor pool size and subsequent decrease in the number of mature myeloid, B-, natural killer (NK) and dendritic cells (DC) (Mackarehtschian *et al.*, 1995).

SCF is another cytokine that has been shown to be essential for HSC survival, proliferation and differentiation (Broudy, 1997; Hartman *et al.*, 2001). Loss of SCF or its receptor c-kit results in severe macrocytic anaemia (Ogawa *et al.*, 1991). Moreover, following the administration of a neutralising antibody against c-kit, adult mice show complete depletion of progenitor cells and, eventually, absence of all mature myeloid and erythroid cells within the BM (Ogawa *et al.*, 1991).

By growing daughter cells of granulocyte-macrophage colony forming cells in the presence of either G- or GM-CSF, Metcalf and Burgess observed the differentiation of these cells into granulocytes and macrophages, respectively (Metcalf and Burgess, 1982). Unexpectedly, null mutations of GM-CSF or its corresponding receptor had no consequence in the normal numbers of myeloid progenitor cells (Stanley *et al.*, 1994; Nishinakamura *et al.*, 1995; Robb *et al.*, 1995). On the other hand, mice with a null mutation in both G-CSF and its receptor were shown to present incomplete granulopoiesis, characterised by chronic neutropenia, a decrease in mature myeloid cells in the BM and decreased neutrophil release from the BM (Lieschke *et al.*, 1994; Liu *et al.*, 1996; Semerad *et al.*, 2002). However, mice with ablated G-CSF receptor only presented a mild reduction in the proportion of committed myeloid progenitors. These observations promoted the discussion that cytokines have the ability to act compensatively upon the complete absence of a certain molecule, as deletion of a single cytokine or receptor never results in the complete absence of a specific haematopoietic lineage (Brown *et al.*, 2018).

## 1.2 Acute Myeloid Leukaemia

### 1.2.1 Overview

Acute myeloid leukaemia (AML) is a malignant blood cancer characterised by the clonal expansion of abnormal myeloid cells, firstly within the BM, with subsequent exfiltration of these cells through the bloodstream into other organs, such as spleen and liver. As the name suggests, this is an extremely fast-growing disease, with symptoms appearing following a few weeks of disease development. These can include fatigue, fever, recurrent infections, persisting bruising and bleeding, amongst others. AML is the most common acute leukaemia in adults (Khwaja *et al.*, 2016); though rare in comparison with the most prevalent cancer types. For instance, whilst approximately 40,000 people are diagnosed every year in the UK with colorectal cancer, only about 3,000 are diagnosed with AML (Cancer Research UK, 2021). AML is diagnosed based on the accumulation of myeloid blasts within the BM, which are subjected to immunophenotyping, cytogenetic and molecular characterisation to distinguish between AML subtypes (1.2.3). Currently, the presence of abnormal molecular and cytogenetic features at diagnosis are considered the most important prognostic factors and are highly predictive of complete remission (CR) rates, disease-free survival, risk of relapse and overall survival (OS) (Döhner *et al.*, 2017). However, the fact that these cells often exhibit several molecular abnormalities, contributes to disease heterogeneity (Ley *et al.*, 2013; Lindsley *et al.*, 2015). Even though significant improvements have been made in the treatment of this disease, prognosis in elder patients, who account for the majority of the new cases, remains poor (Shah *et al.*, 2013a).

### 1.2.2 Pathophysiology of AML

AML develops as a result of genetic and epigenetic changes in myeloid progenitor cells. In most cases, there is no predisposing factor that makes individuals more susceptible to developing AML. Several models have shown that more than one cooperating abnormality is necessary to develop AML (Grisolano *et al.*, 2003; Schessl *et al.*, 2005). Originally, the developmental process of AML was thought of as a '2-hit' model, in which it was suggested that the clonal expansion of malignant cells required the acquisition and cooperation of, at least, two mutations from different protein/gene classes (Conway O'Brien *et al.*, 2014; Gilliland and Griffin, 2002). According to this model, Class I mutations, which result in activated signalling

pathways that regulate cell proliferation and survival, must occur in combination with Class II mutations, involving TFs implicated in the regulation of cell differentiation and self-renewal (Takahashi, 2011; Kihara *et al.*, 2014). Common Class I mutations include *FLT3*, kirsten- or neuroblastoma-rat sarcoma virus gene (*K/NRAS*, respectively), tumour protein p53 (*TP53*) and tyrosine-protein kinase KIT (*c-KIT*). Interestingly, these mutations often occur in sub-clonal cellular fractions, occurring late in pathogenesis of disease, suggesting that these are a result of late clonal events; and whilst they may impact treatment response, the evidence that they initiate disease is scarce (Papaemmanuil *et al.*, 2016). Further studies have also highlighted the role of signal transducer and activator of transcription 3 (*STAT3*) in the stimulation of cellular proliferation and survival (Cook *et al.*, 2014; Ghoshal Gupta *et al.*, 2008; Yamada and Kawauchi, 2013). Enhanced tyrosine phosphorylation of *STAT3* occurs as a result of increased secretion of cytokines, such as IL-6, or due to mutations in receptor tyrosine kinases, seen in up to 50% of AML cases and associated with a worse prognosis (Schuringa *et al.*, 2000; Steensma *et al.*, 2006).

Significant Class II mutations include nucleophosmin-1 (*NPM1*) and *CEBPA*, which are found in ~ 27% and 6% of cases, respectively (Ley *et al.*, 2013) (**Table 1.1**), as well as the expression of the fusion proteins RUNX1-ETO (aka AML1-ETO or RUNX1-RUNX1T1) and PLM-RARA. However, in recent years, this has been shown to be an over-simplification of the leukaemogenic process, with the identification of mutations within epigenetic regulators, classified as Class III mutations (Sun *et al.*, 2018). These have been shown to possess downstream effects on both cellular differentiation and proliferation, and include mutations in DNA-methylation and post-translational histone modification genes, such as DNA methyltransferase 3 alpha (*DMNT3A*), Tet methylcytosine dioxygenase 2 (*TET2*), Wilms tumour 1 (*WT1*), and isocitrate dehydrogenase (NADP(+)) 1 / 2 (*IDH1 / IDH2*) (Ley *et al.*, 2013; Patel *et al.*, 2012).

A low number of coding sequence mutations are found in AML, compared to most solid epithelial tumours (Ley *et al.*, 2013). As a result, mutations associated with the development of AML are well defined, and their presence is important when classifying disease subtype (**1.2.3**), infer prognosis (**1.2.4**) and respective treatment (**1.2.5**).

**Table 1.1 – Recurrent mutations in AML**

Table outlining recurrent molecular and cytogenetic aberrations are frequently observed in AML. Adapted from (Grove and Vassiliou, 2014).

<b>Type of mutation</b>	<b>Gene</b>	<b>Frequency</b>
<b>Signal transduction Genes</b>	<i>FLT3, NRAS, c-KIT, PTPN11</i>	59%
<b>DNA modification Genes</b>	<i>DNMT3A, TET2, IDH1/2</i>	44%
<b>Chromatin Modifiers</b>	MLL-fusions, <i>ASXL1, EZH2</i>	30%
<i>NPM1</i>	-	27%
<b>Fusion Genes</b>	<i>PML-RARA, MYH11-CBFB, RUNX1-ETO</i>	25%
<b>Myeloid transcription factors</b>	<i>CEBPA, RUNX1</i>	22%
<b>Tumour Suppressor Genes</b>	<i>TP53, WT1, PHF6</i>	16%
<b>Spliceosome Genes</b>	<i>SF3B1, SRSF2, U2AF1</i>	14%
<b>Cohesins</b>	<i>SMC1A, SMC3, RAD1, STAG2</i>	13%

Collectively, the most frequent mutations in AML can be found in genes related to signalling pathways, such as *FLT3*, presented as internal tandem duplications (ITD), or tyrosine kinase domain mutations (TKD); *c-KIT* and *RAS*, present in over half of AML patients (**Table 1.1**). Additionally, mutations in epigenetic regulators occur c40% of patients (**Table 1.1**). For instance, *DNMT3A* mutations are observed in 20-25% of AML patients (Ley *et al.*, 2010), and mouse knockout (KO) studies have shown that this gene plays a crucial role in limiting the self-renewal potential of haematopoietic stem and progenitor cells (HSPC) and in regulating myeloid differentiation (Challen *et al.*, 2011). However, despite their prevalence in AML, these mutations alone are not sufficient to promote disease development. Progression to leukaemia often requires the acquisition of driver mutations, such as the expression of the fusion protein PML-RARA, not detected in healthy individuals (Abelson *et al.*, 2018; Welch *et al.*, 2012). Chromosomal aberrations have in fact been widely described in AML, and can include inversions (inv), such as inv(16), responsible for the expression of CBF $\beta$ -MYH11, or reciprocal translocations (t), including t(8;21) and t(15;17), resulting in the formation of the chimeric proteins RUNX1-ETO and PML-RARA, respectively, which can alter the cells differentiation process.

Even though several studies have attempted to understand the leukaemogenic process that leads to the development of AML, much remains to be understood. As suggested by the model described above, the pathogenesis and behaviour of AML relies on the interaction between different somatic mutations and chromosomal rearrangements, as well as epigenetic alterations. For instance, the *c-KIT* mutation has been associated with t(8;21) and inv(16), and its presence/absence significantly influences disease prognosis (De Kouchkovsky and Abdul-Hay, 2016). Similarly, mutations in *NPM1* frequently occur with *FLT3-ITD*, or with mutations in the epigenetic genes *DNMT3A* and *IDH-1* or *IDH-2* (Patel *et al.*, 2012).

### 1.2.2.1 Leukaemic Stem Cells

The existence and involvement of leukaemic stem cells (LSCs) in the development of AML has been discussed for several decades (Fialkow, 1974; Griffin and Löwenberg, 1986; Lapidot *et al.*, 1994). This is based on the hypothesis that AML populations may mimic the hierarchical developmental process observed in normal haematopoiesis; just as normal HSC can give rise to progressively more mature cells, malignant cells can similarly generate a bulk population of AML cell blasts. However, whilst the normal stem cell has the ability to differentiate into the haematopoietic lineage, a mutated stem cell, even though it possesses similar properties to the

normal stem cell, can either divide carrying defects; or remain undifferentiated, accumulating as immature progenitor cells, or blasts (Jordan, 2007). These cells can undergo self-renewal, are multipotent, and highly proliferative. However, the chemotherapeutic agents used to effectively eradicate blast cells, have little, if any effect in the LSC population (Jordan, 2007; Hanekamp *et al.*, 2017).

Phenotypically, LSC are described as having an HSC-like phenotype, expressing the CD34 marker, with the ability to initiate leukaemia when xenografted into immunosuppressed mice (Bonnet and Dick, 1997). This was examined by Bonnet and Dick by transplanting patient samples with an immature (CD34<sup>+</sup>CD38<sup>-</sup>) and more mature phenotypic cells (CD34<sup>+</sup>CD38<sup>+</sup>) into immunodeficient mice (Bonnet and Dick, 1997). Mice transplanted with more immature cells developed symptoms of overt disease, not observed in recipients of the more mature cellular population (Bonnet and Dick, 1997). Moreover, CD34<sup>+</sup>CD38<sup>-</sup> cells were able maintain their leukaemia initiation effect upon serial transplantations (Bonnet and Dick, 1997). These studies demonstrated the importance of LSC in the development and progression of AML. Subsequent studies have corroborated these finding by showing that LSCs can initiate leukaemia when transplanted into non-obese diabetic/severe combined immunodeficiency (NOD/SCID) mice through xenotransplantation assays (Feuring-Buske *et al.*, 2003). Moreover, these cells were shown to possess unlimited self-renewal capacity, in a similar way to normal HSCs, associated with a higher proliferative potential compared to normal haematopoietic cells (Reinisch *et al.*, 2015). Altogether, these findings indicate that LSCs can contribute to disease progression. However, due to their heterogeneity, there is no unique phenotype that defines this specific cellular subtype. For instance, even though the CD34 is a well-established LSC surface marker, c30% of patients have no CD34 expression in leukaemic blasts, possibly due to their origin in CD34<sup>-</sup> haematopoietic progenitors (Ng *et al.*, 2016; Sarry *et al.*, 2011; Quek *et al.*, 2016; Taussig *et al.*, 2010). Several additional cell surface markers have been identified as upregulated in LSCs, including CD44 (Jin *et al.*, 2006), CD123 (Jin *et al.*, 2009), and CD47 (Jaiswal *et al.*, 2009); a more detailed phenotypical analysis of these cells will undoubtedly contribute to the development of new therapies targeted against this specific cell subtype, essential for achieving long-term remission.

### 1.2.3 Classification of AML

Several classification systems have been developed and employed to improve prognosis and treatment of AML, based on aetiology, morphology, immunophenotyping and genetics. In the 1970s, AML was categorized according to the French-American-British (FAB) classification system, mainly based on morphology and immunophenotypic criteria (Neame *et al.*, 1986) (**Table 1.2**). This classified AML into eight major subtypes, ranging from FAB-M0, corresponding to immature AML cells, up to FAB-M7, described as acute megakaryoblastic leukaemia, displaying the most differentiated blasts (Bennett *et al.*, 1976). However, to incorporate further AML characteristics, such as clinical features, morphology, immunophenotyping, cytogenetics and molecular genetics, the World Health Organization (WHO) developed a new classification system in 2008, with new revised versions released in 2016 (Vardiman *et al.*, 2009; Arber *et al.*, 2016). According to this model, AML can be subclassified into six categories (**Table 1.3**). AML with recurrent genetic abnormalities, includes AML with balanced translocations/inversions, as well as AML with gene mutations, accounting for approximately 20-30% of patients. Core binding factor (CBF) AML, which include the t(8;21)(q22;q22.1) and inv(16)(p13.1q22), are present in this category.

In 2017, a new European LeukemiaNet (ELN) classification system for the diagnosis and management of adult patients diagnosed with AML was developed, which stratified patients into three different outcome groups, based on the cytogenetics and mutation status of the *ASXL1*, *CEBPA*, *FLT3*, *NPM1*, *RUNX1*, and *TP53* genes: favourable, intermediate or adverse (Döhner *et al.*, 2017) (**Table 1.4**). For instance, the CBF leukaemias described above have relatively favourable outcomes. Similarly, patients with biallelic *CEBPA* mutations have been classified in the favourable risk group. On the other hand, poorer outcomes are associated with complex chromosomal alterations, including those involving the deletion of chromosome arms 5q, 7q and/or 17p, frequently occurring in combination with *TP53* mutations.

**Table 1.2 – FAB classification of AML**

AML subtypes based on FAB classification. Adapted from (Kabel *et al.*, 2017).

<b>FAB subtype</b>	<b>Description</b>
<b>M0</b>	Undifferentiated acute myeloblast leukaemia
<b>M1</b>	Acute myeloblast leukaemia with minimal maturation
<b>M2</b>	Acute myeloblastic leukaemia with maturation
<b>M3</b>	Acute promyelocytic leukaemia
<b>M4</b>	Acute myelomonocytic leukaemia
<b>M4<sub>EO</sub></b>	Acute myelomonocytic leukaemia with BM eosinophilia
<b>M5</b>	Acute monocytic leukaemia
<b>M6</b>	Acute erythroid leukaemia
<b>M7</b>	Acute megakaryoblastic leukaemia

**Table 1.3 – WHO (2016) classification of AML**

AML subtypes based on the WHO (2016) stratification. Adapted from (Döhner *et al.*, 2017).

**AML and related neoplasms****AML with recurrent genetic abnormalities**

AML with t(8;21) (q22;q22.1); *RUNX1-RUNX1T1*

AML with inv(16) (p13.1q22) or t(16;16) (p13.1;q22); *CBFβ-MYH11*

APL with *PML-RARα*

AML with t(9;11) (p21.3;q23.3); *MLLT3-KMT2A*

AML with t(6;9) (p23;q34.1); *DEK-NUP214*

AML with inv(3) (q21.3q26.2) or t(3;3)(q21.3;q26.2); *GATA2, MECOM*

AML (megakaryoblastic) with t(1;22) (p13.3;q13.3); *RBM15-MKL1*

Provisional entity: AML with *BCR-ABL1*

AML with mutated *NPM1*

AML with biallelic mutations of *CEBPA*

Provisional entity: AML with mutated *RUNX1*

**AML with myelodysplasia-related changes****Therapy-related myeloid neoplasms****AML, not otherwise specified (NOS)**

AML with minimal differentiation

AML without maturation

AML with maturation

Acute monoblastic/monocytic leukaemia

Pure erythroid leukaemia

Acute megakaryoblastic leukaemia

Acute basophilic leukaemia

Acute panmyelosis with myelofibrosis

**Myeloid sarcoma****Myeloid proliferations related to Down syndrome**

Transient abnormal myelopoiesis (TAM)

Myeloid leukaemia associated with Down Syndrome

**APL** – Acute Promyelocytic Leukaemia

Even though cytogenetic analysis remains a critical tool for the diagnosis and stratification of AML patients, around 45% of all diagnosed cases present a normal karyotype (Grimwade *et al.*, 2001). In general, normal karyotype AML is considered intermediate risk type; however, this depends on the degree of heterogeneity observed in these patients, hence why further molecular analysis is necessary for prognostic outlook and patient stratification. *FLT3-ITD* mutations are classified as a high risk AML, associated with increased risk of relapse (Yanada *et al.*, 2005). *NPM1* mutations, on the other hand, are associated with a favourable outcome (Döhner *et al.*, 2005). However, patients diagnosed with a combined *NPM1*, and *FLT3-IDT* mutations correlate with intermediate factor risk.

#### **1.2.4 Epidemiology, diagnosis, and prognosis of AML**

AML is one of the most common types of leukaemia in adults, with a slightly higher incidence in men, as compared to women (56% vs. 44%; (Cancer Research UK, 2021)). Even though AML can occur in any age group, it is predominant in older adults, with an average age at diagnosis of 68 years old (Short *et al.*, 2018). Since the early 1990s, AML incidence has increased by 29% in the UK (Cancer Research UK, 2021), partially due to the increase of therapy-related AML cases (McNerney *et al.*, 2017). This incidence has remained stable in patients between the ages of 0-59 years; however, these have progressively increased in the age groups 60-69, 70-79 and >80 years old, by 17%, 36% and 72%, respectively (Cancer Research UK, 2021).

Furthermore, incidence of AML is higher in Caucasian people, as compared to Hispanics, African and Asian/Pacific Islanders (Kirtane and Lee, 2017). Several environmental risk factors have the ability to predispose individuals to the development of AML (Short *et al.*, 2018), although no genetic factors have been identified to date. The risk of developing AML increases upon exposure to DNA damaging-agents, including benzene, cigarette smoke, ionising radiation and cytotoxic chemotherapy (Khwaja *et al.*, 2016). An increased risk of developing AML can also be observed in first-degree relatives of patients diagnosed with several types of haematological disorders. Moreover, specific inherited disorders carry a particularly high risk of AML development, including Down syndrome, Fanconi anaemia and Bloom syndrome (Seif, 2011).

**Table 1.4 – ELN 2017 risk stratification of AML**

AML classification system based on the ELN stratification, from 2017. Adapted from (Döhner *et al.*, 2017).

<b>Risk Category</b>	<b>Genetic Abnormality</b>
<b>Favourable</b>	t(8;21) (q22;q22.1); <i>RUNX1-RUNX1T1</i> inv(16) (p13.1q22) or t(16;16)(p13.1;q22); <i>CBFB-MYH11</i> Mutated <i>NPM1</i> without <i>FLT3-ITD</i> or with <i>FLT3-ITD</i> <sup>low</sup> (allelic ratio < 0.5) Biallelic mutated <i>C/EBPα</i>
<b>Intermediate</b>	Mutated <i>NPM1</i> and <i>FLT3-ITD</i> <sup>high</sup> (allelic ratio ≥ 0.5) Wild-type <i>NPM1</i> without <i>FLT3-ITD</i> or with <i>FLT3-ITD</i> <sup>low</sup> (allelic ratio < 0.5)* t(9;11) (p21.3;q23.3); <i>MLLT3-KMT2A</i> Cytogenetic abnormalities not classified as favourable or adverse
<b>Adverse</b>	t(6;9) (p23;q34.1); <i>DEK-NUP214</i> t(v;11q23.3); <i>KMT2A</i> rearranged t(9;22) (q34.1;q11.2); <i>BCR-ABL1</i> inv(3) (q21.3q26.2) or t(3;3) (q21.3;q26.2); <i>GATA2, MECOM</i> -5 or del(5q); -7; -17/abn(17p) Complex karyotype (3 or more chromosomal abnormalities), monosomal karyotype Wild-type <i>NPM1</i> and <i>FLT3-ITD</i> <sup>high</sup> (allelic ratio ≥ 0.5) Mutated <i>RUNX1</i> (if not co-occurring with favourable AML subtypes) Mutated <i>ASXL1</i> (if not co-occurring with favourable AML subtypes) Mutated <i>TP53</i> (associated with AML complex and monosomal karyotype)

\* Without adverse-risk lesions

The presence of >20% myeloblasts in the peripheral blood, including myeloblasts, monoblasts or megakaryoblasts, undoubtedly indicates a leukaemia. Immunophenotypic analysis of the patient samples can further facilitate the subclassification into the different AML subtypes (Bennett *et al.*, 1976; Bene *et al.*, 1995). Confirmation of AML arises following the expression of two of the following markers, detected in myeloblasts: MPO, CD13, CD33, CDw65 and CD117. Lymphoid antigens are detectable in ~25% of individuals with AML. The T cell antigen CD7 has been reported in 10–30% of patients (Jha *et al.*, 2013), whilst the B cell antigen CD19 is abnormally expressed in 5-34% of patients (Shorbagy *et al.*, 2016). Cytogenetic analysis and screening for common gene mutations and rearrangements are an essential part of not only the diagnostic process, but also for prognosis and treatment options. In fact, for three cytogenic groups, AML diagnosis can be based on cytogenetic analysis alone, regardless of blood count - t(8;21)(q22;q22), inv(16)(p13.1q22) and t(15;17)(q22;q12). Patients diagnosed with t(15;17) are associated with a good prognosis, along with those with t(8;21) and inv(16), whilst cytogenetically normal patients usually present an intermediate prognosis. Patients with a complex karyotype (three or more chromosomal abnormalities), inv(3) or t(6;9), on the other hand, are associated with an extremely poor prognosis (Döhner *et al.*, 2017).

### 1.2.5 Treatment of AML

Even though conventional treatment for AML has remained similar over the last few decades, more targeted and specific therapies have emerged based on the patient's disease subtype. An initial assessment is made based on the patients' characteristics, which include age, performance status and pre-treatment comorbidities. The treatment strategy used for AML is often divided into induction, consolidation, and maintenance therapy. The main aim of induction therapy is to achieve CR, by using cytotoxic chemotherapeutic agents. In patients suitable for intensive treatment, standard chemotherapy agents include the use of a combination of infused cytarabine for 7 days, followed by 3 days of anthracycline, commonly daunorubicin (Short *et al.*, 2018). In younger patients (<60 years), CR is observed in approximately 60-85% of patients; however, only 40-60% of older patients (>60 years) achieve remission; these represent almost 50% of all patients diagnosed and are, in general, unsuitable for intensive chemotherapy, due to comorbidities and poor performance status. Alternatively, these patients might be suited for less intensive regimens, including low dose cytarabine or hypomethylating agents (HMA), such as Azacitidine (Dombret *et al.*, 2015; Seymour *et al.*, 2017). CR is

achieved once the tumour burden has been reduced, with <5% blasts in the BM, the absence of Auer rods and extramedullary leukaemia, a neutrophil count of >1,000/ $\mu$ L and a platelet count >100,000/ $\mu$ L (Döhner *et al.*, 2015).

Having achieved CR, consolidation chemotherapy is necessary to reduce the risk of relapse (Burnett *et al.*, 2011a), generally centred on a cytarabine-based regimen. Alternatively, patients can be eligible for an allogenic HSC transplant. Even though this procedure usually improves the outcome of intermediate and poor-risk AML patients, it is highly dependent on patient selection, donor and optimal regimen (Short *et al.*, 2018). Relapsed disease and leukaemia-associated complications are the most common causes of death. AML relapse is often associated with a substantial increase in molecular complexity, with multiple new subclones and mutations identified at the time of relapse, leading to increased resistance to cytotoxic chemotherapy (Short *et al.*, 2018). For this reason, there is a constant need to develop new and more effective therapies, especially for older AML patients.

#### 1.2.5.1 Targeted therapies

Treatment options for AML have been completely revolutionised by the development of a targeted therapy against acute promyelocytic leukaemia (APL). Currently, these patients are treated with a combination of all-*trans* retinoic acid (ATRA) and arsenic trioxide (ATO), which induce PML-RARA degradation and cell differentiation (reviewed in (Tomita *et al.*, 2013)). Approximately 80-90% of patients diagnosed with APL are expected to achieve CR, associated with a long-term survival of over 98% (Khwaja *et al.*, 2016; Sanz *et al.*, 2019).

Since the discovery and approval of ATRA, several efforts have been made into developing new targeted therapies against specific mutations, namely *FLT3*. In patients, mutant *FLT3* is constitutively active, leading to the induction of cell proliferation, whilst suppressing differentiation (Meshinchi and Appelbaum, 2009). Inhibition of tyrosine kinase receptors has been successfully used for the treatment of other blood cancers, including Philadelphia-chromosome positive leukaemias (De Kouchkovsky and Abdul-Hay, 2016). For the treatment of AML, first-generation targeted therapies were initially developed, including Sorafenib (Zhang *et al.*, 2008b; Borthakur *et al.*, 2011; Crump *et al.*, 2010) and Midostaurin (Stone *et al.*, 2005; Fischer *et al.*, 2010). However, these are associated with a high incidence of off-target effects (De Kouchkovsky and Abdul-Hay, 2016; Short *et al.*, 2018), which led to the

development of second-generation drugs, such as Quizartinib (Zarrinkar *et al.*, 2009), Crenolanib (Galanis *et al.*, 2014) and Gilteritinib (Ueno *et al.*, 2014). A recurrent concern with FLT3-targeted therapies is the development of secondary mutations, mostly associated with its kinase domain (Daver *et al.*, 2015). To overcome this, conventional FLT3 inhibitors are often used with other treatments, including HMA (Chang *et al.*, 2016; Williams *et al.*, 2013), Venetoclax (Ma *et al.*, 2019; Singh Mali *et al.*, 2021) or proteasome inhibitors (Larrue *et al.*, 2016; Walker *et al.*, 2016; Saliba *et al.*, 2017). Several novel FLT3 inhibitors are currently being developed, such as multikinase inhibitors, including Ponatinib (Shah *et al.*, 2013b), Cabozantinib (Lu *et al.*, 2016; Fathi *et al.*, 2018), Pexidartinib (Smith *et al.*, 2020) and Ibrutinib (Wu *et al.*, 2016); and next-generation inhibitors, for instance FF-10101 (Yamaura *et al.*, 2018).

AML poses as the ideal target for monoclonal antibody-based immunotherapy, as cells are largely present in the blood and BM, tissues readily accessible to antibodies. However, applying this has been challenging as not many leukaemia-specific antigens have been identified to date. AML cells typically express antigens found on normal myeloid progenitor and differentiated cells, such as macrophages and monocytes, particularly CD45 (expressed in 97.2% cells), CD33 (95.3%), and CD13 (94.3%) (Khalidi *et al.*, 1998). Additionally, CD56 expression in t(8;21) AML has been previously associated with a higher rate of relapse (Iriyama *et al.*, 2013), making this marker a potential therapeutic target. The first clinically viable monoclonal antibody to be approved in haematological malignancies was Gemtuzumab ozogamicin (Mylotarg<sup>®</sup>), which targets CD33. By combining Mylotarg<sup>®</sup> with conventional induction chemotherapy on the first day of therapy of previously untreated AML patients, the MRC AML15 trial reported a significant survival benefit without increased toxicity in younger patients with favourable cytogenetics, particularly CBF leukaemias (Burnett *et al.*, 2011b). Currently, this drug is used for the treatment of newly diagnosed CD33-positive AML in adults and for the treatment of relapsed or refractory CD33-positive AML in adults and paediatric patients above the age of 2 (Williams *et al.*, 2019).

Patients unable to tolerate standard chemotherapy have profited from Venetoclax, a Bcl-2 inhibitor, often used in combination with the HMA decitabine or azacytidine (Williams *et al.*, 2019). In 2018, the US food and drug administration (FDA) approved the combination of Venetoclax with either low dose cytarabine or demethylation therapy for older or unfit AML patients following successful clinical trials (Wei *et al.*, 2019; DiNardo *et al.*, 2019). Even

though there were improvements in remission rates and OS, resistance to targeted therapies remains a challenge in the treatment of AML (DiNardo *et al.*, 2019). *FLT3-ITD* and *TP53* defects were shown to correlate with adaptive resistance to Venetoclax alone as well as in combination with other therapies (DiNardo *et al.*, 2019).

IDH inhibitors are used in patients with AML associated with either *IDH1* or *IDH2* mutations, including Ivosidenib (Popovici-Muller *et al.*, 2018), an IDH1 inhibitor, and Enasidenib (Stein *et al.*, 2017; Yen *et al.*, 2017) targeting IDH2. Recently, a new therapeutic agent, CPX-351, has been approved for the treatment of AML with myelodysplasia-related changes and therapy-related AML, based on a liposomal formulation of cytarabine and daunorubicin, showing improved OS as compared to traditional '3+7' treatment regimens (Lancet *et al.*, 2014; Lancet *et al.*, 2018).

Alternatively, treatment of AML has the potential to benefit from a less harmful approach, using chimeric antigen receptor (CAR) T cells, based on genetically engineered T cells able to recognise specific antigen epitopes. Following activation, CAR T cells are able to secrete several anti-tumour cytokines, resulting in the recruitment of other immune cells, tumour cell elimination and in the inhibition of tumour relapse (Marofi *et al.*, 2021; Zhang and Xu, 2017). However, due to the challenge in identifying suitable cell surface markers in AML blasts, this therapy has encountered many challenges. Nevertheless, several markers have been identified as of interest in these cells, including CD33 (Kenderian *et al.*, 2015; Wang *et al.*, 2015b), CD38 (Yoshida *et al.*, 2016) and CD123 (Gill *et al.*, 2014; Thokala *et al.*, 2016).

All these studies show that genomic profiling is of the utmost importance upon treatment design in AML. Currently, there are several treatment options based on different strategies, including monoclonal antibodies, as well as checkpoint inhibitors and cellular therapies, in a clinical trial stage (reviewed in (Short *et al.*, 2018)). However, due to the heterogeneity observed in AML patients, these options are only available for a reduced cohort of patients.

### 1.3 Pathophysiology of RUNX1-ETO

The t(8;21)(q22;q22) reciprocal translocation is the most frequent chromosomal abnormality observed in AML and is frequently associated with the FAB-M2 subtype of AML (Rowley, 1984) (1.2.3). Cells with this translocation are usually associated with impaired

granulocytic maturation and expression of the cell surface markers CD13, CD19, CD34 and CD36 (Bitter *et al.*, 1987). Molecular studies identified the rearrangement of the *RUNX1* gene, located at chromosome 21q22, with the *ETO* gene, at chromosome 8q22, thus generating the fusion protein RUNX1-ETO (aka AML1-ETO, RUNX1-RUNX1T1) (Miyoshi *et al.*, 1993; Miyoshi *et al.*, 1991). AML patients with t(8;21) are often associated with good prognosis (**Table 1.4**), as compared to other subtypes of AML. However, a significant proportion of these patients eventually relapse, which can be attributed to a high heterogeneity within RUNX1-ETO leukaemia (Qin *et al.*, 2017).

### 1.3.1 RUNX1

#### 1.3.1.1 *The RUNX family of transcription factors*

Runt-related TF (RUNX) proteins belong to a family of TFs described as master regulators. Within the last decades, members of this family have been implicated in several processes, including proliferation, cell differentiation, apoptosis, and lineage determination. Furthermore, RUNX family members have been linked to the development of oncogenic processes and signalling pathways associated with cancer development (Otálora-Otálora *et al.*, 2019).

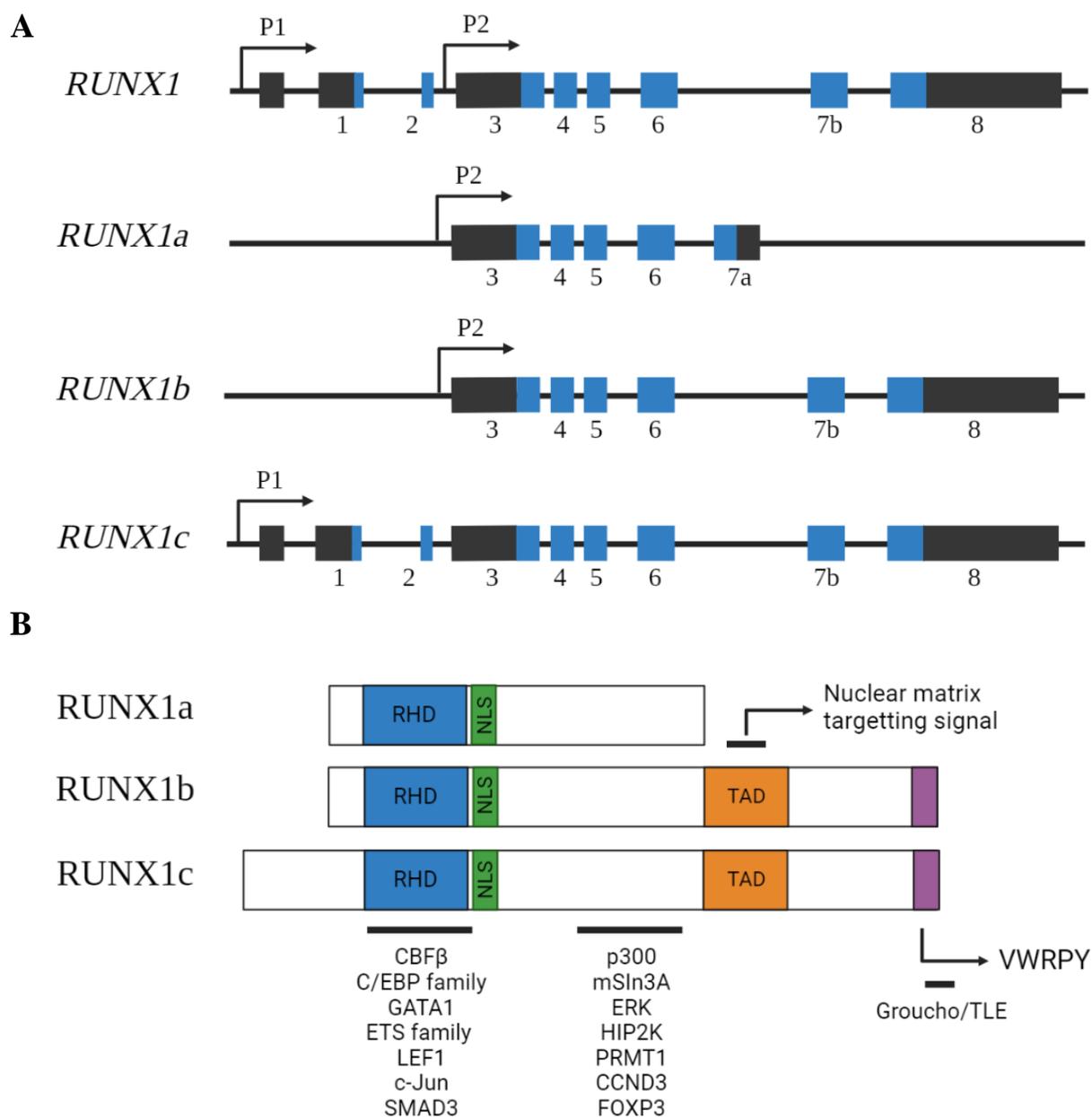
RUNX genes were firstly identified in *Drosophila melanogaster*, where the *runt* gene was found to be essential for early embryonic segmentation (Gergen and Butler, 1988; Nüsslein-Volhard and Wieschaus, 1980). In mammals, there are three *RUNX* genes, each with distinct expression patterns subject to tissue-specification: *RUNX1*, *RUNX2* and *RUNX3*. *RUNX1* is localised on human chromosome 21q22 and is functionally important for haematopoietic cell differentiation (Yamagata *et al.*, 2005; Dowdy *et al.*, 2010). *RUNX2* (or *AML3*) is located on chromosome 6p21 and has been described to be essential in osteogenesis (Lian and Stein, 2003; Lian *et al.*, 2004; Komori, 2003). Lastly, *RUNX3* (or *AML2*) is responsible for regulating gastric epithelial growth, and its located on chromosome 1p36 (Fukamachi, 2006). In cancer, *RUNX* genes have been associated with both oncogenic and tumour suppressor roles (Blyth *et al.*, 2005; Blyth *et al.*, 2010; Kilbey *et al.*, 2008).

#### 1.3.1.2 *RUNX1 structure and function*

All *RUNX* genes share a similar genomic structure (Levanon and Groner, 2004), and their expression is regulated by two promoters, one distal (P1) and one proximal (P2). Depending on the cells stage of differentiation, the promoters are triggered and generate distinct RUNX

isoforms (Ito, 2008; Levanon and Groner, 2004). Moreover, RUNX transcript alternative splicing can also lead to the expression of isoforms with altered properties (Bae *et al.*, 1994; Miyoshi *et al.*, 1995; Tanaka *et al.*, 1995b; Stewart *et al.*, 1997; Bangsow *et al.*, 2001). This leads to distinct developmental expression patterns observed for all RUNX isoforms. At the transcriptional level, *RUNX1* is the largest gene with nine exons and three extensively described isoforms: *RUNX1a*, *RUNX1b* and *RUNX1c* (**Figure 1.4A**). The proximal promoter P2 is responsible for the transcription of isoforms *RUNX1a* and *RUNX1b*, whilst *RUNX1c* is under the control of the distal promoter P1 (Miyoshi *et al.*, 1995). Structurally, all RUNX1 protein isoforms comprise a highly conserved DNA binding domain (also known as Runt domain), a 128-amino acid sequence located near the N-terminus region, responsible for DNA binding at the consensus ‘PyGPyGGTPy’ RUNX motif (Kamachi *et al.*, 1990; Wang and Speck, 1992), protein interactions (Lilly *et al.*, 2016; Nagata *et al.*, 1999), as well as for the nuclear localization of RUNX factors (Michaud *et al.*, 2002; Telfer *et al.*, 2004) (**Figure 1.4B**). *RUNX1a* lacks the transcriptional regulatory domains found in the C-terminal domain of the other RUNX1 isoforms (Tsuzuki *et al.*, 2007). Moreover, the *RUNX1c* isoform possesses an additional five amino acid motif (VWRPY), essential for the recruitment of the Groucho/TLE family of co-repressors (Levanon *et al.*, 1998; Seo *et al.*, 2012b; Yarmus *et al.*, 2006). All RUNX1 proteins further possess a conserved nuclear matrix-targeting signal sequence, important for the regulation of its activity and nuclear localisation (Zaidi *et al.*, 2001; Zeng *et al.*, 1998). This region represents the basis for the functional diversity observed in RUNX proteins, including their ability to function as regulators of transcription (Chuang *et al.*, 2013).

The multiple RUNX1 isoforms have been shown to play specific roles in HSC development and in the regulation of embryonic haematopoiesis. The *RUNX1a* isoform, the shorter of the three, lacks the transactivation domain and is thought to act as a dominant negative (Levanon *et al.*, 2001), promoting haematopoietic commitment (Ran *et al.*, 2013) and increasing HSC renewal (Tsuzuki *et al.*, 2007; Tsuzuki and Seto, 2012). In adult haematopoiesis, *RUNX1c* is the dominant isoform, whilst *RUNX1b* is exclusive to progenitor sub-populations of granulocytes/macrophages, megakaryocytes and lymphoid lineages (Bee *et al.*, 2009; Draper *et al.*, 2016; Telfer and Rothenberg, 2001). In fact, terminal differentiation of granulocytes/macrophages and lymphoid cells relies on the downregulation of *RUNX1b*, and its expression correlates with increased proliferation and colony-forming unit-culture activity in MEPs (Draper *et al.*, 2016).



**Figure 1.4 – Structural representation of RUNX1 isoforms genes and proteins**

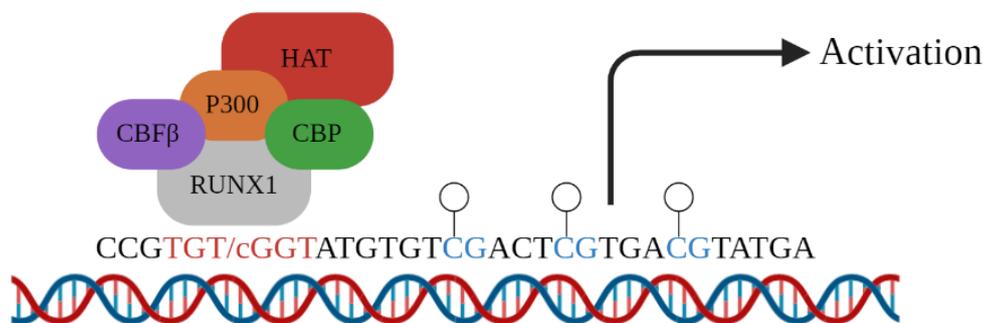
(A) Genomic organisation of the human *RUNX1* locus. The *RUNX1a* and *RUNX1b* isoforms are transcribed from the proximal promoter P2, whilst *RUNX1c* is under the control of the distal promoter P1 and contains a unique N-terminal sequence encoded by exons 1 and 2. Black boxes correspond to untranslated regions, whilst blue boxes represent coding regions. Adapted from (van der Kouwe and Staber, 2019); (B) Schematic representation of the RUNX1a, RUNX1b and RUNX1c proteins and their functional domains. RUNX1a solely contains the RHD domain, whereas RUNX1b and RUNX1c contain the RHD, TAD and VWRPY motif. Below are represented TF and regulators that interact with RUNX1, and the corresponding regions they interact with. Adapted from (van der Kouwe and Staber, 2019; Lam and Zhang, 2012).

**NLS** – Nuclear localisation signal; **RHD** – Runt homology domain; **TAD** – Transactivation domain.

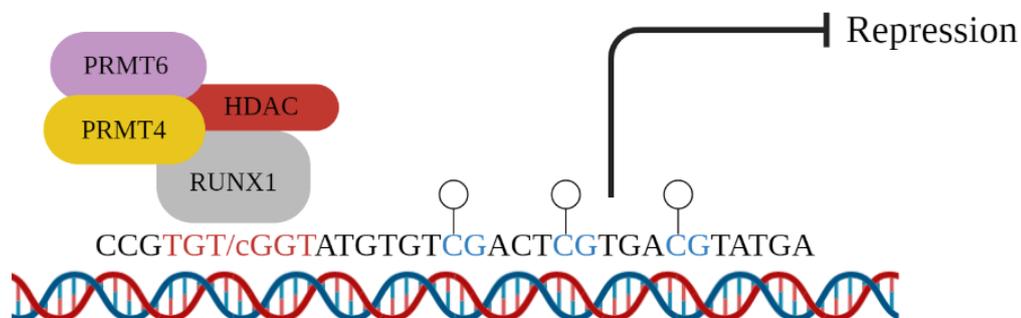
### 1.3.1.3 Transcriptional regulation by RUNX1

The RUNX family of genes represent the  $\alpha$  subunit of a heterodimeric complex formed by dimerization with the core binding factor subunit  $\beta$  (CBF $\beta$ ), ubiquitously expressed and encoded by a single gene in mammals, which in itself is unable to bind to DNA (Kamachi *et al.*, 1990; Ogawa *et al.*, 1993; Wang *et al.*, 1993). However, upon binding to members of the RUNX family, CBF $\beta$  increases the RUNX DNA-binding affinity and promotes complex stabilization, by triggering flexible DNA-recognition loops (Bravo *et al.*, 2001; Huang *et al.*, 2001; Tang *et al.*, 2000; Yan *et al.*, 2004). The cooperative process between RUNX1 and CBF $\beta$  is responsible for not only regulating ubiquitin-mediated degradation of RUNX1 (Huang *et al.*, 2001), but also enhancing RUNX1 phosphorylation/acetylation responsible for a decreased interaction with transcriptional repressors (Wee *et al.*, 2008). The RUNX1-CBF $\beta$  complex has been shown to interact with PU.1, C/EBP $\alpha$ , histone acetyltransferase p300 (p300), mammalian Sin3a (mSin3a) and Friend leukaemia integration 1 (FLI1) (Zhang *et al.*, 1996; Petrovick *et al.*, 1998; Kitabayashi *et al.*, 1998b; Imai *et al.*, 2004; Huang *et al.*, 2009b). Furthermore, it also possesses the ability to regulate several molecules, including the growth factors GM-CSF, MPO and IL-1; the surface receptors T Cell Receptor Alpha and Beta Locus (TCRA and TCRB, respectively), M-CSF receptor and FLT3; the signalling molecule Cyclin Dependent Kinase Inhibitor 1A (CDKN1A); transcriptional activators STAT3 and MYC; and, lastly, proliferation and survival regulators BLK and BCL-2 (Michaud *et al.*, 2003; Ito, 2004). In addition to its interaction with CBF $\beta$ , RUNX1 has been shown to interact with other TFs and transcriptional co-regulators, including ETS Proto-Oncogene 1 (ETS1), with which it coordinates transcriptional activity through the Runt domain within RUNX1, thus eliminating the requirement for CBF $\beta$  and resulting in enhanced DNA-binding ability of the two proteins and the synergistic activation of a promoter (Ito, 2008; Kim *et al.*, 1999). Even though, on their own, RUNX TFs are often characterised as weak, their interaction with other TF help regulate its target genes in a tissue-specific manner by promoting chromatin de-condensation, allowing the recruitment of other transcriptional regulators (Zaret and Carroll, 2011) (**Figure 1.5**).

A



B



**Figure 1.5 – Transcriptional regulation mediated by RUNX1**

RUNX1 can act both as a transcriptional activator or repressor, depending on the presence of coactivators/corepressor at a specific time point. (A) CBF $\beta$  stabilizes RUNX1 and protects it from degradation. Histone acetyltransferases (HATs) are recruited via p300 and CBP proteins, activating gene transcription through histone acetylation. Adapted from (Duque-Afonso *et al.*, 2014; Brettingham-Moore *et al.*, 2015). (B) RUNX1 can also recruit corepressors and other epigenetic modifiers, such as PRMT6 and PRMT4, as well as histone deacetylases (HDAC), which result in the repression of RUNX1 activity. Open circles represent methylation sites [unmethylated]. Adapted from (Brettingham-Moore *et al.*, 2015)

### 1.3.1.4 Post-translational modifications

In addition to being transcriptionally regulated, RUNX1 is also under the control of several post-translational mechanisms, which influence its protein activity, subcellular localisation and stability, correlated with genes essential for myeloid and lymphoid differentiation (Zhao *et al.*, 2008; Seo *et al.*, 2012a). These processes are generally associated with chromatic modifiers, co-factors and other TFs targeting certain regulatory regions.

Transcriptional activation relies in the recruitment of co-activator proteins to certain promoters, which often possess histone acetyltransferase activity. RUNX proteins, due to their abundance in lysine residues, are often modified by lysine acetyltransferases (KAT), which stimulate their transcriptional activity (Blumenthal *et al.*, 2017). P300 is a member of this family, and its mediated acetylation of RUNX1 increases the latter's binding to DNA and its transcriptional activation (Yamaguchi *et al.*, 2004). Lysine acetyltransferase 6A (MOZ) is another RUNX1 co-activator that, upon binding to the RUNX1 C-terminal transactivation domain promotes the expression of genes involved in monocyte/macrophage differentiation (Kitabayashi *et al.*, 2001).

The transcriptional activity of RUNX1 is further regulated by methyltransferases. The mixed lineage leukaemia (MLL) lysine methyltransferase physically interacts with the N-terminal region of RUNX1, stabilising it by inhibiting its poly-ubiquitination (Huang *et al.*, 2011). The protein arginine N-methyltransferase-1 (PRMT1) disrupts the association between RUNX1 and the co-repressor SIN3A, thus enhancing RUNX1 transcriptional activity and promoting its binding to target gene promoters (Zhao *et al.*, 2008).

The most studied post-translational mechanism to control RUNX1 activity is phosphorylation. This is mediated by kinases activated by haematopoietic cytokines and growth factors, as well as cell cycle regulatory proteins. Extracellular signal-regulated kinases (ERK) phosphorylate the C-terminus domain of RUNX1, thus preventing RUNX1 from interacting with SIN3A and enhancing RUNX1-mediated transcription (Imai *et al.*, 2004). Additionally, RUNX1 is phosphorylated by the homeodomain interacting protein kinase 2 (HIPK2), thus inducing p300 phosphorylation and subsequent transcriptional activation (Wee *et al.*, 2008; Aikawa *et al.*, 2006). During cell cycle, the G<sub>1</sub> to S transition is directly regulated by RUNX1, following which the RUNX1 protein is ubiquitously degraded during the G<sub>2</sub>/M phase, a process triggered by cyclin-dependent kinase-1 (CDK1) and CDK6 (Biggs *et al.*, 2006). Furthermore, CDK1/2/6

also phosphorylate RUNX1 (Zhang *et al.*, 2008a), reducing its interaction with histone deacetylase-1 (HDAC1) and HDAC3, further promoting transcriptional activation (Guo and Friedman, 2011). Moreover, upon cytokine stimulation, the PIM1 kinase possesses the ability to interact with RUNX1, increasing its transactivation activity (Aho *et al.*, 2006).

In addition to increasing RUNX1 transcriptional abilities, post-translational modifications can also negatively impact transcription, through deacetylation, methylation and phosphorylation processes. Physiologically, HDAC complexes participate in several processes, including chromatin remodelling and gene expression, and have been shown to interact with RUNX1, in particular HDAC1, SIN3A and Gro/TLE (Levanon *et al.*, 1998; Imai *et al.*, 2004). Interestingly, an increased recruitment of co-repressors is observed in chromosomal translocations involving RUNX1, as compared to wildtype RUNX1, suggesting the repression of its target genes (Guidez *et al.*, 2000).

Additionally, the methyltransferase PRMT4, highly expressed in HSCs, leads to the assembly of a repressive complex that influences myeloid differentiation (Vu *et al.*, 2013). RUNX1 has the ability to form a repressor complex with PRMT6, thus mediating the repression of genes preceding megakaryocytic differentiation (Herglotz *et al.*, 2013). Similarly, RUNX1 phosphorylation can impact its gene expression. Megakaryocytic differentiation has been shown to be repressed by RUNX1 tyrosine phosphorylation (Huang *et al.*, 2012). RUNX1's interaction with c-SRC and the tyrosine phosphatase non-receptor type 11 (SHP2) leads to the inhibition of the interaction between RUNX1 and CBF $\beta$  and, consequently, GATA-1 and FLI1 (Huang *et al.*, 2012).

#### 1.3.1.5 *The role of RUNX1 in normal haematopoiesis*

RUNX1 is expressed in all haematopoietic cells except for mature erythrocytes (North *et al.*, 1999; North *et al.*, 2002). Absence of RUNX1 was shown to result in foetal death due to severe haemorrhaging of the central nervous system, due to defective angiogenesis, as well as lack of definite haematopoiesis (Okuda *et al.*, 1996; Wang *et al.*, 1996). However, continuous expression of RUNX1 is not necessary for survival in adult mice, indicated by absence of lethality (Cai *et al.*, 2011; Growney *et al.*, 2005; Ichikawa *et al.*, 2004; Jacob *et al.*, 2010; Putz *et al.*, 2006). In this model, RUNX1 deletion resulted in either the expansion (Growney *et al.*, 2005; Ichikawa *et al.*, 2004) or exhaustion (Jacob *et al.*, 2010) of phenotypic HSCs. However,

studies suggest that deletion of RUNX1 results in a significant expansion of the HSPC compartment, as well as reduced apoptosis and ribosome biogenesis, possible contributing to a pre-leukaemic condition (Cai *et al.*, 2011; Growney *et al.*, 2005; Ichikawa *et al.*, 2004; Jacob *et al.*, 2010; Cai *et al.*, 2015). Moreover, absence of RUNX1 in adult mice also resulted in decreased numbers of B and T cells, as well as lower platelet counts (Growney *et al.*, 2005; Ichikawa *et al.*, 2004). Additionally, a role for RUNX1 in balancing megakaryocytic differentiation has been proposed, through its interaction with AP-1, p300, GATA and ETS TFs (Elagib *et al.*, 2003; Pencovich *et al.*, 2013; Pencovich *et al.*, 2011). RUNX1 has further been found to act as a transcriptional co-repressor, suggesting that its role is highly context dependant, relying on the complexes in which it functions (**Figure 1.5**).

In addition to its role in haematopoiesis, RUNX1 has also been associated with several immune cell processes, including T lymphocyte development in the thymus, in a RUNX1-dependent manner (Ikawa *et al.*, 2004; Kawamoto *et al.*, 1999; Krueger and von Boehmer, 2007; Perry *et al.*, 2004; Petrie, 2007). RUNX1 induces *BCL11b* expression which, consequently, promotes the expression of T cell-lineage specific genes, including *Thpok* (*Zbtb7b*) and *RUNX3* (Kojo *et al.*, 2017; Liu *et al.*, 2010). Conversely, loss of RUNX1 results in a block in normal T cell development (Egawa *et al.*, 2007; Sato *et al.*, 2003). Furthermore, RUNX1 expression is essential for the development of natural killer T cells (Tachibana *et al.*, 2011). In the BM, loss of *RUNX1* results in defects in early B cell development, as this is required, together with CBF $\beta$ , to cooperate with the TF EBF for progression into the pro-B cell stage (Maier *et al.*, 2004; Seo *et al.*, 2012a).

#### 1.3.1.6 The role of RUNX1 in AML

Considering the role *RUNX1* plays in haematopoietic development (**1.3.1.5**), it is not surprising that this gene has been found to be a recurrent target for mutations and chromosomal abnormalities in haematological disorders. RUNX1 has been shown to be implicated in more than 50 chromosomal translocations reported in paediatric acute lymphoblastic leukaemia (ALL), AML and myelodysplastic syndrome (MDS) (De Braekeleer *et al.*, 2011). The most common chromosomal translocations involving RUNX1 are the t(8;21)(q22;q22), resulting in the expression of the fusion protein RUNX1-ETO (**1.3.3**); the t(12;21)(p13;q22), leading to the expression of ETV6-RUNX1, present in about 25% of patients with pre-B cell ALL (Golub *et al.*, 1995); and t(3;21)(q26;q22), detected in 3% of therapy-related MDS and AML (Rubin *et al.*, 1995).

*al.*, 1990; Rubin *et al.*, 1987), in which the N-terminal portion of RUNX1 is fused with one of three genes present on chromosome 3, including EVI, MDS1 or EAP (Nucifora *et al.*, 1994). However, even though these translocations represent a major risk factor, additional cooperating mutations are necessary for leukaemogenesis to occur.

Another phenomenon related to cancer development involving the *RUNX1* gene is the occurrence of somatic point mutations, identified in approximately 15% of *de novo* AML, especially in the FAB-M0 subtype, and in 3% of paediatric AML (Tang *et al.*, 2009; Greif *et al.*, 2012; Mendler *et al.*, 2012; Skokowa *et al.*, 2014). Even though mutations are also common in therapy-related MDS and AML (16-40%) (Harada *et al.*, 2003; Christiansen *et al.*, 2004), these are considered a rare event in leukaemia (Preudhomme *et al.*, 2000; Harada *et al.*, 2004). Somatic mutations can either be mono-allelic or bi-allelic, with the latter being predominant in undifferentiated AML (Osato, 2004). Mono-allelic mutations are usually localised in the Runt domain and influence the DNA-binding ability of RUNX1 (Mangan and Speck, 2011; Song *et al.*, 1999; Matheny *et al.*, 2007). Traditionally, these are defined as loss-of-function mutations; however, some mutants have shown gain of function. For instance, expression of DNA-binding RUNX1 mutants in BM progenitors led to an increased replating efficiency, even as compared to RUNX1-deficient cells (Cammenga *et al.*, 2007). Furthermore, other mutations in the *RUNX1* N-terminal domain have been observed, that do not result in clear transactivation defects, and are particularly associated with epigenetic modifications in RUNX1 target genes (Huang *et al.*, 2011). Mutations are further associated with older age, male gender, and poor disease prognosis, as compared to wild-type RUNX1 (Tang *et al.*, 2009; Greif *et al.*, 2012; Mendler *et al.*, 2012). RUNX1 mutations are often observed in combination with *FLT3-ITD*, *FLT3-TKD* or mutations in driver genes, such as *CEBPA*, *DNMT3A*, *NRAS*, *KIT*, *IDH1*, *IDH2*, and *WT1* (Gaidzik *et al.*, 2011; Schnittger *et al.*, 2011; Greif *et al.*, 2012; Mendler *et al.*, 2012). Interestingly, mutations in the *RUNX1* and *NPM1* genes are suggested to be mutually exclusive (Greif *et al.*, 2012; Mendler *et al.*, 2012; Gaidzik *et al.*, 2011).

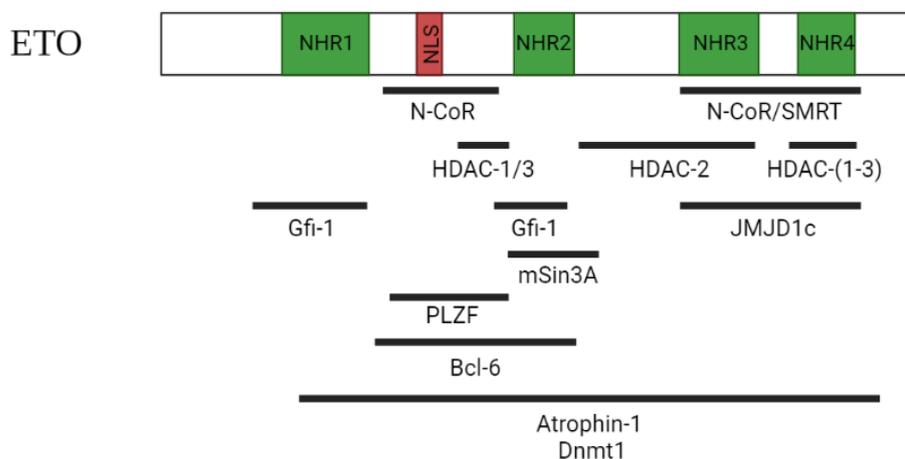
### 1.3.2 ETO

The ETO family of genes comprises three members: the first member to be identified, eight twenty-one (*ETO*, or *MTG8*); the myeloid transforming gene related protein-1 (*MTGR1*); and the myeloid transforming gene chromosome 16 (*MTG16*) (Miyoshi *et al.*, 1993; Fracchiolla *et al.*, 1998; Gamou *et al.*, 1998; Kitabayashi *et al.*, 1998a). Until its identification as RUNX1's

fusion partner and member of the RUNX1-ETO fusion protein, the ETO protein was virtually unknown (Erickson *et al.*, 1992; Miyoshi *et al.*, 1993; Nisson *et al.*, 1992). Its gene is located on chromosome 8q22 and consists of 11 exons (Wolford and Prochazka, 1998) (**Figure 1.6A**). The ETO protein consists of four evolutionary conserved domains, termed *nervy* homology regions (NHR) 1-4, or zinc-finger motifs (**Figure 1.6B**). Whilst NHR1 and NHR2 are responsible for the interaction with other protein motifs (Lutterbach *et al.*, 1998; Davis *et al.*, 1999; McGhee *et al.*, 2003), NHR4, consisting of two non-classical zinc fingers, mediates protein interactions, but its unable to bind DNA (Lutterbach *et al.*, 1998; Wang *et al.*, 1998). Additionally, NHR3 is suggested to be involved with specific co-repressors (Hildebrand *et al.*, 2001). Under physiological conditions, ETO is found in the nucleus, due to the presence of a non-canonical nuclear localisation signal (NLS), found between NHR1 and NHR2 (Davis *et al.*, 1999; Odaka *et al.*, 2000; Sacchi *et al.*, 1998). ETO is expressed in several tissues, but its most abundant in the heart, brain, lungs and testis (Erickson *et al.*, 1994; Wolford and Prochazka, 1998). Moreover, similarly to RUNX1, ETO is expressed in HSPC, but absent from differentiated leukocytes (Erickson *et al.*, 1996).

ETO generally functions as a transcriptional co-repressor (Hiebert *et al.*, 2001). In support of this, earlier studies aiming at isolating the human nuclear receptor co-repressor (N-CoR) were performed using ETO as bait (Wang *et al.*, 1998). This complex had been previously shown to interact with DNA-bound nuclear receptors, thus repressing normal gene transcription through the recruitment of HDAC complexes (Hörlein *et al.*, 1995). NH4 has been shown to be sufficient to promote the interaction of ETO with the N-CoR complex, as deletion of NHR 1-3 motifs abolished any binding between ETO and N-CoR, whilst a portion of ETO containing the NHR4 domain was enough to promote binding (Wang *et al.*, 1998). Subsequent studies by Lutterbach *et al.* confirmed these observations and added that ETO can bind to the central portion of the N-CoR complex (Lutterbach *et al.*, 1998).

Moreover, follow-up studies have established a relationship between ETO and the silencing mediator of retinoic acid and thyroid hormone receptors (SMRT) complex (Gelmetti *et al.*, 1998; Chen and Evans, 1995), as well with HDACs (Wang *et al.*, 1996; Rao *et al.*, 1996; Wang *et al.*, 1998; Lutterbach *et al.*, 1998), the latter in combination with mSin3A, leading to the establishment of a N-CoR/mSin3A/HDAC1 complex (Wang *et al.*, 1998).

**A****B**

**Figure 1.6 – Structural representation of the ETO gene and protein**

(A) Genomic organisation of the human *ETO* gene. Black boxes correspond to untranslated regions, whilst yellow boxes represent coding regions. Adapted from (Lin *et al.*, 2017). (B) Schematic representation of the ETO protein and its functional domains. The ETO protein comprises 4 evolutionary conserved *nervy* homology domains (NHR), as well as an NLS. Below are represented regulators and complexes that interact with ETO, and the corresponding regions they interact with. Adapted from (Davis *et al.*, 2003; Lam and Zhang, 2012).

**NHR** – Nery Homology Region; **NLS** – Nuclear Localisation Signal

These observations suggest that ETO is a component of several co-repressor complexes *in vivo*. However, several authors have demonstrated that, even though NHR4 is necessary to promote the interactions between ETO and both the SMRT and the N-CoR complex *in vitro*, it is not sufficient, and requires the presence of other NHR domains (Zhang *et al.*, 2001; Hildebrand *et al.*, 2001).

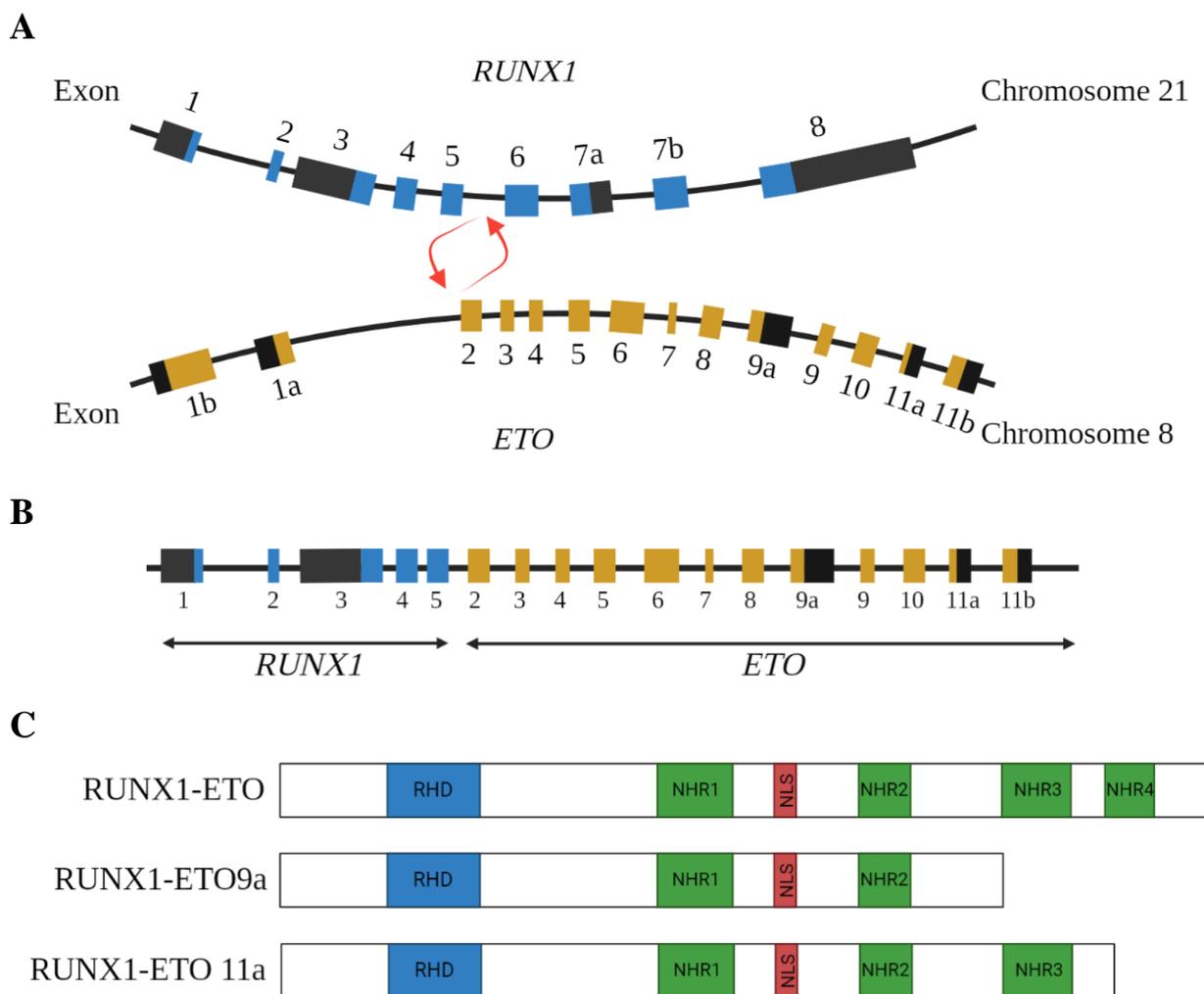
ETO has been suggested to interact with sequence-specific DNA binding proteins, including the transcriptional repressors promyelocytic leukaemia zinc-finger (PLZF) and GFI1 (Melnick *et al.*, 2000b; Costoya and Pandolfi, 2001; Grimes *et al.*, 1996a; McGhee *et al.*, 2003). PLZF is usually found in haematopoietic cells and has been shown to be downregulated as cells differentiate into the myeloid lineage (Costoya and Pandolfi, 2001). ETO binds PLZF, promoting the activity of histone deacetylases and enhancing the transcriptional repression induced by this protein (Melnick *et al.*, 2000b). GFI1 expression is elevated in the thymus, spleen, BM and testis, and has been shown to play a role in haematopoietic cell differentiation and survival (Grimes *et al.*, 1996a; Grimes *et al.*, 1996b; Tong *et al.*, 1998). Similarly to ETO, GFI1 is nuclear matrix-attached, and generally interacts with HDAC-containing complexes (McGhee *et al.*, 2003). Altogether, these observations support the notion that ETO can interact with several co-repressor proteins and complexes, suggesting a role in haematopoietic development, specifically cell differentiation and proliferation.

ETO is involved in a non-random chromosomal translocation that binds it to the RUNX1 gene on chromosome 21 (Erickson *et al.*, 1992; Gamou *et al.*, 1998). Briefly, ETO contributes to the ability of RUNX1-ETO to bind to histone-deacetylase complexes, playing an essential role in the development of t(8;21) leukaemia (Hiebert *et al.*, 2001) (**1.3.3**).

### **1.3.3 RUNX1-ETO**

#### *1.3.3.1 RUNX1-ETO structure*

The chromosomal breakpoints that generate RUNX1-ETO occur within intron 5 of the *RUNX1* gene and in intron 1 of the *ETO* gene (Tighe and Calabi, 1995; Zhang *et al.*, 2002) (**Figure 1.7A-B**). The exact mechanism through which this translocation occurs remains unclear; however, Wnt/ $\beta$ -catenin signalling has been suggested to be an instigator of the genomic proximity between the *RUNX1* and *ETO* genes, facilitating translocation events (Ugarte *et al.*, 2015).



**Figure 1.7 – RUNX1-ETO structure and isoforms**

(A) Genomic structure of *RUNX1* on chromosome 21 and *ETO* on chromosome 8. The translocation occurs between exons 5 and 6 of the *RUNX1* gene, and exons 1 and 3 of *ETO*. Black boxes indicate untranslated regions, whilst blue/yellow boxes indicate coding sequences. Adapted from (Lin *et al.*, 2017). (B) Structure of the fusion gene *RUNX1-ETO*, containing exons 1-5 from the *RUNX1* gene, and exons 2-11 derived from *ETO*. Adapted from (Swart and Heidenreich, 2021). (C) Protein structure of the full-length *RUNX1-ETO* and derived isoforms. *RUNX1-ETO* possesses the RHD domain (blue box) from *RUNX1*, responsible for DNA binding and heterodimerisation with  $\text{CBF}\beta$ , as well as four NHR domain (green boxes), derived from the *ETO* domain, involved in the binding of co-repressor complexes. *RUNX1-ETO* also contains an NLS, represented in red. *RUNX1-ETO9a* lacks NHR3 and 4 domains, whilst *RUNX1-ETO* lacks solely the latter. Adapted from (Lin *et al.*, 2017).

**NHR** – Nery homology region; **NLS** – Nuclear localisation signal; **RHD**- Runt homology domain.

RUNX1-ETO encompasses the N-terminal region of the RUNX1 protein, containing the Runt domain, and the majority of the ETO protein, including the four NHR regions, resulting in a 752 amino-acid fusion protein (Miyoshi *et al.*, 1993) (**Figure 1.7C**). The Runt domain present in the RUNX1 portion is responsible for mediating DNA binding, as well as the heterodimerisation with the co-factor CBF $\beta$  (**1.3.1.5**). Within the ETO portion, NHR2 has been shown to be important for leukaemic development, as it's responsible for the recruitment on the NCoR-SIN3A co-repressor complex, together with HDAC (Liu *et al.*, 2006; Kwok *et al.*, 2009), and its disruption results in a decrease in the self-renewal ability of RUNX1-ETO-expressing HSCs (Byrd *et al.*, 2002). NHR4 is responsible for the recruitment of SMRT, as well as SIN3 and HDACs, via NCoR (Hug and Lazar, 2004), assisted by NHR3. Furthermore, between NHR1 and 2, the ETO domain contains an NLS, essential for the RUNX1-ETO nuclear localisation (Odaka *et al.*, 2000; Barseguian *et al.*, 2002).

In addition to the full-length RUNX1-ETO, containing exons 1-5 from RUNX1, and exons 2-11 from ETO, several isoforms of the RUNX1-ETO transcript have been identified. These include the alternative exons 9a and 11a, providing a stop codon following the amino acid encoded by exon 9 and 11, respectively (Wolford and Prochazka, 1998). From the 9a exon, a C-terminal truncated protein is generated, termed RUNX1-ETO9a, lacking both NHR3 and 4 (**Figure 1.7C**). This protein has a decreased ability to repress RUNX1-mediated gene activation (DeKelver *et al.*, 2013b); however, expression of RUNX1-ETO9a promotes faster leukaemic development as compared to the unspliced form in mice (Yan *et al.*, 2006), a reason why it is often used for RUNX1-ETO studies, in detriment of the full-length protein. RUNX1-ETO11a equally leads to the generation of a truncated RUNX1-ETO isoform, lacking the NHR4 domain, with reduced transcriptional activity (Kozu *et al.*, 2005) (**Figure 1.7C**). However, the exact mechanism through which this isoform contributes to leukemogenesis has not been determined yet.

### 1.3.3.2 Cooperation between RUNX1 and RUNX1-ETO

A central mechanism for RUNX1-ETO-induced leukaemogenesis relies on the dominant inhibition of native RUNX1. Studies examining the consequences of RUNX1-ETO expression in mice have been essential to further the understanding of the leukaemogenic process observed in AML t(8;21). Similarly to previous observations in RUNX1 KO mice (**1.3.1.5**), RUNX1-ETO heterozygous mice display early embryonic lethality and identical haematopoietic defects (Yergeau *et al.*, 1997; Okuda *et al.*, 1998).

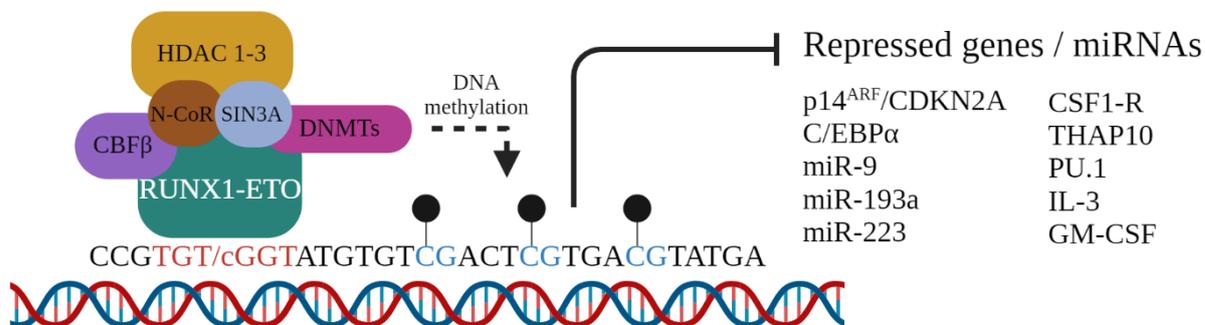
This suggests that RUNX1-ETO is able to block RUNX1 function, in a dominant-negative approach, thus influencing normal haematopoietic processes and leading to the development of a preleukaemic condition. Several studies have shown that the mechanisms through which RUNX1-ETO represses RUNX1-induced gene activation is through the recruitment of co-repressors by the ETO moiety (Gelmetti *et al.*, 1998; Lutterbach *et al.*, 1998; Wang *et al.*, 1998; Amann *et al.*, 2001; Zhang *et al.*, 2001; Hiebert *et al.*, 2001) (**Figure 1.8A**). To bypass embryonic death, a conditional RUNX1-ETO knock-in model was developed, which allowed the expansion of myeloid progenitor cells (Higuchi *et al.*, 2002). However, even though these cells presented aberrant growth and differentiation phenotypes, RUNX1-ETO alone did not lead to the development of AML, suggesting that additional molecular abnormalities are necessary for leukaemogenesis (**1.2.2**).

In human cord blood (CB) derived CD34<sup>+</sup> HSPC, ectopic expression of RUNX1-ETO resulted in the suppression of both erythroid and granulocytic growth, with an increased self-renewal capacity (Mulloy *et al.*, 2002; Tonks *et al.*, 2003; Tonks *et al.*, 2004). Subsequent studies have further validated these observations, in which RUNX1-ETO was responsible for blocking normal myeloid development in embryonic stem cells, as a result of cell cycle arrest and through the interference with RUNX1 chromatin binding (Nafria *et al.*, 2020). Similarly, these studies demonstrated that RUNX1-ETO alone is insufficient to induce leukaemia (Rhoades *et al.*, 2000; Yuan *et al.*, 2001; Higuchi *et al.*, 2002; Mulloy *et al.*, 2003; Tonks *et al.*, 2004) and additional genetic abnormalities are necessary for this process to occur (Schessl *et al.*, 2005; Wang *et al.*, 2011b; Zhao *et al.*, 2014; Goyama *et al.*, 2016).

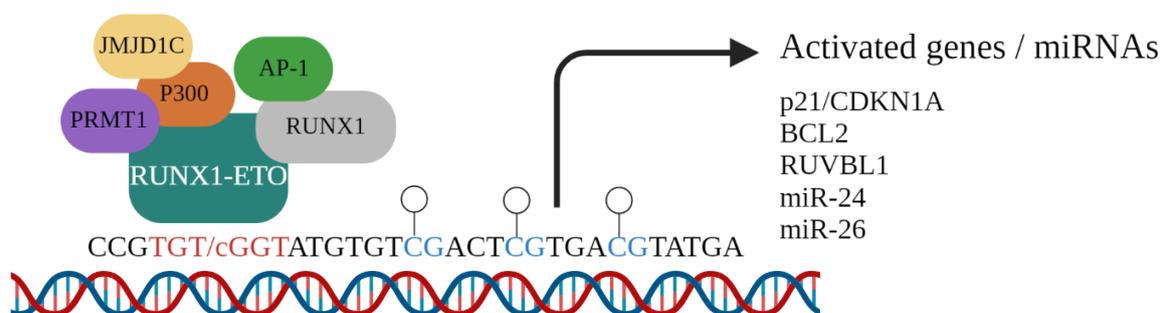
Subsequent studies have shown that there is more complexity to this process than initially thought. RUNX1 has previously been shown to be mutated in several types of AML (**1.3.1.6**); however, no inactivating mutations are observed in t(8;21) AML, suggesting that functional RUNX1 is necessary to promote the leukaemogenic process induced by RUNX1-ETO (Goyama and Mulloy, 2011). Later studies have supported this hypothesis, in which knockdown (KD) of RUNX1 resulted in growth inhibition and decreased survival of RUNX1-ETO leukaemic cells (Ben-Ami *et al.*, 2013; Goyama *et al.*, 2013).

Mechanistically, chromatin immunoprecipitation sequencing (ChIP-Seq) and RNA-sequencing (RNASeq) studies have identified RUNX1 as a member of the TF complex containing RUNX1-ETO, controlling gene activation or repression based on chromatin binding (Li *et al.*, 2016).

A



B



**Figure 1.8 – Mechanisms of transcriptional dysregulation induced by RUNX1-ETO**

(A) For transcriptional repression, RUNX1-ETO recruits corepressors, including N-CoR, HDACs, mSin3A, further interacting with DNMT proteins no promote DNA methylation, resulting in the repression of target genes. (B) For transcriptional activation, RUNX1-ETO and native RUNX1, coupled with CBFβ, bind to RUNX1-motifs, promoting the recruitment of AP-1. RUNX1-ETO also interacts with the coactivators p300, PRMT1 and JMJD1C to activate target gene transcription. Representative target genes/miRNAs regulated by RUNX1-ETO are also shown. Open/full circles represent methylation sites (unmethylated and methylated, respectively). Adapted from (Lin *et al.*, 2017; Duque-Afonso *et al.*, 2014).

Additionally, KD of RUNX1 or RUNX1-ETO results in ‘inverse’ gene expression profiles, suggesting opposite roles in gene regulatory processes (Ben-Ami *et al.*, 2013). Further studies using siRNA have further elucidated the transcriptomic program driven by RUNX1-ETO (Ptasinska *et al.*, 2014; Ptasinska *et al.*, 2012). This fusion protein has been shown to significantly reprogramme the transcriptional network and maintain leukaemia by impairing the expression of approximately 1,400 genes linked to myeloid and granulocytic differentiation (Ptasinska *et al.*, 2014; Ptasinska *et al.*, 2012). Approximately 50% of all genes affected were shown to either maintain normal expression or be upregulated by RUNX1-ETO, and were associated with several processes, including cell cycle progression, glycolysis, oxidative phosphorylation, MTOR signalling, RNA processing, and ribosome biogenesis (Ptasinska *et al.*, 2014; Ptasinska *et al.*, 2012).

### 1.3.3.3 Identification of RUNX1-ETO-mediated transcriptional de-regulation

RUNX1-ETO regulates gene expression through interactions with several different proteins via its Runt domain, through binding to CBF $\beta$ , a feature that is maintained in RUNX1-ETO expressing cells. In normal haematopoiesis, the RUNX1/CBF $\beta$  complex is essential for the emergence of haematopoietic cells and previous studies have shown that loss of either one of the factors results in embryonic death with lack of definitive haematopoiesis (1.3.1.5). Similarly, CBF $\beta$  promotes the binding of RUNX1-ETO to DNA, (Roudaia *et al.*, 2009), although the exact nature of this mechanism remains to be elucidated (Gu *et al.*, 2000; Kanno *et al.*, 1998).

RUNX1-ETO requires the recruitment of complexes comprised of multiple proteins to specific target genes. Earlier studies have described RUNX1-ETO as a dominant negative regulator, acting upon RUNX1 transcriptional target genes, thus inhibiting normal myeloid differentiation (Meyers *et al.*, 1995). Through its interaction with NCoR and mSin3A, ETO recruits HDACs, thus leading to changes in normal chromosome structure to allow a closer conformation on RUNX1-ETO target promoters (Lutterbach *et al.*, 1998; Wang *et al.*, 1998; Gelmetti *et al.*, 1998; Amann *et al.*, 2001; Hildebrand *et al.*, 2001). In addition to HDACs, RUNX1-ETO has also been shown to recruit DNA methyltransferases (DNMTs), thus promoting DNA methylation (Liu *et al.*, 2005). Subsequent studies performed in the human t(8;21) cell line Kasumi-1 have identified an endogenous complex containing RUNX1-ETO, as well as several other haematopoietic TF, including CBF $\beta$ , E proteins (HEB and E1A), LYL1, LMO22 and its binding partner LDB1 (Sun *et al.*, 2013). KD of any of the members of this

complex resulted in a decrease in the protein level of other components, suggesting a mutual stabilisation mechanism.

In addition to acting as a transcriptional repressor, RUNX1-ETO can induce the expression of several target genes through the interaction with different transcriptional activators (**Figure 1.8B**). The p300 acetyltransferase increases RUNX1-ETO transcriptional activation ability by promoting its acetylation and recruiting transcriptional pre-initiation complexes (Wang *et al.*, 2011a). The methyltransferase PRMT1, on the other hand, promotes residue methylation and contributes to gene activation (Shia *et al.*, 2012). Similarly, RUNX1-ETO directly recruits the histone demethylase JMJD1C, enhancing transcriptional upregulation (Chen *et al.*, 2015). Moreover, several haematopoietic TF have been shown to interact with the Runt domain of this fusion protein, including *GATA1*, *CEBPA*, and *PU.1* (Pabst *et al.*, 2001).

Transcriptomic analysis has revealed significantly altered gene expression profiles as a result of the expression of RUNX1-ETO in t(8;21) AML (Ross *et al.*, 2004; Valk *et al.*, 2004). Genome-wide studies performed in the t(8;21) cell lines Kasumi-1 and SKNO-1, as well as t(8;21) patient samples, revealed that 60-80% of RUNX1-ETO binding sites are shared with RUNX1 (Gardini *et al.*, 2008; Ptasińska *et al.*, 2012; Li *et al.*, 2016; Ben-Ami *et al.*, 2013). As described above (**1.3.3.2**), a basal level of RUNX1 is necessary for maintaining the cell growth induced by RUNX1-ETO, as the complete absence of its regulated differentiation genes is not feasible (Gardini *et al.*, 2008). This further results in the recruitment of protein complexes with opposing roles. For instance, whilst RUNX1-ETO recruits HDACs, RUNX1 cooperates with p300 on genes repressed by this fusion protein (Ptasińska *et al.*, 2014). Another example is the recruitment of DNMT1 and DNMT3A by RUNX1-ETO, whilst RUNX1 recruits TET histone demethylases (Gu *et al.*, 2014) (**Figure 1.8A**). However, not all RUNX1-ETO binding sites contain RUNX1 motifs, suggesting that there are other factors that contribute to the binding of this fusion protein to DNA (Ptasińska *et al.*, 2012; Maiques-Diaz *et al.*, 2012). This binding can occur through E proteins, including HEB and E2A, as previous ChIP-Seq analyses identified the overrepresentation of E-boxes in RUNX1-ETO binding regions (Gardini *et al.*, 2008; Zhang *et al.*, 2004a; Sun *et al.*, 2013). Moreover, genome-wide analysis identified ERG and FLI1, members of the ETS family of TFs, as facilitators in the RUNX1-ETO binding process towards DNA, further occupying similar genomic regions as to this fusion protein (Martens *et al.*, 2012). Similarly, SP1 has been shown to interact with RUNX1-ETO and

promote DNA binding through SP1-binding sites; in fact, SP1 inhibition has been demonstrated to attenuate RUNX1-ETO-mediated transcriptional repression of target genes (Maiques-Diaz *et al.*, 2012). Further studies have associated RUNX1-ETO binding to chromatic accessibility (Saeed *et al.*, 2012; Maiques-Diaz *et al.*, 2012). However, most of these studies have primarily focused on transformed AML cell lines. To address this, Tonks *et al.* performed microarray analysis on normal human haematopoietic progenitor cells transduced with RUNX1-ETO. This allowed the identification dysregulated genes when RUNX1-ETO was expressed as a single abnormality in human CD34<sup>+</sup> cells (Tonks *et al.*, 2007). This study further promoted the identification of several genes dysregulated in RUNX1-ETO HSPC, and their potential involvement in the leukaemogenic process, including CD200 and  $\gamma$ -catenin.

#### 1.3.3.4 Molecular mechanisms of RUNX1-ETO-induced leukaemia

The process through which RUNX1-ETO can induce cell proliferation involves several factors. HIF1 $\alpha$  is a TF known to be overexpressed in t(8;21) patients, and the interaction between this factor and RUNX1-ETO have been shown to be necessary to drive proliferation in both *in vivo* and *in vitro* models (Peng *et al.*, 2008). Mechanistically, HIF1 $\alpha$  interacts with DNMT3a, leading to DNA hypermethylation (Gao *et al.*, 2015). RUNX1-ETO further leads to the upregulation of Pontin, an ATPase involved in several cellular functions, including cell cycle progression and proliferation (Breig *et al.*, 2014). Dysregulated signalling pathways, such as Wnt and RAS, have been shown to drive cell proliferation through its constitutive activation in t(8;21) patients (Kuchenbauer *et al.*, 2006; Bacher *et al.*, 2006). In HSC, expression of RUNX1-ETO results in an enhanced self-renewal potential, coupled with a block in normal cell differentiation.

RUNX1-ETO directly alters gene expression by interfering with the normal function of myeloid TF. By interacting with PU.1 and C/EBP $\alpha$ , RUNX1-ETO reduces their DNA binding activity, thus influencing their normal expression (Pabst *et al.*, 2001; Vangala *et al.*, 2003). Furthermore, RUNX1-ETO influences GATA1 expression, a major erythroid TF, by preventing its acetylation and normal transcriptional activity (Choi *et al.*, 2006).

Additionally, RUNX1-ETO has been proposed to possess a role in the regulation of apoptosis by inhibiting normal apoptotic processes, through the upregulation of the anti-apoptotic proteins BCL-1 and BCL-XL (Klampfer *et al.*, 1996; Chou *et al.*, 2012). Subsequent

depletion of RUNX1-ETO results in an increase in apoptosis and cell cycle arrest. Further, RUNX1-ETO has been demonstrated to downregulate the expression of genes involved in DNA repair, including *POLE* and *OGG1*, leading to an increase in DNA damage (Alcalay *et al.*, 2003; Krejci *et al.*, 2008).

Dysregulation of tumour suppressor genes has also been observed in RUNX1-ETO induced leukaemia. RUNX1-ETO has been shown to transcriptionally repress p14<sup>ARF</sup> (aka *CDKN2A*) and NF1 (Linggi *et al.*, 2002; Yang *et al.*, 2005). RUNX3, another member of the RUNX family and a tumour-suppressor, is also suppressed by RUNX1-ETO (Cheng *et al.*, 2008). Conversely, RUNX1-ETO also has the ability to upregulate the expression of cyclin-dependent kinase inhibitor 1A (p21<sup>WAF1</sup> or *CDKN1A*) (Peterson *et al.*, 2007). Transcriptional regulation of genes involved in signalling pathways has been shown to be modulated by RUNX1-ETO, including through the expression of  $\gamma$ -catenin (*JUP*) and Cox-2, which in turn activate Wnt signalling (Müller-Tidow *et al.*, 2004; Zheng *et al.*, 2004b; Zhang *et al.*, 2013; Yeh *et al.*, 2009; Tonks *et al.*, 2007). Suppression of the leukaemogenic potential in RUNX1-ETO was achieved upon the use of Cox inhibitors, and by deleting  $\beta$ - or  $\gamma$ -catenin, indicating a critical role of the Wnt pathway in the maintenance of LSC in RUNX1-ETO cells (Zhang *et al.*, 2013; Yeh *et al.*, 2009). The TPO/MPL signalling pathway has been identified as a key pathway involved in the survival and leukemogenesis process induced by RUNX1-ETO through the upregulation of the anti-apoptotic protein Bcl-xL, as well as promotion of the PI3K/AKT and JAK/STAT pathways (Pulikkan *et al.*, 2012; Chou *et al.*, 2012). Further pathways that have been associated with RUNX1-ETO include NF- $\kappa$ B signalling (Nakagawa *et al.*, 2011) and UBASH3B/CBL (Goyama *et al.*, 2015).

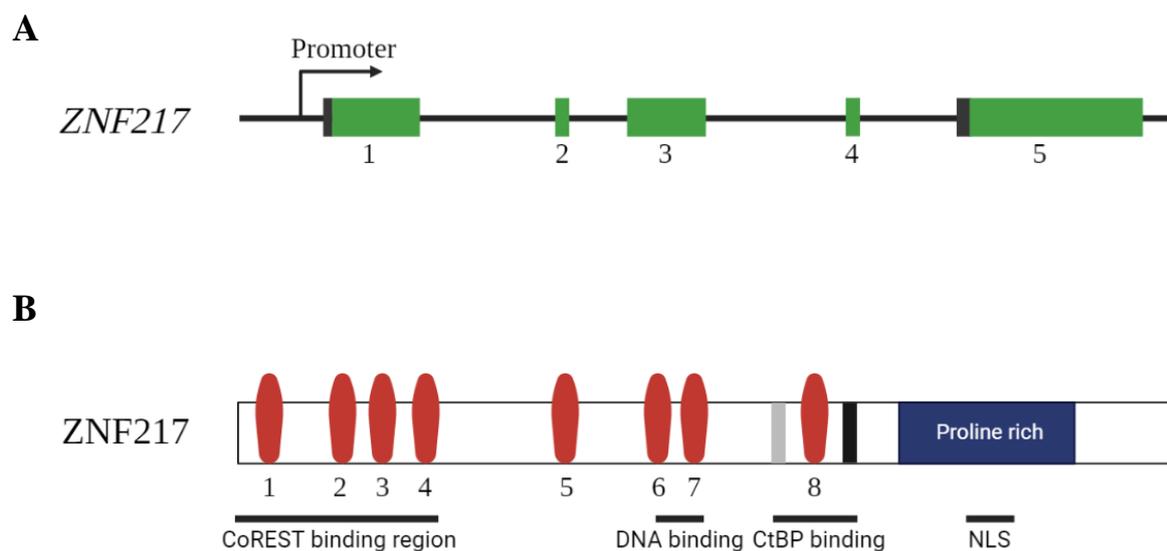
Recently, through the integration of several techniques, including RNASeq, ChIP-Seq and DNase I hypersensitive site-se1 experiments, it was possible to identify *CCND2* as a target of interest in RUNX1-ETO-driven leukaemogenesis, a member of the CDK6/*CCND2* kinase complex (Martinez-Soria *et al.*, 2018). Depletion of either RUNX1-ETO or *CCND2*, as well as pharmacological inhibition of CDK6, resulted in cell senescence, suggesting that RUNX1-ETO is responsible for driving cell cycle progression through the G<sub>1</sub> phase by promoting the expression of both CDK6 and *CCND2* (Byrd *et al.*, 2002; Martinez-Soria *et al.*, 2018).

Altogether, these observations indicate that RUNX1-ETO leads to significant changes in normal transcriptional processes, likely mediated through changes in TF expression. Whilst

several transcriptomic studies have been performed, these usually rely on an unsupervised analysis, without focusing on a specific gene class. Additionally, whilst transcriptomic analysis can be used as a strategy for target identification, it is not a powerful predictor of protein expression, reason why these techniques are often combined with alternative technologies, focused on analysing the cells' proteomic profile. In line with this, Singh *et al.* aimed at identifying the RUNX1-ETO target proteins that could lead to novel insights into the pathogenesis of RUNX1-ETO-induced leukaemia on a post-genomic functional level. This was achieved by inducing the expression of RUNX1-ETO in a Tet-off-inducible U937 cell line coupled with mass spectrometry analysis (Singh *et al.*, 2010). The authors showed that in these cells, the cells' protein profile is drastically changed due to the expression of RUNX1-ETO, and were able to identify several changing proteins, including NM23 and HSP27. However, several limitations are observed, as these studies have primarily focused on transformed AML cell lines, and in total transcript/protein expression profiles.

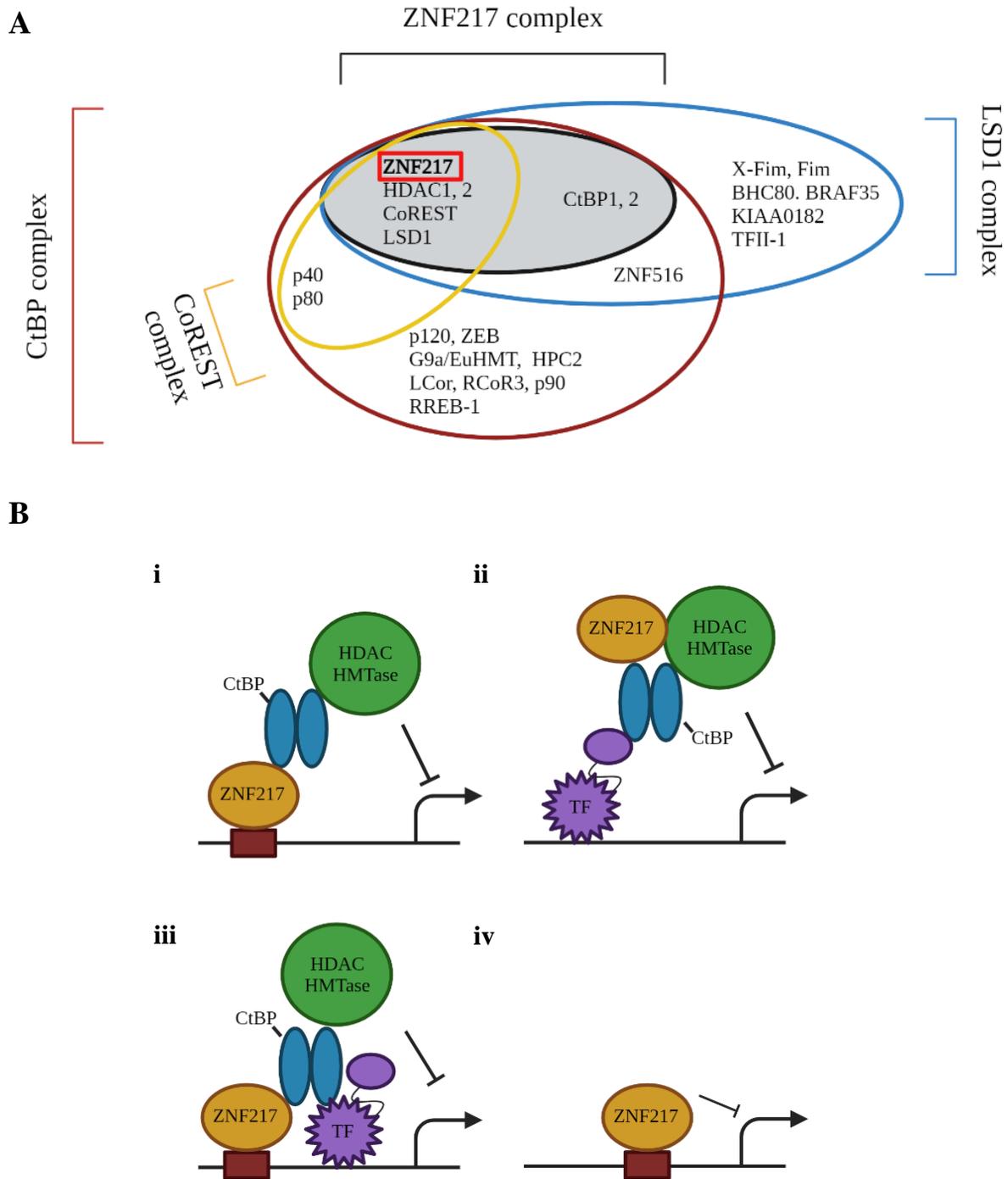
#### 1.4 ZNF217

In Chapter 3, I identified zinc finger protein 217 (*ZNF217*) as a significantly upregulated gene at the transcriptional level in RUNX1-ETO expressing CD34<sup>+</sup> HSPC (**Chapter 3**). *ZNF217* belongs to the Kruppel-like family of TFs, and contains eight C2H2 zinc finger motifs, as well as a proline-rich region (Kallioniemi *et al.*, 1994) (**Figure 1.9**). The *ZNF217* gene is located on chromosome 20q13, seen to be frequently amplified in human tumours (Collins *et al.*, 1998; Tabach *et al.*, 2011). *ZNF217* binds to DNA sequences that regulate gene expression (Cowger *et al.*, 2007; Krig *et al.*, 2007), is a component of the HDAC complex (CoREST-HDAC) (Cowger *et al.*, 2007; Thillainadesan *et al.*, 2012; Thillainadesan *et al.*, 2008), and can be found in complexes with the transcriptional co-repressor C-terminal binding protein 1 (CtBP1) (Shi *et al.*, 2003), lysine-specific histone demethylase 1A (LSD1) (Hakimi *et al.*, 2003; Lee *et al.*, 2005; Shi *et al.*, 2004) and KDM5B/JARID1B/ PLU-1, and the methyltransferases G9a and EZH2 (Quinlan *et al.*, 2007; Cowger *et al.*, 2007; Thillainadesan *et al.*, 2012; Thillainadesan *et al.*, 2008; Quinlan *et al.*, 2006) (**Figure 1.10A**).



**Figure 1.9 – Structural representation of the *ZNF217* gene and protein**

(A) Genomic organisation of the human *ZNF217* gene. Black boxes correspond to untranslated regions, whilst green boxes represent coding regions. (B) Schematic representation of the *ZNF217* protein and its functional domains. The *ZNF217* protein comprises eight C2H2 zinc finger regions and a proline-rich domain, including an NLS motif. Below are represented the main functional domains of *ZNF217*. Adapted from (Quinlan *et al.*, 2007).



**Figure 1.10 – ZNF217 mediated transcriptional repression**

(A) Schematic figure representing ZNF217-containing transcriptional repressor complexes identified through Co-IP studies. (B) Four models have been proposed for ZNF217 recruitment and function as a component of transcriptional repressor complexes: (i) ZNF217 binds DNA directly and orientates an associated repressor complex; (ii) Another TF recruits ZNF217 and its partners, resulting in transcriptional repression; (iii) ZNF217 cooperates with other TF in the binding and recruitment of corepressors; (iv) ZNF217 binds directly to DNA and represses transcription through intrinsic enzymatic activity. Adapted from (Quinlan *et al.*, 2007).

Even though ZNF217 was initially thought to act as a transcriptional co-repressor (Quinlan *et al.*, 2007), several studies have shown that it can also positively regulate the expression of specific genes (Krig *et al.*, 2007; Thillainadesan *et al.*, 2008; Krig *et al.*, 2010; Vendrell *et al.*, 2012).

The exact mechanism through which ZNF217 is able to repress gene expression is not fully understood; however, it has been speculated that this occurs through multiple mechanisms (Krig *et al.*, 2007) (**Figure 1.10B**). The first hypothesis is based on the direct binding of ZNF217 to promoters through its zinc finger domains 6 and 7, thus recruiting the CtBP co-repressor complex through direct protein-protein interactions (Krig *et al.*, 2007). The second model relies on the recruitment of other DNA binding factors responsible for mediating the interaction between the ZNF217/CtBP complex and other promoter regions (Krig *et al.*, 2007). It has also been hypothesised that ZNF217 and other TF are recruited to the same promoters and later interact with CtBP to inhibit gene expression (Krig *et al.*, 2007). The last theory relies on ZNF217 repressing gene expression by itself, through an unknown mechanism (Krig *et al.*, 2007). Furthermore, several studies have shown that ZNF217 gene expression levels do not often correlate with the corresponding protein expression, due to the interference of microRNAs (miRNA) (Li *et al.*, 2015; Bai *et al.*, 2014; Szczyrba *et al.*, 2013) and promoter methylation (Renner *et al.*, 2013; Etcheverry *et al.*, 2010).

#### 1.4.1 ZNF217 in haematopoietic development

The exact role of ZNF217 in the haematopoietic process has not been described; however, some studies have examined the role of ZNF217 in the epithelial-mesenchymal transition (EMT). This occurs during normal embryonic development and is important for lineage determination. During this process, the cell undergoes multiple biochemical changes that allow it to assume a mesenchymal phenotype, with increased migratory ability, invasiveness and increased apoptotic resistance (Kalluri and Neilson, 2003). However, this process has been shown to be involved in tumour progression with metastatic expansion, through the generation of tumour cells with stem cell properties, which play a significant role in drug resistance associated with cancer treatment (Nieto *et al.*, 2016; Lambert *et al.*, 2017; Moustakas and de Herreros, 2017).

ZNF217 has been reported as a promoter of EMT in human mammary epithelial cells (HMEC) and was found to negatively correlate with the expression of several epithelial

markers, such as E-cadherin, a direct target for ZNF217 (Krig *et al.*, 2007; Vendrell *et al.*, 2012). Furthermore, this process was followed by an increase in the expression of mesenchymal proteins and increase on the mRNA levels of TFs identified as promoters of EMT, such as *Snail1/2*, *Twist1/2* and *ZEB1/2* (Vendrell *et al.*, 2012). Additionally, in human breast cancer cells and tumours, the promoters of *Snail1* and *Snail2* genes were found to be enriched by ZNF217 (Krig *et al.*, 2007). Furthermore, overexpression of ZNF217 in the same model was shown to result in the sustained activation of the Transforming growth factor  $\beta$  (TGF- $\beta$ ) pathway through its binding to the *TGFB2* and *TGFB3* promoters (Vendrell *et al.*, 2012). Conversely, inhibition of the TGF- $\beta$  pathway resulted in the reversal of the ZNF217-induced EMT, suggesting that this process incorporates the transcriptional repression of E-cadherin coupled with the constitutive activation of the TGF- $\beta$ -activated signalling pathway (Vendrell *et al.*, 2012). In late-stage tumours, several studies have described a redirection of TGF- $\beta$  signalling from suppressing cell proliferation to EMT activation (Pirozzi *et al.*, 2011).

#### 1.4.2 ZNF217 as an oncogene

Due to its localisation, ZNF217 has been extensively studied as a potential oncogene and biomarker of disease. In fact, several studies have shown that ZNF217 has the ability to interfere with multiple processes associated with the ‘hallmarks of cancer’. These include sustained proliferative signals, evasion from growth suppressors, replicative immortality, resistance to cell death, cancer stem cell enrichment, and activation of invasion and metastasis (Hanahan and Weinberg, 2011; Hanahan and Weinberg, 2000). Moreover, several tumours have shown to possess increased copies of the *ZNF217* gene (Plevova *et al.*, 2010; Fang *et al.*, 2010; Rahman *et al.*, 2012), in variable frequency according to the type of tumour, and have been linked to a poorer outcome in some studies (Quinlan *et al.*, 2007; Peiró *et al.*, 2002; Rooney *et al.*, 2004; Ginzinger *et al.*, 2000). Even though several roles have been proposed for ZNF217 in the development of different cancers, including breast and ovarian, no relationship has been established between the role of this TF and leukaemogenesis.

##### 1.4.2.1 Proliferation

One of the most fundamental characteristics of cancer cells relates to the cells ability to sustain chronic proliferation. ZNF217 has been described as having a major role in this process through the dysregulation of signals which allow progression through cell cycle and growth,

as well as through the disruption of anti-proliferative signalling. In breast and ovarian cancer cells, overexpression of *ZNF217* was shown to promote cell proliferation (Li *et al.*, 2014; Thollet *et al.*, 2010), whilst silencing it led to a reduction in cell growth in prostate, colorectal, ovarian and breast cancer cells (Szczyrba *et al.*, 2013; Li *et al.*, 2015; Rahman *et al.*, 2012; Sun *et al.*, 2008a; Thollet *et al.*, 2010). *In vivo* studies agreed with these observations, as constitutive expression of *ZNF217* in breast and ovarian cancer cells promoted the growth and rate of tumour formation in mice (Vendrell *et al.*, 2012; Li *et al.*, 2014; Thollet *et al.*, 2010; Littlepage *et al.*, 2012). This was associated with a substantial increase in the proportion of cells in S phase (Li *et al.*, 2014), coupled with the abnormal expression of several cyclins and, post-translationally, through the upregulation of Aurora kinase A (Aurora-A) (Thollet *et al.*, 2010). One of the genes thought to play a role in *ZNF217*-induced proliferation is *ErbB3*, positively regulated by the TF, both *in vivo* and in human breast cancer cells (Krig *et al.*, 2010). Furthermore, in normal HMEC and in breast cancer cell lines, ectopic expression of *ZNF217* induced *ErbB3* protein overexpression (Krig *et al.*, 2010; Vendrell *et al.*, 2012), as well as its heterodimer *ErbB2*. Both *ErbB* proteins function as a ‘oncogenic unit’ and result in the activation of the PI3K/Akt and MAPK survival pathways (Krig *et al.*, 2010; Littlepage *et al.*, 2012).

An additional mechanism for the sustained proliferative ability of *ZNF217*-overexpressing cells has been proposed, involving the p15<sup>ink4b</sup> tumour suppressor, which inhibits cell cycle progression at the G<sub>1</sub>/S transition (Kim and Sharpless, 2006). The *ZNF217*/CoREST complex is able to inhibit p15<sup>ink4b</sup> activity through the recruitment of the methyltransferase DNMT3A, which in turn leads to promoter hypermethylation (Thillainadesan *et al.*, 2012). By preventing the recruitment of co-factors responsible for the demethylation of p15<sup>ink4b</sup>, overexpression of *ZNF217* results in the block of the anti-proliferative TGF- $\beta$  signalling pathway (Thillainadesan *et al.*, 2012), shown to promote malignancy (Massagué, 2008; Ikushima and Miyazono, 2010).

#### 1.4.2.2 Replicative immortality

One of the first studies that supported the hypothesis that *ZNF217* may function as an oncogene relied on the transduction of finite life-span HMEC with the *ZNF217* gene, giving rise to an immortalised cell line, with increased telomerase activity and resistance to TGF- $\beta$ -induced anti-proliferative action (Nonet *et al.*, 2001). Subsequent studies demonstrated that, in an SV40 Tag/tag expressing, p53/pRB-deficient, human ovarian surface epithelial cell line

(IOSE), ZNF217 overexpression promoted neoplastic progression, associated with high telomerase activity and genomic changes linked to ovarian carcinomas (Li *et al.*, 2007). The oncogenic translation elongation factor *EEF1A2* has been shown to be upregulated in ZNF217-overexpressing IOSE cells (Li *et al.*, 2007; Abbas *et al.*, 2015), and its overexpression in the same cells was shown to promote neoplastic progression (Sun *et al.*, 2008b). Additionally, both immortalised HMEC and IOSE cells showed a stabilization of telomere length, with increased activity, features associated with bypassing senescence (Nonet *et al.*, 2001; Li *et al.*, 2007). By overexpressing ZNF217, cells exhibited a decrease in the normal apoptotic rate, as a consequence of functionally compromised telomeres (Huang *et al.*, 2005). These observations suggest that ZNF217 acts by promoting the de-repression of telomerases, thus allowing pre-malignant cells to overcome senescence due to telomere dysfunction (Nonet *et al.*, 2001).

#### 1.4.2.3 Drug resistance

Common therapeutic treatment options often include the combination of different drugs for enhanced effect. These can include doxorubicin, which inhibits topoisomerase II and induces double-strand DNA breaks, leading to ATM-dependent p53-mediated apoptosis, or taxanes, such as paclitaxel, which lead to cell cycle arrest and apoptosis through microtubule-stabilising processes. Ectopic expression of ZNF217 has been shown to contribute to both doxorubicin and paclitaxel resistance in breast cancer cells, whilst silencing it resulted in the opposite effect, with increased drug sensitivity (Thollet *et al.*, 2010; Huang *et al.*, 2005). This is thought to occur through the interference of apoptotic signals induced by the drugs, due to ZNF217 expression (Thollet *et al.*, 2010; Huang *et al.*, 2005). In fact, ZNF217-mediated resistance to paclitaxel in breast cancer cells resulted in significant changes in intrinsic mitochondrial apoptotic pathways, associated with the dysregulation of the anti-apoptotic proteins BCL-2 and BCL-xL, and the pro-apoptotic proteins Bad, Bak and Bax (Thollet *et al.*, 2010). Moreover, the anti-apoptotic potential of the ZNF217 protein has also been found to be mediated through the PI3K/Akt pathway, as ectopic expression of this protein led to the activation of the pathway, whilst silencing ZNF217 resulted in decreased Akt phosphorylation (Krig *et al.*, 2010; Huang *et al.*, 2005).

#### 1.4.2.4 Cell differentiation

Identically to LSCs, cancer stem cells (CSCs) possess multilineage differentiation ability, as well as an increased self-renewal potential, and are thought to be the promoters of

tumorigenesis and metastases (Nguyen *et al.*, 2012). Several studies have shown that ZNF217 binds to the promoter region of several genes involved in the differentiation process and organ development, resulting in a block in cell differentiation, as well as CSC maintenance (Krig *et al.*, 2007; Vendrell *et al.*, 2012; Littlepage *et al.*, 2012). In primary mammary epithelial cells, as well as in breast cancer cells, overexpression of ZNF217 led to an increase in the formation of mammospheres with self-renewal potential (Vendrell *et al.*, 2012; Littlepage *et al.*, 2012; Nguyen *et al.*, 2014). Moreover, this was associated with a repression of adult stem cell expression signature, frequently found downregulated in cancer (Littlepage *et al.*, 2012). Lastly, following treatment with retinoic acid, a pro-differentiation agent, ZNF217 expression was repressed in embryonal cells, suggesting that, in differentiated adult cells, aberrant expression of ZNF217 might suppress differentiation, thus leading to tumorigenesis (Krig *et al.*, 2007).

Glioma stem cells (GSCs) are thought to be the driving population leading to the development of glioblastoma multiforme, the most common and aggressive type of primary brain tumour (Singh *et al.*, 2004). These cells have been shown to significantly upregulate ZNF217 expression, as compared to non-GSCs (Mao *et al.*, 2011). Moreover, downregulation of ZNF217 was observed upon GSCs forced differentiation, whilst knocking it down in these cells led to a decrease in cell proliferation and reduction in the stem-like cell population (Mao *et al.*, 2011). A possible mechanism for this is through the involvement of Aurora-A, suggested to regulate self-renewal potential through the stabilization and activation of Wnt/ $\beta$ -catenin signalling (Xia *et al.*, 2013). Additionally, the GSC niche requires hypoxic conditions, and ZNF217 has been demonstrated to be regulated by several HIF both through direct and indirect mechanisms (Mao *et al.*, 2011). These observations suggest that ZNF217 may promote the hypoxia-induced stemness observed in GSCs through its role as a downstream target of HIF.

## 1.5 C/EBP $\beta$

In Chapter 3, I identified the C/EBP $\beta$  protein as a significantly downregulated protein in the nuclear compartment of RUNX1-ETO expressing CD34<sup>+</sup> HSPC (**Chapter 3**). The C/EBP family comprises six structurally and functionally homologous TFs (C/EBP $\alpha$ , C/EBP $\beta$ , C/EBP $\delta$ , C/EBP $\gamma$ , C/EBP $\epsilon$ , and CHOP) responsible for regulating several physiological processes, including cell proliferation, differentiation, inflammation and metabolism, as well

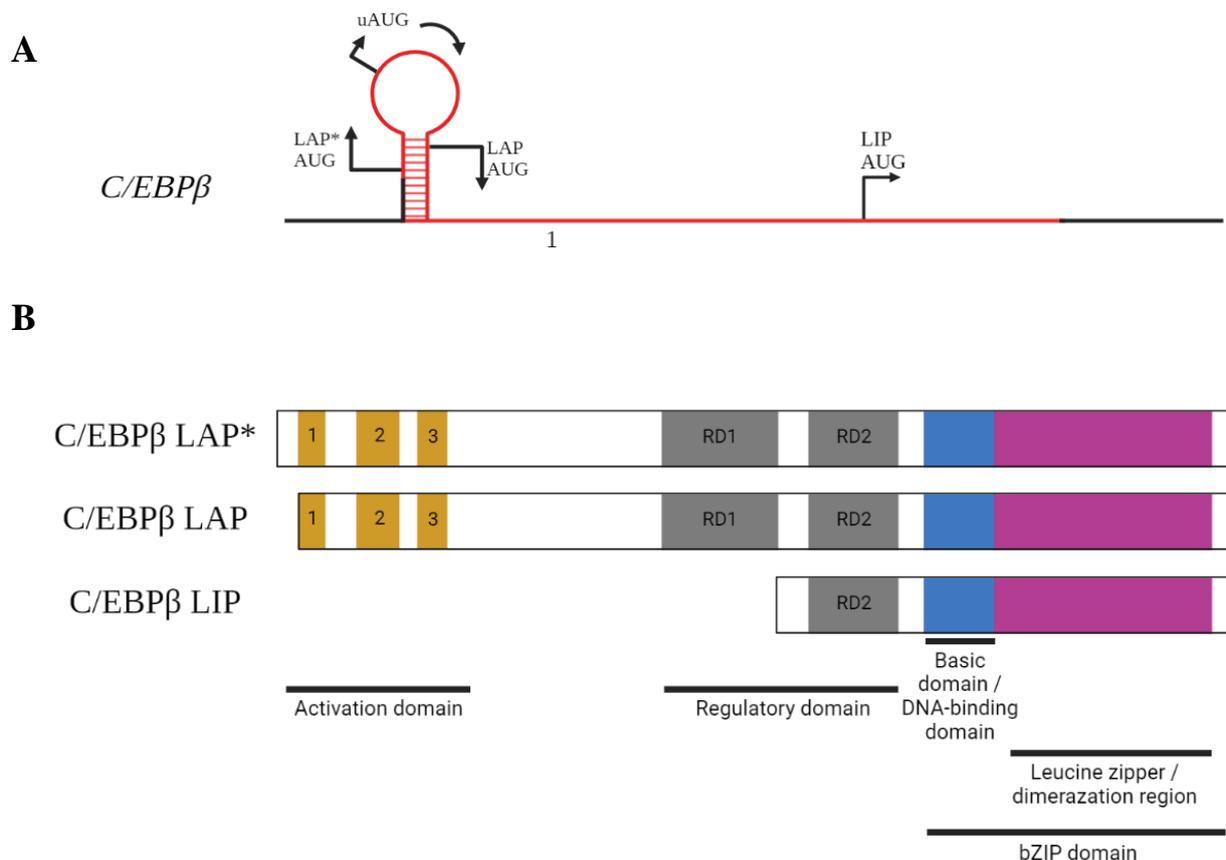
as oncogene-induced senescence and tumorigenesis (Cao *et al.*, 1991; Diehl, 1998; Poli, 1998; Zahnw, 2002; Ramji and Foka, 2002; Sebastian and Johnson, 2006). All members of this family share a highly conserved C-terminal leucine-zipper dimerization domain (bZIP) involved in DNA binding (Ramji and Foka, 2002), whilst the N-terminal region is less conserved, except for three activation domains, common to all C/EBP family members (Kowenz-Leutz *et al.*, 1994; Williams *et al.*, 1995; Angerer *et al.*, 1999; Williamson *et al.*, 1998), responsible for promoting interactions with transcriptional co-activators (Mink *et al.*, 1997).

The first protein to be identified was C/EBP $\alpha$ , isolated from rat liver and suggested to play an essential role in adipocyte differentiation (Johnson *et al.*, 1987). With time, other C/EBP proteins were identified and attributed to multiple processes. C/EBP $\beta$  was first identified in 1990 and was initially described as a bZIP structured factor binding to the IL-1-reponsive element within the IL-6 promoter (Akira *et al.*, 1990). C/EBP $\beta$  is highly expressed across several tissues, including liver, lung, spleen and kidneys, as well as in myelomonocytic cells and macrophages (Williams *et al.*, 1991; Katz *et al.*, 1993; Haas *et al.*, 2010; Gutsch *et al.*, 2011). Several external factors can contribute to the activation or inhibition of C/EBP $\beta$ , including differentiation- or proliferation-inducing agents, hormones, cytokines and inflammatory molecules, through the activation of specific pathways (Ramji and Foka, 2002). Moreover, several additional mechanisms can contribute to C/EBP $\beta$ 's expression, including transcriptional processes, mTOR-mediated alternative translation, post-translational modifications, and protein-protein interactions (Zahnw, 2009; Nerlov, 2007; Tsukada *et al.*, 2011). The expression of cytokines or chemokines (and their respective receptors), pro-inflammatory genes, pro-proliferative or differentiation-related genes, as well as metabolic enzymes are regulated through the activation of C/EBP $\beta$  (Ramji and Foka, 2002). Subsequently, several cellular functions are influenced by C/EBP $\beta$  expression, including proliferation (Haas *et al.*, 2010; Gutsch *et al.*, 2011), differentiation (Katz *et al.*, 1993; Gutsch *et al.*, 2011; Pham *et al.*, 2007), metabolic regulation (Liu *et al.*, 1999; Croniger *et al.*, 2001), and orchestration of the immune response (Poli, 1998; Zwergal *et al.*, 2006; Cappello *et al.*, 2009). Additionally, C/EBP $\beta$  is also involved in the pathogenesis of several diseases, such as cancer hyper- or hypo-inflammation, and bacterial/viral infections (Ramji and Foka, 2002; Liu *et al.*, 2009).

### 1.5.1 Structure and function

*CEBPB* mRNA expression is regulated by several TF, through its TATA box and different binding sites present in its promotor region, including C/EBP $\beta$  itself (Chang *et al.*, 1995; Niehof *et al.*, 2001a; Foka *et al.*, 2001), STAT3 (Niehof *et al.*, 2001b), SP1 (Berrier *et al.*, 1998), members of the cAMP responsive element binding protein (CREB)/activating TF family (Berrier *et al.*, 1998; Niehof *et al.*, 1997), early growth response 2 (EGR2) (Chen *et al.*, 2005), Fos related antigen 2 (Fra-2) (Chang *et al.*, 2004), sterol-regulatory element-binding protein 1c (SREBP1c) (Le Lay *et al.*, 2002), myoblastosis TF (Myb) (Mink *et al.*, 1999), and RAR $\alpha$  (Duprez *et al.*, 2003). *CEBPB* is an introneless gene, consisting of one single exon and located in chromosome 20q13.13, coding for the production of a single mRNA molecule (Calkhoven *et al.*, 2000; Xiong *et al.*, 2001) (**Figure 1.11A**). However, due to the presence of three initiation codons on different sites (Calkhoven *et al.*, 2000; Duprez *et al.*, 2003), it can generate three protein isoforms: liver-enriched activating protein\* (LAP\* or full-length [FL]), liver-enriched activating protein (LAP) and liver-enriched inhibitory protein (LIP) (Zahnow, 2009; Sears and Sealy, 1994; Condamine *et al.*, 2015) (**Figure 1.11B**). LAP\* is a 44 kDa protein consisting of 345 amino acids; LAP is slightly shorter, with 42 kDa and 322 amino acids, whilst LIP represents a truncated version of C/EBP $\beta$ , with 20 kDa and 145 amino acids (Zahnow, 2009; Sears and Sealy, 1994).

As described above (**1.5**), the C/EBP $\beta$  protein structure is characterised by an N-terminal region containing up to three transactivation domains (TAD) and two regulatory domains (RD1 and RD2) (Ramji and Foka, 2002; Williams *et al.*, 1995). On the C-terminal region, C/EBP $\beta$  contains a basic DNA-binding domain adjacent to the bZIP dimerization domain (Ramji and Foka, 2002; Zahnow, 2009; Lekstrom-Himes and Xanthopoulos, 1998; Wedel and Ziegler-Heitbrock, 1995; Tsukada *et al.*, 2011). Both LAP\* and LAP contain all three TAD and both regulatory domains, whilst LIP lacks all TAD regions, possesses a full-length RD2 and a truncated form of RD1 (Ramji and Foka, 2002; Williams *et al.*, 1995). Several factors influence dimerization, cellular localization, DNA binding, and the transactivation activity of C/EBP $\beta$ , such as post-translational modifications, phosphorylation, acetylation, methylation, sumoylation, and proteolysis (Ramji and Foka, 2002; Zahnow, 2009; Tsukada *et al.*, 2011; Hattori *et al.*, 2003).



**Figure 1.11 – Structural representation of the *C/EBPβ* gene and protein isoforms**

(A) Genomic organisation of the human *CEBPB* gene. A cis-regulatory uORF in the single exon *CEBPB* transcript regulates initiation of translation at conserved in-frame start sites to generate three *C/EBPβ* isoforms. (B) Schematic representation of the three *C/EBPβ* protein isoforms and their functional domains. LAP\* and LAP isoforms contain three transactivation domains, as well as two regulatory regions RD1 and RD2. LIP, on the other hand, lacks the transactivation domain, and only contains the regulatory region RD2. All three isoforms possess a DNA binding domain, couples with a leucine zipper region, making up the bZIP domain. Adapted from (Zahnow, 2002).

Moreover, the different C/EBP $\beta$  isoforms themselves can regulate the protein's transactivation ability, through the presence (LAP\* and LAP) or absence (LIP) of intrinsic transactivation potential or through the appropriate dimerization/interaction partners (Ramji and Foka, 2002; Zahnw, 2009; Tsukada *et al.*, 2011). Furthermore, the ratio of LIP to LAP\*/LAP is an important factor in C/EBP $\beta$  function (Descombes and Schibler, 1991; Ossipow *et al.*, 1993; Li *et al.*, 2008) (1.5.3). Additionally, C/EBP $\beta$  target gene activation also relies on the composition of target gene promoters and localisation (Ramji and Foka, 2002; Zahnw, 2009; Tsukada *et al.*, 2011).

LAP\*, LAP and LIP possess distinctive functions. Regarding 'pure' transactivation activity, LAP\* is a weaker activator of gene transcription, as compared to LAP, due to the formation of an additional disulphide bond (Williams *et al.*, 1995; Su *et al.*, 2003). However, upon the recruitment of the switch/sucrose nonfermentable' (SWI/SNF) nucleosome remodelling complex by its N-terminal domain, LAP\* promotes not only the activation of genes otherwise silenced, but also the interaction with the myeloid-specific TF c-Myb (Kowenz-Leutz and Leutz, 1999). Through the interaction with other TF and regulators, LAP\* represents a more effective (co)regulator of myeloid genes than LAP (Kowenz-Leutz and Leutz, 1999). LIP, on the other hand, is considered a dominant inhibitor of transcriptionally active C/EBPs, including LAP\* and LAP (Descombes and Schibler, 1991). However, a recent study has demonstrated the ability of LIP to act compensatively in mice lacking the *CEBPB* gene, suggesting that the classification of LIP as a dominant inhibitor relies on temporal and cellular contexts, in the adipogenic differentiation process (Bégay *et al.*, 2018). However, the exact mechanism underlying this process remains to be understood.

C/EBP $\beta$  further possesses the ability to interact with proteins outside the C/EBP family, thus increasing the accessible DNA binding motifs and changing C/EBP protein functions and specificities (Tsukada *et al.*, 2011). These include other bZIP-structured TF, such as members of the Jun/Fos (Hsu *et al.*, 1994) and the CREB/ATF families (Vallejo *et al.*, 1993), as well as non-bZIP proteins, including nuclear factor- $\kappa$ B (NF- $\kappa$ B) subunits p65 (Zwergal *et al.*, 2006) and p50 (LeClair *et al.*, 1992), glucocorticoid receptor (Savoldi *et al.*, 1997), E2F proteins (Sebastian *et al.*, 2005; Johnson, 2005), c-Myb (Kowenz-Leutz and Leutz, 1999), PRMT4 (Kowenz-Leutz *et al.*, 2010), CBP (Guo *et al.*, 2001), PU.1 (Pham *et al.*, 2007; Tissières *et al.*, 2009), death-associated protein 6 (Daxx) (Wethkamp and Klempnauer, 2009), in addition to CBF and RUNX proteins (Tahirov *et al.*, 2001).

### 1.5.2 C/EBP $\beta$ in haematopoietic development

C/EBP $\beta$  is highly expressed in myelomonocytic cells, as well as monocytes and macrophages (1.5). In these cells, C/EBP $\beta$  controls several processes such as proliferation and differentiation through the activation or repression of specific target genes. In myelomonocytic cells, C/EBP $\beta$  is most importantly responsible for regulating proliferation and cell cycle progression (Friedman, 2007; Sato *et al.*, 2020). In HSCs transduced with a FLT3-wild type receptor, FLT3 ligand induced the expression of the truncated C/EBP $\beta$  protein LIP, thus leading to a decrease in the LAP/LIP ratio, associated with proliferation (Haas *et al.*, 2010). Furthermore, LIP-deficient murine embryonic fibroblasts show a significant reduction in cell duplication and a decrease in the expression of the proliferative marker's cyclins A1, A2, B1, E1 and E2 (Wethmar *et al.*, 2010). The consequences arising from the absence of C/EBP $\beta$  on the proliferative process, however, are time-dependent, as cells in early developmental stages respond differently to more differentiated cells. For instance, absence of C/EBP $\beta$  in murine BM-derived progenitor cells, led to a significant decrease in the number of myeloid colonies, grown in the presence of G-CSF, GM-CSF and IL-3, and in the number of generated cells, as compared to control cells (Hirai *et al.*, 2006). In more differentiated cells from the same model, on the other hand, C/EBP $\beta$  KO increased cell proliferation; *in vivo*, C/EBP $\beta$ -KO mice presented hyperplastic haematopoiesis and hypermyeloproliferation (Screpanti *et al.*, 1995) whilst *in vitro* studies showed that macrophages lacking the *CEBPB* gene display an increased proliferative rate, as well as enhanced proportion of cells exhibiting S or G<sub>2</sub>/M cell cycle markers (Gutsch *et al.*, 2011). Although usually regarded as transcriptional activators, LAP\* and LAP are thought to be the mediators of this process by promoting the repression of c-Myc in monocytes (Gutsch *et al.*, 2011; Zhang *et al.*, 2011a). By doing this, C/EBP $\beta$  represses the transcription of its pro-proliferative target genes, such as cyclin D, whilst promoting the expression of cell cycle inhibitors, such as p27, usually repressed by c-Myc (Gutsch *et al.*, 2011). Subsequent studies have suggested that C/EBP $\beta$  interacts with the RB-E2F protein complex, leading to RB-dependent cell cycle arrest and repression of E2F target genes (Sebastian *et al.*, 2005; Johnson, 2005). In summary, in differentiated monocytic cells, the predominant expression of the LAP\* and/or LAP results in an arrest in cell cycle, specifically in the G<sub>0</sub>/G<sub>1</sub> phase, and a reduced proliferative ability (Gutsch *et al.*, 2011).

C/EBP $\beta$  isoforms have previously been shown to contribute to monocytic differentiation (Friedman, 2007). This process is characterised by both a significant upregulation of the LAP\* and LAP isoforms (Katz *et al.*, 1993; Gutsch *et al.*, 2011; Pham *et al.*, 2007; Natsuka *et al.*, 1992; Pan *et al.*, 1999; Ji and Studzinski, 2004; Zhang *et al.*, 2011a) and an increase in the LAP/LIP ratio (Gutsch *et al.*, 2011) is also observed in the differentiation process of other cells, including hepatocytes (Descombes and Schibler, 1991; Buck *et al.*, 1994) and adipocytes (Zhu *et al.*, 2007; Vigilanza *et al.*, 2011). Additionally, C/EBP $\beta$  is involved in the regulation of the expression of several differentiation-associated genes, including CD14 (Pan *et al.*, 1999; Ji and Studzinski, 2004; Zhang *et al.*, 2011a; Xu *et al.*, 2008), Fc $\gamma$  receptor II (Fc $\gamma$ RII) (Gorgoni *et al.*, 2002), monocyte-specific esterase (MSE) (Ji and Studzinski, 2004), 1 $\alpha$ -hydroxylase (Stoffels *et al.*, 2006), and the cytoplasmic proline-rich tyrosine kinase 2 (Pyk2) (Park *et al.*, 2008). Moreover, C/EBP $\beta$  has been shown to interact with PU.1 as a transcriptional cofactor (Pham *et al.*, 2007; Tissières *et al.*, 2009). At specific time-points, LAP\* and LAP have the potential to promote c-Myc-mediated transcription by facilitating chromatic opening, through the recruitment of histone acetylating cofactors of the p300/CBP family (by LAP\* and LAP) and/or the SWI/SNF chromatin remodelling complex (exclusively by LAP\*) (Kowenz-Leutz *et al.*, 2010; Plachetka *et al.*, 2008). As a consequence of the activation of differentiation-associated genes, there is a decrease in the cells' proliferative rate, coupled with morphological changes (Gutsch *et al.*, 2011; Zhang *et al.*, 2011a) and increased antimicrobial ability (Zhang *et al.*, 2011a). Even though C/EBP $\beta$ -KO mice have the ability to give rise to macrophage-like cells, these showed a reduced functional potential, since their response to external stimuli is impaired (Screpanti *et al.*, 1995; Tanaka *et al.*, 1995a). Subsequently, macrophages derived from C/EBP $\beta$ -KO mice shown a reduction in the expression of differentiation markers (Koschmieder *et al.*, 2009), as well as atypical morphological characteristics (Gutsch *et al.*, 2011). Moreover, C/EBP $\beta$  has been suggested to be required for the survival of peripheral blood monocytes, whilst it's not necessary for the development and maintenance of mature cells of other lineages at steady state (Tamura *et al.*, 2015a).

In addition to monocytic development, LAP\* and LAP have been shown to promote granulocytic differentiation (Hirai *et al.*, 2006; Popernack *et al.*, 2001). In murine primary BM cells, C/EBP $\beta$  overexpression resulted in a reduction in the proportion of myeloid progenitor cells and promoted granulocytic differentiation (Popernack *et al.*, 2001). Moreover, in response to external stimuli, C/EBP $\beta$  has the ability to promote both proliferation and differentiation of

HSPC to supply granulocytes in demand (Satake *et al.*, 2012). In the absence of C/EBP $\beta$ , however, mice infected with *C. albicans* showed a reduced response in terms of emergency granulopoiesis (Hirai *et al.*, 2006). Moreover, in patients with congenital neutropenia, addition of G-CSF led to a significant upregulation of C/EBP $\beta$  in BM-derived myeloid cells, resulting in urgent granulocytic development (Skokowa and Welte, 2009). Osteoclast formation, on the other hand, is regulated by LIP, but inhibited by LAP\* or LAP, through the regulation of the osteoclastogenesis inhibitor MafB (Smink *et al.*, 2009; Smink *et al.*, 2012). Similarly, uORF-deficient mice incapable of inducing LIP expression display increased levels of MafB, resulting in impaired osteoclastogenesis (Wethmar *et al.*, 2010). Recent studies have shown that C/EBP $\beta$  is upregulated in HSPC as a response to stress (Sato *et al.*, 2020), and this is essential for the cells to regenerate, thus meeting the increasing myeloid cell demand. Furthermore, the expression of the different isoforms is dependent on cell state: first, upregulation of LIP in an early regenerative stage allows the proliferation of LT-HSC and their differentiation into MPP, whilst LAP\*/LAP induce subsequent myeloid differentiation at a later stage of regeneration (Sato *et al.*, 2020).

### 1.5.3 C/EBP $\beta$ as an oncogene

Due to their ability to block cell growth and respond to DNA damage, C/EBP proteins are generally identified as tumour suppressor factors. However, upon certain stimuli and depending on cell type and isoform present, these can elicit opposite effects. C/EBP $\beta$  function highly resembles that of C/EBP $\alpha$ , since it has been shown to promote cell differentiation, suppress tumorigenesis and block proliferation. When not required, C/EBP $\beta$  is maintained in a latent state through the action of auto-inhibitory factors aimed at suppressing its DNA binding and transactivation functions (Williams *et al.*, 1995; Kowenz-Leutz *et al.*, 1994).

#### 1.5.3.1 Cell survival

The role of C/EBP $\beta$  in cancer is thought to be mediated through its role in regulating cell survival and apoptosis. For instance, upon induced DNA damage, C/EBP $\beta$  plays an essential role in the survival process of hepatic stellate cells (Buck *et al.*, 2001) and in macrophages, C/EBP $\beta$  is required for survival in response to Myc-Raf transformation (Wessells *et al.*, 2004). Moreover, C/EBP $\beta$ -induced cell survival in response to DNA damage has been shown to be

mediated through a reduction in p53 expression and activity (Yoon *et al.*, 2007; Ewing *et al.*, 2008).

In addition to its involvement in apoptosis, several studies have showed that C/EBP $\beta$  has a role in oncogene-induced senescence (OIS). Under normal circumstances, senescence is characterised by a state of irreversible growth arrest that can act as a barrier to malignant transformation. Forced expression of C/EBP $\beta$  LAP has been shown to result in cell cycle arrest in hepatocarcinoma cells (Buck *et al.*, 1994), keratinocytes (Zhu *et al.*, 1999) and fibroblasts (Johnson, 2005). Moreover, cooperation between C/EBP $\beta$  and RB-E2F proteins has been shown to result in an irreversible cell cycle arrest at the G<sub>1</sub>/S phase, through Ras-induced cellular senescence (Sebastian *et al.*, 2005). Lastly, several C/EBP $\beta$ -induced pro-inflammatory cytokines and chemokines have been shown to contribute to OIS (Acosta *et al.*, 2008; Kuilman *et al.*, 2008). An additional association between Ras and C/EBP $\beta$  was described as studies showed that in the context of C/EBP $\beta$  deficiency, v-HA-Ras transgenic mice showed reduced skin-tumorigenic potential, suggesting that C/EBP $\beta$  plays an oncogenic role in the downstream Ras signalling pathway (Poli, 1998). Moreover, gene expression analysis in human cancers suggested that C/EBP $\beta$  is involved in cyclin D1-induced oncogenic signature (Zahnow *et al.*, 1997). Increased expression of C/EBP $\beta$ , and its association with oncogenic development has been reported in several cancer types, including breast, ovarian, colorectal renal and gastric tumours. However, the exact role of C/EBP $\beta$  in the regulation of survival, apoptosis and senescence is highly context specific.

#### 1.5.3.2 Tumour aggressiveness

In general, few mutations have been identified in C/EBP $\beta$ , and those identified are not thought to contribute to the development of epithelial cancers (Vegesna *et al.*, 2002). However, the region in which the C/EBP $\beta$  gene is found has been shown to be amplified in a small proportion of human cancers, and it has been associated with lobular carcinoma *in situ* of the breast (Mastracci *et al.*, 2006). An increase in *CEBPB* mRNA levels results in an increase in protein translation, associated with a rise in C/EBP $\beta$  isoform expression and the imbalance of the LAP/LIP ratio, with the upregulation of the LIP isoform, observed in oestrogen-receptor-negative, aneuploid, highly proliferative breast tumours, with poor prognosis (Zahnow *et al.*, 1997; Milde-Langosch *et al.*, 2003). Dysregulation in the normal LAP/LIP balance has also been related to a defective TGF- $\beta$ -dependent cyostatic response in metastatic breast cancer

(Gomis *et al.*, 2006). However, forced expression of LAP\*/LAP, thus normalising the LAP/LIP ratio was able to restore TGF- $\beta$  cyostatic response and reduce the proliferative ability of metastatic cells (Gomis *et al.*, 2006).

The increase in *CEBPB* levels has been observed in a more-aggressive subset of tumours, associated with metastatic breast cancer (van de Vijver *et al.*, 2002), a high tumour grade (van 't Veer *et al.*, 2002; Ma *et al.*, 2004; Finak *et al.*, 2008) and overall poorer prognosis (van de Vijver *et al.*, 2002), as compared to less-aggressive tumours. Additionally, the expression of C/EBP $\beta$  was found to be heightened in malignant ovarian tumours, as compared to borderline or benign tumours (Sundfeldt *et al.*, 1999). These observations suggest that the transcriptional control or regulation of mRNA stability may influence *CEBP $\beta$*  expression in more aggressive tumour types.

#### 1.5.3.3 Drug resistance

C/EBP $\beta$  has been suggested to contribute to multi-drug resistance in breast cancer (Combates *et al.*, 1994; Conze *et al.*, 2001; Chen *et al.*, 2004), through the P-glycoprotein transporter, encoded by *MDR1*, further associated with poor prognosis (Leonessa and Clarke, 2003). Further, the *MDR1* gene is regulated by C/EBP $\beta$  in HepG2 hepatoma cells (Combates *et al.*, 1994) and in MCF7 breast cancer cells (Conze *et al.*, 2001). Similar observations were made in ovarian cancer, in which C/EBP $\beta$  was shown to promote the expression of drug resistance genes (Liu *et al.*, 2018), through its interaction with the telomeric silencing 1-like (DOT1L) protein. Inhibition or KD of DOT1L led to a decrease in the promotion of the expression of drug resistance genes by C/EBP $\beta$ , and actually reversed cisplatin resistance (Liu *et al.*, 2018). These observations suggest that expression of C/EBP $\beta$  influences treatment outcome and the development of drug-resistance events.

### 1.5.4 C/EBP $\beta$ in leukaemogenesis

Of all the members of the C/EBP family, the role of C/EBP $\alpha$  in the leukaemogenic process has by far been the most extensively studied (Nerlov, 2004; Koschmieder *et al.*, 2009). Nonetheless, C/EBP $\beta$  has been shown to be related to the development of several myelomonocytic leukaemias (Wethmar *et al.*, 2010; Sebastian and Johnson, 2006; Duprez, 2004). C/EBP $\beta$  has been shown to be dysregulated in FLT3-ITD-mutated AML patients, with an increased concentration of LIP observed in these patients (Haas *et al.*, 2010). Accordingly,

an increased LIP expression has been shown in several ITD-positive monocytic cell lines, including MV4;11, MOLM-13 and PL21 (Haas *et al.*, 2010). Furthermore, LIP is also highly expressed in the erythroleukaemia-like cell line, K562 (Wall *et al.*, 1996), as well as in the anaplastic large cell lymphoma cell lines SUDHL-1, Ki-JK, and Karpas 299 (Quintanilla-Martinez *et al.*, 2006). LAP\* and LAP, on the other hand, are generally reduced in rapidly proliferating cells, such as in BM-derived cells from CML patients in blast crisis (Guerzoni *et al.*, 2006).

Upregulation of LAP\* or LAP has been shown to reduce cell proliferation and induce the differentiation of leukaemic cell lines, and suppressing leukaemogenesis in *in vivo* models (Guerzoni *et al.*, 2006). In the NB4 AML cell line, upregulation of LAP\*/LAP following ATRA treatment resulted in the differentiation of less developed myelomonocytic cells into granulocytes (Duprez *et al.*, 2003). However, this was not observed in ATRA-resistant cells, in which it was not possible to induce the expression of any C/EBP $\beta$  isoform (Duprez *et al.*, 2003). Moreover, treatment of primary AML cells with deltanoids (Studzinski *et al.*, 2005), U937 cells with vitamin D3 (Koschmieder *et al.*, 2007) or ATRA (Chen *et al.*, 2009), and HL-60 cells with 1,25-Dihydroxyvitamin D3 [1,25(OH) $_2$ D $_3$ ] (Studzinski *et al.*, 2005) also resulted in an increase in C/EBP $\beta$  protein levels associated with the differentiation of these cells into monocytes. Similarly, expression of C/EBP $\beta$  in the BCR-ABL cell line 32D was found to be repressed, as this fusion protein negatively regulates C/EBP $\beta$  (Guerzoni *et al.*, 2006). However, this was restored upon treatment with imatinib, following which C/EBP $\beta$  was able to induce differentiation and inhibit the proliferation of these cells (Guerzoni *et al.*, 2006). Subsequently, treatment of AML patients with vitamin D3 or derivatives might represent an alternative therapeutic approach for patients that do not respond to classical chemotherapeutic agents, through the induction of C/EBP $\beta$ -supported cell differentiation (Hughes *et al.*, 2010). However, the role of C/EBP $\beta$  in the development of t(8;21) leukaemia has not been determined.

## 1.6 Aims and objectives

The t(8;21) translocation, resulting in the expression of the fusion protein RUNX1-ETO, is a common chromosomal translocation in AML. Even though it is associated with a favourable prognosis, the exact mechanism underlying the leukaemogenic process remains unknown.

Identifying new therapeutic targets is therefore important if we are to progress the treatment of patients with AML.

Previous studies performed within our group have shown that RUNX1-ETO is able to inhibit the development of haematopoietic cells, promote the growth of immature cells and increase the cells' self-renewal ability (Tonks *et al.*, 2003; Tonks *et al.*, 2004). Transcriptomic analysis of these cells identified hundreds of genes impacted by RUNX1-ETO expression (Tonks *et al.*, 2007). However, this analysis was performed using an unsupervised approach, focusing on the highest changes observed in RUNX1-ETO expressing cells. Given that transcriptional regulation is likely mediated by changes in TF expression, this study will re-analyse transcriptomic data obtained from Tonks *et al.*, and focus specifically on changes observed in this gene class (Tonks *et al.*, 2007). Further, even though transcriptomic analysis has proved to be an essential strategy for target identification, it is not a powerful predictor of protein expression. These approaches are often combined with alternative technologies, aiming at analysing the cells' proteomic profile. In addition, no study has so far quantified the cytosolic and nuclear protein expression profile of CD34<sup>+</sup> HSPC expressing RUNX1-ETO. I hypothesize that the early differentiation block observed upon the expression of RUNX1-ETO is mediated by changes in TF expression, and that by identifying these, it will be possible to reverse the phenotype and promote normal cell development (**Figure 1.12**). This hypothesis will be investigated through the following objectives:

**To analyse existing transcriptomic data to identify novel changes in TF mRNA expression arising from RUNX1-ETO expression**

Pre-existing transcriptomic data relating to RUNX1-ETO-induced changes in gene expression in human CD34<sup>+</sup> HSPC will be analysed to identify changes in TF mRNA expression (**Chapter 3**).

**To quantify changes in protein expression in human CD34<sup>+</sup> cord-blood derived HSPC expressing RUNX1-ETO**

Cytosolic and nuclear protein fractions will be extracted from human RUNX1-ETO-expressing and control CD34<sup>+</sup> HSPC and quantitative proteomic analysis will be performed,

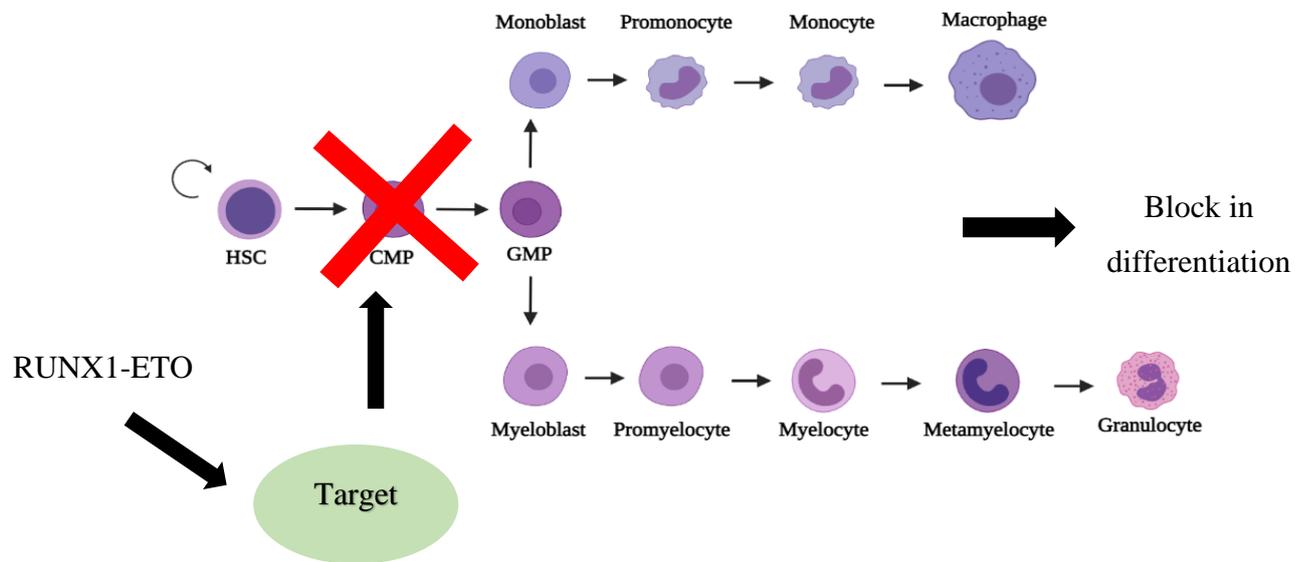
using Sequential Window Acquisition of all Theoretical Mass Spectra (SWATH-MS), further allowing the identification of a target of interest for subsequent studies (**Chapter 3**).

**Determine the effects of ZNF217 overexpression or KD on normal human myeloid development**

Functional analysis will be performed using overexpression and KD (shRNA) vectors to determine the role of ZNF217 in normal human haematopoiesis. Specifically, I will determine whether ZNF217 disrupts myeloid cell proliferation and differentiation. Furthermore, the effect of ZNF217 on proliferation and apoptosis in AML cell lines will be assessed using flow cytometry (**Chapter 4**).

**Determine the effects of C/EBP $\beta$  overexpression or KD on normal human myeloid development**

Functional analysis will be performed using overexpression and KD (shRNA) vectors to determine the role of C/EBP $\beta$  in normal human haematopoiesis. Specifically, I will determine whether C/EBP $\beta$  disrupts myeloid cell proliferation and differentiation. Furthermore, the effect of C/EBP $\beta$  on proliferation and apoptosis in AML cell lines will be assessed using flow cytometry (**Chapter 5**).



**Figure 1.12 – Hypothesis of the study**

This study hypothesizes that RUNX-ETO induces changes in TF expression, which subsequently lead to a block in normal cell differentiation. By identifying these TF, it will be possible to modulate their expression and promote normal haematopoietic development. These changes will be evaluated by analysing previously generated transcriptomic data, as well as newly performed proteomic analysis.

# Chapter 2

---

## Materials and Methods

## 2.1 Molecular Biology

### 2.1.1 Plasmids used in the study

Retroviruses, along with lentiviruses, are RNA viruses capable of infecting eukaryotic cells and are widely used to deliver nucleic acids to many cell types in a variety of experimental systems. They do this by using the reverse transcriptase enzyme to produce DNA within the host. This DNA is incorporated into the host's genome using an integrase enzyme and is then replicated with the host's DNA. A key feature of both viral vectors is that they produce replication-defective, or self-inactivating, particles, allowing the delivery of the desired sequence, without continued viral replication in the target cells (Cooray *et al.*, 2012). Whereas transfection of nucleic acids results only in transient transgene expression, the activity of the viral integrase in retroviral and lentiviral-based systems allows for stable integration of the transgene which is then inherited and continuously expressed over repeated cell divisions. Lentiviruses, unlike retroviruses, are the only retroviral family member that can integrate in non-dividing cells, making them one of the most efficient gene delivery systems (Dufait *et al.*, 2012). The constructs used in the project are outlined in **Table 2.1** and **Figure 2.1**.

### 2.1.2 Bacterial growth medium and selective agar plates

Selective Luria-Bertani (LB) broth media was prepared by dissolving 20g of LB broth (ThermoFisher Scientific, Loughborough, UK) in 1 L of deionised water (dH<sub>2</sub>O) and autoclaved. LB Agar plates were prepared by adding 35g of LB agar (Sigma-Aldrich, Poole, UK) to 1L of dH<sub>2</sub>O, and autoclaved. Medium were supplemented with 100 µg/mL ampicillin (Sigma-Aldrich) for selective growth.

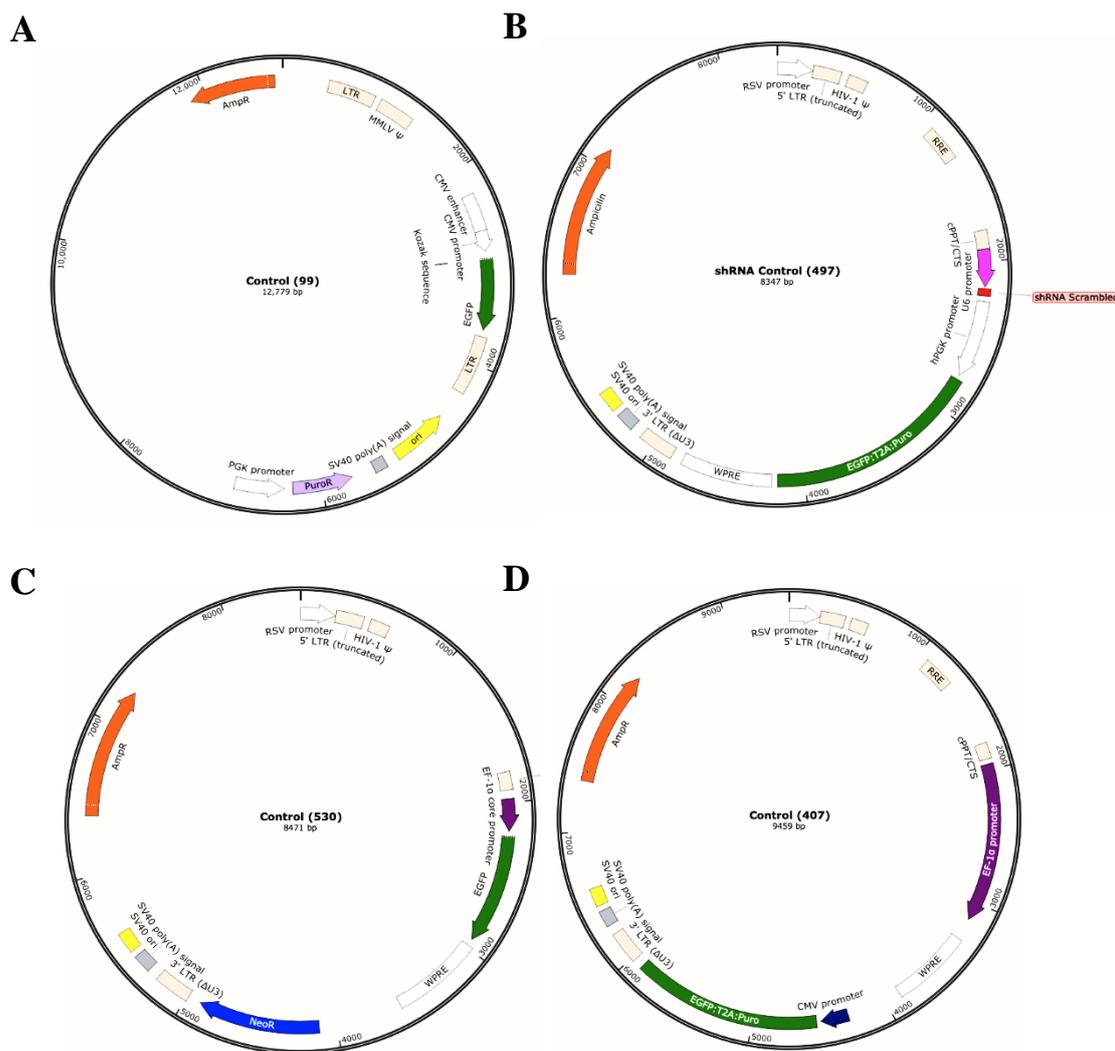
Previously stored glycerol stocks of transformed *E.coli* cells were spread over ampicillin selective LB agar plates and incubated overnight at 37°C. A single colony was aseptically picked and grown in 5 mL LB broth supplemented with ampicillin (100 µg/mL) at 37°C on a rotary shaker at 225 revolutions *per* minute (rpm) for 5-6 h. The culture was subsequently diluted by transferring 150 µL to 150 mL of LB broth supplemented with ampicillin (100 µg/mL) and cultured overnight at 37°C, on a rotary shaker (225 rpm). Prior to plasmid DNA isolation and quantification (**2.1.3**), 850 µL of inoculated broth were added to 150 µL of glycerol and stored at -80 °C.

Table 2.1 – Plasmid vectors used in this study

Target	#	Plasmid	Studies	Vector Name	Target sequence	Match region	Marker(s)	Host bacteria <sup>¥</sup>	Source
Control	99	Pinco	Overexpression <b>Chapter 3</b>	Control	n/a	n/a	GFP	<i>Stbl3</i> (ThermoFisher)	(Tonks <i>et al.</i> , 2003; Grignani <i>et al.</i> , 1998)
	407	pLV	Overexpression <b>Chapter 5</b>	Control	n/a	n/a	Puromycin; GFP	<i>Stbl3</i>	VectorBuilder
	497 <sup>1</sup>	pLV	Knockdown <b>Chapter 4/5</b>	shRNA Control	Scrambled	n/a	Puromycin; GFP	<i>DH5α</i>	VectorBuilder
	530	pLV	Overexpression <b>Chapter 4</b>	Control	n/a	n/a	Neomycin; GFP	<i>Stbl3</i>	VectorBuilder
RUNX1-ETO	1	Pinco	Overexpression <b>Chapter 3</b>	RUNX1-ETO	RUNX1-ETO	n/a	GFP	<i>Stbl3</i>	(Tonks <i>et al.</i> , 2003)
ZNF217	543	pLV	Overexpression <b>Chapter 4</b>	ZNF217 OE	ZNF217 (NM_006526.2)	n/a	GFP	<i>VB UltraStable</i>	VectorBuilder
	552	pLV	Knockdown <b>Chapter 4</b>	shZNF217 #1	TRCN0000364558	CDS	Puromycin; GFP	<i>VB UltraStable</i>	VectorBuilder
	554	pLV	Knockdown <b>Chapter 4</b>	shZNF217 #2	TRCN0000013060	CDS	Puromycin; GFP	<i>VB UltraStable</i>	VectorBuilder
	556	pLV	Knockdown <b>Chapter 4</b>	shZNF217 #3	TRCN0000364635	3'UTR	Puromycin; GFP	<i>VB UltraStable</i>	VectorBuilder
	558	pLV	Knockdown <b>Chapter 4</b>	shZNF217 #4	TRCN0000364634	CDS	Puromycin; GFP	<i>VB UltraStable</i>	VectorBuilder

<b>C/EBP<math>\beta</math></b>	570	pLV	Overexpression <b>Chapter 5</b>	C/EBP $\beta$ OE	C/EBP $\beta$ (NM_005194.3)	n/a	Puromycin; GFP	<i>VB UltraStable</i>	VectorBuilder
	565	pLV	Knockdown <b>Chapter 5</b>	shC/EBP $\beta$ #1	TRCN000007440	3'UTR	Puromycin; GFP	<i>VB UltraStable</i>	VectorBuilder
	565	pLV	Knockdown <b>Chapter 5</b>	shC/EBP $\beta$ #2	TRCN0000364533	CDS	Puromycin; GFP	<i>VB UltraStable</i>	VectorBuilder
	567	pLV	Knockdown <b>Chapter 5</b>	shC/EBP $\beta$ #3	TRCN0000364495	CDS	Puromycin; GFP	<i>VB UltraStable</i>	VectorBuilder
<b>Other</b>	158	pMD2	n/a	pMD2	Addgene_12259	n/a	n/a	<i>Stbl3</i>	Addgene
	189	pMD2	n/a	pMD2	Addgene_12260	n/a	n/a	<i>DH5<math>\alpha</math></i>	Addgene

# indicates laboratory plasmid reference number; <sup>1</sup>DNA sequence with no mammalian target; <sup>2</sup>all plasmids contain an *ampicillin* resistant gene; n/a - not applicable; **shRNA** – short hairpin RNA; **TRCN** – The RNAi Consortium Number; **GFP** – Green Fluorescent Protein; Plasmids were purchased from VectorBuilder (Guangzhou, China) or Addgene (Massachusetts, USA).



**Figure 2.1 – Plasmid DNA constructs used in this study**

Representative plasmid maps of (A) PINCO control vector, with a EGFP selectable marker; (B) shRNA with EGFP:T2A:Puro selectable marker and U6 promoter; (C) lentiviral control vector with a EF-1 $\alpha$  promoter and EGFP/Neomycin selectable markers; (D) lentiviral control with EGFP:T2A:Puro selectable markers and EF-1 $\alpha$  promoter. All vectors were created using SnapGene Software (GSL Biotech LLC, USA).

$\Delta$ U3/3' LTR - truncated HIV-1 3' long terminal repeat;  $\Delta$ 5' LTR/5'UTR - truncated HIV-1 5' long terminal repeat; AmpR - ampicillin resistance gene; CMV enhancer – human cytomegalovirus (CMV) immediate early enhancer; CMV promoter - CMV immediate early promoter; cPPT/CTS - central polypurine tract; EF-1 $\alpha$  core promoter - core promoter for human elongation factor EF-1 $\alpha$ ; EGFP – Green fluorescent protein; EGFP:T2A:Puro - EGFP and Puromycin linked by T2A; HIV-1  $\Psi$  - HIV-1 packaging signal; MMLV  $\Psi$  - packaging signal of Moloney murine leukaemia virus (MMLV); Neor – neomycin resistance gene; PGK promoter - mouse phosphoglycerate kinase 1 promoter; PuroR - puromycin resistance gene; RRE - Rev response element; RSV promoter - Rous sarcoma virus promoter; SV40 poly(A) signal - SV40 polyadenylation signal; SV40 ori - SV40 origin of replication; U6 - RNA polymerase III U6 promoter; WPRE - woodchuck hepatitis virus post-transcriptional regulatory element.

### 2.1.3 Isolation and quantification of recombinant plasmid DNA from bacteria

Following overnight bacterial growth, plasmid DNA was isolated using the QIAGEN HiSpeed® Plasmid Maxi Kit (QIAGEN, Manchester, UK) according to the manufacturer's instructions. Briefly, grown bacterial culture was centrifuged at 6,000 xg for 15 min, at 4°C. Following supernatant discard, pellet was resuspended in 10 mL of P1 buffer, containing RNase I (resuspension buffer), followed by the addition of 10 mL of P2 buffer (lysis buffer). The solution was mixed by slowly inverting the tube and incubated at room temperature (RT) for 5 min, before adding 10 mL of P3 buffer (neutralization buffer). Lysate was transferred onto a QIA filter cartridge and incubated at RT for 10 min. In the meantime, a HiSpeed Maxi tip was equilibrated by adding 10 mL of QBT buffer (equilibration buffer). Lysate was added to the column and allowed to pass through the resin by gravity flow, before being washed with 60 mL of QC buffer (wash buffer). Plasmid DNA was eluted by adding 15 mL of QF buffer and precipitated by adding 10.5 mL of isopropanol (Sigma-Aldrich). Following a 5 min incubation, the eluted DNA was passed through a QIA precipitator using a 30 mL syringe and washed with 2 mL of 70% v/v ethanol (Sigma-Aldrich). The plasmid DNA, retained in the precipitator, was air-dried and collected by passing 1 mL of dH<sub>2</sub>O, through the syringe. The elute was then passed through the precipitator once more to ensure maximum DNA recovery.

Isolated plasmid DNA was quantified using a NanoDrop OneC Spectrophotometer (ThermoFisher Scientific), according to the manufacturer's instructions. To assess DNA purity, the ratio of absorbance at 260/280 nm was used. A ratio between 1.8-2 is considered good quality; proteins absorb at 280 nm, hence lower ratios indicate the presence of high concentration of proteins. This was combined with the analysis of the A<sub>260</sub>/A<sub>230</sub> ratio, that assesses sample purity in terms of salts and other contaminants, which absorb at 230 nm. A sample with a ratio within 2-2.2 is considered "pure". Pure DNA samples were used to generate recombinant virus, as described in 2.4.

### 2.1.4 DNA Sequencing

To validate DNA sequences, Sanger sequencing was performed by Eurofins Genomics, using the Eurofins Genomic's SmartSeq kit sequencing service. For this, 50-100 ng of plasmid DNA, eluted in water, was made up to a volume of 15 µL and combined with 10 µM of the desired primer (**Table 2.2**). Sequences were verified by performing a pairwise sequence

alignment using ClustalW2 (<https://www.ebi.ac.uk/Tools/msa/clustalw2/>). Vectors purchased from VectorBuilder were not used to perform sequencing, as the company provides this data as a quality control check.

**Table 2.2 – Primers used for direct sequencing**

Table outlining primer sequence (5'-3') used for sequencing PINCO vectors above (2.1.4).

Primer	Vector	Strand	Sequence
P1	PINCO	Reverse	5' TTA TGT AAC GCG GAA CTC CA 3'
P2	PINCO	Forward	5' TAG AAC CTC GCT GGA AAG GA 3'

## 2.2 Cell culture

All tissue culture work was performed in a Class II biosafety cabinet and the waste disinfected with bleach and/or autoclaved. Cells were cultured under aseptic conditions in a humidified incubator at 37°C and 5% CO<sub>2</sub>, unless otherwise specified.

### 2.2.1 Culture of cell lines

#### 2.2.1.1 Cell culture of non-adherent cell lines

The origin and specific culture conditions of non-adherent cells used in this study are outlined in **Table 2.3**. All lines were maintained in the appropriate culture media supplemented with 10% (v/v) heat-inactivated foetal bovine serum (FBS; 77133, Labtech International Ltd, Sussex, UK), 1% (v/v) L-glutamine (Invitrogen, California, USA) and 20 µg/mL Gentamycin (Life Technologies, California, USA), unless otherwise specified. Cells were maintained in log phase growth, at a density of 1-8x10<sup>5</sup> cells/mL, in the corresponding culture media.

#### 2.2.1.2 Cell culture of adherent cell lines

The characteristics of adherent cells used in this study have been described in **Table 2.4**. Cell lines were maintained with Dulbecco's Modified Eagle Medium (DMEM; Merck Life Science, Gillingham, UK) supplemented with 10% (v/v) FBS, 1% (v/v) L-glutamine and 20 µg/mL Gentamycin. Adherent cells were passaged when reaching 90% confluence.

**Table 2.3 – Culture conditions for non-adherent cells**

All cell lines were cultured in the presence of 10% v/v FBS, 2mM L-Glutamine and 20 µg/mL gentamicin, unless otherwise stated; all cell lines are derived from human patients, unless otherwise stated.

Cell line	Description	FAB	Molecular characteristics	Media	Seeding density (cells/mL)	Doubling time	Source
<b>HEL</b>	Erythroleukaemia	M6	JAK2 mutation (V617F)	RPMI 1640	1x10 <sup>5</sup> - 1x10 <sup>6</sup>	36h	ECACC
<b>HL-60</b>	APL	M2	t(5;17)(p11;q11)	RPMI 1640	5x10 <sup>5</sup> – 1.5x10 <sup>6</sup>	40h	ATCC
<b>K562</b>	CML	-	t(9;22)(q34;q11) - BCR-ABL1	RPMI 1640	1x10 <sup>6</sup>	30-40h	ECACC
<b>Kasumi-1</b>	AML	M2	t(8;21) (q22;q22) – RUNX1-ETO	RPMI 1640; 20% v/v FBS	3x10 <sup>5</sup>	48-72h	DSMZ
<b>KG1</b>	Erythroleukaemia	M1	-	RPMI 1640	2x10 <sup>5</sup> – 1x10 <sup>6</sup>	35-40h	ATCC
<b>MV-4-11</b>	Myelomonocytic Leukaemia	M5	t(4;11)(q21;q23) – KMT2A-AFF1	IMDM	1x10 <sup>5</sup> - 1x10 <sup>6</sup>	50h	ATCC
<b>NB4</b>	APL	M3	t(15;17)(q22;q11-12.1) – PML-RARA	RPMI 1640	5x10 <sup>5</sup> - 1x10 <sup>6</sup>	35-45h	ATCC
<b>NOMO-1</b>	AML	M5	t(9;11)(p22;q23) – KMT2A-MLLT3	RPMI 1640	5x10 <sup>5</sup>	35h	DSMZ
<b>OCI-AML2</b>	AML	M4	DNMT3A mutation (R635W)	αMEM; 20% v/v FBS	7x10 <sup>5</sup> - 1x10 <sup>6</sup>	40-50h	DSMZ
<b>OCI-AML5</b>	AML	M4	t(1;19)(p13;p13)	αMEM; 20% v/v FBS; 10 ng/mL GM-CSF	5x10 <sup>5</sup> - 1.5x10 <sup>6</sup>	30-40h	DSMZ
<b>PLB985</b>			Clone of HL-60	RPMI 1640	5x10 <sup>5</sup> – 1x10 <sup>6</sup>	30h	ATCC
<b>SKNO-1</b>	AML	M2	t(8;21) (q22;q22) – RUNX1-ETO	RPMI 1640; 10 ng/mL GM-CSF	2x10 <sup>5</sup> - 2x10 <sup>6</sup>	35-50h	DSMZ
<b>TF1</b>	Erythroleukaemia	M6	-	RPMI 1640	5x10 <sup>5</sup> – 1x10 <sup>6</sup>	70h	ATCC
<b>THP-1</b>	Acute Monocytic Leukaemia	M5	(9;11)(p21;q23) leading to KMT2A-MLLT3 (MLL-MLLT3; MLL-AF9)	RPMI 1640	2x10 <sup>5</sup> - 4x10 <sup>5</sup>	40-50h	ECACC
<b>TK6</b>	CML	-	t(9;22)(q34;q11) - BCR-ABL1	RPMI 1640	2x10 <sup>5</sup> - 1x10 <sup>6</sup>	12h	Prof Allan, Newcastle Uni.

---

<b>U937</b>	Histiocytic Lymphoma	M5	t(10;11)(p12;q14) – MLLT10-PICALM	RPMI 1640	$1 \times 10^5$ - $2 \times 10^6$	30-40h	ATCC
-------------	----------------------	----	-----------------------------------	-----------	-----------------------------------	--------	------

---

**αMEM** - Minimum Essential Medium Eagle- Alpha Modification; **APL** – Acute Promyelocytic Leukaemia; **ATCC** – American Type Culture Collection (USA); **CML** – Chronic Myeloid Leukaemia; **DMEM** - Dulbecco's Modified Eagle's Medium; **DSMZ** - German Collection of Microorganisms and Cell Cultures (Germany); **ECACC** - European Collection of Authenticated Cell Cultures (UK); **FBS** – Heat-Inactivated Foetal Bovine Serum (Labtech, East Sussex, UK); **Gent** – Gentamycin (Gibco, ThermoFisher Scientific); **IMDM** - Iscove Modified Dulbecco's Medium; **L-Glu** – L-Glutamine (Gibco, ThermoFisher Scientific); **RPMI** - Roswell Park Memorial Institute-1640 Medium.

Confluent cell monolayers grown were detached from the flask by incubation with 3 mL (F25) or 5 mL (F75) of pre-warmed trypsin (500 µg/mL; Fisher Scientific) at RT for 3 min. Trypsin was neutralized by adding an equal volume of media, following which cells were disaggregated from the wall of the flask and transferred to a universal container (UC). Following centrifugation at 270 xg for 10 min, supernatant was discarded, and cell pellets resuspended in 20 mL of culture media. Cells were counted, as described in 2.2.2, and seeded at the density required for subsequent assays.

**Table 2.4 – Culture conditions for adherent cells**

All cell lines were cultured in DMEM, in the presence of 10% (v/v) FBS, 1% (v/v) L-Glutamine and 20 µg/mL gentamicin.

Cell line	Description	Seeding Density (cells/mL)	Doubling time	Source
<b>HEK 293T</b>	Human embryonic kidney cells	Ratio 1:6 as necessary or every 72h	24-30h	ECACC
<b>HeLa</b>	Human epithelial adenocarcinoma cells	Ratio: 1:4 as necessary or every 3-5 days	40-48h	ATCC
<b>Phoenix</b>	Human epithelial kidney cells	Ratio 1:3 as necessary or every 36h	24-30h	Prof Nolan, Stanford University

### 2.2.2 Determination of cell density

Cell density was verified before each passage and experiment using a Neubauer chamber (Hawksley, Brighton, UK). To do this, 10 µL of cell suspension were added to the chamber, and cell count was performed under the microscope. Density was determined using the following equation:

$$Cell\ density = \frac{\sum Cells\ per\ quadrant}{4} \times 10^4 / mL$$

Alternative, counting by flow cytometry was performed by adding 50 µL of cell suspension to a mini flow tube and acquiring 10 µL of cells through an Accuri C6 Plus flow cytometer (2.7). After excluding cell debris, based on forward scattered light (FSC) and side scattered light (SSC), the number of events was multiplied by 100 to obtain the correct cell density

(cells/mL). To ensure proper counting, beads seeded at a known density were acquired, in triplicate, before determining the density of each cell culture.

### 2.2.3 Determination of cell viability

TO-PRO™-3 Iodide (ThermoFisher) is a viability stain that is impermeant to live cells but can pierce through the membrane of dead cells, allowing their identification within a cell culture. This stain exhibits far-red fluorescence with excitation at 642 nm and emission at 661 nm. Cell viability was assessed by adding 1  $\mu$ L of TO-PRO™-3 (5  $\mu$ M) to 50  $\mu$ L of cell suspension and analysed using an Accuri Plus C6 cytometer, as described in 2.7.

### 2.2.4 Cryopreservation and thawing of cell lines

For cryopreservation, cell suspensions were transferred to a UC and centrifuged at 270  $\times$ g for 10 min. Supernatants were discarded, and the pellets resuspended in 500  $\mu$ L/vial of appropriate culture media (**Table 2.3**, **Table 2.4**). An equal volume of freezing media (Iscove Modified Dulbecco's Medium (IMDM, Gibco) supplemented with 30% (*v/v*) FBS and 20% (*v/v*) Dimethyl Sulfoxide (DMSO; Sigma-Aldrich)), was added to the cells, following which samples were immediately transferred to 1.8 mL cryovials and placed in controlled refrigerated containers (Cool Cell™ Cell Freezing Container, BioCision, California, USA). These were placed in a -80°C freezer overnight to allow the cells to equilibrate. For long-term storage, cells were transferred to liquid nitrogen containers.

Cryopreserved cells were thawed by adding 1 mL of pre-warmed FBS, followed by rapid thawing at 37°C in a water bath. Cells were transferred to a UC containing 5 mL of fresh media, followed by a 5 min centrifugation at 270  $\times$  g. Supernatant was discarded, pellet was resuspended in the appropriate media and transferred to a new F25 tissue culture flask, to allow recovery overnight. The following day, cells were transferred to a new F75 flask and supplemented with fresh media (**Table 2.3**; **Table 2.4**). HSPC were recovered as described in 2.3.2.

## 2.3 Primary cell culture

### 2.3.1 Isolation of mononuclear cells from human neonatal cord blood

Human neonatal CB was obtained from healthy full-term pregnancies at the Maternity Unit of the University Hospital of Wales in Cardiff, with the informed consent of mothers undertaking elective caesareans. The permission for the use of CB was authorized through South East Wales local research ethics committee (licence number 06/WSE03/6). In addition, fresh CB was purchased from the NHS Bone Marrow and Transplant Unit in London. Mononuclear cells were isolated from cord blood by density gradient centrifugation using Ficoll-Paque™ (GE Healthcare, Little Chalfont, UK). Briefly, cord blood was diluted by adding an equal volume of Hank's Balanced Salt Solution (HBSS; Gibco) supplemented with 10 mg/mL heparin, 20 µg/mL gentamycin and 25 mM HEPES solution (Gibco). Approximately 7 to 8 mL of diluted cord blood were gently layered on 5 mL of Ficoll-Paque™ and centrifuged at 400 xg for 40 min, with slow acceleration and brake off. The mononuclear cells found at the interface between the plasma and the Ficoll-Paque™ solution were carefully aspirated and transferred to a UC containing CB Wash (CBW; RPMI 1640, 5% v/v FBS, 20 µg/mL gentamycin and 10 mg/mL heparin), followed by a 10 min centrifugation at 270 xg. The mononuclear cells were washed again until supernatant was clear ensuring no platelet contamination. Cells were resuspended in 20 mL of CBW, from which 10 µL were removed and added to 190 µL of CBW and 1.5 µL of zaponin (Beckman-Coulter, High Wycombe, UK) for cell counting (2.2.2). Following centrifugation cells were frozen as described in 2.2.4 at a density of  $5 \times 10^7$  mononuclear cells/vial. Once thawed, CD34<sup>+</sup> cells were isolated as described in 2.3.2. Estimation of CD34 positivity within each sample was performed as described in 2.7.2.

### 2.3.2 Isolation of CD34<sup>+</sup> HSPC from mononuclear cells

Following mononuclear cell isolation, CD34<sup>+</sup> HSPC were isolated using the MiniMACS® Indirect CD34 MicroBead Kit (Miltenyi Biotec, Woking, UK) according to the manufacturer's instructions. Briefly, to each vial of frozen cells, 1 mL of FBS and 20 µL of DNase (10 µg/mL; Sigma-Aldrich) were added and rapidly thawed in a 37°C water bath. Following viral content transfer to a labelled UC, an equal volume of a 1x Phosphate Buffer Saline (PBS) solution supplemented with 5 mM magnesium chloride (MgCl<sub>2</sub>; Sigma-Aldrich) at RT was added

dropwise to the culture, over 3 minutes, resulting in the doubling of the volume. Two subsequent doubling dilution steps were performed. Cells were gently centrifuged and an additional 10  $\mu\text{L}$  of DNase (10  $\mu\text{g}/\text{mL}$ ) were added. Subsequent volumes were determined *per*  $1 \times 10^8$  cells, based on previous counts. Centrifuged pellets were resuspended in 400  $\mu\text{L}$  of cold column buffer (MACS buffer; 1x PBS; 1% *v/v* Bovine Serum Albumin [BSA; Biosera]; 5 mM  $\text{MgCl}_2$ ) and incubated at 4°C for 15 min with 100  $\mu\text{L}$  hapten-conjugated monoclonal CD34 antibody (reagent A1) in the presence of FcR blocking agent (reagent A2). Cells were washed with 5 mL of MACS buffer to stop the reaction and centrifuged at 270 *xg* for 5 min. Washed cells were resuspended in 400  $\mu\text{L}$  MACS buffer and incubated at 4°C with 100  $\mu\text{L}$  of anti-hapten microbeads (reagent B) for 15 min. Subsequently, 5 mL of MACS buffer were added, and the cell suspension flowed through a cell 40  $\mu\text{m}$  cell strainer, following which cells were centrifuged at 270 *xg* for 5 min. The cell pellet was later resuspended in 500  $\mu\text{L}$  of MACS buffer and transferred to a pre-equilibrated MS/LS column (depending on desired cell number) (Miltenyi Biotec). Once column flow stopped, three wash steps were performed by adding 500  $\mu\text{L}$  of MACS buffer to the column, following which CD34<sup>+</sup> HSPC enriched population was eluted in 1 mL MACS buffer. The eluted cells were transferred to a second column and previously described wash steps performed to maximise enrichment. Following culture elution, cells were counted (2.2.2) and subcultured at  $2 \times 10^5$  cells/mL in appropriate medium (2.3.3). Isolation of CD34<sup>+</sup> HSCPC is defined as ‘day 0’ in growth and differentiation experiments (2.5, 2.7, 2.8).

### 2.3.3 Culture of CD34<sup>+</sup> HSPC

Human CD34<sup>+</sup> HSPC were cultured in IMDM containing 20% *v/v* FBS, 1% *v/v* BSA, 45  $\mu\text{M}$   $\beta$ -mercaptoethanol ( $\beta$ -ME; Sigma-Aldrich), 360  $\mu\text{g}/\text{mL}$  of 30% iron saturated human transferrin (Roche Diagnostics; Switzerland) and 20  $\mu\text{g}/\text{mL}$  gentamycin. Basal media was subsequently supplemented with specific combinations of cytokines, based on cell requirements (Table 2.5). After isolation, CD34<sup>+</sup> HSPC were cultured in cytokine rich media, defined as 3S<sup>high</sup>, to promote recovery and cell cycle progression. On day 3 of culture, and following retro- or lentiviral transduction (2.5), cells were maintained in 3S<sup>low</sup>G/GM (Table 2.5) to support myeloid cell differentiation. Media was replenished every 2-3 days to guarantee appropriate cytokine concentrations. Cells were seeded at different densities, according to developmental stages.

**Table 2.5 – Cell culture media used for the growth and development of CD34<sup>+</sup> HSPC**

Human CD34<sup>+</sup> HSPC were cultured in IMDM supplemented with 20% v/v FBS, 1% v/v BSA, human transferrin (30 mg/mL),  $\beta$ -ME (9 mM) and gentamicin (20  $\mu$ g/mL). Different combinations of cytokines were added to the growth according to developmental stage.

<b>Days</b>	<b>Application</b>	<b>Cytokine*Concentration</b>	<b>Media</b>	
<b>0-3</b>	<b>Cell recovery</b>	IL-3	50 ng/mL	<b>36S<sup>high</sup></b>
		IL-6	25 ng/mL	
		SCF	50 ng/mL	
		G-CSF	25 ng/mL	
		GM-CSF	25 ng/mL	
		Flt3	50 ng/mL	
<b>3-13</b>	<b>Long-term culture</b>	IL-3	5 ng/mL	<b>3S<sup>low</sup>G/GM</b>
		SCF	5 ng/mL	
		G-CSF	5 ng/mL	
		GM-CSF	5 ng/mL	

\* All cytokines were purchased from Biolegend (London, UK); **IL-3** – human interleukin-3; **IL-6** - human interleukin-6; **SCF** – human Stem Cell Factor; **G-CSF** - human Granulocyte Colony-Stimulating Factor; **GM-CSF** - human Granulocyte-Macrophage Colony Stimulating Factor; **Flt3** - Human Fms-like tyrosine kinase-3 ligand.

Within the 3<sup>rd</sup> and 6<sup>th</sup> days of culture, cells were seeded at a  $0.5 \times 10^5$  cells/mL, to allow maximum proliferative rate. From days 6 through 10, cells were seeded at  $2 \times 10^5$  cells/mL, and  $3 \times 10^5$  cells/mL from 10 onwards.

### 2.3.4 Colony forming efficiency

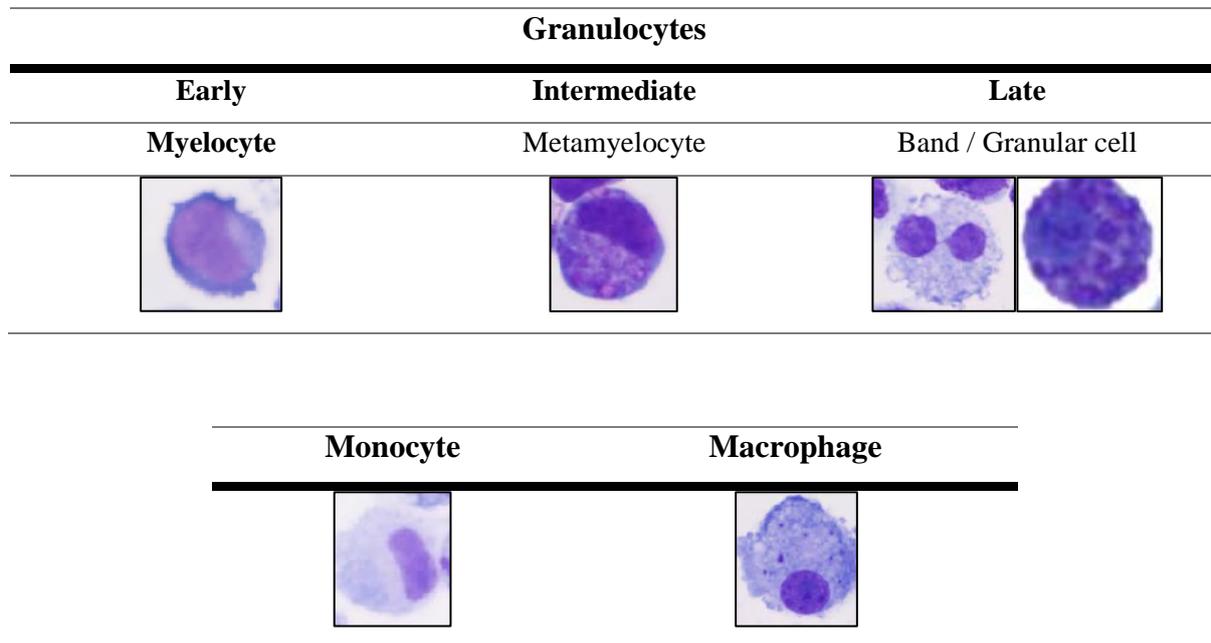
To determine cloning-forming cell frequency, a limiting dilution colony assay was performed on Fluorescently Activated Cell Sorted (FACS) CD34<sup>+</sup> HSPC (2.7.3) transduced with different viral plasmids (2.5.2), in 96-U plates, at a density of 0.3 cells/well, to ensure the presence of a cell within every third well, and incubated at 37°C with 5% CO<sub>2</sub>. Individual colonies (>50 cells) and clusters (>5 cells) were scored by counting via light microscopy following 7 days of growth. To assess self-renewal efficiency, identified colonies were harvested, counted, and re-plated at a density of 1 cell/well, and cultured for an additional week before re-scoring, as mentioned above.

### 2.3.5 Assessment of cell morphology

To analyse cell morphology,  $3 \times 10^4$  cells were transferred to a pre-assembled cytopsin slide chamber with a glass slide and centrifuged 60 xg for 5 min, using a Cytospin 3 (Fisher Scientific). Slides were subsequently stained with May-Grünwald-Giemsa for morphology examination and scanned using Zeiss AxioScan Z1 Slide Scanner (Carl Zeiss, New York, USA) at 20x magnification. Differential counts were performed using QuPath (Edinburgh, UK). Cells were defined as being in an early, intermediate, or late phase of development if they presented as myeloblasts/promyelocytes, myelocytes/metamyelocytes and band/segmented cells, respectively, as described in **Figure 2.2**.

## 2.4 Virus transfection

To optimize viral transduction of cell lines and CD34<sup>+</sup> HSPC, thus ensuring high titre virus, two transfection methods were tested, as described below. Lipofection (2.4.1) was performed to generate lentivirus, using HEK 293T cells, whilst the calcium-phosphate method proved to be more efficient when producing retroviral particles using the Phoenix cell line (2.4.2).



**Figure 2.2 – Morphological analysis of HSPC**

Assessment of cell morphology was performed on sorted primary cells, on day 17 of culture, using May-Grünwald-Giemsa staining. Cells were scored as granulocytes, in different stages of differentiation, and identified in early, intermediate, or late developmental stages. Additionally, cells were scored as monocytes and macrophages.

### 2.4.1 Transfection using Lipofection

Lipofection was achieved the Lipofectamine<sup>®</sup> 3000 Kit (Thermo Fisher Scientific) according to the manufacturer's instructions. Briefly, cells were counted and seeded the day before transfection at  $5 \times 10^6$  (T25) or  $15 \times 10^6$  cells (T75) *per* flask in a previously Poly-D-Lysine (Trevigen<sup>®</sup>, Abingdon, UK) coated flask. Flasks were incubated overnight at 37°C, in 5% CO<sub>2</sub> incubator to allow monolayer formation. The following day, cell confluence was assessed, and two tubes were prepared as follows.

For retroviral production, tube A was prepared by adding 8 µg (T25) or 24 µg (T75) of the desired vector to 667 µL (T25) or 2 mL (T75) of Opti-MEM Media (Gibco) and 16µL (T25) or 48 µL (T75) of the Lipofectamine 3000 Reagent. Tube B was prepared by adding 667 µL (T25) or 2 mL (T75) of Opti-MEM Media and 10 µL (T25) or 57 µL (T75) of the P3000 Enhancer Reagent. To prepare lipid-DNA complexes, contents of Tube A were transferred to Tube B and incubated for 15-20 min at RT. Following incubation, media in each flask was reduced to 2 mL (T25) or 6 mL (T75), and the lipid-DNA complex was carefully added to the cells. Flasks were incubated for 6 h at 37°C, following which time all media was removed, and replaced with fresh. Cells were incubated for 24 h, following which time supernatant was removed and discarded, and fresh media was added. Retrovirus harvesting was performed 48- and 72h post-transfection, with media being changed between each harvest. Viral supernatants were centrifuged at 270 xg for 10 min to remove cell debris, and aliquoted into 1.8mL cryovials, following with they were snap frozen in liquid nitrogen. For long-term storage, cryovials were kept at -80°C and thawed when necessary.

For lentiviral transfection, two additional plasmids were added to Tube A: 2.2 µg (T25) or 6.6 µg (T75) of pMD2, an envelope plasmid, and 4 µg (T25) or 12 µg (T75) of psPAX2, a packaging plasmid, as well as 2.2 µg (T25) or 6.6 µg (T75) of the desired vector; tube B was prepared as described for above. For viral harvest, lentiviral supernatant was collected 24- and 48h post-transfection and pooled (stored overnight on each collection at 4°C). Once both viral harvests were performed and combined, viral supernatants were centrifuged at 270 xg for 10 min and aliquoted/stored as described above.

To assess target protein expression, cells were collected, and proteins extracted, according to 2.6. Furthermore, where cells were transfected with virus encoding for GFP expression, visual inspection by fluorescent microscope was performed.

#### 2.4.2 Transfection using calcium-phosphate precipitation

To improve transduction efficiencies in human CD34<sup>+</sup> HSPC, an alternative method of virus transfection was used. For this, Phoenix virus-packaging cells were transfected using the Calcium phosphate transfection kit (Invitrogen™) according to the manufacturer's instructions (this method was only used for retroviral production). Briefly, cells were counted and initially seeded at  $8 \times 10^6$  cells *per* T75 flask. The following morning, media was replaced with 14 mL of fresh pre-warmed media. By the afternoon, tissue grade sterile water was mixed with 45 µg of plasmid DNA and 45 µL of CaCl<sub>2</sub> (2.5M), followed by the dropwise addition of 450 µL of 2x HEPES-buffered saline (HBS; 50mM), whilst the solution was bubbled with a pipette. Following vortexing, the solution was left to precipitate for 20 min at RT. Prior to the transfection, 15 µL of chloroquine (25 µM) was added to the cultures, followed by the dropwise addition of DNA to the medium. Cells were incubated 37°C and 5% CO<sub>2</sub> until the following day, by which growth media was changed and cells were incubated at 33°C to maximise retroviral stability. Viral supernatant was harvested at 48 and 72 h post-transfection as previously described (2.4.1).

#### 2.4.3 Virus titration

Virus titration was performed using K562 cells. 96-well untreated plates were coated with Retronectin® (30 µg/mL; Takara Biotech Inc., Shiga, Japan), a recombinant human fibronectin peptide fragment that promotes binding between cells and viruses (Tonks *et al.*, 2005), and incubated at 4°C overnight. The following day, Retronectin® was removed and washed with blocking buffer (PBS with 1% (v/v) BSA) for 30 min at RT. Following blocking, 100 µL of K562 cells at a density of  $1 \times 10^5$  cells/mL were added to each well, before adding 50 µL of viral supernatant (2.3.4) to each well; retroviral and lentiviral aliquots of known titre (standards) were also included to standardise the data. Cells were cultured for two days at 37°C and 5% CO<sub>2</sub>. Following incubation, media was carefully aspirated and washed to remove residual virus, by gently adding PBS without disturbing the cells attached to the wells. Cells

were resuspended in Staining Buffer (SB; 1x PBS supplemented with 0.02% (v/v) sodium azide and 1% (v/v) BSA) and analysed by flow cytometry for GFP expression (2.7).

## 2.5 Transduction of cell lines and human CD34<sup>+</sup> HSPC

### 2.5.1 Transduction of cell lines

For overexpression studies, using either retro- or lentiviral vectors, 12-well non-tissue culture plates were coated Retronectin<sup>®</sup> (30 µg/mL) and incubated at 4°C overnight. The following day, Retronectin<sup>®</sup> was aspirated and replaced with 1% v/v BSA as previously described (2.4.3). The BSA was aspirated and replaced with 1 mL of retro- or lentiviral supernatant. The culture plate was centrifuged at 3000 xg for 120 minutes at RT, following which the supernatant was removed from the wells and replaced with the appropriate number of cells (4x10<sup>5</sup> cells/mL), subsequently incubated at 37°C overnight. Cells expressing GFP were assessed alongside parental cell cultures to measure infection rates, the day following infection, by flow cytometry (2.7).

For KD studies using shRNA constructs, pre-coating of cells was unnecessary given the higher titre virus generated from these constructs. Instead, cells were resuspended at 4x10<sup>5</sup> cells/mL and seeded into a 12-well plate, and 1 mL of culture was added along with 0.5-1 mL of viral supernatant. Cells were assessed for GFP expression as previously described. Following successful cell transduction, cultures were subjected to antibiotic selection to enrich the culture for transduced cells and ensure a pure population. To this end, transduced cells were incubated with 10 µg/mL puromycin (Sigma-Aldrich; P8833), until parental non-transduced cultures were no longer viable. Following selection, transduced cell lines were maintained in normal growth media until confluency was achieved, by which point cells were cryopreserved (2.2.4) and total protein lysates were collected to validate protein expression by western blot analysis (2.6). KD studies were performed within 3 weeks of generating the cell lines.

### 2.5.2 Transduction of human CD34<sup>+</sup> HSPCs

For overexpression and KD studies, plates were coated Retronectin<sup>®</sup> (30 µg/mL) and incubated at 4°C overnight, as previously described (2.5.1). Following BSA washing, viral supernatants were added to the corresponding wells and plates centrifuged at 3000 xg for 120 minutes at RT. Following centrifugation, the viral supernatant was removed and CD34<sup>+</sup> cells,

previously seeded at  $2 \times 10^5$  cells/mL for overnight recovery following isolation (2.3.2), were added to the wells. For overexpression studies, the following day, cells were temporarily transferred to a UC and placed at  $37^\circ\text{C}$  in the incubator, whilst another round of infection was performed as above. For KD studies, however, a second viral coating step was not performed due to high viral titre. Instead, cells were incubated with the lentiviral particles for 2 consecutive days, before being harvested and analysed on day 3 of culture. Cells were subsequently analysed through flow cytometry and used to perform the assays described in 2.7.2-2.7.6, 2.3.4 and 2.8.

## 2.6 Determination of protein expression by western blot

### 2.6.1 Total protein extraction

To generate total protein lysates, the appropriate number of cells to be fractionated ( $1 \times 10^6$  AML cells;  $2 \times 10^6$  HSPC) was centrifuged at  $270 \times g$  for 10 min. Supernatant was discarded, and cell pellets were washed using 10 mL of ice-cold tris-buffered saline (TBS; 50 mM Tris [Sigma-Aldrich] dissolved in  $\text{dH}_2\text{O}$  and adjusted to pH 7.6 with HCl; 150 mM NaCl [Sigma - Aldrich] dissolved in  $\text{dH}_2\text{O}$ ). Supernatant was discarded and pellets snap frozen in liquid nitrogen before protein extraction. When required, frozen cell pellets were thawed on ice in the presence of 1 mg/mL DNase for 5 min and resuspended in 50  $\mu\text{L}$  of homogenisation buffer (0.25 M sucrose [Sigma-Aldrich], 10 mM HEPES-KOH pH7.2 [Gibco], 1mM magnesium acetate [Sigma-Aldrich], 0.5 mM EDTA [Sigma-Aldrich], 0.5 mM EGTA [Sigma-Aldrich], 12.6 M BME [Sigma-Aldrich], 1 tablet of EDTA free-protease inhibitor [Roche Diagnostics; Basel, Switzerland], 1% v/v x100-Triton [Sigma-Aldrich]). Cells were incubated on ice for 30 min and subjected to occasional vortexing. Subsequently, cell suspension was transferred to ice cold Eppendorf tubes (Eppendorf, Stevenage, UK) and centrifuged for at  $16,000 \times g$  for 5min at  $4^\circ\text{C}$ , using a Biofuge fresco Heraeus centrifuge (ThermoFisher Scientific). Supernatant was collected transferred to a new pre-chilled Eppendorf and stored at  $-80^\circ\text{C}$  until required. Protein concentration was determined by Bradford assay (2.6.3).

### 2.6.2 Fractionated protein extraction

To generate fractionated protein cell lysates, extractions were performed using the Biovision Nuclear/Cytosol Fractionation Kit (Cambridge Bioscience, Cambridge, UK), following manufacturer's instructions. Briefly, the appropriate number of cells ( $2 \times 10^6$  AML cells;  $5 \times 10^6$

HSPC) was centrifuged at 270 xg for 10 min, following which supernatant was discarded and pellets washed with 20 mL of ice-cold TBS. Subsequently, cell pellets were resuspended in 200  $\mu$ L of Cytosol Extraction Buffer A, supplemented with Protein Inhibitor Cocktail (PIC) and Dithiothreitol (DTT) and vortexed, followed by a 10-min incubation on ice. Additionally, 11  $\mu$ L of Cytosolic Extraction Buffer B were added, cells were vortexed again and incubated on ice for a further minute. Cell pellets were vortexed and centrifuged at 10,000 xg for 8 min, at 4°C. The supernatant, containing the cytosolic fraction was collected and stored at -80°C. The remaining cell pellet was washed with 500  $\mu$ L of PBS supplemented with 5 mM MgCl<sub>2</sub> and centrifuged for 3 min at 10,000 xg. Supernatant was discarded and the pellet snap frozen in liquid nitrogen. Following 2-3 cycles of rapid freezing and thawing in the vapour phase on liquid nitrogen, thus allowing the rupture of the nuclear membrane, 2  $\mu$ L/million cells of benzonase (25 U/ $\mu$ L; Merck Millipore, USA) was added, and cells were incubated on ice for 30 min, subjected to occasional vortexing. Following incubation, 50-100  $\mu$ L of TEAB lysis buffer (0.5 M Triethylammonium bicarbonate buffer; 0.05% SDS [ThermoFisher Scientific]; 0.1% Protease Inhibitor Cocktail [Sigma-Aldrich]; 0.1% Phosphatase Inhibitor Cocktail [Sigma-Aldrich], dH<sub>2</sub>O) was added, and cells were kept on ice for another 30 min, with occasional vortexing. Following incubation, cell lysates were centrifuged at 10,000 xg for 10 min, at 4°C and supernatant collected to a fresh tube. Protein quantification was carried out as described below (2.6.3).

### 2.6.3 Protein quantification using Bradford assay

Protein quantification was performed using Bradford Assay (Ernst and Zor, 2010). Total protein extract samples were diluted 1:100 using ddH<sub>2</sub>O, whilst cytosolic/nuclear samples were diluted 1:50. Each sample was assayed in duplicate by adding 10  $\mu$ L of diluted lysates alongside 10  $\mu$ L protein standards, ranging between 0-100  $\mu$ g/mL (Sigma-Aldrich), which were also added in duplicate. Bradford's reagent (Sigma-Aldrich; B6916) was diluted 1:1 in dH<sub>2</sub>O, and 190  $\mu$ L of diluted reagent was added to each well. The plate was then incubated for 5 at RT, protected from light. Absorbances were measured at 595 nm using a plate reader. The protein standards were used to generate a standard curve from which the concentration of protein in each sample was determined.

#### 2.6.4 Protein electrophoresis and electroblotting

All reagents and equipment were purchased from Invitrogen and were prepared according to the manufacturer's instructions, unless otherwise stated. Prior to SDS-polyacrylamide gel electrophoresis, samples were denatured in a 70°C water bath in the presence of NuPAGE™ 4x Lithium Dodecyl Sulphate Sample Buffer (LDS), NuPAGE™ 10x Sample Reducing Agent and dH<sub>2</sub>O. Protein gel electrophoresis was performed using a NuPAGE™ XCell SureLock™ Mini-cell and XCell II blot module system. NuPAGE™ 4-12% Bis-Tris gels (12- or 17-well) were placed in a module to which 1x NuPAGE™ MOPS SDS Running Buffer, supplemented with 500 µL of NuPAGE™ antioxidant was added. Between 10 – 20 µg of protein were added to each well, to a maximum volume of 10 µL (in a 17-well gel) to 20 µL (in a 12-well gel). Samples were loaded alongside MagicMarker XP protein ladder (1/10 dilution in 4x LDS sample buffer and dH<sub>2</sub>O), to allow visualisation of protein molecular weight. Following sample loading, the outer chamber of the gel tank was filled with approximately 600 mL of 1x MOPS buffer. Proteins were separated through electrophoresis at 200V for 50 min.

During this period, 1x NuPAGE™ Transfer buffer supplemented with 1 mL NuPAGE™ Antioxidant and 10% *v/v* methanol (20% *v/v* if blotting two gels) was prepared and used to pre-soak blotting pads and filter paper. Polyvinylidene difluoride (PVDF) membranes were soaked for 10 min in methanol prior to electroblotting. Following electrophoresis, the gels were removed from their cassettes and assembled between the pre-soaked PVDF membrane and filter paper, followed by the pre-soaked blotting pads. Following module assembly, 1x transfer buffer was used to fill the transfer cell until the pads were completely submerged, and 600 mL dH<sub>2</sub>O was used to fill the outer chamber. Membranes were electroblotted for 1 h at 30V.

Following transfer, electroblotted membranes were washed twice with dH<sub>2</sub>O using a plate shaker (1 revolution / sec) for 5 min. Proteins were visualised for both equal loading and to facilitate cutting of membrane using 20 mL of Ponceau S solution (Sigma-Aldrich), for 30 sec. The Ponceau S stain was removed using dH<sub>2</sub>O before membranes were blocked with 2.5% (*w/v*) powdered milk diluted in TBS supplemented with 0.1% (*v/v*) Tween-20 (TBS-T; Sigma-Aldrich) for 30 min at RT on a rotary shaker. Membranes were subsequently washed for 15 min in TBS-T, followed by 3 x 5-min cycles of further TBS-T washes. Membranes were used for immunoblotting (2.6.5), as described below.

### 2.6.5 Immunoblotting and protein detection

Following membrane blocking and subsequent washes, membranes were incubated with primary antibodies as described in **Table 2.6**, in 2.5% (*w/v*) powdered milk diluted TBS-T overnight at 4°C. The following day, membranes were washed with TBS-T, as described above (**2.6.4**), and incubated with the appropriate anti-rabbit or anti-mouse horseradish peroxidase (HRP) conjugated secondary antibody, diluted in 1% (*w/v*) powdered milk diluted in TBS-T for 1 h at RT. Following incubation, membranes were washed with TBS-T as previously described (**2.6.4**). HRP-conjugated antibodies were exposed through a chemiluminescence reaction using Amersham™ ECL™ Prime (GE HealthCare Life Sciences), according to the manufacturer's instructions. Briefly, equal volumes of Amersham™ ECL™ Luminol Enhancer solution (solution A) and Amersham™ ECL™ peroxide solution (solution B) were added to the membranes, to a combined volume of 4 mL. Following a 5 min incubation at RT, protected from light, excess substrate was removed using filter paper. Membranes were visualized using a LAS-3000 digital imager (Fujifilm UK Ltd, Bedfordshire, UK), using an exposure time of up to 30 min.

Equal loading was further assessed using Glyceraldehyde 3-Phosphate Dehydrogenase (GAPDH) or Histone H1 antibodies for total/cytosolic and nuclear protein extracts, respectively, using fluorescent antibodies. For this, following chemiluminescent imaging, membranes were washed with TBS-T, as described above (**2.6.4**). Subsequently, antibodies were diluted in 2.5% (*w/v*) powdered milk diluted TBS-T and incubated for 1h at RT, protected from light, following which membranes were washed for 5 min using TBS-T. Membranes were imaged directly after probing using an Odyssey® FC Imaging System (LI-COR Biosciences, Nebraska, USA) for 30s. Image files were saved as .TIFF Images and analysed using Adobe Photoshop CS4. Densitometric analysis was performed as described in **2.6.7**.

### 2.6.6 Membrane and gel protein staining

Alternatively, following protein transfer, to confirm protein loading and sample integrity, membranes were stained with SYPRO® Ruby protein blot stain (ThermoFisher Scientific) for 15 min in a rotary shaker. Following incubation, these were washed 4-6 times with dH<sub>2</sub>O for 1 min and visualized in an Odyssey® Fc Imaging System (LI-COR Biosciences, Cambridge, UK). These membranes were not suitable for subsequent antibody incubation.

**Table 2.6 – List of primary antibodies used in this study**

Table outlining the source, clone, conjugate, and dilution of the antibodies used in western blot analysis (2.6.5). Primary antibodies were stored at -20°C and secondary antibodies at 4°C.

Antibody	Cat. #	Species	Conjugate	Conditions	Dilution	Source	RRID
Primary antibodies							
AML1 (RUNX1)	#4334	Rabbit	-	4°C · O/N · Chemi	1:5000	CST	AB_2184099
ARID5B	PA5- 78597	Rabbit	-	4°C · O/N · Chemi	1:1000	Fisher Scientific	AB_2735478
C/EBPβ	#90081	Rabbit	-	4°C · O/N · Chemi	1:1000	CST	Not defined
IRF7	PA5- 79519	Rabbit	-	4°C · O/N · Chemi	1:1000	Fisher Scientific	AB_2746635
IRF9	D2T8M	Rabbit	-	4°C · O/N · Chemi	1:1000	CST	AB_2799885
ZNF217	#720352	Rabbit	-	4°C · O/N · Chemi	1:1000	Fisher Scientific	AB_2716919
Secondary Antibodies							
Anti- mouse	NA931	Donkey	HRP	RT · 1h	1:5000	GE Healthcare	AB_772210
Anti- rabbit	NA934	Donkey	HRP	RT · 1h	1:5000	GE Healthcare	AB_772206
Loading Controls							
GAPDH DyLight® 680	GA1R	Mouse	DyLight® 680	RT · 1h · Fluor	1:10000	Fisher Scientific	AB_2537657
Histone H1 AF® 790	AE-4	Mouse	AF790	RT · 1h · Fluor	1:5000	SCBT	AB_675641

**AF90** – Alexa Fluor 790; **Chemi** – Chemiluminescence; **CST** – Cell Signalling Technologies (London, UK); **Fluor** – Fluorescence; **HRP** – Horseradish peroxidase; **O/N** – Overnight; **RRID** – Research Resource Identifier; **SCBT** – Santa Cruz Biotechnology (Heidelberg, Germany)

### 2.6.7 Densitometry analysis

Relative changes in protein expression were semi-quantitatively measured using Image J software (Fiji; v.2.0.0.71; U. S. National Institutes of Health, Maryland, USA). Briefly, a region of interest (ROI) was drawn around a band of interest. From this, a histogram of peak intensity was generated, and a baseline of background intensity was set from the area surrounding the band within the ROI. Subsequently, the area under the curve was calculated to give an arbitrary intensity value. The band intensities of the protein of interest were normalised to the band intensity of the loading control for each sample. Furthermore, to compare protein expression between a test sample and the corresponding control, the first was normalised to the latter.

Unless otherwise stated, due to the limited availability of cord blood, and the high number of cells necessary for proteomic analysis, western blot analysis was only performed once (n=1).

## 2.7 Flow cytometry

### 2.7.1 Equipment and Analysis

For this study, a benchtop Accuri C6 Plus Flow Cytometer™ (Accuri Cytometers, BD) was used. The technical specifications of this instrument are outlined in **Table 2.7**. Data analysis was performed using FCS Express® ver.6 software (De Novo Software, Pasadena, USA).

### 2.7.2 Estimation of CD34<sup>+</sup> cell count

The percentage of CD34<sup>+</sup> in human cord blood samples was determined by performing an immunophenotyping assay, in parallel with the isolation protocol (**2.3.2**). Briefly, 100 µL of undiluted blood were stained with 10 µL of CD34- phycoerythrin (PE) and 5 µL of CD45-allophycocyanin (APC), in parallel with a control sample stained with 10 µL of IgG1-PE and 5 µL of CD45-APC (**Table 2.8**). Following an incubation at 4°C for 30 min, 5 mL of FACSTM Lysis Buffer (BD, Berkshire, UK) was added to each tube and incubated at RT for 10 min. The cells were washed in 5 mL 1x PBS, followed by a centrifugation at 200 x g for 5 min. and supernatant discarded. Cells were resuspended in residual supernatant and analysed by flow cytometry according to the gating strategy described in **Figure 2.3**.

**Table 2.7 – Technical specifications of the Accuri™ C6 Plus used in this study**

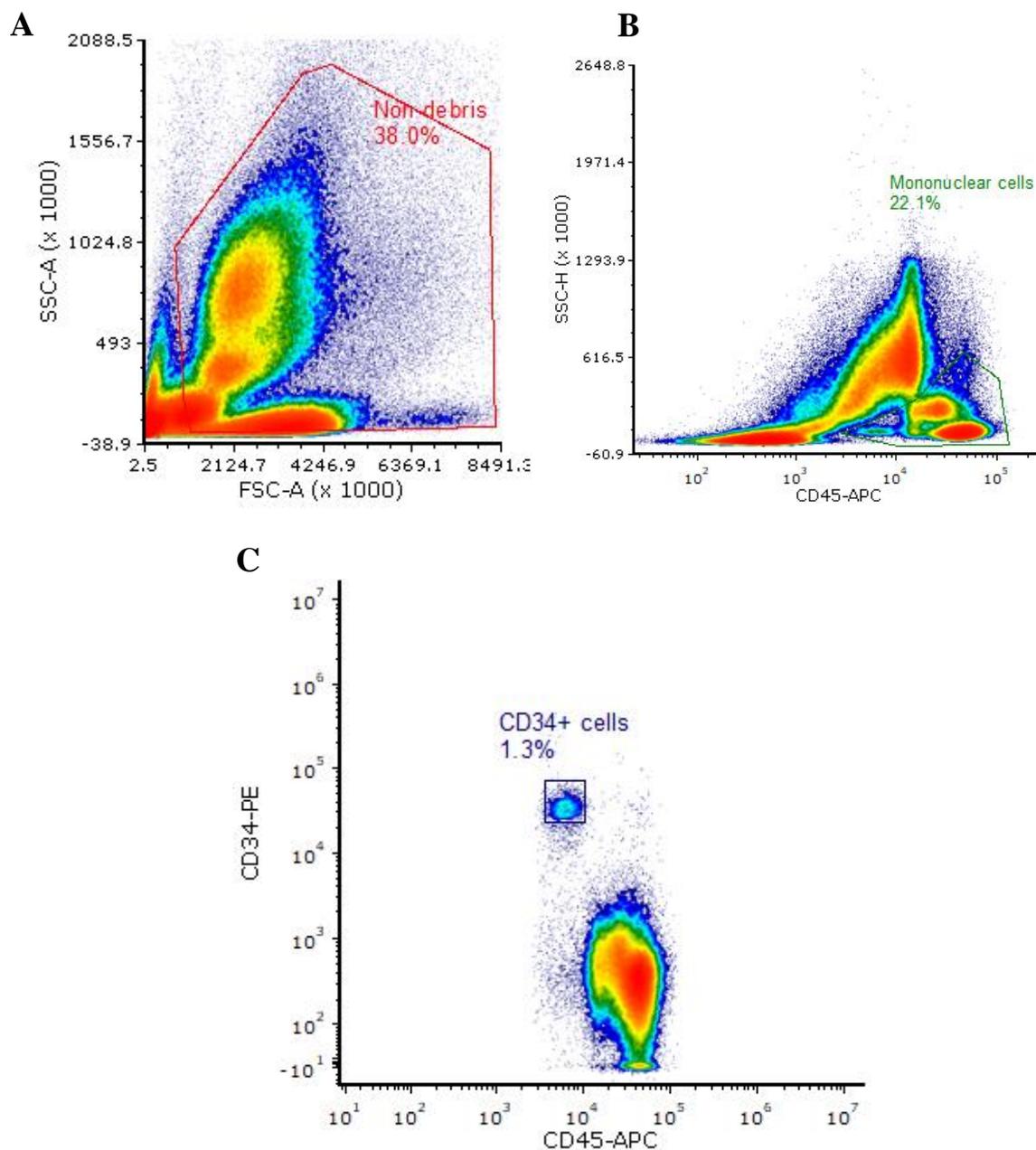
Fluorophore	Characteristics
Laser excitation	Red – 640 nm Blue – 488 nm
Emission Detection	FITC/GFP – 533/30 nm (FL1) PE/PI – 585/40 nm (FL2) PerCP-Cy™5.5 – > 670 nm (FL3) APC – 675/25 nm (FL4)
Flow rate	10-100 µL/min

**APC** - Allophycocyanin; **FITC** - Fluorescein isothiocyanate; **GFP** – Green fluorescent protein; **PE** – Phycoerythrin; **PI** – Propidium iodide; **PerCP-Cy5.5** – Peridinin chlorophyll protein – Cyanine5.5

**Table 2.8 – List of antibodies used for cord blood immunophenotyping**

Antibody	Fluorescent conjugate	Clone	Dilution	Source	RRID
IgG1	PE	MOPC-21	1/10	Biologend	AB_2847829
CD34	PE	581	1/10	Biologend	AB_1731862
CD45	APC	HI30	1/20	Biologend	AB_314400

**APC** - Allophycocyanin; **PE** – Phycoerythrin; **RRID** – Research Resource Identifier.



**Figure 2.3 – Estimation of CD34<sup>+</sup> HSPC in cord blood by flow cytometry**

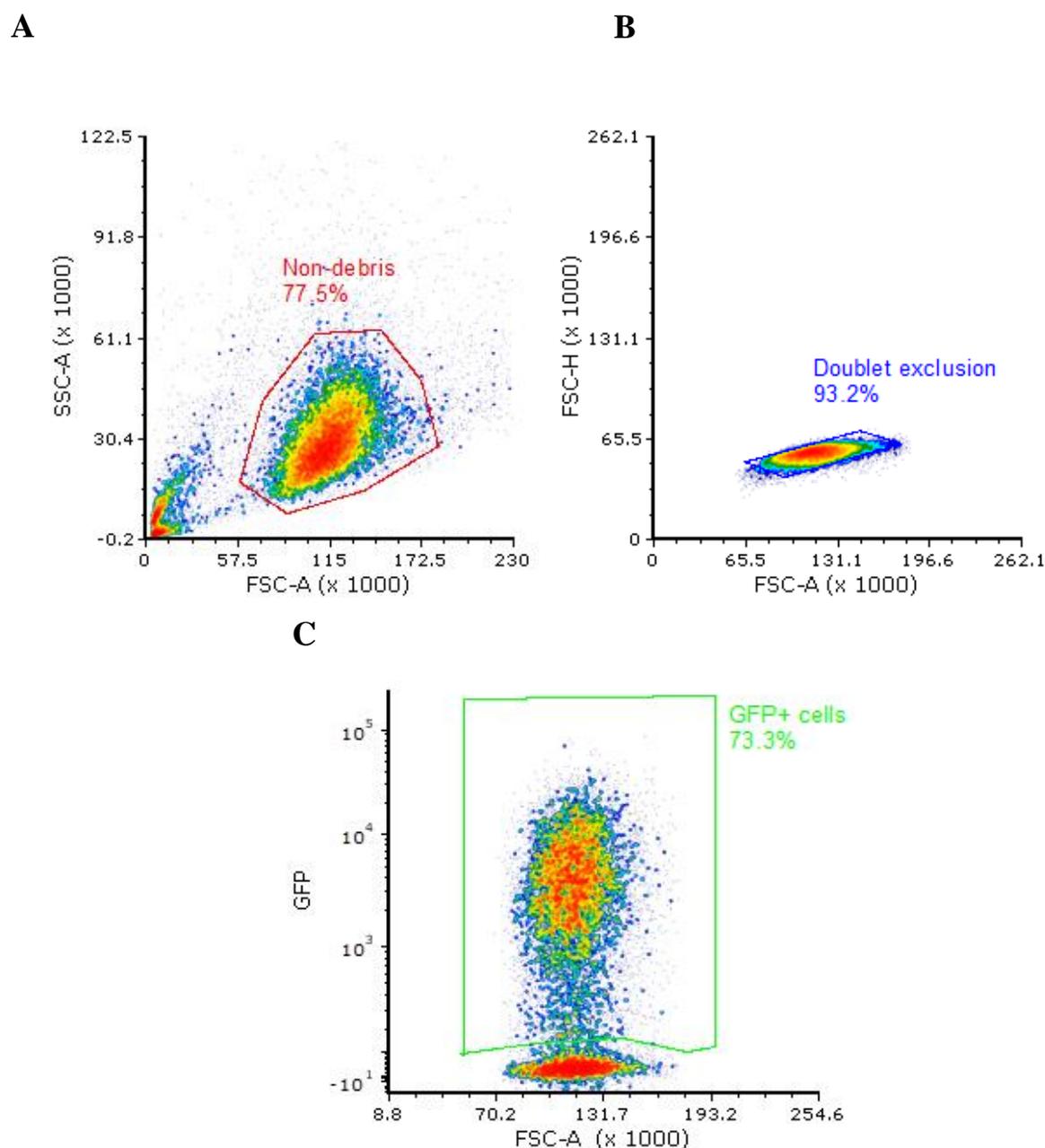
Representative bivariate dot plots showing the gating strategy used for the estimation of CD34<sup>+</sup> HSPC in human cord blood samples. (A) Example gate used to exclude cell debris, based on FSC and SSC; (B) Using the “Non-debris” gate, the expression of CD45 was used to identify mononuclear cells (CD45<sup>+</sup>); (C) The percentage of CD34<sup>+</sup> cells were estimated from within the mononuclear cell population. Background auto fluorescence was established using CD34<sup>+</sup> cells stained with the isotype control IgG1-PE (Table 2.8).

### 2.7.3 Fluorescence Activated Cell Sorting

Following transduction (2.5.2), CD34<sup>+</sup> HSPC were subjected to FACS for GFP positivity to ensure an enriched population (Figure 2.4), using a BD FACSAriaIII™, within the Central Biotechnology Services (CBS) facilities, Cardiff University. Cells were counted and centrifuged at 270 xg for 10 min. Supernatant was discarded and cells were resuspended at a final density of 1 x 10<sup>6</sup> cells/mL in PBS containing 1% BSA and 5 ng/mL IL-3, to minimize background. Resuspended cells were passed through a 40 µM cell strainer into 15 mL falcon tubes. Uninfected (mock) cells were used to set background autofluorescence. Cells were sorted at RT, using the 100 µm nozzle, at a low flow rate. Sorted cells were collected in 15 mL falcon tubes containing 8 mL of IMDM supplemented with 10% (v/v) FBS and 20 mg/mL gentamycin. Following sorting, cells were transferred to UCs and centrifuged for 5 min at 270 xg, at RT. Following centrifugation, supernatant was discarded, and cells resuspended in 1 mL 3S<sup>low</sup>G/GM (Table 2.5) for cell enumeration by flow cytometry (2.2.2). Cell density was adjusted as appropriate, for a final concentration of 2 x 10<sup>5</sup> cells/mL. Sorted cells were subsequently used in colony assays (2.3.4) and morphology slides (2.3.4).

### 2.7.4 Immunophenotypic analysis of transduced CD34<sup>+</sup> HSPCs

To assess myeloid cell differentiation, immunostaining was carried at different time-points using a panel of cell surface markers (Table 2.9). CD13-APC, in combination with CD36-biotin was used for lineage discrimination, along with streptavidin PerCP-Cy™5.5, in a two-step antibody labelling. Furthermore, cells were incubated with one on the following PE-labelled markers: IgG1, CD11b, CD14, CD15 and CD34. Briefly, 2.5 - 5 x 10<sup>5</sup> cells were washed and resuspended in 75 µL SB. Antibodies were added to a U-bottom 96-well plate, followed by the addition of 15 µL of cell suspension. Plates were briefly centrifuged to bring down the contents of each well, and carefully vortexed prior to incubation at 4°C for 30 min. Subsequently, 150 µL of SB was added to each well and its contents transferred to mini flow tubes. Cells were centrifuged at 200 x g for 5 min and supernatant discarded, following which samples were labelled with streptavidin-peridinin chlorophyll protein (SA-PerCP-Cy™5.5; BD Pharmingen, USA) for 30 min at 4°C. Stained cells were washed and resuspended in 100 µL of SB for analysis using the Accuri™C6 Plus cytometer (2.7.1), using the gating strategy outlined in Figure 2.5 and the appropriate compensation settings (Supplementary Table 1).



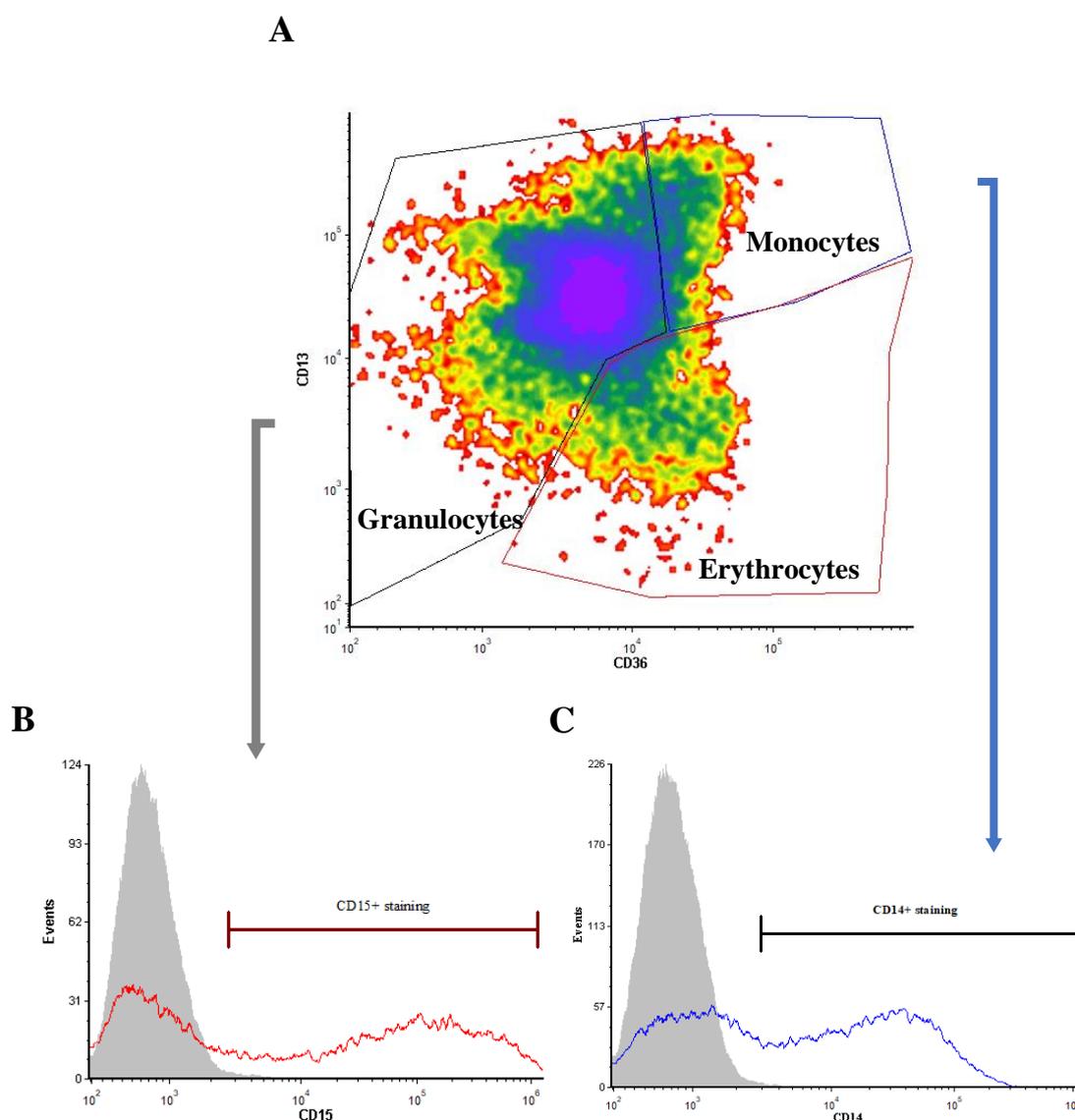
**Figure 2.4 – Gating strategy used to sort GFP-expressing CD34<sup>+</sup> HSPC**

Representative bivariate dot plots showing FACS-enriched CD34<sup>+</sup> HSPC cultures for transduced cells, according to GFP positivity, using FACSariaIII™. **(A)** Example gate used to exclude cell debris, based on FSC and SSC; **(B)** Example gating used to exclude doublets, based on the previously established “Non-debris”. **(C)** The percentage of GFP<sup>+</sup> cells within a bulk culture was established based on the “Non-debris” and “Doublet exclusion” gates shown in A and B. Background auto fluorescence was established using cells subjected to the equivalent viral infection procedure but in the absence of virus (for CD34<sup>+</sup> cultures – mock culture).

**Table 2.9 – Antibodies used for immunophenotypic analysis of transduced CD34+ HSPC**

Antibody	Fluorescent conjugate	Clone	Dilution	Source	RRID
IgG1	PE	MOPC-21	1/17	Biologend	AB_2847829
CD11b	PE	ICRF44	1/17	Biologend	AB_314158
CD13	APC	WM15	1/17	Biologend	AB_314182
CD14	PE	HCD14	1/17	Biologend	AB_830679
CD15	PE	W6D3	1/17	Biologend	AB_756012
CD34	PE	581	1/17	Biologend	AB_1731862
CD36	Biotin	SMO	1/100	Ancell Corporation	Not defined
PerCP-Cy <sup>™</sup> 5.5 Streptavidin	PerCP-Cy <sup>™</sup> 5.5	-	1/100	BD Biosciences	AB_2868934

**APC** - Allophycocyanin; **PE** – Phycoerythrin; **PerCP-Cy5.5** – Peridinin chlorophyll protein – Cyanine5.5; **RRID** – Research Resource Identifier.



**Figure 2.5 – Immunophenotypic analysis of CD34<sup>+</sup> HSPC**

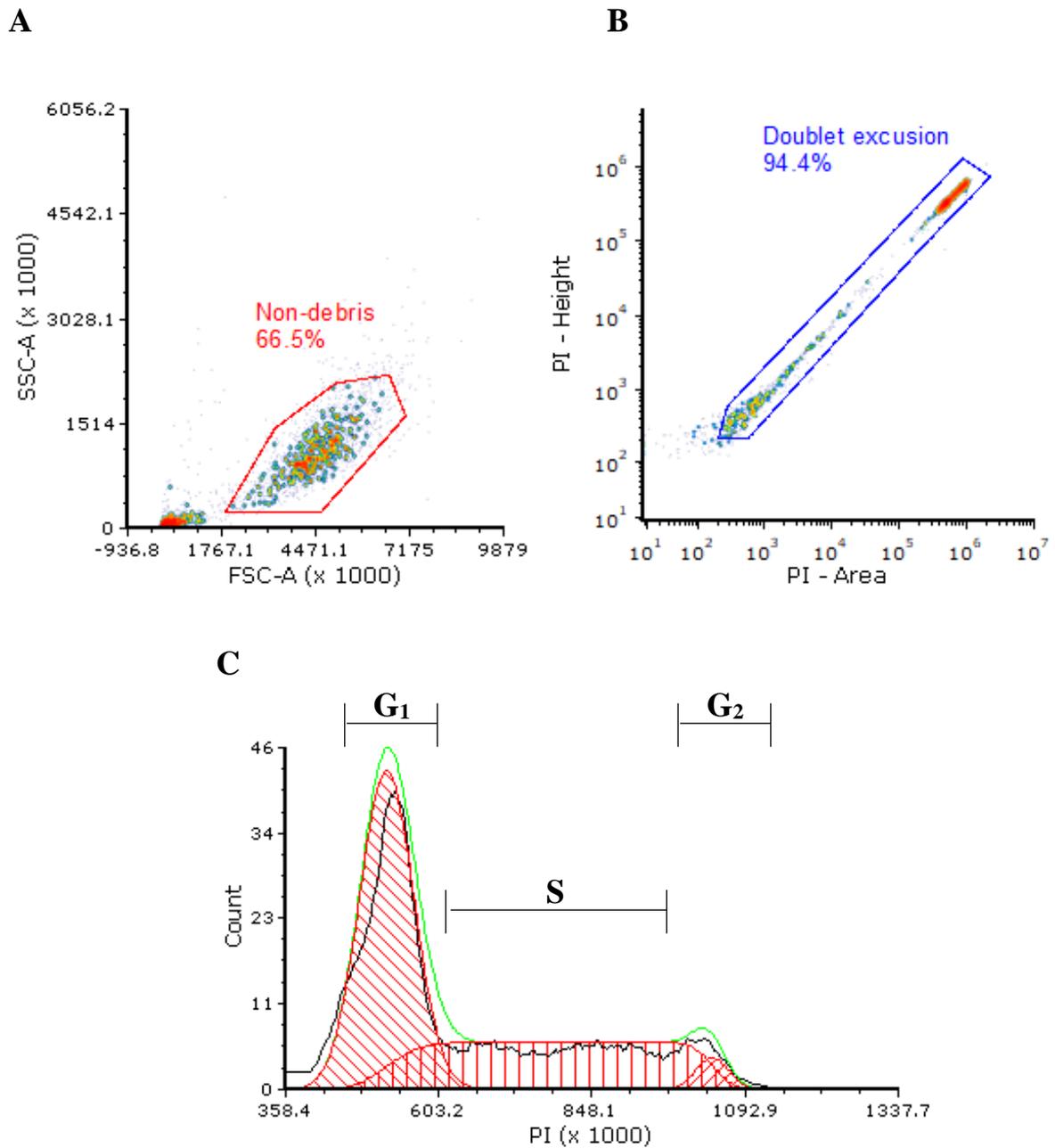
Representative bivariate plots and histograms outlining the gating strategy used for immunophenotypic analysis of CD34<sup>+</sup> HSPC throughout haematopoietic development. Cells were gated according to the “Non-debris” gate showed in **Figure 2.4A**. Appropriate compensation was applied before the analysis. **(A)** The CD13 and CD36 markers were used to discriminate between the granulocytic, monocytic and erythroid lineages within GFP<sup>+</sup> population (**Figure 2.4C**). **(B-C)** Each subpopulation was further examined for differentiation markers, such as the granulocytic marker CD15 or monocytic marker CD14 (**Table 2.9**) (Isotype - filled histogram; CD14/CD15 – open histogram).

### 2.7.5 Cell cycle analysis

To determine changes in CD34<sup>+</sup> HSPC cell cycle, propidium iodide (PI, Sigma-Aldrich) was used to label intracellular DNA. For this analysis,  $5 \times 10^4$  cells were transferred to a UC and washed with 20 mL of 1x PBS. Following centrifugation, supernatant was discarded, and cultures resuspended in 300  $\mu$ L of PBS, following which cells were fixed for 30 min on ice by adding 700  $\mu$ L absolute ethanol. After fixation, samples were stored at  $-20^{\circ}\text{C}$  for an O/N incubation. The following day, cells were washed with 10 mL of 1x PBS and centrifuged at 270 xg for 10 min. Supernatant was discarded and cells were resuspended in 75  $\mu$ L 1x PBS. Subsequently, cells were incubated with 25  $\mu$ L staining solution, containing 40  $\mu\text{g}/\mu\text{L}$  PI and 0.1 mg/mL RNase (Sigma-Aldrich) diluted in 1x PBS, at  $37^{\circ}\text{C}$  for 30 min. Samples were acquired within 20 min following incubation using the Accuri<sup>TM</sup>C6 Plus cytometer (2.7.1). Cell cycle analysis was performed using FCS Express's Multicycle AV DNA analysis tool plug-in, according to the gating strategy outlined in **Figure 2.6**.

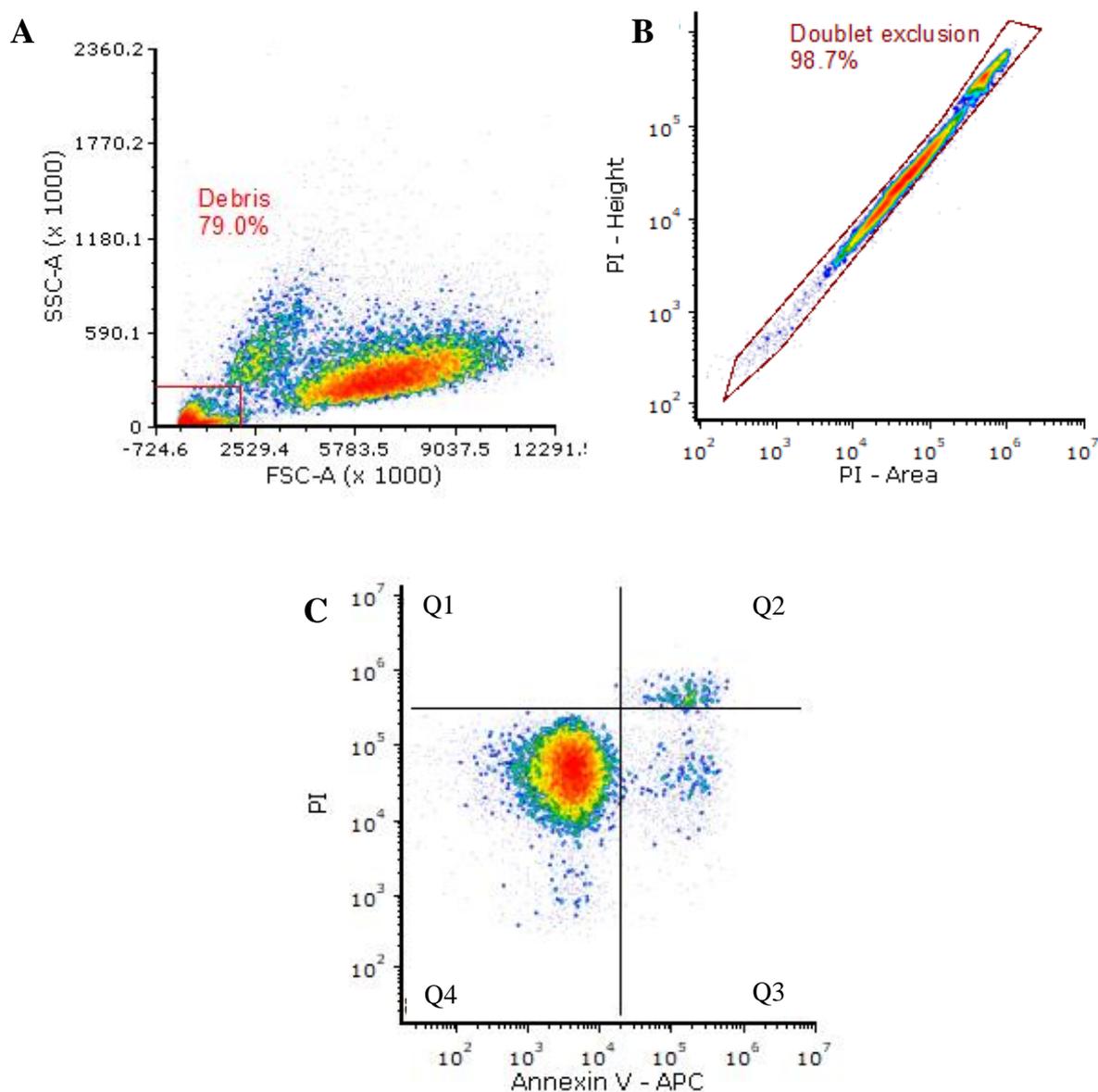
### 2.7.6 Apoptosis

To determine the proportion of apoptotic cells, an apoptosis assay using Annexin V (ThermoFisher) was performed. Annexin V is a calcium-dependent phospholipid-binding protein, that binds to phosphatidylserine (PS), usually located in the inner leaflet of the plasma membrane. When a cell becomes apoptotic, PS is translocated to the extracellular membrane leaflet, and is detected by the fluorescent labelled Annexin V in a calcium-dependent manner. To perform this assay,  $5 \times 10^4$  cells were transferred onto a labelled UC and washed with 1x PBS, followed by an additional wash with 1x Annexin V Binding Buffer (ThermoFisher). Supernatant was discarded and cells resuspended in 50  $\mu$ L Binding Buffer. Afterwards, 5  $\mu$ L of APC-conjugated Annexin V was added and cells incubated for 10-15 min at RT. Following incubation, cells were washed with 1x Binding Buffer and resuspended in 100  $\mu$ L of the same solution. 5  $\mu$ L of PI or 7-AAD were added to the UCs and cells analysed by flow cytometry (2.7.1), using the setting described in **Figure 2.7**.



**Figure 2.6 – Gating strategy used to perform cell cycle analysis**

Representative bivariate plots and histogram describing the gating strategy used for the analysis of cell cycle status by flow cytometry. **(A)** Gating strategy used to exclude cell debris based on FSC and SSC; **(B)** Gating strategy used to exclude doublet and cell aggregates based on Height and Area parameters. **(C)** Histogram showing DNA content analysis, and determination of the proportion of cells within the G<sub>1</sub>, S and G<sub>2</sub> phases on cell cycle.



**Figure 2.7 –Gating strategy used to analyse apoptosis**

Representative bivariate dot plots showing (A) gating strategy used to identify cell debris and exclude them for further analysis, based on FSC and SSC; (B) gating strategy used to exclude doublet cells, based on PI-height and PI-area; (C) quadrant gating used to determine viable (Q4), pre-apoptotic (Q3) and late apoptotic cells (Q2).

### 2.7.7 Growth and proliferation assay

To measure growth and proliferation of cell lines, a proliferation/viability assay was performed. Cells in log phase growth were pelleted in a UC and washed in 10 mL of serum free medium. Cells were seeded at a concentration of  $1 \times 10^5$  cells/mL in a 24-well plate. Following seeding, 50  $\mu$ L of cell suspension was recovered and transferred to FACS tubs to confirm cell density. In addition, 1  $\mu$ L of 5  $\mu$ M TOPRO-3 was added to each tube to check cell viability, as described in 2.2.3.

The following day, cells in each well were resuspended and 50  $\mu$ L of cell suspension were transferred to FACS tubes. To each tube, 1  $\mu$ L of TO-PRO-3 was added, and suspension mixed. The Accuri Plus B6 flow cytometer was used to count the cells and cell's viability was calculated for each sample (2.2.2). This process was repeated for 5 consecutive days.

## 2.8 Proteomic analysis using SWATH<sup>®</sup>

Sample preparation was performed at Cardiff University. Sample processing was performed during a visit to Manchester University, whilst analysis was performed in the Stoller Biomarker Discovery Centre in the University of Manchester by Dr. Andrew Pierce and Dr. Bethany Geary, in Prof. Anthony Whetton's group.

### 2.8.1 Sample preparation

Human cord blood-derived CD34<sup>+</sup> HSPC were isolated as described in 2.3.2. Subsequently, cells were transduced with either a control (PINCO GFP) or a RUNX1-ETO PINCO vector (n=3), according to 2.5.2. On the third day of cell culture, efficiency of infection was assessed through GFP expression (2.7), following which nuclear and cytosolic proteins were extracted. Following protein quantification (2.6.3), samples were stored at -80°C until processing at the University of Manchester (2.8.2).

### 2.8.2 Tryptic digestion

Samples were prepared by diluting 10  $\mu$ g of extract 30  $\mu$ L of 30 nM Ammonium Bicarbonate (AmBic). Disulphide bonds were reduced by adding 3  $\mu$ L of 50mM TCEP, followed by an incubation at 60°C for 1 h. Subsequently, proteins were alkylated by adding 3.3  $\mu$ L of 60mM iodoacetamide and incubated at RT for 30 min in the dark. Lastly, proteins were

digested by adding trypsin at a 10:1 substrate:enzyme ratio (i.e. by adding 10 µg of trypsin per 100 µg of protein). Samples were incubated overnight at 37°C.

### 2.8.3 Strong Cation Exchange Clean-up

To remove contaminating solutions from the digestion process and to purify the sample, a UltraMicroSpin™ (The Nest Group, Massachusetts, USA; [URL: <https://www.nestgrp.com>]) column based on Strong Cation Exchange (SCX) was used, according to the manufacturers protocol. Columns were conditioned by adding 100 µL of 100% methanol and centrifuged for 1 min at 110 x g. The columns were subsequently washed with 50 µL of water and centrifuged 1 min at 110 x g. After discarding the liquid, 100 µL of a 0.2M monosodium phosphate supplemented with 0.3M sodium acetate solution (pH 3-6.5) was added, and columns centrifuged for 10 sec to allow for the solution to be in contact with the filter. Columns were left for 1 h at RT, following which the solution was eluted by spinning the columns for 1 min at 110 xg. The final step of conditioning consisted of washing the columns with 100 µL of water, followed by a new step of centrifugation. The columns were then equilibrated by adding 100 µL of 25mM ammonium formate in 85% acetonitrile and centrifuged for 1 min at 110 xg.

In the meantime, samples were placed in a miVac DNA concentrator for 1 h to allow cell pellets to dry and solvents to evaporate. Tryptic peptides were afterwards resuspended in 100 µL of 25 mM ammonium formate in 85% acetonitrile and added to the column, which was centrifuged for 1 min at 110 xg. After discarding the supernatant, columns were washed by adding 50 µL of 25mM ammonium formate in 85% acetonitrile. Samples were eluted by placing the columns onto new 2mL microcentrifuge tubes, and by adding 30 µL of 15 mM formic acid, followed by a 1 min centrifugation at 110 xg. Elution was repeated one more time to maximize sample yield.

### 2.8.4 Relative quantification by SWATH acquisition

Samples were analysed using SCIEX (Framingham, Boston, MA, USA) TripleTOF 6600 instrumentation coupled on-line to an Ultimate 3000 HPLC (Dionex), according to the standard operating procedures of the Stoller Biomarker Discovery Centre (Manchester, UK). The analytical column used was an Eksigent LC system with a nanoLC 400 autosampler and nanoLC 425 pump module. Peptides were eluted using a gradient from a 98% water, 2%

acetonitrile and 0.1% formic acid buffer to an 80% (*v/v*) acetonitrile, 20% (*v/v*) water, 0.1% (*v/v*) formic acid buffer over 120 minutes at 0.3uL/min. Nuclear and cytosolic samples from control and RUNX1-ETO cells were analysed using data-independent acquisition (DIA) method. Raw spectral data files were analysed using openSWATH (version 2.0.0) using the Pan Human Library (Rosenberger *et al.*, 2014) followed by scoring and filtering using pyProphet (version 0.18.3). Resultant output files were combined using the MSproteomicstools feature alignment script. Further filtering and protein abundance summarisation and normalisation were performed in R (version 3.4.1) using the packages SWATH2stats and MSstats from Bioconductor (release 3.5). Statistical significance was analysed according to section 2.8.5, and pathway analysis according to section 2.9.2.

### 2.8.5 SWATH-MS statistical analysis

Partek<sup>®</sup> Genomics Suite<sup>®</sup> (v. 7.0, St. Louis, USA) was used to perform statistical analysis on the protein expression dataset. To do this, detected proteins were partitioned according to their sub-cellular localisation (cytoplasmic or nuclear). Subsequently, only proteins that were commonly detected in all 3 biological replicates (for each condition) were used to perform statistical analysis. An ANOVA was used to determine statistically dysregulated protein changes between RUNX1-ETO expressing and control cells, and a cut-off value of  $p < 0.05$  was applied to the analysis. Mean values of protein expression were used for calculation of fold-change (FC). No fold-change cut offs were applied in this analysis. Cytosolic and nuclear proteins were then combined to generate a new list for pathway and network analysis, performed using Metacore<sup>™</sup>, according to (2.10).

## 2.9 Analysis of existing mRNA expression data

### 2.9.1 Analysis of microarray data of RUNX1-ETO induced changes in mRNA expression in CD34<sup>+</sup> HSPC

Previously, the effects of aberrant transcriptional activity of RUNX1-ETO were analysed by Affymetrix<sup>®</sup> microarray gene expression in a study performed by Tonks *et al.* (Tonks *et al.*, 2007). Dysregulated genes arising from this analysis were downloaded from <https://www.nature.com/articles/2404961#supplementary-information>, Supplementary Table 2, accession number E-MEXP-583. Dysregulated probe-sets were converted to Entrez Gene ID using the DAVID tool (Database for Annotation, Visualization and Integrated Discovery ver.

6.7) (Huang *et al.*, 2009a). Pathway and Interactome analysis were performed using Metacore™, as described below (2.10).

## 2.9.2 Publicly available databases

### 2.9.2.1 Bloodspot

mRNA expression in human haematopoietic cells for the different targets mentioned in this study was accessed through the online database tool Bloodspot (Bagger *et al.*, 2016) using the ‘Normal haematopoiesis with AML’ dataset, and the corresponding probesets (URL: <http://servers.binf.ku.dk/bloodspot/>) (GSE42519; (Rapin *et al.*, 2014)).

### 2.9.2.2 cBioPortal

In addition, mRNA expression was also assessed in AML patients through cBioPortal (Cerami *et al.*, 2012; Gao *et al.*, 2013) (URL: <https://cbioportal.org/>). Analysis was performed based on the Cancer Genome Atlas (TCGA), NEJM 2013 dataset, comprising of whole-genome or whole-exome sequencing of 163 adult *de novo* AML samples (Ley *et al.*, 2013). For patient survival analyses, patient samples were excluded based on (i) untreated patients and/or (ii) diagnosed with APL. The remaining patients were stratified according to mRNA expression of the gene of interest into the upper (high expression) or lower (low expression) quartiles. Furthermore, patients were analysed based in clinical attributes associated with high and low mRNA target expression, including FAB classification, cytogenetic or molecular abnormalities, and overall-survival. All mRNA expression data within cBioPortal is represented as RNASeq RSEM.

## 2.10 Meta-analysis of mRNA and proteomic data

Metacore™ (Clarivate™ Analytics, UK), is a web-based bioinformatics tool used for the analysis of multiple omics data, including microarray, RNASeq and proteomics datasets. This analysis is based on published peer-reviewed literature combined into a curated knowledge database, *MetaBase*, covering over 2700 scientific journals. In all the studies described below, the following elements were selected for the input dataset: for the mRNA analysis, Entrez Gene ID identifiers were used, as well as their corresponding fold-change values; for the protein analysis, the protein name and the corresponding p-value were used.

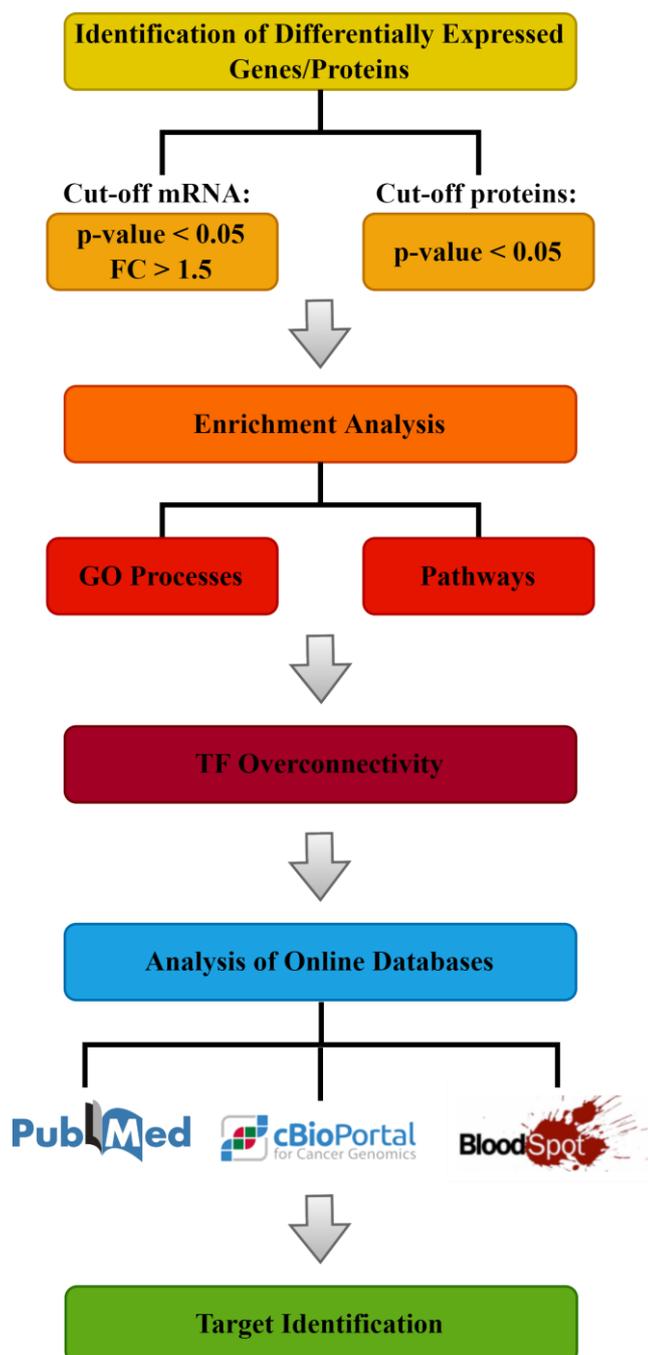
Ingenuity Pathway Analysis (IPA) is another web-based software application, alike Metacore™. Similarly, it allows the analysis and interpretation of microarray and proteomics data, by performing not only pathway analysis, but also by identifying key regulators of a specific dataset, predicting downstream effects, and providing potentially useful target data. IPA is based on the Ingenuity Knowledge Database, derived from manually curated literature from over 300 scientific journals, and abstracts from more than 3600 peer-reviewed journals. The workflow used to analyse both datasets has been described in **Figure 2.8**.

### 2.10.1 Pathway analysis using Metacore

The Metacore™ Pathway Analysis tool was used to perform an Enrichment analysis of the dataset of dysregulated genes/proteins. The background of the analysis of the microarray dataset, defined as the base gene list in which all analyses was performed, was defined as “Affy // U133A”, corresponding to the Microarray GeneChip® previously used for the mRNA analysis. The resulting pathways were ranked according to their p-value, based on a hypergeometric distribution. Metacore™ also uses p-value calculation to evaluate a network’s relevance to Gene Ontology biological processes classification. The p-value is calculated as follows:

$$pVal(r, n, R, N) = \sum_{i=\max(r, R+n-N)}^{\min(n, R)} P(i, n, R, N) = \frac{R!n!(N-R)!(N-n)!}{N!} \sum_{i=\max(r, R+n-N)}^{\min(n, R)} \frac{1}{i!(R-i)!(n-i)!(N-R-n+i)!}$$

The set of all nodes in the Metacore™ database of interactions represent the “Global network” of size **N**. Respectively, a set of nodes corresponding to our dataset is **R** and set of nodes in the network under evaluation is **n**, considering that **r** nodes in **n** turned out to be “marked” by association with user’s list. **R** and **r** could also represent nodes marked by their association with some other list, such as “process” category from Gene Ontology. For the evaluation of statistical significance, Metacore™ considers the null hypothesis which states that the subsets **R** and **n** are independent and, therefore, the size of their intersection follows the hypergeometric distribution. The alternative hypothesis states that there is positive correlation between the subsets. Based on these assumptions, Metacore™ is able to calculate a p-value as the probability that intersection of two randomly selected sub-sets of **N** would have having the size of **n** or larger.



**Figure 2.8 - Strategy for the analysis of differentially expressed genes/protein to study the effects of RUNX1-ETO overexpression in human CD34<sup>+</sup> HSPC**

Diagram illustrating the strategy employed to analyse the consequences of RUNX1-ETO expression in CD34<sup>+</sup> HSPC. Differentially expressed genes had been previously filtered according to Tonks *et al.* using p-value < 0.05 and a fold-change cut-off of 1.5 (Tonks *et al.*, 2007); differentially expressed proteins were identified by applying a cut-off of p-value < 0.05. Enrichment analysis was performed using Metacore and IPA (for mRNA analysis); TF overconnectivity was performed using Metacore.

The p-value was set with a False Discovery Rate (FDR) of <0.05 (no more than 5% false positives) and used to estimate the probability of a random intersection between the input probe set IDs with ontology gene IDs within the Metacore™ database. The lower p-value, the higher relevance of the gene ontology IDs to the dataset, showing as a higher rating for the given probe sets of the corresponding pathway.

### 2.10.2 Interactome analysis using Metacore

The Metacore™ Interactome tool was used to identify genes/proteins overrepresented within the given datasets. This tool provides an estimation of the level of interconnectivity within the input dataset based on published literature. This algorithm estimates statistically significant interactions in the set, and enrichment of the dataset according to certain genes/proteins.

In this study, to identify statistically significant TF responsible for the regulation of other genes/proteins, the algorithm “Interactome Transcription Factor” was used. The results were ranked by p-value using the basic formula for a hypergeometric distribution, as described in 2.10.1. Metacore™ provides an alternative statistical analysis based on the level of over- and under-connectivity of each TF, designated z-score. The z-score measures the difference between the obtained number of proteins and the expected average number of proteins corresponding to genes expressed in units of standard dispersion, and is calculates as below:

$$Z - \text{score} = \frac{r - n \frac{R}{N}}{\sqrt{n \left(\frac{R}{N}\right) \left(1 - \frac{R}{N}\right) \left(1 - \frac{n-1}{N-1}\right)}}$$

Where **r** represents the number of proteins derived from the current protein list that have interactions with given protein; **R** represents the total number of proteins in the GeneGo global network that have interactions with the given protein; **n** refers to the total number of proteins in given protein list and **N** is based on the total number of proteins in the GeneGo global network. All analyses were performed considering an FDR of < 0.05

---

### **2.10.3 Pathway analysis and identification of transcription factors using Ingenuity Pathway Analysis**

Similarly, IPA was used to analyse the microarray dataset. For the microarray analysis, the background was defined before any analysis was made as Affy // U133A. To identify TF, Upstream regulator analysis algorithm was used, allowing to predict the regulators responsible for the changes in gene expression. Once more, genes were ranked according to their p-value, calculated using a Right-Tailed Fisher's Exact Test, reflecting the likelihood that the association or overlap between a set of significant molecules from our dataset and a given process/pathway/transcription neighbourhood is due to random chance. The smaller the p-value the less likely that the association is random.

## **2.11 Statistical analysis**

Statistical analysis was performed using a paired sample t-test, or one-way ANOVA with Bonferroni's multiple test correction, unless otherwise stated. Values of  $p < 0.05$  were considered statistically significant. All statistical analyses were conducted using GraphPad Prism (ver. 8) (GraphPad Software, California, USA). Details regarding the statistical tests used are labelled in figure legends.

# Chapter 3

---

## **RUNX1-ETO Induces Changes in mRNA and Protein Expression of Transcription Factors**

### 3.1 Introduction

AML is characterised by a block in terminal myeloid cell differentiation resulting in the accumulation of non-differentiated leukaemic cells in the BM and PB. It is widely accepted that abnormal activity of critical genes, such as TFs, which are involved in haematopoietic cell development, survival, proliferation, and maturation contribute to the development of this disease (Saultz and Garzon, 2016; De Kouchkovsky and Abdul-Hay, 2016). Chromosomal translocations are one of the most common cytogenetic abnormalities in AML and result in the expression of abnormal oncogenic fusion proteins (**1.2.2**). These translocations almost invariably consist of at least one TF protein (or a shortened truncated form) frequently involved in the regulation of myeloid cell differentiation (Lee *et al.*, 2006). Consequently, AML-associated fusion proteins function as aberrant transcriptional regulators with the potential to interfere with the normal processes of haematopoietic differentiation.

One of the most common translocations in AML is the t(8;21), which results in the expression of the oncogenic fusion protein RUNX1-ETO (Rowley, 1990; Reikvam *et al.*, 2011). Expression of this protein is the product of the fusion of *RUNX1* (chromosome 21), a critical TF in haematopoietic development, with *ETO* (chromosome 8) (**1.3.3**). To investigate the *in vivo* mechanisms through which RUNX1-ETO contributes to leukaemogenesis, Okuda *et al.*, generated a murine model containing a single allele of RUNX1-ETO whose expression was regulated by the endogenous transcriptional regulatory elements of murine RUNX1 (Okuda *et al.*, 1998). The study showed that RUNX1-ETO not only disrupts normal RUNX1-mediated transcriptional activity, but also contributes to the generation of abnormal haematopoietic progenitor cells. In addition, mice that are heterozygous for the RUNX1-ETO allele are foetally incompatible with life due to haemorrhaging in the central nervous system and exhibit a severe block in foetal liver haematopoiesis (Yergeau *et al.*, 1997). *In vitro* studies using an experimental model based on normal human primary HSPC (CD34<sup>+</sup> cord blood derived) enabled the analysis of the effects of RUNX1-ETO expression as a single abnormality on normal human haematopoiesis. These studies showed that ectopic expression of this fusion protein was able to inhibit the development of haematopoietic cells, promote the growth of immature blood cells and increase self-renewal – hallmarks of AML (Tonks *et al.*, 2003; Tonks *et al.*, 2004; Schnerch *et al.*, 2012; Mulloy *et al.*, 2002).

Increasing data has shown that developmental abnormalities observed in the presence of the RUNX1-ETO fusion protein disturbs one or more transcriptional processes that regulate haematopoietic development (1.3.3.4). To understand the mechanism through which this fusion protein blocks haematopoietic development, transcriptomic analysis was performed using DNA oligonucleotide microarrays (Tonks *et al.*, 2007). Using the human cord blood-derived model, Tonks *et al.*, found 380 genes to be dysregulated in the presence of the RUNX1-ETO fusion protein when compared to control cells. The analysis identified several dysregulated genes known to be implicated in AML and in the arrest of cell development, including *PU.1* and *CEBPA*; both widely described in the literature as master regulators of haematopoiesis (Nerlov and Graf, 1998; Koschmieder *et al.*, 2005). Furthermore, gene changes were also observed throughout myeloid cell development, including monocytic, granulocytic and erythroid lineages. These dysregulated genes included *Sox4*, *CD200* and *γ-catenin*, suggesting that the effect of RUNX1-ETO on gene expression in human haematopoietic progenitor cells is highly context-dependent (Tonks *et al.*, 2007; Morgan *et al.*, 2013; Damiani *et al.*, 2015; Lu *et al.*, 2017).

Further studies have since been performed to try to understand the molecular mechanism through which this fusion protein contributes to leukaemia development, by analysing the transcriptome or proteome of RUNX1-ETO expressing cells. Ptasinska *et al.* performed a genome-wide association study in t(8;21) AML patients coupled with RUNX1-ETO expressing cell lines (Kasumi-1 and SKNO-1) to determine the alterations in the epigenetic landscape caused by the expression of this protein (Ptasinska *et al.*, 2012). The authors were able to demonstrate that the expression of genes involved in haematopoietic differentiation and self-renewal is controlled by RUNX1-ETO; these genes also included the well-established TF *C/EBPα*, *PU.1* and *NFE2*. The same study showed that knockdown of RUNX1-ETO using siRNA, led to an epigenetic reprogramming and a genome-wide redistribution of RUNX1 binding, influencing processes involved in myeloid differentiation, proliferation, and self-renewal. Furthermore, RUNX1-ETO knockdown in primary t(8;21) patient cells and cell lines resulted in the upregulation of genes involved in myeloid differentiation, whilst downregulated genes were shown to be involved in cell proliferation and cell cycle progression, resulting in the inversion of the typical t(8;21) phenotype. An additional study performed by Martinez-Soria *et al.* using an RNA interference (RNAi) approach identified Cyclin D2 (*CCND2*) as a crucial transmitter of RUNX1-ETO-driven leukaemic propagation (Martinez-Soria *et al.*,

2018). Moreover, Singh *et al.* identified RUNX1-ETO mediated changes in protein expression by inducing the expression of this oncoprotein in a Tet-off-inducible U937 cell line coupled with mass spectrometry analysis (Singh *et al.*, 2010). The authors showed that in these cells, the protein profile is drastically changed due to the expression of RUNX1-ETO, and were able to identify several changing proteins, including NM23 and HSP27. These specific changes were also observed in AML patient samples (Singh *et al.*, 2010).

However, the above studies have primarily focused on transformed AML cell lines, and in general gene/protein expression, using an unsupervised analysis of transcriptomic/proteomic changes. Given that transcriptional regulation is likely mediated by changes in TF expression, this study will re-analyse transcriptomic data obtained from microarray studies performed by Tonks *et al.*, and use a supervised approach to focus specifically on changes observed in TF. Furthermore, whilst mRNA abundance can be used as a strategy for target identification, it is not a powerful predictor of protein expression. For this reason, this study will use both TF transcriptomics and proteomics expression data to identify potential targets of the phenotype observed in t(8;21), by performing quantitative MS. In addition, no study has specifically examined the cytosolic and nuclear protein expression profile of CD34<sup>+</sup> HSPC upon the expression of the RUNX1-ETO fusion protein, focusing instead on changes to total protein expression. Therefore, MS analysis will be employed on cytosolic and nuclear fractions of RUNX1-ETO expressing human CD34<sup>+</sup> HSPC to identify, firstly, changes in TF expression, and secondly, aberrant subcellular protein localisation that could help explain the phenotypic changes observed in AML t(8;21). Analysis of transcriptomic and proteomic data sets will be used to investigate similarities between differentially expressed genes and proteins, as a result of the expression of the RUNX1-ETO fusion protein. However, it is important to take into consideration the well-recognised delay between mRNA expression and protein production, alongside the short half-life of mRNA against the extended half-life of some proteins. This understanding leads to the expectation that, despite both analysis being performed following three days of HSPC transduction, some measure of discrepancy will inherently remain.

### **3.2 Hypothesis and Aims**

I hypothesise that the RUNX1-ETO fusion protein dysregulates TF that mediate the block in haematopoietic cell differentiation. The overall aim of this chapter is to identify changes in

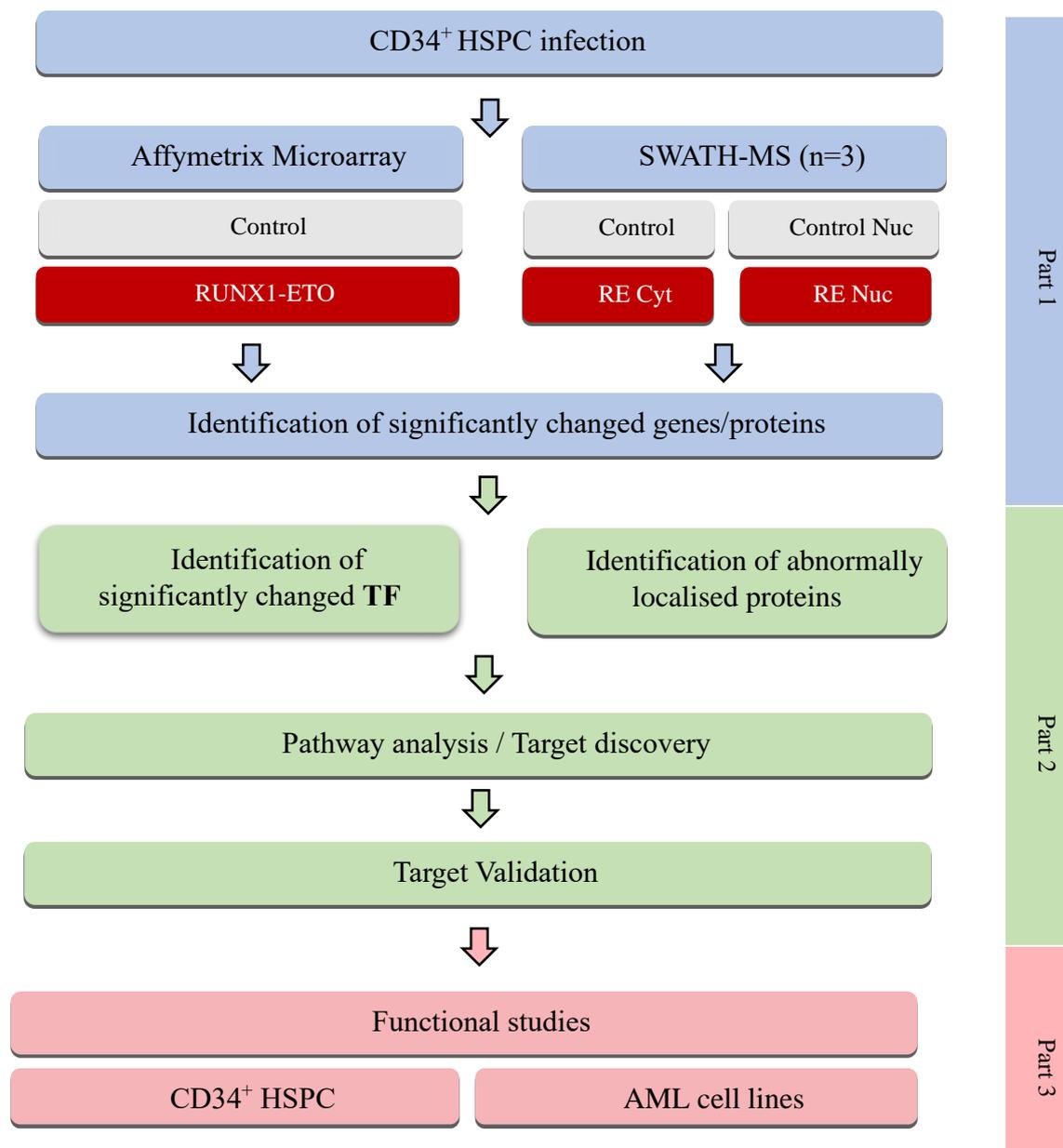
the expression of TFs relevant in the context of t(8;21) AML by analysing the transcriptome and proteome of human CD34<sup>+</sup> HSPC expressing RUNX1-ETO (**Figure 3.1**). The specific objectives are:

**To determine RUNX1-ETO induced changes in TF mRNA expression using normal human CD34<sup>+</sup> HSPC.**

Pre-existing Affymetrix transcriptome data relating to RUNX1-ETO-induced changes in mRNA gene expression in human CD34<sup>+</sup> HSPC is available (Tonks *et al.*, 2007). This dataset will be analysed to identify changes in TF mRNA abundance in CD34<sup>+</sup> HSPC expressing RUNX1-ETO, compared to normal CD34<sup>+</sup> HSPC controls. This will be achieved using pathway analysis software Metacore™ and IPA®, which are manually curated databases used to analyse omics data. In addition, identification of dysregulated genes will be compared to publicly available transcriptomic data from AML patients to identify clinically relevant changing TF mRNA.

**To determine RUNX1-ETO induced changes in TF protein expression using normal human CD34<sup>+</sup> HSPC.**

To quantify significant changes in TF protein expression as a result of RUNX1-ETO expression, SWATH-MS analysis (2.8) will be performed on cytosolic and nuclear extracts of human CD34<sup>+</sup> HSPC expressing RUNX1-ETO and compared to control. This will allow the identification of TFs inappropriately localised in either subcellular compartment, thus interfering with normal cellular processes. Pathway analysis using Metacore™ will be used to identify mis-regulated proteins arising from the expression of RUNX1-ETO. Furthermore, examination of the cytoplasmic to nuclear protein expression ratio in control and RUNX1-ETO CD34<sup>+</sup> HSPC will allow the identification of differential protein expression with regard to these subcellular compartments.



**Figure 3.1 – Strategy for the identification of changes in gene/protein expression arising from the expression of the fusion protein RUNX1-ETO**

Flow diagram showing the experimental strategy used to identify key target genes/proteins in RUNX1-ETO expressing cells, compared to control, for further study. Human CD34<sup>+</sup> HSPC were isolated from neonatal cord blood and transduced with retrovirus co-expressing either RUNX1-ETO and GFP or GFP alone (control). Transduced cells were assessed for transduction efficiency and RNA/proteins extracted on day 3 of cell culture. Samples were analysed by performing Affymetrix Microarray (n=4) or SWATH-MS (n=3) analysis to determine mRNA and protein expression, respectively. Differentially expressed genes/proteins were uploaded into pathway analysis software for pathway analysis and target identification (Part 2), with a specific focus on TF. Additionally, aberrantly localised proteins were analysed in the cytoplasmic and nuclear compartments. Following identification of a target(s) of interest, these was used to perform overexpression and/or KD functional studies in CD34<sup>+</sup> HSPC and AML cell lines (Part 3; Chapter 4, Chapter 5).

### 3.3 Results

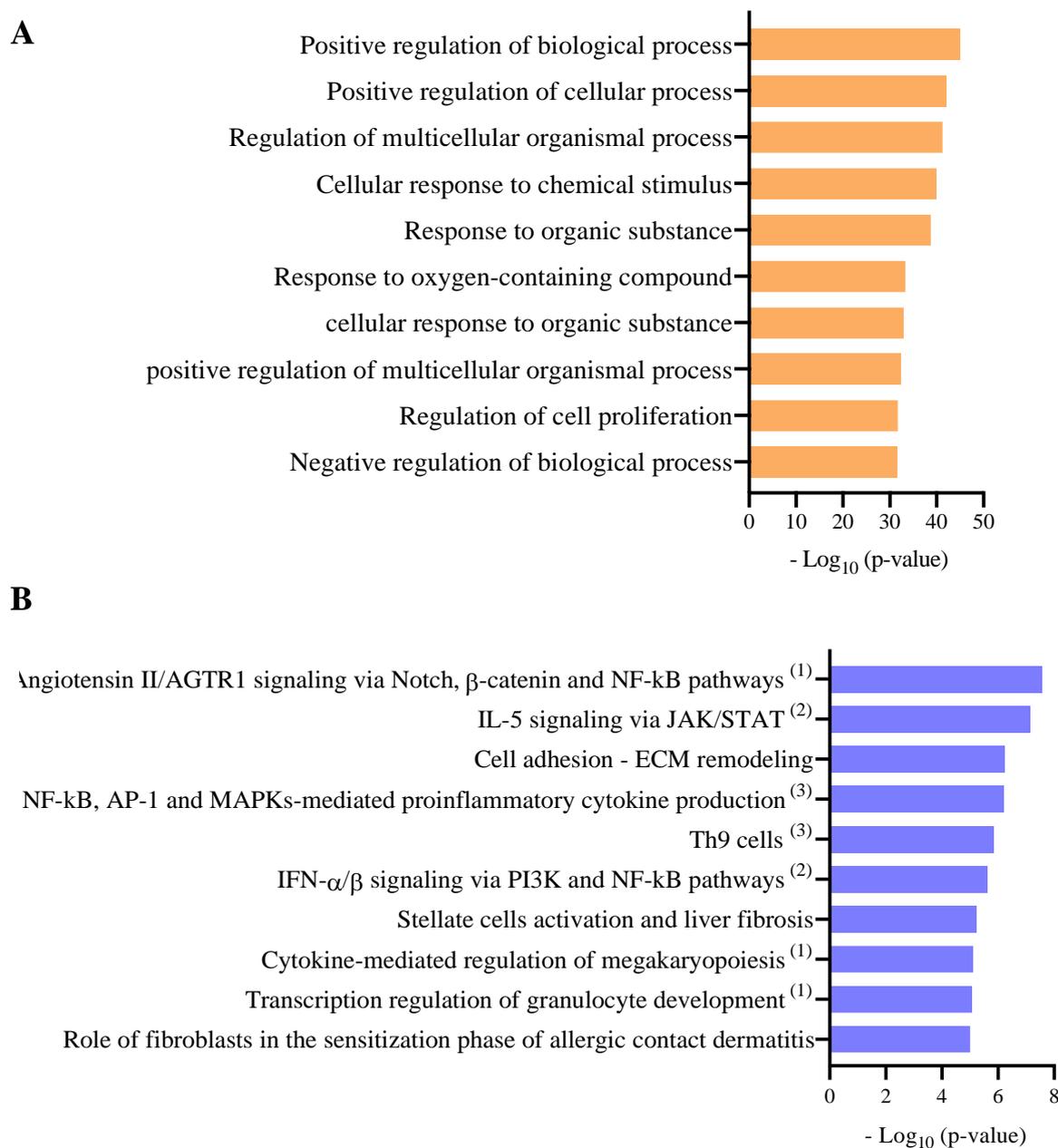
#### 3.3.1 Transcriptional dysregulation mediated by RUNX1-ETO

##### 3.3.1.1 *RUNX1-ETO* expression in HSPC alters TF gene expression and is associated with GO processes that regulate transcription

Previously, the effects of RUNX1-ETO expression on the cell's transcriptome was analysed in cord blood-derived CD34<sup>+</sup> HSPC by contrasting control against a RUNX1-ETO overexpression vector (Tonks *et al.*, 2007). However, the analysis did not consider an assessment of changes to cellular pathways or gene expression networks but rather focused on highest fold changing transcripts. A more focused approach based on changes observed in TF coupled with 'Pathway' analysis will refine this analysis, leading to the identification of novel changes in transcriptional changes.

To understand the biological implications of RUNX1-ETO expression, functional 'Enrichment Analysis' was performed using Metacore™. Based upon a fold change value of  $\pm 1.5$ , significantly changed probe-sets (465) were matched to 380 unique genes and subsequent Gene Ontology (GO) analysis was performed. Several GO Processes were found to be significantly disrupted as a result of the expression of RUNX1-ETO, including 'positive and negative cell regulation processes' and 'cell proliferation' (**Figure 3.2A**), consistent with the hallmarks of AML (Niebuhr *et al.*, 2008). This approach is effective in identifying processes that are involved in RUNX1-ETO-induced gene dysregulation; however, it does not consider the directionality of expression changes, which is an inherent limitation of overrepresentation analysis, such as the one presented in this study. Nevertheless, this is a well-recognised and widely utilised technique for identifying the most broad processes which have been perturbed in this setting.

Alternative analysis examining Metacore™ curated Pathways Maps identified several disrupted cell developmental processes, of which 'transcription regulation of granulocytic development' was found to be significantly over-represented, supporting previous observations that RUNX1-ETO expression impairs haematopoietic differentiation (**Figure 3.2B**) (Niebuhr *et al.*, 2008; Tonks *et al.*, 2004). Additionally, 'immune response' related pathways were found to be disrupted as a consequence of RUNX1-ETO expression (Vangala *et al.*, 2003; Choi *et al.*, 2006).



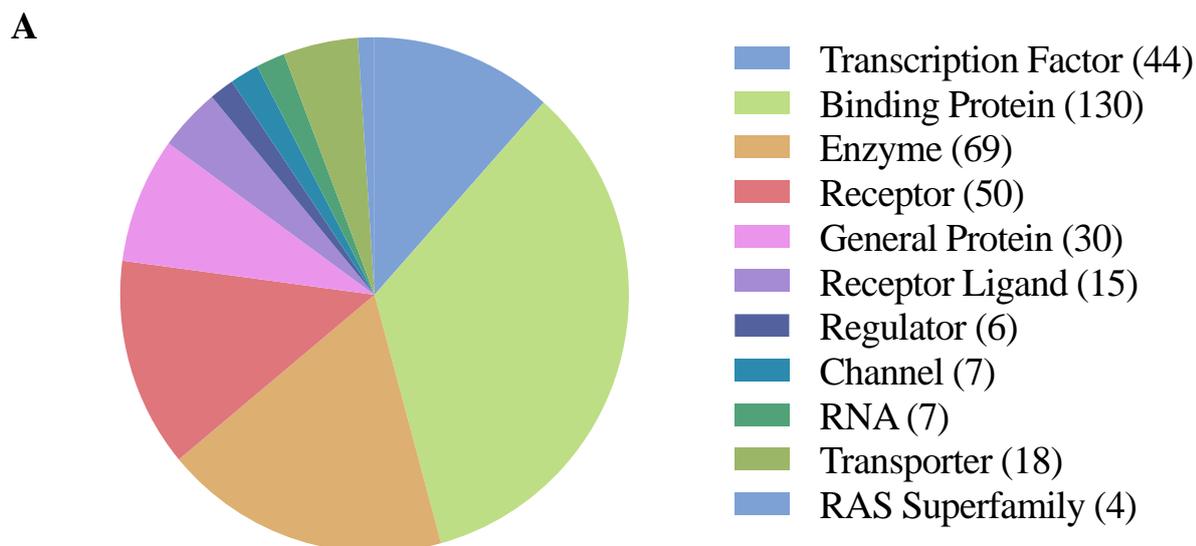
**Figure 3.2 – Functional Enrichment Analysis of mRNA changes observed in human CD34<sup>+</sup> HSPC expressing RUNX1-ETO**

Enrichment Analysis of the 380 significantly dysregulated genes (Tonks *et al.*, 2007) in CD34<sup>+</sup> HSPC expressing RUNX1-ETO, when compared to control, using Metacore™. Top 10 Significantly dysregulated (A) GO processes and (B) Pathway Maps, ranked by  $-\text{Log}(\text{p-Value})$ . Processes can further be clustered <sup>(1)</sup> Developmental process, <sup>(2)</sup> Immune response-related process and <sup>(3)</sup> Asthma-related process. A False Discovery Rate (FDR) of 0.05 was applied to both analyses.

Changes observed in the transcriptome of RUNX1-ETO expressing CD34<sup>+</sup> HSPC, when compared to control, correlate with the phenotypical features observed *in vitro* and in t(8;21) AML patients, in which cells display an arrest in normal development, as well as corroborating that multiple physiological processes are disrupted as a consequence of the expression of this fusion protein.

To determine the proportion of functional classes of genes dysregulated as a consequence of RUNX1-ETO expression, a classification for all 380 dysregulated genes was performed using Metacore™ (**Figure 3.3A**). In order to statistically support these observations, an Enrichment Analysis by Protein Function using Metacore™ was performed. TF were identified as the most over-represented group of genes (determined by Z-score) whose expression was significantly disrupted by RUNX1-ETO expression (48 network objects, corresponding to 44 TFs), followed by receptor ligands and general enzymes (**Figure 3.3B**). Changes in TF expression supports the pathway analysis described previously in which transcriptional and development processes are significantly associated with RUNX1-ETO expression. **Table 3.1** shows that 22 TFs were found to be upregulated in RUNX1-ETO cells compared to control, whilst 22 were found to be downregulated under the same conditions.

Since developmental disruption is likely mediated through changes in TF regulation, we attempted to use the 380 dysregulated gene list to determine the TF which regulated the highest number of genes within this list. As different pathway analysis programs can rely on different databases, two parallel approaches were undertaken using Metacore™ and IPA® to maximise the resulting targets. Initially, Metacore™, in combination with its built-in ‘Interactome Transcription Factors’ algorithm was used to identify the upstream TF(s) responsible for the changing mRNA expression of downstream genes. This algorithm achieves this by determining the density and statistical significance of the interactions of each TF with other genes from the dataset. This approach resulted in the identification of 25 TFs (**Table 3.1**). The most significantly connected TF included *PUL1*, which is a master regulator of myeloid cells (Nerlov and Graf, 1998; Koschmieder *et al.*, 2005); *ERG*, a member of the erythroblast transformation-specific family, responsible for regulating haematopoiesis and the differentiation and maturation of megakaryocytic cells (Knudsen *et al.*, 2015); and *CEBPA*, a coordinator of proliferation arrest and differentiation of myeloid progenitors (Koschmieder *et al.*, 2005).



**B**

Protein class	Actual	N	R	N	p-value	z-score	% in dataset
Other	219	402	32582	40433	$7.663 \times 10^{-33}$	-13.3	54.48%
TF	48	402	1249	40433	$1.632 \times 10^{-15}$	10.31	11.94%
Receptor	50	402	1701	40433	$9.642 \times 10^{-12}$	8.262	12.44%
Receptor Ligand	17	402	555	40433	$4.995 \times 10^{-5}$	4.946	4.23%
Enzyme	44	402	2868	40433	$2.91 \times 10^{-3}$	3.024	10.95%
Kinase	12	402	685	40433	$4.298 \times 10^{-2}$	2.016	2.99%
Phosphatase	5	402	241	40433	$9.378 \times 10^{-2}$	1.696	1.24%
Protease	9	402	623	40433	$1.709 \times 10^{-1}$	1.142	2.24%

**Figure 3.3 – TF were identified as the most significantly represented functional class due to RUNX1-ETO expression in CD34<sup>+</sup> HSPC**

(A) Pie chart indicating the classification of the differentially expressed genes using Metacore™. 380 genes were found to be dysregulated in RUNX1-ETO expressing cells when compared to control. Numbers after each class represent the number of genes identified. (B) Table indicating the classification of ‘network objects’ dysregulated upon the expression of RUNX1-ETO. Metacore™ identified TF as the most significantly changed class of genes, followed by receptors. However, it failed to identify over 50% of all ‘network objects’ present in the dataset. (**Actual** - number of network objects from the dataset(s) for a given protein class; **n** - number of network objects in the activated dataset; **R** - number of network objects of a given protein class in the complete database or background list; **N** - total number of network objects in the complete database or background list.). As z-scores can be negative, the higher z-scores correspond to shorter ‘distance’ and most meaningful results.

**Table 3.1 – Identification of RUNX1-ETO induced TF dysregulation in CD34<sup>+</sup> HSPC**

TF expression had been previously determined by performing microarray analysis on CD34<sup>+</sup> HSPC transfected either with a RUNX1-ETO retroviral vector, or a control plasmid, on day 3 of cell culture. To identify significantly overconnected genes within the dataset, the Metacore™ ‘Interactome Transcription Factors’ and the IPA® ‘Upstream Regulators’ algorithms were used. ✓ and ✗ denotes if the gene was identified or not (respectively) by the algorithm used. The results from both approaches were used to generate a list of 35 genes for subsequent analysis. p-values were calculated by both software as described in 2.10.

RNASeq data from the TCGA 2013 dataset (Ley *et al.*, 2013) (GSE13159) was used to determine mRNA expression in AML M2 patients, divided into t(8;21), expressing the fusion protein RUNX1-ETO, and non-t(8;21) patients. mRNA expression in normal human HSC were obtained from (Rapin *et al.*, 2014) (GSE42519). Each gene has a corresponding Affymetrix ID unique to the Hu133A GeneChip® (Affymetrix®).

Fold-change (FC) values represent the regulation of each gene when compared to control cells (green – upregulated; red – repressed, when compared to control cells). (↑) denotes upregulated or (↓) downregulated gene expression according to each analysis.

\* Genes are represented by multiple Affymetrix IDs. For genes with multiple Affymetrix IDs, values of fold-change were averaged. (n/a denotes not applicable; CDS denotes coding sequence)

Gene	FC	Affymetrix ID	Metacore™		IPA®		Literature review	AML M2 t(8;21) vs non-t(8;21)	AML t(8;21) vs. HSC	CDS (kbp)
			Interactome TF	p-value	Upstream Regulators	p-value				
<i>EGR2</i>	3.69	205249_at	✓	3.65 x 10 <sup>-13</sup>	✓	2.61 x 10 <sup>-4</sup>	-	↓	n/a	n/a
<i>HOXB6</i>	3.05	205366_s_at	✗	-	✗	-	n/a	n/a	n/a	n/a
<i>IDI1</i>	2.99	208937_s_at	✗	-	✓	2.89 x 10 <sup>-4</sup>	(Wang <i>et al.</i> , 2015a)	n/a	n/a	n/a
<i>CREM*</i>	2.75	209967_s_at 214508_x_at	✓	2.39 x 10 <sup>-4</sup>	✓	5.53 x 10 <sup>-5</sup>	-	↓	n/a	n/a
<i>ARID5B</i>	2.59	212614_at	✗	-	✓	4.11 x 10 <sup>-2</sup>	-	↑	↑	~2.9
<i>KLF6</i>	2.58	208961_s_at	✓	4.38 x 10 <sup>-4</sup>	✗	-	(DeKolver <i>et al.</i> , 2013a)	n/a	n/a	n/a

Gene	FC	Affymetrix ID	Metacore™		IPA®		Literature review	AML M2 t(8;21) vs non-t(8;21)	AML t(8;21) vs. HSC	CDS (kbp)
			Interactome TF	p-value	Upstream Regulators	p-value				
<b>ZBTB20*</b>	2.51	222357_at 213158_at	X	-	✓	1.01 x 10 <sup>-2</sup>	-	↓	n/a	n/a
<b>AHR</b>	2.42	205749_at	✓	8.84 x 10 <sup>-10</sup>	X	-	(Ly <i>et al.</i> , 2019)	n/a	n/a	n/a
<b>NFKB2</b>	2.12	207535_s_at	✓	2.56 x 10 <sup>-7</sup>	✓	3.77 x 10 <sup>-6</sup>	(Nakagawa <i>et al.</i> , 2011; Zhou <i>et al.</i> ,	n/a	n/a	n/a
<b>RUNX1</b>	1.91	1440878_at	✓	6.9 x 10 <sup>-34</sup>	✓	1.57 x 10 <sup>-13</sup>	n/a	n/a	n/a	n/a
<b>HLF</b>	1.88	204753_s_at	X	-	X	-	n/a	n/a	n/a	n/a
<b>IRF9</b>	1.78	203882_at	X	-	✓	4.65 x 10 <sup>-3</sup>	-	↑	↑	~ 1.4
<b>MXI1</b>	1.73	202364_at	X	-	✓	2.73 x 10 <sup>-4</sup>	(Weng <i>et al.</i> , 2017)	n/a	n/a	n/a
<b>SOX4</b>	1.65	201416_at	✓	1.62 x 10 <sup>-4</sup>	✓	2.09 x 10 <sup>-4</sup>	(Nafria <i>et al.</i> , 2020; Fernando <i>et al.</i> , 2017)	n/a	n/a	n/a
<b>IRF7</b>	1.64	208436_s_at	✓	3.06 x 10 <sup>-4</sup>	✓	9.07 x 10 <sup>-4</sup>	-	No change	↑	~ 1.5
<b>E2F5</b>	1.62	221586_s_at	✓	1.38 x 10 <sup>-5</sup>	✓	1.40 x 10 <sup>-4</sup>	(Huang <i>et al.</i> , 2020)	n/a	n/a	n/a
<b>TCF7L2</b>	1.62	216035_x_at	✓	6.26 x 10 <sup>-20</sup>	✓	2.46 x 10 <sup>-3</sup>	(Saenz <i>et al.</i> , 2020; Zhao <i>et al.</i> , 2019)	n/a	n/a	n/a
<b>ERG</b>	1.61	213541_s_at	✓	2.31 x 10 <sup>-18</sup>	✓	1.28 x 10 <sup>-7</sup>	(Martens <i>et al.</i> , 2012; Mandoli <i>et al.</i> , 2016)	n/a	n/a	n/a
<b>SMARCA1</b>	1.55	203874_s_at	X	-	X	-	n/a	n/a	n/a	n/a
<b>HBPI</b>	1.54	209102_s_at	✓	1.10 x 10 <sup>-2</sup>	X	-	-	↑	↓	n/a
<b>HOXA5</b>	1.54	213844_at	X	-	X	-	n/a	n/a	n/a	n/a
<b>ZNF217</b>	1.5	203739_at	X	-	✓	3.79 x 10 <sup>-3</sup>	-	No change	↑	~ 3.2
<b>EGR3</b>	-1.52	206115_at	✓	3.69 x 10 <sup>-3</sup>	X	-	(Cheng <i>et al.</i> , 2015)	n/a	n/a	n/a
<b>IRF8</b>	-1.52	204057_at	✓	4.88 x 10 <sup>-8</sup>	✓	2.57 x 10 <sup>-9</sup>	(Pogosova-Agadjanyan <i>et al.</i> , 2013; Liss <i>et al.</i> ,	n/a	n/a	n/a

Gene	FC	Affymetrix ID	Metacore™		IPA®		Literature review	AML M2 t(8;21) vs non-t(8;21)	AML t(8;21) vs. HSC	CDS (kbp)
			Interactome TF	p-value	Upstream Regulators	p-value				
<i>JUND</i>	-1.53	203752_s_at	✓	8.5 x 10 <sup>-13</sup>	✓	3.77 x 10 <sup>-6</sup>	(Ptasinska <i>et al.</i> , 2019; Martinez-Soria <i>et al.</i> ,	n/a	n/a	n/a
<i>E2F8</i>	-1.55	219990_at	✗	-	✗	-	n/a	n/a	n/a	n/a
<i>TSC22D1</i>	-1.56	215111_s_at	✗	-	✗	-	n/a	n/a	n/a	n/a
<i>CBFA2T3</i>	-1.57	208056_s_at	✓	1.39 x 10 <sup>-2</sup>	✓	2.88 x 10 <sup>-2</sup>	(Steinauer <i>et al.</i> , 2019; Jakobczyk <i>et al.</i> , 2021)	n/a	n/a	n/a
<i>RARA</i>	-1.57	203749_s_at	✓	1.58 x 10 <sup>-5</sup>	✗	-	(De Braekeleer <i>et al.</i> , 2014)	n/a	n/a	n/a
<i>ID2</i>	-1.58	213931_at	✗	-	✓	3.56 x 10 <sup>-13</sup>	(Zhou <i>et al.</i> , 2017; Deb and Somerville, 2016)	n/a	n/a	n/a
<i>MYCN</i>	-1.59	209757_s_at	✗	-	✓	1.47 x 10 <sup>-5</sup>	(Kawagoe <i>et al.</i> , 2007; Liu <i>et al.</i> , 2017)	n/a	n/a	n/a
<i>ZBTB16</i>	-1.61	205883_at	✓	6 x 10 <sup>-7</sup>	✓	3.15 x 10 <sup>-6</sup>	(Ptasinska <i>et al.</i> , 2012; Melnick <i>et al.</i> , 2000a)	n/a	n/a	n/a
<i>C/EBPε</i>	-1.64	214523_at	✓	1.06 x 10 <sup>-6</sup>	✓	3.61 x 10 <sup>-5</sup>	(Shimizu <i>et al.</i> , 2000; Li <i>et al.</i> , 2019)	n/a	n/a	n/a
<i>ZFP36L2</i>	-1.65	201367_s_at	✗	-	✗	-	n/a	n/a	n/a	n/a
<i>CIAO1</i>	-1.66	217501_at	✗	-	✗	-	n/a	n/a	n/a	n/a
<i>RUNX3*</i>	-1.69	204197_s_at 204198_s_at	✓	9 x 10 <sup>-10</sup>	✓	1.06 x 10 <sup>-2</sup>	*	n/a	n/a	n/a
<i>GATA2</i>	-1.76	209710_at	✓	2.25 x 10 <sup>-30</sup>	✓	1.06 x 10 <sup>-14</sup>	(Sood <i>et al.</i> , 2016; Loke <i>et al.</i> , 2017)	n/a	n/a	n/a
<i>PU.1</i>	-1.76	209454_s_at	✓	1.18 x 10 <sup>-35</sup>	✓	5.58 x 10 <sup>-11</sup>	(Vangala <i>et al.</i> , 2003; Eder <i>et al.</i> , 2016)	n/a	n/a	n/a
<i>GATA1</i>	-1.81	210446_at	✗	-	✓	8.62 x 10 <sup>-20</sup>	(Elagib <i>et al.</i> , 2003; Choi <i>et al.</i> , 2006)	n/a	n/a	n/a
<i>GFI1B</i>	-1.94	208501_at	✓	5.46 x 10 <sup>-5</sup>	✓	5.74 x 10 <sup>-3</sup>	(Möröy and Khandanpour, 2019;	n/a	n/a	n/a
<i>TEAD3</i>	-2.09	209454_s_at	✗	-	✗	-	n/a	n/a	n/a	n/a
<i>C/EBPα</i>	-2.34	204039_at	✓	1.57 x 10 <sup>-22</sup>	✓	2.08 x 10 <sup>-9</sup>	(Pabst <i>et al.</i> , 2001; Grossmann <i>et al.</i> ,	n/a	n/a	n/a

Gene	FC	Affymetrix ID	Metacore™		IPA®		Literature review	AML M2 t(8;21) vs non-t(8;21)	AML t(8;21) vs. HSC	CDS
			Interactome TF	p-value	Upstream Regulators	p-value				(kbp)
<i>KLF1</i>	-2.59	210504_at	✓	1.05 x 10 <sup>-5</sup>	✓	8.47 x 10 <sup>-3</sup>	(Kuvardina <i>et al.</i> , 2015; Mansoor <i>et al.</i> , 2012; Wang <i>et al.</i> , 2010a)	n/a	n/a	n/a
<i>NF-E2</i>	-2.64	209930_s_at	✓	1.68 x 10 <sup>-7</sup>	✓	3.43 x 10 <sup>-5</sup>		n/a	n/a	n/a

Other overconnected TF included *EGR2* and *JUND*, known for protecting cells from p53-dependent senescence and apoptosis, reflecting the diversity of genes and processes influenced by the RUNX1-ETO fusion protein (Morita *et al.*, 2016; Hernandez *et al.*, 2008).

The second approach used the IPA<sup>®</sup> algorithm ‘Upstream Regulators’. Similarly, this method aims to identify the cascade of upstream transcriptional regulators based on how many known downstream targets of each transcription regulator are represented in the dataset. This approach identified 35 upstream transcriptional regulators. However, transcriptional regulators refers to both TFs and regulators of transcription; after filtering non-relevant results, 31 TF were identified for further investigation, as shown in **Table 3.1**. The results obtained from both proprietary software approaches were combined, resulting in the generation of a list containing 35 overconnected TF, of which 22 were identified by both approaches, suggesting concordance between the underlying databases. Subsequent steps of analyses were based on the combined 35-gene list.

### 3.3.1.2 Identification of aberrantly expressed TFs driving transcriptional change in AML

To identify novel genes that could be involved in the arrest in cell differentiation observed upon the expression on the RUNX1-ETO fusion protein, subsequent analysis filtered the 35-TF based on key literature search criteria, including novelty, significance in terms of its role in AML t(8;21) and model system used. Several genes identified had already been extensively studied in AML t(8;21), including *PU.1* (*SPI1*), *CEBPA*, *NFKB2* and *NFE2* (Vangala *et al.*, 2003; Pabst *et al.*, 2001; Zhou *et al.*, 2015; Wang *et al.*, 2010a; Jutzi *et al.*, 2019). Genes that lacked the novelty factor were excluded from further analysis.

Subsequently, to determine whether the pattern of expression of the TF identified above correlated with mRNA expression observed in patients, this study examined the transcriptome of AML blast samples from AML M2 patients harbouring the (8;21) translocation, and compared them to non-t(8;21) patients. In order to do this, the publicly available TCGA dataset generated by RNA-sequencing was analysed for the mRNA expression pattern of each gene (Ley *et al.*, 2013). Moreover, an additional analysis examined gene expression patterns in undifferentiated HSC and compared them to that of patients diagnosed with t(8;21) AML (Ley *et al.*, 2013; Rapin *et al.*, 2014). In order to refine the gene list, the dataset was filtered by only examining genes for which expression was concordant between analysed studies (genes with

no change in one of the studies, but same trend in the other as the CD34<sup>+</sup> cord blood-derived model were still considered) (**Table 3.1**).

Furthermore, a feasibility analysis was performed to ensure downstream analysis was possible; for instance, genes with a coding sequence bigger than 3.5 kb were excluded due to an inability to overexpress it in our model system (**Table 3.1**). Based on all these factors, four gene targets were chosen for further study: *ARID5B*, *IRF7*, *IRF9* and *ZNF217*. Given that the expression or function of these genes has not yet been described in AML, this study initially assessed their protein expression in RUNX1-ETO expressing CD34<sup>+</sup> HSPC by western blot.

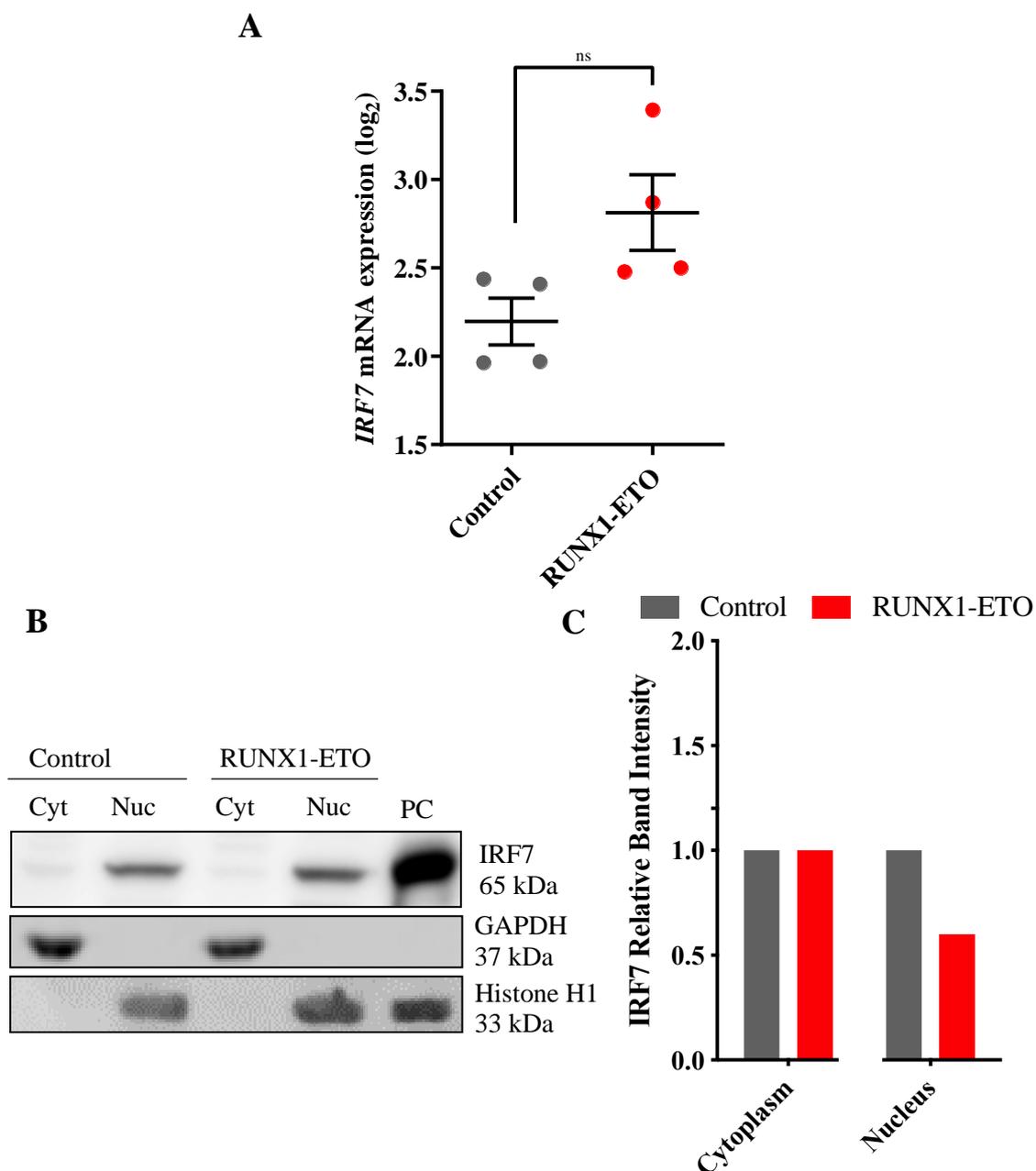
*IRF7* is a master regulator of Type I IFN-dependant immune response, and its dysregulated expression has been described in prostate and breast cancer (**3.4**) (Yang *et al.*, 2016; Tara *et al.*, 2018). Increased expression of *IRF7*-regulated genes has also been associated with prolonged metastatic-free survival (Savitsky *et al.*, 2010; Takaoka *et al.*, 2008). **Figure 3.4** shows *IRF7* mRNA was 1.6-fold upregulated in CD34<sup>+</sup> HSPC expressing RUNX1-ETO, when compared to control cells; however, no difference was observed at the protein level. Another member of the IRF family, *IRF9* has been described as having a role in Type-I IFN-mediated cellular response and has been shown to stimulate the p53 pathway, preventing oncogene-induced malignant cell transformation and in inducing DNA damage-induced apoptosis (Savitsky *et al.*, 2010; Takaoka *et al.*, 2008). Even though other *IRF* genes have been linked to myeloid cell differentiation (Tamura *et al.*, 2015b; Wang and Morse, 2009), this has not been described in AML. As shown in **Figure 3.5**, *IRF9* mRNA was 2.7-fold upregulated in CD34<sup>+</sup>-RUNX1-ETO expressing cells compared to control, whilst a 3-fold increase in *IRF9* protein expression was observed in the cytoplasm of CD34<sup>+</sup> HSPC expressing the fusion protein RUNX1-ETO. A 1.5-fold increase in the expression of this protein in the nucleus of RUNX1-ETO expressing cells was also observed.

*ARID5B* has been identified for its role in embryonic development, cell type-specific gene expression and cell growth regulation (Patsialou *et al.*, 2005). Its aberrant expression or mutations have been widely described in ALL and APL, however its role in AML remains unidentified. mRNA analysis of CD34<sup>+</sup> HSPC expressing RUNX1-ETO showed a 4-fold upregulation *ARID5B* in these cells when compared to control (**Figure 3.6**). *ARID5B* protein was found to be detected in the nucleus of both RUNX1-ETO and control CD34<sup>+</sup> HSPC; however, expression level was the same. Moreover, no *ARID5B* protein was detected in the cytoplasm of either control or RUNX1-ETO expressing CD34<sup>+</sup> cells.

Lastly, ZNF217 has been widely described as having a crucial role in the development of multiple types of cancer, including breast, ovarian and prostate, and its increased expression has been linked to a poorer outcome in this disease (see **0**) (Rahman *et al.*, 2012; Rooney *et al.*, 2004). Furthermore, ZNF217 has been implicated in the maintenance of an undifferentiated phenotype in glioma and breast primary cultures (Cohen *et al.*, 2015). ZNF217 mRNA and protein is upregulated by 1.7 and >10 fold respectively in cells expressing the RUNX1-ETO fusion protein (**Figure 3.7**).

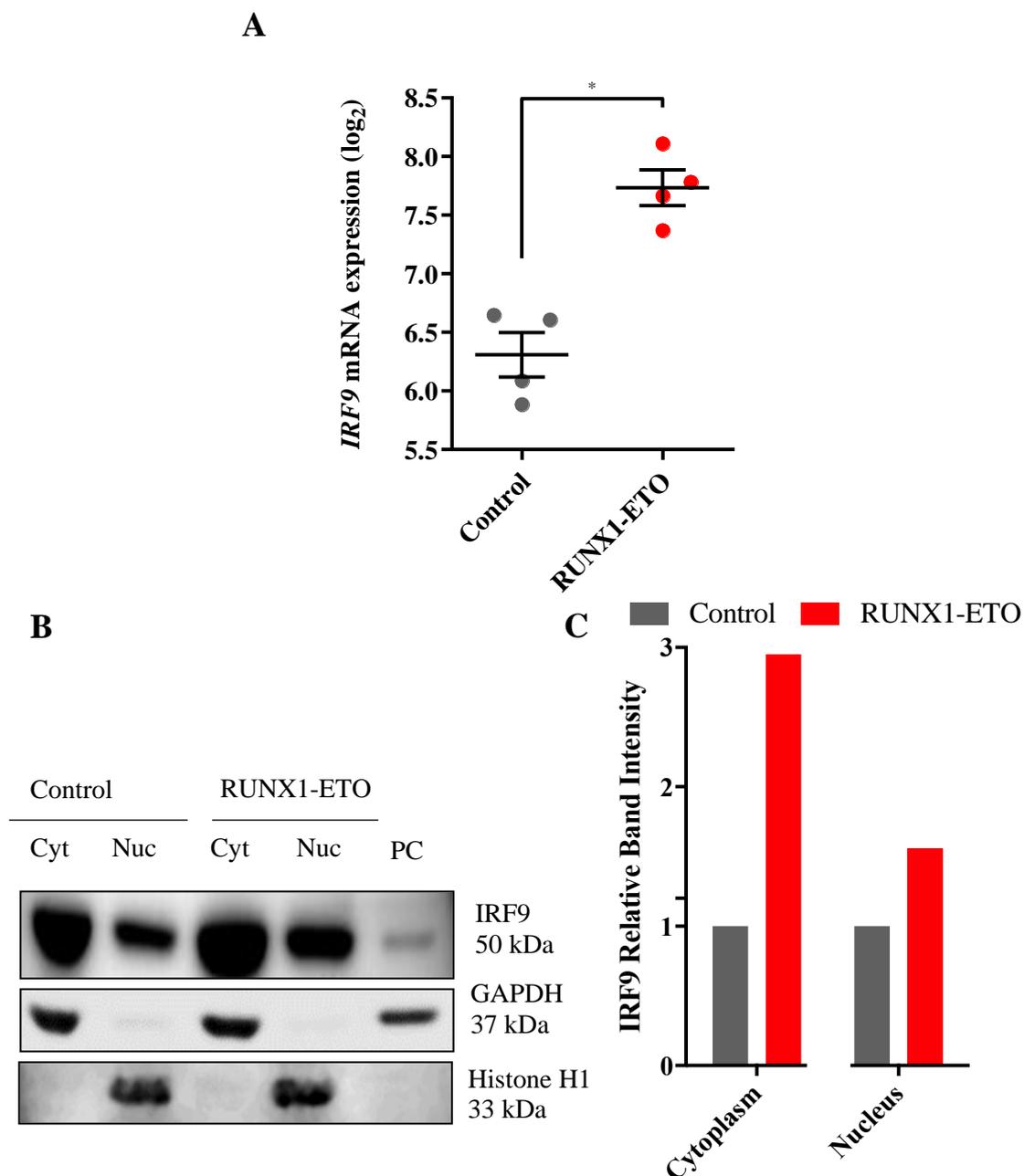
Additionally, SWATH-MS analysis was performed to determine the proteomic profile of RUNX1-ETO CD34<sup>+</sup> HSPC, as compared to control (see **3.3.2** for a complete description of this analysis). Regarding these 4 targets, SWATH-MS was able to detect 3 of the potentially relevant proteins, but failed to identify IRF7, detectable through western blotting technique. Nevertheless, this analysis showed that the IRF9 protein was found to be exclusively expressed in the nuclear compartment of RUNX1-ETO cells, with no detectable expression in control CD34<sup>+</sup> HSPC (**Figure 3.8A**). SWATH-MS was able to detect the ARID5B protein, previously found in the nucleus of both control and RUNX1-ETO CD34<sup>+</sup> HSPC; however, this was identified as being expressed in the cytoplasm of these cells (**Figure 3.8B**). Lastly, the ZNF217 protein was detected as being overexpressed in the nucleus of RUNX1-ETO cells, as compared to control (**Figure 3.8C**). Even though both SWATH-MS and western blotting analysis were performed on identical samples, both techniques measured distinct protein expression levels, and were not concordant between each other. The discrepancies between SWATH-MS and western blotting results will be further discussed in section **0**.

In summary, this study was able to identify four potentially relevant targets in the development of AML t(8;21) (**Table 3.2**). Even though several targets were found to be upregulated at the mRNA level, these changes did not equate to the corresponding protein expression. For this reason, additional methods for analysing the proteomic profile of these cells are necessary.



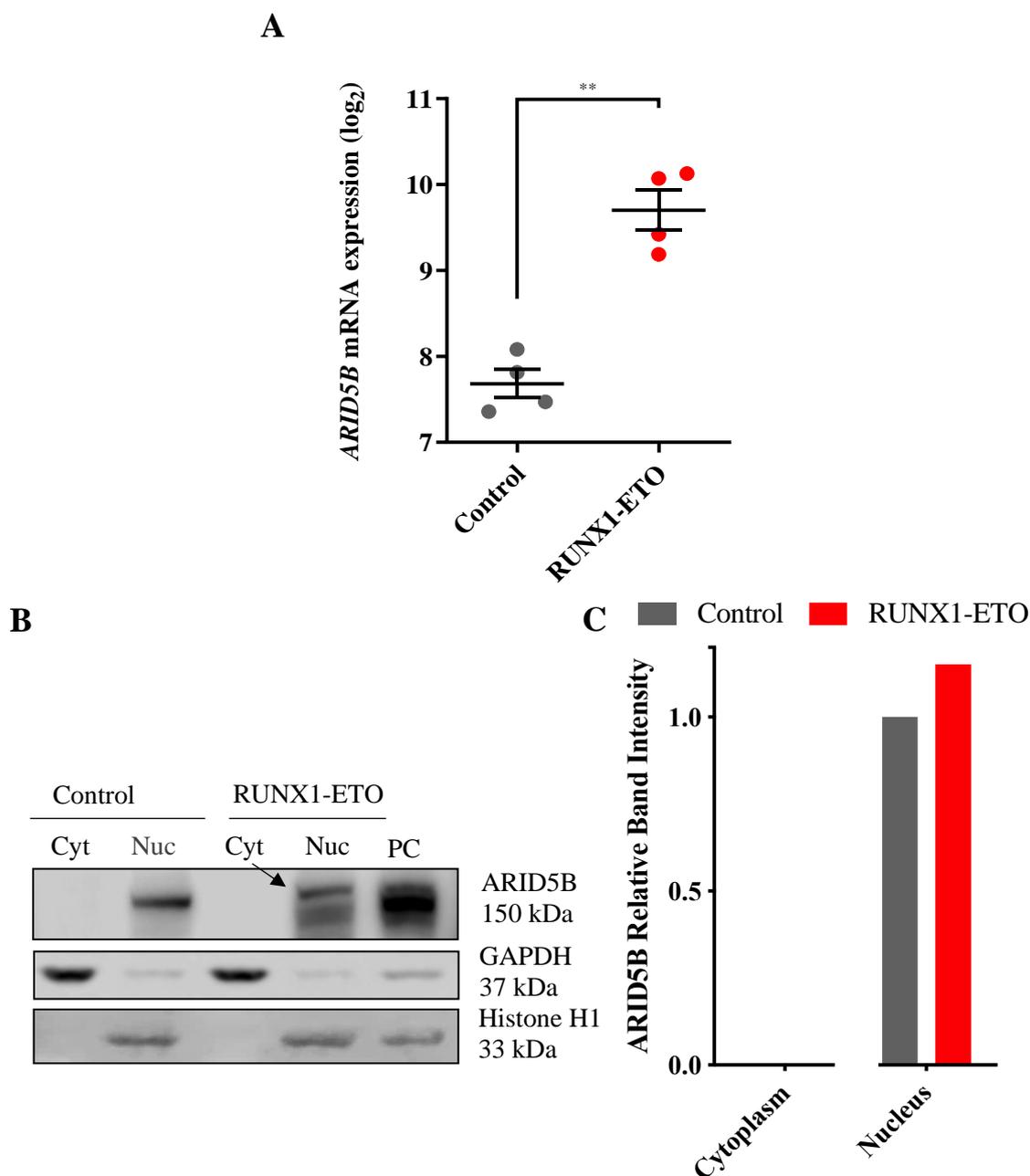
**Figure 3.4 – RUNX1-ETO does not influence the expression of IRF7 in CD34<sup>+</sup> HSPC**

(A) *IRF7* mRNA expression following retroviral transfection (day 3 of cell culture), in CD34<sup>+</sup> HSPC transduced with a control or RUNX1-ETO vector (Tonks *et al.*, 2007) (n=4) (Paired T-test; ns - not significant). (B) Western blot analysis of IRF7 expression in CD34<sup>+</sup> HSPC transduced with either a control or a RUNX1-ETO vector, on day 3 of cell culture (n=1). The OCI-AML2 cell line was used as a positive control (PC). GAPDH was used as a loading control for cytosolic fractions; Histone H1 was used as a loading control for nuclear fractions. Both proteins were used as a verification of subcellular fraction. (C) Western blot quantification by densitometry analysis using ImageJ. Relative protein expression was determined by normalising each cytosolic and nuclear sample to GAPDH or Histone H1, respectively, and to each corresponding ‘control’ (n=1).



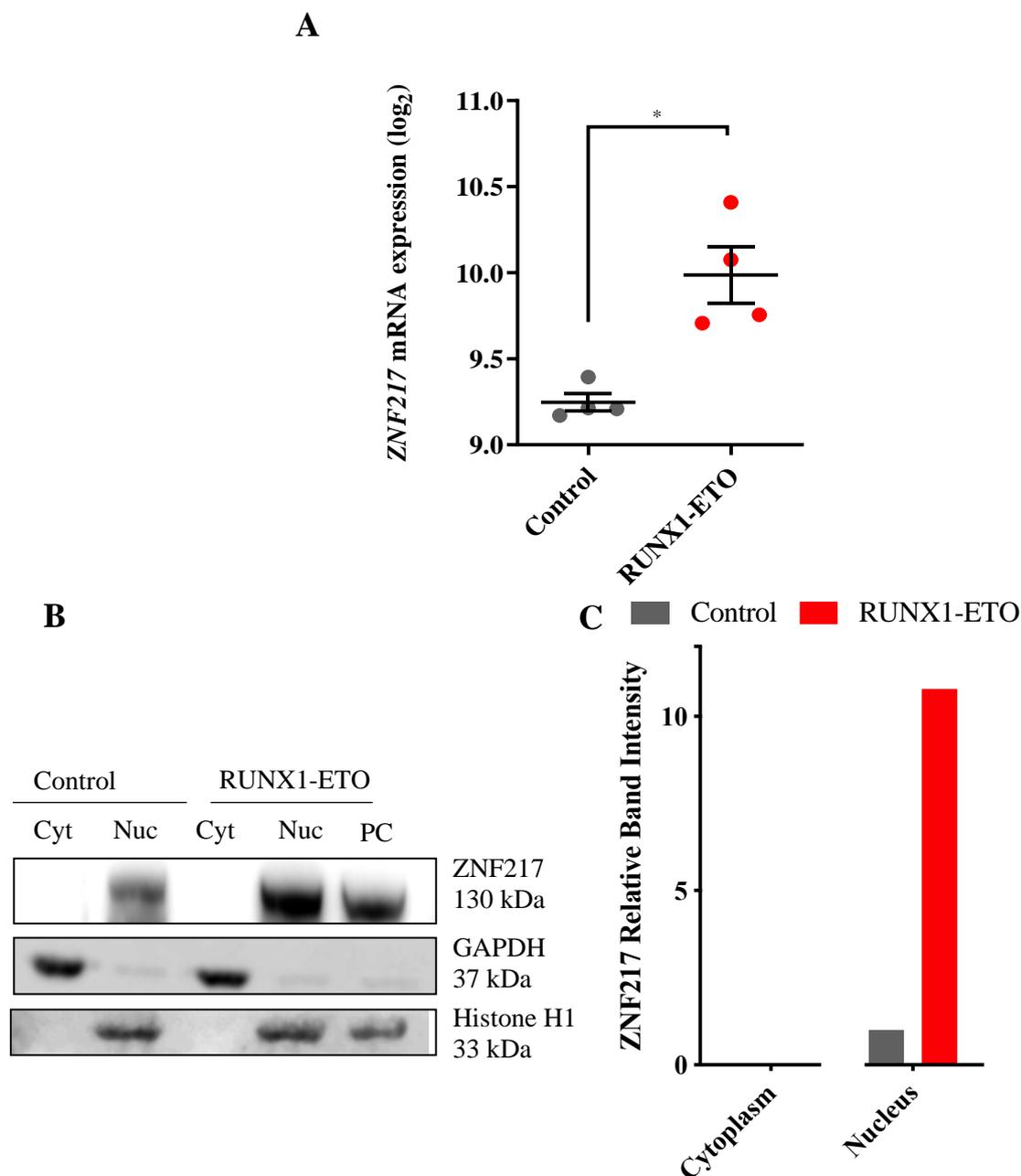
**Figure 3.5 – RUNX1-ETO leads to the upregulation of IRF9 in CD34<sup>+</sup> HSPC**

(A) *IRF9* mRNA expression following retroviral transfection (day 3 of cell culture), in CD34<sup>+</sup> HSPC transduced with a control or RUNX1-ETO vector (Tonks *et al.*, 2007) (n=4). Significant differences between RUNX1-ETO and control cells was analysed using paired t-test; \* denotes p<0.05. (B) Western blot analysis of IRF9 expression in CD34<sup>+</sup> HSPC transduced with either a control or a RUNX1-ETO vector, on day 3 of cell culture (n=1). The OCI-AML2 cell line was used as a positive control (PC). GAPDH was used as a loading control for cytosolic fractions; Histone H1 was used as a loading control for nuclear fractions. Both proteins were used as a verification of subcellular fraction. (C) Western blot quantification by densitometry analysis using ImageJ. Relative protein expression was determined by normalising each cytosolic and nuclear sample to GAPDH or Histone H1, respectively, and to each corresponding ‘control’ (n=1).



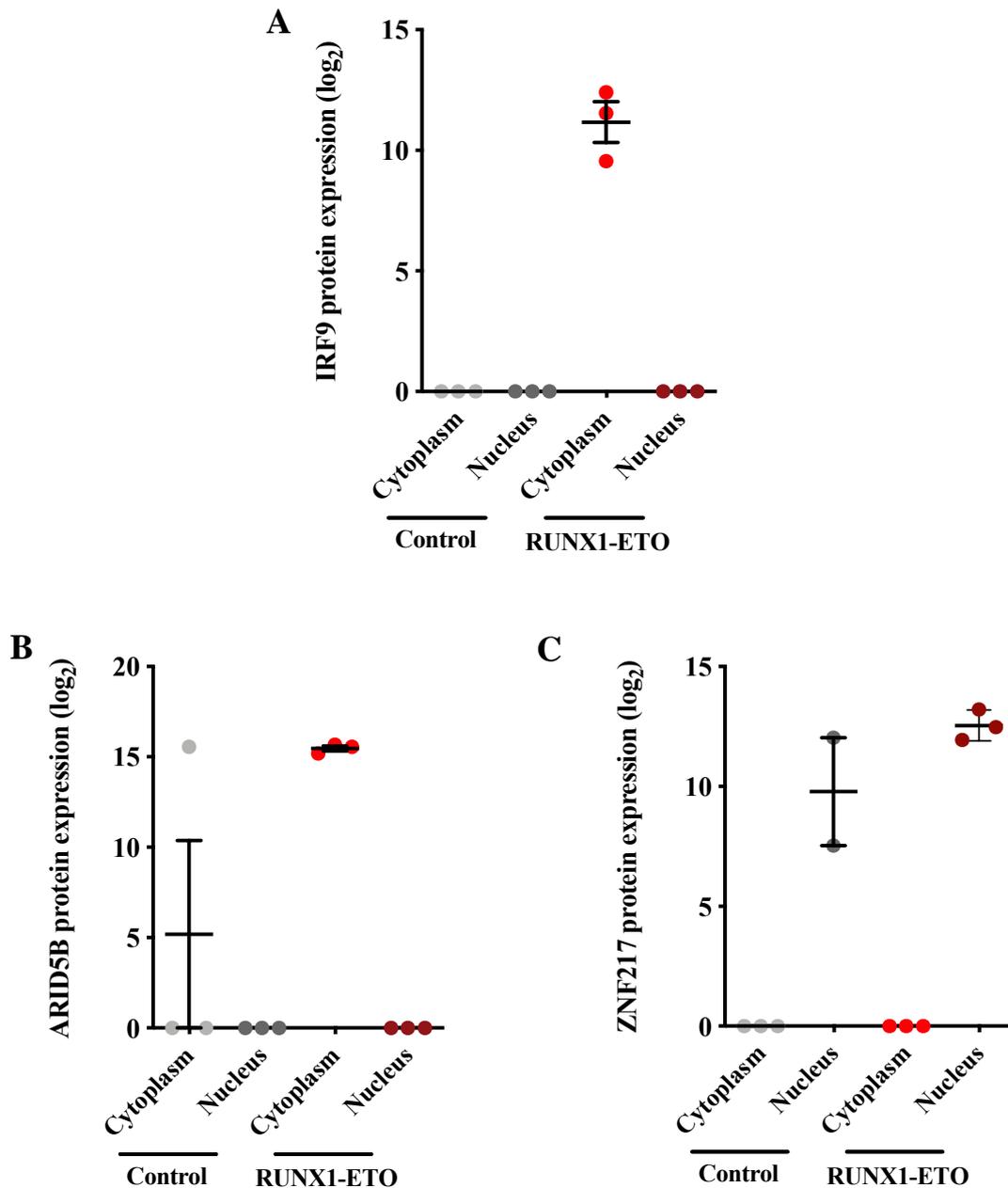
**Figure 3.6 – RUNX1-ETO leads to the upregulation of ARID5B in CD34<sup>+</sup> HSPC**

(A) *ARID5B* mRNA expression following retroviral transfection (day 3 of cell culture), in CD34<sup>+</sup> HSPC transduced with a control or RUNX1-ETO vector (Tonks *et al.*, 2007) (n=4). Significant differences between RUNX1-ETO and control cells was analysed using paired t-test; \*\* denotes p<0.01. (Appropriate band is indicated by the arrow). (B) Western blot analysis of ARID5B expression in CD34<sup>+</sup> HSPC transduced with either a control or a RUNX1-ETO vector, on day 3 of cell culture (n=1). The HeLa cell line was used as a positive control (PC). GAPDH was used as a loading control for cytosolic fractions; Histone H1 was used as a loading control for nuclear fractions. Both proteins were used as a verification of subcellular fraction. (C) Western blot quantification by densitometry analysis using ImageJ. Relative protein expression was determined by normalising each cytosolic and nuclear sample to GAPDH or Histone H1, respectively, and to each corresponding ‘control’ (n=1).



**Figure 3.7 – RUNX1-ETO leads to the upregulation of ZNF217 in CD34<sup>+</sup> HSPC**

(A) *ZNF217* mRNA expression following retroviral transfection (day 3 of cell culture), in CD34<sup>+</sup> HSPC transduced with a control or RUNX1-ETO vector (Tonks *et al.*, 2007) (n=4). Significant differences between RUNX1-ETO and control cells was analysed using paired t-test; \* denotes p<0.05. (B) Western blot analysis of ZNF217 expression in CD34<sup>+</sup> HSPC transduced with either a control or a RUNX1-ETO vector, on day 3 of cell culture (n=1). The K562 cell line was used as a positive control (PC). GAPDH was used as a loading control for cytosolic fractions; Histone H1 was used as a loading control for nuclear fractions. Both proteins were used as a verification of subcellular fraction. (C) Western blot quantification by densitometry analysis using ImageJ. Relative protein expression was determined by normalising each cytosolic and nuclear sample to GAPDH or Histone H1, respectively, and to each corresponding ‘control’ (n=1).



**Figure 3.8 – RUNX1-ETO leads to the upregulation of protein targets, in CD34<sup>+</sup> HSPC**

SWATH-MS analysis was performed on CD34<sup>+</sup> HSPC transduced with either a RUNX1-ETO overexpression retroviral vector, or a control plasmid. Cytosolic and nuclear proteins were extracted and quantified on day 3 of cell culture. Data indicates Mean  $\pm$  1SD (where applicable) (n=3). Null expression values refer to protein samples not being detected in the SWATH-MS run for each corresponding protein.

**Table 3.2 – Summary showing the effects of RUNX1-ETO expression in CD34<sup>+</sup> HSPC on gene and protein expression in prioritised targets**

§ mRNA expression was compared between RUNX1-ETO and control CD34<sup>+</sup> HSPC by microarray analysis, performed on day 3 of cell culture (adapted from (Tonks *et al.*, 2007))

‡ Protein expression was compared between RUNX1-ETO and control CD34<sup>+</sup> HSPC by performing western blot/SWATH-MS analysis on cytosolic and nuclear extracts, on day 3 of cell culture.

(↑) (↓) denotes upregulation or downregulation of gene/protein in RUNX1-ETO cells, respectively, as compared to control HSPC.

\* Fold-change was not calculated for these samples due to the fact that protein was not detected in the control samples. (n/a denotes not applicable; CDS denotes coding sequence)

Gene / Protein	CDS (kbp)	mRNA expression in CD34 <sup>+</sup> HSPC §		Protein expression in CD34 <sup>+</sup> HSPC ‡			
		Regulation	Loc.	Western Blot		SWATH-MS	
				Regulation	Loc.	Regulation	Loc.
<b>IRF7</b>	~ 1.5	↑ 1.6-fold	n/a	Unaltered	C+N	Not detected	
<b>IRF9</b>	~ 1.4	↑ 2.7-fold	n/a	↑ 3-fold	C	↑ *	C
<b>ARID5B</b>	~ 2.9	↑ 4-fold	n/a	↓ 1.5-fold	N	↑ *	C
<b>ZNF217</b>	~ 3.2	↑ 1.7-fold	n/a	Unaltered	N	↑ 3-fold	N

Loc. – localisation within the cell; C – cytoplasm, N - nucleus

### 3.3.2 Quantitative SWATH-MS analysis reveals proteomic changes arising from RUNX1-ETO expression, in CD34<sup>+</sup> HSPC

#### 3.3.2.1 Generation of human CD34<sup>+</sup> HSPC expressing RUNX1-ETO for SWATH-MS analysis

This study used a primary human CD34<sup>+</sup> HSPC model to investigate the effects of RUNX1-ETO on the dysregulation of protein expression. Changes in global protein expression were quantitated using SWATH-MS, a data-independent acquisition (DIA) technique that allows comprehensive quantification of thousands of detectable analytes in a sample (2.8.5). To generate the samples for MS analysis, three replicate pairs of CD34<sup>+</sup> HSPC from biologically distinct sources were infected with either a control or a RUNX1-ETO-overexpression vector (2.5.2), following which proteins were separated and enriched according to their subcellular localisation: cytoplasmic or nuclear. Following MS analysis, statistical analysis was used to identify significantly changed proteins as a consequence of RUNX1-ETO expression, compared to control. Metacore™ was used to perform pathway analysis and to identify targets of interest.

#### 3.3.2.2 Assessment of CD34<sup>+</sup> HSPC nuclear and cytosolic extracts shows sample purity and integrity

Proteomic analysis requires high quantities of extremely pure protein. As CD34<sup>+</sup> HSPC account for such a small proportion of total cord blood and with the requirement to extract proteins by day 3 of culture, thereby limiting the growth potential, the method had to be adapted to maximise recovery where possible. To isolate sufficient quantities of nuclear proteins from transduced cells, it was decided that if sufficient infection rate was achieved (>60%) (**Figure 3.9**), cells would not be sorted, thereby avoiding the concern surrounding stress response to the sorting process and further cell loss. This does, however, raise two important limitations. Firstly, the existence of a GFP expression range suggests varied incorporation rates within CD34<sup>+</sup> HSPC, which may suggest differing overexpression levels. In addition, the remaining 40% of cells are not transduced with the plasmid and do not experience overexpression. Despite this, the strict protein requirements for SWATH-MS necessitated such compromises. Following infection, cytosolic and nuclear proteins from RUNX1-ETO and control CD34<sup>+</sup> HSPC were extracted. Prior to SWATH-MS analysis, the overall protein quality and fractionation efficiency of each sample was assessed by performing QC assays to ensure no cross-contamination between cellular compartments. Firstly, protein extract quality was

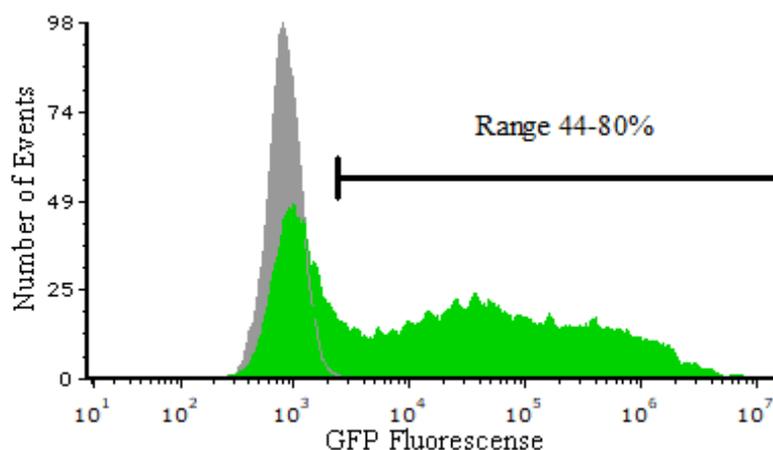
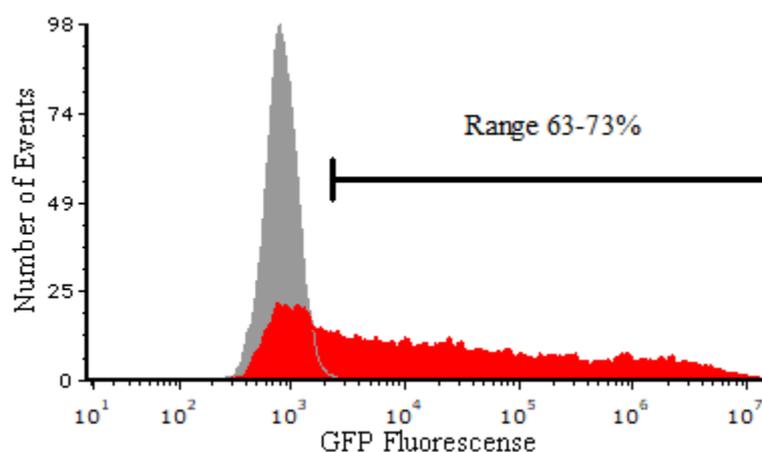
assessed using in gel SYPRO™ Ruby Protein Blot Stain (2.6.6). This analysis showed little or undetectable cytoplasmic contamination in the nuclear fraction, and undetectable nuclear contamination in the cytoplasmic samples (**Figure 3.10A - Figure 3.11**). In the nuclear samples, it was possible to identify the histone proteins (~ 30 kDa), exclusive to this cell compartment, whilst their presence in the cytoplasm was absent.

In addition, in gel protein staining showed no degradation of any sample, with a clear separation of proteins according to size. Moreover, the RUNX1-ETO fusion protein was solely detected in the nucleus of RUNX1-ETO-transduced CD34<sup>+</sup> HSPC whilst no protein was detected in the control samples (**Figure 3.10B**). Based on these assays, the samples were appropriate for SWATH-MS and downstream analysis.

### 3.3.2.3 SWATH-MS analysis detected over 4000 proteins in CD34<sup>+</sup> HSPC

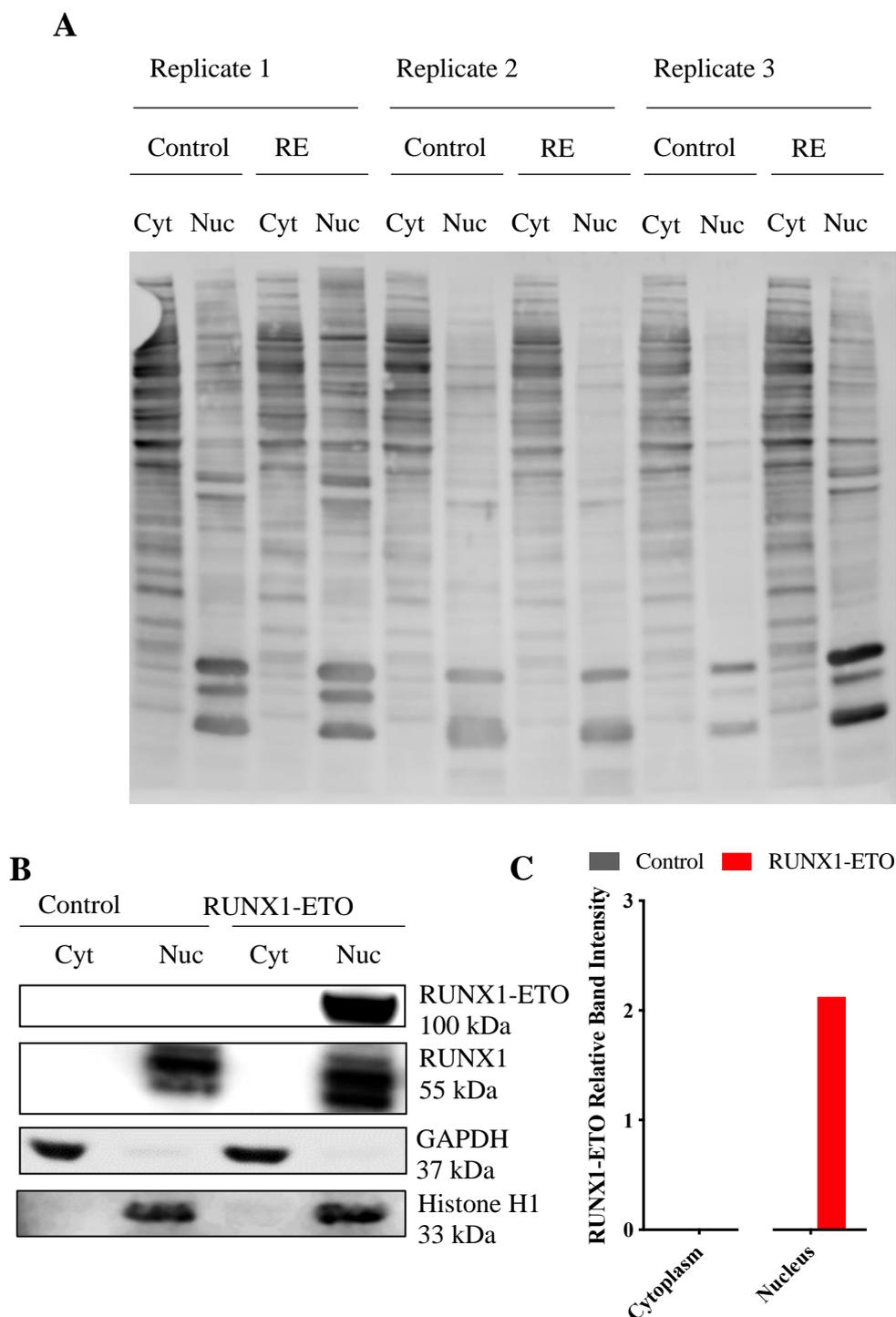
This study initially examined ‘housekeeping’ proteins exclusively expressed within the cytoplasmic or nuclear compartment to further confirm fractionation efficiency. As expected, GAPDH protein was found almost exclusively in the cytoplasm of both control and RUNX1-ETO expressing cells, whilst Histone H1 was solely detected in the nuclear compartment (**Figure 3.11**).

The protein library provides coverage for 50.9% of all human proteins annotated by UniProtKB/Swiss-Prot, which accounts for approximately 13,000 proteins (Rosenberger *et al.*, 2014). Unfortunately, as RUNX1-ETO is a fusion protein formed through abnormal translocations, it is not detectable in this setting; however, RUNX1 levels may be used as an approximate indicator for RUNX1-ETO levels. RUNX1 was expressed in the nucleus with levels significantly higher in RUNX1-ETO infected cells compared to control (**Figure 3.12**). It should be noted, this is not a definitive conclusion, as both endogenous RUNX1 and the RUNX1 portion of the fusion protein RUNX1-ETO can be detected.

**A****B**

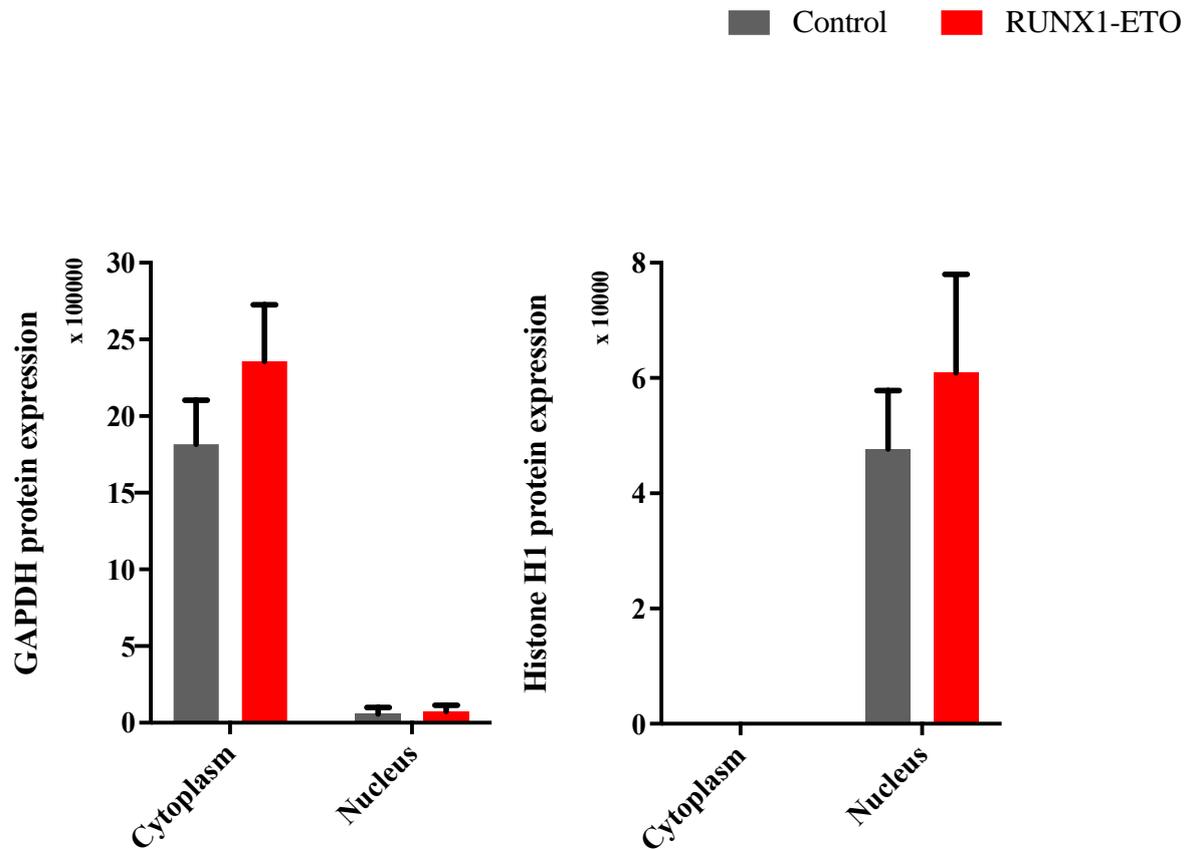
**Figure 3.9 – Viral infection of CD34<sup>+</sup> cells. CD34<sup>+</sup> HSPC cells were analysed for GFP expression on day 3 of cell culture**

Representative flow cytometry histograms showing percentage of HSPC expressing GFP in (A) Control (green); (B) RUNX1-ETO (red) transduced cells. Histograms were gated to exclude cell debris and dead cells based on FCS/SSC. Background auto-fluorescence was established using HSPC subjected to the equivalent retroviral infection procedure but in the absence of retrovirus (mock) (grey). Percentage within the black marker line represents proportion of cells showing fluorescence greater than background autofluorescence at < 0.01% (n=3).



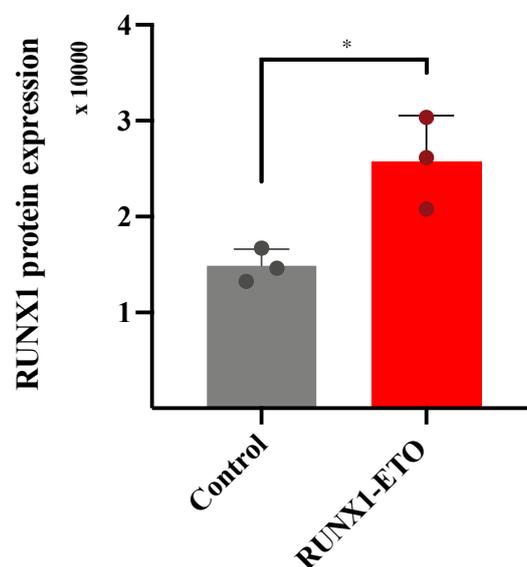
**Figure 3.10 – Quality control and characterisation of CD34<sup>+</sup> HSPC overexpressing RUNX1-ETO**

(A) Purity of the cytoplasmic (Cyt) and nuclear (Nuc) fractionated samples was assessed by immunoblotting 1  $\mu$ g of sample onto a PVDF membrane stained with SYPRO<sup>TM</sup> Ruby Protein Blot Stain. (B) Overexpression of RUNX1-ETO protein in CD34<sup>+</sup> cells transduced with RUNX1-ETO, when compared to control. Each lane was loaded with 10  $\mu$ g of protein. GAPDH and histone H1 were used as a loading control and as a verification of subcellular fraction for cytosolic and nuclear samples, respectively. This blot was a reprobe of **Figure 3.5B**, hence the same GAPDH and Histone H1 control images were used. (C) Western blot quantification by densitometry analysis using ImageJ. Relative protein expression was determined by normalising each cytosolic and nuclear sample to GAPDH or Histone H1, respectively, and to each corresponding ‘control’ (n=1).



**Figure 3.11 – GAPDH and Histone H1 protein quantitation by SWATH-MS supports acceptable subcellular fraction efficiencies for cytosol and nuclear protein isolation**

SWATH-MS analysis performed on cytosolic and nucleus fractions of CD34<sup>+</sup> HSPC allowed the detection of the GAPDH protein almost exclusively in the cell's cytoplasm, whilst Histone H1 was only detected in the cell's nucleus, on both control and RUNX1-ETO expressing cells. The Y-axis represents the protein expression value; the X-axis represents the different cell compartments analysed (n=3).



**Figure 3.12 – RUNX1 is significantly overexpressed in RUNX1-ETO transfected CD34<sup>+</sup> HSPC**

SWATH-MS analysis showed RUNX1 upregulation in the nucleus of CD34<sup>+</sup> HSPC transfected with a RUNX1-ETO retroviral vector, on day 3 of cell culture, as compared to the nuclear compartment of control cells. The Y-axis represents RUNX1 protein expression value (arbitrary units) (n=3). Significant differences between RUNX1-ETO and control cells was analysed using paired t-test; \* denotes p<0.05.

Data acquisition, filtering, and annotation was performed at the University of Manchester by Professor Anthony Whetton's group. Three approaches were undertaken in order to study protein dysregulation as a result of RUNX1-ETO expression, in CD34<sup>+</sup> HSPC.

Firstly, proteins were separated according to sub-cellular localisation, cytoplasmic or nuclear. For each compartment, proteins were filtered to those that were detected in all biological replicates in both control and RUNX1-ETO expressing cells (3/3 replicates), resulting in 4,635 detectable proteins, of which 2,787 were identified in the cytoplasm and 1,848 in the nucleus (explored in **3.3.2.4 -3.3.2.7**).

The second approach consisted of analysing proteins exclusively expressed in RUNX1-ETO cells, without detectable protein in control cells. This analysis is possible as SWATH-MS is unable to detect certain proteins, either because there is no expression, expression level is undetectable or due errors in the detection process (discussed in **3.4**). To prevent the latter from happening and influencing the results, this analysis will take into consideration proteins that failed to be detected in all three biological replicates of either control or RUNX1-ETO. This allowed the identification of 52 aberrantly expressed proteins in RUNX1-ETO cells, from which 29 were localised in the cytoplasm and 23 in the nucleus (**3.3.2.8**).

Lastly, the third approach will analyse all proteins detected in both the cytoplasm and nucleus of control and RUNX1-ETO cells to identify proteins that could have suffered a shift in normal protein localisation as a result of RUNX1-ETO expression. This analysis will consider the expression of 926 proteins, explored in **3.3.2.9**.

Taken together, the above data demonstrates highly efficient sub-cellular fractionations in control or RUNX1-ETO expressing cells. In addition, SWATH-MS was able to quantify over 4,500 proteins suitable for downstream targets analysis.

#### *3.3.2.4 Exploratory data analysis*

To assess the relationship between control and RUNX1-ETO samples, two methods were employed. Firstly, principal component analysis (PCA) is a useful tool which allows the visualisation of patterns from large and complex datasets. Additionally, hierarchical clustering, consisting of a statistical method used to assign objects into groups, or clusters, allows the highlighting of patterns in the protein expression dataset, based on sample type (RUNX1-ETO

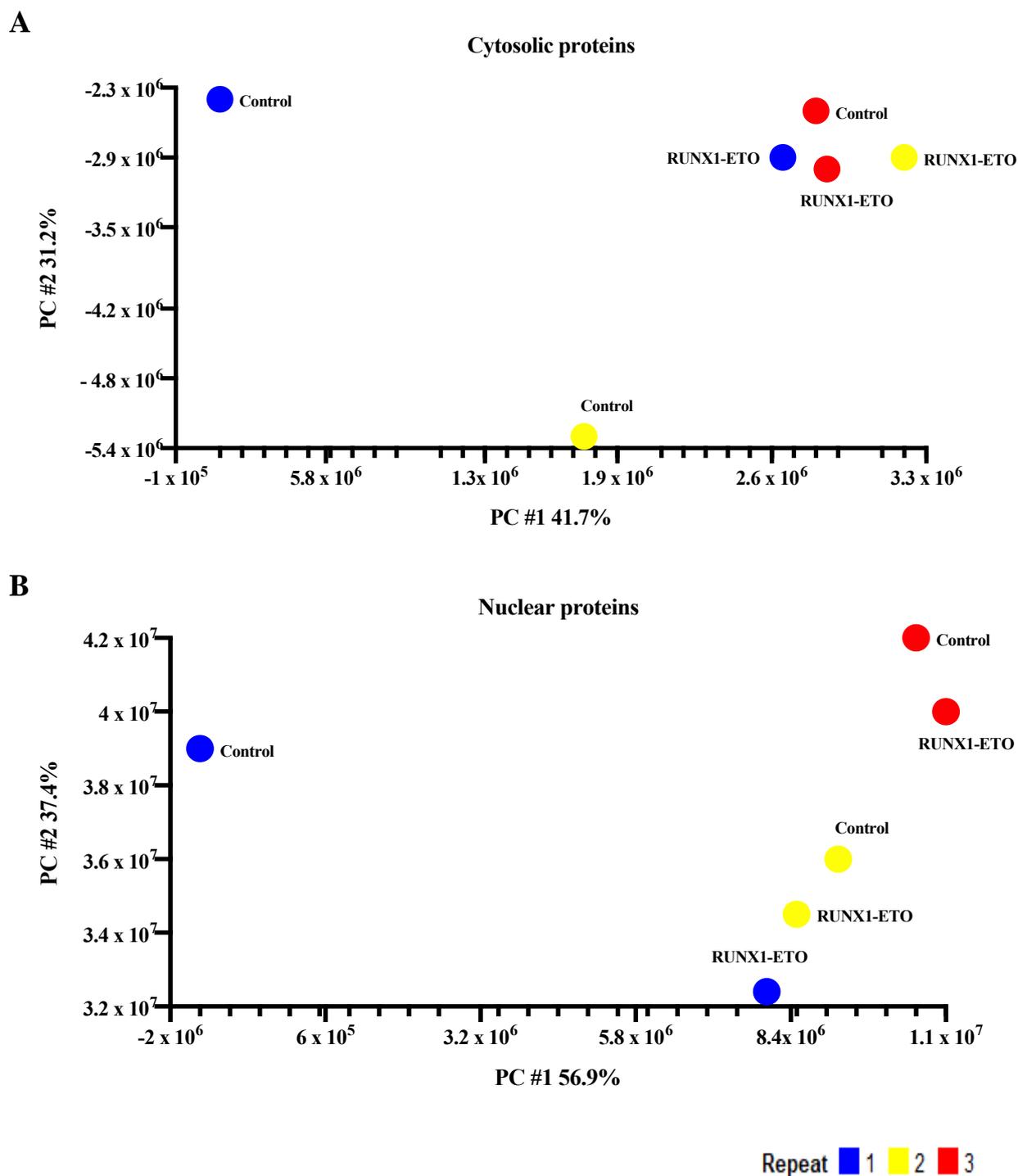
vs. control CD34<sup>+</sup> HSPC). This analysis was based on protein expression according to their sub-cellular localisation (4,635 proteins – 2,768 cytosolic and 1,848 nuclear proteins).

Regarding cytosolic proteins, **Figure 3.13A** and **Figure 3.14A** show the clustering of RUNX1-ETO samples, indicating that these cells have a protein expression profile similar to each other, whilst the control samples cluster separately. Nuclear protein expression data, on the other hand, show a similarity between 2/3 RUNX1-ETO samples and, once more, no relationship between control samples.

Altogether, this analysis indicates a distinction between control and RUNX1-ETO-expressing CD34<sup>+</sup> HSPC; however, control samples did not cluster together, indicating variable patterns of expression, along with sample variability, which might be sourced from the fact that samples were prepared from different cord blood donations. Moreover, the fact that RUNX1-ETO samples were shown to cluster together suggests that this is a result of 'oncogenic' induction, considering that all samples were subjected to the same viral transduction. For this reason, a paired parametric statistical test with ANOVA was performed to determine statistically significant changing proteins, arising from the expression of the fusion protein RUNX1-ETO.

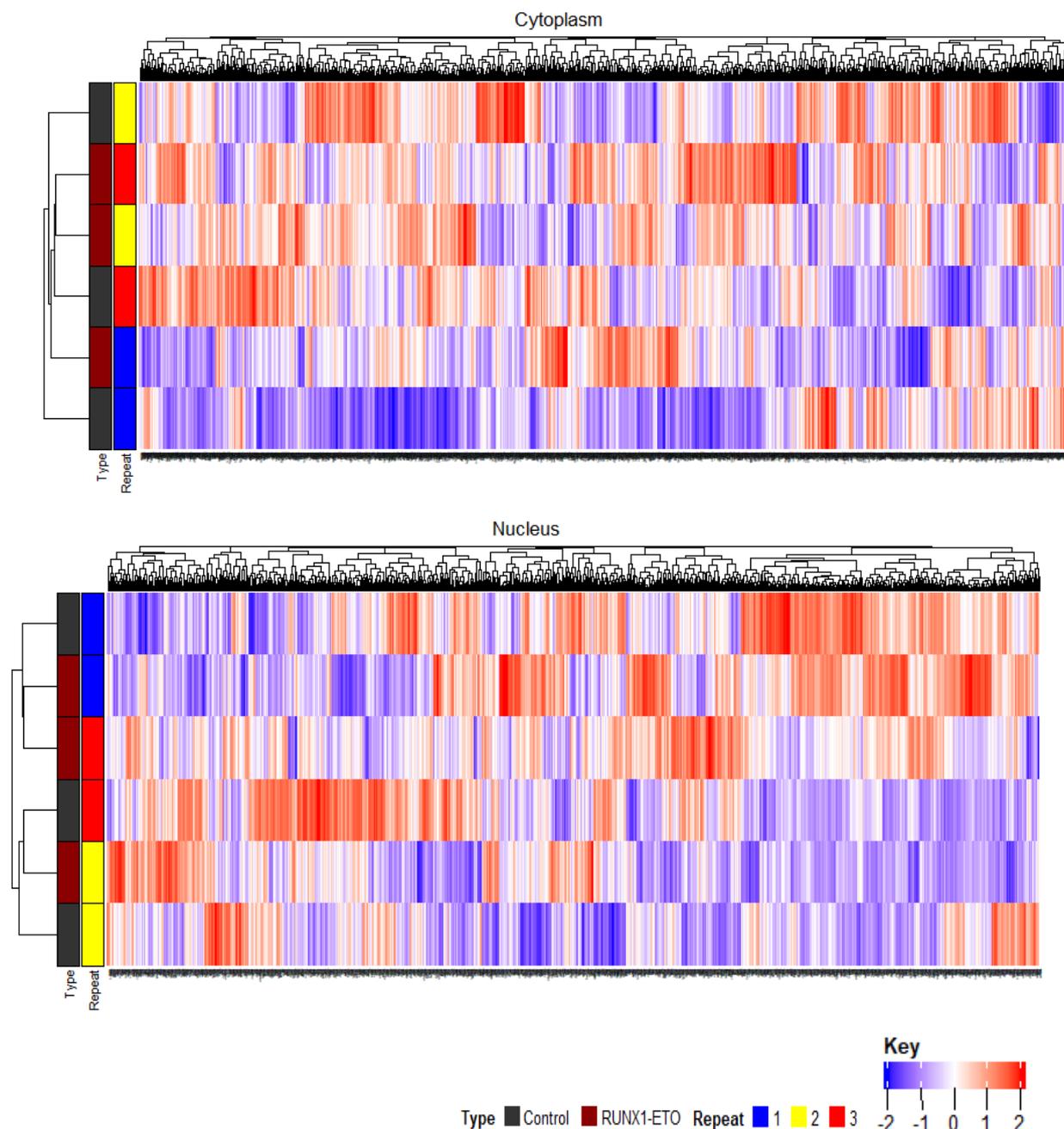
#### 3.3.2.5 *Statistical model used for data analysis*

Following normalization ( $\log_2$ ) of the protein expression dataset, a one-way ANOVA statistical test was employed to identify significantly changed proteins in the cytoplasm and nucleus of CD34<sup>+</sup> HSPC, arising from the expression of the fusion protein RUNX1-ETO. Uncorrected p-values were considered as this study represents an exploratory investigation due to low sample size. An uncorrected cut-off p-value of  $< 0.05$  was applied to generate two lists of significantly changing cytosolic and nuclear proteins between CD34<sup>+</sup> HSPC expressing RUNX1-ETO and control cells (**Supplementary Table 2**). This resulted in a combined list of 257 proteins significantly dysregulated as a consequence of RUNX1-ETO expression, from which 183 were present in the cytoplasm and 74 were found in the nucleus (**Figure 3.15A-B**).



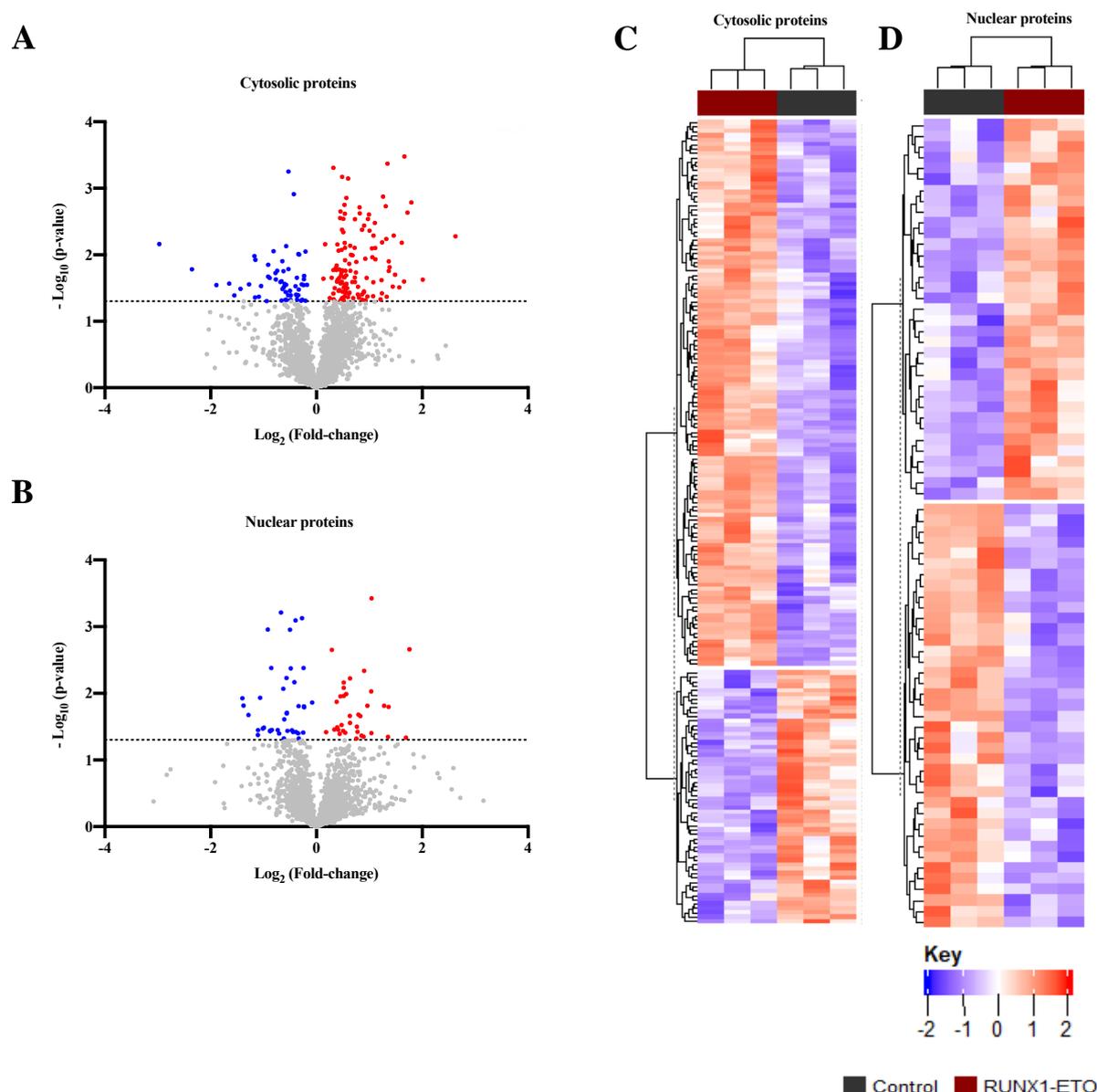
**Figure 3.13 – Principal Component Analysis (PCA) of the protein expression data**

Protein expression data was analysed based on its cellular localisation: (A) cytoplasm or (B) nucleus. Protein expression from each sample is represented by coloured dots, according to each biological replicate. Each dot represents individual cord blood samples transduced with retroviral vectors. For each replicate of the experiment, different cord blood was used, and three replicates were performed per condition (n=3).



**Figure 3.14 – Hierarchical clustering of proteins expressed in the cytoplasmic and nuclear compartment of CD34<sup>+</sup> HSPC**

Heatmap of Control and RUNX1-ETO-expressing CD34<sup>+</sup> HSPC biological replicate samples (n=3). Protein expression values were uploaded into Partek<sup>®</sup> and a Euclidean correlation was applied as the distance metric to cluster the proteins. Columns represent proteins; rows represent samples, defined by experimental condition and replicate number. Dendrogram trees show the hierarchy of clusters for both samples and genes. The colour bar (key) represents the z-score from mean expression, calculated for each protein. The heatmaps were divided into cytosolic and nuclear proteins. Red indicates upregulation and blue indicates downregulation of proteins in RUNX1-ETO CD34<sup>+</sup> HSPC when compared to control cells.



**Figure 3.15 – Differential protein expression analysis between  $\text{CD34}^+$  HSPC expressing RUNX1-ETO and control**

*Left panel:* Volcano plots of proteins differentially expressed in the cytosol and nucleus of RUNX1-ETO cells, when compared to control. The Y-axis represents statistical significance ( $-\text{Log}_{10} \text{p-value}$ ), whilst the X-axis indicates fold-change ( $\text{Log}_2$  fold-change) of protein expression in cord blood derived  $\text{CD34}^+$  cells transfected with either a RUNX1-ETO or a control vector ( $n=3$ ). (A) 183 cytosolic proteins were found to be statistically significantly changed in RUNX1-ETO cells compared to control, whereas (B) 74 proteins in the nucleus were identified using the same approach. The significance cut-off ( $p < 0.05$ ) is highlighted with a dotted line. Colour coding is based on fold change, in which blue indicates repressed proteins in RUNX1-ETO  $\text{CD34}^+$  HSPC when compared to control, whilst proteins in red were found to be overexpressed in the same cells.

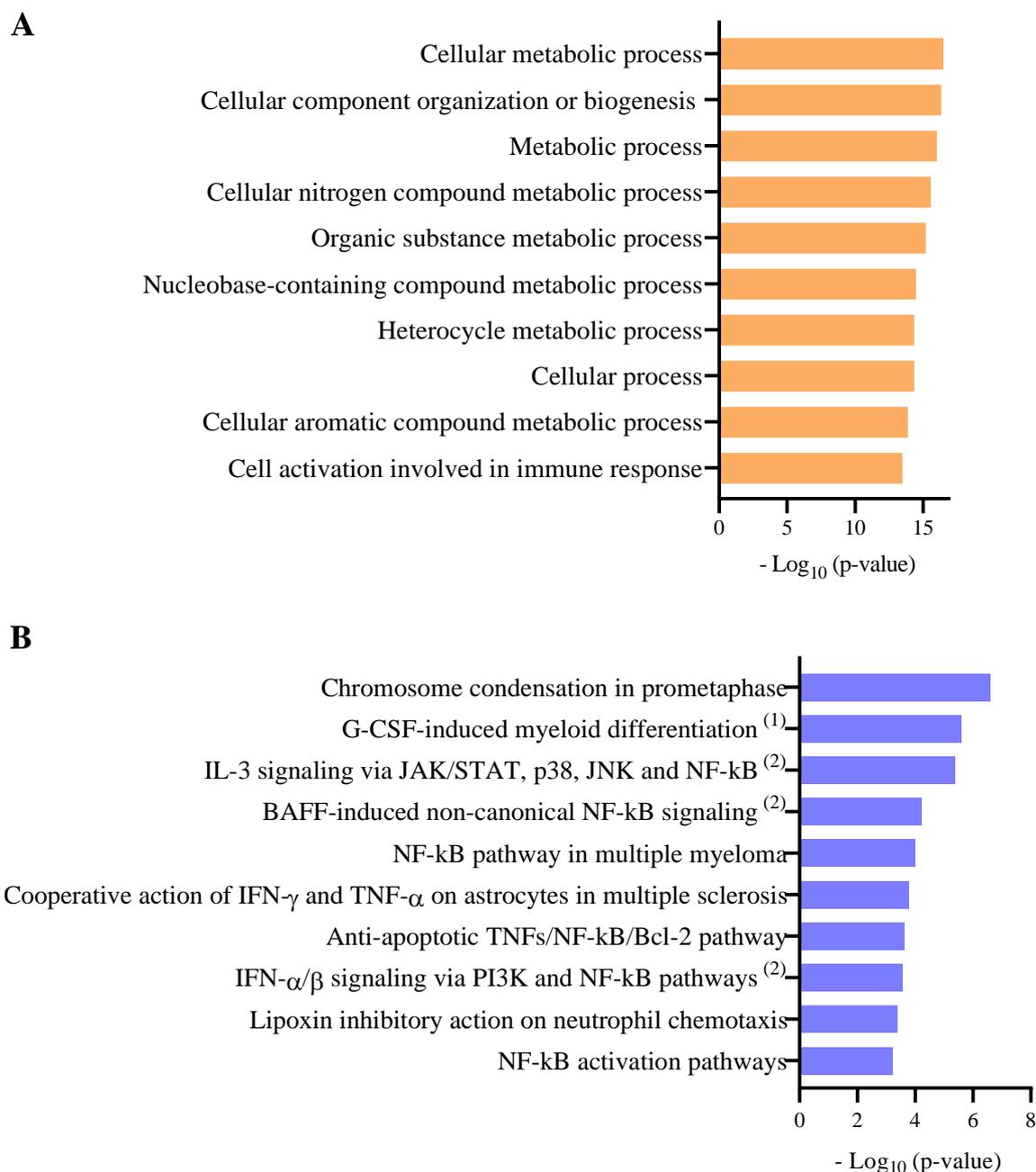
*Right panel:* A hierarchical clustering algorithm was used to cluster differentially expressed proteins in control and RUNX1-ETO  $\text{CD34}^+$  HSPC. Columns represent samples, defined by experimental condition; rows represent proteins. Dendrogram trees show the hierarchy of clusters for both samples and genes. The colour bar (key) represents the z-score from mean expression, as it's been calculated for each protein. Red indicates up-regulation, blue indicates down-regulation of a specific protein. Heatmap showing differential expression of (C) 183 cytosolic and (D) 74 nuclear proteins.

### 3.3.2.6 Expression of RUNX1-ETO disrupts normal protein expression in CD34<sup>+</sup> cord blood derived HSPC

Comparably to the analysis performed to identify differentially expressed genes changes in RUNX1-ETO CD34<sup>+</sup> HSPC, the 257-protein list was uploaded into Metacore™ for pathway and downstream target analysis. Firstly, to understand the underlying biological processes and pathways which might be regulated by RUNX1-ETO, functional Enrichment Analysis was performed (**3.3.1.1**). GO analysis identified several altered processes in cells overexpressing RUNX1-ETO (**Figure 3.16A**). Interestingly, these included ‘Metabolic processes’, ‘Biogenesis’ and ‘Immune response’. Even though disrupted metabolic processes have not been widely described in AML t(8;21), leukaemic cell proliferation has been shown to require the up-regulation and rewiring of metabolic pathways to promote anabolic cell growth (Rashkovan and Ferrando, 2019; Hanahan and Weinberg, 2011; Robinson *et al.*, 2021).

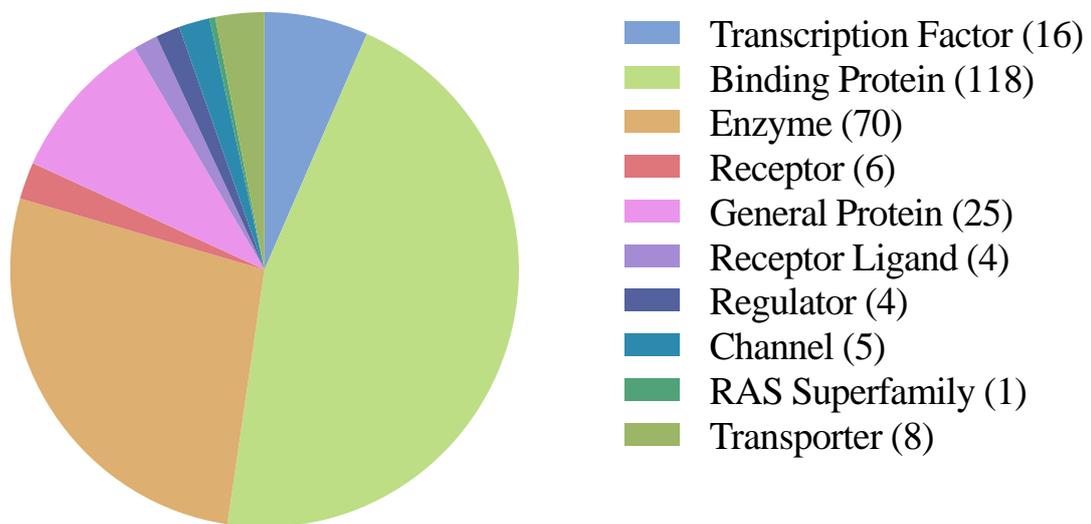
To support GO analysis, an alternative approach determining the most significant dysregulated pathways was performed (**Figure 3.16B**). The most significantly influenced processes included ‘Cell cycle regulation’, ‘Myeloid differentiation’, and changes in the NF- $\kappa$ B pathway. Aberrant activation of the NF- $\kappa$ B pathway has been described in AML, as its been shown to regulate cell survival and apoptosis (Zhou *et al.*, 2015). These processes are similar to the changes observed in the transcriptome of RUNX1-ETO cells, described above, and agree with phenotypical changes observed *in vitro* and in AML t(8;21) patients (**3.3.1.1**). Furthermore, this analysis demonstrates the wide range of processes implicated in the development of AML, arising as a consequence of the expression of the RUNX1-ETO fusion protein.

To clarify the classes of proteins dysregulated by RUNX1-ETO, proteins were analysed for their attributed function in Metacore™: TF, binding proteins, enzymes, receptors, general proteins, receptor ligands, regulators, channel proteins, RAS superfamily and transporters (**Figure 3.17**). Binding proteins and enzymes were found to represent more than half of the total number of proteins significantly altered by the expression of RUNX1-ETO. This supports my analysis of RUNX1-ETO dysregulated mRNA (**3.3.1.1**).



**Figure 3.16 – Functional Enrichment Analysis of protein changes observed in normal haematopoietic development vs. RUNX1-ETO expressing CD34<sup>+</sup> HSPC**

Enrichment Analysis using Metacore™ of 250 proteins identified to be significantly dysregulated in CD34<sup>+</sup> HSPC expressing RUNX1-ETO when compared to control. Top 10 Significantly dysregulated (A) GO processes and (B) Pathway Maps, ranked by  $-\text{Log}(\text{p-value})$ . Processes can further be clustered into <sup>(1)</sup> Developmental, <sup>(2)</sup> Immune response-related or <sup>(3)</sup> Asthma-related processes. A False Discovery Rate (FDR) of 0.05 was applied to both analyses.



**Figure 3.17 – Classification of differentially expressed proteins due to the expression of RUNX1-ETO in CD34<sup>+</sup> HSPC**

Pie chart indicating the classification of the differentially expressed genes using Metacore™. 257 proteins were found to be significantly dysregulated in RUNX1-ETO expressing cells when compared to control.

### 3.3.2.7 Identification of aberrantly expressed TF in CD34<sup>+</sup> HSPC expressing RUNX1-ETO

Pathway analysis described above can be supported by changes arising from TF dysregulated expression. This study was able to identify 16 statistically significantly dysregulated TF proteins arising from the expression of the RUNX1-ETO fusion protein (**Table 3.3**). From these, 11 TFs were found to be upregulated in CD34<sup>+</sup> HSPC expressing RUNX1-ETO, as compared to control, whilst 5 were shown to be downregulated under the same conditions. Most of the TFs were localised in the nucleus, as expected; however, 4 were found to have their expression significantly dysregulated in the cytoplasm of CD34<sup>+</sup> cells (**Table 3.3**). To identify the level of TF connectivity with other proteins from the dataset, this study used the Metacore™ algorithm ‘Interactome Transcription Factors’ on the significantly perturbed proteins (257 proteins). This analysis identified 4 overconnected TF, all present in the nucleus of CD34<sup>+</sup> HSPC, detailed in **Table 3.3**. CBFβ and RUNX1 were found to be two of the TF with the highest number of interactions with other proteins from the dataset. Both proteins belong to the CBF complex and are responsible not only for promoting the expression of other TF involved in cell differentiation, but also growth factors and proliferation and survival regulators. *RUNX1* was shown to be equally overexpressed in RUNX1-ETO-expressing CD34<sup>+</sup> HSPC, on day 3 of cell culture, as compared to control HSPC. Even though there was an increase in CBFβ protein expression arising from the expression of RUNX1-ETO, there was no difference in *CBFβ* mRNA expression levels under the same conditions, as compared to control CD34<sup>+</sup> HSPC.

This analysis was also able to identify PU.1, a master regulator of myeloid cells, as a TF responsible for the higher number of changes observed at the protein level in CD34<sup>+</sup> HSPC upon the expression of RUNX1-ETO compared to control cells. PU.1, a known gene/protein to be dysregulated in t(8;21), was also identified in the mRNA analysis (**3.3.1.2**). In this context, *PU.1* mRNA expression was concordant with protein expression levels, showing similar levels of downregulation in RUNX1-ETO-expressing CD34<sup>+</sup> HSPC.

This analysis also identified C/EBPβ as a TF with the highest influence on other proteins from the dataset. This protein is a member of the CCAT enhancer-binding protein (C/EBP) family, which have been known to play important roles in proliferation and differentiation, including the suppression of myeloid leukemogenesis (Ramji and Foka, 2002). C/EBPβ specifically has been shown to reprogram B cells to the myeloid lineage, in mouse splenic and bone marrow cells (Heavey *et al.*, 2003).

**Table 3.3 – Identification of RUNX1-ETO induced TF protein dysregulation in CD34<sup>+</sup> HSPC**

TF protein expression was determined by performing SWATH-MS analysis on cytosolic and nuclear extracts of CD34<sup>+</sup> HSPC transfected either with a RUNX1-ETO retroviral vector, or a control plasmid, on day 3 of cell culture. To identify significantly overconnected genes within the dataset, the Metacore™, ‘Interactome Transcription Factors’ algorithm was used. ✓ and ✗ denotes if the gene was identified or not (respectively) by the algorithm. p-values were as described in 2.10.

§ Normalised microarray expression (log<sub>2</sub>) was compared between RUNX1-ETO and control CD34<sup>+</sup> HSPC by microarray analysis, performed on day 3 of cell culture (adapted from (Tonks *et al.*, 2007)). (↑) (↓) denotes upregulation or downregulation of corresponding gene in RUNX1-ETO CD34<sup>+</sup> HSPC, respectively, as compared to control.

RNASeq data from the TCGA 2013 dataset (Ley *et al.*, 2013) (GSE13159) was used to determine mRNA expression in AML M2 patients, divided into t(8;21), expressing the fusion protein RUNX1-ETO, and non-t(8;21) patients. mRNA expression in normal human HSC were obtained from (Rapin *et al.*, 2014) (GSE42519). Each gene has a corresponding Affymetrix ID unique to the Hu133A GeneChip® (Affymetrix®).

Fold-change (FC) values represent the regulation of each gene when compared to control cells (green – upregulated; red – repressed, when compared to control cells). (↑) denotes upregulated or (↓) downregulated gene expression in patients with t(8;21) as compared to normal undifferentiated HSC. (n/a denotes not applicable; CDS denotes coding sequence; Loc. denotes localisation within the cell; C denotes cytoplasm, N denotes nucleus).

Protein	UniProtID	p-value	FC	Loc.	Metacore™		CD34 <sup>+</sup> HSPC§		Literature Review	AML M2 t(8;21) vs non t(8;21)	AML t(8;21) vs. HSC	CDS (kbp)
					Interactome TF	p-Value	Reg.	FC				
NFKB2	Q00653	3.30 x 10 <sup>-4</sup>	3.2	C	✗	-	n/a		n/a	n/a	n/a	
TRIM28	Q13263	7.16 x 10 <sup>-4</sup>	1.5	C	✗	-	n/a		n/a	n/a	n/a	
PU.1	P17947	1.11 x 10 <sup>-3</sup>	-1.9	N	✓	1.453 x 10 <sup>-10</sup>	↓	-1.76	(Vangala <i>et al.</i> , 2003; Eder <i>et al.</i> , 2016)	n/a	n/a	
KMT2A	Q03164	5.97 x 10 <sup>-3</sup>	1.6	N	✗	-	n/a		n/a	n/a	n/a	
ZBTB48	P10074	6.90 x 10 <sup>-3</sup>	1.4	N	✗	-	n/a		n/a	n/a	n/a	
LRRFIP1	Q32MZ4	7.01 x 10 <sup>-3</sup>	1.31	C	✗	-	n/a		n/a	n/a	n/a	
EDF1	O60869	1.10 x 10 <sup>-3</sup>	1.4	N	✗	-	n/a		n/a	n/a	n/a	

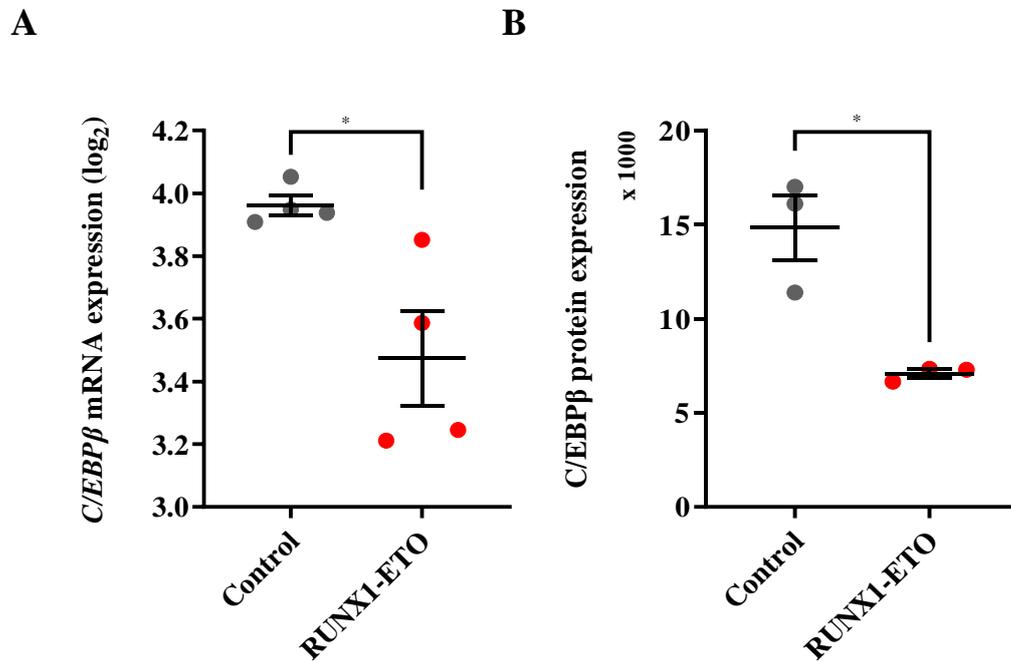
Protein	UniProtID	p-value	FC	Loc.	Metacore™		CD34 <sup>+</sup> HSPC <sup>s</sup>		Literature Review	AML M2 t(8;21) vs non t(8;21)	AML t(8;21) vs. HSC	CDS (kbp)
					Interactome TF	p-Value	Reg.	FC				
<b>CEBPβ</b>	P17676	1.17 x 10 <sup>-2</sup>	-2.1	N	✓	1.39 x 10 <sup>-4</sup>	↓	-1.6	-	↓	↑	~ 1
<b>RB1</b>	P06400	1.54 x 10 <sup>-2</sup>	2.4	N	✗	-	n/a		n/a	n/a		n/a
<b>RUNX1</b>	Q01196	2.08 x 10 <sup>-2</sup>	1.7	N	✓	1.53 x 10 <sup>-32</sup>	↑	1.91	n/a	n/a		n/a
<b>CBFβ</b>	Q13951	2.99 x 10 <sup>-2</sup>	1.4	N	✓	1.571 x 10 <sup>-5</sup>	No change		(Roudaia <i>et al.</i> , 2009)	n/a		n/a
<b>BAZ2A</b>	Q9UIF9	3.66 x 10 <sup>-2</sup>	1.4	N	✗	-	n/a		n/a	n/a		n/a
<b>TTF2</b>	Q9UNY4	3.71 x 10 <sup>-2</sup>	-1.8	N	✗	-	n/a		n/a	n/a		n/a
<b>HIRA</b>	P54198	3.80 x 10 <sup>-2</sup>	-1.4	N	✗	-	n/a		n/a	n/a		n/a
<b>ZBTB49</b>	Q6ZSB9	3.89 x 10 <sup>-2</sup>	-1.2	N	✗	-	n/a		n/a	n/a		n/a
<b>TSHZ1</b>	Q6ZSZ6	4.49 x 10 <sup>-2</sup>	1.6	C	✗	-	n/a		n/a	n/a		n/a

However, no studies have been performed to determine the role of this protein in the development of AML t(8;21). In CD34<sup>+</sup> HSPC, this gene was found to be repressed both at the mRNA and the protein level, by 1.6- and 2.1-fold, respectively, upon the expression of RUNX1-ETO (**Figure 3.18**). Moreover, the *CEBPB* gene was seen to be significantly downregulated in AML patients harbouring the t(8;21), as compared to other FAB-M2 diagnosed patients (**Table 3.3**). For this reason, C/EBP $\beta$  has been identified as an interesting target and its role in the development of AML t(8;21) will be determined in **Chapter 5**.

### 3.3.2.8 *RUNX1-ETO leads to the aberrant expression of 53 proteins*

As described above, this study identified proteins whose expression was found to be significantly (statistically) dysregulated in RUNX1-ETO expressing CD34<sup>+</sup> HSPC, as compared to control. However, statistical analysis is not possible in all samples in which there is no detectable protein expression due to missing expression values. Hence, an additional method of analysis was applied based on proteins solely detected in RUNX1-ETO cells, with no protein being detected in control cells. In total, 53 proteins, of which 29 were expressed in the cytoplasm and 24 in the nucleus, were identified in RUNX1-ETO expressing CD34<sup>+</sup> HSPC but could not be detected in control cells under the same conditions.

Of these proteins, 5 were found to be TFs or regulators of transcription, 4 of which detected in the cytoplasm and 1 in the nucleus of RUNX1-ETO CD34<sup>+</sup> HSPC (**Table 3.4**). The fact that these proteins were found to be in cytoplasm, as opposed to the nucleus, could suggest that their normal function is somehow impaired in these cells; however, they were not detected in control cells at all, indicating that are either being expressed earlier in RUNX1-ETO cells, given that analysis was performed on day 3 of blood cell development, or RUNX1-ETO is somehow promoting the expression of these proteins (**Figure 3.19A**). Whilst it's appreciated that TF would only be expected to have transcriptomic consequences within the nucleus, aberrant expression of a TF, which is not normally expressed, could lead to mis-appropriate shuttling in certain situations. However, as protein is generated within the cytoplasmic compartment of the cell, a measure of background detection of nuclear proteins can occur. Nevertheless, as these TF were not detected in control cells, they are aberrant in expression in RUNX1-ETO AML. ATXN7 and MIER1 were identified as the most highly expressed in RUNX1-ETO CD34<sup>+</sup> HSPC.



**Figure 3.18 – C/EBPβ expression profile in CD34<sup>+</sup> HSPC expressing RUNX1-ETO**

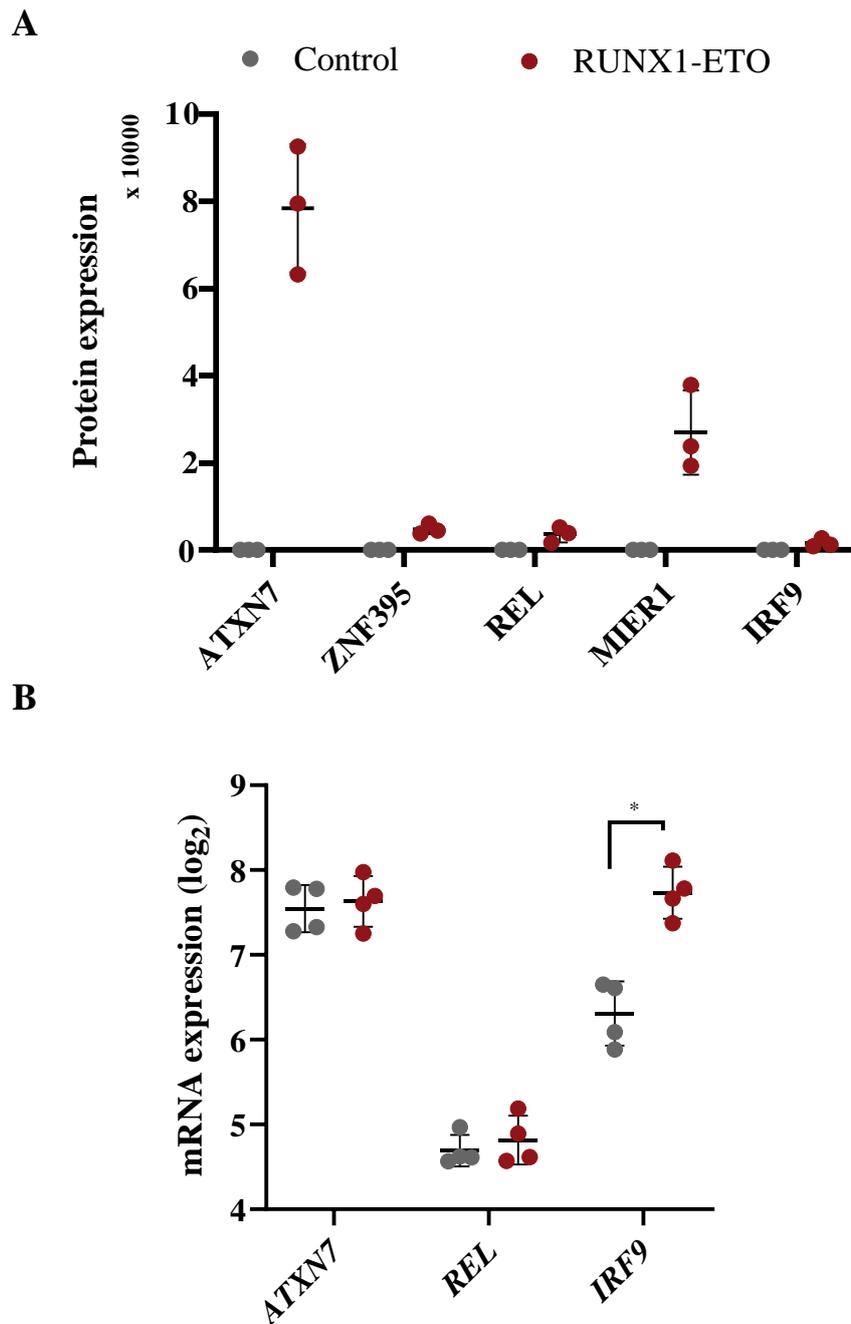
(A) Normalised microarray data showing log<sub>2</sub> *CEBPB* mRNA expression in normal human CD34<sup>+</sup> HSPC expressing RUNX1-ETO compared to control on day 3 of cell culture (n=4) (adapted from (Tonks *et al.*, 2007)). Significant differences between RUNX1-ETO and control cells was analysed using paired t-test; \* denotes p<0.05. (B) C/EBPβ protein expression profile was obtained by performing SWATH-MS analysis on CD34<sup>+</sup> HSPC transfected with either a control or a RUNX1-ETO; this protein was found to be expressed solely in the nucleus (Paired T-test; \* denotes p < 0.05)

**Table 3.4 – TF detected exclusively in RUNX1-ETO expressing CD34<sup>+</sup> HSPC**

§ mRNA expression (log<sub>2</sub>) was compared between RUNX1-ETO and control CD34<sup>+</sup> cells by microarray analysis, performed on day 3 of blood cell development (adapted from (Tonks *et al.*, 2007).

‡ Protein expression profiles were obtained by performing SWATH-MS analysis on CD34<sup>+</sup> HSPC transfected with either a control or a RUNX1-ETO expressing vector. Cytosolic and nuclear proteins were extracted on day 3 of blood cell development. ‘Detected’ indicates protein detection in RUNX1-ETO CD34<sup>+</sup> HSPC. (TReg denotes transcriptional regulator).

Protein	UniProtID	Localisation	Classification	mRNA expression in CD34 <sup>+</sup> HSPC §	Protein expression in CD34 <sup>+</sup> HSPC ‡	CDS kbp
<b>ATXN7</b>	O15265	Cytoplasm	TReg	No change	Detected	~ 2.7
<b>ZNF395</b>	Q9H8N7	Cytoplasm	TReg	Not detected	Detected	~ 1.4
<b>REL</b>	Q04864	Cytoplasm	TF	No change	Detected	~ 1.8
<b>IRF9</b>	Q00978	Cytoplasm	TF	↑	Detected	~ 1.4
<b>MIER1</b>	Q8N108	Nucleus	TReg	Absent from chip	Detected	~ 1.7



**Figure 3.19 – Target protein and mRNA expression profiles of TF found exclusively in RUNX1-ETO CD34<sup>+</sup> HSPC**

(A) Protein expression profiles were obtained by performing SWATH-MS analysis on CD34<sup>+</sup> HSPC transfected with either a control or a RUNX1-ETO expressing vector. Cytosolic and nuclear proteins were extracted on day 3 of blood cell development. Data indicates Mean  $\pm$  1SD (n=3). (B) Normalised microarray data showing  $\log_2$  mRNA expression of the identified target genes, where data was available. mRNA expression in normal human CD34<sup>+</sup> HSPC and cells transfected with a RUNX1-ETO expressing vector was measured on day 3 of blood cell development (n=4) (adapted from (Tonks *et al.*, 2007)). Significant differences between RUNX1-ETO and control cells was analysed using paired t-test; \* denotes  $p < 0.05$ .

Given the differences observed in protein expression between RUNX1-ETO and control CD34<sup>+</sup> HSPC, this study analysed the corresponding mRNA expression levels to determine if these differences could be observed at the transcriptomic level, or if these are transcriptionally driven. This analysis was unable to detect two genes, *ZNF385* and *MIER1*: even though the microarray chip used contained the probe-set for the *ZNF385* transcript, this was not detected. Moreover, no probe-set for the *MIER1* transcript can be found on the chip used (Hu133A). Previously, *IRF9* was shown to be significantly upregulated in RUNX1-ETO cells, as compared to control CD34<sup>+</sup> HSPC (3.3.1.2); however, this analysis showed that whilst *IRF9* mRNA was present in control CD34<sup>+</sup> HSPC, no protein was detected in these cells. Furthermore, no differences were observed in the mRNA levels of the *ATXN7* and *REL* genes in RUNX1-ETO expressing CD34<sup>+</sup> HSPC, when compared to control cells (**Figure 3.19B**).

A subsequent feasibility assay was performed in order to ensure that overexpression/knockdown studies were possible, in the model systems used. All genes analysed presented a CDS below 3kbp, making them suitable for subsequent functional studies (**Table 3.4**).

This analysis showed significant differences between mRNA and resulting protein expression, thus strengthening previous observations that mRNA analysis can be insufficient to infer protein expression. The fact that these proteins were found to be solely detected in RUNX1-ETO cells, even though its mRNA can also be detected in control cells, could suggest that changes are post-transcriptionally driven; however, it is unlikely that these are direct transcriptional targets of RUNX1-ETO. This study found that no gene expression correlated with its corresponding protein, except IRF9. Interestingly, IRF9 had already been identified as a target of interest when analysing the transcriptomic profile of CD34<sup>+</sup> HSPC expressing the fusion protein RUNX1-ETO (3.3.1.2); the fact that it was identified as being expressed solely in RUNX1-ETO cells suggests that this protein may play a role in the development of AML t(8;21). Nevertheless, all targets are of interest, as their expression seems to be promoted either directly or indirectly by RUNX1-ETO.

#### 3.3.2.9 *RUNX1-ETO leads to a shift in protein sub-cellular localisation in CD34<sup>+</sup> HSPC.*

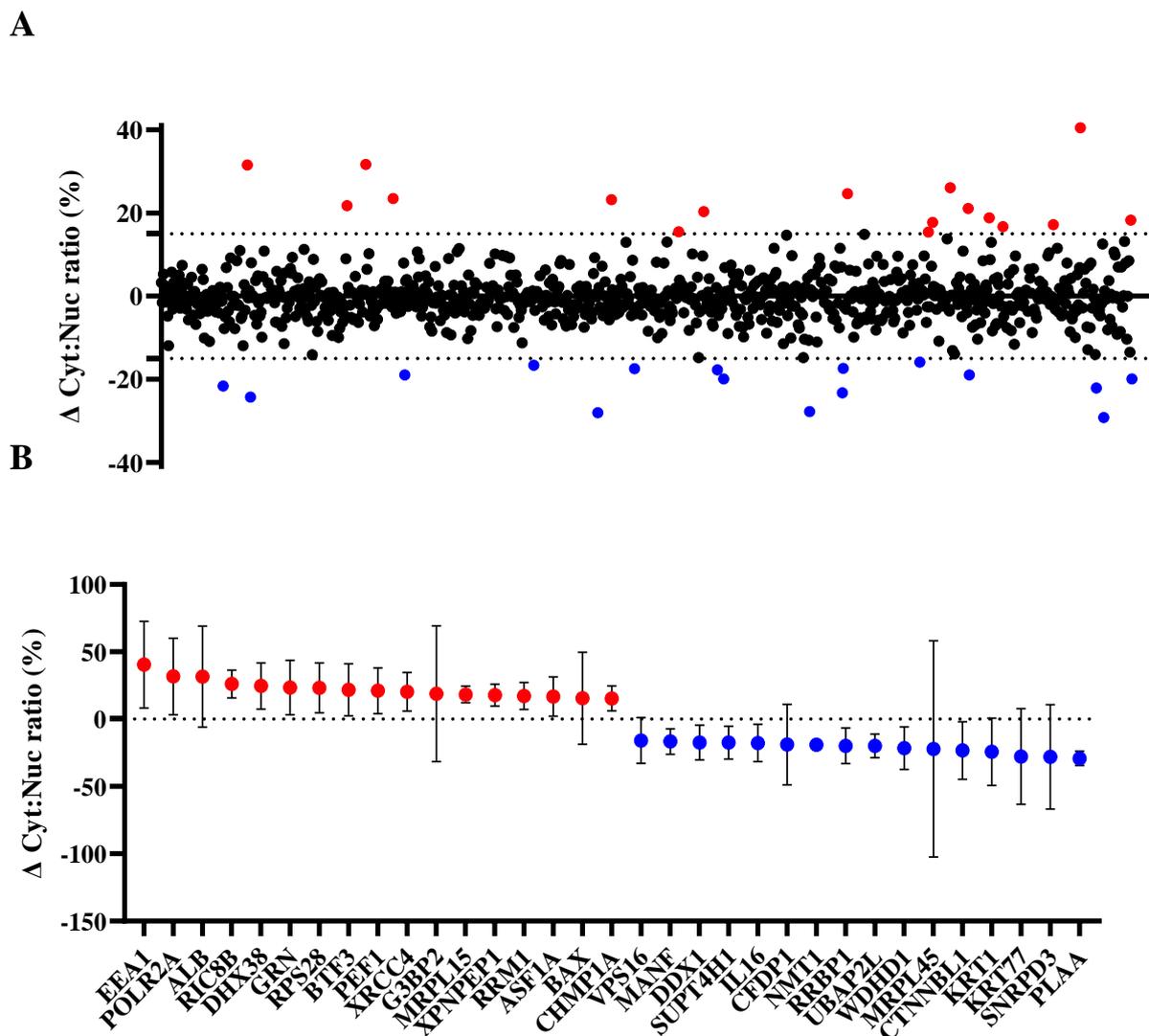
I next examined whether the expression of RUNX1-ETO was associated with a change in protein sub-cellular localisation in CD34<sup>+</sup> HSPC. A total of 926 proteins were identified as

being expressed in both the cytoplasm and the nucleus of control and RUNX1-ETO-expressing CD34<sup>+</sup> HSPC. Considering the total amount of protein and its relative sub-cellular localisation, 33 proteins were determined to be localised differently to proteins in normal CD34<sup>+</sup> HSPC as a consequence of RUNX1-ETO expression (**Figure 3.20A**).

The relative expression of 17 proteins was found to be higher in the cytoplasm of RUNX1-ETO-expressing cells, as compared to the cytoplasm of control cells. The EEA1, POLR2A, ALB and RIC8B proteins exhibited the greatest abundance in the cytosolic compartment of RUNX1-ETO cells, whilst in control cells, these proteins were found predominantly in the nucleus (**Figure 3.20B**). Conversely, 16 proteins were identified as being more abundant in the nucleus of RUNX1-ETO cells, compared to control, possibly indicating a translocation from the cytoplasm to the nucleus, as a consequence of the expression of this fusion protein.

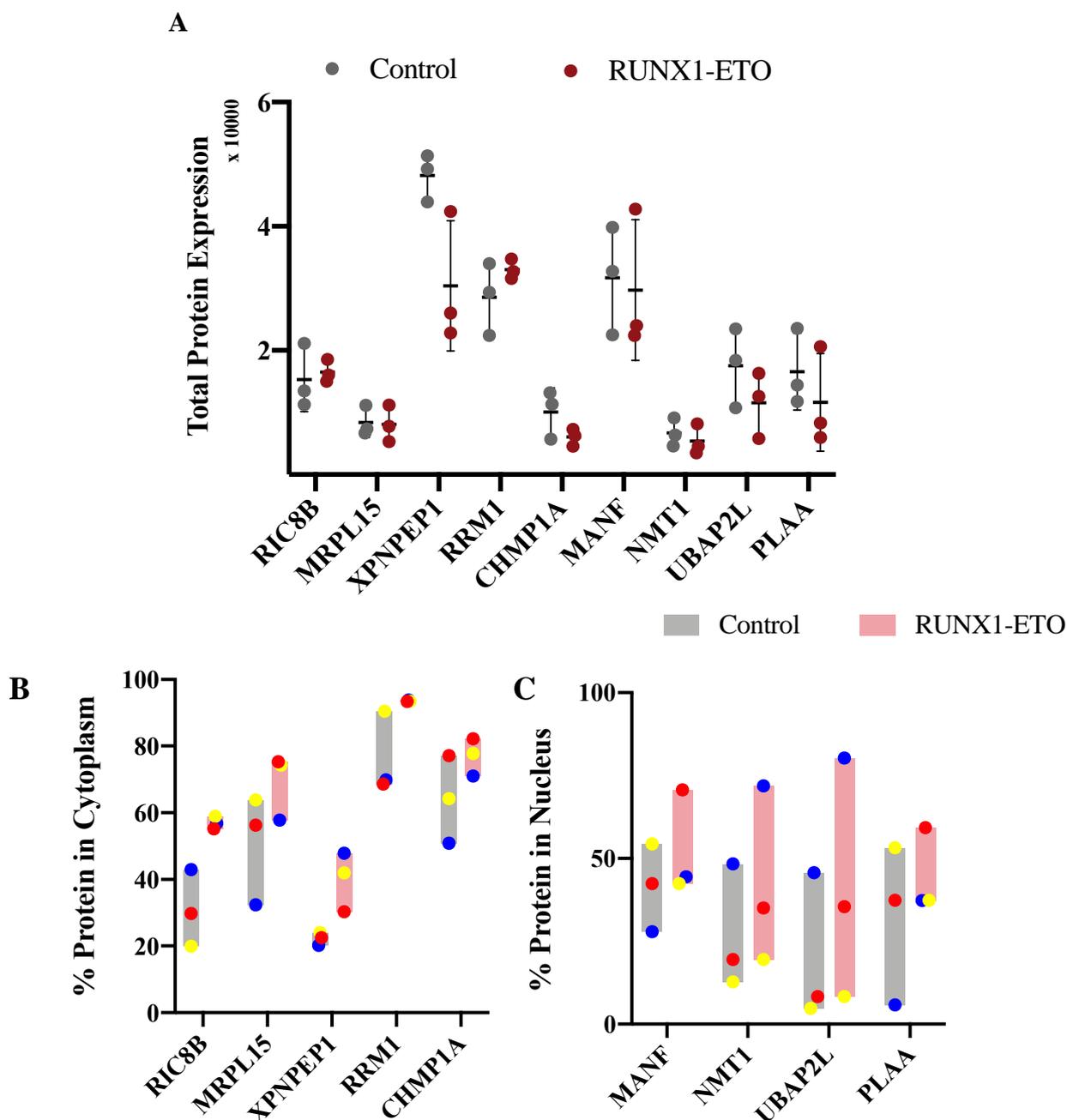
Of these, KRT77, SNRPD3 and PLAA showed the lowest detectable levels in the cytoplasm of RUNX1-ETO cells, when comparing to normal protein expression (**Figure 3.20B**). Due to the increased variability arising from these studies, of the 33 proteins, this study examined the ones that presented with the lowest variation between replicates. A cut off SD of  $\pm 10\%$  was applied, resulting in 9 proteins that presented the lowest variation between all three replicates of RUNX1-ETO-expressing and control CD34<sup>+</sup> HSPC. These proteins showed no significant difference in total protein expression between RUNX1-ETO and control CD34<sup>+</sup> HSPC (**Figure 3.21A**), indicating that, even though the same protein is being expressed at similar levels, the sub-cellular localisation was different. Of these, 5 proteins were possibly translocated to the cytoplasm as a result of RUNX1-ETO expression, whilst 4 presented a higher prevalence in the nucleus of the same cells, when compared to control (**Figure 3.21B-C**). Mechanisms underlying these changes will be discussed in **3.4**.

Most of the proteins possibly mis-localised because of RUNX1-ETO expression were identified as enzymes and binding proteins; however, no abnormal TF protein shifts were observed in this analysis (**Table 3.5**). Nevertheless, these might be involved in protein synthesis and post-translational modification processes. The fact that these proteins were shown to be translocated into a different subcellular compartment, as a result of RUNX1-ETO expression, could indicate that their expression is altered, and promoting the expression of genes involved in the leukaemogenic process.



**Figure 3.20 – Proteins are mis-localised to the nucleus or cytoplasm as a consequence of RUNX1-ETO expression, in CD34<sup>+</sup> HSPC**

For each detected protein, the relative percentage of cytosolic and nuclear protein was quantified based on the total amount of protein, both in control and RUNX1-ETO expressing CD34<sup>+</sup> HSPC. To determine the difference in protein expression between RUNX1-ETO and control cells in the cytoplasmic compartment, the percentage of cytosolic proteins present in RUNX1-ETO cells was subtracted from the equivalent in control cells; the same approach was used to analyse nuclear protein localisation. **(A)** Protein expression was analysed for 926 molecules, and differences in RUNX1-ETO-expressing compared to control cells calculated for each one; lines denote a 15% cut-off, above which there is a possible translocation of at least 15% of the total amount of protein to the cytoplasm of RUNX1-ETO cells (Barretina *et al.*), whereas proteins identified below the line indicate an increase of protein expression in the nucleus of RUNX1-ETO cells, as compared to control. Each dot indicates mean (n=3). **(B)** Proteins with a shift of at least 15% were plotted in an independent graph; as shown in (A), red proteins indicate a protein shift to the cytoplasm in RUNX1-ETO-expressing cells, while blue proteins show a higher protein expression in the nucleus of these cells, compared to normal CD34<sup>+</sup> HSPC. Data indicates mean  $\pm$  1SD (n=3).



**Figure 3.21 – RUNX1-ETO leads to a shift in localisation of 9 proteins**

Protein expression profiles were obtained by performing SWATH-MS analysis on CD34<sup>+</sup> HSPC transfected with either a control or a RUNX1-ETO expressing vector. Cytosolic and nuclear proteins were extracted on day 3 of blood cell development. (A) To calculate total protein expression, cytosolic and nuclear protein expression values were combined, and expression analysed between RUNX1-ETO and control cells, showing no significant differences between the two conditions. However, when analysing each compartment, it was possible to see that (B) 5 proteins presented a higher prevalence in the cytoplasm of RUNX1-ETO cells, as compared to control, indicating a possible shift in normal protein localisation. Conversely, (C) 4 proteins were shown to be more expressed in the nucleus of RUNX1-ETO cells, as opposed to control. Dots represent each one of the three biological replicates. Line indicates Mean  $\pm$  1SD (n=3).

**Table 3.5 – Identification of ‘shifted’ proteins as a result of RUNX1-ETO expression**

§ mRNA expression (log<sub>2</sub>) was compared between RUNX1-ETO and control CD34<sup>+</sup> cells by microarray analysis, performed on day 3 of blood cell development (adapted from (Tonks *et al.*, 2007)).

‡ Protein expression profiles were obtained by performing SWATH-MS analysis on CD34<sup>+</sup> HSPC transfected with either a control or a RUNX1-ETO expressing vector. Cytosolic and nuclear proteins were extracted on day 3 of blood cell development.

(↑) or (↓) denotes upregulation or downregulation of gene/protein in RUNX1-ETO cells, respectively, as compared to control HSPC.

Protein	UniProtID	Classification	Function	mRNA expression in CD34 <sup>+</sup> HSPC §	Protein expression in CD34 <sup>+</sup> HSPC ‡
<b>RIC8B</b>	Q9NVN3	GEF	Intracellular signalling pathways	Not detected	No change
<b>MRPL15</b>	Q9P015	Mitochondrial protein	Protein synthesis in the mitochondria	Not detected	No change
<b>XPNPEP1</b>	Q9NQW7	Enzyme	Maturation/degradation of peptides	No change	↓
<b>RRM1</b>	P23921	Enzyme	Cell proliferation/migration, tumour/metastasis development	No change	↑
<b>CHMP1A</b>	Q9HD42	Binding protein	Multivesicular bodies formation and sorting	No change	↓
<b>MANF</b>	P55145	Binding protein	Protects cells against ER-stress induced damage	No change	↓
<b>NMT1</b>	P30419	Enzyme	Post-translational modifications	No change	No change
<b>UBAP2L</b>	Q14157	Binding protein	Ubiquitination	No change	↓
<b>PLAA</b>	Q9Y263	Binding protein	Ubiquitination	No change	↓

GEF – guanine nucleotide exchange factor; ER – endoplasmic reticulum

In summary, this analysis demonstrated that some proteins might be mis-localised and a shifted to a different subcellular compartment, arising from the expression of RUNX1-ETO. However, the number of proteins identified was small, and the changes in localisation were not significant. Even though it is unlikely that these are a major mechanism of action by RUNX1-ETO, it is possible that these genes/proteins contribute to the t(8;21) phenotype.

In conclusion, this chapter was identified two targets of interest for further study and possible mediators of the RUNX1-ETO phenotype. The roles of ZNF217 and C/EBP $\beta$  will be explored in subsequent **Chapter 4** and **Chapter 5**.

### 3.4 Discussion

Expression of RUNX1-ETO leads to widespread disruption in gene and protein function; however, the exact mechanisms through which this fusion protein is able to promote leukaemogenesis remains unclear. Even though this AML subtype presents a favourable outcome, additional target discovery and a deeper understanding of the underlying processes that lead to the development of the disease are necessary. This chapter's main aim was to identify changes in gene transcription and/or protein expression as a consequence of RUNX1-ETO expression in human CD34<sup>+</sup> HSPC. The expression of this fusion protein results in a block in normal cell differentiation; since this is a transcriptionally regulated process, I specifically focused on TF dysregulation. Transcriptome analysis identified ZNF217, also confirmed by SWATH-MS, as a potential target of interest and possible mediator of the block in differentiation observed in RUNX1-ETO expressing cells. Given that mRNA alone does not reliably predict changes in protein expression, quantitative proteomics was also performed in parallel. SWATH-MS allowed the quantitation of changes in protein expression in the cytosol and nucleus of RUNX1-ETO expressing cells. This approach identified C/EBP $\beta$  as a further target of interest in the context of t(8;21).

#### 3.4.1 Transcriptional dysregulation imposed by RUNX1-ETO expression

Previously, microarray studies performed by Tonks *et al.* identified 380 dysregulated genes as a consequence of RUNX1-ETO expression in human CD34<sup>+</sup> HSPC (Tonks *et al.*, 2007). However, this analysis was performed using an unsupervised approach, focusing exclusively on the highest fold changing genes, and not on global changes to pathways. This current study was able to improve upon the previous analysis and identify further transcriptomic changes, by

analysing changes within pathway contexts and focusing specifically on TFs. The analysis used two online platforms: Metacore™ and IPA®. One of Metacore's™ most useful tools is the Enrichment Analysis workflow, which calculates enriched p-values in different types of gene sets within the uploaded dataset. These gene sets originate from curated pathways, networks or related genes derived primarily from literature evaluation and from the GO lexicon (Cirillo *et al.*, 2017). An 'Enrichment Analysis' performed on the 380 differentially expressed genes identified them as being implicated in multiple and complex cellular pathways, including those regulating cell proliferation. In addition, this study identified several differentially expressed pathways involved in the progression of AML. These included changes in the  $\beta$ -catenin pathway, which have previously been described in AML t(8;21) - several studies have shown that  $\beta$ -catenin is essential for the clonogenic growth of RUNX1-ETO-expressing HSPC and leukaemic cells (Zhang *et al.*, 2013; Ysebaert *et al.*, 2006). Furthermore,  $\beta$ -catenin has been described as playing a key role in the development of AML (Wang *et al.*, 2010b), and cooperates with several chromosomal aberrations driving this disease (Müller-Tidow *et al.*, 2004). Dysregulated pathways also included the NF- $\kappa$ B pathway, previously described in AML to be associated with disrupted apoptosis in several cell lines (Testa and Riccioni, 2007). Moreover, constitutive activation of NF- $\kappa$ B has been observed in AML CD34<sup>+</sup> HSPC, contrasting to normal bone marrow cells (Birkenkamp *et al.*, 2004; Braun *et al.*, 2006; Guzman *et al.*, 2001), leading to changes in tumour formation and maintenance (Bassères and Baldwin, 2006). Lastly, pathway analysis identified significant changes in the normal transcriptional regulation of granulocytic development as reported by Tonks *et al.* in CD34<sup>+</sup> HSPC (Tonks *et al.*, 2004).

In order to determine the most significantly represented gene/protein classes disrupted due to RUNX1-ETO expression, this study used the Enrichment Analysis in Metacore™. Using the z-score (a bidirectional measure of a difference from the mean of involved proteins within a pathway) (Shi *et al.*, 2010), this analysis demonstrated that TFs were found to be the most represented gene class to be dysregulated as a result of RUNX1-ETO expression, following 'Others'. The latter was not considered to be of interest as it is too generic and does not represent a specific gene class. The reason the 'Others' category was found to be the most significant, relies on the fact that Metacore™ was unable to classify 221 genes, which represent 55% of the dataset. As z-scores can be negative, since it's the absolute value that indicates the significance of the analysis, the greatest deviant z-scores from the mean correspond to the most meaningful results. For this reason, this analysis alone can lead to misleading results, as the

‘Others’ category will contain genes from different classes, hence a more detailed classification of genes is often necessary.

Having determined that TFs were one of the most represented gene class dysregulated as a consequence of RUNX1-ETO expression, this study aimed to identify specific TF whose activity could be enhanced due to the presence of RUNX1-ETO. Similarly to Metacore™, IPA® allows the identification of key transcriptional regulators according to overall gene expression, based on the Ingenuity Knowledge Base. To generate a more robust analysis, considering that Metacore™ and IPA® are established on different databases, this study combined both approaches, thus generating a 35-gene list. Metacore™ was able to identify 25 overconnected TF, whilst IPA® identified 31; 22 TF were shown to be identified by both approaches. The fact that some TF were only identified by one of the programs relies on the use of different databases within the programmes.

Among the TF identified were *PU.1*, a known master regulator required for the development of mature myeloid and lymphoid cells, and *CEBPA*, a gene that has been demonstrated to play a role in normal granulocytic development (Pang *et al.*, 2018; Tenen, 2001). Both of these TFs have been extensively studied in AML, and validate the approach undertaken here to identify overconnected TF. Importantly, this approach identified other TFs not yet described as having a leukaemogenic role in AML. These included *ZBTB20*, which normally acts as a transcriptional repressor and is generally associated with B-cell lymphoma, not AML (Peterson *et al.*, 2019). Another example is *SMARCA1*, expressed in a broad range of normal tissues, and reported to modulate the Wnt/ $\beta$ -catenin pathway. Recently, it has been shown that *SMARCA1* contributes to the pathogenesis of leiomyosarcomas, possible due to epigenetic modifications (Patil *et al.*, 2018), as well as to that of gastric carcinomas (Takeshima *et al.*, 2015). However, its role in the development of AML t(8;21) has not been described to date. Taken together, this analysis identified the differential expression of genes whose activities could possibly be modulated and consequently the expression of genes involved in cell differentiation and proliferation.

Following identification of TF of interest, the TCGA (Ley *et al.*, 2013) and MILE (Kohlmann *et al.*, 2008) studies were analysed to determine the mRNA expression of each target in patients diagnosed with AML t(8;21), as compared to normal HSCs, in an attempt to refine the 35-gene list. Genes that presented contradictory patterns of expression, shown to be

upregulated in our model but downregulated in t(8;21) patients, or vice-versa, were excluded from further analysis. Furthermore, technical aspects were also considered. These included sequence size, as overexpression studies are more difficult to perform using genes with a bigger CDS, as well as its clinical significance. Using this approach, I identified four potential targets that might be implicated in the development of the phenotype observed in RUNX1-ETO expressing cells – *IRF7*, *IRF9*, *ARID5B* and *ZNF217*. Since mRNA levels are not powerful predictors of corresponding protein expression, this was assessed through western blotting. To further validate protein expression analysis, SWATH-MS results were analysed for each of targets (see below).

This approach identified two members of the Interferon Regulatory Factor (IRF) family, *IRF7* and *IRF9*, implicated in mediating type I interferon (IFN) response. *IRF7* has been shown to be involved in antibacterial and antiviral innate immunity. In the CD34<sup>+</sup> HSPC overexpressing RUNX1-ETO, *IRF7* was shown to be upregulated at the mRNA level, as well as in t(8;21) AML patient samples. The IRF7 protein was equally found to be upregulated in RUNX1-ETO cells, as compared to control, verified by western blotting. However, this protein was not detected in the SWATH-MS analysis, hence it was eliminated from further analysis.

*IRF9* is another member of IRF family, and has been shown to be upregulated in both CD34<sup>+</sup> HSPC and t(8;21) AML patient samples. Similarly to IRF7, IRF9 is a TF that mediates the type I IFN response by regulating downstream expression of interferon-stimulated genes, namely STAT1 and 2, with which it forms a complex that shuttles to the nucleus to allow IRF9 to act as a TF (Tsuno *et al.*, 2009). IRF9 has also been shown to be involved in regulating cell proliferation (Weihua *et al.*, 1997) and tumour formation (Luker *et al.*, 2001). High IRF9 expression has been shown to display anti-proliferative effect in ovarian adenocarcinoma (Provance and Lewis-Wambi, 2019); the contrary effect has been proven in cells where the expression of IRF9 is absent or low, in prostate cancer (Erb *et al.*, 2013). In breast cancer, IRF9 upregulation has shown to confer resistance to the chemotherapeutic agent paclitaxel (Conze *et al.*, 2001). Even though its role in breast cancer has been well established, whether IRF9 plays an important part in the pathogenesis of AML is still unclear. I found it to be upregulated both at the mRNA and protein level, validated through western blotting and SWATH-MS.

*ARID5B*, or AT-rich domain 5B, is a chromatin-modifier transcriptional regulator present in chromosome 10 found to be significantly upregulated in CD34<sup>+</sup> HSPC expressing RUNX1-ETO, by 4-fold. Expression of the ARID5B protein, however, showed no difference in

RUNX1-ETO cells, as compared to control HSPC, contrasting with SWATH-MS analysis where ARID5B was shown to be detected in cells expressing the fusion protein RUNX1-ETO, whilst no protein was detected in control samples. Under normal circumstances, the ARID5B protein recognizes the core DNA sequence motif AAT(C/T), influencing cell growth and differentiation of B-lymphocyte progenitors. (Whitson *et al.*, 1999; Lahoud *et al.*, 2001). ARID5B knockout mice exhibit abnormalities in B-lymphocyte development and ARID5B mRNA expression has been shown to be upregulated in other haematological malignancies such as acute promyelocytic leukaemia and acute megakaryoblast leukaemia (Lahoud *et al.*, 2001; Paulsson *et al.*, 2010; Bourquin *et al.*, 2006). In addition, increasing evidence has shown that mutations and single nucleotide polymorphisms of the ARID5B gene are associated with the development of acute lymphoblastic leukaemia, influencing treatment outcome, and may also be associated with paediatric AML. (Lin *et al.*, 2014; Emerenciano *et al.*, 2014)

Lastly, ZNF217 is an oncogenic protein that has numerous functions in various human cancers (discussed below), especially considering that it is located in a chromosomal region frequently amplified in human cancers, in 20q13 (Collins *et al.*, 1998; Tabach *et al.*, 2011). This study confirmed that ZNF217 was upregulated in CD34<sup>+</sup> HSPC expressing RUNX1-ETO, compared to control. Upregulation was confirmed at the protein level by western blotting and SWATH-MS. ZNF217 possesses multiple roles, such as binding to specific DNA sequences to regulate target gene expression, is a component of the human histone deacetylase complex CoREST-HDAC and can be found in complexes with other transcriptional co-repressors including the C-terminal binding protein and the methyltransferase G9a (Cowger *et al.*, 2007; Krig *et al.*, 2007; Thillainadesan *et al.*, 2012; Thillainadesan *et al.*, 2008). Even though ZNF217 mainly acts in transcriptional repressor complexes, it positively regulates the expression of specific target genes, making it a double-faced transcriptional regulator (Cohen *et al.*, 2015). Several studies indicate that ZNF217 interferes with multiple intracellular signalling networks to reprogram cancer cells, by altering cell cycle, cell growth, and disrupting anti-proliferative signalling (Cohen *et al.*, 2015). Moreover, ZNF217 has been shown to interfere with normal apoptotic pathways at early stages of tumour progression and conferring resistance to chemotherapy in breast cancer cells (Thollet *et al.*, 2010; Huang *et al.*, 2005). Overexpression of ZNF217 in breast and ovarian cancer promoted cell proliferation, whereas the opposite effect was shown when silencing it in prostate, colorectal, ovarian and breast cancer (Li *et al.*, 2014; Thollet *et al.*, 2010; Szczyrba *et al.*, 2013; Li *et al.*, 2015). In primary cells, ZNF217 has been shown to interfere with the expression of genes involved in differentiation and organ

development, while at the same time dysregulating genes involved in the repression of cell differentiation and maintenance of cancer stem cells (Krig *et al.*, 2007; Vendrell *et al.*, 2012; Littlepage *et al.*, 2012). In primary mammary cells, overexpression of ZNF217 increased the formation of cells displaying self-renewal ability, an important feature observed in leukemic stem cells (Vendrell *et al.*, 2012; Littlepage *et al.*, 2012).

This study used both western blotting and SWATH-MS to measure protein expression in CD34<sup>+</sup> HSPC; however, results were not concordant, with results being different between the two approaches. Even though both approaches were able to identify an increased or detectable expression of three targets in CD34<sup>+</sup> HSPC (IRF9, ARID5B and ZNF217), levels of regulation varied significantly; of these, only ZNF217 had a similar pattern of protein expression. Even though its level of upregulation wasn't the same (10-fold upregulation in western blot vs. 3-fold upregulation detected through SWATH-MS, in RUNX1-ETO cells compared to control), the trend was the same. The inconsistencies observed between the two techniques may rely on the fact that western blot is typically considered a semi-quantitative technique, as it provides a relative comparison of protein levels, usually to a loading control, and not an absolute measure of quantity. Moreover, it could be related to the quality of the antibodies used. This poses the main advantage in using SWATH-MS instead of western blotting, as it provides a level of specificity unavailable when performing western blot. Furthermore, SWATH-MS allows the detection of thousands of proteins in a single run, maximizing the information obtained per sample, while significantly decreasing the time used for each experiment (Jayasena *et al.*, 2016). The main disadvantage of SWATH-MS is cost, which can constrain the use of MS-based techniques.

In summary, ZNF217 was the only target that consistently showed to be overexpressed, both at the mRNA and protein level. Furthermore, based on literature research, overexpression of ZNF217 has been linked to several hallmarks of cancer, but its dysregulation has not been explored in the context of haematopoiesis and AML. This study will, therefore, address this in **Chapter 4**.

### **3.4.2 Dysregulation of the proteome arising from RUNX1-ETO expression**

To date, studies performed to identify putative target genes in RUNX1-ETO expressing cells have mainly focused on analysing the cells transcriptomic profiles (Nafria *et al.*, 2020; Ptasinska *et al.*, 2019; Martinez-Soria *et al.*, 2018; Ptasinska *et al.*, 2012). However, proteomic

profile analysis presents a more realistic picture of cell physiology, as compared to genomic studies (Singh and Sharma, 2020).

Over the past few years, several studies have been performed using MS techniques in an attempt to elucidate the mechanisms underlying RUNX1-ETO-mediated AML. Schoenherr *et al.* applied MS analysis to the Kasumi-1 cell line, which expresses the fusion protein RUNX1-ETO, to analyse the incidence of proteolytic cleavage of cathepsin G and neutrophil elastase and to identify contributing proteases on a proteome-wide scale (Schoenherr *et al.*, 2019). Previously, gene expression studies performed by Lo *et al.* allowed the identification of *RASSF2* as a potential target gene shown to be downregulated in RUNX1-ETO leukaemic blasts, compared to control (Lo *et al.*, 2012). Following up on this study, Stoner *et al.* used MS analysis to identify the *RASSF2*- proximal proteome in the Kasumi-1 cell line, revealing its association with Rac GTPase-related proteins (Stoner *et al.*, 2020). Lastly, a study performed by Singh *et al.* aimed at using an MS-based approach to identify changes in protein expression using the U937 cell line, with inducible RUNX1-ETO expression. The authors showed that RUNX1-ETO leads to drastic change in the cell's proteomic profile, and focused on the increased expression of NM23 (Singh *et al.*, 2010). The authors showed that NM23 expression is normally suppressed by C/EBP proteins; however, upregulation of RUNX1-ETO blocks this effect, leading to increased levels of NM23, which in turn promoted a block in differentiation and increased proliferation (Singh *et al.*, 2010). Even though these studies contributed to a further understating of the mechanism through which RUNX1-ETO leads to the development of AML, their main focus relied on AML cell lines, and not on primary cell material.

For this study, CD34<sup>+</sup> HSPC were transduced with a control or a RUNX1-ETO-overexpressing vector, following which cytosolic and nuclear proteins were extracted and quantified. Before proceeding, extracts were validated for fraction purity, to ensure little or no cross contamination between cytosolic and nuclear fractions, as well as for RUNX1 expression, an indicative of RUNX1-ETO overexpression.

In order to detect all the proteins within our samples, the Pan Human Library was used, which allows the detection and quantification of 50.9% of all human proteins annotated by UniProtKB/Swiss-Prot, corresponding to approximately 13,000 proteins (Rosenberger *et al.*, 2014). Within this study, SWATH-MS analysis detected over 6,000 proteins in the model system used. From these, 4,635 were found to be detected in all three replicates in cytoplasm/nucleus of control and RUNX1-ETO cells, making them suitable for statistical

analysis. The remaining detected proteins were missing expression values, making statistical analysis impossible. This can be attributed to two factors: failure of detection or true protein absence. It is inappropriate to assume that because a protein was not detected by SWATH-MS, is it not being expressed in a cell at the time of analysis. The reasons for samples ‘missing’ some, but not all, protein expression values, can be technical or biological. Technical, and stochastic, reasons can include instrument variation, protein expression levels too close to the noise level, or pre-analytical variation in sample preparation (Collins *et al.*, 2017).

Nevertheless, this study statistically analysed 4,635 proteins, of which 2,787 were detected in the cytoplasm of CD34<sup>+</sup> HSPC and 1,848 were present in the nucleus. An exploratory analysis was performed using PCA plots and heat maps, comparing general protein expression between control and RUNX1-ETO-expressing cells. These showed a certain degree of clustering, more noticeable in the cytosolic proteins, as compared to the nuclear samples, in RUNX1-ETO-transduced cells. Variations in sample clustering can be attributed to different factors. Even though cord blood is an invaluable source of material for the study of blood-related diseases, it is a source of variability. Furthermore, cells originated from different donors have been shown to grow at different rates, depending on a number of donor-related factors (Siegel *et al.*, 2013; Kang *et al.*, 2018; Yap *et al.*, 2000). Moreover, these differences can also be attributed to background noise, due to low expression of certain proteins, thus resulting in very low signal. Nonetheless, of the 4,635 proteins, 257 were found to be statistically dysregulated ( $p < 0.05$ ) because of RUNX1-ETO expression; 183 were localised to the cytoplasm and 74 to the nucleus. This analysis was performed in an exploratory manner and without performing statistical corrections, as the number of samples ( $n=3$ ) is insufficient to confer corrected statistical significance. Downstream analysis was performed on the 257 differentially expressed proteins with a specific focus on TF changes, as previously performed for the transcriptomic dataset. Once more, the first step of analysis consisted of performing an Enrichment Analysis using Metacore™ (**Figure 3.16**). This analysis identified the significantly dysregulated proteins as being implicated in several cellular processes, including ‘Myeloid differentiation’ (Nafria *et al.*, 2020; Ichikawa *et al.*, 2013), ‘NF- $\kappa$ B pathway’ (Zhou *et al.*, 2015), ‘Anti-apoptotic’ (Goyama *et al.*, 2013) and ‘Neutrophil chemotaxis’ pathways (Tonks *et al.*, 2004; Tijchon *et al.*, 2019), known to be implicated in AML t(8;21) pathophysiology. Analysis of changes in the transcriptome and proteome of CD34<sup>+</sup> HSPC, as a consequence of RUNX1-ETO expression, yielded similar results, as cell differentiation and apoptosis were highly represented in both datasets.

Following pathway analysis, this study analysed and clustered the 257 significantly dysregulated proteins previously described. Comparably to the analysis performed above (3.4.1), TFs were one of the classes with the highest representation within the dataset. Subsequently, this study focused on TF changes arising from RUNX1-ETO expression, by analysing their overconnectivity with other proteins from the dataset, leading to the identification of 4 overconnected TF: CBF $\beta$ , C/EBP $\beta$ , PU.1 and RUNX1. As expected, RUNX1 was identified in this analysis, potentially as a result of RUNX1-ETO expression, known to influence normal gene expression. Subsequently, the algorithm also identified CBF $\beta$ , a TF regulating the transcription of several haematopoietic-specific targets, including myeloperoxidase, CSF1 receptor, IL-1 and GM-CSF (Downing, 2001). Upon binding to RUNX1, this complex is responsible for not only the transcriptional activation in early HSC, but also expression of growth factors and proliferation and survival regulators. Similarly to the analysis performed above (3.4.1), PU.1 was once more identified as one of the most overconnected TF within the dataset.

Lastly, C/EBP $\beta$  was identified as one of the TFs with the highest influence on other dysregulated proteins, as identified in 3.3.2.7. Another member of the CEBP family, this TF has been shown to play an important role in macrophage development (Lekstrom-Himes and Xanthopoulos, 1998; Poli, 1998; Ramji and Foka, 2002). The process through which C/EBP $\beta$  is activated is complex, since it is regulated upon different levels, including transcriptionally, translationally and post-translationally. Firstly, the *CEBPB* gene does not possess any introns, and its mRNA can originate three distinct protein isoforms based on three alternative AUG translation initiation sites (Descombes and Schibler, 1991; Timchenko *et al.*, 1999). C/EBP $\beta$  LAP\* (also named full length, or C/EBP $\beta$ -1), C/EBP $\beta$  LAP (C/EBP $\beta$ -2) and C/EBP $\beta$  LIP (C/EBP $\beta$ -3). The latter is thought to function as a transcriptional repressor, since it lacks the transcription activating domain but still processes the ability to dimerize and bind to DNA (Descombes and Schibler, 1991). Alternatively, CEBP proteins can be regulated through protein-protein interactions with proteins important for cell proliferation, including Cdk2, Cdk4 or the Rb protein (Ji and Studzinski, 2004; Charles *et al.*, 2001; Wang *et al.*, 2001). Lastly, C/EBP $\beta$  has several phosphorylation sites, which are important for intracellular localisation and transcriptional activity (Metz and Ziff, 1991; Chinery *et al.*, 1997; Piwien Pilipuk *et al.*, 2003; Berg *et al.*, 2005).

To study the role of C/EBP $\beta$  in differentiation, several studies have used HL-60 cells, an APL cell line, and 1,25D, a drug capable of promoting the differentiation of these cells. These groups noted an increase in C/EBP $\beta$  expression once the drug was added to the cells, contributing to the cell's differentiation into the monocytic lineage, but no difference was observed in the expression levels of other members of the CEBP family, including C/EBP $\alpha$  and C/EBP $\epsilon$  (Ji and Studzinski, 2004; Studzinski *et al.*, 2005; Marcinkowska *et al.*, 2006). Furthermore, the LAP and LIP isoforms of C/EBP $\beta$  seem to increase more rapidly in forced-differentiating HL-60, as compared to control HL-60 cells, suggesting that these isoforms have dominant roles over LAP\*. Moreover, even though LIP functions as a transcriptional repressor, the authors suggest that C/EBP $\beta$  LIP plays a role in the repression of genes unnecessary in differentiating cells, such as those involved in cell proliferation (Marcinkowska *et al.*, 2006). The same trend was observed by Pham *et al.* in peripheral blood monocytes (Pham *et al.*, 2007). The authors showed that human macrophages contain high constitutive levels of functional C/EBP $\beta$ , unlike peripheral monocytes. In addition, PMA-treated THP-1 cells were shown to induce *CEBPB* expression; however, even though the authors observed that these transitions were not associated with major changes in *CEBPB* mRNA levels, which remained stable throughout the differentiation process, there was an increase in C/EBP $\beta$  protein levels, particularly regarding the LAP and LAP\* isoforms. In this study, both C/EBP $\beta$  mRNA and protein were downregulated in cells expressing RUNX1-ETO suggesting that this fusion protein is inhibiting *CEBPB* at the transcriptional level. Currently there is no information whether this change in C/EBP $\beta$  expression could impact growth and development of normal haematopoietic cells, this will therefore be addressed in **Chapter 5**.

#### 3.4.2.1 Alternative analysis of RUNX1-ETO-induced proteomic changes in CD34<sup>+</sup> HSPC

An alternative approach analysed proteins exclusively expressed in RUNX1-ETO cells, but undetectable in control samples. These were not identified above as it was not possible to perform a statistical analysis on undetected samples. This study found that 53 proteins were found to be exclusively expressed in RUNX1-ETO cells, from which 29 were found exclusively in the cytoplasm and 24 in the nucleus.

Of these, this study identified 4 TFs found to be expressed in the cytoplasm, including ATXN7, ZNF395, REL and, interestingly, IRF9. The *IRF9* transcript had previously been identified as a target of interest (3.3.1.2) as it had been identified as one of the most overconnected genes within the mRNA dataset. However, western blotting analysis detected

nuclear levels of IRF9 protein, even if no nuclear protein was detected by SWATH-MS. ATXN7, is a member of the STP3/TAF9/GCN5 histone acetyltransferase (STAGA) complex. Even though some additional studies have been performed in colorectal and clear cell renal carcinomas, the functional consequences arising from mutations in this gene have not been determined yet (Gotoh *et al.*, 2014; Kalvala *et al.*, 2016). Gene expression studies have found that the expression of *ZNF395* is significantly increased in several cancer types, including renal cell carcinoma, osteosarcoma, and Ewing sarcoma (Skubitz *et al.*, 2006; Tsukahara *et al.*, 2004; Dalgin *et al.*, 2007). REL is a proto-oncogene known to play a role in cell differentiation by being part of the NF- $\kappa$ B complex, and its role has already been extensively studied in the context of AML (Guzman *et al.*, 2001; Nakagawa *et al.*, 2011; Zhou *et al.*, 2015). However, since these TFs were identified in the cytoplasm, it is unlikely that they possess any transcriptional roles. Furthermore, additional western blotting analysis would be necessary to assess nuclear protein expression.

Moreover, one TF was identified in the nucleus. MIER1 is a transcriptional regulator differentially expressed in breast carcinoma cell lines and tumours (Paterno *et al.*, 1998; McCarthy *et al.*, 2008). However, these targets were not pursued in this current study, as this posed as an alternative way of analysing the proteomic changes induced by RUNX1-ETO expression.

The last approach analysed proteins that were detected in both the cytoplasm and nucleus of control and RUNX1-ETO CD34<sup>+</sup> HSPC, by analysing relative protein expression in each compartment, regarding to total protein expression, to identify mis-localised proteins. These shifts in proportion of expression at each subcellular localisation may help identify proteins differently localised between RUNX1-ETO and control cells. Previously, several proteins have been shown to promote cancer development due to their abnormal localisation within a cell such as EGFR and  $\beta$ -catenin, normally localised in the plasma membrane, that upon translocation to the nucleus leads to the development of pancreatic cancer (Cohen, 1983). Therefore, this study analysed the expression levels of 926 proteins, and looked at their relative sub-cellular localisation. Of these, 33 proteins were shown to be localised differently in RUNX1-ETO cells, as compared to control, by at least 15% of relative protein expression, indicating that, in RUNX1-ETO cells, there was an accumulation of certain proteins in the cytoplasm or nucleus, as compared to the same compartment of control cells. This cut-off was based on potential validation experiments through western blotting techniques, since a smaller

difference might be harder to detect. Additionally, due to the high variability arising from the source of cord blood material and viral preps, an additional cut off SD of  $\pm 10\%$  was applied ( $n=3$ ). This led to the identification of 9 proteins potentially translocated/retained as a consequence of RUNX1-ETO expression. Of these, 5 proteins showed a higher expression in the cytoplasm of RUNX1-ETO cells, possible indicating protein retention in this compartment, as compared to control cells; and 4 were identified as being more abundant in the nucleus of RUNX1-ETO cells. However, most of these proteins were identified as enzymes and binding proteins, as no TFs were identified using this approach. For this reason, proteins were not examined further. Even so, mis-localisation of these proteins can lead to TF significant dysregulation and, therefore, impaired normal cellular process and leukaemogenic development, as a result of RUNX1-ETO expression.

It is also important to acknowledge that the current proteomics study represents an exploratory pilot programme and, as such, the small sample size and inability to study the consequences of RUNX1-ETO in the cells' proteomic profile is recognised due to cost concerns. The sample size is sufficient to examine changes to inform further research and day 3 of culture was selected as the consequence of RUNX1-ETO is a differentiation block in early developmental stages. Secondly, due to the scarcity of umbilical cord blood donations as a result of the COVID-19 pandemic, it was not possible to validate these findings within the same CD34<sup>+</sup> context. However, owed to the proteomic approach, there is less concern regarding validation as mRNA-based approaches would suffer.

In conclusion, this chapter's main aim was to identify changes in TF expression arising from the expression of the fusion protein RUNX1-ETO, in CD34<sup>+</sup> HSPC. This aim was divided into two main objectives. The first one consisted of analysing previously generated microarray data, to identify changes at the mRNA level that could explain the dysregulation observed in these cells. This led to the identification of *ZNF217* as a target of interest. Its role in the haematopoietic process and in the development of AML will be explored in **Chapter 4**. However, since mRNA is not always an accurate predictor of protein expression, I also performed SWATH-MS analysis on the same cells, to study the cells' proteomic profile. This analysis identified C/EBP $\beta$ , a member of C/EBP family of TF, as a potential mediator of the leukaemogenic process of AML t(8;21). The effects of its dysregulation in haematopoiesis will be addressed in **Chapter 5**.

# Chapter 4

---

## **ZNF217 Promotes Normal Human Myeloid Development**

## 4.1 Introduction

Previously, I identified *ZNF217* as being overexpressed in RUNX1-ETO expressing cells compared to control (**Chapter 3**). *ZNF217* is a TF whose normal function is to regulate gene expression through binding to multiple regulatory complexes, including the human histone deacetylase complex (CoREST-HDAC) (Cowger *et al.*, 2007; Thillainadesan *et al.*, 2012; Thillainadesan *et al.*, 2008). However, this protein has a known oncogenic role in solid tumours, often attributed to the fact that its gene is located in a region frequently amplified in human cancers (20q13 chromosome) (Collins *et al.*, 1998; Tabach *et al.*, 2011). Increased copies of *ZNF217* have been reported in multiple tumours, including breast, colorectal and ovarian cancer (Plevova *et al.*, 2010; Fang *et al.*, 2010; Rahman *et al.*, 2012), and have been linked to a poorer outcome in colon and ovarian carcinomas (Peiró *et al.*, 2002; Rooney *et al.*, 2004).

In solid tumours, *ZNF217* has been shown to interfere with several cell processes including disruption of cell proliferation (Li *et al.*, 2014; Thollet *et al.*, 2010) through changes in TGF- $\beta$ -dependent anti-proliferative signalling (Massagué, 2008; Ikushima and Miyazono, 2010), cell cycle (Li *et al.*, 2014) and apoptosis (Huang *et al.*, 2005). For example, *ZNF217* overexpression was able to confer drug resistance, as well as promote cell proliferation *in vitro* and increase tumour growth *in vivo*, in a breast cancer model (Thollet *et al.*, 2010). Conversely, downregulation of *ZNF217* resulted in a decrease in cell proliferation, increased drug sensitivity, decreased colony formation efficiency and cell migration and invasion in breast, ovarian and colorectal cell lines (Sun *et al.*, 2008a; Thollet *et al.*, 2010; Li *et al.*, 2015). *ZNF217* is also able to bind to the promotor of genes involved in cell differentiation and organ development, resulting in an arrest in the cell differentiation process and maintenance of breast cancer stem cells (Krig *et al.*, 2007; Vendrell *et al.*, 2012; Littlepage *et al.*, 2012). In normal primary mammary epithelial cells, *ZNF217* overexpression led to an increase in the formation of mammospheres with an increased self-renewal potential and was associated with the repression of an adult stem cell expression signature (Littlepage *et al.*, 2012; Vendrell *et al.*, 2012; Nguyen, 2018). Moreover, Mao *et al.* described that *ZNF217* was upregulated in glioma stem cells (GSC) when compared to non-GSC. This study showed that forced differentiation of GSC led to a decrease in the levels of *ZNF217* and its knockdown inhibited their proliferation, while reducing stem-like populations (Mao *et al.*, 2011).

Altogether, these studies have established a link between the tumorigenic process and the overexpression of ZNF217 in solid tumours. However, whether ZNF217 has a role in perturbing haematopoietic growth and development has not been determined.

## 4.2 Hypothesis and Aims

This study hypothesises that ZNF217 expression plays an important role in haematopoiesis and that modulating its expression will disrupt normal human myeloid development, thus leading to AML initiation. This chapter's main objective is to understand the role of ZNF217 in haematopoietic development and its possible contribution to the development of AML, using a normal human primary HSPC model and AML cell lines. This will be achieved by performing the following aims:

### **Determine the expression of ZNF217 during normal human myeloid development and in AML patient blasts**

Analysis of publicly available transcriptomic datasets will be performed to determine ZNF217 mRNA expression in human HSPC subsets and across different AML subtypes.

### **Determine the effects of ZNF217 overexpression and knockdown on myeloid colony forming ability and self-renewal**

Limiting-dilution colony forming assays will be performed on transduced human CD34<sup>+</sup> HSPC and compared to control. Replating assays will be performed to assess the cells' self-renewal potential.

### **Determine the effect of ZNF217 overexpression or knockdown on normal human haematopoietic growth, differentiation, and development.**

CD34<sup>+</sup> HSPC will be transduced with a ZNF217-overexpression or -knockdown vector and grown in bulk liquid culture. Changes in monocytic and granulocytic growth will be assessed by analysing the cells' lineage and cell surface markers using multicolour flow cytometry.

### **To determine the effect of ZNF217 overexpression or knockdown on AML cell growth, proliferation and apoptosis.**

AML cell lines will be transduced with a ZNF217-overexpression or -knockdown vectors and grown in bulk liquid culture. Changes in cell growth will be evaluated by following the cells' proliferative ability over time. Furthermore, transduced cells will be assayed to determine if ZNF217 dysregulated expression results in changes in cell cycle and apoptotic frequency.

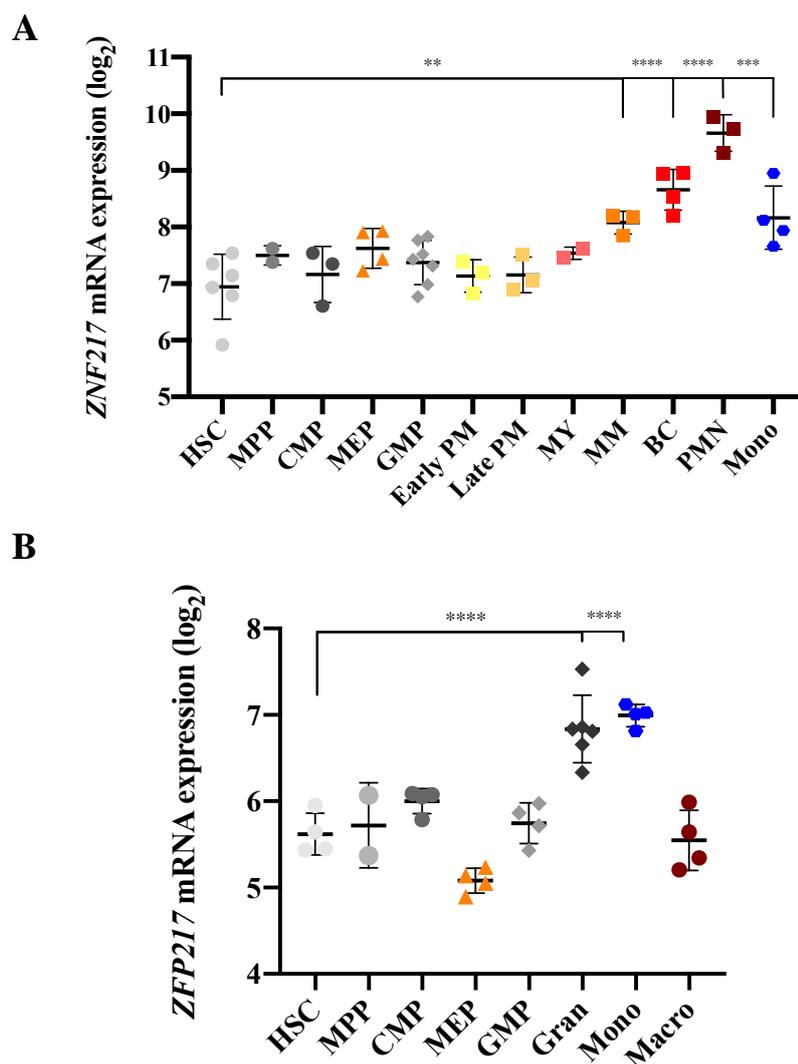
## 4.3 Results

### 4.3.1 Expression of *ZNF217* mRNA increases throughout myeloid cell development

In order to determine the expression of *ZNF217* throughout haematopoietic development, mRNA expression data for normal human haemopoietic cell populations was analysed. *ZNF217* mRNA expression significantly increases throughout granulocytic development, particularly in metamyelocytes, band cells and polymorphonuclear cells, as compared to immature HSC, by 2.2-, 3.3- and 6.6-fold, respectively (**Figure 4.1A**). Regarding monocytic differentiation, *ZNF217* mRNA levels significantly increase by 2.3-fold in mature monocytes, as compared to HSC, though levels were lower compared to granulocytes (**Figure 4.1A**). Analysis of the mouse homolog *Zfp217* also showed upregulation in mature granulocytes and monocytes, with a significant increase of 2.6-fold observed in both populations (**Figure 4.1B**). These observations suggest that *ZNF217* is transcriptionally regulated during blood cell differentiation, and differentially expressed according to specific cell types.

### 4.3.2 *ZNF217* is variably expressed across AML subtypes

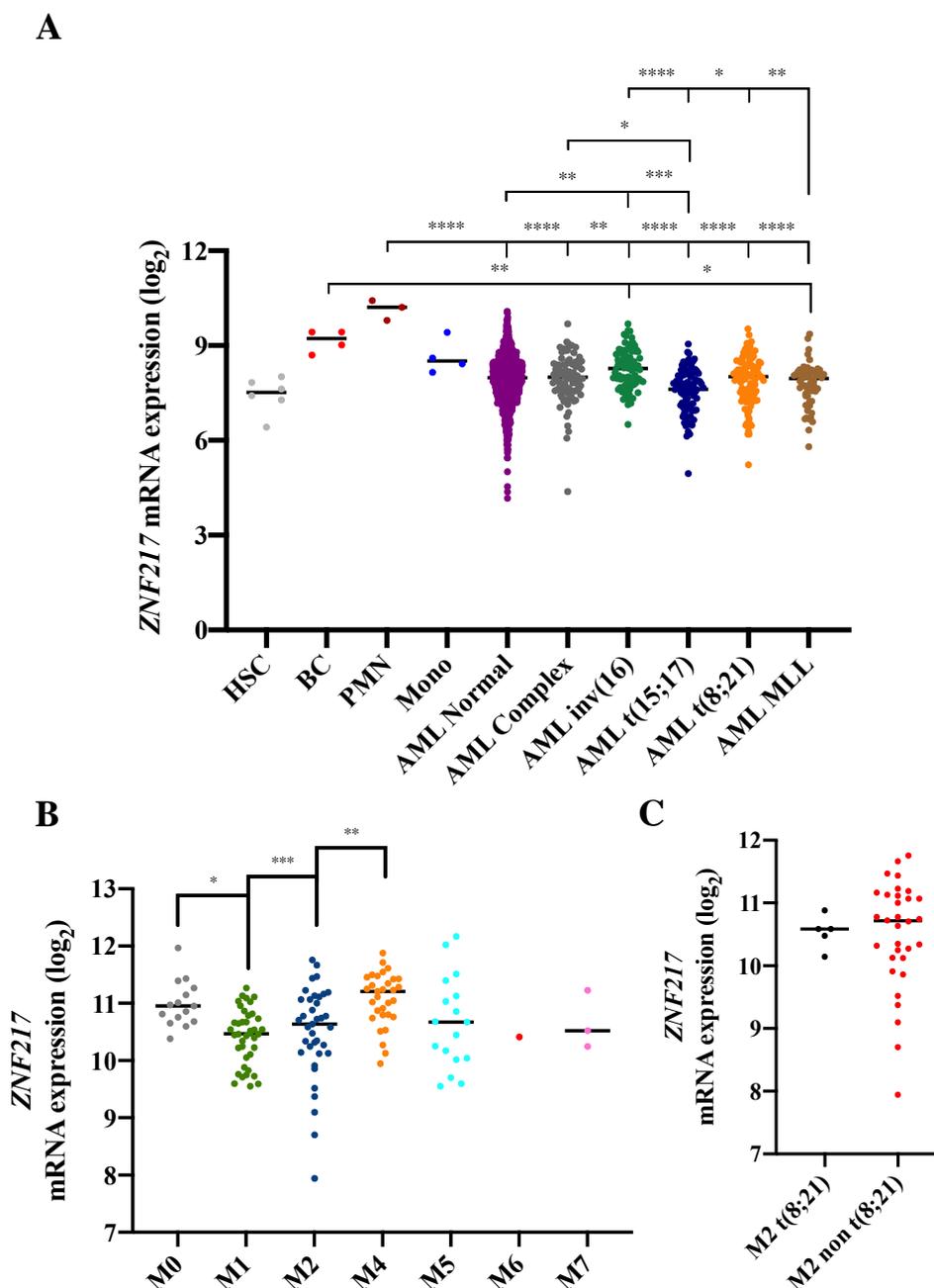
As previously introduced, AML is a highly heterogenous disease characterised by multiple molecular and cytogenetic abnormalities (**1.2.2**). To determine *ZNF217* mRNA expression in different subtypes of AML, publicly available transcriptomic datasets were analysed. Firstly, this study compared *ZNF217* mRNA expression in different subsets of AML with that of cell subsets at different stages of haematopoietic development. As **Figure 4.2A** shows, *ZNF217* expression is highly variable across AML. There were no significant differences in *ZNF217* mRNA expression between undifferentiated HSC and the multiple AML subtypes analysed. However, this expression was significantly lower in all AML subtypes as compared to polymorphonuclear cells, suggesting that expression of *ZNF217* might influence normal granulocytic development.



**Figure 4.1 – ZNF217 mRNA expression in normal haematopoiesis**

(A) Normalised microarray data (log<sub>2</sub>) showing ZNF217 mRNA expression in distinct human haematopoietic cells subsets based on cell surface marker expression. Normal human haematopoiesis data derived from GSE24759 (Rapin *et al.*, 2014; Svendsen *et al.*, 2016) Data indicates Mean ± 1SD. Statistical analysis was performed using one-way ANOVA with Bonferroni's multiple comparisons test; \*\* denotes p<0.01; \*\*\* denotes p<0.001; \*\*\*\* denotes p<0.0001, vs HSC (Probeset 203739\_at). (B) Normalised microarray data (log<sub>2</sub>) showing ZFP217 mRNA expression, ZNF217 mouse homolog, across different murine haematopoietic cell types. Mouse normal haematopoiesis data derived from GSE60101 (Lara-Astiaso *et al.*, 2014). Data indicates Mean ± 1SD. Statistical analysis was performed using one-way ANOVA with Bonferroni's multiple comparisons test; \*\* denotes p<0.01; \*\*\* denotes p<0.001; \*\*\*\* denotes p<0.0001, vs HSC (Probeset 1437414\_at)

**HSC** – Haematopoietic Stem Cell; **MPP** – Multipotential Progenitors; **CMP** – Common Myeloid Progenitor cell; **MEP** – Megakaryocyte / Erythroid Progenitor cell; **GMP** – Granulocyte / Monocyte Progenitors; **Early PM** - Early Promyelocyte; **Late PM** - Late Promyelocyte; **MY** - Myelocyte; **MM** – Metamyelocytes; **BC** – Band cell; **PNM** – Polymorphonuclear cell; **Mono** – Monocytes; **Gran** – Granulocytes; **Macro** – Macrophages.



**Figure 4.2 – *ZNF217* mRNA expression levels are variable across AML subtypes**

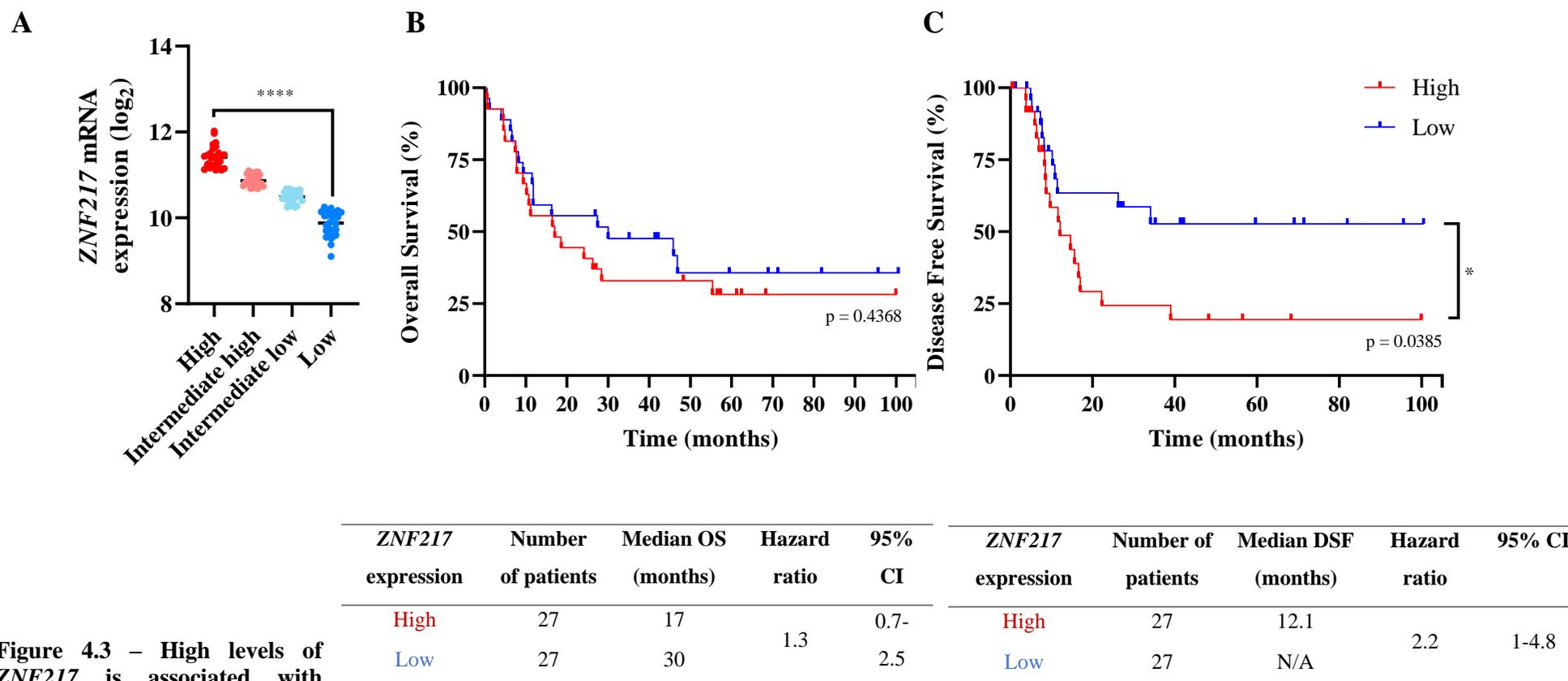
(A) Normalised microarray data ( $\log_2$ ) showing *ZNF217* mRNA expression in normal human haematopoietic developmental subsets vs. AML subtypes. Normal human haematopoiesis data derived from **GSE422519** (Rapin *et al.*, 2014; Svendsen *et al.*, 2016); human AML cells derived from **GSE13159** (Haferlach *et al.*, 2010; Kohlmann *et al.*, 2008), **GSE15434** (Klein *et al.*, 2009), **GSE61804** (Metzelder *et al.*, 2015), **GSE14468** (Wouters *et al.*, 2009; Taskesen *et al.*, 2015; Taskesen *et al.*, 2011) and **TCGA** (Ley *et al.*, 2013) (data analysed using Bloodspot's algorithm Bloodpool (Bagger *et al.*, 2016)). Normalised microarray data ( $\log_2$ ) showing *ZNF217* mRNA expression in (B) AML according to FAB subtype and (C) AML FAB2, categorised into t(8;21) and non-t(8;21), within the **TCGA** dataset (Ley *et al.*, 2013) All studies were analysed using the 203739\_at probeset. Data indicates mean  $\pm$  1SD. Significant differences were analysed by one-way ANOVA with Bonferroni's multiple comparisons correction; \* denotes  $p < 0.05$ ; \*\* denotes  $p < 0.01$ ; \*\*\* denotes  $p < 0.001$ ; \*\*\*\* denotes  $p < 0.0001$ . **AML normal** – AML with normal karyotype; **AML complex** – AML with complex karyotype

Interestingly, analysis of transcriptomic data from the TCGA dataset according to AML FAB subtype, shows that *ZNF217* expression, whilst highly variable, remains similarly expressed across different developmental AML subtypes (**Figure 4.2B**). Furthermore, no difference was observed in AML M2 patients expressing the fusion protein RUNX1-ETO (**Figure 4.2C**).

Subsequently, I analysed the expression of *ZNF217* in patients diagnosed with *de novo* AML using the TCGA 2013 RNAseq dataset (Ley *et al.*, 2013), and stratified them according to high and low quartiles, to represent high and low *ZNF217* expression, respectively (**Figure 4.3A**). For this analysis, patients with APL and those who did not receive standard chemotherapy treatment were excluded from the dataset. Having established these parameters, this study analysed the OS and disease-free survival of patients with high and low *ZNF217* expression. Additionally, the association between differential *ZNF217* expression and disease subtypes, molecular and cytogenetic abnormalities, and patient characteristics were assessed.

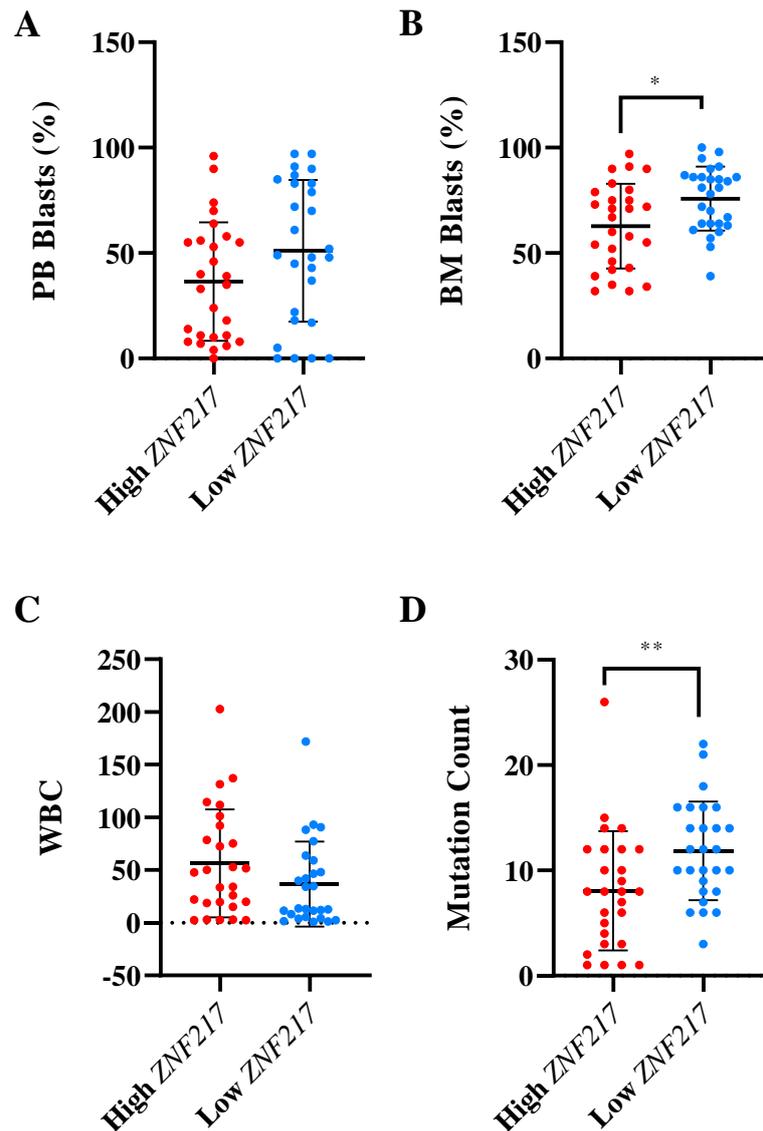
Regarding patient outcome, this analysis showed that patients with higher *ZNF217* expression were associated with a lower OS, as compared to patients with lower mRNA expression, with a median survival of 17 compared to 30 months., although this was deemed not significant (**Figure 4.3B**). Furthermore, patients with high *ZNF217* expression were found to possess a significantly reduced disease-free survival, accompanied by a hazard ratio of 2.2 (95% CI[1-4.8]), suggesting increased risk of relapse (**Figure 4.3C**).

Having determined that higher *ZNF217* expression is associated with reduced disease-free survival, this study aimed at determining the relationship between disease characteristics, such as FAB subtype or molecular abnormalities, as well as patient attributes, and *ZNF217* expression. Firstly, *ZNF217* expression was found to not correlate with PB blasts or WBC in AML patients (**Figure 4.4A, C**). A significantly higher proportion of BM blasts was detected in patients with lower *ZNF217* expression (**Figure 4.4B**), as well as a higher mutation count (**Figure 4.4D**). Both these factors are indicative of a poorer prognosis, despite lower *ZNF217* expression being associated with better disease-free survival. Moreover, dysregulated *ZNF217* expression was significantly correlated with AML FAB subtype (**Figure 4.5A**). Lastly, high *ZNF217* was not associated with patient characteristics, such as age or gender (**Figure 4.5B-C**). Altogether, these results suggest a potential role for *ZNF217* mis-regulation in the development of AML, regardless of subtype.



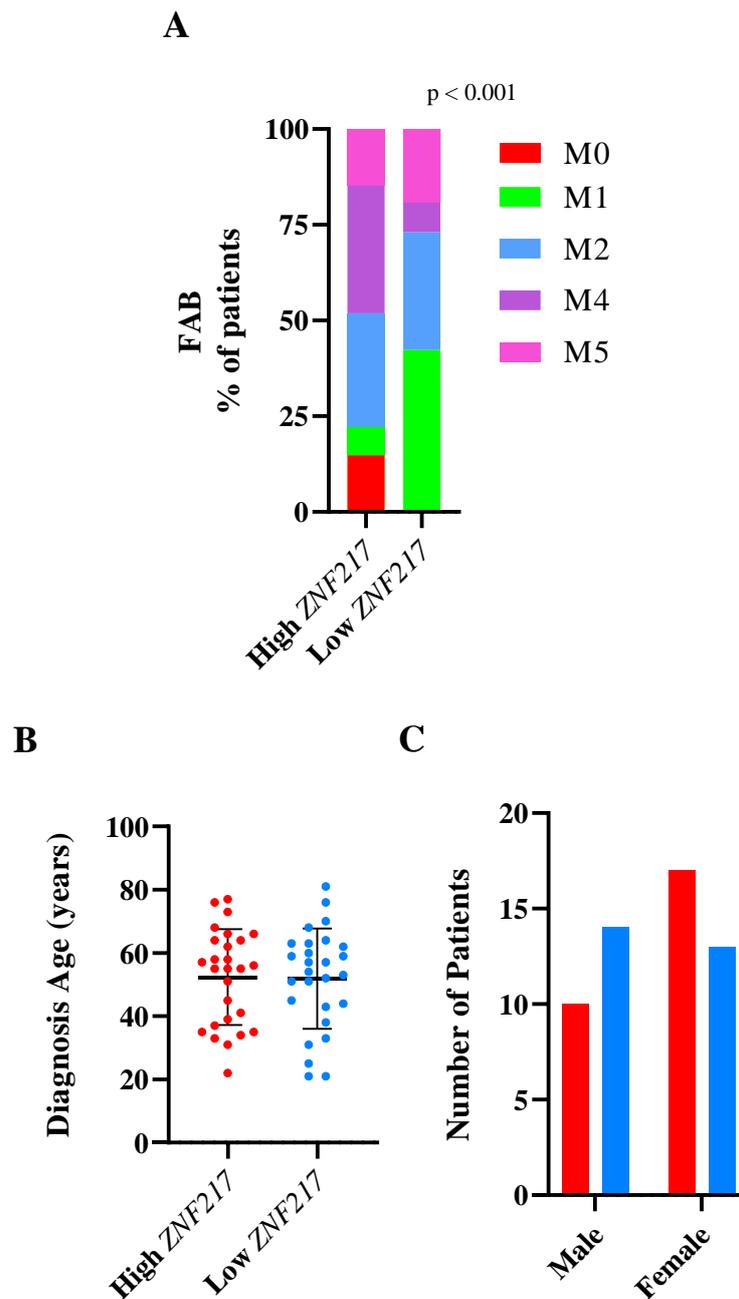
**Figure 4.3 – High levels of *ZNF217* is associated with poorer disease-free survival of AML patients**

(A) Normalised ( $\log_2$ ) *ZNF217* mRNA expression according to upper (n=27) and lower quartiles (n=27). Significance denoted by One-way ANOVA; \*\*\*\*  $p < 0.0001$ . (B) Overall survival and (C) disease-free survival analysis of AML patients stratified according to upper and lower *ZNF217* mRNA expression quartiles. For overall survival curves, *ZNF217* upper quartile n=27; *ZNF217* lower quartile n=27. For disease-free survival curves, *ZNF217* upper quartile n=26; *ZNF217* lower quartile n=27. Untreated patients and t(15;17) AML patients (that present a different treatment regime compared to other AML subtypes) were excluded from this analysis. Statistical analysis was performed using the Long-Rank test between high and low *ZNF217* expression groups. Data obtained from TCGA. (Ley *et al.*, 2013) using cBioPortal ([www.cbioportal.org](http://www.cbioportal.org)).



**Figure 4.4 – High *ZNF217* expression is associated with decreased BM blasts and lower mutation count**

Relationship between high and low *ZNF217* expression and several clinical attributes of AML patients. Data obtained from TCGA (Ley et al., 2013) using cBioPortal ([www.cbioportal.org](http://www.cbioportal.org)). *ZNF217* upper quartile (n=27); *ZNF217* lower quartile (n=27). Graph showing the percentage of (A) peripheral blood (PB) blasts; (B) BM blasts; (C) white blood count (WBC); and (D) mutation count of AML patients according to *ZNF217* high (n=27) and low (n=27) expression. Statistical differences between high and low *ZNF217* expression in AML patients was analysed using Mann-Whitney test; Data represents mean  $\pm$  1SD; \* denotes  $p < 0.05$ ; \*\* denotes  $p < 0.01$ .



**Figure 4.5 – Dysregulated *ZNF217* expression is not associated with patient characteristics**

Relationship between high and low *ZNF217* expression and AML patient characteristics. Data obtained from TCGA (Ley *et al.*, 2013) using cBioPortal ([www.cbioportal.org](http://www.cbioportal.org)). *ZNF217* upper quartile (n=27); *ZNF217* lower quartile (n=27). Graph showing the (A) number of patients across the different FAB subtypes of AML (M0-M7), (B) age at diagnosis (years) and (C) proportion of male and female patients according to *ZNF217* high (n=27) and low (n=27) expression. Statistical differences between high and low *ZNF217* expression in AML patients was analysed using Mann-Whitney test; Data represents mean  $\pm$  1SD; \* denotes  $p < 0.05$ .

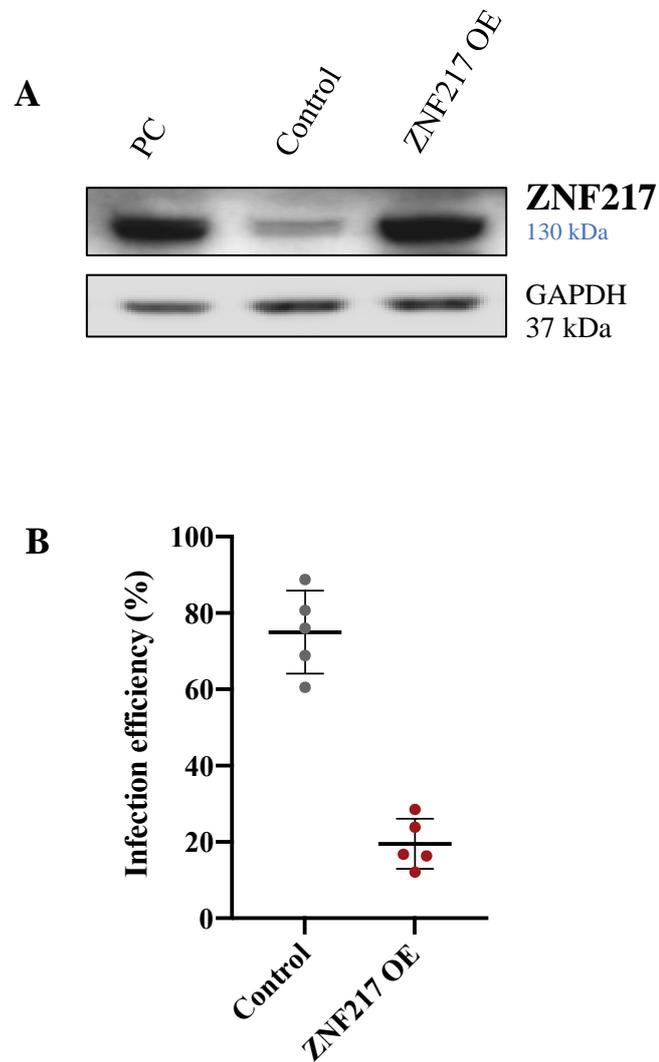
### 4.3.3 Overexpression of ZNF217 promotes human myeloid development

The above data shows that *ZNF217* expression increases with myeloid cell development and the previous chapter demonstrated that this gene and its corresponding protein are upregulated in RUNX1-ETO-expressing HSPCs compared to control vector. To determine the effect of ZNF217 overexpression on myeloid growth, differentiation, and self-renewal, I ectopically expressed this protein as a single abnormality in the CD34<sup>+</sup> HSPC *in vitro* model. Knockdown studies were subsequently performed in 4.3.4.

#### 4.3.3.1 Generation of ZNF217 overexpressing CD34<sup>+</sup> HSPC

A lentiviral vector harbouring the ZNF217 CDS and co-expressing GFP as a selectable marker was used to infect CD34<sup>+</sup> HSPC. To confirm ZNF217 protein expression, western blot was performed on transfected HEK 293T cells. As shown in **Figure 4.6A**, cells transfected with ZNF217 had higher ZNF217 protein expression compared to cells transduced with control vector. A significantly lower infection rate was observed in CD34<sup>+</sup> HSPC transduced with a ZNF217-overexpression plasmid, primarily due to plasmid size (**Figure 4.6B**). For this reason, it was not technically feasible to isolate sufficient nuclear protein from HSPC for western blot.

Assays which required a pure population, such as colony forming efficiency (2.3.4), as well as cell cycle (2.7.5), apoptosis (2.7.6) and morphological analysis (2.3.5), necessitated FACS sorting to ensure non-transduced cells wouldn't influence the interpretation of the data. However, for assays which facilitate sub-selection of the population of interest (i.e. flow cytometry based), it is possible to gate through the GFP population prior to analysis (2.7.4). It is appreciated that this mixed culture could enable non-transduced cells to influence transduced cells; however, sorting this population would not allow the recovery of enough cells to perform such assays. Additionally, puromycin selection was not feasible in these cells, as this process takes approximately 48-72h, in which analysis of the cells would not be possible.



**Figure 4.6 – Transduction of ZNF217 in HEK 293T cells and human HSPC**

(A) Western blot showing ZNF217 protein expression in HEK 293T cells transduced with ZNF217 plasmid or control plasmid (n=1). The K562 cell line was used as a positive control (PC). GAPDH was used as a loading control (n=1). (B) Summary data showing percentages GFP positivity in control and ZNF217-overexpression CD34<sup>+</sup> HSPC cultures, on day 3 of culture. Data indicates Mean  $\pm$  1SD (n=5).

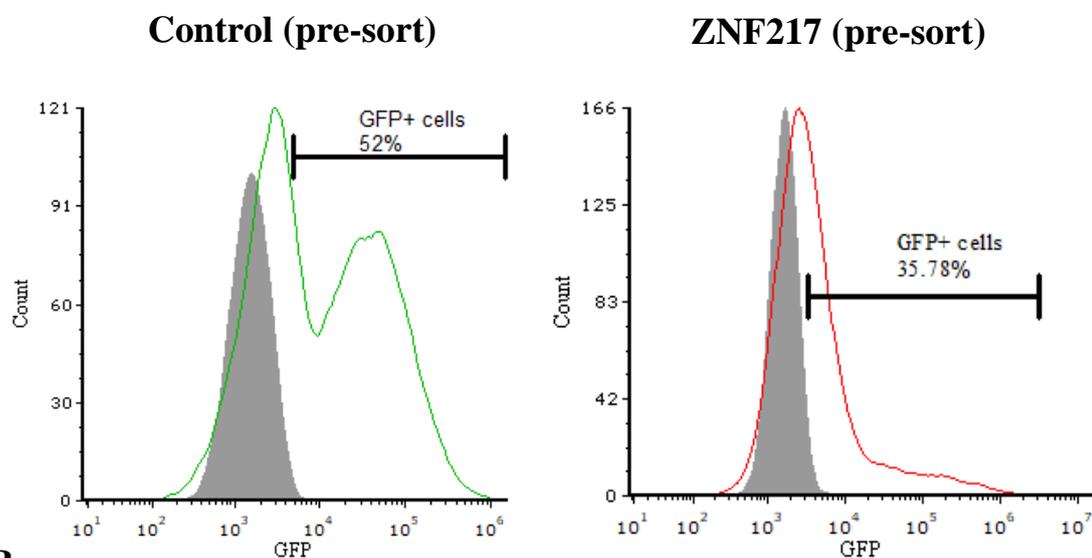
#### 4.3.3.2 *ZNF217 overexpression suppresses myeloid colony formation and self-renewal*

Previous studies have shown that ZNF217 overexpression in primary mammary epithelial or breast cancer cells is able to promote the formation of mammospheres with increased self-renewal potential (Vendrell *et al.*, 2012; Littlepage *et al.*, 2012). Initially, to determine the effect of ZNF217 overexpression on myeloid colony forming ability and self-renewal potential, a colony assay, followed by subsequent replating was performed on GFP<sup>+</sup> cells isolated by FACS (**Figure 4.7**), thus ensuring a pure transduced population. Following 7 days of culture, ZNF217 overexpression led to a significant 35% reduction colony-forming ability compared to control cells (**Figure 4.8A**). Furthermore, a subsequent serial replating assay showed a significant 60% reduction (**Figure 4.8B**) in the cells' self-renewal potential as a result of ZNF217 expression, as compared to control cells. These results indicate that ZNF217 impairs myeloid colony formation, with a reduction in the cells' self-renewal ability. However, it is important to note that the current assay does not allow the discrimination between different types of colonies, including monocytic- and granulocytic-derived cells. As the media used does not support erythroid cell development, it is possible to affirm that these are not present in culture. Additional assays using semi-solid culture medium, such as methylcellulose, would be necessary to definitively make further conclusions.

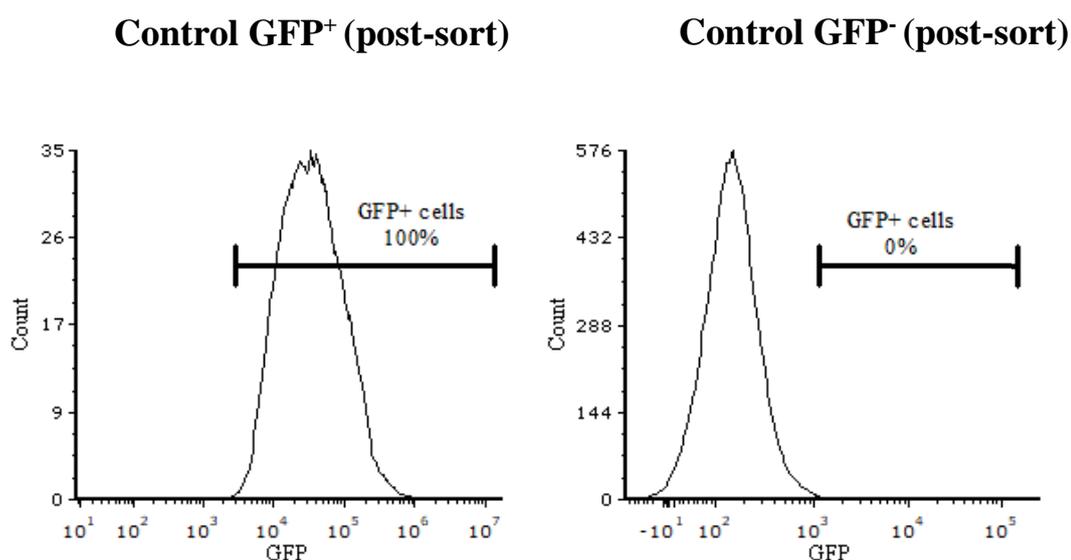
#### 4.3.3.3 *ZNF217 overexpression suppresses the growth of myeloid cells*

Previous studies have shown that upregulated ZNF217 expression leads to the suppression of breast, colon and teratocarcinoma cancer cell differentiation, leading to tumorigenesis (Krig *et al.*, 2007). I hypothesise that ZNF217 overexpression may have an anti-differentiation phenotype on myelopoiesis. To analyse the effect of ZNF217 overexpression in myeloid development, specific time-points were selected at an appropriate frequency to fully capture these changes, previously optimised within the research group. Moreover, cells were re-seeded at the appropriate density following each time-point, as described in **2.3.3**. Overexpression of ZNF217 led to a 2.2-fold decrease in the growth of myeloid cells following 13 days of culture (**Figure 4.9A**). To analyse the impact on different myeloid lineages, the CD13 and CD36 cell surface markers were used to discriminate between monocytic, granulocytic and erythroid populations. ZNF217 caused a significant decrease in monocytic cell growth, compared to control cells (3.5-fold on day 13) (**Figure 4.9B**). A slower growth rate was equally observed in the granulocytic lineage, in which ZNF217-overexpressing cells showed a 2-fold decrease in growth compared to control. (**Figure 4.9C**).

A

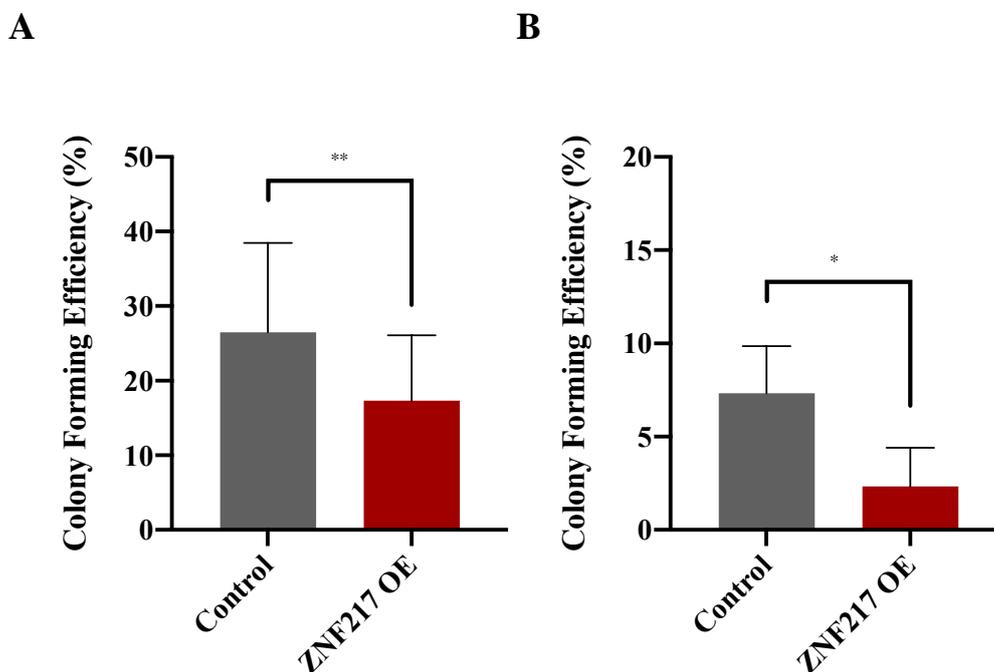


B



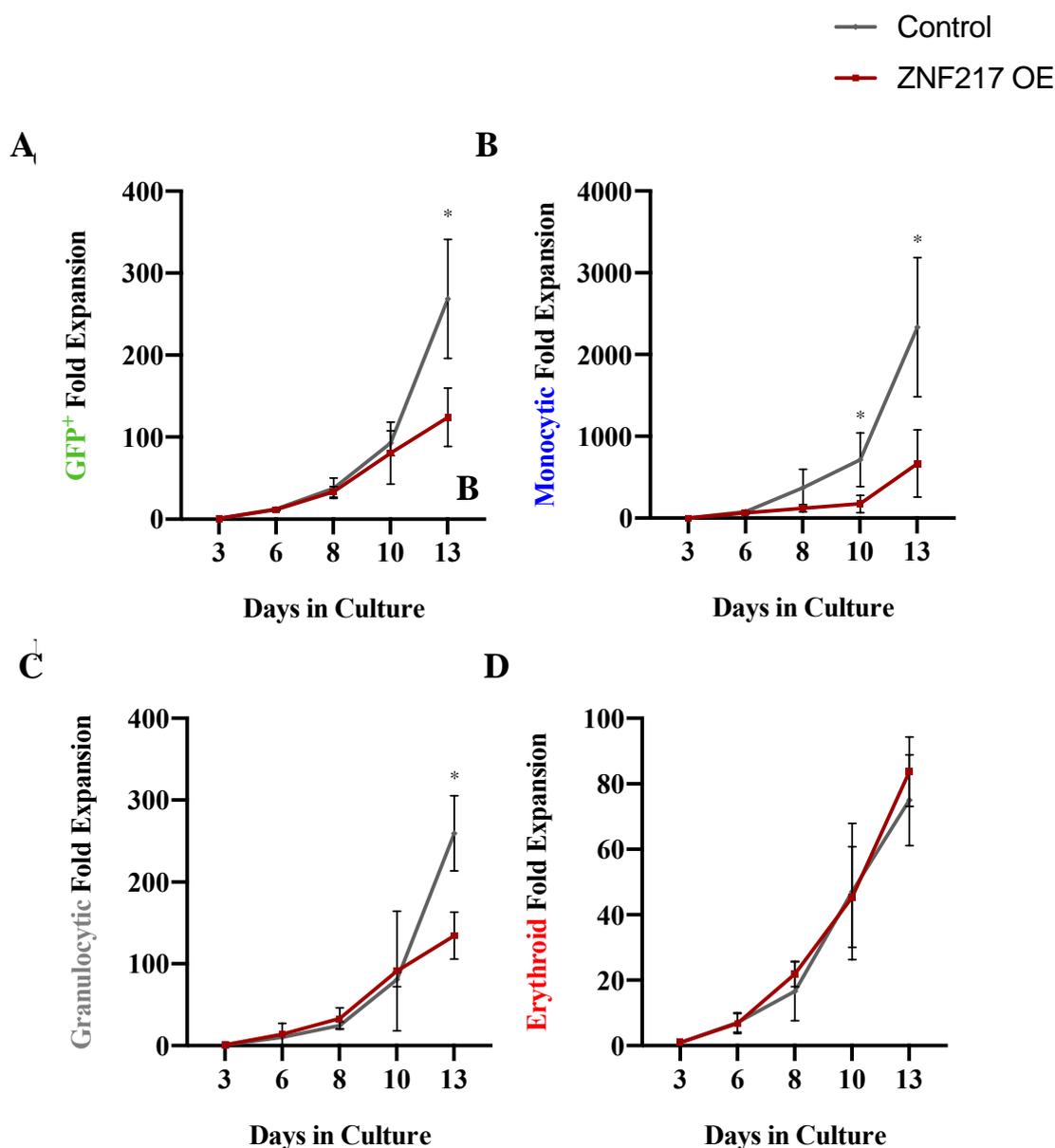
**Figure 4.7 – Expression and enrichment of GFP<sup>+</sup> populations in human CD34<sup>+</sup> HSPC**

(A) Representative flow cytometric histograms showing GFP expression in control and ZNF217 transduced cultures (pre-sorted cultures on day 3 of culture). Plots generated based on the “Non-debris” gate described in **Figure 2.4**. Background auto fluorescence was established using CD34<sup>+</sup> cells subjected to the equivalent viral infection procedure but in the absence of virus (mock). Mock HSPC – grey; Transduced HSPC - green (control) / red (ZNF217 OE) (B) Representative histogram for control culture post sorting for GFP<sup>+</sup> and GFP<sup>-</sup> cell populations. Plots generated based on the “Non-debris” gate described in **Figure 2.4**.



**Figure 4.8 – ZNF217 overexpression inhibits myeloid colony formation and decreases self-renewal capacity**

(A) Colony forming efficiency of control and ZNF217-overexpression cultures following 7 days of growth in liquid culture containing IL-3, SCF, G- and GM-CSF (2.3.4). Data indicates Mean  $\pm$  1SD (n=5). Significant difference was analysed by paired T-test; \*\* denotes  $p < 0.01$ . (B) Self-renewal potential assessed by a single replating round of control and ZNF217 cultures, in the same conditions as previously. Data indicates Mean  $\pm$  1SD (n=3). Significant difference between ZNF217-expressing cells and control was analysed by paired t-test; \* denotes  $p < 0.05$ ; \*\* denotes  $p < 0.01$ .



**Figure 4.9 – ZNF217 overexpression inhibits monocytic and granulocytic growth during myeloid development**

(A) Cumulative fold-expansion of control and ZNF217-overexpressing GFP<sup>+</sup> cells grown over 13 days in culture medium containing IL-3, SCF, G- and GM-CSF (**Supplementary Figure 2**). (B) Cumulative fold-expansion of monocytic cells (CD13<sup>+</sup> CD36<sup>+</sup>) transduced with a control of ZNF217-overexpressing vector. (C) Cumulative fold-expansion of granulocytic cells (CD13<sup>+/−</sup> CD36<sup>−</sup>) transduced with a control of ZNF217-overexpressing vector. (D) Cumulative fold-expansion of erythroid cells (CD13<sup>−</sup> CD36<sup>+</sup>) transduced with a control of ZNF217-overexpressing vector. Cells were re-seeded at the appropriate density following each time-point, as described in 2.3.3. Gating strategies were applied as described in **Figure 2.4** and **Figure 2.5**. Data indicates Mean  $\pm$  1SD (n $\geq$ 3) Significant differences between ZNF217-expressing cells and control at each time-point were analysed by paired t-test; \* denotes p<0.05; (this data has also been represented as dot plots, within the supplementary section [**Supplementary Figure 30**]).

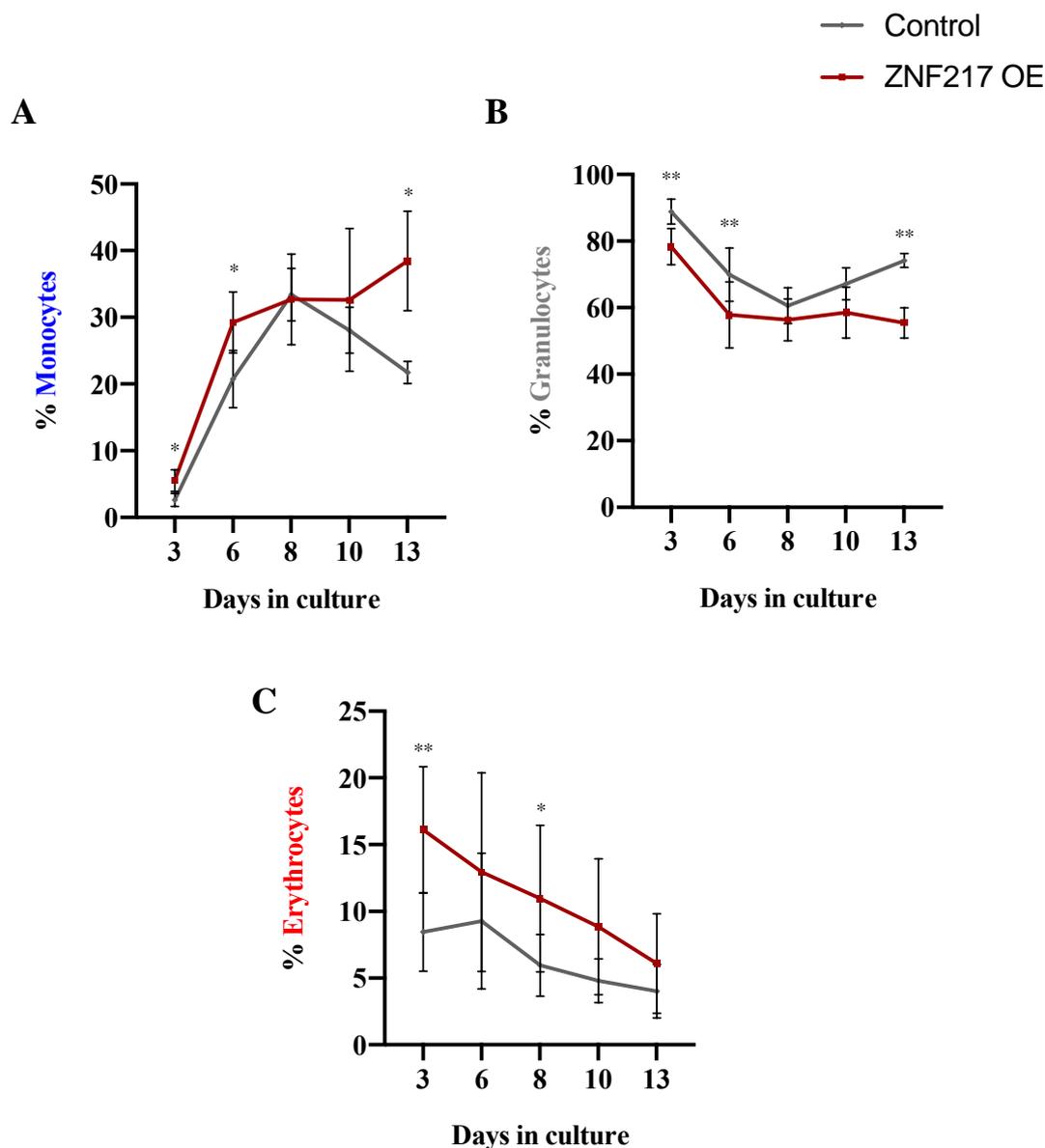
Whilst statistical significance is achieved late (10/13 days of cell culture) the trend is consistent throughout the experiment, suggesting a biological effect is present from onset.

However, the growth in erythroid cells was not significantly altered by ZNF217 overexpression in these cells (**Figure 4.9D**). This was not unexpected as the media in which these cells are grown promotes monocytic and granulocytic development, but not erythroid, due to the absence of the EPO, essential for terminal erythroid differentiation. These observations suggest that ZNF217 overexpression influences the growth of myeloid cells, particularly the cells committed to the monocytic and granulocytic lineages.

#### 4.3.3.4 *ZNF217 overexpression promotes myeloid differentiation*

To determine the consequences of ZNF217 overexpression on myeloid development, I analysed the expression of differentiation cell surface markers in combination with lineage discriminators over time using flow cytometry (**2.7.4**). As shown in **Figure 4.10A**, ZNF217 overexpression led to a significant 1.6-fold (day 13) increase in the percentage of monocytic cells in culture compared to control. Concomitantly, the proportion of granulocytic cells was significantly lower throughout development (**Figure 4.10B**). Conversely, there was an increase in erythroid-committed cells in the ZNF217 culture compared to control, within the same timeframe (**Figure 4.10C**). Altogether, these observations suggest that the balance between lineages in culture is influenced by ZNF217 overexpression, by reducing the granulocytic population at the expense of monocytes, during normal myeloid development of human HSPC.

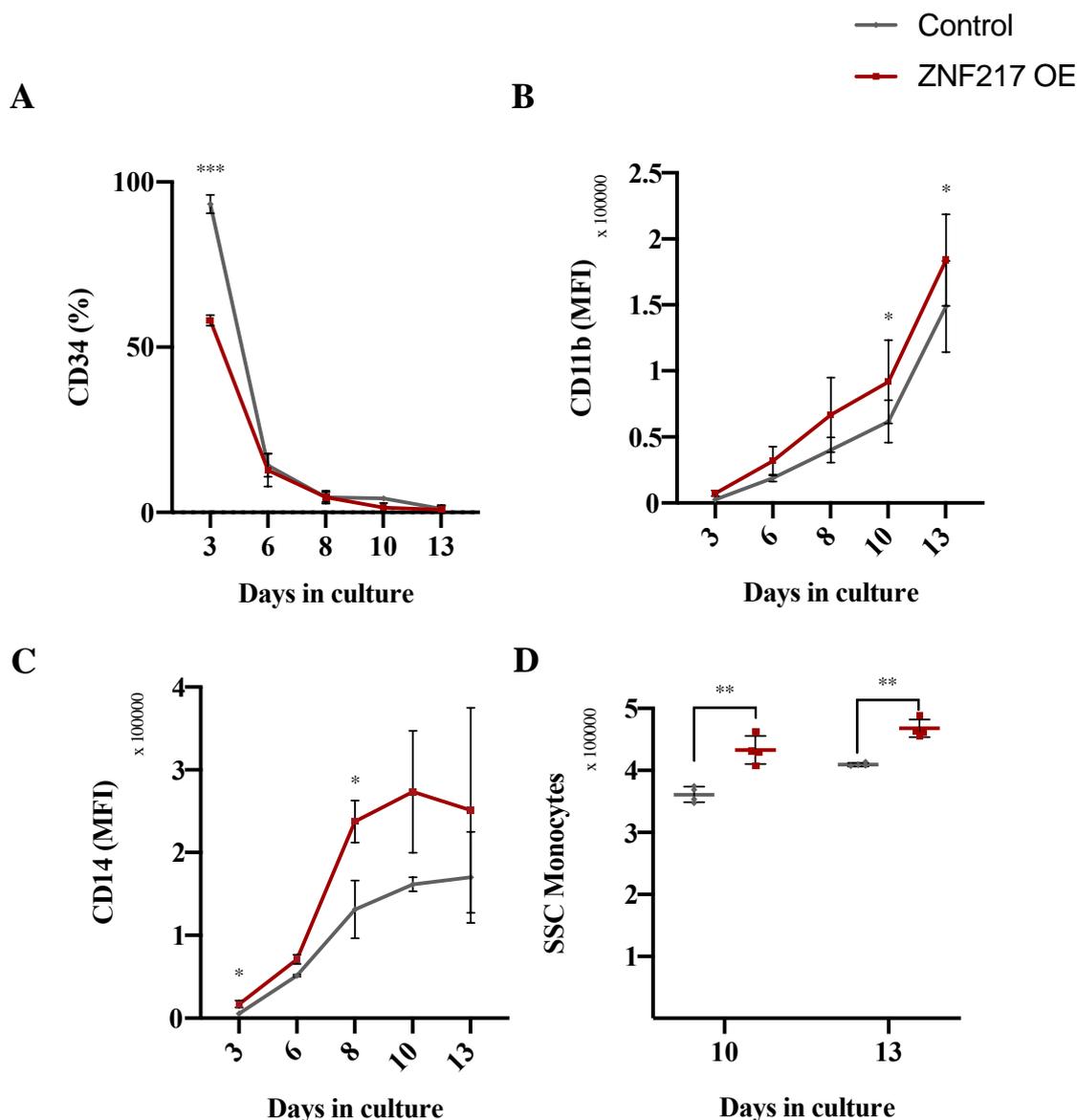
In order to analyse the consequences of ZNF217 overexpression on monocytic and granulocytic differentiation, this study assessed the expression of the differentiation cell surface markers CD34, CD11b, CD14 and CD15. CD34 is expressed in HSPC and progressively declines as HSPC differentiate (Caux *et al.*, 1989). Moreover, as myeloid cell development progresses, there is an increase in the expression of CD11b (Tonks *et al.*, 2004). Monocytic cells upregulate CD14 expression during development, whilst granulocytes show an increase in CD15 (Tonks *et al.*, 2004).



**Figure 4.10 – ZNF217 overexpression disrupts the balance between granulocytic and erythroid populations during myeloid cell development.**

Summary data showing the proportion of (A) monocytic (CD13<sup>+</sup> CD36<sup>+</sup>), (B) granulocytic (CD13<sup>+</sup> CD36<sup>-</sup>) and (C) erythroid cells (CD13<sup>-</sup> CD36<sup>+</sup>) in control and ZNF217-overexpressing cultures. Cells were re-seeded at the appropriate density following each time-point, as described in 2.3.3. Gating strategies were applied as described in Figure 2.4 and Figure 2.5. Data indicates Mean  $\pm$  1SD (n $\geq$ 4). Significant differences between ZNF217-expressing cells and control at each time-point were analysed by paired t-test; \* denotes p<0.05; (this data has also been represented as dot plots, within the supplementary section [Supplementary Figure 31]).

In monocytes, ZNF217 overexpression significantly altered monocytic development (**Figure 4.11**), resulting in a 40% reduction in the percentage of cells expressing the CD34 marker on day 3 (**Figure 4.11A**). Upregulation of CD11b was increased in ZNF217 overexpressing monocytic cells throughout development and, by day 10, it was 1.5-fold higher compared to control (**Figure 4.11B**). CD14 was also upregulated throughout development (**Figure 4.11C**). Additionally, a significant increase in cell granularity (SSC) on days 10 and 13 of development was further observed (**Figure 4.11D**).



**Figure 4.11 – ZNF217 overexpression promotes monocytic differentiation**

(A) Summary data of CD34<sup>+</sup> expression in terms of percentage in monocytic cells (CD13<sup>+</sup> CD36<sup>+</sup>) over time for control and ZNF217-overexpressing cells. Summary data showing (B) CD11b and (C) CD14 expression in terms of MFI in monocytic cells over time for control and ZNF217-overexpressing cells. (D) Complexity of monocytic cells was measured by changes in SSC for control and ZNF217-overexpressing cells on days 10 and 13 of myeloid differentiation. Cells were re-seeded at the appropriate density following each time-point, as described in 2.3.3. Gating strategies were applied as described in Figure 2.4 and Figure 2.5. Data indicates Mean  $\pm$  1SD (n $\geq$ 3). Significant differences between ZNF217-expressing cells and control at each time-point were analysed by paired t-test; \* denotes p<0.05, \*\* denotes p<0.01, \*\*\* denotes p<0.001; (for representative flow cytometry plots please refer to Supplementary Figure 3, Supplementary Figure 4, Supplementary Figure 5; this data has also been represented as dot plots, within the supplementary section [Supplementary Figure 32])

Analysis of morphological features showed an increase in the percentage of cells displaying late monocytic features, consistent with immunophenotypic analysis (**Figure 4.13**). These results suggest that overexpression of ZNF217 promotes monocytic differentiation which may give rise to the decreased proliferation of this lineage observed in **Figure 4.10A**.

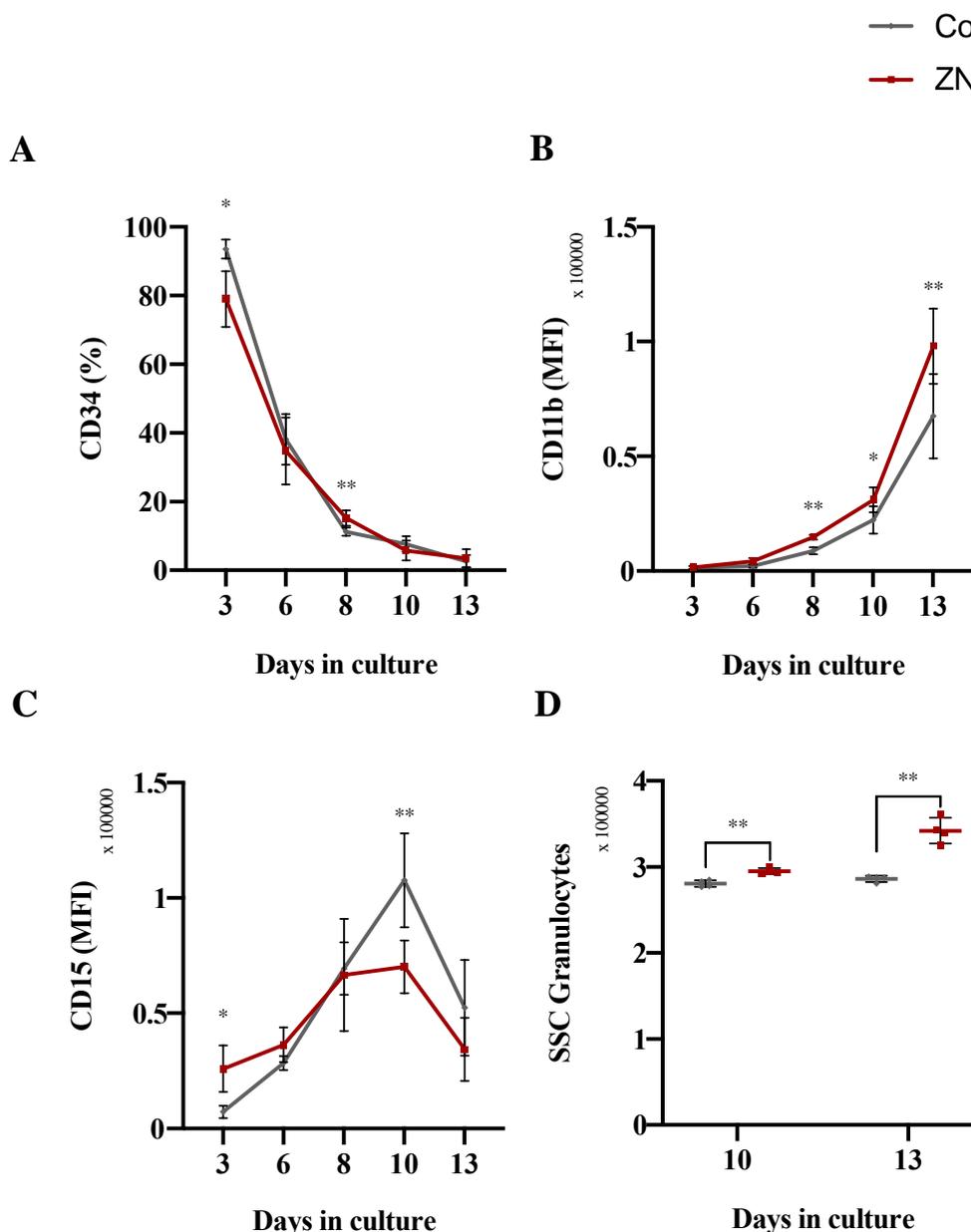
ZNF217 overexpression was shown to also perturb granulocytic lineage development (**Figure 4.12**). CD34<sup>+</sup> was only modestly impacted by ZNF217 (**Figure 4.12A**). CD11b was shown to be upregulated in ZNF217 cells (1.3-fold on day 13; **Figure 4.12B**) suggesting that differentiation was again promoted in this lineage an observation supported by an increase in cell granularity (SSC) (**Figure 4.12D**), however, ZNF217-overexpressing cells had a lower expression level of CD15 (1.5-fold decrease by day 13), compared to control cells (**Figure 4.12C**) confounding this interpretation. Additionally, morphological analysis of granulocytic cells showed no difference in ZNF217-expressing cultures, as compared to control, with an equal proportion of cells displaying late morphological features (**Figure 4.13**).

Taken together, these results suggest that the overexpression of ZNF217 in CD34<sup>+</sup> HSPC leads to a pro-differentiation phenotype, as evidenced by upregulation of monocytic and, to some extent, granulocytic developmental markers, and morphology.

#### 4.3.3.5 *ZNF217 overexpression in HSPC induces changes in cell cycle*

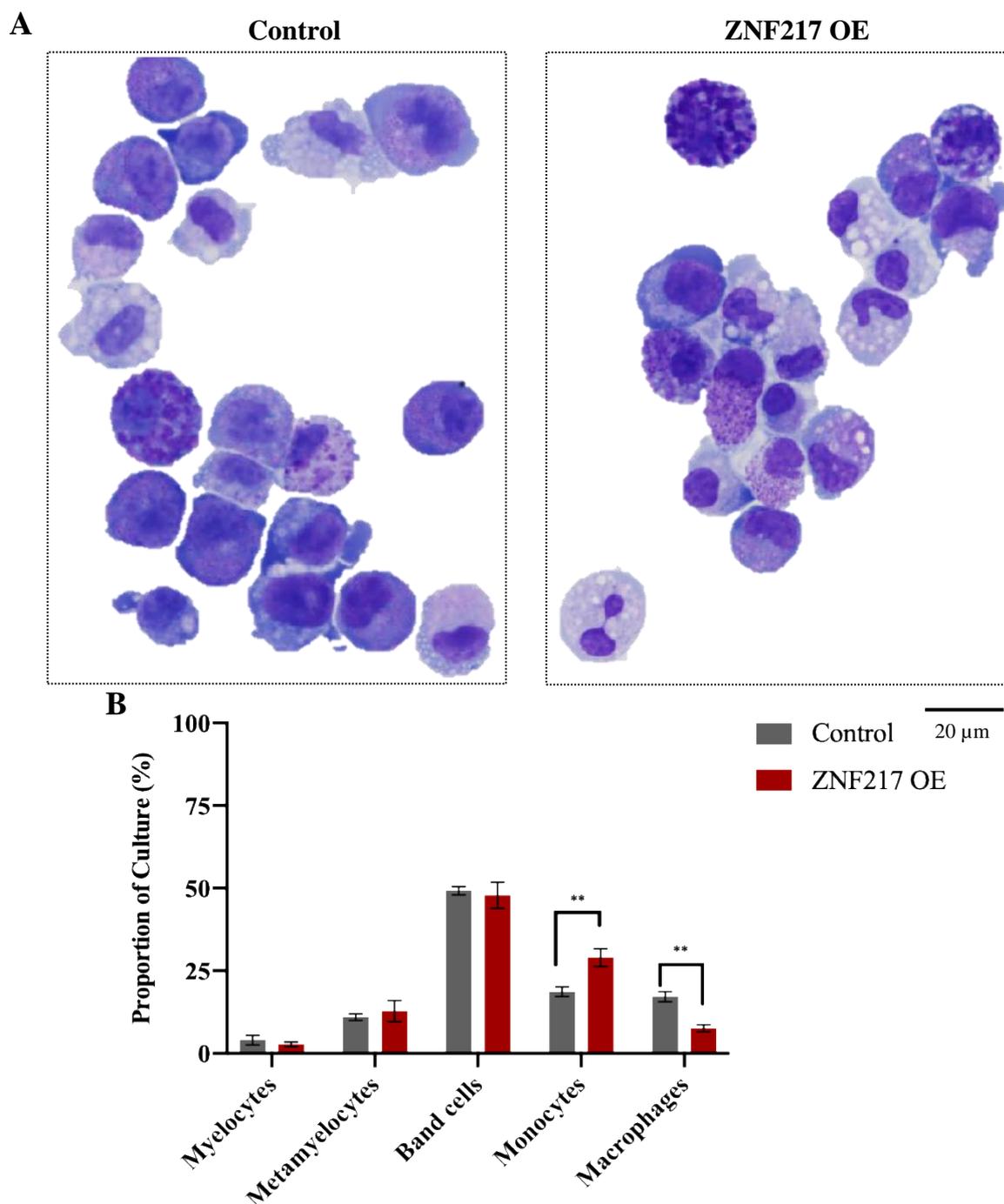
Having determined that ZNF217 leads to a decrease in colony forming ability and self-renewal potential, two possible explanations for this were tested: decreased proliferative capacity and/or increased apoptosis. ZNF217-transduced cells showed a significant decrease in the proportion of cells in cycle (S + G<sub>2</sub>/M), indicating an accumulation of cells in G<sub>1</sub> phase supporting the hypothesis that ZNF217 impacts proliferative capacity (**Figure 4.14A**).

In order to assess whether ZNF217 overexpression led to an increase in apoptosis, an Annexin V-based assay was performed. No significant difference in apoptosis was observed for ZNF217-transduced cells compared to control in the days examined (**Figure 4.14B**). These results indicate that the reduced colony forming ability of ZNF217-expressing cells and reduced growth ability of HSPC is not a result of cell death, but a consequence of the decrease in the proportion of cells entering cell cycle, suggesting that these cells are dividing at a slower rate, as compared to control cells.



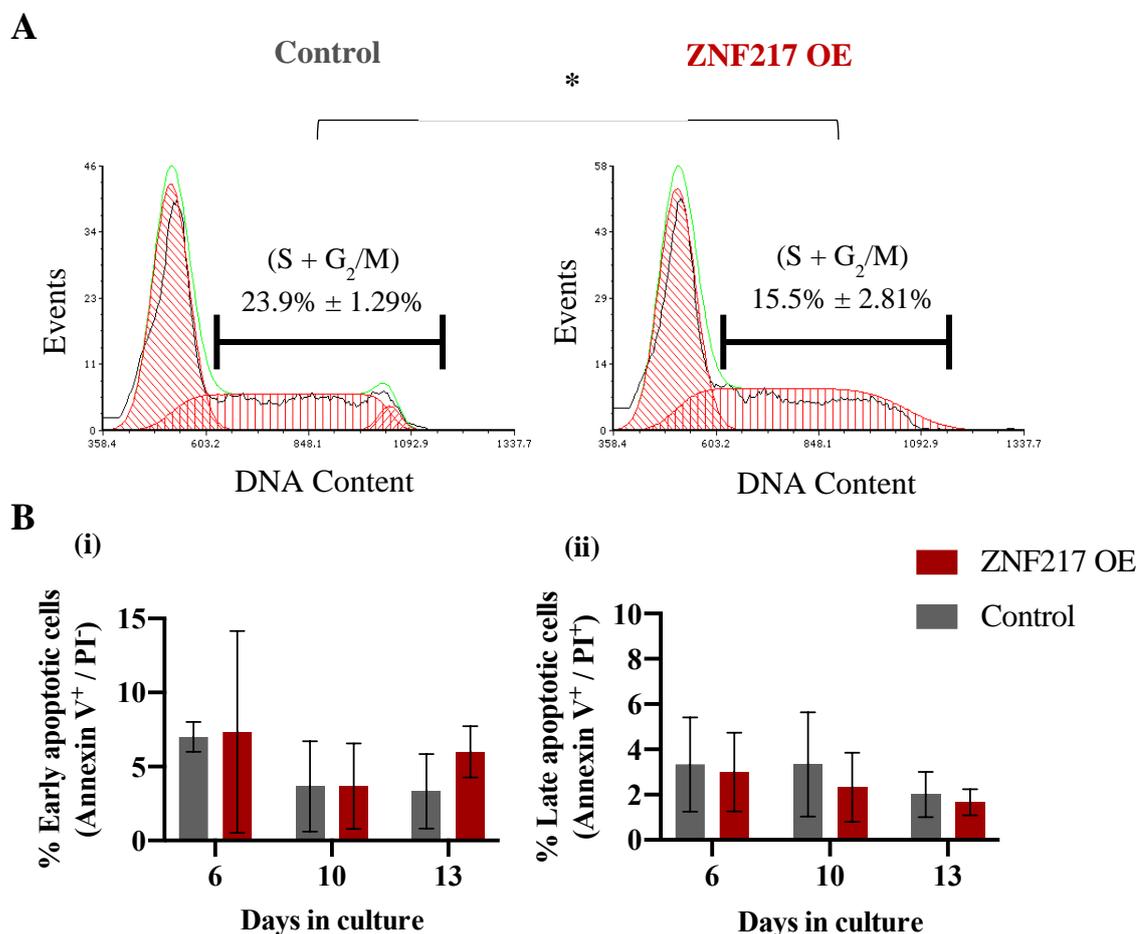
**Figure 4.12 – ZNF217 overexpression promotes granulocytic differentiation by upregulating monocytic markers**

(A) Summary data of CD34<sup>+</sup> expression in granulocytic cells (CD13<sup>+</sup> CD36<sup>-</sup>). Summary data showing (B) CD11b and (C) CD15 expression in granulocytic cells. (D) side scatter analysis (SSC) of transduced cells on days 10 and 13 of myeloid differentiation. Cells were re-seeded at the appropriate density following each time-point, as described in 2.3.3. Gating strategies were applied as described in Figure 2.4 and Figure 2.5. Data indicates Mean  $\pm$  1SD (n $\geq$ 3). Significant differences between ZNF217-expressing cells and control at each time-point were analysed by paired t-test; \* denotes p<0.05, \*\* denotes p<0.01. (for representative flow cytometry plots please refer to Supplementary Figure 6. Supplementary Figure 7, Supplementary Figure 8; this data has also been represented as dot plots, within the supplementary section [Supplementary Figure 33]).



**Figure 4.13 – ZNF217 overexpression alters cell morphology during myeloid development**

(A) Representative images showing the morphology of HSPC transduced with control and ZNF217 OE cultures were analysed on day 13 of differentiation by cytospinning  $3 \times 10^4$  cells and staining them with May-Grünwald-Giemsa differential stain. (B) Differential counts of all cultures with morphology categorised into granulocytic and monocytic-committed cells. Approximately 500 cells were counted in total and each population count normalised to total number of cells, according to **Figure 2.2**. Granulocytic cells were further divided into myelocytes, metamyelocytes and band/segmented cells. Data indicates mean  $\pm$  1SD (n=3). Significant difference between ZNF217 OE cultures and control were analysed by paired T-test; \*\* denotes  $p < 0.01$ .



**Figure 4.14 – ZNF217 overexpression in CD34<sup>+</sup> HSPC significantly disrupts normal cell cycle.**

(A) Representative histograms showing cell cycle distribution of normal HSPC (Control) and ZNF217-expressing cells after 6 days of culture. Initial gating was applied as described in **Figure 2.6** The percentage of cells in S + G<sub>2</sub>/M is indicated by mean ± 1SD (n=3) Significant difference between ZNF217-expressing cells and control was analysed by paired t-test; \* denotes p<0.05. (B) Summary data showing the effect of ZNF217 overexpression in apoptosis, on CD34<sup>+</sup> HSPC following 6, 10 and 13 days of culture. Initial gating was applied as described in **Figure 2.7** (i) Early apoptotic cells are characterised by Annexin V<sup>+</sup> / PI<sup>-</sup>; (ii) Apoptotic cells are characterised by Annexin V<sup>+</sup> / PI<sup>+</sup>. Data indicates Mean ± 1SD (n=3).

Whilst feasible, earlier time-points would not be practical for three main reasons. Firstly, an apoptosis assay requires a minimum of 50,000 cells and hence would use a substantial proportion, if not all, of sorted cells. Secondly, the CD34<sup>+</sup> model exhibits the greatest growth potential between days 3 and 6 of culture, in which 15-20-fold growth is frequently observed; taking cells on day 3 would substantially reduce the available cells for subsequent analysis. Lastly, as cells have only just been collected from cell sorting, it is likely that the stresses incurred by this process would mask any biological effects. For these reasons, day 6 is the earliest suitable time-point for the apoptotic assay.

Altogether, these results indicate that ZNF217 has a pro-differentiation role, in concordance with the mRNA analysis described above (4.3.1), in which the expression of ZNF217 increases with myeloid differentiation. Overexpressing ZNF217 in HSPC suggests that this favours monocytic development, although granulocytic development might also be influenced by its expression. More importantly, given these observations, it is unlikely that overexpression of ZNF217 contributes to leukaemogenic development through change in normal myeloid cell differentiation.

#### 4.3.4 Knockdown of ZNF217 does not impact myeloid differentiation

The data above shows that ZNF217 overexpression was shown to promote normal myeloid development by significantly increasing monocytic and granulocytic cell differentiation. Previously, several authors have also addressed the effect of ZNF217 downregulation in solid tumours. Sun *et al.* used shRNA-mediated KD vectors in ovarian cancer cell lines, and showed that ZNF217-KD cells presented a decrease in cell growth and in colony formation efficiency (Sun *et al.*, 2008a). Similar observations were made by Li *et al.* in colorectal cancer, in which the ZNF217-siRNA mediated KD lead to a decrease in cell proliferation and impaired cell migration and invasion in the colorectal cancer cell line SW480 (Li *et al.*, 2015). Altogether these results indicate that the ZNF217-mediated cell proliferation and invasion can be attenuated by employing KD. However, the consequences of ZNF217 KD in haematopoiesis have not been determined yet. Therefore, this study aimed at determining the effect of ZNF217 KD on myeloid growth and differentiation of CD34<sup>+</sup> HSPC, using shRNA constructs.

##### 4.3.4.1 Selection and validation of ZNF217 shRNA constructs

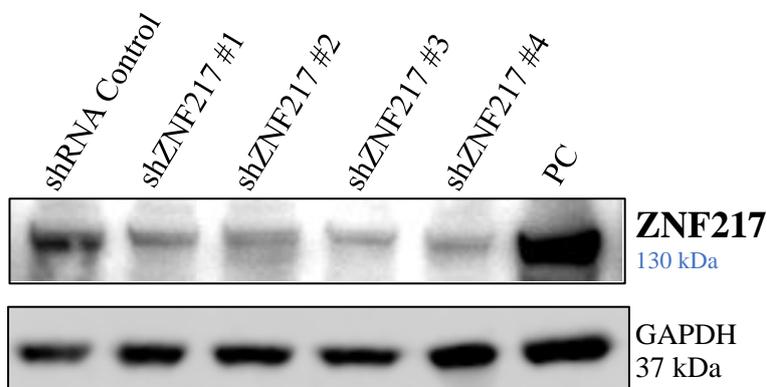
To reduce the levels of ZNF217 expression in normal human CD34<sup>+</sup> HSPC, lentiviral vectors based on shRNA constructs targeting ZNF217, encoding both a GFP and puromycin were used (**Table 2.1**). To initially validate KD, ZNF217 was KD in K562 cells. The level of KD achieved was similar across all constructs used, with a reduction of at least 50% in the levels on endogenous ZNF217 protein expressed in these cells (**Figure 4.15A-B**).

In CD34<sup>+</sup> HSPC, it was possible to achieve over 60% infection efficiency for the different shRNA constructs (**Figure 4.16**). For subsequent assays, shZNF217 #4 was discarded, as its expression resulted in cell death (data not shown), possibly due to off-target effects.

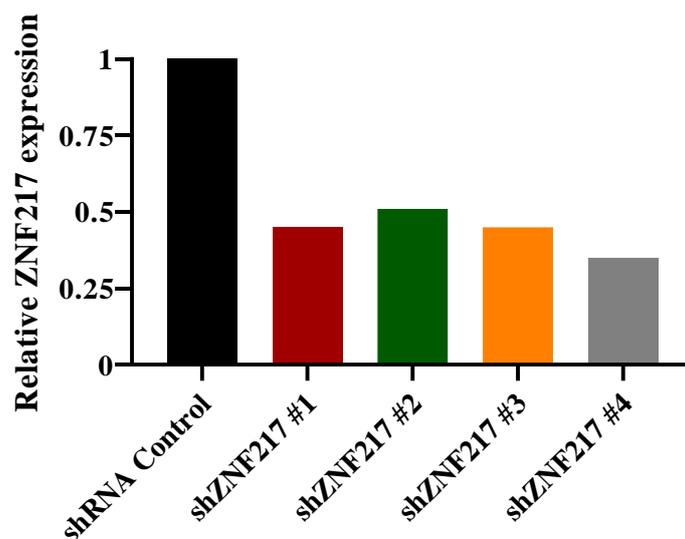
##### 4.3.4.2 Knockdown of ZNF217 leads to a decrease in CD34<sup>+</sup> HSPC colony forming ability

To determine the effect of ZNF217 KD on colony forming ability of myeloid cells, transduced CD34<sup>+</sup> cells were sorted based on GFP<sup>+</sup> expression and plated by limiting dilution, as previously described (**4.3.3**). Following 7 days of culture, all three ZNF217-KD cultures had a significant decrease in colony forming ability of approximately 50%, compared to control cells (**Figure 4.17A**).

A

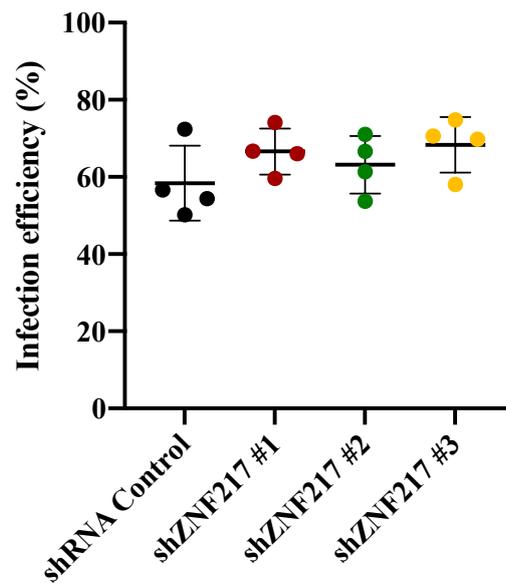


B



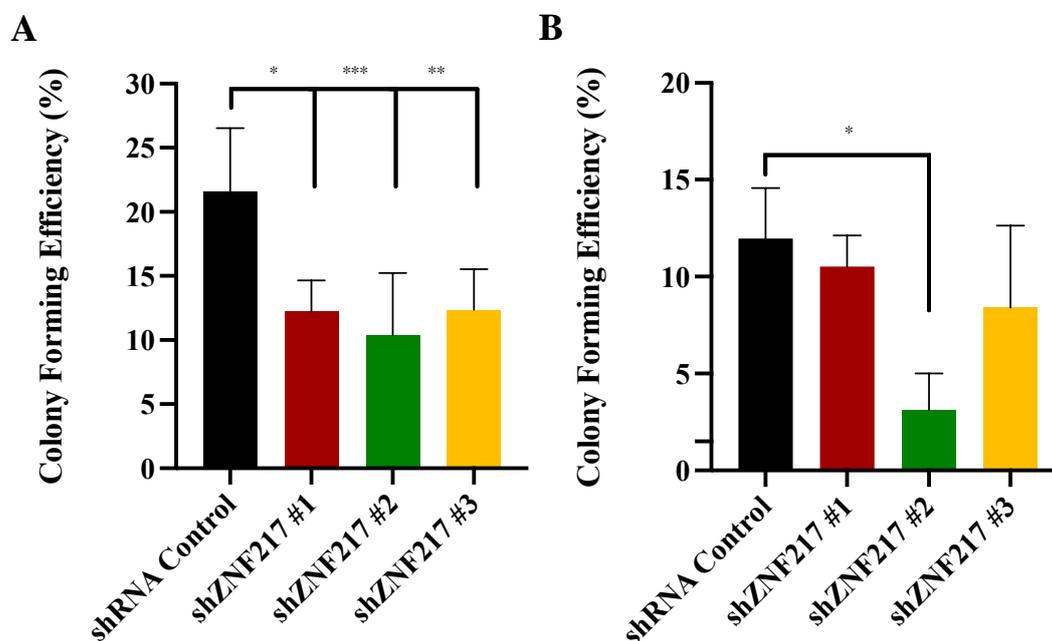
**Figure 4.15 – ZNF217 shRNA constructs successfully reduced the levels of endogenous ZNF217 in K562 cells**

(A) Western blot of ZNF217 total protein levels in K562 cells transduced with a scrambled shRNA vector, as well as three different ZNF217 shRNA constructs (n=1). K562 parental cells were used as a positive control (PC). GAPDH was used as a loading control. TRCN numbers have been described in **Table 2.1**. (B) Western blot semi-quantification by densitometry using ImageJ. Relative protein expression was determined by normalising each sample to GAPDH and to shRNA control.



**Figure 4.16 – Infection efficiency of transduced CD34<sup>+</sup> HSPC using shRNA lentiviral constructs**

Summary data showing the percentage of GFP<sup>+</sup> cells in control and ZNF217 KD CD34<sup>+</sup> HSPC cultures, analysed on day 3 of culture. Gating strategies were applied as described in **Figure 2.4**. Data indicates Mean  $\pm$  1SD (n=4).



**Figure 4.17 – ZNF217 knockdown disrupts myeloid colony formation of HSPC**

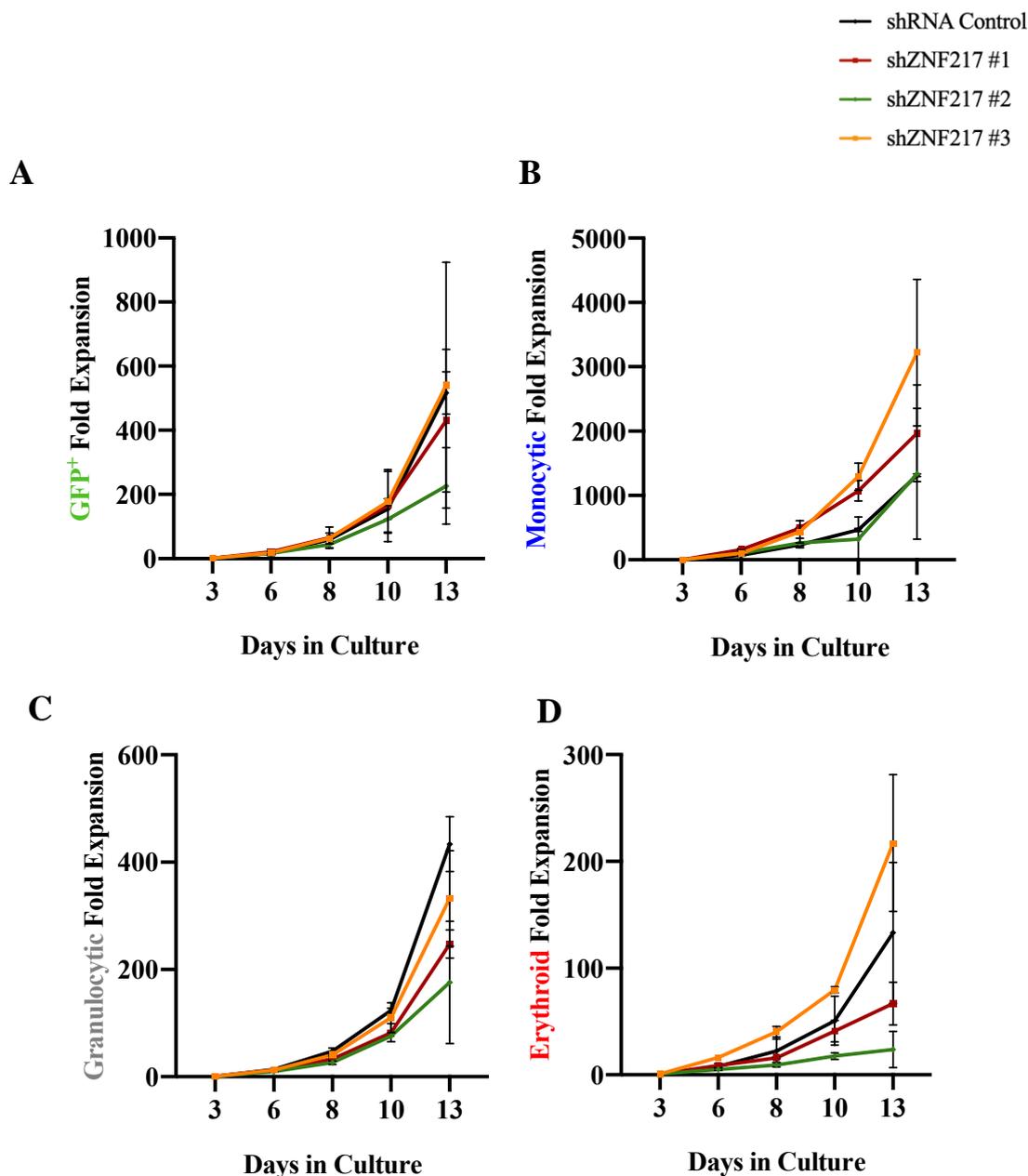
(A) Colony forming efficiency of control and three ZNF217-knockdown cultures following 7 days of growth in liquid culture containing IL-3, SCF, G- and GM-CSF (2.3.4). Data indicates Mean  $\pm$  1SD (n=4). Significant difference between shZNF217 cultures and control was analysed by one-way ANOVA with Bonferroni's multiple comparisons correction; \* denotes  $p < 0.05$ ; \*\* denotes  $p < 0.01$ , \*\*\* denotes  $p < 0.001$ . (B) Self-renewal potential was assessed by a single replating round of control and ZNF217-KD colonies, in the same conditions as previously. Data indicates Mean  $\pm$  1SD (n=3). Significant differences between shZNF217 cultures and control were analysed by one-way ANOVA with Bonferroni's multiple comparisons correction; \* denotes  $p < 0.05$ ; \*\* denotes  $p < 0.01$ ; \*\*\* denotes  $p < 0.001$ .

Furthermore, a replating assay showed no difference in self-renewal ability between control cells and two ZNF217 KD cultures; however, shZNF217 #2 showed a 75% reduction in colony forming-ability (**Figure 4.17B**). Altogether, reduction of ZNF217 levels led to a decrease in the cells' colony forming ability, with no loss of self-renewal potential for two of the constructs used.

#### 4.3.4.3 *Knockdown of ZNF217 disrupts myeloid cell growth and development*

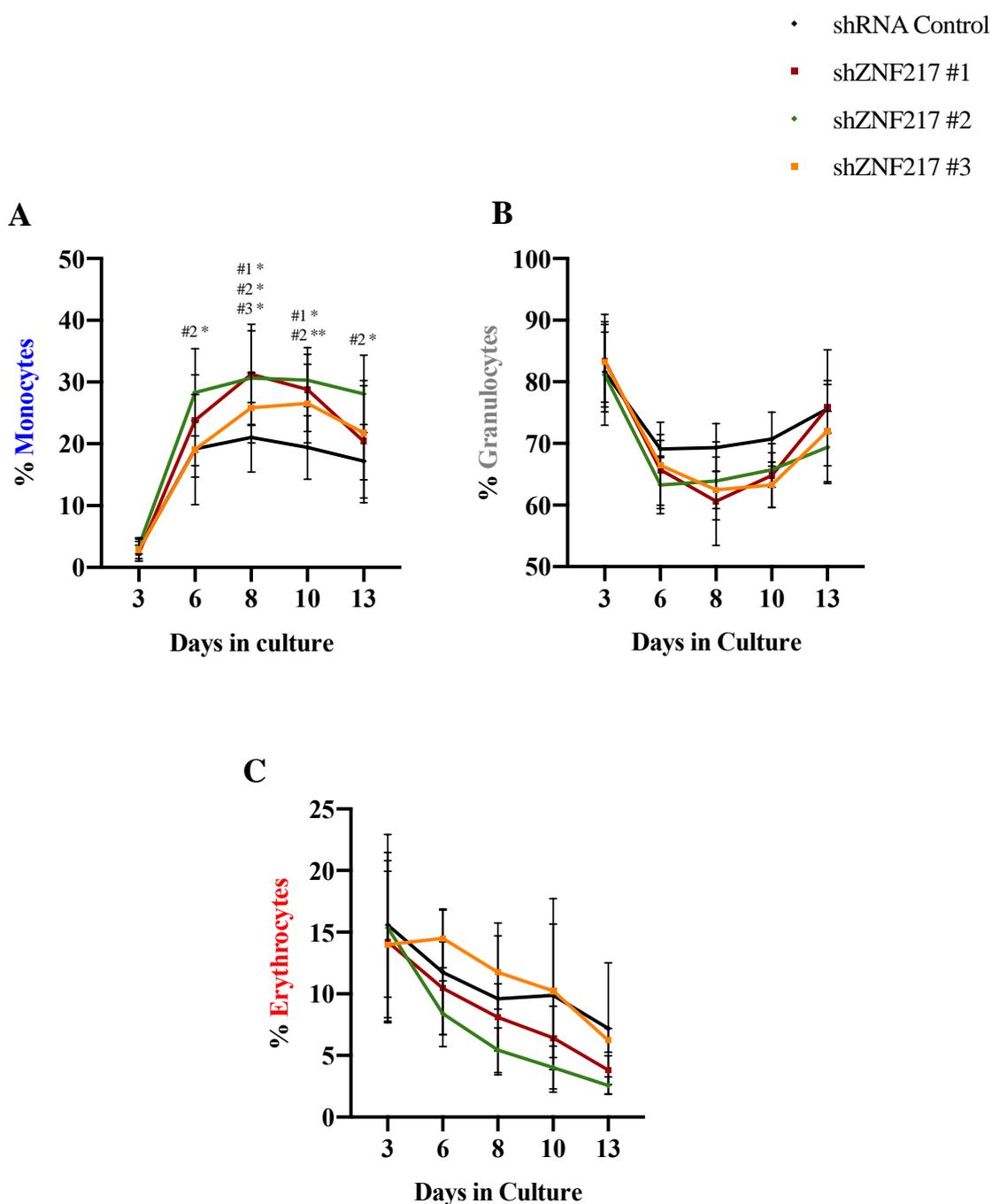
To determine the consequences of reducing endogenous levels of ZNF217 in CD34<sup>+</sup> HSPC, cells were cultured as described above (**4.3.3.3**) and subsequent growth, differentiation, and survival of control and ZNF217 KD cultures were analysed throughout 13 days of culture. Having established that ZNF217 overexpression suppressed the growth of HSPC (**Figure 4.9**), this study hypothesised that reducing ZNF217 expression would increase the proliferative ability of these cells. **Figure 4.18** shows the effect of ZNF217 KD on myeloid cell growth, representing two biological replicates (see **4.4.3** for more details). KD of ZNF217 in these cells failed to induce significant differences in the growth of cells in bulk liquid culture (**Figure 4.18A**). As previously described, the CD13 and CD36 cell surface markers were used to discriminate between monocytic, granulocytic and erythroid populations, and to analyse the growth and differentiation of each lineage. KD of ZNF217 led to an increase in monocytic cell growth, in cultures transduced with shZNF217 #1 and #3, by 1.5- and 2.5-fold on day 13, respectively (**Figure 4.18B**). No meaningful differences were noted in both granulocytic and erythroid growth (**Figure 4.18C-D**). Taken together, these results suggest that KD of ZNF217 disrupts the normal growth of monocytic cells, potentially associated with an immature phenotype.

Similarly to the approach described above (**4.3.3.4**), cell surface markers were used to analyse individual lineages. No significant changes were observed in terms of lineage balance regarding the granulocytic or erythroid lineages, following ZNF217 KD (**Figure 4.19B-C**). However, there was a significant increase in the proportion of monocytic cells in ZNF217 KD cultures as compared to control throughout culture (1.6-fold for shZNF217 #2, on day 13) (**Figure 4.19A**).



**Figure 4.18 – Knockdown of ZNF217 promotes monocytic cell growth in myeloid development**

(A) Cumulative fold-expansion of control and three ZNF217-KD constructs in terms of GFP positivity in culture medium containing IL-3, SCF, G- and GM-CSF, grown over 13 days (**Supplementary Figure 9**). (B) Cumulative fold-expansion of monocytic cells (CD13<sup>+</sup> CD36<sup>+</sup>) transduced with a control or ZNF217-KD vectors. (C) Cumulative fold-expansion of granulocytic cells (CD13<sup>+/−</sup> CD36<sup>−</sup>) transduced with a control or ZNF217-KD vector. (D) Cumulative fold-expansion of erythroid cells (CD13<sup>−</sup> CD36<sup>+</sup>) transduced with a control or ZNF217-KD vector. Cells were re-seeded at the appropriate density following each time-point, as described in 2.3.3. Gating strategies were applied as described in **Figure 2.4** and **Figure 2.5**. Data indicates Mean  $\pm$  1SD (n=2) (no statistical test was employed); (this data has also been represented as dot plots, within the supplementary section [**Supplementary Figure 34**]).



**Figure 4.19 – ZNF217 knockdown disrupts the balance between the monocytic and granulocytic populations in CD34<sup>+</sup> HSPC during myeloid cell development**

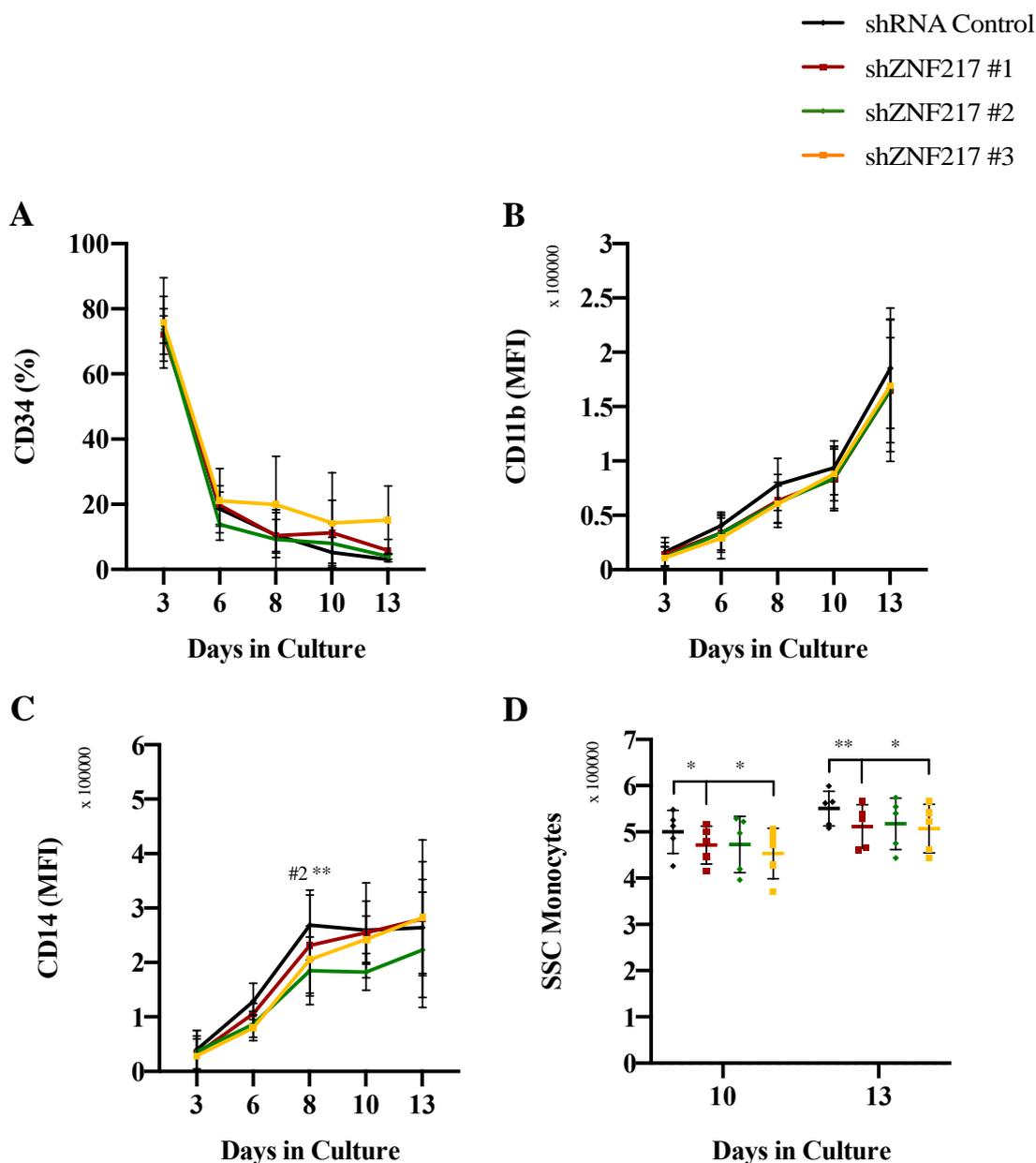
Summary data showing percentage of (A) monocytic (CD13<sup>+</sup> CD36<sup>+</sup>), (B) granulocytic (CD13<sup>+</sup> CD36<sup>-</sup>) and (C) erythroid cells (CD13<sup>-</sup> CD36<sup>+</sup>) in control and three ZNF217-KD cultures. Cells were re-seeded at the appropriate density following each time-point, as described in 2.3.3. Gating strategies were applied as described in Figure 2.4 and Figure 2.5. Data indicates Mean ± 1SD (n ≥ 4). Significant differences between shZNF217 cultures and control at each time-point were analysed by one-way ANOVA with Bonferroni's multiple comparisons correction; \* denotes p < 0.05; \*\* denotes p < 0.01; (this data has also been represented as dot plots, within the supplementary section [Supplementary Figure 35]).

Following the assessment of the proportion of cells committed to each lineage in each condition, the effect of ZNF217 KD on the normal differentiation of monocytes and granulocytes was further assessed, using the cell surface markers previously described (4.3.3.4). **Figure 4.20** shows the effect of ZNF217 KD in the monocytic-committed cells. No significant changes were observed for any of the parameters analysed, including expression of CD34, CD11b and CD14. In contrast, a significant decrease in cell granularity was observed in ZNF217 KD cells, on days 10 and 13 of cell culture. However, even though this was statistically significant, biologically this was very modest and unlikely to show a real effect. Similarly, no significant differences were observed regarding granulocytic development upon KD of ZNF217 (**Figure 4.21**)

Altogether, these results indicate that KD of ZNF217 impairs the colony forming ability of myeloid cells; however, it does not play an endogenous role in myeloid development, and the reduced monocytic growth is not a result of altered differentiation. Moreover, even though ZNF217 expression was shown to promote monocytic differentiation, expression of this TF is not essential for monocytic development.

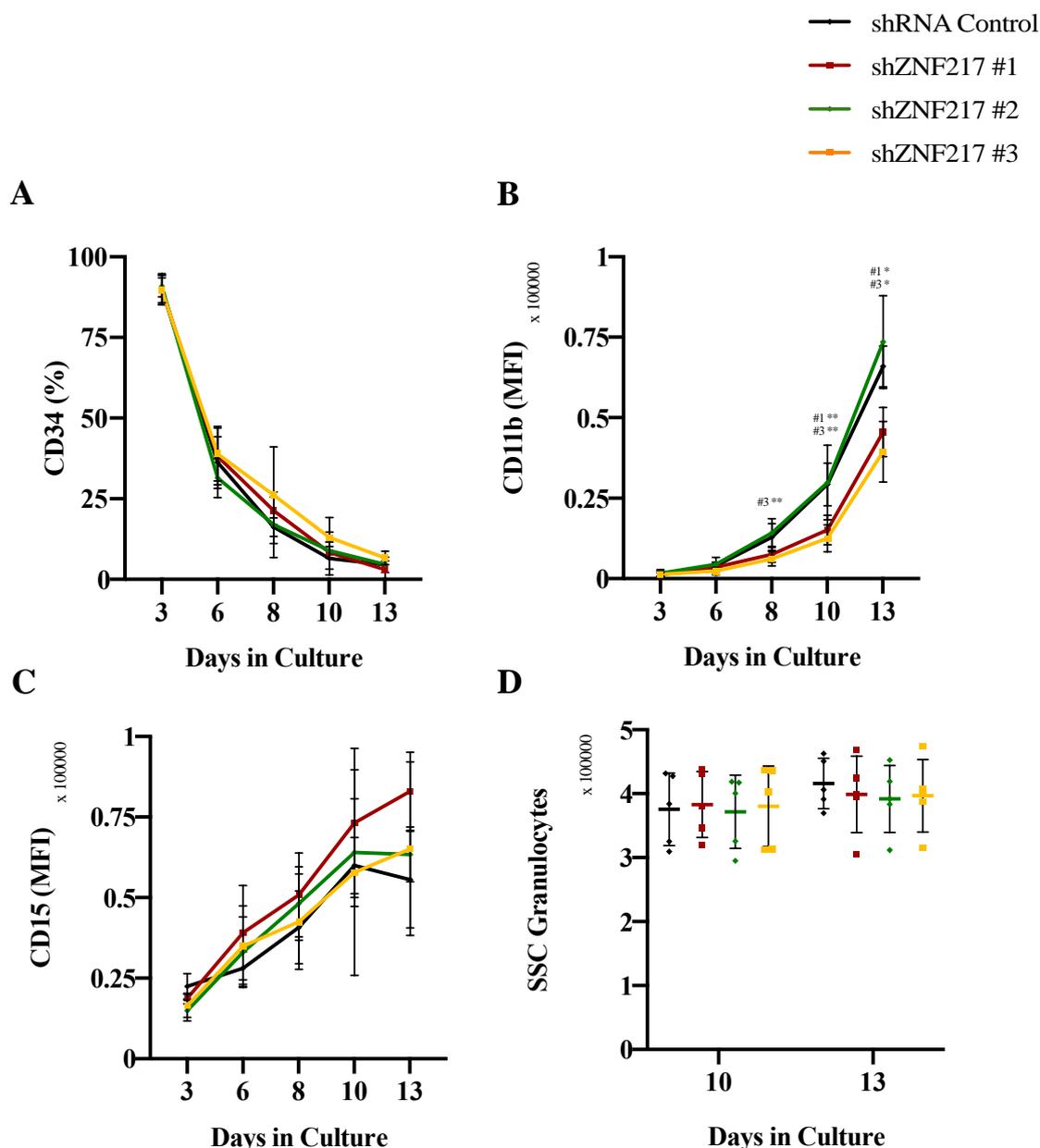
#### 4.3.4.4 *Knockdown of ZNF217 does not affect cell cycle and apoptosis*

Lastly, to determine the effects of ZNF217 KD on cell cycle, this study analysed FACSorted transduced HSPC, on day 6 of culture. No difference between control and ZNF217 KD cells, was observed (**Figure 4.22A**). I next determined if the decrease in the cells' colony forming ability was a result of increased apoptosis. ZNF217-KD cultures showed no difference in apoptosis on days 10 and 13 of development. These results indicate that the impact of ZNF217 KD on the cells colony forming ability and growth is not supported by changes in cells' cell cycle or apoptotic events (**Figure 4.22B**).



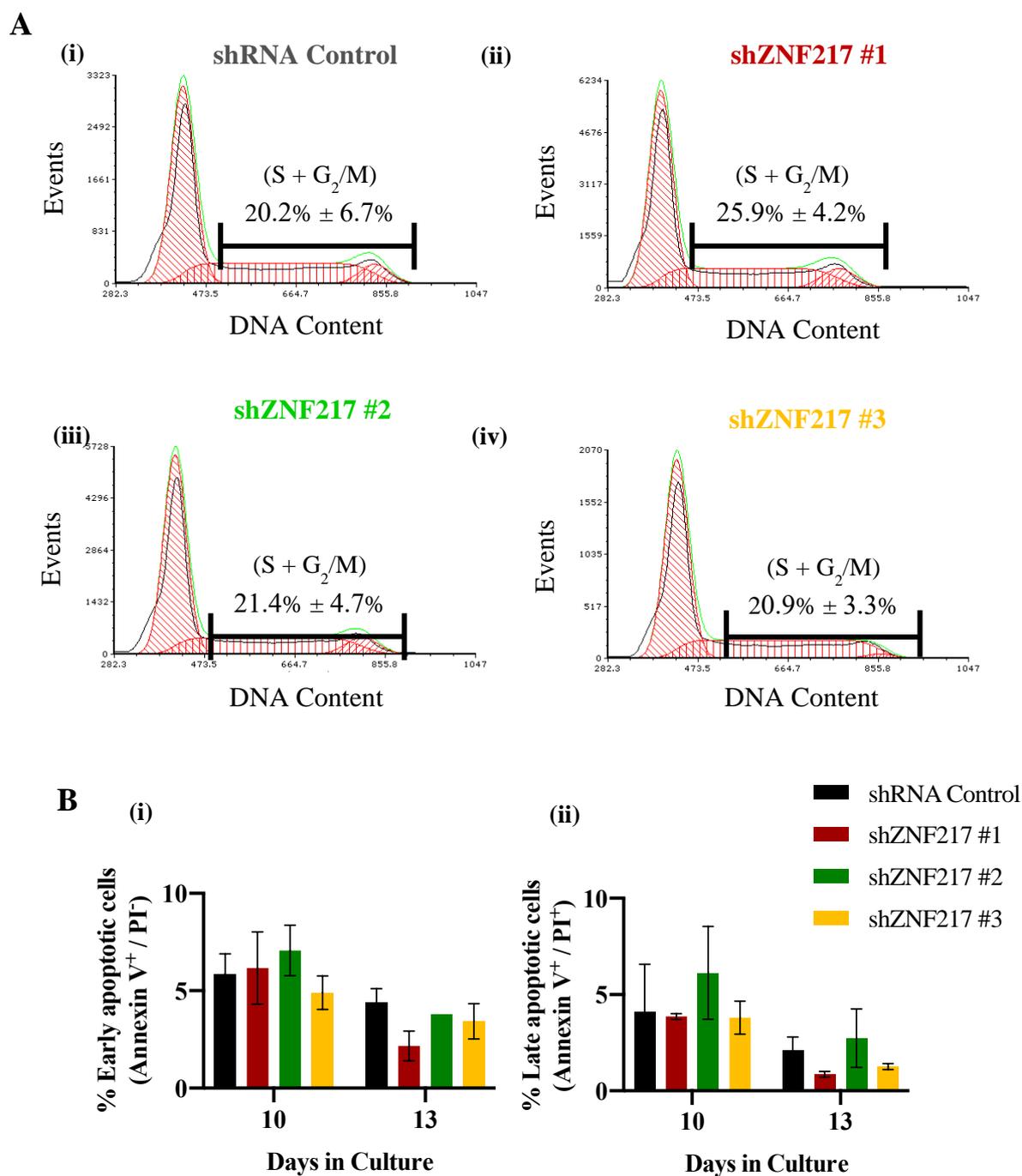
**Figure 4.20 – Knockdown of ZNF217 does not significantly alter monocytic development**

(A) Summary data of CD34<sup>+</sup> expression in terms of percentage in monocytic cells (CD13<sup>+</sup> CD36<sup>+</sup>) over time for control and ZNF217-KD cells. Summary data showing (B) CD11b and (C) CD14 expression in terms of MFI in monocytic cells over time for control and ZNF217-KD cells. (D) Complexity of monocytic cells was measured by changes in SSC for control and ZNF217-KD cultures on days 10 and 13 of myeloid differentiation. Cells were re-seeded at the appropriate density following each time-point, as described in 2.3.3. Gating strategies were applied as described in Figure 2.4 and Figure 2.5. Data indicates Mean  $\pm$  1SD (n=4). Significant differences between shZNF217 cultures and control at each time-point were analysed by one-way ANOVA with Bonferroni's multiple comparisons correction; \* denotes  $p < 0.05$ ; \*\* denotes  $p < 0.01$ ; (for representative flow cytometry plots please refer to Supplementary Figure 10, Supplementary Figure 11, Supplementary Figure 12; this data has also been represented as dot plots, within the supplementary section [Supplementary Figure 36]).



**Figure 4.21 – Knockdown of ZNF217 does not significantly alter granulocytic development**

(A) Summary data of CD34<sup>+</sup> expression in terms of percentage in granulocytic cells (CD13<sup>±</sup> CD36<sup>-</sup>) over time for control and ZNF217-KD cells. Summary data showing (B) CD11b and (C) CD15 expression in terms of MFI in granulocytic cells over time for control and ZNF217-KD cells. (D) Complexity of granulocytic cells was measured by changes in SSC for control and ZNF217-KD cultures on days 10 and 13 of myeloid differentiation. Cells were re-seeded at the appropriate density following each time-point, as described in 2.3.3. Gating strategies were applied as described in Figure 2.4 and Figure 2.5. Data indicates Mean ± 1SD (n≥4). Significant differences between shZNF217 cultures and control at each time-point were analysed by one-way ANOVA with Bonferroni's multiple comparisons correction; \* denotes p<0.05; \*\* denotes p<0.01; (for representative flow cytometry plots please refer to Supplementary Figure 13, Supplementary Figure 14, Supplementary Figure 15; this data has also been represented as dot plots, within the supplementary section [Supplementary Figure 37]).



**Figure 4.22 – Knockdown of ZNF217 has no effect on normal cell cycle or apoptosis**

(A) Representative histograms showing cell cycle distribution of normal HSPC (Control) and three ZNF217-knockdown cultures after 6 days of culture. Initial gating was applied as described in **Figure 2.6**. The percentage of cells in S + G<sub>2</sub>/M is indicated by mean ± 1SD (n=3). (B) Summary data showing the effect of ZNF217 knockdown in apoptosis, on HSPC following 10 and 13 days of culture. Initial gating was applied as described in **Figure 2.7**; (i) Early apoptotic cells are characterised by Annexin V<sup>+</sup> / PI<sup>-</sup> and (ii) Apoptotic cells are characterised by Annexin V<sup>+</sup> / PI<sup>+</sup>. Data indicates Mean ± 1SD (n=3).

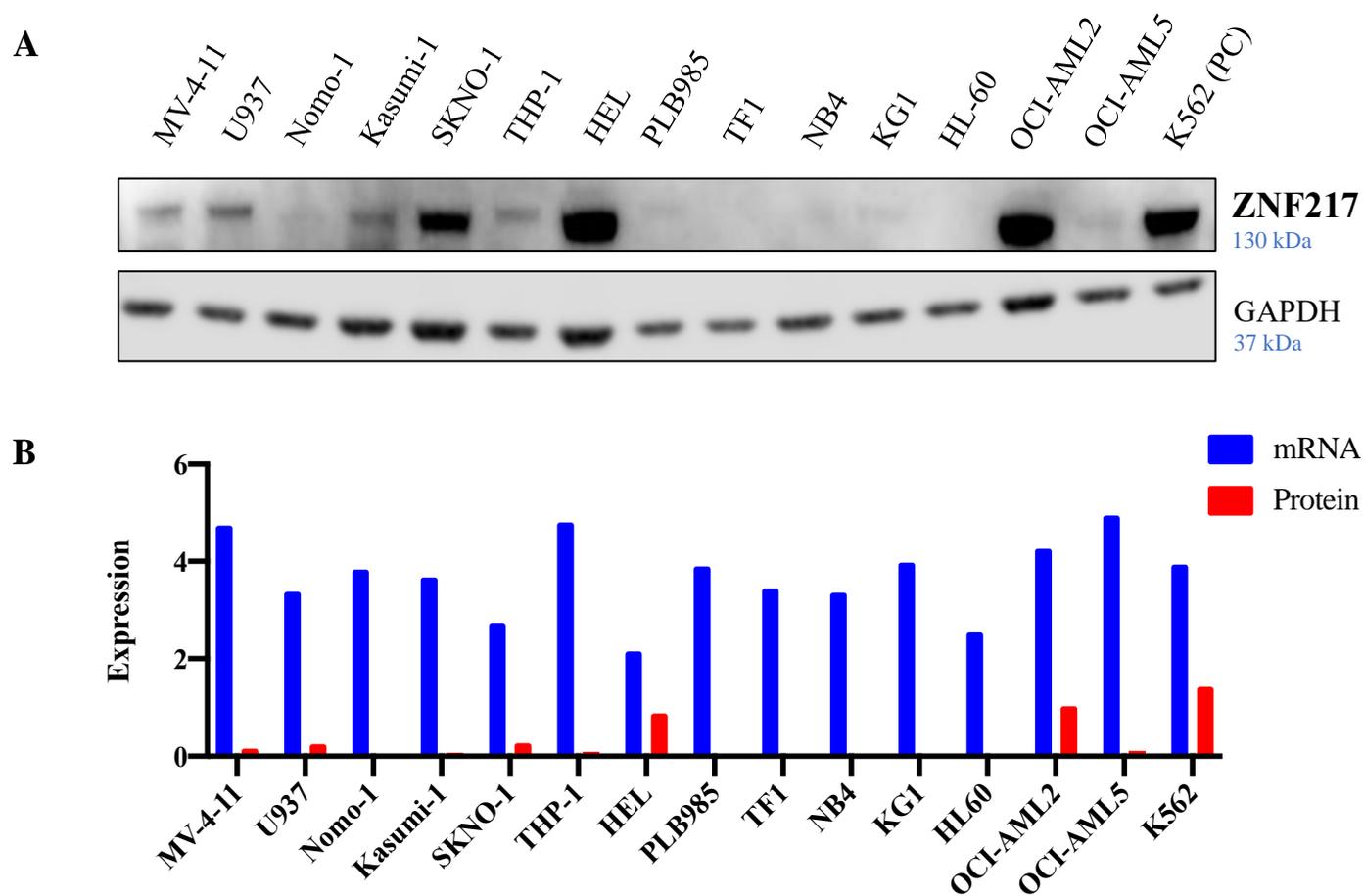
### 4.3.5 The role of ZNF217 in AML cell lines

#### 4.3.5.1 ZNF217 expression is variable across AML cell lines

Using the Cancer Cell Line Encyclopaedia (CCLE) (Barretina *et al.*, 2012), the expression of *ZNF217* mRNA in a cohort of AML cell lines was determined. **Figure 4.3A** shows *ZNF217* transcript is highly expressed across all lines, including in the t(8;21) cell lines Kasumi-1 and SKNO-1. However, previous studies have reported that ZNF217 protein expression does not correlate with *ZNF217* gene amplification levels (Li *et al.*, 2015; Szczyrba *et al.*, 2013; Renner *et al.*, 2013; Etcheverry *et al.*, 2010). For this reason, this study determined ZNF217 protein expression by western blot. ZNF217 protein expression level was shown to be variable across the AML cell lines tested. The expression level of ZNF217 was undetectable in Nomo-1, PLB985, TF1, NB4, HL-60 and OCI-AML5 cells but was strongly expressed in SKNO-1, HEL, OCI-AML2 and K562 cells (**Figure 4.3B**) making these cell lines suitable for overexpression and knockdown studies respectively. Additionally, ZNF217 was heterogeneously expressed in t(8;21) cell lines, showing a substantially lower expression in the Kasumi-1 cell line, as compared to SKNO-1; furthermore, there was no distinct pattern of expression between t(8;21) cell lines and non-t(8;21) cell lines.

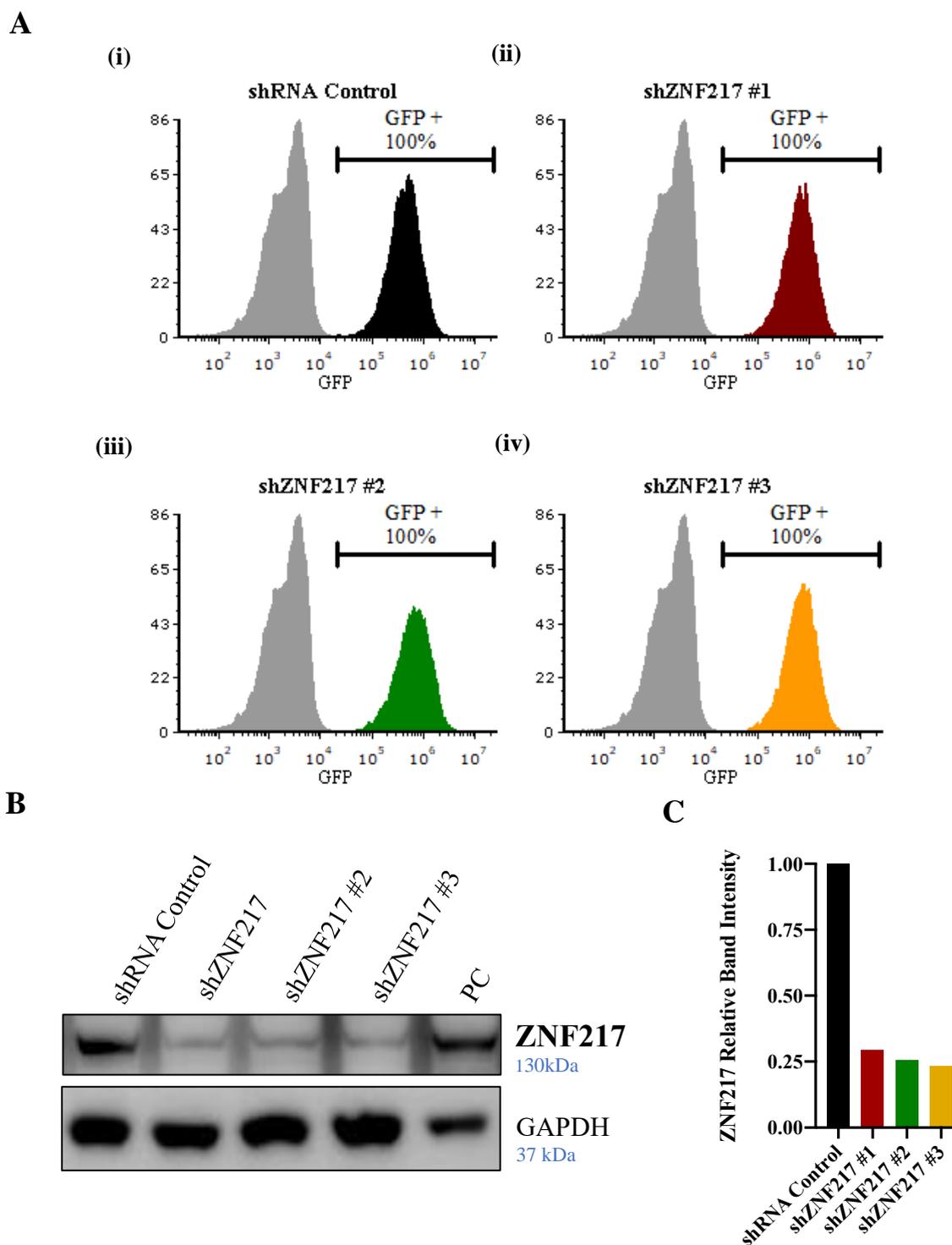
#### 4.3.5.2 Generation of ZNF217-knockdown AML cell lines

Using the same shRNA constructs (**4.3.4**), ZNF217 was knocked-down in the leukaemic cell lines HEL and SKNO-1. which presented the highest levels of endogenous ZNF217 protein. Transfected cell lines were selected using puromycin, confirmed through GFP positivity (**Figure 4.24-Figure 4.25A**) and validated by western blot (**Figure 4.24-Figure 4.25B-C**). Different levels of knockdown were achieved based on the cell line used. In the HEL cell line, all three constructs used were shown to possess a similar level of KD, with endogenous levels of ZNF217 reduced by c75%. In the SKNO-1 line, however, shZNF217 #1 had the weakest level of KD, with a modest c40% reduction in ZNF217 expression; the other two constructs showed similar reduction levels to HEL cells.



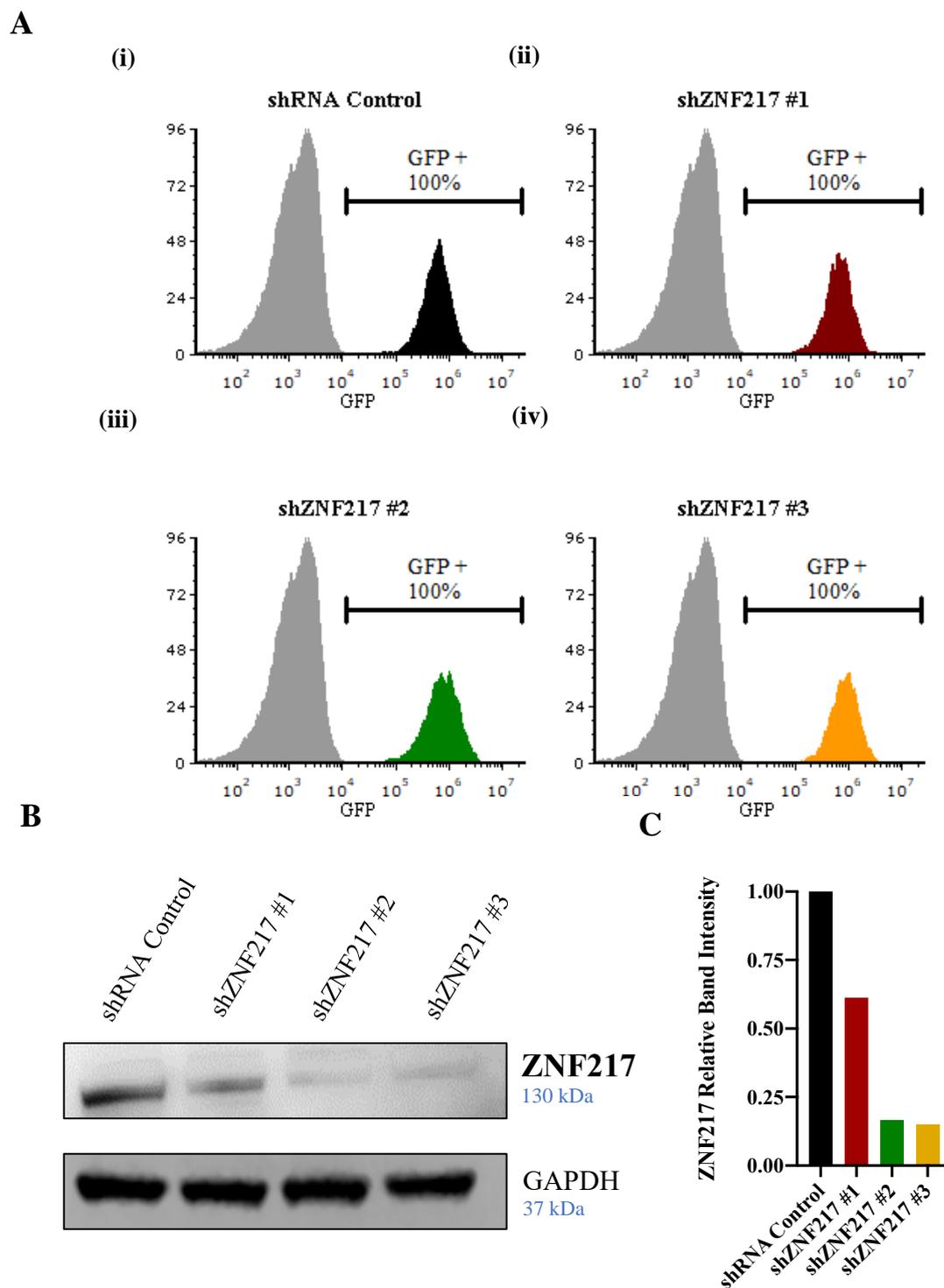
**Figure 4.23 – ZNF217 expression in leukaemic cell lines**

(A) Western blot analysis of ZNF217 endogenous expression in a cohort of AML cell lines (n=1). The K562 cell line was used as a positive control (PC); GAPDH was used as a loading control. (B) Normalised (log<sub>2</sub>) ZNF217 mRNA expression in a cohort of leukaemic cell lines. mRNA expression data derived from <https://portals.broadinstitute.org/ccle/page?gene=ZNF217> (Barretina *et al.*, 2012). Null protein values refer to undetectable levels of endogenous ZNF217.



**Figure 4.24 – ZNF217 is knocked-down in the erythroleukaemia cell line HEL**

(A) Representative flow cytometric plots showing transduction efficiency in the HEL cell line transfected with (i) an shRNA control and (ii-iv) three ZNF217 knockdown constructs (shZNF217 #1, shZNF217 #2, shZNF217 #3), following puromycin selection. (B) Western blot analysis showed successful ZNF217 knockdown in all three shRNA constructs (n=1). GAPDH was used as a loading control.; K562 cells were used as a positive control (PC) (C) Western blot quantification by densitometry analysis using ImageJ. Relative protein expression was determined by normalising each sample firstly to the loading control, then to shRNA control (n=1).



**Figure 4.25 – ZNF217 is knocked-down in the t(8;21) cell line SKNO-1**

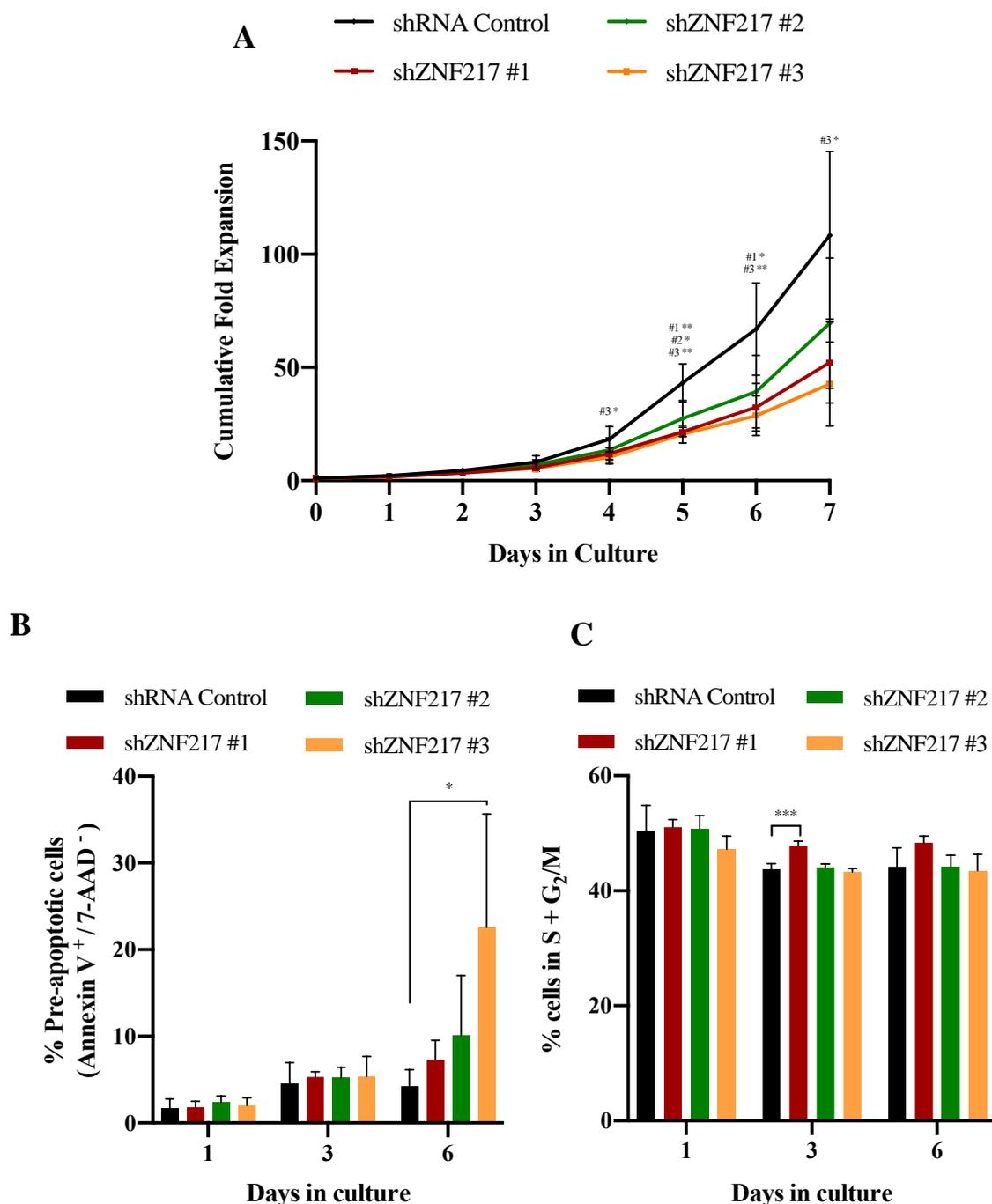
(A) Representative flow cytometric plots showing transduction efficiency in the SKNO-1 cell line transfected with (i) an shRNA control and (ii-iv) three ZNF217 knockdown constructs (shZNF217 #1, shZNF217 #2, shZNF217 #3), following puromycin selection. (B) Western blot analysis showed successful ZNF217 knockdown in all three shRNA constructs (n=1). GAPDH was used as a loading control. (C) Western blot quantification by densitometry analysis using ImageJ. Relative protein expression was determined by normalising each sample firstly to the loading control, then to shRNA control (n=1).

#### 4.3.5.3 *ZNF217 knockdown in AML cell lines reduces cell growth by promoting changes in cell cycle and apoptosis*

I next determined the effect of ZNF217 KD on leukaemic cell growth. Growth was significantly reduced in the HEL cell line following day 4 of culture (**Figure 4.26A**), by 1.5-fold (with shZNF217 #3). This reduction in growth remained consistently lower until day 7, in which a 2.1- and 2.5-fold reduction was observed upon transduction with shZNF217 #1 and #3, respectively. Similarly, even though not statistically significant, a reduction in cell growth was also observed in the SKNO-1 cell line (**Figure 4.27A**).

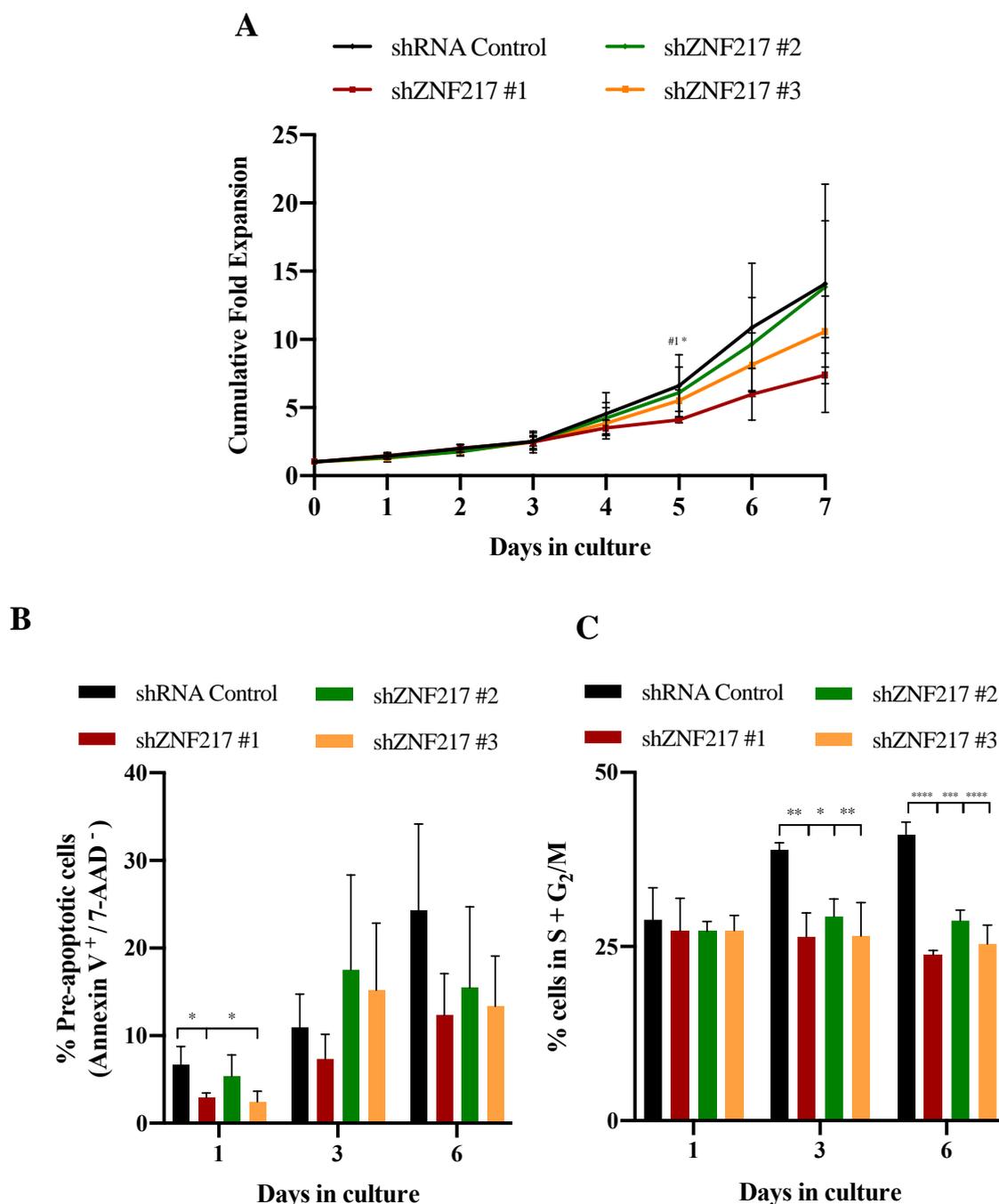
In order to determine the reason behind this decreased proliferative ability, these lines were assayed for apoptosis using Annexin V binding in combination with 7-AAD exclusion. In HEL cells, the decrease in growth observed for shZNF217 #3 can be attributed to an increase in pre-apoptotic cells, on day 6 of culture, by 5-fold, as compared to control (**Figure 4.26B**). However, an analysis of the apoptotic rate on day 4 of culture would be necessary to further strengthen these observations. No significant differences were observed in SKNO-1 cells (**Figure 4.27B**). These data suggest that ZNF217 expression performs an anti-apoptotic role in HEL cells.

Additionally, the cell cycle distribution of these cells was analysed using the DNA dye PI. In the HEL cell line, ZNF217 KD did not lead to significant changes in the normal cells' cell cycle (**Figure 4.26C**). However, loss of this protein resulted in substantial changes in SKNO-1 cells' cell cycle, with a significant increase in the proportion of cells in S+G<sub>2</sub> phase, throughout culture (**Figure 4.27C**). Taken together, the decrease in proliferative potential observed upon ZNF217 KD can be a result of not only an increase in apoptosis, but also due to an arrest in the (S + G<sub>2</sub>) phase, potentially enhancing cell death, suggesting ZNF217 is necessary for AML cells' proliferation and survival. However, it is suggested that this effect occurs independently of RUNX1-ETO expression.



**Figure 4.26 – ZNF217 knockdown significantly decreases HEL cell proliferation**

Summary data showing (A) the growth of HEL cells with three ZNF217 knockdown constructs compared to control over 7 days of growth. Data indicates Mean  $\pm$  1SD (n=5) (2-way ANOVA statistical test with Bonferroni's multiple comparison test - \*\* denotes  $p < 0.01$ ; \*\*\* denotes  $p < 0.001$ ; \*\*\*\* denotes  $p < 0.0001$ ); (B) the effect of ZNF217 knockdown on Annexin V staining in HEL cells following 1, 3 and 6 days of cell growth. Bars indicate percentage of pre-apoptotic cells (Annexin V<sup>+</sup> / 7-AAD<sup>-</sup>). Data indicates Mean  $\pm$  1SD (n=3) (One-way ANOVA statistical test with Bonferroni's multiple comparison test - \* denotes  $p < 0.05$ ); (C) the effect of ZNF217 knockdown in the HEL cells' cell cycle following 1, 3 and 6 days of cell growth. Data indicates Mean  $\pm$  1SD (n=3) (One-way ANOVA statistical test with Bonferroni's multiple comparison test - \* denotes  $p < 0.05$ ; \*\* denotes  $p < 0.01$ ; \*\*\* denotes  $p < 0.001$ ; \*\*\*\* denotes  $p < 0.0001$ ).



**Figure 4.27 – ZNF217 knockdown significantly decreases SKNO-1 cell proliferation**

Summary data showing (A) the growth of SKNO-1 cells with three ZNF217 knockdown constructs compared to control over 7 days of growth. Data indicates Mean  $\pm$  1SD (n=5) (2-way ANOVA statistical test with Bonferroni's multiple comparison test - \*\* denotes  $p < 0.01$ ; \*\*\* denotes  $p < 0.001$ ; \*\*\*\* denotes  $p < 0.0001$ ); (B) the effect of ZNF217 knockdown on Annexin V staining in SKNO-1 cells following 1, 3 and 6 days of cell growth. Bars indicate percentage of pre-apoptotic cells (Annexin V<sup>+</sup> / 7-AAD<sup>-</sup>). Data indicates Mean  $\pm$  1SD (n=3) (One-way ANOVA statistical test with Bonferroni's multiple comparison test - \* denotes  $p < 0.05$ ); (C) the effect of ZNF217 knockdown in the SKNO-1 cells' cell cycle following 1, 3 and 6 days of cell growth. Data indicates Mean  $\pm$  1SD (n=3) (One-way ANOVA statistical test with Bonferroni's multiple comparison test - \* denotes  $p < 0.05$ ; \*\* denotes  $p < 0.01$ ; \*\*\* denotes  $p < 0.001$ ; \*\*\*\* denotes  $p < 0.0001$ ).

## 4.4 Discussion

ZNF217 is a TF belonging to the Kruppel-like family of TF. The gene is located on chromosome 21q13, in a region frequently amplified in various human cancers, including breast and ovarian, as compared to normal tissue and epithelial cells (Collins *et al.*, 1998; Littlepage *et al.*, 2012). Whilst several studies have been performed in solid tumours, the role of ZNF217 in the haematopoietic process and in the development of AML has not yet been described. The main aim of this chapter was to describe the normal expression of ZNF217 during haematopoiesis and its role in normal human myeloid development. Overexpression of ZNF217 in HSPC was found to significantly inhibit the growth of myeloid progenitor cells, associated with an increase in their differentiation potential, therefore suggesting that ZNF217 promotes myeloid development. Moreover, KD of ZNF217 had no impact in both monocytic and granulocytic maturation, suggesting that, even though expression of ZNF217 promotes differentiation, it is not essential for this process to occur.

To understand whether overexpression of *ZNF217* might contribute to leukaemic development, this study aimed at analysing its expression level during normal haematopoiesis. *ZNF217* gradually increases as cells differentiate into mature granulocytes and, to a lesser extent, monocytes. Expression of the mouse homologue *Zfp217* revealed a similar trend, with mature monocytes and granulocytes expressing the higher levels of *Zfp217* as compared to HSC. However, a study based on embryonic stem cells identified *Zfp217* (homologue) as essential for maintaining the undifferentiated state cells present, associated with epigenetic regulation and m<sup>6</sup>A RNA modifications (Aguilo *et al.*, 2015). These high expression levels tightly correlate with the regulation of several pluripotency genes, including *NANOG*, *POU5F1* and *SOX2* and upregulation of these has previously been described in other cancers (Aguilo *et al.*, 2015). For instance, *SOX2* upregulation has not only been identified in several cancers, but it also correlated with poor disease prognosis (Tam and Ng, 2014; Weina and Utikal, 2014; Wuebben and Rizzino, 2017). Ultimately, there are few studies regarding the role of ZNF217 in the haematopoietic process and often have conflicting data. These might be related to different models/techniques used.

#### 4.4.1 High *ZNF217* in AML is associated with increased risk of relapse

Analysis aimed at exploring the association between *ZNF217* expression and leukaemogenesis showed that the expression of *ZNF217* mRNA is highly variable across AML subtypes, but is generally lower in AML as compared to fully mature granulocytes, suggesting that *ZNF217* is promoting a block in granulocytic differentiation.

To determine the relationship between *ZNF217* expression and clinical attributes, the TCGA dataset was used to stratify patients according to high and low expression quartiles (Ley *et al.*, 2013). Disease-free survival was found to be significantly lower in patients with high levels of *ZNF217*, indicating that these patients possess a higher risk of relapse. Similar observations have previously been made in solid tumours. Overexpression of *ZNF217* levels was observed in both primary prostate carcinoma (Szczyrba *et al.*, 2013) and colorectal carcinoma (Zhang *et al.*, 2015), the latter further associating with poor clinical features. Similar results were observed in glioma samples, in which high *ZNF217* was found to relate to poor outcome (Mao *et al.*, 2011). However, the majority of studies regarding overexpression of *ZNF217* in tumour development are associated with breast cancer, in which high *ZNF217* has been documented in breast tumour samples, as compared to normal epithelial cells (Krig *et al.*, 2010; Littlepage *et al.*, 2012; Nguyen *et al.*, 2014; Nonet *et al.*, 2001; Plevova *et al.*, 2010; Thollet *et al.*, 2010; Vendrell *et al.*, 2012; Collins *et al.*, 1998). Similarly to the present study, analysis of breast cancer patients showed that high levels of *ZNF217* were associated with decreased relapse-free survival (Vendrell *et al.*, 2012; Littlepage *et al.*, 2012; Nguyen *et al.*, 2014), further correlating with poor OS (Littlepage *et al.*, 2012; Nguyen *et al.*, 2014) and prognosis (Nguyen *et al.*, 2014; Vendrell *et al.*, 2012), and with the development of metastases (Vendrell *et al.*, 2012). Equally, high levels of *ZNF217* protein have been identified in cervical (Zhu *et al.*, 2009) and colorectal (Zhang *et al.*, 2015; Li *et al.*, 2015) cancers, as well as gastric carcinoma (Shida *et al.*, 2014). These were found to be associated with more aggressive disease subtypes (Zhang *et al.*, 2015; Li *et al.*, 2015), as well as poorer OS (Li *et al.*, 2015) and relapse-free survival (Shida *et al.*, 2014). Given that this study identified *ZNF217* as overexpressed in RUNX1-ETO-expressing cells, it is hypothesised that the block in normal cell differentiation might be mediated by *ZNF217*.

#### 4.4.2 ZNF217 overexpression of promotes myeloid development in HSPC

Previous studies have demonstrated that increased levels of ZNF217 contribute to tumour development in several cancers, including breast and ovarian (Littlepage *et al.*, 2012; Nonet *et al.*, 2001; Plevova *et al.*, 2010; Li *et al.*, 2014; Li *et al.*, 2007). However, its role in the development of leukaemia remains largely unknown. A recent study attempting to unravel the role of this TF in paediatric B-ALL has identified ZNF217 as a crucial mediator of the oncogenic role observed in these patients, and essential for the survival of B-ALL cells (Qin *et al.*, 2019).

To determine whether ZNF217 equally contributes to the pathogenesis of AML by disrupting normal myeloid development, functional studies were performed. To determine if overexpression ZNF217 in HSPC affected haematopoiesis and cell survival, this study performed a colony forming assay. ZNF217 was shown to significantly reduce myeloid colony formation under clonal conditions, suggesting that this TF influences the proliferative ability of these cells. A subsequent replating assay yielded similar results, in which ZNF217-overexpressing cells generated significantly less colonies as compared to control cultures. These observations indicate that expression of ZNF217 in HSPC inhibits colony formation and decreases the self-renewal potential of myeloid progenitor cells, potentially associated with a more mature phenotype. Previous studies have shown that overexpression of ZNF217 in primary mammary epithelial cells resulted in enhanced formation of mammospheres associated with an increased self-renewal ability (Vendrell *et al.*, 2012; Littlepage *et al.*, 2012; Nguyen *et al.*, 2014). However, this study's results contrast with these observations and suggest, instead, that ZNF217 promotes HSPC differentiation in this system. However, once more, the models used to perform these experiments were not equivalent.

Subsequently, this study analysed the consequences of ZNF217 overexpression on myeloid development. In this context, overexpression of ZNF217 significantly reduced HSPC growth, especially within the monocytic and granulocytic populations, suggesting that ZNF217 does not act as an oncogene in a haematopoietic setting. ZNF217 has been previously associated with increased proliferation of ovarian and breast cancer cells *in vitro* (Li *et al.*, 2014; Thollet *et al.*, 2010), with similar observations demonstrated *in vivo*, where constitutive expression of this TF promoted the growth and rate of tumour formation (Vendrell *et al.*, 2012; Li *et al.*, 2014; Thollet *et al.*, 2010; Littlepage *et al.*, 2012), which this study did not observe. A reduction in cell growth can be a result of several factors, including changes in cell cycle, increased

apoptotic rate or differentiation processes. In other models, the increased proliferation of ZNF217-overexpressing cells was accompanied by an increase in the S phase of the cell cycle (Li *et al.*, 2014). In this model, the decreased growth of the cells was actually associated with a significant reduction in the proportion of cells in the S / G<sub>2</sub> phase, suggesting an arrest in normal cell cycle arising from the expression of ZNF217.

The consequences on cell differentiation upon ZNF217 expression were assessed through the expression of both monocytic and granulocytic maturation markers. Overexpression of ZNF217 promoted significant increases in the expression of monocytic markers CD11b and CD14, accompanied by a significant reduction in CD34 in early time points, indicating that the reduced proliferation of monocytic progenitors is a result of the differentiation process induced by the expression of ZNF217. The impact on granulocytic development, on the other hand, was less clear. Even though expression of ZNF217 resulted in the upregulation of CD11b in these cells, increase of CD15 was not observed, hence drawing a meaningful conclusion is not possible and additional analysis of other cell surface markers would need to be conducted. Altogether, these results suggest that ZNF217 is mainly involved in monocytic differentiation and does not contribute to leukaemic development. Additionally, these results do not agree with the analysis described above (4.4.1), which suggested an important role for ZNF217 in granulocytic development. Moreover, a role for ZNF217 in differentiation directly contrasts with preceding studies performed in other tumours. For instance, overexpression of ZNF217 in breast cancer cells was found to promote the repression of an adult stem cell expression signature (Littlepage *et al.*, 2012). Additionally, expression of this TF in CSC was shown to be associated with the dysregulation of genes involved in the differentiation and maintenance of these cells (Krig *et al.*, 2007; Vendrell *et al.*, 2012; Littlepage *et al.*, 2012). In summary, overexpression of ZNF217 in HSPC resulted in an increased differentiation potential, associated with a reduced proliferative ability and colony forming efficiency, leading to the hypothesis that overexpression of this TF is actually promoting haematopoietic development and not contributing the leukaemogenesis. These observations have been summarised in **Table 4.1**.

**Table 4.1 - Summary of the main findings regarding the role of ZNF217 in normal human haematopoiesis**

Table summary detailing the consequences of ZNF217 modulation vs control.

\* denotes significant differences in ZNF217 overexpression / knockdown, as compared to normal control cells.

	ZNF217 Overexpression	ZNF217 Knockdown
<b>Monocytes</b>		
<b>Growth</b>	<b>Decrease *</b>	No change
<b>Markers</b>	<b>Increase in CD11b and CD14 *</b>	No change
<b>Granulocytes</b>		
<b>Growth</b>	<b>Decrease *</b>	No change
<b>Markers</b>	<b>Increase in CD11b and CD15 *</b>	No change
<b>Apoptosis</b>	No change	No change
<b>Cell Cycle</b>	<b>Increase in G1 *</b>	No change
<b>CFU</b>	<b>Decrease *</b>	<b>Decrease *</b>

#### 4.4.3 ZNF217 knockdown in HSPC does not impact myeloid development

To determine if ZNF217 was necessary for myeloid development, an shRNA-based approach was used by lentivirally transducing HSPC with a scrambled shRNA control or three ZNF217 shRNA constructs. Since overexpression of ZNF217 promoted monocytic and granulocytic differentiation, it was hypothesised that knocking it down in the same cells would lead to a block in normal myeloid development associated with an immature phenotype. KD of ZNF217 significantly inhibited myeloid colony formation, associated with a decreased self-renewal potential. These suggest that loss of ZNF217 does not promote the retention of an immature phenotype.

The effects of ZNF217 KD on myeloid development were also assessed in liquid bulk culture. Significant differences in the overall growth were observed, with HSPC derived from frozen units presenting a 10-fold increase in growth, as compared to cells originated from fresh blood. For this reason, growth analysis solely reflects the results obtained from two sets of experiments, which did not allow statistical analysis. KD of ZNF217 was shown to exclusively impact monocytic growth, consistent with a role for this TF in monocytic development. Expression of cell surface markers showed minimal impact throughout myeloid development, suggesting that, even though ZNF217 might possess a role in myeloid differentiation, this is not critical for maturation to occur. These observations have been summarised in **Table 4.1**.

In conclusion, this study hypothesised the overexpression of ZNF217 observed in RUNX1-ETO expressing cells resulted in the disruption of normal myeloid development. This hypothesis was strengthened by the fact that overexpression of ZNF217 had previously been observed in several cancer types, and the assumption that ZNF217 played a key role in tumorigenesis by orchestrating tumour progression. This included the involvement of ZNF217 in the major hallmarks of cancer, including cell proliferation, replicative immortality, resistance to cell death or changes in cell cycle. However, this study found that, instead of possessing a pro-leukaemic phenotype, overexpression of ZNF217 in HSPC resulted in monocytic differentiation, inconsistent with previous studies. Moreover, even though KD of ZNF217 impacted monocytic growth, no impact was observed regarding their differentiation ability, suggesting that this TF is not essential for this process to occur.

# Chapter 5

---

## **Identification of C/EBP $\beta$ Suppression by RUNX1- ETO as a Potential Mediator of a Block in Differentiation**

## 5.1 Introduction

The t(8;21), responsible for the expression of the fusion protein RUNX1-ETO, occurs in approximately 12% of all AML cases. Ectopic expression of RUNX1-ETO in a primary human haematopoietic model has been shown to result in a block in myeloid development, associated with an increase in the cells' self-renewal potential (Tonks *et al.*, 2003; Tonks *et al.*, 2004) (1.3.3.4). A subsequent study by our group demonstrated that expression of this fusion gene leads to dysregulation in the transcriptome of HSPC (Tonks *et al.*, 2007). However, whilst mRNA studies are routinely used for target identification, it is not a powerful predictor of protein expression (Singh and Sharma, 2020). Therefore, this current study performed SWATH-MS to determine quantitative changes in the proteome of human CD34<sup>+</sup> HSPC expressing RUNX1-ETO compared to normal control cells (**Chapter 3**). Expression of RUNX1-ETO in these cells resulted in the dysregulated expression of several TF, including PU.1 and CBF $\beta$ , confirming previous studies (Vangala *et al.*, 2003; Kwok *et al.*, 2009; Roudaia *et al.*, 2009). However, RUNX1-ETO expression was also shown to result in the repression of C/EBP $\beta$ , a member of the C/EBP family of TF (**Figure 3.18**), an observation not previously reported.

Generally, members of the C/EBP family are viewed as tumour suppressor genes, due to their ability to repress cell growth and respond upon DNA damage events (1.5.1). However, in certain cell types, or depending on the protein isoform present, members of this family can possess opposing effects on the cells' proliferative ability. Dysregulation of C/EBP $\alpha$  has been demonstrated in several solid tumours, including breast and lung cancer (Lourenço and Coffey, 2017). Its role in hepatocellular carcinoma and hepatoblastoma, however, has been the subject of multiple studies, as discordant results have been observed. Whilst some studies show that expression of C/EBP $\alpha$  is downregulated in these cancer cells and is associated with a worst OS (Tomizawa *et al.*, 2003; Tseng *et al.*, 2009), other studies claim that a poorer OS is linked to its overexpression (Lu *et al.*, 2015; Gray *et al.*, 2000). This suggests that the role of C/EBP $\alpha$  is context dependent and may have a tumour suppressor or promoter role (at least in liver cancers). In other solid tumours, C/EBP $\beta$  has been shown to be frequently upregulated in aggressive types of breast cancer, correlating with poor prognosis, metastasis, and high tumour grade (van de Vijver *et al.*, 2002; Zahnow, 2009). This was further associated with an increased ratio of C/EBP $\beta$  LIP:LAP isoform ratio (Zahnow, 2009; Arnal-Estapé *et al.*, 2010) (1.5.3). C/EBP $\delta$ , on the other hand, is a tumour promoter in brain cancer, overexpressed in

glioblastoma cells and associated with poor prognosis (Cooper *et al.*, 2012; Carro *et al.*, 2010). *C/EBP $\delta$*  mRNA levels have been shown to be overexpressed in cancer cells expressing high levels of STAT3 which is frequently hyperactivated in cancer (Zhang *et al.*, 2007; Silva, 2004). Additionally, expression of *C/EBP $\gamma$*  has previously been shown to inhibit cellular senescence and promote cell proliferation in several solid tumours, further associating with poor clinical prognosis (Huggins *et al.*, 2013).

Despite their roles as tumour promoter genes, some members of the C/EBP family have been shown to be important in normal haematopoiesis, including myeloid development. For instance, *C/EBP $\alpha$*  has been shown to be important for granulocytic and monocytic differentiation (Suh *et al.*, 2006), and its dysregulation has been shown to block the transition from CMPs to GMPs, resulting in the loss of mature granulocytes (Zhang *et al.*, 1997). This was subsequently shown to be associated with an increase in proliferation, increase in number of functional LT-HSC, and repopulating ability (Zhang *et al.*, 2004b). *C/EBP $\beta$* , on the other hand, is essential for macrophage and B-cell differentiation (Ruffell *et al.*, 2009; Chen *et al.*, 1997). Additionally, myeloid differentiation is preceded by the overexpression of *C/EBP $\delta$* , followed by the upregulation of G-CSF and repression of C-Myc (Agrawal *et al.*, 2007; Wang and Friedman, 2002). In fact, mature granulocytes express high levels of *C/EBP $\delta$*  (Scott *et al.*, 1992; Gery *et al.*, 2005). This is similar to macrophages, in which this TF is responsible for the regulation of genes associated with M1 macrophage polarisation (Banerjee *et al.*, 2013). Lastly, *C/EBP $\epsilon$*  is exclusively expressed in myeloid cells, and is required for the terminal differentiation of both neutrophilic and eosinophilic cells; however, expression of this protein and different isoforms has the ability to reprogram myeloid lineage commitment and differentiation (Koike *et al.*, 1997).

Several studies have reported the dysregulation of members of the C/EBP family and the development of AML. Downregulation of *C/EBP $\alpha$*  has been previously shown to play a role in leukaemogenic development, and has been reported in AML, leading to a block in myeloid differentiation (Pabst *et al.*, 2001; Rosenbauer and Tenen, 2007; Perrotti *et al.*, 2002; Zheng *et al.*, 2004a). Furthermore, mutations in the *CEBPA* gene have equally been associated with the development of AML (Paz-Priel and Friedman, 2011). Hypermethylation of another member of this family, *C/EBP $\delta$* , has previously been associated with its low expression levels in AML (Agrawal *et al.*, 2007). No association has been established to date regarding loss of *C/EBP $\epsilon$*  and the development of AML, even though allelic loss of this TF was detected in AML patients

(Koike *et al.*, 1997). Lastly, whilst different C/EBP $\beta$  protein isoforms have been shown to possess multiple functions, relying on their ratio to each other, there is little information regarding the role of C/EBP $\beta$  in AML (1.5.4). C/EBP $\beta$  regulates macrophage function, as well as emergency granulopoiesis and neoplastic transformation (Hirai *et al.*, 2006; Tanaka *et al.*, 1995a; Natsuka *et al.*, 1992). In APL, C/EBP $\beta$  promotes ATRA-induced cell differentiation. These studies suggest that high levels of this protein result in transcriptional changes able to overcome the block in the differentiation process (Duprez *et al.*, 2003). Similarly, ectopic expression of C/EBP $\beta$  in BCR-ABL-expressing CML cells promotes granulocytic cell differentiation (Guerzoni *et al.*, 2006). However, the pathophysiological role of C/EBP $\beta$  in AML, including t(8;21), as well as its role in haematopoietic development remains largely unknown.

## 5.2 Hypothesis and Aims

This study hypothesises that downregulation of C/EBP $\beta$  expression will disrupt normal human myeloid development. This chapter's main aim is to understand the role of C/EBP $\beta$  in haematopoietic development and its possible contribution to the development of AML, using a normal human primary HSPC model, as well as AML cell lines. This will be achieved by performing the following objectives:

**Determine the expression of *CEBPB* during normal human myeloid development and in AML patient blasts.**

Analysis of publicly available transcriptomic datasets will be performed to determine *CEBPB* mRNA expression in normal human HSPC subsets and across different AML subtypes.

**Determine the effects of C/EBP $\beta$  overexpression and KD on myeloid colony forming ability and self-renewal.**

Limiting-dilution colony forming assays will be performed on transduced human CD34<sup>+</sup> HSPC and compared to control. Replating assays will be performed to assess the cells' self-renewal potential.

**Determine the effect of C/EBP $\beta$  overexpression or KD on normal human haematopoietic growth, differentiation, and development.**

CD34<sup>+</sup> HSPC will be transduced with a C/EBP $\beta$ -overexpression or -KD vectors and cultured in bulk liquid culture. Changes in monocytic and granulocytic growth will be assessed by analysing the cells' lineage and cell surface markers using multicolour flow cytometry.

**To determine the effect of C/EBP $\beta$  overexpression or KD on AML cell growth, proliferation, and apoptosis.**

AML cell lines will be transduced with a C/EBP $\beta$ -overexpression or -KD vectors and changes in cell growth will be evaluated by following the cells' proliferative ability over time. Furthermore, transduced cells will be assayed to determine if C/EBP $\beta$  dysregulated expression results in changes in cell cycle and apoptotic frequency.

## 5.3 Results

### 5.3.1 Expression of *CEBPB* mRNA increases throughout myeloid cell development

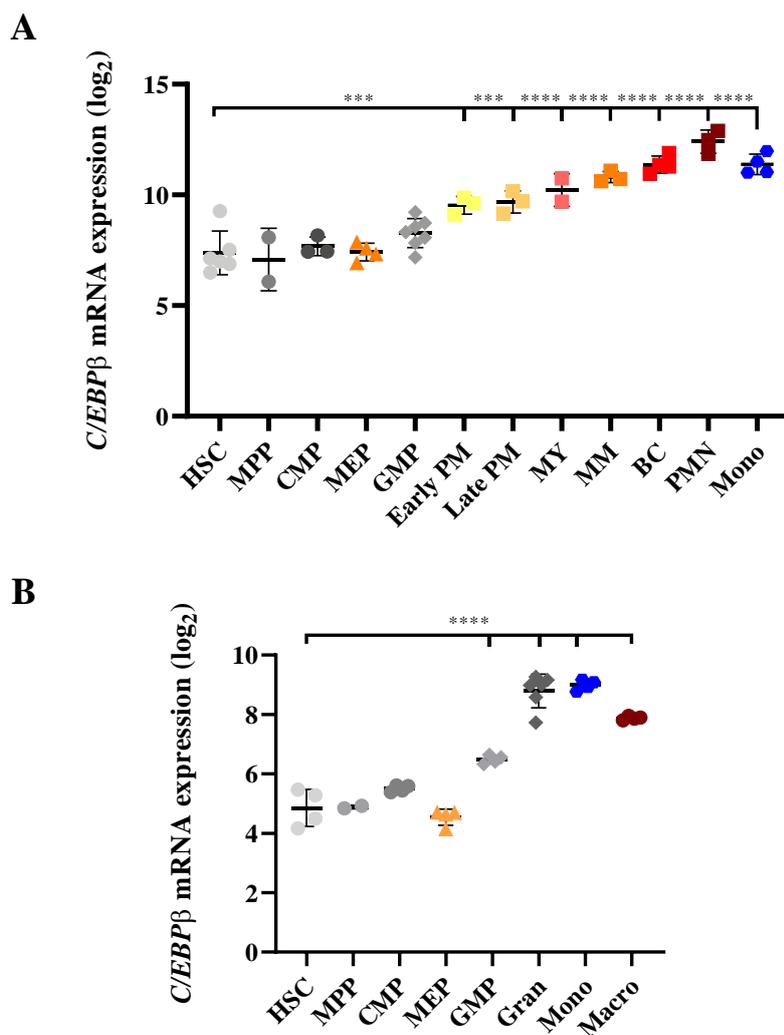
Members of the C/EBP family of TF, including C/EBP $\alpha$ , have been shown to play important roles in haematopoietic development (1.5.2); however, the expression and function of C/EBP $\beta$  in myeloid cell development has not yet been fully established. In order to determine the expression of *CEBPB* in human haematopoietic cells, publicly available microarray data was analysed using the online repository Bloodspot (Bagger *et al.*, 2016) (2.9.2.1). **Figure 5.1A** shows that the expression of *CEBPB* is lower in less differentiated cells, including HSC, MPP and CMP than differentiated cell types (mature granulocytes and monocytes). This data is supported by RNASeq analysis of the *CEBPB* mouse homologue, which showed a similar increase in mRNA expression in differentiated granulocytes, monocytes and macrophages compared to immature HSC (**Figure 5.1B**). These observations suggest the transcriptional regulation of *CEBPB* during haematopoietic development, and a possible role in myeloid differentiation.

To validate mRNA expression data, C/EBP $\beta$  protein expression was determined in human cord blood derived HSPC throughout myeloid development (2.3.3). Total protein lysates harvested over 13 days of myeloid culture were generated and western blot analysis performed. As shown in **Figure 5.2A-B**, the expression of the longer C/EBP $\beta$  isoform LAP\* (1.5.1) was shown to be higher in early myeloid progenitors (day 3 of culture), decreasing over time. Moreover, expression of C/EBP $\beta$  LAP was solely detected on day 3 lysates (**Figure 5.2A**), whilst LIP was not detected through myeloid development. These observations contrast with

mRNA expression. However, these lysates were derived from a combination of all haematopoietic populations, including monocytic, granulocytic and erythroid, and these are not equally distributed within a bulk population. To address this, C/EBP $\beta$  expression in each lineage was assessed in total protein lysates from monocytic-, granulocytic- and erythroid-committed progenitor cells. These cell subpopulations were FACS based on the expression on the cell surface markers CD14 and CD36 (2.7.4). This analysis showed that, on day 6 of culture, C/EBP $\beta$  is expressed in monocytes and erythrocytes and, to a lesser extent, in granulocytes (Figure 5.2C-D). However, day 6 granulocytic-committed cells most likely correspond to late promyelocytes which have lower C/EBP $\beta$  expression as compared to monocytes. However, these results do not correspond to the mRNA analysis, and the fact that C/EBP $\beta$  expression is shown to decline over 13 days cannot be accounted for (5.4.1). Nevertheless, this analysis suggests that C/EBP $\beta$  is differentially expressed in the different progenitor committed populations, suggesting a specific role for this protein in myeloid cell development.

### 5.3.2 *CEBPB* is variably expressed across AML subtypes

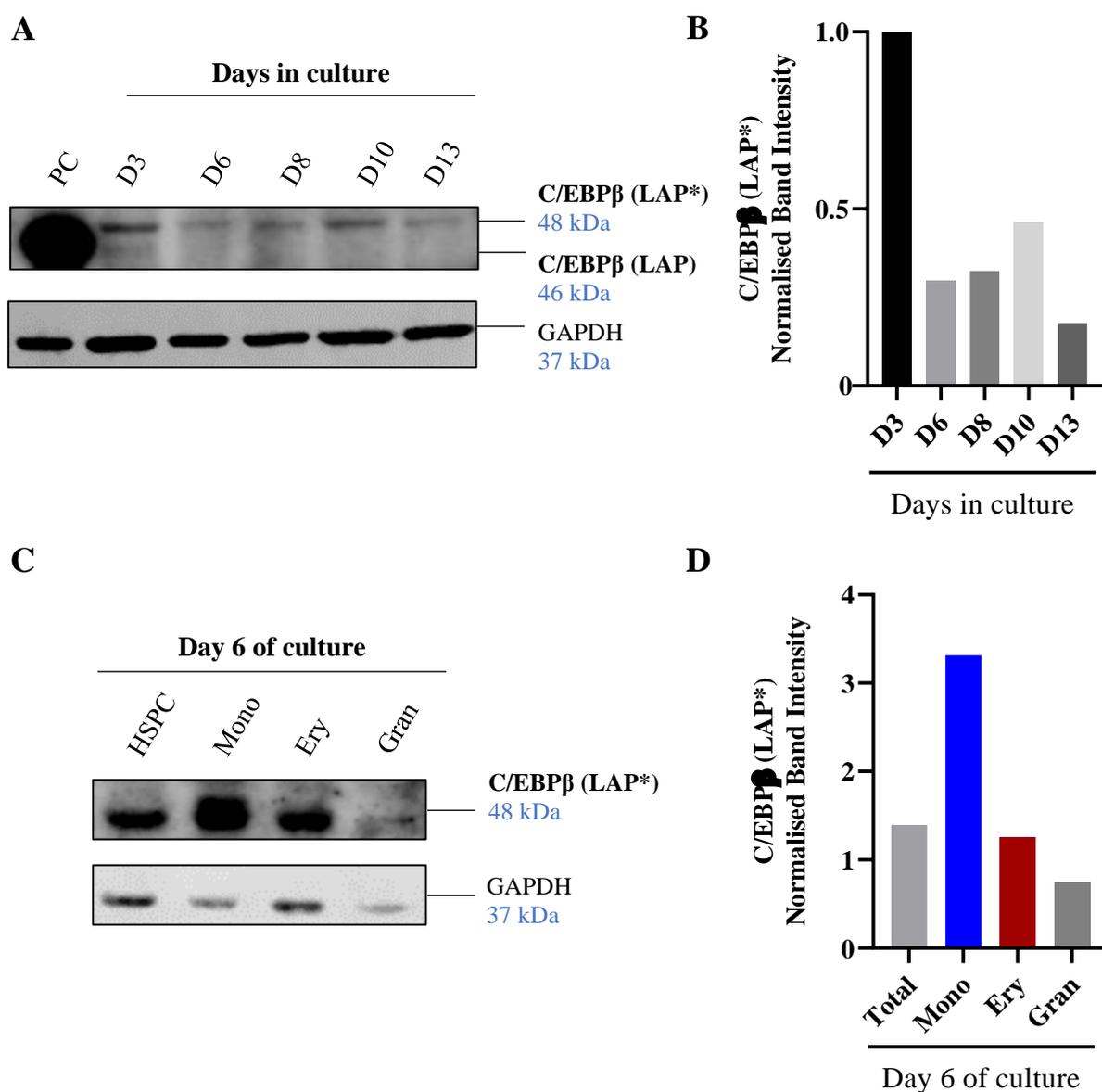
Heterogeneity observed in AML is generally associated with molecular and chromosomal abnormalities (1.2.2). To determine the expression of *CEBPB* mRNA in different subtypes of AML, this study analysed publicly available transcriptomic datasets, and compared them with normal undifferentiated HSC and mature granulocytes/monocytes. Figure 5.3A shows that *CEBPB* expression is significantly increased in all subtypes of AML studied, as compared to normal HSC. However, mRNA expression was found to be significantly downregulated compared to normal terminally differentiated myeloid cells, suggesting that reduced *CEBPB* expression might contribute to a block in normal haematopoietic differentiation. *CEBPB* expression was further observed to be heterogeneously expressed across FAB disease sub-types (Figure 5.3B). Interestingly, within the M2 AML subtype, *CEBPB* expression was shown to be significantly lower in t(8;21) patients as compared to non-t(8;21) FAB-M2 patients (Figure 5.3C), suggesting a specific role for this gene in the development of AML t(8;21).



**Figure 5.1 – *CEBPB* mRNA expression in normal haematopoiesis**

Normalised microarray data (log<sub>2</sub>) showing *CEBPB* mRNA expression in distinct human haematopoietic cells subsets based on cell surface marker expression. Normal human haematopoiesis data derived from (A) **GSE24759** (Rapin *et al.*, 2014; Svendsen *et al.*, 2016). (B) Normalised microarray data (log<sub>2</sub>) showing *CEBPB* mRNA expression across different murine haematopoietic cell types. Mouse normal haematopoiesis data derived from **GSE60101** (Lara-Astiaso *et al.*, 2014). Data indicates Mean ± 1SD. Statistical analysis was performed using one-way ANOVA with Bonferroni's multiple comparisons test; \*\*\* denotes p<0.001; \*\*\*\* denotes p<0.0001, vs HSC (Probeset Cebpb)

**HSC** – Haematopoietic Stem Cell; **MPP** – Multipotential Progenitors; **CMP** – Common Myeloid Progenitor cell; **MEP** – Megakaryocyte / Erythroid Progenitor cell; **GMP** – Granulocyte / Monocyte Progenitors; **Early PM** - Early Promyelocyte; **Late PM** - Late Promyelocyte; **MY** - Myelocyte; **MM** – Metamyelocytes; **BC** – Band cell; **PMN** – Polymorphonuclear cell; **Mono** – Monocytes; **Gran** – Granulocytes; **Macro** – Macrophages.

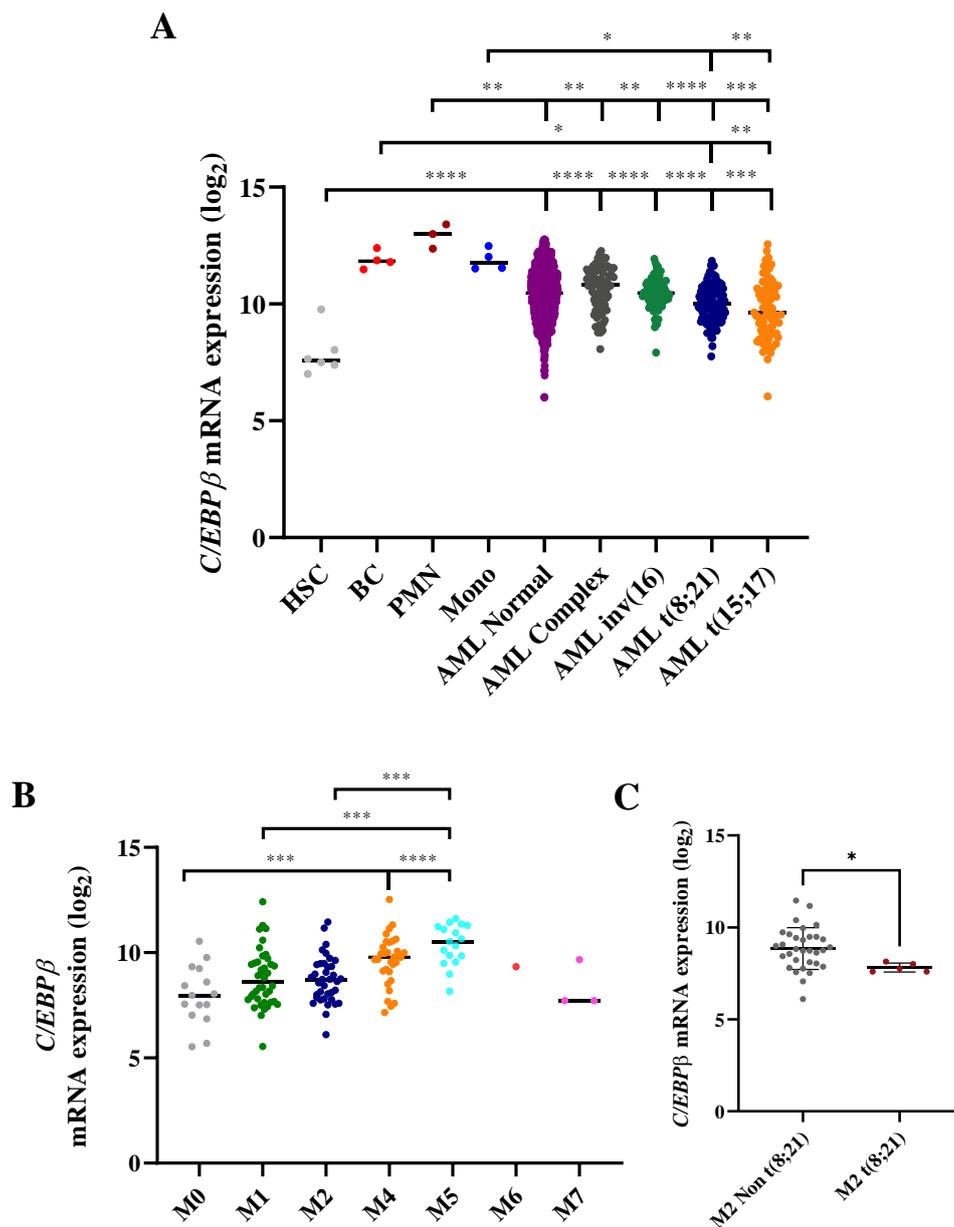


**Figure 5.2 – C/EBPβ endogenous expression throughout haematopoiesis**

(A) Western blot analysis showing C/EBPβ expression in HSPC over 13 days in culture in IMDM supplemented with IL-3 SCF, G- and GM-CSF (3S<sup>low</sup>G/GM) (n=1). HEK 293T cells transduced with an overexpression C/EBPβ vector were used as a positive control (PC). GAPDH was used as a loading control. (B) Bar chart showing densitometry analysis of C/EBPβ expression in HSPC over 13 days of culture derived from (A). C/EBPβ expression was quantified using ImageJ, normalised to loading control (GAPDH) and to HSPC day 3. (C) Western blot analysis showing C/EBPβ expression in human CD34<sup>+</sup> HSPC, and in monocytic (Mono, CD14<sup>high</sup>); erythroid (Ery; CD14<sup>low</sup>CD36<sup>high</sup>) and granulocytic (Gran; CD14<sup>low</sup>CD36<sup>low</sup>) progenitors (Tonks *et al.*, 2007) (n=1). GAPDH was used as a loading control. *Reprobed membrane provided by Dr. Rachael Nicholson.* (D) Bar chart showing densitometry analysis of C/EBPβ expression in HSPC, and respective myeloid populations derived from (C). C/EBPβ expression was quantified using ImageJ, normalised to loading control (GAPDH); PC – Positive control.

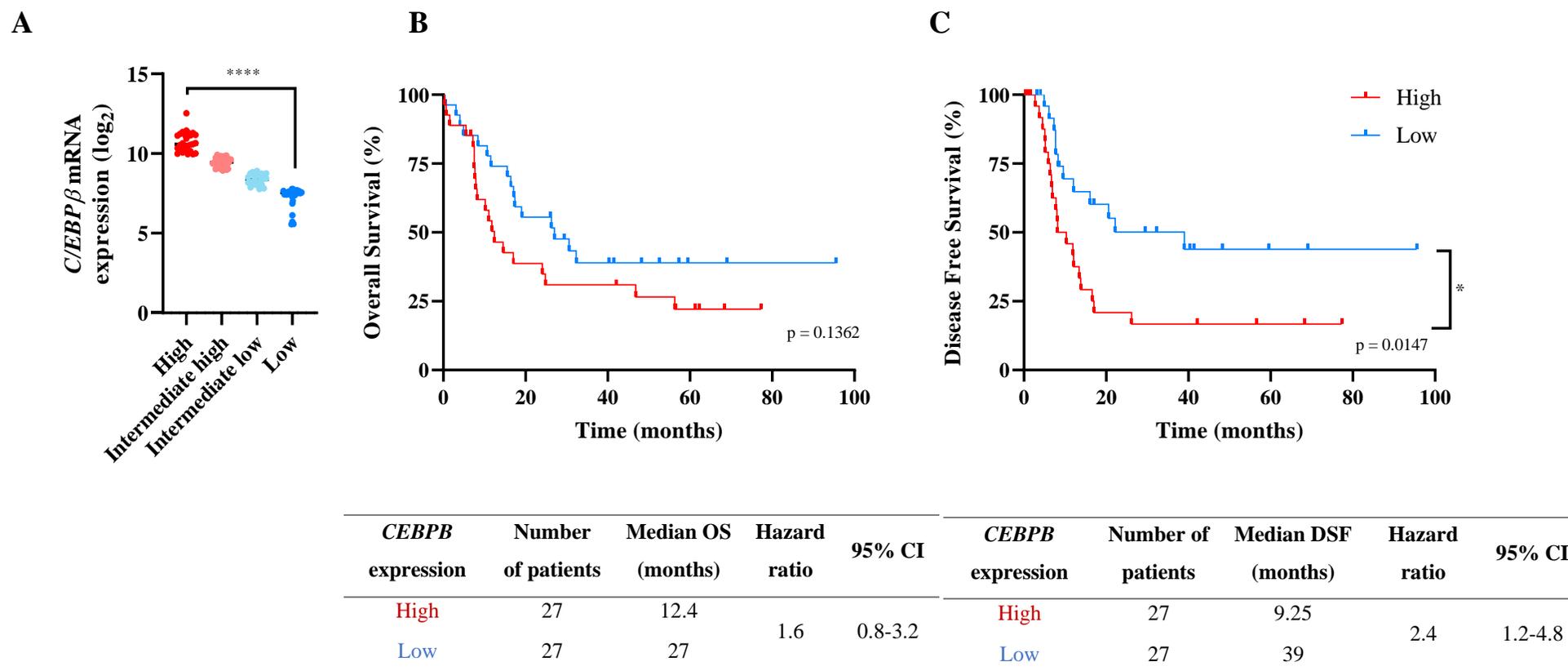
This study next analysed the TCGA 2013 RNASeq dataset (Ley *et al.*, 2013) to determine OS and disease-free survival of AML patients with high and low *CEBPB* expression. Patients were stratified according to *CEBPB* expression using the upper and lower quartiles (high and low *CEBPB* expression, respectively) (**Figure 5.4A**). Patients with APL and those who did not receive standard treatment were excluded from this analysis. High *CEBPB* expression patients had a median survival of 12.4 months compared to 27 months in patients with low *CEBPB* expression (**Figure 5.4B**); however, this difference was not statistically significant. High *CEBPB* expression was, however, associated with a significant decrease in disease-free survival (DFS of 9.25 months, compared with 39 months for patients with low *CEBPB* expression) (**Figure 5.4C**). This was further associated with an increased risk of relapse, suggested by a hazard ratio of 2.4 (95% CI[1.2-4.8]).

To further elucidate the relationship between *CEBPB* expression and clinical outcome, the co-occurrence with disease characteristics and known prognostic indicators was analysed. High *CEBPB* expression was found to be associated with a lower percentage of PB blasts (**Figure 5.5A**), as well as increased WBC (**Figure 5.5C**), although no differences were observed in BM blast proportion (**Figure 5.5B**). Moreover, patients with high *CEBPB* expression were found to be significantly associated with a lower mutation count, as compared to low *CEBPB* expression (**Figure 5.5D**). Additionally, high *CEBPB* mRNA expression significantly correlated with higher age at diagnosis (**Figure 5.6A**). There were no significant differences in patient gender (**Figure 5.6B**). Interestingly, t(8;21) was solely linked to low expression levels (**Figure 5.7A**), which was also significantly associated with mutations in both the *RUNX1* and *ETO* genes (**Figure 5.7B**). These observations further strengthen the hypothesis that *C/EBPβ* downregulation plays a role in the development of AML t(8;21).



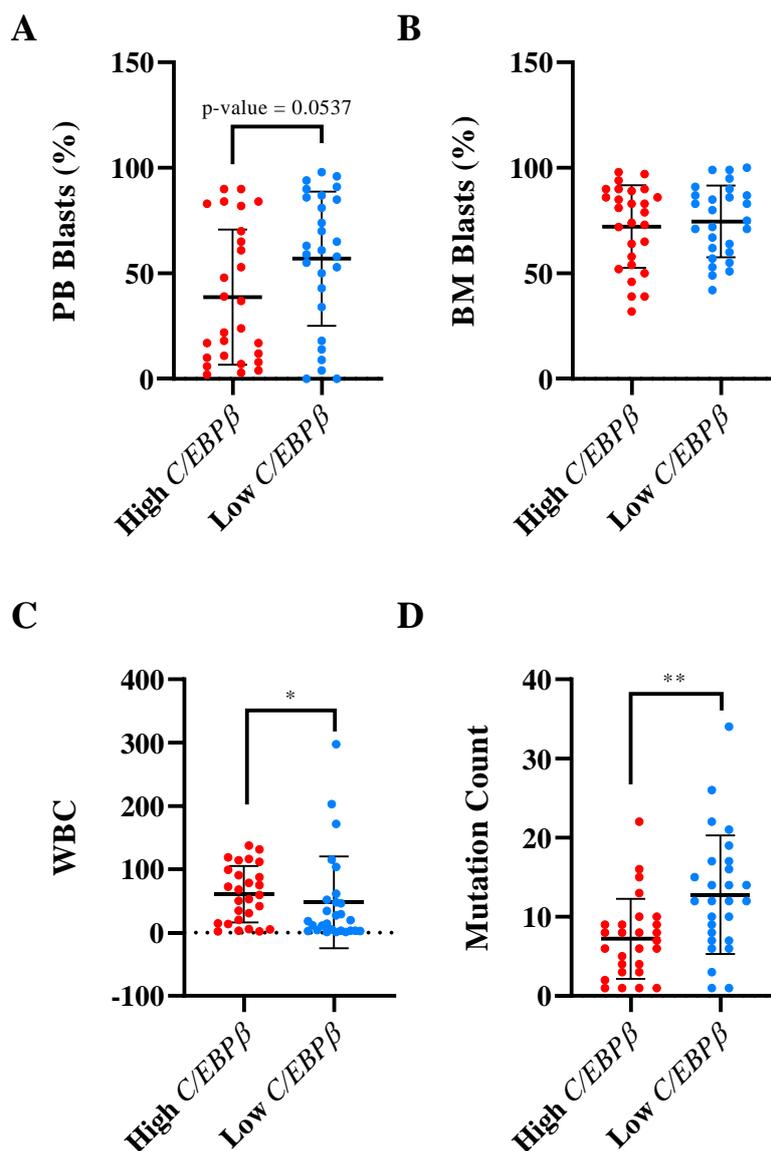
**Figure 5.3 – *CEBPB* mRNA expression levels are variable across AML subtypes**

(A) Normalised microarray data ( $\log_2$ ) showing *CEBPB* mRNA expression in normal human haematopoietic developmental subsets vs. AML subtypes. Normal human haematopoiesis data derived from **GSE422519** (Rapin *et al.*, 2014; Svendsen *et al.*, 2016); human AML cells derived from **GSE13159** (Haferlach *et al.*, 2010; Kohlmann *et al.*, 2008), **GSE15434** (Klein *et al.*, 2009), **GSE61804** (Metzelder *et al.*, 2015), **GSE14468** (Wouters *et al.*, 2009; Taskesen *et al.*, 2015; Taskesen *et al.*, 2011) and **TCGA** (Ley *et al.*, 2013) (data analysed using Bloodspot's algorithm Bloodpool (Bagger *et al.*, 2016)). Normalised microarray data ( $\log_2$ ) showing *CEBPB* mRNA expression in (B) AML according to FAB subtype and (C) AML FAB2, categorised into t(8;21) and non-t(8;21), within the **TCGA** dataset (Ley *et al.*, 2013) All studies were analysed using the 212501\_at probeset. Data indicates mean  $\pm$  1SD. Significant differences were analysed by one-way ANOVA with Bonferroni's multiple comparisons correction; \* denotes  $p < 0.05$ ; \*\* denotes  $p < 0.01$ ; \*\*\* denotes  $p < 0.001$ ; \*\*\*\* denotes  $p < 0.0001$ . **AML normal** – AML with normal karyotype; **AML complex** – AML with complex karyotype



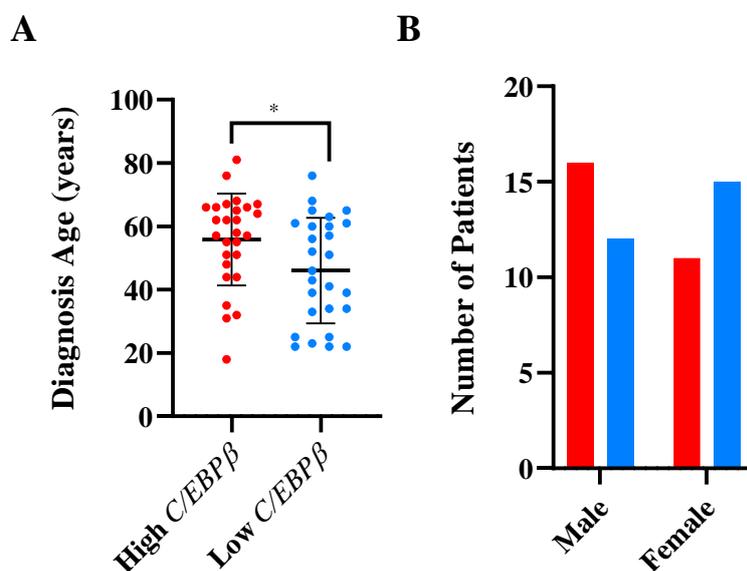
**Figure 5.4 – High levels of *CEBPB* is associated with poorer disease-free survival of AML patients**

(A) Normalised ( $\log_2$ ) *CEBPB* mRNA expression according to upper (n=27) and lower quartiles (n=27). Significance determined by one-way ANOVA; \*\*\*\* denotes  $p < 0.0001$ . (B) Overall survival and (C) disease-free survival analysis of AML patients stratified according to upper and lower *CEBPB* mRNA expression quartiles. Untreated patients and t(15;17) AML patients (that present a different treatment regime compared to other AML subtypes) were excluded from this analysis. Statistical analysis was performed using the Long-Rank test between high and low *CEBPB* expression groups. Data obtained from TCGA. (Ley *et al.*, 2013) using cBioPortal ([www.cbioportal.org](http://www.cbioportal.org)).



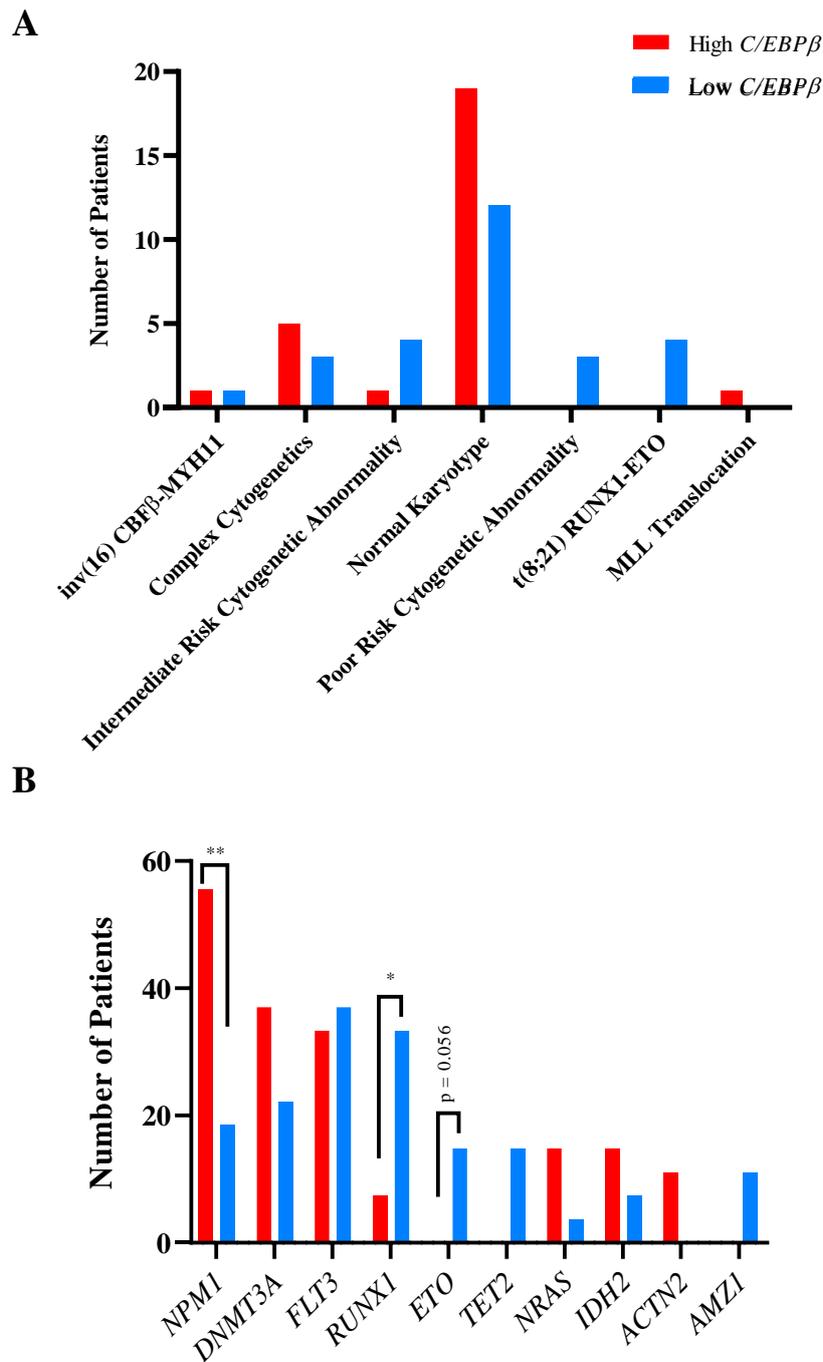
**Figure 5.5 – High *CEBPB* expression is associated with increased WBC and lower mutation count**

Relationship between high and low *CEBPB* expression and several clinical attributes of AML patients. Data obtained from TCGA (Ley *et al.*, 2013) using cBioPortal ([www.cbioportal.org](http://www.cbioportal.org)). *CEBPB* upper quartile (n=27); *CEBPB* lower quartile (n=27). Graph showing the percentage of (A) peripheral blood (PB) blasts; (B) BM blasts; (C) white blood count (WBC); and (D) mutation count of AML patients according to *CEBPB* high (n=27) and low (n=27) expression. Statistical differences between high and low *CEBPB* expression in AML patients was analysed using Mann-Whitney test; Data represents mean  $\pm$  1SD; \* denotes  $p < 0.05$ ; \*\* denotes  $p < 0.01$ .



**Figure 5.6 – Dysregulated *CEBPB* expression is associated with patient characteristics**

Relationship between high and low *CEBPB* expression and AML patient characteristics. Data obtained from TCGA (Ley *et al.*, 2013) using cBioPortal ([www.cbioportal.org](http://www.cbioportal.org)). *CEBPB* upper quartile (n=27); *CEBPB* lower quartile (n=27). Graph showing the (A) age at diagnosis (years) and (B) proportion of male and female patients according to *CEBPB* high (n=27) and low (n=27) expression. Statistical differences between high and low *CEBPB* expression in AML patients was analysed using Mann-Whitney test; Data represents mean  $\pm$  1SD; \* denotes  $p < 0.05$ .



**Figure 5.7 – *CEBPB* expression correlates with cytogenetic and molecular abnormalities**

Relationship between increased and reduced *CEBPB mRNA* expression and the occurrence of specific (A) cytogenetic or (B) molecular abnormalities in AML patients, according to *CEBPB* high (n=27) and low (n=27) quartiles. Data obtained from TCGA (Ley et al., 2013) using cBioPortal ([www.cbioportal.org](http://www.cbioportal.org)). Statistical analysis was performed using one-sided Fisher Exact test. \* denotes  $p < 0.05$ ; \*\* denotes  $p < 0.01$ .

### 5.3.3 Knockdown of C/EBP $\beta$ influences normal myeloid development

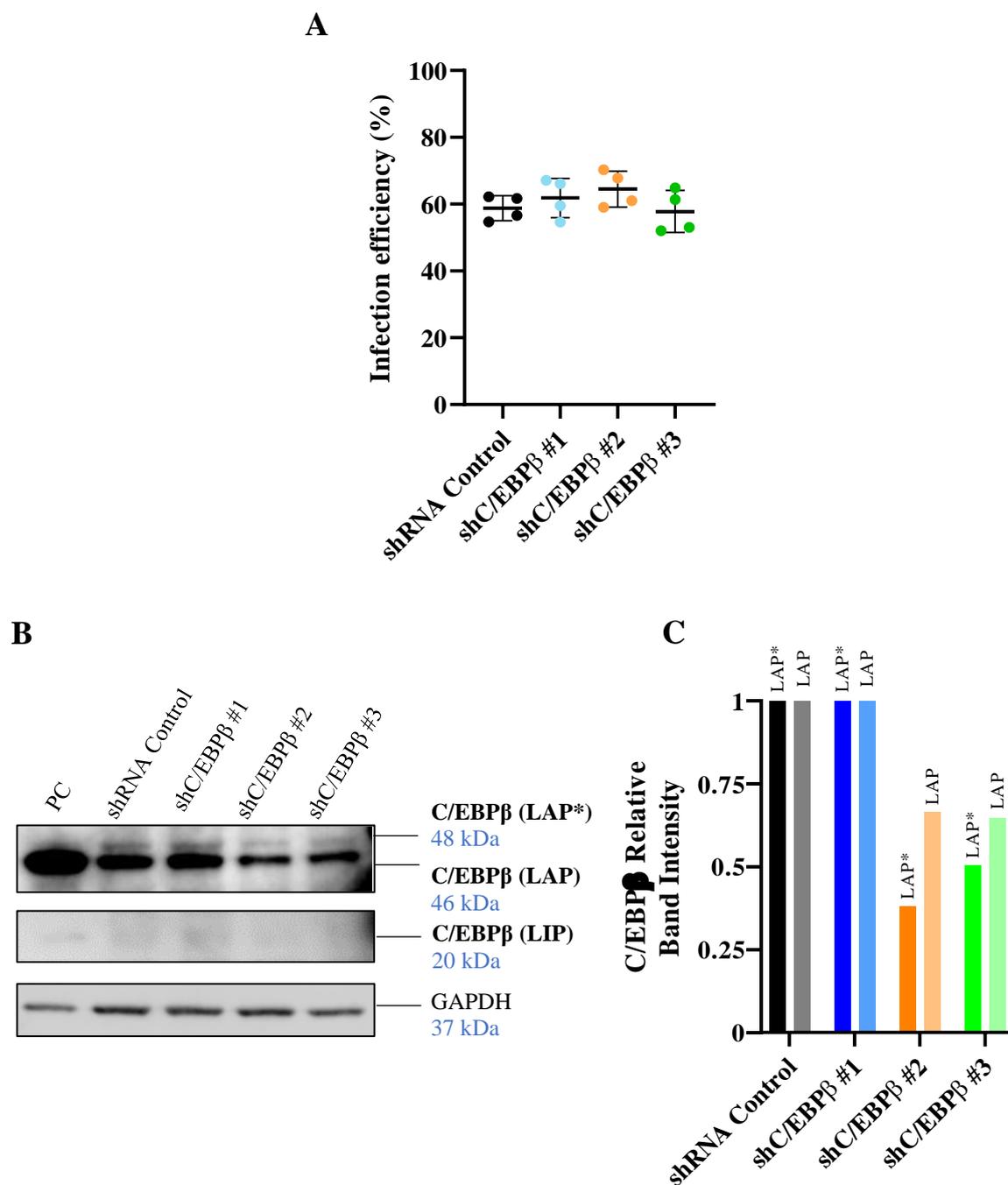
The above observations show that *CEBPB* expression increases throughout blood cell development and lower levels of *CEBPB* are associated with t(8;21) in AML patients. *In vitro* studies have shown that both the C/EBP $\beta$  mRNA and protein are downregulated in RUNX1-ETO-expressing CD34<sup>+</sup> HSPC, as compared to control (**Figure 3.18**). For these reasons, this study next determined if C/EBP $\beta$  is required for normal haematopoietic development. C/EBP $\beta$  was knocked down in normal human cord blood derived CD34<sup>+</sup> HSPC and the effect on myeloid cell growth, differentiation and self-renewal was assessed. Moreover, ectopic expression of this protein for overexpression studies was also performed, as described in **5.3.4**.

#### 5.3.3.1 Generation of C/EBP $\beta$ -knockdown CD34<sup>+</sup> HSPC

To assess the effect of C/EBP $\beta$  KD in CD34<sup>+</sup> HSPC, lentiviral vectors based on shRNA constructs, in combination with a GFP selectable marker were used (**Table 2.1**). Transduction efficiency in HSPC using the four shRNA vectors was assessed using GFP expression. As shown in **Figure 5.8A**, transduction frequencies ranged between 60-70% across all constructs. To validate protein KD by western blot, cells were sorted by FACS for GFP expression. On day 6 of culture, C/EBP $\beta$  LAP\* and LAP isoforms were detected in normal human myeloid cells, with no detectable levels of LIP (**Figure 5.8B**). KD with shC/EBP $\beta$  #1 had no detectable effect on endogenous levels of the C/EBP $\beta$  protein, whilst shC/EBP $\beta$  #2 and #3 showed a reduction of 40-50% of LAP\* protein expression, and a 45% reduction in LAP, as compared to shRNA control (no mammalian target) (**Figure 5.8C**).

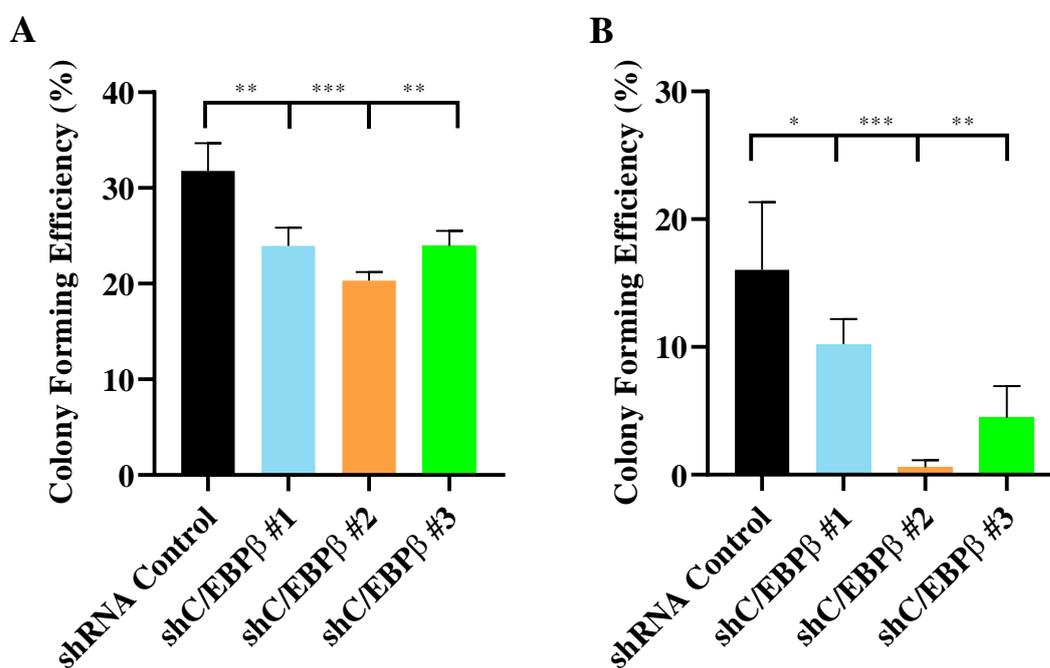
#### 5.3.3.2 Knockdown of C/EBP $\beta$ leads to a decrease in HSPC colony forming ability

To determine the effect of C/EBP $\beta$  KD on myeloid colony forming efficiency and self-renewal potential, a colony assay, followed by subsequent replating was performed on GFP<sup>+</sup> sorted cells (**2.3.4**). Following one week of culture, all three shC/EBP $\beta$  cultures had significantly less myeloid colonies as compared to control; showing a reduction of 25%-36% (**Figure 5.9A**). Moreover, all three cultures showed a significant decrease in self-renewal capacity (**Figure 5.9B**). The fact that one of the knockdown constructs, shC/EBP $\beta$  #1, was shown to inhibit colony formation, even though no level of KD was detected (**Figure 5.8**), suggests off-target effects, further discussed in **5.4**. Nevertheless, these observations suggest that C/EBP $\beta$  KD reduces myeloid colony formation with decreased self-renewal potential.



**Figure 5.8 – C/EBPβ is knocked-down in human CD34<sup>+</sup> HSPC**

(A) Summary data showing percentages of GFP positivity in control and C/EBPβ KD CD34<sup>+</sup> HSPC populations, analysed on day 3 of culture. Gating strategies were applied as described in **Figure 2.4**. Data indicates mean  $\pm$  1SD (n=4). (B) Western blot analysis showing C/EBPβ protein expression in FACS HSPC transduced with a control or C/EBPβ-targeted shRNA construct (**Table 2.1**), on day 6 of culture (n=1). HEK 293T transfected with a C/EBPβ-overexpression construct was used as a positive control (PC); GAPDH was used as a loading control. (C) Bar chart showing densitometry analysis regarding C/EBPβ expression in HSPC transduced with three shC/EBPβ constructs. C/EBPβ expression was quantified using ImageJ, normalised to loading control (GAPDH) and to shRNA control (n=1).



**Figure 5.9 – C/EBPβ KD disrupts myeloid colony formation of HSPC**

**(A)** Myeloid colony forming efficiency of control and three shC/EBPβ cultures following 7 days of growth in liquid culture containing IL-3, SCF, G- and GM-CSF (2.3.4). Data indicates mean  $\pm$  1SD (n=3). Significant difference between shC/EBPβ cultures and control was analysed by one-way ANOVA with Bonferroni's multiple comparisons correction; \*\* denotes  $p < 0.01$ , \*\*\* denotes  $p < 0.001$ . **(B)** Self-renewal potential was assessed by a single replating round of control and shC/EBPβ cultures, in the same conditions as previously. Data indicates mean  $\pm$  1SD (n=3). Significant differences between shC/EBPβ cultures and control were analysed by one-way ANOVA with Bonferroni's multiple comparisons correction; \* denotes  $p < 0.05$ ; \*\* denotes  $p < 0.01$ ; \*\*\* denotes  $p < 0.001$ .

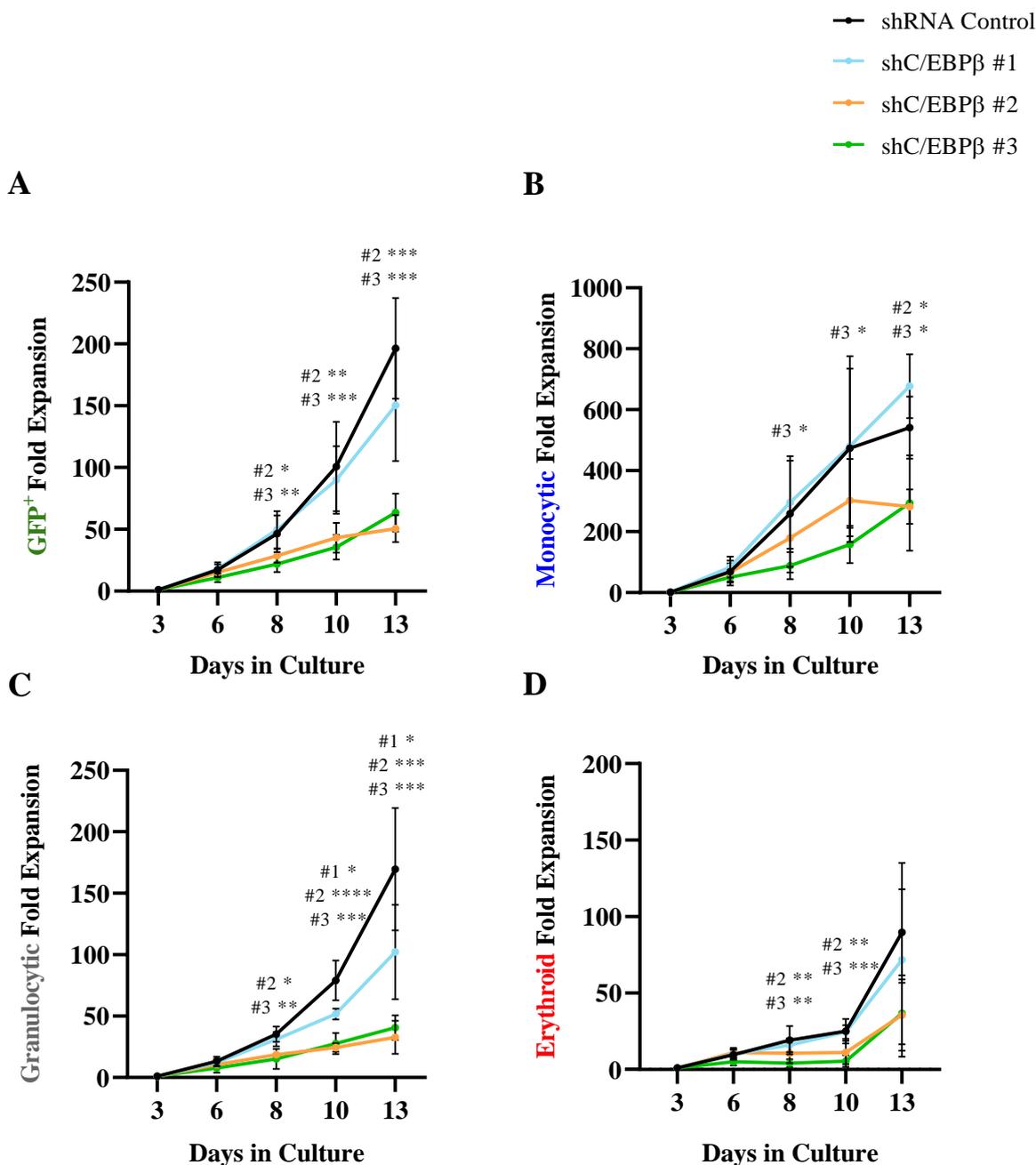
### 5.3.3.3 Knockdown of *C/EBPβ* suppresses myeloid cell growth

This study next determined the consequences of reduced *C/EBPβ* expression on cell proliferation and differentiation. With no influence on *C/EBPβ* expression levels, sh*C/EBPβ* #1 had little effect on myeloid cell growth in a bulk population setting (**Figure 5.10A**). In contrast, there was up to a significant 2.8-fold reduction in HSPC growth transduced with other sh*C/EBPβ*. Monocytic and granulocytic cumulative cell growth was significantly impaired by approximately 2- and 5-fold, respectively, when analysed on day 13 of culture in KD #2 and #3 cells (**Figure 5.10B-C**). Even though erythroid growth showed significant reduction in cells transduced with sh*C/EBPβ* #2 and #3, this was deemed not significant by day 13 of culture (**Figure 5.10D**). However, these cultures are not designed to examine erythroid growth given the lack of EPO. These observations suggest that KD of *C/EBPβ* influences the growth of myeloid cells, particularly in cells committed to the monocytic and granulocytic lineages.

### 5.3.3.4 Knockdown of *C/EBPβ* influences myeloid differentiation

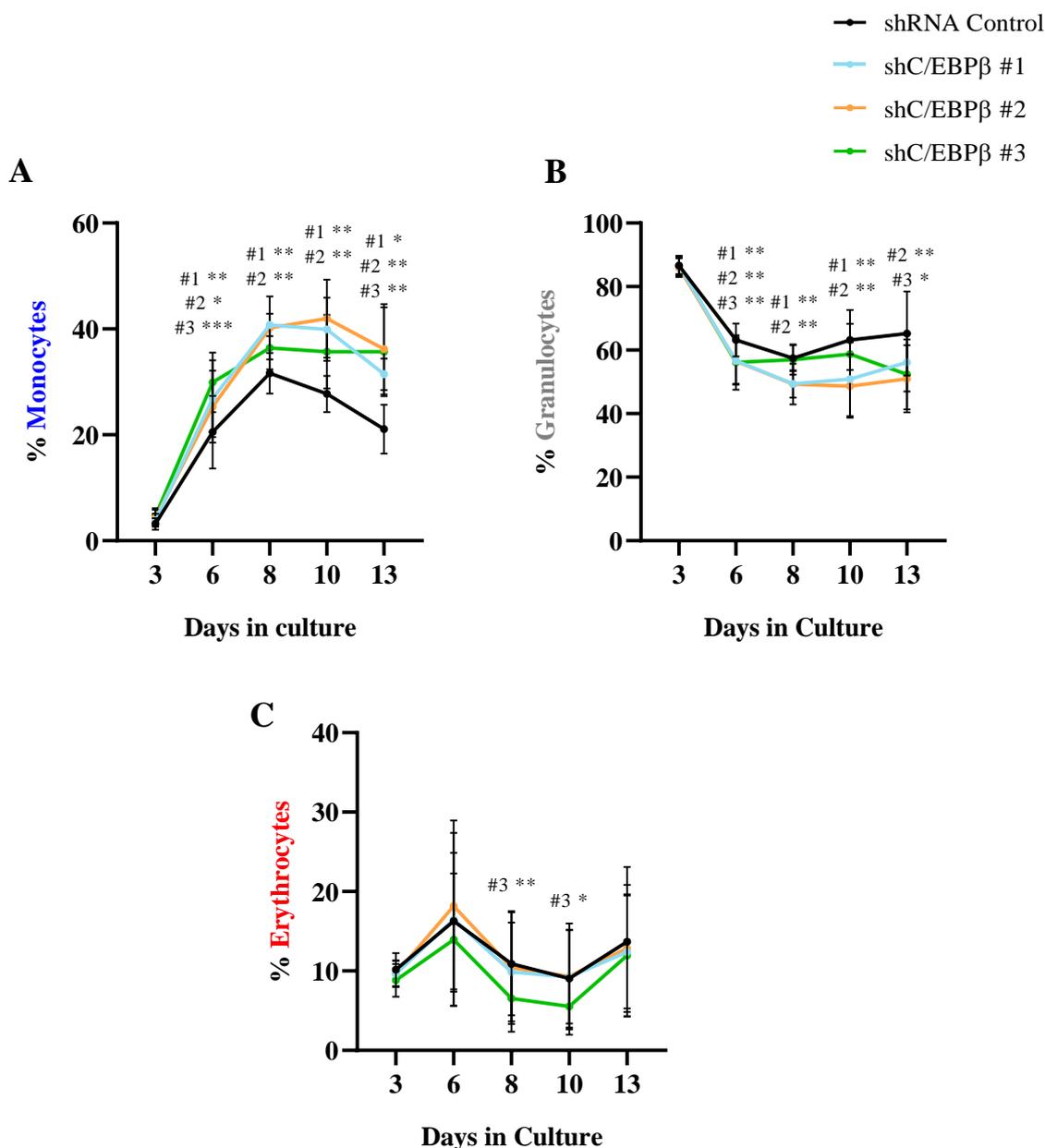
Differentiation in each myeloid lineage was determined using myeloid differentiation cell surface markers, including CD34, CD11b, CD14 and CD15 (**4.3.3.4**). Over 13 days of culture, KD of *C/EBPβ* resulted in a significant increase in the proportion of monocytic cells during myeloid development, culminating in an increase of 10%, 15% and 14% observed with sh*C/EBPβ* #1, #2 and #3, respectively, as compared to control cells (**Figure 5.11A**). This increase was likely at the expense of significant decreases in the granulocytic population, reduced on day 13 of development by 9%, 14% and 13% in cultures transduced with in sh*C/EBPβ* #1, #2 and #3, correspondingly (**Figure 5.11B**). No significant differences were observed in the erythroid-committed population (**Figure 5.11C**); even though these cells have the ability to grow within the same media as the other haematopoietic constituents, terminal erythroid differentiation is impaired due to lack of EPO.

Regarding monocytic development, even though there was a significant impairment of growth in cells transduced with sh*C/EBPβ* #2 and #3, analysis of monocytic differentiation markers showed a modest increase of 1.3 and 1.6-fold in the expression of the cell surface markers CD11b and CD14, respectively (**Figure 5.12A-C**).



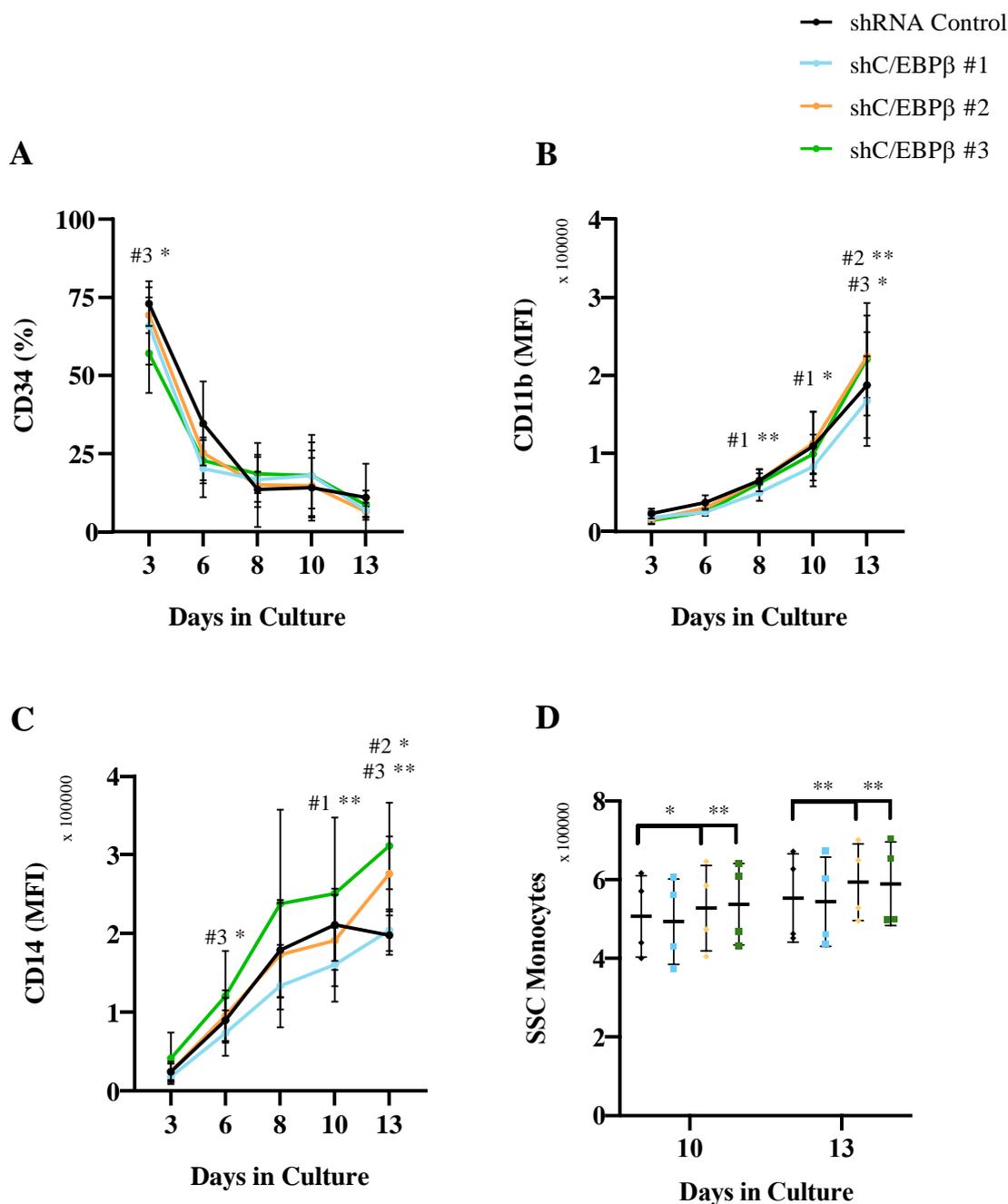
**Figure 5.10 – KD of C/EBPβ decreases granulocytic growth in myeloid development**

(A) Cumulative fold-expansion of control and three shC/EBPβ constructs for GFP positive cells in culture medium containing IL-3, SCF, G- and GM-CSF, grown over 13 days (**Supplementary Figure 16**). (B) Cumulative fold-expansion of monocytic cells (CD13<sup>+</sup> CD36<sup>+</sup>) transduced with a control or shC/EBPβ vectors. (C) Cumulative fold-expansion of granulocytic cells (CD13<sup>+</sup> CD36<sup>-</sup>) transduced with a control or shC/EBPβ vector. (D) Cumulative fold-expansion of erythroid cells (CD13<sup>-</sup> CD36<sup>+</sup>) transduced with a control of shC/EBPβ vector. Cells were re-seeded at the appropriate density following each time-point, as described in 2.3.3. Gating strategies were applied as described in **Figure 2.4** and **Figure 2.5**. Data indicates mean  $\pm$  1SD (n=4). Significant differences between shC/EBPβ cultures and control at each time-point were analysed by one-way ANOVA with Bonferroni's multiple comparisons correction; \* denotes p<0.05; \*\* denotes p<0.01; \*\*\* denotes p<0.001; \*\*\*\* denotes p<0.0001; (this data has also been represented as dot plots, within the supplementary section [**Supplementary Figure 38**]).



**Figure 5.11 – C/EBPβ KD disrupts the balance between the monocytic and granulocytic populations in CD34<sup>+</sup> HSPC during myeloid cell development**

Summary data showing percentage of (A) monocytic (CD13<sup>+</sup> CD36<sup>+</sup>), (B) granulocytic (CD13<sup>+</sup> CD36<sup>-</sup>) and (C) erythroid cells (CD13<sup>-</sup> CD36<sup>+</sup>) in control and three shC/EBPβ cultures. Cells were re-seeded at the appropriate density following each time-point, as described in 2.3.3. Gating strategies were applied as described in Figure 2.4 and Figure 2.5. Data indicates mean  $\pm$  1SD (n=4). Significant differences between shC/EBPβ cultures and control at each time-point were analysed by one-way ANOVA with Bonferroni's multiple comparisons correction; \* denotes p<0.05; \*\* denotes p<0.01; \*\*\* denotes p<0.001; (this data has also been represented as dot plots, within the supplementary section [Supplementary Figure 39]).



**Figure 5.12 – KD of C/EBPβ has little impact on normal monocytic differentiation**

(A) Summary data of CD34<sup>+</sup> expression in terms of percentage in monocytic cells (CD13<sup>+</sup> CD36<sup>+</sup>) over time for control and shC/EBPβ cells. Summary data showing (B) CD11b and (C) CD14 expression in terms of MFI in monocytic cells over time for control and shC/EBPβ cells. (D) Granularity of monocytic cells was measured by changes in SSC for control and shC/EBPβ cultures on days 10 and 13 of myeloid differentiation. Cells were re-seeded at the appropriate density following each time-point, as described in 2.3.3. Gating strategies were applied as described in Figure 2.4 and Figure 2.5. Data indicates mean ± 1SD (n=4). Significant differences between shC/EBPβ cultures and control at each time-point were analysed by one-way ANOVA with Bonferroni's multiple comparisons correction; \* denotes p<0.05; \*\* denotes p<0.01; (for representative flow cytometry plots please refer to Supplementary Figure 17, Supplementary Figure 18, Supplementary Figure 19; this data has also been represented as dot plots, within the supplementary section [Supplementary Figure 40]).

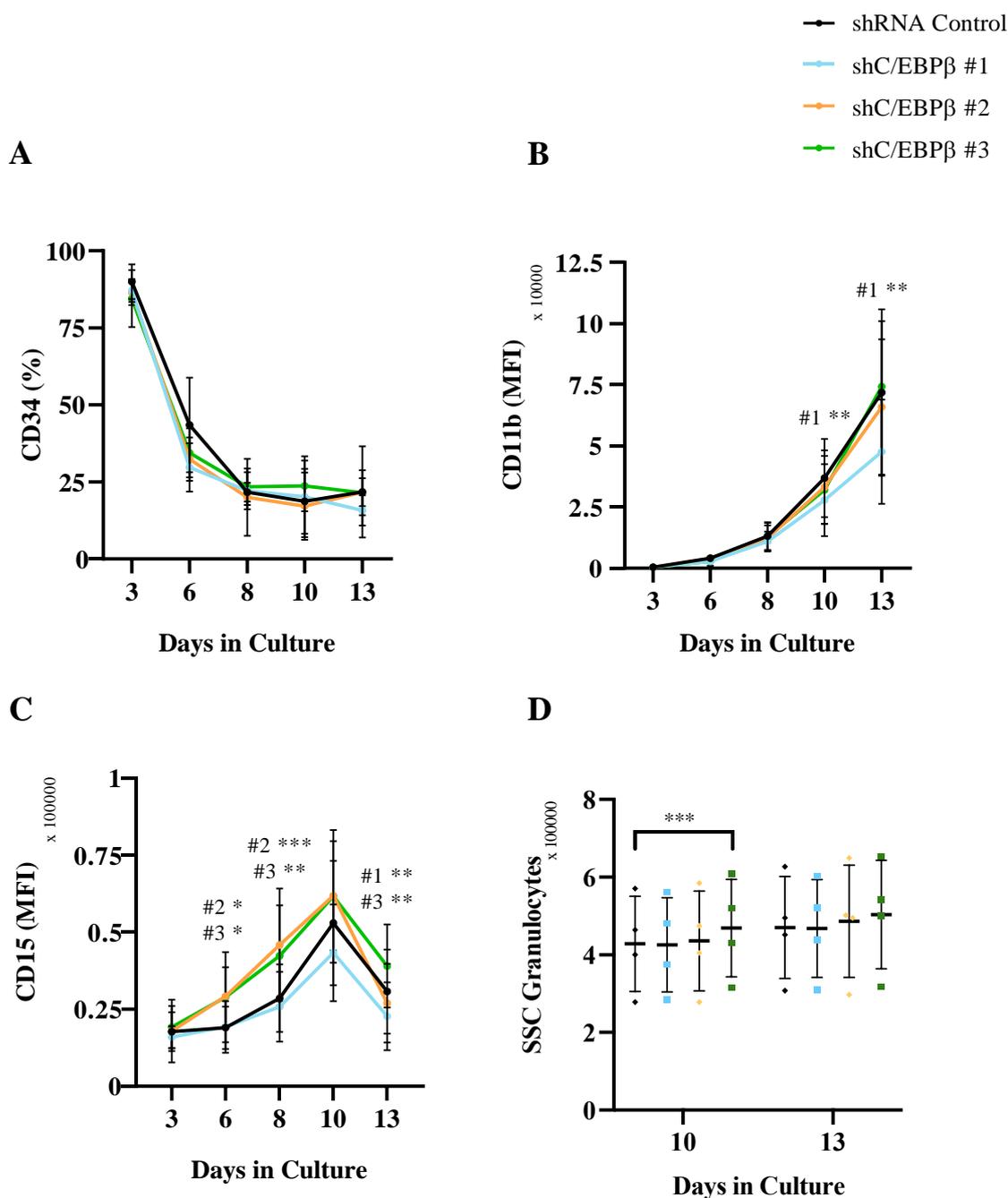
Moreover, monocytic cell surface marker upregulation was accompanied by a significant increase in the SSC of these cells, implying cells transduced with constructs #2 and #3 have an increased granularity and complexity, as compared to shRNA control cells (**Figure 5.12D**). Altogether, these observations suggest that the reduction in monocytic proliferation maybe a result of increased differentiation; however, this increase was modest.

Regarding granulocytic development, KD of C/EBP $\beta$  did not significantly influence the expression of the CD34 maker nor CD11b (**Figure 5.13A-B**). However, shC/EBP $\beta$  #2 and #3-transduced cells showed a significant upregulation of the granulocytic marker CD15 throughout myeloid development, culminating in an increase of 1.3-fold observed in shC/EBP $\beta$  #3 (**Figure 5.13C**). No differences were observed in terms of granularity and complexity upon KD of C/EBP $\beta$  (**Figure 5.13D**). These results indicate that normal levels of C/EBP $\beta$  are not critical for granulocytic development.

Morphological assessment of transduced cells was performed using May-Grünwald-Giemsa staining of cells cultured under myeloid conditions for 17 days (**Figure 5.14**). This analysis scores post-mitotic developmental changes which are difficult to measure by flow cytometry. The proportion of granulocytes was markedly increased in all conditions, as compared to monocytes. shC/EBP $\beta$  cultures showed a decrease in the percentage of terminally differentiated granulocytic cells. These suggest that KD of C/EBP $\beta$  may impact late granulocytic differentiation.

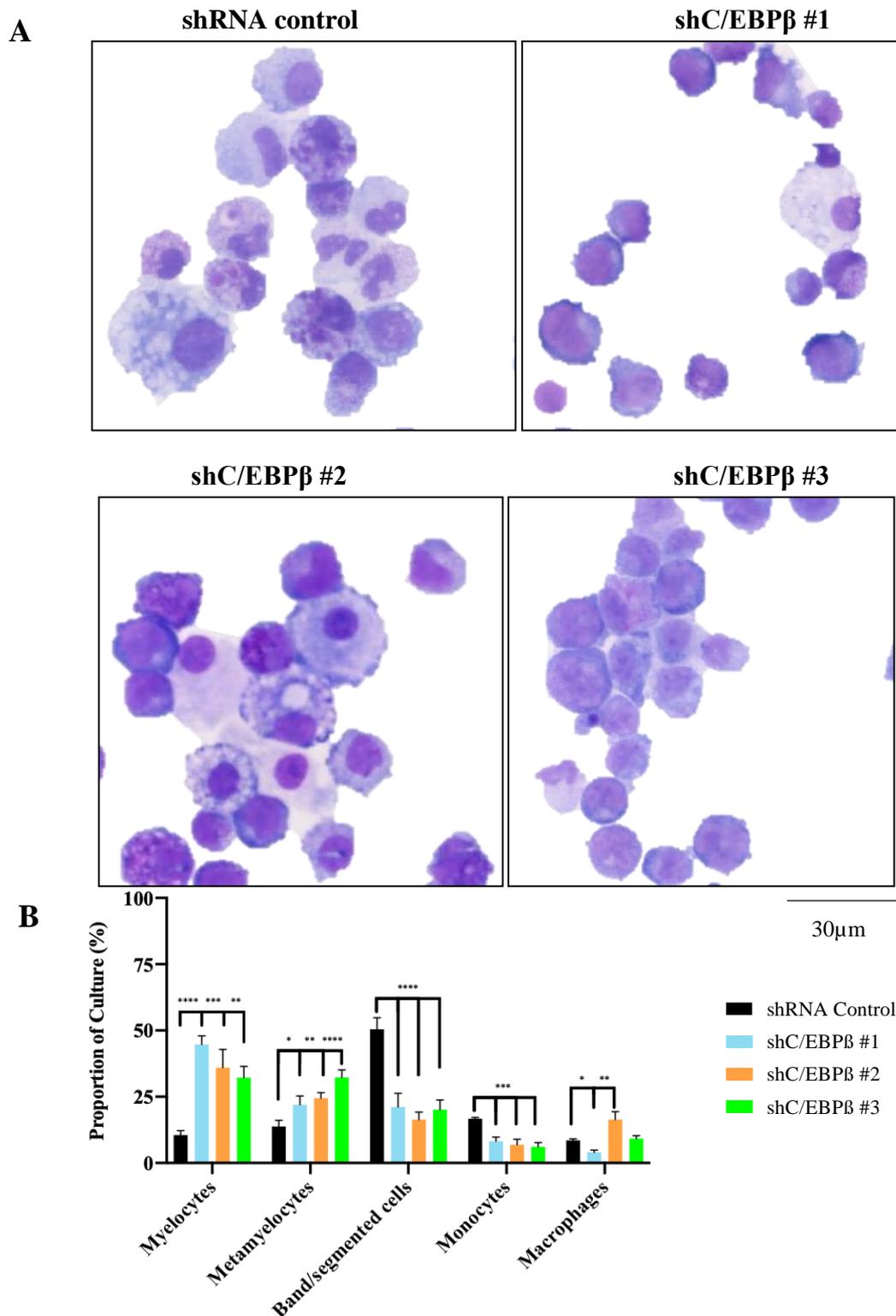
shC/EBP $\beta$  cultures presented a significant decrease in the percentage of monocytic population, in accordance with lineage analysis (**Figure 5.11A-B**). These observations support the hypothesis that KD of C/EBP $\beta$  has a role in promoting monocytic differentiation, as fully differentiated cells will lose their self-renewal potential, reaching terminal maturation earlier than shRNA control cells.

Overall, these observations do not agree with the mRNA results described above (**5.3.1**), which suggested a role for C/EBP $\beta$  in haematopoietic development. However, this maybe a result of several factors, including poor level of KD, as well as the interplay with other members of the C/EBP family of TF, further discussed in section **5.4**.



**Figure 5.13 – KD of C/EBPβ promotes granulocytic development**

(A) Summary data of CD34<sup>+</sup> expression in terms of percentage in granulocytic cells (CD13<sup>+/−</sup> CD36<sup>−</sup>) over time for control and shC/EBPβ cells. Summary data showing (B) CD11b and (C) CD15 expression in terms of MFI in granulocytic cells over time for control and shC/EBPβ cells. (D) Granularity of granulocytic cells was measured by changes in SSC for control and shC/EBPβ cultures on days 10 and 13 of myeloid differentiation. Cells were re-seeded at the appropriate density following each time-point, as described in 2.3.3. Gating strategies were applied as described in Figure 2.4 and Figure 2.5. Data indicates mean ± 1SD (n=4). Significant differences between shC/EBPβ cultures and control at each time-point was analysed by one-way ANOVA with Bonferroni's multiple comparisons correction; \* denotes p<0.05; \*\* denotes p<0.01; \*\*\* denotes p<0.001; (for representative flow cytometry plots please refer to Supplementary Figure 20, Supplementary Figure 21, Supplementary Figure 22; this data has also been represented as dot plots, within the supplementary section [Supplementary Figure 41]).



**Figure 5.14 – C/EBPβ KD affects cell morphology during myeloid development**

(A) Representative images showing control and shC/EBPβ cultures, analysed on day 17 of differentiation, by cytopinning  $3 \times 10^4$  cells and staining them with May-Grünwald-Giemsa differential stain. (B) Differential counts of all cultures with morphology categorised into granulocytic and monocytic-committed cells, according to **Figure 2.2**. Approximately 500 cells were counted in total and each population count normalised to total number of cells. Granulocytic cells were further divided into myelocytes, metamyelocytes and band/segmented cells. Data indicates mean  $\pm$  1SD (n=3). Significant difference between shC/EBPβ cultures and control was analysed by one-way ANOVA with Bonferroni's multiple comparisons correction; \* denotes  $p < 0.05$ ; \*\* denotes  $p < 0.01$ .

### 5.3.3.5 *Knockdown of C/EBP $\beta$ does not result in significant cell cycle or apoptotic changes*

Having determined that C/EBP $\beta$  KD significantly reduces myeloid cell proliferation, this study next determined if the reduction in proliferation could be a result of changes to the cells' normal cell cycle, or due to an increase in apoptosis. Only one shC/EBP $\beta$  culture showed a modest accumulation of cells in the G<sub>1</sub> phase of the cycle (5% in shC/EBP $\beta$  #2) (**Figure 5.15A**), suggesting a slight reduction in rate of cell cycle progression.

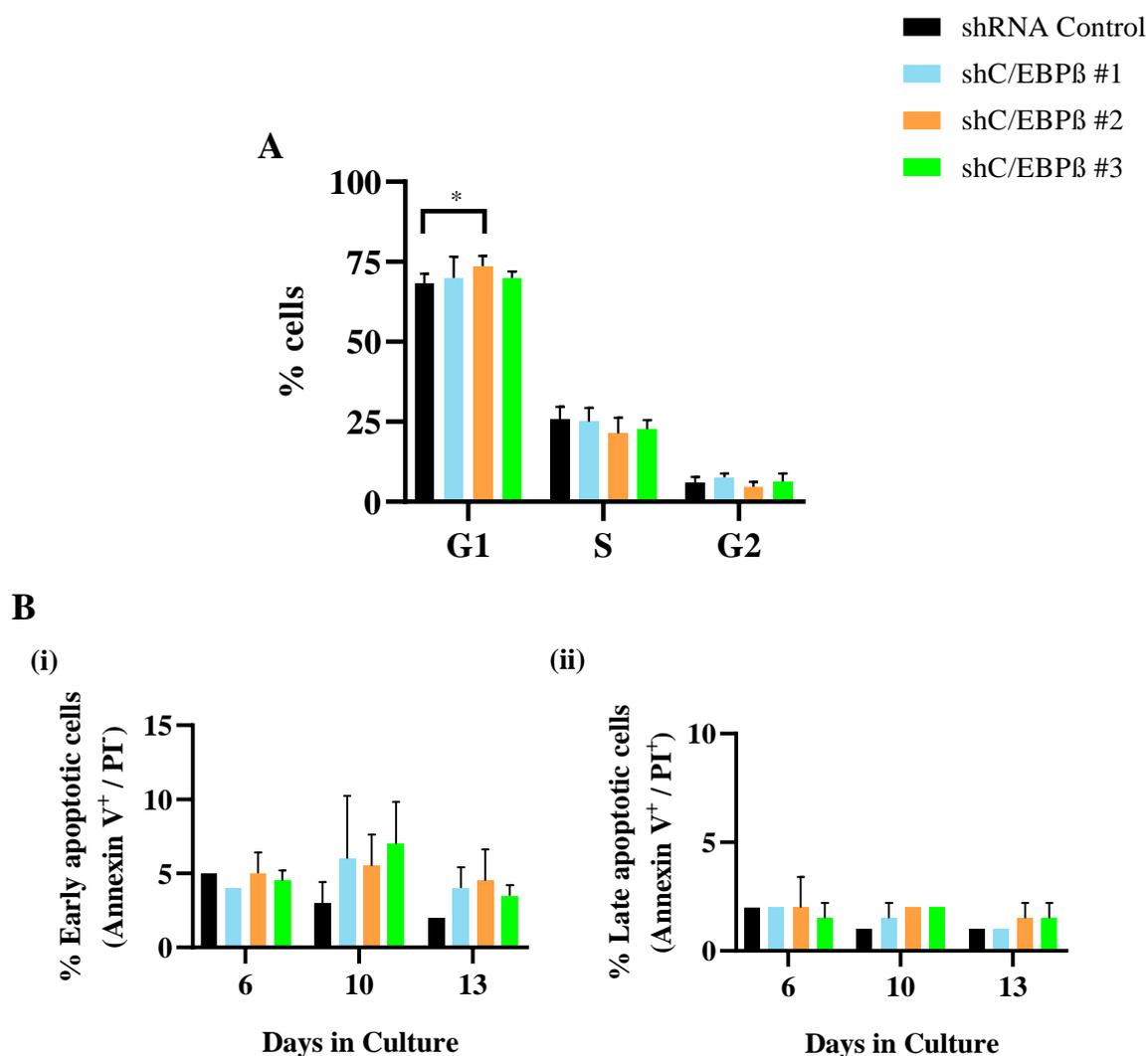
An Annexin-V assay was further performed to assess whether C/EBP $\beta$  induced apoptosis in these cells. No changes were observed in the apoptotic frequency of HSPC, upon KD of C/EBP $\beta$  (**Figure 5.15B**). Altogether, the reduced proliferation observed in shC/EBP $\beta$  cultures is not a result of either significant changes to the cells' normal cycle nor due to apoptotic events.

## 5.3.4 **Overexpression of C/EBP $\beta$ influences normal myeloid development**

Having determined the effect of C/EBP $\beta$  KD on normal human haematopoiesis, this study then aimed at analysing the consequences of C/EBP $\beta$  overexpression using the same human cord blood model.

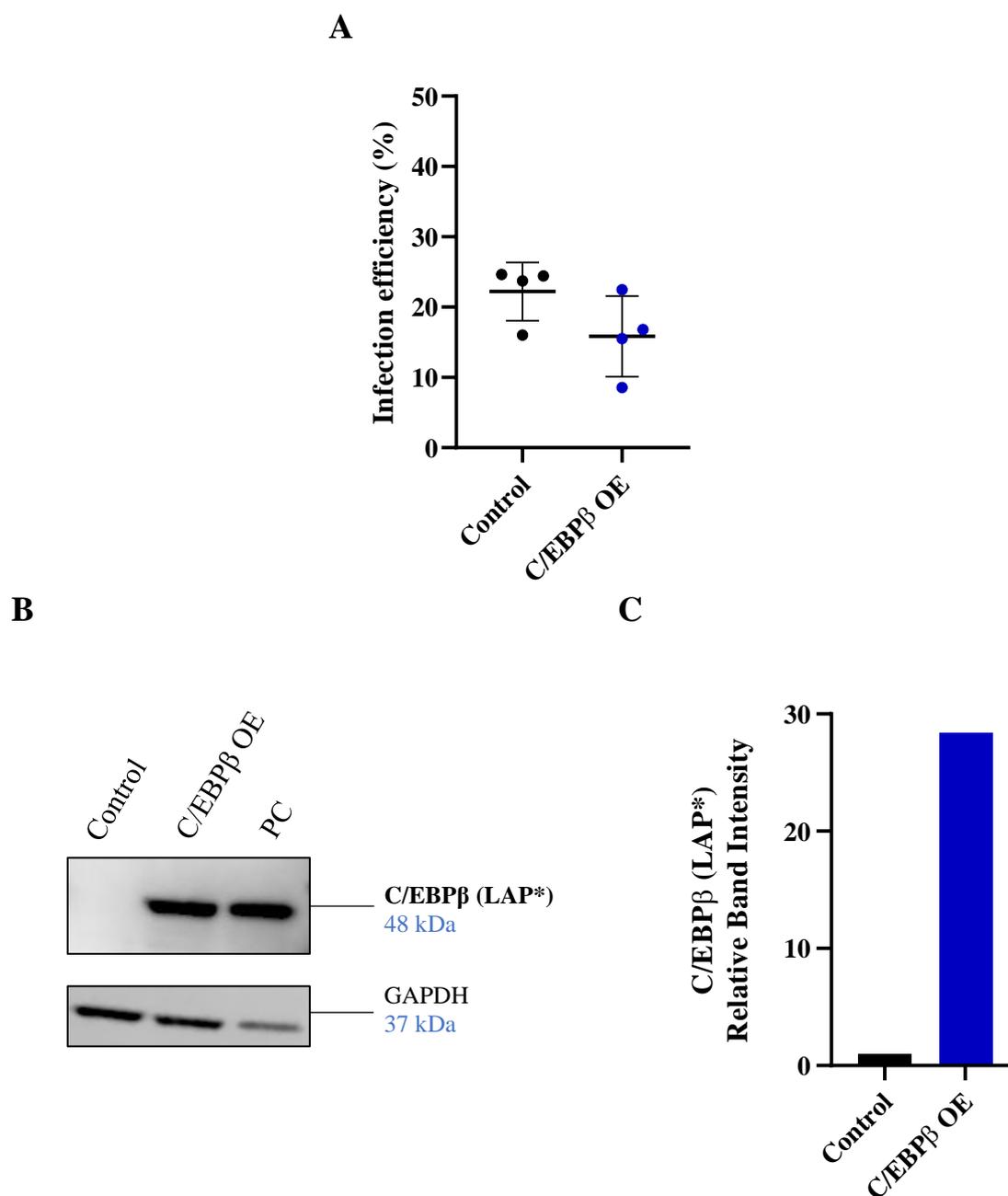
### 5.3.4.1 *Generation of C/EBP $\beta$ -overexpressing HSPC*

To promote the ectopic expression of the C/EBP $\beta$  LAP\* protein in HSPC, a C/EBP $\beta$  overexpression construct was used (**Table 2.1**), coupled with a GFP selectable marker, to allow for cell sorting. HSPC were transduced with the control or overexpression construct and, on day 3 of cell culture, efficiency of infection was evaluated by flow cytometry through GFP expression. Efficiency of transduction of these cultures was c25% for cells transduced with a control vector and c18% of GFP positivity for the C/EBP $\beta$  overexpression construct (**Figure 5.16A**). To validate C/EBP $\beta$  overexpression in these cells through western blot analysis, cells were sorted by FACS based on their GFP expression. C/EBP $\beta$  transduced cells strongly overexpressed the C/EBP $\beta$  LAP\* protein, as compared to the control cells (**Figure 5.16B-C**).



**Figure 5.15 – KD of C/EBP $\beta$  has no significant effect on normal cell cycle or apoptosis**

(A) Graph showing cell cycle distribution of normal CD34<sup>+</sup> HSPC (Control) and three shC/EBP $\beta$  cultures after 6 days of culture. Initial gating was applied as described in **Figure 2.6**. Data indicates mean  $\pm$  1SD (n=3). (B) Summary data showing the effect of C/EBP $\beta$  KD on apoptosis, in HSPC following 6, 10 and 13 days of culture. Initial gating was applied as described in **Figure 2.7**; (i) Early apoptotic cells are characterised by Annexin V<sup>+</sup> / PI<sup>+</sup> and (ii) Apoptotic cells are characterised by Annexin V<sup>+</sup> / PI<sup>+</sup>. Data indicates Mean  $\pm$  1SD (n=3). Significant difference between shC/EBP $\beta$  cultures and control was analysed by one-way ANOVA with Bonferroni's multiple comparisons correction; \* denotes  $p < 0.05$ .



**Figure 5.16 – C/EBP $\beta$  overexpression in human HSPC**

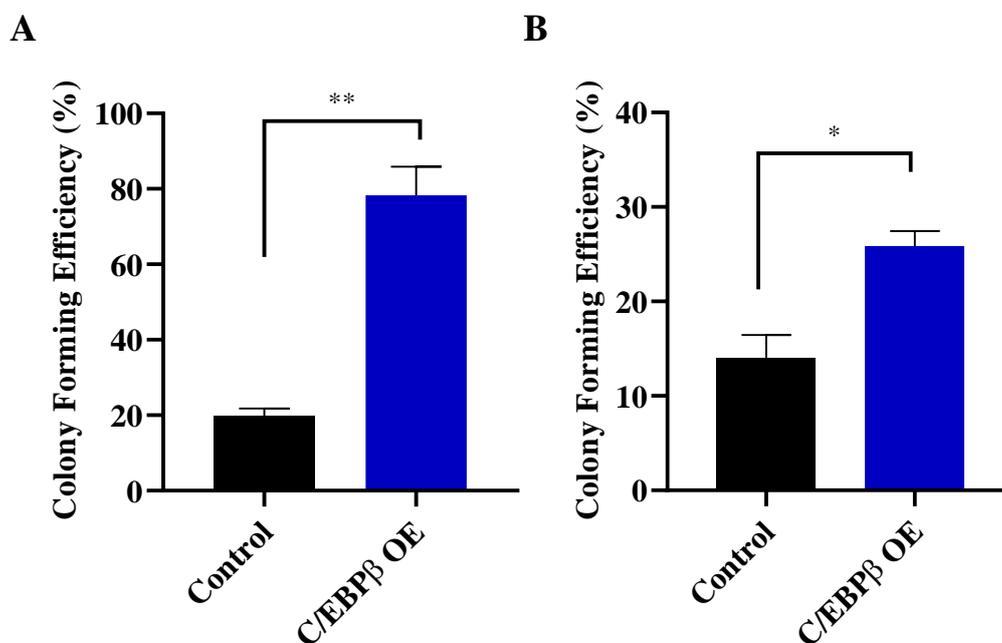
(A) Summary data showing percentages of GFP positivity in control and C/EBP $\beta$  overexpression CD34<sup>+</sup> HSPC populations, analysed on day 3 of culture. Gating strategies were applied as described in **Figure 2.4**. Data indicates mean  $\pm$  1SD (n=4). (B) Western blot analysis showing C/EBP $\beta$  protein expression in sorted cells transduced with a control or C/EBP $\beta$ -overexpression construct, on day 15 of culture (n=1). HEK 293T transfected with a C/EBP $\beta$ -overexpression construct was used as a positive control (PC); GAPDH was used as a loading control. (C) Bar chart showing C/EBP $\beta$  expression in cells transduced with a C/EBP $\beta$ -overexpression construct. C/EBP $\beta$  expression was quantified using ImageJ, normalised to loading control (GAPDH) and to control (n=1).

#### 5.3.4.2 *C/EBP $\beta$ overexpression promotes myeloid colony formation and self-renewal ability*

Having generated and validated the overexpression of C/EBP $\beta$  in human HSPC, the consequences on myeloid colony forming capacity and self-renewal ability were examined by performing a limiting dilution assay. To do this, GFP<sup>+</sup> sorted cells were cultured in the presence of IL-3, SCF, G- and GM-CSF (2.3.4). Under clonal conditions, the colony-forming efficiency was 4-fold higher in C/EBP $\beta$  cultures compared to control cells (**Figure 5.17A**). Serial replating of colony forming cells showed that C/EBP $\beta$ -overexpressing cells were able to generate 2-fold more colonies compared to control cells (**Figure 5.17B**). These results indicate that overexpression of C/EBP $\beta$  increases myeloid colony formation with and suggest an enhanced self-renewal potential.

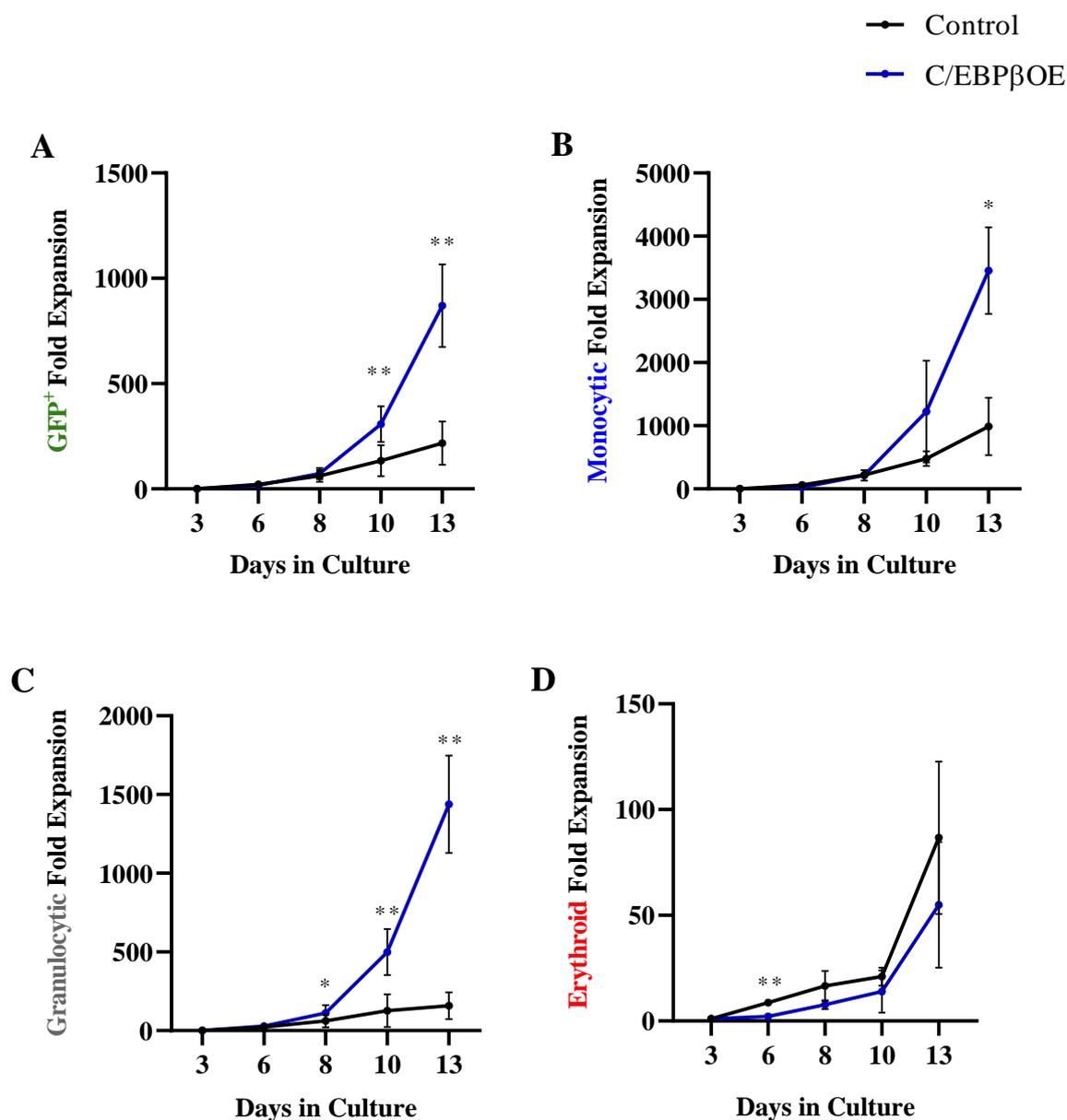
#### 5.3.4.3 *C/EBP $\beta$ promotes myeloid cell growth*

Having determined that overexpression of C/EBP $\beta$  promoted myeloid colony formation, the growth and differentiation of HSPC was followed over 13 days. C/EBP $\beta$  overexpression promoted the expansion of myeloid cells (4-fold) compared to control cells, in bulk liquid culture (**Figure 5.18A**). Using the lineage discriminator markers CD13 and CD36 to discriminate haematopoietic lineages, C/EBP $\beta$ -overexpressing monocytes displayed a 3.5-fold increase in normal growth, compared to cells transduced with a control vector (**Figure 5.18B**). This increase was more evident in the granulocytic population, in which cell expansion was increased by 9-fold in C/EBP $\beta$  overexpression cells compared to control (**Figure 5.18C**). As expected, since the media used for these experiments does not support terminal erythroid differentiation, the growth of erythroid progenitor cells was not significantly altered (**Figure 5.18D**).



**Figure 5.17 – C/EBPβ overexpression promotes myeloid colony formation of HSPC**

(A) Myeloid colony forming efficiency of control and C/EBPβ -overexpression (OE) cultures following 7 days of growth in liquid culture containing IL-3, SCF, G- and GM-CSF (2.3.4). Data indicates mean  $\pm$  1SD (n=3). Significant difference between C/EBPβ-overexpression and control cultures were analysed by paired t-test; \*\* denotes  $p < 0.01$ . (B) Self-renewal potential was assessed by a single replating round of control and C/EBPβ-OE cultures, in the same conditions as previously. Data indicates mean  $\pm$  1SD (n=3) Significant difference between C/EBPβ-overexpression and control cultures was analysed paired t-test; \* denotes  $p < 0.05$ .



**Figure 5.18 – Overexpression of C/EBP $\beta$  promotes monocytic and granulocytic growth in myeloid development**

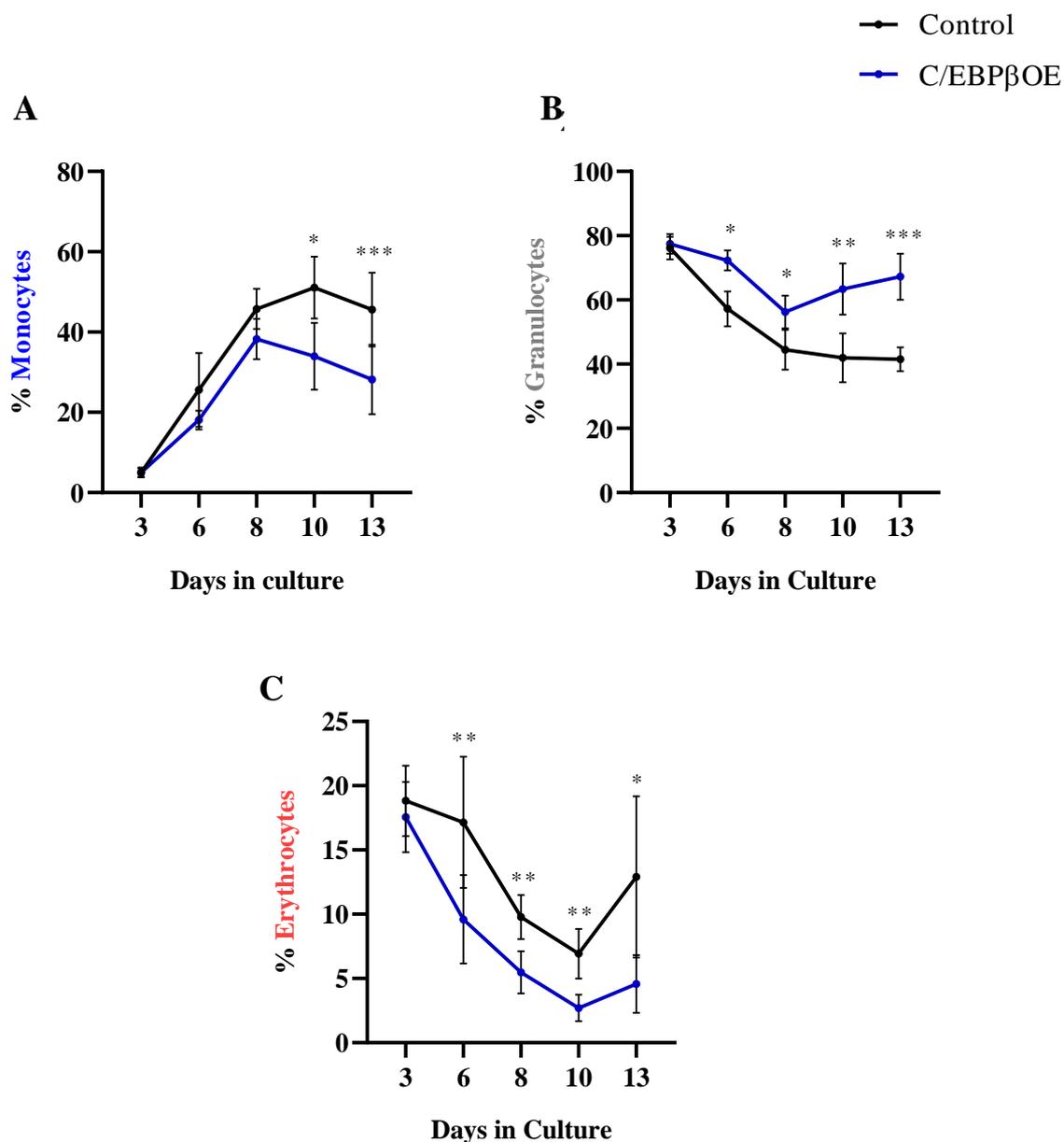
(A) Cumulative fold-expansion of control and C/EBP $\beta$  -overexpression constructs in terms of GFP positivity in culture medium containing IL-3, SCF, G- and GM-CSF, grown over 13 days (**Supplementary Figure 23**). (B) Cumulative fold-expansion of monocytic cells (CD13<sup>+</sup> CD36<sup>+</sup>) transduced with a control or C/EBP $\beta$ -OE vector. (C) Cumulative fold-expansion of granulocytic cells (CD13<sup>+</sup> CD36<sup>-</sup>) transduced with a control or C/EBP $\beta$ -OE vector. (D) Cumulative fold-expansion of erythroid cells (CD13<sup>-</sup> CD36<sup>+</sup>) transduced with a control of C/EBP $\beta$ -OE vector. Cells were re-seeded at the appropriate density following each time-point, as described in 2.3.3. Gating strategies were applied as described in **Figure 2.4** and **Figure 2.5**. Data indicates mean  $\pm$  1SD (n=4). Significant differences between C/EBP $\beta$ -overexpression and control cultures at each time-point were analysed by paired T-test; \* denotes p<0.05; \*\* denotes p<0.01; (this data has also been represented as dot plots, within the supplementary section [**Supplementary Figure 42**]).

#### 5.3.4.4 *C/EBP $\beta$ expression promotes myeloid differentiation*

KD of C/EBP $\beta$  disrupted the balance of haematopoietic populations, by increasing the proportion of monocytes, accompanied by a decrease in granulocytes (**Figure 5.11**). Overexpression of C/EBP $\beta$ , on the other hand, had a contrasting effect, significantly decreasing the proportion of the monocytic population, especially in the latter stages of myeloid development, coupled with an increase in the percentage of granulocytic committed cells throughout myeloid development (**Figure 5.19A-B**). Furthermore, C/EBP $\beta$  also contributed to a significant decrease in erythroid population, even though this was not the focus of the study (**Figure 5.19C**). Since cell proliferation and differentiation are often inversely correlated, it is possible that the increase in cell proliferation arising from C/EBP $\beta$  overexpression is due to an inhibition in HSPC differentiation in the monocytic and granulocytic lineages.

Immunophenotypic analysis of monocyte-progenitor cells showed no difference in the percentage of cells expressing the haematopoietic stem-cell marker CD34 (**Figure 5.20A**). However, a consistent increase above control in the expression of CD11b was observed in C/EBP $\beta$  expressing cells, culminating in a 1.5-fold upregulation in the expression of this marker compared to control (**Figure 5.20B**). Upregulation of the monocytic marker CD14 was further observed in these cells; however, this trend was solely detected within the first 8 days of myeloid growth, decreasing to similar levels to those of control cells afterwards (**Figure 5.20C**). No significant differences were observed in terms of granularity or complexity of C/EBP $\beta$ -overexpressing monocytes (**Figure 5.20D**). These observations suggest that, despite the significant increase in monocytic growth, C/EBP $\beta$  appears to promote monocytic differentiation. This phenotype further agrees with the mRNA and protein analysis, in which C/EBP $\beta$  expression increased with myeloid cell differentiation, particularly towards monocytic committed cells (**Figure 5.1**, **Figure 5.2**).

In contrast, C/EBP $\beta$  overexpression did not significantly impact the expression of CD34 or CD11b in granulocytic progenitor cells (**Figure 5.21A-B**). Overexpression of C/EBP $\beta$  in this population did, however, lead to the upregulation of the granulocytic marker CD15, up until day 10 of culture, in which a 1.5-fold upregulation was observed compared to control cells (**Figure 5.21C**). Similarly, to the monocytic population, no significant differences were observed in terms of cell granularity / complexity in these cells (**Figure 5.21D**).



**Figure 5.19 – C/EBP $\beta$  overexpression disrupts the balance between myeloid population during haematopoietic development**

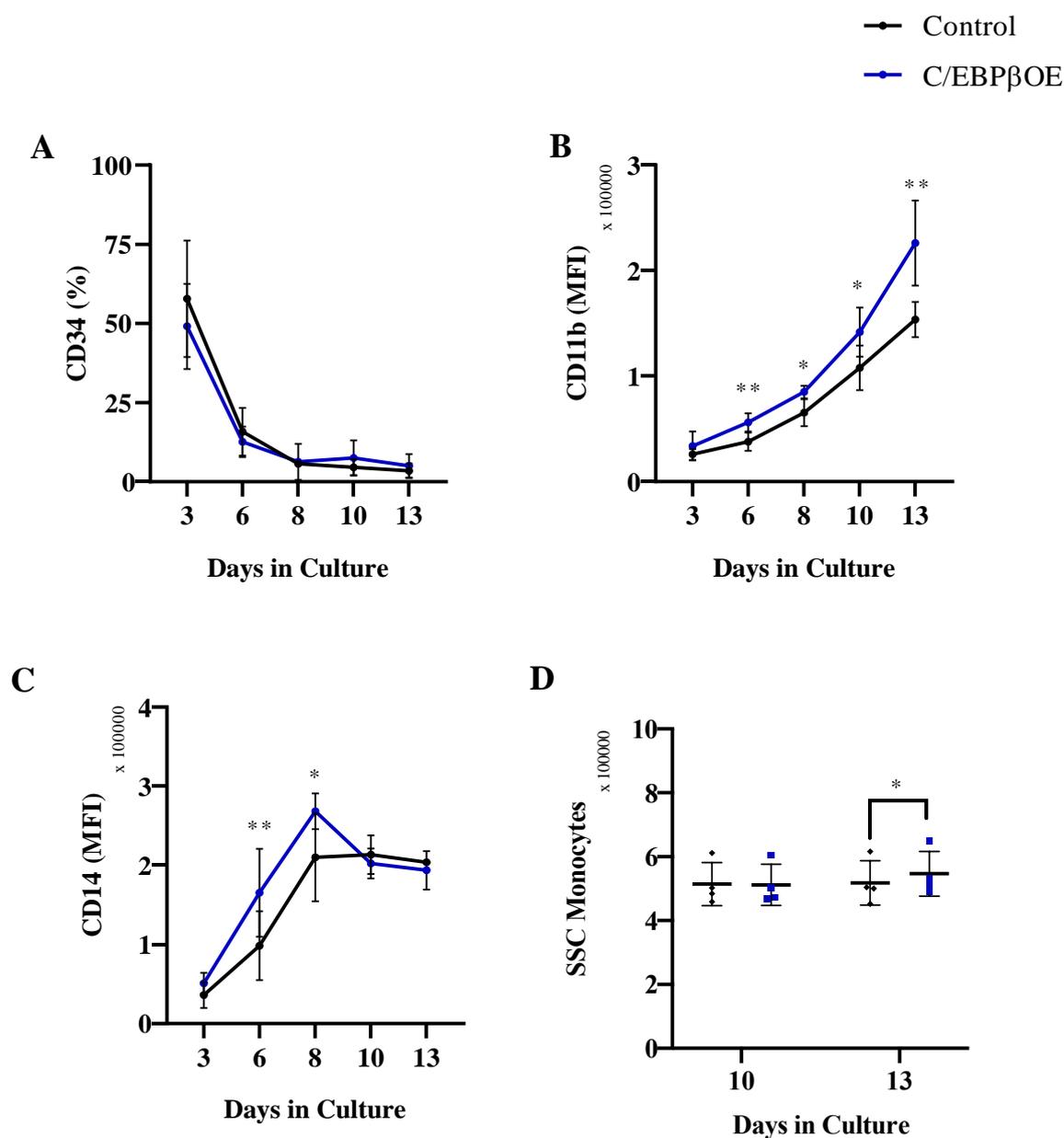
Summary data showing the proportion of (A) monocytic (CD13<sup>+</sup> CD36<sup>+</sup>), (B) granulocytic (CD13<sup>+</sup>/CD36<sup>-</sup>) and (C) erythroid cells (CD13<sup>-</sup> CD36<sup>+</sup>) in control and C/EBP $\beta$ -overexpressing cultures. Cells were re-seeded at the appropriate density following each time-point, as described in 2.3.3. Gating strategies were applied as described in Figure 2.4 and Figure 2.5. Data indicates mean  $\pm$  1SD (n=4). Significant differences between C/EBP $\beta$ -expressing cells and control at each time-point were analysed by paired t-test; \* denotes  $p < 0.05$ , \*\* denotes  $p < 0.01$ , \*\*\* denotes  $p < 0.001$ ; (this data has also been represented as dot plots, within the supplementary section [Supplementary Figure 43]).

This study next sought to assess terminal stages of development, difficult to score by flow cytometry techniques only. This analysis showed that overexpression of C/EBP $\beta$  promotes granulocytic development, with a significant increase in the percentage of cells showing mature granulocytic traits, such as segmented nuclei and cytoplasmic granules (**Figure 5.22**). Moreover, the percentage of monocytes and macrophages in the C/EBP $\beta$ -overexpressing culture was significantly lower compared to the control population, in agreement with the lineage imbalance described above (**Figure 5.19**).

Altogether, these results suggest that ectopic expression of C/EBP $\beta$  LAP\* in HSPC promotes monocytic and granulocytic differentiation, although the latter to a lesser extent. Moreover, the fact that C/EBP $\beta$  LAP\* overexpression in HSPC promoted differentiation suggests that its upregulation is not likely to promote leukaemogenesis through the perturbation of myeloid cell development.

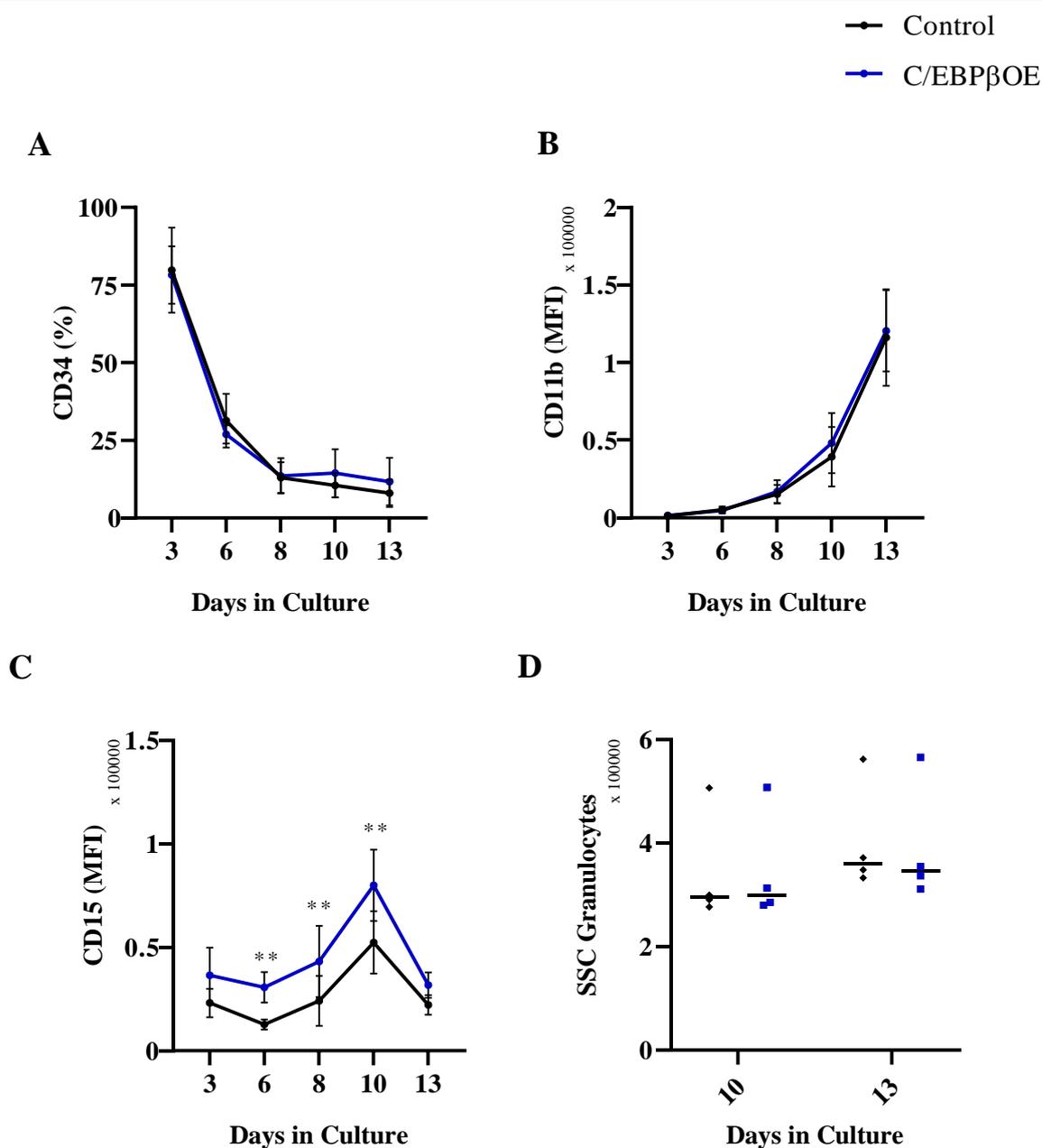
#### *5.3.4.5 Overexpression of C/EBP $\beta$ induces change in the cell cycle of HSPC*

Having determined that ectopic C/EBP $\beta$  expression leads to an increase in cell growth, not associated with an arrest in differentiation, this study aimed at analysing if these changes partially occurred as a result of variations to the normal cells' cell cycle and apoptotic frequency. Cell cycle analysis showed that overexpression of C/EBP $\beta$  led to a significant decrease by 12% in the percentage of cells in the G<sub>1</sub> phase of the cell cycle and, conversely, to an increase of cell in S/G<sub>2</sub> (**Figure 5.23A**). In addition, no significant differences were observed regarding apoptosis (**Figure 5.23B**). In summary, this data suggest that overexpression of C/EBP $\beta$  in human HSPC promotes the expansion of monocytes and granulocytes, whilst promoting the differentiation of both lineages in a myeloid setting.



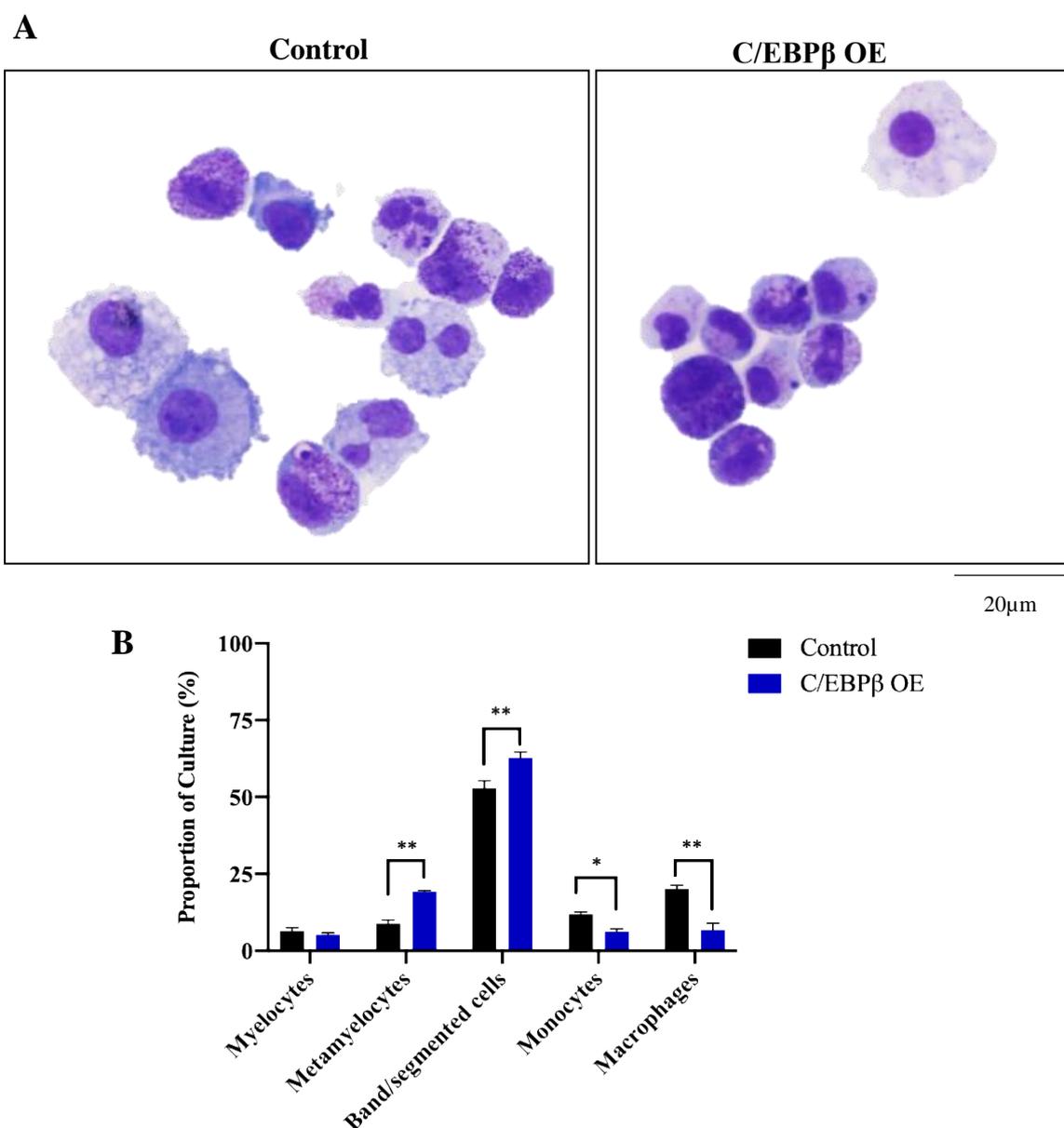
**Figure 5.20 – C/EBP $\beta$  overexpression promotes the upregulation of monocytic markers in monocytic progenitors**

(A) Summary data of CD34<sup>+</sup> expression in terms of percentage in monocytic cells (CD13<sup>+</sup> CD36<sup>+</sup>) over time for control and C/EBP $\beta$ -overexpressing cells. Summary data showing (B) CD11b and (C) CD14 expression in terms of MFI in monocytic cells over time for control and C/EBP $\beta$ -overexpressing cells. (D) Complexity of monocytic cells was measured by changes in SSC for control and C/EBP $\beta$ -overexpressing cells on days 10 and 13 of myeloid differentiation. Cells were re-seeded at the appropriate density following each time-point, as described in 2.3.3. Gating strategies were applied as described in Figure 2.4 and Figure 2.5. Data indicates mean  $\pm$  1SD (n=4). Significant differences between C/EBP $\beta$ -expressing cells and control at each time-point were analysed by paired t-test; \* denotes  $p < 0.05$ , \*\* denotes  $p < 0.01$ ; (for representative flow cytometry plots please refer to Supplementary Figure 24, Supplementary Figure 25, Supplementary Figure 26; this data has also been represented as dot plots, within the supplementary section [Supplementary Figure 44]).



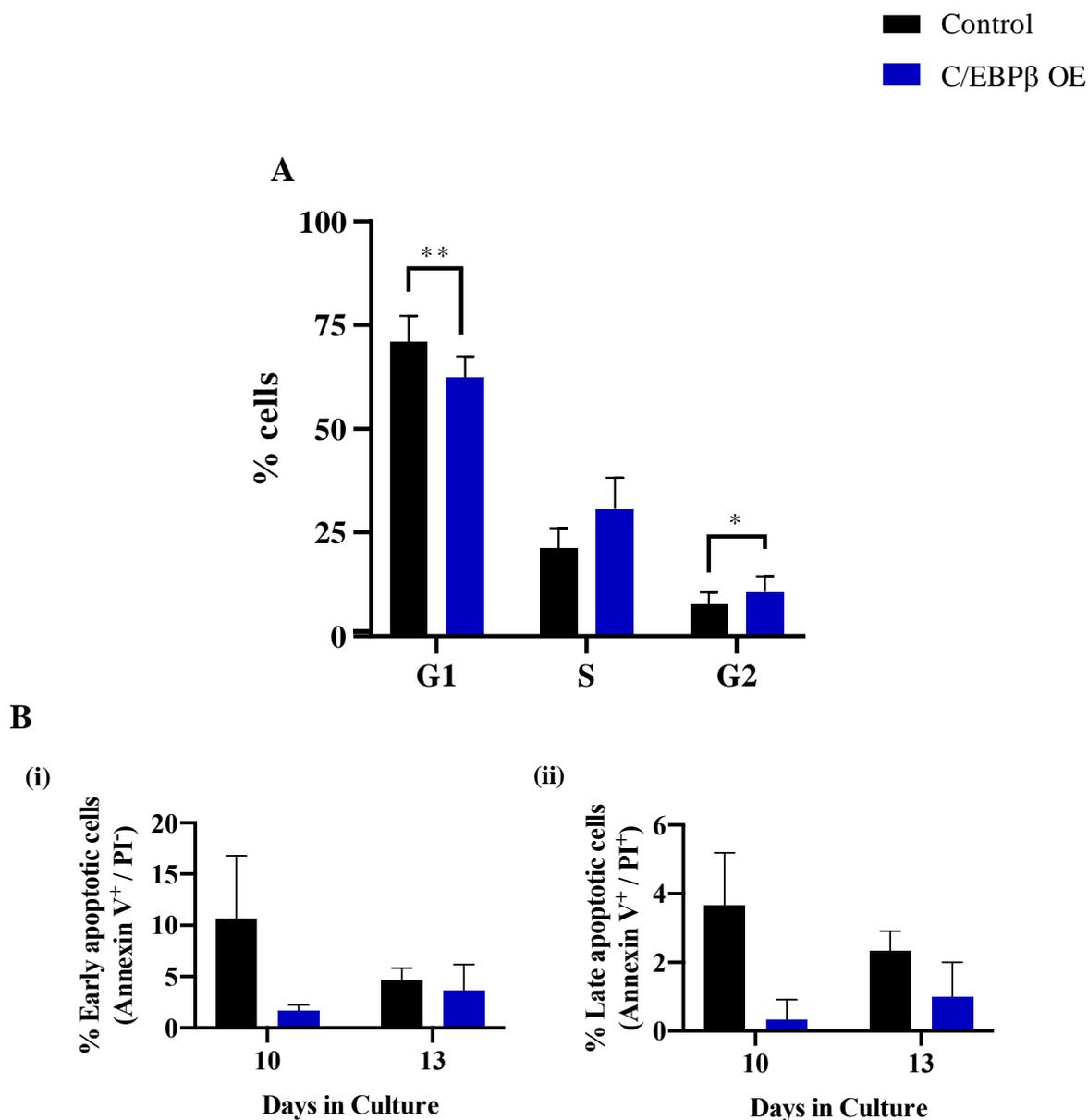
**Figure 5.21 – C/EBP $\beta$  overexpression promotes the upregulation of granulocytic markers in granulocytic progenitors**

(A) Summary data of CD34<sup>+</sup> expression in terms of percentage in granulocytic cells (CD13<sup>+/−</sup> CD36<sup>−</sup>) over time for control and C/EBP $\beta$ -overexpressing cells. Summary data showing (B) CD11b and (C) CD15 expression in terms of MFI in granulocytic cells over time for control and C/EBP $\beta$ -overexpressing cells. (D) Complexity of granulocytic cells was measured by changes in SSC for control and C/EBP $\beta$ -overexpressing cells on days 10 and 13 of myeloid differentiation. Cells were re-seeded at the appropriate density following each time-point, as described in 2.3.3. Gating strategies were applied as described in Figure 2.4 and Figure 2.5. Data indicates mean  $\pm$  1SD (n=4). Significant differences between C/EBP $\beta$ -expressing cells and control at each time-point were analysed by paired t-test; \*\* denotes  $p < 0.01$ ; (for representative flow cytometry plots please refer to Supplementary Figure 27, Supplementary Figure 28, Supplementary Figure 29; this data has also been represented as dot plots, within the supplementary section [Supplementary Figure 45]).



**Figure 5.22 – C/EBPβ overexpression alters cell morphology during myeloid development**

(A) Representative images showing control and C/EBPβ OE cultures, analysed on day 17 of differentiation, by cytospinning  $3 \times 10^4$  cells and staining them with May-Grünwald-Giemsa differential stain. (B) Differential counts of all cultures with morphology categorised into granulocytic and monocytic-committed cells, according to **Figure 2.2**. Approximately 500 cells were counted in total and each population count normalised to total number of cells. Granulocytic cells were further divided into myelocytes, metamyelocytes and band/segmented cells. Data indicates mean  $\pm$  1SD (n=3). Significant differences between C/EBPβ-expressing cells and control were analysed by paired t-test; \* denotes  $p < 0.05$ , \*\* denotes  $p < 0.01$ .



**Figure 5.23 – Overexpression of C/EBP $\beta$  influences normal cell cycle**

(A) Bar chart showing cell cycle distribution of normal CD34<sup>+</sup> HSPC (Control) and C/EBP $\beta$ -overexpression cultures after 6 days of culture. Initial gating was applied as described in **Figure 2.6**. The percentage of cells in each phase of the cell cycle is indicated by mean  $\pm$  1SD (n=3). Significant difference between C/EBP $\beta$ -overexpression and control cultures were analysed paired t-test; \* denotes  $p < 0.05$ ; \*\* denotes  $p < 0.01$ . (B) Summary data showing the effect of C/EBP $\beta$  overexpression on apoptosis in HSPC following 10 and 13 days of culture. Initial gating was applied as described in **Figure 2.7**; (i) Early apoptotic cells are characterised by Annexin V<sup>+</sup> / PI<sup>-</sup> and (ii) Apoptotic cells are characterised by Annexin V<sup>+</sup> / PI<sup>+</sup>. Data indicates mean  $\pm$  1SD (n=3).

### 5.3.5 The effect of C/EBP $\beta$ mis-regulation in AML cell lines

#### 5.3.5.1 Expression of C/EBP $\beta$ is variable across AML cell lines

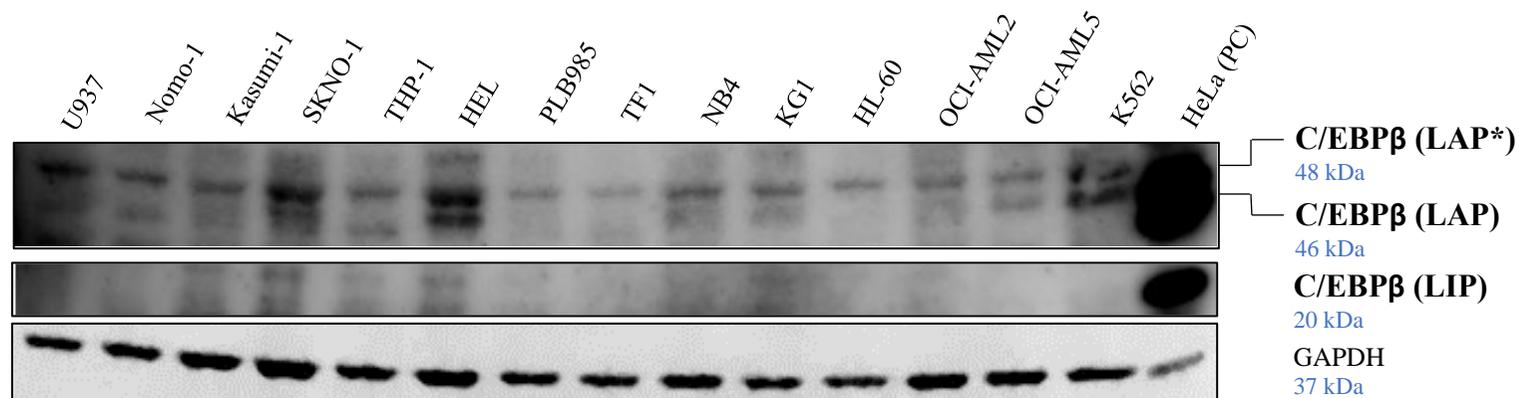
In order to determine the expression of C/EBP $\beta$  in AML cell lines, thus establishing whether these were suitable to study the biological role of this protein in AML, endogenous mRNA and protein levels were assessed in a cohort of AML cell lines. The CCLE database was firstly employed to determine the expression of *CEBPB* mRNA in AML cell lines. Since the *CEBPB* transcript consists of a single exon, giving rise to a single mRNA sequence, it is not possible to determine the mRNA expression of each of the three C/EBP $\beta$  isoforms. *CEBPB* was found to be heterogeneously expressed in all the cell lines tested, including in the t(8;21) cell lines Kasumi-1 and SKNO-1 (**Figure 5.24A**). Subsequently, C/EBP $\beta$  protein expression was determined using western blotting. **Figure 5.24B** shows three cell lines (U937, HEL and SKNO-1) as suitable for KD studies, and an additional three lines (THP-1, OCI-AML 2 and OCI-AML5) with lower levels of endogenous C/EBP $\beta$  appropriate for overexpression studies. Moreover, this analysis allowed the detection of two C/EBP $\beta$  isoforms, LAP\* and LAP, in almost all the AML cell lines tested. Additional comparison between C/EBP $\beta$  mRNA and protein showed poor association between the two (**Figure 5.24A**). In summary, it was possible to identify suitable AML cell lines for to study the consequences of modulating C/EBP $\beta$  expression using overexpression, which could not be achieved in this study (**5.4.5**), or KD systems (**5.3.5.2**).

#### 5.3.5.2 Generation of C/EBP $\beta$ -knockdown cell lines

The consequences of knocking down C/EBP $\beta$  in AML cell lines on proliferation and survival was determined. Using the same shRNA constructs (shRNA control and three shC/EBP $\beta$  vectors), three cell lines were selected for the generation of C/EBP $\beta$ -KD cultures based on the endogenous levels of C/EBP $\beta$  expression. These included the t(8;21) SKNO-1, the erythroleukaemia HEL and the leukaemic U937 cell lines. Successful cell line transduction was confirmed through GFP expression, once bulk cell population achieved 100% GFP positivity (**Figure 5.25A**; **Figure 5.26A**; **Figure 5.27A**), without the need to select them.

KD of C/EBP $\beta$  in the generated cell lines was validated through western blotting. In all three cell lines, cells transduced with shC/EBP $\beta$  #1 showed the lowest level of KD, similarly to that observed in HSPC (**5.3.3.1**) (**Figure 5.25B**; **Figure 5.26B**; **Figure 5.27B**).

A



B

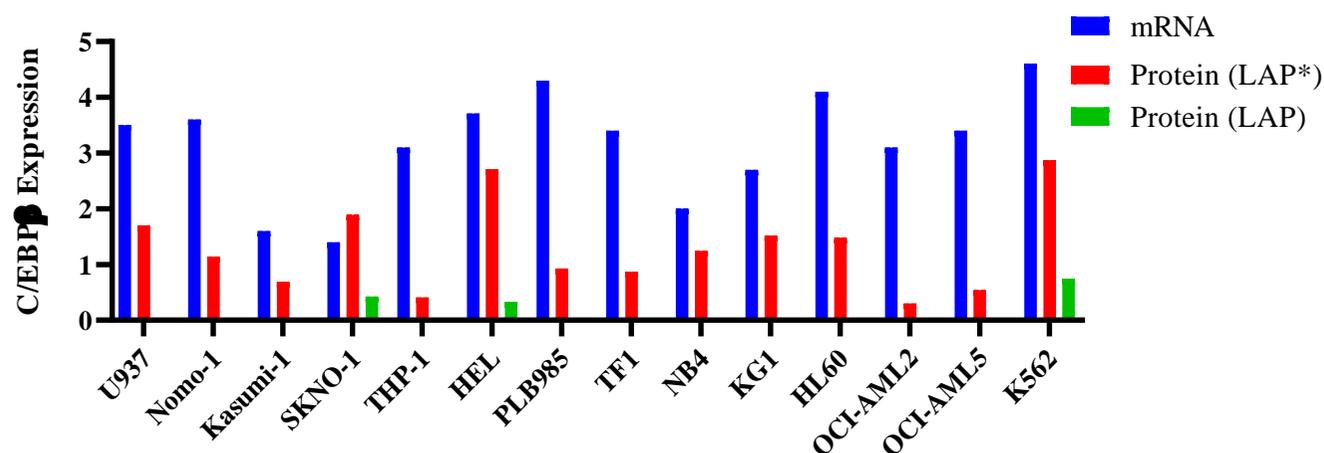
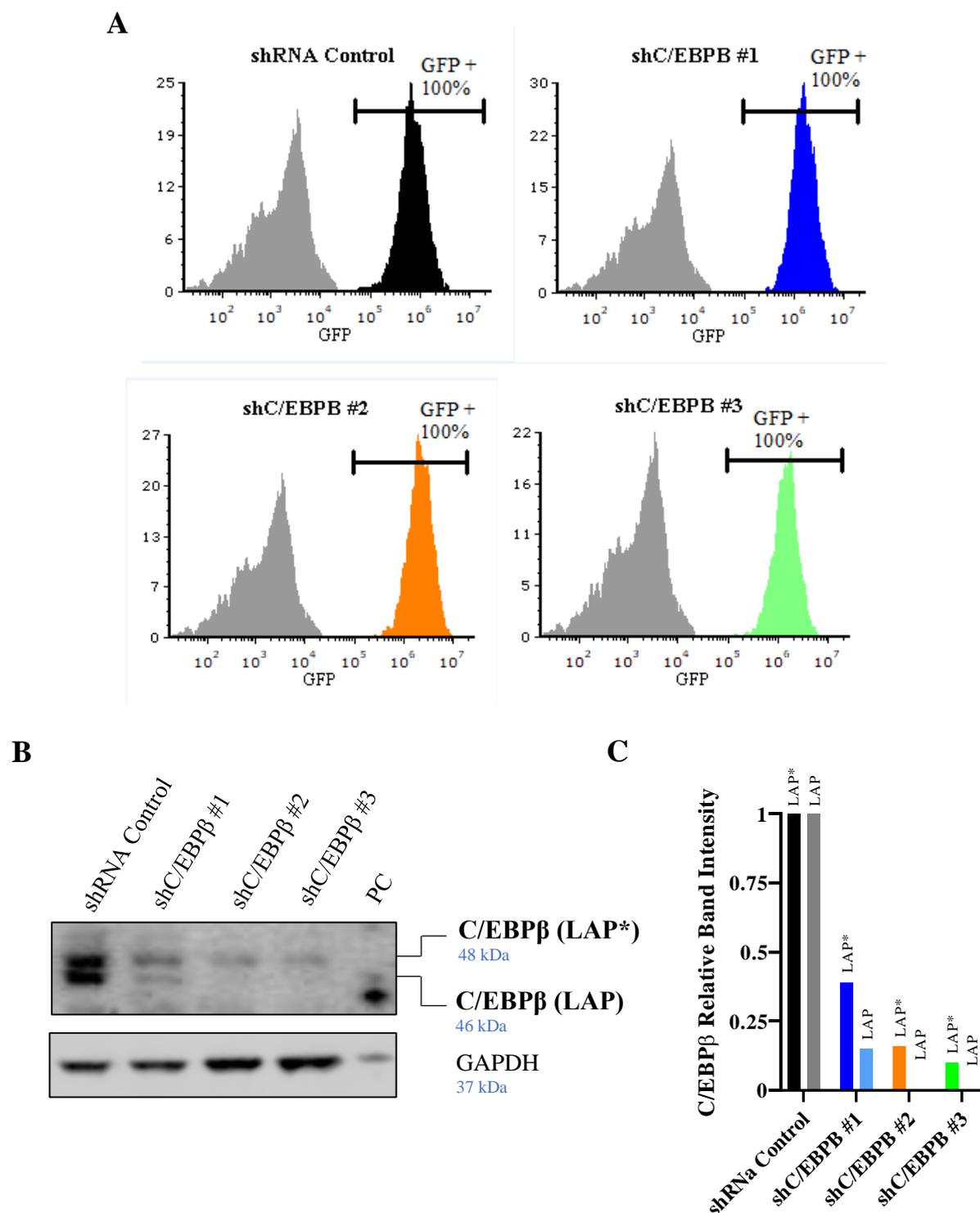


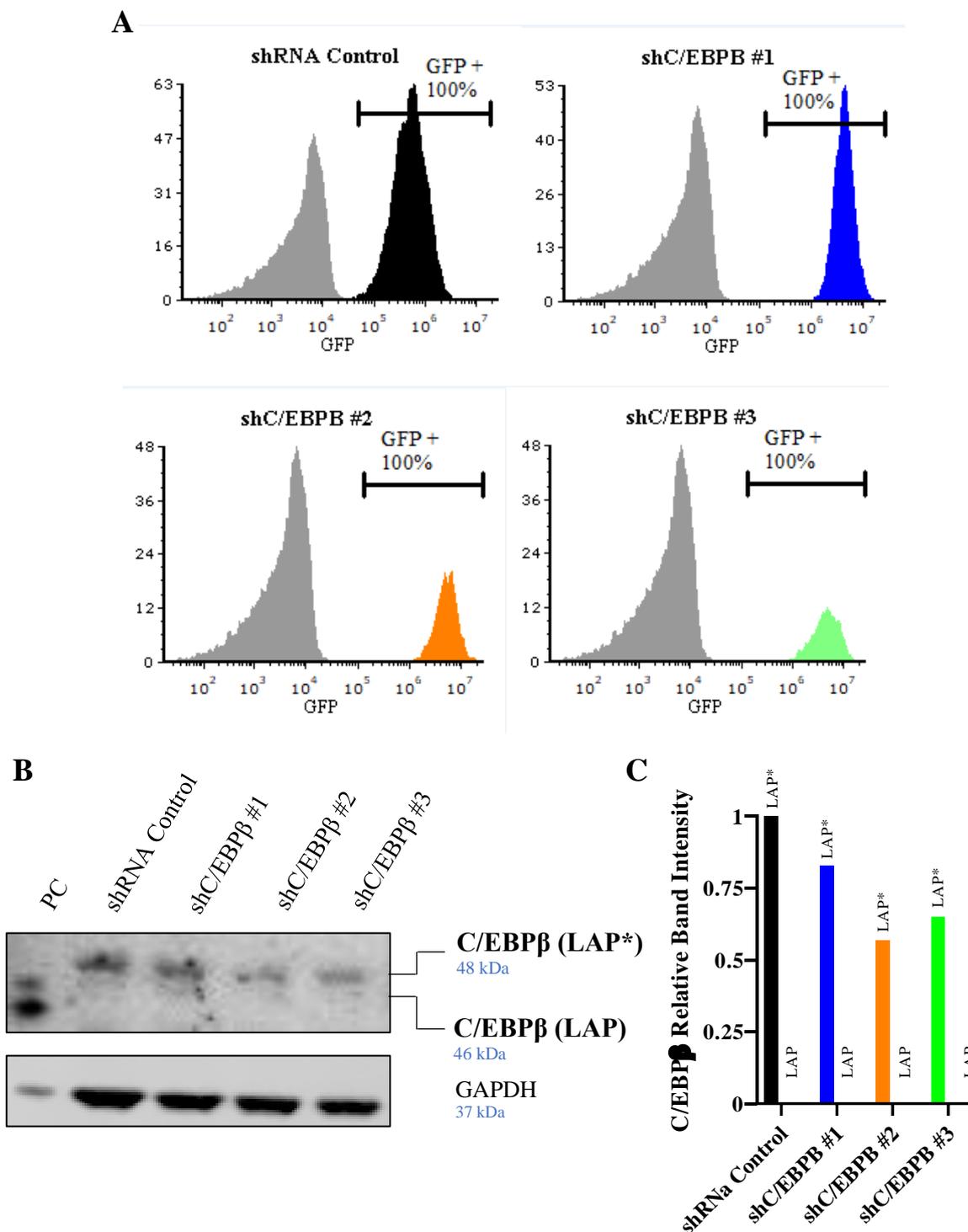
Figure 5.24 – C/EBP $\beta$  expression in leukaemic cell lines

(A) Western blot analysis of C/EBP $\beta$  endogenous expression in a cohort of AML cell lines ( $n=1$ ). The HeLa cell line was used as a positive control (PC); GAPDH was used as a loading control. (B) Normalised ( $\log_2$ ) *CEBPB* mRNA expression in a cohort of leukaemic cell lines. mRNA expression data derived from <https://depmap.org/portal/gene/CEBPB?tab=characterization> (Barretina *et al.*, 2012). Null protein values refer to undetectable levels of endogenous C/EBP $\beta$ , derived from (A).



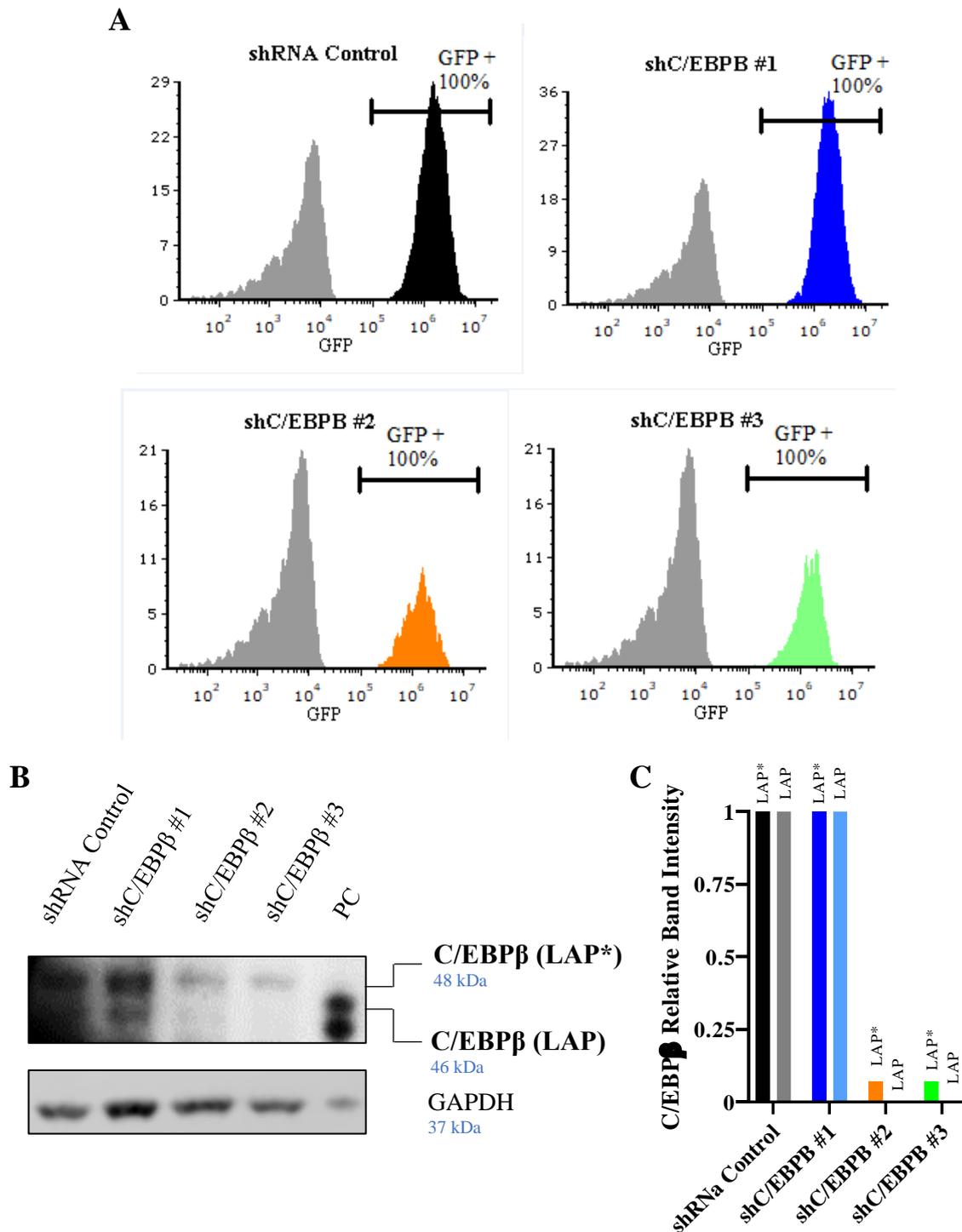
**Figure 5.25 – C/EBPβ KD in the t(8;21) cell line SKNO-1**

(A) Representative flow cytometric plots showing transduction efficiency in the SKNO-1 cell line transduced with an shRNA control and three shC/EBPβ constructs (shC/EBPβ #1, shC/EBPβ #2, shC/EBPβ #3). (B) Western blot analysis showing successful C/EBPβ KD in all three shRNA constructs (n=1). GAPDH was used as a loading control.; HeLa cells were used as a positive control (PC) (C) Western blot quantification by densitometry analysis using ImageJ. Relative protein expression was determined by normalising each sample firstly to the loading control (GAPDH), then to shRNA control (n=1).



**Figure 5.26 – C/EBPβ KD in the erythroleukaemia cell line HEL**

(A) Representative flow cytometric plots showing transduction efficiency in the HEL cell line transduced with an shRNA control and three shC/EBPβ constructs (shC/EBPβ #1, shC/EBPβ #2, shC/EBPβ #3). (B) Western blot analysis showing successful C/EBPβ KD in all three shRNA constructs (n=1). GAPDH was used as a loading control.; HeLa cells were used as a positive control (PC) (C) Western blot quantification by densitometry analysis using ImageJ. Relative protein expression was determined by normalising each sample firstly to the loading control (GAPDH), then to shRNA control (n=1).



**Figure 5.27 – C/EBPβ KD in the leukaemic cell line U937**

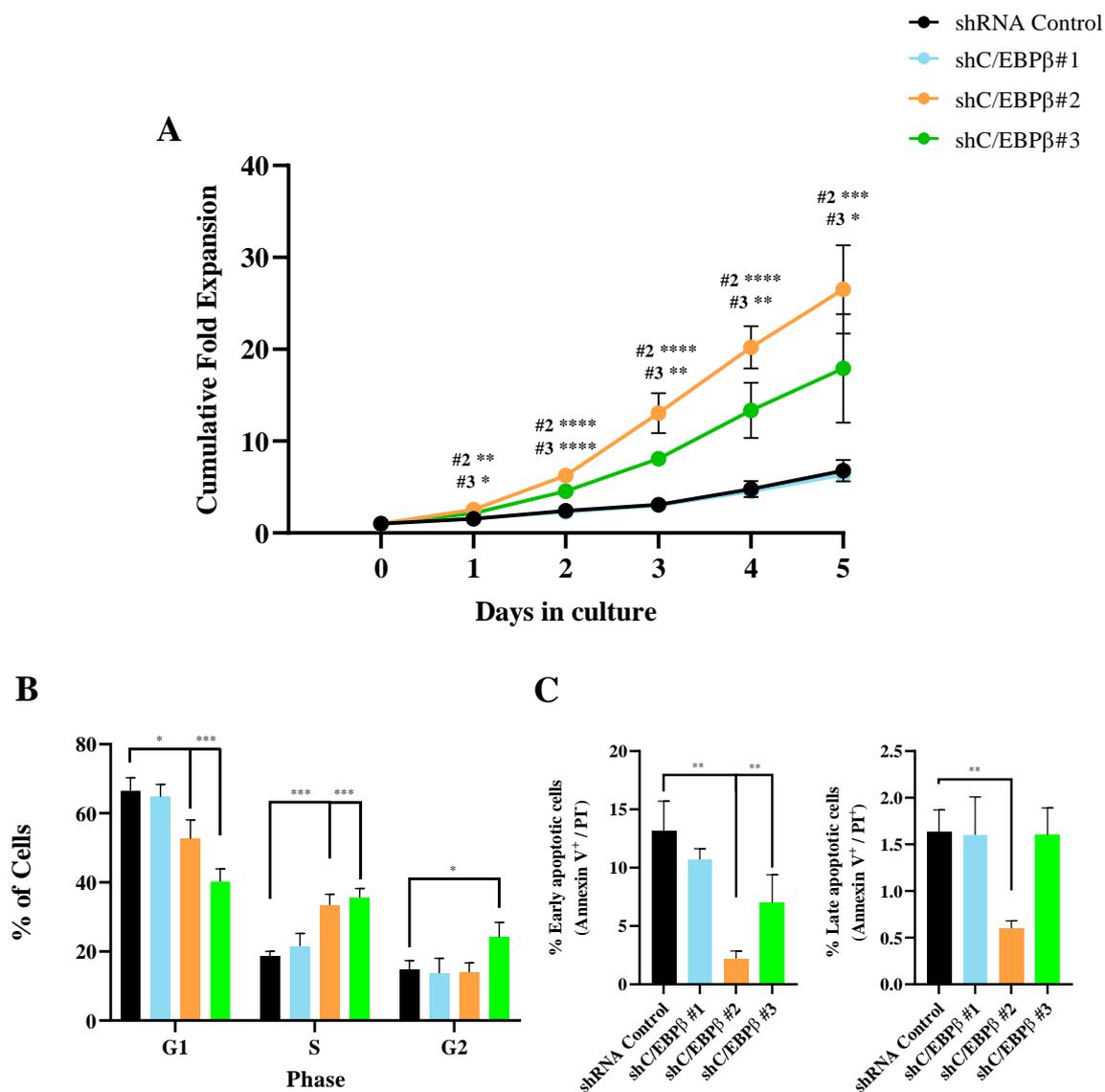
(A) Representative flow cytometric plots showing transduction efficiency in the U937 cell line transduced with an shRNA control and three shC/EBPβ constructs (shC/EBPβ #1, shC/EBPβ #2, shC/EBPβ #3). (B) Western blot analysis showing successful C/EBPβ KD in all three shRNA constructs (n=1). GAPDH was used as a loading control.; HeLa cells were used as a positive control (PC) (C) Western blot quantification by densitometry analysis using ImageJ. Relative protein expression was determined by normalising each sample firstly to the loading control (GAPDH), then to shRNA control (n=1).

Constructs #2 and #3 showed similar levels of KD, with a decrease of c75%, c25% and c90% in the detectable levels of C/EBP $\beta$  LAP\* in SKNO-1, HEL and U937 cells, respectively (**Figure 5.25C**; **Figure 5.26C**; **Figure 5.27C**), as compared to cells transduced with a control shRNA vector. The LAP isoform was shown to be completely KD in SKNO-1 and U937 cells; additionally, this isoform was not detected in the HEL cell line. The C/EBP $\beta$  LIP isoform was not detected in any of the cell lines tested.

#### 5.3.5.3 *The impact of C/EBP $\beta$ knockdown on the AML t(8;21) cell line SKNO-1*

Having demonstrated that C/EBP $\beta$  was successfully knocked-down in all three cell lines, the phenotypic impact of this, including cell growth and differentiation, were assessed by flow cytometry. Firstly, this study analysed the t(8;21) cell line SKNO-1. In these cells, KD of C/EBP $\beta$  resulted in a 4- and 2.6-fold significant increase in cell growth, observed in cells transduced with shC/EBP $\beta$  #2 and #3, respectively (**Figure 5.28A**). Given the significant impact of C/EBP $\beta$  KD in these cells, this study further analysed their DNA content, to determine cell cycle progression. This showed an increase in the proportion of cells in the S/G<sub>2</sub> phase of the cell cycle, as compared to control (**Figure 5.28B**). Subsequently, these lines were assayed for apoptosis using Annexin V and 7-ADD, showing a significant decrease in the percentage of cells in early and late apoptosis in cultures transduced with shC/EBP $\beta$  #2 and #3, compared to control (**Figure 5.28C**). Altogether, these results indicate that the increase in cell proliferation observed upon KD of C/EBP $\beta$  can be attributed to changes in normal cell cycle, together with a decreased apoptotic rate.

To determine if KD of C/EBP $\beta$  had any effect on the differentiation status of transformed lines, shC/EBP $\beta$ -transduced cells were analysed using immunophenotypic assays. A broad panel of markers associated both with immature HSPC (such as CD34), as well as markers of monocytic (CD11b, CD14), granulocytic (CD11b, CD15) and erythroid (GlyA) differentiation was used. According to this analysis, SKNO-1 control cells are typically CD13<sup>+</sup>CD33<sup>+</sup>CD38<sup>+</sup>CD45RA<sup>+</sup>, resembling a GMP phenotype (**Figure 1.1**) (**Figure 5.29**). However, upon KD of C/EBP $\beta$ , these cells lose the expression of all myeloid markers, being defined as CD13<sup>-</sup>CD33<sup>-</sup>CD38<sup>-</sup>CD45RA<sup>-</sup>, suggesting a more immature phenotype associated with HSC or MPP.



**Figure 5.28 – C/EBPβ KD promotes cell growth in SKNO-1 cells, associated with changes in normal cell cycle and apoptotic rate**

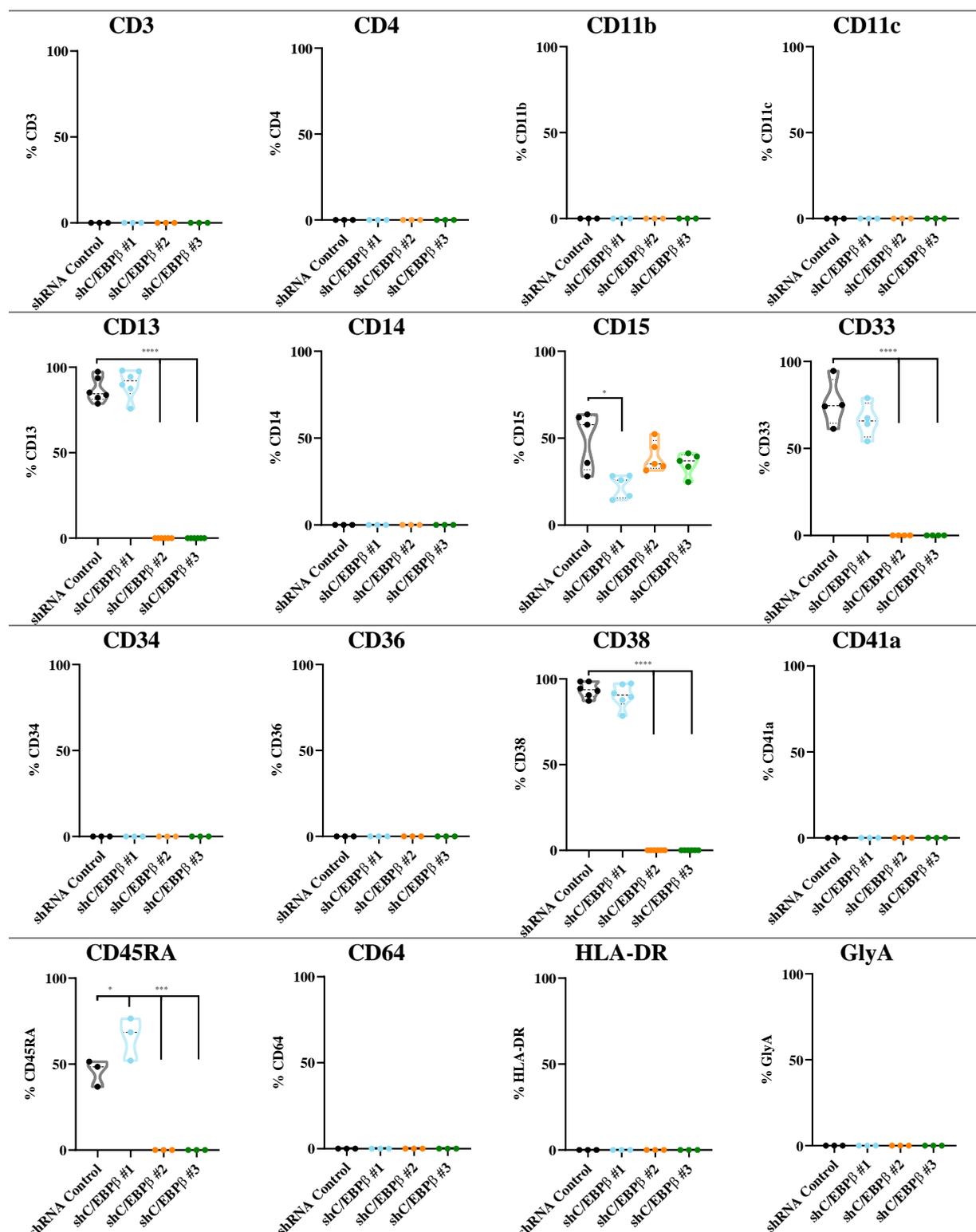
Summary data showing (A) the growth of SKNO-1 cells with three shC/EBPβ constructs compared to control, over 5 days of culture. Data indicates mean  $\pm$  1SD (n=4); (B) the effect of C/EBPβ KD in normal cell cycle. Data indicates Mean  $\pm$  1SD (n=3); (C) the effect of C/EBPβ KD on normal apoptotic rate using Annexin V staining. Bars indicate percentage of pre-apoptotic cells (Annexin V<sup>+</sup> / 7-AAD<sup>-</sup>) and late apoptotic cells (Annexin V<sup>+</sup> / 7-AAD<sup>+</sup>). Data indicates Mean  $\pm$  1SD (n=3). Significant differences between shC/EBPβ cultures and control were analysed using one-way ANOVA with Bonferroni's multiple comparisons correction; \* denotes p<0.05; \*\* denotes p<0.01.; \*\*\* denotes p<0.001; \*\*\*\* denotes p<0.0001.

Following immunophenotypic analysis, it was hypothesised that the increase in cell growth, associated with the loss of myeloid surface marker expression could be correlated with the promotion of a more undifferentiated phenotype. Undifferentiated cells are usually bigger and less internally complex than differentiated cells, thus the morphology of transduced SKNO-1 cells was assessed. This showed that SKNO-1 cells transduced with shC/EBP $\beta$  #2 and #3 were larger than control cells, or cells transduced with shC/EBP $\beta$  #1, the weakest construct (**Figure 5.30A**). This was further confirmed by analysing the cells average area, as well as through the forward and side scatter properties of these cells, with all three parameters showing a significant increase in cells transduced with shC/EBP $\beta$  #2 and #3, as compared to the shRNA control vector (**Figure 5.30B-D**).

Altogether, these results suggest that KD of C/EBP $\beta$  in the AML t(8;21) cell line SKNO-1 promotes cell growth, associated with an increase of cells in the S phase of the cell cycle, as well a decrease in the proportion of cells in early and late apoptotic stages. Moreover, KD of this protein was further associated with the ablation of myeloid cell surface marker expression which, in combination with morphological analysis, suggests that these cells show a regression in their differentiation state. However, additional cell surface markers would have to be analysed to further strengthen this hypothesis. Furthermore, subsequent studies in other t(8;21) cell lines would be necessary to definitively conclude how KD of C/EBP $\beta$  influences the cell's development and differentiation status, upon the expression of RUNX1-ETO. Even though this was attempted in the t(8;21) cell line Kasumi-1, no meaningful results were obtained, due to technical difficulties (**Supplementary Figure 46**).

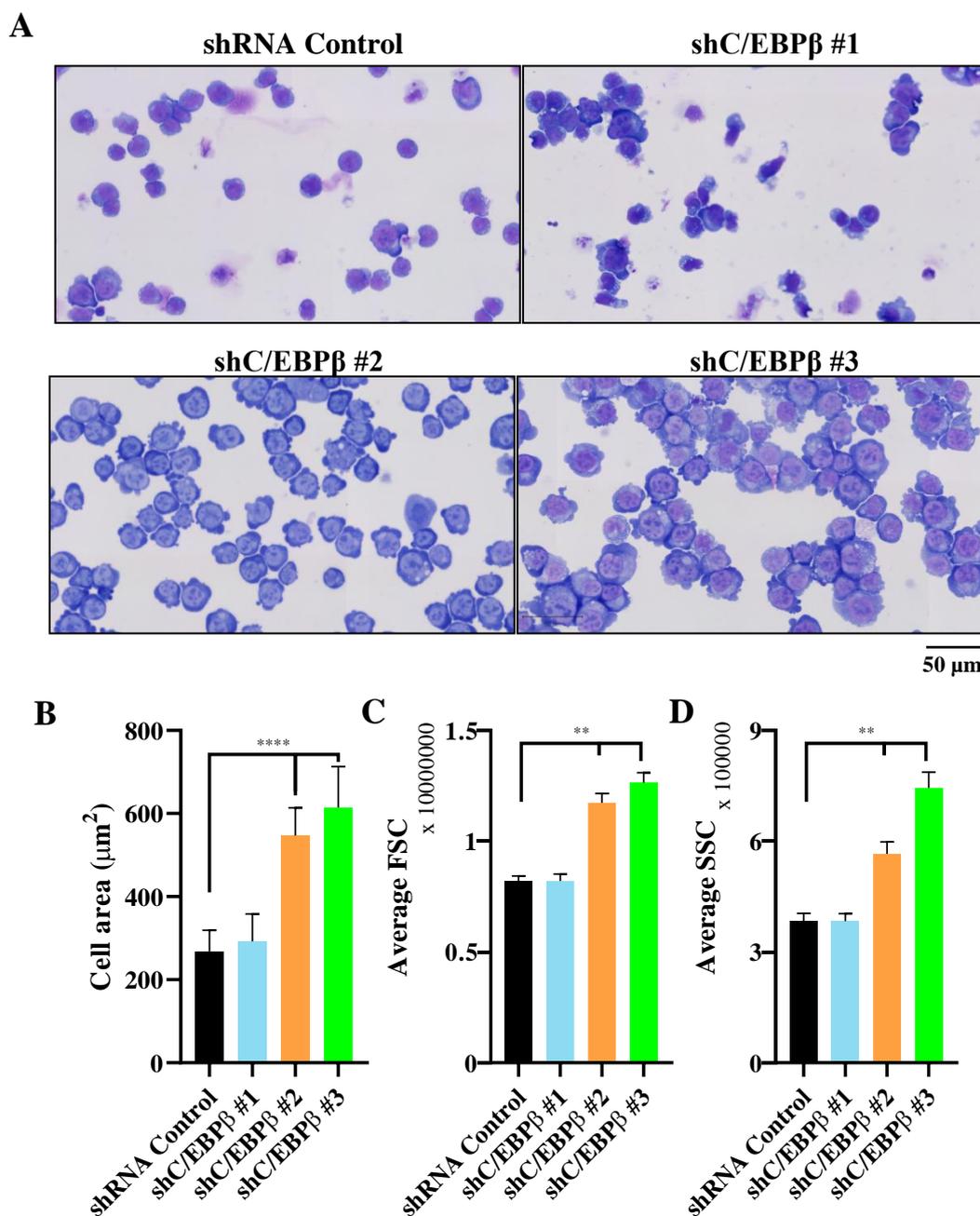
#### *5.3.5.4 The impact of C/EBP $\beta$ knockdown on AML non-t(8;21) cell lines.*

Previous results have suggested that low levels of C/EBP $\beta$  are associated with the development of t(8;21), and that this protein might interact with the RUNX1-ETO in the context of this disease (**5.3.2**). Having determined that, in the RUNX1-ETO-expressing cell line SKNO-1, KD of C/EBP $\beta$  promoted cell proliferation and was associated with an immature phenotype, this study additionally analysed the consequences of knocking-down C/EBP $\beta$  in the context on other AML subtypes, by applying the same KD vectors to HEL and U937 cells. In contrast to the effect on SKNO-1 cells, KD of C/EBP $\beta$  significantly reduced cell growth in both cell lines (**Figure 5.31A**; **Figure 5.34A**).



**Figure 5.29 – The effect of C/EBPβ KD on cell surface marker expression in SKNO-1 cells**

Expression of cell surface markers in SKNO-1 cells transduced with a control or shC/EBPβ constructs. An isotype control (IgG) was used to define negative marker expression ( $n \geq 3$ ). Significant difference between shC/EBPβ cultures and control were analysed using one-way ANOVA with Bonferroni's multiple comparisons correction; \* denotes  $p < 0.05$ ; \*\*\* denotes  $p < 0.001$ ; \*\*\*\* denotes  $p < 0.0001$ .



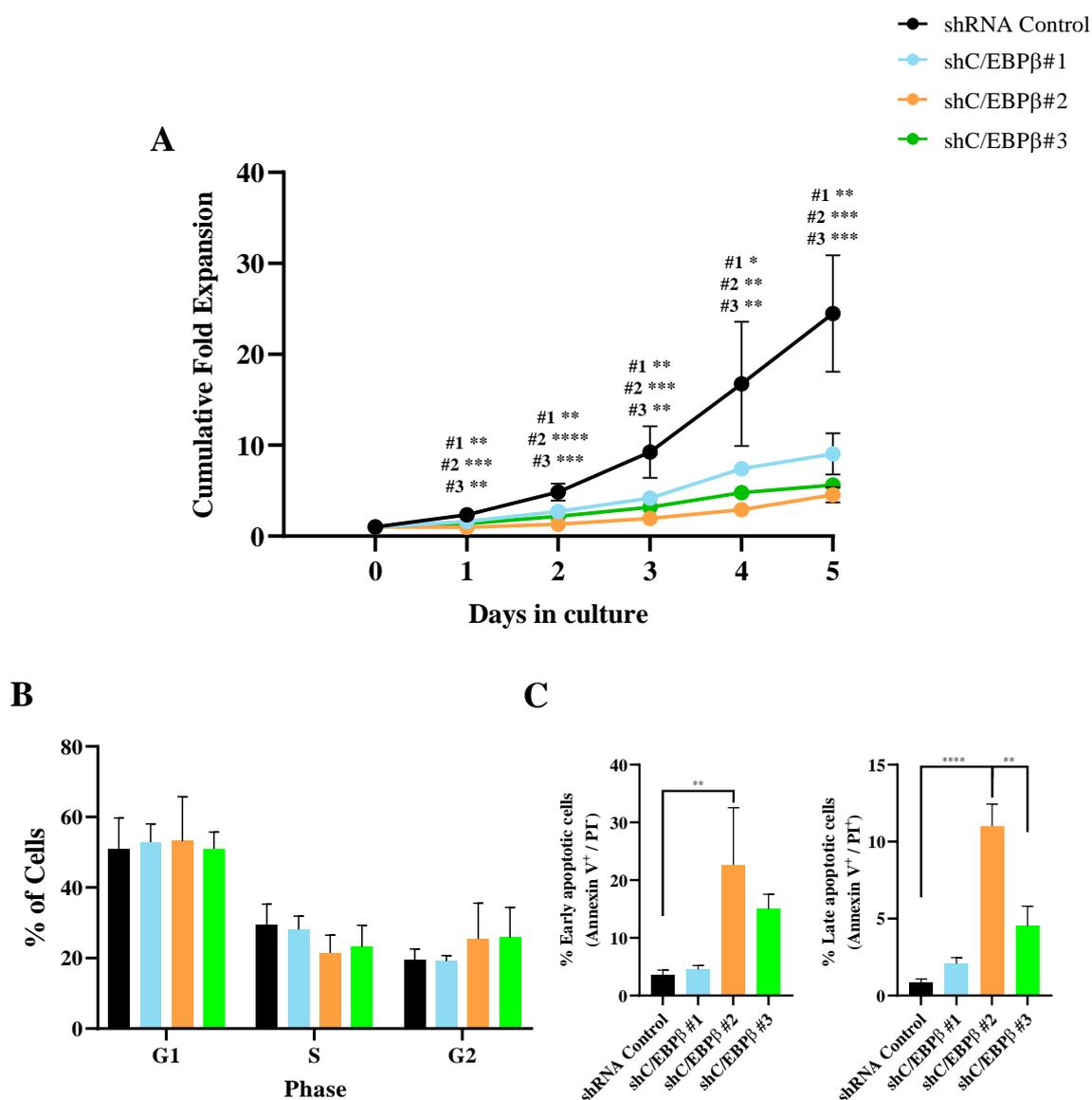
**Figure 5.30 – Morphological effect of C/EBPβ KD in SKNO-1 cells**

(A) Images showing the morphology of SKNO-1 cells transduced with a control or shC/EBPβ constructs (Table 2.1). Cells were stained with a combination of May-Grunwald and Gisma differential stains and scanned using a Zeiss Axioscan Z1 slide scanner, at 20x magnification (2.3.5). Scale bar indicates 50 μm. Bar charts showing (B) the area of SKNO-1 cells, as calculated using ImageJ; (C) FSC and (D) SSC of SKNO-1 cells transduced with a control or shC/EBPβ constructs. Data indicates mean ± 1SD (n=3). Significant difference between shC/EBPβ cultures and control were analysed using one-way ANOVA with Bonferroni's multiple comparisons correction; \*\* denotes p<0.01; \*\*\*\* denotes p<0.0001.

These changes were further associated with an increase in the proportion of cells in the G<sub>1</sub> phase of the cell cycle, although these changes were deemed not significant (**Figure 5.31B**; **Figure 5.34B**). Moreover, both HEL and U937 cells presented a significant increase in the percentage of cells in early and late apoptotic stages, specifically for cultures transduced with shC/EBP $\beta$  #2 and #3 (**Figure 5.31C**; **Figure 5.34C**).

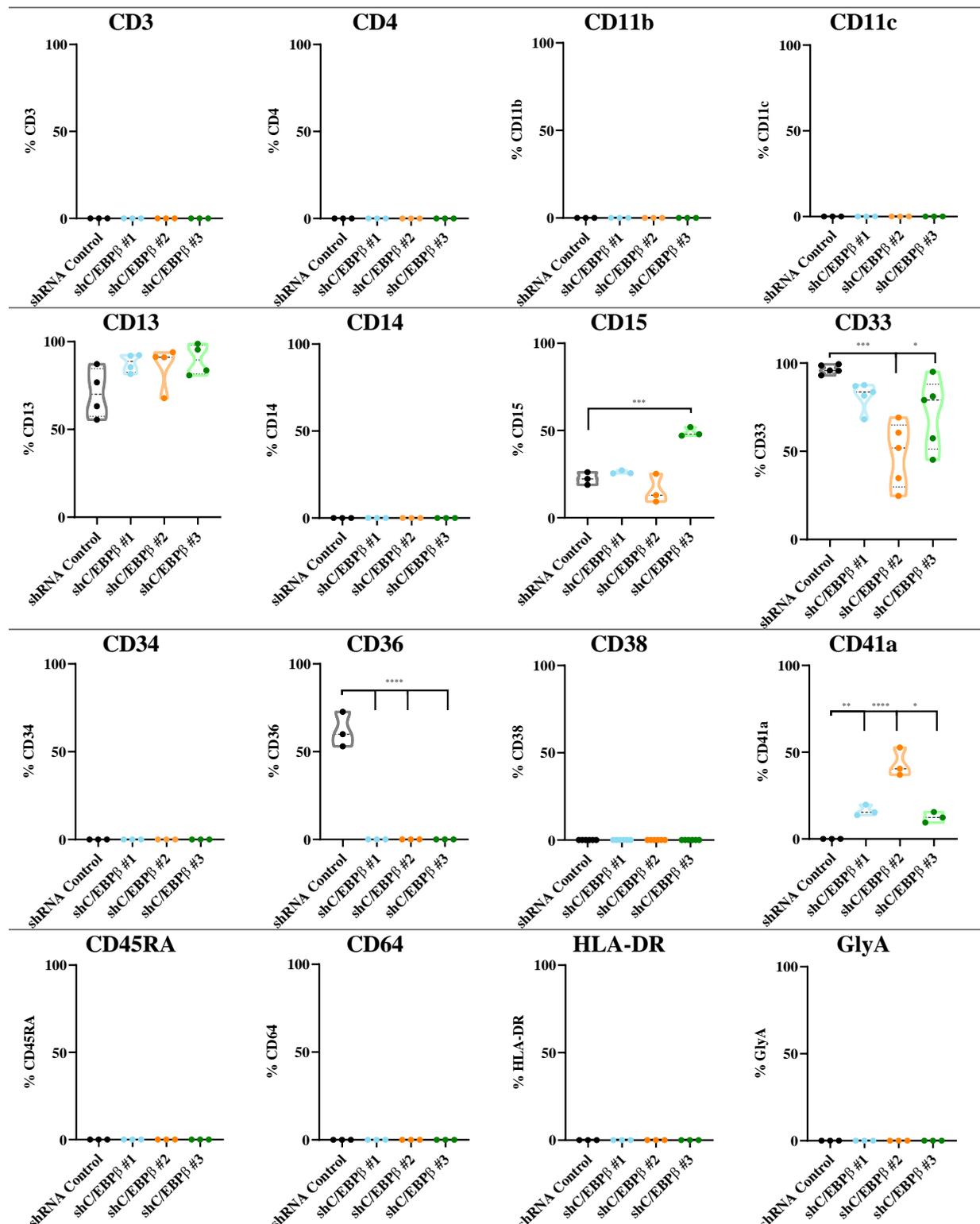
Subsequently, immunophenotypic analysis was performed to determine the effect of C/EBP $\beta$  KD on cell surface marker expression. The HEL cell line, derived from an erythroleukaemia patient, was found to be characterised as CD13<sup>+</sup>CD33<sup>+</sup>CD36<sup>+</sup> (**Figure 5.32**). KD populations were found to equally express the CD13 marker, however, shC/EBP $\beta$  #2 and #3 cultures showed a significant decreases in the expression of CD33 and CD36 and increased expression of CD41a. In the histiocytic lymphoma cell line U937, this study found that control cells were CD13<sup>+</sup> CD45RA<sup>+</sup>, with a heterogeneous expression of CD14, CD15, CD33, CD38 and CD64 (**Figure 5.35**). Upon C/EBP $\beta$  KD, these cells showed a significant decrease in the expression of CD14, CD45RA and CD38, with the simultaneous increase of CD15, CD36, CD64 and HLA-DR.

Subsequently, these cells were subjected to morphological analysis. An increase in cell size and complexity is usually observed as the differentiation process occurs. Both HEL and U937 cells showed an increase in cell size, further associated with a significant increase in the forward scatter) and side scatter properties of these cells (**Figure 5.33**; **Figure 5.36**). Additionally, KD of C/EBP $\beta$  in HEL cells, associated with shC/EBP $\beta$  #2 and #3, resulted in the development of multinucleated cells, suggestive of megakaryocytic differentiation, supported by increased expression of the CD41 marker. U937 cells, on the other hand, particularly cells transduced with shC/EBP $\beta$  #3, showed features associated with macrophage differentiation, but this was not supported by the immunophenotype of these cells.



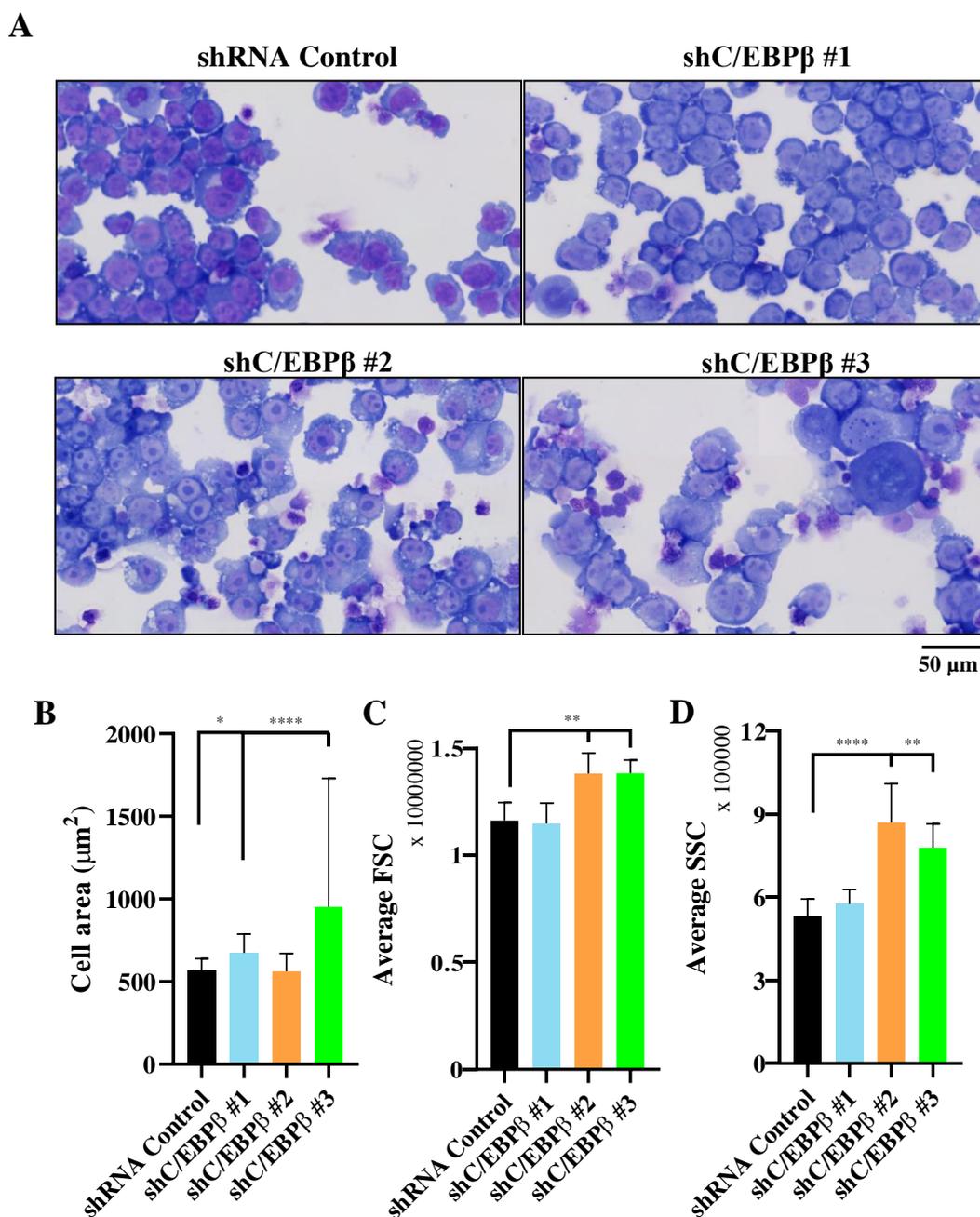
**Figure 5.31 – C/EBPβ KD significantly reduces cell growth in HEL cells, associated with changes in apoptotic frequency**

Summary data showing (A) the growth of HEL cells with three shC/EBPβ constructs compared to control, over 5 days of culture. Data indicates mean  $\pm$  1SD (n=4); (B) the effect of C/EBPβ KD in normal cell cycle. Data indicates Mean  $\pm$  1SD (n=3); (C) the effect of C/EBPβ KD on apoptosis using Annexin V staining. Bars indicate percentage of pre-apoptotic cells (Annexin V<sup>+</sup> / 7-AAD<sup>-</sup>) and late apoptotic cells (Annexin V<sup>+</sup> / 7-AAD<sup>+</sup>). Data indicates Mean  $\pm$  1SD (n=3). Significant differences between shC/EBPβ cultures and control were analysed using one-way ANOVA with Bonferroni's multiple comparisons correction; \* denotes p<0.05; \*\* denotes p<0.01; \*\*\* denotes p<0.001; \*\*\*\* denotes p<0.0001.



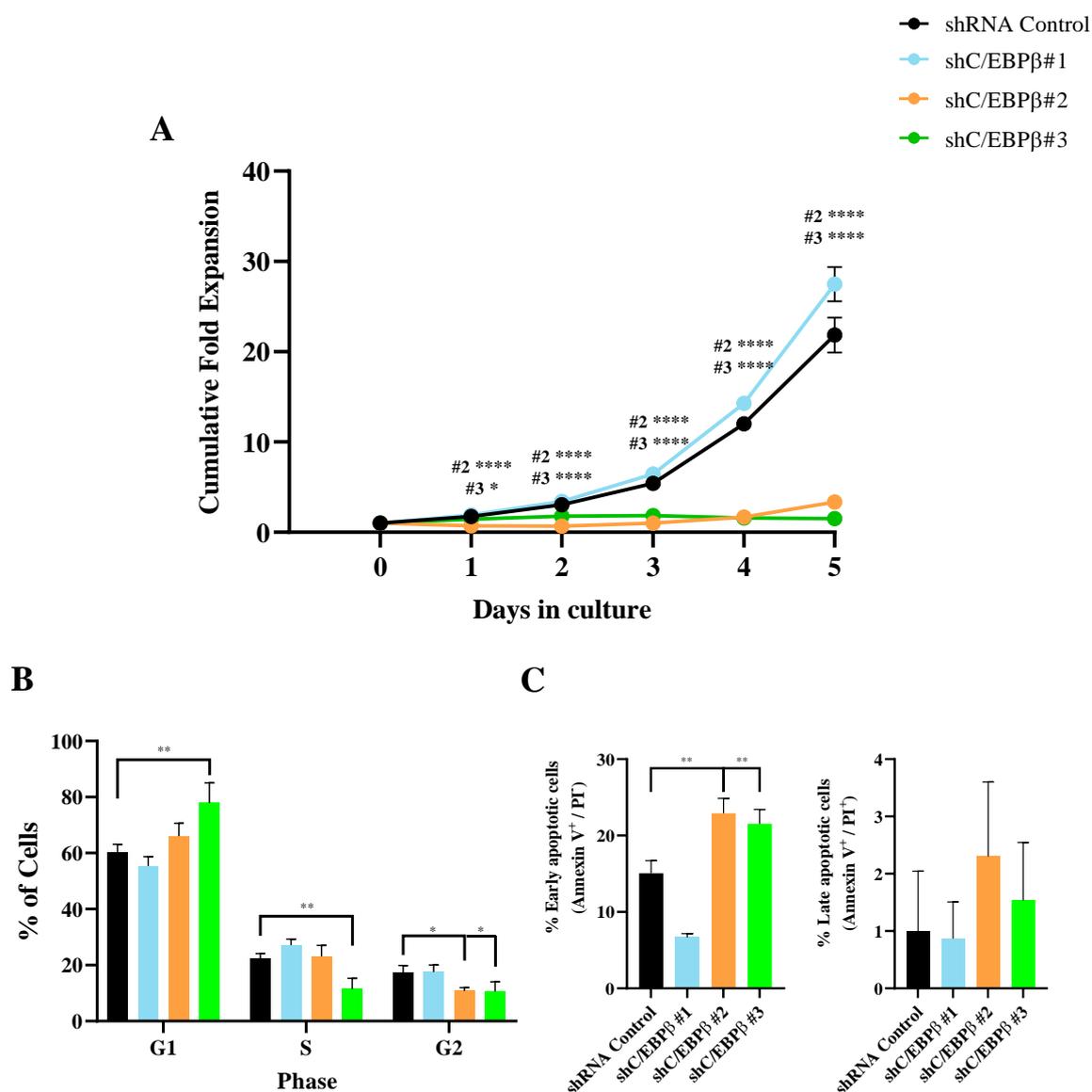
**Figure 5.32 – The effect of C/EBPβ KD on cell surface marker expression in HEL cells**

Expression of cell surface markers in HEL cells transduced with a control or C/EBPβ-KD constructs. An isotype control (IgG) was used to define negative marker expression ( $n \geq 3$ ). Significant difference between shC/EBPβ cultures and control were analysed using one-way ANOVA with Bonferroni's multiple comparisons correction; \* denotes  $p < 0.05$ ; \*\*\* denotes  $p < 0.001$ ; \*\*\*\* denotes  $p < 0.0001$ .



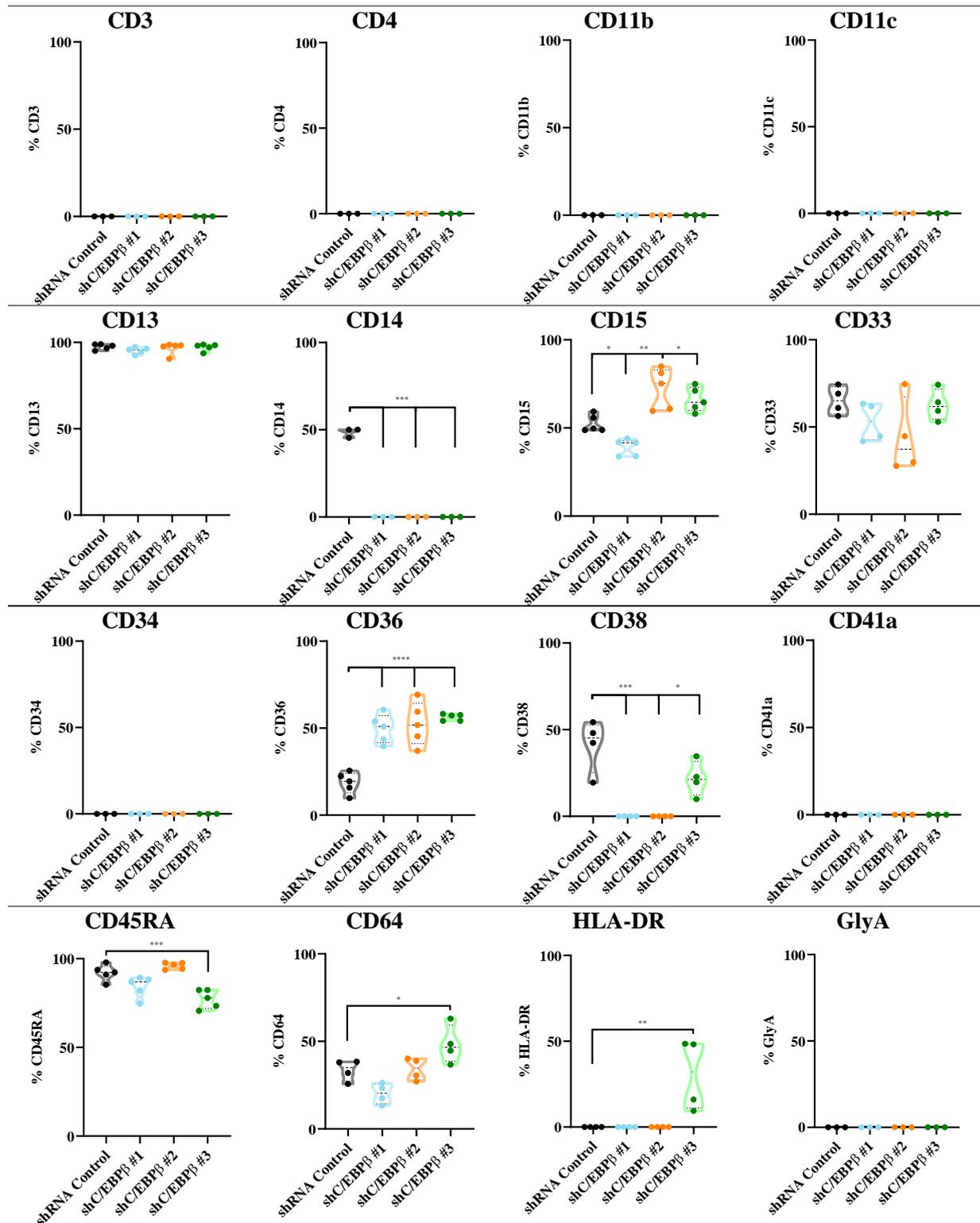
**Figure 5.33 – Morphological effect of C/EBPβ KD in HEL cells**

(A) Images showing the morphology of HEL cells transduced with a control or shC/EBPβ constructs (Table 2.1). Cells were stained with a combination of May-Grunwald-Gisma differential stains and scanned using a Zeiss Axioscan Z1 slide scanner, at 20x magnification (2.3.5). Scale bar indicates 50 μm. Bar charts showing (B) the area of HEL cells, as calculated using ImageJ; (C) FSC and (D) SSC of HEL cells transduced with a control or shC/EBPβ constructs. Data indicates mean ± 1SD (n=3). Significant difference between shC/EBPβ cultures and control were analysed using one-way ANOVA with Bonferroni's multiple comparisons correction; \*\* denotes p<0.01; \*\*\*\* denotes p<0.0001.



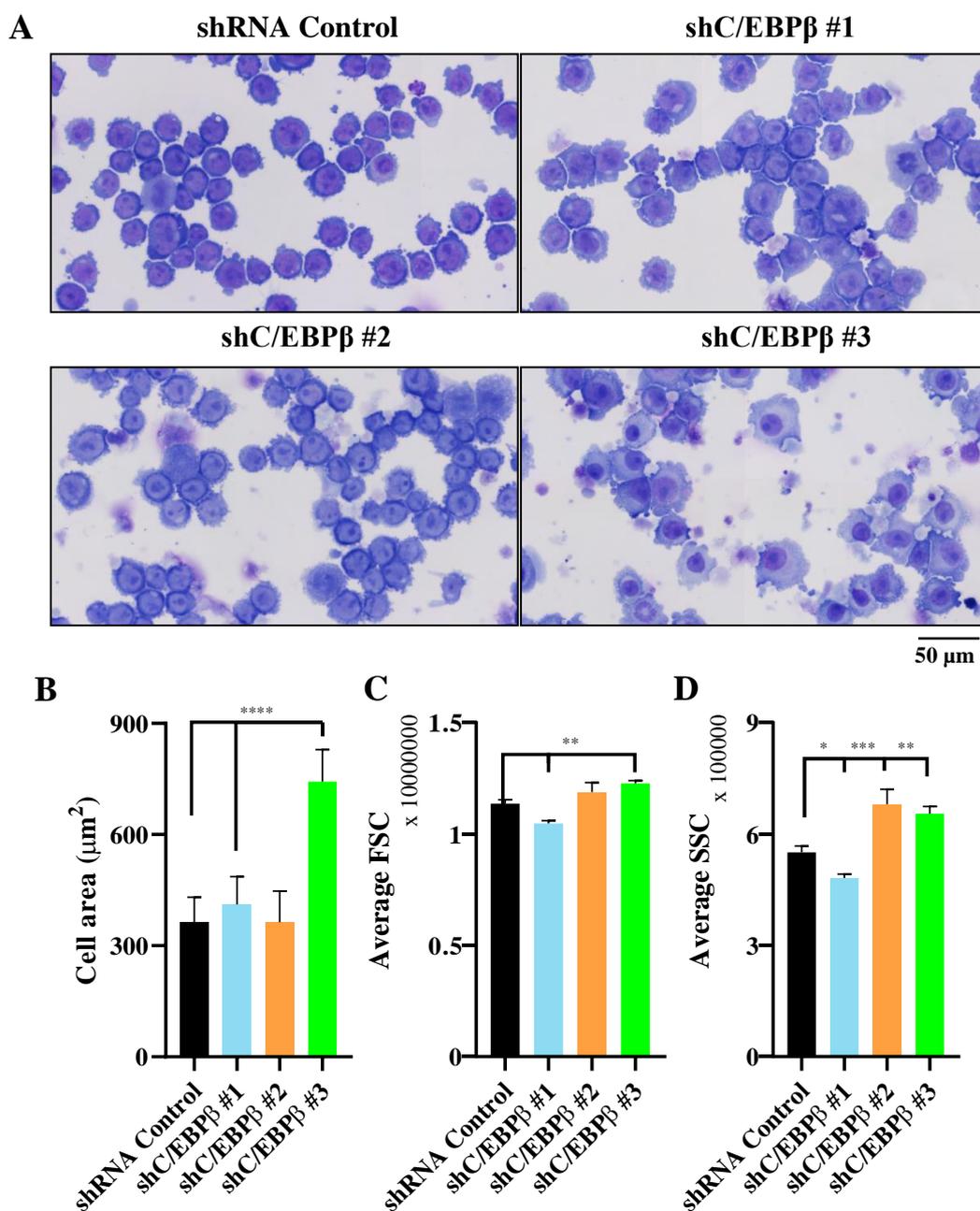
**Figure 5.34 – C/EBPβ KD significantly reduces cell growth in U937 cells, associated with changes in normal cell cycle and apoptotic rate**

Summary data showing (A) the growth of U937 cells with three shC/EBPβ constructs compared to control, over 5 days of culture. Data indicates mean  $\pm$  1SD (n=4); (B) the effect of C/EBPβ KD in normal cell cycle. Data indicates Mean  $\pm$  1SD (n=3); (C) the effect of C/EBPβ KD on normal apoptotic rate using Annexin V staining. Bars indicate percentage of pre-apoptotic cells (Annexin V<sup>+</sup> / 7-AAD<sup>-</sup>) and late apoptotic cells (Annexin V<sup>+</sup> / 7-AAD<sup>+</sup>). Data indicates Mean  $\pm$  1SD (n=3). Significant differences between shC/EBPβ cultures and control were analysed using one-way ANOVA with Bonferroni's multiple comparisons correction; \* denotes  $p < 0.05$ ; \*\* denotes  $p < 0.01$ ; \*\*\* denotes  $p < 0.001$ ; \*\*\*\* denotes  $p < 0.0001$ .



**Figure 5.35 – The effect of C/EBPβ KD on cell surface marker expression in U937 cells**

Expression of cell surface markers in U937 cells transduced with a control or C/EBPβ-KD constructs. An isotype control (IgG) was used to define negative marker expression ( $n \geq 3$ ). Significant difference between shC/EBPβ cultures and control were analysed using one-way ANOVA with Bonferroni's multiple comparisons correction; \* denotes  $p < 0.05$ ; \*\*\* denotes  $p < 0.001$ ; \*\*\*\* denotes  $p < 0.0001$ .



**Figure 5.36 – Morphological effect of C/EBPβ KD in U937 cells**

(A) Images showing the morphology of U937 cells transduced with a control or shC/EBPβ constructs (Table 2.1). Cells were stained with a combination of May-Grunwald and Gisma differential stains and scanned using a Zeiss Axioscan Z1 slide scanner, at 20x magnification (2.3.5). Scale bar indicates 50 μm. Bar charts showing (B) the area of U937 cells, as calculated using ImageJ; (C) FSC and (D) SSC of U937 cells transduced with a control or shC/EBPβ constructs. Data indicates mean ± 1SD (n=3). Significant difference between shC/EBPβ cultures and control were analysed using one-way ANOVA with Bonferroni's multiple comparisons correction; \*\* denotes p<0.01; \*\*\*\* denotes p<0.0001.

## 5.4 Discussion

C/EBP $\beta$  is a member of the C/EBP family of TF, and is involved in regulating several processes, such as cell growth, differentiation, inflammation, metabolism, survival and tumorigenesis (Sebastian and Johnson, 2006; Cao *et al.*, 1991; Poli, 1998; Zahnow, 2002; Diehl, 1998; Ramji and Foka, 2002). Previous studies have shown that, depending on cell context, the regulatory function of C/EBP $\beta$  is variable (Sebastian and Johnson, 2006). For instance, KO of C/EBP $\beta$  in mice resulted in a decrease in the number of B lymphocytes with reduced cell expansion properties (Regalo *et al.*, 2016; Chen *et al.*, 1997). Furthermore, C/EBP $\beta$  has been found to be essential for liver regeneration, as C/EBP $\beta$  KO mice displayed reduced hepatocytic proliferation, indicating that C/EBP $\beta$  is necessary for liver regeneration (Greenbaum *et al.*, 1998). The same trend has been observed in several cancer models. In the BCR-ABL expressing cell line 32D, forced expression of C/EBP $\beta$  led to a decrease in cell proliferation and promoted differentiation (Guerzoni *et al.*, 2006). Similarly, expression of C/EBP $\beta$  in the APL cell line NB4, upon ATRA treatment, was shown to result in decreased proliferative ability, associated with cell differentiation (Duprez *et al.*, 2003). In glioblastoma and gastric cancer cells, on the other hand, decreased expression of C/EBP $\beta$  led to the inhibition of cell growth (Regalo *et al.*, 2016; Aguilar-Morante *et al.*, 2011).

In a haematopoietic setting, C/EBP $\beta$  is highly expressed in cells committed to the myelomonocytic lineage, including monocytes and macrophages (Williams *et al.*, 1991; Katz *et al.*, 1993; Haas *et al.*, 2010; Gutsch *et al.*, 2011). Several roles have been attributed to C/EBP $\beta$  in haematopoietic development. Depending on the system used, studies have shown C/EBP $\beta$  is involved in monocytic differentiation (Friedman, 2007), eosinophil maturation (Nerlov *et al.*, 1998) or granulocytic differentiation (Popernack *et al.*, 2001; Iwama *et al.*, 2002; Yusenko *et al.*, 2021). In this study, RUNX1-ETO expression in CD34<sup>+</sup> HSPC resulted in the significant downregulation of C/EBP $\beta$  mRNA and protein, as compared to control cells. The exact role of C/EBP $\beta$  in the development of t(8;21), however, has not been previously described. This chapter's main aim was to determine the role of C/EBP $\beta$  in normal haematopoietic development, using a cord blood-derived model, and in leukaemogenic development, using AML cell lines. In summary, C/EBP $\beta$  overexpression promoted monocytic and granulocytic differentiation in HSPC. However, no significant impact of C/EBP $\beta$  KD was observed. This latter observation could be explained by the lack of complete KD/KO of C/EBP $\beta$ , or due to redundancy in the context of other members of the C/EBP family, further

discussed in 5.4.3. In AML cell lines, the effect of C/EBP $\beta$  KD was found to be highly context-dependent: in the t(8;21) line, SKNO-1, decreased C/EBP $\beta$  levels resulted in an increased proliferative ability and reduced myeloid marker expression. In non-t(8;21) cells, however, KD of C/EBP $\beta$  reduced cell growth and induced more variable effects on the expression of markers of differentiation.

#### 5.4.1 *CEBPB* expression increases during haematopoietic development

To further enlighten the role of C/EBP $\beta$  on myeloid development, specifically in the monocytic and granulocytic maturation process, *CEBPB* expression was determined using publicly available transcriptomic datasets in human haematopoietic progenitor cells. This analysis showed that *CEBPB* expression increased specifically in fully mature granulocytes and monocytes, as compared to undifferentiated HSCs, suggesting that *CEBPB* is involved in the development of these lineages. These observations were further corroborated with the analysis of the *CEBPB* expression in the mouse and agree with Tamura *et al.*, who showed that *CEBPB* was upregulated in mouse mature monocytes (Tamura *et al.*, 2017). Additionally, the authors further demonstrated that C/EBP $\beta$  is involved in the survival of non-classic (Ly6C<sup>-</sup>) monocytes as well as being essential for monocytic development (Tamura *et al.*, 2017). However, since a single mRNA transcript generates three C/EBP $\beta$  protein isoforms, LAP\*, LAP and LIP (1.5.1) transcriptional analysis cannot discriminate between the roles of these isoforms, whose activity may also be influenced by post-translational modifications, including phosphorylation, acetylation, methylation and sumoylation (Ramji and Foka, 2002; Zahnow, 2009; Tsukada *et al.*, 2011; Hattori *et al.*, 2003). Further, Pham *et al.* reported that whilst *CEBPB* mRNA was unchanged during differentiation process, nuclear levels of C/EBP $\beta$  increased during the maturation process of monocytes into macrophages (Pham *et al.*, 2007).

Since transcriptomic analysis alone may not be an accurate guide to C/EBP $\beta$  protein expression, this study determined the expression of C/EBP $\beta$  protein in HSPC by western blotting analysis of cells cultured in conditions that favoured monopoiesis and granulopoiesis. In this context, C/EBP $\beta$  LAP\* was shown to decrease over time. Expression of LAP was only detectable on day 3, whilst LIP was undetected throughout. These observations directly contrast with the mRNA analysis described above. Analysis of purified lineage subsets showed that C/EBP $\beta$  LAP\* was highest in monocytic-progenitor cells, followed by erythroid-committed cells and granulocytic cells; however, these observations are made early in the myeloid development and later time-points would be necessary to determine the pattern

C/EBP $\beta$  protein expression throughout the development of these lineages. Unfortunately, there was no time to perform these experiments.

C/EBP $\beta$  LAP\* and LAP have previously been shown to be involved in the differentiation process of monocytes (Friedman, 2007), and upregulation of these two C/EBP $\beta$  isoforms is necessary for this process to occur (Katz *et al.*, 1993; Gutsch *et al.*, 2011; Pham *et al.*, 2007; Natsuka *et al.*, 1992; Pan *et al.*, 1999; Studzinski *et al.*, 2005; Zhang *et al.*, 2011a). C/EBP $\beta$  expression has also been shown to be essential for granulopoiesis under stress conditions (Hirai *et al.*, 2006; Popernack *et al.*, 2001) with steady-state granulocytic development being regulated by another member of the C/EBP family, C/EBP $\alpha$  (Suh *et al.*, 2006; Radomska *et al.*, 1998).

#### **5.4.2 Increased *CEBPB* expression is associated with poor prognosis in AML**

Expression of *CEBPB* was found to be elevated across AML subtypes, but lower in t(8;21). Furthermore, increased *CEBPB* expression was associated with poor prognosis; however, since multivariate analysis was not conducted, its status as a prognostic factor was not established in this study. Similar observations had been previously made in solid tumours. High *CEBPB* expression in breast cancer was associated with the most aggressive tumour types, including metastatic breast cancer (van de Vijver *et al.*, 2002), tumours associated with a high tumour grade (van 't Veer *et al.*, 2002; Ma *et al.*, 2004; Finak *et al.*, 2008) and an overall poorer prognosis (van de Vijver *et al.*, 2002). Similarly, in human gastric carcinoma, C/EBP $\beta$  protein expression was found to correlate with poorer prognosis, associated with lower median survival, and the development of metastasis (Du *et al.*, 2013). However, these observations need to be considered carefully, as C/EBP $\beta$  expression is also regulated through post-translational mechanisms (5.4.1). In fact, Kurzejamska *et al.* reported that loss of C/EBP $\beta$  protein promoted metastatic development of mouse mammary tumour cells (Kurzejamska *et al.*, 2014) and in a cohort of 137 breast cancer patients decreased C/EBP $\beta$  expression correlated with poorer OS and development of metastasis within the lymph nodes (Kurzejamska *et al.*, 2014).

#### **5.4.3 Knockdown of C/EBP $\beta$ in HSPC does not influence normal myeloid development**

To establish whether the reduced C/EBP $\beta$  expression seen in t(8;21) patients could contribute to the pathogenesis of AML, functional studies were performed using KD of

C/EBP $\beta$  in human cord blood-derived HSPC. Under clonal conditions, KD of C/EBP $\beta$  resulted in a significant decrease in myeloid colony forming efficiency and self-renewal potential. In bulk liquid culture, KD of C/EBP $\beta$  repressed HSPC monocytic and granulocytic proliferation. In BM-derived progenitor cells from mice, absence of C/EBP $\beta$  also resulted in a significant reduction in the percentage of colonies formed, as well as in the number of generated cells, in accordance with this study's results. Interestingly, however, KO of C/EBP $\beta$  in more differentiated cells within the same model led to an increase in cell proliferation (Gutsch *et al.*, 2011; Screpanti *et al.*, 1995). It is plausible that the effects observed would have been different if C/EBP $\beta$  had been knocked-down in more differentiated cells, as opposed to HSCs.

Developmentally, KD of C/EBP $\beta$  in HSPC did not significantly alter the monocytic differentiation markers CD11b and CD14. Moreover, this study found that C/EBP $\beta$  was not required for the survival of mature monocytes, as suggested by Tamura *et al.* (Tamura *et al.*, 2017). However, Tamura *et al.* applied a mouse C/EBP $\beta$ <sup>-/-</sup> model of peripheral blood monocyte survival, in contrast with the *in vitro* HSPC model used in this study. In granulocytic development, KD of C/EBP $\beta$  promoted CD15 expression, but did not affect the expression of other developmental markers in the panel. The study of additional granulocytic markers, such as CD16 or CD32 (Attar, 2014), would be necessary to unequivocally determine the exact role of C/EBP $\beta$  in granulocytic development. Together, this data suggests that C/EBP $\beta$  is mainly involved in monocytic development; however, it is not essential for this process to occur.

Reduced growth can be a result of perturbed cell cycle progression or increased apoptotic events. KD of C/EBP $\beta$  in HSPC was found to result in the accumulation of cells within the G<sub>1</sub> phase of the cell cycle. Opposingly to this study, Gutsch *et al.* showed that, in macrophage-like cells, KO of C/EBP $\beta$  resulted in the increase of cells displaying S and G<sub>2</sub>/M phase markers (Gutsch *et al.*, 2011), associated with an increase in proliferation. In human non-small cell lung cancer tissues, KD of C/EBP $\beta$  was associated with an increase in the proportion of cells arrested at the G<sub>2</sub>/M phase of the cell cycle (Lee *et al.*, 2019); however, in this study, the delay in cell cycle progression was further associated with the inhibition of cell proliferation.

Overall, KD of C/EBP $\beta$  had a modest impact in the HSCP model used in this study. Several factors can explain these results. Firstly, levels of C/EBP $\beta$  were knocked-down by c50%, and it is possible that these levels are enough to support haematopoietic development. Moreover, other C/EBP isoforms might be acting compensatively in the events of C/EBP $\beta$  KD. In the haematopoietic system, C/EBP $\alpha$ , C/EBP $\beta$ , C/EBP $\delta$ , and C/EBP $\epsilon$  are all expressed within the

myeloid lineage (Tolomeo and Grimaudo, 2020). Moreover, the overlapping expression of the C/EBP family members suggests that myeloid development is regulated by different combinations of C/EBP homo- and heterodimers (Tolomeo and Grimaudo, 2020).

Several studies have demonstrated that C/EBP $\beta$  is only necessary for granulocytic development in stress-derived conditions (Hirai *et al.*, 2006; Hayashi *et al.*, 2013), but it is necessary for the normal function and differentiation of myelomonocytic cells and macrophages (Pham *et al.*, 2007; Cain *et al.*, 2013). KD studies, however, did not demonstrate this, as the differentiation of both granulocytes and monocytes was not impaired upon KD of this TF. These observations have been summarised in **Table 5.1**.

#### **5.4.4 C/EBP $\beta$ overexpression in HSPC promotes myeloid cell growth and development**

Next, this study performed functional studies by overexpressing the longer C/EBP $\beta$  isoform LAP\*, within the same HSPC model. Under clonal conditions, C/EBP $\beta$  significantly promoted myeloid colony formation. A single round of colony replating suggested that this may have, at least in part, be driven by promotion of self-renewal. Overexpression of C/EBP $\beta$  also resulted in an increase in both monocytic and granulocytic growth in bulk liquid culture with an increase in the proportion of granulocytic- cells predominating over that of monocytes.

Phenotypically, overexpression of C/EBP $\beta$  upregulated both the CD11b and CD14 markers on monocytic cells. This agrees with previous studies, which showed that C/EBP $\beta$  controls CD14 expression (Pan *et al.*, 1999; Ji and Studzinski, 2004; Zhang *et al.*, 2011a; Xu *et al.*, 2008) and that C/EBP $\beta$  promotes monocytic development (Pham *et al.*, 2007; Zhang *et al.*, 2011b; Tamura *et al.*, 2017).

Analysis of the granulocytic population showed that overexpression of C/EBP $\beta$  resulted in the upregulation of the cell surface marker CD15 during the differentiation process, suggesting promotion of granulocytic development. Assessment of cell morphology also revealed that overexpression of C/EBP $\beta$  promotes granulocytic differentiation, with an increase in the development of band cells (mature granulocytes). These observations have been summarised in **Table 5.1**.

**Table 5.1 – Summary of the main findings regarding the role of C/EBP $\beta$  in normal human haematopoiesis**

Table summary detailing the consequences of C/EBP $\beta$  modulation *vs* control.

\* denotes significant differences in C/EBP $\beta$  overexpression / knockdown, as compared to normal control cells.

	<b>C/EBP<math>\beta</math> Overexpression</b>	<b>C/EBP<math>\beta</math> Knockdown</b>
<b>Monocytes</b>		
<b>Growth</b>	<b>Decrease</b>	No change
<b>Markers</b>	<b>Increase in CB11b and CD14 *</b>	No change
<b>Granulocytes</b>		
<b>Growth</b>	<b>Decrease *</b>	<b>Decrease *</b>
<b>Markers</b>	<b>Increase in CD15 *</b>	<b>Increase in CD15 *</b>
<b>Apoptosis</b>	No change	No change
<b>Cell Cycle</b>	<b>Decrease in G1 *</b>	No change
<b>CFU</b>	<b>Increase *</b>	<b>Decrease *</b>

### 5.4.5 The consequences of C/EBP $\beta$ KD in AML cell lines is context-dependant

The final part of this chapter consisted of analysing the consequences of C/EBP $\beta$  expression in AML by modulating its expression in a cohort of AML cell lines, to determine whether it influenced leukaemogenic development. KD of C/EBP $\beta$  was found to significantly promote the fold expansion on the t(8;21) cell line SKNO-1 accompanied by a decrease in the proportion of cells in the G<sub>1</sub> phase of the cell cycle and decrease in normal apoptotic frequency, all consistent with a pro-proliferative AML phenotype. Analysis of cell surface markers showed ablation of myeloid marker expression upon KD of C/EBP $\beta$  indicating that the developmental status of these cells was also affected. Interestingly, contrary results were observed in cell lines that do not express the fusion protein RUNX1-ETO. In these cells (HEL and U937), KD of C/EBP $\beta$  was found to suppress cell growth, associated with an increase in the proportion of cells in the G<sub>1</sub> phase of the cycle and increased apoptotic frequency. Immunophenotypic marker expression was again perturbed in these lines but without any clear pattern emerging.

Previous studies have showed that forced expression of C/EBP $\beta$  promotes monocytic development of AML cell lines. For instance, addition of 1,25D to HL-60 cells resulted in forced monocytic differentiation, associated with a significant increase in the expression of C/EBP $\beta$  LAP and LIP isoforms (Marcinkowska *et al.*, 2006). The fact that the dominant-negative LIP isoform was found to be increased in parallel with LAP makes the authors suggest that this isoform is responsible for repressing genes unnecessary for the differentiation process, including those involved in sustained proliferation. In K562 cells, a CML cell line, expression of C/EBP $\beta$  was shown to induce granulocytic development, with changes to cell morphology and promotion of gene expression associated with this cell type (Duprez *et al.*, 2003). In these cells, the process was initiated by C/EBP $\beta$ , which led to the expression of another member of the C/EBP family, C/EBP $\epsilon$ , known to play an important role in granulocytic differentiation (Park *et al.*, 1999). Additionally, in APL cells, increased expression of C/EBP $\beta$  was shown to be necessary in ATRA-induced cell differentiation, and its inhibition resulted in a decreased response to ATRA (Duprez *et al.*, 2003). Similar observations were made by Yusenko *et al.*, who reported that ectopic expression of C/EBP $\beta$  in the APL cell line NB4 led to a decrease in cell proliferation (Yusenko *et al.*, 2021). Therefore, this study hypothesised that KD of C/EBP $\beta$  in AML cell lines would promote cell proliferation associated with the expression of cell surface markers associated with immature cells. Ours results, however, conflict with the above-described studies, where overexpression, opposed to KD, of C/EBP $\beta$  resulted in cell

differentiation and decreased proliferative ability. However, a few studies have been published that describe the downregulation of C/EBP $\beta$  as a potential therapeutic target. With this in line, it has been shown that C/EBP $\beta$  cooperates with MYB and the co-activator p300 to promote the expression of myeloid associated genes (Burk *et al.*, 1993; Ness *et al.*, 1993; Mink *et al.*, 1996; Yusenko *et al.*, 2021), and that MYB is highly expressed in AML cells, making it an interesting therapeutic target (Hess *et al.*, 2006; Somervaille *et al.*, 2009; Jin *et al.*, 2010; Zuber *et al.*, 2011). In a recent study, Yusenko *et al.* showed that inhibiting the binding of MYB to C/EBP $\beta$  resulted in the upregulation of differentiation associated genes and increased apoptosis in AML cell lines (Yusenko *et al.*, 2021), similarly to this study's results. Equally, expression of C/EBP $\beta$  in multiple myeloma cell lines and patient samples was shown to be increased as compared to normal B cells, and silencing it resulted in the downregulation of several TF, including those involved in the proliferative process, such as IRF4, XBP1 and BLIMP1, and apoptotic signalling, including BCL2 (Pal *et al.*, 2009).

Additionally, this study attempted to generate stable C/EBP $\beta$  LAP\* overexpression cell lines. However, this was not successful due to rapid loss of transduced cells, as determined by GFP positivity; this suggests that overexpression of C/EBP $\beta$  LAP\* is incompatible with AML cell lines.

In conclusion, the data presented in this chapter shows that KD of C/EBP $\beta$  in HSPC did not result in an arrest in haematopoietic development. However, a limitation of the application of shRNA is the inability to completely remove C/EBP $\beta$  expression; it is possible that, despite the reduced levels, a far greater level of KD would be necessary to demonstrate any dependency. Additionally, compensation from alternate C/EBP $\beta$  family members could be acting as a mechanism of redundancy. Whilst the former could be further explored with CRISPR-Cas9, which would result in complete KO, the latter concern would require a complex study design to fully consider all the C/EBP family members. Furthermore, this study demonstrated that KD of C/EBP $\beta$  is context-dependant with regards to the differentiation status of the AML cell lines. These findings indicate that t(8;21) might be an important context, but attempts to reproduce this in another t(8;21) line failed, making it hard to draw a meaningful conclusion (**Supplementary Figure 46**). In contrast, this study provides some evidence that the overexpression seen in AML might have a role in developmental inhibition.

# **Chapter 6**

---

## **Conclusion and Future Directions**

## 6.1 Conclusions

AML is a disease characterised by the clonal expansion of primitive HSCs which are incapable of terminal differentiation (Döhner *et al.*, 2015). Furthermore, AML is a highly heterogenous disease, arising from the development of several molecular abnormalities, including mutations and chromosomal rearrangements. A common abnormality is the t(8;21), occurring in approximately 12% of all AML cases, leading to the expression of the fusion protein RUNX1-ETO. Despite the fact that t(8;21) AML is associated with a good prognosis, with the majority of patients achieving CR following standard chemotherapy treatment, a significant proportion of these are deemed not suitable to receive intensive chemotherapy, especially due to their age. Consequently, half of the patients will relapse within 2.5 years of achieving CR (Marcucci *et al.*, 2005; Hospital *et al.*, 2014). The development of new targeted therapies, such as ATRA/ATO applied in APL (reviewed in (Sanz *et al.*, 2019)), or Venetoclax, targeting BCL-2 (DiNardo *et al.*, 2019), has significantly improved CR rates and OS rates of AML patients. Traditional treatment regimens for AML t(8;21) patients, however, have remained unchanged for several years. Despite the fact that these patients would potentially be eligible for targeted therapies, where appropriate, there is a pressing need to develop new therapies and improve the remission rate of t(8;21) patients.

Previous studies have demonstrated that RUNX1-ETO is able to inhibit normal myeloid differentiation in mice and patient samples (Okuda *et al.*, 1998; Pabst *et al.*, 2001; Nafria *et al.*, 2020). Furthermore, ectopic expression of RUNX1-ETO in human HSPCs was shown to block myeloid development and promote their self-renewal, associated with changes in gene expression patterns (Mulloy *et al.*, 2002; Tonks *et al.*, 2007; Tonks *et al.*, 2003; Tonks *et al.*, 2004). Despite the progress in understanding the biological implications of t(8;21), the mechanism through which this fusion protein leads to the development of AML is not yet fully understood. Moreover, a better understanding of the mechanism behind this process would allow the identification of novel therapeutic targets for the treatment of this disease.

The basis for this study's original hypothesis relied on the identification of TF, both at the mRNA and at the protein level, that could be responsible for the developmental disruption observed in t(8;21) patients, upon the expression of RUNX1-ETO. The original goal to apply proteomic analysis techniques to investigate targets of interest was achieved, as it allowed the identification of several targets known to be involved in the development of AML, including

PU.1 and CBFb, with numerous other novel targets being equally detected. Analysis of transcriptomic data yielded similar results, with the identification of well-established TF, thus validating the approach and model used in this study. Ultimately, it was possible to identify two targets, *ZNF217* and C/EBP $\beta$ , derived from the transcriptomic and proteomic analysis, respectively, as targets of interest in the context of t(8;21).

### 6.1.1 RUNX1-ETO-induced transcriptomic and proteomic changes in HSPC

The initial aim of this study was to re-analyse a previously generated microarray dataset, following the ectopic expression of RUNX1-ETO in human HSPCs (Tonks *et al.*, 2007). The previous approach used an unsupervised analysis and identified changes in all classes for genes, including membrane bound proteins (CD200) and signalling molecules ( $\gamma$ -Catenin) (Tonks *et al.*, 2007). I performed an ‘Enrichment Analysis’ using Metacore™ and showed that RUNX1-ETO overexpression significantly impaired megakaryopoiesis and granulocytic development, consistent with previously reported studies (Niebuhr *et al.*, 2008; Tonks *et al.*, 2004). Since developmental alterations are primarily a result of changes in the transcriptional process, I focused on TF dysregulation, to identify regulators of transcription that had been altered upon the expression of RUNX1-ETO. This analysis identified not only TF known to be involved in the development of AML, such as *PU.1* (Nerlov and Graf, 1998; Koschmieder *et al.*, 2005) and *CEBPA* (Koschmieder *et al.*, 2005), but also TF not yet extensively described in either haematopoietic development, or AML. This included the TF *ZNF217*, overexpressed in RUNX1-ETO-expressing cells, and studied in **Chapter 4**.

Whilst most studies have focused on RUNX1-ETO-mediated transcriptional changes (Nafria *et al.*, 2020; Ptasinska *et al.*, 2019; Martinez-Soria *et al.*, 2018; Ptasinska *et al.*, 2012; Tonks *et al.*, 2007), there is a paucity regarding analysis of the cells’ proteomic profile. Additionally, whilst transcriptomic analysis can infer corresponding protein expression, there are several factors that influence its corresponding protein expression including translational or post-translational modifications, epigenetic regulators or protein half-life. For this reason, analysis of the cells proteomic profile represents a more direct insight into the changes cells undergo under certain conditions. To address this, human HSPCs were transduced with a control or RUNX1-ETO-expressing vector, and proteins fractionated according to nuclear or cytosolic localisation (**Figure 6.1A**). From this analysis, this study identified 256 significantly dysregulated proteins, involved in several pathways previously identified in AML, including

dysregulation of normal myeloid differentiation (Nafria *et al.*, 2020; Ichikawa *et al.*, 2013) and changes in the NF- $\kappa$ B pathway (Zhou *et al.*, 2015). Similarly to the transcriptomic analysis described above, I determined the candidate regulators of the observed phenotype, focusing on TF. I identified, a member of the C/EBP family of TF repressed in RUNX1-ETO cells and predicted to have the greatest influence on other dysregulated proteins. C/EBP $\beta$  was studied in **Chapter 5**.

### 6.1.2 ZNF217 does not contribute to the pathogenesis of AML

Following the identification of *ZNF217* overexpression (**Chapter 3**), this study focused on examining its role in normal haematopoietic development. Given the role of *ZNF217* in the development of several human cancers (**1.4.2**), it was hypothesised that it would act as an oncogene in haematopoietic cells by disrupting normal myeloid growth and development.

*ZNF217* mRNA expression was shown to be highly variable across AML subtypes, but this was lower as compared to mature cells, suggesting a potential developmental role. Moreover, higher levels of *ZNF217* were associated with reduced disease-free survival, indicating a possible role for *ZNF217* in promoting relapse; however, no multivariate analysis was performed to conclusively determine this. Indeed, increased *ZNF217* expression has previously been associated with worse patient outcome and prognosis in several solid tumours, including prostate (Szczyrba *et al.*, 2013), colorectal (Zhang *et al.*, 2015; Li *et al.*, 2015), breast (Krig *et al.*, 2010; Littlepage *et al.*, 2012; Nguyen *et al.*, 2014; Nonet *et al.*, 2001; Plevova *et al.*, 2010; Thollet *et al.*, 2010; Vendrell *et al.*, 2012) and ovarian (Zhu *et al.*, 2009), amongst others, highlighting the prognostic value of *ZNF217* in malignant disease. However, even though the role of *ZNF217* has been extensively studied in solid cancers, its role in AML development has not been studied.

In normal haematopoiesis, *ZNF217* mRNA levels increased throughout myeloid differentiation, with highest expression in mature granulocytes, suggesting a role in granulocytic development. Initially, this study sought to determine the role of *ZNF217* on normal myeloid development using a human primary HSPC model. These studies showed that overexpression of *ZNF217* significantly impaired myeloid colony formation, coupled with a decreased self-potential capacity. In bulk liquid culture, *ZNF217* expression promoted monocytic differentiation, associated with an increased expression of the differentiation markers CD11b and CD14. These results indicate that, in contrast to its role in solid tumours,

it is unlikely that ZNF217 possesses a role in leukaemogenesis, which is characterised by a block in myeloid development. Nevertheless, my data suggests that ZNF217 might have a developmental role in normal haematopoiesis. To address this, KD studies were performed. Although KD of ZNF217 significantly impaired HSPC myeloid colony formation and self-renewal potential, no effect was observed during myeloid growth and differentiation of these cells. This suggests that, although ZNF217 can promote myeloid differentiation, this is not essential for the process to occur (**Figure 6.1B**).

### 6.1.3 KD of C/EBP $\beta$ does not contribute to the pathogenesis of AML

**Chapter 5** focused on analysing the importance of C/EBP $\beta$  expression and dysregulation during normal myeloid development. In **Chapter 3**, analysis of proteomic data revealed that C/EBP $\beta$  is repressed upon the expression of RUNX1-ETO, in human HSPCs. Even though some studies have been performed in an attempt to determine its role in normal human haematopoiesis, its role in the development of AML t(8;21) has not been described to date.

This study first determined the association between *CEBPB* mRNA expression and the pathogenesis of AML. This analysis showed that *CEBPB* expression is highly variable in different AML subtypes, however, interestingly, decreased expression of *CEBPB* was consistently observed in t(8;21) patients within the FAB M2 AML subtype, as compared to non-t(8;21) patients, suggesting an association between RUNX1-ETO and *CEBPB* in the leukaemogenic process. However, mRNA studies related to *CEBPB* expression need to be interpreted with caution. *CEBPB* mRNA results in the synthesis of three C/EBP $\beta$  isoforms: LAP\*, LAP and LIP due to the existence of alternative start codons (Zahnow, 2009; Sears and Sealy, 1994). For this reason, is it impossible to infer expression of each isoform in AML patients and in normal haematopoiesis from mRNA analysis alone.

Analysis of *CEBPB* mRNA expression in normal human haematopoiesis revealed that is expressed across all haematopoietic lineages, with higher expression levels in mature monocytes and granulocytes. Analysis of C/EBP $\beta$  protein that LAP\* is highly expressed in monocytic progenitor cells, however expression was relatively low in granulocyte progenitors. This analysis did not detect either of the remaining two isoforms LAP and LIP. However, this analysis was performed on immature cells, and additional studies on more differentiated sub-populations would be necessary to give a more complete understanding of C/EBP $\beta$  protein expression during development. Nevertheless, these observations suggested a role for C/EBP $\beta$

in myeloid development, particularly in the monocytic population. In agreement with these observations, upregulation of C/EBP $\beta$  has been observed in myelomonocytic cells and macrophages (Williams *et al.*, 1991; Katz *et al.*, 1993; Haas *et al.*, 2010; Gutsch *et al.*, 2011), where it influences cell proliferation (Friedman, 2007). Even though this study was only able to detect low levels of C/EBP $\beta$  in granulocyte-progenitor cells, previous studies have shown that, whilst C/EBP $\alpha$ , another member of the C/EBP family, is important for normal granulopoiesis (Hirai *et al.*, 2006; Zhang *et al.*, 1997; Zhang *et al.*, 2004b; Radomska *et al.*, 1998), C/EBP $\beta$  is essential for stress-induced granulopoiesis (Akagi *et al.*, 2008; Zhang *et al.*, 2010; Hall *et al.*, 2012). In the present study, KD of C/EBP $\beta$  in human HSPC failed to significantly impact myeloid cell growth and development. This observation is likely due to low levels of KD achieved in these cells. Additionally, functional redundancy with other members of the C/EBP family can further account for these results, as previous studies have shown that these TF can act compensatively in the absence of a certain member (Tanaka *et al.*, 1995a; Screpanti *et al.*, 1995; Hu *et al.*, 1998). In line with this, studies have suggested that C/EBP $\alpha$  and C/EBP $\beta$  might collaborate with each other, in an attempt to ensure a proper supply of granulocytes at all times, through the regulation of common target genes linked to granulocytic differentiation (Jones *et al.*, 2002). Moreover, whilst some authors have claimed that KO of C/EBP $\beta$  in early progenitor cells give rise to a lower number of colonies and cells (Hirai *et al.*, 2006), in more differentiated cells, absence of C/EBP $\beta$  results in a pro-proliferative phenotype (Gutsch *et al.*, 2011; Screpanti *et al.*, 1995) (**Figure 6.2**).

Next, I determined the consequences of C/EBP $\beta$  LAP\* overexpression as a single abnormality on the myeloid development on human HSPCs. Overexpression of this isoform increased progenitor cell colony forming ability, as well as self-renewal potential observations consistent with an inhibition of differentiation. However, analysis of granulocytic and monocytic cell surface markers indicated C/EBP $\beta$  LAP\* overexpression promoted the differentiation of these cells. Is it plausible that the increase in self-renewal potential is, in fact, a result of arrested terminal differentiation. Of note, increased expression of C/EBP $\beta$  in both APL and CML cells following treatment with ATRA or Imatinib, respectively, was shown to induce differentiation of AML cell lines (Duprez, 2004; Duprez *et al.*, 2003; Guerzoni *et al.*, 2006), whilst its KD in APL cells reduced the differentiation potential of these cells. These observations suggest that overexpression of the LAP\* isoform promotes haematopoietic development in HSPC.

As KD of C/EBP $\beta$  in HSPC is unlikely to contribute to the pathogenesis of AML through the disruption of normal myeloid development, this study assessed the consequences of knocking down C/EBP $\beta$  in AML cell lines. Interestingly, KD of C/EBP $\beta$  yielded distinct results depending on the context of the transformed cell, likely a result of the cells' developmental stage. In the t(8;21) cell line SKNO-1, KD of C/EBP $\beta$  resulted in an increase in the cells' proliferative ability, combined with the ablation of myeloid cell surface marker expression. However, similar experiments performed in HEL and U937 demonstrated a growth inhibition associated with the increased expression of markers of differentiation. Interestingly, a recent study has reported that inhibition of C/EBP $\beta$  activity actually promoted the expression of differentiation-associated cell surface markers and genes (Yusenko *et al.*, 2021) (**Figure 6.2**).

In conclusion, this study has shown that C/EBP $\beta$  downregulation is associated with AML t(8;21). However, KD of C/EBP $\beta$  in HSPCs did not support an oncogenic role for this TF in normal haematopoiesis, most likely due to insufficient level of KD, suggesting that even low levels of C/EBP $\beta$  are enough for myeloid development to occur. Moreover, the absence of a differentiation block in these cells might be a result of the interplay between C/EBP $\beta$  and other members of the C/EBP family (discussed in ).

#### **6.1.4 Limitations of the study**

The novel targets identified by applying the latest proteomic analysis platform known as SWATH-MS represent the advantage of leveraging a next-generational technique; however, early adoption requires certain concessions. Whilst a great advantage of SWATH-MS is that expression results are direct measures of protein expression, rather than mRNA level, which can be disassociated from actual protein expression due to post-translational regulation, this requires substantially greater cell number for sample preparation, limiting the analysis of rare cell subsets. Further, SWATH-MS relies upon the construction of a library for protein identification. This is often generated by each respective research site, as no standard has been adopted widely, increasing the input requirements, limiting the number of proteins considered, and potential affecting reproducibility. In this study, such requirements precluded the ability to FACS a population as insufficient material would have been recovered. Lastly, integration of proteomic and transcriptomic data at specified timepoints is complicated by the time delay between mRNA expression and protein generation.

The CD34 model used within this study was further limited by access to cord-blood material, restricting the number of CD34<sup>+</sup> HSPC as input for experiments. Due to this, different protocols had to be implemented, depending on the experiment outcome necessary. If the majority of the population had been successfully transduced, these were deemed suitable for analysis albeit with the acknowledgement that untransduced cells were perhaps influencing those with overexpression or knockdown, potentially masking a phenotype. For certain plasmids, this proved a greater challenge as the substantial plasmid size heightened the difficulty of transducing CD34, as seen in overexpression studies. The difficulty was exacerbated in preliminary investigation of double infections, which were deemed infeasible as the proportion of transduced cells harbouring both constructs were too few for subsequent studies.

A subsequent limitation of the current study relies on the fact that certain proteins of interest were not investigated in all of the desired settings, or in multiple time-points. For instance, ZNF217 was not examined in the context of normal haematopoiesis which was due to the scarcity of the material. For the same reason, it was not possible to analyse each target (ZNF217 and C/EBP $\beta$ ) expression throughout the 13-day experiments performed (OE and KD). This was a consequence of the COVID-19 pandemic which prevented cord-blood donations.

A limitation of the cell line approach is the inherent separation from the clinical context of primary material. However, access to such material was not possible and cell lines represented the next best model. Unfortunately, limited cell lines harbour the t(8;21) translocation limiting the sample sizes within this study. Whilst Kasumi-1 is another t(8;21) line which could have been used, this line is very slow growing and very resistant to transduction, limiting the usefulness to this study.

Cell lines were investigated in the context of knockdown by shRNA. A more conclusive endpoint could have been reached by application of CRISPR-Cas9 which is able to generate complete knockout of target genes. However, CRISPR-Cas9 requires tremendous cost both monetarily and in labour to generate a complete knockout clone. As this study wished to examine numerous targets in multiple contexts, shRNA provided the best means to achieve this.

The colony formation assay presented herein is an in-house solution to overcome the poor reproducibility, cost, and access to material associated with Matrigel assays. However, colonies were not examined for their specific subtypes – erythroid, myeloid, granulocytic, macrophage

– as these did not represent an endpoint of interest; rather, colony formation ability was focused on the capacity of self-renewal at the second plating round.

Together, these data represent the first in-depth examination of two novel targets of interest in RUNX1-ETO AML. Whilst sample sizes and access to material limited the study, these data were generated in an attempt to provide a broad understanding of the target mechanisms, a goal which was achieved.

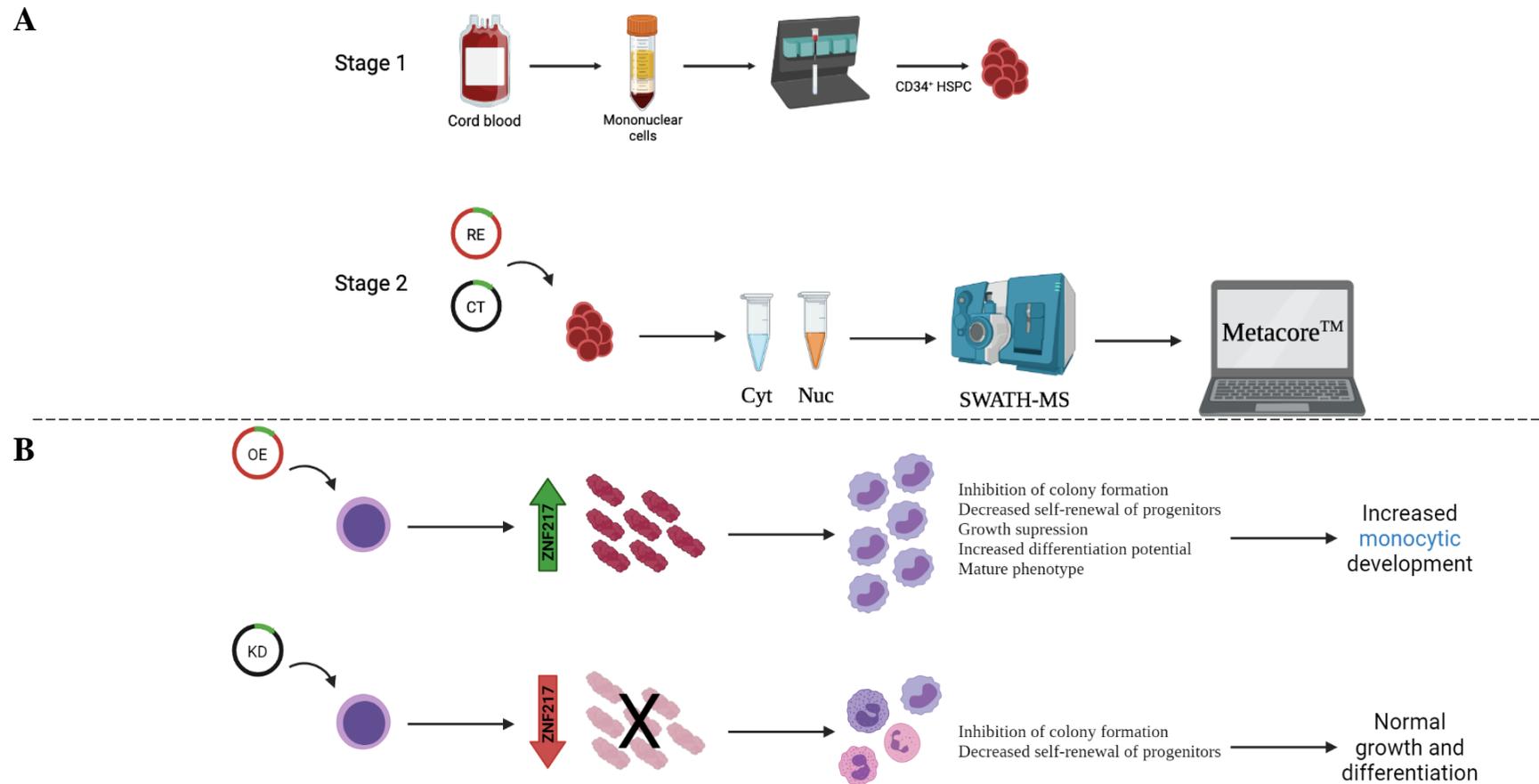
## 6.2 Future directions

Even though the current study provide insights into the role of both ZNF217 and C/EBP $\beta$  in normal haematopoietic development, additional studies would be necessary to support the role of these TF in t(8;21) AML development. Expression of ZNF217 should be knocked-down/out in HSPCs expressing RUNX1-ETO, whilst C/EBP $\beta$  should be overexpressed in the same model, to determine if either of these factors can reverse RUNX1-ETO-induced phenotype. Additionally, the fact that ZNF217, as a single abnormality, was shown to promote myeloid development, suggests that RUNX1-ETO might be inhibiting its normal function. For this reason, ChIP-Seq analysis of RUNX-ETO cells, co-transduced with a ZNF217-overexpressing vector should be performed, to determine chromatin co-occupancy of both RUNX1-ETO and ZNF217.

Regarding C/EBP $\beta$  studies, this study hypothesised that the failure to induce a significant block in myeloid differentiation was a result of two separate factors, and additional experiments would be necessary to test them. On one side, it was not possible to achieve efficient KD. To address this, CRISPR-Cas9 experiments could be applied in the same HSPC model, to try to achieve complete KO. Furthermore, considering the functional redundancy within the C/EBP family, KD, or KO, of other C/EBP proteins, such as C/EBP $\alpha$  or C/EBP $\epsilon$ , in combination with C/EBP $\beta$  should be employed to determine the extent of the interactions between all the family members in haematopoietic development.

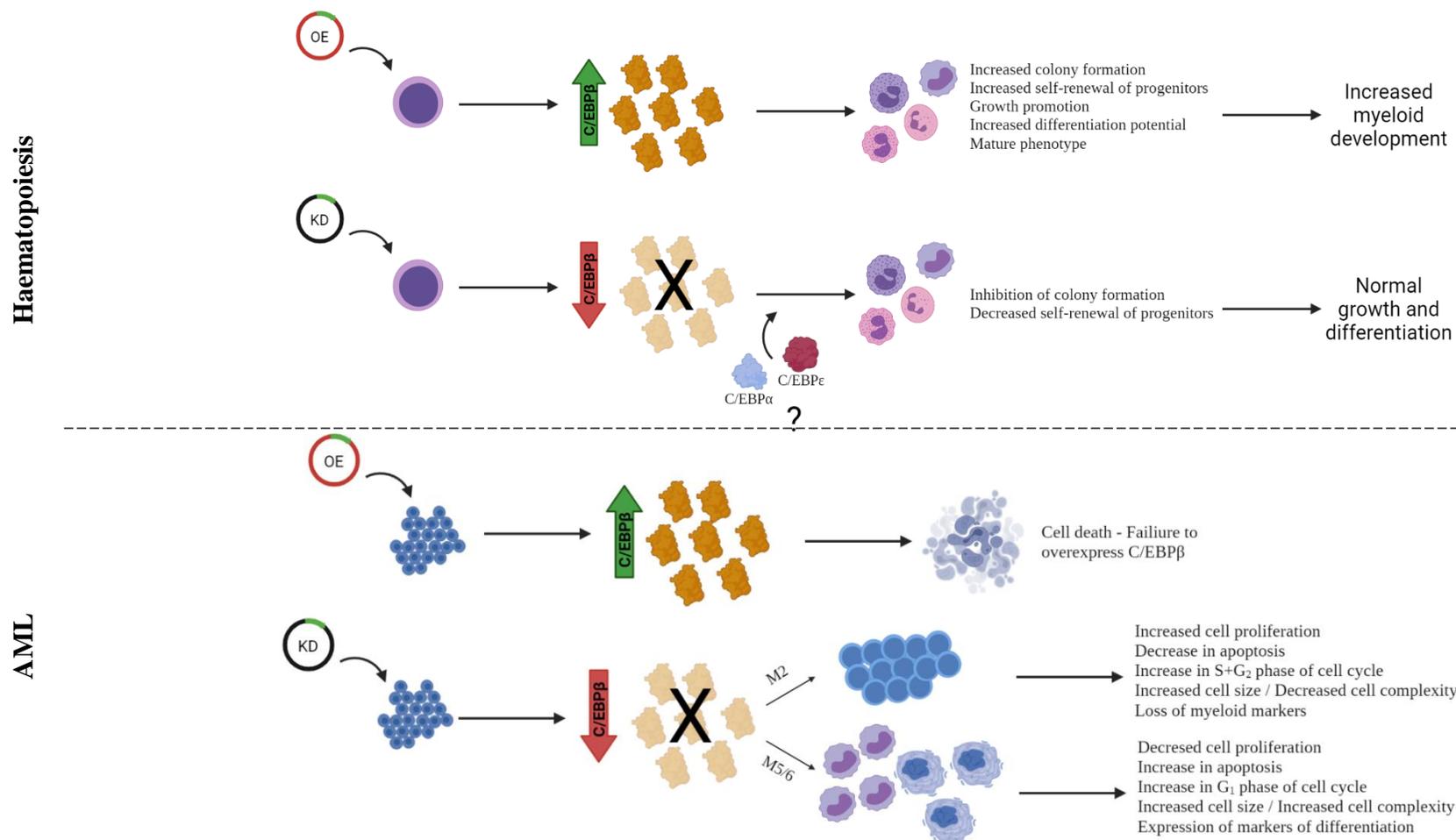
Furthermore, studies in additional models would allow the complete representation of both ZNF217 and C/EBP $\beta$  in other contexts. These would include *in vivo* murine studies, to match the *in vitro* experiments in HSPCs; patient-derived primary material, particularly those with t(8;21), with the intent to replicate the AML cell lines result; and, lastly, induced pluripotent stem cells (iPSCs), to determine the role of these TF in other contexts, for instance, in fully

differentiated monocytes or granulocytes. Additionally, it would be beneficial to target ZNF217 and C/EBP $\beta$  with clinically transferable inhibitors and activators, respectively, in the context of AML t(8;21). Unfortunately, no agents currently exist in case of C/EBP $\beta$  and the only inhibitor of ZNF217 has been abandoned (Collins *et al.*, 2007). Furthermore, other targets could be investigated, such as *ARID5B* or *IRF9*, described in **Chapter 3**, to determine if these would contribute to the leukaemogenic development.



**Figure 6.1 – Summary of the main findings described on Chapters 3 and 4**

(A) Human CD34<sup>+</sup> HSPC were isolated from neonatal cord blood (Stage 1) and transduced with retrovirus co-expressing either RUNX1-ETO and GFP (RE), or GFP alone (CT). Transduced cells were fractionated into cytosolic and nuclear protein fractions, and analysed by performing SWATH-MS to determine protein expression. Differentially expressed proteins were uploaded into Metacore for pathway analysis and target identification (Stage 2). (B) To determine the role of ZNF217 in normal human haematopoiesis, CD34<sup>+</sup> HSPC were isolated as described in (A). Following this, functional studies using overexpression and KD constructs were applied to study the cells myeloid development.



**Figure 6.2 – Summary of the main findings described on Chapter 5**

To determine the role of C/EBP $\beta$  in normal human haematopoiesis, CD34<sup>+</sup> HSPC were isolated as described in **Figure 6.1A**. Following this, functional studies using overexpression and KD constructs were applied to study the cells myeloid development. Additionally, to determine the role of C/EBP $\beta$  in leukaemogenic development, the same KD constructs were applied to AML cell lines, with different characteristics and stages of differentiation: SKNO-1 (M2), HEL (M6) and U937 (M5).

**Supplementary Tables****Supplementary Table 1 – Compensation matrix**

Compensation was performed computationally post-acquisition according to the following matrix:

	<b>FITC</b>	<b>PE</b>	<b>PERCP</b>	<b>APC</b>
<b>FITC</b>	1	-0.11	0.011	0
<b>PE</b>	-0.024	1	-0.22	0
<b>PERCP</b>	-0.002	-0.0001	1	-0.15
<b>APC</b>	0	0	-0.06	1

**Supplementary Table 2 – Identification of significantly dysregulated proteins, as a result of RUNX1-ETO expression**

SWATH-MS identified 256 significantly dysregulated proteins in HSPC, as a result of increased RUNX1-ETO expression.

Fold-change (FC) values represent the regulation of each protein when compared to control cells (green – upregulated; red – repressed, when compared to control cells). Statistical analysis was performed using one-way ANOVA, without p-value correction (n=3).

<b>Protein</b>	<b>Localisation</b>	<b>p-value</b>	<b>FC</b>
<b>CA2</b>	Cytoplasm	0.0053	6.19
<b>GBF1</b>	Cytoplasm	0.0237	4.03
<b>MAGED2</b>	Cytoplasm	0.0016	3.47
<b>COPZ1</b>	Nucleus	0.0022	3.38
<b>ALDH1L1</b>	Cytoplasm	0.0023	3.30
<b>RNH1</b>	Nucleus	0.0462	3.24
<b>NFKB2</b>	Cytoplasm	0.0003	3.17
<b>GRN</b>	Cytoplasm	0.0253	3.16
<b>ALDH1A1</b>	Cytoplasm	0.0066	3.06
<b>CMPK2</b>	Cytoplasm	0.0308	2.96
<b>MAN1A1</b>	Cytoplasm	0.0199	2.80
<b>SERPINB9</b>	Cytoplasm	0.0052	2.75
<b>CLPP</b>	Cytoplasm	0.0301	2.72
<b>KLC1</b>	Cytoplasm	0.0153	2.60
<b>ANXA2</b>	Cytoplasm	0.0126	2.58
<b>KTI12</b>	Cytoplasm	0.0175	2.58
<b>INTS12</b>	Nucleus	0.0159	2.57
<b>NMT1</b>	Nucleus	0.0451	2.55
<b>PPP4R1</b>	Cytoplasm	0.0004	2.54
<b>EVI2B</b>	Cytoplasm	0.0428	2.51
<b>COL4A3BP</b>	Cytoplasm	0.0058	2.50
<b>NT5C</b>	Cytoplasm	0.0019	2.48
<b>RB1</b>	Nucleus	0.0154	2.43
<b>NIT1</b>	Cytoplasm	0.0013	2.40
<b>HSPB1</b>	Cytoplasm	0.0065	2.36
<b>HPGDS</b>	Cytoplasm	0.0381	2.36
<b>ICAM1</b>	Cytoplasm	0.0242	2.35

<b>JAK1</b>	Cytoplasm	0.0468	2.32
<b>RAB2A</b>	Cytoplasm	0.0117	2.17
<b>RNASEH2B</b>	Cytoplasm	0.0079	2.16
<b>NUCB1</b>	Cytoplasm	0.0033	2.15
<b>PDCD4</b>	Cytoplasm	0.0051	2.11
<b>ZNF142</b>	Cytoplasm	0.0259	2.10
<b>LAIR1</b>	Cytoplasm	0.0419	2.10
<b>TXN</b>	Cytoplasm	0.0109	2.08
<b>BRCC3</b>	Cytoplasm	0.0080	2.07
<b>NKAP</b>	Nucleus	0.0004	2.06
<b>ARRDC1</b>	Nucleus	0.0396	2.06
<b>CENPF</b>	Nucleus	0.0094	2.05
<b>NEDD8</b>	Cytoplasm	0.0043	2.00
<b>PYM1</b>	Cytoplasm	0.0025	1.99
<b>ZNF428</b>	Cytoplasm	0.0029	1.97
<b>SPTAN1</b>	Nucleus	0.0154	1.95
<b>USP47</b>	Cytoplasm	0.0424	1.95
<b>FAM91A1</b>	Cytoplasm	0.0223	1.93
<b>DNASE2</b>	Cytoplasm	0.0485	1.90
<b>MARCKSL1</b>	Cytoplasm	0.0115	1.90
<b>RNF214</b>	Cytoplasm	0.0036	1.87
<b>FKBP8</b>	Nucleus	0.0046	1.87
<b>ISYNA1</b>	Cytoplasm	0.0356	1.86
<b>HIST1H1B</b>	Cytoplasm	0.0365	1.86
<b>MED12</b>	Nucleus	0.0447	1.86
<b>B2M</b>	Cytoplasm	0.0474	1.83
<b>ap1g2</b>	Cytoplasm	0.0395	1.82
<b>EMG1</b>	Cytoplasm	0.0042	1.82
<b>NDUFA4</b>	Nucleus	0.0430	1.80
<b>PRKAR2A</b>	Nucleus	0.0220	1.78
<b>MRPS6</b>	Cytoplasm	0.0024	1.77
<b>ARFIP2</b>	Cytoplasm	0.0019	1.76
<b>IKBKG</b>	Cytoplasm	0.0443	1.75
<b>RUNX1</b>	Nucleus	0.0208	1.73
<b>BICD2</b>	Cytoplasm	0.0303	1.72
<b>STX7</b>	Cytoplasm	0.0396	1.72

<b>HSP90AB4P</b>	Nucleus	0.0378	1.71
<b>CCT8</b>	Nucleus	0.0318	1.70
<b>PTER</b>	Cytoplasm	0.0128	1.69
<b>DBNL</b>	Nucleus	0.0478	1.68
<b>IFI16</b>	Cytoplasm	0.0185	1.68
<b>RAC1</b>	Cytoplasm	0.0029	1.66
<b>ITGB1</b>	Cytoplasm	0.0258	1.66
<b>SH3BGRL</b>	Cytoplasm	0.0219	1.65
<b>LGALS9</b>	Cytoplasm	0.0079	1.65
<b>TSHZ1</b>	Cytoplasm	0.0449	1.62
<b>DDAH2</b>	Cytoplasm	0.0102	1.61
<b>PLEKHA7</b>	Cytoplasm	0.0257	1.60
<b>EIF6</b>	Cytoplasm	0.0100	1.57
<b>SERPINH1</b>	Cytoplasm	0.0174	1.56
<b>TM9SF2</b>	Cytoplasm	0.0073	1.56
<b>KMT2A</b>	Nucleus	0.0060	1.55
<b>RPS27L</b>	Nucleus	0.0278	1.55
<b>RPS2</b>	Nucleus	0.0218	1.55
<b>ARHGDIB</b>	Cytoplasm	0.0370	1.54
<b>SEC23A</b>	Cytoplasm	0.0137	1.53
<b>SET</b>	Cytoplasm	0.0381	1.53
<b>TRIM28</b>	Cytoplasm	0.0007	1.51
<b>NADSYN1</b>	Cytoplasm	0.0433	1.51
<b>DHX38</b>	Cytoplasm	0.0318	1.50
<b>NDRG3</b>	Cytoplasm	0.0272	1.48
<b>FKBP3</b>	Cytoplasm	0.0014	1.48
<b>RPL13A</b>	Nucleus	0.0102	1.48
<b>SERPINB6</b>	Cytoplasm	0.0203	1.47
<b>NAMPT</b>	Cytoplasm	0.0282	1.47
<b>TMEM263</b>	Cytoplasm	0.0351	1.46
<b>LASP1</b>	Cytoplasm	0.0173	1.46
<b>ATP6V1C1</b>	Cytoplasm	0.0377	1.46
<b>FKBP1A</b>	Cytoplasm	0.0309	1.46
<b>MPI</b>	Nucleus	0.0391	1.46
<b>CSNK2A2</b>	Cytoplasm	0.0121	1.46
<b>ACBD3</b>	Cytoplasm	0.0079	1.45

<b>IRGQ</b>	Cytoplasm	0.0024	1.45
<b>VARS</b>	Cytoplasm	0.0287	1.45
<b>LAMP1</b>	Cytoplasm	0.0066	1.45
<b>RRM1</b>	Cytoplasm	0.0018	1.45
<b>PRDX1</b>	Cytoplasm	0.0219	1.44
<b>RPE</b>	Cytoplasm	0.0425	1.44
<b>ZBTB48</b>	Nucleus	0.0069	1.43
<b>RPLP2</b>	Nucleus	0.0083	1.43
<b>CDC73</b>	Cytoplasm	0.0218	1.43
<b>EDF1</b>	Nucleus	0.0110	1.43
<b>PHAX</b>	Cytoplasm	0.0311	1.43
<b>GLB1</b>	Cytoplasm	0.0480	1.43
<b>rbm12</b>	Cytoplasm	0.0046	1.42
<b>STMN1</b>	Cytoplasm	0.0029	1.42
<b>BAZ2A</b>	Nucleus	0.0366	1.42
<b>CAP1</b>	Cytoplasm	0.0358	1.41
<b>TMED10</b>	Cytoplasm	0.0171	1.40
<b>SNX6</b>	Cytoplasm	0.0007	1.40
<b>AGPAT1</b>	Cytoplasm	0.0449	1.40
<b>APBB1IP</b>	Cytoplasm	0.0042	1.40
<b>ARHGAP32</b>	Cytoplasm	0.0260	1.40
<b>PPIA</b>	Cytoplasm	0.0084	1.39
<b>SEC23IP</b>	Cytoplasm	0.0255	1.39
<b>CBFB</b>	Nucleus	0.0299	1.39
<b>ACTR3</b>	Cytoplasm	0.0182	1.39
<b>CFL1</b>	Cytoplasm	0.0201	1.38
<b>ETFB</b>	Cytoplasm	0.0481	1.38
<b>POTEF</b>	Cytoplasm	0.0269	1.37
<b>INPP5D</b>	Nucleus	0.0111	1.37
<b>GSTO1</b>	Cytoplasm	0.0022	1.37
<b>PGK1</b>	Cytoplasm	0.0028	1.36
<b>RPL7</b>	Nucleus	0.0405	1.35
<b>MTHFD2</b>	Cytoplasm	0.0348	1.35
<b>PSMF1</b>	Cytoplasm	0.0338	1.35
<b>PSMC3</b>	Cytoplasm	0.0166	1.35
<b>UROD</b>	Cytoplasm	0.0224	1.34

<b>AMPD2</b>	Cytoplasm	0.0087	1.34
<b>PLIN3</b>	Cytoplasm	0.0346	1.33
<b>NDUFV2</b>	Nucleus	0.0352	1.33
<b>MMS22L</b>	Cytoplasm	0.0238	1.32
<b>LRRFIP1</b>	Cytoplasm	0.0070	1.31
<b>NUP93</b>	Nucleus	0.0323	1.31
<b>CCT6A</b>	Cytoplasm	0.0254	1.31
<b>MRPS30</b>	Cytoplasm	0.0145	1.30
<b>PHIP</b>	Nucleus	0.0134	1.30
<b>URM1</b>	Cytoplasm	0.0168	1.30
<b>THUMPD1</b>	Cytoplasm	0.0169	1.29
<b>HEXA</b>	Cytoplasm	0.0157	1.29
<b>CASP6</b>	Cytoplasm	0.0500	1.27
<b>PFDN1</b>	Cytoplasm	0.0318	1.27
<b>CDC42BPB</b>	Nucleus	0.0350	1.26
<b>ADH5</b>	Cytoplasm	0.0467	1.25
<b>KNTC1</b>	Cytoplasm	0.0005	1.25
<b>ABCF1</b>	Cytoplasm	0.0382	1.25
<b>LPXN</b>	Cytoplasm	0.0169	1.25
<b>RPS8</b>	Nucleus	0.0022	1.22
<b>ABHD10</b>	Cytoplasm	0.0221	1.22
<b>EWSR1</b>	Cytoplasm	0.0448	1.19
<b>RBM10</b>	Nucleus	0.0382	1.13
<b>PSMB2</b>	Cytoplasm	0.0069	1.12
<b>TUFM</b>	Cytoplasm	0.0229	1.10
<b>GAR1</b>	Nucleus	0.0138	-1.06
<b>FAM96A</b>	Cytoplasm	0.0283	-1.13
<b>RPA3</b>	Cytoplasm	0.0090	-1.15
<b>NUDT5</b>	Cytoplasm	0.0489	-1.16
<b>STAG2</b>	Nucleus	0.0162	-1.17
<b>RPS21</b>	Cytoplasm	0.0208	-1.18
<b>POLR2H</b>	Nucleus	0.0157	-1.18
<b>NCBP1</b>	Cytoplasm	0.0233	-1.18
<b>CUL1</b>	Cytoplasm	0.0285	-1.18
<b>SUPT5H</b>	Nucleus	0.0042	-1.19
<b>zbtb49</b>	Nucleus	0.0389	-1.19

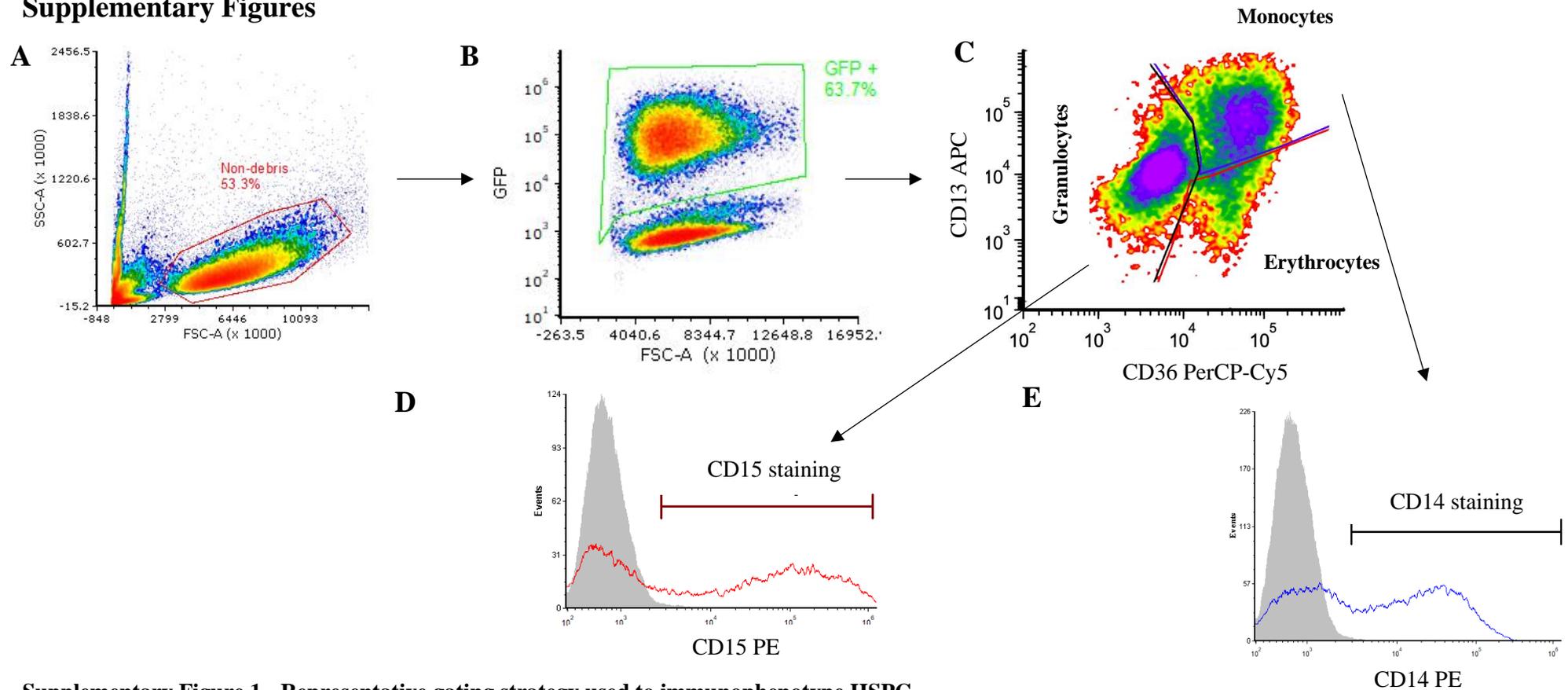
<b>PKM</b>	Cytoplasm	0.0474	-1.21
<b>RIF1</b>	Nucleus	0.0007	-1.21
<b>SNRPD3</b>	Cytoplasm	0.0278	-1.22
<b>PTPN23</b>	Cytoplasm	0.0491	-1.22
<b>RETSAT</b>	Cytoplasm	0.0099	-1.25
<b>SERPINC1</b>	Cytoplasm	0.0334	-1.26
<b>ESF1</b>	Nucleus	0.0476	-1.26
<b>EXOSC9</b>	Nucleus	0.0156	-1.26
<b>CORO1A</b>	Cytoplasm	0.0097	-1.27
<b>NUP88</b>	Cytoplasm	0.0405	-1.27
<b>ITIH2</b>	Nucleus	0.0397	-1.27
<b>NDUFB11</b>	Cytoplasm	0.0468	-1.29
<b>TOR1AIP2</b>	Cytoplasm	0.0221	-1.29
<b>POLA2</b>	Cytoplasm	0.0404	-1.30
<b>TARS2</b>	Cytoplasm	0.0397	-1.31
<b>SCAF11</b>	Nucleus	0.0381	-1.31
<b>INCENP</b>	Nucleus	0.0008	-1.31
<b>MCMBP</b>	Cytoplasm	0.0297	-1.31
<b>c8orf33</b>	Nucleus	0.0068	-1.34
<b>UCK2</b>	Cytoplasm	0.0012	-1.35
<b>HIRA</b>	Nucleus	0.0380	-1.36
<b>LIMA1</b>	Nucleus	0.0358	-1.38
<b>ACSL3</b>	Cytoplasm	0.0406	-1.40
<b>HIRIP3</b>	Nucleus	0.0042	-1.40
<b>PYCARD</b>	Cytoplasm	0.0476	-1.40
<b>CYFIP1</b>	Cytoplasm	0.0347	-1.41
<b>EXOSC4</b>	Nucleus	0.0011	-1.41
<b>KPNA2</b>	Cytoplasm	0.0006	-1.44
<b>RPS3</b>	Cytoplasm	0.0164	-1.44
<b>PTPN6</b>	Cytoplasm	0.0405	-1.45
<b>KIF20B</b>	Nucleus	0.0366	-1.46
<b>MCM7</b>	Nucleus	0.0196	-1.47
<b>ARRB2</b>	Nucleus	0.0200	-1.48
<b>CHAF1B</b>	Nucleus	0.0059	-1.48
<b>TBC1D5</b>	Cytoplasm	0.0478	-1.48
<b>ITGB2</b>	Cytoplasm	0.0074	-1.49

<b>LAMTOR2</b>	Cytoplasm	0.0283	-1.49
<b>CPSF3</b>	Cytoplasm	0.0364	-1.50
<b>CAPG</b>	Cytoplasm	0.0473	-1.50
<b>EFHD2</b>	Cytoplasm	0.0274	-1.52
<b>MCM4</b>	Nucleus	0.0245	-1.53
<b>KIF22</b>	Nucleus	0.0474	-1.53
<b>MARS2</b>	Cytoplasm	0.0310	-1.54
<b>ATL3</b>	Cytoplasm	0.0176	-1.54
<b>LSP1</b>	Nucleus	0.0086	-1.54
<b>SLC1A4</b>	Cytoplasm	0.0328	-1.56
<b>NCF4</b>	Cytoplasm	0.0125	-1.56
<b>SNRPE</b>	Cytoplasm	0.0250	-1.58
<b>TAGLN2</b>	Cytoplasm	0.0487	-1.58
<b>KRI1</b>	Nucleus	0.0006	-1.59
<b>SLC16A3</b>	Cytoplasm	0.0258	-1.60
<b>MKI67</b>	Nucleus	0.0402	-1.63
<b>CKAP2</b>	Nucleus	0.0356	-1.66
<b>CARNMT1</b>	Cytoplasm	0.0174	-1.67
<b>GALE</b>	Cytoplasm	0.0235	-1.70
<b>MPO</b>	Cytoplasm	0.0189	-1.71
<b>ALOX5AP</b>	Cytoplasm	0.0089	-1.76
<b>GFAP</b>	Nucleus	0.0355	-1.79
<b>TOP2A</b>	Nucleus	0.0042	-1.81
<b>TTF2</b>	Nucleus	0.0371	-1.83
<b>MORF4L2</b>	Nucleus	0.0362	-1.84
<b>LGALS1</b>	Cytoplasm	0.0223	-1.85
<b>NDUFS6</b>	Cytoplasm	0.0213	-1.88
<b>SRP19</b>	Cytoplasm	0.0141	-1.88
<b>SPI1</b>	Nucleus	0.0011	-1.89
<b>DOK2</b>	Cytoplasm	0.0497	-1.92
<b>XPNPEP1</b>	Nucleus	0.0327	-1.99
<b>XRCC4</b>	Nucleus	0.0339	-2.02
<b>OXSM</b>	Cytoplasm	0.0296	-2.06
<b>CEBPB</b>	Nucleus	0.0117	-2.09
<b>RAD51C</b>	Cytoplasm	0.0429	-2.13
<b>MPO</b>	Nucleus	0.0357	-2.15

---

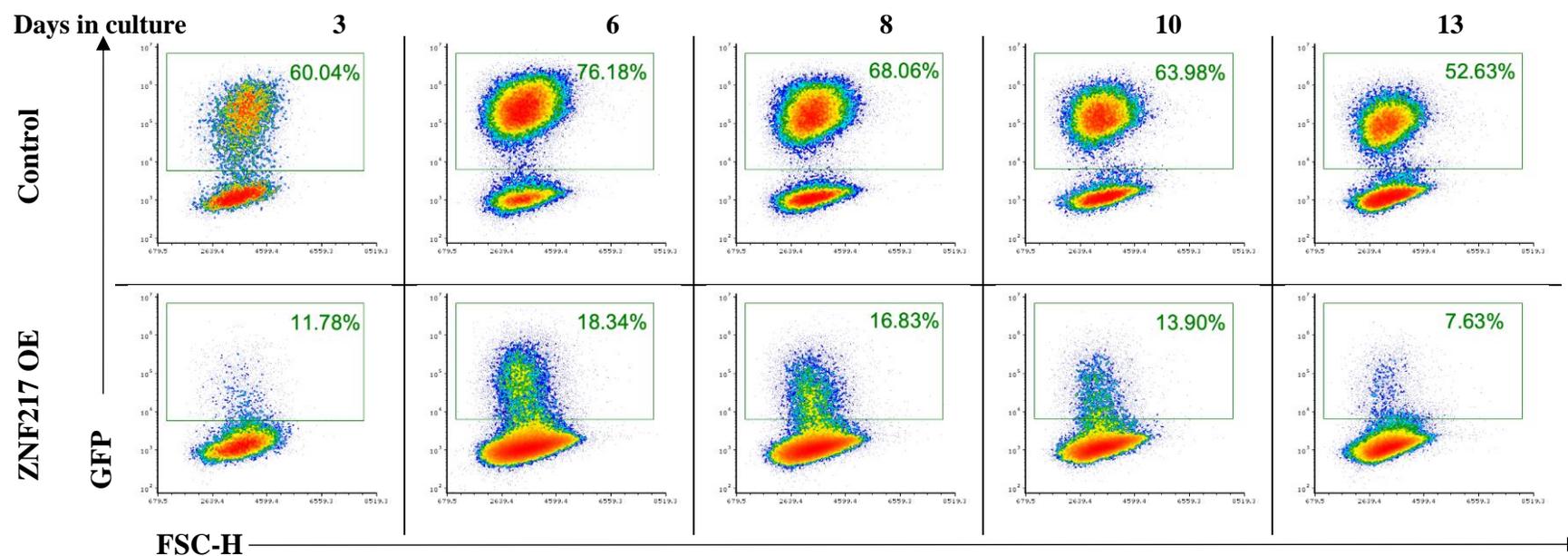
<b>CTSG</b>	Nucleus	0.0422	-2.15
<b>SSR1</b>	Cytoplasm	0.0119	-2.22
<b>WBP11</b>	Cytoplasm	0.0439	-2.24
<b>MRPS2</b>	Cytoplasm	0.0105	-2.25
<b>SAP18</b>	Cytoplasm	0.0279	-2.42
<b>AURKA</b>	Nucleus	0.0212	-2.44
<b>PRG2</b>	Nucleus	0.0153	-2.60
<b>MINA</b>	Nucleus	0.0119	-2.64
<b>CHTF18</b>	Cytoplasm	0.0326	-2.71
<b>CTSG</b>	Cytoplasm	0.0407	-2.94
<b>KRT10</b>	Cytoplasm	0.0271	-3.14
<b>APOBR</b>	Cytoplasm	0.0286	-3.71
<b>CLC</b>	Cytoplasm	0.0165	-5.11
<b>RNASE3</b>	Cytoplasm	0.0069	-7.84

## Supplementary Figures



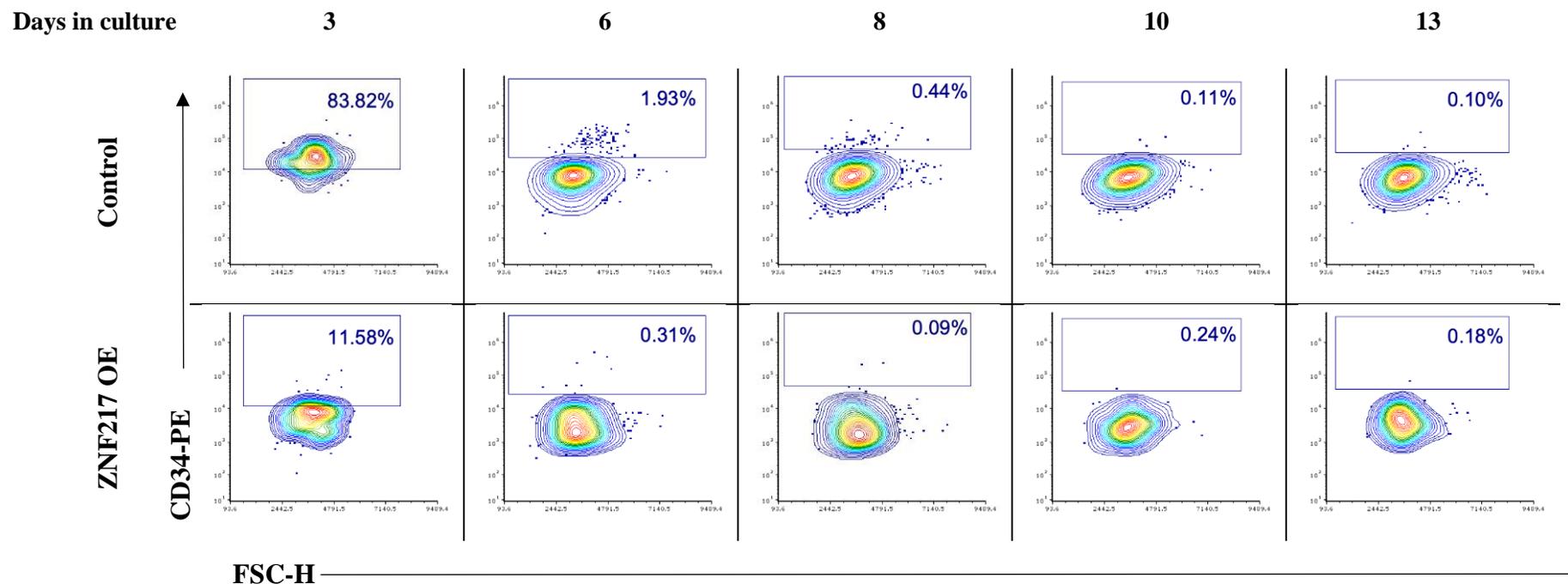
**Supplementary Figure 1 - Representative gating strategy used to immunophenotype HSPC**

Representative density plots and histograms outlining the gating strategy used for the immunophenotypic analysis of HSPC, as described in XX. **(A)** Example gate used to exclude cell debris, based on FSC and SSC; **(B)** The percentage of GFP<sup>+</sup> cells within a bulk population was established based on the 'non-debris' gate shown in **(A)**; **(C)** The CD13-APC and CD36-PerCP-Cy5 markers were used to discriminate between the granulocytic, monocytic and erythroid lineages within the GFP<sup>+</sup> population. Background auto fluorescence was established using cells subjected to the equivalent viral infection procedure but in the absence of virus (mock culture). **(D-E)** Each subpopulation was subsequently examined for differentiation markers, such as the granulocytic marker CD15 or the monocytic marker CD14. Cells stained the isotype control IgG (grey) were used to set the appropriate gating strategies.



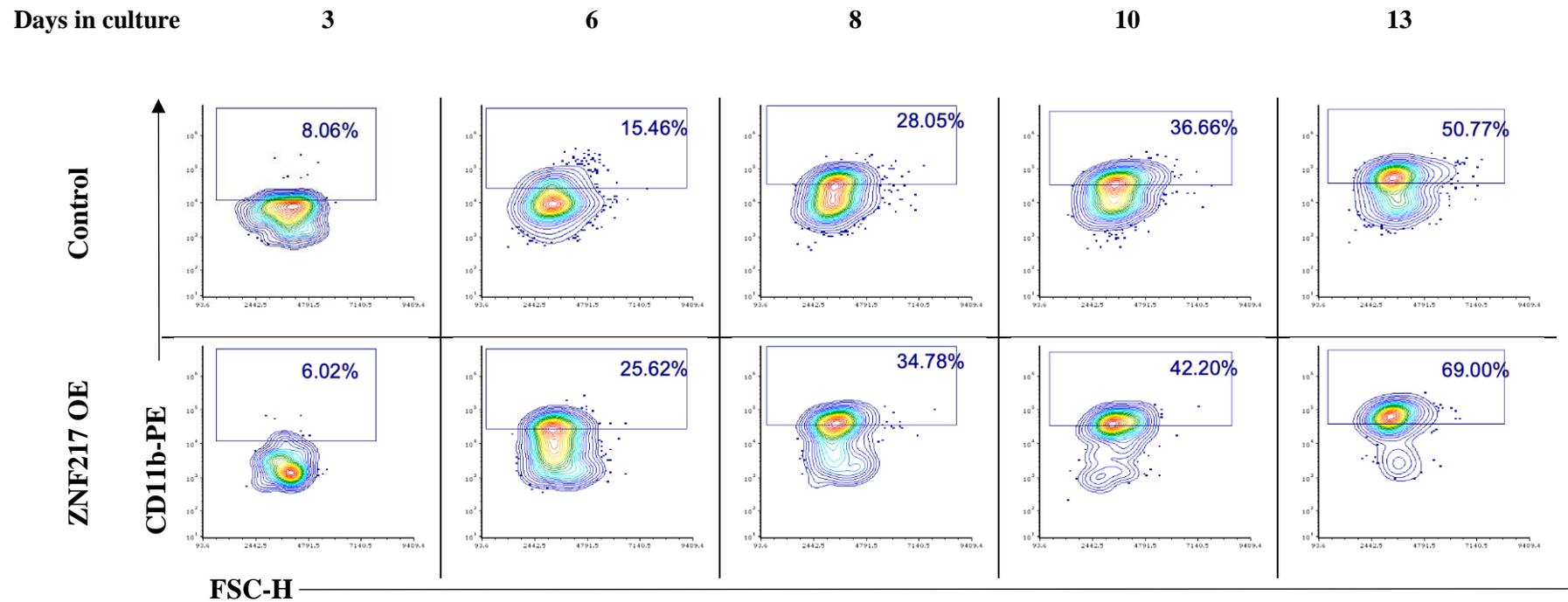
**Supplementary Figure 2 – Expression of control GFP and ZNF217 in human HSPC**

Representative density plots showing the expression of GFP in control and ZNF217 overexpressing cultures, throughout 13 days of culture, as measured by flow cytometry. An initial gating to exclude cell debris was applied, as described in **Supplementary Figure 1A**. Background autofluorescence was established using cells subjected to the equivalent viral infection procedure but in the absence of virus (mock culture).



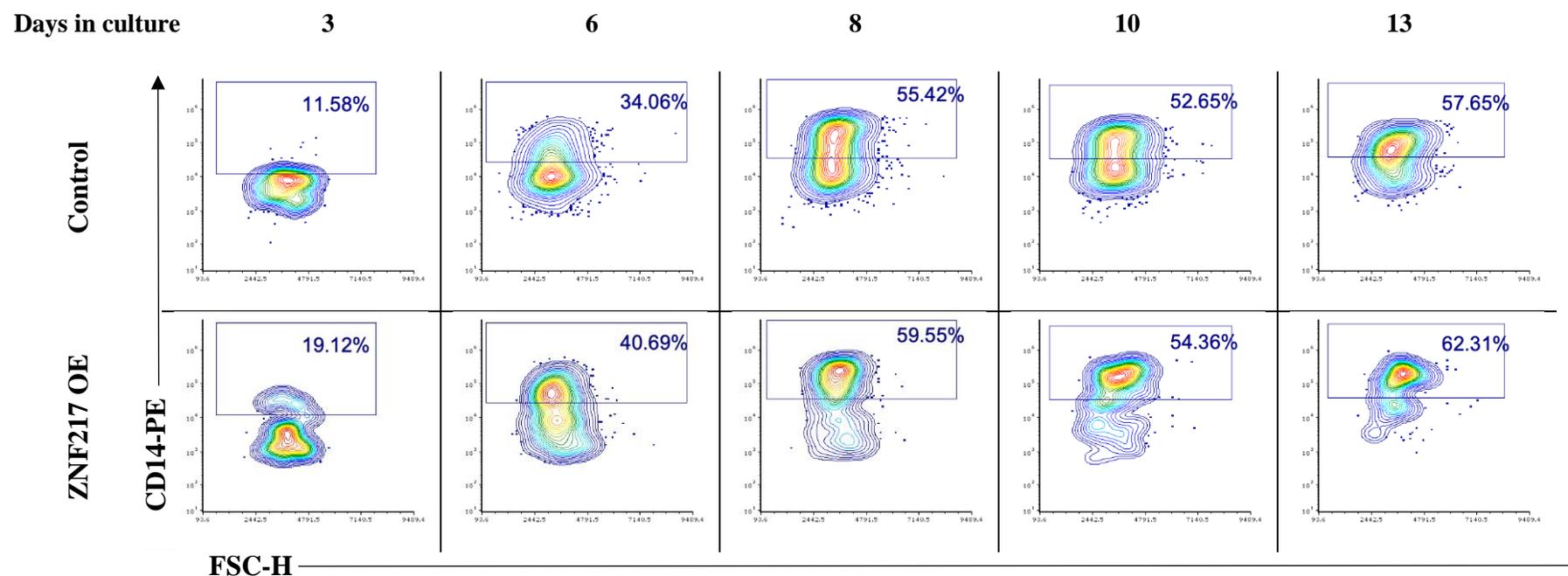
**Supplementary Figure 3 – Expression of the stem cell marker CD34 in monocytes during differentiation, following ZNF217 overexpression**

Representative contour plots showing the expression of CD34 in monocytes (CD13<sup>+</sup> CD36<sup>+</sup>) transduced with a control or a ZNF217 overexpression vector throughout 13 days of culture, as measured by flow cytometry. An initial gating to exclude cell debris was applied, followed by the selection of GFP<sup>+</sup> cells, as described in **Supplementary Figure 1**. Transduced haematopoietic populations were subsequently delimited through the myeloid markers CD13 and CD36. Basal expression was determined using the corresponding IgG control.



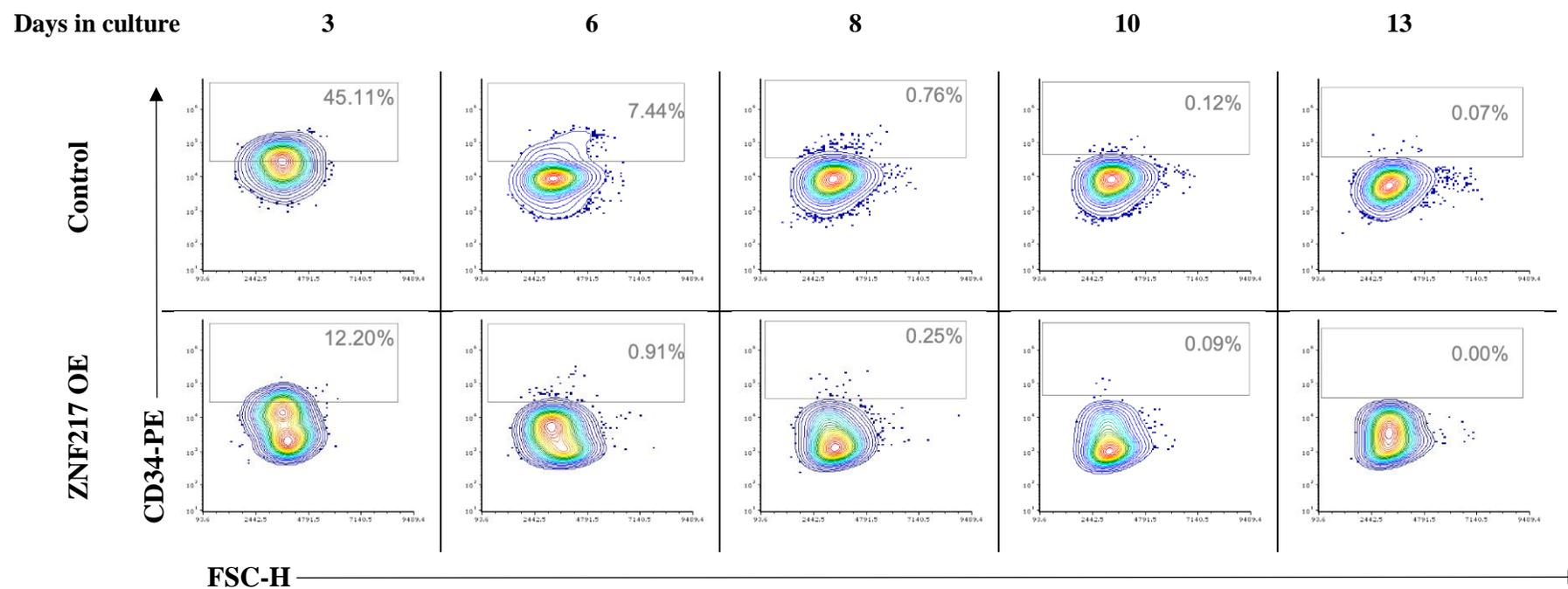
**Supplementary Figure 4 – Expression of the myeloid marker CD11b in monocytes during differentiation, following ZNF217 overexpression**

Representative contour plots showing the expression of CD11b in monocytes (CD13<sup>+</sup> CD36<sup>+</sup>) transduced with a control or a ZNF217 overexpression vector throughout 13 days of culture, as measured by flow cytometry. An initial gating to exclude cell debris was applied, followed by the selection of GFP<sup>+</sup> cells, as described in **Supplementary Figure 1**. Transduced haematopoietic populations were subsequently delimited through the myeloid markers CD13 and CD36. Basal expression was determined using the corresponding IgG control.



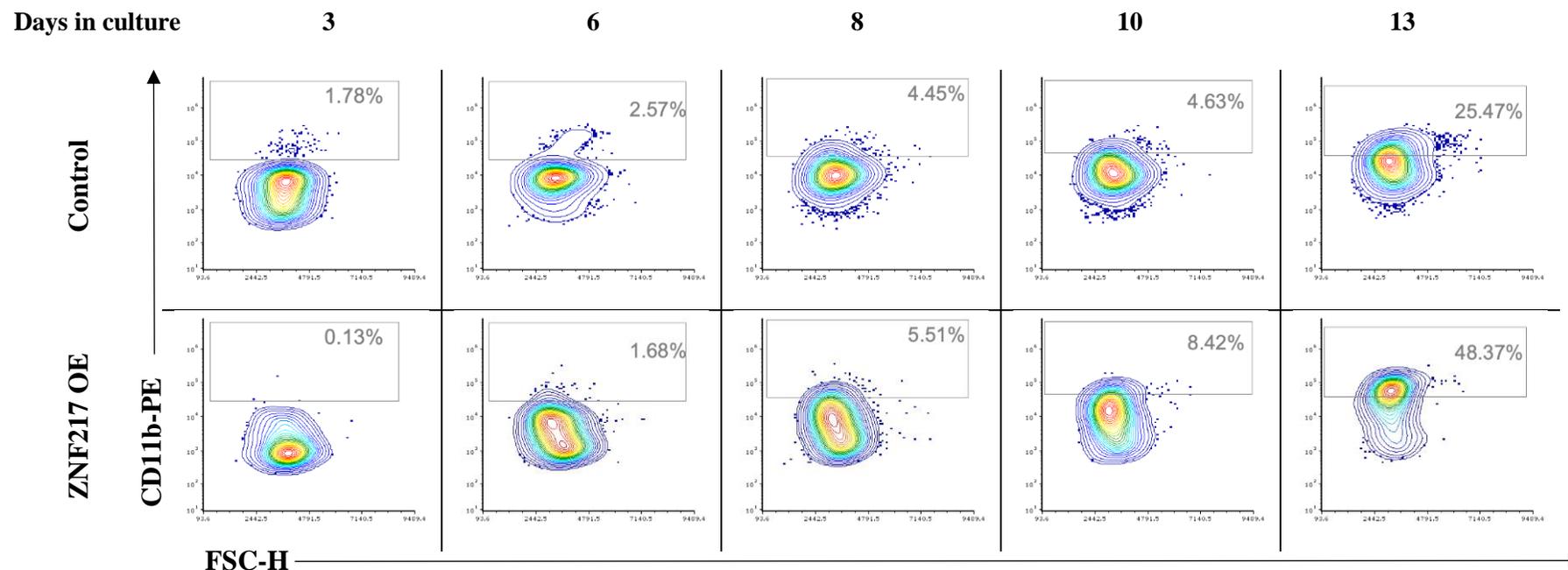
**Supplementary Figure 5 – Expression of the monocytic marker CD14 in monocytes during differentiation, following ZNF217 overexpression**

Representative contour plots showing the expression of CD14 in monocytes ( $CD13^+ CD36^+$ ) transduced with a control or a ZNF217 overexpression vector throughout 13 days of culture, as measured by flow cytometry. An initial gating to exclude cell debris was applied, followed by the selection of  $GFP^+$  cells, as described in **Supplementary Figure 1**. Transduced haematopoietic populations were subsequently delimited through the myeloid markers CD13 and CD36. Basal expression was determined using the corresponding IgG control.



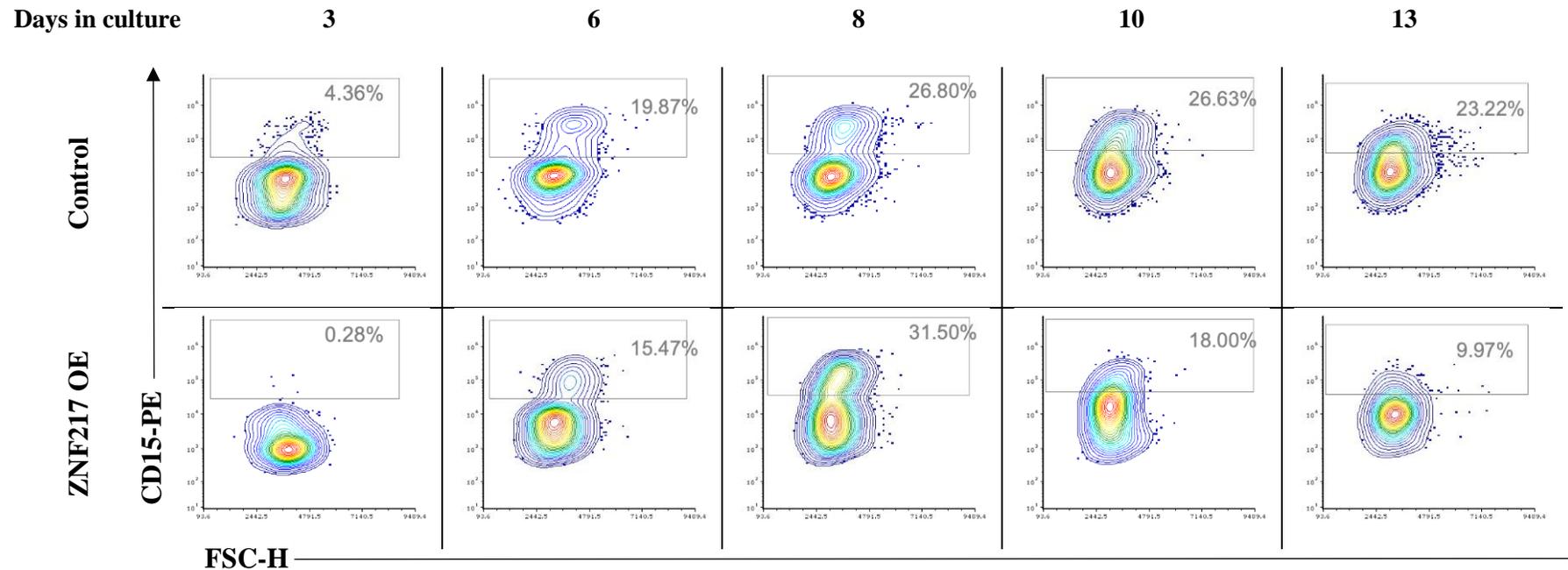
**Supplementary Figure 6 – Expression of the stem cell marker CD34 in granulocytes during differentiation, following ZNF217 overexpression**

Representative contour plots showing the expression of CD34 in granulocytes (CD13<sup>+</sup> CD36<sup>-</sup>) transduced with a control or a ZNF217 overexpression vector throughout 13 days of culture, as measured by flow cytometry. An initial gating to exclude cell debris was applied, followed by the selection of GFP<sup>+</sup> cells, as described in **Supplementary Figure 1**. Transduced haematopoietic populations were subsequently delimited through the myeloid markers CD13 and CD36. Basal expression was determined using the corresponding IgG control.



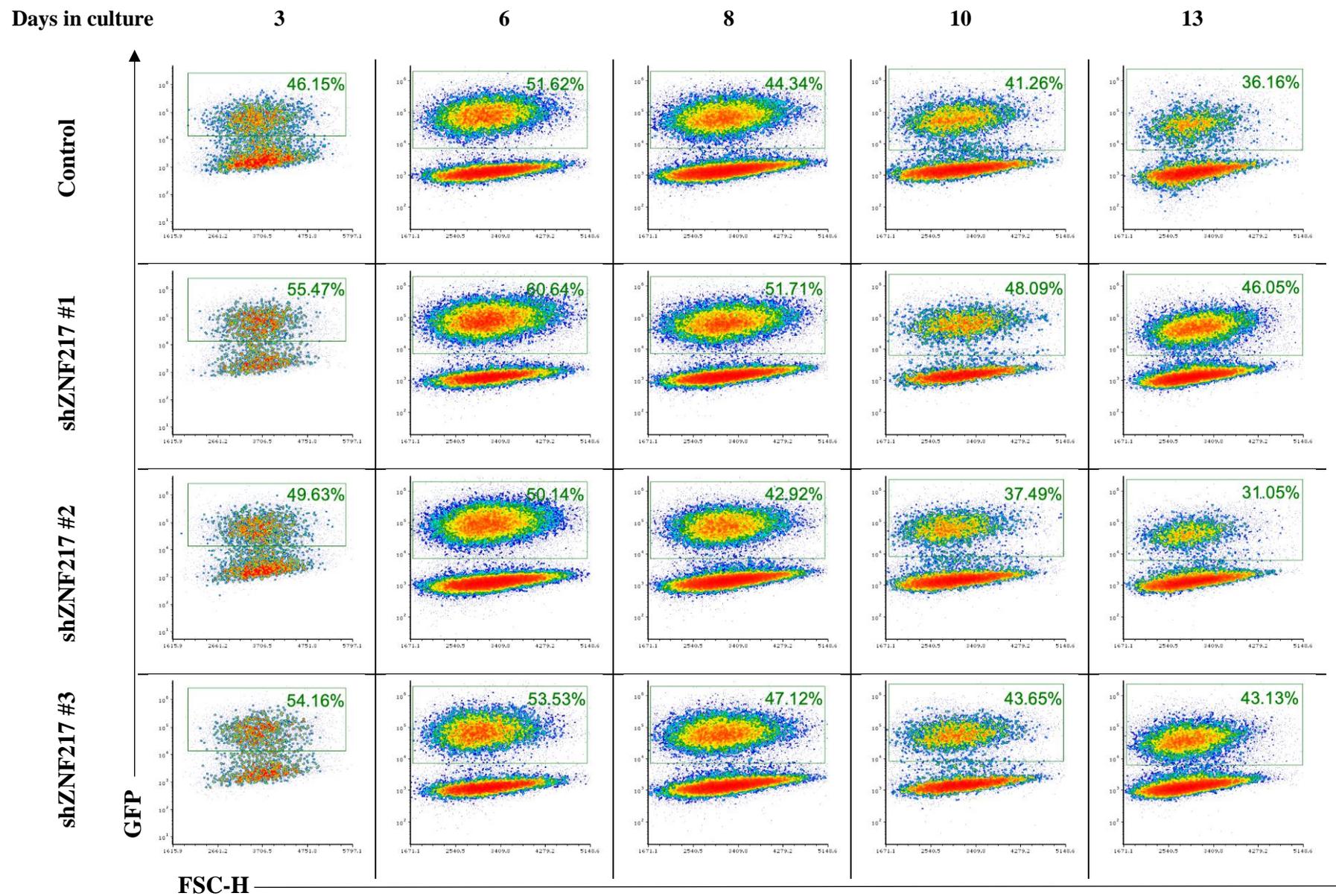
**Supplementary Figure 7 – Expression of the myeloid marker CD11b in granulocytes during differentiation, following ZNF217 overexpression**

Representative contour plots showing the expression of CD11b in granulocytes (CD13<sup>+/-</sup> CD36<sup>-</sup>) transduced with a control or a ZNF217 overexpression vector throughout 13 days of culture, as measured by flow cytometry. An initial gating to exclude cell debris was applied, followed by the selection of GFP<sup>+</sup> cells, as described in **Supplementary Figure 1**. Transduced haematopoietic populations were subsequently delimited through the myeloid markers CD13 and CD36. Basal expression was determined using the corresponding IgG control.



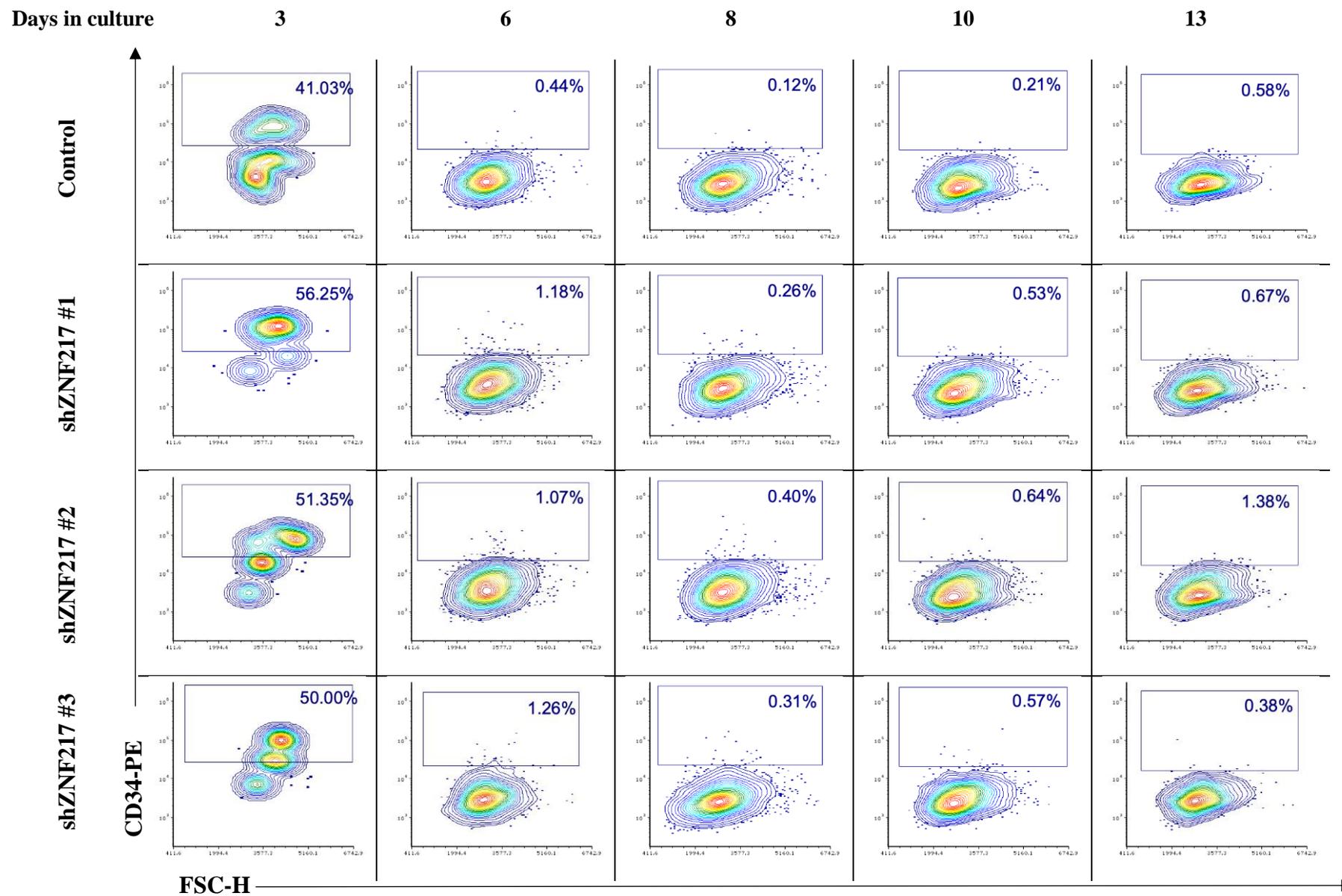
**Supplementary Figure 8 – Expression of the granulocytic marker CD15 in granulocytes during differentiation, following ZNF217 overexpression**

Representative contour plots showing the expression of CD15 in granulocytes ( $CD13^{+/-} CD36^{-}$ ) transduced with a control or a ZNF217 overexpression vector throughout 13 days of culture, as measured by flow cytometry. An initial gating to exclude cell debris was applied, followed by the selection of  $GFP^{+}$  cells, as described in **Supplementary Figure 1**. Transduced haematopoietic populations were subsequently delimited through the myeloid markers CD13 and CD36. Basal expression was determined using the corresponding IgG control.



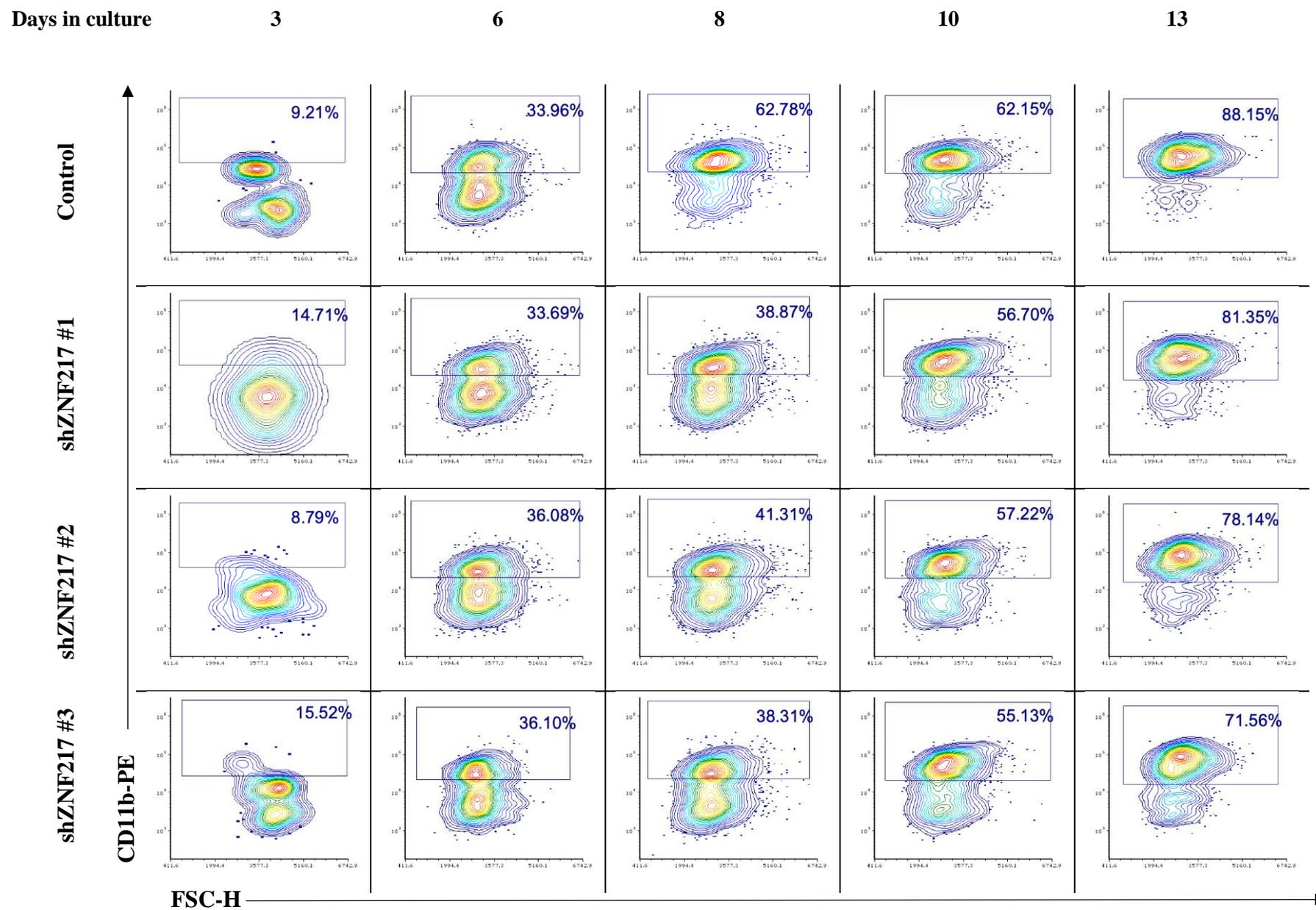
**Supplementary Figure 9 – Expression of control GFP and shZNF217 KD in human HSPC**

Representative density plots showing the expression of GFP in control and shZNF217 KD cultures, throughout 13 days of culture, as measured by flow cytometry. An initial gating to exclude cell debris was applied, as described in **Supplementary Figure 1A**. Background autofluorescence was established using cells subjected to the equivalent viral infection procedure but in the absence of virus (mock culture).



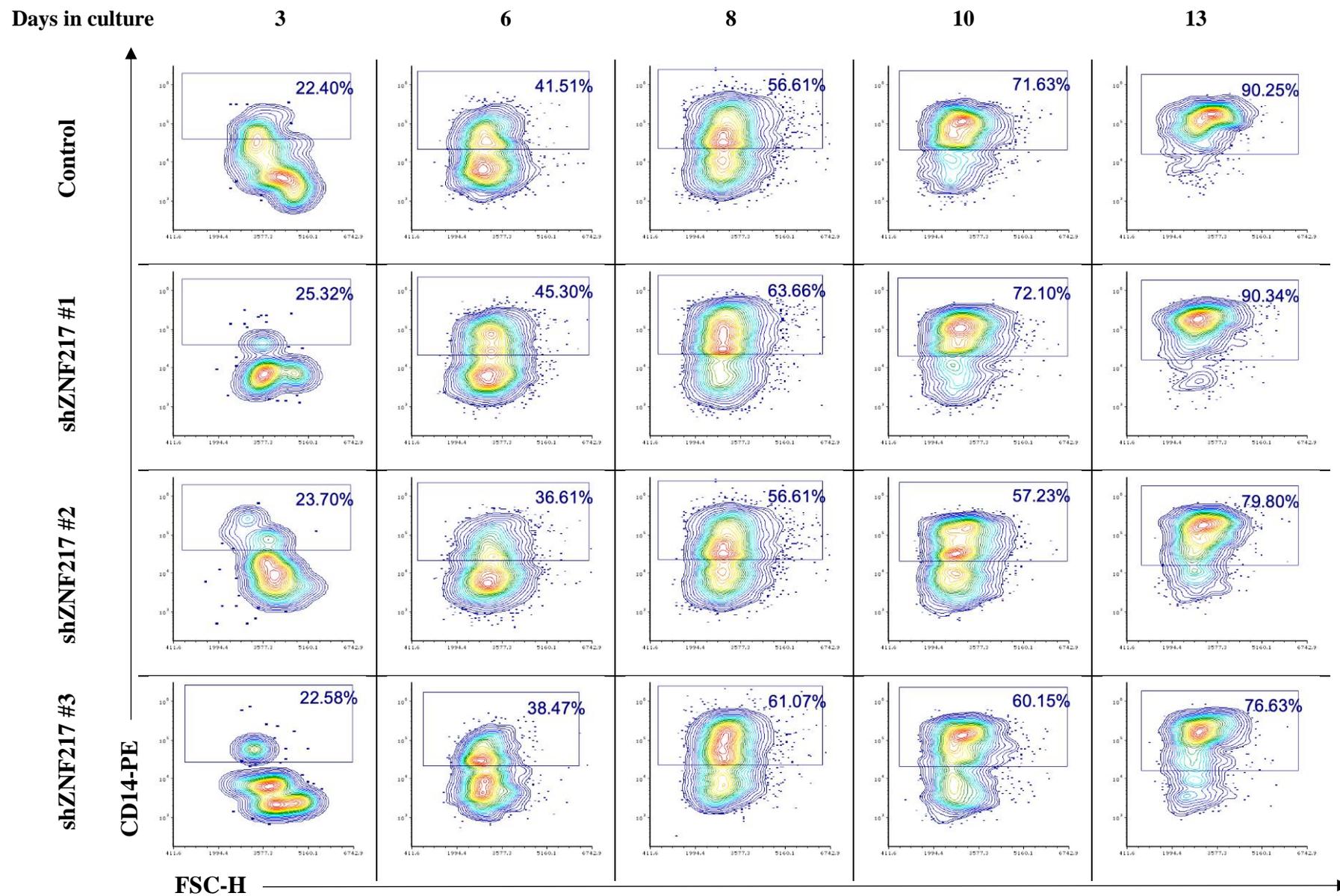
**Supplementary Figure 10 – Expression of the stem cell marker CD34 in monocytes during differentiation, following ZNF217 KD**

Representative contour plots showing the expression of CD34 in monocytes (CD13<sup>+</sup> CD36<sup>+</sup>) transduced with a control or a shZNF217 KD construct throughout 13 days of culture, as measured by flow cytometry. An initial gating to exclude cell debris was applied, followed by the selection of GFP<sup>+</sup> cells, as described in **Supplementary Figure 1**. Transduced haematopoietic populations were subsequently delimited through the myeloid markers CD13 and CD36. Basal expression was determined using the corresponding IgG control.



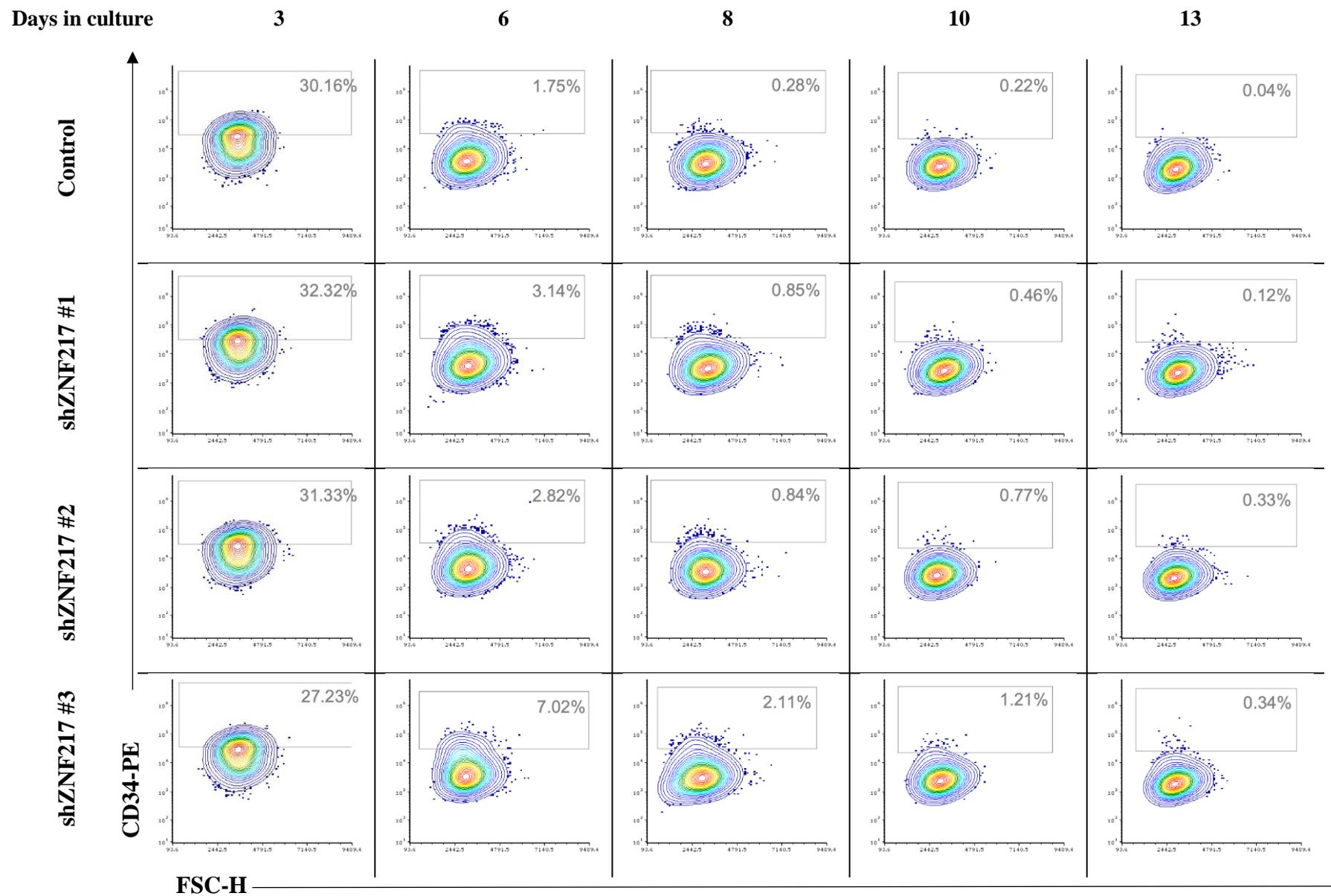
**Supplementary Figure 11 – Expression of the myeloid marker CD11b in monocytes during differentiation, following ZNF217 KD**

Representative contour plots showing the expression of CD11b in monocytes (CD13<sup>+</sup> CD36<sup>+</sup>) transduced with a control or a shZNF217 KD construct throughout 13 days of culture, as measured by flow cytometry. An initial gating to exclude cell debris was applied, followed by the selection of GFP<sup>+</sup> cells, as described in **Supplementary Figure 1**. Transduced haematopoietic populations were subsequently delimited through the myeloid markers CD13 and CD36. Basal expression was determined using the corresponding IgG control.



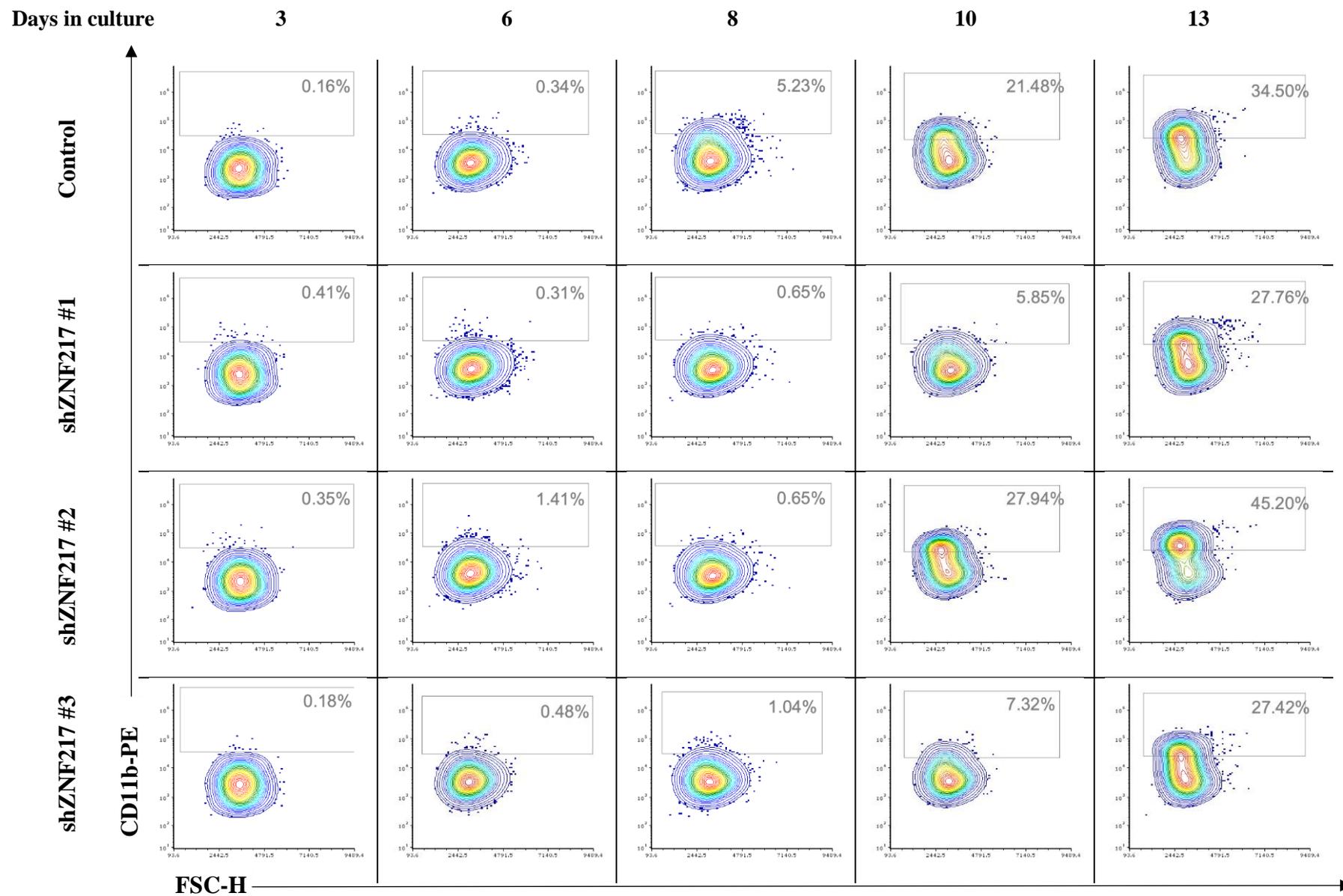
**Supplementary Figure 12 – Expression of the monocytic marker CD14 in monocytes during differentiation, following ZNF217 KD**

Representative contour plots showing the expression of CD14 in monocytes (CD13<sup>+</sup> CD36<sup>+</sup>) transduced with a control or a shZNF217 KD construct throughout 13 days of culture, as measured by flow cytometry. An initial gating to exclude cell debris was applied, followed by the selection of GFP<sup>+</sup> cells, as described in **Supplementary Figure 1**. Transduced haematopoietic populations were subsequently delimited through the myeloid markers CD13 and CD36. Basal expression was determined using the corresponding IgG control.



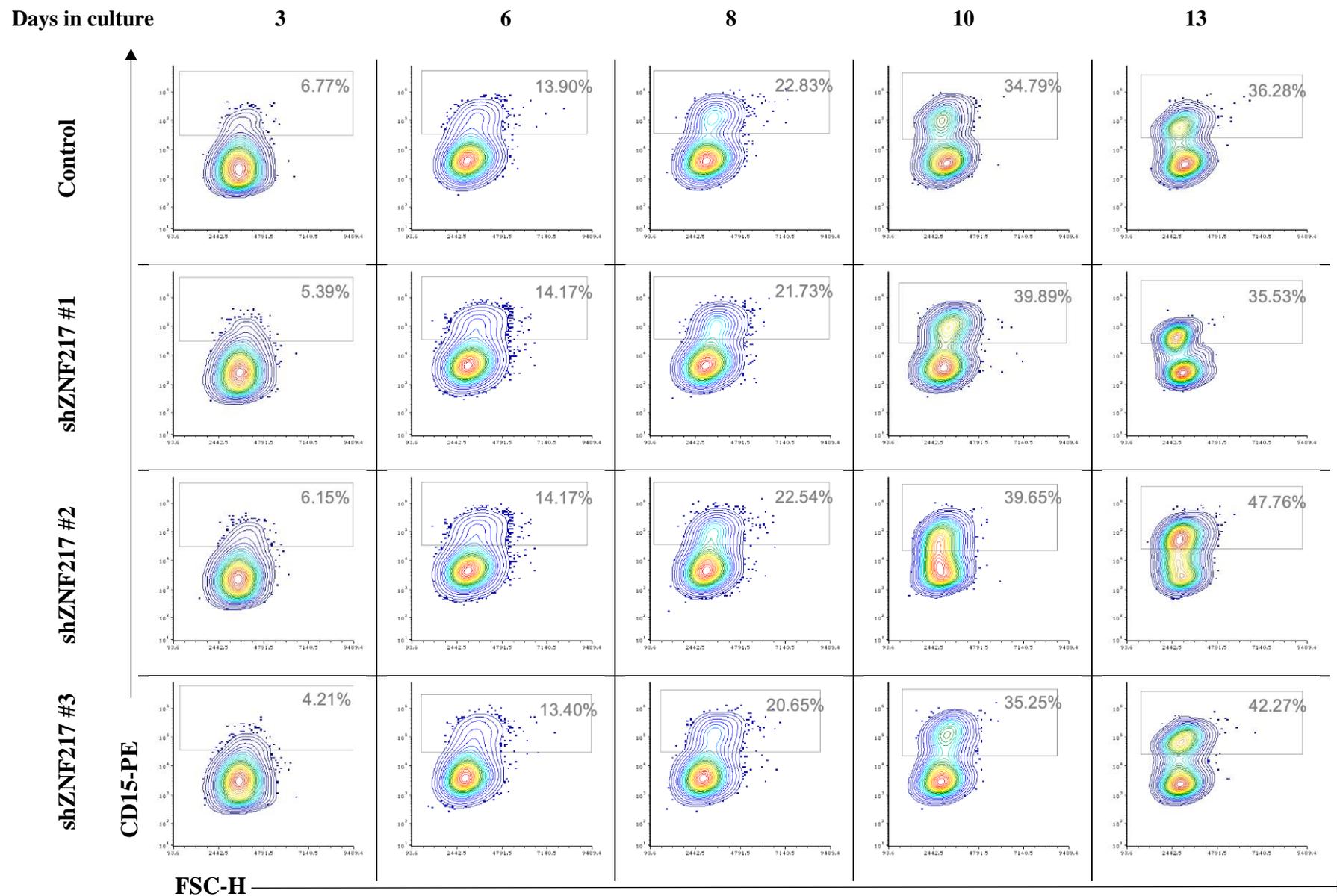
**Supplementary Figure 13 – Expression of the stem cell marker CD34 in granulocytes during differentiation, following ZNF217 KD**

Representative contour plots showing the expression of CD34 in granulocytes (CD13<sup>+/−</sup> CD36<sup>−</sup>) transduced with a control or a shZNF217 KD construct throughout 13 days of culture, as measured by flow cytometry. An initial gating to exclude cell debris was applied, followed by the selection of GFP<sup>+</sup> cells, as described in **Supplementary Figure 1**. Transduced haematopoietic populations were subsequently delimited through the myeloid markers CD13 and CD36. Basal expression was determined using the corresponding IgG control.



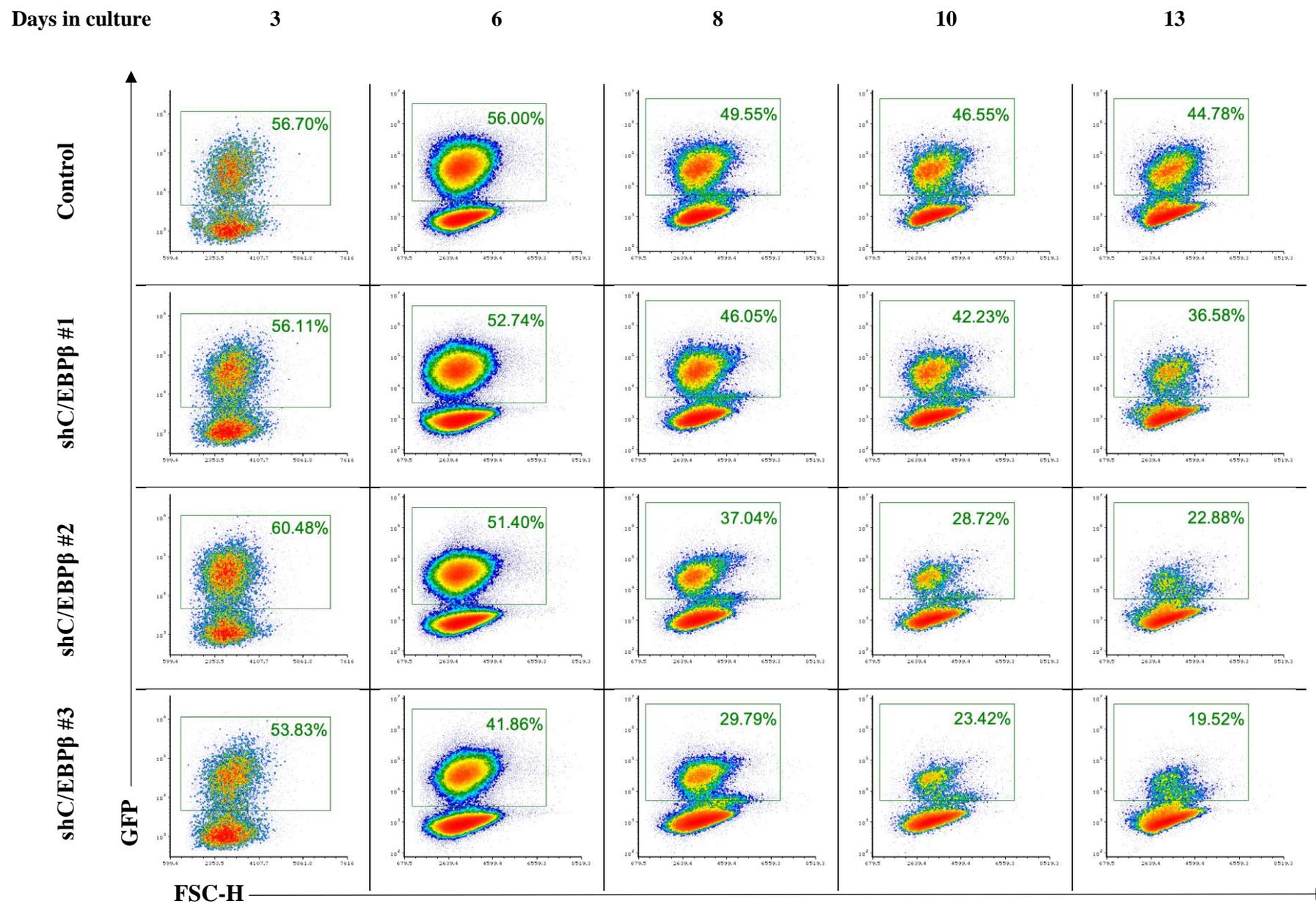
**Supplementary Figure 14 – Expression of the myeloid marker CD11b in granulocytes during differentiation, following ZNF217 KD**

Representative contour plots showing the expression of CD11b in granulocytes (CD13<sup>+/+</sup> CD36<sup>-</sup>) transduced with a control or a shZNF217 KD construct throughout 13 days of culture, as measured by flow cytometry. An initial gating to exclude cell debris was applied, followed by the selection of GFP<sup>+</sup> cells, as described in **Supplementary Figure 1**. Transduced haematopoietic populations were subsequently delimited through the myeloid markers CD13 and CD36. Basal expression was determined using the corresponding IgG control.



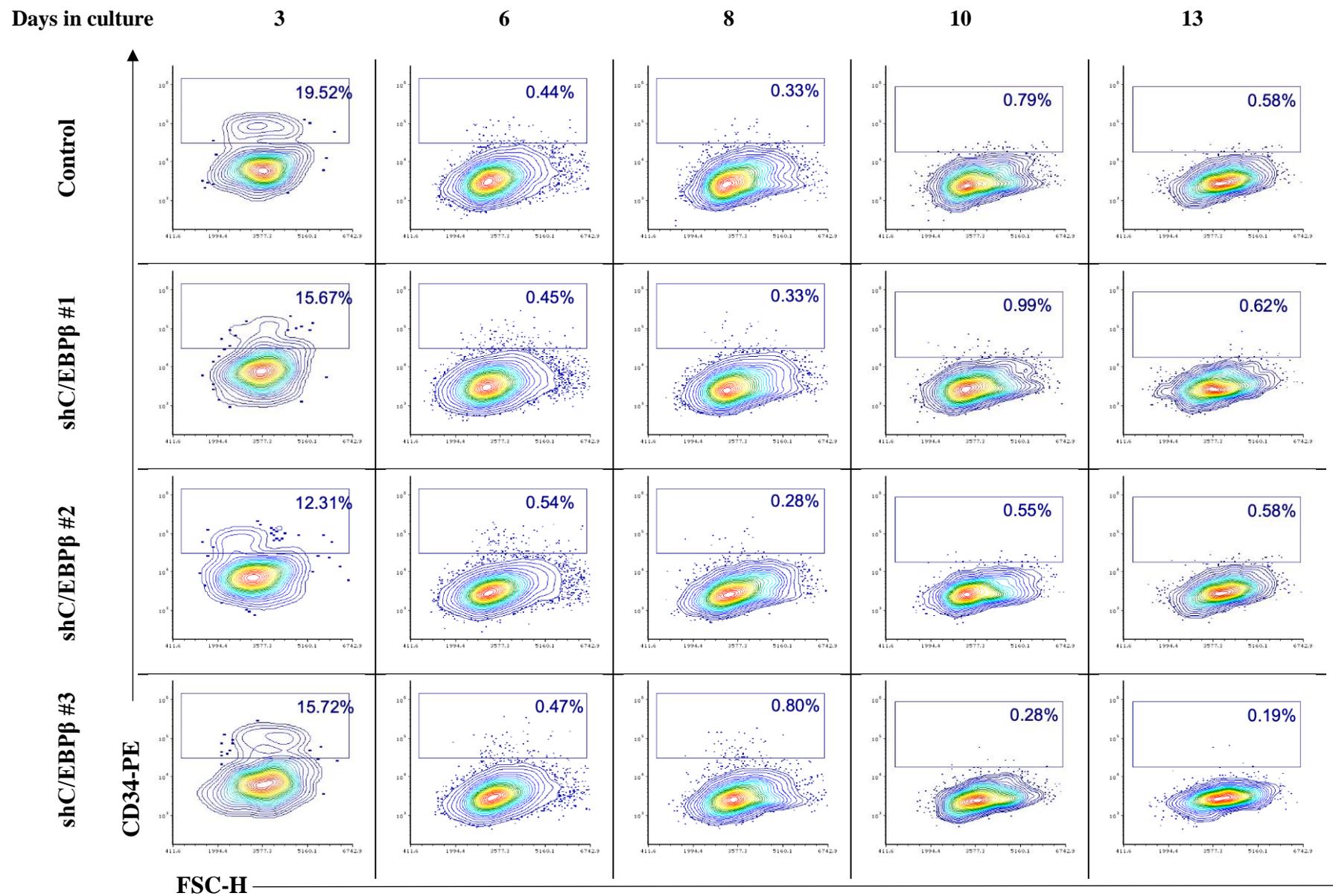
**Supplementary Figure 15 – Expression of the granulocytic marker CD15 in granulocytes during differentiation, following ZNF217 KD**

Representative contour plots showing the expression of CD15 in granulocytes (CD13<sup>+/−</sup> CD36<sup>−</sup>) transduced with a control or a shZNF217 KD construct throughout 13 days of culture, as measured by flow cytometry. An initial gating to exclude cell debris was applied, followed by the selection of GFP<sup>+</sup> cells, as described in **Supplementary Figure 1**. Transduced haematopoietic populations were subsequently delimited through the myeloid markers CD13 and CD36. Basal expression was determined using the corresponding IgG control.



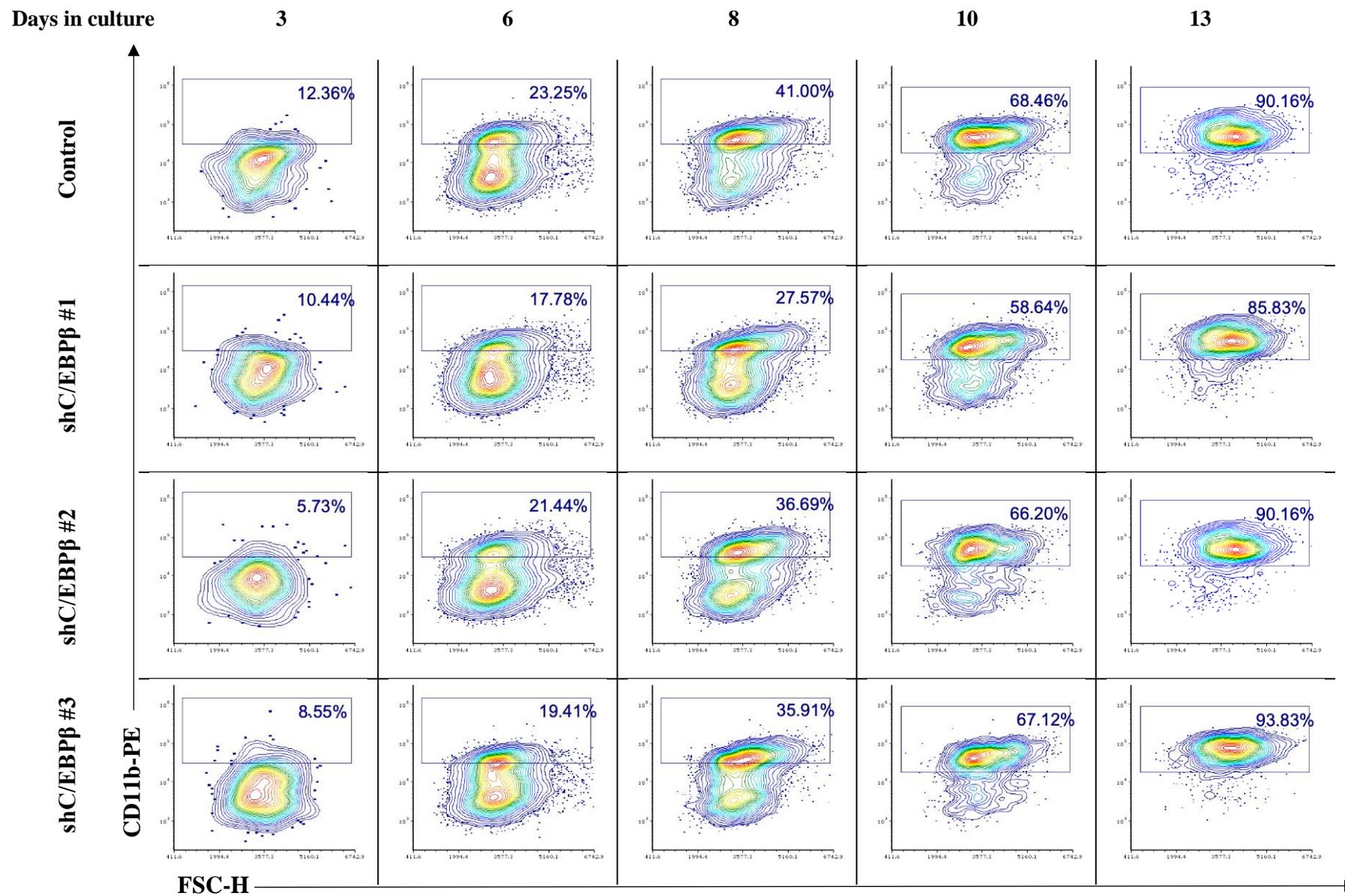
**Supplementary Figure 16 – Expression of control GFP and shC/EBP $\beta$  KD in human HSPC**

Representative density plots showing the expression of GFP in control and shC/EBP $\beta$  KD cultures, throughout 13 days of culture, as measured by flow cytometry. An initial gating to exclude cell debris was applied, as described in **Supplementary Figure 1A**. Background autofluorescence was established using cells subjected to the equivalent viral infection procedure but in the absence of virus (mock culture).



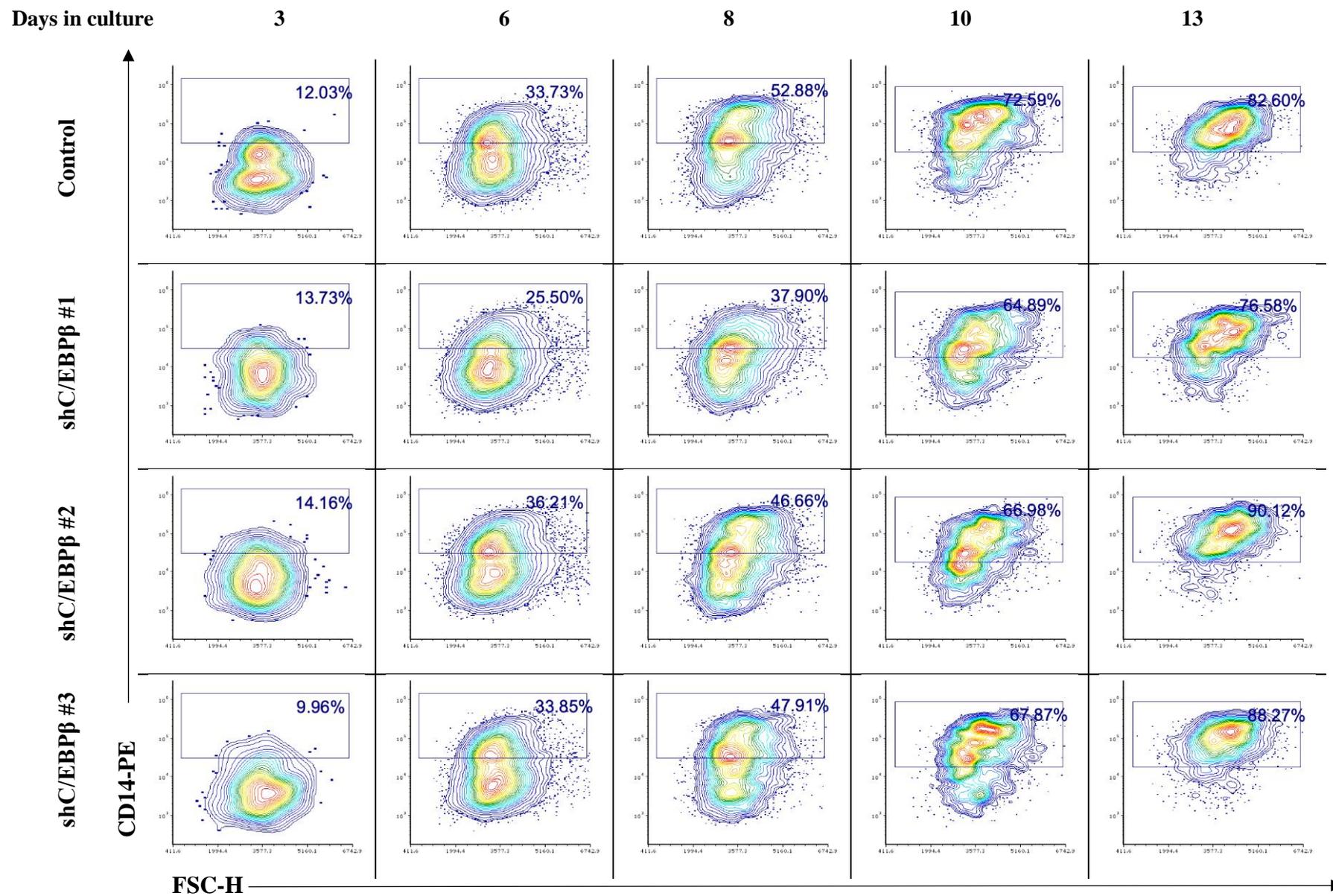
**Supplementary Figure 17 – Expression of the stem cell marker CD34 in monocytes during differentiation, following C/EBP $\beta$  KD**

Representative contour plots showing the expression of CD34 in monocytes (CD13<sup>+</sup> CD36<sup>+</sup>) transduced with a control or a shC/EBP $\beta$  KD construct throughout 13 days of culture, as measured by flow cytometry. An initial gating to exclude cell debris was applied, followed by the selection of GFP<sup>+</sup> cells, as described in **Supplementary Figure 1**. Transduced haematopoietic populations were subsequently delimited through the myeloid markers CD13 and CD36. Basal expression was determined using the corresponding IgG control.



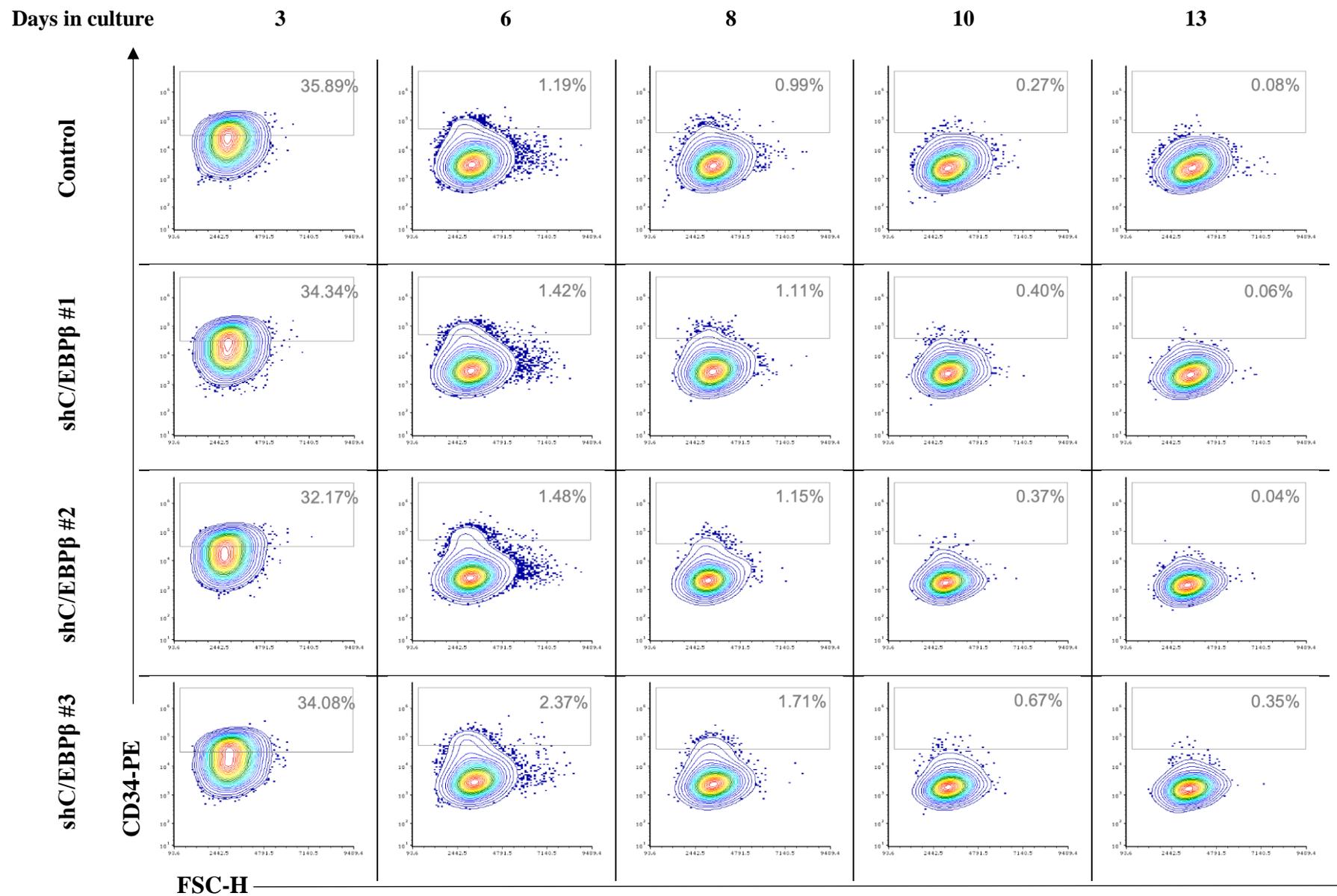
**Supplementary Figure 18 – Expression of the myeloid marker CD11b in monocytes during differentiation, following C/EBP $\beta$  KD**

Representative contour plots showing the expression of CD11b in monocytes (CD13<sup>+</sup> CD36<sup>+</sup>) transduced with a control or a shC/EBP $\beta$  KD construct throughout 13 days of culture, as measured by flow cytometry. An initial gating to exclude cell debris was applied, followed by the selection of GFP<sup>+</sup> cells, as described in **Supplementary Figure 1**. Transduced haematopoietic populations were subsequently delimited through the myeloid markers CD13 and CD36. Basal expression was determined using the corresponding IgG control.



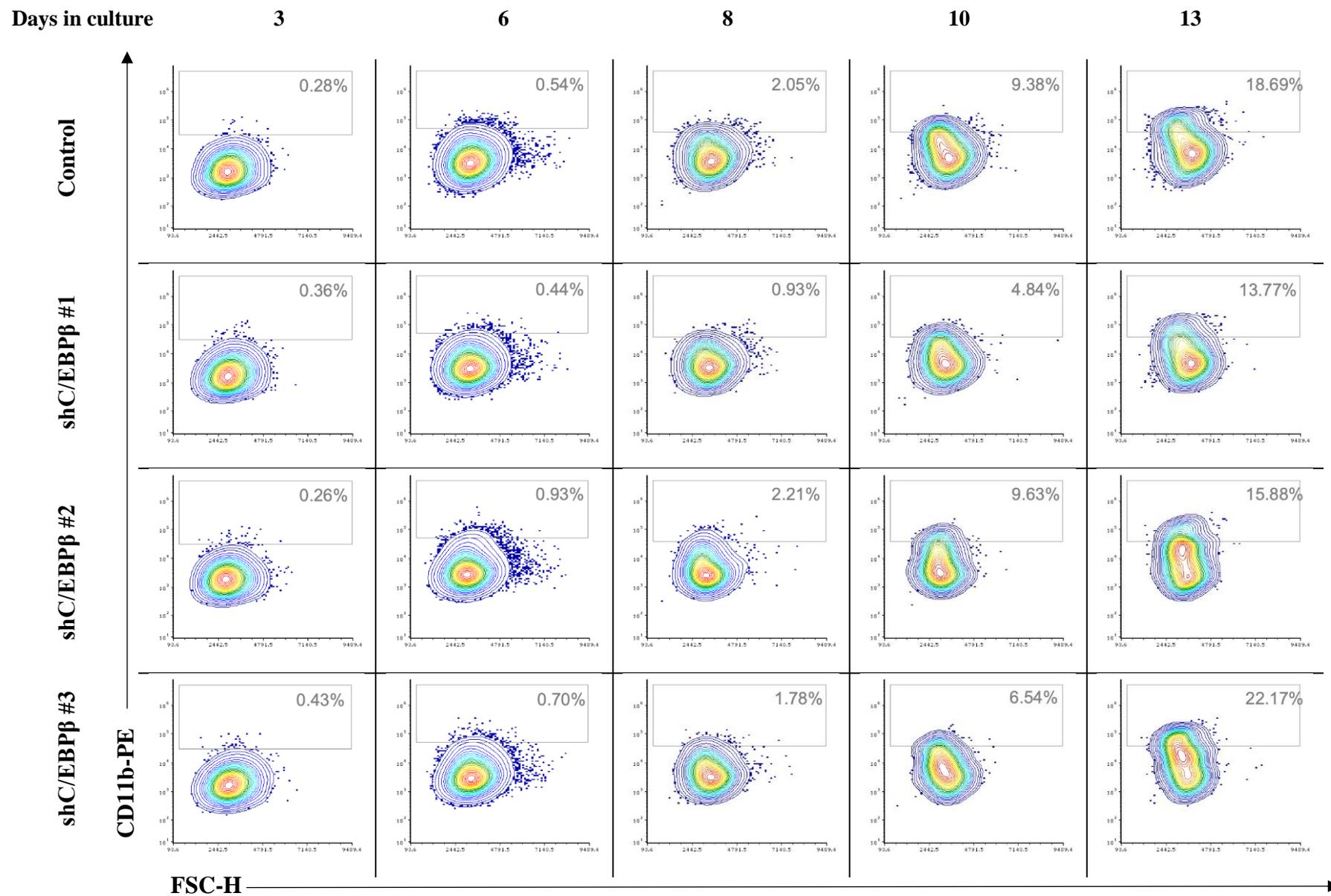
**Supplementary Figure 19 – Expression of the monocytic marker CD14 in monocytes during differentiation, following C/EBP $\beta$  KD**

Representative contour plots showing the expression of CD14 in monocytes (CD13<sup>+</sup> CD36<sup>+</sup>) transduced with a control or a shC/EBP $\beta$  KD construct throughout 13 days of culture, as measured by flow cytometry. An initial gating to exclude cell debris was applied, followed by the selection of GFP<sup>+</sup> cells, as described in **Supplementary Figure 1**. Transduced haematopoietic populations were subsequently delimited through the myeloid markers CD13 and CD36. Basal expression was determined using the corresponding IgG control.



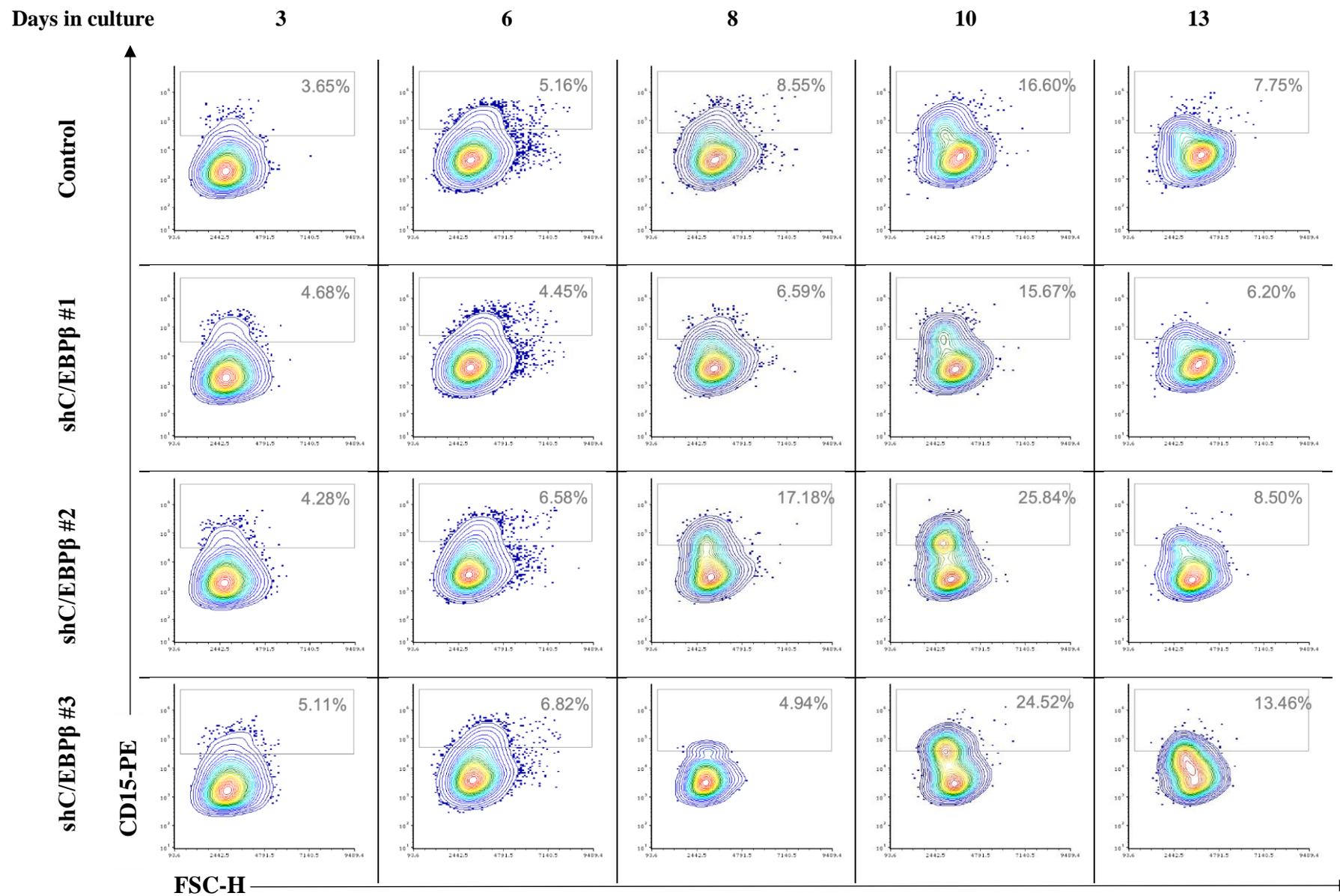
**Supplementary Figure 20 – Expression of the stem cell marker CD34 in granulocytes during differentiation, following C/EBP $\beta$  KD**

Representative contour plots showing the expression of CD34 in granulocytes (CD13<sup>+/−</sup> CD36<sup>−</sup>) transduced with a control or a shC/EBP $\beta$  KD construct throughout 13 days of culture, as measured by flow cytometry. An initial gating to exclude cell debris was applied, followed by the selection of GFP<sup>+</sup> cells, as described in **Supplementary Figure 1**. Transduced haematopoietic populations were subsequently delimited through the myeloid markers CD13 and CD36. Basal expression was determined using the corresponding IgG control.



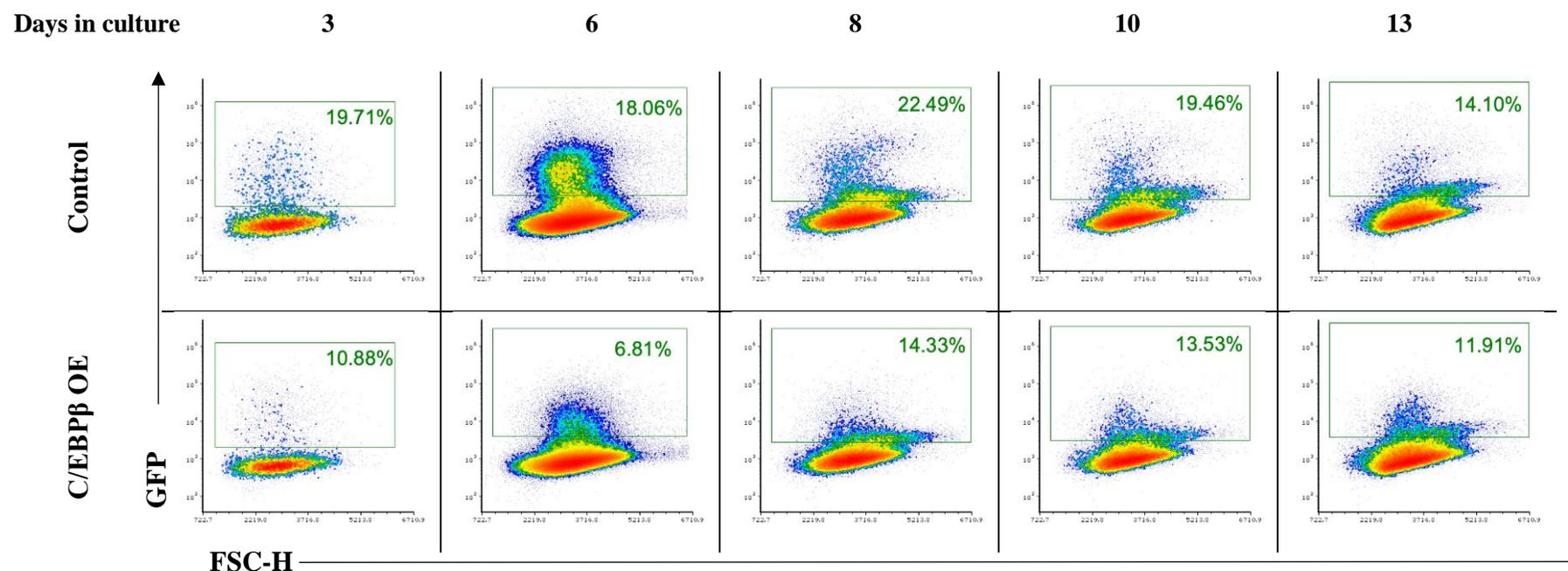
**Supplementary Figure 21 – Expression of the myeloid marker CD11b in granulocytes during differentiation, following C/EBP $\beta$  KD**

Representative contour plots showing the expression of CD11b in granulocytes (CD13<sup>+/+</sup> CD36<sup>-</sup>) transduced with a control or a shC/EBP $\beta$  KD construct throughout 13 days of culture, as measured by flow cytometry. An initial gating to exclude cell debris was applied, followed by the selection of GFP<sup>+</sup> cells, as described in **Supplementary Figure 1**. Transduced haematopoietic populations were subsequently delimited through the myeloid markers CD13 and CD36. Basal expression was determined using the corresponding IgG control.



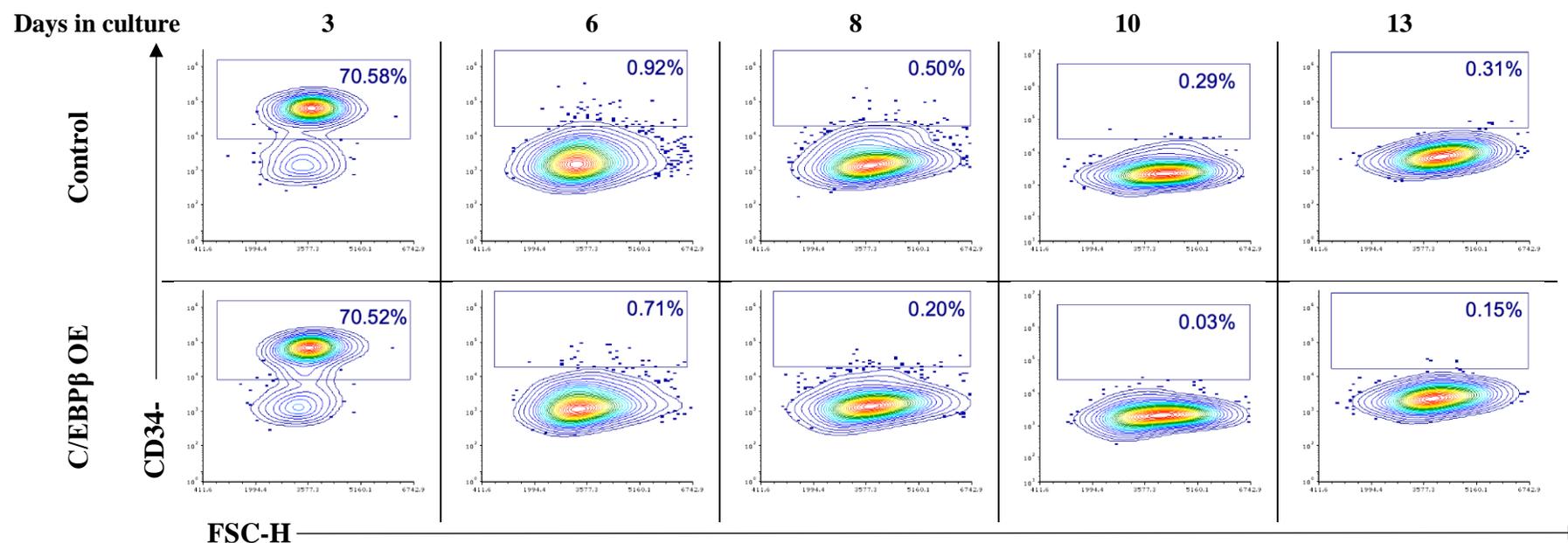
**Supplementary Figure 22 – Expression of the granulocytic marker CD15 in granulocytes during differentiation, following C/EBP $\beta$  KD**

Representative contour plots showing the expression of CD15 in granulocytes (CD13<sup>+</sup> CD36<sup>-</sup>) transduced with a control or a shC/EBP $\beta$  KD construct throughout 13 days of culture, as measured by flow cytometry. An initial gating to exclude cell debris was applied, followed by the selection of GFP<sup>+</sup> cells, as described in **Supplementary Figure 1**. Transduced haematopoietic populations were subsequently delimited through the myeloid markers CD13 and CD36. Basal expression was determined using the corresponding IgG control.



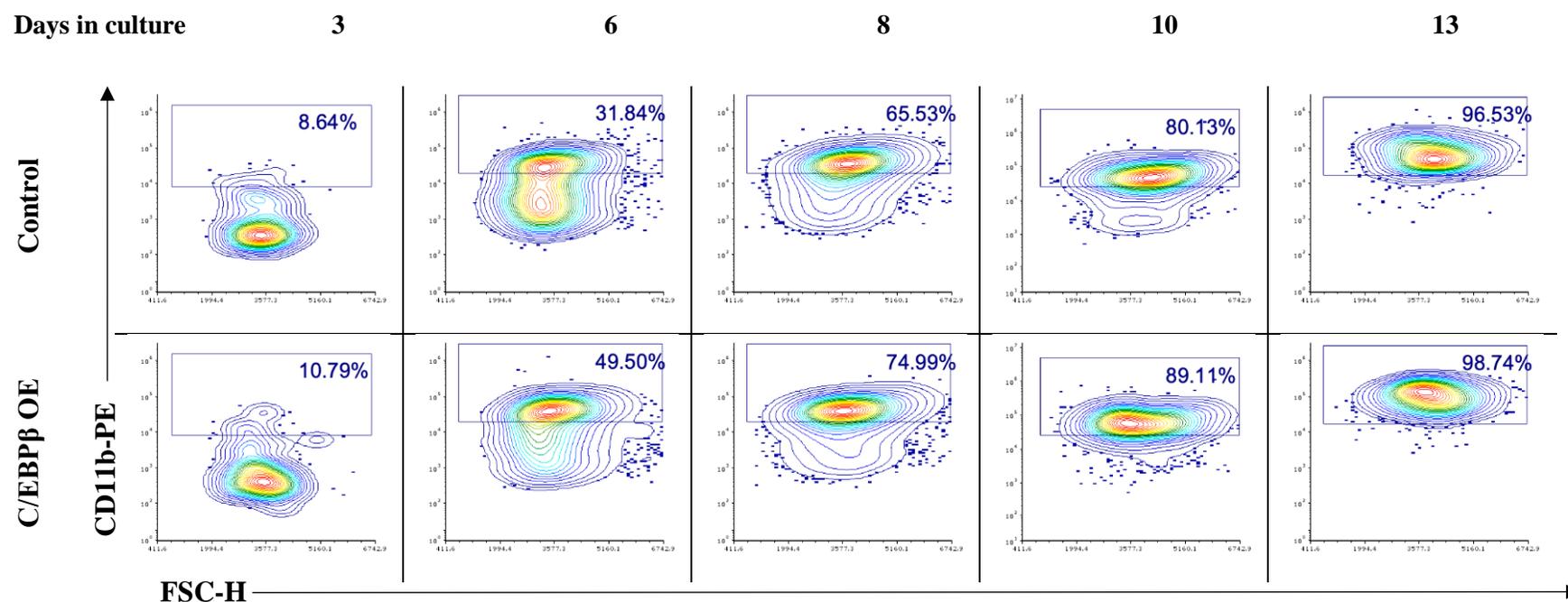
**Supplementary Figure 23 – Expression of control GFP and C/EBP $\beta$  in human HSPC**

Representative density plots showing the expression of GFP in control and ZNF217 overexpressing cultures, throughout 13 days of culture, as measured by flow cytometry. An initial gating to exclude cell debris was C/EBP $\beta$ , as described in **Supplementary Figure 1A**. Background autofluorescence was established using cells subjected to the equivalent viral infection procedure but in the absence of virus (mock culture).



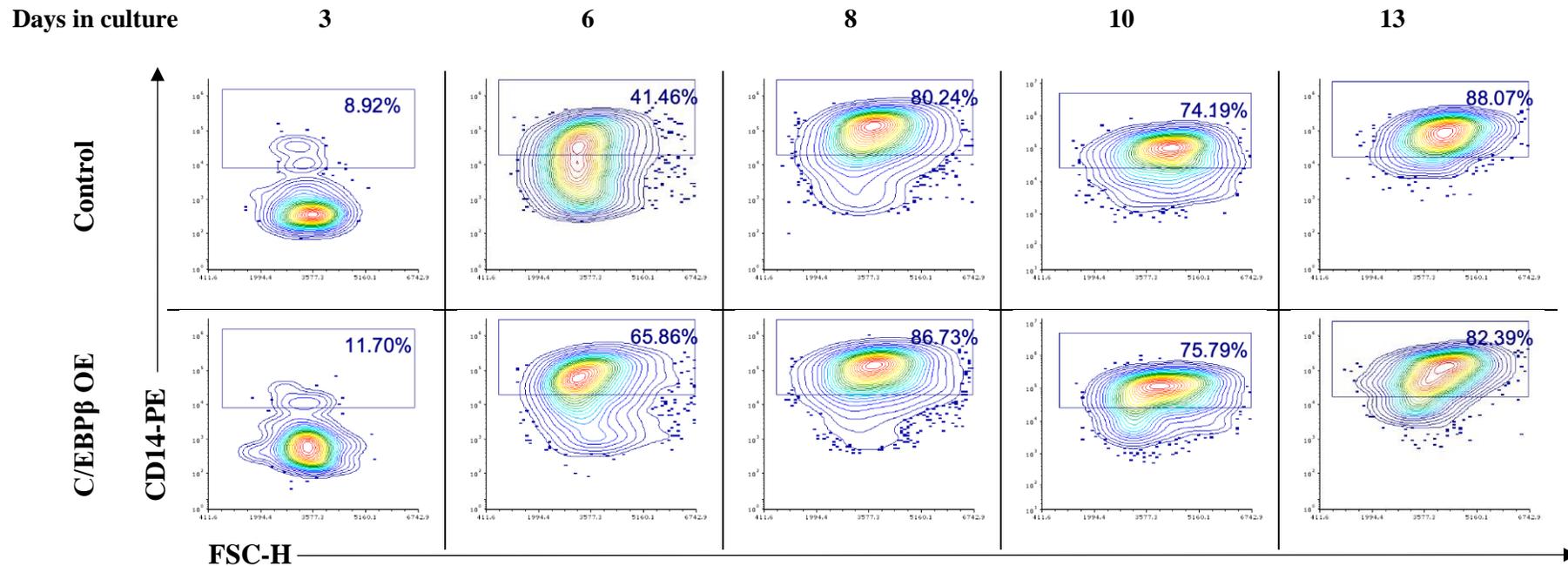
**Supplementary Figure 24 – Expression of the stem cell marker CD34 in monocytes during differentiation, following C/EBPβ overexpression**

Representative contour plots showing the expression of CD34 in monocytes (CD13<sup>+</sup> CD36<sup>+</sup>) transduced with a control or a C/EBPβ overexpression vector throughout 13 days of culture, as measured by flow cytometry. An initial gating to exclude cell debris was applied, followed by the selection of GFP<sup>+</sup> cells, as described in **Supplementary Figure 1**. Transduced haematopoietic populations were subsequently delimited through the myeloid markers CD13 and CD36. Basal expression was determined using the corresponding IgG control.



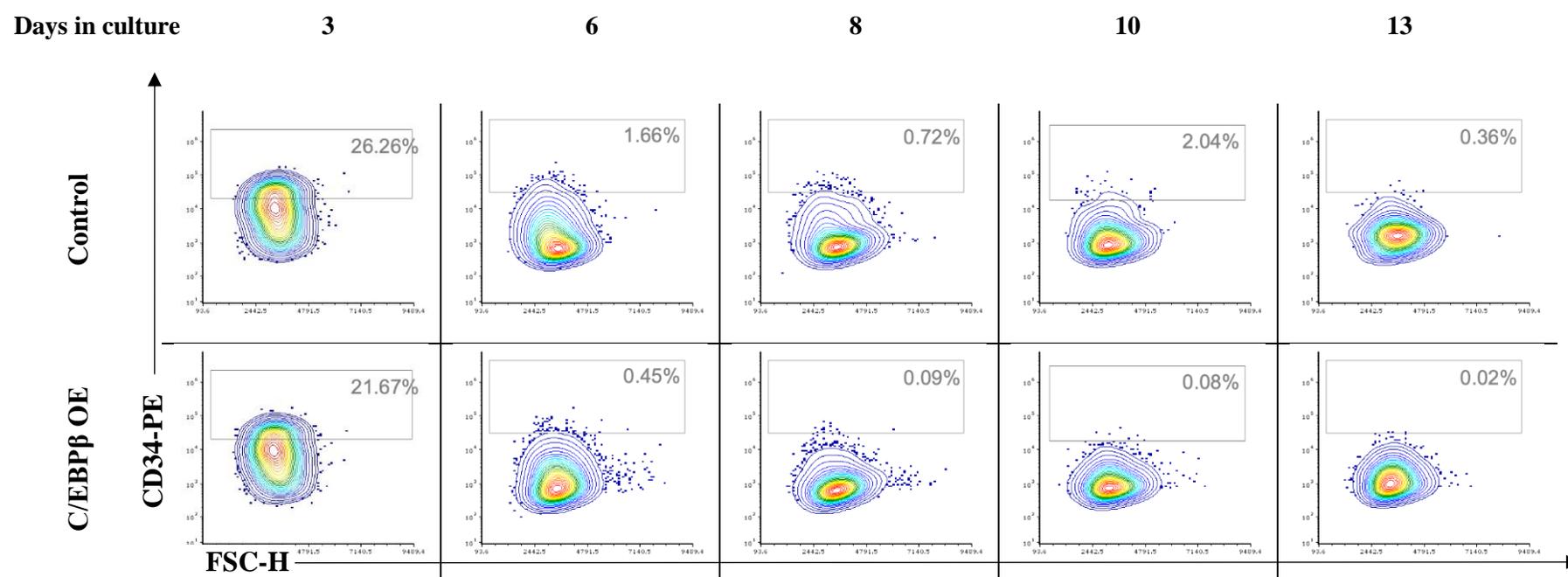
**Supplementary Figure 25 – Expression of the myeloid marker CD11b in monocytes during differentiation, following C/EBPβ overexpression**

Representative contour plots showing the expression of CD11b in monocytes (CD13<sup>+</sup> CD36<sup>+</sup>) transduced with a control or a C/EBPβ overexpression vector throughout 13 days of culture, as measured by flow cytometry. An initial gating to exclude cell debris was applied, followed by the selection of GFP<sup>+</sup> cells, as described in **Supplementary Figure 1**. Transduced haematopoietic populations were subsequently delimited through the myeloid markers CD13 and CD36. Basal expression was determined using the corresponding IgG control.



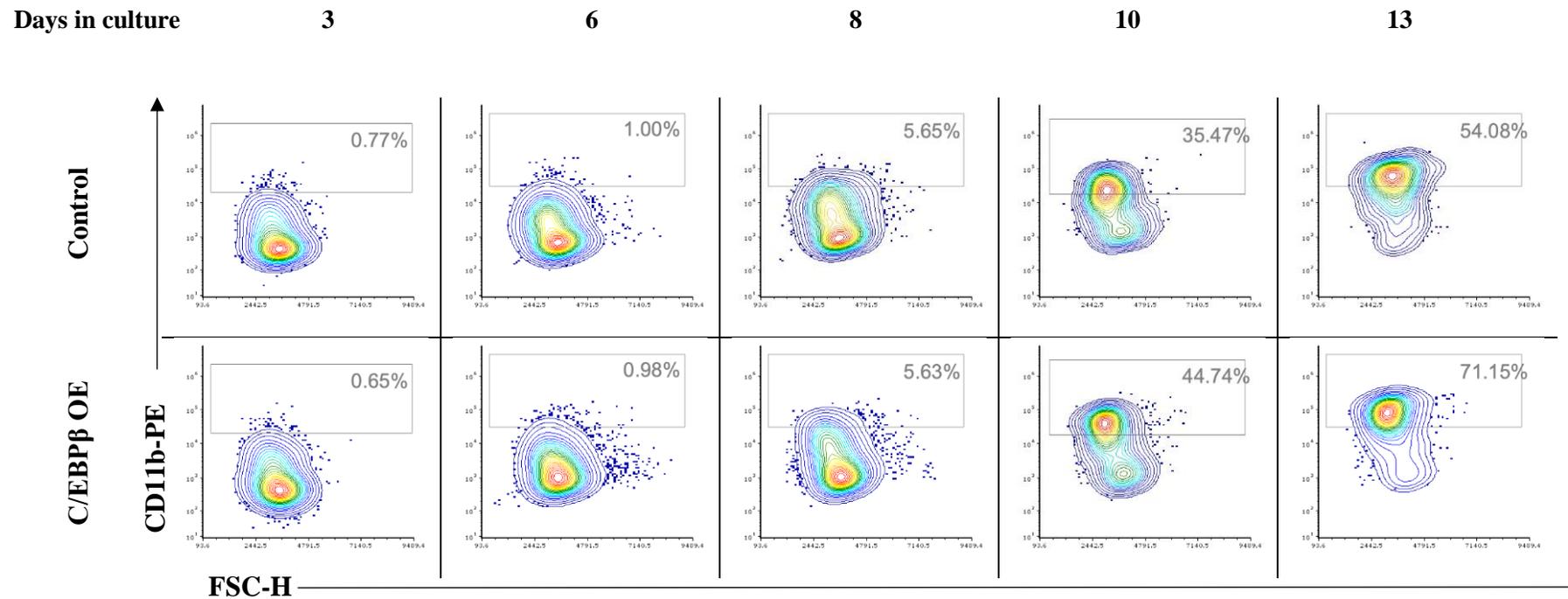
**Supplementary Figure 26 – Expression of the myeloid marker CD14 in monocytes during differentiation, following C/EBPβ overexpression**

Representative contour plots showing the expression of CD14 in monocytes (CD13<sup>+</sup> CD36<sup>+</sup>) transduced with a control or a C/EBPβ overexpression vector throughout 13 days of culture, as measured by flow cytometry. An initial gating to exclude cell debris was applied, followed by the selection of GFP<sup>+</sup> cells, as described in **Supplementary Figure 1**. Transduced haematopoietic populations were subsequently delimited through the myeloid markers CD13 and CD36. Basal expression was determined using the corresponding IgG control.



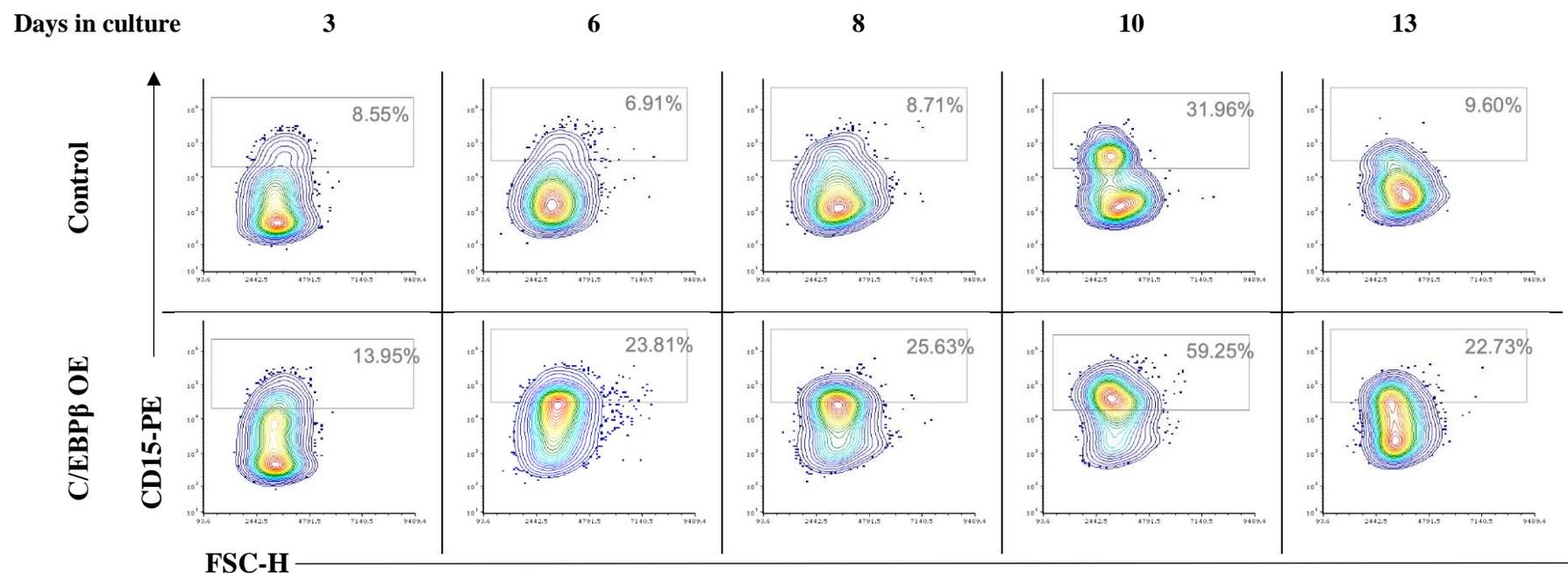
**Supplementary Figure 27 – Expression of the stem cell marker CD34 in granulocytes during differentiation, following C/EBPβ overexpression**

Representative contour plots showing the expression of CD34 in granulocytes (CD13<sup>+</sup> CD36<sup>-</sup>) transduced with a control or a C/EBPβ overexpression vector throughout 13 days of culture, as measured by flow cytometry. An initial gating to exclude cell debris was applied, followed by the selection of GFP<sup>+</sup> cells, as described in **Supplementary Figure 1**. Transduced haematopoietic populations were subsequently delimited through the myeloid markers CD13 and CD36. Basal expression was determined using the corresponding IgG control.



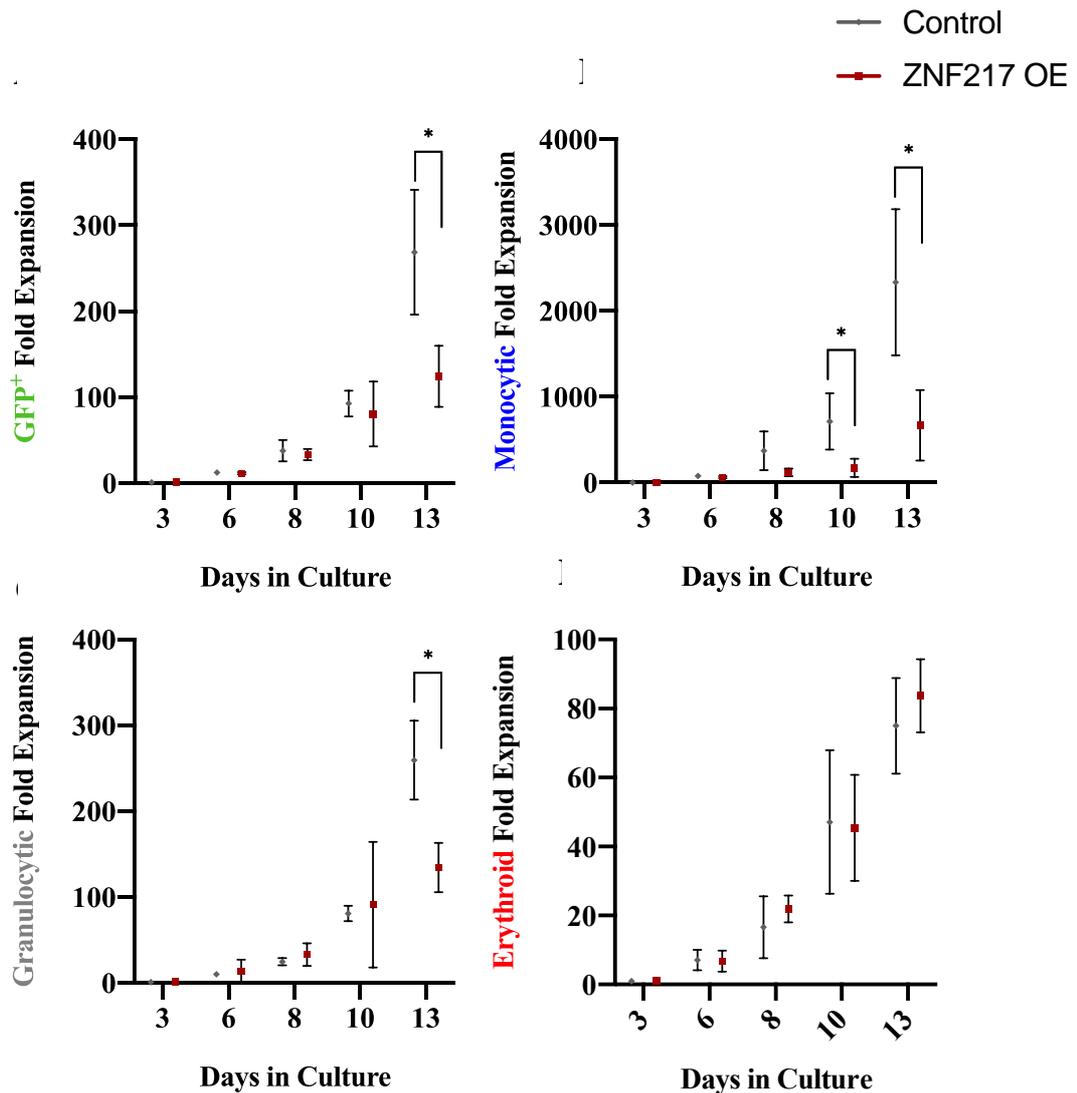
**Supplementary Figure 28 – Expression of the myeloid marker CD11b in granulocytes during differentiation, following C/EBPβ overexpression**

Representative contour plots showing the expression of CD11b in granulocytes (CD13<sup>+</sup> CD36<sup>-</sup>) transduced with a control or a C/EBPβ overexpression vector throughout 13 days of culture, as measured by flow cytometry. An initial gating to exclude cell debris was applied, followed by the selection of GFP<sup>+</sup> cells, as described in **Supplementary Figure 1**. Transduced haematopoietic populations were subsequently delimited through the myeloid markers CD13 and CD36. Basal expression was determined using the corresponding IgG control.



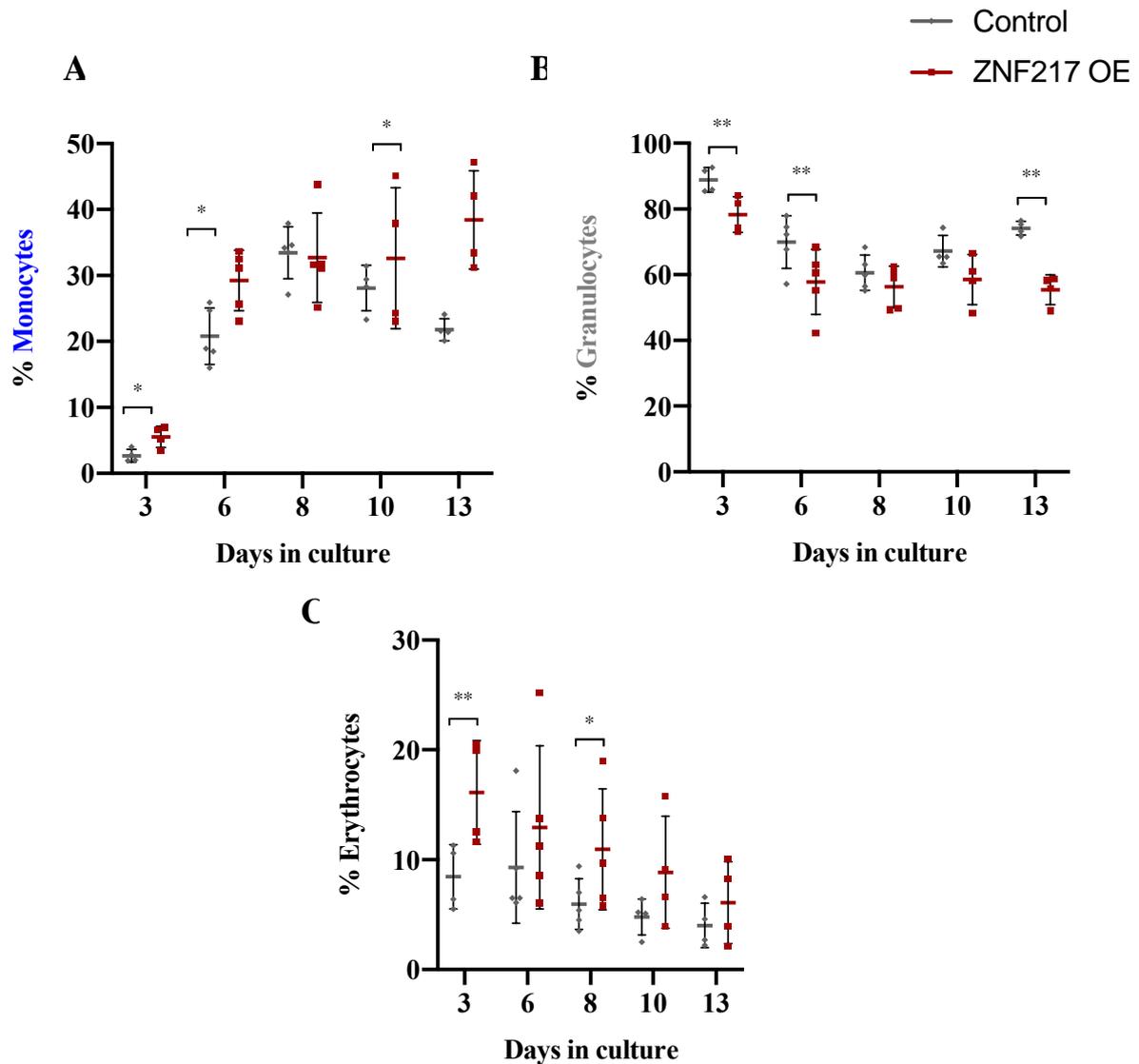
**Supplementary Figure 29 – Expression of the granulocytic marker CD15 in granulocytes during differentiation, following C/EBPβ overexpression**

Representative contour plots showing the expression of CD15 in granulocytes (CD13<sup>+/-</sup> CD36<sup>-</sup>) transduced with a control or a C/EBPβ overexpression vector throughout 13 days of culture, as measured by flow cytometry. An initial gating to exclude cell debris was applied, followed by the selection of GFP<sup>+</sup> cells, as described in **Supplementary Figure 1**. Transduced haematopoietic populations were subsequently delimited through the myeloid markers CD13 and CD36. Basal expression was determined using the corresponding IgG control.



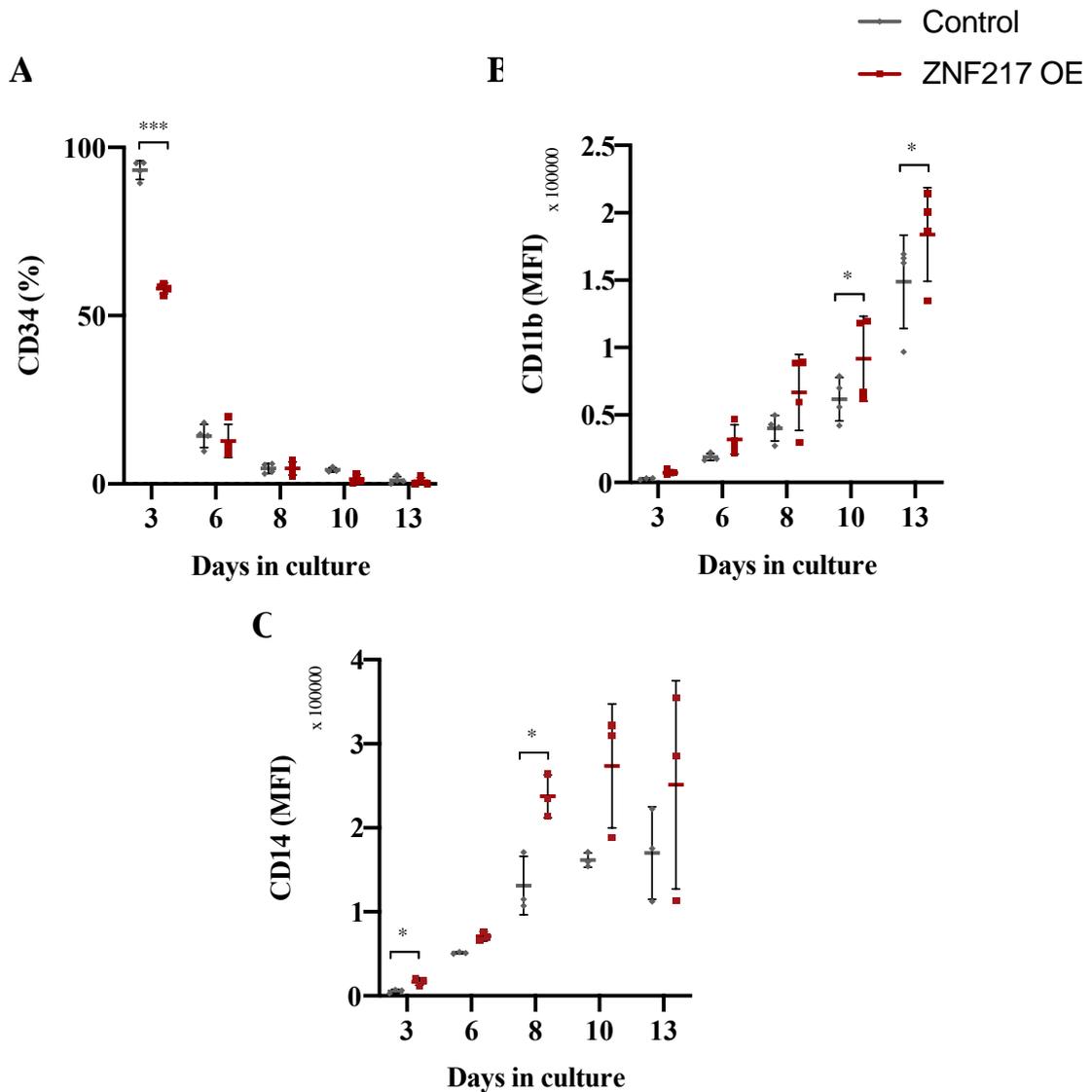
### Supplementary Figure 30 – ZNF217 overexpression inhibits monocytic and granulocytic growth during myeloid development

Summary dot plots showing: **(A)** Cumulative fold-expansion of control and ZNF217-overexpressing HSP GFP<sup>+</sup> cells grown over 13 days in culture medium containing IL-3, SCF, G- and GM-CSF. **(B)** Cumulative fold-expansion of monocytic cells (CD13<sup>+</sup> CD36<sup>+</sup>) transduced with a control of ZNF217-overexpressing vector. **(C)** Cumulative fold-expansion of granulocytic cells (CD13<sup>+/−</sup> CD36<sup>−</sup>) transduced with a control of ZNF217-overexpressing vector. **(D)** Cumulative fold-expansion of erythroid cells (CD13<sup>−</sup> CD36<sup>+</sup>) transduced with a control of ZNF217-overexpressing vector. Cells were re-seeded at the appropriate density following each time-point, as described in 2.3.3. Gating strategies were applied as described in Supplementary Figure 1. Data indicates Mean ± 1SD (n≥3) Significant differences between ZNF217-expressing cells and control at each time-point were analysed by paired t-test; \* denotes p<0.05.



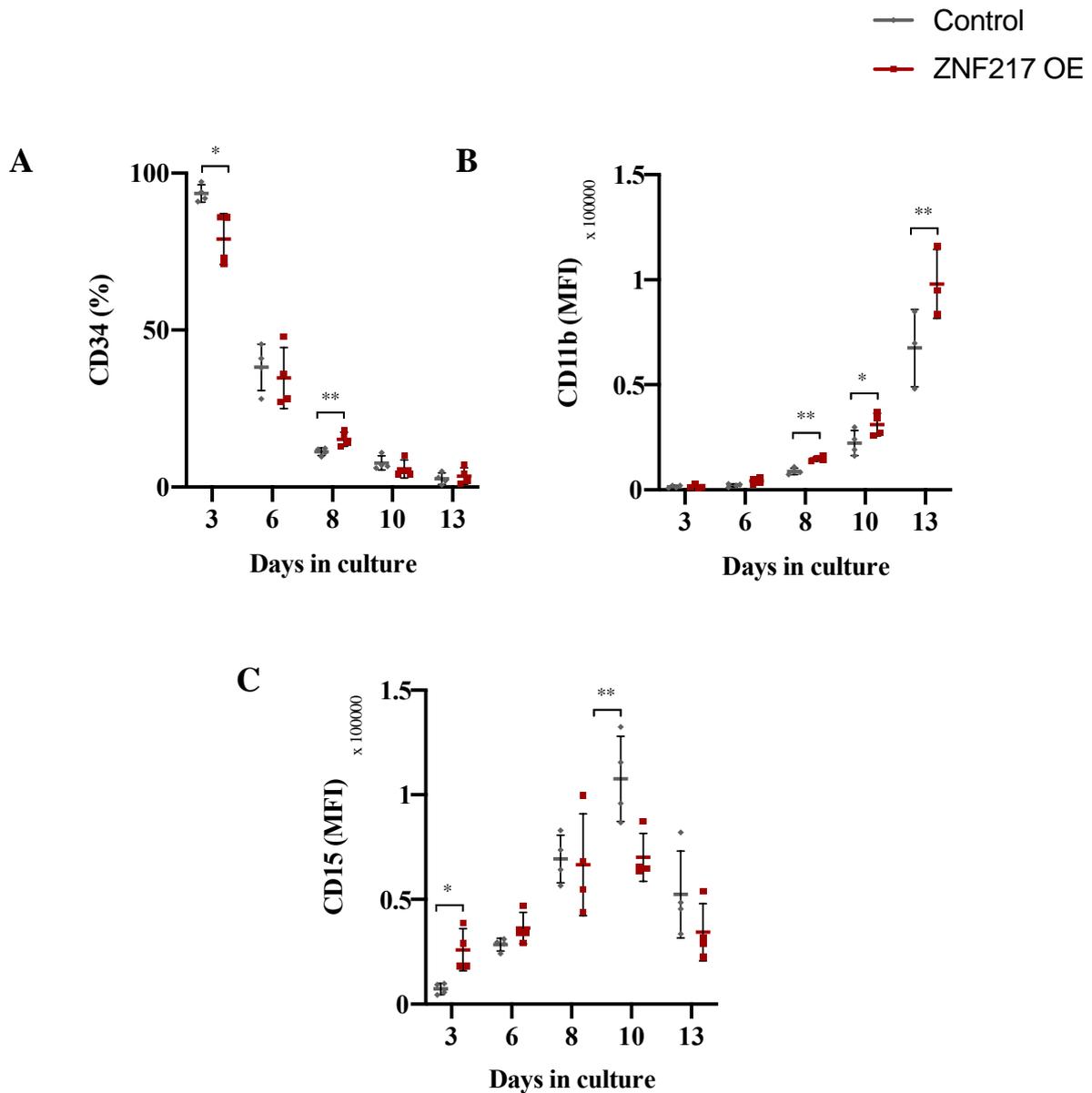
### Supplementary Figure 31– ZNF217 overexpression disrupts the balance between granulocytic and erythroid populations during myeloid cell development

Summary dot plots showing the proportion of (A) monocytic (CD13<sup>+</sup> CD36<sup>+</sup>), (B) granulocytic (CD13<sup>+/−</sup> CD36<sup>−</sup>) and (C) erythroid cells (CD13<sup>−</sup> CD36<sup>+</sup>) in control and ZNF217-overexpressing cultures. Cells were re-seeded at the appropriate density following each time-point, as described in 2.3.3. Gating strategies were applied as described in Supplementary Figure 1. Data indicates Mean ± 1SD (n ≥ 4). Significant differences between ZNF217-expressing cells and control at each time-point were analysed by paired t-test; \* denotes p < 0.05.



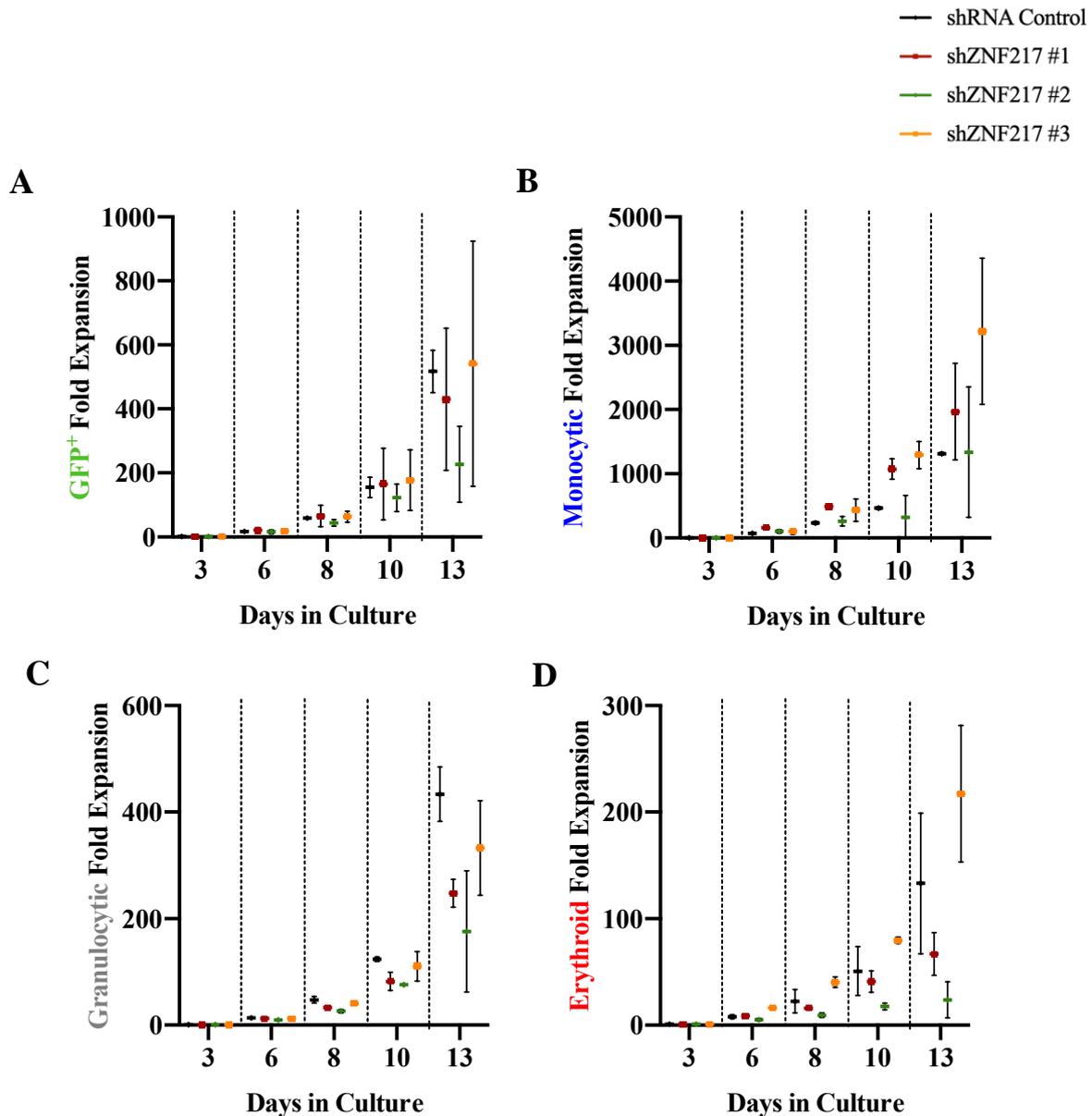
### Supplementary Figure 32 – ZNF217 overexpression promotes monocytic differentiation

Summary dot plots showing (A) CD34<sup>+</sup> expression in terms of percentage in monocytic cells (CD13<sup>+</sup> CD36<sup>+</sup>) over time for control and ZNF217-overexpressing cells. Summary data showing (B) CD11b and (C) CD14 expression in terms of MFI in monocytic cells over time for control and ZNF217-overexpressing cells. Cells were re-seeded at the appropriate density following each time-point, as described in 2.3.3. Gating strategies were applied as described in **Supplementary Figure 1**. Data indicates Mean  $\pm$  1SD (n $\geq$ 3). Significant differences between ZNF217-expressing cells and control at each time-point were analysed by paired t-test; \* denotes p<0.05, \*\* denotes p<0.01, \*\*\* denotes p<0.001.



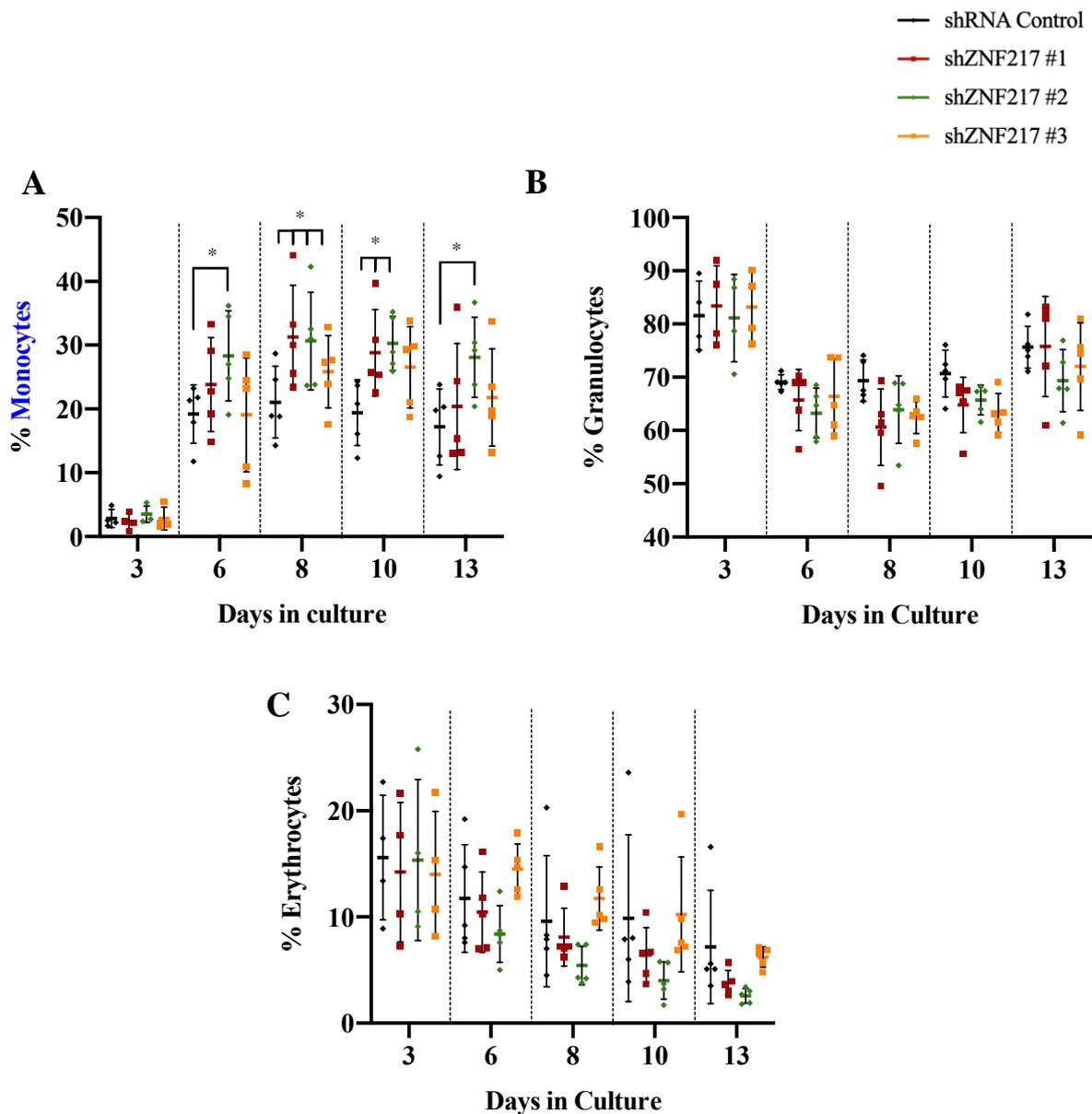
**Supplementary Figure 33 – ZNF217 overexpression promotes granulocytic differentiation by upregulating monocytic markers**

Summary dot plots showing (A) CD34<sup>+</sup> expression in granulocytic cells (CD13<sup>+/−</sup> CD36<sup>−</sup>). Summary data showing (B) CD11b and (C) CD15 expression in terms of MFI in granulocytic cells over time for control and ZNF217-overexpressing cells. Cells were re-seeded at the appropriate density following each time-point, as described in 2.3.3. Gating strategies were applied as described in **Supplementary Figure 1**. Data indicates Mean  $\pm$  1SD (n $\geq$ 3). Significant differences between ZNF217-expressing cells and control at each time-point were analysed by paired t-test; \* denotes p<0.05, \*\* denotes p<0.01.



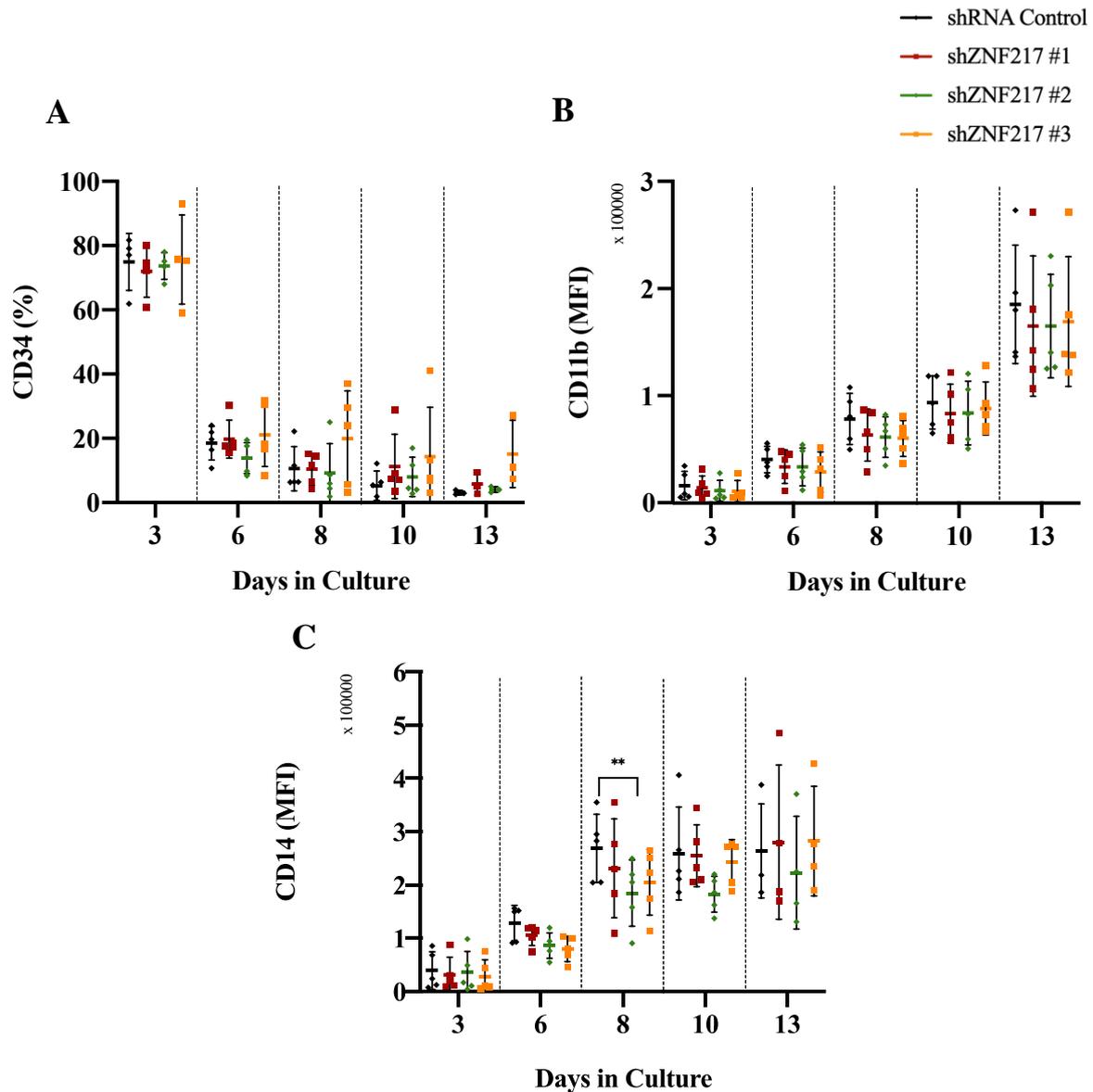
### Supplementary Figure 34 – Knockdown of ZNF217 promotes monocytic cell growth in myeloid development

Summary dot plot showing the (A) cumulative fold-expansion of control and three ZNF217-KD constructs in terms of GFP positivity in culture medium containing IL-3, SCF, G- and GM-CSF, grown over 13 days. (B) Cumulative fold-expansion of monocytic cells (CD13<sup>+</sup> CD36<sup>+</sup>) transduced with a control or ZNF217-KD vectors. (C) Cumulative fold-expansion of granulocytic cells (CD13<sup>+</sup> CD36<sup>-</sup>) transduced with a control or ZNF217-KD vector. (D) Cumulative fold-expansion of erythroid cells (CD13<sup>-</sup> CD36<sup>+</sup>) transduced with a control or ZNF217-KD vector. Cells were re-seeded at the appropriate density following each time-point, as described in 2.3.3. Gating strategies were applied as described in Supplementary Figure 1. Data indicates Mean  $\pm$  1SD (n=2) (no statistical test was employed).



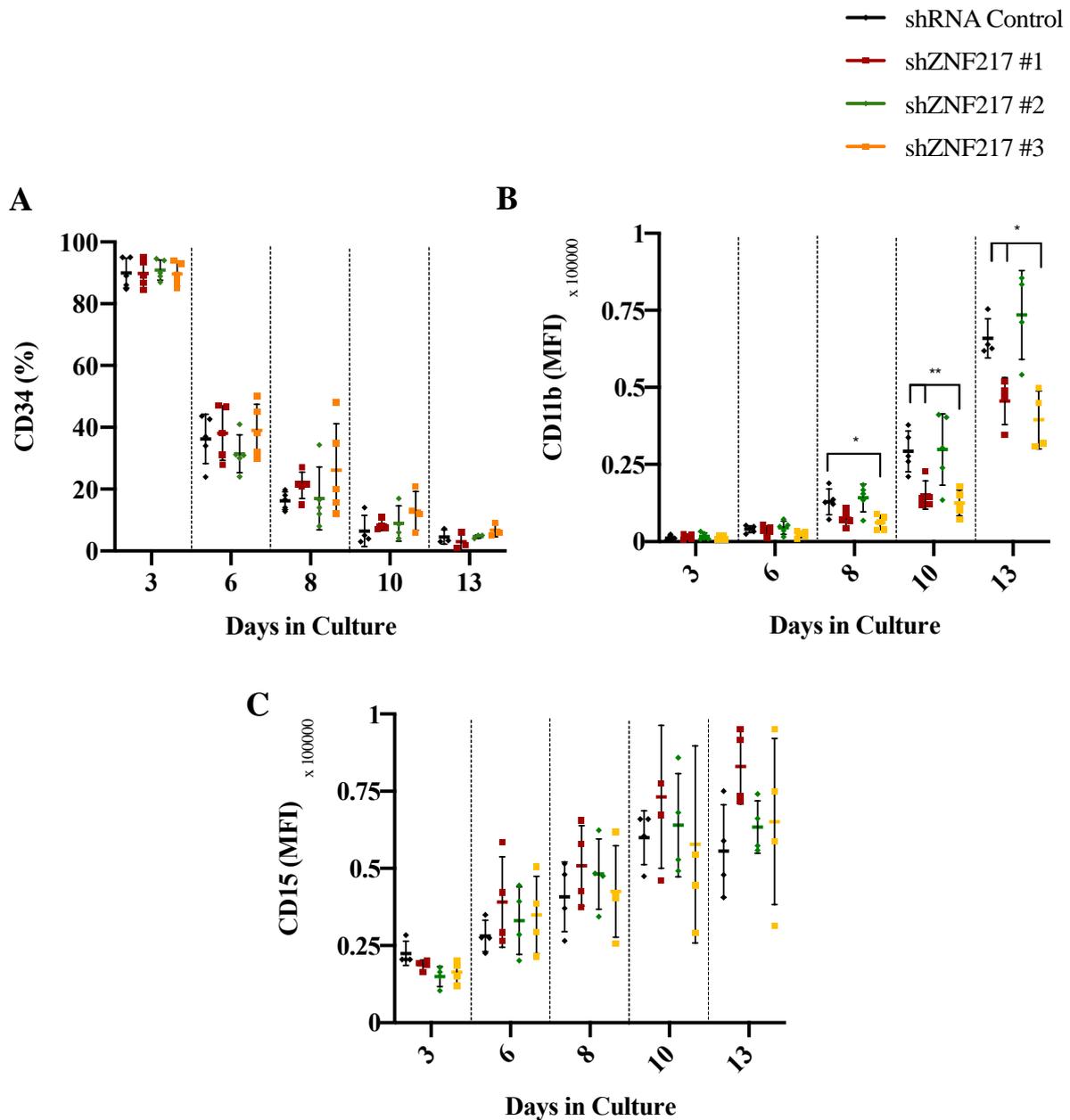
**Supplementary Figure 35 – ZNF217 knockdown disrupts the balance between the monocytic and granulocytic populations in CD34<sup>+</sup> HSPC during myeloid cell development**

Summary dot plot showing percentage of (A) monocytic (CD13<sup>+</sup> CD36<sup>+</sup>), (B) granulocytic (CD13<sup>+</sup> CD36<sup>-</sup>) and (C) erythroid cells (CD13<sup>-</sup> CD36<sup>+</sup>) in control and three ZNF217-KD cultures. Cells were re-seeded at the appropriate density following each time-point, as described in 2.3.3. Gating strategies were applied as described in **Supplementary Figure 1**. Data indicates Mean  $\pm$  1SD (n $\geq$ 4). Significant differences between shZNF217 cultures and control at each time-point were analysed by one-way ANOVA with Bonferroni's multiple comparisons correction; \* denotes p<0.05; \*\* denotes p<0.01.



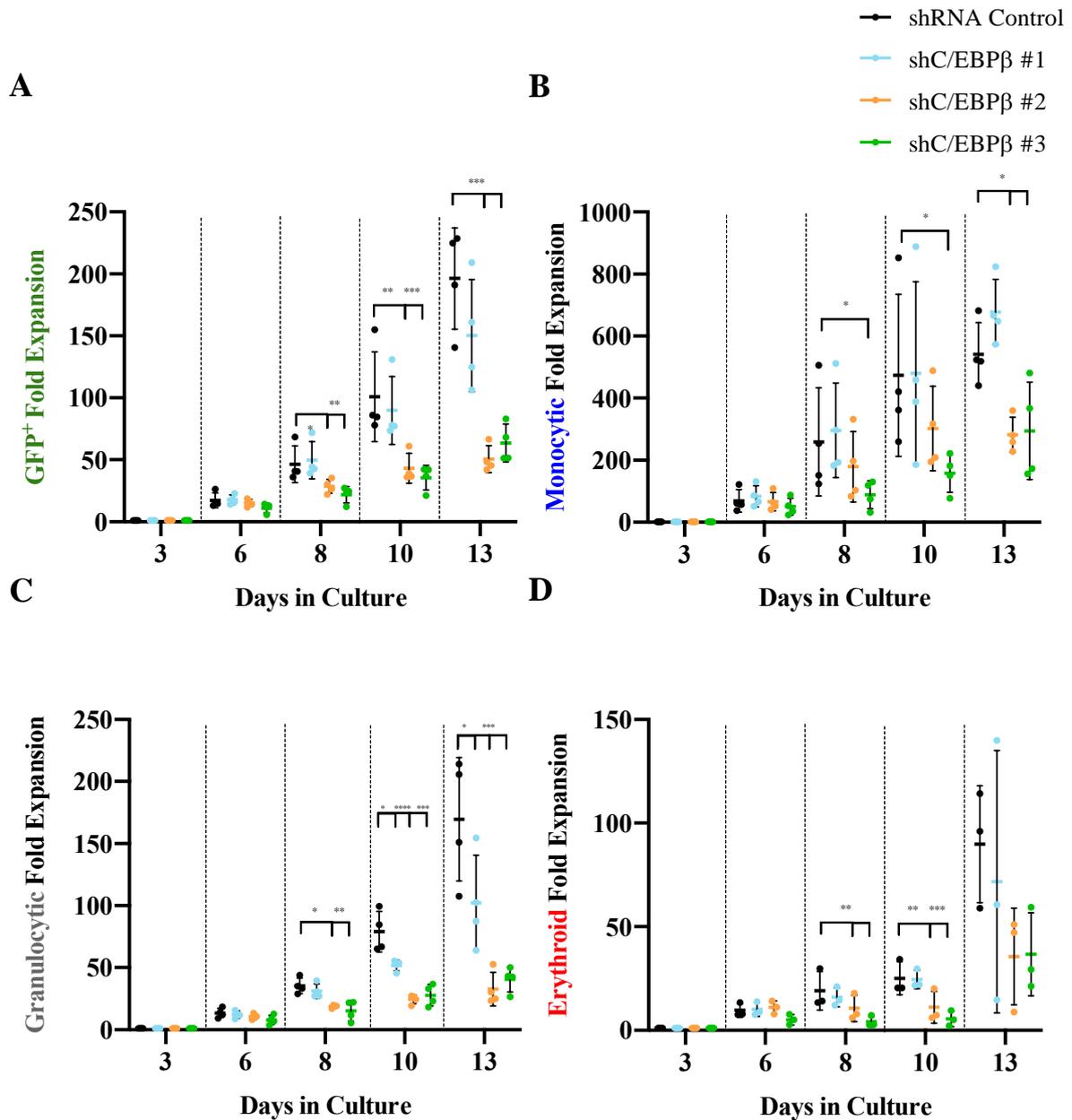
### Supplementary Figure 36 – Knockdown of ZNF217 does not significantly alter monocytic development

Summary dot plots showing (A) CD34<sup>+</sup> expression in terms of percentage in monocytic cells (CD13<sup>+</sup> CD36<sup>+</sup>) over time for control and ZNF217-KD cells. Summary data showing (B) CD11b and (C) CD14 expression in terms of MFI in monocytic cells over time for control and ZNF217-KD cells. Cells were re-seeded at the appropriate density following each time-point, as described in 2.3.3. Gating strategies were applied as described in **Supplementary Figure 1**. Data indicates Mean  $\pm$  1SD (n=4). Significant differences between shZNF217 cultures and control at each time-point were analysed by one-way ANOVA with Bonferroni's multiple comparisons correction; \* denotes p<0.05; \*\* denotes p<0.01.



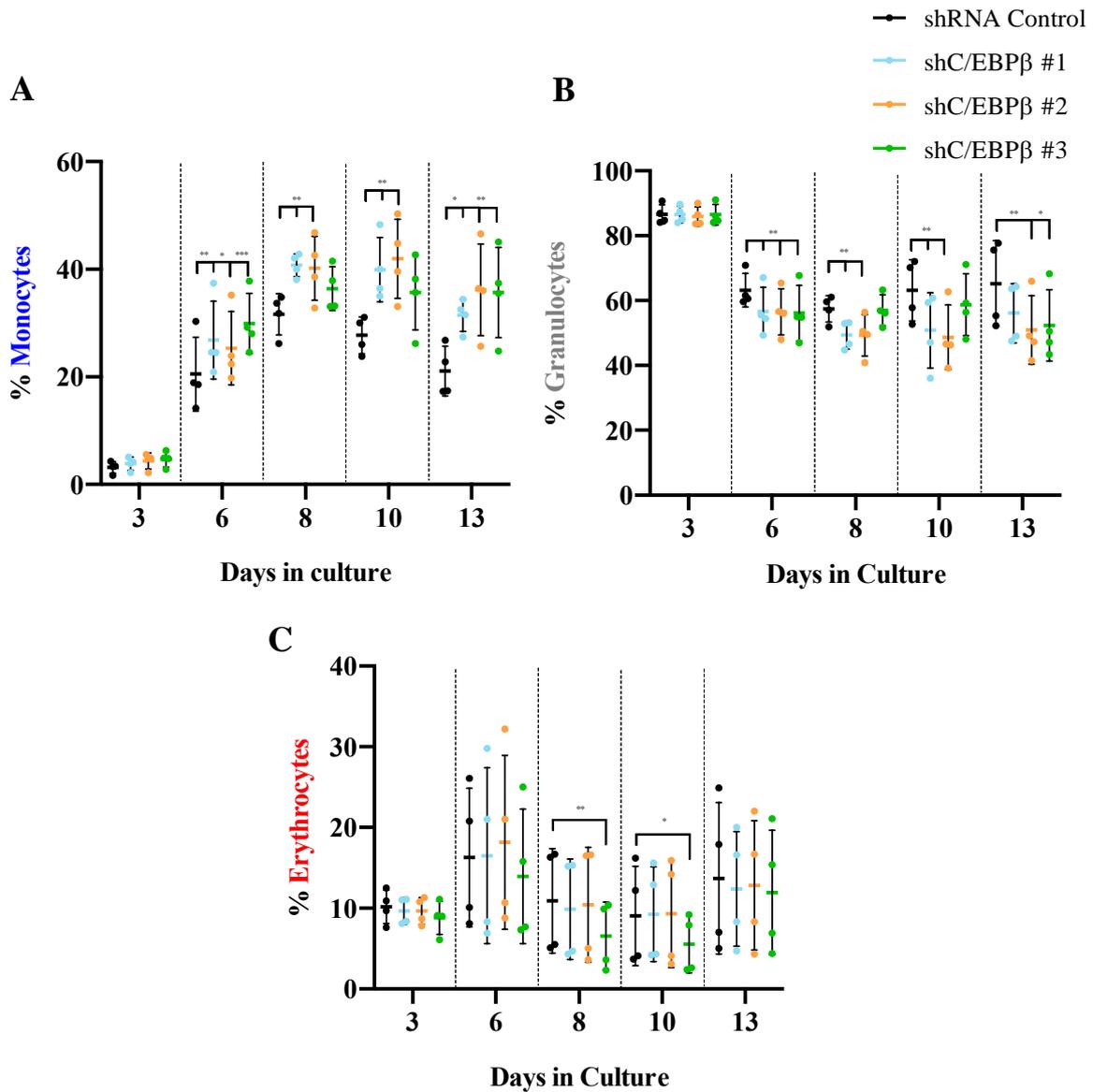
### Supplementary Figure 37 – Knockdown of ZNF217 does not significantly alter granulocytic development

**Summary dot plots showing (A)** CD34<sup>+</sup> expression in terms of percentage in granulocytic cells (CD13<sup>+</sup> CD36<sup>-</sup>) over time for control and ZNF217-KD cells. Summary data showing **(B)** CD11b and **(C)** CD15 expression in terms of MFI in granulocytic cells over time for control and ZNF217-KD cells. Cells were re-seeded at the appropriate density following each time-point, as described in 2.3.3. Gating strategies were applied as described in **Supplementary Figure 1**. Data indicates Mean ± 1SD (n≥4). Significant differences between shZNF217 cultures and control at each time-point were analysed by one-way ANOVA with Bonferroni's multiple comparisons correction; \* denotes p<0.05; \*\* denotes p<0.01.



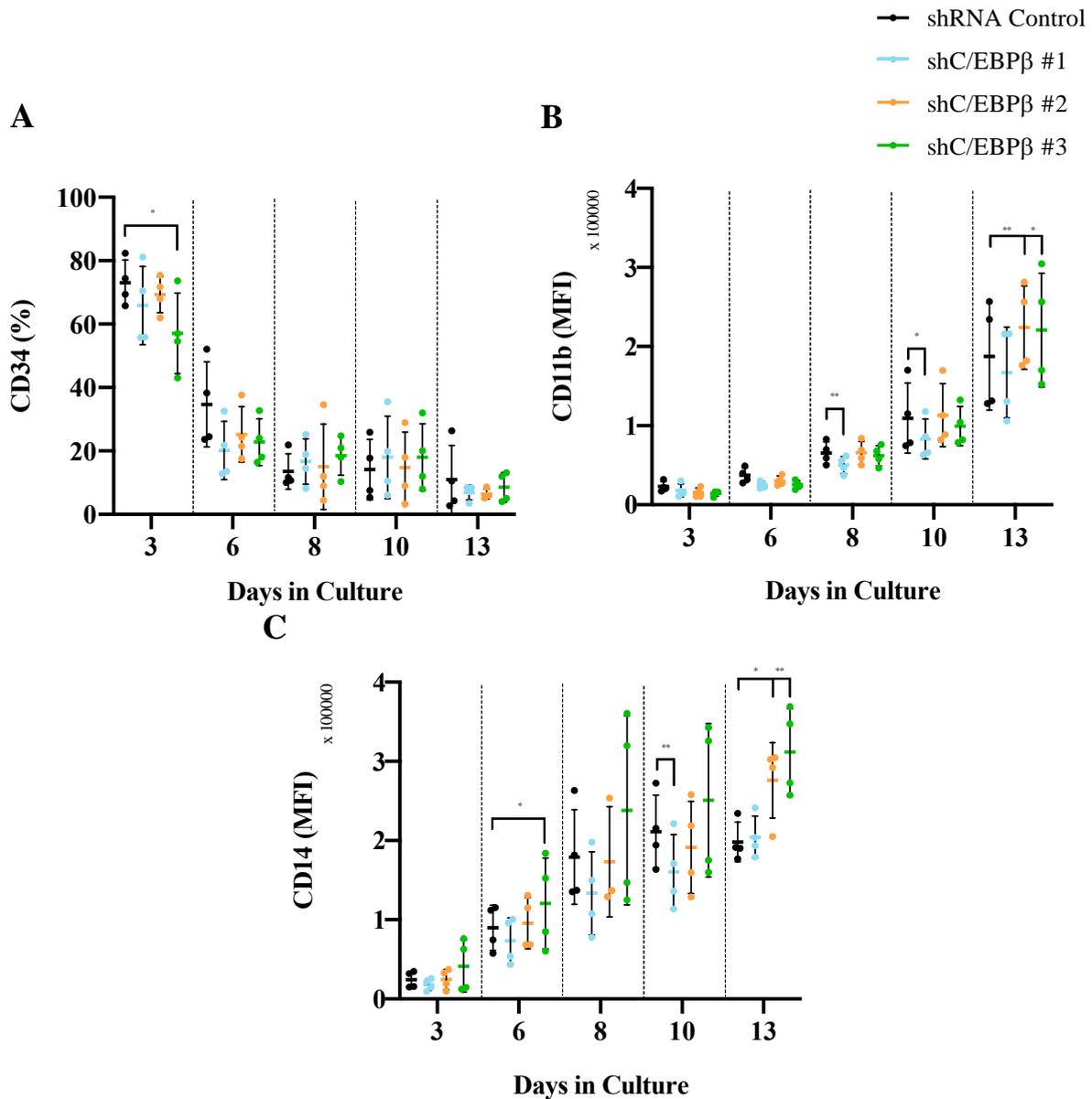
### Supplementary Figure 38 – KD of C/EBPβ decreases granulocytic growth in myeloid development

Summary dot plots showing the (A) cumulative fold-expansion of control and three shC/EBPβ constructs for GFP positive cells in culture medium containing IL-3, SCF, G- and GM-CSF, grown over 13 days. (B) Cumulative fold-expansion of monocytic cells (CD13<sup>+</sup> CD36<sup>+</sup>) transduced with a control or shC/EBPβ vectors. (C) Cumulative fold-expansion of granulocytic cells (CD13<sup>+</sup> CD36<sup>-</sup>) transduced with a control or shC/EBPβ vector. (D) Cumulative fold-expansion of erythroid cells (CD13<sup>-</sup> CD36<sup>+</sup>) transduced with a control or shC/EBPβ vector. Cells were re-seeded at the appropriate density following each time-point, as described in 2.3.3. Gating strategies were applied as described in **Supplementary Figure 1**. Data indicates mean ± 1SD (n=4). Significant differences between shC/EBPβ cultures and control at each time-point were analysed by one-way ANOVA with Bonferroni's multiple comparisons correction; \* denotes p<0.05; \*\* denotes p<0.01; \*\*\* denotes p<0.001; \*\*\*\* denotes p<0.0001.



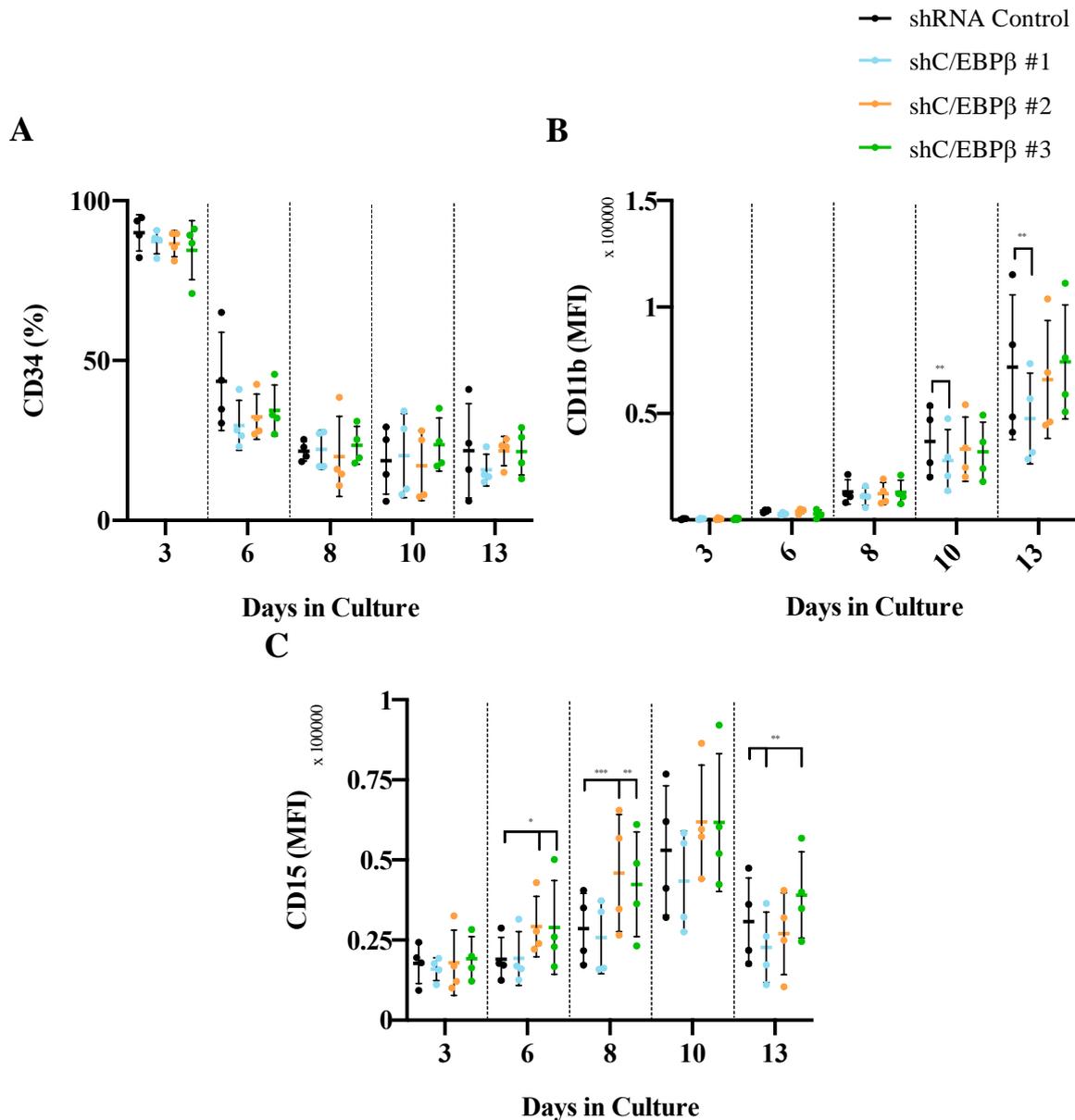
**Supplementary Figure 39 – C/EBPβ KD disrupts the balance between the monocytic and granulocytic populations in CD34<sup>+</sup> HSPC during myeloid cell development**

Summary dot plots showing percentage of (A) monocytic (CD13<sup>+</sup> CD36<sup>+</sup>), (B) granulocytic (CD13<sup>+</sup> CD36<sup>-</sup>) and (C) erythroid cells (CD13<sup>-</sup> CD36<sup>+</sup>) in control and three shC/EBPβ cultures. Cells were re-seeded at the appropriate density following each time-point, as described in 2.3.3. Gating strategies were applied as described in Supplementary Figure 1. Data indicates mean ± 1SD (n=4). Significant differences between shC/EBPβ cultures and control at each time-point were analysed by one-way ANOVA with Bonferroni's multiple comparisons correction; \* denotes p<0.05; \*\* denotes p<0.01; \*\*\* denotes p<0.001.



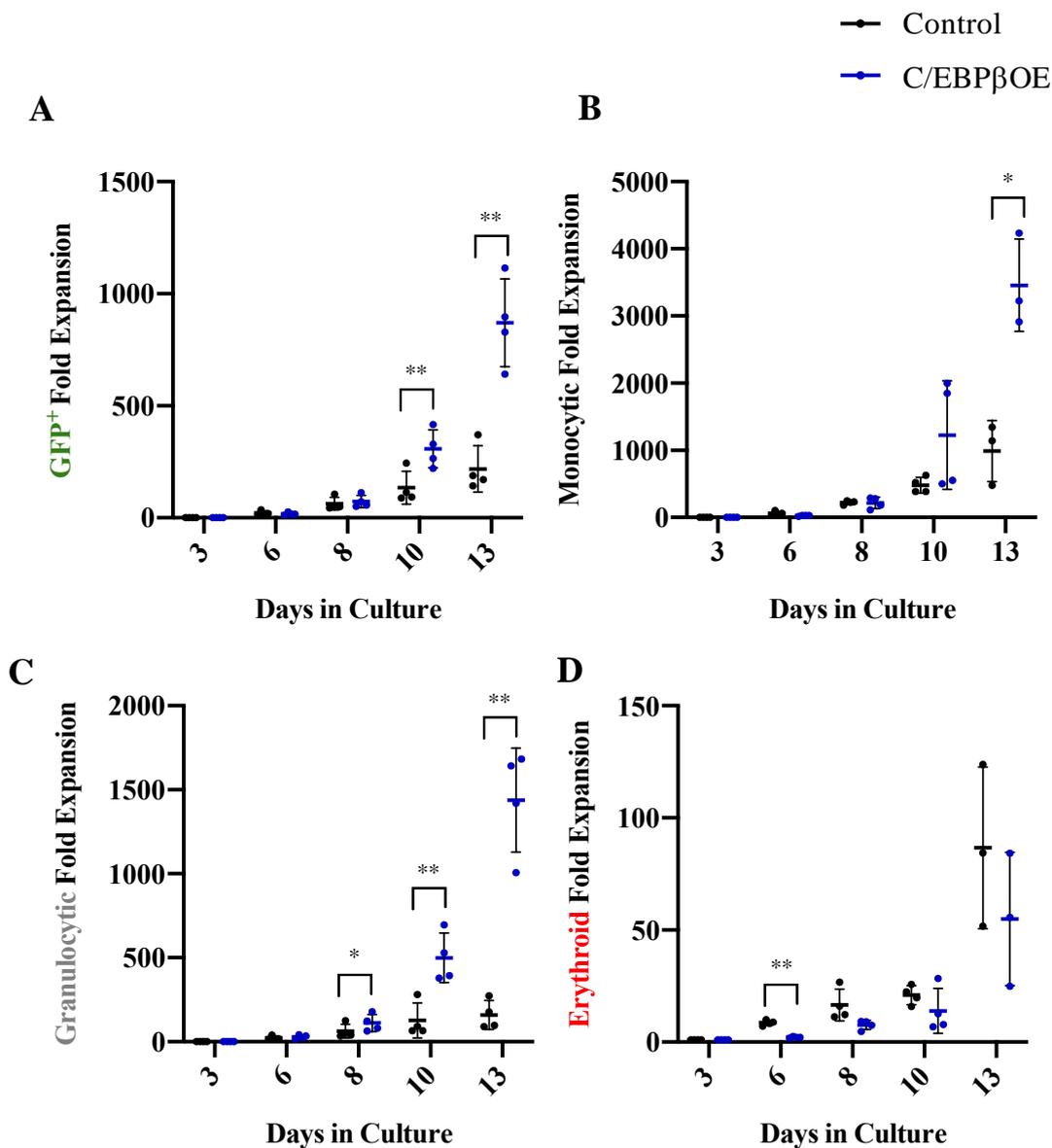
### Supplementary Figure 40 – KD of C/EBP $\beta$ has little impact on normal monocytic differentiation

Summary dot plots showing (A) CD34<sup>+</sup> expression in terms of percentage in monocytic cells (CD13<sup>+</sup> CD36<sup>+</sup>) over time for control and shC/EBP $\beta$  cells. Summary data showing (B) CD11b and (C) CD14 expression in terms of MFI in monocytic cells over time for control and shC/EBP $\beta$  cells. Cells were re-seeded at the appropriate density following each time-point, as described in 2.3.3. Gating strategies were applied as described in **Supplementary Figure 1**. Data indicates mean  $\pm$  1SD (n=4). Significant differences between shC/EBP $\beta$  cultures and control at each time-point were analysed by one-way ANOVA with Bonferroni's multiple comparisons correction; \* denotes p<0.05; \*\* denotes p<0.01.



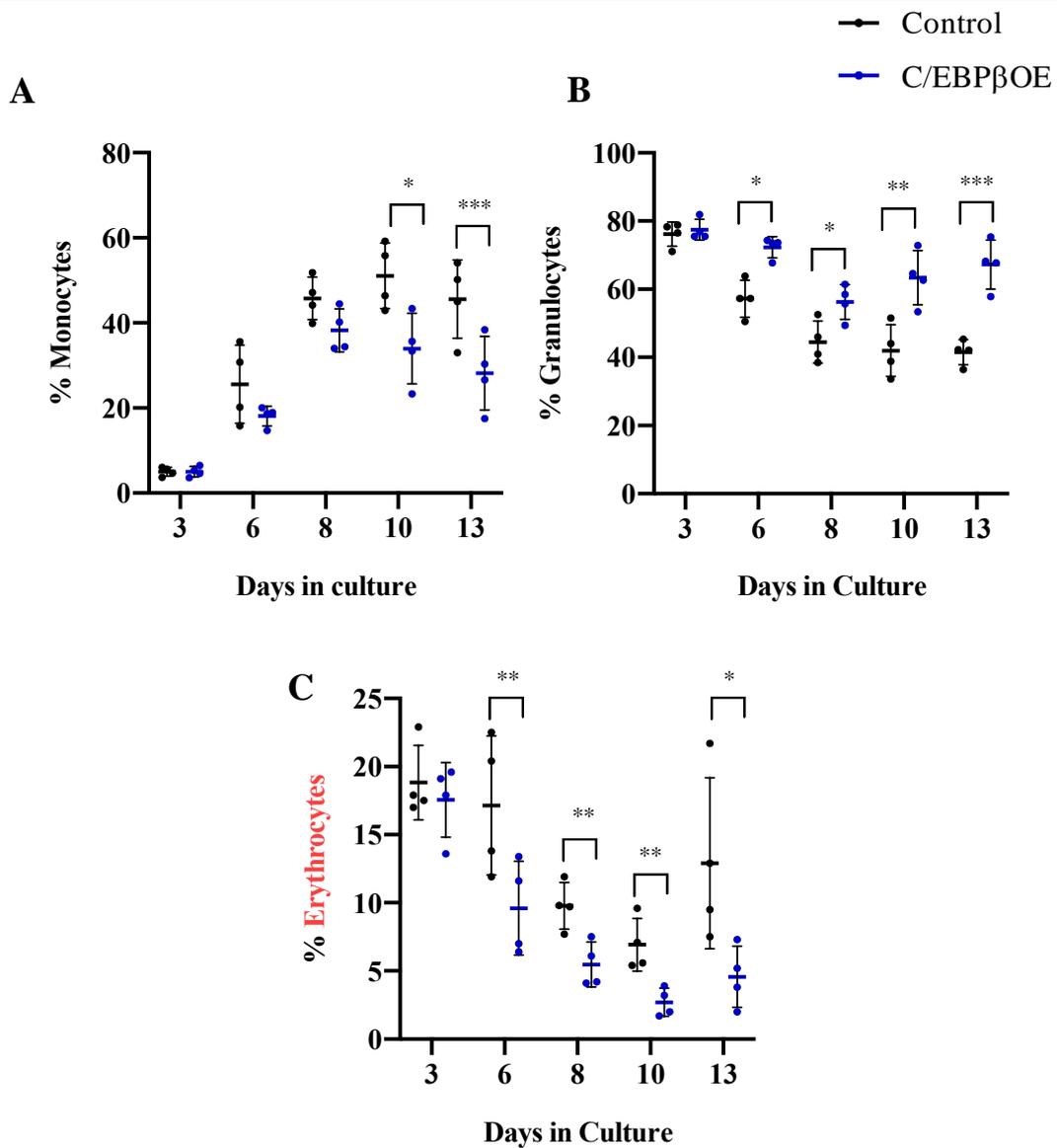
**Supplementary Figure 41 – KD of C/EBPβ promotes granulocytic development**

**Summary dot plots showing (A)** CD34<sup>+</sup> expression in terms of percentage in granulocytic cells (CD13<sup>+</sup> CD36<sup>-</sup>) over time for control and shC/EBPβ cells. Summary data showing **(B)** CD11b and **(C)** CD15 expression in terms of MFI in granulocytic cells over time for control and shC/EBPβ cells. Cells were re-seeded at the appropriate density following each time-point, as described in 2.3.3. Gating strategies were applied as described in **Supplementary Figure 1**. Data indicates mean ± 1SD (n=4). Significant differences between shC/EBPβ cultures and control at each time-point was analysed by one-way ANOVA with Bonferroni's multiple comparisons correction; \* denotes p<0.05; \*\* denotes p<0.01; \*\*\* denotes p<0.001.



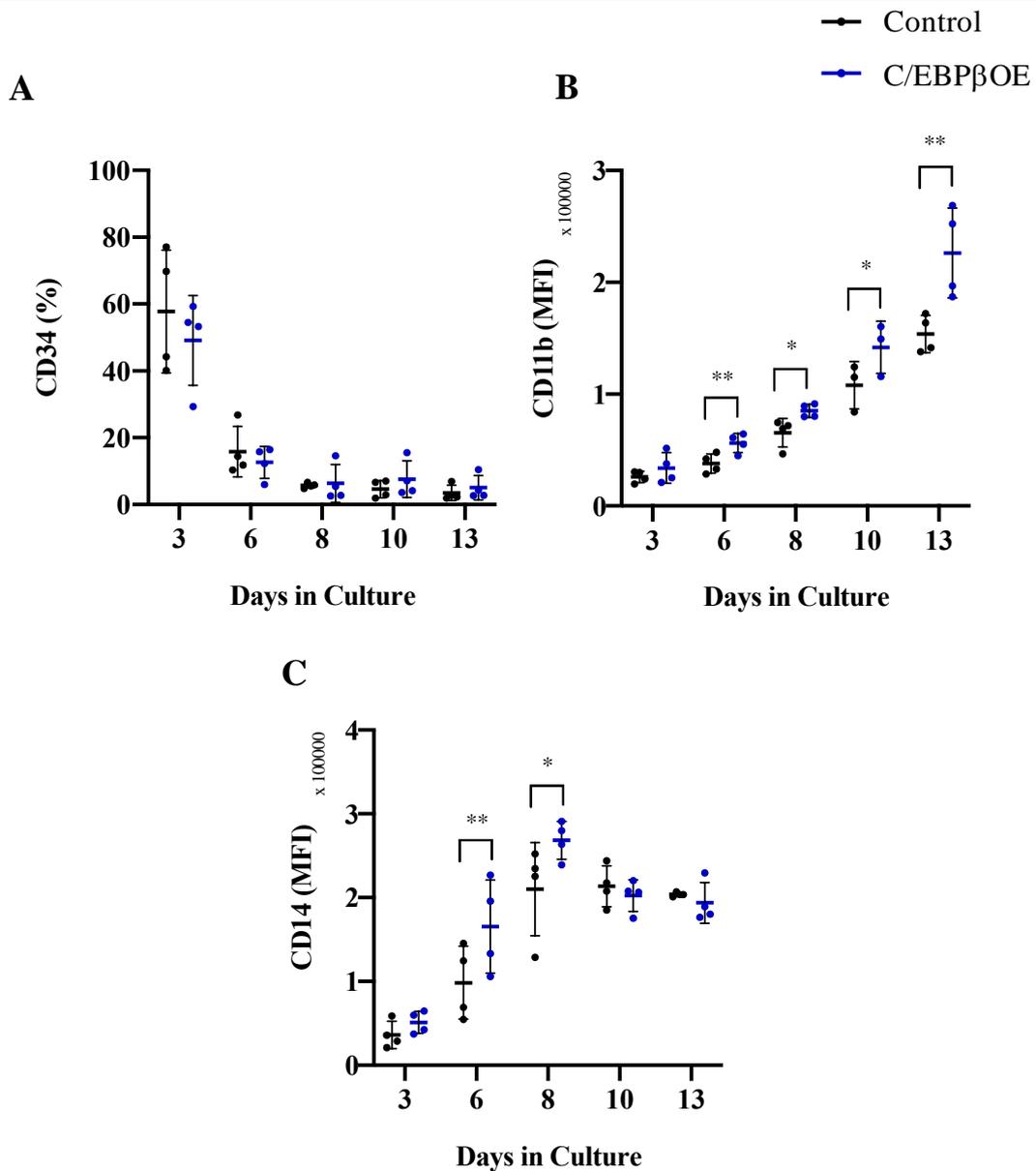
### Supplementary Figure 42 – Overexpression of C/EBPβ promotes monocytic and granulocytic growth in myeloid development

Summary dot plots showing the (A) cumulative fold-expansion of control and C/EBPβ-overexpression constructs in terms of GFP positivity in culture medium containing IL-3, SCF, G- and GM-CSF, grown over 13 days. (B) Cumulative fold-expansion of monocytic cells (CD13<sup>+</sup> CD36<sup>+</sup>) transduced with a control or C/EBPβ-OE vector. (C) Cumulative fold-expansion of granulocytic cells (CD13<sup>+</sup> CD36<sup>+</sup>) transduced with a control or C/EBPβ-OE vector. (D) Cumulative fold-expansion of erythroid cells (CD13<sup>-</sup> CD36<sup>+</sup>) transduced with a control or C/EBPβ-OE vector. Cells were re-seeded at the appropriate density following each time-point, as described in 2.3.3. Gating strategies were applied as described in Supplementary Figure 1. Data indicates mean ± 1SD (n=4). Significant differences between C/EBPβ-overexpression and control cultures at each time-point were analysed by paired T-test; \* denotes p<0.05; \*\* denotes p<0.01.



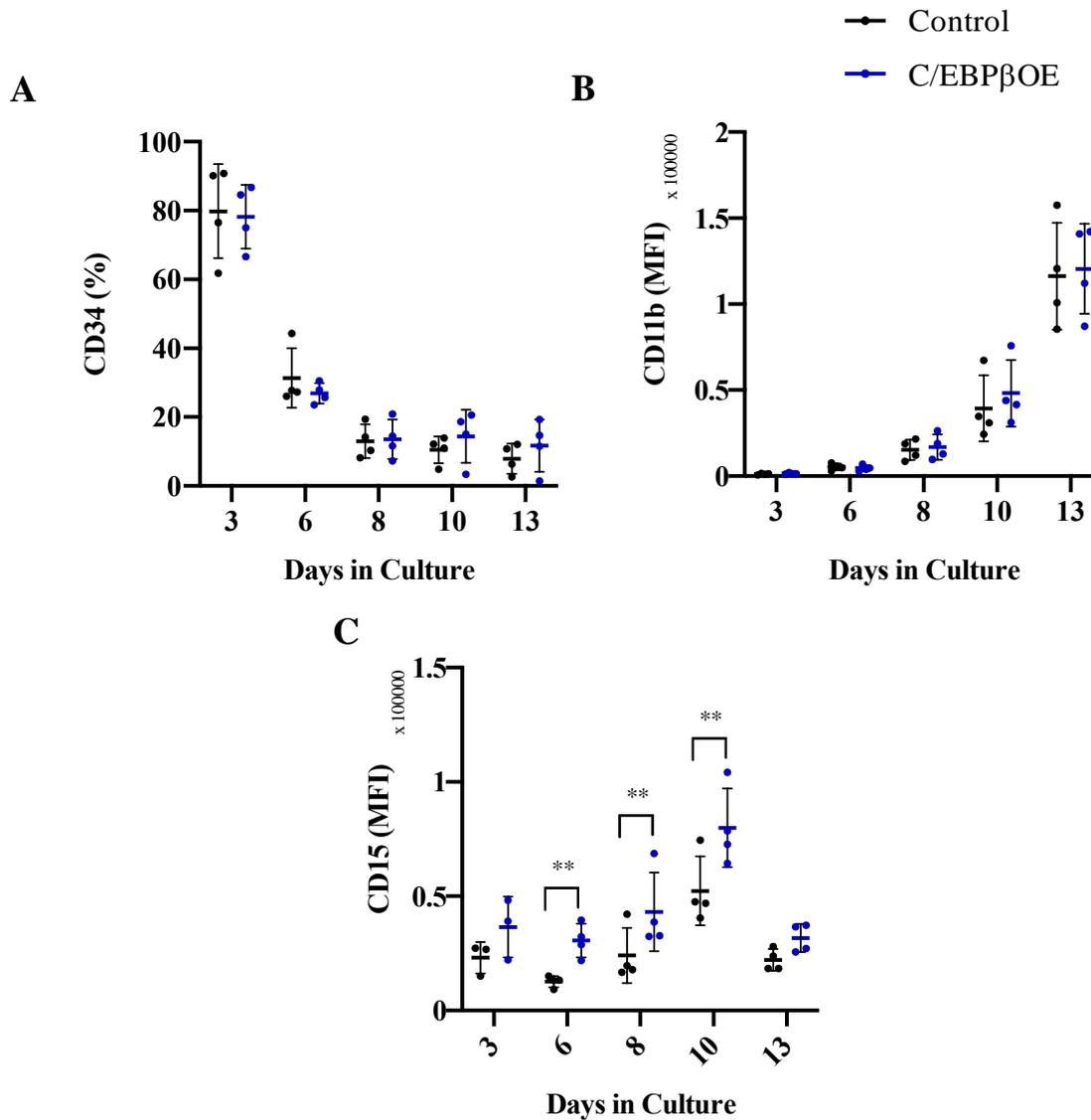
**Supplementary Figure 43 – C/EBPβ overexpression disrupts the balance between myeloid population during haematopoietic development**

Summary dot plots showing the proportion of (A) monocytic (CD13<sup>+</sup> CD36<sup>+</sup>), (B) granulocytic (CD13<sup>+</sup> CD36<sup>-</sup>) and (C) erythroid cells (CD13<sup>-</sup> CD36<sup>+</sup>) in control and C/EBPβ-overexpressing cultures. Cells were re-seeded at the appropriate density following each time-point, as described in 2.3.3. Gating strategies were applied as described in Supplementary Figure 1. Data indicates mean ± 1SD (n=4). Significant differences between C/EBPβ-expressing cells and control at each time-point were analysed by paired t-test; \* denotes p<0.05, \*\* denotes p<0.01, \*\*\* denotes p<0.001.



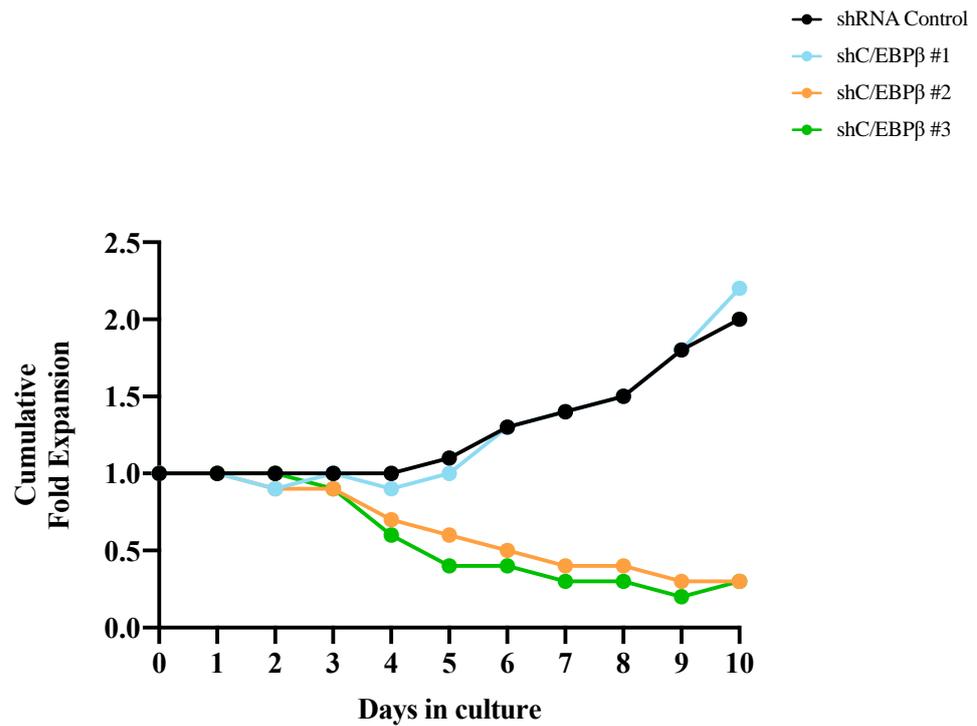
**Supplementary Figure 44 – C/EBP $\beta$  overexpression promotes the upregulation of monocytic markers in monocytic progenitors**

Summary dot plots showing (A) CD34<sup>+</sup> expression in terms of percentage in monocytic cells (CD13<sup>+</sup> CD36<sup>+</sup>) over time for control and C/EBP $\beta$ -overexpressing cells. Summary data showing (B) CD11b and (C) CD14 expression in terms of MFI in monocytic cells over time for control and C/EBP $\beta$ -overexpressing cells. Cells were re-seeded at the appropriate density following each time-point, as described in 2.3.3. Gating strategies were applied as described in **Supplementary Figure 1**. Data indicates mean  $\pm$  1SD (n=4). Significant differences between C/EBP $\beta$ -expressing cells and control at each time-point were analysed by paired t-test; \* denotes p<0.05, \*\* denotes p<0.01.



**Supplementary Figure 45 – C/EBPβ overexpression promotes the upregulation of granulocytic markers in granulocytic progenitors**

Summary dot plots showing (A) CD34<sup>+</sup> expression in terms of percentage in granulocytic cells (CD13<sup>+/−</sup>CD36<sup>−</sup>) over time for control and C/EBPβ-overexpressing cells. Summary data showing (B) CD11b and (C) CD15 expression in terms of MFI in granulocytic cells over time for control and C/EBPβ-overexpressing cells. Cells were re-seeded at the appropriate density following each time-point, as described in 2.3.3. Gating strategies were applied as described in **Supplementary Figure 1**. Data indicates mean ± 1SD (n=4). Significant differences between C/EBPβ-expressing cells and control at each time-point were analysed by paired t-test; \*\* denotes p<0.01.



**Supplementary Figure 46 – C/EBPβ KD in the Kasumi-1 cell line**

Growth of Kasumi-1 cells with three shC/EBPβ constructs compared to control, over 10 days of culture. (n=1). Cells were selected for GFP positivity on day 2 of culture, using the puromycin selectable marker.

---

## References

- Abbas, W., Kumar, A. & Herbein, G. **2015**. The eEF1A Proteins: At the Crossroads of Oncogenesis, Apoptosis, and Viral Infections. *Front Oncol*, 5, 75.
- Abelson, S., Collord, G., Ng, S. W. K., Weissbrod, O., Mendelson Cohen, N., Niemeyer, E., *et al.* **2018**. Prediction of acute myeloid leukaemia risk in healthy individuals. *Nature*, 559, 400-404.
- Acosta, J. C., O'loghlen, A., Banito, A., Guijarro, M. V., Augert, A., Raguz, S., *et al.* **2008**. Chemokine signaling via the CXCR2 receptor reinforces senescence. *Cell*, 133, 1006-18.
- Adolfsson, J., Borge, O. J., Bryder, D., Theilgaard-Mönch, K., Astrand-Grundström, I., Sitnicka, E., *et al.* **2001**. Upregulation of Flt3 expression within the bone marrow Lin(-)Sca1(+)c-kit(+) stem cell compartment is accompanied by loss of self-renewal capacity. *Immunity*, 15, 659-69.
- Adolfsson, J., Månsson, R., Buza-Vidas, N., Hultquist, A., Liuba, K., Jensen, C. T., *et al.* **2005**. Identification of Flt3+ lympho-myeloid stem cells lacking erythro-megakaryocytic potential a revised road map for adult blood lineage commitment. *Cell*, 121, 295-306.
- Agrawal, S., Hofmann, W. K., Tidow, N., Ehrich, M., Van Den Boom, D., Koschmieder, S., *et al.* **2007**. The C/EBPdelta tumor suppressor is silenced by hypermethylation in acute myeloid leukemia. *Blood*, 109, 3895-905.
- Aguilar-Morante, D., Cortes-Canteli, M., Sanz-Sancristobal, M., Santos, A. & Perez-Castillo, A. **2011**. Decreased CCAAT/enhancer binding protein  $\beta$  expression inhibits the growth of glioblastoma cells. *Neuroscience*, 176, 110-9.
- Aguilo, F., Zhang, F., Sancho, A., Fidalgo, M., Di Cecilia, S., Vashisht, A., *et al.* **2015**. Coordination of m(6)A mRNA Methylation and Gene Transcription by ZFP217 Regulates Pluripotency and Reprogramming. *Cell Stem Cell*, 17, 689-704.
- Aho, T. L., Sandholm, J., Peltola, K. J., Ito, Y. & Koskinen, P. J. **2006**. Pim-1 kinase phosphorylates RUNX family transcription factors and enhances their activity. *BMC Cell Biol*, 7, 21.
- Aikawa, Y., Nguyen, L. A., Isono, K., Takakura, N., Tagata, Y., Schmitz, M. L., *et al.* **2006**. Roles of HIPK1 and HIPK2 in AML1- and p300-dependent transcription, hematopoiesis and blood vessel formation. *EMBO J*, 25, 3955-65.
- Akagi, T., Saitoh, T., O'Kelly, J., Akira, S., Gombart, A. F. & Koeffler, H. P. **2008**. Impaired response to GM-CSF and G-CSF, and enhanced apoptosis in C/EBPbeta-deficient hematopoietic cells. *Blood*, 111, 2999-3004.
- Akashi, K., Traver, D., Miyamoto, T. & Weissman, I. L. **2000**. A clonogenic common myeloid progenitor that gives rise to all myeloid lineages. *Nature*, 404, 193-7.
- Akira, S., Isshiki, H., Sugita, T., Tanabe, O., Kinoshita, S., Nishio, Y., *et al.* **1990**. A nuclear factor for IL-6 expression (NF-IL6) is a member of a C/EBP family. *EMBO J*, 9, 1897-906.
- Alcalay, M., Meani, N., Gelmetti, V., Fantozzi, A., Fagioli, M., Orleth, A., *et al.* **2003**. Acute myeloid leukemia fusion proteins deregulate genes involved in stem cell maintenance and DNA repair. *J Clin Invest*, 112, 1751-61.
- Alexander, W. S., Roberts, A. W., Nicola, N. A., Li, R. & Metcalf, D. **1996**. Deficiencies in progenitor cells of multiple hematopoietic lineages and defective megakaryocytopoiesis in mice lacking the thrombopoietic receptor c-Mpl. *Blood*, 87, 2162-70.
- Amann, J. M., Nip, J., Strom, D. K., Lutterbach, B., Harada, H., Lenny, N., *et al.* **2001**. ETO, a target of t(8;21) in acute leukemia, makes distinct contacts with multiple histone deacetylases and binds mSin3A through its oligomerization domain. *Mol Cell Biol*, 21, 6470-83.

- Angerer, N. D., Du, Y., Nalbant, D. & Williams, S. C. **1999**. A short conserved motif is required for repressor domain function in the myeloid-specific transcription factor CCAAT/enhancer-binding protein epsilon. *J Biol Chem*, 274, 4147-54.
- Arai, F., Hirao, A., Ohmura, M., Sato, H., Matsuoka, S., Takubo, K., *et al.* **2004**. Tie2/angiopoietin-1 signaling regulates hematopoietic stem cell quiescence in the bone marrow niche. *Cell*, 118, 149-61.
- Arber, D. A., Orazi, A., Hasserjian, R., Thiele, J., Borowitz, M. J., Le Beau, M. M., *et al.* **2016**. The 2016 revision to the World Health Organization classification of myeloid neoplasms and acute leukemia. *Blood*, 127, 2391-405.
- Arnal-Estapé, A., Tarragona, M., Morales, M., Guiu, M., Nadal, C., Massagué, J., *et al.* **2010**. HER2 silences tumor suppression in breast cancer cells by switching expression of C/EBP $\beta$  isoforms. *Cancer Res*, 70, 9927-36.
- Attar, A. **2014**. Changes in the Cell Surface Markers During Normal Hematopoiesis: A Guide to Cell Isolation. *Global Journal of Hematology and Blood Transfusion*, 1, 20-28.
- Bacher, U., Haferlach, T., Schoch, C., Kern, W. & Schnittger, S. **2006**. Implications of NRAS mutations in AML: a study of 2502 patients. *Blood*, 107, 3847-53.
- Bae, S. C., Ogawa, E., Maruyama, M., Oka, H., Satake, M., Shigesada, K., *et al.* **1994**. PEBP2 alpha B/mouse AML1 consists of multiple isoforms that possess differential transactivation potentials. *Mol Cell Biol*, 14, 3242-52.
- Bagger, F. O., Sasivarevic, D., Sohi, S. H., Laursen, L. G., Punthir, S., Sønderby, C. K., *et al.* **2016**. BloodSpot: a database of gene expression profiles and transcriptional programs for healthy and malignant haematopoiesis. *Nucleic Acids Res*, 44, D917-24.
- Bai, W. D., Ye, X. M., Zhang, M. Y., Zhu, H. Y., Xi, W. J., Huang, X., *et al.* **2014**. MiR-200c suppresses TGF- $\beta$  signaling and counteracts trastuzumab resistance and metastasis by targeting ZNF217 and ZEB1 in breast cancer. *Int J Cancer*, 135, 1356-68.
- Baldrige, M. T., King, K. Y., Boles, N. C., Weksberg, D. C. & Goodell, M. A. **2010**. Quiescent haematopoietic stem cells are activated by IFN-gamma in response to chronic infection. *Nature*, 465, 793-7.
- Banerjee, S., Xie, N., Cui, H., Tan, Z., Yang, S., Icyuz, M., *et al.* **2013**. MicroRNA let-7c regulates macrophage polarization. *J Immunol*, 190, 6542-9.
- Bangsow, C., Rubins, N., Glusman, G., Bernstein, Y., Negreanu, V., Goldenberg, D., *et al.* **2001**. The RUNX3 gene--sequence, structure and regulated expression. *Gene*, 279, 221-32.
- Banu, N. & Williams, N. **1995**. Immunochemical characterisation of megakaryocyte potentiator activity from mouse bone marrow. *J Cell Physiol*, 163, 486-92.
- Barretina, J., Caponigro, G., Stransky, N., Venkatesan, K., Margolin, A. A., Kim, S., *et al.* **2012**. The Cancer Cell Line Encyclopedia enables predictive modelling of anticancer drug sensitivity. *Nature*, 483, 603-7.
- Barseguian, K., Lutterbach, B., Hiebert, S. W., Nickerson, J., Lian, J. B., Stein, J. L., *et al.* **2002**. Multiple subnuclear targeting signals of the leukemia-related AML1/ETO and ETO repressor proteins. *Proc Natl Acad Sci U S A*, 99, 15434-9.
- Bassères, D. S. & Baldwin, A. S. **2006**. Nuclear factor-kappaB and inhibitor of kappaB kinase pathways in oncogenic initiation and progression. *Oncogene*, 25, 6817-30.
- Batard, P., Monier, M. N., Fortunel, N., Ducos, K., Sansilvestri-Morel, P., Phan, T., *et al.* **2000**. TGF-(beta)1 maintains hematopoietic immaturity by a reversible negative control of cell cycle and induces CD34 antigen up-modulation. *J Cell Sci*, 113 ( Pt 3), 383-90.

- Baum, C. M., Weissman, I. L., Tsukamoto, A. S., Buckle, A. M. & Peault, B. **1992**. Isolation of a candidate human hematopoietic stem-cell population. *Proc Natl Acad Sci U S A*, 89, 2804-8.
- Bee, T., Liddiard, K., Swiers, G., Bickley, S. R., Vink, C. S., Jarratt, A., *et al.* **2009**. Alternative Runx1 promoter usage in mouse developmental hematopoiesis. *Blood Cells Mol Dis*, 43, 35-42.
- Bégay, V., Baumeier, C., Zimmermann, K., Heuser, A. & Leutz, A. **2018**. The C/EBP $\beta$  LIP isoform rescues loss of C/EBP $\beta$  function in the mouse. *Sci Rep*, 8, 8417.
- Ben-Ami, O., Friedman, D., Leshkowitz, D., Goldenberg, D., Orlovsky, K., Pencovich, N., *et al.* **2013**. Addiction of t(8;21) and inv(16) acute myeloid leukemia to native RUNX1. *Cell Rep*, 4, 1131-43.
- Bene, M. C., Castoldi, G., Knapp, W., Ludwig, W. D., Matutes, E., Orfao, A., *et al.* **1995**. Proposals for the immunological classification of acute leukemias. European Group for the Immunological Characterization of Leukemias (EGIL). *Leukemia*, 9, 1783-6.
- Bennett, J. M., Catovsky, D., Daniel, M. T., Flandrin, G., Galton, D. A., Gralnick, H. R., *et al.* **1976**. Proposals for the classification of the acute leukaemias. French-American-British (FAB) co-operative group. *Br J Haematol*, 33, 451-8.
- Benveniste, P., Frelin, C., Janmohamed, S., Barbara, M., Herrington, R., Hyam, D., *et al.* **2010**. Intermediate-term hematopoietic stem cells with extended but time-limited reconstitution potential. *Cell Stem Cell*, 6, 48-58.
- Berg, T., Didon, L., Barton, J., Andersson, O. & Nord, M. **2005**. Glucocorticoids increase C/EBP $\beta$  activity in the lung epithelium via phosphorylation. *Biochem Biophys Res Commun*, 334, 638-45.
- Berrier, A., Siu, G. & Calame, K. **1998**. Transcription of a minimal promoter from the NF-IL6 gene is regulated by CREB/ATF and SP1 proteins in U937 promonocytic cells. *J Immunol*, 161, 2267-75.
- Bhatia, M., Bonnet, D., Wu, D., Murdoch, B., Wrana, J., Gallacher, L., *et al.* **1999**. Bone morphogenetic proteins regulate the developmental program of human hematopoietic stem cells. *J Exp Med*, 189, 1139-48.
- Bhatia, M., Wang, J. C., Kapp, U., Bonnet, D. & Dick, J. E. **1997**. Purification of primitive human hematopoietic cells capable of repopulating immune-deficient mice. *Proc Natl Acad Sci U S A*, 94, 5320-5.
- Biggs, J. R., Peterson, L. F., Zhang, Y., Kraft, A. S. & Zhang, D. E. **2006**. AML1/RUNX1 phosphorylation by cyclin-dependent kinases regulates the degradation of AML1/RUNX1 by the anaphase-promoting complex. *Mol Cell Biol*, 26, 7420-9.
- Birkenkamp, K. U., Geugien, M., Schepers, H., Westra, J., Lemmink, H. H. & Vellenga, E. **2004**. Constitutive NF- $\kappa$ B DNA-binding activity in AML is frequently mediated by a Ras/PI3-K/PKB-dependent pathway. *Leukemia*, 18, 103-12.
- Bitter, M. A., Le Beau, M. M., Rowley, J. D., Larson, R. A., Golomb, H. M. & Vardiman, J. W. **1987**. Associations between morphology, karyotype, and clinical features in myeloid leukemias. *Hum Pathol*, 18, 211-25.
- Blumenthal, E., Greenblatt, S., Huang, G., Ando, K., Xu, Y. & Nimer, S. D. **2017**. Covalent Modifications of RUNX Proteins: Structure Affects Function. *Adv Exp Med Biol*, 962, 33-44.
- Blyth, K., Cameron, E. R. & Neil, J. C. **2005**. The RUNX genes: gain or loss of function in cancer. *Nat Rev Cancer*, 5, 376-87.
- Blyth, K., Vaillant, F., Jenkins, A., McDonald, L., Pringle, M. A., Huser, C., *et al.* **2010**. Runx2 in normal tissues and cancer cells: A developing story. *Blood Cells Mol Dis*, 45, 117-23.

- Bonnet, D. & Dick, J. E. **1997**. Human acute myeloid leukemia is organized as a hierarchy that originates from a primitive hematopoietic cell. *Nat Med*, 3, 730-7.
- Borthakur, G., Kantarjian, H., Ravandi, F., Zhang, W., Konopleva, M., Wright, J. J., *et al.* **2011**. Phase I study of sorafenib in patients with refractory or relapsed acute leukemias. *Haematologica*, 96, 62-8.
- Bosma, G. C., Custer, R. P. & Bosma, M. J. **1983**. A severe combined immunodeficiency mutation in the mouse. *Nature*, 301, 527-30.
- Bourquin, J. P., Subramanian, A., Langebrake, C., Reinhardt, D., Bernard, O., Ballerini, P., *et al.* **2006**. Identification of distinct molecular phenotypes in acute megakaryoblastic leukemia by gene expression profiling. *Proc Natl Acad Sci U S A*, 103, 3339-44.
- Boyer, S. W., Schroeder, A. V., Smith-Berdan, S. & Forsberg, E. C. **2011**. All hematopoietic cells develop from hematopoietic stem cells through Flk2/Flt3-positive progenitor cells. *Cell Stem Cell*, 9, 64-73.
- Braun, T., Carvalho, G., Fabre, C., Grosjean, J., Fenaux, P. & Kroemer, G. **2006**. Targeting NF-kappaB in hematologic malignancies. *Cell Death Differ*, 13, 748-58.
- Bravo, J., Li, Z., Speck, N. A. & Warren, A. J. **2001**. The leukemia-associated AML1 (Runx1)--CBF beta complex functions as a DNA-induced molecular clamp. *Nat Struct Biol*, 8, 371-8.
- Breig, O., Bras, S., Martinez Soria, N., Osman, D., Heidenreich, O., Haenlin, M., *et al.* **2014**. Pontin is a critical regulator for AML1-ETO-induced leukemia. *Leukemia*, 28, 1271-9.
- Brettingham-Moore, K. H., Taberlay, P. C. & Holloway, A. F. **2015**. Interplay between Transcription Factors and the Epigenome: Insight from the Role of RUNX1 in Leukemia. *Frontiers in Immunology*, 6, 499.
- Broudy, V. C. **1997**. Stem cell factor and hematopoiesis. *Blood*, 90, 1345-64.
- Brown, G., Tsapogas, P. & Ceredig, R. **2018**. The changing face of hematopoiesis: a spectrum of options is available to stem cells. *Immunol Cell Biol*, 96, 898-911.
- Buck, M., Poli, V., Hunter, T. & Chojkier, M. **2001**. C/EBPbeta phosphorylation by RSK creates a functional XEXD caspase inhibitory box critical for cell survival. *Mol Cell*, 8, 807-16.
- Buck, M., Turler, H. & Chojkier, M. **1994**. LAP (NF-IL-6), a tissue-specific transcriptional activator, is an inhibitor of hepatoma cell proliferation. *EMBO J*, 13, 851-60.
- Buenrostro, J. D., Corces, M. R., Lareau, C. A., Wu, B., Schep, A. N., Aryee, M. J., *et al.* **2018**. Integrated Single-Cell Analysis Maps the Continuous Regulatory Landscape of Human Hematopoietic Differentiation. *Cell*, 173, 1535-1548.e16.
- Burk, O., Mink, S., Ringwald, M. & Klempnauer, K. H. **1993**. Synergistic activation of the chicken mim-1 gene by v-myb and C/EBP transcription factors. *EMBO J*, 12, 2027-38.
- Burnett, A., Wetzler, M. & Löwenberg, B. **2011a**. Therapeutic advances in acute myeloid leukemia. *J Clin Oncol*, 29, 487-94.
- Burnett, A. K., Hills, R. K., Milligan, D., Kjeldsen, L., Kell, J., Russell, N. H., *et al.* **2011b**. Identification of patients with acute myeloblastic leukemia who benefit from the addition of gemtuzumab ozogamicin: results of the MRC AML15 trial. *J Clin Oncol*, 29, 369-77.
- Butler, J. M., Nolan, D. J., Vertes, E. L., Varnum-Finney, B., Kobayashi, H., Hooper, A. T., *et al.* **2010**. Endothelial cells are essential for the self-renewal and repopulation of Notch-dependent hematopoietic stem cells. *Cell Stem Cell*, 6, 251-64.

- Byrd, J. C., Mrózek, K., Dodge, R. K., Carroll, A. J., Edwards, C. G., Arthur, D. C., *et al.* **2002**. Pretreatment cytogenetic abnormalities are predictive of induction success, cumulative incidence of relapse, and overall survival in adult patients with de novo acute myeloid leukemia: results from Cancer and Leukemia Group B (CALGB 8461). *Blood*, 100, 4325-36.
- Cai, X., Gao, L., Teng, L., Ge, J., Oo, Z. M., Kumar, A. R., *et al.* **2015**. Runx1 Deficiency Decreases Ribosome Biogenesis and Confers Stress Resistance to Hematopoietic Stem and Progenitor Cells. *Cell Stem Cell*, 17, 165-77.
- Cai, X., Gaudet, J. J., Mangan, J. K., Chen, M. J., De Obaldia, M. E., Oo, Z., *et al.* **2011**. Runx1 loss minimally impacts long-term hematopoietic stem cells. *PLoS One*, 6, e28430.
- Cain, D. W., O'koren, E. G., Kan, M. J., Womble, M., Sempowski, G. D., Hopper, K., *et al.* **2013**. Identification of a tissue-specific, C/EBP $\beta$ -dependent pathway of differentiation for murine peritoneal macrophages. *J Immunol*, 191, 4665-75.
- Calkhoven, C. F., Müller, C. & Leutz, A. **2000**. Translational control of C/EBP $\alpha$  and C/EBP $\beta$  isoform expression. *Genes Dev*, 14, 1920-32.
- Calvi, L. M., Adams, G. B., Weibrecht, K. W., Weber, J. M., Olson, D. P., Knight, M. C., *et al.* **2003**. Osteoblastic cells regulate the haematopoietic stem cell niche. *Nature*, 425, 841-6.
- Cammenga, J., Niebuhr, B., Horn, S., Bergholz, U., Putz, G., Buchholz, F., *et al.* **2007**. RUNX1 DNA-binding mutants, associated with minimally differentiated acute myelogenous leukemia, disrupt myeloid differentiation. *Cancer Res*, 67, 537-45.
- Cancer Research UK 2021. Acute myeloid leukaemia (AML) incidence statistics. Cancer Research UK.
- Cao, Z., Umek, R. M. & Mcknight, S. L. **1991**. Regulated expression of three C/EBP isoforms during adipose conversion of 3T3-L1 cells. *Genes Dev*, 5, 1538-52.
- Cappello, C., Zwergal, A., Kanclerski, S., Haas, S. C., Kandemir, J. D., Huber, R., *et al.* **2009**. C/EBP $\beta$  enhances NF- $\kappa$ B-associated signalling by reducing the level of I $\kappa$ B $\alpha$ . *Cell Signal*, 21, 1918-24.
- Carro, M. S., Lim, W. K., Alvarez, M. J., Bollo, R. J., Zhao, X., Snyder, E. Y., *et al.* **2010**. The transcriptional network for mesenchymal transformation of brain tumours. *Nature*, 463, 318-25.
- Carver-Moore, K., Broxmeyer, H. E., Luoh, S. M., Cooper, S., Peng, J., Burstein, S. A., *et al.* **1996**. Low levels of erythroid and myeloid progenitors in thrombopoietin-and c-mpl-deficient mice. *Blood*, 88, 803-8.
- Caux, C., Favre, C., Saeland, S., Duvert, V., Mannoni, P., Durand, I., *et al.* **1989**. Sequential loss of CD34 and class II MHC antigens on purified cord blood hematopoietic progenitors cultured with IL-3: characterization of CD34-, HLA-DR+ cells. *Blood*, 74, 1287-94.
- Cerami, E., Gao, J., Dogrusoz, U., Gross, B. E., Sumer, S. O., Aksoy, B. A., *et al.* **2012**. The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. *Cancer Discov*, 2, 401-4.
- Challen, G. A., Sun, D., Jeong, M., Luo, M., Jelinek, J., Berg, J. S., *et al.* **2011**. Dnmt3a is essential for hematopoietic stem cell differentiation. *Nat Genet*, 44, 23-31.
- Chang, C. J., Shen, B. J. & Lee, S. C. **1995**. Autoregulated induction of the acute-phase response transcription factor gene, *agp/ebp*. *DNA Cell Biol*, 14, 529-37.
- Chang, E., Ganguly, S., Rajkhowa, T., Gocke, C. D., Levis, M. & Konig, H. **2016**. The combination of FLT3 and DNA methyltransferase inhibition is synergistically cytotoxic to FLT3/ITD acute myeloid leukemia cells. *Leukemia*, 30, 1025-32.

- Chang, M. K., Raggatt, L. J., Alexander, K. A., Kuliwaba, J. S., Fazzalari, N. L., Schroder, K., *et al.* **2008**. Osteal tissue macrophages are intercalated throughout human and mouse bone lining tissues and regulate osteoblast function in vitro and in vivo. *J Immunol*, 181, 1232-44.
- Chang, W., Rewari, A., Centrella, M. & Mccarthy, T. L. **2004**. Fos-related antigen 2 controls protein kinase A-induced CCAAT/enhancer-binding protein beta expression in osteoblasts. *J Biol Chem*, 279, 42438-44.
- Charles, A., Tang, X., Crouch, E., Brody, J. S. & Xiao, Z. X. **2001**. Retinoblastoma protein complexes with C/EBP proteins and activates C/EBP-mediated transcription. *J Cell Biochem*, 83, 414-25.
- Chen, J. D. & Evans, R. M. **1995**. A transcriptional co-repressor that interacts with nuclear hormone receptors. *Nature*, 377, 454-7.
- Chen, K. G., Sale, S., Tan, T., Ermoian, R. P. & Sikic, B. I. **2004**. CCAAT/enhancer-binding protein beta (nuclear factor for interleukin 6) transactivates the human MDR1 gene by interaction with an inverted CCAAT box in human cancer cells. *Mol Pharmacol*, 65, 906-16.
- Chen, M., Zhu, N., Liu, X., Laurent, B., Tang, Z., Eng, R., *et al.* **2015**. JMJD1C is required for the survival of acute myeloid leukemia by functioning as a coactivator for key transcription factors. *Genes Dev*, 29, 2123-39.
- Chen, S., Han, Y. H., Zheng, Y., Zhao, M., Yan, H., Zhao, Q., *et al.* **2009**. NDRG1 contributes to retinoic acid-induced differentiation of leukemic cells. *Leuk Res*, 33, 1108-13.
- Chen, X., Liu, W., Ambrosino, C., Ruocco, M. R., Poli, V., Romani, L., *et al.* **1997**. Impaired generation of bone marrow B lymphocytes in mice deficient in C/EBPbeta. *Blood*, 90, 156-64.
- Chen, Z., Torrens, J. I., Anand, A., Spiegelman, B. M. & Friedman, J. M. **2005**. Krox20 stimulates adipogenesis via C/EBPbeta-dependent and -independent mechanisms. *Cell Metab*, 1, 93-106.
- Cheng, C. K., Li, L., Cheng, S. H., Lau, K. M., Chan, N. P., Wong, R. S., *et al.* **2008**. Transcriptional repression of the RUNX3/AML2 gene by the t(8;21) and inv(16) fusion proteins in acute myeloid leukemia. *Blood*, 112, 3391-402.
- Cheng, H., Hao, S., Liu, Y., Pang, Y., Ma, S., Dong, F., *et al.* **2015**. Leukemic marrow infiltration reveals a novel role for Egr3 as a potent inhibitor of normal hematopoietic stem cell proliferation. *Blood*, 126, 1302-13.
- Chinery, R., Brockman, J. A., Dransfield, D. T. & Coffey, R. J. **1997**. Antioxidant-induced nuclear translocation of CCAAT/enhancer-binding protein beta. A critical role for protein kinase A-mediated phosphorylation of Ser299. *J Biol Chem*, 272, 30356-61.
- Choi, Y., Elagib, K. E., Delehanty, L. L. & Goldfarb, A. N. **2006**. Erythroid inhibition by the leukemic fusion AML1-ETO is associated with impaired acetylation of the major erythroid transcription factor GATA-1. *Cancer Res*, 66, 2990-6.
- Chou, F. S., Griesinger, A., Wunderlich, M., Lin, S., Link, K. A., Shrestha, M., *et al.* **2012**. The thrombopoietin/MPL/Bcl-xL pathway is essential for survival and self-renewal in human preleukemia induced by AML1-ETO. *Blood*, 120, 709-19.
- Christensen, J. L. & Weissman, I. L. **2001**. Flk-2 is a marker in hematopoietic stem cell differentiation: a simple method to isolate long-term stem cells. *Proc Natl Acad Sci U S A*, 98, 14541-6.
- Christiansen, D. H., Andersen, M. K. & Pedersen-Bjergaard, J. **2004**. Mutations of AML1 are common in therapy-related myelodysplasia following therapy with alkylating agents and are significantly associated with deletion or loss of chromosome arm 7q and with subsequent leukemic transformation. *Blood*, 104, 1474-81.
- Chuang, L. S., Ito, K. & Ito, Y. **2013**. RUNX family: Regulation and diversification of roles through interacting proteins. *Int J Cancer*, 132, 1260-71.

- Cirillo, E., Parnell, L. D. & Evelo, C. T. **2017**. A Review of Pathway-Based Analysis Tools That Visualize Genetic Variants. *Front Genet*, 8, 174.
- Civin, C. I., Strauss, L. C., Brovall, C., Fackler, M. J., Schwartz, J. F. & Shaper, J. H. **1984**. Antigenic analysis of hematopoiesis. III. A hematopoietic progenitor cell surface antigen defined by a monoclonal antibody raised against KG-1a cells. *J Immunol*, 133, 157-65.
- Cohen, P. A., Donini, C. F., Nguyen, N. T., Lincet, H. & Vendrell, J. A. **2015**. The dark side of ZNF217, a key regulator of tumorigenesis with powerful biomarker value. *Oncotarget*, 6, 41566-81.
- Cohen, S. **1983**. The epidermal growth factor (EGF). *Cancer*, 51, 1787-91.
- Collins, B. C., Hunter, C. L., Liu, Y., Schilling, B., Rosenberger, G., Bader, S. L., *et al.* **2017**. Multi-laboratory assessment of reproducibility, qualitative and quantitative performance of SWATH-mass spectrometry. *Nat Commun*, 8, 291.
- Collins, C., Huang, G. & Gray, J. W. **2007**. *Potentiation of cancer therapies by ZNF217 inhibition*. USA patent application US 2007/0037202 A1.
- Collins, C., Rommens, J. M., Kowbel, D., Godfrey, T., Tanner, M., Hwang, S. I., *et al.* **1998**. Positional cloning of ZNF217 and NABC1: genes amplified at 20q13.2 and overexpressed in breast carcinoma. *Proc Natl Acad Sci U S A*, 95, 8703-8.
- Combates, N. J., Rzepka, R. W., Chen, Y. N. & Cohen, D. **1994**. NF-IL6, a member of the C/EBP family of transcription factors, binds and trans-activates the human MDR1 gene promoter. *J Biol Chem*, 269, 29715-9.
- Condamine, T., Mastio, J. & Gabrilovich, D. I. **2015**. Transcriptional regulation of myeloid-derived suppressor cells. *J Leukoc Biol*, 98, 913-22.
- Conneally, E., Cashman, J., Petzer, A. & Eaves, C. **1997**. Expansion in vitro of transplantable human cord blood stem cells demonstrated using a quantitative assay of their lympho-myeloid repopulating activity in nonobese diabetic-scid/scid mice. *Proc Natl Acad Sci U S A*, 94, 9836-41.
- Conway O'brien, E., Prideaux, S. & Chevassut, T. **2014**. The epigenetic landscape of acute myeloid leukemia. *Adv Hematol*, 2014, 103175.
- Conze, D., Weiss, L., Regen, P. S., Bhushan, A., Weaver, D., Johnson, P., *et al.* **2001**. Autocrine production of interleukin 6 causes multidrug resistance in breast cancer cells. *Cancer Res*, 61, 8851-8.
- Cook, A. M., Li, L., Ho, Y., Lin, A., Li, L., Stein, A., *et al.* **2014**. Role of altered growth factor receptor-mediated JAK2 signaling in growth and maintenance of human acute myeloid leukemia stem cells. *Blood*, 123, 2826-2837.
- Cooper, L. A., Gutman, D. A., Chisolm, C., Appin, C., Kong, J., Rong, Y., *et al.* **2012**. The tumor microenvironment strongly impacts master transcriptional regulators and gene expression class of glioblastoma. *Am J Pathol*, 180, 2108-19.
- Cooray, S., Howe, S. J. & Thrasher, A. J. **2012**. Retrovirus and lentivirus vector design and methods of cell conditioning. *Methods Enzymol*, 507, 29-57.
- Costoya, J. A. & Pandolfi, P. P. **2001**. The role of promyelocytic leukemia zinc finger and promyelocytic leukemia in leukemogenesis and development. *Curr Opin Hematol*, 8, 212-7.
- Cowger, J. J., Zhao, Q., Iovic, M. & Torchia, J. **2007**. Biochemical characterization of the zinc-finger protein 217 transcriptional repressor complex: identification of a ZNF217 consensus recognition sequence. *Oncogene*, 26, 3378-86.

- Croniger, C. M., Millward, C., Yang, J., Kawai, Y., Arinze, I. J., Liu, S., *et al.* **2001**. Mice with a deletion in the gene for CCAAT/enhancer-binding protein beta have an attenuated response to cAMP and impaired carbohydrate metabolism. *J Biol Chem*, 276, 629-38.
- Crump, M., Hedley, D., Kamel-Reid, S., Leber, B., Wells, R., Brandwein, J., *et al.* **2010**. A randomized phase I clinical and biologic study of two schedules of sorafenib in patients with myelodysplastic syndrome or acute myeloid leukemia: a NCIC (National Cancer Institute of Canada) Clinical Trials Group Study. *Leuk Lymphoma*, 51, 252-60.
- Dalgin, G. S., Holloway, D. T., Liou, L. S. & Delisi, C. **2007**. Identification and characterization of renal cell carcinoma gene markers. *Cancer Inform*, 3, 65-92.
- Damiani, D., Tiribelli, M., Raspadori, D., Sirianni, S., Meneghel, A., Cavallin, M., *et al.* **2015**. Clinical impact of CD200 expression in patients with acute myeloid leukemia and correlation with other molecular prognostic factors. *Oncotarget*, 6, 30212-21.
- Daver, N., Cortes, J., Ravandi, F., Patel, K. P., Burger, J. A., Konopleva, M., *et al.* **2015**. Secondary mutations as mediators of resistance to targeted therapy in leukemia. *Blood*, 125, 3236-45.
- Davis, J. N., Mcghee, L. & Meyers, S. **2003**. The ETO (MTG8) gene family. *Gene*, 303, 1-10.
- Davis, J. N., Williams, B. J., Herron, J. T., Galiano, F. J. & Meyers, S. **1999**. ETO-2, a new member of the ETO-family of nuclear proteins. *Oncogene*, 18, 1375-83.
- De Braekeleer, E., Douet-Guilbert, N. & De Braekeleer, M. **2014**. RARA fusion genes in acute promyelocytic leukemia: a review. *Expert Rev Hematol*, 7, 347-57.
- De Braekeleer, E., Douet-Guilbert, N., Morel, F., Le Bris, M. J., Férec, C. & De Braekeleer, M. **2011**. RUNX1 translocations and fusion genes in malignant hemopathies. *Future Oncol*, 7, 77-91.
- De Kouchkovsky, I. & Abdul-Hay, M. **2016**. 'Acute myeloid leukemia: a comprehensive review and 2016 update'. *Blood Cancer J*, 6, e441.
- De Sauvage, F. J., Carver-Moore, K., Luoh, S. M., Ryan, A., Dowd, M., Eaton, D. L., *et al.* **1996**. Physiological regulation of early and late stages of megakaryocytopoiesis by thrombopoietin. *J Exp Med*, 183, 651-6.
- Deb, G. & Somervaille, T. C. P. **2016**. Antagonizing Self-Renewal in Acute Myeloid Leukemia: ID2 Takes the Stage. *Cancer Cell*, 30, 5-7.
- Dekelver, R. C., Lewin, B., Lam, K., Komeno, Y., Yan, M., Rundle, C., *et al.* **2013a**. Cooperation between RUNX1-ETO9a and novel transcriptional partner KLF6 in upregulation of Alox5 in acute myeloid leukemia. *PLoS Genet*, 9, e1003765.
- Dekelver, R. C., Yan, M., Ahn, E. Y., Shia, W. J., Speck, N. A. & Zhang, D. E. **2013b**. Attenuation of AML1-ETO cellular dysregulation correlates with increased leukemogenic potential. *Blood*, 121, 3714-7.
- Descombes, P. & Schibler, U. **1991**. A liver-enriched transcriptional activator protein, LAP, and a transcriptional inhibitory protein, LIP, are translated from the same mRNA. *Cell*, 67, 569-79.
- Diehl, A. M. **1998**. Roles of CCAAT/enhancer-binding proteins in regulation of liver regenerative growth. *J Biol Chem*, 273, 30843-6.
- Dinardo, C. D., Pratz, K., Pullarkat, V., Jonas, B. A., Arellano, M., Becker, P. S., *et al.* **2019**. Venetoclax combined with decitabine or azacitidine in treatment-naive, elderly patients with acute myeloid leukemia. *Blood*, 133, 7-17.

- Ding, L. & Morrison, S. J. **2013**. Haematopoietic stem cells and early lymphoid progenitors occupy distinct bone marrow niches. *Nature*, 495, 231-5.
- Ding, L., Saunders, T. L., Enikolopov, G. & Morrison, S. J. **2012**. Endothelial and perivascular cells maintain haematopoietic stem cells. *Nature*, 481, 457-62.
- Döhner, H., Estey, E., Grimwade, D., Amadori, S., Appelbaum, F. R., Büchner, T., *et al.* **2017**. Diagnosis and management of AML in adults: 2017 ELN recommendations from an international expert panel. *Blood*, 129, 424-447.
- Döhner, H., Weisdorf, D. J. & Bloomfield, C. D. **2015**. Acute Myeloid Leukemia. *N Engl J Med*, 373, 1136-52.
- Döhner, K., Schlenk, R. F., Habdank, M., Scholl, C., Rücker, F. G., Corbacioglu, A., *et al.* **2005**. Mutant nucleophosmin (NPM1) predicts favorable prognosis in younger adults with acute myeloid leukemia and normal cytogenetics: interaction with other gene mutations. *Blood*, 106, 3740-6.
- Dombret, H., Seymour, J. F., Butrym, A., Wierzbowska, A., Selleslag, D., Jang, J. H., *et al.* **2015**. International phase 3 study of azacitidine vs conventional care regimens in older patients with newly diagnosed AML with >30% blasts. *Blood*, 126, 291-9.
- Dowdy, C. R., Xie, R., Frederick, D., Hussain, S., Zaidi, S. K., Vradii, D., *et al.* **2010**. Definitive hematopoiesis requires Runx1 C-terminal-mediated subnuclear targeting and transactivation. *Hum Mol Genet*, 19, 1048-57.
- Downing, J. R. **2001**. AML1/CBFbeta transcription complex: its role in normal hematopoiesis and leukemia. *Leukemia*, 15, 664-5.
- Draper, J. E., Sroczyńska, P., Tsoulaki, O., Leong, H. S., Fadlullah, M. Z., Miller, C., *et al.* **2016**. RUNX1B Expression Is Highly Heterogeneous and Distinguishes Megakaryocytic and Erythroid Lineage Fate in Adult Mouse Hematopoiesis. *PLoS Genet*, 12, e1005814.
- Driggers, P. H., Ennist, D. L., Gleason, S. L., Mak, W. H., Marks, M. S., Levi, B. Z., *et al.* **1990**. An interferon gamma-regulated protein that binds the interferon-inducible enhancer element of major histocompatibility complex class I genes. *Proc Natl Acad Sci U S A*, 87, 3743-7.
- Du, Y. X., Zhang, L. H., Wang, X. H., Xing, X. F., Cheng, X. J., Du, H., *et al.* **2013**. [Expression of CCAAT/enhancer binding protein beta in human gastric carcinoma and its clinical significance]. *Zhonghua Wei Chang Wai Ke Za Zhi*, 16, 179-82.
- Dufait, I., Liechtenstein, T., Lanna, A., Bricogne, C., Laranga, R., Padella, A., *et al.* **2012**. Retroviral and lentiviral vectors for the induction of immunological tolerance. *Scientifica (Cairo)*, 2012.
- Duprez, E. **2004**. A new role for C/EBPbeta in acute promyelocytic leukemia. *Cell Cycle*, 3, 389-90.
- Duprez, E., Wagner, K., Koch, H. & Tenen, D. G. **2003**. C/EBPbeta: a major PML-RARA-responsive gene in retinoic acid-induced differentiation of APL cells. *EMBO J*, 22, 5806-16.
- Duque-Afonso, J., Lübbert, M. & Cleary, M. L. **2014**. Epigenetic Modifications Mediated by the AML1/ETO and MLL Leukemia Fusion Proteins. In: Lübbert, M. & Jones, P. A. (eds.) *Epigenetic Therapy of Cancer: Preclinical Models and Treatment Approaches*. Berlin, Heidelberg: Springer Berlin Heidelberg.
- Eder, M., Schoenherr, C., Wohlan, K., Frömling, F., Hilfiker-Kleiner, D., Ganser, A., *et al.* **2016**. Stable Silencing of RUNX1/ETO Induces Expression of a Shortened PU.1 Variant in t(8;21) Kasumi Cells. *Blood*, 128, 2716-2716.
- Egawa, T., Tillman, R. E., Naoe, Y., Taniuchi, I. & Littman, D. R. **2007**. The role of the Runx transcription factors in thymocyte differentiation and in homeostasis of naive T cells. *J Exp Med*, 204, 1945-57.

- Elagib, K. E., Racke, F. K., Mogass, M., Khetawat, R., Delehanty, L. L. & Goldfarb, A. N. **2003**. RUNX1 and GATA-1 coexpression and cooperation in megakaryocytic differentiation. *Blood*, 101, 4333-41.
- Emerenciano, M., Barbosa, T. C., Lopes, B. A., Blunck, C. B., Faro, A., Andrade, C., *et al.* **2014**. ARID5B polymorphism confers an increased risk to acquire specific MLL rearrangements in early childhood leukemia. *BMC Cancer*, 14, 127.
- Erb, H. H., Langlechner, R. V., Moser, P. L., Handle, F., Casneuf, T., Verstraeten, K., *et al.* **2013**. IL6 sensitizes prostate cancer to the antiproliferative effect of IFN $\alpha$ 2 through IRF9. *Endocr Relat Cancer*, 20, 677-89.
- Erickson, P., Gao, J., Chang, K. S., Look, T., Whisenant, E., Raimondi, S., *et al.* **1992**. Identification of breakpoints in t(8;21) acute myelogenous leukemia and isolation of a fusion transcript, AML1/ETO, with similarity to Drosophila segmentation gene, runt. *Blood*, 80, 1825-31.
- Erickson, P. F., Dessev, G., Lasher, R. S., Philips, G., Robinson, M. & Drabkin, H. A. **1996**. ETO and AML1 phosphoproteins are expressed in CD34+ hematopoietic progenitors: implications for t(8;21) leukemogenesis and monitoring residual disease. *Blood*, 88, 1813-23.
- Erickson, P. F., Robinson, M., Owens, G. & Drabkin, H. A. **1994**. The ETO portion of acute myeloid leukemia t(8;21) fusion transcript encodes a highly evolutionarily conserved, putative transcription factor. *Cancer Res*, 54, 1782-6.
- Ernst, O. & Zor, T. **2010**. Linearization of the Bradford protein assay. *J Vis Exp*.
- Essers, M. A., Offner, S., Blanco-Bose, W. E., Waibler, Z., Kalinke, U., Duchosal, M. A., *et al.* **2009**. IFN $\alpha$  activates dormant haematopoietic stem cells in vivo. *Nature*, 458, 904-8.
- Etcheverry, A., Aubry, M., De Tairac, M., Vauleon, E., Boniface, R., Guenot, F., *et al.* **2010**. DNA methylation in glioblastoma: impact on gene expression and clinical outcome. *BMC Genomics*, 11, 701.
- Ewing, S. J., Zhu, S., Zhu, F., House, J. S. & Smart, R. C. **2008**. C/EBP $\beta$  represses p53 to promote cell survival downstream of DNA damage independent of oncogenic Ras and p19(Arf). *Cell Death Differ*, 15, 1734-44.
- Fang, Z., Xiong, Y., Zhang, C., Li, J., Liu, L., Li, M., *et al.* **2010**. Coexistence of copy number increases of ZNF217 and CYP24A1 in colorectal cancers in a Chinese population. *Oncol Lett*, 1, 925-930.
- Fathi, A. T., Blonquist, T. M., Hernandez, D., Amrein, P. C., Ballen, K. K., McMasters, M., *et al.* **2018**. Cabozantinib is well tolerated in acute myeloid leukemia and effectively inhibits the resistance-conferring FLT3/tyrosine kinase domain/F691 mutation. *Cancer*, 124, 306-314.
- Fernando, T. R., Contreras, J. R., Zampini, M., Rodriguez-Malave, N. I., Alberti, M. O., Anguiano, J., *et al.* **2017**. The lncRNA CASC15 regulates SOX4 expression in RUNX1-rearranged acute leukemia. *Mol Cancer*, 16, 126.
- Ferrara, F. & Schiffer, C. A. **2013**. Acute myeloid leukaemia in adults. *Lancet*. England: 2013 Elsevier Ltd.
- Feuring-Buske, M., Gerhard, B., Cashman, J., Humphries, R. K., Eaves, C. J. & Hogge, D. E. **2003**. Improved engraftment of human acute myeloid leukemia progenitor cells in beta 2-microglobulin-deficient NOD/SCID mice and in NOD/SCID mice transgenic for human growth factors. *Leukemia*, 17, 760-3.
- Fialkow, P. J. **1974**. The origin and development of human tumors studied with cell markers. *N Engl J Med*, 291, 26-35.
- Finak, G., Bertos, N., Pepin, F., Sadekova, S., Souleimanova, M., Zhao, H., *et al.* **2008**. Stromal gene expression predicts clinical outcome in breast cancer. *Nat Med*, 14, 518-27.

- Fischer, T., Stone, R. M., Deangelo, D. J., Galinsky, I., Estey, E., Lanza, C., *et al.* **2010**. Phase IIB trial of oral Midostaurin (PKC412), the FMS-like tyrosine kinase 3 receptor (FLT3) and multi-targeted kinase inhibitor, in patients with acute myeloid leukemia and high-risk myelodysplastic syndrome with either wild-type or mutated FLT3. *J Clin Oncol*, 28, 4339-45.
- Foka, P., Kousteni, S. & Ramji, D. P. **2001**. Molecular characterization of the Xenopus CCAAT-enhancer binding protein beta gene promoter. *Biochem Biophys Res Commun*, 285, 430-6.
- Forsberg, E. C., Serwold, T., Kogan, S., Weissman, I. L. & Passegué, E. **2006**. New evidence supporting megakaryocyte-erythrocyte potential of flk2/flt3+ multipotent hematopoietic progenitors. *Cell*, 126, 415-26.
- Foudi, A., Hochedlinger, K., Van Buren, D., Schindler, J. W., Jaenisch, R., Carey, V., *et al.* **2009**. Analysis of histone 2B-GFP retention reveals slowly cycling hematopoietic stem cells. *Nat Biotechnol*, 27, 84-90.
- Fox, N., Priestley, G., Papayannopoulou, T. & Kaushansky, K. **2002**. Thrombopoietin expands hematopoietic stem cells after transplantation. *J Clin Invest*, 110, 389-94.
- Fracchiolla, N. S., Colombo, G., Finelli, P., Maiolo, A. T. & Neri, A. **1998**. EHT, a new member of the MTG8/ETO gene family, maps on 20q11 region and is deleted in acute myeloid leukemias. *Blood*, 92, 3481-4.
- Friedman, A. D. **2007**. Transcriptional control of granulocyte and monocyte development. *Oncogene*, 26, 6816-28.
- Fukamachi, H. **2006**. Runx3 controls growth and differentiation of gastric epithelial cells in mammals. *Dev Growth Differ*, 48, 1-13.
- Fukuda, T. **1973**. Fetal hemopoiesis. I. Electron microscopic studies on human yolk sac hemopoiesis. *Virchows Arch B Cell Pathol*, 14, 197-213.
- Gaidzik, V. I., Bullinger, L., Schlenk, R. F., Zimmermann, A. S., Röck, J., Paschka, P., *et al.* **2011**. RUNX1 mutations in acute myeloid leukemia: results from a comprehensive genetic and clinical analysis from the AML study group. *J Clin Oncol*, 29, 1364-72.
- Galán-Díez, M., Cuesta-Domínguez, Á. & Kousteni, S. **2018**. The Bone Marrow Microenvironment in Health and Myeloid Malignancy. *Cold Spring Harb Perspect Med*, 8.
- Galanis, A., Ma, H., Rajkhowa, T., Ramachandran, A., Small, D., Cortes, J., *et al.* **2014**. Crenolanib is a potent inhibitor of FLT3 with activity against resistance-conferring point mutants. *Blood*, 123, 94-100.
- Gamou, T., Kitamura, E., Hosoda, F., Shimizu, K., Shinohara, K., Hayashi, Y., *et al.* **1998**. The partner gene of AML1 in t(16;21) myeloid malignancies is a novel member of the MTG8(ETO) family. *Blood*, 91, 4028-37.
- Gao, J., Aksoy, B. A., Dogrusoz, U., Dresdner, G., Gross, B., Sumer, S. O., *et al.* **2013**. Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. *Sci Signal*, 6, p11.
- Gao, X. N., Yan, F., Lin, J., Gao, L., Lu, X. L., Wei, S. C., *et al.* **2015**. AML1/ETO cooperates with HIF1 $\alpha$  to promote leukemogenesis through DNMT3a transactivation. *Leukemia*, 29, 1730-40.
- Gardini, A., Cesaroni, M., Luzi, L., Okumura, A. J., Biggs, J. R., Minardi, S. P., *et al.* **2008**. AML1/ETO oncoprotein is directed to AML1 binding regions and co-localizes with AML1 and HEB on its targets. *PLoS Genet*, 4, e1000275.
- Gelmetti, V., Zhang, J., Fanelli, M., Minucci, S., Pelicci, P. G. & Lazar, M. A. **1998**. Aberrant recruitment of the nuclear receptor corepressor-histone deacetylase complex by the acute myeloid leukemia fusion partner ETO. *Mol Cell Biol*, 18, 7185-91.

- Gergen, J. P. & Butler, B. A. **1988**. Isolation of the *Drosophila* segmentation gene *runt* and analysis of its expression during embryogenesis. *Genes Dev*, 2, 1179-93.
- Gery, S., Tanosaki, S., Hofmann, W. K., Koppel, A. & Koeffler, H. P. **2005**. C/EBPdelta expression in a BCR-ABL-positive cell line induces growth arrest and myeloid differentiation. *Oncogene*, 24, 1589-97.
- Ghoshal Gupta, S., Baumann, H. & Wetzler, M. **2008**. Epigenetic regulation of signal transducer and activator of transcription 3 in acute myeloid leukemia. *Leuk Res*, 32, 1005-14.
- Gill, S., Tasian, S. K., Ruella, M., Shestova, O., Li, Y., Porter, D. L., *et al.* **2014**. Preclinical targeting of human acute myeloid leukemia and myeloablation using chimeric antigen receptor-modified T cells. *Blood*, 123, 2343-54.
- Gilliland, D. G. & Griffin, J. D. **2002**. The roles of FLT3 in hematopoiesis and leukemia. *Blood*, 100, 1532-42.
- Ginzinger, D. G., Godfrey, T. E., Nigro, J., Moore, D. H., Suzuki, S., Pallavicini, M. G., *et al.* **2000**. Measurement of DNA copy number at microsatellite loci using quantitative PCR analysis. *Cancer Res*, 60, 5405-9.
- Golub, T. R., Barker, G. F., Bohlander, S. K., Hiebert, S. W., Ward, D. C., Bray-Ward, P., *et al.* **1995**. Fusion of the TEL gene on 12p13 to the AML1 gene on 21q22 in acute lymphoblastic leukemia. *Proc Natl Acad Sci U S A*, 92, 4917-21.
- Gomis, R. R., Alarcón, C., Nadal, C., Van Poznak, C. & Massagué, J. **2006**. C/EBPbeta at the core of the TGFbeta cytostatic response and its evasion in metastatic breast cancer cells. *Cancer Cell*, 10, 203-14.
- Goode, D. K., Obier, N., Vijayabaskar, M. S., Lie-a-Ling, M., Lilly, A. J., Hannah, R., *et al.* **2016**. Dynamic Gene Regulatory Networks Drive Hematopoietic Specification and Differentiation. *Dev Cell*, 36, 572-87.
- Gorgoni, B., Maritano, D., Marthyn, P., Righi, M. & Poli, V. **2002**. C/EBP beta gene inactivation causes both impaired and enhanced gene expression and inverse regulation of IL-12 p40 and p35 mRNAs in macrophages. *J Immunol*, 168, 4055-62.
- Gotoh, M., Ichikawa, H., Arai, E., Chiku, S., Sakamoto, H., Fujimoto, H., *et al.* **2014**. Comprehensive exploration of novel chimeric transcripts in clear cell renal cell carcinomas using whole transcriptome analysis. *Genes Chromosomes Cancer*, 53, 1018-32.
- Goyama, S., Huang, G., Kurokawa, M. & Mulloy, J. C. **2015**. Posttranslational modifications of RUNX1 as potential anticancer targets. *Oncogene*, 34, 3483-92.
- Goyama, S. & Mulloy, J. C. **2011**. Molecular pathogenesis of core binding factor leukemia: current knowledge and future prospects. *Int J Hematol*, 94, 126-133.
- Goyama, S., Schibler, J., Cunningham, L., Zhang, Y., Rao, Y., Nishimoto, N., *et al.* **2013**. Transcription factor RUNX1 promotes survival of acute myeloid leukemia cells. *J Clin Invest*, 123, 3876-88.
- Goyama, S., Schibler, J., Gasilina, A., Shrestha, M., Lin, S., Link, K. A., *et al.* **2016**. UBASH3B/Sts-1-CBL axis regulates myeloid proliferation in human preleukemia induced by AML1-ETO. *Leukemia*, 30, 728-39.
- Grassinger, J., Haylock, D. N., Storan, M. J., Haines, G. O., Williams, B., Whitty, G. A., *et al.* **2009**. Thrombin-cleaved osteopontin regulates hemopoietic stem and progenitor cell functions through interactions with alpha9beta1 and alpha4beta1 integrins. *Blood*, 114, 49-59.
- Grassinger, J., Haylock, D. N., Williams, B., Olsen, G. H. & Nilsson, S. K. **2010**. Phenotypically identical hemopoietic stem cells isolated from different regions of bone marrow have different biologic potential. *Blood*, 116, 3185-96.

- Gray, S. G., Kytölä, S., Matsunaga, T., Larsson, C. & Ekström, T. J. **2000**. Comparative genomic hybridization reveals population-based genetic alterations in hepatoblastomas. *Br J Cancer*, 83, 1020-5.
- Greenbaum, L. E., Li, W., Cressman, D. E., Peng, Y., Ciliberto, G., Poli, V., *et al.* **1998**. CCAAT enhancer-binding protein beta is required for normal hepatocyte proliferation in mice after partial hepatectomy. *J Clin Invest*, 102, 996-1007.
- Greif, P. A., Konstandin, N. P., Metzeler, K. H., Herold, T., Pasalic, Z., Ksienzyk, B., *et al.* **2012**. RUNX1 mutations in cytogenetically normal acute myeloid leukemia are associated with a poor prognosis and up-regulation of lymphoid genes. *Haematologica*, 97, 1909-15.
- Griffin, J. D. & Löwenberg, B. **1986**. Clonogenic cells in acute myeloblastic leukemia. *Blood*, 68, 1185-95.
- Grignani, F., Kinsella, T., Mencarelli, A., Valtieri, M., Riganelli, D., Lanfrancone, L., *et al.* **1998**. High-efficiency gene transfer and selection of human hematopoietic progenitor cells with a hybrid EBV/retroviral vector expressing the green fluorescence protein. *Cancer Res*, 58, 14-9.
- Grimes, H. L., Chan, T. O., Zweidler-Mckay, P. A., Tong, B. & Tsichlis, P. N. **1996a**. The Gfi-1 proto-oncoprotein contains a novel transcriptional repressor domain, SNAG, and inhibits G1 arrest induced by interleukin-2 withdrawal. *Mol Cell Biol*, 16, 6263-72.
- Grimes, H. L., Gilks, C. B., Chan, T. O., Porter, S. & Tsichlis, P. N. **1996b**. The Gfi-1 protooncoprotein represses Bax expression and inhibits T-cell death. *Proc Natl Acad Sci U S A*, 93, 14569-73.
- Grimwade, D., Walker, H., Harrison, G., Oliver, F., Chatters, S., Harrison, C. J., *et al.* **2001**. The predictive value of hierarchical cytogenetic classification in older adults with acute myeloid leukemia (AML): analysis of 1065 patients entered into the United Kingdom Medical Research Council AML11 trial. *Blood*, 98, 1312-20.
- Grisolano, J. L., O'neal, J., Cain, J. & Tomasson, M. H. **2003**. An activated receptor tyrosine kinase, TEL/PDGFBetaR, cooperates with AML1/ETO to induce acute myeloid leukemia in mice. *Proc Natl Acad Sci U S A*, 100, 9506-11.
- Grossmann, V., Bacher, U., Kohlmann, A., Butschalowski, K., Roller, A., Jeromin, S., *et al.* **2012**. Expression of CEBPA is reduced in RUNX1-mutated acute myeloid leukemia. *Blood Cancer J*, 2, e86.
- Grove, C. S. & Vassiliou, G. S. **2014**. Acute myeloid leukaemia: a paradigm for the clonal evolution of cancer? *Dis Model Mech*, 7, 941-51.
- Grover, A., Mancini, E., Moore, S., Mead, A. J., Atkinson, D., Rasmussen, K. D., *et al.* **2014**. Erythropoietin guides multipotent hematopoietic progenitor cells toward an erythroid fate. *J Exp Med*, 211, 181-8.
- Growney, J. D., Shigematsu, H., Li, Z., Lee, B. H., Adelsperger, J., Rowan, R., *et al.* **2005**. Loss of Runx1 perturbs adult hematopoiesis and is associated with a myeloproliferative phenotype. *Blood*, 106, 494-504.
- Gu, T. L., Goetz, T. L., Graves, B. J. & Speck, N. A. **2000**. Auto-inhibition and partner proteins, core-binding factor beta (CBFbeta) and Ets-1, modulate DNA binding by CBFalpha2 (AML1). *Mol Cell Biol*, 20, 91-103.
- Gu, X., Hu, Z., Ebrahim, Q., Crabb, J. S., Mahfouz, R. Z., Radivoyevitch, T., *et al.* **2014**. Runx1 regulation of Pu.1 corepressor/coactivator exchange identifies specific molecular targets for leukemia differentiation therapy. *J Biol Chem*, 289, 14881-95.
- Guenechea, G., Gan, O. I., Dorrell, C. & Dick, J. E. **2001**. Distinct classes of human stem cells that differ in proliferative and self-renewal potential. *Nat Immunol*, 2, 75-82.
- Guerzoni, C., Bardini, M., Mariani, S. A., Ferrari-Amorotti, G., Neviani, P., Panno, M. L., *et al.* **2006**. Inducible activation of CEBPB, a gene negatively regulated by BCR/ABL, inhibits proliferation and promotes differentiation of BCR/ABL-expressing cells. *Blood*, 107, 4080-9.

- Guidez, F., Petrie, K., Ford, A. M., Lu, H., Bennett, C. A., Macgregor, A., *et al.* **2000**. Recruitment of the nuclear receptor corepressor N-CoR by the TEL moiety of the childhood leukemia-associated TEL-AML1 oncoprotein. *Blood*, 96, 2557-61.
- Guilliams, M., Mildner, A. & Yona, S. **2018**. Developmental and Functional Heterogeneity of Monocytes. *Immunity*, 49, 595-613.
- Guo, H. & Friedman, A. D. **2011**. Phosphorylation of RUNX1 by cyclin-dependent kinase reduces direct interaction with HDAC1 and HDAC3. *J Biol Chem*, 286, 208-15.
- Guo, S., Cichy, S. B., He, X., Yang, Q., Ragland, M., Ghosh, A. K., *et al.* **2001**. Insulin suppresses transactivation by CAAT/enhancer-binding proteins beta (C/EBPbeta). Signaling to p300/CREB-binding protein by protein kinase B disrupts interaction with the major activation domain of C/EBPbeta. *J Biol Chem*, 276, 8516-23.
- Gurney, A. L., Carver-Moore, K., De Sauvage, F. J. & Moore, M. W. **1994**. Thrombocytopenia in c-mpl-deficient mice. *Science*, 265, 1445-7.
- Gutsch, R., Kandemir, J. D., Pietsch, D., Cappello, C., Meyer, J., Simanowski, K., *et al.* **2011**. CCAAT/enhancer-binding protein beta inhibits proliferation in monocytic cells by affecting the retinoblastoma protein/E2F/cyclin E pathway but is not directly required for macrophage morphology. *J Biol Chem*, 286, 22716-29.
- Guzman, M. L., Neering, S. J., Upchurch, D., Grimes, B., Howard, D. S., Rizzieri, D. A., *et al.* **2001**. Nuclear factor-kappaB is constitutively activated in primitive human acute myelogenous leukemia cells. *Blood*, 98, 2301-7.
- Haas, S. C., Huber, R., Gutsch, R., Kandemir, J. D., Cappello, C., Krauter, J., *et al.* **2010**. ITD- and FL-induced FLT3 signal transduction leads to increased C/EBPbeta-LIP expression and LIP/LAP ratio by different signalling modules. *Br J Haematol*, 148, 777-90.
- Haferlach, T., Kohlmann, A., Wiczorek, L., Basso, G., Kronnie, G. T., Béné, M. C., *et al.* **2010**. Clinical utility of microarray-based gene expression profiling in the diagnosis and subclassification of leukemia: report from the International Microarray Innovations in Leukemia Study Group. *J Clin Oncol*, 28, 2529-37.
- Hakimi, M. A., Dong, Y., Lane, W. S., Speicher, D. W. & Shiekhatar, R. **2003**. A candidate X-linked mental retardation gene is a component of a new family of histone deacetylase-containing complexes. *J Biol Chem*, 278, 7234-9.
- Hall, C. J., Flores, M. V., Oehlers, S. H., Sanderson, L. E., Lam, E. Y., Crosier, K. E., *et al.* **2012**. Infection-responsive expansion of the hematopoietic stem and progenitor cell compartment in zebrafish is dependent upon inducible nitric oxide. *Cell Stem Cell*, 10, 198-209.
- Hanahan, D. & Weinberg, R. A. **2000**. The hallmarks of cancer. *Cell*, 100, 57-70.
- Hanahan, D. & Weinberg, Robert a. **2011**. Hallmarks of Cancer: The Next Generation. *Cell*, 144, 646-674.
- Hanekamp, D., Cloos, J. & Schuurhuis, G. J. **2017**. Leukemic stem cells: identification and clinical application. *Int J Hematol*, 105, 549-557.
- Harada, H., Harada, Y., Niimi, H., Kyo, T., Kimura, A. & Inaba, T. **2004**. High incidence of somatic mutations in the AML1/RUNX1 gene in myelodysplastic syndrome and low blast percentage myeloid leukemia with myelodysplasia. *Blood*, 103, 2316-24.
- Harada, H., Harada, Y., Tanaka, H., Kimura, A. & Inaba, T. **2003**. Implications of somatic mutations in the AML1 gene in radiation-associated and therapy-related myelodysplastic syndrome/acute myeloid leukemia. *Blood*, 101, 673-80.

- Hartman, M., Piliponsky, A. M., Temkin, V. & Levi-Schaffer, F. **2001**. Human peripheral blood eosinophils express stem cell factor. *Blood*, 97, 1086-91.
- Hattori, T., Ohoka, N., Inoue, Y., Hayashi, H. & Onozaki, K. **2003**. C/EBP family transcription factors are degraded by the proteasome but stabilized by forming dimer. *Oncogene*, 22, 1273-80.
- Hayashi, Y., Hirai, H., Kamio, N., Yao, H., Yoshioka, S., Miura, Y., *et al.* **2013**. C/EBP $\beta$  promotes BCR-ABL-mediated myeloid expansion and leukemic stem cell exhaustion. *Leukemia*, 27, 619-28.
- Haylock, D. N., Williams, B., Johnston, H. M., Liu, M. C., Rutherford, K. E., Whitty, G. A., *et al.* **2007**. Hemopoietic stem cells with higher hemopoietic potential reside at the bone marrow endosteum. *Stem Cells*, 25, 1062-9.
- Heavey, B., Charalambous, C., Cobaleda, C. & Busslinger, M. **2003**. Myeloid lineage switch of Pax5 mutant but not wild-type B cell progenitors by C/EBP $\alpha$  and GATA factors. *EMBO J*, 22, 3887-97.
- Herglotz, J., Kuvardina, O. N., Kolodziej, S., Kumar, A., Hussong, H., Grez, M., *et al.* **2013**. Histone arginine methylation keeps RUNX1 target genes in an intermediate state. *Oncogene*, 32, 2565-75.
- Hernandez, J. M., Floyd, D. H., Weilbaecher, K. N., Green, P. L. & Boris-Lawrie, K. **2008**. Multiple facets of junD gene expression are atypical among AP-1 family members. *Oncogene*, 27, 4757-67.
- Hess, J. L., Bittner, C. B., Zeisig, D. T., Bach, C., Fuchs, U., Borkhardt, A., *et al.* **2006**. c-Myb is an essential downstream target for homeobox-mediated transformation of hematopoietic cells. *Blood*, 108, 297-304.
- Hiebert, S. W., Lutterbach, B. & Amann, J. **2001**. Role of co-repressors in transcriptional repression mediated by the t(8;21), t(16;21), t(12;21), and inv(16) fusion proteins. *Curr Opin Hematol*, 8, 197-200.
- Higuchi, M., O'Brien, D., Kumaravelu, P., Lenny, N., Yeoh, E. J. & Downing, J. R. **2002**. Expression of a conditional AML1-ETO oncogene bypasses embryonic lethality and establishes a murine model of human t(8;21) acute myeloid leukemia. *Cancer Cell*, 1, 63-74.
- Hildebrand, D., Tiefenbach, J., Heinzl, T., Grez, M. & Maurer, A. B. **2001**. Multiple regions of ETO cooperate in transcriptional repression. *J Biol Chem*, 276, 9889-95.
- Hirai, H., Zhang, P., Dayaram, T., Hetherington, C. J., Mizuno, S., Imanishi, J., *et al.* **2006**. C/EBP $\beta$  is required for 'emergency' granulopoiesis. *Nat Immunol*, 7, 732-9.
- Ho, T. T., Warr, M. R., Adelman, E. R., Lansinger, O. M., Flach, J., Verovskaya, E. V., *et al.* **2017**. Autophagy maintains the metabolism and function of young and old stem cells. *Nature*, 543, 205-210.
- Hock, H., Hamblen, M. J., Rooke, H. M., Schindler, J. W., Saleque, S., Fujiwara, Y., *et al.* **2004**. Gfi-1 restricts proliferation and preserves functional integrity of haematopoietic stem cells. *Nature*, 431, 1002-7.
- Hock, H., Hamblen, M. J., Rooke, H. M., Traver, D., Bronson, R. T., Cameron, S., *et al.* **2003**. Intrinsic requirement for zinc finger transcription factor Gfi-1 in neutrophil differentiation. *Immunity*, 18, 109-20.
- Hock, H. & Orkin, S. H. **2006**. Zinc-finger transcription factor Gfi-1: versatile regulator of lymphocytes, neutrophils and hematopoietic stem cells. *Curr Opin Hematol*, 13, 1-6.
- Holtschke, T., Löhler, J., Kanno, Y., Fehr, T., Giese, N., Rosenbauer, F., *et al.* **1996**. Immunodeficiency and chronic myelogenous leukemia-like syndrome in mice with a targeted mutation of the ICSBP gene. *Cell*, 87, 307-17.
- Hörlein, A. J., Näär, A. M., Heinzl, T., Torchia, J., Gloss, B., Kurokawa, R., *et al.* **1995**. Ligand-independent repression by the thyroid hormone receptor mediated by a nuclear receptor co-repressor. *Nature*, 377, 397-404.

- Hosokawa, K., Arai, F., Yoshihara, H., Iwasaki, H., Nakamura, Y., Gomei, Y., *et al.* **2010**. Knockdown of N-cadherin suppresses the long-term engraftment of hematopoietic stem cells. *Blood*, 116, 554-63.
- Hospital, M. A., Prebet, T., Bertoli, S., Thomas, X., Tavernier, E., Braun, T., *et al.* **2014**. Core-binding factor acute myeloid leukemia in first relapse: a retrospective study from the French AML Intergroup. *Blood*, 124, 1312-9.
- Hsu, W., Kerppola, T. K., Chen, P. L., Curran, T. & Chen-Kiang, S. **1994**. Fos and Jun repress transcription activation by NF-IL6 through association at the basic zipper region. *Mol Cell Biol*, 14, 268-76.
- Hu, H. M., Baer, M., Williams, S. C., Johnson, P. F. & Schwartz, R. C. **1998**. Redundancy of C/EBP alpha, -beta, and -delta in supporting the lipopolysaccharide-induced transcription of IL-6 and monocyte chemoattractant protein-1. *J Immunol*, 160, 2334-42.
- Huang, D. W., Sherman, B. T. & Lempicki, R. A. **2009a**. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc*, 4, 44-57.
- Huang, G., Krig, S., Kowbel, D., Xu, H., Hyun, B., Volik, S., *et al.* **2005**. ZNF217 suppresses cell death associated with chemotherapy and telomere dysfunction. *Hum Mol Genet*, 14, 3219-25.
- Huang, G., Shigesada, K., Ito, K., Wee, H. J., Yokomizo, T. & Ito, Y. **2001**. Dimerization with PEBP2beta protects RUNX1/AML1 from ubiquitin-proteasome-mediated degradation. *EMBO J*, 20, 723-33.
- Huang, G., Zhao, X., Wang, L., Elf, S., Xu, H., Sashida, G., *et al.* **2011**. The ability of MLL to bind RUNX1 and methylate H3K4 at PU.1 regulatory regions is impaired by MDS/AML-associated RUNX1/AML1 mutations. *Blood*, 118, 6544-52.
- Huang, H., Woo, A. J., Waldon, Z., Schindler, Y., Moran, T. B., Zhu, H. H., *et al.* **2012**. A Src family kinase-Shp2 axis controls RUNX1 activity in megakaryocyte and T-lymphocyte differentiation. *Genes Dev*, 26, 1587-601.
- Huang, H., Yu, M., Akie, T. E., Moran, T. B., Woo, A. J., Tu, N., *et al.* **2009b**. Differentiation-dependent interactions between RUNX-1 and FLI-1 during megakaryocyte development. *Mol Cell Biol*, 29, 4103-15.
- Huang, Y., Wei, J., Huang, X., Zhou, W., Xu, Y., Deng, D.-H., *et al.* **2020**. Comprehensively analyze the expression and prognostic role for ten-eleven translocations (TETs) in acute myeloid leukemia. *Translational Cancer Research*, 9, 7259-7283.
- Hug, B. A. & Lazar, M. A. **2004**. ETO interacting proteins. *Oncogene*, 23, 4270-4274.
- Huggins, C. J., Malik, R., Lee, S., Salotti, J., Thomas, S., Martin, N., *et al.* **2013**. C/EBP $\gamma$  suppresses senescence and inflammatory gene expression by heterodimerizing with C/EBP $\beta$ . *Mol Cell Biol*, 33, 3242-58.
- Hughes, P. J., Marcinkowska, E., Gocek, E., Studzinski, G. P. & Brown, G. **2010**. Vitamin D3-driven signals for myeloid cell differentiation--implications for differentiation therapy. *Leuk Res*, 34, 553-65.
- Ichikawa, M., Asai, T., Saito, T., Seo, S., Yamazaki, I., Yamagata, T., *et al.* **2004**. AML-1 is required for megakaryocytic maturation and lymphocytic differentiation, but not for maintenance of hematopoietic stem cells in adult hematopoiesis. *Nat Med*, 10, 299-304.
- Ichikawa, M., Yoshimi, A., Nakagawa, M., Nishimoto, N., Watanabe-Okochi, N. & Kurokawa, M. **2013**. A role for RUNX1 in hematopoiesis and myeloid leukemia. *Int J Hematol*, 97, 726-34.
- Ihle, J. N. **1992**. Interleukin-3 and hematopoiesis. *Chem Immunol*, 51, 65-106.

- Ikawa, T., Masuda, K., Lu, M., Minato, N., Katsura, Y. & Kawamoto, H. **2004**. Identification of the earliest prethymic T-cell progenitors in murine fetal blood. *Blood*, 103, 530-7.
- Ikushima, H. & Miyazono, K. **2010**. TGFbeta signalling: a complex web in cancer progression. *Nat Rev Cancer*, 10, 415-24.
- Imai, Y., Kurokawa, M., Yamaguchi, Y., Izutsu, K., Nitta, E., Mitani, K., *et al.* **2004**. The corepressor mSin3A regulates phosphorylation-induced activation, intranuclear location, and stability of AML1. *Mol Cell Biol*, 24, 1033-43.
- Iriyama, N., Hatta, Y., Takeuchi, J., Ogawa, Y., Ohtake, S., Sakura, T., *et al.* **2013**. CD56 expression is an independent prognostic factor for relapse in acute myeloid leukemia with t(8;21). *Leuk Res*, 37, 1021-6.
- Ito, K., Turcotte, R., Cui, J., Zimmerman, S. E., Pinho, S., Mizoguchi, T., *et al.* **2016**. Self-renewal of a purified Tie2+ hematopoietic stem cell population relies on mitochondrial clearance. *Science*, 354, 1156-1160.
- Ito, Y. **2004**. Oncogenic potential of the RUNX gene family: 'overview'. *Oncogene*, 23, 4198-208.
- Ito, Y. **2008**. RUNX genes in development and cancer: regulation of viral gene expression and the discovery of RUNX family genes. *Adv Cancer Res*, 99, 33-76.
- Ivanovs, A., Rytsov, S., Ng, E. S., Stanley, E. G., Elefanty, A. G. & Medvinsky, A. **2017**. Human haematopoietic stem cell development: from the embryo to the dish. *Development*, 144, 2323-2337.
- Iwama, A., Osawa, M., Hirasawa, R., Uchiyama, N., Kaneko, S., Onodera, M., *et al.* **2002**. Reciprocal roles for CCAAT/enhancer binding protein (C/EBP) and PU.1 transcription factors in Langerhans cell commitment. *J Exp Med*, 195, 547-58.
- Jacob, B., Osato, M., Yamashita, N., Wang, C. Q., Taniuchi, I., Littman, D. R., *et al.* **2010**. Stem cell exhaustion due to Runx1 deficiency is prevented by Evi5 activation in leukemogenesis. *Blood*, 115, 1610-20.
- Jagannathan-Bogdan, M. & Zon, L. I. **2013**. Hematopoiesis. *Development*, 140, 2463-7.
- Jaiswal, S., Jamieson, C. H., Pang, W. W., Park, C. Y., Chao, M. P., Majeti, R., *et al.* **2009**. CD47 is upregulated on circulating hematopoietic stem cells and leukemia cells to avoid phagocytosis. *Cell*, 138, 271-85.
- Jakobczyk, H., Debaize, L., Soubise, B., Avner, S., Rouger-Gaudichon, J., Commet, S., *et al.* **2021**. Reduction of RUNX1 transcription factor activity by a CBFA2T3-mimicking peptide: application to B cell precursor acute lymphoblastic leukemia. *J Hematol Oncol*, 14, 47.
- Jayasena, T., Poljak, A., Braidy, N., Zhong, L., Rowlands, B., Muenchhoff, J., *et al.* **2016**. Application of Targeted Mass Spectrometry for the Quantification of Sirtuins in the Central Nervous System. *Sci Rep*, 6, 35391.
- Jha, R., Grover, G. & Bose, P. **2013**. Lymphoid associated antigen expression in new cases of Acute Myeloid Leukemia. *Journal of Pathology of Nepal*, 3, 487-490.
- Ji, Y. & Studzinski, G. P. **2004**. Retinoblastoma protein and CCAAT/enhancer-binding protein beta are required for 1,25-dihydroxyvitamin D3-induced monocytic differentiation of HL60 cells. *Cancer Res*, 64, 370-7.
- Jin, L., Hope, K. J., Zhai, Q., Smadja-Joffe, F. & Dick, J. E. **2006**. Targeting of CD44 eradicates human acute myeloid leukemic stem cells. *Nat Med*, 12, 1167-74.
- Jin, L., Lee, E. M., Ramshaw, H. S., Busfield, S. J., Peoppl, A. G., Wilkinson, L., *et al.* **2009**. Monoclonal antibody-mediated targeting of CD123, IL-3 receptor alpha chain, eliminates human acute myeloid leukemic stem cells. *Cell Stem Cell*, 5, 31-42.

- Jin, S., Zhao, H., Yi, Y., Nakata, Y., Kalota, A. & Gewirtz, A. M. **2010**. c-Myb binds MLL through menin in human leukemia cells and is an important driver of MLL-associated leukemogenesis. *J Clin Invest*, 120, 593-606.
- Johnson, P. F. **2005**. Molecular stop signs: regulation of cell-cycle arrest by C/EBP transcription factors. *J Cell Sci*, 118, 2545-55.
- Johnson, P. F., Landschulz, W. H., Graves, B. J. & Mcknight, S. L. **1987**. Identification of a rat liver nuclear protein that binds to the enhancer core element of three animal viruses. *Genes Dev*, 1, 133-46.
- Jones, L. C., Lin, M. L., Chen, S. S., Krug, U., Hofmann, W. K., Lee, S., *et al.* **2002**. Expression of C/EBPbeta from the C/ebpalpha gene locus is sufficient for normal hematopoiesis in vivo. *Blood*, 99, 2032-6.
- Jordan, C. T. **2007**. The leukemic stem cell. *Best Pract Res Clin Haematol*, 20, 13-8.
- Jung, Y., Wang, J., Song, J., Shiozawa, Y., Havens, A., Wang, Z., *et al.* **2007**. Annexin II expressed by osteoblasts and endothelial cells regulates stem cell adhesion, homing, and engraftment following transplantation. *Blood*, 110, 82-90.
- Jutzi, J. S., Basu, T., Pellmann, M., Kaiser, S., Steinemann, D., Sanders, M. A., *et al.* **2019**. Altered NFE2 activity predisposes to leukemic transformation and myeloid sarcoma with AML-specific aberrations. *Blood*, 133, 1766-1777.
- Kabel, A. M., Zamzami, F., Al-Talhi, M., Al-Dwila, K. & Hamza, R. **2017**. Acute Myeloid Leukemia: A focus on Risk Factors, Clinical Presentation, Diagnosis and Possible Lines of Management. *Journal of Cancer Research and Treatment*, 5, 62-67.
- Kallioniemi, A., Kallioniemi, O. P., Piper, J., Tanner, M., Stokke, T., Chen, L., *et al.* **1994**. Detection and mapping of amplified DNA sequences in breast cancer by comparative genomic hybridization. *Proc Natl Acad Sci U S A*, 91, 2156-60.
- Kalluri, R. & Neilson, E. G. **2003**. Epithelial-mesenchymal transition and its implications for fibrosis. *J Clin Invest*, 112, 1776-84.
- Kalvala, A., Gao, L., Aguila, B., Dotts, K., Rahman, M., Nana-Sinkam, S. P., *et al.* **2016**. Rad51C-ATXN7 fusion gene expression in colorectal tumors. *Mol Cancer*, 15, 47.
- Kamachi, Y., Ogawa, E., Asano, M., Ishida, S., Murakami, Y., Satake, M., *et al.* **1990**. Purification of a mouse nuclear factor that binds to both the A and B cores of the polyomavirus enhancer. *J Virol*, 64, 4808-19.
- Kang, I., Lee, B. C., Choi, S. W., Lee, J. Y., Kim, J. J., Kim, B. E., *et al.* **2018**. Donor-dependent variation of human umbilical cord blood mesenchymal stem cells in response to hypoxic preconditioning and amelioration of limb ischemia. *Exp Mol Med*, 50, 35.
- Kanno, T., Kanno, Y., Chen, L. F., Ogawa, E., Kim, W. Y. & Ito, Y. **1998**. Intrinsic transcriptional activation-inhibition domains of the polyomavirus enhancer binding protein 2/core binding factor alpha subunit revealed in the presence of the beta subunit. *Mol Cell Biol*, 18, 2444-54.
- Karamitros, D., Stoilova, B., Aboukhalil, Z., Hamey, F., Reinisch, A., Samitsch, M., *et al.* **2018**. Single-cell analysis reveals the continuum of human lympho-myeloid progenitor cells. *Nat Immunol*, 19, 85-97.
- Karsunky, H., Mende, I., Schmidt, T. & Möröy, T. **2002a**. High levels of the onco-protein Gfi-1 accelerate T-cell proliferation and inhibit activation induced T-cell death in Jurkat T-cells. *Oncogene*, 21, 1571-9.
- Karsunky, H., Zeng, H., Schmidt, T., Zevnik, B., Kluge, R., Schmid, K. W., *et al.* **2002b**. Inflammatory reactions and severe neutropenia in mice lacking the transcriptional repressor Gfi1. *Nat Genet*, 30, 295-300.

- Katz, S., Kowenz-Leutz, E., Müller, C., Meese, K., Ness, S. A. & Leutz, A. **1993**. The NF-M transcription factor is related to C/EBP beta and plays a role in signal transduction, differentiation and leukemogenesis of avian myelomonocytic cells. *EMBO J*, 12, 1321-32.
- Kawagoe, H., Kandilci, A., Kranenburg, T. A. & Grosveld, G. C. **2007**. Overexpression of N-Myc rapidly causes acute myeloid leukemia in mice. *Cancer Res*, 67, 10677-85.
- Kawamoto, H., Ohmura, K., Fujimoto, S. & Katsura, Y. **1999**. Emergence of T cell progenitors without B cell or myeloid differentiation potential at the earliest stage of hematopoiesis in the murine fetal liver. *J Immunol*, 162, 2725-31.
- Kenderian, S. S., Ruella, M., Shestova, O., Klichinsky, M., Aikawa, V., Morrissette, J. J., *et al.* **2015**. CD33-specific chimeric antigen receptor T cells exhibit potent preclinical activity against human acute myeloid leukemia. *Leukemia*, 29, 1637-47.
- Khalidi, H. S., Medeiros, L. J., Chang, K. L., Brynes, R. K., Slovak, M. L. & Arber, D. A. **1998**. The immunophenotype of adult acute myeloid leukemia: high frequency of lymphoid antigen expression and comparison of immunophenotype, French-American-British classification, and karyotypic abnormalities. *Am J Clin Pathol*, 109, 211-20.
- Khwaja, A., Bjorkholm, M., Gale, R. E., Levine, R. L., Jordan, C. T., Ehninger, G., *et al.* **2016**. Acute myeloid leukaemia. *Nat Rev Dis Primers*, 2, 16010.
- Kiel, M. J., Yilmaz, O. H., Iwashita, T., Terhorst, C. & Morrison, S. J. **2005**. SLAM family receptors distinguish hematopoietic stem and progenitor cells and reveal endothelial niches for stem cells. *Cell*, 121, 1109-21.
- Kieran, M. W., Perkins, A. C., Orkin, S. H. & Zon, L. I. **1996**. Thrombopoietin rescues in vitro erythroid colony formation from mouse embryos lacking the erythropoietin receptor. *Proc Natl Acad Sci U S A*, 93, 9126-31.
- Kihara, R., Nagata, Y., Kiyoi, H., Kato, T., Yamamoto, E., Suzuki, K., *et al.* **2014**. Comprehensive analysis of genetic alterations and their prognostic impacts in adult acute myeloid leukemia patients. *Leukemia*, 28, 1586-95.
- Kilbey, A., Terry, A., Cameron, E. R. & Neil, J. C. **2008**. Oncogene-induced senescence: an essential role for Runx. *Cell Cycle*, 7, 2333-40.
- Kim, W. Y. & Sharpless, N. E. **2006**. The regulation of INK4/ARF in cancer and aging. *Cell*, 127, 265-75.
- Kim, W. Y., Sieweke, M., Ogawa, E., Wee, H. J., Englmeier, U., Graf, T., *et al.* **1999**. Mutual activation of Ets-1 and AML1 DNA binding by direct interaction of their autoinhibitory domains. *EMBO J*, 18, 1609-20.
- Kirtane, K. & Lee, S. J. **2017**. Racial and ethnic disparities in hematologic malignancies. *Blood*, 130, 1699-1705.
- Kitabayashi, I., Aikawa, Y., Nguyen, L. A., Yokoyama, A. & Ohki, M. **2001**. Activation of AML1-mediated transcription by MOZ and inhibition by the MOZ-CBP fusion protein. *EMBO J*, 20, 7184-96.
- Kitabayashi, I., Ida, K., Morohoshi, F., Yokoyama, A., Mitsuhashi, N., Shimizu, K., *et al.* **1998a**. The AML1-MTG8 leukemic fusion protein forms a complex with a novel member of the MTG8(ETO/CDR) family, MTGR1. *Mol Cell Biol*, 18, 846-58.
- Kitabayashi, I., Yokoyama, A., Shimizu, K. & Ohki, M. **1998b**. Interaction and functional cooperation of the leukemia-associated factors AML1 and p300 in myeloid cell differentiation. *EMBO J*, 17, 2994-3004.
- Klampfer, L., Zhang, J., Zelenetz, A. O., Uchida, H. & Nimer, S. D. **1996**. The AML1/ETO fusion protein activates transcription of BCL-2. *Proc Natl Acad Sci U S A*, 93, 14059-64.

- Klein, H. U., Ruckert, C., Kohlmann, A., Bullinger, L., Thiede, C., Haferlach, T., *et al.* **2009**. Quantitative comparison of microarray experiments with published leukemia related gene expression signatures. *BMC Bioinformatics*, 10, 422.
- Klemsz, M. J., Mckercher, S. R., Celada, A., Van Beveren, C. & Maki, R. A. **1990**. The macrophage and B cell-specific transcription factor PU.1 is related to the ets oncogene. *Cell*, 61, 113-24.
- Knudsen, K. J., Rehn, M., Hasemann, M. S., Rapin, N., Bagger, F. O., Ohlsson, E., *et al.* **2015**. ERG promotes the maintenance of hematopoietic stem cells by restricting their differentiation. *Genes Dev*, 29, 1915-29.
- Kobayashi, H., Butler, J. M., O'donnell, R., Kobayashi, M., Ding, B. S., Bonner, B., *et al.* **2010**. Angiocrine factors from Akt-activated endothelial cells balance self-renewal and differentiation of haematopoietic stem cells. *Nat Cell Biol*, 12, 1046-56.
- Kohlmann, A., Kipps, T. J., Rassenti, L. Z., Downing, J. R., Shurtleff, S. A., Mills, K. I., *et al.* **2008**. An international standardization programme towards the application of gene expression profiling in routine leukaemia diagnostics: the Microarray Innovations in LEukemia study prephase. *Br J Haematol*, 142, 802-7.
- Koike, M., Chumakov, A. M., Takeuchi, S., Tasaka, T., Yang, R., Nakamaki, T., *et al.* **1997**. C/EBP- $\epsilon$ : Chromosomal mapping and mutational analysis of the gene in leukemia and preleukemia. *Leukemia Research*, 21, 833-839.
- Kojo, S., Tanaka, H., Endo, T. A., Muroi, S., Liu, Y., Seo, W., *et al.* **2017**. Priming of lineage-specifying genes by Bcl11b is required for lineage choice in post-selection thymocytes. *Nat Commun*, 8, 702.
- Komori, T. **2003**. Requisite roles of Runx2 and Cbfb in skeletal development. *J Bone Miner Metab*, 21, 193-7.
- Kondo, M., Scherer, D. C., Miyamoto, T., King, A. G., Akashi, K., Sugamura, K., *et al.* **2000**. Cell-fate conversion of lymphoid-committed progenitors by instructive actions of cytokines. *Nature*, 407, 383-6.
- Kondo, M., Weissman, I. L. & Akashi, K. **1997**. Identification of clonogenic common lymphoid progenitors in mouse bone marrow. *Cell*, 91, 661-72.
- Koschmieder, S., Agrawal, S., Radomska, H. S., Huettner, C. S., Tenen, D. G., Ottmann, O. G., *et al.* **2007**. Decitabine and vitamin D3 differentially affect hematopoietic transcription factors to induce monocytic differentiation. *Int J Oncol*, 30, 349-55.
- Koschmieder, S., Halmos, B., Levantini, E. & Tenen, D. G. **2009**. Dysregulation of the C/EBPalpha differentiation pathway in human cancer. *J Clin Oncol*, 27, 619-28.
- Koschmieder, S., Rosenbauer, F., Steidl, U., Owens, B. M. & Tenen, D. G. **2005**. Role of transcription factors C/EBPalpha and PU.1 in normal hematopoiesis and leukemia. *Int J Hematol*, 81, 368-77.
- Koury, M. J. & Bondurant, M. C. **1990**. Erythropoietin retards DNA breakdown and prevents programmed death in erythroid progenitor cells. *Science*, 248, 378-81.
- Kowenz-Leutz, E. & Leutz, A. **1999**. A C/EBP beta isoform recruits the SWI/SNF complex to activate myeloid genes. *Mol Cell*, 4, 735-43.
- Kowenz-Leutz, E., Pless, O., Dittmar, G., Knoblich, M. & Leutz, A. **2010**. Crosstalk between C/EBPbeta phosphorylation, arginine methylation, and SWI/SNF/Mediator implies an indexing transcription factor code. *EMBO J*, 29, 1105-15.
- Kowenz-Leutz, E., Twamley, G., Ansieau, S. & Leutz, A. **1994**. Novel mechanism of C/EBP beta (NF-M) transcriptional control: activation through derepression. *Genes Dev*, 8, 2781-91.

- Kozu, T., Fukuyama, T., Yamami, T., Akagi, K. & Kaneko, Y. **2005**. MYND-less splice variants of AML1-MTG8 (RUNX1-CBFA2T1) are expressed in leukemia with t(8;21). *Genes Chromosomes Cancer*, 43, 45-53.
- Krejci, O., Wunderlich, M., Geiger, H., Chou, F. S., Schleimer, D., Jansen, M., *et al.* **2008**. p53 signaling in response to increased DNA damage sensitizes AML1-ETO cells to stress-induced death. *Blood*, 111, 2190-9.
- Krig, S. R., Jin, V. X., Bieda, M. C., O'geen, H., Yaswen, P., Green, R., *et al.* **2007**. Identification of genes directly regulated by the oncogene ZNF217 using chromatin immunoprecipitation (ChIP)-chip assays. *J Biol Chem*, 282, 9703-12.
- Krig, S. R., Miller, J. K., Fritze, S., Beckett, L. A., Neve, R. M., Farnham, P. J., *et al.* **2010**. ZNF217, a candidate breast cancer oncogene amplified at 20q13, regulates expression of the ErbB3 receptor tyrosine kinase in breast cancer cells. *Oncogene*, 29, 5500-10.
- Krueger, A. & Von Boehmer, H. **2007**. Identification of a T lineage-committed progenitor in adult blood. *Immunity*, 26, 105-16.
- Kuchenbauer, F., Schnittger, S., Look, T., Gilliland, G., Tenen, D., Haferlach, T., *et al.* **2006**. Identification of additional cytogenetic and molecular genetic abnormalities in acute myeloid leukaemia with t(8;21)/AML1-ETO. *Br J Haematol*, 134, 616-9.
- Kuilman, T., Michaloglou, C., Vredeveld, L. C., Douma, S., Van Doorn, R., Desmet, C. J., *et al.* **2008**. Oncogene-induced senescence relayed by an interleukin-dependent inflammatory network. *Cell*, 133, 1019-31.
- Kurzejamska, E., Johansson, J., Jirström, K., Prakash, V., Ananthashan, S., Boon, L., *et al.* **2014**. C/EBP $\beta$  expression is an independent predictor of overall survival in breast cancer patients by MHCII/CD4-dependent mechanism of metastasis formation. *Oncogenesis*, 3, e125.
- Kuvarina, O. N., Herglotz, J., Kolodziej, S., Kohrs, N., Herkt, S., Wojcik, B., *et al.* **2015**. RUNX1 represses the erythroid gene expression program during megakaryocytic differentiation. *Blood*, 125, 3570-9.
- Kwok, C., Zeisig, B. B., Qiu, J., Dong, S. & So, C. W. **2009**. Transforming activity of AML1-ETO is independent of CBF $\beta$  and ETO interaction but requires formation of homo-oligomeric complexes. *Proc Natl Acad Sci U S A*, 106, 2853-8.
- Lahoud, M. H., Ristevski, S., Venter, D. J., Jermiin, L. S., Bertoncello, I., Zavarsek, S., *et al.* **2001**. Gene targeting of Desrt, a novel ARID class DNA-binding protein, causes growth retardation and abnormal development of reproductive organs. *Genome Res*, 11, 1327-34.
- Lam, K. & Zhang, D. E. **2012**. RUNX1 and RUNX1-ETO: roles in hematopoiesis and leukemogenesis. *Front Biosci (Landmark Ed)*, 17, 1120-39.
- Lambert, A. W., Pattabiraman, D. R. & Weinberg, R. A. **2017**. Emerging Biological Principles of Metastasis. *Cell*, 168, 670-691.
- Lancet, J. E., Cortes, J. E., Hogge, D. E., Tallman, M. S., Kovacovics, T. J., Damon, L. E., *et al.* **2014**. Phase 2 trial of CPX-351, a fixed 5:1 molar ratio of cytarabine/daunorubicin, vs cytarabine/daunorubicin in older adults with untreated AML. *Blood*, 123, 3239-46.
- Lancet, J. E., Uy, G. L., Cortes, J. E., Newell, L. F., Lin, T. L., Ritchie, E. K., *et al.* **2018**. CPX-351 (cytarabine and daunorubicin) Liposome for Injection Versus Conventional Cytarabine Plus Daunorubicin in Older Patients With Newly Diagnosed Secondary Acute Myeloid Leukemia. *J Clin Oncol*, 36, 2684-2692.
- Lansdorp, P. M., Sutherland, H. J. & Eaves, C. J. **1990**. Selective expression of CD45 isoforms on functional subpopulations of CD34+ hemopoietic cells from human bone marrow. *J Exp Med*, 172, 363-6.

- Lapidot, T., Sirard, C., Vormoor, J., Murdoch, B., Hoang, T., Caceres-Cortes, J., *et al.* **1994**. A cell initiating human acute myeloid leukaemia after transplantation into SCID mice. *Nature*, 367, 645-8.
- Lara-Astiaso, D., Weiner, A., Lorenzo-Vivas, E., Zaretsky, I., Jaitin, D. A., David, E., *et al.* **2014**. Immunogenetics. Chromatin state dynamics during blood formation. *Science*, 345, 943-9.
- Larrue, C., Saland, E., Boutzen, H., Vergez, F., David, M., Joffre, C., *et al.* **2016**. Proteasome inhibitors induce FLT3-ITD degradation through autophagy in AML cells. *Blood*, 127, 882-92.
- Laurenti, E., Doulatov, S., Zandi, S., Plumb, I., Chen, J., April, C., *et al.* **2013**. The transcriptional architecture of early human hematopoiesis identifies multilevel control of lymphoid commitment. *Nat Immunol*, 14, 756-63.
- Laurenti, E. & Göttgens, B. **2018**. From haematopoietic stem cells to complex differentiation landscapes. *Nature*, 553, 418-426.
- Le Lay, S., Lefrère, I., Trautwein, C., Dugail, I. & Krief, S. **2002**. Insulin and sterol-regulatory element-binding protein-1c (SREBP-1C) regulation of gene expression in 3T3-L1 adipocytes. Identification of CCAAT/enhancer-binding protein beta as an SREBP-1C target. *J Biol Chem*, 277, 35625-34.
- Leclair, K. P., Blonar, M. A. & Sharp, P. A. **1992**. The p50 subunit of NF-kappa B associates with the NF-IL6 transcription factor. *Proc Natl Acad Sci U S A*, 89, 8145-9.
- Lee, J. H., Sung, J. Y., Choi, E. K., Yoon, H. K., Kang, B. R., Hong, E. K., *et al.* **2019**. C/EBP $\beta$  Is a Transcriptional Regulator of Wee1 at the G<sub>2</sub>/M Phase of the Cell Cycle. *Cells*, 8.
- Lee, M. G., Wynder, C., Cooch, N. & Shiekhhattar, R. **2005**. An essential role for CoREST in nucleosomal histone 3 lysine 4 demethylation. *Nature*, 437, 432-5.
- Lee, S., Chen, J., Zhou, G., Shi, R. Z., Bouffard, G. G., Kocherginsky, M., *et al.* **2006**. Gene expression profiles in acute myeloid leukemia with common translocations using SAGE. *Proc Natl Acad Sci U S A*, 103, 1030-5.
- Lekstrom-Himes, J. & Xanthopoulos, K. G. **1998**. Biological role of the CCAAT/enhancer-binding protein family of transcription factors. *J Biol Chem*, 273, 28545-8.
- Leonessa, F. & Clarke, R. **2003**. ATP binding cassette transporters and drug resistance in breast cancer. *Endocr Relat Cancer*, 10, 43-73.
- Levanon, D., Glusman, G., Bangsow, T., Ben-Asher, E., Male, D. A., Avidan, N., *et al.* **2001**. Architecture and anatomy of the genomic locus encoding the human leukemia-associated transcription factor RUNX1/AML1. *Gene*, 262, 23-33.
- Levanon, D., Goldstein, R. E., Bernstein, Y., Tang, H., Goldenberg, D., Stifani, S., *et al.* **1998**. Transcriptional repression by AML1 and LEF-1 is mediated by the TLE/Groucho corepressors. *Proc Natl Acad Sci U S A*, 95, 11590-5.
- Levanon, D. & Groner, Y. **2004**. Structure and regulated expression of mammalian RUNX genes. *Oncogene*, 23, 4211-9.
- Ley, T. J., Ding, L., Walter, M. J., McLellan, M. D., Lamprecht, T., Larson, D. E., *et al.* **2010**. DNMT3A mutations in acute myeloid leukemia. *N Engl J Med*, 363, 2424-33.
- Ley, T. J., Miller, C., Ding, L., Raphael, B. J., Mungall, A. J., Robertson, A., *et al.* **2013**. Genomic and epigenomic landscapes of adult de novo acute myeloid leukemia. *N Engl J Med*, 368, 2059-74.
- Li, J., Song, L., Qiu, Y., Yin, A. & Zhong, M. **2014**. ZNF217 is associated with poor prognosis and enhances proliferation and metastasis in ovarian cancer. *Int J Clin Exp Pathol*, 7, 3038-47.

- Li, K., Du, Y., Wei, D. Q. & Zhang, F. **2019**. CEBPE expression is an independent prognostic factor for acute myeloid leukemia. *J Transl Med*, 17, 188.
- Li, P., Maines-Bandiera, S., Kuo, W. L., Guan, Y., Sun, Y., Hills, M., *et al.* **2007**. Multiple roles of the candidate oncogene ZNF217 in ovarian epithelial neoplastic progression. *Int J Cancer*, 120, 1863-73.
- Li, P. & Zon, L. I. **2010**. Resolving the controversy about N-cadherin and hematopoietic stem cells. *Cell Stem Cell*, 6, 199-202.
- Li, Y., Bevilacqua, E., Chiribau, C. B., Majumder, M., Wang, C., Croniger, C. M., *et al.* **2008**. Differential control of the CCAAT/enhancer-binding protein beta (C/EBPbeta) products liver-enriched transcriptional activating protein (LAP) and liver-enriched transcriptional inhibitory protein (LIP) and the regulation of gene expression during the response to endoplasmic reticulum stress. *J Biol Chem*, 283, 22443-56.
- Li, Y., Wang, H., Wang, X., Jin, W., Tan, Y., Fang, H., *et al.* **2016**. Genome-wide studies identify a novel interplay between AML1 and AML1/ETO in t(8;21) acute myeloid leukemia. *Blood*, 127, 233-42.
- Li, Z., Du, L., Dong, Z., Yang, Y., Zhang, X., Wang, L., *et al.* **2015**. MiR-203 suppresses ZNF217 upregulation in colorectal cancer and its oncogenicity. *PLoS One*, 10, e0116170.
- Lian, J. B., Javed, A., Zaidi, S. K., Lengner, C., Montecino, M., Van Wijnen, A. J., *et al.* **2004**. Regulatory controls for osteoblast growth and differentiation: role of Runx/Cbfa/AML factors. *Crit Rev Eukaryot Gene Expr*, 14, 1-41.
- Lian, J. B. & Stein, G. S. **2003**. Runx2/Cbfa1: a multifunctional regulator of bone formation. *Curr Pharm Des*, 9, 2677-85.
- Lieschke, G. J., Grail, D., Hodgson, G., Metcalf, D., Stanley, E., Cheers, C., *et al.* **1994**. Mice lacking granulocyte colony-stimulating factor have chronic neutropenia, granulocyte and macrophage progenitor cell deficiency, and impaired neutrophil mobilization. *Blood*, 84, 1737-46.
- Lilly, A. J., Costa, G., Largeot, A., Fadlullah, M. Z., Lie-a-Ling, M., Lacaud, G., *et al.* **2016**. Interplay between SOX7 and RUNX1 regulates hemogenic endothelial fate in the yolk sac. *Development*, 143, 4341-4351.
- Lin, C., Song, W., Bi, X., Zhao, J., Huang, Z., Li, Z., *et al.* **2014**. Recent advances in the ARID family: focusing on roles in human cancer. *Onco Targets Ther*, 7, 315-24.
- Lin, C. S., Lim, S. K., D'agati, V. & Costantini, F. **1996**. Differential effects of an erythropoietin receptor gene disruption on primitive and definitive erythropoiesis. *Genes Dev*, 10, 154-64.
- Lin, S., Mulloy, J. C. & Goyama, S. **2017**. RUNX1-ETO Leukemia. *Adv Exp Med Biol*, 962, 151-173.
- Lindsley, R. C., Mar, B. G., Mazzola, E., Grauman, P. V., Shareef, S., Allen, S. L., *et al.* **2015**. Acute myeloid leukemia ontogeny is defined by distinct somatic mutations. *Blood*, 125, 1367-76.
- Linggi, B., Müller-Tidow, C., Van De Locht, L., Hu, M., Nip, J., Serve, H., *et al.* **2002**. The t(8;21) fusion protein, AML1 ETO, specifically represses the transcription of the p14(ARF) tumor suppressor in acute myeloid leukemia. *Nat Med*, 8, 743-50.
- Liss, F., Frech, M., Wang, Y., Giel, G., Fischer, S., Simon, C., *et al.* **2021**. IRF8 Is an AML-Specific Susceptibility Factor That Regulates Signaling Pathways and Proliferation of AML Cells. *Cancers (Basel)*, 13.
- Littlepage, L. E., Adler, A. S., Kouros-Mehr, H., Huang, G., Chou, J., Krig, S. R., *et al.* **2012**. The transcription factor ZNF217 is a prognostic biomarker and therapeutic target during breast cancer progression. *Cancer Discov*, 2, 638-51.

- Liu, D., Zhang, X. X., Li, M. C., Cao, C. H., Wan, D. Y., Xi, B. X., *et al.* **2018**. C/EBP $\beta$  enhances platinum resistance of ovarian cancer cells by reprogramming H3K79 methylation. *Nat Commun*, 9, 1739.
- Liu, F., Wu, H. Y., Wesselschmidt, R., Kornaga, T. & Link, D. C. **1996**. Impaired production and increased apoptosis of neutrophils in granulocyte colony-stimulating factor receptor-deficient mice. *Immunity*, 5, 491-501.
- Liu, L., Xu, F., Chang, C. K., He, Q., Wu, L. Y., Zhang, Z., *et al.* **2017**. MYCN contributes to the malignant characteristics of erythroleukemia through EZH2-mediated epigenetic repression of p21. *Cell Death Dis*, 8, e3126.
- Liu, P., Li, P. & Burke, S. **2010**. Critical roles of Bcl11b in T-cell development and maintenance of T-cell identity. *Immunol Rev*, 238, 138-49.
- Liu, S., Croniger, C., Arizmendi, C., Harada-Shiba, M., Ren, J., Poli, V., *et al.* **1999**. Hypoglycemia and impaired hepatic glucose production in mice with a deletion of the C/EBPbeta gene. *J Clin Invest*, 103, 207-13.
- Liu, S., Shen, T., Huynh, L., Klisovic, M. I., Rush, L. J., Ford, J. L., *et al.* **2005**. Interplay of RUNX1/MTG8 and DNA methyltransferase 1 in acute myeloid leukemia. *Cancer Res*, 65, 1277-84.
- Liu, Y., Cheney, M. D., Gaudet, J. J., Chruszcz, M., Lukasik, S. M., Sugiyama, D., *et al.* **2006**. The tetramer structure of the Nervy homology two domain, NHR2, is critical for AML1/ETO's activity. *Cancer Cell*, 9, 249-60.
- Liu, Y., Nonnemacher, M. R. & Wigdahl, B. **2009**. CCAAT/enhancer-binding proteins and the pathogenesis of retrovirus infection. *Future Microbiol*, 4, 299-321.
- Lo, M. C., Peterson, L. F., Yan, M., Cong, X., Jin, F., Shia, W. J., *et al.* **2012**. Combined gene expression and DNA occupancy profiling identifies potential therapeutic targets of t(8;21) AML. *Blood*, 120, 1473-84.
- Loke, J., Assi, S. A., Imperato, M. R., Ptasinska, A., Cauchy, P., Grabovska, Y., *et al.* **2017**. RUNX1-ETO and RUNX1-EV11 Differentially Reprogram the Chromatin Landscape in t(8;21) and t(3;21) AML. *Cell Rep*, 19, 1654-1668.
- Lord, B. I., Testa, N. G. & Hendry, J. H. **1975**. The relative spatial distributions of CFUs and CFUc in the normal mouse femur. *Blood*, 46, 65-72.
- Lorenz, E., Uphoff, D., Reid, T. R. & Shelton, E. **1951**. Modification of irradiation injury in mice and guinea pigs by bone marrow injections. *J Natl Cancer Inst*, 12, 197-201.
- Lourenço, A. R. & Coffey, P. J. **2017**. A tumor suppressor role for C/EBP $\alpha$  in solid tumors: more than fat and blood. *Oncogene*, 36, 5221-5230.
- Lu, G. D., Ang, Y. H., Zhou, J., Tamilarasi, J., Yan, B., Lim, Y. C., *et al.* **2015**. CCAAT/enhancer binding protein  $\alpha$  predicts poorer prognosis and prevents energy starvation-induced cell death in hepatocellular carcinoma. *Hepatology*, 61, 965-78.
- Lu, J. W., Hsieh, M. S., Hou, H. A., Chen, C. Y., Tien, H. F. & Lin, L. I. **2017**. Overexpression of SOX4 correlates with poor prognosis of acute myeloid leukemia and is leukemogenic in zebrafish. *Blood Cancer J*, 7, e593.
- Lu, J. W., Wang, A. N., Liao, H. A., Chen, C. Y., Hou, H. A., Hu, C. Y., *et al.* **2016**. Cabozantinib is selectively cytotoxic in acute myeloid leukemia cells with FLT3-internal tandem duplication (FLT3-ITD). *Cancer Lett*, 376, 218-25.
- Luckett, W. P. **1978**. Origin and differentiation of the yolk sac and extraembryonic mesoderm in presomite human and rhesus monkey embryos. *Am J Anat*, 152, 59-97.

- Luker, K. E., Pica, C. M., Schreiber, R. D. & Piwnica-Worms, D. **2001**. Overexpression of IRF9 confers resistance to antimicrotubule agents in breast cancer cells. *Cancer Res*, 61, 6540-7.
- Lutterbach, B., Westendorf, J. J., Linggi, B., Patten, A., Moniwa, M., Davie, J. R., *et al.* **1998**. ETO, a target of t(8;21) in acute leukemia, interacts with the N-CoR and mSin3 corepressors. *Mol Cell Biol*, 18, 7176-84.
- Ly, M., Rentas, S., Vujovic, A., Wong, N., Moreira, S., Xu, J., *et al.* **2019**. Diminished AHR Signaling Drives Human Acute Myeloid Leukemia Stem Cell Maintenance. *Cancer Res*, 79, 5799-5811.
- Ma, J., Zhao, S., Qiao, X., Knight, T., Edwards, H., Polin, L., *et al.* **2019**. Inhibition of Bcl-2 Synergistically Enhances the Antileukemic Activity of Midostaurin and Gilteritinib in Preclinical Models of FLT3-Mutated Acute Myeloid Leukemia. *Clin Cancer Res*, 25, 6815-6826.
- Ma, X. J., Wang, Z., Ryan, P. D., Isakoff, S. J., Barmettler, A., Fuller, A., *et al.* **2004**. A two-gene expression ratio predicts clinical outcome in breast cancer patients treated with tamoxifen. *Cancer Cell*, 5, 607-16.
- Mackarechtschian, K., Hardin, J. D., Moore, K. A., Boast, S., Goff, S. P. & Lemischka, I. R. **1995**. Targeted disruption of the flk2/flt3 gene leads to deficiencies in primitive hematopoietic progenitors. *Immunity*, 3, 147-61.
- Maier, H., Ostraat, R., Gao, H., Fields, S., Shinton, S. A., Medina, K. L., *et al.* **2004**. Early B cell factor cooperates with Runx1 and mediates epigenetic changes associated with mb-1 transcription. *Nat Immunol*, 5, 1069-77.
- Maillard, I., Koch, U., Dumortier, A., Shestova, O., Xu, L., Sai, H., *et al.* **2008**. Canonical notch signaling is dispensable for the maintenance of adult hematopoietic stem cells. *Cell Stem Cell*, 2, 356-66.
- Maiques-Diaz, A., Chou, F. S., Wunderlich, M., Gómez-López, G., Jacinto, F. V., Rodriguez-Perales, S., *et al.* **2012**. Chromatin modifications induced by the AML1-ETO fusion protein reversibly silence its genomic targets through AML1 and Sp1 binding motifs. *Leukemia*, 26, 1329-37.
- Mancini, S. J., Mantei, N., Dumortier, A., Suter, U., Macdonald, H. R. & Radtke, F. **2005**. Jagged1-dependent Notch signaling is dispensable for hematopoietic stem cell self-renewal and differentiation. *Blood*, 105, 2340-2.
- Mandoli, A., Singh, Abhishek a., Prange, Koen h. M., Tijchon, E., Oerlemans, M., Dirks, R., *et al.* **2016**. The Hematopoietic Transcription Factors RUNX1 and ERG Prevent AML1-ETO Oncogene Overexpression and Onset of the Apoptosis Program in t(8;21) AMLs. *Cell Reports*, 17, 2087-2100.
- Mangan, J. K. & Speck, N. A. **2011**. RUNX1 mutations in clonal myeloid disorders: from conventional cytogenetics to next generation sequencing, a story 40 years in the making. *Crit Rev Oncog*, 16, 77-91.
- Mansoor, A., Mansoor, M. O., Patel, J. L., Zhao, S., Natkunam, Y. & Bieker, J. J. **2020**. KLF1/EKLF expression in acute leukemia is correlated with chromosomal abnormalities. *Blood Cells Mol Dis*, 83, 102434.
- Månsson, R., Hultquist, A., Luc, S., Yang, L., Anderson, K., Kharazi, S., *et al.* **2007**. Molecular evidence for hierarchical transcriptional lineage priming in fetal and adult stem cells and multipotent progenitors. *Immunity*, 26, 407-19.
- Mao, X. G., Yan, M., Xue, X. Y., Zhang, X., Ren, H. G., Guo, G., *et al.* **2011**. Overexpression of ZNF217 in glioblastoma contributes to the maintenance of glioma stem cells regulated by hypoxia-inducible factors. *Lab Invest*, 91, 1068-78.
- Marcinkowska, E., Garay, E., Gocek, E., Chrobak, A., Wang, X. & Studzinski, G. P. **2006**. Regulation of C/EBPbeta isoforms by MAPK pathways in HL60 cells induced to differentiate by 1,25-dihydroxyvitamin D3. *Exp Cell Res*, 312, 2054-65.
- Marcucci, G., Mrózek, K., Ruppert, A. S., Maharry, K., Kolitz, J. E., Moore, J. O., *et al.* **2005**. Prognostic factors and outcome of core binding factor acute myeloid leukemia patients with t(8;21) differ from those of patients with inv(16): a Cancer and Leukemia Group B study. *J Clin Oncol*, 23, 5705-17.

- Marofi, F., Motavalli, R., Safonov, V. A., Thangavelu, L., Yumashev, A. V., Alexander, M., *et al.* **2021**. CAR T cells in solid tumors: challenges and opportunities. *Stem Cell Res Ther*, 12, 81.
- Martens, J. H., Mandoli, A., Simmer, F., Wierenga, B. J., Saeed, S., Singh, A. A., *et al.* **2012**. ERG and FLI1 binding sites demarcate targets for aberrant epigenetic regulation by AML1-ETO in acute myeloid leukemia. *Blood*, 120, 4038-48.
- Martinez-Soria, N., Mckenzie, L., Draper, J., Ptasinska, A., Issa, H., Potluri, S., *et al.* **2018**. The Oncogenic Transcription Factor RUNX1/ETO Corrupts Cell Cycle Regulation to Drive Leukemic Transformation. *Cancer Cell*, 34, 626-642.e8.
- Massagué, J. **2008**. TGFbeta in Cancer. *Cell*, 134, 215-30.
- Mastracci, T. L., Shadeo, A., Colby, S. M., Tuck, A. B., O'malley, F. P., Bull, S. B., *et al.* **2006**. Genomic alterations in lobular neoplasia: a microarray comparative genomic hybridization signature for early neoplastic proliferation in the breast. *Genes Chromosomes Cancer*, 45, 1007-17.
- Matheny, C. J., Speck, M. E., Cushing, P. R., Zhou, Y., Corpora, T., Regan, M., *et al.* **2007**. Disease mutations in RUNX1 and RUNX2 create nonfunctional, dominant-negative, or hypomorphic alleles. *EMBO J*, 26, 1163-75.
- Mccarthy, P. L., Mercer, F. C., Savicky, M. W., Carter, B. A., Paterno, G. D. & Gillespie, L. L. **2008**. Changes in subcellular localisation of MI-ER1 alpha, a novel oestrogen receptor-alpha interacting protein, is associated with breast cancer progression. *Br J Cancer*, 99, 639-46.
- Mccune, J. M., Namikawa, R., Kaneshima, H., Shultz, L. D., Lieberman, M. & Weissman, I. L. **1988**. The SCID-hu mouse: murine model for the analysis of human hematolymphoid differentiation and function. *Science*, 241, 1632-9.
- Mcghee, L., Bryan, J., Elliott, L., Grimes, H. L., Kazanjian, A., Davis, J. N., *et al.* **2003**. Gfi-1 attaches to the nuclear matrix, associates with ETO (MTG8) and histone deacetylase proteins, and represses transcription using a TSA-sensitive mechanism. *J Cell Biochem*, 89, 1005-18.
- Mcnerney, M. E., Godley, L. A. & Le Beau, M. M. **2017**. Therapy-related myeloid neoplasms: when genetics and environment collide. *Nat Rev Cancer*, 17, 513-527.
- Melnick, A., Carlile, G. W., McConnell, M. J., Polinger, A., Hiebert, S. W. & Licht, J. D. **2000a**. AML-1/ETO fusion protein is a dominant negative inhibitor of transcriptional repression by the promyelocytic leukemia zinc finger protein. *Blood*, 96, 3939-3947.
- Melnick, A. M., Westendorf, J. J., Polinger, A., Carlile, G. W., Arai, S., Ball, H. J., *et al.* **2000b**. The ETO protein disrupted in t(8;21)-associated acute myeloid leukemia is a corepressor for the promyelocytic leukemia zinc finger protein. *Mol Cell Biol*, 20, 2075-86.
- Méndez-Ferrer, S., Michurina, T. V., Ferraro, F., Mazloom, A. R., Macarthur, B. D., Lira, S. A., *et al.* **2010**. Mesenchymal and haematopoietic stem cells form a unique bone marrow niche. *Nature*, 466, 829-34.
- Mendler, J. H., Maharry, K., Radmacher, M. D., Mrózek, K., Becker, H., Metzeler, K. H., *et al.* **2012**. RUNX1 mutations are associated with poor outcome in younger and older patients with cytogenetically normal acute myeloid leukemia and with distinct gene and MicroRNA expression signatures. *J Clin Oncol*, 30, 3109-18.
- Meshinchi, S. & Appelbaum, F. R. **2009**. Structural and functional alterations of FLT3 in acute myeloid leukemia. *Clin Cancer Res*, 15, 4263-9.
- Metcalf, D. & Burgess, A. W. **1982**. Clonal analysis of progenitor cell commitment of granulocyte or macrophage production. *J Cell Physiol*, 111, 275-83.

- Metz, R. & Ziff, E. **1991**. cAMP stimulates the C/EBP-related transcription factor rNFIL-6 to trans-locate to the nucleus and induce c-fos transcription. *Genes Dev*, 5, 1754-66.
- Metzelder, S. K., Michel, C., Von Bonin, M., Rehberger, M., Hessmann, E., Inselmann, S., *et al.* **2015**. NFATc1 as a therapeutic target in FLT3-ITD-positive AML. *Leukemia*, 29, 1470-7.
- Meyers, S., Lenny, N. & Hiebert, S. W. **1995**. The t(8;21) fusion protein interferes with AML-1B-dependent transcriptional activation. *Mol Cell Biol*, 15, 1974-82.
- Michaud, J., Scott, H. S. & Escher, R. **2003**. AML1 interconnected pathways of leukemogenesis. *Cancer Invest*, 21, 105-36.
- Michaud, J., Wu, F., Osato, M., Cottles, G. M., Yanagida, M., Asou, N., *et al.* **2002**. In vitro analyses of known and novel RUNX1/AML1 mutations in dominant familial platelet disorder with predisposition to acute myelogenous leukemia: implications for mechanisms of pathogenesis. *Blood*, 99, 1364-72.
- Milde-Langosch, K., Löning, T. & Bamberger, A. M. **2003**. Expression of the CCAAT/enhancer-binding proteins C/EBPalpha, C/EBPbeta and C/EBPdelta in breast cancer: correlations with clinicopathologic parameters and cell-cycle regulatory proteins. *Breast Cancer Res Treat*, 79, 175-85.
- Mink, S., Haenig, B. & Klempnauer, K. H. **1997**. Interaction and functional collaboration of p300 and C/EBPbeta. *Mol Cell Biol*, 17, 6609-17.
- Mink, S., Jaswal, S., Burk, O. & Klempnauer, K. H. **1999**. The v-Myb oncoprotein activates C/EBPbeta expression by stimulating an autoregulatory loop at the C/EBPbeta promoter. *Biochim Biophys Acta*, 1447, 175-84.
- Mink, S., Kerber, U. & Klempnauer, K. H. **1996**. Interaction of C/EBPbeta and v-Myb is required for synergistic activation of the mim-1 gene. *Mol Cell Biol*, 16, 1316-25.
- Miyamoto, T., Iwasaki, H., Reizis, B., Ye, M., Graf, T., Weissman, I. L., *et al.* **2002**. Myeloid or lymphoid promiscuity as a critical step in hematopoietic lineage commitment. *Dev Cell*, 3, 137-47.
- Miyoshi, H., Kozu, T., Shimizu, K., Enomoto, K., Maseki, N., Kaneko, Y., *et al.* **1993**. The t(8;21) translocation in acute myeloid leukemia results in production of an AML1-MTG8 fusion transcript. *EMBO J*, 12, 2715-21.
- Miyoshi, H., Ohira, M., Shimizu, K., Mitani, K., Hirai, H., Imai, T., *et al.* **1995**. Alternative splicing and genomic structure of the AML1 gene involved in acute myeloid leukemia. *Nucleic Acids Res*, 23, 2762-9.
- Miyoshi, H., Shimizu, K., Kozu, T., Maseki, N., Kaneko, Y. & Ohki, M. **1991**. t(8;21) breakpoints on chromosome 21 in acute myeloid leukemia are clustered within a limited region of a single gene, AML1. *Proc Natl Acad Sci U S A*, 88, 10431-4.
- Moore, M. A., Williams, N. & Metcalf, D. **1973**. In vitro colony formation by normal and leukemic human hematopoietic cells: characterization of the colony-forming cells. *J Natl Cancer Inst*, 50, 603-23.
- Morad, V., Pevsner-Fischer, M., Barnees, S., Samokovlisky, A., Rousso-Noori, L., Rosenfeld, R., *et al.* **2008**. The myelopoietic supportive capacity of mesenchymal stromal cells is uncoupled from multipotency and is influenced by lineage determination and interference with glycosylation. *Stem Cells*, 26, 2275-86.
- Morgan, R. G., Pearn, L., Liddiard, K., Pumford, S. L., Burnett, A. K., Tonks, A., *et al.* **2013**.  $\gamma$ -Catenin is overexpressed in acute myeloid leukemia and promotes the stabilization and nuclear localization of  $\beta$ -catenin. *Leukemia*, 27, 336-43.
- Morita, K., Okamura, T., Sumitomo, S., Iwasaki, Y., Fujio, K. & Yamamoto, K. **2016**. Emerging roles of Egr2 and Egr3 in the control of systemic autoimmunity. *Rheumatology (Oxford)*, 55, ii76-ii81.

- Möröy, T. & Khandanpour, C. **2019**. Role of GF11 in Epigenetic Regulation of MDS and AML Pathogenesis: Mechanisms and Therapeutic Implications. *Front Oncol*, 9, 824.
- Morrison, S. J. & Scadden, D. T. **2014**. The bone marrow niche for haematopoietic stem cells. *Nature*, 505, 327-34.
- Morrison, S. J., Wandycz, A. M., Hemmati, H. D., Wright, D. E. & Weissman, I. L. **1997**. Identification of a lineage of multipotent hematopoietic progenitors. *Development*, 124, 1929-39.
- Morrison, S. J. & Weissman, I. L. **1994**. The long-term repopulating subset of hematopoietic stem cells is deterministic and isolatable by phenotype. *Immunity*, 1, 661-73.
- Mossadegh-Keller, N., Sarrazin, S., Kandalla, P. K., Espinosa, L., Stanley, E. R., Nutt, S. L., *et al.* **2013**. M-CSF instructs myeloid lineage fate in single haematopoietic stem cells. *Nature*, 497, 239-43.
- Moustakas, A. & De Herreros, A. G. **2017**. Epithelial-mesenchymal transition in cancer. *Mol Oncol*, 11, 715-717.
- Müller-Tidow, C., Steffen, B., Cauvet, T., Tickenbrock, L., Ji, P., Diederichs, S., *et al.* **2004**. Translocation products in acute myeloid leukemia activate the Wnt signaling pathway in hematopoietic cells. *Mol Cell Biol*, 24, 2890-904.
- Mulloy, J. C., Cammenga, J., Berguido, F. J., Wu, K., Zhou, P., Comenzo, R. L., *et al.* **2003**. Maintaining the self-renewal and differentiation potential of human CD34+ hematopoietic cells using a single genetic element. *Blood*, 102, 4369-76.
- Mulloy, J. C., Cammenga, J., Mackenzie, K. L., Berguido, F. J., Moore, M. A. & Nimer, S. D. **2002**. The AML1-ETO fusion protein promotes the expansion of human hematopoietic stem cells. *Blood*, 99, 15-23.
- Na Nakorn, T., Traver, D., Weissman, I. L. & Akashi, K. **2002**. Myeloerythroid-restricted progenitors are sufficient to confer radioprotection and provide the majority of day 8 CFU-S. *J Clin Invest*, 109, 1579-85.
- Nafria, M., Keane, P., Ng, E. S., Stanley, E. G., Elefanty, A. G. & Bonifer, C. **2020**. Expression of RUNX1-ETO Rapidly Alters the Chromatin Landscape and Growth of Early Human Myeloid Precursor Cells. *Cell Rep*, 31, 107691.
- Nagata, T., Gupta, V., Sorce, D., Kim, W. Y., Sali, A., Chait, B. T., *et al.* **1999**. Immunoglobulin motif DNA recognition and heterodimerization of the PEBP2/CBF Runt domain. *Nat Struct Biol*, 6, 615-9.
- Nakagawa, M., Shimabe, M., Watanabe-Okochi, N., Arai, S., Yoshimi, A., Shinohara, A., *et al.* **2011**. AML1/RUNX1 functions as a cytoplasmic attenuator of NF- $\kappa$ B signaling in the repression of myeloid tumors. *Blood*, 118, 6626-37.
- Natsuka, S., Akira, S., Nishio, Y., Hashimoto, S., Sugita, T., Isshiki, H., *et al.* **1992**. Macrophage differentiation-specific expression of NF-IL6, a transcription factor for interleukin-6. *Blood*, 79, 460-6.
- Neame, P. B., Soamboonsrup, P., Browman, G. P., Meyer, R. M., Bengner, A., Wilson, W. E., *et al.* **1986**. Classifying acute leukemia by immunophenotyping: a combined FAB-immunologic classification of AML. *Blood*, 68, 1355-62.
- Nerlov, C. **2004**. C/EBP $\alpha$  mutations in acute myeloid leukaemias. *Nat Rev Cancer*, 4, 394-400.
- Nerlov, C. **2007**. The C/EBP family of transcription factors: a paradigm for interaction between gene expression and proliferation control. *Trends Cell Biol*, 17, 318-24.
- Nerlov, C. & Graf, T. **1998**. PU.1 induces myeloid lineage commitment in multipotent hematopoietic progenitors. *Genes Dev*, 12, 2403-12.

- Nerlov, C., McNagny, K. M., Döderlein, G., Kowenz-Leutz, E. & Graf, T. **1998**. Distinct C/EBP functions are required for eosinophil lineage commitment and maturation. *Genes Dev*, 12, 2413-23.
- Ness, S. A., Kowenz-Leutz, E., Casini, T., Graf, T. & Leutz, A. **1993**. Myb and NF-M: combinatorial activators of myeloid genes in heterologous cell types. *Genes Dev*, 7, 749-59.
- Newcombe, A. A., Gibson, B. E. S. & Keeshan, K. **2018**. Harnessing the potential of epigenetic therapies for childhood acute myeloid leukemia. *Experimental Hematology*.
- Ng, S. W., Mitchell, A., Kennedy, J. A., Chen, W. C., Mcleod, J., Ibrahimova, N., *et al.* **2016**. A 17-gene stemness score for rapid determination of risk in acute leukaemia. *Nature*, 540, 433-437.
- Nguyen, H. H. 2018. Long-Term Stability and Integrity of Plasmid-Based DNA Data Storage. *In*: Park, J., Park, S. J., Lee, C.-S., Hwang, S., Shin, Y.-B., H A, T. H. & Kim, M. (eds.). *Polymers*.
- Nguyen, L. V., Vanner, R., Dirks, P. & Eaves, C. J. **2012**. Cancer stem cells: an evolving concept. *Nat Rev Cancer*, 12, 133-43.
- Nguyen, N. T., Vendrell, J. A., Poulard, C., Györfy, B., Goddard-Léon, S., Bièche, I., *et al.* **2014**. A functional interplay between ZNF217 and estrogen receptor alpha exists in luminal breast cancers. *Mol Oncol*, 8, 1441-57.
- Niebuhr, B., Fischer, M., Täger, M., Cammenga, J. & Stocking, C. **2008**. Gatekeeper function of the RUNX1 transcription factor in acute leukemia. *Blood Cells Mol Dis*, 40, 211-8.
- Niehof, M., Kubicka, S., Zender, L., Manns, M. P. & Trautwein, C. **2001a**. Autoregulation enables different pathways to control CCAAT/enhancer binding protein beta (C/EBP beta) transcription. *J Mol Biol*, 309, 855-68.
- Niehof, M., Manns, M. P. & Trautwein, C. **1997**. CREB controls LAP/C/EBP beta transcription. *Mol Cell Biol*, 17, 3600-13.
- Niehof, M., Streetz, K., Rakemann, T., Bischoff, S. C., Manns, M. P., Horn, F., *et al.* **2001b**. Interleukin-6-induced tethering of STAT3 to the LAP/C/EBPbeta promoter suggests a new mechanism of transcriptional regulation by STAT3. *J Biol Chem*, 276, 9016-27.
- Nieto, M. A., Huang, R. Y., Jackson, R. A. & Thiery, J. P. **2016**. EMT: 2016. *Cell*, 166, 21-45.
- Nilsson, S. K., Johnston, H. M., Whitty, G. A., Williams, B., Webb, R. J., Denhardt, D. T., *et al.* **2005**. Osteopontin, a key component of the hematopoietic stem cell niche and regulator of primitive hematopoietic progenitor cells. *Blood*, 106, 1232-9.
- Nishinakamura, R., Nakayama, N., Hirabayashi, Y., Inoue, T., Aud, D., Mcneil, T., *et al.* **1995**. Mice deficient for the IL-3/GM-CSF/IL-5 beta c receptor exhibit lung pathology and impaired immune response, while beta IL3 receptor-deficient mice are normal. *Immunity*, 2, 211-22.
- Nisson, P. E., Watkins, P. C. & Sacchi, N. **1992**. Transcriptionally active chimeric gene derived from the fusion of the AML1 gene and a novel gene on chromosome 8 in t(8;21) leukemic cells. *Cancer Genet Cytogenet*, 63, 81-8.
- Nonet, G. H., Stampfer, M. R., Chin, K., Gray, J. W., Collins, C. C. & Yaswen, P. **2001**. The ZNF217 gene amplified in breast cancers promotes immortalization of human mammary epithelial cells. *Cancer Res*, 61, 1250-4.
- North, T., Gu, T. L., Stacy, T., Wang, Q., Howard, L., Binder, M., *et al.* **1999**. Cbfa2 is required for the formation of intra-aortic hematopoietic clusters. *Development*, 126, 2563-75.

- North, T. E., De Bruijn, M. F., Stacy, T., Talebian, L., Lind, E., Robin, C., *et al.* **2002**. Runx1 expression marks long-term repopulating hematopoietic stem cells in the midgestation mouse embryo. *Immunity*, 16, 661-72.
- Notta, F., Doulatov, S., Laurenti, E., Poepl, A., Jurisica, I. & Dick, J. E. **2011**. Isolation of single human hematopoietic stem cells capable of long-term multilineage engraftment. *Science*, 333, 218-21.
- Notta, F., Zandi, S., Takayama, N., Dobson, S., Gan, O. I., Wilson, G., *et al.* **2016**. Distinct routes of lineage development reshape the human blood hierarchy across ontogeny. *Science*, 351, aab2116.
- Nucifora, G., Begy, C. R., Kobayashi, H., Roulston, D., Claxton, D., Pedersen-Bjergaard, J., *et al.* **1994**. Consistent intergenic splicing and production of multiple transcripts between AML1 at 21q22 and unrelated genes at 3q26 in (3;21)(q26;q22) translocations. *Proc Natl Acad Sci U S A*, 91, 4004-8.
- Nüsslein-Volhard, C. & Wieschaus, E. **1980**. Mutations affecting segment number and polarity in *Drosophila*. *Nature*, 287, 795-801.
- Odaka, Y., Mally, A., Elliott, L. T. & Meyers, S. **2000**. Nuclear import and subnuclear localization of the proto-oncoprotein ETO (MTG8). *Oncogene*, 19, 3584-97.
- Ogawa, E., Inuzuka, M., Maruyama, M., Satake, M., Naito-Fujimoto, M., Ito, Y., *et al.* **1993**. Molecular cloning and characterization of PEBP2 beta, the heterodimeric partner of a novel *Drosophila* runt-related DNA binding protein PEBP2 alpha. *Virology*, 194, 314-31.
- Ogawa, M., Matsuzaki, Y., Nishikawa, S., Hayashi, S., Kunisada, T., Sudo, T., *et al.* **1991**. Expression and function of c-kit in hemopoietic progenitor cells. *J Exp Med*, 174, 63-71.
- Okuda, T., Cai, Z., Yang, S., Lenny, N., Lyu, C. J., Van Deursen, J. M., *et al.* **1998**. Expression of a knocked-in AML1-ETO leukemia gene inhibits the establishment of normal definitive hematopoiesis and directly generates dysplastic hematopoietic progenitors. *Blood*, 91, 3134-43.
- Okuda, T., Van Deursen, J., Hiebert, S. W., Grosveld, G. & Downing, J. R. **1996**. AML1, the target of multiple chromosomal translocations in human leukemia, is essential for normal fetal liver hematopoiesis. *Cell*, 84, 321-30.
- Onai, N., Obata-Onai, A., Tussiwand, R., Lanzavecchia, A. & Manz, M. G. **2006**. Activation of the Flt3 signal transduction cascade rescues and enhances type I interferon-producing and dendritic cell development. *J Exp Med*, 203, 227-38.
- Osato, M. **2004**. Point mutations in the RUNX1/AML1 gene: another actor in RUNX leukemia. *Oncogene*, 23, 4284-96.
- Ossipow, V., Descombes, P. & Schibler, U. **1993**. CCAAT/enhancer-binding protein mRNA is translated into multiple proteins with different transcription activation potentials. *Proc Natl Acad Sci U S A*, 90, 8219-23.
- Otálora-Otálora, B. A., Henríquez, B., López-Kleine, L. & Rojas, A. **2019**. RUNX family: Oncogenes or tumor suppressors (Review). *Oncol Rep*, 42, 3-19.
- Pabst, T., Mueller, B. U., Harakawa, N., Schoch, C., Haferlach, T., Behre, G., *et al.* **2001**. AML1-ETO downregulates the granulocytic differentiation factor C/EBPalpha in t(8;21) myeloid leukemia. *Nat Med*, 7, 444-51.
- Pal, R., Janz, M., Galson, D. L., Gries, M., Li, S., Jöhrens, K., *et al.* **2009**. C/EBPbeta regulates transcription factors critical for proliferation and survival of multiple myeloma cells. *Blood*, 114, 3890-8.
- Pan, Z., Hetherington, C. J. & Zhang, D. E. **1999**. CCAAT/enhancer-binding protein activates the CD14 promoter and mediates transforming growth factor beta signaling in monocyte development. *J Biol Chem*, 274, 23242-8.

- Pang, S. H. M., De Graaf, C. A., Hilton, D. J., Huntington, N. D., Carotta, S., Wu, L., *et al.* **2018**. PU.1 Is Required for the Developmental Progression of Multipotent Progenitors to Common Lymphoid Progenitors. *Front Immunol*, 9, 1264.
- Papaemmanuil, E., Gerstung, M., Bullinger, L., Gaidzik, V. I., Paschka, P., Roberts, N. D., *et al.* **2016**. Genomic Classification and Prognosis in Acute Myeloid Leukemia. *N Engl J Med*, 374, 2209-2221.
- Park, M. H., Park, S. Y. & Kim, Y. **2008**. Induction of proline-rich tyrosine kinase2 (Pyk2) through C/EBPbeta is involved in PMA-induced monocyte differentiation. *FEBS Lett*, 582, 415-22.
- Patel, J. P., Gönen, M., Figueroa, M. E., Fernandez, H., Sun, Z., Racevskis, J., *et al.* **2012**. Prognostic relevance of integrated genetic profiling in acute myeloid leukemia. *N Engl J Med*, 366, 1079-89.
- Paterno, G. D., Mercer, F. C., Chayter, J. J., Yang, X., Robb, J. D. & Gillespie, L. L. **1998**. Molecular cloning of human er1 cDNA and its differential expression in breast tumours and tumour-derived cell lines. *Gene*, 222, 77-82.
- Patil, P. A., Lombardo, K., Sturtevant, A., Mangray, S. & Yakirevich, E. **2018**. Loss of Expression of a Novel Chromatin Remodeler SMARCA1 in Soft Tissue Sarcoma. *J Cytol Histol*, 9.
- Patsialou, A., Wilsker, D. & Moran, E. **2005**. DNA-binding properties of ARID family proteins. *Nucleic Acids Res*, 33, 66-80.
- Paulsson, K., Forestier, E., Lilljebjörn, H., Heldrup, J., Behrendtz, M., Young, B. D., *et al.* **2010**. Genetic landscape of high hyperdiploid childhood acute lymphoblastic leukemia. *Proc Natl Acad Sci U S A*, 107, 21719-24.
- Paz-Priel, I. & Friedman, A. **2011**. C/EBP $\alpha$  dysregulation in AML and ALL. *Crit Rev Oncog*, 16, 93-102.
- Peiró, G., Diebold, J. & Löhrs, U. **2002**. CAS (cellular apoptosis susceptibility) gene expression in ovarian carcinoma: Correlation with 20q13.2 copy number and cyclin D1, p53, and Rb protein expression. *Am J Clin Pathol*, 118, 922-9.
- Pellin, D., Loperfido, M., Baricordi, C., Wolock, S. L., Montepeloso, A., Weinberg, O. K., *et al.* **2019**. A comprehensive single cell transcriptional landscape of human hematopoietic progenitors. *Nat Commun*, 10, 2395.
- Pencovich, N., Jaschek, R., Dicken, J., Amit, A., Lotem, J., Tanay, A., *et al.* **2013**. Cell-autonomous function of Runx1 transcriptionally regulates mouse megakaryocytic maturation. *PLoS One*, 8, e64248.
- Pencovich, N., Jaschek, R., Tanay, A. & Groner, Y. **2011**. Dynamic combinatorial interactions of RUNX1 and cooperating partners regulates megakaryocytic differentiation in cell line models. *Blood*, 117, e1-14.
- Peng, Z. G., Zhou, M. Y., Huang, Y., Qiu, J. H., Wang, L. S., Liao, S. H., *et al.* **2008**. Physical and functional interaction of Runt-related protein 1 with hypoxia-inducible factor-1alpha. *Oncogene*, 27, 839-47.
- Perrotti, D., Cesi, V., Trotta, R., Guerzoni, C., Santilli, G., Campbell, K., *et al.* **2002**. BCR-ABL suppresses C/EBPalpha expression through inhibitory action of hnRNP E2. *Nat Genet*, 30, 48-58.
- Perry, S. S., Wang, H., Pierce, L. J., Yang, A. M., Tsai, S. & Spangrude, G. J. **2004**. L-selectin defines a bone marrow analog to the thymic early T-lineage progenitor. *Blood*, 103, 2990-6.
- Peterson, J. F., Blackburn, P. R., Webley, M. R., Pearce, K. E., Williamson, C. M., Vasmatazis, G., *et al.* **2019**. Identification of a Novel ZBTB20-JAK2 Fusion by Mate-Pair Sequencing in a Young Adult With B-Lymphoblastic Leukemia/Lymphoma. *Mayo Clin Proc*, 94, 1381-1384.

- Peterson, L. F., Yan, M. & Zhang, D. E. **2007**. The p21Waf1 pathway is involved in blocking leukemogenesis by the t(8;21) fusion protein AML1-ETO. *Blood*, 109, 4392-8.
- Petrie, H. T. **2007**. Early commitment: T cell progenitors in the blood. *Immunity*, 26, 7-8.
- Petrovick, M. S., Hiebert, S. W., Friedman, A. D., Hetherington, C. J., Tenen, D. G. & Zhang, D. E. **1998**. Multiple functional domains of AML1: PU.1 and C/EBPalpha synergize with different regions of AML1. *Mol Cell Biol*, 18, 3915-25.
- Pevny, L., Simon, M. C., Robertson, E., Klein, W. H., Tsai, S. F., D'agati, V., *et al.* **1991**. Erythroid differentiation in chimaeric mice blocked by a targeted mutation in the gene for transcription factor GATA-1. *Nature*, 349, 257-60.
- Pham, T. H., Langmann, S., Schwarzfischer, L., El Chartouni, C., Lichtinger, M., Klug, M., *et al.* **2007**. CCAAT enhancer-binding protein beta regulates constitutive gene expression during late stages of monocyte to macrophage differentiation. *J Biol Chem*, 282, 21924-33.
- Pietras, E. M., Reynaud, D., Kang, Y. A., Carlin, D., Calero-Nieto, F. J., Leavitt, A. D., *et al.* **2015**. Functionally Distinct Subsets of Lineage-Biased Multipotent Progenitors Control Blood Production in Normal and Regenerative Conditions. *Cell Stem Cell*, 17, 35-46.
- Pike, B. L. & Robinson, W. A. **1970**. Human bone marrow colony growth in agar-gel. *J Cell Physiol*, 76, 77-84.
- Pirozzi, G., Tirino, V., Camerlingo, R., Franco, R., La Rocca, A., Liguori, E., *et al.* **2011**. Epithelial to mesenchymal transition by TGFβ-1 induction increases stemness characteristics in primary non small cell lung cancer cell line. *PLoS One*, 6, e21548.
- Piwien Pilipuk, G., Galigniana, M. D. & Schwartz, J. **2003**. Subnuclear localization of C/EBP beta is regulated by growth hormone and dependent on MAPK. *J Biol Chem*, 278, 35668-77.
- Plachetka, A., Chayka, O., Wilczek, C., Melnik, S., Bonifer, C. & Klempnauer, K. H. **2008**. C/EBPbeta induces chromatin opening at a cell-type-specific enhancer. *Mol Cell Biol*, 28, 2102-12.
- Plevova, P., Cerna, D., Balcar, A., Foretova, L., Zapletalova, J., Silhanova, E., *et al.* **2010**. CCND1 and ZNF217 gene amplification is equally frequent in BRCA1 and BRCA2 associated and non-BRCA breast cancer. *Neoplasma*, 57, 325-32.
- Pogosova-Agadjanyan, E. L., Kopecky, K. J., Ostronoff, F., Appelbaum, F. R., Godwin, J., Lee, H., *et al.* **2013**. The prognostic significance of IRF8 transcripts in adult patients with acute myeloid leukemia. *PLoS One*, 8, e70812.
- Poli, V. **1998**. The role of C/EBP isoforms in the control of inflammatory and native immunity functions. *J Biol Chem*, 273, 29279-82.
- Ponomaryov, T., Peled, A., Petit, I., Taichman, R. S., Habler, L., Sandbank, J., *et al.* **2000**. Induction of the chemokine stromal-derived factor-1 following DNA damage improves human stem cell function. *J Clin Invest*, 106, 1331-9.
- Popernack, P. M., Truong, L. T., Kamphuis, M. & Henderson, A. J. **2001**. Ectopic expression of CCAAT/enhancer binding protein beta (C/EBPbeta) in long-term bone marrow cultures induces granulopoiesis and alters stromal cell function. *J Hematother Stem Cell Res*, 10, 631-42.
- Popovici-Muller, J., Lemieux, R. M., Artin, E., Saunders, J. O., Salituro, F. G., Travins, J., *et al.* **2018**. Discovery of AG-120 (Ivosidenib): A First-in-Class Mutant IDH1 Inhibitor for the Treatment of IDH1 Mutant Cancers. *ACS Med Chem Lett*, 9, 300-305.

- Poulos, M. G., Guo, P., Kofler, N. M., Pinho, S., Gutkin, M. C., Tikhonova, A., *et al.* **2013**. Endothelial Jagged-1 is necessary for homeostatic and regenerative hematopoiesis. *Cell Rep*, 4, 1022-34.
- Prada-Arismendy, J., Arroyave, J. C. & Röthlisberger, S. **2017**. Molecular biomarkers in acute myeloid leukemia. *Blood Reviews*, 31, 63-76.
- Preudhomme, C., Warot-Loze, D., Roumier, C., Gardel-Duflos, N., Garand, R., Lai, J. L., *et al.* **2000**. High incidence of biallelic point mutations in the Runt domain of the AML1/PEBP2 alpha B gene in Mo acute myeloid leukemia and in myeloid malignancies with acquired trisomy 21. *Blood*, 96, 2862-9.
- Provance, O. K. & Lewis-Wambi, J. **2019**. Deciphering the role of interferon alpha signaling and microenvironment crosstalk in inflammatory breast cancer. *Breast Cancer Res*, 21, 59.
- Ptasinska, A., Assi, S. A., Mannari, D., James, S. R., Williamson, D., Dunne, J., *et al.* **2012**. Depletion of RUNX1/ETO in t(8;21) AML cells leads to genome-wide changes in chromatin structure and transcription factor binding. *Leukemia*, 26, 1829-41.
- Ptasinska, A., Assi, S. A., Martinez-Soria, N., Imperato, M. R., Piper, J., Cauchy, P., *et al.* **2014**. Identification of a dynamic core transcriptional network in t(8;21) AML that regulates differentiation block and self-renewal. *Cell Rep*, 8, 1974-1988.
- Ptasinska, A., Pickin, A., Assi, S. A., Chin, P. S., Ames, L., Avellino, R., *et al.* **2019**. RUNX1-ETO Depletion in t(8;21) AML Leads to C/EBP $\alpha$ - and AP-1-Mediated Alterations in Enhancer-Promoter Interaction. *Cell Rep*, 28, 3022-3031.e7.
- Pulikkan, J. A., Madera, D., Xue, L., Bradley, P., Landrette, S. F., Kuo, Y. H., *et al.* **2012**. Thrombopoietin/MPL participates in initiating and maintaining RUNX1-ETO acute myeloid leukemia via PI3K/AKT signaling. *Blood*, 120, 868-79.
- Putz, G., Rosner, A., Nuesslein, I., Schmitz, N. & Buchholz, F. **2006**. AML1 deletion in adult mice causes splenomegaly and lymphomas. *Oncogene*, 25, 929-39.
- Qin, X., Su, R., Yang, L., Chan, A., Deng, X., Qing, Y., *et al.* **2019**. Identification of ZNF217 As an Essential Oncogenic Gene in B-Cell Acute Lymphoblastic Leukemia By CRISPR/Cas9-Based Library Screening. *Blood*, 134, 1465-1465.
- Qin, Y. Z., Wang, Y., Xu, L. P., Zhang, X. H., Chen, H., Han, W., *et al.* **2017**. The dynamics of RUNX1-RUNX1T1 transcript levels after allogeneic hematopoietic stem cell transplantation predict relapse in patients with t(8;21) acute myeloid leukemia. *J Hematol Oncol*, 10, 44.
- Qiu, J., Papatsenko, D., Niu, X., Schaniel, C. & Moore, K. **2014**. Divisional history and hematopoietic stem cell function during homeostasis. *Stem Cell Reports*, 2, 473-90.
- Quek, L., Otto, G. W., Garnett, C., Lhermitte, L., Karamitros, D., Stoilova, B., *et al.* **2016**. Genetically distinct leukemic stem cells in human CD34- acute myeloid leukemia are arrested at a hemopoietic precursor-like stage. *J Exp Med*, 213, 1513-35.
- Quinlan, K. G., Nardini, M., Verger, A., Francescato, P., Yaswen, P., Corda, D., *et al.* **2006**. Specific recognition of ZNF217 and other zinc finger proteins at a surface groove of C-terminal binding proteins. *Mol Cell Biol*, 26, 8159-72.
- Quinlan, K. G., Verger, A., Yaswen, P. & Crossley, M. **2007**. Amplification of zinc finger gene 217 (ZNF217) and cancer: when good fingers go bad. *Biochim Biophys Acta*, 1775, 333-40.
- Quintanilla-Martinez, L., Pittaluga, S., Miething, C., Klier, M., Rudelius, M., Davies-Hill, T., *et al.* **2006**. NPM-ALK-dependent expression of the transcription factor CCAAT/enhancer binding protein beta in ALK-positive anaplastic large cell lymphoma. *Blood*, 108, 2029-36.

- Radomska, H. S., Huettner, C. S., Zhang, P., Cheng, T., Scadden, D. T. & Tenen, D. G. **1998**. CCAAT/enhancer binding protein alpha is a regulatory switch sufficient for induction of granulocytic development from bipotential myeloid progenitors. *Mol Cell Biol*, 18, 4301-14.
- Rafii, S., Butler, J. M. & Ding, B. S. **2016**. Angiocrine functions of organ-specific endothelial cells. *Nature*, 529, 316-25.
- Rahman, M. T., Nakayama, K., Rahman, M., Nakayama, N., Ishikawa, M., Katagiri, A., *et al.* **2012**. Prognostic and therapeutic impact of the chromosome 20q13.2 ZNF217 locus amplification in ovarian clear cell carcinoma. *Cancer*, 118, 2846-57.
- Ramji, D. P. & Foka, P. **2002**. CCAAT/enhancer-binding proteins: structure, function and regulation. *Biochem J*, 365, 561-75.
- Ran, D., Shia, W. J., Lo, M. C., Fan, J. B., Knorr, D. A., Ferrell, P. I., *et al.* **2013**. RUNX1a enhances hematopoietic lineage commitment from human embryonic stem cells and inducible pluripotent stem cells. *Blood*, 121, 2882-90.
- Rao, G., Alland, L., Guida, P., Schreiber-Agus, N., Chen, K., Chin, L., *et al.* **1996**. Mouse Sin3A interacts with and can functionally substitute for the amino-terminal repression of the Myc antagonist Mxi1. *Oncogene*, 12, 1165-72.
- Rapin, N., Bagger, F. O., Jendholm, J., Mora-Jensen, H., Krogh, A., Kohlmann, A., *et al.* **2014**. Comparing cancer vs normal gene expression profiles identifies new disease entities and common transcriptional programs in AML patients. *Blood*, 123, 894-904.
- Rashkovan, M. & Ferrando, A. **2019**. Metabolic dependencies and vulnerabilities in leukemia. *Genes & development*, 33, 1460-1474.
- Regalo, G., Förster, S., Resende, C., Bauer, B., Fleige, B., Kemmner, W., *et al.* **2016**. C/EBP $\beta$  regulates homeostatic and oncogenic gastric cell proliferation. *J Mol Med (Berl)*, 94, 1385-1395.
- Reikvam, H., Hatfield, K. J., Kittang, A. O., Hovland, R. & Bruserud, Ø. **2011**. Acute myeloid leukemia with the t(8;21) translocation: clinical consequences and biological implications. *J Biomed Biotechnol*, 2011, 104631.
- Reinisch, A., Chan, S. M., Thomas, D. & Majeti, R. **2015**. Biology and Clinical Relevance of Acute Myeloid Leukemia Stem Cells. *Semin Hematol*, 52, 150-64.
- Renner, M., Wolf, T., Meyer, H., Hartmann, W., Penzel, R., Ulrich, A., *et al.* **2013**. Integrative DNA methylation and gene expression analysis in high-grade soft tissue sarcomas. *Genome Biol*, 14, r137.
- Rhoades, K. L., Hetherington, C. J., Harakawa, N., Yergeau, D. A., Zhou, L., Liu, L. Q., *et al.* **2000**. Analysis of the role of AML1-ETO in leukemogenesis, using an inducible transgenic mouse model. *Blood*, 96, 2108-15.
- Rieger, M. A. & Schroeder, T. **2012**. Hematopoiesis. *Cold Spring Harb Perspect Biol*, 4.
- Robb, L., Drinkwater, C. C., Metcalf, D., Li, R., Köntgen, F., Nicola, N. A., *et al.* **1995**. Hematopoietic and lung abnormalities in mice with a null mutation of the common beta subunit of the receptors for granulocyte-macrophage colony-stimulating factor and interleukins 3 and 5. *Proc Natl Acad Sci U S A*, 92, 9565-9.
- Robinson, A. J., Davies, S., Darley, R. L. & Tonks, A. **2021**. Reactive Oxygen Species Rewires Metabolic Activity in Acute Myeloid Leukemia. *Front Oncol*, 11, 632623.
- Rongvaux, A., Willinger, T., Takizawa, H., Rathinam, C., Auerbach, W., Murphy, A. J., *et al.* **2011**. Human thrombopoietin knockin mice efficiently support human hematopoiesis in vivo. *Proc Natl Acad Sci U S A*, 108, 2378-83.

- Rooney, P. H., Boonsong, A., Mcfadyen, M. C., Mcleod, H. L., Cassidy, J., Curran, S., *et al.* **2004**. The candidate oncogene ZNF217 is frequently amplified in colon cancer. *J Pathol*, 204, 282-8.
- Rosenbauer, F. & Tenen, D. G. **2007**. Transcription factors in myeloid development: balancing differentiation with transformation. *Nat Rev Immunol*, 7, 105-17.
- Rosenberger, G., Koh, C. C., Guo, T., Röst, H. L., Kouvonen, P., Collins, B. C., *et al.* **2014**. A repository of assays to quantify 10,000 human proteins by SWATH-MS. *Sci Data*, 1, 140031.
- Ross, M. E., Mahfouz, R., Onciu, M., Liu, H. C., Zhou, X., Song, G., *et al.* **2004**. Gene expression profiling of pediatric acute myelogenous leukemia. *Blood*, 104, 3679-87.
- Roudaia, L., Cheney, M. D., Manuylova, E., Chen, W., Morrow, M., Park, S., *et al.* **2009**. CBFbeta is critical for AML1-ETO and TEL-AML1 activity. *Blood*, 113, 3070-9.
- Rowley, J. D. **1984**. Biological implications of consistent chromosome rearrangements in leukemia and lymphoma. *Cancer Res*, 44, 3159-68.
- Rowley, J. D. **1990**. Recurring chromosome abnormalities in leukemia and lymphoma. *Semin Hematol*, 27, 122-36.
- Rubin, C. M., Larson, R. A., Anastasi, J., Winter, J. N., Thangavelu, M., Vardiman, J. W., *et al.* **1990**. t(3;21)(q26;q22): a recurring chromosomal abnormality in therapy-related myelodysplastic syndrome and acute myeloid leukemia. *Blood*, 76, 2594-8.
- Rubin, C. M., Larson, R. A., Bitter, M. A., Carrino, J. J., Le Beau, M. M., Diaz, M. O., *et al.* **1987**. Association of a chromosomal 3;21 translocation with the blast phase of chronic myelogenous leukemia. *Blood*, 70, 1338-42.
- Ruffell, D., Mourkioti, F., Gambardella, A., Kirstetter, P., Lopez, R. G., Rosenthal, N., *et al.* **2009**. A CREB-C/EBPbeta cascade induces M2 macrophage-specific gene expression and promotes muscle injury repair. *Proc Natl Acad Sci U S A*, 106, 17475-80.
- Sacchi, N., Tamanini, F., Willemsen, R., Denis-Donini, S., Campiglio, S. & Hoogeveen, A. T. **1998**. Subcellular localization of the oncoprotein MTG8 (CDR/ETO) in neural cells. *Oncogene*, 16, 2609-15.
- Saeed, S., Logie, C., Francoijs, K. J., Frigè, G., Romanenghi, M., Nielsen, F. G., *et al.* **2012**. Chromatin accessibility, p300, and histone acetylation define PML-RAR $\alpha$  and AML1-ETO binding sites in acute myeloid leukemia. *Blood*, 120, 3058-68.
- Saenz, D. T., Fiskus, W., Mill, C. P., Perera, D., Manshouri, T., Lara, B. H., *et al.* **2020**. Mechanistic basis and efficacy of targeting the  $\beta$ -catenin-TCF7L2-JMJD6-c-Myc axis to overcome resistance to BET inhibitors. *Blood*, 135, 1255-1269.
- Safaei, A., Monabati, A., Safavi, M., Atashabparvar, A. & Hosseini, M. **2018**. Additional cytogenetic aberrations in chronic myeloid leukemia: a single-center experience in the Middle East. *Blood Res*, 53, 49-52.
- Saliba, A. N., Boswell, H. S., Cripe, L. D., Abu Zaid, M. I., Weisenbach, J. & Sayar, H. **2017**. Final Results of Phase I/II Study of Combination of Sorafenib, Vorinostat, and Bortezomib in Acute Myeloid Leukemia with FLT3-ITD Mutation or Poor-Risk Cytogenetics. *Blood*, 130, 3897.
- Sanjuan-Pla, A., Macaulay, I. C., Jensen, C. T., Woll, P. S., Luis, T. C., Mead, A., *et al.* **2013**. Platelet-biased stem cells reside at the apex of the haematopoietic stem-cell hierarchy. *Nature*, 502, 232-6.
- Sanz, M. A., Fenaux, P., Tallman, M. S., Estey, E. H., Löwenberg, B., Naoe, T., *et al.* **2019**. Management of acute promyelocytic leukemia: updated recommendations from an expert panel of the European LeukemiaNet. *Blood*, 133, 1630-1643.

- Sarry, J. E., Murphy, K., Perry, R., Sanchez, P. V., Secreto, A., Keefer, C., *et al.* **2011**. Human acute myelogenous leukemia stem cells are rare and heterogeneous when assayed in NOD/SCID/IL2R $\gamma$ c-deficient mice. *J Clin Invest*, 121, 384-95.
- Satake, S., Hirai, H., Hayashi, Y., Shime, N., Tamura, A., Yao, H., *et al.* **2012**. C/EBP $\beta$  is involved in the amplification of early granulocyte precursors during candidemia-induced "emergency" granulopoiesis. *J Immunol*, 189, 4546-55.
- Sato, A., Kamio, N., Yokota, A., Hayashi, Y., Tamura, A., Miura, Y., *et al.* **2020**. C/EBP $\beta$  isoforms sequentially regulate regenerating mouse hematopoietic stem/progenitor cells. *Blood Adv*, 4, 3343-3356.
- Sato, T., Ito, R., Nunomura, S., Ohno, S., Hayashi, K., Satake, M., *et al.* **2003**. Requirement of transcription factor AML1 in proliferation of developing thymocytes. *Immunol Lett*, 89, 39-46.
- Saultz, J. N. & Garzon, R. **2016**. Acute Myeloid Leukemia: A Concise Review. *J Clin Med*, 5.
- Savitsky, D., Tamura, T., Yanai, H. & Taniguchi, T. **2010**. Regulation of immunity and oncogenesis by the IRF transcription factor family. *Cancer Immunol Immunother*, 59, 489-510.
- Savoldi, G., Fenaroli, A., Ferrari, F., Rigaud, G., Albertini, A. & Di Lorenzo, D. **1997**. The glucocorticoid receptor regulates the binding of C/EPB $\beta$  on the alpha-1-acid glycoprotein promoter in vivo. *DNA Cell Biol*, 16, 1467-76.
- Scheller, M., Foerster, J., Heyworth, C. M., Waring, J. F., Löhler, J., Gilmore, G. L., *et al.* **1999**. Altered development and cytokine responses of myeloid progenitors in the absence of transcription factor, interferon consensus sequence binding protein. *Blood*, 94, 3764-71.
- Schessl, C., Rawat, V. P., Cusan, M., Deshpande, A., Kohl, T. M., Rosten, P. M., *et al.* **2005**. The AML1-ETO fusion gene and the FLT3 length mutation collaborate in inducing acute leukemia in mice. *J Clin Invest*, 115, 2159-68.
- Schnerch, D., Yalcintepe, J., Schmidts, A., Becker, H., Follo, M., Engelhardt, M., *et al.* **2012**. Cell cycle control in acute myeloid leukemia. *Am J Cancer Res*, 2, 508-28.
- Schnittger, S., Dicker, F., Kern, W., Wendland, N., Sundermann, J., Alpermann, T., *et al.* **2011**. RUNX1 mutations are frequent in de novo AML with noncomplex karyotype and confer an unfavorable prognosis. *Blood*, 117, 2348-57.
- Schoedel, K. B., Morcos, M. N. F., Zerjatke, T., Roeder, I., Grinenko, T., Voehringer, D., *et al.* **2016**. The bulk of the hematopoietic stem cell population is dispensable for murine steady-state and stress hematopoiesis. *Blood*, 128, 2285-2296.
- Schoenherr, C., Wohlan, K., Dallmann, I., Pich, A., Hegermann, J., Ganser, A., *et al.* **2019**. Stable depletion of RUNX1-ETO in Kasumi-1 cells induces expression and enhanced proteolytic activity of Cathepsin G and Neutrophil Elastase. *PLoS One*, 14, e0225977.
- Schofield, R. **1978**. The relationship between the spleen colony-forming cell and the haemopoietic stem cell. *Blood Cells*, 4, 7-25.
- Schuringa, J. J., Wierenga, A. T., Kruijer, W. & Vellenga, E. **2000**. Constitutive Stat3, Tyr705, and Ser727 phosphorylation in acute myeloid leukemia cells caused by the autocrine secretion of interleukin-6. *Blood*, 95, 3765-70.
- Scott, L. M., Civin, C. I., Rorth, P. & Friedman, A. D. **1992**. A novel temporal expression pattern of three C/EBP family members in differentiating myelomonocytic cells. *Blood*, 80, 1725-35.

- Screpanti, I., Romani, L., Musiani, P., Modesti, A., Fattori, E., Lazzaro, D., *et al.* **1995**. Lymphoproliferative disorder and imbalanced T-helper response in C/EBP beta-deficient mice. *EMBO J*, 14, 1932-41.
- Sears, R. C. & Sealy, L. **1994**. Multiple forms of C/EBP beta bind the EFII enhancer sequence in the Rous sarcoma virus long terminal repeat. *Mol Cell Biol*, 14, 4855-71.
- Sebastian, T. & Johnson, P. F. **2006**. Stop and go: anti-proliferative and mitogenic functions of the transcription factor C/EBPbeta. *Cell Cycle*, 5, 953-7.
- Sebastian, T., Malik, R., Thomas, S., Sage, J. & Johnson, P. F. **2005**. C/EBPbeta cooperates with RB:E2F to implement Ras(V12)-induced cellular senescence. *EMBO J*, 24, 3301-12.
- Seif, A. E. **2011**. Pediatric leukemia predisposition syndromes: clues to understanding leukemogenesis. *Cancer Genet*, 204, 227-44.
- Semerad, C. L., Liu, F., Gregory, A. D., Stumpf, K. & Link, D. C. **2002**. G-CSF is an essential regulator of neutrophil trafficking from the bone marrow to the blood. *Immunity*, 17, 413-23.
- Seo, W., Ikawa, T., Kawamoto, H. & Taniuchi, I. **2012a**. Runx1-Cbfb facilitates early B lymphocyte development by regulating expression of Ebf1. *J Exp Med*, 209, 1255-62.
- Seo, W., Tanaka, H., Miyamoto, C., Levanon, D., Groner, Y. & Taniuchi, I. **2012b**. Roles of VWRPY motif-mediated gene repression by Runx proteins during T-cell development. *Immunol Cell Biol*, 90, 827-30.
- Seymour, J. F., Döhner, H., Butrym, A., Wierzbowska, A., Selleslag, D., Jang, J. H., *et al.* **2017**. Azacitidine improves clinical outcomes in older patients with acute myeloid leukaemia with myelodysplasia-related changes compared with conventional care regimens. *BMC Cancer*, 17, 852.
- Shah, A., Andersson, T. M., Rachet, B., Björkholm, M. & Lambert, P. C. **2013a**. Survival and cure of acute myeloid leukaemia in England, 1971-2006: a population-based study. *Br J Haematol*, 162, 509-16.
- Shah, N. P., Talpaz, M., Deininger, M. W., Mauro, M. J., Flinn, I. W., Bixby, D., *et al.* **2013b**. Ponatinib in patients with refractory acute myeloid leukaemia: findings from a phase 1 study. *Br J Haematol*, 162, 548-52.
- Shi, W., Bessarabova, M., Dosymbekov, D., Dezso, Z., Nikolskaya, T., Dudoladova, M., *et al.* **2010**. Functional analysis of multiple genomic signatures demonstrates that classification algorithms choose phenotype-related genes. *Pharmacogenomics J*, 10, 310-23.
- Shi, Y., Lan, F., Matson, C., Mulligan, P., Whetstine, J. R., Cole, P. A., *et al.* **2004**. Histone demethylation mediated by the nuclear amine oxidase homolog LSD1. *Cell*, 119, 941-53.
- Shi, Y., Sawada, J., Sui, G., Affar, E. B., Whetstine, J. R., Lan, F., *et al.* **2003**. Coordinated histone modifications mediated by a CtBP co-repressor complex. *Nature*, 422, 735-8.
- Shia, W. J., Okumura, A. J., Yan, M., Sarkeshik, A., Lo, M. C., Matsuura, S., *et al.* **2012**. PRMT1 interacts with AML1-ETO to promote its transcriptional activation and progenitor cell proliferative potential. *Blood*, 119, 4953-62.
- Shida, A., Fujioka, S., Kurihara, H., Ishibashi, Y., Mitsumori, N., Omura, N., *et al.* **2014**. Prognostic significance of ZNF217 expression in gastric carcinoma. *Anticancer Res*, 34, 4813-7.
- Shimizu, K., Kitabayashi, I., Kamada, N., Abe, T., Maseki, N., Suzukawa, K., *et al.* **2000**. AML1-MTG8 leukemic protein induces the expression of granulocyte colony-stimulating factor (G-CSF) receptor through the up-regulation of CCAAT/enhancer binding protein epsilon. *Blood*, 96, 288-296.

- Shivdasani, R. A., Mayer, E. L. & Orkin, S. H. **1995**. Absence of blood formation in mice lacking the T-cell leukaemia oncogene protein tal-1/SCL. *Nature*, 373, 432-4.
- Shorbagy, S., Haggag, R., Alazizi, N. & Abouzeid, T. **2016**. CD56 and CD19 antigens expression in acute myeloid leukemia identifies patients with adverse prognosis in Egypt. *Int J Science Res*, 5, 2319-7064.
- Short, N. J., Rytting, M. E. & Cortes, J. E. **2018**. Acute myeloid leukaemia. *Lancet*, 392, 593-606.
- Shultz, L. D., Lyons, B. L., Burzenski, L. M., Gott, B., Chen, X., Chaleff, S., *et al.* **2005**. Human lymphoid and myeloid cell development in NOD/LtSz-scid IL2R gamma null mice engrafted with mobilized human hemopoietic stem cells. *J Immunol*, 174, 6477-89.
- Shultz, L. D., Schweitzer, P. A., Christianson, S. W., Gott, B., Schweitzer, I. B., Tennent, B., *et al.* **1995**. Multiple defects in innate and adaptive immunologic function in NOD/LtSz-scid mice. *J Immunol*, 154, 180-91.
- Siegel, G., Kluba, T., Hermanutz-Klein, U., Bieback, K., Northoff, H. & Schäfer, R. **2013**. Phenotype, donor age and gender affect function of human bone marrow-derived mesenchymal stromal cells. *BMC Med*, 11, 146.
- Signer, R. A., Magee, J. A., Salic, A. & Morrison, S. J. **2014**. Haematopoietic stem cells require a highly regulated protein synthesis rate. *Nature*, 509, 49-54.
- Silva, C. M. **2004**. Role of STATs as downstream signal transducers in Src family kinase-mediated tumorigenesis. *Oncogene*, 23, 8017-23.
- Silver, I. A., Murrills, R. J. & Etherington, D. J. **1988**. Microelectrode studies on the acid microenvironment beneath adherent macrophages and osteoclasts. *Exp Cell Res*, 175, 266-76.
- Simsek, T., Kocabas, F., Zheng, J., Deberardinis, R. J., Mahmoud, A. I., Olson, E. N., *et al.* **2010**. The distinct metabolic profile of hematopoietic stem cells reflects their location in a hypoxic niche. *Cell Stem Cell*, 7, 380-90.
- Singh, A. N. & Sharma, N. **2020**. Quantitative SWATH-Based Proteomic Profiling for Identification of Mechanism-Driven Diagnostic Biomarkers Conferring in the Progression of Metastatic Prostate Cancer. *Front Oncol*, 10, 493.
- Singh Mali, R., Zhang, Q., Defilippis, R. A., Cavazos, A., Kuruvilla, V. M., Raman, J., *et al.* **2021**. Venetoclax combines synergistically with FLT3 inhibition to effectively target leukemic cells in FLT3-ITD+ acute myeloid leukemia models. *Haematologica*, 106, 1034-1046.
- Singh, S. K., Hawkins, C., Clarke, I. D., Squire, J. A., Bayani, J., Hide, T., *et al.* **2004**. Identification of human brain tumour initiating cells. *Nature*, 432, 396-401.
- Singh, S. M., Trivedi, A. K., Lochab, S., Christopheit, M., Hiddemann, W. & Behre, G. **2010**. Proteomics of AML1/ETO Target Proteins: AML1-ETO Targets a C/EBP-NM23 Pathway. *Clinical Proteomics*, 6, 83-91.
- Skokowa, J., Steinemann, D., Katsman-Kuipers, J. E., Zeidler, C., Klimenkova, O., Klimiankou, M., *et al.* **2014**. Cooperativity of RUNX1 and CSF3R mutations in severe congenital neutropenia: a unique pathway in myeloid leukemogenesis. *Blood*, 123, 2229-37.
- Skokowa, J. & Welte, K. **2009**. Dysregulation of myeloid-specific transcription factors in congenital neutropenia. *Ann N Y Acad Sci*, 1176, 94-100.
- Skubitz, K. M., Zimmermann, W., Zimmerman, W., Kammerer, R., Pambuccian, S. & Skubitz, A. P. **2006**. Differential gene expression identifies subgroups of renal cell carcinoma. *J Lab Clin Med*, 147, 250-67.
- Smink, J. J., Bégay, V., Schoenmaker, T., Sterneck, E., De Vries, T. J. & Leutz, A. **2009**. Transcription factor C/EBPbeta isoform ratio regulates osteoclastogenesis through MafB. *EMBO J*, 28, 1769-81.

- Smink, J. J., Tunn, P. U. & Leutz, A. **2012**. Rapamycin inhibits osteoclast formation in giant cell tumor of bone through the C/EBP $\beta$  - MafB axis. *J Mol Med (Berl)*, 90, 25-30.
- Smith, C. C., Levis, M. J., Frankfurt, O., Pagel, J. M., Roboz, G. J., Stone, R. M., *et al.* **2020**. A phase 1/2 study of the oral FLT3 inhibitor pexidartinib in relapsed/refractory FLT3-ITD-mutant acute myeloid leukemia. *Blood Adv*, 4, 1711-1721.
- Solar, G. P., Kerr, W. G., Zeigler, F. C., Hess, D., Donahue, C., De Sauvage, F. J., *et al.* **1998**. Role of c-mpl in early hematopoiesis. *Blood*, 92, 4-10.
- Somervaille, T. C., Matheny, C. J., Spencer, G. J., Iwasaki, M., Rinn, J. L., Witten, D. M., *et al.* **2009**. Hierarchical maintenance of MLL myeloid leukemia stem cells employs a transcriptional program shared with embryonic rather than adult stem cells. *Cell Stem Cell*, 4, 129-40.
- Song, W. J., Sullivan, M. G., Legare, R. D., Hutchings, S., Tan, X., Kufirin, D., *et al.* **1999**. Haploinsufficiency of CBFA2 causes familial thrombocytopenia with propensity to develop acute myelogenous leukaemia. *Nat Genet*, 23, 166-75.
- Sood, R., Hansen, N. F., Donovan, F. X., Carrington, B., Bucci, D., Maskeri, B., *et al.* **2016**. Somatic mutational landscape of AML with inv(16) or t(8;21) identifies patterns of clonal evolution in relapse leukemia. *Leukemia*, 30, 501-4.
- Stanley, E., Lieschke, G. J., Grail, D., Metcalf, D., Hodgson, G., Gall, J. A., *et al.* **1994**. Granulocyte/macrophage colony-stimulating factor-deficient mice show no major perturbation of hematopoiesis but develop a characteristic pulmonary pathology. *Proc Natl Acad Sci U S A*, 91, 5592-6.
- Steensma, D. P., McClure, R. F., Karp, J. E., Tefferi, A., Lasho, T. L., Powell, H. L., *et al.* **2006**. JAK2 V617F is a rare finding in de novo acute myeloid leukemia, but STAT3 activation is common and remains unexplained. *Leukemia*, 20, 971-8.
- Stein, E. M., Dinardo, C. D., Pollyea, D. A., Fathi, A. T., Roboz, G. J., Altman, J. K., *et al.* **2017**. Enasidenib in mutant. *Blood*, 130, 722-731.
- Steinauer, N., Guo, C., Huang, C., Wong, M., Tu, Y., Freter, C. E., *et al.* **2019**. Myeloid translocation gene CBFA2T3 directs a relapse gene program and determines patient-specific outcomes in AML. *Blood Adv*, 3, 1379-1393.
- Stewart, M., Terry, A., Hu, M., O'hara, M., Blyth, K., Baxter, E., *et al.* **1997**. Proviral insertions induce the expression of bone-specific isoforms of PEBP2alphaA (CBFA1): evidence for a new myc collaborating oncogene. *Proc Natl Acad Sci U S A*, 94, 8646-51.
- Stier, S., Ko, Y., Forkert, R., Lutz, C., Neuhaus, T., Grünewald, E., *et al.* **2005**. Osteopontin is a hematopoietic stem cell niche component that negatively regulates stem cell pool size. *J Exp Med*, 201, 1781-91.
- Stoffels, K., Overbergh, L., Giuliatti, A., Verlinden, L., Bouillon, R. & Mathieu, C. **2006**. Immune regulation of 25-hydroxyvitamin-D3-1alpha-hydroxylase in human monocytes. *J Bone Miner Res*, 21, 37-47.
- Stone, R. M., Deangelo, D. J., Klimek, V., Galinsky, I., Estey, E., Nimer, S. D., *et al.* **2005**. Patients with acute myeloid leukemia and an activating mutation in FLT3 respond to a small-molecule FLT3 tyrosine kinase inhibitor, PKC412. *Blood*, 105, 54-60.
- Stoner, S. A., Liu, K. T. H., Andrews, E. T., Liu, M., Arimoto, K. I., Yan, M., *et al.* **2020**. The RUNX1-ETO target gene RASSF2 suppresses t(8;21) AML development and regulates Rac GTPase signaling. *Blood Cancer J*, 10, 16.
- Studzinski, G. P., Wang, X., Ji, Y., Wang, Q., Zhang, Y., Kutner, A., *et al.* **2005**. The rationale for dexamethasone in therapy for myeloid leukemia: role of KSR-MAPK-C/EBP pathway. *J Steroid Biochem Mol Biol*, 97, 47-55.

- Su, W. C., Chou, H. Y., Chang, C. J., Lee, Y. M., Chen, W. H., Huang, K. H., *et al.* **2003**. Differential activation of a C/EBP beta isoform by a novel redox switch may confer the lipopolysaccharide-inducible expression of interleukin-6 gene. *J Biol Chem*, 278, 51150-8.
- Suh, H. C., Gooya, J., Renn, K., Friedman, A. D., Johnson, P. F. & Keller, J. R. **2006**. C/EBPalpha determines hematopoietic cell fate in multipotential progenitor cells by inhibiting erythroid differentiation and inducing myeloid differentiation. *Blood*, 107, 4308-16.
- Sun, G., Zhou, J., Yin, A., Ding, Y. & Zhong, M. **2008a**. Silencing of ZNF217 gene influences the biological behavior of a human ovarian cancer cell line. *Int J Oncol*, 32, 1065-71.
- Sun, X. J., Wang, Z., Wang, L., Jiang, Y., Kost, N., Soong, T. D., *et al.* **2013**. A stable transcription factor complex nucleated by oligomeric AML1-ETO controls leukaemogenesis. *Nature*, 500, 93-7.
- Sun, Y., Chen, B. R. & Deshpande, A. **2018**. Epigenetic Regulators in the Development, Maintenance, and Therapeutic Targeting of Acute Myeloid Leukemia. *Front Oncol*, 8, 41.
- Sun, Y., Wong, N., Guan, Y., Salamanca, C. M., Cheng, J. C., Lee, J. M., *et al.* **2008b**. The eukaryotic translation elongation factor eEF1A2 induces neoplastic properties and mediates tumorigenic effects of ZNF217 in precursor cells of human ovarian carcinomas. *Int J Cancer*, 123, 1761-9.
- Sundfeldt, K., Ivarsson, K., Carlsson, M., Enerbäck, S., Janson, P. O., Brännström, M., *et al.* **1999**. The expression of CCAAT/enhancer binding protein (C/EBP) in the human ovary in vivo: specific increase in C/EBPbeta during epithelial tumour progression. *Br J Cancer*, 79, 1240-8.
- Suter, D. M., Molina, N., Gatfield, D., Schneider, K., Schibler, U. & Naef, F. **2011**. Mammalian genes are transcribed with widely different bursting kinetics. *Science*, 332, 472-4.
- Svendsen, J. B., Baslund, B., Cramer, E. P., Rapin, N., Borregaard, N. & Cowland, J. B. **2016**. MicroRNA-941 Expression in Polymorphonuclear Granulocytes Is Not Related to Granulomatosis with Polyangiitis. *PLoS One*, 11, e0164985.
- Swart, L. E. & Heidenreich, O. **2021**. The RUNX1/RUNX1T1 network: translating insights into therapeutic options. *Exp Hematol*, 94, 1-10.
- Szczyrba, J., Nolte, E., Hart, M., Döll, C., Wach, S., Taubert, H., *et al.* **2013**. Identification of ZNF217, hnRNP-K, VEGF-A and IPO7 as targets for microRNAs that are downregulated in prostate carcinoma. *Int J Cancer*, 132, 775-84.
- Tabach, Y., Kogan-Sakin, I., Buganim, Y., Solomon, H., Goldfinger, N., Hovland, R., *et al.* **2011**. Amplification of the 20q chromosomal arm occurs early in tumorigenic transformation and may initiate cancer. *PLoS One*, 6, e14632.
- Tachibana, M., Tenno, M., Tezuka, C., Sugiyama, M., Yoshida, H. & Taniuchi, I. **2011**. Runx1/Cbfb2 complexes are required for lymphoid tissue inducer cell differentiation at two developmental stages. *J Immunol*, 186, 1450-7.
- Tahirov, T. H., Inoue-Bungo, T., Morii, H., Fujikawa, A., Sasaki, M., Kimura, K., *et al.* **2001**. Structural analyses of DNA recognition by the AML1/Runx-1 Runt domain and its allosteric control by CBFbeta. *Cell*, 104, 755-67.
- Taichman, R. S. & Emerson, S. G. **1994**. Human osteoblasts support hematopoiesis through the production of granulocyte colony-stimulating factor. *J Exp Med*, 179, 1677-82.
- Takahashi, S. **2011**. Current findings for recurring mutations in acute myeloid leukemia. *J Hematol Oncol*, 4, 36.
- Takaoka, A., Tamura, T. & Taniguchi, T. **2008**. Interferon regulatory factor family of transcription factors and regulation of oncogenesis. *Cancer Sci*, 99, 467-78.

- Takeshima, H., Niwa, T., Takahashi, T., Wakabayashi, M., Yamashita, S., Ando, T., *et al.* **2015**. Frequent involvement of chromatin remodeler alterations in gastric field cancerization. *Cancer Lett*, 357, 328-338.
- Takizawa, H., Regoes, R. R., Boddupalli, C. S., Bonhoeffer, S. & Manz, M. G. **2011**. Dynamic variation in cycling of hematopoietic stem cells in steady state and inflammation. *J Exp Med*, 208, 273-84.
- Takubo, K., Nagamatsu, G., Kobayashi, C. I., Nakamura-Ishizu, A., Kobayashi, H., Ikeda, E., *et al.* **2013**. Regulation of glycolysis by Pdk functions as a metabolic checkpoint for cell cycle quiescence in hematopoietic stem cells. *Cell Stem Cell*, 12, 49-61.
- Tam, W. L. & Ng, H. H. **2014**. Sox2: masterminding the root of cancer. *Cancer Cell*, 26, 3-5.
- Tamura, A., Hirai, H., Yokota, A., Kamio, N., Sato, A., Shoji, T., *et al.* **2017**. C/EBP $\beta$  is required for survival of Ly6C. *Blood*, 130, 1809-1818.
- Tamura, A., Hirai, H., Yokota, A., Sato, A., Shoji, T., Kashiwagi, T., *et al.* **2015a**. Accelerated apoptosis of peripheral blood monocytes in Cebpb-deficient mice. *Biochem Biophys Res Commun*, 464, 654-8.
- Tamura, T., Kurotaki, D. & Koizumi, S. **2015b**. Regulation of myelopoiesis by the transcription factor IRF8. *Int J Hematol*, 101, 342-51.
- Tamura, T., Nagamura-Inoue, T., Shmeltzer, Z., Kuwata, T. & Ozato, K. **2000**. ICSBP directs bipotential myeloid progenitor cells to differentiate into mature macrophages. *Immunity*, 13, 155-65.
- Tanaka, T., Akira, S., Yoshida, K., Umemoto, M., Yoneda, Y., Shirafuji, N., *et al.* **1995a**. Targeted disruption of the NF-IL6 gene discloses its essential role in bacteria killing and tumor cytotoxicity by macrophages. *Cell*, 80, 353-61.
- Tanaka, T., Tanaka, K., Ogawa, S., Kurokawa, M., Mitani, K., Nishida, J., *et al.* **1995b**. An acute myeloid leukemia gene, AML1, regulates hemopoietic myeloid cell differentiation and transcriptional activation antagonistically by two alternative spliced forms. *EMBO J*, 14, 341-50.
- Tang, J. L., Hou, H. A., Chen, C. Y., Liu, C. Y., Chou, W. C., Tseng, M. H., *et al.* **2009**. AML1/RUNX1 mutations in 470 adult patients with de novo acute myeloid leukemia: prognostic implication and interaction with other gene alterations. *Blood*, 114, 5352-61.
- Tang, Y. Y., Crute, B. E., Kelley, J. J., Huang, X., Yan, J., Shi, J., *et al.* **2000**. Biophysical characterization of interactions between the core binding factor alpha and beta subunits and DNA. *FEBS Lett*, 470, 167-72.
- Tara, S., Isshiki, Y., Nakajima-Takagi, Y., Oshima, M., Aoyama, K., Tanaka, T., *et al.* **2018**. insufficiency promotes initiation and progression of myelodysplastic syndrome. *Blood*.
- Taskesen, E., Babaei, S., Reinders, M. M. & De Ridder, J. **2015**. Integration of gene expression and DNA-methylation profiles improves molecular subtype classification in acute myeloid leukemia. *BMC Bioinformatics*, 16 Suppl 4, S5.
- Taskesen, E., Bullinger, L., Corbacioglu, A., Sanders, M. A., Erpelinck, C. A., Wouters, B. J., *et al.* **2011**. Prognostic impact, concurrent genetic mutations, and gene expression features of AML with CEBPA mutations in a cohort of 1182 cytogenetically normal AML patients: further evidence for CEBPA double mutant AML as a distinctive disease entity. *Blood*, 117, 2469-75.
- Taussig, D. C., Vargaftig, J., Miraki-Moud, F., Griessinger, E., Sharrock, K., Luke, T., *et al.* **2010**. Leukemia-initiating cells from some acute myeloid leukemia patients with mutated nucleophosmin reside in the CD34(-) fraction. *Blood*, 115, 1976-84.

- Telfer, J. C., Hedblom, E. E., Anderson, M. K., Laurent, M. N. & Rothenberg, E. V. **2004**. Localization of the domains in Runx transcription factors required for the repression of CD4 in thymocytes. *J Immunol*, 172, 4359-70.
- Telfer, J. C. & Rothenberg, E. V. **2001**. Expression and function of a stem cell promoter for the murine CBFalpha2 gene: distinct roles and regulation in natural killer and T cell development. *Dev Biol*, 229, 363-82.
- Tenen, D. G. **2001**. Abnormalities of the CEBP alpha transcription factor: a major target in acute myeloid leukemia. *Leukemia*, 15, 688-9.
- Tenen, D. G. **2003**. Disruption of differentiation in human cancer: AML shows the way. *Nat Rev Cancer*, 3, 89-101.
- Testa, U. & Riccioni, R. **2007**. Deregulation of apoptosis in acute myeloid leukemia. *Haematologica*, 92, 81-94.
- Thillainadesan, G., Chitilian, J. M., Isovich, M., Ablack, J. N., Mymryk, J. S., Tini, M., *et al.* **2012**. TGF- $\beta$ -dependent active demethylation and expression of the p15ink4b tumor suppressor are impaired by the ZNF217/CoREST complex. *Mol Cell*, 46, 636-49.
- Thillainadesan, G., Isovich, M., Loney, E., Andrews, J., Tini, M. & Torchia, J. **2008**. Genome analysis identifies the p15ink4b tumor suppressor as a direct target of the ZNF217/CoREST complex. *Mol Cell Biol*, 28, 6066-77.
- Thivakaran, A., Botezatu, L., Hönes, J. M., Schütte, J., Vassen, L., Al-Matary, Y. S., *et al.* **2018**. Gfi1b: a key player in the genesis and maintenance of acute myeloid leukemia and myelodysplastic syndrome. *Haematologica*, 103, 614-625.
- Thokala, R., Olivares, S., Mi, T., Maiti, S., Deniger, D., Huls, H., *et al.* **2016**. Redirecting Specificity of T cells Using the Sleeping Beauty System to Express Chimeric Antigen Receptors by Mix-and-Matching of VL and VH Domains Targeting CD123+ Tumors. *PLoS One*, 11, e0159477.
- Thollet, A., Vendrell, J. A., Payen, L., Ghayad, S. E., Ben Larbi, S., Grisard, E., *et al.* **2010**. ZNF217 confers resistance to the pro-apoptotic signals of paclitaxel and aberrant expression of Aurora-A in breast cancer cells. *Mol Cancer*, 9, 291.
- Tighe, J. E. & Calabi, F. **1995**. t(8;21) breakpoints are clustered between alternatively spliced exons of MTG8. *Clin Sci (Lond)*, 89, 215-8.
- Tijchon, E., Yi, G., Mandoli, A., Smits, J. G. A., Ferrari, F., Heuts, B. M. H., *et al.* **2019**. The acute myeloid leukemia associated AML1-ETO fusion protein alters the transcriptome and cellular progression in a single-oncogene expressing in vitro induced pluripotent stem cell based granulocyte differentiation model. *PLoS One*, 14, e0226435.
- Till, J. E. & McCulloch, E. A. **1961**. A direct measurement of the radiation sensitivity of normal mouse bone marrow cells. *Radiat Res*, 14, 213-22.
- Timchenko, N. A., Welm, A. L., Lu, X. & Timchenko, L. T. **1999**. CUG repeat binding protein (CUGBP1) interacts with the 5' region of C/EBPbeta mRNA and regulates translation of C/EBPbeta isoforms. *Nucleic Acids Res*, 27, 4517-25.
- Tissières, P., Araud, T., Ochoda, A., Drifte, G., Dunn-Siegrist, I. & Pugin, J. **2009**. Cooperation between PU.1 and CAAT/enhancer-binding protein beta is necessary to induce the expression of the MD-2 gene. *J Biol Chem*, 284, 26261-72.
- Tolomeo, M. & Grimaudo, S. **2020**. The "Janus" Role of C/EBPs Family Members in Cancer Progression. *Int J Mol Sci*, 21.

- Tomita, A., Kiyoi, H. & Naoe, T. **2013**. Mechanisms of action and resistance to all-trans retinoic acid (ATRA) and arsenic trioxide (As<sub>2</sub>O<sub>3</sub>) in acute promyelocytic leukemia. *Int J Hematol*, 97, 717-25.
- Tomizawa, M., Watanabe, K., Saisho, H., Nakagawara, A. & Tagawa, M. **2003**. Down-regulated expression of the CCAAT/enhancer binding protein alpha and beta genes in human hepatocellular carcinoma: a possible prognostic marker. *Anticancer Res*, 23, 351-4.
- Tong, B., Grimes, H. L., Yang, T. Y., Bear, S. E., Qin, Z., Du, K., *et al.* **1998**. The Gfi-1B proto-oncoprotein represses p21WAF1 and inhibits myeloid cell differentiation. *Mol Cell Biol*, 18, 2462-73.
- Tonks, A., Pearn, L., Musson, M., Gilkes, A., Mills, K. I., Burnett, A. K., *et al.* **2007**. Transcriptional dysregulation mediated by RUNX1-RUNX1T1 in normal human progenitor cells and in acute myeloid leukaemia. *Leukemia*, 21, 2495-505.
- Tonks, A., Pearn, L., Tonks, A. J., Pearce, L., Hoy, T., Phillips, S., *et al.* **2003**. The AML1-ETO fusion gene promotes extensive self-renewal of human primary erythroid cells. *Blood*, 101, 624-32.
- Tonks, A., Tonks, A. J., Pearn, L., Mohamad, Z., Burnett, A. K. & Darley, R. L. **2005**. Optimized retroviral transduction protocol which preserves the primitive subpopulation of human hematopoietic cells. *Biotechnol Prog*, 21, 953-8.
- Tonks, A., Tonks, A. J., Pearn, L., Pearce, L., Hoy, T., Couzens, S., *et al.* **2004**. Expression of AML1-ETO in human myelomonocytic cells selectively inhibits granulocytic differentiation and promotes their self-renewal. *Leukemia*, 18, 1238-45.
- Tsapogas, P., Swee, L. K., Nusser, A., Nuber, N., Kreuzaler, M., Capoferri, G., *et al.* **2014**. In vivo evidence for an instructive role of fms-like tyrosine kinase-3 (FLT3) ligand in hematopoietic development. *Haematologica*, 99, 638-46.
- Tseng, H. H., Hwang, Y. H., Yeh, K. T., Chang, J. G., Chen, Y. L. & Yu, H. S. **2009**. Reduced expression of C/EBP alpha protein in hepatocellular carcinoma is associated with advanced tumor stage and shortened patient survival. *J Cancer Res Clin Oncol*, 135, 241-7.
- Tsukada, J., Yoshida, Y., Kominato, Y. & Auron, P. E. **2011**. The CCAAT/enhancer (C/EBP) family of basic-leucine zipper (bZIP) transcription factors is a multifaceted highly-regulated system for gene regulation. *Cytokine*, 54, 6-19.
- Tsukahara, T., Nabeta, Y., Kawaguchi, S., Ikeda, H., Sato, Y., Shimosawa, K., *et al.* **2004**. Identification of human autologous cytotoxic T-lymphocyte-defined osteosarcoma gene that encodes a transcriptional regulator, papillomavirus binding factor. *Cancer Res*, 64, 5442-8.
- Tsuno, T., Mejido, J., Zhao, T., Schmeisser, H., Morrow, A. & Zoon, K. C. **2009**. IRF9 is a key factor for eliciting the antiproliferative activity of IFN-alpha. *J Immunother*, 32, 803-16.
- Tsuzuki, S., Hong, D., Gupta, R., Matsuo, K., Seto, M. & Enver, T. **2007**. Isoform-specific potentiation of stem and progenitor cell engraftment by AML1/RUNX1. *PLoS Med*, 4, e172.
- Tsuzuki, S. & Seto, M. **2012**. Expansion of functionally defined mouse hematopoietic stem and progenitor cells by a short isoform of RUNX1/AML1. *Blood*, 119, 727-35.
- Ueno, Y., Kaneko, N., Saito, R., Kondoh, Y., Shimada, I., Mori, M., *et al.* **2014**. ASP2215, a novel FLT3/AXL inhibitor: Preclinical evaluation in combination with cytarabine and anthracycline in acute myeloid leukemia (AML). *Journal of Clinical Oncology*, 32, 7071-7071.
- Ugarte, G. D., Vargas, M. F., Medina, M. A., León, P., Necuñir, D., Elorza, A. A., *et al.* **2015**. Wnt signaling induces transcription, spatial proximity, and translocation of fusion gene partners in human hematopoietic cells. *Blood*, 126, 1785-9.

- Valk, P. J., Verhaak, R. G., Beijen, M. A., Erpelinck, C. A., Barjesteh Van Waalwijk Van Doorn-Khosrovani, S., Boer, J. M., *et al.* **2004**. Prognostically useful gene-expression profiles in acute myeloid leukemia. *N Engl J Med*, 350, 1617-28.
- Vallejo, M., Ron, D., Miller, C. P. & Habener, J. F. **1993**. C/ATF, a member of the activating transcription factor family of DNA-binding proteins, dimerizes with CAAT/enhancer-binding proteins and directs their binding to cAMP response elements. *Proc Natl Acad Sci U S A*, 90, 4679-83.
- Van 't Veer, L. J., Dai, H., Van De Vijver, M. J., He, Y. D., Hart, A. A., Mao, M., *et al.* **2002**. Gene expression profiling predicts clinical outcome of breast cancer. *Nature*, 415, 530-6.
- Van De Vijver, M. J., He, Y. D., Van't Veer, L. J., Dai, H., Hart, A. A., Voskuil, D. W., *et al.* **2002**. A gene-expression signature as a predictor of survival in breast cancer. *N Engl J Med*, 347, 1999-2009.
- Van Der Kouwe, E. & Staber, P. B. **2019**. RUNX1-ETO: Attacking the Epigenome for Genomic Instable Leukemia. *Int J Mol Sci*, 20.
- Vangala, R. K., Heiss-Neumann, M. S., Rangatia, J. S., Singh, S. M., Schoch, C., Tenen, D. G., *et al.* **2003**. The myeloid master regulator transcription factor PU.1 is inactivated by AML1-ETO in t(8;21) myeloid leukemia. *Blood*, 101, 270-7.
- Vannini, N., Girotra, M., Naveiras, O., Nikitin, G., Campos, V., Giger, S., *et al.* **2016**. Specification of haematopoietic stem cell fate via modulation of mitochondrial activity. *Nat Commun*, 7, 13125.
- Vardiman, J. W., Thiele, J., Arber, D. A., Brunning, R. D., Borowitz, M. J., Porwit, A., *et al.* **2009**. The 2008 revision of the World Health Organization (WHO) classification of myeloid neoplasms and acute leukemia: rationale and important changes. *Blood*, 114, 937-51.
- Varnum-Finney, B., Xu, L., Brashem-Stein, C., Nourigat, C., Flowers, D., Bakkour, S., *et al.* **2000**. Pluripotent, cytokine-dependent, hematopoietic stem cells are immortalized by constitutive Notch1 signaling. *Nat Med*, 6, 1278-81.
- Vegesna, V., Takeuchi, S., Hofmann, W. K., Ikezoe, T., Tavor, S., Krug, U., *et al.* **2002**. C/EBP-beta, C/EBP-delta, PU.1, AML1 genes: mutational analysis in 381 samples of hematopoietic and solid malignancies. *Leuk Res*, 26, 451-7.
- Velten, L., Haas, S. F., Raffel, S., Blaszkiewicz, S., Islam, S., Hennig, B. P., *et al.* **2017**. Human haematopoietic stem cell lineage commitment is a continuous process. *Nat Cell Biol*, 19, 271-281.
- Vendrell, J. A., Thollet, A., Nguyen, N. T., Ghayad, S. E., Vinot, S., Bièche, I., *et al.* **2012**. ZNF217 is a marker of poor prognosis in breast cancer that drives epithelial-mesenchymal transition and invasion. *Cancer Res*, 72, 3593-606.
- Vigilanza, P., Aquilano, K., Baldelli, S., Rotilio, G. & Ciriolo, M. R. **2011**. Modulation of intracellular glutathione affects adipogenesis in 3T3-L1 cells. *J Cell Physiol*, 226, 2016-24.
- Vu, L. P., Perna, F., Wang, L., Voza, F., Figueroa, M. E., Tempst, P., *et al.* **2013**. PRMT4 blocks myeloid differentiation by assembling a methyl-RUNX1-dependent repressor complex. *Cell Rep*, 5, 1625-38.
- Walker, A. R., Wang, H., Walsh, K., Bhatnagar, B., Vasu, S., Garzon, R., *et al.* **2016**. Midostaurin, bortezomib and MEC in relapsed/refractory acute myeloid leukemia. *Leuk Lymphoma*, 57, 2100-8.
- Wall, L., Destroismaisons, N., Delvoye, N. & Guy, L. G. **1996**. CAAT/enhancer-binding proteins are involved in beta-globin gene expression and are differentially expressed in murine erythroleukemia and K562 cells. *J Biol Chem*, 271, 16477-84.

- Wang, H., Iakova, P., Wilde, M., Welm, A., Goode, T., Roesler, W. J., *et al.* **2001**. C/EBPalpha arrests cell proliferation through direct inhibition of Cdk2 and Cdk4. *Mol Cell*, 8, 817-28.
- Wang, H. & Morse, H. C. **2009**. IRF8 regulates myeloid and B lymphoid lineage diversification. *Immunol Res*, 43, 109-17.
- Wang, J., Hoshino, T., Redner, R. L., Kajigaya, S. & Liu, J. M. **1998**. ETO, fusion partner in t(8;21) acute myeloid leukemia, represses transcription by interaction with the human N-CoR/mSin3/HDAC1 complex. *Proc Natl Acad Sci U S A*, 95, 10860-5.
- Wang, J. C., Doedens, M. & Dick, J. E. **1997**. Primitive human hematopoietic cells are enriched in cord blood compared with adult bone marrow or mobilized peripheral blood as measured by the quantitative in vivo SCID-repopulating cell assay. *Blood*, 89, 3919-24.
- Wang, L., Gural, A., Sun, X. J., Zhao, X., Perna, F., Huang, G., *et al.* **2011a**. The leukemogenicity of AML1-ETO is dependent on site-specific lysine acetylation. *Science*, 333, 765-9.
- Wang, L., Man, N., Sun, X. J., Tan, Y., Garcia-Cao, M., Cao, M. G., *et al.* **2015a**. Regulation of AKT signaling by Id1 controls t(8;21) leukemia initiation and progression. *Blood*, 126, 640-50.
- Wang, Q., Stacy, T., Binder, M., Marin-Padilla, M., Sharpe, A. H. & Speck, N. A. **1996**. Disruption of the Cbfa2 gene causes necrosis and hemorrhaging in the central nervous system and blocks definitive hematopoiesis. *Proc Natl Acad Sci U S A*, 93, 3444-9.
- Wang, Q. F. & Friedman, A. D. **2002**. CCAAT/enhancer-binding proteins are required for granulopoiesis independent of their induction of the granulocyte colony-stimulating factor receptor. *Blood*, 99, 2776-85.
- Wang, Q. S., Wang, Y., Lv, H. Y., Han, Q. W., Fan, H., Guo, B., *et al.* **2015b**. Treatment of CD33-directed chimeric antigen receptor-modified T cells in one patient with relapsed and refractory acute myeloid leukemia. *Mol Ther*, 23, 184-91.
- Wang, S., Wang, Q., Crute, B. E., Melnikova, I. N., Keller, S. R. & Speck, N. A. **1993**. Cloning and characterization of subunits of the T-cell receptor and murine leukemia virus enhancer core-binding factor. *Mol Cell Biol*, 13, 3324-39.
- Wang, S. W. & Speck, N. A. **1992**. Purification of core-binding factor, a protein that binds the conserved core site in murine leukemia virus enhancers. *Mol Cell Biol*, 12, 89-102.
- Wang, W., Schwemmers, S., Hexner, E. O. & Pahl, H. L. **2010a**. AML1 is overexpressed in patients with myeloproliferative neoplasms and mediates JAK2V617F-independent overexpression of NF-E2. *Blood*, 116, 254-66.
- Wang, Y., Krivtsov, A. V., Sinha, A. U., North, T. E., Goessling, W., Feng, Z., *et al.* **2010b**. The Wnt/beta-catenin pathway is required for the development of leukemia stem cells in AML. *Science*, 327, 1650-3.
- Wang, Y. Y., Zhao, L. J., Wu, C. F., Liu, P., Shi, L., Liang, Y., *et al.* **2011b**. C-KIT mutation cooperates with full-length AML1-ETO to induce acute myeloid leukemia in mice. *Proc Natl Acad Sci U S A*, 108, 2450-5.
- Warr, M. R., Binnewies, M., Flach, J., Reynaud, D., Garg, T., Malhotra, R., *et al.* **2013**. FOXO3A directs a protective autophagy program in haematopoietic stem cells. *Nature*, 494, 323-7.
- Wedel, A. & Ziegler-Heitbrock, H. W. **1995**. The C/EBP family of transcription factors. *Immunobiology*, 193, 171-85.
- Wee, H. J., Voon, D. C., Bae, S. C. & Ito, Y. **2008**. PEBP2-beta/CBF-beta-dependent phosphorylation of RUNX1 and p300 by HIPK2: implications for leukemogenesis. *Blood*, 112, 3777-87.

- Wei, A. H., Strickland, S. A., Hou, J. Z., Fiedler, W., Lin, T. L., Walter, R. B., *et al.* **2019**. Venetoclax Combined With Low-Dose Cytarabine for Previously Untreated Patients With Acute Myeloid Leukemia: Results From a Phase Ib/II Study. *J Clin Oncol*, 37, 1277-1284.
- Weihua, X., Lindner, D. J. & Kalvakolanu, D. V. **1997**. The interferon-inducible murine p48 (ISGF3gamma) gene is regulated by protooncogene c-myc. *Proc Natl Acad Sci U S A*, 94, 7227-32.
- Weina, K. & Utikal, J. **2014**. SOX2 and cancer: current research and its implications in the clinic. *Clin Transl Med*, 3, 19.
- Weiskopf, K., Schnorr, P. J., Pang, W. W., Chao, M. P., Chhabra, A., Seita, J., *et al.* **2016**. Myeloid Cell Origins, Differentiation, and Clinical Implications. *Microbiol Spectr*, 4.
- Welch, J. S., Ley, T. J., Link, D. C., Miller, C. A., Larson, D. E., Koboldt, D. C., *et al.* **2012**. The origin and evolution of mutations in acute myeloid leukemia. *Cell*, 150, 264-78.
- Weng, S., Matsuura, S., Mowery, C. T., Stoner, S. A., Lam, K., Ran, D., *et al.* **2017**. Restoration of MYC-repressed targets mediates the negative effects of GM-CSF on RUNX1-ETO leukemogenicity. *Leukemia*, 31, 159-169.
- Wessells, J., Yakar, S. & Johnson, P. F. **2004**. Critical prosurvival roles for C/EBP beta and insulin-like growth factor I in macrophage tumor cells. *Mol Cell Biol*, 24, 3238-50.
- Wethkamp, N. & Klempnauer, K. H. **2009**. Daxx is a transcriptional repressor of CCAAT/enhancer-binding protein beta. *J Biol Chem*, 284, 28783-94.
- Wethmar, K., Bégay, V., Smink, J. J., Zaragoza, K., Wiesenthal, V., Dörken, B., *et al.* **2010**. C/EBPbetaDeltaORF mice--a genetic model for uORF-mediated translational control in mammals. *Genes Dev*, 24, 15-20.
- Whitson, R. H., Huang, T. & Itakura, K. **1999**. The novel Mrf-2 DNA-binding domain recognizes a five-base core sequence through major and minor-groove contacts. *Biochem Biophys Res Commun*, 258, 326-31.
- Williams, B. A., Law, A., Hunyadkurti, J., Desilets, S., Leyton, J. V. & Keating, A. **2019**. Antibody Therapies for Acute Myeloid Leukemia: Unconjugated, Toxin-Conjugated, Radio-Conjugated and Multivalent Formats. *J Clin Med*, 8.
- Williams, C. B., Kambhampati, S., Fiskus, W., Wick, J., Dutreix, C., Ganguly, S., *et al.* **2013**. Preclinical and phase I results of decitabine in combination with midostaurin (PKC412) for newly diagnosed elderly or relapsed/refractory adult patients with acute myeloid leukemia. *Pharmacotherapy*, 33, 1341-52.
- Williams, S. C., Baer, M., Dillner, A. J. & Johnson, P. F. **1995**. CRP2 (C/EBP beta) contains a bipartite regulatory domain that controls transcriptional activation, DNA binding and cell specificity. *EMBO J*, 14, 3170-83.
- Williams, S. C., Cantwell, C. A. & Johnson, P. F. **1991**. A family of C/EBP-related proteins capable of forming covalently linked leucine zipper dimers in vitro. *Genes Dev*, 5, 1553-67.
- Williamson, E. A., Xu, H. N., Gombart, A. F., Verbeek, W., Chumakov, A. M., Friedman, A. D., *et al.* **1998**. Identification of transcriptional activation and repression domains in human CCAAT/enhancer-binding protein epsilon. *J Biol Chem*, 273, 14796-804.
- Wilson, A., Laurenti, E., Oser, G., Van Der Wath, R. C., Blanco-Bose, W., Jaworski, M., *et al.* **2008**. Hematopoietic stem cells reversibly switch from dormancy to self-renewal during homeostasis and repair. *Cell*, 135, 1118-29.
- Winkler, I. G., Sims, N. A., Pettit, A. R., Barbier, V., Nowlan, B., Helwani, F., *et al.* **2010**. Bone marrow macrophages maintain hematopoietic stem cell (HSC) niches and their depletion mobilizes HSCs. *Blood*, 116, 4815-28.

- Wolford, J. K. & Prochazka, M. **1998**. Structure and expression of the human MTG8/ETO gene. *Gene*, 212, 103-9.
- Woolthuis, C. M. & Park, C. Y. **2016**. Hematopoietic stem/progenitor cell commitment to the megakaryocyte lineage. *Blood*, 127, 1242-8.
- Wouters, B. J., Löwenberg, B., Erpelinck-Verschueren, C. A., Van Putten, W. L., Valk, P. J. & Delwel, R. **2009**. Double CEBPA mutations, but not single CEBPA mutations, define a subgroup of acute myeloid leukemia with a distinctive gene expression profile that is uniquely associated with a favorable outcome. *Blood*, 113, 3088-91.
- Wu, H., Hu, C., Wang, A., Weisberg, E. L., Wang, W., Chen, C., *et al.* **2016**. Ibrutinib selectively targets FLT3-ITD in mutant FLT3-positive AML. *Leukemia*, 30, 754-7.
- Wu, H., Liu, X., Jaenisch, R. & Lodish, H. F. **1995**. Generation of committed erythroid BFU-E and CFU-E progenitors does not require erythropoietin or the erythropoietin receptor. *Cell*, 83, 59-67.
- Wuebben, E. L. & Rizzino, A. **2017**. The dark side of SOX2: cancer - a comprehensive overview. *Oncotarget*, 8, 44917-44943.
- Xia, Z., Wei, P., Zhang, H., Ding, Z., Yang, L., Huang, Z., *et al.* **2013**. AURKA governs self-renewal capacity in glioma-initiating cells via stabilization/activation of  $\beta$ -catenin/Wnt signaling. *Mol Cancer Res*, 11, 1101-11.
- Xiong, W., Hsieh, C. C., Kurtz, A. J., Rabek, J. P. & Papaconstantinou, J. **2001**. Regulation of CCAAT/enhancer-binding protein-beta isoform synthesis by alternative translational initiation at multiple AUG start sites. *Nucleic Acids Res*, 29, 3087-98.
- Xu, Y., Tabe, Y., Jin, L., Watt, J., Mcqueen, T., Ohsaka, A., *et al.* **2008**. TGF-beta receptor kinase inhibitor LY2109761 reverses the anti-apoptotic effects of TGF-beta1 in myelo-monocytic leukaemic cells co-cultured with stromal cells. *Br J Haematol*, 142, 192-201.
- Yamada, O. & Kawauchi, K. **2013**. The role of the JAK-STAT pathway and related signal cascades in telomerase activation during the development of hematologic malignancies. *JAKSTAT*, 2, e25256.
- Yamagata, T., Maki, K. & Mitani, K. **2005**. Runx1/AML1 in normal and abnormal hematopoiesis. *Int J Hematol*, 82, 1-8.
- Yamaguchi, Y., Kurokawa, M., Imai, Y., Izutsu, K., Asai, T., Ichikawa, M., *et al.* **2004**. AML1 is functionally regulated through p300-mediated acetylation on specific lysine residues. *J Biol Chem*, 279, 15630-8.
- Yamamoto, R., Morita, Y., Ooehara, J., Hamanaka, S., Onodera, M., Rudolph, K. L., *et al.* **2013**. Clonal analysis unveils self-renewing lineage-restricted progenitors generated directly from hematopoietic stem cells. *Cell*, 154, 1112-1126.
- Yamanaka, R., Barlow, C., Lekstrom-Himes, J., Castilla, L. H., Liu, P. P., Eckhaus, M., *et al.* **1997**. Impaired granulopoiesis, myelodysplasia, and early lethality in CCAAT/enhancer binding protein epsilon-deficient mice. *Proc Natl Acad Sci U S A*, 94, 13187-92.
- Yamaura, T., Nakatani, T., Uda, K., Ogura, H., Shin, W., Kurokawa, N., *et al.* **2018**. A novel irreversible FLT3 inhibitor, FF-10101, shows excellent efficacy against AML cells with. *Blood*, 131, 426-438.
- Yan, J., Liu, Y., Lukasik, S. M., Speck, N. A. & Bushweller, J. H. **2004**. CBFbeta allosterically regulates the Runx1 Runt domain via a dynamic conformational equilibrium. *Nat Struct Mol Biol*, 11, 901-6.
- Yan, M., Kanbe, E., Peterson, L. F., Boyapati, A., Miao, Y., Wang, Y., *et al.* **2006**. A previously unidentified alternatively spliced isoform of t(8;21) transcript promotes leukemogenesis. *Nat Med*, 12, 945-9.

- Yanada, M., Matsuo, K., Suzuki, T., Kiyoi, H. & Naoe, T. **2005**. Prognostic significance of FLT3 internal tandem duplication and tyrosine kinase domain mutations for acute myeloid leukemia: a meta-analysis. *Leukemia*, 19, 1345-9.
- Yang, G., Khalaf, W., Van De Locht, L., Jansen, J. H., Gao, M., Thompson, M. A., *et al.* **2005**. Transcriptional repression of the Neurofibromatosis-1 tumor suppressor by the t(8;21) fusion protein. *Mol Cell Biol*, 25, 5869-79.
- Yang, Z., Wenxia, C., Weiliang, Z., Hui, M., Ting, W. & Jian, Z. **2016**. Irf7 as a novel marker for predicting bone metastases of primary prostate cancer. *Journal of Clinical Oncology*, 34, e16581-e16581.
- Yap, C., Loh, M. T., Heng, K. K., Tan, P., Yu, S. L., Chan, S. H., *et al.* **2000**. Variability in CD34+ cell counts in umbilical cord blood: implications for cord blood transplants. *Gynecol Obstet Invest*, 50, 258-9.
- Yarmus, M., Woolf, E., Bernstein, Y., Fainaru, O., Negreanu, V., Levanon, D., *et al.* **2006**. Groucho/transducin-like Enhancer-of-split (TLE)-dependent and -independent transcriptional regulation by Runx3. *Proc Natl Acad Sci U S A*, 103, 7384-9.
- Yeh, J. R., Munson, K. M., Elagib, K. E., Goldfarb, A. N., Sweetser, D. A. & Peterson, R. T. **2009**. Discovering chemical modifiers of oncogene-regulated hematopoietic differentiation. *Nat Chem Biol*, 5, 236-43.
- Yen, K., Travins, J., Wang, F., David, M. D., Artin, E., Straley, K., *et al.* **2017**. AG-221, a First-in-Class Therapy Targeting Acute Myeloid Leukemia Harboring Oncogenic. *Cancer Discov*, 7, 478-493.
- Yergeau, D. A., Hetherington, C. J., Wang, Q., Zhang, P., Sharpe, A. H., Binder, M., *et al.* **1997**. Embryonic lethality and impairment of haematopoiesis in mice heterozygous for an AML1-ETO fusion gene. *Nat Genet*, 15, 303-6.
- Yoon, K., Zhu, S., Ewing, S. J. & Smart, R. C. **2007**. Decreased survival of C/EBP beta-deficient keratinocytes is due to aberrant regulation of p53 levels and function. *Oncogene*, 26, 360-7.
- Yoshida, T., Mihara, K., Takei, Y., Yanagihara, K., Kubo, T., Bhattacharyya, J., *et al.* **2016**. All-trans retinoic acid enhances cytotoxic effect of T cells with an anti-CD38 chimeric antigen receptor in acute myeloid leukemia. *Clin Transl Immunology*, 5, e116.
- Yoshihara, H., Arai, F., Hosokawa, K., Hagiwara, T., Takubo, K., Nakamura, Y., *et al.* **2007**. Thrombopoietin/MPL signaling regulates hematopoietic stem cell quiescence and interaction with the osteoblastic niche. *Cell Stem Cell*, 1, 685-97.
- Ysebaert, L., Chicanne, G., Demur, C., De Toni, F., Prade-Houdellier, N., Ruidavets, J. B., *et al.* **2006**. Expression of beta-catenin by acute myeloid leukemia cells predicts enhanced clonogenic capacities and poor prognosis. *Leukemia*, 20, 1211-6.
- Yuan, Y., Zhou, L., Miyamoto, T., Iwasaki, H., Harakawa, N., Hetherington, C. J., *et al.* **2001**. AML1-ETO expression is directly involved in the development of acute myeloid leukemia in the presence of additional mutations. *Proc Natl Acad Sci U S A*, 98, 10398-403.
- Yusenko, M. V., Trentmann, A., Casolari, D. A., Abdel Ghani, L., Lenz, M., Horn, M., *et al.* **2021**. C/EBP $\beta$  is a MYB- and p300-cooperating pro-leukemogenic factor and promising drug target in acute myeloid leukemia. *Oncogene*, 40, 4746-4758.
- Zahnow, C. A. **2002**. CCAAT/enhancer binding proteins in normal mammary development and breast cancer. *Breast Cancer Res*, 4, 113-21.
- Zahnow, C. A. **2009**. CCAAT/enhancer-binding protein beta: its role in breast cancer and associations with receptor tyrosine kinases. *Expert Rev Mol Med*, 11, e12.

- Zahnow, C. A., Younes, P., Laucirica, R. & Rosen, J. M. **1997**. Overexpression of C/EBPbeta-LIP, a naturally occurring, dominant-negative transcription factor, in human breast cancer. *J Natl Cancer Inst*, 89, 1887-91.
- Zaidi, S. K., Javed, A., Choi, J. Y., Van Wijnen, A. J., Stein, J. L., Lian, J. B., *et al.* **2001**. A specific targeting signal directs Runx2/Cbfa1 to subnuclear domains and contributes to transactivation of the osteocalcin gene. *J Cell Sci*, 114, 3093-102.
- Zaret, K. S. & Carroll, J. S. **2011**. Pioneer transcription factors: establishing competence for gene expression. *Genes Dev*, 25, 2227-41.
- Zarrinkar, P. P., Gunawardane, R. N., Cramer, M. D., Gardner, M. F., Brigham, D., Belli, B., *et al.* **2009**. AC220 is a uniquely potent and selective inhibitor of FLT3 for the treatment of acute myeloid leukemia (AML). *Blood*, 114, 2984-92.
- Zeng, C., Mcneil, S., Pockwinse, S., Nickerson, J., Shopland, L., Lawrence, J. B., *et al.* **1998**. Intranuclear targeting of AML/CBFalpha regulatory factors to nuclear matrix-associated transcriptional domains. *Proc Natl Acad Sci U S A*, 95, 1585-9.
- Zhang, D. E., Hetherington, C. J., Meyers, S., Rhoades, K. L., Larson, C. J., Chen, H. M., *et al.* **1996**. CCAAT enhancer-binding protein (C/EBP) and AML1 (CBF alpha2) synergistically activate the macrophage colony-stimulating factor receptor promoter. *Mol Cell Biol*, 16, 1231-40.
- Zhang, D. E., Zhang, P., Wang, N. D., Hetherington, C. J., Darlington, G. J. & Tenen, D. G. **1997**. Absence of granulocyte colony-stimulating factor signaling and neutrophil development in CCAAT enhancer binding protein alpha-deficient mice. *Proc Natl Acad Sci U S A*, 94, 569-74.
- Zhang, E. & Xu, H. **2017**. A new insight in chimeric antigen receptor-engineered T cells for cancer immunotherapy. *J Hematol Oncol*, 10, 1.
- Zhang, H., Nguyen-Jackson, H., Panopoulos, A. D., Li, H. S., Murray, P. J. & Watowich, S. S. **2010**. STAT3 controls myeloid progenitor growth during emergency granulopoiesis. *Blood*, 116, 2462-71.
- Zhang, J., Hug, B. A., Huang, E. Y., Chen, C. W., Gelmetti, V., Maccarana, M., *et al.* **2001**. Oligomerization of ETO is obligatory for corepressor interaction. *Mol Cell Biol*, 21, 156-63.
- Zhang, J., Kalkum, M., Yamamura, S., Chait, B. T. & Roeder, R. G. **2004a**. E protein silencing by the leukemogenic AML1-ETO fusion protein. *Science*, 305, 1286-9.
- Zhang, L., Fried, F. B., Guo, H. & Friedman, A. D. **2008a**. Cyclin-dependent kinase phosphorylation of RUNX1/AML1 on 3 sites increases transactivation potency and stimulates cell proliferation. *Blood*, 111, 1193-200.
- Zhang, P., Iwasaki-Arai, J., Iwasaki, H., Fenyus, M. L., Dayaram, T., Owens, B. M., *et al.* **2004b**. Enhancement of hematopoietic stem cell repopulating capacity and self-renewal in the absence of the transcription factor C/EBP alpha. *Immunity*, 21, 853-63.
- Zhang, T., He, Y. M., Wang, J. S., Shen, J., Xing, Y. Y. & Xi, T. **2011a**. Ursolic acid induces HL60 monocytic differentiation and upregulates C/EBP $\beta$  expression by ERK pathway activation. *Anticancer Drugs*, 22, 158-65.
- Zhang, W., Konopleva, M., Shi, Y. X., Mcqueen, T., Harris, D., Ling, X., *et al.* **2008b**. Mutant FLT3: a direct target of sorafenib in acute myelogenous leukemia. *J Natl Cancer Inst*, 100, 184-98.
- Zhang, Y., Gao, S., Xia, J. & Liu, F. **2018**. Hematopoietic Hierarchy - An Updated Roadmap. *Trends Cell Biol*, 28, 976-986.
- Zhang, Y., Sif, S. & Dewille, J. **2007**. The mouse C/EBPdelta gene promoter is regulated by STAT3 and Sp1 transcriptional activators, chromatin remodeling and c-Myc repression. *J Cell Biochem*, 102, 1256-70.

- Zhang, Y., Strissel, P., Strick, R., Chen, J., Nucifora, G., Le Beau, M. M., *et al.* **2002**. Genomic DNA breakpoints in AML1/RUNX1 and ETO cluster with topoisomerase II DNA cleavage and DNase I hypersensitive sites in t(8;21) leukemia. *Proc Natl Acad Sci U S A*, 99, 3070-5.
- Zhang, Y., Wang, J., Wheat, J., Chen, X., Jin, S., Sadrzadeh, H., *et al.* **2013**. AML1-ETO mediates hematopoietic self-renewal and leukemogenesis through a COX/ $\beta$ -catenin signaling pathway. *Blood*, 121, 4906-16.
- Zhang, Y. Y., Li, X., Qian, S. W., Guo, L., Huang, H. Y., He, Q., *et al.* **2011b**. Transcriptional activation of histone H4 by C/EBP $\beta$  during the mitotic clonal expansion of 3T3-L1 adipocyte differentiation. *Mol Biol Cell*, 22, 2165-74.
- Zhang, Z. C., Zheng, L. Q., Pan, L. J., Guo, J. X. & Yang, G. S. **2015**. ZNF217 is overexpressed and enhances cell migration and invasion in colorectal carcinoma. *Asian Pac J Cancer Prev*, 16, 2459-63.
- Zhao, Q., Zhao, S., Li, J., Zhang, H., Qian, C., Wang, H., *et al.* **2019**. TCF7L2 activated HOXA-AS2 decreased the glucocorticoid sensitivity in acute lymphoblastic leukemia through regulating HOXA3/EGFR/Ras/Raf/MEK/ERK pathway. *Biomed Pharmacother*, 109, 1640-1649.
- Zhao, S., Zhang, Y., Sha, K., Tang, Q., Yang, X., Yu, C., *et al.* **2014**. KRAS (G12D) cooperates with AML1/ETO to initiate a mouse model mimicking human acute myeloid leukemia. *Cell Physiol Biochem*, 33, 78-87.
- Zhao, X., Jankovic, V., Gural, A., Huang, G., Pardnani, A., Menendez, S., *et al.* **2008**. Methylation of RUNX1 by PRMT1 abrogates SIN3A binding and potentiates its transcriptional activity. *Genes Dev*, 22, 640-53.
- Zheng, R., Friedman, A. D., Levis, M., Li, L., Weir, E. G. & Small, D. **2004a**. Internal tandem duplication mutation of FLT3 blocks myeloid differentiation through suppression of C/EBP $\alpha$  expression. *Blood*, 103, 1883-90.
- Zheng, X., Beissert, T., Kukoc-Zivojnov, N., Puccetti, E., Altschmied, J., Stolz, C., *et al.* **2004b**. Gamma-catenin contributes to leukemogenesis induced by AML-associated translocation products by increasing the self-renewal of very primitive progenitor cells. *Blood*, 103, 3535-43.
- Zhou, B. O., Yue, R., Murphy, M. M., Peyer, J. G. & Morrison, S. J. **2014**. Leptin-receptor-expressing mesenchymal stromal cells represent the main source of bone formed by adult bone marrow. *Cell Stem Cell*, 15, 154-68.
- Zhou, J., Ching, Y. Q. & Chng, W. J. **2015**. Aberrant nuclear factor-kappa B activity in acute myeloid leukemia: from molecular pathogenesis to therapeutic target. *Oncotarget*, 6, 5490-500.
- Zhou, J. D., Ma, J. C., Zhang, T. J., Li, X. X., Zhang, W., Wu, D. H., *et al.* **2017**. High bone marrow. *Oncotarget*, 8, 91979-91989.
- Zhu, M., Lee, G. D., Ding, L., Hu, J., Qiu, G., De Cabo, R., *et al.* **2007**. Adipogenic signaling in rat white adipose tissue: modulation by aging and calorie restriction. *Exp Gerontol*, 42, 733-44.
- Zhu, S., Oh, H. S., Shim, M., Sterneck, E., Johnson, P. F. & Smart, R. C. **1999**. C/EBP $\beta$  modulates the early events of keratinocyte differentiation involving growth arrest and keratin 1 and keratin 10 expression. *Mol Cell Biol*, 19, 7181-90.
- Zhu, X., Lv, J., Yu, L., Wu, J., Zou, S. & Jiang, S. **2009**. Proteomic identification of differentially-expressed proteins in squamous cervical cancer. *Gynecol Oncol*, 112, 248-56.
- Zuber, J., Rappaport, A. R., Luo, W., Wang, E., Chen, C., Vaseva, A. V., *et al.* **2011**. An integrated approach to dissecting oncogene addiction implicates a Myb-coordinated self-renewal program as essential for leukemia maintenance. *Genes Dev*, 25, 1628-40.

Zwergal, A., Quirling, M., Saugel, B., Huth, K. C., Sydlik, C., Poli, V., *et al.* **2006.** C/EBP beta blocks p65 phosphorylation and thereby NF-kappa B-mediated transcription in TNF-tolerant cells. *J Immunol*, 177, 665-72.

