

# **INVESTIGATING THE ROLE OF *CEBPA* & *MYBL2* IN HAEMATOLOGICAL DISORDERS**

**YARA OTHMAN ALYAHYAWI**



**A thesis submitted to University of Birmingham for the degree of  
DOCTOR OF PHILOSOPHY**

**Institute of Cancer and Genomics Science  
Collage of Medical and Dental Science  
University of Birmingham  
August 2023**

UNIVERSITY OF  
BIRMINGHAM

**University of Birmingham Research Archive**

**e-theses repository**

This unpublished thesis/dissertation is copyright of the author and/or third parties. The intellectual property rights of the author or third parties in respect of this work are as defined by The Copyright Designs and Patents Act 1988 or as modified by any successor legislation.

Any use made of information contained in this thesis/dissertation must be in accordance with that legislation and must be properly acknowledged. Further distribution or reproduction in any format is prohibited without the permission of the copyright holder.

## ABSTRACT

Myelodysplastic syndromes (MDS) are a diverse group of clonal haematological disorders that affect the proliferation and differentiation of haematopoietic stem cells, in which the bone marrow fails to produce mature blood cells. There is no cure for MDS; more than half of patients with MDS do not respond to current treatments, and 40% of MDS cases progress to acute myeloid leukaemia (AML). Consequently, there is an urgent clinical need to understand the mechanisms underlying these blood disorders to develop effective diagnostic and therapeutic strategies.

Through this thesis, we studied the contribution of two transcription factors, C/EBP $\alpha$  and MYBL2, to the MDS disease phenotype. The *CEBPA* gene encodes the C/EBP $\alpha$  protein, a transcription factor required for haematopoietic stem cell self-renewal and granulocytic commitment. 8% of patients with MDS have mutations in *CEBPA* at the time of initial diagnosis, and 12% progress from MDS to AML. The main goal of this study is to understand the contribution of the *CEBPA* mutation which disrupts the C-terminal domain to the dysplastic phenotype observed in MDS. To accomplish this, we developed a novel *in vitro* model system that combines human induced pluripotent stem cells (hiPSCs) with CRISPR-Cas9 technology to produce a heterozygous mutation in *CEBPA*, which results in the disruption of the DNA binding domain.

Our data show that the hiPSC-*CEBPA*<sup>+/-mut</sup> could differentiate into hematopoietic stem/progenitor cells (HSPCs) (CD34<sup>+</sup>/CD43<sup>+</sup>, CD34<sup>+</sup>/CD45<sup>+</sup>). HSPCs generated from hiPSC-*CEBPA*<sup>+/-mut</sup> clones differentiated in methylcellulose semi-solid media, albeit at a reduced capacity compared to isogenic control iPSCs. Moreover, consistent with previous studies showing the importance of *CEBPA* in regulating myeloid differentiation, HSPCs derived from hiPSC-*CEBPA*<sup>+/-mut</sup> showed an inability to form granulocytic colony-forming units (CFU-G); they displayed altered expression of myeloid differentiation-required genes,

such as *PUL1*, *GATA2*, and *RUNX1*, and exhibited a high percentage of aberrant myeloid cells indicated by the presence of pseudo Pelger-Huët anomaly. HSPCs derived from hiPSC-*CEBPA*<sup>+/mut</sup> promoted aberrant erythroid differentiation, as evidenced by the elevated expression of *EPO-R* and *TFRC* and the presence of aberrant morphology, such as multi-nucleated erythroblasts.

The other transcription factor studied in this thesis is *MYBL2*, a gene located in the long arm of chromosome 20, commonly deleted in 5-10% of MDS and AML patients (del20q). *MYBL2* is essential for sustaining the balance between self-renewal and the differentiation of haemopoietic stem cells (HSCs). In MDS, *MYBL2* has been found to play a tumour suppressor role. Studies have shown that half of patients with MDS have low levels of *MYBL2*, regardless of del20q cytogenetic abnormalities, indicating that dysregulation of *MYBL2* could significantly affect MDS progression. Attempts to generate hiPSCs with low levels of *MYBL2* using CRISPR-Cas9 technology failed; thus, we generated three different human leukaemia cell lines to downregulate *MYBL2* by lentiviral transduction with doxycycline-inducible *MYBL2*shRNA. After generating and validating these *MYBL2* and isogenic control cell lines, we studied their erythroid lineage differentiation capacity after cytarabine (Ara-C) treatment. The results showed that reducing *MYBL2* in leukaemia cell lines did not affect cell numbers or differentiation when treated with 1  $\mu$ M Ara-C to drive erythroid differentiation. Our work using different experimental systems highlights the advantages of hiPSCs in studying the role of these transcription factors to elucidate the molecular mechanisms underlying the initiation and progression of haematological disorders.

## **ACKNOWLEDGMENTS**

Words cannot express my gratitude to everyone who helped me during my PhD journey.

I am immensely grateful to my God for granting me the strength and blessings that sustained me during times of uncertainty and endowed me with the perseverance required to overcome obstacles.

I thank my supervisor, Dr Paloma García. Her invaluable guidance has propelled me forward from beginning my research project to completing my thesis. I appreciate her patience, constructive feedback, motivation, encouragement during moments of self-doubt, and generosity in providing knowledge and insightful suggestions over the last 4 years.

I am also profoundly grateful to the members of the García group who have been my companions on this journey, past and present (Rachel, Nuria, Ruba, and Ola), for their precious advice and continuous assistance while conducting my research. Their companionship and expertise in sharing knowledge enhanced my research experience.

I would be remiss if I did not acknowledge my parents (Mama and Baba) for their unwavering love, thoughts, prayers, and sacrifices that have been instrumental in my success. I extend my thanks for my backbone, Mama, thank you from the bottom of my heart; I could not have completed this journey without your enormous assistance and emotional support. I thank my brothers (Ahmed, Faisal, Khalid, Abdul-Aziz, Nawaf, and Mohmmad). Their faith in me has kept my spirits and motivation high during these years. I am also grateful to my sweetheart, Tala, for her ability to endure the long distances apart and the moments when I did not exist. Thanks to all my Saudi and non-Saudi friends for their help, encouragement, moments of laughter, and joy, and to my friend Hanin for long phone calls when we were depressed or homesick. Finally, I would like to express my profound appreciation to my home country, Saudi Arabia. My research would not have been possible without the generous funding the Saudi Arabian Culture Bureau provided.

# TABLE OF CONTENTS

|  |           |
|--|-----------|
| <b>Chapter 1: Introduction</b> .....   | <b>1</b>  |
| <b>1.1 Haematopoiesis</b> .....  | <b>2</b>  |
| 1.1.1 Haematopoietic stem cells (HSCs).....  | 3         |
| 1.1.2 Models of Haematopoietic stem cells.....                                       | 3         |
| 1.1.3 Haematopoietic factors.....  | 10        |
| <b>1.2 Myelodysplastic syndromes (MDS)</b> .....                                     | <b>12</b> |
| 1.2.1 Classification of MDS .....  | 13        |
| 1.2.2 Clonal evolution in MDS.....   | 18        |
| 1.2.3 Chromosomal abnormalities in MDS.....  | 22        |
| 1.2.4 Somatic mutations in MDS .....   | 25        |
| 1.2.5 Mutations in transcription factors.....  | 28        |
| <b>1.3 <i>CEBPA</i></b> .....  | <b>29</b> |
| 1.3.1 <i>CEBPA</i> mutations in haematological disorders.....                        | 30        |
| <b>1.4 <i>MYBL2</i></b> .....  | <b>34</b> |
| 1.4.1 Role of <i>MYBL2</i> in the maintenance of stem cells and differentiation..... | 34        |
| 1.4.2 <i>MYBL2</i> in haematological disorders .....                                 | 36        |
| <b>1.5 Therapeutic approaches of MDS</b> .....                                       | <b>37</b> |
| <b>1.6 Human induced pluripotent stem cells (hiPSCs)</b> .....                       | <b>39</b> |
| 1.6.1 Advantages and potential applications of hiPSCs.....                           | 41        |
| <b>1.7 Genome editing using the CRISPR–Cas9 technology</b> .....                     | <b>43</b> |
| 1.7.1 Mechanisms of CRISPR-Cas9 tools.....   | 43        |
| 1.7.2 Application of CRISPR-Cas9 in haematological disorder research.....            | 46        |
| <b>1.8 Aim &amp; hypothesis</b> .....  | <b>49</b> |
| <b>Chapter 2: Materials and Methods</b> .....  | <b>51</b> |
| <b>2.1 Cell culture</b> .....  | <b>52</b> |
| 2.1.1 Human iPSC culture methods .....   | 52        |
| 2.1.2 Sub culturing of human cells.....  | 54        |
| <b>2.2 Lentivirus production</b> .....   | <b>55</b> |
| 2.2.1 Virus preparation and transduction .....                                       | 56        |
| <b>2.3 Treatment of leukaemia cells with Ara-C</b> .....                             | <b>56</b> |
| <b>2.4 Alkaline phosphatase stain</b> .....  | <b>57</b> |
| <b>2.5 Identification of pluripotent markers</b> .....                               | <b>58</b> |
| 2.5.1 Immunofluorescent analysis for cytoplasmic pluripotent marker .....            | 58        |
| 2.5.2 Immunofluorescence staining for nuclear pluripotent markers.....               | 58        |
| <b>2.6 Differentiation of hiPSCs</b> .....   | <b>59</b> |
| 2.6.1 Evaluation of trilineage differentiation of hiPSCs .....                       | 59        |
| 2.6.2 Hematopoietic differentiation (STEMdiff).....                                  | 61        |
| 2.6.3 Clonogenic progenitor assay (Colony forming assay) .....                       | 61        |
| 2.6.4 Erythroid differentiation in liquid culture .....                              | 62        |
| 2.6.5 Myeloid differentiation in liquid culture .....                                | 63        |
| <b>2.7 Cytogenetics</b> .....  | <b>64</b> |
| <b>2.8 Flow cytometric analysis of differentiation markers</b> .....                 | <b>64</b> |
| <b>2.9 Kwik-Diff staining</b> .....  | <b>66</b> |

|   |            |
|---|------------|
| <b>2.10 Genomic DNA extraction .....</b>  | <b>67</b>  |
| <b>2.11 RNA analysis .....</b>  | <b>67</b>  |
| 2.11.1 RNA isolation.....   | 67         |
| 2.11.2 cDNA synthesis .....   | 68         |
| 2.11.3 Quantitative real-time PCR (qPCR) .....  | 68         |
| <b>2.12 CRISPR-Cas9 to modify <i>CEBPA</i> and <i>MYBL2</i> in hiPSCs .....</b>   | <b>69</b>  |
| 2.12.1 Designing single guide RNA (sgRNA) .....   | 69         |
| 2.12.2 Annealing sgRNA oligos.....  | 69         |
| 2.12.3 Cloning of gRNA into the pSpCas9 (BB)-2A-GFP (PX458) .....   | 70         |
| <b>2.13 Competent bacteria transformation.....</b>  | <b>70</b>  |
| <b>2.14 DNA extraction from plasmids (Miniprep and Maxiprep).....</b>   | <b>70</b>  |
| <b>2.15 Electroporation and selection of hiPSCs.....</b>  | <b>72</b>  |
| <b>2.16 Mutational screening following CRISPR-Cas9 editing.....</b>   | <b>73</b>  |
| <b>2.17 T7 Endonuclease I (T7EI) assay .....</b>  | <b>73</b>  |
| <b>2.18 Agarose gel extraction .....</b>  | <b>74</b>  |
| <b>2.19 Western blot analysis.....</b>  | <b>74</b>  |
| 2.19.1 Protein extraction and quantification.....   | 74         |
| 2.19.2 Protein quantification .....   | 75         |
| 2.19.3 SDS-PAGE and immunoblotting .....  | 75         |
| <b>2.20 Statistical analysis .....</b>  | <b>76</b>  |
| <b>Chapter 3: Generation and Characterisation of hiPSC with a <i>CEBPA</i> monoallelic mutation .....</b>   | <b>77</b>  |
| <b>3.1 Introduction.....</b>  | <b>78</b>  |
| <b>3.2 Results .....</b>  | <b>80</b>  |
| 3.2.1 Generation and characterisation of hiPSC clones with a <i>CEBPA</i> mutation (hiPSC- <i>CEBPA</i> <sup>+/-mut</sup> ) .....   | 80         |
| 3.2.2 hiPSC – <i>CEBPA</i> <sup>+/-mut</sup> clones exhibit no changes in C/EBP $\alpha$ protein expression .....   | 88         |
| 3.2.3 hiPSC- <i>CEBPA</i> <sup>+/-mut</sup> clones do not show chromosomal instability .....  | 90         |
| 3.2.4 Pluripotency characterisation of hiPSC controls and hiPSC- <i>CEBPA</i> <sup>+/-mut</sup> clones.....   | 92         |
| <b>3.3 Discussion .....</b>   | <b>102</b> |
| 3.3.1 Successful generation of hiPSC lines harbouring a monoallelic <i>CEBPA</i> mutation disrupting the bZIP domain .....  | 102        |
| 3.3.2 A frameshift mutation in the mid region of one allele of <i>CEBPA</i> does not affect the protein level .   | 104        |
| 3.3.3 hiPSC – <i>CEBPA</i> <sup>+/-mut</sup> revealed a normal karyotype.....   | 105        |
| 3.3.4 The generated hiPSC cells are pluripotent .....   | 106        |
| <b>Chapter 4: Studying the contribution of monoallelic mutation of <i>CEBPA</i> to the myelodysplastic phenotype.....</b>   | <b>109</b> |
| <b>4.1 Introduction.....</b>  | <b>110</b> |
| <b>4.2 Results .....</b>  | <b>111</b> |
| 4.2.1 Successive emergence of haematopoietic stem/progenitor cell (HSPC) populations from hiPSC control and hiPSC- <i>CEBPA</i> <sup>+/-mut</sup> .....                               | 111        |
| 4.2.2 Evaluation of the differentiation potential of HSPC generated from hiPSC control and hiPSC – <i>CEBPA</i> <sup>+/-mut</sup> cells by colony assays in a semi-solid medium ..... | 116        |
| 4.2.3 Defining the role of <i>CEBPA</i> <sup>+/-mut</sup> in the proliferation and ability of self-renewal of HSPCs.....  | 124        |
| 4.2.4 Analysing the myeloid lineage potential of hiPSC – <i>CEBPA</i> <sup>+/-mut</sup> clones.....   | 126        |
| 4.2.5 Investigating <i>CEBPA</i> <sup>+/-mut</sup> target gene expression.....  | 132        |
| 4.2.6 Investigating the capacity of hiPSCs harbouring <i>CEBPA</i> <sup>+/-mut</sup> to differentiate into the erythroid lineage.....   | 135        |

|  |            |
|--|------------|
| <b>4.3 Discussion</b> .....  | <b>143</b> |
| 4.3.1 <i>CEBPA</i> <sup>+/<i>mut</i></sup> have no impact on hiPSC differentiation into HSPCs .....  | 143        |
| 4.3.2 <i>CEBPA</i> <sup>+/<i>mut</i></sup> diminishes the colony-forming potential of the myeloid lineage and increases erythroid lineage differentiation, but does not impact the self-renewal ability of HSPCs ..... | 145        |
| 4.3.3 <i>CEBPA</i> <sup>+/<i>mut</i></sup> affects granulocyte differentiation .....   | 147        |
| 4.3.4 <i>CEBPA</i> <sup>+/<i>mut</i></sup> affects the expression of TFs that regulate myeloid development.....  | 149        |
| 4.3.5 <i>CEBPA</i> <sup>+/<i>mut</i></sup> does not block erythroid differentiation but promotes dyserythropoiesis.....  | 151        |
| <b>Chapter 5: Investigating the role of MYBL2 in haematological disorders</b> .....  | <b>153</b> |
| <b>5.1 Introduction</b> .....  | <b>154</b> |
| <b>5.2 Result:</b> .....   | <b>156</b> |
| 5.2.1 Generation of stable hiPSC with low MYBL2 expression.....  | 156        |
| 5.2.2 Downregulation of MYBL2 by shRNA in leukaemia cell line .....  | 160        |
| 5.2.3 Studying the effect of Ara-C on cell proliferation.....  | 166        |
| 5.2.4 Ara-C induces erythroid differentiation .....  | 173        |
| <b>5.3 Discussion:</b> .....   | <b>184</b> |
| 5.3.1 Engineering human hiPSCs with low <i>MYBL2</i> levels by CRISPR-Cas9 .....   | 184        |
| 5.3.2 Use of human cell lines to study the role of MYBL2 in erythroid and myeloid differentiation .....  | 185        |
| 5.3.3 Ara-C affects cell proliferation.....  | 187        |
| 5.3.4 Ara-C induces differentiation of leukaemia cells .....   | 189        |
| <b>Chapter 6: General discussion</b> .....   | <b>191</b> |
| <b>6.1 iPSCs as a model system for studying haematological disorders</b> .....   | <b>192</b> |
| <b>6.2 <i>CEBPA</i><sup>+/<i>mut</i></sup> block granulocyte differentiation and lead to myeloid and erythroid dysplasia</b> .....   | <b>193</b> |
| <b>6.3 The impact of MYBL2 on cell proliferation and differentiation in leukaemia cells was negligible, whereas Ara-C demonstrated a significant effect</b> .....  | <b>195</b> |
| <b>6.4 Limitations of this study</b> .....   | <b>197</b> |
| <b>6.5 Final conclusion and future work</b> .....  | <b>198</b> |
| <b>REFERENCES</b> .....  | <b>200</b> |
| <b>SUPPLEMENTARY</b> .....   | <b>235</b> |



## LIST OF FIGURES

|  |     |
|--|-----|
| Figure 1.1 Classical models for adult hematopoietic stem cell lineage commitment.....  | 6   |
| Figure 1.2 Advanced model of HSCs. ....  | 9   |
| Figure 1.3 Clonal haematopoiesis of indeterminate potential (CHIP) is identified at hotspots in non-cancer individuals. ....                         | 20  |
| Figure 1.4 C/EBP $\alpha$ protein structure.....   | 30  |
| Figure 1.5 Timelines of evolution of <i>CEBPA</i> mutations in AML. ....   | 33  |
| Figure 1.6 Diagram of iPSC application.....  | 42  |
| Figure 1.7 CRISPR-Cas9 principle.....  | 45  |
| Figure 3.1 <i>CEBPA</i> sequence. ....   | 81  |
| Figure 3.2 Utilising CRISPR-Cas9 in hiPSCs. ....   | 83  |
| Figure 3.3 T7E1 mismatch cleavage assay. ....  | 84  |
| Figure 3.4 Sanger sequencing. ....   | 86  |
| Figure 3.5 Sequence of C/EBP $\alpha$ protein. ....  | 87  |
| Figure 3.6 Western blot analysis of C/EBP $\alpha$ protein.....  | 89  |
| Figure 3.7 Chromosomal spread analyses for hiPSC control, CRISPR control and hiPSC- <i>CEBPA</i> <sup>+mut</sup> clones. ....                        | 91  |
| Figure 3.8 Morphology of human-induced pluripotent stem cells (hiPSCs). ....   | 93  |
| Figure 3.9 The hiPSCs and hiPSC- <i>CEBPA</i> <sup>+mut</sup> clones are positive for alkaline phosphatase. ....                                     | 95  |
| Figure 3.10 The hiPSC and hiPSC - <i>CEBPA</i> <sup>+mut</sup> clones express pluripotent markers.....   | 97  |
| Figure 3.11 Schematic illustration of trilineage differentiation assay. ....   | 99  |
| Figure 3.12 The hiPSC and hiPSC- <i>CEBPA</i> <sup>+mut</sup> clones differentiate into the three germ layers. ....                                  | 101 |
| Figure 4.1 Schematic representation and timeline of HSPC differentiation using the haematopoietic STEMdiff protocol from STEMCELL Technologies. .... | 112 |
| Figure 4.2 HSPC derived from <i>CEBPA</i> <sup>+mut</sup> strongly expressed CD43+. ....   | 114 |
| Figure 4.3 The hiPSC- <i>CEBPA</i> <sup>+mut</sup> can generate definitive HSPC (CD34 <sup>+</sup> CD45 <sup>+</sup> ). ....                         | 115 |
| Figure 4.4 Haematopoietic differentiation potential of HSPC-derived iPSCs in semi-solid medium. ....   | 117 |
| Figure 4.5 Type of colony-forming units in HSPC derived from hiPSC <i>CEBPA</i> <sup>+mut</sup> clones. ....   | 119 |
| Figure 4.6 Morphology of colony-forming units in HSPCs derived from hiPSC <i>CEBPA</i> <sup>+mut</sup> clones. ....                                  | 120 |
| Figure 4.7 Characterisation of colony forming units in hiPSC- control and hiPSC <i>CEBPA</i> <sup>+mut</sup> clones. ....                            | 122 |
| Figure 4.8 Morphological assessment of cells from clonogenic assay.....  | 123 |
| Figure 4.9 Analysis of self-renewal capacity of CFUs.....  | 125 |
| Figure 4.10 Schematic representation method and timeline of myeloid differentiation.....   | 126 |
| Figure 4.11 Analysis of myeloid differentiation of HSPCs derived from hiPSC- <i>CEBPA</i> <sup>+mut</sup> . ....                                     | 128 |
| Figure 4.12 Myeloid cells morphology.....  | 130 |
| Figure 4.14 Morphology analysis of aberrant myeloid cells. ....  | 131 |

|  |     |
|--|-----|
| Figure 4.15 Myeloid and erythroid marker expression in HSPCs and myeloid cells derived from hiPSC- <i>CEBPA</i> <sup>+/-mut</sup> clones. .... | 134 |
| Figure 4.16 Schematic representation method and timeline of erythroid differentiation. ....  | 136 |
| Figure 4.17 Analysis of erythroid differentiation of HSPCs derived from hiPSC- <i>CEBPA</i> <sup>+/-mut</sup> . ....                           | 137 |
| Figure 4.18 Analysis of erythroid differentiation derived from hiPSC- <i>CEBPA</i> <sup>+/-mut</sup> . ....                                    | 138 |
| Figure 4.19 Erythroid cell morphology. ....  | 141 |
| Figure 4.20 Morphology analysis of aberrant erythroid cells. ....  | 142 |
| Figure 5.1 <i>MYBL2</i> sequence. ....   | 157 |
| Figure 5.2 T7E1 mismatch cleavage assay. ....  | 158 |
| Figure 5.3 <i>MYBL2</i> protein expression detection in hiPSC ....   | 159 |
| Figure 5.4 Knockdown of <i>MYBL2</i> in K562, KG1a, and SKM-1. ....  | 161 |
| Figure 5.5 Knockdown of <i>MYBL2</i> in K562. ....   | 163 |
| Figure 5.6 Knockdown of <i>MYBL2</i> in KG1a. ....   | 164 |
| Figure 5.7 Knockdown of <i>MYBL2</i> in SKM-1. ....  | 165 |
| Figure 5.8 Schematic representations of inducing K563, KG1a, and SKM-1 cells with Ara-C. ....  | 166 |
| Figure 5.9 Ara-C treatment affects the proliferation of K652 cells. ....   | 168 |
| Figure 5.10 Ara-C treatment affects the proliferation of KG1a cells. ....  | 169 |
| Figure 5.11 Ara-C treatment affects the proliferation of SKM-1 cells. ....   | 170 |
| Figure 5.12 Ara-C affects cell proliferation. ....   | 172 |
| Figure 5.13 Ara-C induces erythroid differentiation in K562 cells. ....  | 175 |
| Figure 5.14 Ara-C induced erythroid differentiation in KG1a cells. ....  | 176 |
| Figure 5.15 Ara-C does not induce erythroid differentiation in SKM-1 cells. ....   | 177 |
| Figure 5.16 Morphology of Ara-C treated cell line. ....  | 179 |
| Figure 5.17 Erythroid and myeloid marker expression in Ara-C-treated K562 cells. ....  | 181 |
| Figure 5.18 Erythroid and myeloid marker expression in Ara-C-treated KG1a cells. ....  | 182 |
| Figure 5.19 Erythroid and myeloid marker expression in Ara-C-treated SKM-1 cells. ....   | 183 |

## LIST OF TABLES

|  |    |
|--|----|
| Table 1-1 International Prognostic Score System.....                                 | 15 |
| Table 1-2 Revised IPSS (IPSS-R).....   | 15 |
| Table 1-3 Classification of Myelodysplastic Syndrome according to WHO 2016 .....     | 17 |
| Table 1-4 Major driver genes.....  | 28 |
| Table 2-1 Type of cell line and source.....  | 54 |
| Table 2-2 Lentivirus plasmids .....  | 55 |
| Table 2-3 List of buffers used for alkaline phosphatase staining .....               | 57 |
| Table 2-4 List primary antibodies for immunofluorescence staining protocol.....      | 60 |
| Table 2-5 List of secondary antibodies for immunofluorescence staining protocol..... | 60 |
| Table 2-6 Cytokines for erythroid differentiation.....                               | 63 |
| Table 2-7 Cytokines for Myeloid differentiation.....                                 | 63 |
| Table 2-8 List of antibodies used for flow cytometry.....                            | 65 |
| Table 2-9 List of isotype antibodies used for flow cytometry .....                   | 66 |
| Table 2-10 TaqMan® assays (ThermoFisher Scientific).....                             | 69 |
| Table 2-11 KiCqStart™ primers KSPQ12012 (Sigma-Aldrich).....                         | 69 |

## LIST OF ABBREVIATIONS

|                           |  |
|---------------------------|--|
| -7                        | Monosomy 7   |
| +8                        | Trisomy 8  |
| Allogeneic HSCT           | Allogeneic Haematopoietic Stem Cell transplantation                          |
| AML                       | Acute myeloid leukaemia  |
| AP                        | Alkaline phosphatase   |
| Ara-C                     | Cytarabine, cytosine arabinoside   |
| BFU-E                     | Burst forming unit — erythroid   |
| BM                        | Bone marrow  |
| BMPs                      | Bone morphogenetic proteins  |
| BSA                       | Bovine serum albumin   |
| bZIP                      | Basic leucine zipper   |
| C/EBP- $\alpha$ & $\beta$ | CCAAT-enhancer-binding proteins Alfa & beta                                  |
| CDR                       | Critical region of deletion  |
| CFU-E                     | Colony forming unit — erythroid  |
| CFU-G                     | Colony forming unit — granulocytes   |
| CFU-GEMM                  | Colony forming unit—granulocytes/<br>erythrocytes/macrophages/megakaryocytes |
| CFU-GM                    | Colony forming unit—granulocytes/<br>macrophage                              |
| CFU-M                     | Colony forming unit —macrophage  |
| CLPs                      | Common lymphoid progenitors  |

|             |   |
|-------------|---|
| CMPs        | Common myeloid progenitors  |
| ChIP-seq    | Chromatin immunoprecipitation followed by sequencing                                      |
| CHIP        | Clonal haematopoiesis of indeterminate potential  |
| CMML        | Chronic myelomonocytic leukaemia  |
| CMRP        | Common myeloid repopulating progenitors   |
| CRISPR-Cas9 | Clustered regularly interspaced short palindromic repeats and CRISPR-associated protein 9 |
| CSF         | Colony stimulating factor   |
| DBD         | DNA binding domain  |
| del5q       | Chromosome 5q deletion  |
| del 20q     | Chromosome 20q deletion   |
| del 7q      | Chromosome 7q deletion  |
| DMSO        | Dimethylsulfoxide   |
| DNMT        | DNA methyltransferase   |
| DNMT3A      | DNA Methyltransferase 3A  |
| DSB         | Double-strand break   |
| EDTA        | Ethylenediaminetetraacetic acid   |
| EMT         | Epithelial-mesenchymal transition   |
| EPO         | Erythropoietin  |
| EOP-R       | EOP-receptor  |

|           |  |
|-----------|--|
| ESCs      | Embryonic stem cells                             |
| EZH2      | Enhancer of zest homolog 2                       |
| FAB       | French American British system                   |
| FACS      | Fluorescence-activated cell sorting              |
| FGF-4     | Fibroblast growth factor 4                       |
| FLT3      | Fms-like tyrosine kinase 3                       |
| FokI      | Flavobacterium okeanoikoites                     |
| FSC       | Forward Scatter                                  |
| GATA 1& 2 | GATA Binding Protein 1 & 2                       |
| GM-SCF    | Granulocyte-macrophage colony-stimulating factor |
| GMPs      | Granulocyte-macrophage progenitors               |
| HSPCs     | Hematopoietic stem and progenitor cells          |
| GYPA      | Glycophorin A                                    |
| G-CSF     | Granulocyte-colony stimulating factor            |
| Hb A      | Adult haemoglobin                                |
| Hb F      | Foetal haemoglobin                               |
| hiPSCs    | Human induced pluripotent stem cells             |
| HBB       | Human $\beta$ -globin                            |
| HDF       | Human dermal fibroblast                          |
| HMAAs     | Hypomethylating agents                           |
| HSCs      | Hematopoietic stem cell                          |
| HR        | Homologous recombination                         |
| IL-2      | Interleukin-2                                    |
| IL-3      | Interleukin-3                                    |

|         |   |
|---------|---|
| IL-6    | Interleukin-6   |
| IL-7    | Interleukin-7   |
| IPSS    | The International Prostate Symptom Score                  |
| IPSS-R  | Revised IPSS  |
| LMPPs   | Lymphoid-primed multipotent progenitors                   |
| LMPs    | Lymphoid–myeloid progenitors                              |
| LT      | Long term   |
| M-CSF   | Macrophage-colony stimulating factor                      |
| M-MLV   | Moloney murine leukaemia virus                            |
| MDS-MLD | Myelodysplastic syndrome with<br>multilineage dysplasia   |
| MDS-SLD | Myelodysplastic syndrome with single<br>lineage dysplasia |
| MDS-U   | Myelodysplastic syndrome -unclassifiable                  |
| MegNs   | Meganucleases   |
| MEPs    | Megakaryocytic-erythroid progenitors                      |
| (MERPs) | Megakaryocyte- erythroid repopulating<br>progenitors      |
| (MkRPs) | Megakaryocyte repopulating progenitors                    |
| MPL     | Myeloproliferative leukaemia virus<br>oncogene            |
| MPN     | Myeloproliferative neoplasms                              |
| MPPs    | Multipotent progenitors                                   |
| MYBL2   | MYB Proto-Oncogene Like 2                                 |
| c-Myc   | Myelocytomatosis viral oncogene                           |

|              |  |
|--------------|--|
| MDS          | Myeloid dysplastic syndrome                                      |
| NGS          | Next-generation sequencing                                       |
| NHEJ         | Nonhomologous end joining  |
| NK-MDS       | MDS with normal karyotypes                                       |
| NK           | Natural killer   |
| NPM1         | Nucleophosmin-1  |
| NRAS         | Neuroblastoma RAS viral oncogene                                 |
| OSKM         | OCT4, SOX2, KLF4, c-MYC  |
| PB           | Peripheral blood   |
| PBMNC        | Peripheral blood mononuclear cells                               |
| PBS          | Phosphate buffered saline  |
| PAM          | Protospacer-adjacent motif                                       |
| QPCR         | Quantitative Polymerase chain reaction                           |
| RA           | Refractory anaemia   |
| (RAEB-t)     | Refractory anaemia with an excess of the blast in transformation |
| RARS         | Refractory anaemia with ring sideroblasts                        |
| RAEB         | Refractory anaemia with the excess of blasts                     |
| RBCs         | Red blood cells  |
| RAEB1, RAEB2 | Refractory anemia with excess blasts 1, 2                        |
| RCMD         | Refractory cytopenias with multilineage dysplasia                |
| RCUD         | Refractory cytopenia with unilineage dysplasia                   |



|                             |   |
|-----------------------------|---|
| RFP                         | Red fluorescent protein                     |
| Red fluorescent protein     | Rho-associated coiled-coil protein kinase   |
| RS                          | Ring sideroblasts                           |
| RUNX-1                      | Runt-related transcription factor-1         |
| <i>S. pyogenes</i> ; SpCas9 | <i>Streptococcus pyogenes</i>               |
| sAML                        | Secondary Acute myeloid leukaemia           |
| SOX2                        | Sex-determining region Y box-2              |
| ST                          | Short-term                                  |
| SCD                         | Sickle cell disease                         |
| SCF                         | Stem cell factor                            |
| SFs                         | Splicing factors                            |
| Sh-RNA                      | Short hairpin RNA                           |
| SCC                         | Side Scatter                                |
| SNP                         | Single nucleotide polymorphism              |
| SF3B1                       | Splicing Factor 3b Subunit 1                |
| TALENs                      | Transcription activators- like endonuclease |
| TET2                        | Tet methylcytosine dioxygenase 2            |
| TFs                         | Transcription factors                       |
| U2AF1                       | U2 Small Nuclear RNA Auxiliary Factor 1     |
| vEGF                        | Vascular endothelial growth factor          |
| vWf                         | von Willebrand factor                       |
| WHO                         | World health organization                   |
| ZFNs                        | Zinc-finger nucleases                       |

# **Chapter 1: Introduction**

## 1.1 Haematopoiesis

Haematopoiesis is a process through which all blood cellular components are produced during the lifespan of an individual (Jagannathan-Bogdan and Zon 2013). This process occurs within the haematopoietic system, which comprises organs and tissues such as the bone marrow, spleen, and liver (Zhang et al. 2018). It is strictly controlled by homeostatic mechanisms to facilitate adaptation to challenges such as infection and bleeding (Essers et al. 2009; Calvi and Link 2015). In mammals, haematopoiesis occurs in overlapping successive waves during development (Elsaid et al. 2020). Haematopoietic cells arise from the mesoderm, which is known to produce the endothelial and haematopoietic lineages (Gritz and Hirschi 2016). The initial wave of haematopoiesis—primitive haematopoiesis—begins around embryonic day 7 in mice and around day 18–20 in humans to produce unipotent precursors that generate primitive erythrocytes, megakaryocytes, and macrophages (early EMPs) (Hoeffel et al. 2015; Neo et al. 2021; Vink et al. 2022). The second wave of haematopoiesis coincides with the beginning of definitive haematopoiesis and the appearance of both lymphoid–myeloid progenitors (LMPs) and EMPs (late EMPs) (McGrath et al. 2015a; Böiers et al. 2013). For the definitive haematopoietic waves, it is well established that haematopoiesis arises through an endothelial-to-haematopoietic transition from a specialised endothelial subpopulation called hemogenic endothelium (Zovein et al. 2008; Garcia-Alegria et al. 2018; Ottersbach 2019). A similar endothelial-to-haematopoietic transition arising from an HE-like intermediary known as a hemogenic angioblast has also recently been reported for primitive haematopoiesis (Lancrin et al. 2009; Stefanska et al. 2017). Despite the possibility of a shared cellular origin, not all waves begin from the same anatomical location. The initial wave originates extra-embryonically in the yolk sac. Conversely, the second wave of development occurs in the dorsal aorta in aorta gonad mesonephros (AGM) region of the embryo, whereas haematopoietic stem cells (HSCs) originate among the intra-aortic haematopoietic clusters (Boisset et al. 2010; Dzierzak and

Bigas 2018). The HSCs develop and multiply in the foetal liver before settling in the bone marrow (BM) (Neo et al. 2021).

Notably, a recent study by the group of Suda found through genetic tracing in mice that HSCs and defined progenitors are derived from precursor cells in the foetal liver in an HSC-independent manner. The study found that haematopoietic progenitors in foetal liver are not generated by gain of commitment to HSCs. On the contrary, HSCs and intermediate progenitors from different precursors converge to generate the hierarchical structure *in situ*, suggesting that the blueprint for the structure is embedded in the precursor state. This study did not observe any involvement of HSCs in the development of progenitors during late gestation, indicating that most blood cells in the embryo were formed in an HSC-independent manner (Yokomizo et al. 2022).

### **1.1.1 Haematopoietic stem cells (HSCs)**

In adults, HSCs are situated at the top of the haematopoietic hierarchy that develops in BM; HSCs have the ability to self-renew and produce multipotent and restricted-lineage progenitors (Rieger and Schroeder 2012). The self-renewal capacity of HSCs facilitates their repopulation without differentiation. Nevertheless, their multipotency enables them to differentiate into all types of blood cells that regulate homeostatic balance, immune function, and inflammatory response (Seita and Weissman 2010).

### **1.1.2 Models of Haematopoietic stem cells**

#### **1.1.2.1 Classical haematopoietic hierarchical model**

The concept of HSCs was first proposed by Till and McCulloch in 1961. They found that lethally irradiated animals transplanted with mouse BM cells produced colonies of haematopoietic cells in the spleen; their findings revealed that these colonies were derived from differentiated HSCs (Till and McCulloch, 1961; Becker et al., 1963). Subsequently, researchers used various techniques to isolate HSCs from the BM to elucidate their activity and molecular

regulatory network using antibodies, and fluorescence-activated cell sorting (FACS) facilitated the separation of HSCs. In 1988, Weissman and co-workers first identified cells enriched with HSC using various surface markers. (Spangrude et al. 1988). Consequently, numerous researchers have attempted to identify additional surface markers for HSC purification, including Sca-1, c-Kit, CD34, and signalling lymphocyte activation markers such as CD150 and CD 48, which have been widely utilised in isolating HSCs (Ikuta and Weissman 1992; Kiel et al. 2005; Oguro et al. 2013).

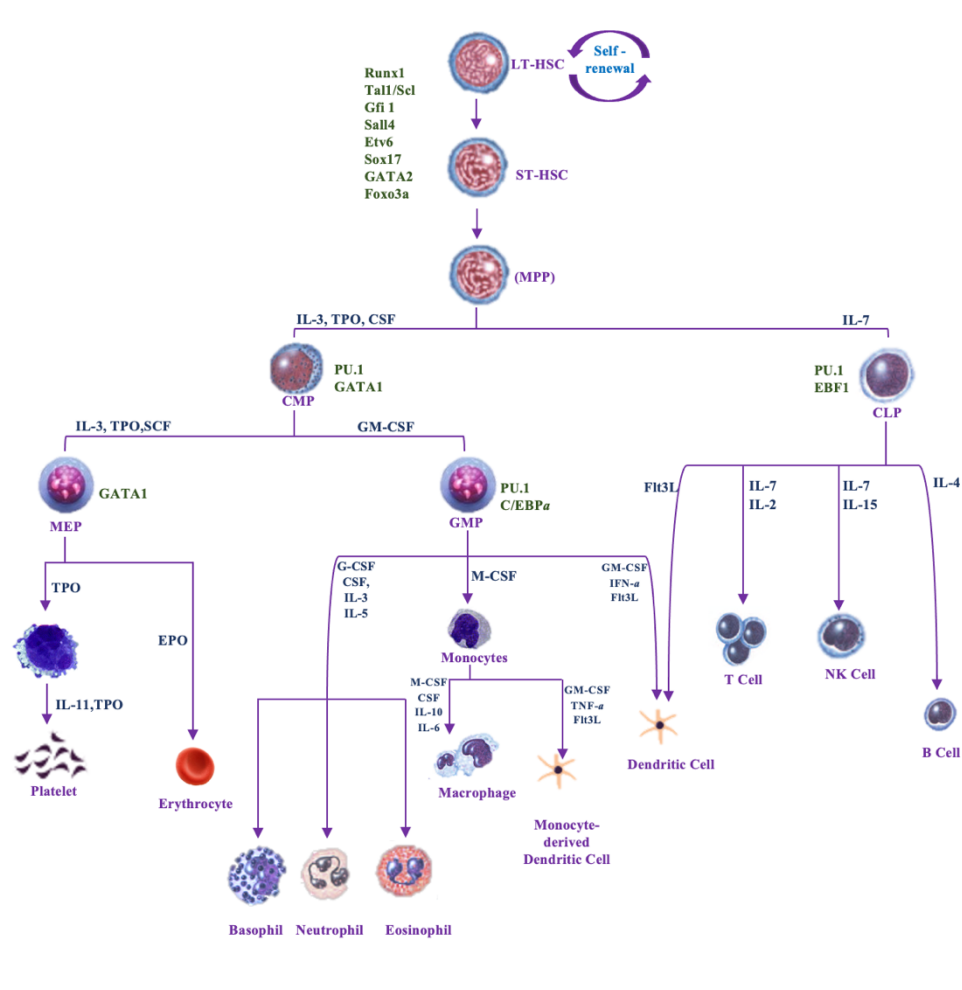
The two fundamental properties of HSCs—self-renewal and multipotent differentiation—enable them to generate blood cells of all lineages. In contrast, progenitors are characterised by their lack of self-renewal ability and commitment to lineage differentiation. (Orkin 2000; Dick 2003).

A group led by Weissman clarified the association between HSCs and their progeny, on one hand, and the sequential differentiation programme, on the other hand to create a tree hierarchy-based immunophenotype using mouse BM (Kondo et al. 1997; Akashi et al. 2000; Manz et al. 2002). In this conventional model, HSCs were categorised into two distinct subgroups based on CD34 expression: (i) CD34<sup>+</sup> short-term (ST) HSCs and (ii) CD34<sup>-</sup> long-term (LT) HSCs. ST HSCs possess only a one-month reconstitution capacity. In contrast, LT HSCs comprise a quiescent population in the BM with a long-term extension to 3–4 months or serial repopulation capacity. Additionally, they differentiate into ST HSCs, which then differentiate into multipotent progenitors (MPPs) that have no discernible capacity for self-renewal (Yang et al. 2005). To further identify the cell populations, CD34 and FLT3 markers were employed to differentiate between LT-HSCs, which were indicated by the expression of CD34<sup>-</sup> and FLT3<sup>-</sup> phenotype, ST-HSCs expressed CD34<sup>+</sup> and FLT3<sup>-</sup> phenotype. At the same time, MPPs were characterised by CD34<sup>+</sup> and FLT3<sup>+</sup> phenotypes and subfractions of the KSL population (Volpe et al. 2015).

MPPs differentiate into (i) common lymphoid progenitors (CLPs), which then differentiate into B cells, T cells, natural killer cells and dendritic cells, and (ii) common myeloid progenitors (CMPs) that give rise to myeloid cells, megakaryocytes and erythroid cells. CMPs are further divided into megakaryocyte–erythrocyte progenitors (MEPs), granulocyte–macrophage progenitors (GMPs) that generate monocytes, and granulocytes. Expression of CD34 and CD16/32 were used to distinguish between MEP, CMP, and GMP: (i) CD34<sup>-</sup> CD16/32<sup>-</sup> for MEPs (ii) CD34<sup>+</sup> CD16/32<sup>+</sup> are expressed in GMPs; and (iii) CD34<sup>-</sup> CD16/32<sup>+</sup> are expressed in CMPs (Akashi et al. 2000).

Figure 1.1 illustrates the hierarchically balanced structure differentiation programme of HSCs, turning them into into mature blood cells regulated by the cytokines and transcription factors (TFs). (Robb 2007; Zhu and Emerson 2002; Seita and Weissman 2010).

Subsequently, advances in genetic mouse models and single-cell technology have challenged the classical model of haematopoietic stem and progenitor cells (HSPCs), leading to the identification and study of new types of HSPCs as described in 1.1.2.2



**Figure 1.1 Classical models for adult hematopoietic stem cell lineage commitment.**

The classical model for hematopoietic development postulates that lineage commitment of long-term hematopoietic stem cells (LT-HSCs) sits at the apex of the hierarchy. LT-HSCs differentiate into ST-HSCs and then into MPPs with diminished capacity for self-renewal. MPPs divide into Common lymphoid progenitors (CLPs) and Common myeloid progenitors (CMPs). CLPs produce lymphocytes and dendritic cells. CMPs produce GMPs and MEPs. MEPs develop megakaryocytes and erythrocytes, while GMP produces granulocytes, macrophages, and dendritic cells. Haematopoietic differentiation programme regulated by transcription factors and cytokines.

### 1.1.2.2 Advances in research on haematopoietic hierarchy

Several research groups have studied intermediate-term HSCs, which are situated between ST- and LT-HSCs and have been found to play a significant role in reconstitution for a duration exceeding eight months after transplantation (Benveniste et al. 2010; Yamamoto et al. 2013). Using viral genetic barcoding and advanced sequencing, Lu and colleagues (2011) identified the heterogeneity of HSC phenotypes *in vivo* and demonstrated that HSCs did not equally give rise to progenitor cell.

A single irradiated mouse recipient exhibited two unique HSC differentiation pathways. The first category consisted of GMPs, MEPs and CLPs, which are progenitor cell populations, whereas the second category comprised mature lymphoid blood cells. Using single-cell transplantation, the group of Toshio Suda observed that phenotypically defined HSCs included self-renewing lineage-restricted progenitors, such as common myeloid repopulating progenitors, megakaryocyte–erythrocyte repopulating progenitors and megakaryocyte repopulating progenitors (Figure 1.2 A) (Yamamoto et al. 2013).

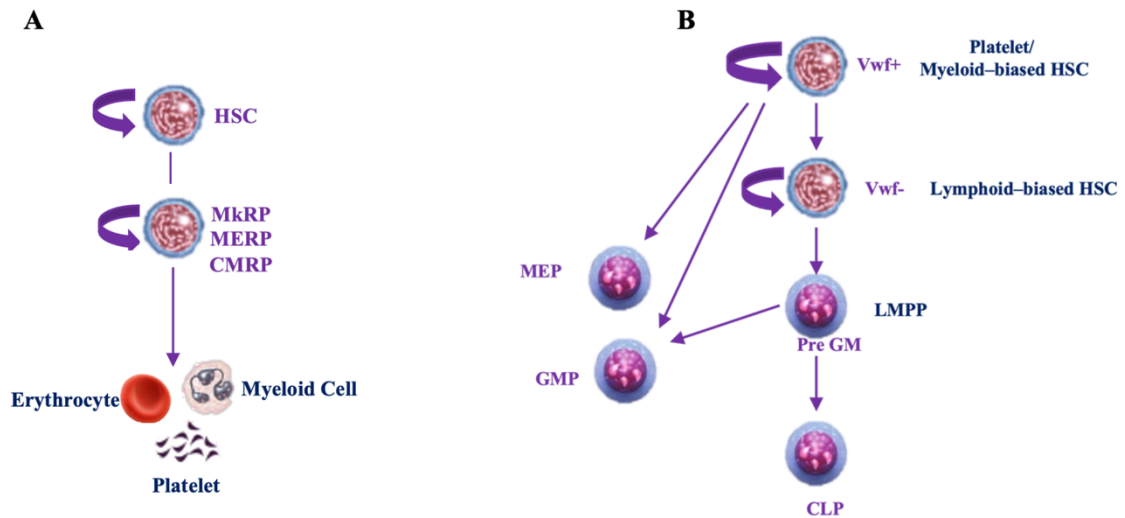
Scientists have questioned the origin of megakaryocytes for many years. A group led by Jacobsen discovered that lymphoid-primed MPPs (LMPPs) cannot generate erythrocyte and megakaryocyte lineages. However, they can differentiate into granulocytes, macrophages, and lymphoid lineages (Figure 1.2 B) (Adolfsson et al. 2005). Previous tracking lineage studies have revealed that LMPPs could give rise to erythrocyte and megakaryocyte lineages.

Subsequently, the expression of the surface receptor c-Kit and the megakaryocyte marker von Willebrand factor (vWF) suggested a platelet-biased yet multipotent HSC subpopulation (Sanjuan-Pla et al. 2013; Shin et al. 2014). Sanjuan-Pla and colleagues (2013) also proved that platelet-primed HSCs were at the top of the haematological hierarchy. Sanjuan-Pla and colleagues (2013) found that vWF was expressed in 25% of LT HSCs and that vWF+ HSCs were primed for platelet-specific gene expression and more likely to restore platelets over the



long term. The vWF<sup>+</sup> platelet-primed HSCs also showed their self-renewal capacity and ability to generate vWF-biased lymphoid HSCs and LT-biased HSCs of the myeloid lineage (Figure 1.2 B) (Sanjuan-Pla et al. 2013). Previous studies of paired cell transplantation have suggested that megakaryocyte progenitors are produced directly from HSCs (Figure 1.2 B) (e.g., Yamamoto et al. 2013). A more recent study has discovered that HSC progenitors committed to megakaryocytes have stem cell properties. According to Haas and colleagues (2015), this group of cells shares numerous characteristics with the HSCs. In situations of inflammatory stress, these cells are activated to efficiently produce platelets, indicating the possibility of a direct pathway from HSCs to megakaryocytes during inflammation.

Through the tracking precursors and mature blood lineage cells derived from transplanted HSCs, Jacobsen's group demonstrated that a unique group of HSCs could efficiently replenish megakaryocytes and platelets but did not contribute to the replenishment of other type of blood cells (Carrelha et al. 2018).



**Figure 1.2 Advanced model of HSCs.**

(A) The LT-HSC population in the myeloid bypass model contains Common myeloid repopulating progenitors (CMRPs), megakaryocyte–erythrocyte repopulating progenitors (MERPs), megakaryocyte repopulating progenitors (MkRPs) created directly from HSCs. (B) Platelet-myeloid-biased HSC (vWF+) set on the top of the hierarchy are capable of differentiating into MEP and GMP. vWF- lymphoid-biased HSCs were found downstream of vWF+ HSCs. Lymphoid-primed multipotent progenitors (LMPPs) can give rise to granulocyte/macrophage and CLP. HSCs can directly produce MEP.

### 1.1.3 Haematopoietic factors

Haematopoiesis is a tightly regulated process that maintains the equilibrium between immature and mature functioning cells while also increasing the production of mature cells in response to injury, stress, damage, inflammation, or infection. It is regulated by both intrinsic and external factors. For instance, activation and silencing of TFs at various stages of haematopoiesis are crucial inherent variables that regulate the differentiation process. Red and white blood cell synthesis is tightly regulated under healthy conditions. Proliferation and self-renewal of these cells are dependent on growth factors. Stem cell factor (SCF), which binds to the c-kit receptor on HSCs, is one of the key players in the self-renewal and growth of HSCs (Broudy 1997). Moreover, erythropoietin (EPO) is required to induce the differentiation of myeloid progenitor cells into the erythrocyte lineage (Tsiftoglou et al. 2009; Dzierzak and Philipsen 2013). In contrast, thrombopoietin (TPO) induces the differentiation of myeloid progenitor cells into megakaryocytes (thrombocyte-forming cells) (Hauke and Tarantolo 2000).

The regulators of progenitor cell development in various committed cell lineages, such as interleukins IL-2, IL-3, IL-6, and IL-7, are other key glycoprotein growth factors that govern cell proliferation and maturation (Ymer et al. 1985; Nitsche et al. 2003; Tie et al. 2019; Cordeiro Gomes et al. 2016). Other factors, known as colony-stimulating factors (CSFs), stimulate the production of committed cells, promote the formation of granulocytes, and are active in either progenitor or end cell production; these CSFs are granulocyte CSF (G-CSF), macrophage CSF (M-CSF) and granulocyte–macrophage CSF (GM-CSF) (Ketley and Newland 1997).

TFs are pleiotropic, playing distinct roles at various phases of development or differentiation of several cell types; the best example is the transcription factor stem cell leukaemia/T-cell acute lymphoblastic leukaemia (SCL/TAL). A review by Porcher et al. emphasises the

importance of SCL/TAL for mesoderm specification during development of haematopoietic cells (Porcher et al. 2017); the review focuses on the critical role of transcription factors in normal haematopoiesis and their aberrant expression or activity in hematologic malignancies. Aberrant expression of SCL/TAL is associated with T-cell acute lymphoblastic leukaemia (Porcher et al. 2017).

Furthermore, RUNX1, one of the three members of the RUNX family of transcription factors, is necessary for the development and normal functioning of definitive HSCs. De Bruijn and Dzierzak have highlighted the importance of RUNX1 by outlining its functions in developmental haematopoiesis and the various functions performed by RUNX1 (de Bruijn and Dzierzak 2017). Moreover, another companion review has discussed the role of RUNX1 factors in haematological disease, ranging from germline mutations that cause familial platelet disorder with associated myeloid malignancies to acquired mutations that induce acute myeloid leukaemia (AML) and numerous other conditions (Sood et al. 2017). Although SCL/TAL and RUNX1 influence the most primitive haematopoietic cells, specific haematopoietic lineages express and utilise different transcription factors, for example, C/EBP $\alpha$ , PU.1, GATA-1, and GATA-2.

C/EBP $\alpha$  has a central function in myeloid cells, and its expression promotes myeloid cell differentiation from MPPs and is essential for neutrophil maturation during normal haematopoiesis. Moreover, it engages with other TFs to regulate neutrophil and eosinophil development (Avellino and Delwel 2017). Furthermore, it regulates the state of quiescence in HSCs. Previous study shows that knockout of *CEBPA* enhanced proliferation, increasing the similarity between adult HSC and foetal HSC (Ye et al. 2013).

GATA-1 is expressed in erythroid, megakaryocytic, dendritic and mast cells and eosinophils and basophils (Martin et al. 1990; Leonard et al. 1993; Gutiérrez et al. 2007). Expression of GATA-1 and GATA-2 partially overlaps, for example, in yolk sac-derived primitive

erythroblasts, megakaryocytes, and eosinophils (Tsai and Orkin 1997; Fujiwara et al. 2004). In other contexts, the expression of these proteins is contradictory or mutually exclusive (Katsumura and Bresnick 2017).

GATA-2 is exclusively expressed in HSPCs and erythroid progenitors before GATA-1. The distinct patterns suggest that these GATA factors are responsible for distinct biological processes. (Weiss et al. 1994; Grass et al. 2003). Lastly, it has been reported that growth factors initiate the activation of transcription factors by activating signal transduction pathways. Depending on the combination of growth factors and the stage of cell differentiation, the effects of growth factors vary; for example, short-term expression of PU.1 results in the production of immature eosinophils, whereas long-term activation of PU.1 induces myeloid differentiation (Engel and Murre 1999).

Disruption of HSC function results in various haematological disorders, such as leukaemia, myeloma, lymphoma, and myelodysplastic syndrome.

## **1.2 Myelodysplastic syndromes (MDS)**

Myelodysplastic syndromes (MDS)—also called myelodysplasia—are a heterogeneous group of HSCs disorders characterised by inefficient haematopoiesis, in which the bone marrow cannot produce mature blood cells, leading to dysplasia, anaemia, and cytopenia despite normal or hypercellular BM (Syed et al. 2020). MDS is caused by disruption of HSC differentiation and maturation and changes in bone marrow stroma (Germing et al. 2013; Steensma 2018). In approximately 40% of patients, MDS progresses to AML (Adès et al. 2014; Pfeilstöcker et al. 2016). According to UK cancer research statistics, the annual incidence of MDS in the UK is approximately 3.7 per 100,000 individuals, predominately in elderly patients (median age during diagnosis 75.7 years) (Killick et al. 2021). Moreover, it is more common in men and smokers (Barzi and Sekeres 2010; Sekeres and Taylor 2022).

### 1.2.1 Classification of MDS

MDS classification facilitates the diagnoses and assists in the selection of treatment protocols. The original French–American–British (FAB) classification system established in 1972 provided guidelines for diagnosing MDS on the basis of dysplastic lineages and percentage of blast cells in the BM and the peripheral blood (Bennett et al. 1976).

The International Prognostic Scoring System (IPSS) established in 1997 by Greenberg and co-workers according to data were collected from patients in the United States, Japan and Europe (Table 1-1) (Greenberg et al. 1997). The predictive value and ease of use of the IPSS led to its widespread adoption. Although it has been the cornerstone of MDS outcome prediction for approximately 15 years, different organisations and academics have identified several drawbacks. The final analysis included MDS with 20–30% BM myeloblasts and chronic myelomonocytic leukaemia (CMML). Subsequently, both are excluded from the MDS classes in the subsequent World Health Organization (WHO) classifications. Moreover, age (more than 60 years) was omitted as a factor contributing to the final score because it has been demonstrated to have predictive importance for overall survival (OS) but not for time of progression to AML (Greenberg et al. 1997). MDS are arguably the most varied of all haematological malignancies. The IPSS classification is according to the percentage of bone marrow blasts, cytogenetic aberrations, and the levels of cytopenia. MDS is categorised into four subgroups based on time of progression to AML and OS: low, intermediate-1, intermediate-2, and high. The reversed IPSS or IPSS-R provides a deeper explanation for cytogenetic abnormalities in the illness characteristics listed in the IPSS (Table 1-2). The IPSS-R was developed using cytogenetic data from 7012 individuals with primary MDS, facilitating their categorisation into five risk groups (very low, low, intermediate, high, and very high) with varying estimates time of progression to AML and OS (Greenberg et al., 2012). Moreover,

platelet counts and levels of haemoglobin were scored according to whether the levels are 50–  
100 × 10<sup>9</sup>/L and 8–10 g/dL, respectively.

**Table 1-1 International Prognostic Score System**

Taken from (Greenberg et al. 1997)

| Variable                  | Parameter                                       | Score | Final Score | Risk Group     | LFS Median (Years) | OS Median (Years) |
|---------------------------|---|-------|-------------|----------------|--------------------|-------------------|
| Blasts in bone marrow (%) | <5  | 0     | 0           | Low            | 9.4                | 5.7               |
|                           | 5–10  | 0.5   |             |                |                    |                   |
|                           | 11–20   | 1.5   |             |                |                    |                   |
|                           | 21–30   | 2     |             |                |                    |                   |
| Cytogenetic aberrations   | Normal, del(5q), del(20q)                       | 0     | 0.5–1       | Intermediate-1 | 3.3                | 3.5               |
|                           | Other alterations                               | 0.5   | 1.5–2       | Intermediate-2 | 1.1                | 1.2               |
|                           | 3 or more alterations, Chromosome 7 aberrations | 1     |             |                |                    |                   |
| Number of cytopenias      | None or 1                                       | 0     | ≥2.5        | High           | 0.2                | 0.4               |
|                           | 2 or 3  | 0.5   |             |                |                    |                   |

LFS, leukaemia free survival; OS, overall survival.

**Table 1-2 Revised IPSS (IPSS-R)**

Taken from (Greenberg et al. 2012)

| Variable                  | Score   |                          | Final score | Risk group   | Median time to AML (years) | OS. median (years) |     |     |
|---------------------------|---|--------------------------|-------------|--------------|----------------------------|--------------------|-----|-----|
| Blasts in bone marrow (%) | <2  | 0                        | ≤1.5        | Very low     | Nr                         | 8.8                |     |     |
|                           | >2 to <5  | 1                        |             |              |                            |                    |     |     |
|                           | 5–10  | 2                        |             |              |                            |                    |     |     |
|                           | >10   | 3                        |             |              |                            |                    |     |     |
| Cytogenetic aberrations   | –y, del(11q)  | 0                        | 2–3         | Low          | 10.8                       | 5.3                |     |     |
|                           | Normal, del(5q), del(12p), del(20q), double including del(5q)                   | 1                        |             |              |                            |                    |     |     |
|                           | Del(7q), +8, +19, i(17q), any other single or double independent clones         | 2                        | 3.5–4.5     | Intermediate | 3.2                        | 3                  |     |     |
|                           | –7, inv(3)/t(3q)/del(3q), double including –7/del(7q), complex: 3 abnormalities | 3                        |             |              |                            |                    |     |     |
|                           | Complex: >3 abnormalities   | 4                        |             |              |                            |                    |     |     |
| Cytopenia                 | Hb (g/dl)   | ≥10                      | 5–6         | High         | 1.4                        | 1.6                |     |     |
|                           |   | 8–10                     |             |              |                            |                    | 1   |     |
|                           |   | <8                       |             |              |                            |                    | 1.5 |     |
|                           | Platelets (×10 <sup>9</sup> /l)   | >100                     | 0           | ≥6.5         | Very high                  | 0.7                | 0.8 |     |
|                           |   | 50–<100                  | 0.5         |              |                            |                    |     |     |
|                           |   | <50                      | 1           |              |                            |                    |     |     |
|                           |   | ANC(×10 <sup>9</sup> /l) | >0.8        |              |                            |                    |     | 0   |
|                           |   |                          | <0.8        |              |                            |                    |     | 0.5 |

OS, overall survival.



Next-generation sequencing (NGS) technologies have facilitated the study of molecular pathogenesis of MDS; therefore, the WHO introduced a new classification system in 2001. This system was based on the FAB structure but incorporated parameters describing the biology of the diseases, including cell counts, dysplastic changes, and genetic defects, which were revised in 2008 and 2016. The 2008 WHO classification of MDS is based on a mixture of clinical, immunophenotypic, genetic, and morphological of BM and peripheral blood characteristics, identifying six subgroups: refractory cytopenia with multilineage dysplasia (RCMD), myelodysplastic syndrome with isolated del5q (5q syndrome), myelodysplastic syndrome not classified refractory cytopenia (RA), refractory anaemia with excess blasts (RAEB1 and RAEB2), , and refractory anaemia with ring sideroblasts (RARS) (Vardiman et al. 2009).

The revised 2016 WHO classification (Table 1–3) removed cytopenic lineage factors (such as refractory anaemia) in favour of dysplastic lineage factors (MDS with single-lineage dysplasia). The most recent revision in the classification system introduced refinements in cytopenia, blast cell percentage, and morphological change as well as the effect of genetic information on MDS classification and diagnosis (Killick et al. 2021).

**Table 1-3 Classification of Myelodysplastic Syndrome according to WHO 2016**

Taken from (Hong and He 2017)

| Type                             | Dysplastic lineages | Cytopenias | Ring sideroblasts in erythroid elements of BM | Blasts                                  | Cytogenetics  |
|----------------------------------|---------------------|------------|---|---|---|
| MDS-SLD                          | 1                   | 1 or 2     | RS<15% (or <5%)                               | PB <1% BM <5% No Auer rods              | Any, unless fulfils criteria for isolated del(5q)                   |
| MDS-MLD                          | 2 or 3              | 1-3        | RS<15% (or <5%)                               | PB <1% BM <5% No Auer rods              | Any, unless fulfils criteria for isolated del(5q)                   |
| MDS-RS<br>MDS-RS-SLD             | 1                   | 1 or 2     | RS $\geq$ 15% (or $\geq$ 5%)                  | PB <1% BM <5% No Auer rods              | Any, unless fulfils criteria for isolated del(5q)                   |
| MDS-RS-MLD                       | 2 or 3              | 1-3        | RS $\geq$ 15% (or $\geq$ 5%)                  | PB <1% BM <5% No Auer rods              | Any, unless fulfils criteria for isolated del(5q)                   |
| MDS with isolated del(5q)        | 1-3                 | 1-2        | None or any                                   | PB <1% BM <5% No Auer rods              | Del(5q) alone or with 1 additional abnormality except -7 or del(7q) |
| MDS-EB<br>MDS-EB-1               | 0-3                 | 1-3        | None or any                                   | PB 2 ~ 4% or BM 5 ~ 9%, no Auer rods    | Any   |
| MDS-EB-2                         | 0-3                 | 1-3        | None or any                                   | PB 5 ~ 19% or BM 10% ~ 19% or Auer rods | Any   |
| MDS-U With 1% PB blast           | 1-3                 | 1-3        | None or any                                   | PB = 1%, BM < 5%, Auer rods             | Any   |
| with SLD and pancytopenia        | 1                   | 3          | None or any                                   | PB <1% BM <5% No Auer rods              | Any   |
| Defining cytogenetic abnormality | 0                   | 1-3        | <15%  | PB <1% BM <5% No Auer rods              | MDS defining abnormality  |
| RCC                              | 1-3                 | 1-3        | None  | PB < 2% BM < 5% No Auer rods            | Any   |

The molecular profile of MDS elucidated using NGS highlights its considerably heterogeneous mutational landscape, which has been considered essential for defining the clinical and pathological characteristics of MDS since its initial description. The NGS data have provided substantial insight into the pathophysiological complexity of MDS, which is characterised by frequent mutations in genes involved in diverse cellular processes such as transcriptional regulation, chromatin modification, RNA splicing, signal transduction, and DNA methylation (Malcovati et al. 2007; Della Porta et al. 2015; Pfeilstöcker et al. 2016; Makishima et al. 2017).

### **1.2.2 Clonal evolution in MDS**

The rapid developments in molecular biology approaches necessitate the re-evaluation of the definition of MDS, including the reassessment of key diagnostic criteria. The term clonal haematopoiesis of indeterminate potential (CHIP) refers to the occurrence of clonal somatic mutations or oncogenes in HSPCs reported to be involved in haematological malignancies, with a variant allele frequency of approximately 2% in the absence or presence of hematologic cancer or other clonal diseases (Corces-Zimmerman and Majeti 2014; Steensma et al. 2015; Valent et al. 2017).

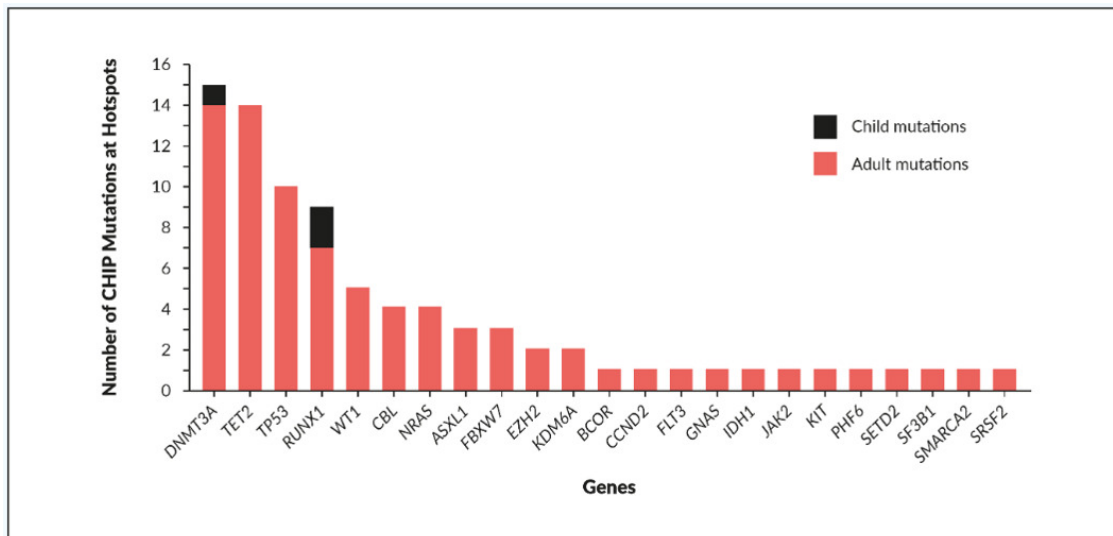
NGS-based cancer research and diagnostic techniques have expanded rapidly in recent years (Barbany et al. 2019). These techniques were used to explore the likely progression of clonal haematopoiesis to MDS in an extensive study of patients with unexplained cytopenia of unclear significance. Individuals with a mutation had an approximately 14-fold higher significant risk of emerging myeloid neoplasms. Various mutations or patterns of mutations have multiple effects on the chances of disease development. Patients harbouring a mutation in one of the epigenetic factors such as, *DNMT3A*, *TET2*, or *ASXL* or harbouring a mutation in one of the spliceosome genes such as, *SF3B1*, *ZRSR2*, *SRSF2*, or *U2AF1* or in conjunction with another mutation had an estimated yearly risk of approximately 20%. In patients with a different

mutation pattern (non-including epigenetic factors or spliceosome genes), the annual risk is around 10% (Malcovati et al. 2017).

Moreover, two retrospective studies investigated the possibility of identifying risk factors for the development of AML. Mutations in spliceosome genes, and *TP53*, *IDH1/2*, were related to a higher risk of progression. (Abelson et al. 2018; Desai et al. 2018).

A recent NGS-based study by Feusier et al. identified 20,141 protein-altering mutations in 7,430 patients with either leukaemia or other hematologic malignancies. Among these, 434 substantially recurring mutation hotspots spanning 85 genes were identified, of which 364 occurred at loci that can be reliably assessed for CHIP. At the time of diagnosis, 134 of 755 component mutations were documented in at least three individuals with hematologic malignancies. This study revealed clones with CHIP mutations identical to those identified in substantial hematologic malignancy hotspots in 83 of 4,530 (1.83%) healthy individuals without cancer (Feusier et al. 2021). In individuals without cancer, CHIP mutations were detected mostly in *DNMT3A*, *TET2*, and *TP53* genes (Figure 1.3). They also reported that 183 cancer-free individuals carried non-hotspot CHIP mutations in *TET2* and *DNMT3A*. Lastly, they observed that most patients with CHIP exhibited a single clonal mutation in one of the genes that govern haematopoiesis, indicating that most CHIP mutations, found at both hotspot and non-hotspot loci, were likely to be the drivers of clonal expansion (Feusier et al. 2021).

Overall, novel genetic research will aid in identifying patient groupings that derive advantages from routine screening for CHIP, implementing preventive measures, and individualising treatment.



**Figure 1.3 Clonal haematopoiesis of indeterminate potential (CHIP) is identified at hotspots in non-cancer individuals.**

Taken from (Feusier et al., 2021).

An in-depth understanding of the clonal dynamics and molecular mechanisms underlying the onset, development, and evolution of MDS is of the utmost importance to facilitate the advancement of novel therapeutic strategies, improve clinical management, and efficiently prevent disease progression. It has been hypothesised that MDS clones acquire mutations sequentially as they develop, with specific clones expanding and dominating as disease progresses. Recent investigations employing single-cell analysis and deep sequencing have highlighted the heterogeneous and complex characteristics of clonal evolution in MDS. (Nakajima 2021). Most likely, the progression from CHIP to MDS involves a complex interaction between epigenetic modifications in HSCs, genetic mutations, and dysfunctional bone marrow microenvironment. For example, mutations in or loss of *TET2* and *DNMT3A* expand the mutant HSC clone and enhance HSC function at the expense of normal polyclonal haematopoiesis during aging (Moran-Crusio et al. 2011; Challen et al. 2012; Li et al. 2013). These findings indicate that genetic lesions initiate MDS and promote self-renewal of HSCs, resulting in advantage in proliferation over normal HSCs and asymptomatic clonal expansion, which ultimately leads to disease progression. Previous studies have calculated allele

frequencies using bulk sequencing and single-cell sequencing, reporting that splicing factors *SF3B1* and *SRSF2* and epigenetic modifiers, especially *DNMT3A* and *TET2*, tend to show early mutations in the progression of MDS. *TET2* and *DNMT3A* mutations have previously been reported to occur as CHIP mutations, suggesting that CHIP partially corresponds with MDS-initiating mutations (Papaemmanuil et al., 2013; Hughes et al., 2014; Makishima et al., 2017). In contrast, mutations in the transcription factors *CUX1*, *GATA2*, and *RUNX1* are detected in early or late events (Papaemmanuil et al. 2013; Chen et al. 2019). Moreover, it has been found that disease progression frequently results in the acquisition of additional sub-clonal mutations, typically in genes related to cell signalling and transcription regulation, such as *FLT3*, *NPM1*, and *TP53*. These mutations increase genomic instability and the risk of AML transformation (Papaemmanuil et al. 2016; Chen et al. 2019).

One of the key challenges to clinically manage of patients with MDS is preventing the progress of the disease to its high-risk or secondary AML (sAML). The progresses of the disease owing to the occurrence of new clones that have acquired additional mutations or different genetic abnormalities. Makishima et al. identified the genetic basis for the clonal evolution of MDS by sequencing DNA samples from 2,250 patients who showed progression from lower-risk MDS to sAML. According to their findings, the number of genetic mutations identified in the BM cells of patients increased as the disease progressed. In addition, mutations were classified as type 1 or type 2 mutations, based on the transition from MDS to AML and the progression from lower-risk to higher-risk MDS, respectively. Type 1 mutations were linked to quicker progress to sAML and less OS time than type 2. Type 1 mutations are observed in *IDH1*, *PTPN11*, *FLT3*, *WT1*, *NPM1*, *NRAS*, and *IDH2*, whereas type 2 mutations were observed in *ZRSR2*, *STAG2*, *TP53*, *GATA2*, *TET2*, *KRAS*, *ASXL1*, and *RUNX1*. It was also demonstrated that patients carrying *SF3B1* mutations had a reduced likelihood of AML transformation compared to those with type 1 mutations, who had a poorer prognosis. These findings indicate

that each mutation plays a distinct role in the pathogenesis and progression of MDS and that the order and type of mutations significantly affect the clinical characteristics of MDS. (Makishima et al. 2017).

In summary, the correlation between different recurrent genetic mutations associated with MDS and their clinical manifestations, including diagnostic classification, therapeutic response, and prognosis, have been elucidated using NGS.

### **1.2.3 Chromosomal abnormalities in MDS**

Acquired cytogenetic abnormalities are observed in 40%–50% of MDS cases, such as chromosome 5q, 7q, and 20q deletions, trisomy 8, and complex karyotypes (Tefferi and Vardiman 2009). Complex karyotypes are defined as the occurrence of three or more chromosomal abnormalities or at least one structural aberration, excluding monosomy 7 clonal evolution cases (Göhring et al. 2010). Identifying the clinical consequences of each unique karyotype is essential to understand disease progression and therapy (Haase et al. 2007; Schanz et al. 2011).

Deletions of the long arm of chromosome 5 (Del [5q]) is one of the most prevalent cytogenetic abnormalities in MDS. Del (5q) is the first genetic mutation to be included in the WHO classification of MDS, defining a distinct subtype of MDS (List et al. 2018). Approximately 15% of identified MDS cases are associated with del (5q) (Giagounidis et al. 2005; Haase et al. 2007). MDS with del (5q) displays a heterogeneous clinical pattern divided into sub-group classifications based on clinicopathological characteristics, patient prognosis, and therapeutic responsiveness. One of the del (5q) subtypes develops following treatment with cytotoxic chemotherapy such as alkylating drugs or radiation exposure, and it is frequently associated with other chromosomal abnormalities and *TP53* mutations; moreover, it is associated with a greater risk of leukemic transformation and reduced survival (6–17 months, depending on the number or severity of other abnormalities) (Kantarjian et al. 2009; Mollgard et al. 2011). In

contrast, patients with isolated del (5q) have a better prognosis and a lower likelihood of developing AML (Scharenberg et al. 2017). Studies have shown that the loss of specific genes in the critical region of deletion (CDR) leads to impaired ribosome biogenesis, decreased protein synthesis, and activation of the *p53* pathway (McGowan et al. 2011).

The reduced ribosomal capacity and increased *p53* activity have been shown to lead to increased apoptosis and reduced proliferation of erythroid progenitor cells, resulting in the characteristic anaemia observed in del (5q) MDS and AML (McGowan et al. 2011). Moreover, the dysregulation of other haematopoietic stem cell genes such as *TP53*, *FLT3*, and *NFI* has been shown to be associated with the pathogenesis of del (5q) MDS. Del (5q) MDS is associated with a relatively favourable prognosis compared with other cytogenetic abnormalities in MDS. The use of erythropoietin-stimulating agents and lenalidomide, an immunomodulatory drug, has been shown to improve anaemia and reduce transfusion dependence in del (5q) MDS (Platzbecker et al. 2021). The mechanism of action of lenalidomide in del (5q) MDS is believed to be the upregulation of the *RPS14* gene, which is deleted in del (5q) MDS (Martinez-Høyer and Karsan 2020).

Chromosome 7 anomalies (often monosomy 7 or 7q deletion) have been found in approximately 10% of patients with de novo MDS and in up to 50% of patients with therapy-related MDS (Christiansen et al. 2004; Haase et al. 2007). Moreover, chromosome 7 aberrations are associated with a poor prognosis and decreased overall survival in MDS and other myeloid cancers such as AML (Schanz et al. 2012). *EZH2*, located in 7q36, encodes a histone methyltransferase that regulates gene expression and is involved in myeloid cell development (Göllner et al. 2017; Stomper et al. 2021). This chromatin remodeler is mutated in approximately 6% of MDS cases and is associated with poor prognosis (Bejar et al., 2011). Nevertheless, it has been shown that 7q deletions do not lead to a loss of the *EZH2* gene (Ernst



et al. 2010). Overall, 7q deletions are large, and haploinsufficiency of numerous genes in the deleted regions contributes to the pathogenesis of MDS (Pellagatti and Boulwood 2015).

Trisomy 8 (+8), the most prevalent chromosomal aberration in MDS, occurs in approximately 11% of *de novo* MDS cases with an aberrant karyotype (Solé et al. 2000; Drevon et al. 2018). Moreover, Individuals with MDS who have isolated trisomy 8 are associated with an intermediate prognosis risk (median overall survival of 23 months) (Schanz et al. 2011). However, +8 is not specific to MDS but is also observed in CMML and AML (Tang et al. 2014). Previous studies show that +8 MDS cells manifest elevated amounts of survivin (anti-apoptotic protein) and show increased resistance to apoptotic stimuli such as gamma-ray irradiation (Sloand et al. 2007). Knockdown of this anti-apoptotic protein eliminates the survival benefit of the +8 MDS clone, suggesting a potential therapy target for this class of patients (Sloand et al. 2007). In the same study, patients with +8 MDS showed a high rate of response to immunosuppressive therapy, demonstrating an underlying immunological pathology mainly associated with trisomy 8.

Deletion of the long arm of chromosome 20 (del [20q]), correlated with increased sensitivity to cytotoxic treatment, occurs in 5%–10% of MDS cases (Yin et al. 2015). However, it is not specific to MDS and is also found in myeloproliferative neoplasms (MPN), MDS/MPN overlap disorders, and AML (Bench et al. 2000a; Hofmann et al. 2004; Haase et al. 2007). Unlike del (5q), del (20q) is not recognised as a distinct factor in the WHO classification of MDS. Del (20q) is observed in the BM specimens of individuals lacking diagnostic morphological characteristics of any myeloid neoplasms and those with non-myeloid malignancies or unexplained cytopenia (Sperling et al. 2017). Compared with other types of MDS, MDS with isolated del (20q) is associated with a lower likelihood of progression to AML and longer survival (Braun et al., 2011). Patients with only del (20q) without other mutations show a low

risk for progression to myeloid neoplasms; in contrast, approximately one-third of patients with other mutations show progression to a myeloid neoplasm (Ravindran et al. 2020).

It has been shown that del (20q) pre-initiating clones evolve and acquire additional mutations (especially in splicing factor genes); these secondary clones were eliminated by chemotherapy, but del (20q) pre-initiating clones resisted therapy, leading to relapse (Hirsch et al. 2016).

A recent study sought to identify the distinct characteristics of MDS patients with this type of deletion. The data were collected from 69 patients with MDS carrying del20q and compared with 502 MDS patients with normal karyotypes (NK-MDS). Patients harbouring del20q were predominantly male and older than those in the NK-MDS group. Additionally, they had a greater BM blast percentage BM (5%), higher proportions of low and int-1 risks, as well as a lower median platelet count, according to the IPSS score. After a median of 28 months following diagnosis, nine patients developed AML. The median OS for the entire cohort was 60.6 months, with a 5-year cumulative OS of 55.9%. Their study hypothesised that MDSs with del20q constitute a discrete biological entity with distinct clinical and prognostic characteristics (Campagna et al. 2022).

Several studies have found a CDR of at least 3 Mb that spans chromosome 20q12–20q13. *L3MBTL1* and *MYBL2* are the primary transcriptional regulators involved in the aetiology of del (20q)-associated cancers (Bench et al. 2000b; Bacher et al. 2014). It has been demonstrated that *L3MBTL1* silencing biases erythroid development in human CD34+ cells (Perna et al. 2010). The role of *MYBL2* is explained in 1.4 section.

#### **1.2.4 Somatic mutations in MDS**

Somatic mutations may arise spontaneously or be induced by environmental factors, such as exposure to radiation or chemicals (Paul et al. 2019). Until the early 2000s, alterations in only a few genes, including *NRAS* (Hirai et al., 1987), *TP53* (Sugimoto et al., 1993), *RUNX1* (Osato et al. 1999), and *ATRX* (Steensma et al. 2004), had been recognised to be involved in MDS

(Ogawa 2019). Subsequently, several new mutational targets in MDS, including *FLT3*, *KRAS*, *KIT*, *PTPN11*, *EZH2*, *CBL*, *KRAS*, *MPL*, *ASXL1*, and *TET2*, were found using high-throughput capillary sequencing and SNP array karyotyping (Pardanani et al. 2006; Sanada et al. 2009; Delhommeau et al. 2009; Nakajima 2021). However, during the subsequent decade, the use of modern sequencing technologies identified the involvement of somatic mutations in the myeloid lineage in MDS, which completely revised the understanding of MDS genomes (Grinfeld et al. 2018). In MDS, somatic mutations are found in genes that regulate haematopoiesis, such as *DNMT3A*, *TET2*, *ASXL1*, and *SF3B1*. These genes are involved in epigenetic regulation, chromatin remodelling, and mRNA splicing, and their dysregulation leads to aberrant haematopoiesis and the development of MDS (Jaiswal et al. 2014; Mason et al. 2016). MDS cases harbour, on average, two or three driver mutations; (MDS-EB and MDS with multilineage dysplasia), high-risk MDS and chronic myelomonocytic leukaemia are likely to have more driver mutations than lower-risk MDS, such as MDS with single-lineage dysplasia with or without increased ring sideroblasts (RS) and MDS with isolated del (5q) (Ogawa 2019). The target driver genes are classified according to their association with distinct functional pathways; these genes and their functions are found in Table 1.4. *SF3B1*, *TP53*, *RUNX1*, *EZH2*, *ASXL1*, and *NRAS* were previously found to have an independent prognostic impact in MDS and affect OS. These genes, except *SF3B1*, have a deleterious effect on survival and may actively induce disease progression to AML (Yoshida et al. 2011; Bejar et al. 2011; Malcovati et al. 2014). The tumour suppressor gene *TP53* is necessary for DNA repair processes, cell-cycle arrest and cellular differentiation and apoptosis initiation in response to genetic damage (Carr and Jones 2016). *TP53* mutations are observed in 10%–15% of individuals with MDS, and it predicts a poor clinical prognosis and diminished therapeutic response (Nazha et al., 2017; Fenaux et al., 2020). Moreover, it is usually observed in patients with MDS with complex karyotypes, who have the shortest overall survival (less than 6

months) (Sebaa et al. 2012; Haase et al. 2019). *ASXL1* mutations are found in 13%–21% of MDS cases, and according to the IPSS, their presence is a predictor of worse survival in patients with low or intermediate-1 risk (Bejar et al. 2011). *SF3B1* is an RNA splicing machinery member; it is the sole gene whose acquired lesion is defined by the WHO category (MDS with ring sideroblasts). (Arber et al. 2016). Furthermore, *SF3B1* mutations have an independent prognostic impact on survival and probability of progression to AML. (Palomo and Solé 2020).

**Table 1-4 Major driver genes**

Taken from (Ogawa 2019)

| <b>Pathway/functions</b>                 | <b>Driver genes</b>  |
|--|--|
| <b>DNA methylation</b>                   | <i>DNMT3A, TET2, IDH1, IDH2, and WT1</i>                                   |
| <b>Chromatin remodellers</b>             | <i>EZH2, SUZ12, EED, JARID2, ASXL1, KMT2, KDM6A, ARID2, PHF6, and ATRX</i> |
| <b>RNA splicing</b>                      | <i>SF3B1, SRSF2, U2AF1, U2AF2, ZRSR2, SF1, PRPF8, LUC7L2</i>               |
| <b>Cohesion complex</b>                  | <i>STAG2, RAD21, SMC3, and SMC1A (PDS5B, CTCF, NIPBL, and ESCO2)</i>       |
| <b>Transcription</b>                     | <i>RUNX1, ETV6, GATA2, IRF1, CEBPA, BCOR, BCORL1, NCOR2, and CUX1</i>      |
| <b>Cytokine receptor/tyrosine kinase</b> | <i>FLT3, KIT, JAK2, and MPL, CALR, and CSF3R</i>                           |
| <b>RAS signalling</b>                    | <i>PTPN11, NF1, NRAS, KRAS, and CBL (RIT1 and BRAF)</i>                    |
| <b>Other signalling</b>                  | <i>GNAS, GNB1, FBWX7, and PTEN</i>   |
| <b>Checkpoint/cell cycle</b>             | <i>TP53 and CDKN2A</i>   |
| <b>DNA repair</b>                        | <i>ATM, BRCC3, and FANCL</i>   |
| <b>Others</b>                            | <i>NPM1, SETBP1, and DDX41</i>   |

### 1.2.5 Mutations in transcription factors

Myeloid transcription factors such as *RUNX1*, *ETV6*, and *GATA2*, which regulate gene expression by binding to particular DNA sequences, are an additional class of key mutational targets in MDS (Papaemmanuil et al. 2013). Approximately 10%–15% of patients with MDS have somatic mutations in these genes, but these mutations may also be found as germline variants (Churpek and Bresnick 2019).

*RUNX1* is a critical component of normal haematopoiesis, contributing to the development and differentiation of HSCs. It is mutated in approximately 10% of MDS cases and is frequently identified in patients who show progression to AML (Harada et al. 2004; Papaemmanuil et al. 2013). Furthermore, patients lacking this mutation survive longer and exhibit a longer time of progression to AML (Della Porta et al. 2016; Nazha et al. 2019; Bejar et al. 2012; Papaemmanuil et al. 2013).

The *ETV6* gene also encodes a TF involved in bone marrow haematopoiesis (Wang et al., 1998). It is a rare mutation in MDS, arising in 1%–3% of cases. *ETV6* mutations are

accompanied by poor prognosis, as evidenced by a decreased median OS and a worse OS when controlling for prognostic factors (Bejar et al. 2011) However, this finding needs to be verified by further research.

GATA1/2 are protein transcription factors that regulate the growth and proliferation of immature megakaryocytes and red blood cells. *GATA1* mutations, both inherited and acquired, have been linked to malignant or abnormal haematopoiesis, whereas *GATA2* mutations have been linked to mononuclear cytopenia, also in MDS, and AML (Crispino 2005; Dickinson et al. 2014)

### **1.3 *CEBPA***

Human *CEBPA* is an intron-less gene that is found on chromosome 19.q13.1, which encodes for the CCAAT/enhancer-binding protein, a basic leucine zipper (LZ) TF. C/EBP $\alpha$  is essential for the differentiation of multipotent HSCs into myeloid progenitors and is involved in the differentiation of granulocytes (Radomska et al. 1998). *CEBPA* mRNA is translated in two different translation site to produce two isoforms into a full-length 42-kDa isoform (p42) and a shortened 30-kDa isoform (p30), which generate from the downstream of translational starting site and can form a homodimer or a heterodimer with other C/EBP $\alpha$  proteins to regulate genes associated with cell differentiation, survival, growth and metabolism (Calkhoven et al. 2000; Keeshan et al. 2003).

The C/EBP $\alpha$  protein comprises three transactivation domains (TADs) at its N-terminus, a DNA-binding domain (DBD) and a dimerisation domain (bZIP domain) at its C-terminus (Nerlov 2004a). The p30 isoform lacks two of the N-terminal TADs (Figure 1.4) (Friedman and McKnight 1990; Nerlov 2004a). The ratio of p42 to p30 affects myeloid differentiation; for example, elevated p30 levels prevent the terminal differentiation of granulocytes as it interacts with E2F, which increases proliferation potential (Nerlov 2004a). C/EBP $\alpha$  activates the expression of the myeloid gene programme during cell fate decisions by enhancing genes

and binding to the promoters involved in the myeloid process, for example, *IL-6R*, *CEBPE*, *GFI-1* and *CSF3R* (Zhang et al. 1998; Friedman 2007; Ma et al. 2014).



**Figure 1.4 C/EBP $\alpha$  protein structure.**

Schematic of C/EBP $\alpha$ , depicting the location of trans-activation domains (TADs), and binding leucine zipper domain (bZIP) that consists of a basic region DNA binding domain (DBD) and the leucine zipper domain (LZ).

*CEBPA* is expressed in non-haematopoietic tissues such as the liver, respiratory epithelium, and adipose tissue (Koschmieder et al. 2009). During early embryogenesis, C/EBP $\alpha$  has both essential and redundant functions in conjunction with C/EBP $\beta$  (Smink and Leutz 2010). C/EBP $\alpha$  is required for the CMP to GMP transition and for the growth of the liver and lungs and plays a role in postnatal maintenance of systemic energy homeostasis and fat storage (Müller et al. 2004). Moreover, it plays a significant role in downregulating gene expression, which is responsible for maintaining undifferentiated and proliferative cells. This downregulation occurs through the suppression of *E2F1*, which is necessary for *CEBPA* to trigger the terminal differentiation of adipocytes and granulocytes. In turn, *E2F1* prevents adipocyte differentiation by binding to specific promoters and inhibits *CEBPA* from binding to the promoters of its target genes (Müller et al. 2004).

### 1.3.1 *CEBPA* mutations in haematological disorders

The most prevalent *CEBPA* mutations in AML, observed in 10%–15% of patients, frequently are N-terminus frameshift mutations that cause premature truncation of the p42 isoform while maintaining the expression of the p30 isoform (Leroy et al. 2005). Additionally, *CEBPA* mutations are in-frame insertions or deletions in the C-terminal bZIP domain, which inhibit DNA binding and dimerization (Grossmann et al. 2011; Tawana et al. 2015). *CEBPA* mutations in human neoplasia were first reported in 2001 by the group of Mueller with *CEBPA* mutations found in 10 out of 137 individuals with AML (Pabst et al. 2001). In 2002, another study

retrospectively evaluated the prognostic impact of *CEBPA* mutations in another cohort of 130 patients with heterogeneous types of leukaemia based on the FAB classification criteria, except for acute promyelocytic leukaemia and determined that *CEBPA* mutations in 15 out of 135 patients were an independent indicator of favourable long-term outcomes (Preudhomme et al. 2002).

In 2009, a research group from the Netherlands highlighted the predictive significance of a *CEBPA* single mutation and *CEBPA* double mutations (Wouters et al. 2009). They discovered that patients harbouring *CEBPA* double mutations had a higher survival rate and unique gene expression patterns. However, when patients with a *CEBPA* single mutation were compared to patients with the wild-type *CEBPA* gene, their gene expression and results were comparable (Wouters et al. 2009). Subsequently, several studies have corroborated the favourable prognosis of AML in patients carrying double mutations of *CEBPA* (Pabst and Mueller 2009; Dufour et al. 2010; Green et al. 2010; Taskesen et al. 2011). Therefore, AML with *CEBPA* double mutations is recognised as a unique entity in the 2016 revision of the WHO classification of myeloid neoplasms and AML due to its unique biological and diagnostic features (Arber et al. 2016).

Patients with MDS exhibit a single *CEBPA* mutation, whereas a double mutation have been detected in AML cases that arise as a consequence of MDS (sAML) (Wen et al. 2015).

The co-occurrence of a single *CEBPA* mutation with mutations in other genes was accompanying with an unfavourable prognosis in MDS (Shih et al. 2005; Kato et al. 2011; Armes et al. 2022). Furthermore, *CEBPA* methylation has been recognised as a common occurrence in MDS (Wen et al. 2015; Yu, Li, Li, et al. 2020).

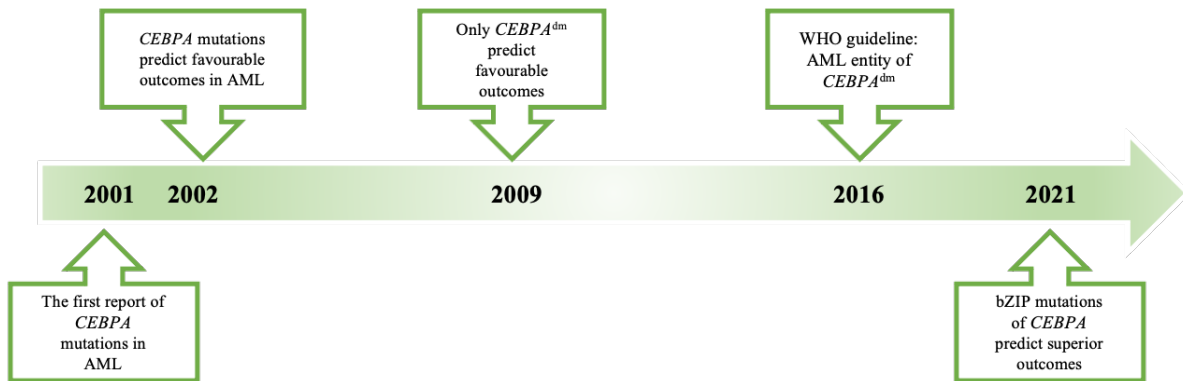
Shih and colleagues found that the prevalence of *CEBPA* gene mutations was 8% at the time of MDS diagnosis and 12% upon progressing from MDS to AML, suggesting that *CEBPA* mutations play a role in the pathogenesis of a subcategory of MDS cases with disease



progression. (Shih et al. 2005). *CEBPA* mutations are prevalent in sAML, indicating that they are acquired later during the progression of the disease to AML (Kato et al. 2011; Armes et al. 2022).

Recent research has indicated that the classification of single and double mutations is necessary to adequately capture the biological basis and clinical implications of AML. For example, in their recent study involving 4,708 adult patients (spanning all age groups) diagnosed with AML and secondary AML after MDS, Taube and colleagues (2022) reported that patients who had disrupted bZIP of *CEBPA* (*CEBPA*<sup>bZIP</sup>) or those carrying double mutations exhibited the same gene expression pattern and clinical characteristics, such as a higher leukocyte count upon diagnosis, a younger age and a higher survival rate compared to patients carrying a *CEBPA* single mutation in a TAD (Taube et al., 2022). Moreover, Taube and colleagues (2022) found that patients with in-frame mutations in bZIP, regardless of single or double mutations, had higher clinical and molecular features and better OS (Taube et al., 2022). A favourable prognosis for *CEBPA*<sup>bZIP</sup> was also reported in a separate group of 1,028 patients with AML. The presence of *CEBPA*<sup>bZIP</sup> is a significant marker of improved survival, a higher likelihood of achieving CR and reduced susceptibility to relapse (Wakita et al. 2022).

The significant advances in research on *CEBPA*-mutant AML over the last two decades are summarised in Figure 1.5.



**Figure 1.5** Timelines of evolution of *CEBPA* mutations in AML.

## 1.4 MYBL2

The MYB transcription factor family consists of three members: MYB (c-Myb), MYBL1 (A-Myb), and MYBL2 (B-Myb). *MYB* is the mammalian homolog of the retroviral v-Myb oncogene, which induces acute leukaemia and modulates avian haematopoietic cells. *MYB* was the first family member found and is the mammalian equivalent of the v-Myb oncogene (Roussel et al. 1979; Ness 2003). MYB is expressed in HSCs, colon crypt cells, and brain stem cells in mammals (Ness 2003). MYBL1 is predominantly expressed in the nervous and reproductive systems and germinal B-lymphocytes (Trauth et al. 1994). *MYBL1* deletion results in viable mice, and *MYB* deletion results in late embryonal death owing to a lack of erythropoiesis (Mucenski et al. 1991; Toscani et al. 1997).

*MYBL2* (B-myb; v-Myb avian myeloblastosis viral oncogene homolog-like 2) is located on the long arm of chromosome 20 (20q13); it is a highly conserved member of the MYB transcription factor family. MYBL2 is ubiquitously expressed in rapidly dividing cells including embryogenic stem cells and haematopoietic cells (Liang et al. 2017) which may account for the lethal phenotype of *Mybl2* knockout mice exhibiting early embryonal death due to a defective G1/S transition in the cell cycle (Tanaka et al. 1999). According to the findings of meta-analysis studies, MYBL2 is needed for mammalian development and is a physiological regulator of progressive cell cycles, cell proliferation, cell differentiation, and cell survival (Sala 2005; Whitfield et al. 2006; Martinez and DiMaio 2011). In agreement with its role in cell proliferation, MYBL2 is observed to be upregulated in many solid cancers such as ovarian, lung, and breast cancer, leading to the deregulation of cell survival, cell cycle progression and differentiation (Ren et al. 2015).

### 1.4.1 Role of MYBL2 in the maintenance of stem cells and differentiation

Several studies indicate that MYBL2 maintains the self-renewal capacity of stem cells and thus plays a critical role in regulating stem cell division, differentiation, and self-renewal (Zhan et

al. 2012; Baker et al. 2014). It has been reported that constitutive MYBL2 expression inhibited retinoic acid-induced neural differentiation; in contrast, downregulation of MYBL2 during retinoic acid treatment induced neural and glial differentiation in neuroblastoma cells (Raschell et al. 1995). In agreement with this finding, studies in several other cell types such as male gonocytes, leukemic cell lines, keratinocytes, and intestinal epithelial cells have also demonstrated a role for MYBL2 in maintaining cells in an undifferentiated state (Bies et al. 1996; Latham et al. 1996; Papetti and Augenlicht 2011; Maruyama et al. 2014). It has been postulated that MYBL2 regulates a transcriptional network that regulate cell cycle progression and cell fate decisions to maintain the self-renewal capacity and pluripotency of embryonic stem cells (ESCs) and induced pluripotent stem cells (iPSCs) (Zhan et al. 2012; Lorvellec et al. 2010; Ward et al. 2018). *MYBL2* maintains pluripotency by directly regulating the expression of crucial genes associated with differentiation and pluripotency in ESCs (*POU5F1*, *SOX2*, and *NANOG*) (Tarasov et al. 2008; Zhan et al. 2012). Likewise, it regulates the self-renewal and differentiation of HSCs by downregulating the genes responsible for cell differentiation, such as *CEBPA* and *ID1* and upregulating genes encoding proliferation factors, such as *GATA2* (Briegel et al. 1993; Baker et al. 2014).

Moreover, a previous study by our group has shown that MYBL2 is critical for genome stability and plays a role in suppressing replication stress in ESCs by activating ATM. ATM is protein involved in the DNA damage response (DDR) (Marechal and Zou 2013). In this study they found that loss of MYBL2 or inhibition of ATM or Mre11, another protein involved in DDR (Livak and Schmittgen 2001) in mouse ESCs leads to raised origin firing, replication fork slowing and increased fork stalling, all indicators of replication stress. The study also discovered that inhibiting CDC7 activity (a protein involved in DNA replication initiation) could alleviate replication stress triggered by the loss of MYBL2 and inhibition of ATM. This

suggests that the uncontrolled firing of new replication origins may be the underlying cause of the replication stress phenotype. (Blakemore et al. 2021).

#### **1.4.2 MYBL2 in haematological disorders**

Previous research has emphasised the significant role of *MYBL2* in the haematopoietic system in sustaining the balance between self-renewal and differentiation of HSCs (Baker et al. 2014). It was found that *MYBL2* deletion leads to the depletion of HSCs and thus the loss of mature cells. Clarke et al. reported a significant association between low expression of *MYBL2* and lowered expression of further genes involved in checkpoint control and DNA replication pathways in CD34<sup>+</sup> BM cells of patients with RAEB2 MDS (Clarke et al. 2013). This finding was verified in another study, demonstrating that regardless of del (20q), 50% of patients with MDS exhibit reduced levels of *MYBL2*. The published data suggest that gene within del (20q) affects MDS progression more significantly than previously hypothesised (Heinrichs et al. 2013).

Previous study in our lab has identified a novel function for *MYBL2* in DNA double-strand break (DSB) repair in HSC. This study discovered a correlation between low levels of *MYBL2* and the disruption of DNA repair genes in individuals with MDS following exposure to ionising radiation. Additionally, stem and progenitor cells from these patients exhibit poor repair kinetics for DSB. Moreover, the findings revealed that haploinsufficiency of *Mybl2* in mice impaired the repair of DSBs caused by ionising radiation in HSC, which was characterised by unstable phosphorylation of the ATM substrate KAP1 and fragile telomeres, suggesting that *MYBL2* levels are a key regulator of DSB repair, and a biological biomarker to define the DNA repair capacity of HSCs from patients with MDS, also, is considered a biomarker to guide the selection of patients for treatments that target DNA repair (Bayley et al. 2018).

## 1.5 Therapeutic approaches of MDS

Owing to the diverse and varied features of patients with MDS, the approach to therapy is customised based on the expected prognosis for each patient. Therefore, good prognostic prediction is a crucial component of comprehensive patient care. The IPSS-R classification is frequently used for assessing the likelihood of MDS progression to leukaemia and guiding treatment such as chemotherapy and allogeneic haematopoietic stem cell transplantation (Allogeneic HSCT) for younger patients with high-risk MDS.

Overall, MDS require both curative and supportive care/therapy options. In the case of low-risk MDS (IPSS low/intermediate-1 risk; IPSS-R very low, low, and intermediate risk), treatment focuses primarily on ameliorating cytopenia to minimise consequences such as bleeding and severe infections, reducing transfusion load, and enhancing quality of life (Platzbecker 2019).

Moreover, one of the main goals of patients with MDS is to maintain a haemoglobin level of 8–10 g/dL. Patients with low-risk or intermediate-1-risk MDS with platelet counts between 50,000-100,000/L and haemoglobin levels >10 g/dL, typically not requiring blood transfusion, are frequently treated using a 'watch and wait' approach with monitoring (Ria et al. 2009).

Erythropoiesis-stimulating agents like recombinant humanised erythropoietin or the longer-acting erythropoietin darbepoetin alfa is regarded as the standard first-line treatment for anaemia in patients with low-risk MDS (Saygin and Carraway 2021; Volpe et al. 2022). The immunomodulatory medicine lenalidomide is recommended therapy for 5q-type MDS with or without additional cytogenetic anomalies resulting in erythroid responses in around 70 % of the cases (Scalzulli et al. 2021; Volpe et al. 2022).

On the basis of clinical evaluation of symptoms related to anaemia and the presence of secondary infections, patients with high-risk MDS can be cured with supportive therapy, including blood transfusion to replenish platelets and red blood cells (Ria et al. 2009). In

addition to administering immunosuppressive medicines, curing infectious diseases, iron chelation therapy and management of haemorrhagic consequences are employed to mitigate cytopenia and improve quality of life. (Sanchez 2011). The potential for iron overload in elderly patients resulting from transfusion may require the use of iron-chelating agents in conjunction with the oral administration of deferasirox (Exjade®:20–40 mg/kg per day). In individuals with transfusion-dependent myelodysplastic syndrome (MDS), this therapeutic approach led to a notable decrease in mean serum ferritin levels after 12 months of treatment (List et al., 2012). Demethylating medicines are the mainstay treatment for patients with higher-risk MDS. They are a class of chemotherapeutic medicines that induce short-term hypomethylation of DNA (Garcia-Manero 2008). The FDA has licenced two DNA-hypomethylating drugs, including 5-azacitidine and 5-aza-2'-deoxycytidine (decitabine), for treatment of MDS. These medications are highly effective in high-risk MDS and are widely used in the USA and Europe (Silverman et al. 2002; Diesch et al. 2016). In contrast with lenalidomide, azacytidine or decitabine is frequently the treatment of choice for patients with non-del (5q) MDS. However, it has been seen in recent years that the combination of lenalidomide and decitabine appears to be an effective and safe therapy for MDS patients with del (5q) and excess blasts, with no discernible adverse effects (Serin et al. 2020). Patients with high-risk MDS and who are ineligible for intensive chemotherapy are approved to receive hypomethylating agents, such as azacytidine and decitabine. although, half of these group of patients respond to this alternative therapy, its efficacy is limited to less than 2 years (Saygin and Carraway 2021; Volpe et al. 2022).

Recently, Inqovi®—also called ASTX727—which combines decitabine with cedazuridine, an inhibitor of cytidine deaminase (CDA3), has been approved by FDA for the treatment of adults with several FAB subtypes of MDS (Savona et al. 2020).

Allogeneic HSCT is still considered to be the most effective preventative treatment against progression of MDS to AML, and it is considered for patients with high-risk MDS as soon as possible after diagnosis (Sekeres and Taylor 2022).

In recent years, significant progress has been made in the development of novel therapeutic approaches to treat MDS. Unfortunately, there is currently no definitive treatment for MDS. Various medications are still being assessed in clinical trials; for example, the telomerase inhibitor imetelstat (Steensma et al. 2021) and monotherapy drugs such as ivosidenib and enasidenib are being assessed for use in treating patients with high-risk MDS harbouring *IDH1/2* mutations (Stein et al. 2020). Although the efficacy of these drugs is promising, it should still be considered that the path from clinical trials to approved therapies may be long and challenging.

In conclusion, although treatment options for MDS are limited, genetic research using various model systems is crucial to understanding disease progression and would help inform clinical decision-making. These may have a significant impact and ultimately lead to a cure for this complex disease, and identifying these mutations will potentially improve treatment outcomes.

### **1.6 Human induced pluripotent stem cells (hiPSCs)**

iPSCs are stem cells derived from somatic cells co-expressing certain pluripotency-associated factors. Similar to ESCs, iPSCs can proliferate indefinitely and self-renew *in vitro* and are pluripotent, as they can differentiate into all three germ layers: (i) ectoderm, (ii) mesoderm, (iii) and endoderm (Takahashi and Yamanaka 2006; Takahashi et al. 2007). In 2006, iPSCs were generated from mouse fibroblasts by Takahashi and Yamanaka at Kyoto University, Japan. (Takahashi and Yamanaka 2006). In 2007, the same researchers successfully generated iPSCs using human cells (Takahashi et al. 2007) They are generated by the utilised retroviral transduction of four transcription factors (*Oct4*, *Sox2*, *Klf4*, and *c-Myc*), collectively referred to as OSKM, into fibroblasts. The key advantage of using retroviral vectors is their high



transduction efficiency (Takahashi and Yamanaka 2006). Nevertheless, retroviral vectors can only integrate into dividing cells, a limitation overcome by lentiviral vectors, which can integrate into both dividing and non-dividing cells (Okita et al. 2007). However, the propensity of these vectors to integrate into the host genome increases the risk of insertional mutagenesis, which may lead to tumorigenesis (Papapetrou and Schambach, 2016).

Nevertheless, the need to use multiple lentiviral vectors to drive the expression of TFs has been replaced by the development of polycistronic viral vectors such as the STEMCCA lentiviral vector, in which loxP sites flank a polycistronic OSKM lentiviral vector (Sommer et al. 2009; Carey et al. 2009). The cells of interest are infected with the lentivirus to generate iPSCs. The OSKM genes are removed from the genome using Cre recombinase. Thus, STEMCCA vectors increase the reprogramming efficiency by driving the expression of all OSKM factors with a single promoter (Carey et al., 2009; Sommer et al., 2009).

In addition, the use of episomal vectors to generate iPSCs offers a non-viral and non-integrating approach as an alternative to integration-free viruses. Episomes are extrachromosomal DNA molecules within a cell that can replicate independently of chromosomal DNA. As plasmid or minicircle DNA, episomal vectors can carry OSKM factors directly into somatic cells (Okita et al., 2010). In 2009, the research group led by Dr Thomson was the first to derive hiPSCs using non-integrating episomal vectors. This discovery was significant because the approach removes one of the barriers to the clinical use of hiPSCs, as the host cell genome remains intact; the finding demonstrated that somatic cells do not need genomic integration or exogenous reprogramming factors' continued presence (Yu et al. 2009).

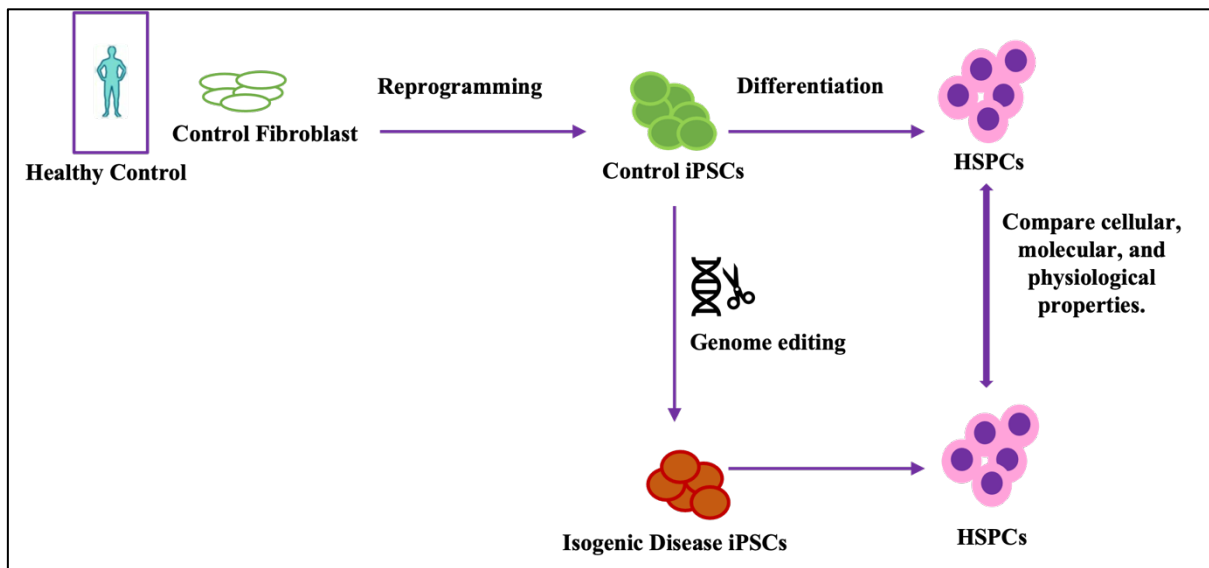
Another approach for reprogramming human somatic cells into hiPSCs was proposed by Guan and colleagues using chemical methods. This method led to the production of cell characteristics similar to those of ESCs. This new reprogramming method involves chemical stimulation. This chemical reprogramming method targets cellular signalling pathways and

epigenetic modifier genes. Small molecules are advantageous because of their non-integrative nature, controllability, and ease of standardisation. Moreover, unlike previous lineage reprogramming methods, they unlock the human epigenetic landscape by inducing dedifferentiation. Furthermore, researchers have recognised the importance of inhibition of the JNK pathway, which is crucial for promoting cell plasticity and initiating a regeneration-like program by reducing pro-inflammatory pathways, as they found that the JNK pathway is a major barrier to chemical reprogramming (Guan et al. 2022).

### **1.6.1 Advantages and potential applications of hiPSCs**

The iPSC discovery can be used to tailor therapeutic approaches and to perform drug screening (Yamanaka 2012). Moreover, using iPSCs offers a solution to bioethical issues because this approach does not require the use of embryos, a primary ethical concern associated with the use of ESCs. Thus, iPSCs provide an ethically acceptable alternative for stem cell research (Zheng 2016). Moreover, iPSCs derived from the somatic cells of individual patients hold the potential to overcome the problem of immune rejection in transplantation (Taylor et al., 2019). Furthermore, iPSCs provide an *in vitro* model of human developmental processes and cell lineage differentiation to investigate the timing of gene expression, cell fate, and regulation of the niche environment at various developmental stages. Differentiating iPSCs also serve as substrates for examining single or several gene roles and their physiological functions *in vitro* by introducing genes, deletions, mutations, or replacements (Figure 1.5) (Xue et al. 2009; Hou et al. 2013; Bhatia et al. 2013).

In summary, iPSCs are currently considered an excellent alternative source of pluripotent stem cells, combining the use of these cells with genome editing tools such as the Clustered, regularly interspaced, short palindromic repeats–CRISPR-associated protein 9 (CRISPR–Cas9) system has the potential to provide feasible disease models *in vitro* to study the function of different genes in disease processes and progression and drug evaluation (Wattanapanitch 2019; Moradi et al. 2019).



**Figure 1.6 Diagram of iPSC application**

An example of applying iPSCs in a disease model through gene editing of normal iPSCs and differentiation of iPSCs into HSPCs *in vitro*.

## **1.7 Genome editing using the CRISPR–Cas9 technology**

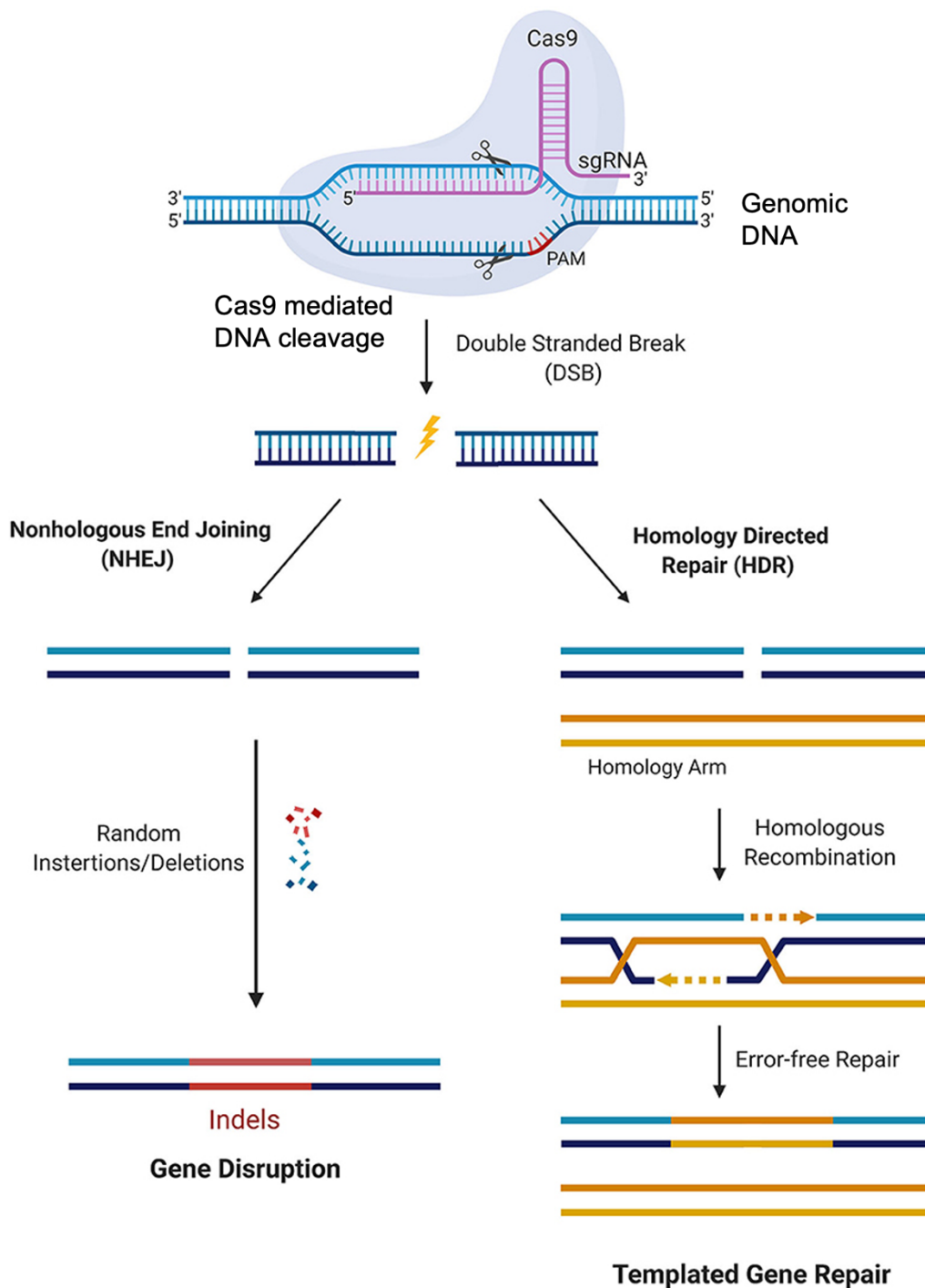
Genome-editing methods have been widely used in scientific research. Previous studies have used targeted genome editing tools, for instance, transcription activator-like effector nucleases (TALENs), meganucleases (MegNs), and zinc finger nucleases (ZFNs) (Alagoz and Kherad 2020; Khalil 2020). The primary drawbacks of ZFNs are their low targeting specificity and efficacy (Ghosh et al. 2021), whereas MegNs have limited design flexibility and are expensive (Abdallah et al. 2015). TALENs have been shown to be both specific and efficient; however, their construction and design are more challenging, which has limited their application despite their superior performance (Feng et al. 2014). Subsequently, the CRISPR-Cas9 system captured the attention of researchers and has since become a popular method for genome or gene editing. Clustered, regularly interspaced, short palindromic repeats (CRISPR)–CRISPR-associated protein 9 (CRISPR-Cas9) is a powerful genome modifying tool that has led to significant advancements in the field of gene therapy owing to its adaptability and ease of use.

### **1.7.1 Mechanisms of CRISPR-Cas9 tools**

The CRISPR-Cas9 system, which constitutes a part of the adaptive immune system in prokaryotes, was first identified in 1987 by a team of Japanese scientists led by Ishino, who observed unusual repetitive palindromic DNA sequences interrupted by spacers in *Escherichia coli* while analysing a gene responsible for alkaline phosphatase isozyme conversion (Ishino et al. 1987). Subsequently, this system has been repurposed for genome engineering in a wide range of living organisms (Knott and Doudna 2018). Its application is more versatile and convenient than previously designed nucleases (Siva et al. 2021; Zhang et al. 2021). The two primary components of the CRISPR-Cas9 system are Cas9 endonuclease and guide RNA (gRNA)—a small RNA molecule designed to bind to a specific target sequence in the DNA. The gRNA contains a 20-nucleotide sequence complementary to the target DNA sequence and serves as a guide for the Cas9 protein to cleave the DNA. The Cas9 protein is a bacterial protein

that acts as molecular scissors to cleave the DNA at the target site, inducing DSBs. The Cas9 protein has two domains: a nuclease domain that cleaves the DNA and a binding domain that recognises the gRNA. A CRISPR-Cas9 system employing *Streptococcus pyogenes* Cas9 endonuclease is the most commonly used form (Le Rhun et al. 2019). Cas9 generates a DSB at the target site upon detection of the protospacer-adjacent motif (PAM), which is a short DNA sequence (three nucleotides: NGG) required for Cas9 to recognise and bind to the target DNA (Brunner et al. 2019; Zhao et al. 2021). In brief, the gRNA directs Cas9 to the target area of the genome where the DSB would arise, thereby activating the DNA repair mechanisms of the cells. DSBs induced in cells are repaired through two main DNA repair processes: non-homologous end joining (NHEJ), which is favoured for gene knockout, and homology-directed repair (HDR), which is essential for gene knock-in. Typically, HDR is a precise process that utilises homologous donor DNA to repair DNA damage, whereas NHEJ is an error-prone process that joins the broken ends of DNA (Klaver-Flores et al. 2021). The principle of CRISPR-Cas9 is depicted in Figure 1.7.

The ability to modify and manipulate specific DNA and RNA sequences in living cells from several species, including mammalian cells, has been transformed using the CRISPR-Cas9 genome-editing tool. Although the CRISPR-Cas9 technology for genetic engineering is a remarkable innovation for gene therapy, it may pose safety issues such as inadequate specificity associated with off-target effects and on-target but undesirable mutations, immunogenicity, and adverse bio-distribution. Recently, base, and prime editing have been established in CRISPR-Cas-based genome editing technology, facilitating DNA modification at the target site without generating (DSBs), thus reducing the likelihood of genomic rearrangements (Kantor et al. 2020; Antoniou et al. 2021).



**Figure 1.7 CRISPR-Cas9 principle.**

Schematic represents CRISPR-Cas9-mediated genome editing. Cas9 protein and sgRNA attach to the target location of genomic DNA, causing a double-strand break (DSB) at 3 or 4 nucleotides upstream of the PAM sequence. Repair DSB via non-homologous end joining (NHEJ) or homologous recombination (HR). Random insertions and deletions are inserted into the genome during NHEJ. HR integrates precise alterations into the target genomic locus by supplying a donor sequence with homology arms with the DSB site. Adapted from (Uddin et al., 2020)

### 1.7.2 Application of CRISPR-Cas9 in haematological disorder research

The CRISPR-Cas9 system has numerous significant medical applications such as identification of genes involved in various illnesses, development of models for multiple diseases, establishment of diagnostic and therapeutic techniques, drug screening and cancer immunotherapy (Baliou et al. 2018; Li et al. 2020). One of the main uses of CRISPR-Cas9 in human health care is its potential for treating blood diseases (Zhang and McCarty 2016; Daniel-Moreno et al. 2019).

Several studies have investigated potential treatment strategies for individuals with the most prevalent inherited blood condition,  $\beta$ -thalassaemia, which is caused by a mutation in the human  $\beta$ -globin (*HBB*) gene, resulting in severe anaemia or ineffective erythropoiesis (Ribeil et al. 2013; Gambari et al. 2015). The only therapeutic option for this condition is autologous stem cell transplant. However, this treatment approach is limited to patients who have found suitable matching donors and has an increased risk of graft-versus-host disease, graft rejection, and transplantation-related mortality (King and Shenoy 2014).

Over the past few years, CRISPR-Cas9 has been utilised to correct *HBB* gene mutations in patients through HDR, restoring normal erythropoiesis. Several studies have used CRISPR-Cas9 to correct  $\beta$ -thalassaemia mutations in patient-derived iPSCs (Xie et al., 2014; Song et al., 2015); iPSCs that have been genetically modified restore *HBB* expression with minimal off-target effects. In another study, the fibroblasts of a patient with  $\beta$ -thalassaemia were reprogrammed to obtain transgene-free naïve-state iPSCs; the results demonstrated that the CRISPR-Cas9 genome editing system had a substantially higher targeting efficiency than primed iPSCs (Yang et al. 2016).

A recent study by Li and co-workers recruited a patient with thalassaemia with a  $\beta$  41-42 mutation (TCTT deletion) in the *HBB* gene and an Hb-WS mutation ( $\alpha\alpha$ WS/ $\alpha\alpha$ ) in the human haemoglobin alpha 2 gene; these mutations were repaired using linearised donor DNA and

HDR mediated by CRISPR-Cas9. The gene correction was validated by sequencing, and these gene-corrected hiPSCs maintained normal pluripotency (Li et al. 2022). Moreover, they differentiated into haematopoietic progenitors, demonstrating that they retain the ability for multilineage differentiation (Li et al. 2022).

Another accomplishment in the field of hemoglobinopathies was reported in research on sickle cell disease (SCD), which— similar to thalassemia—is a genetic illness associated with changes in haemoglobin. The mutation in *HBB* affects red blood cells, causing these cells to form a sickle shape and obstruct blood vessels, and is associated with symptoms including ischaemia, anaemia, haemolysis, and multi-organ damage. (Gewin 2015; Hoban et al. 2016). Significant research efforts have been made to correct the SCD mutation using ZFN- and TALEN-induced HDR in patient-derived iPSCs (Zou et al. 2011; Sun and Zhao 2014). Additionally, the CRISPR-Cas9 technology has demonstrated higher efficiency than other nucleases in targeting the endogenous *HBB* locus in human iPSCs derived from patients with SCD (Huang et al. 2015).

The combination of patient-derived hiPSCs and CRISPR-Cas9 genome editing has been used for advances in MDS research. Kotini and colleagues elucidated the haematopoietic phenotype of a patient with del7q MDS, which included the inhibition of haematopoietic differentiation. They successfully repaired the cells' capacity to differentiate using CRISPR-Cas9 to include the del7q region, resulting in a normal karyotype (Kotini et al. 2015). Subsequently, iPSCs were combined with CRISPR-Cas9 to identify the *SRSF2*<sup>P95L</sup> mutation associated with the MDS phenotype. Their study indicated that *SRSF2*<sup>P95L</sup> did not affect the number of haematopoietic cells in iPSCs harbouring this mutation, which subsequently differentiated into the haematopoietic lineage. However, the cells displayed dysplastic features, highlighting the critical role of the *SRSF2*<sup>P95L</sup> mutation in inducing dysplasia (Chang et al. 2018).



Moreover, two research groups derived iPSCs from patients with germline *GATA2* mutations, who had a familial predisposition to MDS/AML. The research group led by Papapetrou observed that iPSC derived from patients carrying *GATA2* mutations exhibited moderate haematopoietic defects associated with preleukemic conditions ((Kotini et al. 2017). Notably, in their study, Kotini and colleagues successfully generated a phenotypic progression model from low-risk or high-risk MDS in their study by introducing further genetic lesions using CRISPR-Cas9 to delete chromosome 7q or inactivate another allele of *GATA2*. Another study, led by Thomas Winkler, reported a reduction in the haematopoietic differentiation potential, but this phenotype could not be reliably attributed to mutations (Jung et al. 2018).

Overall, CRISPR-Cas9 is a successful tool for genome modifying and complete gene silencing. Application of CRISPR and iPSCs have facilitated the construction of a dynamic model of haematological disorder diseases such as MDS and AML progression, offering a platform to study disease evolution and identify molecular changes to facilitate the development of effective therapies targeting biomarkers of disease progression.

## 1.8 Aim & hypothesis.

MDS is a heterogeneous group of clonal haematological disorders that impair the proliferation and differentiation of HSCs, leading to deficiencies in mature blood cells. Cytopenia, anaemia, and dysplasia can arise from these disorders. There is no treatment for MDS, and the disease progresses to AML in 30% of the MDS cases. However, the causes of therapeutic failure and progression to AML remain unclear. Therefore, it is crucial to create novel model systems to better understand the molecular mechanisms causing this disorder and improve diagnostic and treatment techniques for MDS. We hypothesise that using iPSCs and CRISPR-Cas9 will facilitate the understanding of the role of two transcription factors, namely C/EBP $\alpha$  and MYBL2, in the development and progression of MDS.

A review of MDS by WHO in 2016 included two diseases related to *CEBPA*: AML with biallelic *CEBPA* mutations and AML with germline mutations in *CEBPA*. *CEBPA* mutations have been extensively researched in multiple studies, and several reviews have discussed their molecular mechanisms and clinical implications. However, in the present study, we examine the unique elements of biology and prognosis in monoallelic *CEBPA* mutant subtypes in human iPSCs and utilising molecular and cellular techniques to investigate the contribution of disruption of the *CEBPA* C-terminal domain to the dysplastic phenotype observed in MDS.

Moreover, few studies have demonstrated the vital role of *MYBL2* as a tumour suppressor gene. Approximately 50% of patients with high-risk MDS express a low level of MYBL2. In turn, low MYBL2 levels affect DNA damage response, leading to defects in HSCs, and changes in DNA damage response in HSCs have been associated with numerous blood diseases (Mohrin et al., 2010). Therefore, the present study aims to assess the role of MYBL2 in haematological disorders, especially MDS, using iPSCs as a model system. Also, we used immortalised myeloid cell lines.

The aim will be achieved by accomplishing the subsequent objectives and sub-objectives:

Determine the contribution of *CEBPA* to the MDS phenotype by:

1. Generation of hiPSC harbouring heterozygous mutations in *CEBPA* using CRISPR-Cas9
2. Characterisation of *CEBPA* mutant hiPSC by assessing pluripotency: alkaline phosphatase activity, pluripotent markers, and the capacity to differentiate into the three germline layers.
3. Validate the disease phenotype by differentiating *CEBPA* mutant hiPSC into HSPC and studying their proliferation and differentiating potential.

Determine the contribution of *MYBL2* to MDS development by:

1. Generation of stable hiPSCs with a modified *MYBL2* allele using CRISPR-Cas9
2. Differentiation of hiPSC with *MYBL2* deletion towards HSPC and the myeloid and erythroid lineages to determine how low *MYBL2* levels affect their differentiation.
3. Examine the involvement of *MYBL2* in myeloid and erythroid lineage differentiation in immortalised myeloid leukaemia cell lines (K562, KG1a, and SKM-1) in which *MYBL2* is downregulated by shRNA.
4. Investigating the effects of *MYBL2* downregulation on erythroid and myeloid differentiation in these cell lines following drug-induced erythroid differentiation (Ara-C), using different assay.

## **Chapter 2: Materials and Methods**

## **2.1 Cell culture**

### **2.1.1 Human iPSC culture methods**

Human iPSC BU3.10 cells were a generous gift from Dr George Murphy's lab at the Center for Regenerative Medicine, Boston University, Boston. These human iPSCs were generated from peripheral blood mononuclear cells by a single floxed-excisable lentiviral vector constitutively expressing the 4 factors and transduced with the STEMCCA Cre lentivirus (Sommer et al. 2012). The human iPSC lines were cultured in a 6-well plate coated with Matrigel (Corning Matrigel hESC qualified Matrix, 54277) under serum-free and feeder-independent conditions.

#### **2.1.1.1 Matrigel**

Prior to aliquoting the Matrigel, the required materials (tubes and tips) were kept at -20 °C to prevent solidification. The Matrigel bottle was submerged in an ice bucket overnight and thawed in a 4°C refrigerator. After thawing, all procedures were conducted on ice inside a sterile cabinet hood. The Matrigel bottle was swirled to ensure homogeneity before being aliquoting the Matrigel into sterile ice-cold Eppendorf tubes using cold sterile pipette tips.

The aliquoted volume was calculated using the dilution factor specified in the certificate of analysis. The dilution factor was determined based on the protein concentrations observed in each batch, 0.5 mg per 6-well plate and 1.0 mg per 2×6-well plate. The Matrigel was aliquoted through chilled pipette tips changed to avoid Matrigel clotting.

To thaw and coat a plate with Matrigel, a 12.5 mL cold DMEM/F-12 medium (Gibco, 11540446) was used to dilute a 1.0 mg Matrigel aliquot, carefully pipetting to thaw the Matrigel. Then, 1 mL of diluted Matrigel was added to each well of a six-well plate and incubated at room temperature for 1 h. The mixture was then promptly aspirated, and the wells were rinsed with phosphate-buffered saline (PBS, Sigma, D8537) before transferring the cells to the plate. If the plates were not used immediately, 1 mL hiPSC medium was added to each

well and storage at 37°C to avoid dehydration. The plates were covered with parafilm and maintained at 2–8°C for approximately 10 days or at 37°C for 5-7 days in an incubator.

#### **2.1.1.2 Thawing human iPSCs on Matrigel**

The hiPSCs were removed from liquid nitrogen and defrosted at 37°C until a tiny ice pellet remained visible in the cryovial. After gentle pipetting, the cells were transferred to a sterile 15 mL conical tube of StemFlex medium (Gibco, A3349401) and 100 µg/mL Primocin (Invivogen, ant-pm-2). The cells were then centrifuged at 150g for 2 min and delicately resuspended in 1 mL of StemFlex medium. Cells were plated on one well of six well-coated Matrigel plate; the plate was then slightly agitated back and forth and then left to stand for approximately 3 min to let the cells settle before incubation. The plate was incubated for 24 h at 37°C, in 5% CO<sub>2</sub>. The medium was replaced daily, and 70–90% of confluent cells were passaged every 5–6 days at a 1:6 ratio.

Colonies were detected first by light microscopy to passage the cells, and areas indicating cell differentiation were removed from the culture using a sterile scalp or P20 pipette inside a sterile cabinet hood. Colonies were cut into approximately 100 µm in size by mechanical passaging tools (StemPro EZPassage tool) (Invitrogen, 23181010).

#### **2.1.1.3 Cryopreservation of iPSCs**

StemPro EZPassage was used to cut cells at 70 to 90% confluency from the 6-well plate. The cells were pelleted by centrifugation at 150g for 2 min and the pellet was carefully resuspended in 1 mL of PSC Cryopreservation medium (Gibco, A2644601) and transferred to a marked cryovial tube. The tube was immediately placed into a Mr Frosty freezing container containing 100% (v/v) isopropanol and stored overnight at -80 °C to transferred it to liquid nitrogen for long-term storage.

### 2.1.2 Sub culturing of human cells

All acute myeloid leukaemia (AML) cell lines (SKM-1, K562 and KG1a) (Table 2-1), were grown in RPMI 1640 medium (Gibco, 21875034) containing 15% (v/v) heat-inactivated foetal bovine serum (FBS) (Sigma-Aldrich, F7524), 200 mM L-glutamine (Gibco, 25030024) and 100 U/mL penicillin-streptomycin (Gibco, 15070063).

For adherent cells, 293T cells were cultured in Dulbecco's modified Eagle's medium (DMEM; Gibco, 11995065) supplemented with 10% FBS, 200 mM L-glutamine and 100 U/mL penicillin-streptomycin. The cells were maintained in a humidified incubator at 37°C with a 5% CO<sub>2</sub> concentration. The MycoAlert™ mycoplasma detection kit (LT07-218, Lonza) was used at regular intervals to check the cells for mycoplasma contamination.

**Table 2-1 Type of cell line and source**

| <b>Cell line</b> | <b>Cell phenotype</b>                                     | <b>Source reference</b>                                     |
|------------------|---|---|
| <b>SKM-1</b>     | Human AML (M5) derived from MDS                           | DMSZ – German Collection of Microorganisms and Cell Culture |
| <b>K562</b>      | Erythrocytic cell derived from chronic myeloid leukaemia  | ATCC  |
| <b>KG1a</b>      | Promyeloblast macrophage cell line                        | Provide by group of Prof. Constanze Bonifier                |
| <b>293T</b>      | An epithelial-like cell that was isolated from the kidney | ATCC  |

The exponential growth of AML cell lines (SKM-1, K562, KG1a) was maintained by passaging every 2 to 3 days at  $2 \times 10^5$  cells/mL. The adherent 293T cell line was subcultured every 2–3 days at (1:10) ratio, or when they achieved 80% confluence. Briefly, the cells were detached by removing the cell culture medium, and washed with 1 mL of PBS, and incubating cells for 5 min at 37°C with 1 mL of 0.25% trypsin in 1 mM ethylenediaminetetraacetic acid

(EDTA) (Thermo Fisher Scientific,11560626). After incubation, the cells were tapped gently, and complete detachment was evaluated by light microscopy. The trypsin was inactivated by adding 5–10 mL of fresh medium, the cells were pelleted by centrifugation, and the matching dilution was carried out.

### 2.1.2.1 Cell quantification

The number of cells was calculated using haemocytometer counting chambers containing enhanced trypan blue (ThermoFisher, 15250061). In brief, 10  $\mu$ L of cell suspension was aliquoted from culture and pipetted immediately under the coverslip of a counting chamber, and number of cells was counted by light microscopy. The cellularity of the sample in cells/mL was calculated by multiplying the average number of cells per cuboid by  $1 \times 10^4$ .

### 2.1.2.2 Cell Cryopreservation

A sample containing  $1-10 \times 10^6$  cells were harvested by centrifugation at 300g for 4 min and resuspended in freezing medium consisting of 90% complete growth medium and 10% (v/v) DMSO. The cryovial was immediately placed into a Mr Frosty container, which was kept at  $-80^\circ\text{C}$ , and the tubes were transported to liquid nitrogen for long-term storage.

## 2.2 Lentivirus production

All plasmids used to knock down *MYBL2* in the leukaemia cell lines were purchased from Horizon Discovery and are listed in Table 2-2

**Table 2-2 Lentivirus plasmids**

| Primers  | Catalogue number  |
|--|-------------------|
| TRIPZ Inducible Lentiviral non-silencing shRNA Control | RHS4743           |
| TRIPZ Human MYBL2 shRNA                                | RHS4696-200679039 |
| TRIPZ Human MYBL2 shRNA                                | RHS4696-200703587 |



### **2.2.1 Virus preparation and transduction**

Prior to transfection, a total of  $1 \times 10^7$  293T cells were seeded in a 15 cm plate in 10 mL complete DMEM medium. After 24 h, the culture medium was carefully removed and refreshed with new DMEM medium. A mixture containing 90  $\mu$ L of Trans-IT 293 trans (Mirus, Mir2700)-IT/DMEM, including a mixture of 30  $\mu$ g DNA plasmid transfection mix, was transferred using the following plasmid proportions into a sterile 1.5 mL Eppendorf tube as per the manufacturer's instructions: (24  $\mu$ g backbone, 1.2  $\mu$ g tat, 1.2  $\mu$ g rev, 1.2  $\mu$ g gag/pol, 2.4:svs-g). The mixture was vortexed, left to stand for 15 min at 25°C, and then 2 mL of trans-IT/DNA/DMEM mixture was added dropwise to each plate and gently mixed with a back-and-forth motion in two directions, followed by incubation at 37°C in 5% CO<sub>2</sub>, for 24 h. At 48 h after transfection, supernatants were collected 3 times every 12 h in a 50 mL Falcon tube and stored at 4°C. The collected viral supernatant was filtered through a 0.45  $\mu$ m polyvinylidene fluoride (PVDF) filter, and viral particles were concentrated by ultracentrifugation at 4°C for 90 min at 17K (48960g on a Beckman SW32 rotor). The supernatant was removed, and the concentrated virus was aliquoted and frozen at -80°C within a sealed container.

Leukaemia cells ( $2.5 \times 10^5$  cells) were resuspended in 1 mL RPMI and transferred to 6-well plates for transduction. Polybrene (8  $\mu$ g/mL) (Sigma-Aldrich, TR-1003) and 70  $\mu$ L of concentrated lentivirus were added, and the plate was centrifuged for 1.5 h at 350 g at 32 °C for 'spinoculation'. The cells were then incubated for 2 h at 37°C in 5% CO<sub>2</sub>, washed twice with PBS, and resuspended in RPMI medium. The cells were left to recover for 48 h, then puromycin (1  $\mu$ g/mL) was added to the cell cultures for 5 days. Fresh new medium with antibiotics was added every two days.

### **2.3 Treatment of leukaemia cells with Ara-C**

Cytarabine (Ara-c) (Sigma-Aldrich) was prepared as a 1 mM stock solution in PBS and stored at -20°C. The drug was assessed by seeding SKM-1, K562, and KG1a cells in triplicate at

2.5×10<sup>5</sup> cells/mL and treated with 1 μM Ara-C for 48 h before incubating at 37°C in 5% CO<sub>2</sub>.

The cells were then harvested for further experiments.

## 2.4 Alkaline phosphatase stain

Human iPSCs were cultured in a 6-well plate until they reached 60% confluence, and most of the colonies were compact and displayed multi-layering in the centre. The fixation of cells was performed on ice in 1 mL of chilled neutral formalin solution containing 10% (v/v) formaldehyde for 15 min (Table 2-3), washed once after fixative removal and left for 15 min in cold distilled water (dH<sub>2</sub>O). After removing the dH<sub>2</sub>O, the cells were stained for 40 min at room temperature with 1 mL/well of AP stain. After removal of the staining solution, the cells were gently rinsed with deionised water and allowed to air dry. The cells were imaged using a Primovert microscope (ZEISS) and a Canon EOS 600D camera to capture the red, positively stained colonies.

**Table 2-3 List of buffers used for alkaline phosphatase staining**

| Buffer                      | Content   | Company           | Catalogue number |
|-----------------------------|---|-------------------|------------------|
| 10% neutral formalin buffer | 0.11M Na <sub>2</sub> HPO <sub>4</sub>                        | VWR               | 102494C          |
|                             | 25.6 mM of NaH <sub>2</sub> PO <sub>4</sub> .H <sub>2</sub> O | VWR               | 125330           |
|                             | 4% of PFA   | Thermo Scientific | 28906            |
| AP solution                 | 0.7mM Fast red violet LB salt                                 | Sigma-aldrich     | F3381            |
|                             | 0.2M of Tris-HCL  | Sigma-aldrich     | 10812846001      |
|                             | 27.3 mM N, N-Dimethylformamide                                | Sigma-aldrich     | 227056           |
|                             | 0.1m M Naphthol Phosphate                                     | Sigma-aldrich     | 70485            |

## **2.5 Identification of pluripotent markers**

### **2.5.1 Immunofluorescent analysis for cytoplasmic pluripotent marker**

Before staining, the cells were transferred to 35 mm plates (Corning Life Science, 430165) containing a sterile glass coverslip (18 mm × 18 mm, Thermo Scientific, 11798681) coated with Matrigel. For immunofluorescence staining of the TRA1-8 surface marker, the cells were rinsed twice with StemFlex medium before incubating with primary mouse TRA-81 antibody (Table 2-4) at 37°C for 1 h. After 3 washes with StemFlex medium, the cells were incubated at 37°C for 1 h with secondary antibody (Alexa Fluor-labelled goat anti-mouse-IgM) (Table 2-5). Following three washes with StemFlex medium, the cells were fixated for 10 minutes at room temperature with 2% methanol-free paraformaldehyde (Thermo Scientific, 28906).

The cells were then rinsed twice and mounted with prolonged medium containing DAPI (Invitrogen, P36941). Images were captured with a Leica microscope (Leica DM6000, Leica Microsystems) and processed with Leica software.

### **2.5.2 Immunofluorescence staining for nuclear pluripotent markers**

Cells were washed with PBS before fixation with 4% paraformaldehyde (Fisher Scientific, 28906) for 10 min. The cells were then washed twice with PBS, and 50 mM NH<sub>4</sub>Cl (Sigma, 254134) was added to the cells for 5 min to quench the free aldehyde groups. The cells were then permeabilised with Triton X-100 (0.5%, (v/v); Sigma, 11332481001) in PBS for 15 min at room temperature. The cells were blocked with PBS containing 1% bovine serum albumin (Sigma-Aldrich, A1933), 0.3% Triton X-100, 10% (v/v) FBS, and 1% (v/v) goat serum (Sigma, G9023) for 1 h at room temperature, then incubated overnight at 4 °C with primary anti-SOX2 and anti-NANOG antibodies diluted in blocking buffer (Table 2.4). After washing cells with PBS + 0.1% (v/v) Tween-20 (Sigma, P9416) (PBST), the cells were incubated with Anti-mouse Alexa Fluor® 633 and Anti-Goat Alexa Fluor® 488 secondary antibodies (Table 2.5) for 1 h at room temperature. Then, the cells were washed three times with PBST prior to mounting

with prolonged medium containing DAPI (Invitrogen, P36941). Images were captured with a Leica microscope and processed with Leica software.

## **2.6 Differentiation of hiPSCs**

### **2.6.1 Evaluation of trilineage differentiation of hiPSCs**

The three germ layers were differentiated using a Trilineage differentiation kit (STEMCELL Technologies, 05230), following the manufacturer's instructions. Briefly, 70% confluent hiPSCs were trypsinised into single-cell suspensions using a gentle cell dissociation reagent (StemCell Technology 07174). Cells were seeded at  $4 \times 10^5$  cells per well into 12-well Matrigel-coated plates (Corning Life Science, 3513) to evaluate ectoderm and endoderm differentiation and at  $1 \times 10^5$  cells per well to evaluate mesoderm differentiation. Cells were plated with STEMdiff Trilineage Ectoderm medium + 10  $\mu$ M of ROCK inhibitor (Y-27632) (Adooq Bioscience, A11001) for ectoderm differentiation and StemFlex + 10  $\mu$ M ROCK inhibitor (Y-27632) for endoderm and mesoderm differentiation and incubated at 37°C. Each cell type was then fed the corresponding differentiation medium on day 1. The medium was refreshed daily, and the cultures were kept in differentiation medium for 7 days for the ectoderm and 5 days for the mesoderm and endoderm. After that time, the expression of lineage-specific markers was analysed by immunofluorescence, as explained in section 2.5.2, using specific unconjugated antibodies (R&D Systems) for ectoderm (Otx2), mesoderm (Brachyury), and endoderm (SOX17) primary and secondary antibodies as listed in (Table 2-4 and Table 2-5).

**Table 2-4 List primary antibodies for immunofluorescence staining protocol**

| <b>Antibody</b> | <b>Immunoglobulin</b> | <b>Clone</b> | <b>Company</b> | <b>Catalogue number</b> | <b>Stock concentration(<math>\mu\text{g/ml}</math>)</b> | <b>Dilution</b> |
|-----------------|-----------------------|--------------|----------------|-------------------------|---|-----------------|
| TRA1-81         | Mouse IgM             | C1.261       | Invitrogen     | MA1-024                 | 100   | 1:50            |
| NANOG           | Goat IgG              | ABZ92376     | R&D            | AF1997                  | 100   | 1:100           |
| SOX2            | Mouse IgG             | 245610       | R&D            | AF2018                  | 100   | 1:100           |
| OTX2            | Goat IgG              | 246826       | R&D            | AF1979                  | 100   | 1:100           |
| Brachyury       | Goat IgG              | 1161B        | R&D            | AF2085                  | 100   | 1:100           |
| SOX17           | Goat IgG              | 245013       | R&D            | AF1924                  | 100   | 1:100           |

**Table 2-5 List of secondary antibodies for immunofluorescence staining protocol**

| <b>Antibody</b>                     | <b>Immunogen</b> | <b>Company</b> | <b>Catalogue number</b> | <b>Stock concentration(mg/ml)</b> | <b>Dilution</b> |
|-------------------------------------|------------------|----------------|-------------------------|-----------------------------------|-----------------|
| Goat anti-mouse Alexa<br>Fluor 488  | Mouse IgM        | Invitrogen     | A10684                  | 0.5                               | 1:50            |
| Goat anti-mouse Alexa<br>Fluor 633  | Goat IgG         | Invitrogen     | A21052                  | 1                                 | 1:200           |
| Donkey anti-goat Alexa<br>Fluor 488 | Goat IgG         | Invitrogen     | A32841                  | 1                                 | 1:200           |

### **2.6.2 Hematopoietic differentiation (STEMdiff)**

hiPSCs were differentiated into haematopoietic progenitor cells using a STEMdiff Haematopoietic kit (STEMCELL Technologies, 05310) following the manufacturer's instructions. Cells were cut into fragments using Ez passage roller. Then, approximately 16–20 colony fragments (100–200  $\mu\text{m}$  in size) were seeded into each well of a Matrigel-coated 12-well plate and incubated at 37°C for 24 h. The following day, the plate was cleaned to eliminate fragments from colonies larger than 200  $\mu\text{m}$  in diameter or if they exceeded 20 attached fragmented clones. The cells were grown in medium A 1.5 mL per well for 2 days, with half of the medium changed. The medium was changed on day 3 to medium B over 11 days, and 500  $\mu\text{L}$  of medium was replaced every other day. The presence and number of haematopoietic progenitor cells were determined at several time points (days 7, 10, 12, and 14) by flow cytometry (see section 2.8).

### **2.6.3 Clonogenic progenitor assay (Colony forming assay)**

Haematopoietic clonogenic assays were conducted according to the MethoCult protocol (STEMCELL Technologies). MethoCult H4435 semi-solid medium (STEMCELL Technologies, 04435) was thawed at 4°C overnight, and 250  $\mu\text{L}$  penicillin-streptomycin (Gibco, 15140122) was added before vigorous shaking and left for approximately 1 h to settle either at 4°C or at room temperature. The medium was then aliquoted and store at -20°C as a full MethoCult medium. To initiate the colony forming assay,  $1 \times 10^4$  haematopoietic cells from day 12 of STEMdiff differentiation were seeded in 35-mm dishes with 1.2 mL of full MethoCult H4435 medium per dish in duplicate. Because of the high viscosity of the methylcellulose medium, the required volume was measured with syringes and 18G sharp needles (BD 30518). The 35 mm plastic dishes were placed on a 150 mm plate (Corning Life Science, 430599) to incubate at 37°C in 5%  $\text{CO}_2$ , with an open 35 mm dish containing PBS to maintain humidity. Technical duplicates were always performed. Colonies were evaluated after

incubation for 14 days. The colonies were counted and scored in duplicate based on morphology, size, and colour criteria from STEMCELL Technologies; erythrocytes (CFU-E), granulocytes, erythrocytes, macrophages, megakaryocytes colonies (CFU-GEMM), granulocyte, macrophage colonies (CFU-GM), granulocytes (CFU- G) and macrophage colonies (CFU-M).

For the AML cell lines, 500 cells from each cell line were cultured in duplicate on 1 mL methylcellulose-based medium (MethoCult, STEMCELL Technologies) with 1  $\mu$ M Ara-C. Thereafter, cells were captured after 4 days using an EVOS microscope at 4 $\times$  or 10 $\times$  magnification.

#### **2.6.4 Erythroid differentiation in liquid culture**

Erythroid cells were differentiated according to a previously reported methodology (Bayley et al., 2018) with some amendments. Briefly, single haematopoietic cells (day 10 from STEMdiff protocol) were plated in a 12-well plate at a density of  $1-2 \times 10^5$  cells/mL and cultured for 18 days in Iscove's Modified Dulbecco's medium IMDM (Sigma, 12440053) comprising human transferrin (500 g/mL; R&D, 2914-HT- 001G), 3 mg/mL human plasma (Finished Product, QE Pharmacy), insulin (10 g/mL; Sigma, 19278), heparin (3 U/mL; Sigma, H3149), and 3% (w/v) human serum albumin (Sigma, H4522). The following cytokines were used to promote erythroid differentiation of haematopoietic cells (Table 2-6):

**From Day 0 to Day 8:** SCF, IL-3 and EPO.

**From Day 8 to Day 11:** SCF and EPO.

**From Day 12 to Day 18:** only supplemented with EPO.

Every four days, flow cytometry examination of erythroid markers (CD71-APC and CD235a-PE) (see section 2.8) was conducted. Kwik-diff staining was performed for morphological analyses (see section 2.9).

**Table 2-6 Cytokines for erythroid differentiation**

| <b>Cytokine</b> | <b>Concentration</b> | <b>Company</b> | <b>Catalogue number</b> |
|-----------------|----------------------|----------------|-------------------------|
| SCF             | 100 ng/ml            | Peprtech       | 255-SC-010              |
| IL-3            | 5 ng /ml             | Peprtech       | 203-IL-010              |
| EPO             | 3 U/ml               | Peprtech       | 100-64                  |

**2.6.5 Myeloid differentiation in liquid culture**

Myeloid differentiation was performed in accordance with a previously described technique (Hansen et al., 2018) with a few adjustments. Single haematopoietic cells (day 12 from the STEMdiff protocol) were plated at a  $1 \times 10^5$  cells/mL density on a 12-well plate and cultured for 7 days in Stemline II® haematopoietic stem cell growth medium enriched with 1% penicillin/streptomycin and the following cytokines (listed in Table 2-7) to induce myeloid cell formation:

**Day 0 to Day 3:** IL-3, GM-CSF, G-CSF, FLT-3, hSCF and G-CSF.

**Day 3 to Day 7:** G-CSF.

On days 4 and 7, flow cytometry analysis of myeloid markers (CD11b-PECy7 and CD14-APCCy7) was performed to assess myeloid differentiation. Myeloid cells were analysed morphologically following Kwik-Diff staining.

**Table 2-7 Cytokines for Myeloid differentiation**

| <b>Cytokine</b> | <b>Concentration</b> | <b>Company</b> | <b>Catalogue number</b> |
|-----------------|----------------------|----------------|-------------------------|
| SCF             | 50 ng/ml             | Peprtech       | 255-SC-010              |
| IL-3            | 10 ng /ml            | Peprtech       | 203-IL-010              |
| GM-SCF          | 10ng/ml              | Peprtech       | 300-03                  |
| FLT-3           | 50ng/ml              | Peprtech       | 300-19                  |
| G-CSF           | 30ng/ml              | Peprtech       | 300-23                  |



## **2.7 Cytogenetics**

Haematopoietic stem/progenitor cells (HSPCs) were harvested on day 12 of differentiation, and 0.02 g/mL of KaryoMAX Colcemid Solution (ThermoFisher, 15212012) was added directly to the cells before being placed at 37°C for 2 h to facilitate cell cycle arrest at metaphase. The cells were then centrifugated at 300g for 8 min and then swollen for 20 min at 37°C in a warmed hypotonic lysis solution (0.075 M potassium chloride). Cells were fixed for 20 min at room temperature in cold fixative buffer (3:1 [v/v] methanol/glacial acetic acid), centrifuged at 300 g for 5 min, rinsed twice and incubated in fixative at 4°C for 30 min. The cells were washed, resuspended in the cold fixative buffer, dropped onto chilled, humidified slides, and left to dry. The mitotic index/metaphase spread quality was assessed by phase contrast microscopy. The slides were then aged on a heating block overnight set to 80°C. Chromosomes were stained with Giemsa-modified solution, rinsed with distilled water for 30 s, and observed and counted at 100× by brightfield microscopy with an oil immersion lens (Leica DM6000, Leica Microsystems). 30 mitotic cells per cell clone were chosen randomly from three independent experiments, and chromosomes were captured using (Leica DM6000, Leica Microsystems) and counted.

## **2.8 Flow cytometric analysis of differentiation markers**

The ability of hiPSCs to differentiate into HSPCs was determined by flow cytometry. Haematopoietic cells were isolated at various stages of the STEMdiff erythroid and myeloid differentiation protocols. Cells were centrifuged at 300 g for 5 min. The cells were resuspended in 120 µL of PBS containing 2% FBS for surface staining performed in 96-well plates (Costar, 3367) on ice. The cells were then incubated with 4 µL of Fc block; eBioscience, 14916173, Clone AB468581) for 20 min, followed by a 1 h incubation with several antibody combinations (Table 2-8). HSPCs were identified using the haemato-endothelial marker CD34 and the haematopoietic stem and progenitor markers CD43 and CD45. The CD71 and CD235a markers

were used to characterise **erythroid differentiation**, and the CD11b and CD14 markers were used to characterise **myeloid lineage differentiation**. The cocktail of isotype controls is detailed in Table 2-9.

After antibody incubation for 1h, the cells were rinsed with 100 µL of PBS containing 2% FBS and centrifuged at 300 g for 2 min. After discarding the supernatant, the cells were resuspended in 300 µL PBS containing 2% FBS, transferred to flow cytometry tubes after being filtered using sterile single-pack CellTrics® filters (Sysmex,04-004-2323), and analysed using a Beckman Coulter CyAn™ADP flow cytometer. The flow cytometry data were analysed using FlowJo v10.6.2 software.

**Table 2-8 List of antibodies used for flow cytometry**

| <b>Antibody</b> | <b>Fluorochrome</b> | <b>Clone</b> | <b>Stock concentration (mg/ml)</b> | <b>Company</b> | <b>Catalogue number</b> | <b>Dilution</b> |
|-----------------|---------------------|--------------|------------------------------------|----------------|-------------------------|-----------------|
| CD34            | PE                  | 8G12         | 0.2                                | BD Pharmingen  | 550619                  | 1:100           |
| CD43            | APC                 | 1G10         | 0.2                                | BD Pharmingen  | 560198                  | 1:100           |
| CD45            | FITC                | HI30         | 0.2                                | ThermoFisher   | 11045942                | 1:100           |
| CD14            | APC-CY7             | 61D3         | 0.2                                | ThermoFisher   | 47014942                | 1:100           |
| CD71            | APC                 | OKT-9        | 0.2                                | ThermoFisher   | 17071941                | 1:100           |
| CD235a          | PE                  | GA-R2        | 0.2                                | ThermoFisher   | 12-9987-80              | 1:100           |
| CD11b           | PE-CY7              | ICRF44       | 0.2                                | ThermoFisher   | 15518356                | 1:100           |

**Table 2-9 List of isotype antibodies used for flow cytometry**

| <b>Antibody</b>   | <b>Fluorochrome</b> | <b>Clone</b> | <b>Stock concentration (mg/ml)</b> | <b>Company</b> | <b>Catalogue number</b> | <b>Dilution</b> |
|-------------------|---------------------|--------------|------------------------------------|----------------|-------------------------|-----------------|
| Mouse-IgG1-Kappa  | PE                  | MOPC-31C     | 0.2                                | BD Pharmingen  | 550617                  | 1:100           |
| Mouse IgG2b-Kappa | PE                  | Ebm2b        | 0.2                                | ThermoFisher   | 12473281                | 1:100           |
| Mouse-IgG1-Kappa  | APC                 | P3.6.2.8.1   | 0.2                                | ThermoFisher   | 17471482                | 1:100           |
| Mouse-IgG1-Kappa  | FITC                | P3.6.2.8.1   | 0.2                                | ThermoFisher   | 11471481                | 1:100           |
| Mouse-IgG1-Kappa  | APC-CY7             | P3.6.2.8.1   | 0.2                                | ThermoFisher   | 47471482                | 1:100           |
| Mouse-IgG1-Kappa  | PE-CY7              | P3.6.2.8.1   | 0.2                                | ThermoFisher   | 11520627                | 1:100           |

## **2.9 Kwik-Diff staining**

Kwik-Diff staining (ThermoFisher, 9990700) was used to determine the erythrocyte and myeloid morphology. Briefly, cells were centrifuged for 5 min at 300 g, placed on slides, air-dried, and then immersed in fixative solution (solution I) for 30 s, followed by immediate transfer into eosin solution (solution II) for 30 s and then into methylene blue (Solution III) for 20 s. The slides were washed thoroughly with water and allowed to dry completely, then

mounted using histological mounting medium (Scientific Laboratory Supplies, NAT1310) for examination. The images captured by brightfield microscopy (Leica DM6000, Leica Microsystems).

## **2.10 Genomic DNA extraction**

DNA was extracted according to the manufacturer's protocol for the DNeasy blood and tissue kit (Qiagen, 69504). Briefly,  $5 \times 10^6$  single hiPSC cells were resuspended in 200  $\mu$ L PBS and 20  $\mu$ L proteinase K and incubated in lysis buffer (AL) at 56°C for 10 minutes. Subsequently, 200  $\mu$ l of 100% ethanol was added to clean the sample thoroughly, the entire sample was placed on a micro spin column, and the DNA was purified using several washing buffers (AW1&AW2). Lastly, 200  $\mu$ L of buffer AE was used to elute the DNA then the DNA was measured using a Nanodrop 2000c spectrophotometer (ThermoFisher).

## **2.11 RNA analysis**

### **2.11.1 RNA isolation**

RNA was extracted from cells by the RNeasy kit (Qiagen, 74106) according to the manufacturer's instructions. Cells were washed with PBS and lysed with 350  $\mu$ L RLT lysis buffer and 350 $\mu$ l of 70% ethanol, placed on a mini spin column to bind RNA and centrifuged. The RNA samples were treated with RNase-free DNase (Qiagen, 79254) to remove residual DNA. Briefly, 10  $\mu$ L of DNase solution was added to 70  $\mu$ L buffer RDD and mixed gently. The mixture was added to the mini spin column membrane and incubated at room temperature for 15 min. Then the column was washed with different buffers to remove biomolecules, such as protein. RNA-free water was used to elute the RNA into a sterile microcentrifuge tube. The quantity of RNA was determined by measuring absorbance ratios (A260/280) with a Nanodrop 2000c spectrophotometer (ThermoFisher Scientific).

### **2.11.2 cDNA synthesis**

For reverse transcription, 1–5 µg of RNA from each sample was incubated with 0.5 µg oligo dT 15 primers (Promega, C1101) and heated at 70°C for 5 min, followed by cooling down for 5 min on ice. Subsequently, 1.25 µL dNTP mix (10 mM; Invitrogen, 10297018), 1 µL RNase OUT (Invitrogen, 10777019), 5X Moloney murine leukemia virus (M-MLV) reaction buffer, and 1 µL M-MLV reverse transcriptase enzyme (200 U/µL; Promega, M1701) were added and the mixture was incubated at 42°C for 1 h and stored at -20 °C.

### **2.11.3 Quantitative real-time PCR (qPCR)**

Real-time qPCR (using either SYBRGreen or TaqMan) was conducted using a MicroAmp optical 96-well reaction plate (Life Technology, 430673). For TaqMan real-time qPCR, 10 µL of TaqMan universal master mix II (ThermoFisher Scientific, 4426710), 8 µL of RNase/DNase-free water, and 1 µL Taqman oligos (Table 2-10) were mixed and 1 µL of cDNA was added to the mixture in the plate. Each reaction was performed in triplicate (technical replicates). The TaqMan qPCR, was carried out in Stratagene Mx3005P (Agilent Technologies), the qPCR programme included a denaturation step (50°C for 2 min, 90°C for 10 min) followed by 40 amplification cycles (95°C for 15 s, 60°C for 1 min).

For SYBR Green-based PCR, the reaction consisted of 10 µL 2X SYBR™ Green master mix reagent (Invitrogen, 4309155), 1 µL of cDNA, forward and reverse oligos at a 100 nM concentration (Table 2-11); RNase/DNase-free water (Promega, P1193) was added to a final volume of 20 µL. The thermocycler for SYBR green was carried out in Stratagene Mx3005P (Agilent Technologies), and the programme was consisted of an initial hot start cycle at 95 °C for 1 min, followed by 40 amplification cycles at 95 °C for 30 s, 58 °C for 1 min, 72°C for 1 min and ending with a dissociation cycle (95°C for 1 min, 58°C for 30 s and 95°C for 30 s).

The Ct values were calculated using MxPro 3000 Stratagene software. The calculation of relative gene expression values against GAPDH was performed using the  $\Delta\Delta C_t$  mathematical model (Livak and Schmittgen 2001).

**Table 2-10 TaqMan® assays (ThermoFisher Scientific)**

| Oligo       | Catalogue number |
|-------------|------------------|
| CEBPA       | Hs00269972_s1    |
| RUNX1       | Hs01021970_m1    |
| GATA2       | Hs00231119_m1    |
| EPO         | Hs00959427_m1    |
| SPI1 (PU.1) | Hs02786711_m1    |
| GAPDH       | Hs02758991_g1    |

**Table 2-11 KiCqStart™ primers KSPQ12012 (Sigma-Aldrich)**

| Oligo | Catalogue number |
|-------|------------------|
| GYPA  | NM_002099        |
| TFRC  | NM_001128148     |
| ITGAM | NM_00063224-25   |

## 2.12 CRISPR-Cas9 to modify *CEBPA* and *MYBL2* in hiPSCs

### 2.12.1 Designing single guide RNA (sgRNA)

This protocol was described previously by Ran et al. (2013). An sgRNA was designed using an online tool (Trust Sanger Institute Editing database) to target the C-terminal domain of *CEBPA* gene on chromosome 19 at position 13.11 and *MYBL2* on chromosome on 20q13.12.

A pair of complementary oligonucleotides, consisting of a forward and reverse sequence, were used as the sgRNA. The sgRNA possessed a homology of 18–20 base pairs with the target sequence. The nucleotide sequences 'CACC' and 'AAAC' were added to the 5' ends of the sgRNA forward and reverse pair and sent to Sigma-Aldrich for synthesis.

### 2.12.2 Annealing sgRNA oligos

Forward gRNA and reverse gRNA (100  $\mu$ M each) were mixed in a PCR tube with 10 U of T4 polynucleotide kinase (PNK, Promega, M4101) and 1 $\times$  T4 PNK buffer. The annealing reaction was performed using a PCR machine with the following programme: 37°C for 30 min,

followed by 95°C for 5 min, and the temperature was reduced to 25°C at 5°C per min. The annealed gRNA was then diluted 1:200 to final concentrations of 0.1 µM for subsequent cloning into the PX548 vector.

### **2.12.3 Cloning of gRNA into the pSpCas9 (BB)-2A-GFP (PX458)**

This step requires digestion of the PX458 and ligating gRNA into the vector in a single step. First, 100 ng of PX458 vector (Addgene, 48138), 1× Tango buffer (Thermo Fisher, By5), 1 µL Fast digest BbsI (Thermo Fisher, FD1014), 10 mM DTT (Invitrogen, Y00147), 1500U T7 DNA Ligase (NEB M0318S), 10 mM ATP (Invitrogen, 55082) and 0.2 ng dsRNA oligo were mixed and then incubated in a thermocycler at 37°C for 5 min, and 21°C for 5 min; this cycle was repeated 6 times. A Plasmid Safer Exonuclease kit (Lucigen, E3101K) was then used to digest any residual linearised DNA. Briefly, the complete ligation PCR product from the preceding step was treated with 1U of plasmid-safe exonuclease, 10 mM ATP, and 1x plasmid-safer buffer. Then, the entire reaction was incubated in the PCR machine at 37 °C for 30 min and 70 °C for 30 min.

### **2.13 Competent bacteria transformation**

The ligated DNA plasmid (4 µL) was mixed with 100 µL of one-shot Stbl3 chemically competent *E. coli* (Life Technology, C7373-03). This bacterial strain has a reduced frequency of homologous recombination. The mixture was placed in a 42°C water bath for 30 sec before being placed on ice for 5 min. The bacteria were cultured overnight at 37°C on a Lysogeny broth (LB) agar plate with 50 µg/mL ampicillin (Sigma, 10835242001). On the following day, approximately 15 colonies were picked for further growth in 4 mL of LB containing 100 µg/mL ampicillin.

### **2.14 DNA extraction from plasmids (Miniprep and Maxiprep)**

Miniprep of bacterial colonies was performed using A QIAprep Spin Miniprep kit (Qiagen, 12123).

Bacterial culture (2 ml) was sedimented at room temperature by centrifugation at 6800 g for 3 min. The pellet then resuspended in buffer P1, before being transferred to a microcentrifuge tube and lysed by adding P2 buffer and gently inverting the tube 3–7 times. After thoroughly mixing with P3 buffer, the solution was centrifuged for 10 min at 17,900 g. The DNA was then purified by passing the supernatant through a spin column. Following washing the column with PE buffer, the DNA was eluted with EB buffer into microcentrifuge tube. The DNA concentration was determined using a Nanodrop instrument and stored at -20° C. The DNA from colonies was sequenced using the U6 promoter primer (5'TACGATACAAGGCTGTTAGA'3) to determine the cloning success. The sequencing was performed at the Nottingham-based Source Bioscience Laboratory.

The maxiprep for the plasmid isolated from a positive colony was performed using an endotoxin-free plasmid plus maxi kit (QIAGEN, 12362). A pre-culture of 4 mL of bacteria was inoculated into 400 mL of LB and incubated overnight at 37°C on a shaker at 250 rpm.

The following day, the bacterial culture was transferred to a 500 mL tube and centrifuged for 30 min at 3000 g at 4°C. After discarding the supernatant, the pellet was resuspended in 10 mL of P1 buffer. Then the bacterial suspension was mixed with 10 ml of lysis buffer P2 buffer and incubated at room temperature for 5 min. Subsequently, the tube was then inverted 4-5 times after adding 10 mL of neutralisation buffer (P3 buffer). The bacterial lysate was then poured into the QIAfilter Cartridge and left at room temperature for 10 minutes to allow the sediment to float to the top of the solution.

Once the sediment had risen to the top, the plunger was gently inserted to transfer the lysate to conical tube via QIAfilter Maxi Cartridge.

Subsequently, the filtered lysate was placed in a 50 mL tube with ER buffer and incubated on ice for 30 minutes. The lysate was then loaded onto a QIAGEN-tip column until the whole sample was passed through. After loading the lysate onto a QIAGEN-tip column the column



was washed with QC buffer, and the DNA was eluted with 15 mL QN buffer into a fresh 50 mL tube. The DNA was precipitated with 10.5 mL isopropanol, and immediately centrifuged at 15,000 g for 30 min at 4°C. The pellet was washed with 70% ethanol, redissolved in Tris EDTA buffer, and the DNA yield was determined with a Nanodrop instrument. The samples were stored at -20°C.

### **2.15 Electroporation and selection of hiPSCs**

The electroporation protocol was performed using the P3 Amaxa kit (Lonza, V4XP-3024), as described in the manufacturer's instructions. First, the nucleofector reagent (solution P3 + supplement) was added to a suspension of  $1 \times 10^6$  hiPSCs, followed by 1 µg PX458+ gRNA of *CEBPA* or gRNA of *MYBL2* and 1 µg of PX458 plasmid alone (as a control). The mixture was transferred to a nucleocuvette and inserted within a 4D nucleofection system (Lonza). Nucleofection was performed using the DS-150 programme. After nucleofection, 1 mL of warm medium containing 5µM ROCK inhibitor (Y-27632, LKT laboratory, Y1000) was added to the cuvette and incubated at 37 °C for 10 min. The cells were then transferred to two wells of a 6-well Matrigel-coated plate in medium containing rock inhibitor and grown for 24 h prior to sorting. For sorting, the transfected cells were detached from their plates by incubation with TrypLE (Thermo Fisher, 12604013) for 10 min at 37 °C. The cells were then collected from the TrypLE medium, and 2 mL Stemflex medium was added, and the cells were centrifuged at 300 g for 5 min. The cell pellet was resuspended in StemFlex medium and passed through sterile single-pack CellTrics® filters. The GFP+ cells were sorted on a BD FACSAria™ Fusion instrument (BD Bioscience) and collected into tubes containing StemFlex medium. The sorted cells were then cultured in Matrigel-coated plates at a low-density dilution. Seven days after sorting, clones were picked and expanded for further analysis.

## 2.16 Mutational screening following CRISPR-Cas9 editing

The *CEBPA* or *MYBL2* targeted area was amplified using Q5 Hot start High-Fidelity 2X DNA polymerase (NEB, M0494S). Each reaction contained 1X Q5 High-Fidelity DNA polymerase, 0.5  $\mu$ M of forward and reverse oligo, and 100 ng DNA. The reaction was performed in a 25  $\mu$ L volume. The PCR programme used for amplification was 30 s at 98 °C, followed by 30 cycles at 98 °C for 10 s, 64 °C for 30 s, and 72 °C for 30 s, and a final extension at 72 °C for 2 min.

Primers used for PCR:

CEBPA forward primer 5'GGCCTCTCCCTTACCAGCC'3

CEBPA reverse primer 5'CTGGTCAGCTCCAGCACCTT'3

MYBL2 forward primer 5'TCAGGTGGATGTGAAGGGCT'3

MYBL2 reverse primer 5'CACCCTTCCTCCTGACTATGCAT'3

The DNA amplification was verified by agarose gel electrophoresis performed by dissolving 1% (w/v) agarose (Sigma, A9539, w/v) in 1 $\times$  TAE buffer (40 mM Tris pH 7.6, 20 mM acetic acid and 1 mM EDTA) by heating for 3 min on high power in a microwave oven. Ethidium bromide (0.05  $\mu$ g/mL Sigma, E8751) was added, and the mixture was poured into a gel casting tray. The gel was allowed to solidify for approximately 30 min at room temperature and then transferred to a horizontal electrophoresis tank and coated with 1 $\times$  TAE buffer. Each PCR sample was loaded onto the gel with 5  $\mu$ L of 100 bp DNA ladder using a 6 $\times$  loading buffer (Orange G) (NEB, N0551G). The gels were run at 100 V for approximately 1 h, then the gel was removed and placed into the ChemiDoc MP imaging system (Bio-Rad) to assess the plasmid sizes to the DNA ladder. The expected band size for wild type pf *CEBPA* was 456 bp and 420 bp for *MYBL2*. The positive clones were then subjected to an endonuclease I test utilising a T7 endonuclease I.

## 2.17 T7 Endonuclease I (T7EI) assay

The PCR products were analysed with a T7 endonuclease I (T7EI) assay and the Alt-R genome editing detection kit (IDT, 1075932) following the manufacturer's protocol. The PCR products encoding the *CEBPA* or *MYBL2* target area were incubated for 60 min at 37°C with 1 U/L of

T7EI to determine the editing efficiency of sgRNA at recognising and cleaving the non-perfectly matched DNA generated by gRNA. Two positive controls (Control A, a homoduplex control, and Control B, a heteroduplex control) were employed to assess T7EI assay function. As indicated in Section 2.18, the DNA PCR products were analysed on a 2% agarose gel and sent for Sanger sequencing to identify the mutation type.

## **2.18 Agarose gel extraction**

DNA was extracted from the agarose gels and purified using a gel extraction kit (QIAGEN, 28704). The required DNA band was excised from agarose gel using a clean, sharp scalpel under blue light. QG buffer (3× the gel weight) was added, and the sample was incubated at 50°C for 10 min to dissolve the gel slice completely. After adding one volume of isopropanol to the sample, the mixture was passed through a QIAquick column to precipitate DNA. PE buffer was then used to wash the column, and the DNA was eluted with EB buffer into 1.5 mL microcentrifuge tube. The DNA concentration was determined using a Nanodrop 2000c (Thermo Fisher) instrument. The purified DNA samples were sent to the Source Bioscience facility in Nottingham, UK, for sequencing using the primers described below.

CEBPA forward primer 5'GGCCTCTCCCTTACCAGCC'3  
CEBPA reverse primer 5'CTGGTCAGCTCCAGCACCTT'3  
MYBL2 forward primer 5'TCAGGTGGATGTGAAGGGCT'3  
MYBL2 reverse primer 5'CACCCTTCCTCCTGACTATGCAT'3

## **2.19 Western blot analysis**

### **2.19.1 Protein extraction and quantification**

Protein was extracted as described by Lorvellec et al. (2009). Cells at 80% confluency or  $2 \times 10^6$  cells were used. Cells were washed by centrifuging at 250 g at 4°C for 5 min, resuspending in 1mL pre-cooled PBS, and centrifuging again at 250 g at 4°C for 5 min. The washed cells were lysed in ice-cold RIPA buffer containing 10 mM EDTA, 20 mM Tris pH 7.4, 1 mM sodium fluoride, 1 mM  $\beta$ -glycerophosphate, 100 mM NaCl, 1% (v/v) Triton-100, 1 mM ethylene glycol tetra-acetic acid, 5 mM polypropiolate sodium, 1 mM phenylmethylsulfonyl

fluoride (Sigma, P7626) and 1 mM sodium orthovanadate (Sigma, S6508), and 1x halt protease inhibitor cocktail (ThermoFisher, 78430) on ice. The lysates were incubated for 30 min on ice and then centrifuged at 800 g at 4°C for 10 min. The supernatant from each sample was collected into three new microcentrifuge tubes. One tube was used for protein concentration measurements, and the other two were snap-frozen and stored at -80°C.

### **2.19.2 Protein quantification**

The Bradford absorbance assay was conducted to determine each sample's protein concentration. Initially, a series of bovine serum albumin (BSA) concentrations (0, 1, 2, 4, 6, 8, and 10 µg/µL) were prepared from a 1 µg/µL of BSA (Sigma-Aldrich, 1076192). Then, 800 µL of water was added to BSA and mixed with 200 µL of the protein assay dye reagent (Bio-Rad, 5000006) to generate a standard curve. A 1 µL aliquot of each protein sample was treated in the same manner. The absorbance of all samples was measured at 595 nm using a spectrophotometer, and a graph of absorbance versus concentration was plotted. The absorbance of each sample was compared with the standard curve to determine the protein concentration (µg/µL).

### **2.19.3 SDS-PAGE and immunoblotting**

For protein identification, 50 µg of protein lysate was added to 4X Laemmli buffer consisting of 0.04% (w/v) bromophenol blue, 40% (v/v) glycerol, 8% (w/v) sodium dodecyl sulphate (SDS), and 5% (v/v) β-mercaptoethanol in 240 mM Tris/HCl pH 6.8. The samples were denatured at 95°C for 10 min. Readymade precast 10% acrylamide SDS-PAGE gels (Bio-Rad) were placed in a tank filled with 1× SDS-PAGE running buffer (25 mM Tris, 190 mM glycine and 0.1% (w/v) SDS, pH 8.3). To load the samples, 8 µL of pre-stained protein marker (P7719, New England Biolabs) and 50 µg of total protein were loaded into the wells. The gel was run at 100 V for approximately 50 min. The separated proteins were transferred to a PVDF membrane (Millipore) using a wet transfer system (Bio-Rad transblot turbo) at 100 Amp for

90 min. The membranes were then blocked with 5% (w/v) milk powder in Tris-buffered saline containing Tween20 (TBST; 20 mM Tris pH 7.5, 150 mM NaCl, plus 0.1% [v/v] Tween20) on a rocking platform for 1 h. The membrane was then incubated with B-MYB primary antibody (1:500 dilution; N-19, Santa Cruz) or C/EBP $\alpha$  primary antibody (Rabbit, 1:1000 dilution; ab10548, Abcam) overnight at 4°C on a rolling shaker. The following day, the membrane was washed three times with 10 mL of TBST for about 15 mins each time. The membrane was then incubated with 6 mL blocking buffer containing horseradish peroxidase (HRP)-linked secondary antibody (1:5000 dilution; Amersham, NA9340) on a rocking platform for 1 h. The membrane was washed three times with TBST to minimise background and remove unbound secondary antibodies. The protein bands (42 and 30 kDa for C/EBP $\alpha$  and 72 kDa for MYBL2) were then visualised using Pierce™ ECL Western Blotting Substrate (Thermo Scientific, 32106). A ChemiDoc MP imaging system (Bio-Rad) was used to develop chemiluminescence onto the photographic film. Blotting for the loading control was performed by washing the membrane 3 times with TBST, incubating it with HRP-conjugated antibody to the housekeeping protein  $\beta$ -actin (CST 5125, Rabbit, Cell Signalling Technology) for 1 h, washing, and visualising using the same protocol.

## **2.20 Statistical analysis**

GraphPad Prism 9 was used for statistical analysis (GraphPad Prism version 9.0, GraphPad Software, San Diego, California, USA). The data are presented as the mean  $\pm$ Standard Error (SEM). Statistical significance was set at  $p < 0.05$ , and significant figures were labelled with an asterisk (\*). The statistical analyses employed were two-way ANOVA, one-way ANOVA, and unpaired Student's t-test.

## **Chapter 3: Generation and Characterisation of hiPSC with a *CEBPA* monoallelic mutation**

### 3.1 Introduction

MDS is characterised by bone marrow failure and a propensity to evolve into acute myeloid leukaemia (AML). Approximately 30% of MDS cases progress to AML (Menssen and Walter 2020). Hypomethylating agents (HMAs) are effective licenced therapies for individuals with MDS; however, a number of patients do not benefit from this treatment, and almost all cases eventually cease to respond to HMAs (Cogle et al. 2017). Therefore, developing effective diagnostic and treatment techniques for blood diseases requires a deeper understanding of their underlying mechanisms. One gene involved in controls quiescence of haematopoietic stem cells and their granulocytic commitment is *CEBPA* (Ye et al. 2013). At initial diagnosis, 8% of MDS patients show monoallelic *CEBPA* mutations, and 12% of them develop AML (Renneville et al. 2009).

Although the leukaemogenicity mechanisms of the biallelic *CEBPA* mutation are well understood, the significance of a monoallelic *CEBPA* mutation in MDS and AML remains unclear. A single *CEBPA* mutation likely results in a proliferative and developmentally aberrant “preleukaemic” haematopoietic stem cell that can progress to AML by acquiring subsequent mutations or epigenetic alterations (Mendoza et al. 2021). Fortunately, patients with a bZIP *CEBPA* mutation have a favourable prognosis, according to a study on *CEBPA* mutations in paediatric AML (Yu, Li, Zhang, et al. 2020).

Previous work in our lab, was aimed at investigating the mechanisms underlying the progression from low-risk to high-risk MDS. For these studies, isogenic iPSC lines from MDS longitudinal samples were generated from a patient who had suffered disease progression from low-risk MDS (harbouring somatic mutations in three genes) to high-risk MDS (by acquiring an additional mutation, a heterozygous *CEBPA* mutation in the mid region of the protein, disrupting the bZIP domain (performed by Ruba Almaghrabi, preprints work). The goal of the present research was to determine the effect of disrupting the C-terminal bZIP domain of

*CEBPA* on haematopoiesis under normal conditions. In this chapter, I described the generation of a novel *in vitro* model system (hiPSC) harbouring a heterozygous *CEBPA* mutation in the C-terminal domain (*CEBPA*<sup>+mut</sup>). Moreover, using various approaches, the pluripotency of the isolated clones was examined to ascertain the potential impact of CRISPR-Cas9 on pluripotency.



## 3.2 Results

### 3.2.1 Generation and characterisation of hiPSC clones with a *CEBPA* mutation (hiPSC- *CEBPA*<sup>+ /mut</sup>)

CRISPR-Cas9 technology was employed to investigate the impact of disrupting the *CEBPA* DNA-binding domain (DBD) on haematopoiesis. Cas9 was targeted to the mid region (before the bZIP domain) of *CEBPA* by searching *CEBPA* gRNA using an online tool (Trust Sanger Institute Editing Database). A specific gRNA was chosen because it was located in the desired region and had no hits with two or three mismatched nucleotides (Figure 3.1).

First, the gRNA primers were modified prior to their synthesis. In particular, the forward primer sequence was modified by adding CACC at the 3' end and AAAC before the guide's reverse complement primer. These modifications were necessary for cloning the gRNA into the pSpCas9(BB)-2A-GFP (PX458) vector (Addgene) at the BbsI restriction site (Ran et al. 2013).

The PX458 vector includes Cas9, the GFP reporter marker, and the U6 promoter to drive the gRNA expression. The gRNA primers were annealed, and the double-strand gRNA was cloned into the PX458 vector. The plasmid PX458 containing the *CEBPA* gRNA, or an empty vector were transfected into the hiPSCs (BU3.10) via Amaxa nucleofection. At 24 h post-nucleofection, the cells were sorted by FACS based on the GFP reporter gene expression. Approximately 250,000 cells were recovered and replated at low density to allow the isolation of single gene mutated clones. Three weeks post-nucleofection, eighteen clones had expanded, and the genomic DNA was extracted from nine clones. The gene-editing efficiency of CRISPR-Cas9 was evaluated by conducting a T7EI assay. In this assay, the T7 endonuclease can recognise and cleave non-perfectly matched DNA. Sanger sequencing was used to identify the type of mutation. A schematic of the complete procedure is summarised in Figure 3.2.

### WT *CEBPA* sequencing

```
TGTTCCAGCACAGCCGGCAGCAGGAGAAGGCCAAGGCGGCCGTGGGCCCCACGGGCGGGCGGGCGGGCGGGG
GGCGACTTTGACTACCCGGGCGCGCCCGGGCCCCGGCGGGCCCGTCATGCCCGGGGAGCGCACGG
GCCCCCGCCCGGCTACGGCTGCGCGGCCCGCCGGCTACCTGGACGGCAGGCTGGAGCCCCTGTACGAGC
GCGTCGGGGCGCCGGCGCTGCGGCCGCTGGTGATCAAGCAGGAGCCCCGCGAGGAGGATGAAGCCAAG
CAGCTGGCGCTGGCCGGCCTCTCCCTACCAGCCGCCGCCGCCGCCGCCCTCGCACCCGCACCCG
CACCCGCCGCCCGCGCACCTGGCCGCCCCGACCTGCAGTTCCAGATCGCGCACTGCGGCCAGACCACC
ATGCACCTGCAGCCCGGTACCCCCACGCCGCCGCCACGCCCGTGCCAGCCCGCACCCCGCGCCCGCG
CTCGGTGCCGCCGGCTGCCGGGCCCTGGCAGCGCGCTCAAGGGGCTGGGCGCCGCGCACCCCGACCTC
CGCGCGAGTGGCGGCAGCGGCCGCGGGCAAGGCCAAGAAGTCGGTGGACAAGAACAGCAACGAGTACC
GGGTGCGGGCGAGCGCAACAACATCGCGGTGCGCAAGAGCCGCGACAAGGCCAAGCAGCGCAACGT
GGAGACGCAGCAGAAGGTGCTGGAGCTGACCAGTGACAATGACCGCCTGCGCAAGCGGGTGAACAGC
TGAGCCGCGAACTGGACACGCTGCGGGGCATCTTCCGCCAGCTGCCAGAGAGCTCCTTGGTCAAGGCCA
TGGGCAACTGCGCGTGAGGCGCGCGGCTGTGGGACCGCCCTGGGCCAGCCTCCGGCGGGGACCCAGGG
AGTGGTTTGGGGTGCCTCGGATCTCGAGGCTTGCCCGAGCCGTGCGAGCCAGGACTAGGAGATTCCGGTG
CCTCCTGAAAGCCTGGCCTGCTCCGCGTGTCCCTCCCTCCTCTGCGCCGGAAGTGGTGGCTTAAGAT
GAGGGGGCCAGGCGGTGGCTTCTCCCTGCGAGGAGGGGAGAATTCTGGGGCTGAGCTGGGAGCCCCG
CAACTTAGTATTTAGGATAACCTTGTGCCTTGAAATGCAAACCTCACCGCT
```

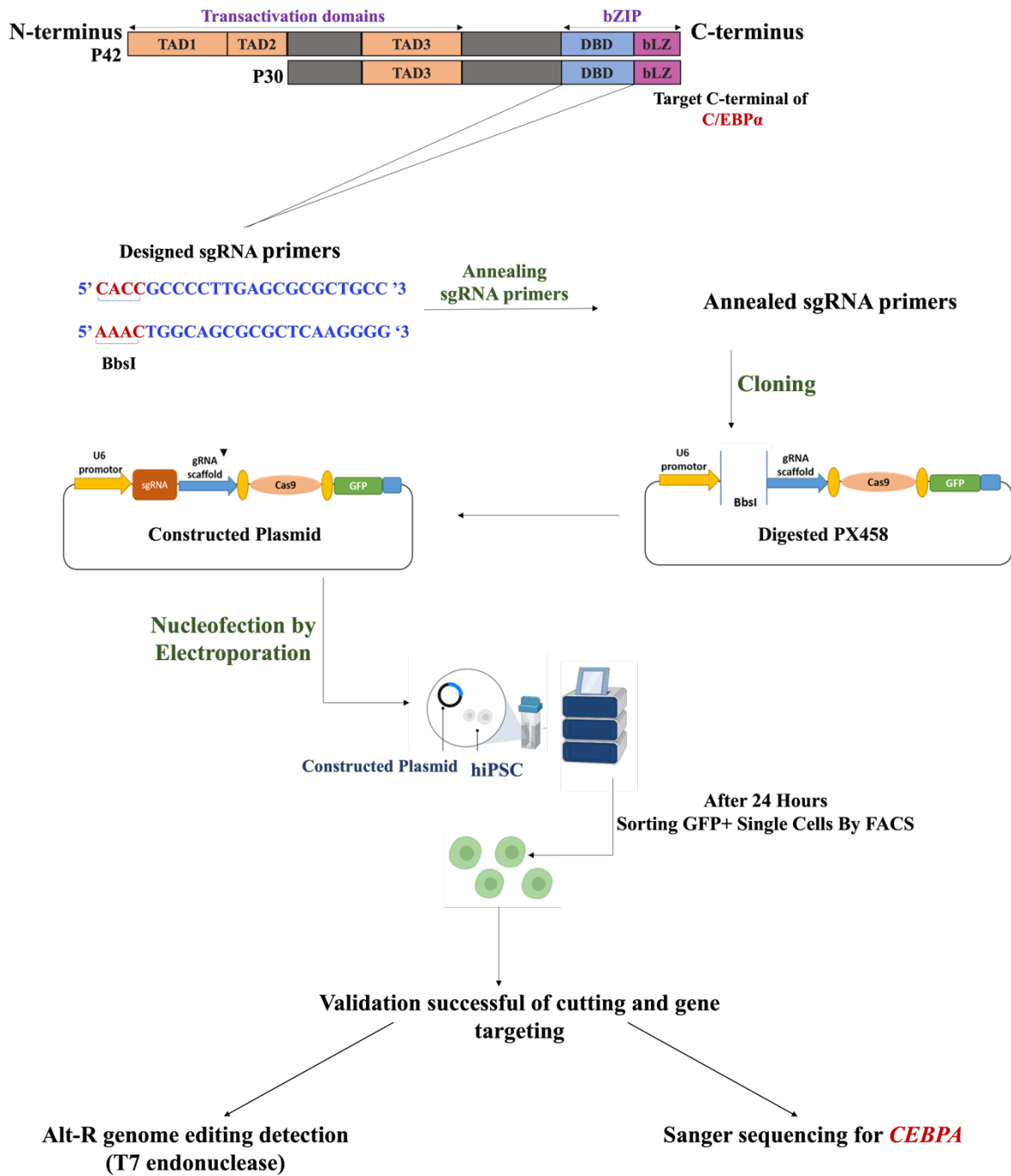
 Primers to amplify target region

 gRNA position

### Figure 3.1 *CEBPA* sequence.

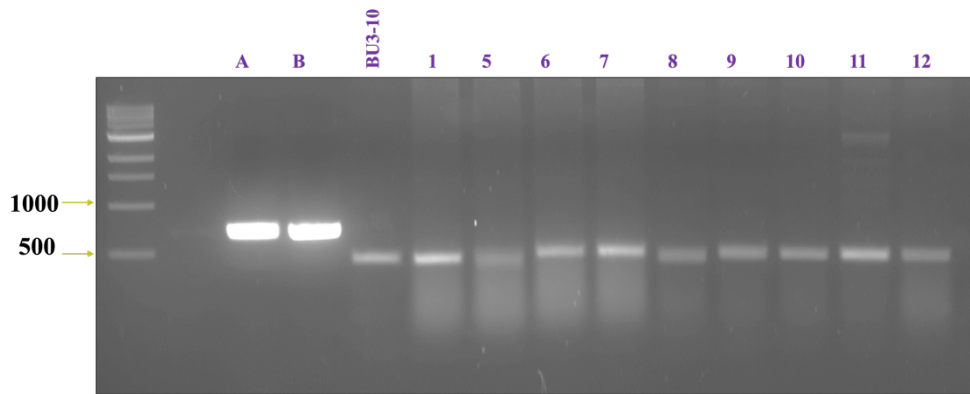
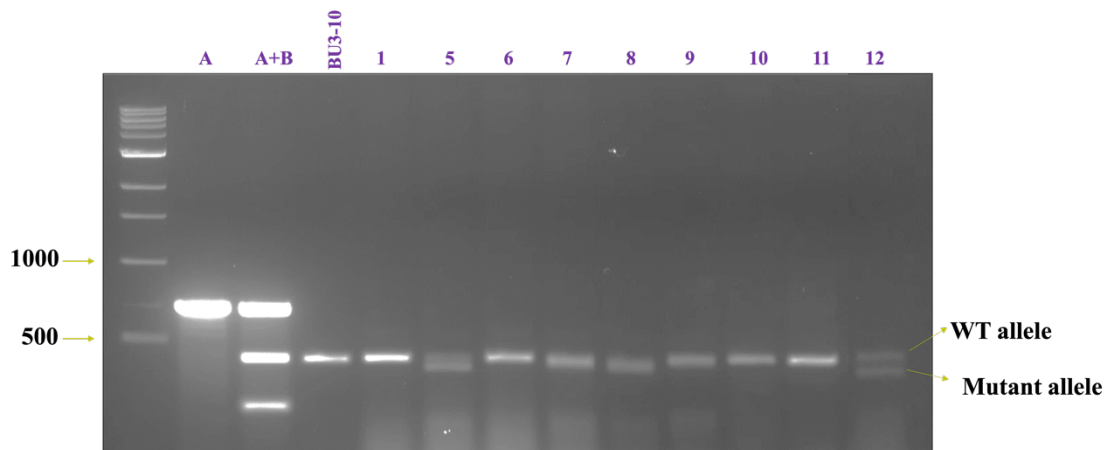
Snapshot of *CEBPA* sequence. The position of gRNA (green) and primers used to amplify the targeted region (yellow) are highlighted.

For the T7EI assay, PCR primers were first designed to amplify a region flanking the targeted site of *CEBPA* (see section 2.16). PCR amplification was successful in the clones, as indicated by the agarose gel results (Figure 3.3, A). The DNA from these selected clones, as well as the parental WT iPSC line (BU3-10), was then treated with T7EI enzyme, along with a homoduplex PCR product control (A) (for homozygous mutation) and a heteroduplex PCR product control (A+B) (for heterozygous mutation) to recognise and cleave non-perfectly matched DNA. The results were visualised on agarose gels. Two clones (C5 and 12) were identified that exhibited a heterozygous mutation, as indicated by the presence of two DNA fragments after T7 endonuclease reaction. The upper DNA fragment had the expected size (456 bp) based on an NCBI-published sequence, while the presence of a smaller lower DNA fragment indicated that a near mismatch had occurred in the amplified region due to CRISPR-Cas9 (Figure 3.3, B).



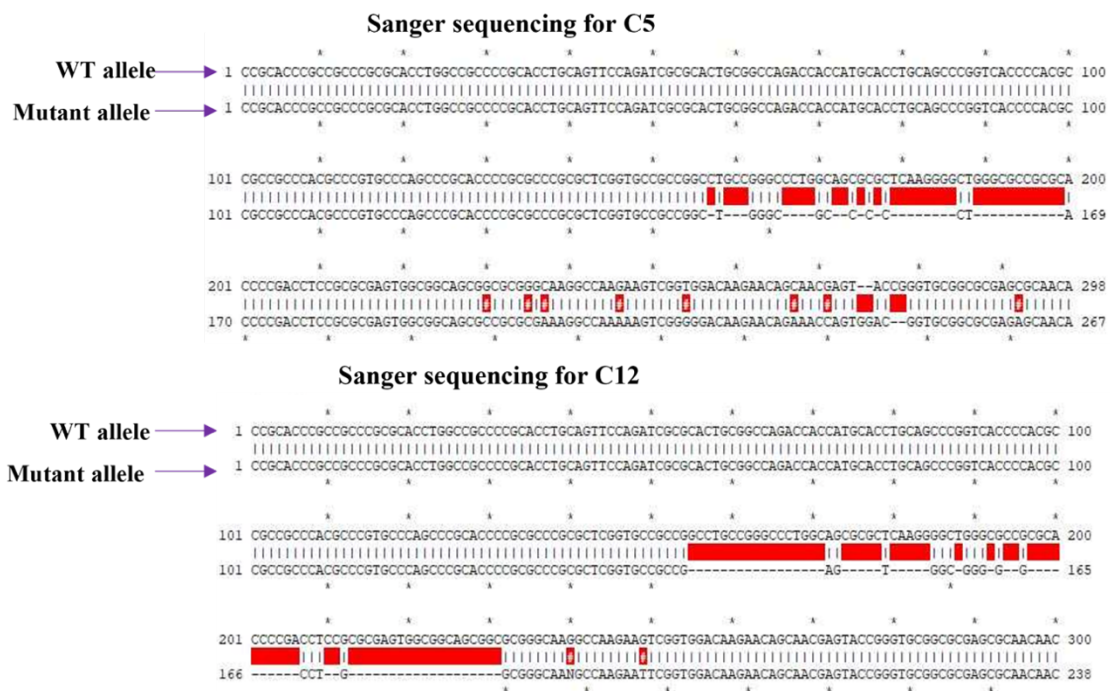
**Figure 3.2 Utilising CRISPR-Cas9 in hiPSCs.**

Schematic illustration of *CEBPA* mutation engineering in iPSCs using the PX458 Cas9 plasmid.

**A****Amplification of the target part of *CEBPA*****B****Mismatch cleavage assay with T7E1****Figure 3.3 T7E1 mismatch cleavage assay.**

(A) Following CRISPR editing, the gDNA was extracted from the hiPSC (BU310) and hiPSC- *CEBPA*<sup>+mut</sup> selected clones. PCR was performed to amplify DNA flanking the target region (456bp) using primers designed around the target area. The PCR results were visualised on an agarose gel (B) T7E1 was used to digest the PCR products from step A and digestion reactions were visualised on an agarose gel. Sample 1 contains Control A homoduplex PCR product and sample 2 contains Control A and B homoduplex and heteroduplex PCR products. The size of the WT allele (456) and mutant allele are indicated.

The T7EI assay only allowed us to establish which clones were heterozygous for the *CEBPA* mutation; however, ascertaining where the mutation had occurred and whether it would disrupt the DBD of the *CEBPA* were essential for understanding the gene function. Therefore, clones C5 and C12 were sent for Sanger sequencing to detect potential insertions, deletions, or mismatches introduced by CRISPR-Cas9 alteration in the target region. Sanger sequencing revealed that each of these clones had acquired deletions at the target region in one allele of the *CEBPA* gene (Figure 3.4). According to the sequencing data, hiPSC- *CEBPA*<sup>+/-mut</sup> C5 had a 31 bp deletion, while hiPSC- *CEBPA*<sup>+/-mut</sup> C12 had a 62 bp deletion in the required region. Protein sequencing of C/EBP $\alpha$  also indicated the presence of deletions in the desired region leading to a change in the open reading frame in *CEBPA*, suggesting that alteration in the protein sequence would modify the DNA binding domain of *CEBPA* and probably impair its functionality (Figure 3.5). Both clones were maintained and expanded for utilisation in subsequent experiments. In particular, four clones were used for each experiment: (i) a hiPSC control, used as a normal iPSC; (ii) hiPSC+PX458, used as a CRISPR-Cas9 control; (iii) hiPSC- *CEBPA*<sup>+/-mut</sup> C5, and (iiii) hiPSC- *CEBPA*<sup>+/-mut</sup> C12 (the two clones containing the monoallelic *CEBPA* mutation).



**Figure 3.4 Sanger sequencing.**

Sequence of the targeted region for the hiPSC- *CEBPA*<sup>+/mut</sup> clones C5 and C12. Sanger sequencing was performed by SourceBioscience using primers flanking the target region. The sequence was aligned using ApE-plasmid editor software.

**WT full C/EBP $\alpha$  protein sequence**

MESADFYEAEP RPPMSSHLQSPHPAPSSAAF GFPRGAGPAQPPAPPAAPEPLGGIC  
EHETSIDISAYIDPAAFNDEF LADLFQHSRQQEKAKAAVGPTGGGGGGDFDYPGA  
PAGPGGAVMPGGAHGPPPGYGCAAAGYLDGRLEPLYERVGAPALRPLVIKQEPR  
EED EAKQLALAGLFPYQPPPPPPSHPHPHPPPAHLAAPHLQFQIAHCGQTTMHLQ  
PGHPTPPPTPVPSHPAPALGA **AGLPGPGSALKGLGAAHPDLRASGGSGAGKAKK**  
**SVDKNSNEYRVRRRERNIAVRKSRDKAKQRNVETQQKVLELTSNDRLRKRVE**  
**QLSRELDTLRGIFRQLPESSLVKAMGNCA**

**C5 C/EBP $\alpha$  protein sequence**

PHPPPAHLAAPHLQFQIAHCGQTTMHLQPGHPTPPPTPVPSHPAPALGA  
**AGWAPPTPTSARVAAAPRRPKSRGTRTETSGRCGARATXARCARPTRA**  
**SNPT\*RRTRC\*N\*PEG**

**C12 C/EBP $\alpha$  protein sequence**

PHPPPAHLAAPHLQFQIAHCGQTTMHLQPGHPTPPPTPVPSHPAPALGA  
**AEWRGPGGQXQEFGGQEQQRVPGAARAQQHRGAQEPRQGQAAQRGDAPEG**  
**AGADQXRDKAXXRNXXXXXXXXXXLTR**



**WT Peptide**



**Mutant Peptide**

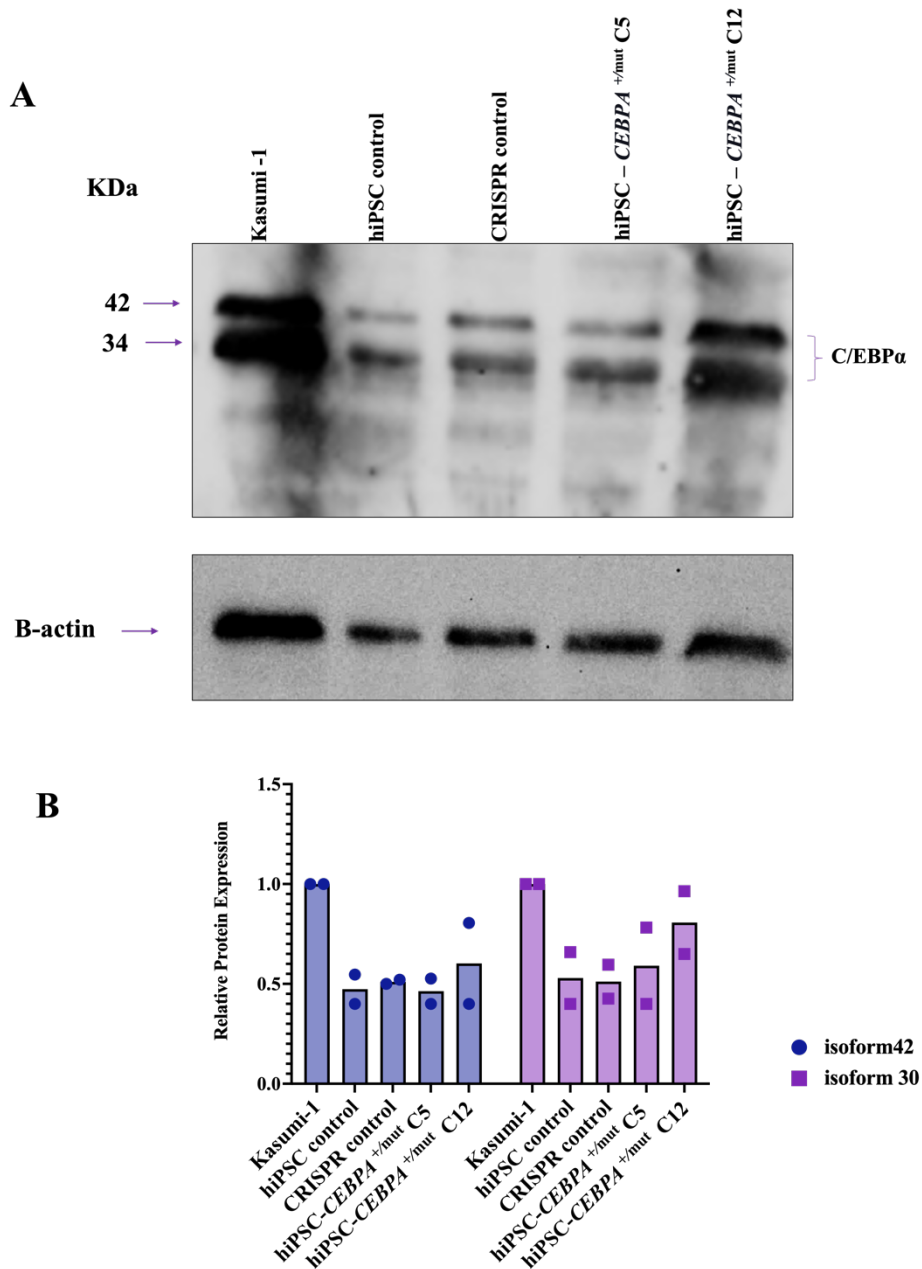
**Figure 3.5 Sequence of C/EBP $\alpha$  protein.**

Sequence for C/EBP $\alpha$  protein in WT hiPSC and the C5 and C12 clones. The deletion of the desired region by CRISPR-Cas9 led to a changing open reading frame (C5 and C12), subsequently changing the sequence of mutated clones. Protein sequencing was determined using the Basic Local Alignment Search Tool (BLAST).



### 3.2.2 hiPSC – *CEBPA*<sup>+/-mut</sup> clones exhibit no changes in C/EBP $\alpha$ protein expression

After identifying changes in the C/EBP $\alpha$  protein sequence from the hiPSC- *CEBPA*<sup>+/-mut</sup> clones C5 and C12, the next crucial step was to determine whether the CRISPR-induced deletion would impact the protein levels or cause an imbalance in the expression of the two isoforms of C/EBP $\alpha$ . Therefore, western blotting was used to compare C/EBP $\alpha$  protein expression in the hiPSC control, CRISPR control (hiPSC+PX458), hiPSC- *CEBPA*<sup>+/-mut</sup> C5 and hiPSC- *CEBPA*<sup>+/-mut</sup> C12 cells. The protein extract derived from Kasumi-1 cells, a human acute myeloid leukaemia cell line characterised by the presence of a translocation (8;21), exhibits the expression of both isoforms of C/EBP $\alpha$  (Larizza et al., 2005), was utilised as a positive control. No notable change was detected in C/EBP $\alpha$  protein expression between the WT hiPSC lines and the CRISPR control, as they expressed both isoforms (Figure 3.6A), with a slightly higher expression of the p30 over the p42 isoform (Figure 3.4B). Both isoforms of C/EBP $\alpha$  were also detected in the hiPSC- *CEBPA*<sup>+/-mut</sup> clones (Figure 3.6 A-B).

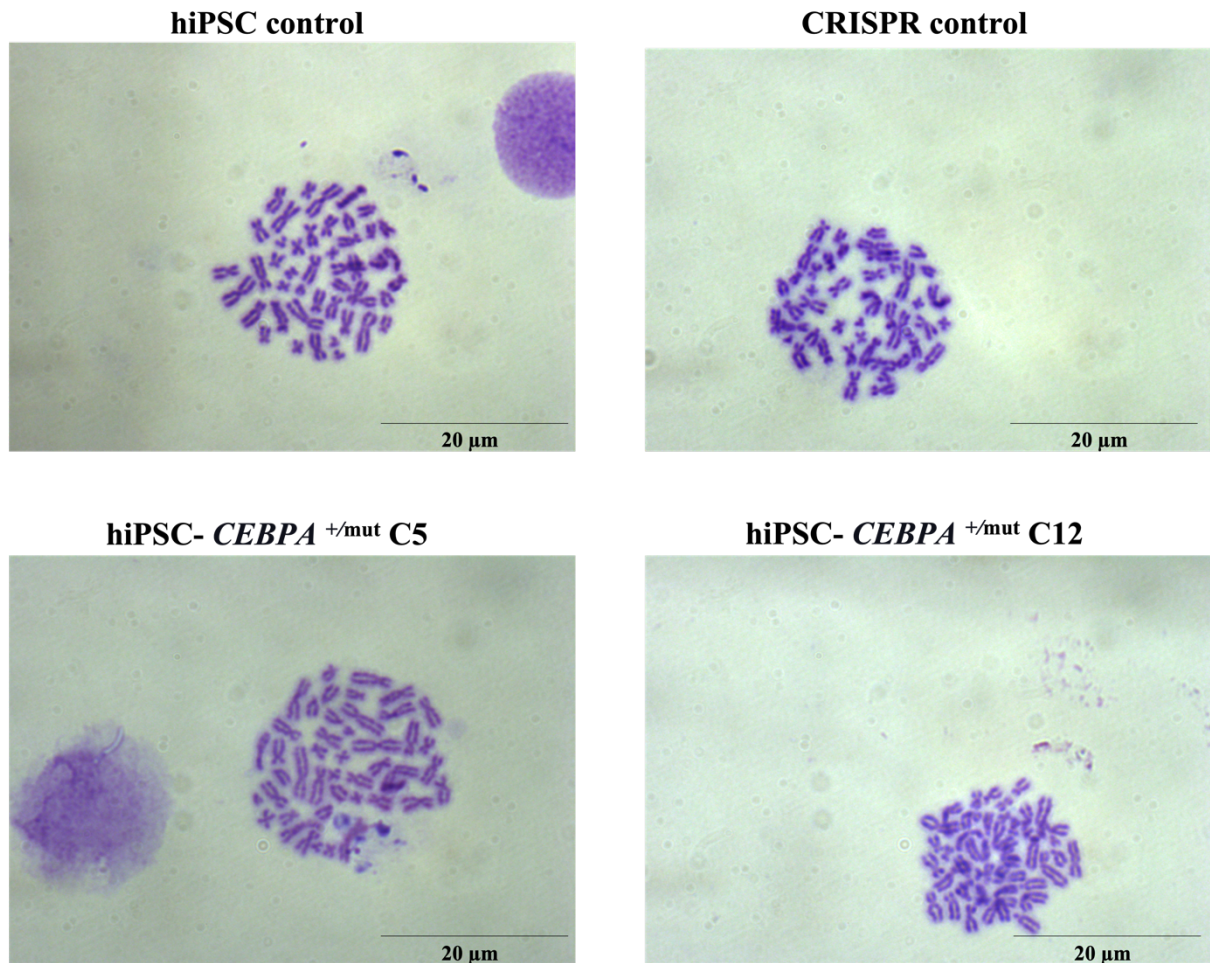


**Figure 3.6 Western blot analysis of C/EBP $\alpha$  protein.**

(A) Protein (50  $\mu$ g) from the Kasumi-1, hiPSC Control, CRISPR control, hiPSC- *CEBPA*<sup>+/mut</sup> C5 and hiPSC-*CEBPA*<sup>+/mut</sup> C12 was subjected to SDS-PAGE, and C/EBP $\alpha$  protein expression was detected using antibodies against C/EBP $\alpha$ . The left side of the blot show the relevant molecular weight standards, The arrows on the right side of the picture indicate distinct bands corresponding to each isoform. (B) The C/EBP $\alpha$  protein expression was quantified by normalising to  $\beta$ -actin and comparing the band intensity using ImageJ software. Data represent the means of 2 independent experiments.

### **3.2.3 hiPSC- *CEBPA*<sup>+mut</sup> clones do not show chromosomal instability**

We examined whether CRISPR-Cas9 affected chromosome stability due to chromosome fusions and whether the generation of new lines from single cells led to the growth of chromosomally unstable cells with proliferative advantage. Human cells have 46 chromosomes; therefore, a gain or loss in this number would indicate chromosome instability. Cells were cultured and treated with colcemid for 2 h, a drug that prevents spindle formation during mitosis, to block the cells at the metaphase stage (Stoyka et al. 1978; Sackett and Varma 1993). Chromosome preparations were then obtained, and the numbers of chromosomes on the metaphase spreads were counted in 20 to 30 metaphase cells per condition. The controls, hiPSC BU3.10 and CRISPR control cells were also analysed. In all cases, 46 chromosomes were counted (Figure 3.7). This result shows that employing CRISPR-Cas9 to disrupt DBD in *CEBPA* did not affect stability of chromosome in the hiPSC clones.

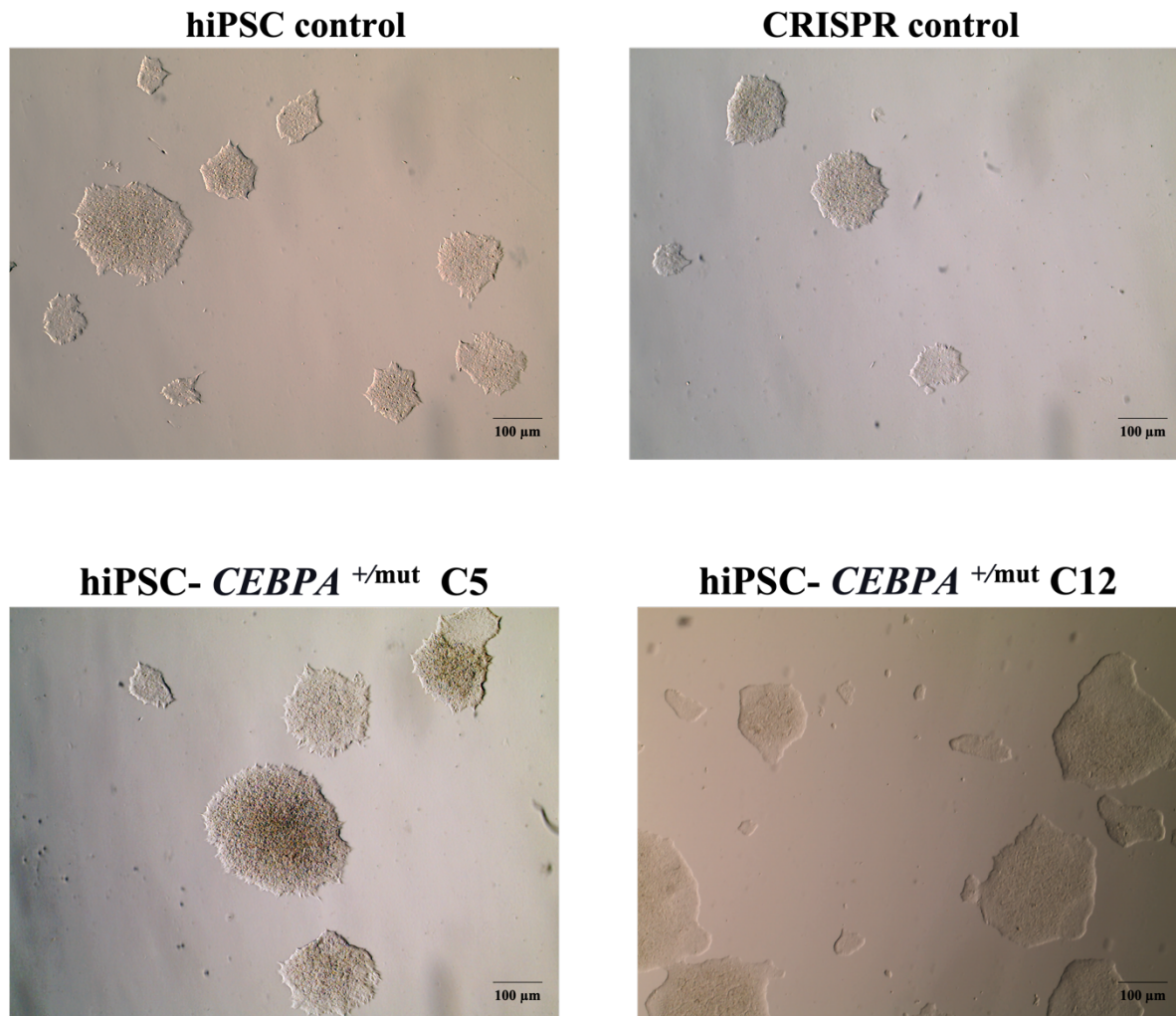


**Figure 3.7 Chromosomal spread analyses for hiPSC control, CRISPR control and hiPSC-*CEBPA*<sup>+/mut</sup> clones.**

Cells were treated for 2 h with 0.02 g/mL colcemid. Subsequently, they were subjected to swelling using a 0.075 M KCl solution. Following this, the cells were fixed using cold Carnoy's solution, placed onto humidified chilled glass slides, and stained with Giemsa-modified solution stain. The cells were captured at 100 × magnification using a Leica DM6000 light microscope. More than 25 metaphase cells per sample were examined per experiment. Scale bar=20 μm, representative images of n=3.

### 3.2.4 Pluripotency characterisation of hiPSC controls and hiPSC- *CEBPA*<sup>+/-mut</sup> clones

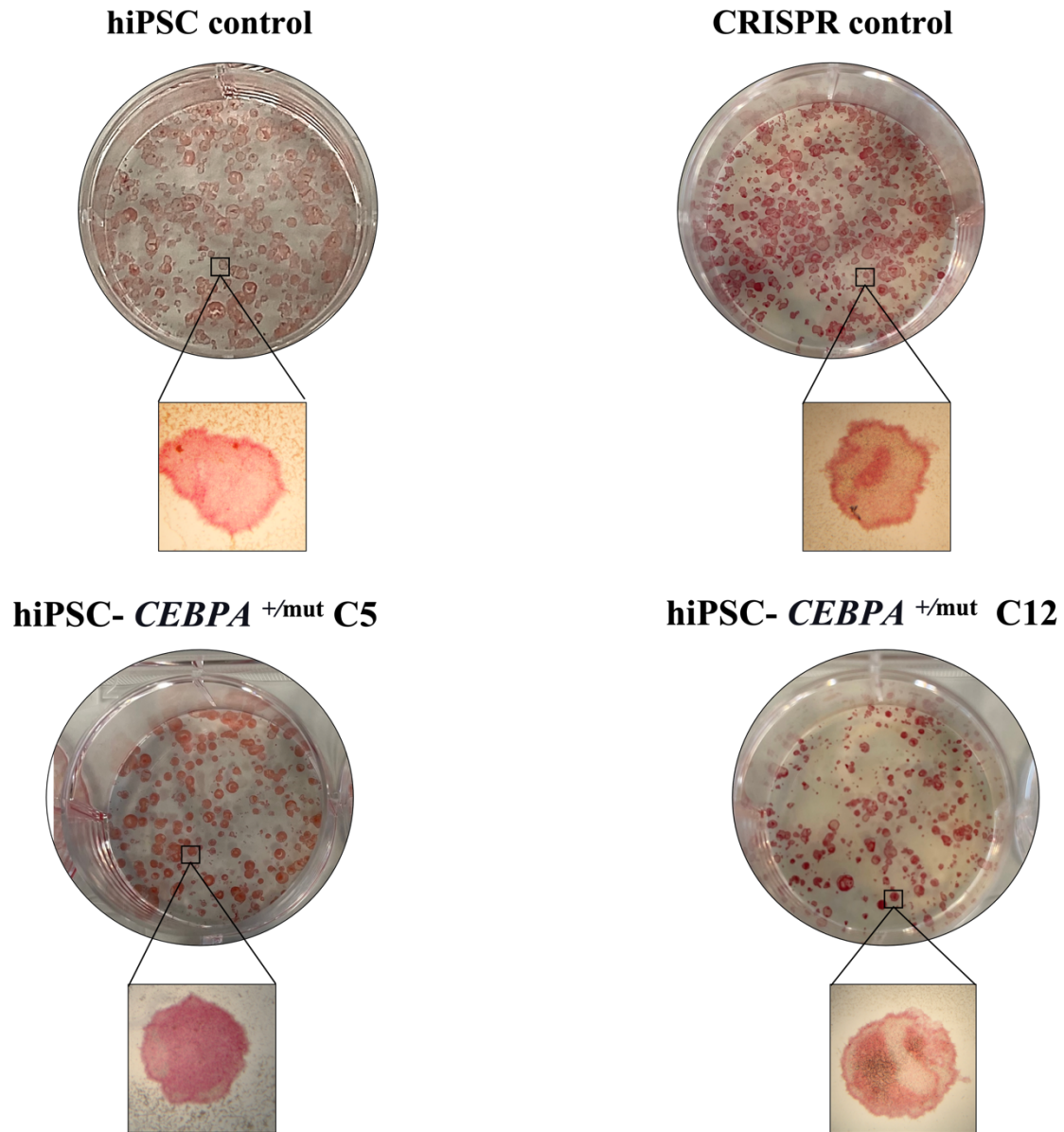
After determining that the four hiPSC lines exhibited a normal karyotype, next, we aimed to characterise the pluripotency potential of each clone. Daily phenotypic examination under phase contrast microscopy revealed that both hiPSC- *CEBPA*<sup>+/-mut</sup> clones and the CRISPR control cells had identical morphology similar to that of the hiPSC control parental line (BU3.10). They typically formed flat and tightly packed colonies, and the cells had scant cytoplasm and large nuclei (Figure 3.8) (Takahashi et al. 2007). The pluripotency of the CRISPR control (hiPSC+PX458), and hiPSC- *CEBPA*<sup>+/-mut</sup> clones was tested by (i) alkaline phosphatase staining; (ii) assessment of the expression of pluripotent markers; and (iii) determination of the capacity of the cells to differentiate into the three germ layers.



**Figure 3.8 Morphology of human-induced pluripotent stem cells (hiPSCs).**

The hiPSC control, CRISPR control, and hiPSC- *CEBPA*<sup>+/-mut</sup> (C5, C12) cells were seeded in 6-well plates coated with Corning Matrigel. Images were captured after 3 days in culture at 4× magnification using the EVOS cell imaging system (ThermoFisher Scientific). Scale bar = 100μm.

The hiPSC control and hiPSC- *CEBPA*<sup>+/-mut</sup> clones were grown in a 6-well plate coated with Matrigel until the cells reached 80% confluency. The cells were then stained to measure alkaline phosphatase activity, as increasing the AP enzyme activity level is considered a marker of pluripotent embryonic stem cells (Zhao et al. 2012). As shown in Figure 3.9, the CRISPR control (hiPSC+PX458), hiPSC- *CEBPA*<sup>+/-mut</sup> C5, and C12 colonies presented a pink-red cytoplasm, indicating strongly positive AP activity. This result suggested that the hiPSC lines generated by CRISPR-Cas9 were maintaining pluripotent features similar to those of the parental BU3.10 iPSC line.



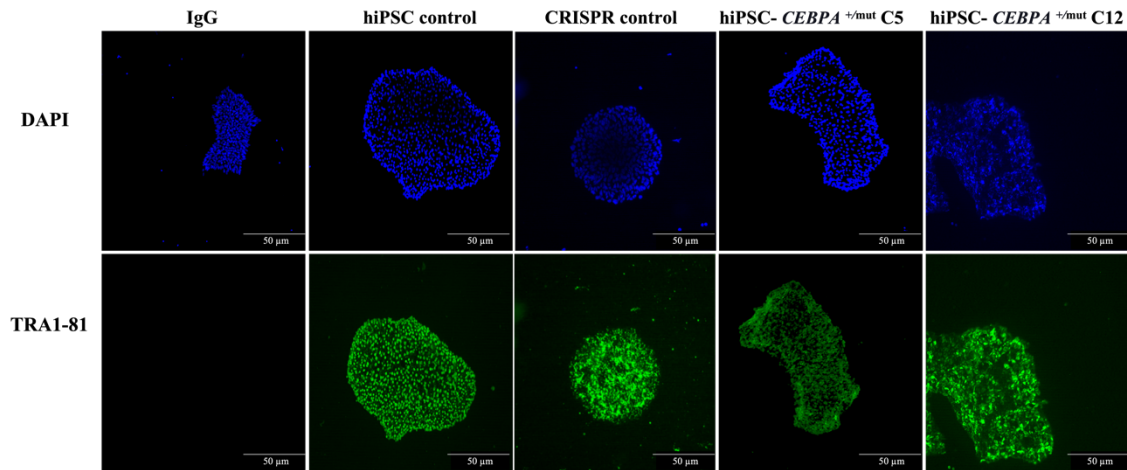
**Figure 3.9 The hiPSCs and hiPSC- *CEBPA*<sup>+/mut</sup> clones are positive for alkaline phosphatase.**

The hiPSC Control, CRISPR control and hiPSC- *CEBPA*<sup>+/mut</sup> clones (C5, C12) were grown on 6-well plates coated with Matrigel and stained for alkaline phosphatase activity for 40 min. The hiPSC clones positive for alkaline phosphatase activity are stained red in the images. The images were captured with a Primovert microscope (ZEISS) at a magnification of 10×, representative images of n=3.

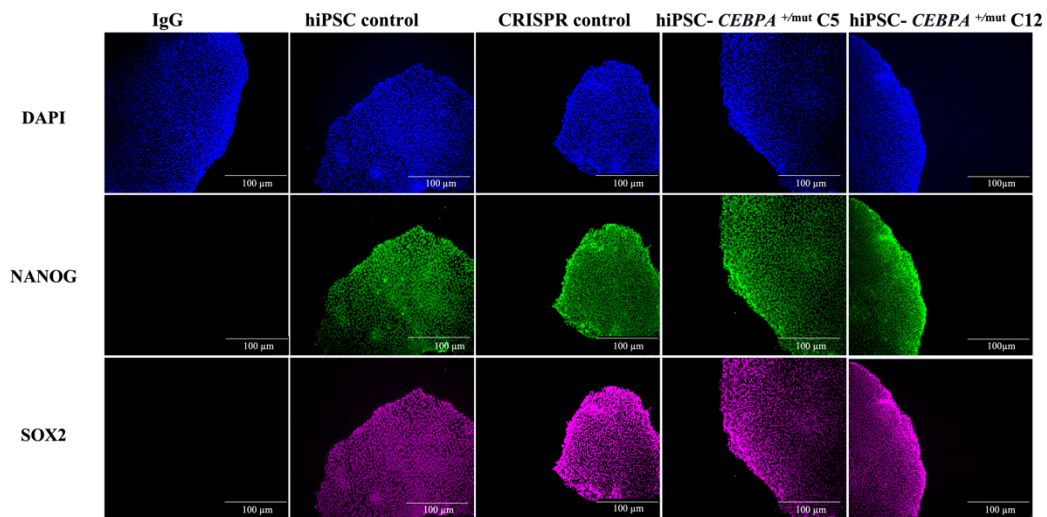


The expression of pluripotent protein markers was also analysed by culturing all hiPSC parental BU3.10, CRISPR control, and hiPSC- *CEBPA*<sup>+/-mut</sup> clones C5 and C12 on coverslips coated with Matrigel. Each clone was then stained with antibodies against the pluripotency surface marker TRA1-81 (Schopperle and DeWolf 2007; Trusler et al. 2018) or against the intracellular markers (nuclear) NANOG and SOX2, as these are essential for maintaining the pluripotent embryonic stem cell phenotype (Luo et al. 2013; Swain et al. 2020). All hiPSC- *CEBPA*<sup>+/-mut</sup> clones (C5, C12) and hiPSC control and hiPSC + PX458 were strongly positive for the surface marker TRA1-81 (Figure 3.10 A). The expressions of NANOG and SOX2 in CRISPR control and hiPSC- *CEBPA*<sup>+/-mut</sup> C5 and C12 were similar to the expression in the parental line, further validating that nucleofection of the CRISPR-Cas9 did not affect the pluripotency of the hiPSCs (Figure 3.10 B).

A



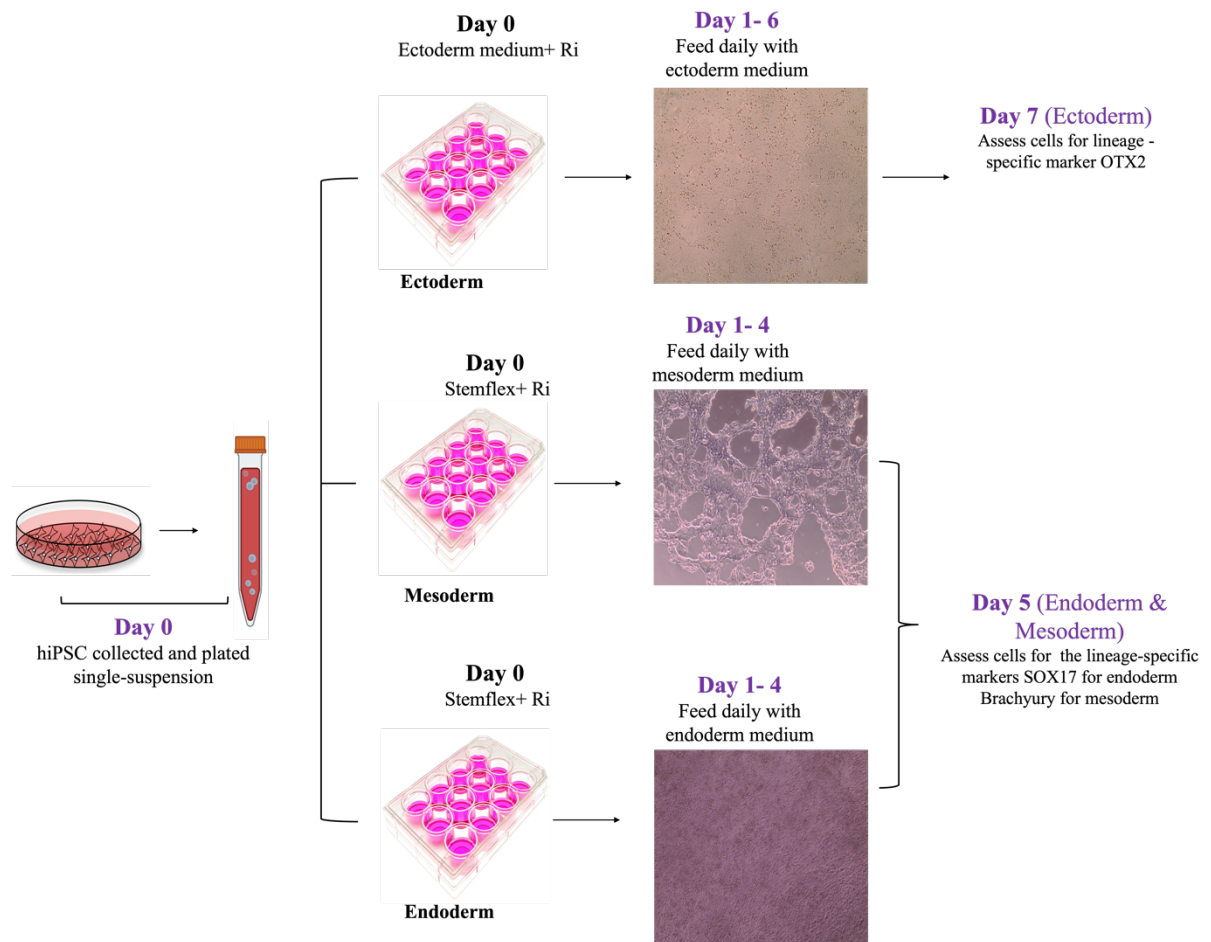
B



**Figure 3.10 The hiPSC and hiPSC - *CEBPA*<sup>+mut</sup> clones express pluripotent markers.**

(A) The hiPSCs control (BU3.10), CRISPR control, hiPSC- *CEBPA*<sup>+mut</sup> C5, and hiPSC- *CEBPA*<sup>+mut</sup> C12 clones were cultured on coverslips coated with Matrigel. (A) Cells showed positive immunostaining for Tra1-81 (green). The cells were stained with mouse Alexa Fluor 488 IgM, fixed, and counterstained with DAPI (blue). (B) Cells were fixed and permeabilised before incubation with NANOG (green) and SOX2 (purple) antibodies, followed by anti-goat Alexa Fluor 488 and anti-mouse Alexa 633, respectively, and counterstained with DAPI (blue). Scale bars = 50 μm and 100 μm, 20× magnification, Leica DM6000, representative images of n=3.

Finally, pluripotent stem cells were characterised by their potential to differentiate into the three germ layers: ectoderm, endoderm and mesoderm (Romito and Cobellis 2016). The Trilineage differentiation capacity of mutant lines was assessed to determine whether they were genuinely pluripotent stem cells. To do so, hiPSC- *CEBPA*<sup>+mut</sup> C5 and C12, besides the hiPSC control and CRISPR control, were induced to differentiate into three germ layers *in vitro* using the STEMdiff Trilineage differentiation kit (STEMCELL Technologies) After induction, lineage commitment was assessed by immunofluorescence staining and subsequent imaging using antibodies targeting markers associated with each germ layer (Figure 3.11).

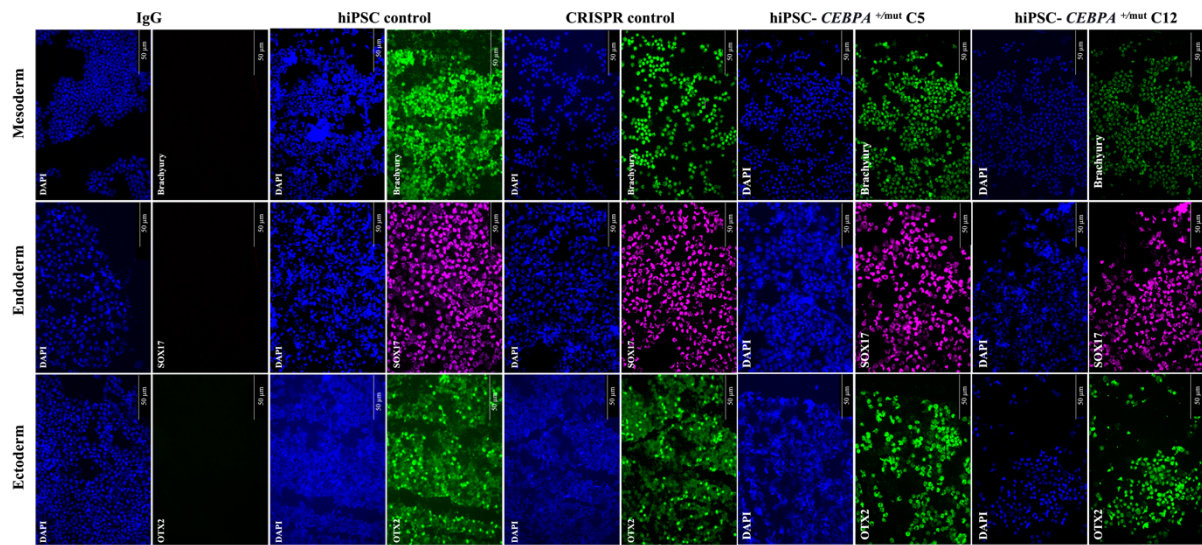


**Figure 3.11 Schematic illustration of trilineage differentiation assay.**

Schematic diagram showing the method used to differentiate hiPSC lines into the three germ layers using the STEMdiff Trilineage differentiation kit (STEMCELL Technologies). Ri=Rock inhibitor. Times are depicted in the figure.

As shown in Figure 3.12, the iPSC control, CRISPR control, and hiPSC- *CEBPA*<sup>+mut</sup> C5 and C12 expressed OTX2 on day 7 of ectoderm differentiation, indicating a commitment to an early ectodermal lineage (Gammill and Sive, 1997). Similarly, by day 5 of mesoderm and endoderm differentiation, all hiPSCs showed positive expression of brachyury (Chen et al. 2020) or SOX17 (Schroeder et al., 2012), respectively. Thus, these results indicate that the hiPSC generated could successfully differentiate into all three germ layers.

Overall, our data demonstrated the successful generation of hiPSCs harbouring a monoallelic mutation in the DBD of *CEBPA*, as these cells sustained the same karyotype and exhibited the following characteristics of pluripotent stem cells: (i) they are morphologically undifferentiated, (ii) they are positive for alkaline phosphatase, (iii) they express pluripotent markers, and (iv) they can differentiate into the three germ layers.



**Figure 3.12** The hiPSC and hiPSC- *CEBPA*<sup>+mut</sup> clones differentiate into the three germ layers.

Cells from each germ line were centrifuged at 300g for 7 min onto glass slides, and then fixed, permeabilised, and stained with antibodies specific to each lineage. Ectoderm marker: Otx2 (green), Endoderm marker: SOX17 (purple), Mesoderm marker: Brachyury (green). DAPI (blue) was used to stain nuclei. Scale bar=50μm, 20× magnification, Leica DM6000, representative images of n=3.

### 3.3 Discussion

#### 3.3.1 Successful generation of hiPSC lines harbouring a monoallelic *CEBPA* mutation disrupting the bZIP domain

Experiments involving genetic mutations, such as gene deletions and gene insertions, have provided growing evidence that transcription factors are essential to the execution of cell differentiation programmes. Gene disruption experiments have helped to establish the developmental role of various critical transcription factors in regulating haematopoiesis. One of these key transcription factors is C/EBP $\alpha$ .

*CEBPA* mutations have been reported in 5–14% of *de novo* AML cases, with the majority displaying two distinct mutations on different alleles and a favourable prognosis (Mannelli et al. 2017). Kato et al. (2011) suggested that, in most therapy-related AML or MDS cases, *CEBPA* mutations are present only in one allele and that AML progressed from MDS carrying a *CEBPA* mutation. *CEBPA* is one of the pivotal haematopoietic transcription factors, and mutations in this gene impede myeloid development. Most AMLs with *CEBPA* mutations have two mutations (*CEBPA* double-mut), commonly biallelic, whereas single heterozygous mutations (*CEBPA* single-mut) are less common. The relationship between single *CEBPA* mutation and MDS remains an open question. Therefore, the aim of this chapter was to induce a monoallelic mutation in *CEBPA* that would disrupt the C-terminal domain, a region responsible for dimerization and binding to DNA (DiNardo and Cortes 2016; Su et al. 2018). Since their discovery by Dr Shinya Yamanaka in 2006, iPSCs have become essential research platforms for studying different human diseases (Singh et al. 2015; Barak et al. 2022). The use of CRISPR-Cas9 technology to modify the DNA of mammalian cells is another major advance in biomedical research (You et al. 2019). Although various studies have been conducted to improve techniques for gene editing in iPSCs, such as TALENs and ZFNs, the discovery of CRISPR-Cas9 helped overcome several issues, such as excessive time and labour

consumption, low editing efficiency, and potential off-target effects (H.-X. Zhang et al. 2019). In the present study, we were able to obtain CRISPR-Cas9 edited iPSC cell lines harbouring a monoallelic mutation in the C-terminal region of *CEBPA*.

Detection of this mutation was accomplished by size separation of the products by agarose gel electrophoresis as a quick method for screening the C-terminal of *CEBPA* (van Waalwijk van Doorn-Khosrovani et al. 2003). We utilised T7 endonuclease to determine targeted clones because this is an inexpensive, simple, affordable and suitable method when only a few samples need to be screened. In addition, as the PCR product is rarely purified prior to treatment with T7 endonuclease, the data can be produced in a few hours (Sentmanat et al. 2018). By performing this assay, we were able to define two clones, C5 and C12, that show a DNA mismatch that leads to cleavage of the original band. However, this assay can only detect the presence these mismatches.

Therefore, to assess the status of this specific *CEBPA* mutation, we utilised the most general method routinely used either directly or after cloning: Sanger sequencing of the PCR data. The advantage of nucleotide sequencing is that it detects all forms of alterations, whether heterozygous, homozygous, point mutations, insertions or deletions (Li et al. 2020). The results of Sanger sequencing revealed the occurrence of a deletion of 31 bp of C5 and 62 bp of C12. Despite the larger deletion in C12 relative to C5, the differences in bp deletions did not significantly affect the outcomes of subsequent experimental analyses conducted on the clones, as their results were quite similar when compared to those of the control group. This may indicate that, regardless of the difference in the numbers of deleted nucleotides between these clones, targeting this part of the gene produces a similar phenotype (see Chapter 4). Our data also revealed that deletions in the desired region led to a change in the open reading frame of *CEBPA*.



Overall, applying CRISPR-Cas9 technology is a powerful, straightforward and cost-effective method of establishing new hiPSC clones carrying monoallelic *CEBPA* mutations. Our results revealed that successfully targeting the mid region of *CEBPA* in human iPSCs led to a heterozygous mutation in *CEBPA* affecting the bZIP domain.

### **3.3.2 A frameshift mutation in the mid region of one allele of *CEBPA* does not affect the protein level**

C/EBP $\alpha$  interacts with DNA via the basic-region LZ domain as a heterodimer or a homodimer to control gene expression. The utilisation of two distinct translation initiation sites leads to the formation of two isoforms of C/EBP $\alpha$  protein: (i) 42 kDa known as p42 and (ii) 30 kDa called p30 (Alan D. Friedman 2015). In AML patients, mutations in two hotspots areas of the *CEBPA* gene have been discovered. One type of mutation occurs in the C-terminal, where in-frame deletions or insertions in the fork region within the LZ or between the LZ and the basic region cause changes in the structure of the basic region, resulting in a disruption of the DNA binding, as well as a possible breakage of the helical phase of the bZIP domain. The other type is an N-terminal frame-shift mutation between the first and the second start codons that produce loss expression of p42 (Gombart et al. 2002; Heyes et al. 2021).

After introducing the in-frame mutation in C-terminal, we assessed the protein concentration of C/EBP $\alpha$  by western blotting. This revealed the expression of two isoforms of the C/EBP $\alpha$  in hiPSC- *CEBPA*<sup>+mut</sup> C5 and C12 and CRISPR control cells. Previous studies have reported that N-terminal frame-shift mutations lead to translation of a 30-kDa protein only (Leroy et al. 2005). The frame-shift mutations in AML also suppress the translation of the full-length p42 isoform, leaving solo expression of the shorter p30 without the 2 transactivation elements (Leroy et al. 2005). Conversely, C-terminal in-frame mutations in the basic zip region reduce DNA binding and dimerisation (Gombart et al. 2002; Asou et al. 2003; DiNardo and Cortes 2016; Su et al. 2018). However, although we found the same level of protein, we cannot discard

the possibility that the modification disrupted the protein function. Leroy and co-workers discovered that insertions /deletions in particular parts of the C/EBP $\alpha$  protein have no effect on protein synthesis but can impair C/EBP $\alpha$  protein functionality (Leroy et al. 2005). Similarly, another possibility is that the same situation arose in our study. The differences in protein sequences between WT control and hiPSC- *CEBPA*<sup>+mut</sup> C5 and C12, together with the results for the cellular assays conducted on these clones, could potentially verify whether the protein function has been impeded. However, confirming the possibility that we have disrupted the protein function would require further experiments, such as Chromatin immunoprecipitation followed by sequencing (ChIP-Seq).

### **3.3.3 hiPSC – *CEBPA*<sup>+mut</sup> revealed a normal karyotype**

The CRISPR-Cas9 technique is reported as likely to result in unintended chromosomal modifications that affect the genetic structure. Previous research by Papathanasiou et al. (2021) examined the effects on the karyotype of mouse embryos after CRISPR-Cas9 genome editing. They combined imaging and single-cell genome sequencing to examine embryos at the 8-cell stage over the first three stages of development. Their results revealed unwanted DNA damage during mitosis, appearing as micronuclei and chromosome bridges that can lead to significant structural changes in chromosomes and to whole chromosome loss. This indicates that using CRISPR-Cas9 can produce unwanted on-target side effects that have further undesired consequences in development.

Based on these studies, we used a basic karyotype technique to assess whether hiPSC control, CRISPR control, and hiPSC- *CEBPA*<sup>+mut</sup> C5 and C12 cells exhibited substantial chromosomal abnormalities (i.e., changes in the number of chromosomes). Chromosome counting indicated that none of the clones showed evidence of chromosomal instability following CRISPR-Cas9 editing. The number of chromosomes was determined using our in-house facility. Alternative commercial services are easily accessible for the detection of off-target mutations or the

identification of chromosomal alteration. For instance, G-binding and fluorescence *in situ* hybridisation is a commonly utilised cytogenetic technique for diagnostic applications and cytogenetic research. It can also be used following the application of the CRISPR-Cas9 method (Rayner et al. 2019).

The application of these techniques to determine chromosomal instability would have provided a more robust indication of the genome stability in our generated clones. Array Comparative Genomic Hybridisation is another highly effective molecular method for identifying karyotypic abnormalities. Its advantages over previous techniques are its higher resolution, accurate localisation, and improved sensitivity and specificity (Pinkel and Albertson 2005).

### **3.3.4 The generated hiPSC cells are pluripotent**

We were keen to determine if the CRISPR-Cas9 system influenced the pluripotency of the picked iPSCs clones (BU310+PX458 as CRISPR control, hiPSC- *CEBPA*<sup>+/-mut</sup> C5, and C12) because they were generated from single cells after sorting. The pluripotency was characterised using well-established techniques: the expression of pluripotent markers, alkaline phosphatase staining, and capability for trilineage differentiation (Castaño et al. 2017; Hsu et al. 2019).

The hiPSC colonies remained flat with a tightly packed morphology and produced compact colonies with prominent and well-defined edges composed of cells with large nuclei and scant cytoplasm (Takahashi et al. 2007; Robinton and Daley 2012; Wakao et al. 2012). The AP enzyme showed positive activity in the hiPSC- *CEBPA*<sup>+/-mut</sup> clones, providing further validation of the pluripotency of the cells, as its membrane-bound localisation and high activity are associated with pluripotent stem cell morphology and its expression is downregulated in committed differentiating cells (Martins et al. 2014; Štefková et al. 2015).

As previously mentioned in the methods section, our parental cell model (BU310) was generated using a highly effective and consistent reprogramming technique, STEMCCA lentivirus, and Yamanaka factors were removed by transient transfection of Cre-recombinase

(Sommer et al. 2010). Therefore, SOX2 expression (one of Yamanaka's factors) in the BU310 cell line is a valid way to determine pluripotency, together with expression of NANOG and TRA1-81.

The CRISPR control and hiPSC- *CEBPA*<sup>+mut</sup> clones expressed the pluripotent markers SOX2, NANOG, and TRA1-81. Significant expression of these genes has been reported in undifferentiated mouse and human ESCs/iPSCs and embryonic germ cells (Brafman et al. 2013). These genes also play a vital role in self-renewal and maintaining cells in a pluripotent state (Rodda et al. 2005; Fong et al. 2008; De Los Angeles et al. 2012; Abujarour et al. 2013). SOX2 and NANOG expression in iPSCs has been shown to facilitate iPSC pluripotency and restrict differentiation-related gene expression (Boyer et al., 2005; Pan and Thomson, 2007; Zhang, 2014). Therefore, not surprisingly, the hiPSC- *CEBPA*<sup>+mut</sup> clones were positive for both intracellular and extracellular pluripotent proteins (NANOG, SOX2 and TRA1-81).

Trilineage differentiation capacity (ectoderm, mesoderm and endoderm) of our hiPSCs further corroborated the pluripotency of the edited cells. Upon subjecting the iPSC towards ectoderm differentiation, all cells became positive for the expression of OTX2, a feature marker of the ectoderm layer (Gammill and Sive 2001; Mortensen et al. 2015). All cells also expressed SOX17 when differentiated into endoderm; this marker is necessary for generating definitive endoderm (Kanai-Azuma et al., 2002; D'Amour et al., 2005). Brachyury was expressed when the cells differentiated into mesoderm (Lam et al., 2014). The use of the STEMdiff Trilineage Differentiation Kit to promote cell differentiation into the three germ layers (endoderm, mesoderm and ectoderm) to evaluate the pluripotency of hiPSC has been reported previously (Mawaribuchi et al. 2019). Overall, this method is rapid, robust, and direct for detecting the potential ability of undifferentiated hiPSC to differentiate into the three germ layers (Sekine et al. 2020).

Taken together, our findings allow us to conclude that our data is in accordance with previous research showing that altering the genome with CRISPR-Cas9 does not affect the ability of iPSCs to remain pluripotent and differentiate into various cell types (Geng et al., 2020). This suggests that combining the use of iPSCs and CRISPR-Cas9 editing could help to generate very useful isogenic cell lines (De Masi et al. 2020). In the present study, the combined use of our parental hiPSC BU3.10 and CRISPR-Cas9 technology allowed the generation of isogenic cell lines that could help to understand the effects of the *CEBPA* monoallelic bZIP disruption on the phenotypic abnormalities that contribute to MDS.

**Chapter 4: Studying the contribution of monoallelic mutation of *CEBPA* to the myelodysplastic phenotype.**

## 4.1 Introduction

Chapter 3 verified the successful use of CRISPR-Cas9 for the generation of hiPSC clones harbouring monoallelic *CEBPA* mutations in the DBD. Several techniques were used to determine and validate the pluripotency of the new hiPSC clones generated after the mutations were induced. Although *CEBPA* double mutations have been extensively studied, and several molecular mechanisms have been published (Green et al. 2010; Ahn et al. 2016; Mannelli et al. 2017; Tien et al. 2018), single mutations remain to be explored. With the new experimental system created here, we were eager to clarify the relationship between *CEBPA* mutations disrupting the C-terminal DBD and the phenotype that features MDS. Consequently, here in Chapter 4, our aim was to collect data that would assist in filling these gaps by attaining the following objectives:

- To study the capacity of hiPSC- *CEBPA*<sup>+/*mut*</sup> to differentiate towards HSPC.
- To utilise colony assays to define the differentiation potential of HSPCs derived from hiPSC- *CEBPA*<sup>+/*mut*</sup> clones.
- To understand the effect of hiPSC- *CEBPA*<sup>+/*mut*</sup> on the proliferation and self-renewal capacity of HSPCs.
- To use flow cytometry and morphological analysis to determine the erythroid and myeloid lineage differentiation of HSPCs generated from hiPSC- *CEBPA*<sup>+/*mut*</sup> in liquid culture.

## 4.2 Results

### 4.2.1 Successive emergence of haematopoietic stem/progenitor cell (HSPC) populations from hiPSC control and hiPSC- *CEBPA*<sup>+mut</sup>

The successful generation of hiPSC harbouring a monoallelic mutation in *CEBPA* disrupting the bZIP domain required an examination of whether the *CEBPA*<sup>+mut</sup> might impact the ability of iPSCs to differentiate into HSPCs. We therefore examined the differentiation of the hiPSC control, hiPSC+PX458 and hiPSC- *CEBPA*<sup>+mut</sup> clones (C5 and C12) into HSPCs using the STEMdiff Haematopoietic kit from STEMCELL Technologies (Figure 4.1). After the hiPSCs reached 80% confluence, 16 to 20 chunks of hiPSC fragments with a specific size of roughly 100  $\mu\text{m}$  were seeded; this chunk size was essential for survival and efficient differentiation. The first stage of differentiation was achieved by growing the cells in Medium A, which contains VEGFA, BMP4 and bFGF and promotes mesoderm differentiation. Medium B (containing SCF, VEGFA, TPO, BMP4, FLT3, and bFGF) was then utilised to induce haematopoietic cells. Induction of mesoderm layer, presence of hemogenic endothelium and the emergence of single floating cells indicated successful differentiation into haematopoietic cells. The differentiation potential of the hiPSC control and mutant clones was assessed by evaluating the expression of HSPC immunophenotypic markers using flow cytometry on different days of haematopoietic differentiation (days 7, 10, 12, and 14). The following markers were assessed to ascertain the dynamics of haematopoietic differentiation: CD34 (a haemato-endothelial marker), CD43 (an early haematopoietic marker that remains in differentiating precursor cells) (Vodyanik et al., 2006) and CD45 (the significant feature of haematopoietic cells) (McKinney-Freeman et al. 2009). The approach used for gating is depicted in Supplemental Figure 1.



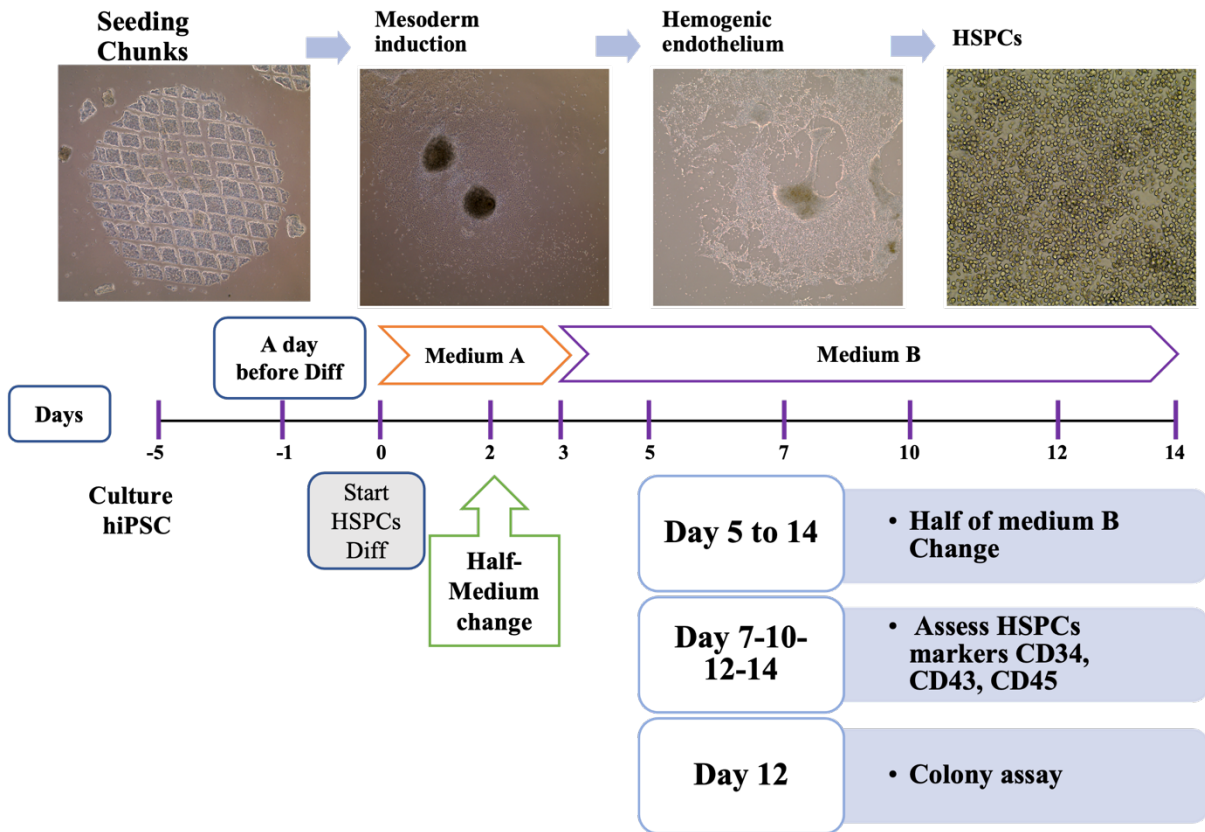
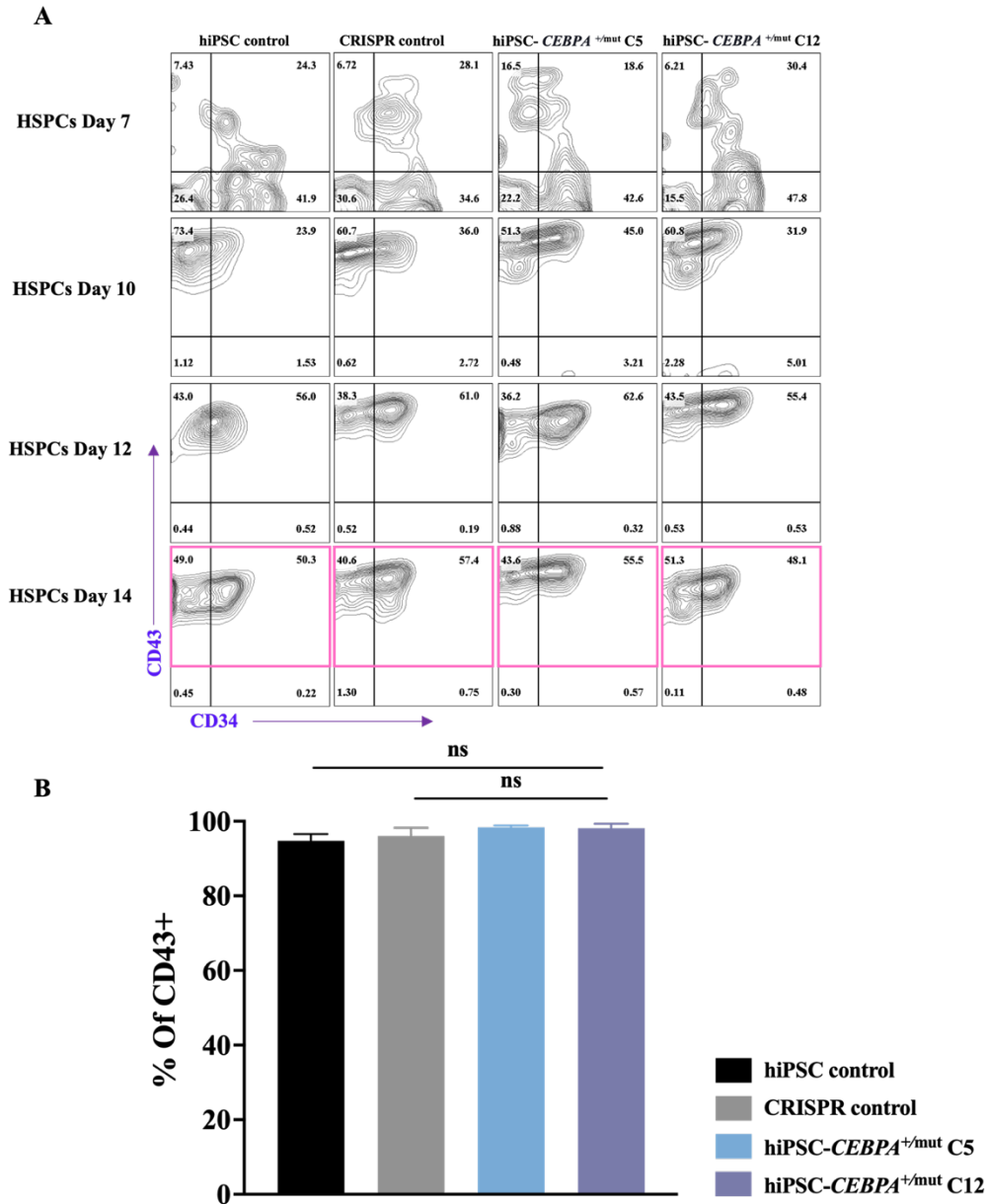


Figure 4.1 Schematic representation and timeline of HSPC differentiation using the haematopoietic STEMdiff protocol from STEMCELL Technologies.

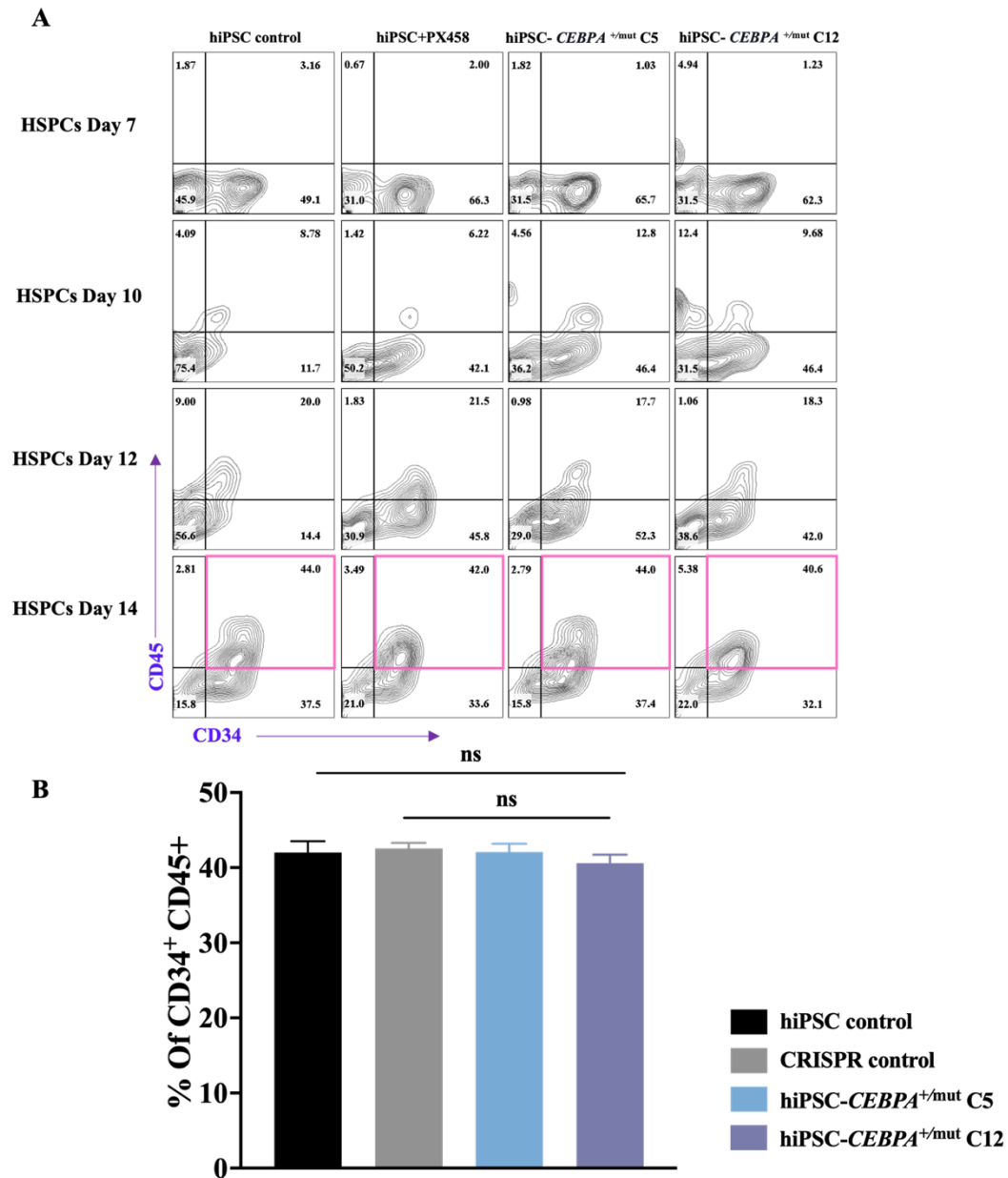
Both hiPSC control with hiPSC- *CEBPA*<sup>+ /mut</sup> clones performed similarly during the time course of differentiation. On day 7, the population of haematopoietic cells expressing CD43 was approximately 30–40% under all conditions, and this was followed by an increase at days 10 and 12 of differentiation, with 90% of the cells becoming CD43<sup>+</sup> for the hiPSC control, hiPSC+PX458, C5 and C12 clones. This percentage remained the same at day 14 (Figure 4.2, A). Statistical analysis showed no significant differences in the proportion of CD43<sup>+</sup> cells obtained from the hiPSCs studied, irrespective of their genotypes (Figure 4.2, B).

The percentage of cells co-expressing CD34<sup>+</sup> and CD45<sup>+</sup> was minimal on day 7 (approximately 1–3%). Subsequently, the percentage of cells co-expressing CD34<sup>+</sup> and CD45<sup>+</sup> gradually increased during differentiation to reach 6–13% on day 10, 17–20% on day 12 and approximately 40% by day 14 for all hiPSC lines. Statistical analysis confirmed no significant difference between the generated hiPSCs (parental and CRISPR control) and both hiPSC- *CEBPA*<sup>+ /mut</sup> clones, suggesting that the cloned cells retained their ability to differentiate into HSPCs. The transient transfection of the empty PX458 vector had no effect on the cells' differentiation potential. (Figure 4.3 A, B).

In summary, our data demonstrate that the STEMdiff™ HSPC differentiation protocol efficiently differentiated the hiPSC control, CRISPR control into primitive and definitive haematopoietic cells. In addition, these findings suggest that creation of the monoallelic *CEBPA* mutation did not affect the production of early and late HSPCs (CD34<sup>+</sup> CD43<sup>+</sup> and CD34<sup>+</sup> CD45<sup>+</sup>).



**Figure 4.2 HSPC derived from *CEBPA*<sup>+mut</sup> strongly expressed CD43<sup>+</sup>.** (A) Contour plot showing the ability of cells derived from hiPSC-*CEBPA*<sup>+mut</sup> to differentiate into early HSPC. Phenotypic markers were assessed over multiple days of differentiation using flow cytometry. (B) Bar graph depicting the mean percentage of (CD43<sup>+</sup>) on day 14. The results are displayed as mean ± standard error of the mean (SEM) and were evaluated using one-way ANOVA with multiple comparisons. *P*-values are shown as ns = not significant. Data represent three independent experiments.

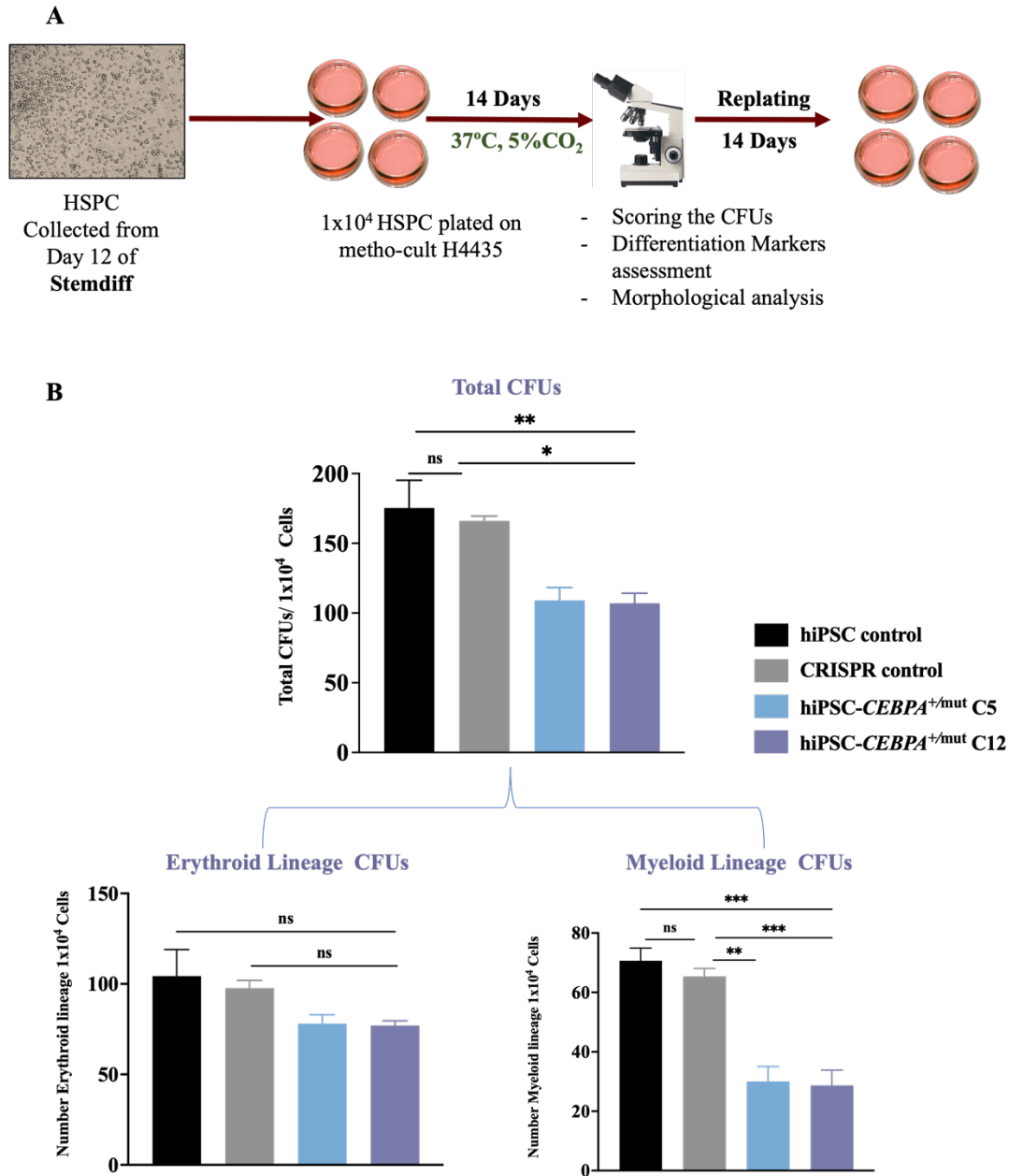


**Figure 4.3 The hiPSC-*CEBPA*<sup>+mut</sup> can generate definitive HSPC (CD34<sup>+</sup> CD45<sup>+</sup>).**  
**(A)** Contour plot showing the ability of cells derived from hiPSC-*CEBPA*<sup>+mut</sup> to differentiate into late HSPC markers. Phenotypic markers were assessed over multiple days of differentiation by flow cytometry.  
**(B)** Bar graph depicting the mean percentage of (CD34<sup>+</sup> / CD45<sup>+</sup>) on day 14. The results are displayed as mean ± standard error of the mean (SEM) and were evaluated using one-way ANOVA with multiple comparisons. *P*-values are shown as ns = not significant. Data represent three independent experiments.

#### **4.2.2 Evaluation of the differentiation potential of HSPC generated from hiPSC control and hiPSC – *CEBPA*<sup>+/-mut</sup> cells by colony assays in a semi-solid medium**

After confirming that the hiPSC- *CEBPA*<sup>+/-mut</sup> clones could differentiate into HSPCs, we examined whether the monoallelic *CEBPA* mutation affected the HSPC differentiation potential. For this, hiPSC control, hiPSC+ PX458 and hiPSC- *CEBPA*<sup>+/-mut</sup> clones were differentiated with STEMdiff, and at day 12 of differentiation,  $1 \times 10^4$  HSPCs were plated in methylcellulose based medium supplemented with recombinant cytokines that support myeloid haematopoietic differentiation. This methylcellulose formulation stimulates differentiation into mixed lineage containing granulocytes, erythrocytes, macrophages, megakaryocytes (CFU-GEMM), myeloid lineage progenitors (CFU-GM, CFU-M, CFU-G) and erythroid (BFU-E). After 14 days of culture, the types and the number of colonies were counted according to their phenotypic features (Figure 4.4 A).

Across the three independent experiments, the whole number of colonies was 30% lower for HSPCs generated from hiPSC- *CEBPA*<sup>+/-mut</sup> (C5 and C12) cells than from hiPSC parental control and CRISPR control cells. HSPCs from hiPSC control and CRISPR control cells were able to produce myeloid and erythroid CFUs, while the clones harbouring the *CEBPA* mutation displayed a significant decrease in their clonogenic capacity, showing a 50% reduction compared to controls. This reduction, as expected, affects mainly the myeloid lineage. Our result obtained using our new model supports previous studies stating that *CEBPA* is the primary regulator of granulopoiesis (Figure 4.4 B).

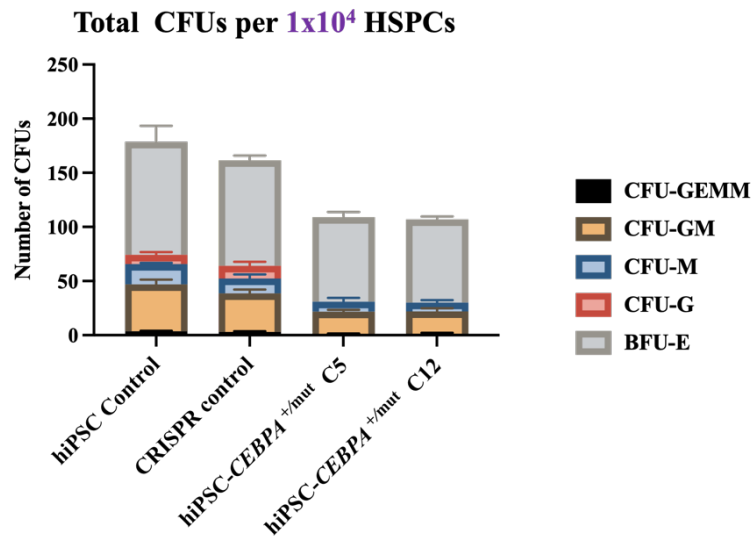


**Figure 4.4 Haematopoietic differentiation potential of HSPC-derived iPSCs in semi-solid medium.**

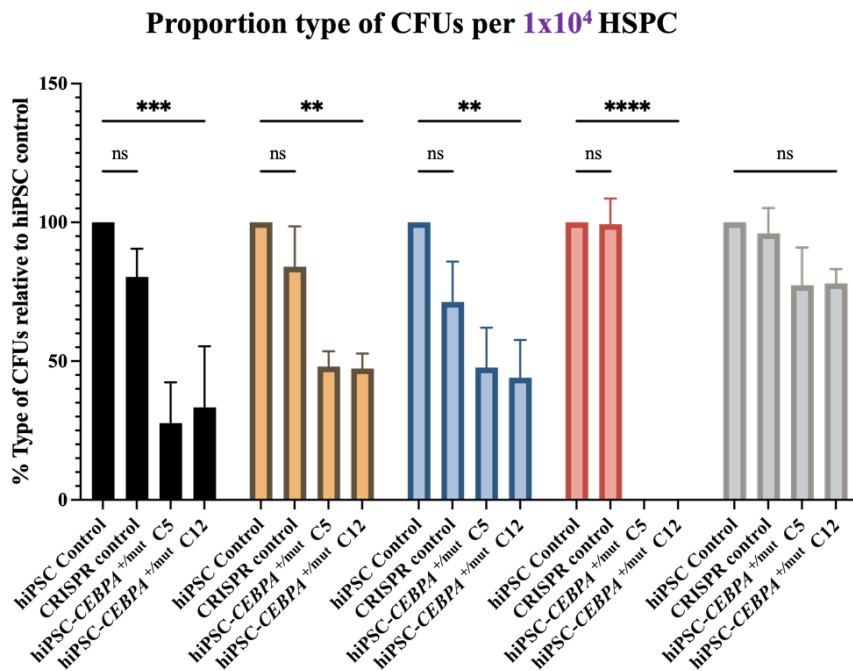
(A) Schematic depiction of haematopoietic potential of day 12 differentiating cells evaluated using a colony-forming assay. (B) After 14 days in a methylcellulose-based medium, the total number of CFUs was scored. The total number of CFUs was then used to identify the erythroid and myeloid lineages of CFUs. Statistical results are presented as the mean  $\pm$  SEM and were analysed using one-way ANOVA with multiple comparisons. *P*-values for the total CFU number of hiPSC- *CEBPA*<sup>+mut</sup> compared with hiPSC control were \*\* 0.0098 (C5) and 0.0083 (C12), and CRISPR control was \* 0.0217 (C5) and 0.0182 (C12), respectively. *P*-values for myeloid lineage CFUs of hiPSC- *CEBPA*<sup>+mut</sup> clones compared with hiPSC controls were \*\*\* 0.0005 (C5) and 0.0004 (C12), and CRISPR control was \*\* 0.0012(C5) and \*\*\* 0.0010 (C12). The *P*-value for total CFUs erythroid lineage derived from hiPSC- *CEBPA*<sup>+mut</sup> was not significant when compared with the hiPSC controls (ns = not significant). The data represent three independent experiments.

Scoring the different types of CFUs from each clone revealed that the mutant clones had lost their potential to differentiate into CFU-G and were now only capable of differentiating into CFU-M and erythroid colonies (BFU-E) (Figure 4.5 A). In addition, the proportion of each colony type was lower in the mutant clones than in the hiPSC control and CRISPR control clones; however, no significant change was observed for BFU-E (Figure 4.5, B). Despite the reduction in colony formation in the mutant HSPC clones, no discernible variations were observed in the morphology and size of the colonies when compared to the colonies derived from control HSPCs (Figure 4.6).

A



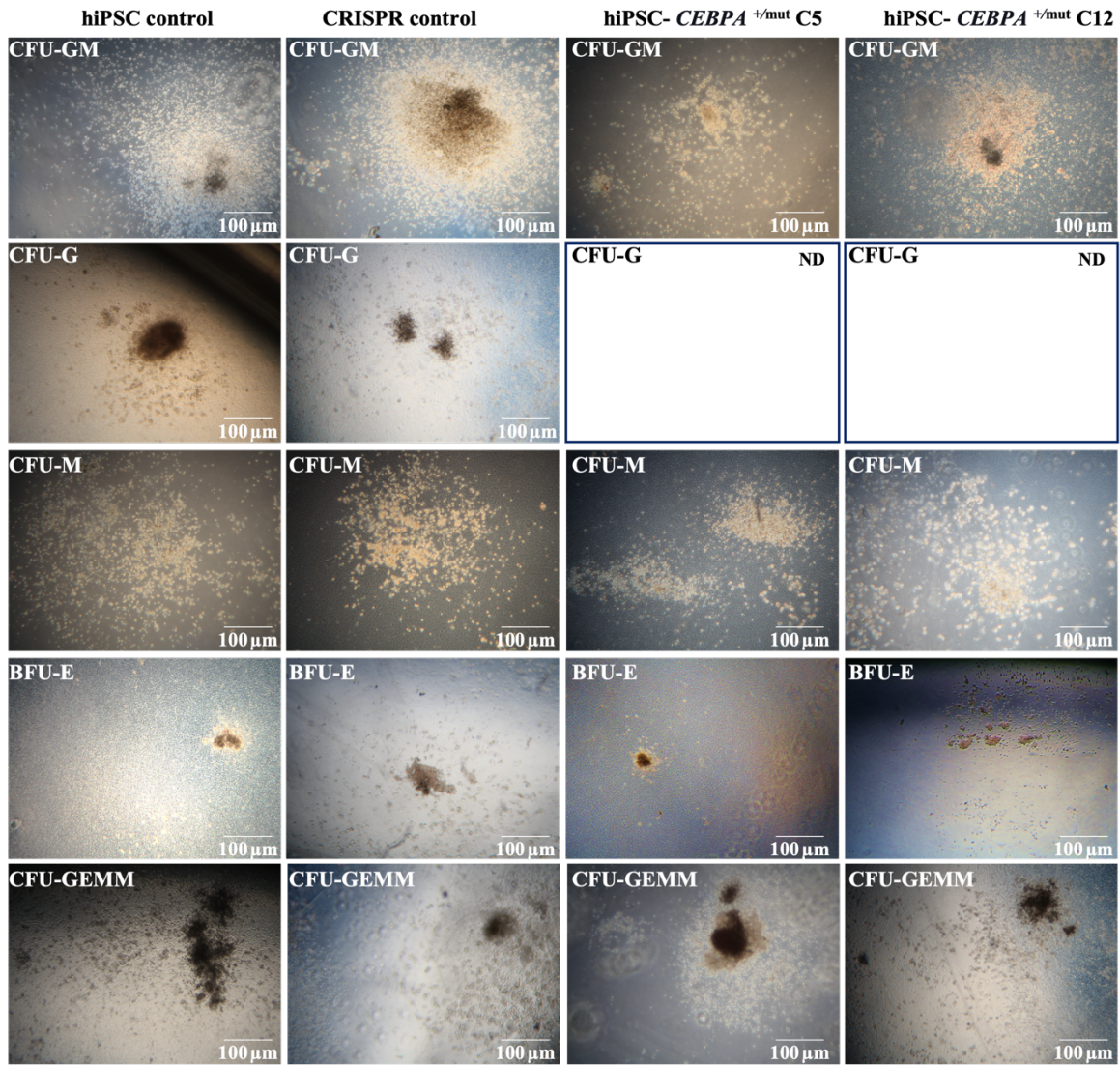
B



**Figure 4.5 Type of colony-forming units in HSPC derived from hiPSC *CEBPA* $^{\pm}/mut$  clones.**

(A) The histogram represents the mean number of CFUs per type scored after seeding  $1 \times 10^4$  cells and culturing for 14 days on semi-solid medium. Granulocytes, macrophages, erythrocytes, and megakaryocytes are present in the CFU-GEMM colony. A CFU-GM colony contains granulocytes and macrophages, whereas a CFU-M colony only contains macrophages, granulocytic colony-forming units (CFU-G); and burst-forming units-erythroid (BFU-E), colony contains erythrocytes. (B) Histogram shows the proportion of each type of CFU generated from  $1 \times 10^4$  HSPCs after 14 days of culture on semi-solid medium. Statistical results are presented as mean  $\pm$  SEM and analysed using two-way ANOVA with multiple comparisons; the *p*-values for mutant cells compared to hiPSC were \*\*\* GEMM were 0.003 and (GM) were \*\* 0.0047 (C5) and 0.0042 (C12). For (CFU-M), the *p*-values were \*\* 0.0044 (C5), 0.0023 (C12), and < 0.0001 (CFU-G). Data represent three independent experiments.





**Figure 4.6 Morphology of colony-forming units in HSPCs derived from hiPSC *CEBPA*<sup>+mut</sup> clones.**

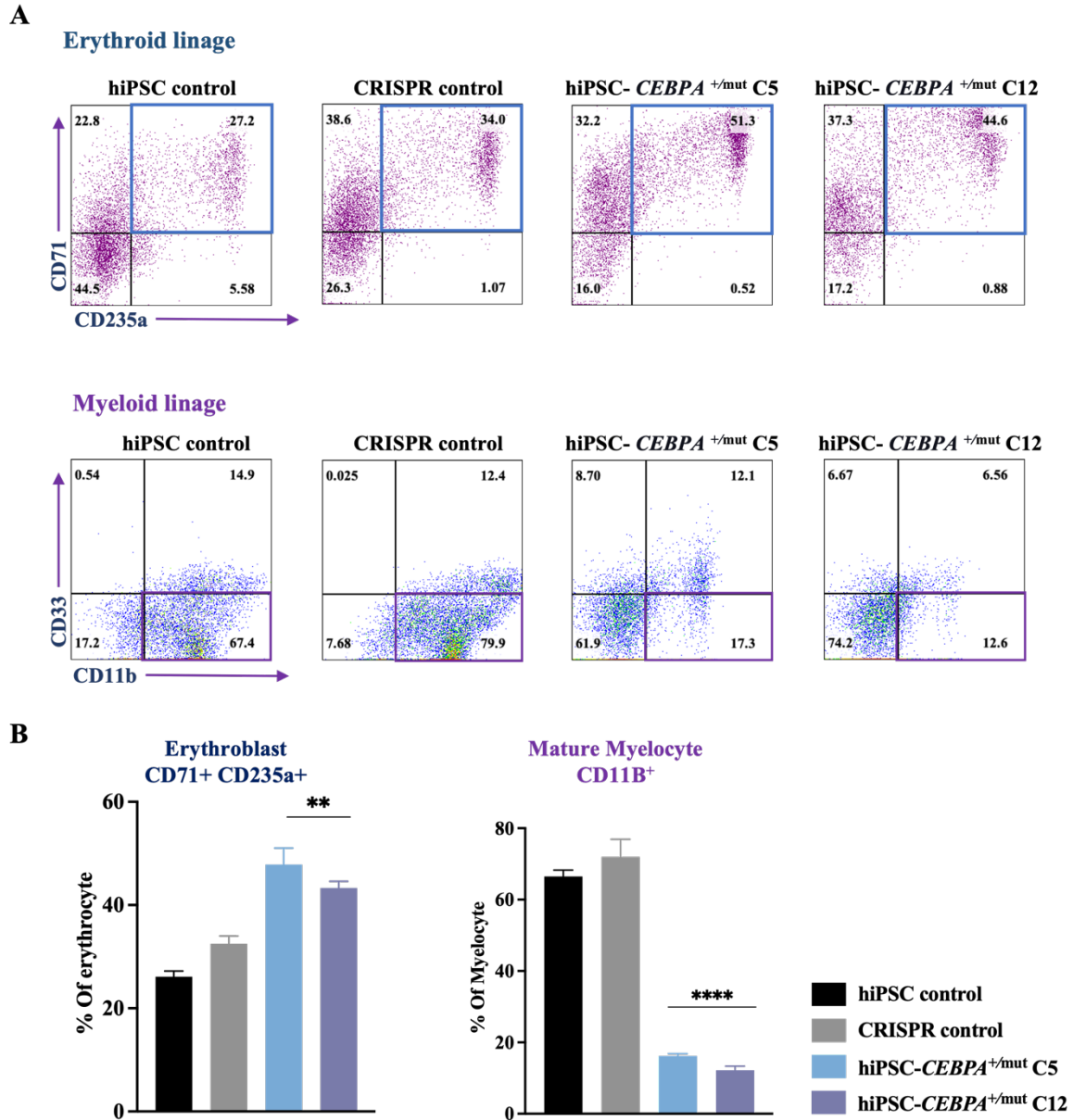
Phase contrast images exhibit the morphology of CFUs for hiPSC control, hiPSC CRISPR control and both hiPSC- *CEBPA*<sup>+mut</sup> C5 and C12 clones after 14 days of culture on methocellulose culture (Metho-cult) medium; Images were captured using a Primovert microscope (ZEISS) with a canon camera at 4× magnification. Scale bar = 100 μm. representative images of n=3.

A more accurate determination of the proportion of cells differentiating into erythroid and myeloid lineages was obtained by collecting cells from the methylcellulose medium, washing several times with PBS, staining for erythroid markers (CD71 and CD235a) and myeloid markers (CD33, CD11b), and assessing expression by flow cytometry. This analysis indicated that hiPSC- *CEBPA*<sup>+/-mut</sup> clones (C5 and C12) generated a higher percentage of erythroid cells in methylcellulose culture, with around 45 to 50% erythroblast cells (CD71<sup>+</sup>CD235a<sup>+</sup>, blue square), compared to 27–34% in hiPSC-control and hiPSC CRISPR control clones (blue square) (Figure 4.7 A and B).

The flow cytometry analysis further revealed an impairment in the formation of mature myeloid CD11b<sup>+</sup> cells in the hiPSC- *CEBPA*<sup>+/-mut</sup> C5 and C12 clones (17% and 12%, respectively) compared to the parental control and CRISPR control hiPSC clones (67.4 and 79.9%, respectively) (purple rectangle) (Figure 4.7, A and B).

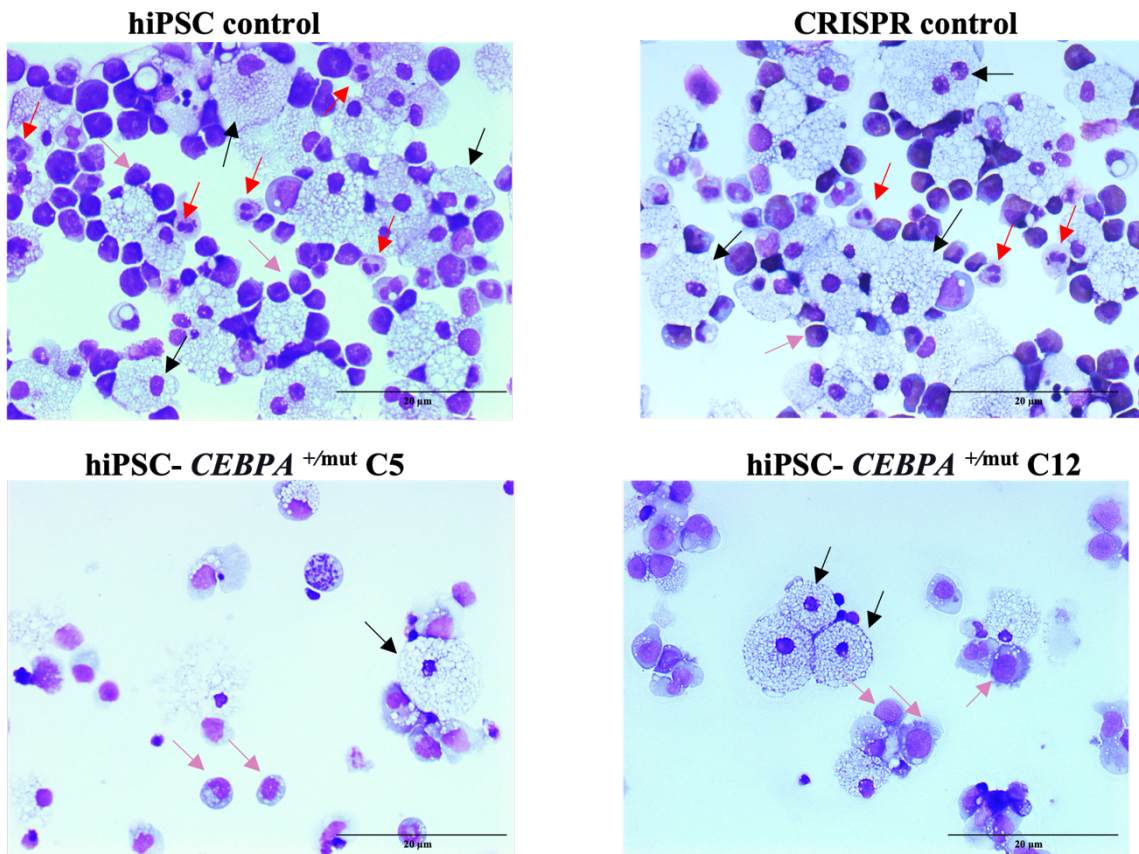
Morphological analysis of the CFUs after 14 days in methylcellulose-based media and phenotypical analysis by flow cytometry demonstrated that HSPCs from the hiPSC control and hiPSCs CRISPR control clones could differentiate and produce a heterogenous population of erythroid and myeloid cells. By contrast, a similar morphological analysis of hiPSC- *CEBPA*<sup>+/-mut</sup> (C5 and C12) clones showed mostly erythroblast cells and monocytes (Figure 4.8).

In conclusion, our findings show that the new HSPCs derived from hiPSCs with a monoallelic *CEBPA* mutation can differentiate in methylcellulose semi-solid medium, albeit with a limited transition to the myeloid lineage, and that they primarily differentiate into erythroblast cells. These results support the critical role of *CEBPA* in regulating myeloid differentiation, in agreement with previous studies (Alan D Friedman 2015; Avellino and Delwel 2017).



**Figure 4.7 Characterisation of colony forming units in hiPSC- control and hiPSC *CEBPA*<sup>+/mut</sup> clones.**

(A) Flow cytometry dot plots showing the expression of erythroid markers (CD71 and CD235a) and myeloid markers (CD33 and CD11b) of cells after 14 days of differentiation on semi-solid medium. The strategy of gating is shown in supplementary figure 2. A blue square defines erythroblast cells (CD71<sup>+</sup>CD235a<sup>+</sup>) and a purple rectangle represents mature myeloid cells. (B) Histogram showing the mean percentage of erythroblast markers (CD71<sup>+</sup>CD235a<sup>+</sup>), and mature myeloid cells (CD11b<sup>+</sup>). Data are presented as mean± standard error of the mean (SEM) and evaluating using one-way ANOVA with multiple comparisons. The *p*-values for erythroblasts were \*\* 0.0033 (C5) and 0.0079 (C12) compared with hiPSC control cells and the *p*-value was \*\*\*\* <0.0001 for mature myelocytes from hiPSC- *CEBPA*<sup>+/mut</sup> compared with hiPSC and hiPSC CRISPR control cells. Data represent three independent experiments.



**Figure 4.8 Morphological assessment of cells from clonogenic assay.**

Cytospin of CFUs after 14 days shows granulocytes (**red arrows**) only for the hiPSC control and CRISPR control cells, the erythroid cells (**pink arrows**) and monocytes (**black arrows**). The images were captured with a Leica DM6000 camera at 40× magnification, Scale bar = 20 μm, representative images of n=3.

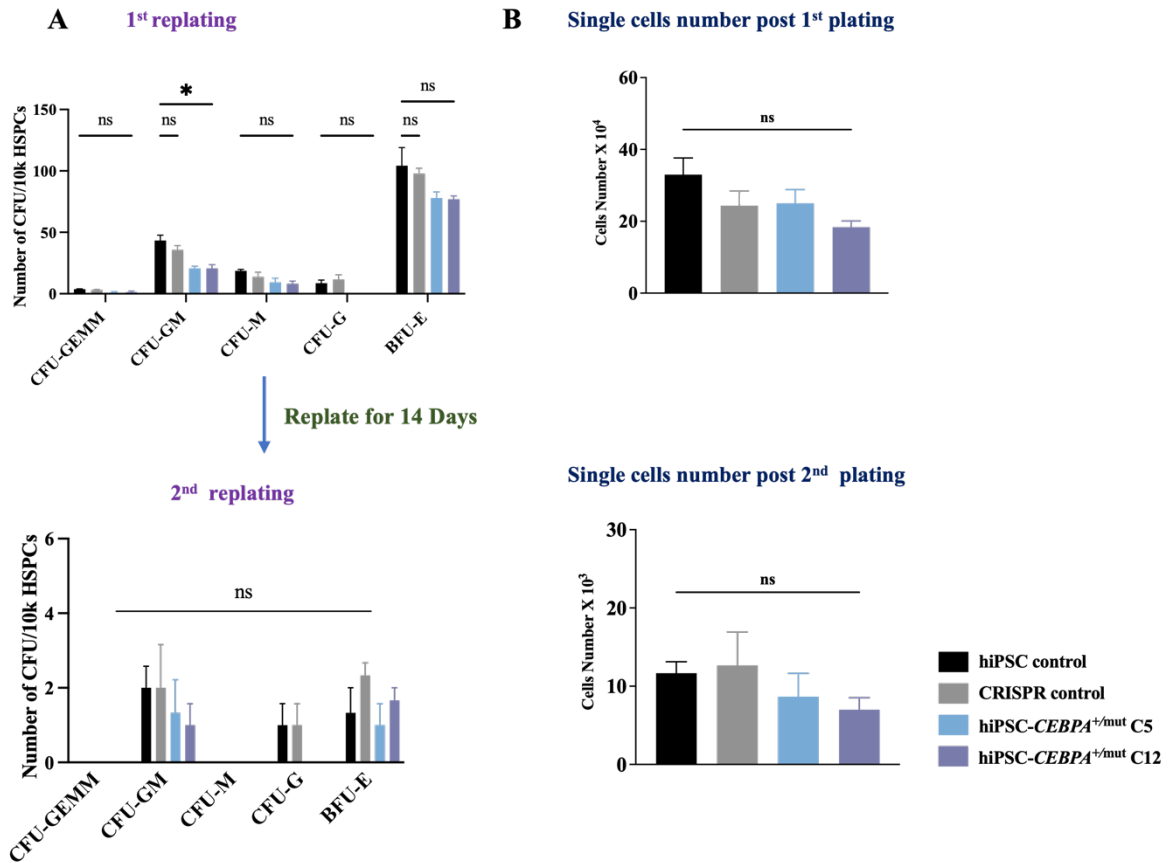
### 4.2.3 Defining the role of *CEBPA*<sup>+mut</sup> in the proliferation and ability of self-renewal of HSPCs

We further evaluated the effect of the *CEBPA*<sup>+mut</sup> on self-renewal and proliferation capacity by conducting serial re-plating in methylcellulose medium. In the initial plating,  $1 \times 10^4$  HSPCs were plated in methylcellulose, and after 14 days, colonies were scored based on their morphology. As in previous experiments, a decrease in the number of CFU-G and CFU-GM colonies was observed in the *CEBPA* mutant clones (Figure 4.9A). Secondary re-plating was conducted by collecting all the single cells from the initial plating, counting them, seeding them in fresh methylcellulose medium for another 14 days, and then scoring and counting the colonies again.

Human hiPSC control, CRISPR control, and hiPSC- *CEBPA*<sup>+mut</sup> clones had negligible re-plating capacity, with an occurrence of 1 to 2 colonies of CFU-GM formed by hiPSC- *CEBPA*<sup>+mut</sup> clones compared to 3 to 4 formed by hiPSC control and CRISPR control clones. The BFU-E colony scoring revealed that both the mutant and control HSPC-iPSCs produced one or two BFU-E colonies. (Figure 4.9A, B).

Cells were also collected at each re-plating and counted. A marked decrease was observed in the number of cells between the first and second re-plating of both hiPSC- *CEBPA*<sup>+mut</sup> clones (20-fold reduction). Similar results were observed with the other hiPSC cell lines (hiPSC control and CRISPR control), as the total number of cells was again dramatically reduced between the first and second re-plating (Figure 4.9B).

Overall, employing an *in vitro* hiPSC model confirms the fact that monoallelic of the *CEBPA* bZIP domain induces myeloid maturation defects, which were previously attributed to *CEBPA*. Our results also indicate that this monoallelic mutation does not increase the proliferation or self-renewal capacity of HSPCs.

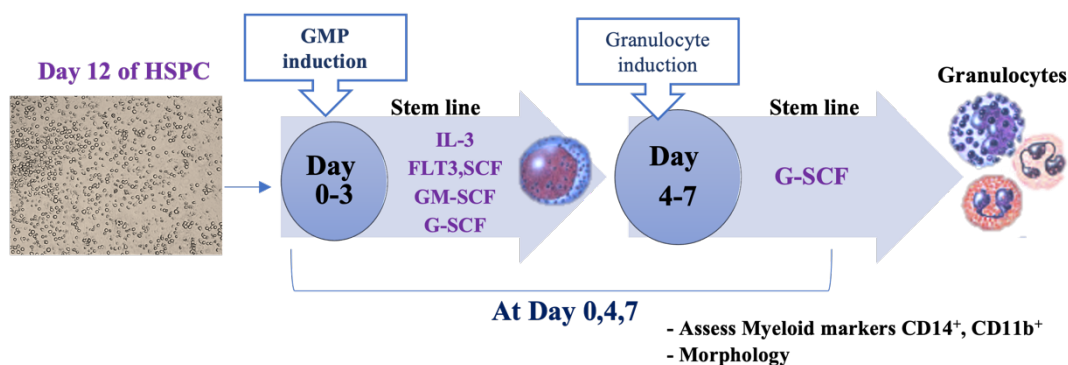


**Figure 4.9 Analysis of self-renewal capacity of CFUs.**

(A) CFUs derived from hiPSC-control, CEISPR control and hiPSC- *CEBPA*<sup>+/mut</sup> clones were replated every 14 days. The *p*-values for the 1<sup>st</sup> plating compared with the hiPSC control (GM) were \* (C5) and (C12), (E) and (GEMM) were ns. The *p*-values for the 2<sup>nd</sup> plating were all ns for the hiPSC control, hiPSC CRISPR control and hiPSC- *CEBPA*<sup>+/mut</sup> CFUs. (B) Bar graphs showing the number of single cells obtained after each experiment. Statistical results are presented as mean  $\pm$  SEM and analysed using one-way ANOVA with multiple comparisons: The *p*-values for the 1st and 2nd re-plating hiPSC- *CEBPA*<sup>+/mut</sup> clones compared with the WT control and CRISPR control were ns. Data are presented as the mean  $\pm$  SEM and were evaluated using two-way ANOVA with multiple comparisons and represent three independent experiments.

#### 4.2.4 Analysing the myeloid lineage potential of hiPSC – *CEBPA*<sup>+mut</sup> clones

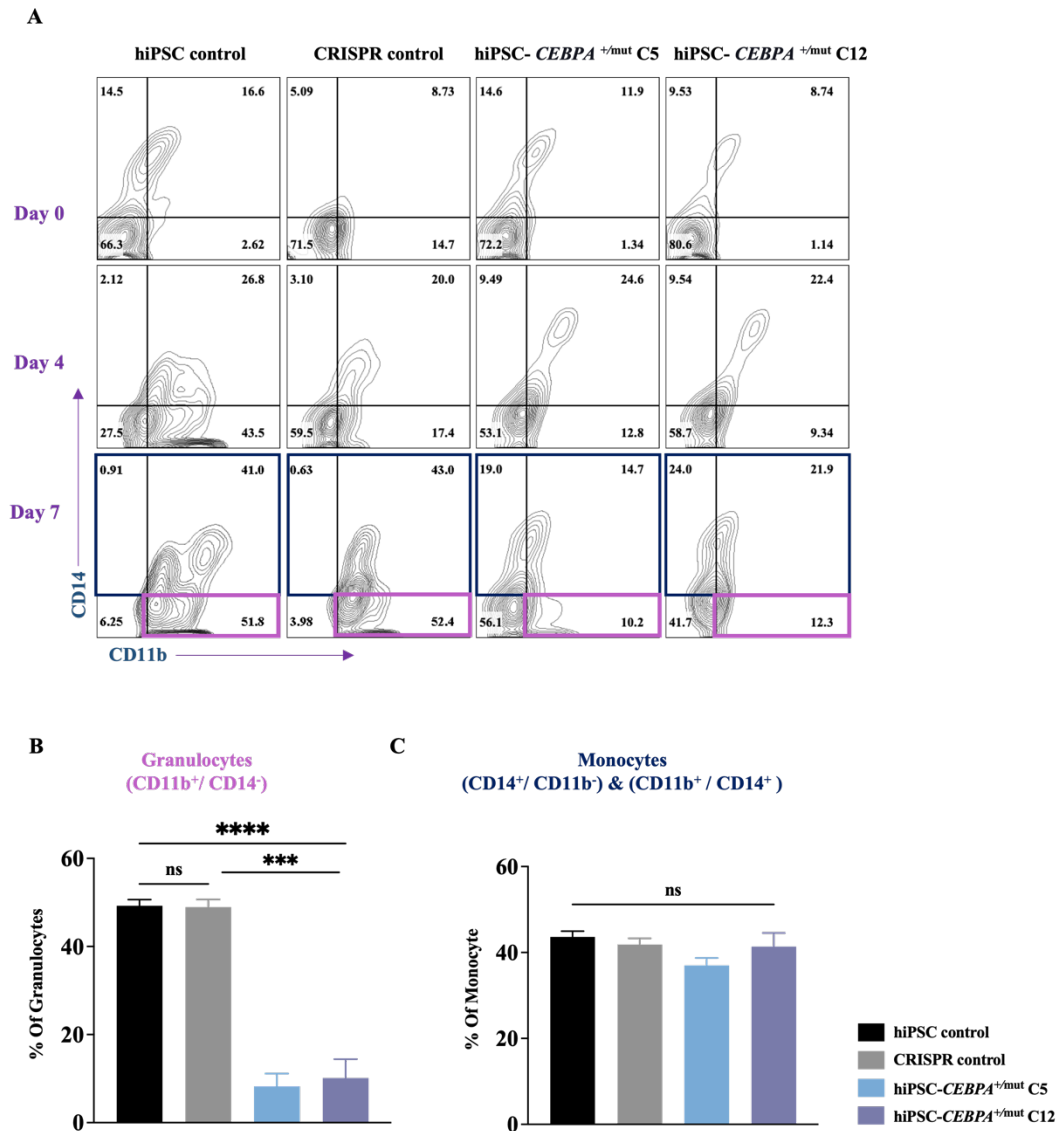
As discussed previously, a monoallelic disruption of the *CEBPA* bZIP domain alters the myeloid differentiation potential of HSPCs growing in semi-solid cultures (colony assay). We decided to perform a myeloid differentiation assay in liquid culture to allow us to evaluate the phenotypic changes in the cells at different times during the progression of the differentiation. Myeloid differentiation was initiated by culturing HSPCs (taken on day 12) in Stemline II® medium containing a cocktail of myeloid-inducing cytokines (IL-3, SCF, FLT3 GM-CSF, and G-CSF). After growing the HSPCs for three days in this medium, the cells were grown for an additional four days with only G-CSF to drive the differentiation to granulocytes (Chakraborty and Tweardy, 1998; Rossetti et al., 2010). At different times during the differentiation process (Days 0, 4, and 7), the cells were harvested and both antigen expression and morphological analysis were determined (Figure 4.10).



**Figure 4.10 Schematic representation method and timeline of myeloid differentiation.**

Flow cytometry analysis was used to evaluate the emergence of granulocytes and monocyte. Cells positive for the CD11b marker but negative for CD14 (CD11b<sup>+</sup>CD14<sup>-</sup>) represent granulocytic cells, whereas cells positive for the CD14 marker represent monocytic cells, regardless of their CD11b expression (Nielsen et al. 1994). The approach used for gating is depicted in Supplemental Figure 3. Our results revealed that at day 7 of myeloid differentiation, 50% of the cells had differentiated into granulocytes (CD11b<sup>+</sup>CD14<sup>-</sup>) in both control lines (parental and CRISPR control) (Figure 4.11.A). Conversely, the hiPSC- *CEBPA*<sup>+/-mut</sup> clones (C5 and C12) had significantly reduced differentiation into granulocytes, at 10.2% and 12.3%, respectively (Figure 4.11 B), although the differentiation into monocytes was similar in all the cases (Figure 4.11 C).





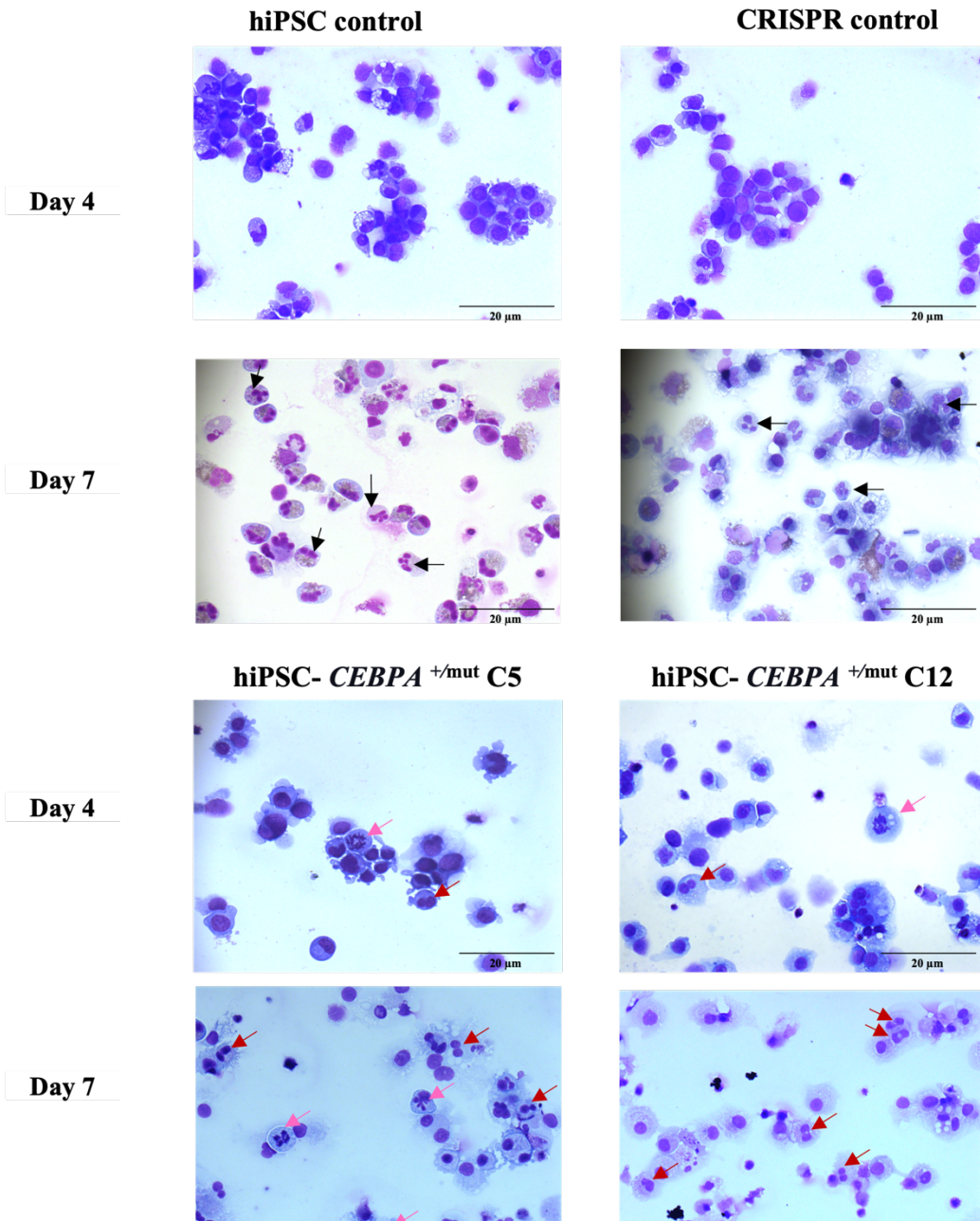
**Figure 4.11 Analysis of myeloid differentiation of HSPCs derived from hiPSC-*CEBPA*<sup>+/-mut</sup>.**

(A) Contour plots showing the percentages of CD11b<sup>+</sup>, CD14<sup>-</sup>, CD11b<sup>+</sup> and CD14<sup>+</sup> tracked during differentiation on days 0, 4 and 7. (B) Bar graph depicting the mean percentage of (CD11b<sup>+</sup> / CD14<sup>-</sup>) (granulocytes) on day 7. For granulocytes compared to the hiPSC control, hiPSC-*CEBPA*<sup>+/-mut</sup> C5 and hiPSC-*CEBPA*<sup>+/-mut</sup> C12, the *p*-values were \*\*\*\**p* < 0.0001; compared to the hiPSC CRISPR control, the *p*-values were \*\*\* (0.0002) C5 and (0.0003) C12. (C) Bar graph depicting the mean percentage of (CD11b<sup>+</sup> / CD14<sup>+</sup>) (monocytes) on day 7. The *p*-values for monocytes compared to the hiPSC control were not significant (ns). Statistical results are presented as mean ± standard error of the mean (SEM) and were analysed using one-way ANOVA with multiple comparisons. Data represent three independent experiments.

We also evaluated the morphology of the myeloid cells by light microscopy at different times during differentiation to obtain a better insight into characteristic maturation features, which included the granule content as well as the size and shape of the nucleus. The presence of neutrophils, identified by their segmented nuclei and the presence of granules in the cytoplasm, was determined within the myeloid population of the hiPSC control and CRISPR control clones. By contrast, no normal granulocyte morphology was observed in the myeloid cultures obtained from both hiPSC- *CEBPA*<sup>+mut</sup> C5 and C12 clones, confirming the phenotypic abnormalities observed by flow cytometry (Figure 4.12). Strikingly, erythrocyte precursors with bilobed nuclei were seen only in the mutant clones.

Interestingly, the morphological results for myeloid cells in mutant clones revealed a dysplastic morphology, including hypo-segmentations, agranular segmented neutrophils, and hyper-segmented and Pseudo Pelger Huët anomaly (Figure 4.13 A). The percentage of abnormal morphology was obtained by scoring the abnormal morphology in 500 cells at various time intervals, and the percentage of hiPSC- *CEBPA*<sup>+mut</sup> clones with aberrant morphology relative to the abnormal morphology of the hiPSC control was determined at days 4 and 7. As shown in Figure 4.13 B, the percentage of aberrant cells was significantly higher (50%) for hiPSC- *CEBPA*<sup>+mut</sup> clones than for WT and CRISPR control hiPSC clones at days 4 and 7 of differentiation.

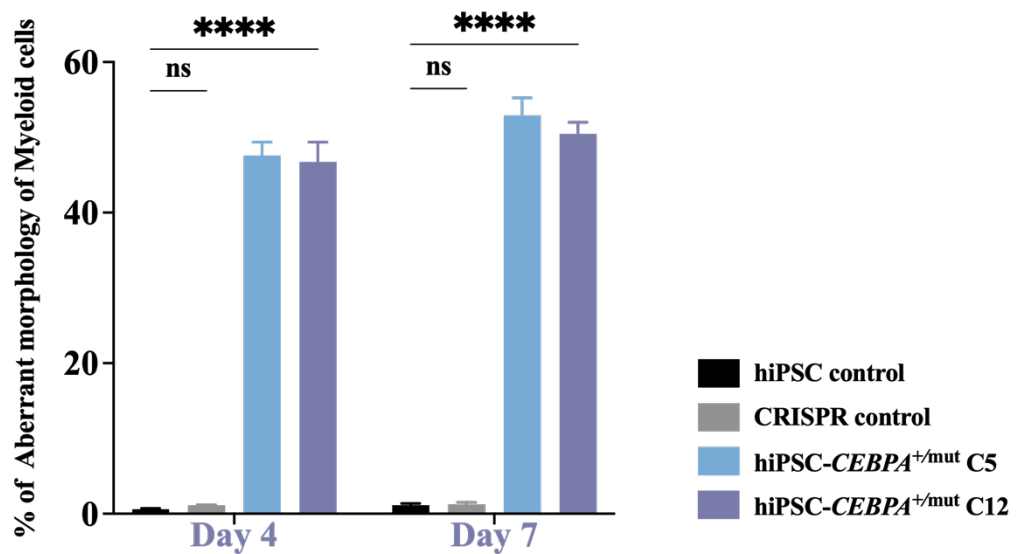
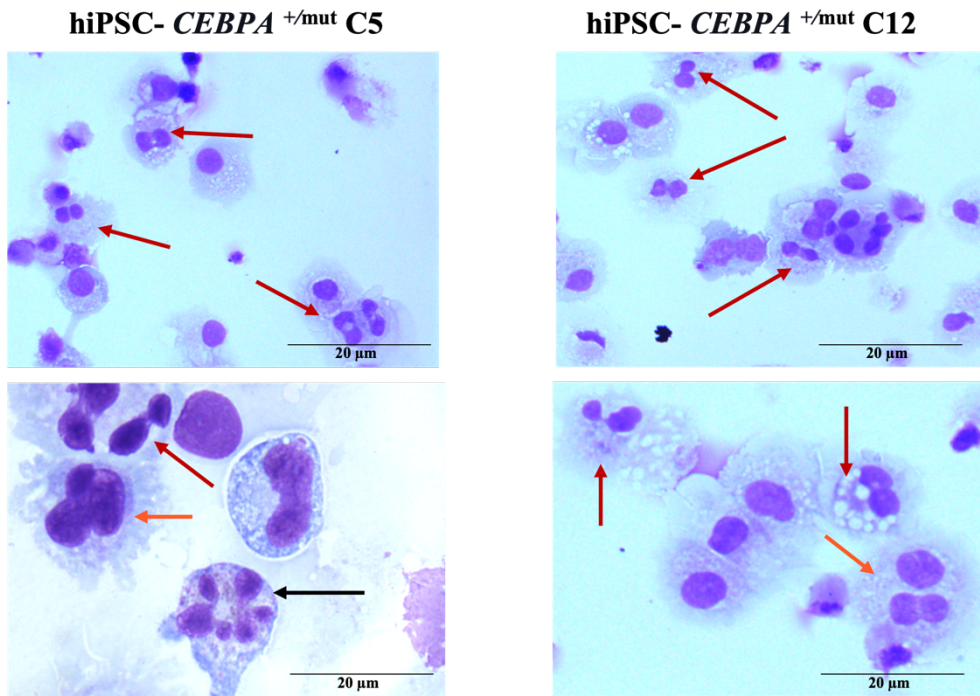
In summary, our findings demonstrated that the hiPSC- *CEBPA*<sup>+mut</sup> mutation diminishes the myeloid differentiation into the granulocytic lineage, as shown in previous studies (Pabst and Mueller 2009). In addition, significant dysplastic morphology was observed in the myeloid cells. Our findings, derived from our novel *in vitro* model system (hiPSCs), corroborated previous research highlighting the significant role of *C/EBPα* as a potent modulator of differentiation in the myeloid lineage and that it plays a role in activating myeloid cell programming.



**Figure 4.12 Myeloid cells morphology.**

Myeloid cells show the presence of granulocytic cells (**black arrows**) in hiPSC control and CRISPR control clones, while hiPSC- *CEBPA*<sup>+mut</sup> C5 and C12 clones show hypo-segmented neutrophils (**red arrow**) and bilobed erythrocyte precursors (**pink arrows**). Images were captured with a Leica DM6000 microscope at 40× magnification. Scale bar = 20 μm, representative images of n=3.

A



**Figure 4.13 Morphology analysis of aberrant myeloid cells.**

(A) hiPSC-CEBPA<sup>+/mut</sup> exhibited aberrant morphology, including agranular segmented neutrophils (orange arrow), pseudo-Pelger-Huet anomaly (bi-lobed neutrophils) (red arrow) and hyper-segmented neutrophils (black arrow). The images were captured with a Leica DM6000 microscope at 100× magnification. Scale bar = 20 μm, representative images of n=3. (B) Histogram represents the ratio of the percentage of aberrant cells of hiPSC-CEBPA<sup>+/mut</sup> clones to the aberrant cells of the hiPSC control. Results are presented as mean ± SEM and analysed using two-way ANOVA with multiple comparisons. \*\*\*\* *p*-value <0.0001. The data represent three independent experiments.

#### 4.2.5 Investigating *CEBPA*<sup>+mut</sup> target gene expression

Our myeloid differentiation data showed a profound suppressive effect of a *CEBPA* monoallelic mutation in the DBD on granulocyte differentiation. The fact behind this is that the *CEBPA* mutation in the DNA-binding domain altered the gene functionality. To confirm this, we determined the expression levels of various genes known to be regulated by *CEBPA* and that are critical for haematopoiesis differentiation.

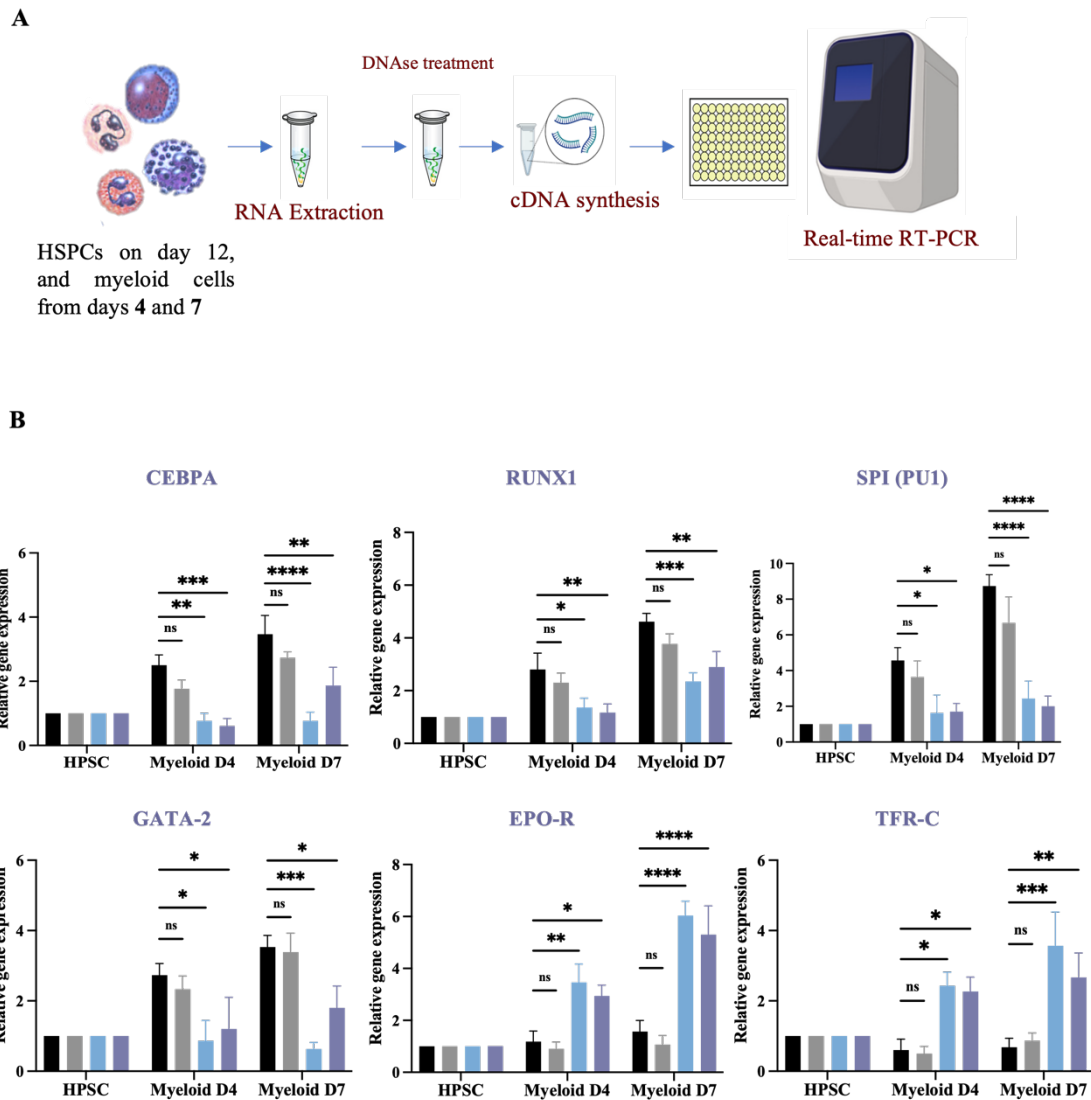
RNA was extracted from HSPCs on day 12, as well as from myeloid cells obtained on days 4 and 7 from hiPSC control, CRISPR control, and hiPSC- *CEBPA*<sup>+mut</sup> C5 and C12 clones. The RNA samples were used to synthesise cDNA, which was subjected to gene expression analysis using qRT-PCR (Figure 4.14 A).

According to our findings, the expression of *CEBPA* mRNA was reduced in myeloid cells from hiPSC- *CEBPA*<sup>+mut</sup> clones on day 4 compared to myeloid cells from hiPSC control and CRISPR control clones. This reduction persisted until day 7 for the C5 clone and was slightly greater for the C12 clone (Figure 4.14 B). Compared to the controls, the myeloid cells derived from the hiPSC- *CEBPA*<sup>+mut</sup> clones showed reduced levels of *RUNX1*. This reduction could potentially be attributed to a decline in the expression of *CEBPA*.

*PU.1* and *GATA2* play crucial roles as regulators of haematopoiesis and myeloid differentiation. *CEBPA* and other transcription factors engage in interactions to govern the regulation of transcriptional programmes during the development of myeloid cells. Evaluation of the transcription factors *PU.1* and *GATA2* by qPCR from myeloid and HSPC cells generated from hiPSCs revealed lower expression of *PU.1* in myeloid cells generated from the hiPSC- *CEBPA*<sup>+mut</sup> C5 and C12 clones. The expression of *GATA2* was also reduced in myeloid cells derived from hiPSC – *CEBPA*<sup>+mut</sup> on day 4, although a marginal increase was observed in C12 compared to C5 on day 7.

Erythrocyte precursors were predominantly present in the myeloid liquid cultures of the mutant clones. Measurement of erythroid markers confirmed the induction of erythroid differentiation by *EPO-R*, as demonstrated in previous studies (Koury et al. 1989). By day 4 of myeloid differentiation, the level of *EPO-R* was higher in both hiPSC – *CEBPA*<sup>+mut</sup> clones compared to the iPSC control lines. This expression was further increased by day 7. The *TFRC* expression levels in *CEBPA* mutant clones paralleled those of *EPO-R* levels at day 4, whereas the *TFRC* level in myeloid cells was lower in C12 than in C5 at day 7.

Our findings indicate that disruption of the *CEBPA* bZIP domain significantly affects myeloid cell differentiation by decreasing myeloid-inducing genes, while it causes elevations in erythroid markers.



**Figure 4.14 Myeloid and erythroid marker expression in HSPCs and myeloid cells derived from hiPSC-*CEBPA*<sup>+/*mut*</sup> clones.**

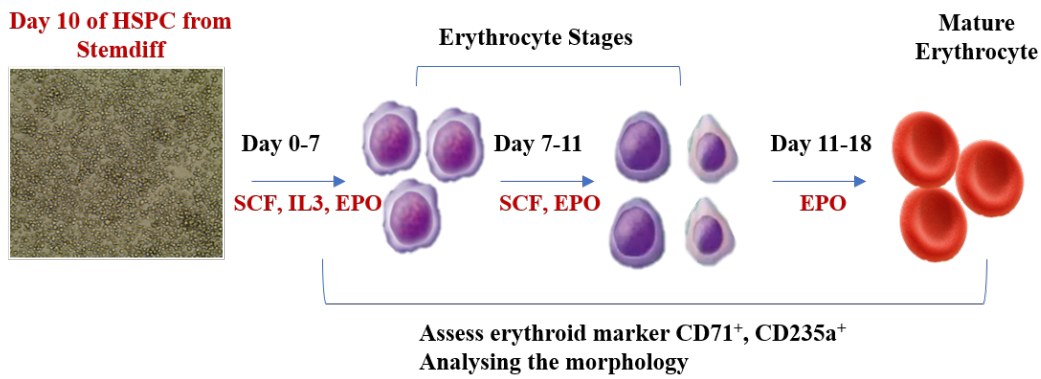
(A) Schematic diagram showing the use of qPCR to measure myeloid and erythroid markers in HSPCs, as well as myeloid cells. (B) Histogram graphs show the findings from the qRT-PCR performed on HSPC and myeloid cells derived from the hiPSC control, hiPSC, CRISPR control and hiPSC-*CEBPA*<sup>+/*mut*</sup> C5, and hiPSC-*CEBPA*<sup>+/*mut*</sup> C12. The relative expression was normalised to *GAPDH* and expressed relative to HSPC. Statistical results are presented as mean  $\pm$  standard error of the mean (SEM) and were analysed using two-way ANOVA with multiple comparisons. ns = not significant, \* $p < 0.05$  and \*\* $p < 0.01$ , \*\*\* $p < 0.001$ , \*\*\*\* $p < 0.0001$ . Data represent three independent experiments.

#### **4.2.6 Investigating the capacity of hiPSCs harbouring *CEBPA*<sup>+/-mut</sup> to differentiate into the erythroid lineage**

The results of the colony assays demonstrated that the hiPSC- *CEBPA*<sup>+/-mut</sup> clones were able to produce erythroid colonies (Figure 4.15), suggesting the possibility that the maturation of erythrocytes from these clones might be influenced in the same way as observed for granulocyte maturation. Therefore, we assessed the ability of these cells to differentiate into the erythroid lineage, as well as the potential emergence of aberrant erythroid cells in liquid culture in the presence of a cocktail of cytokines that enhance erythroid differentiation.

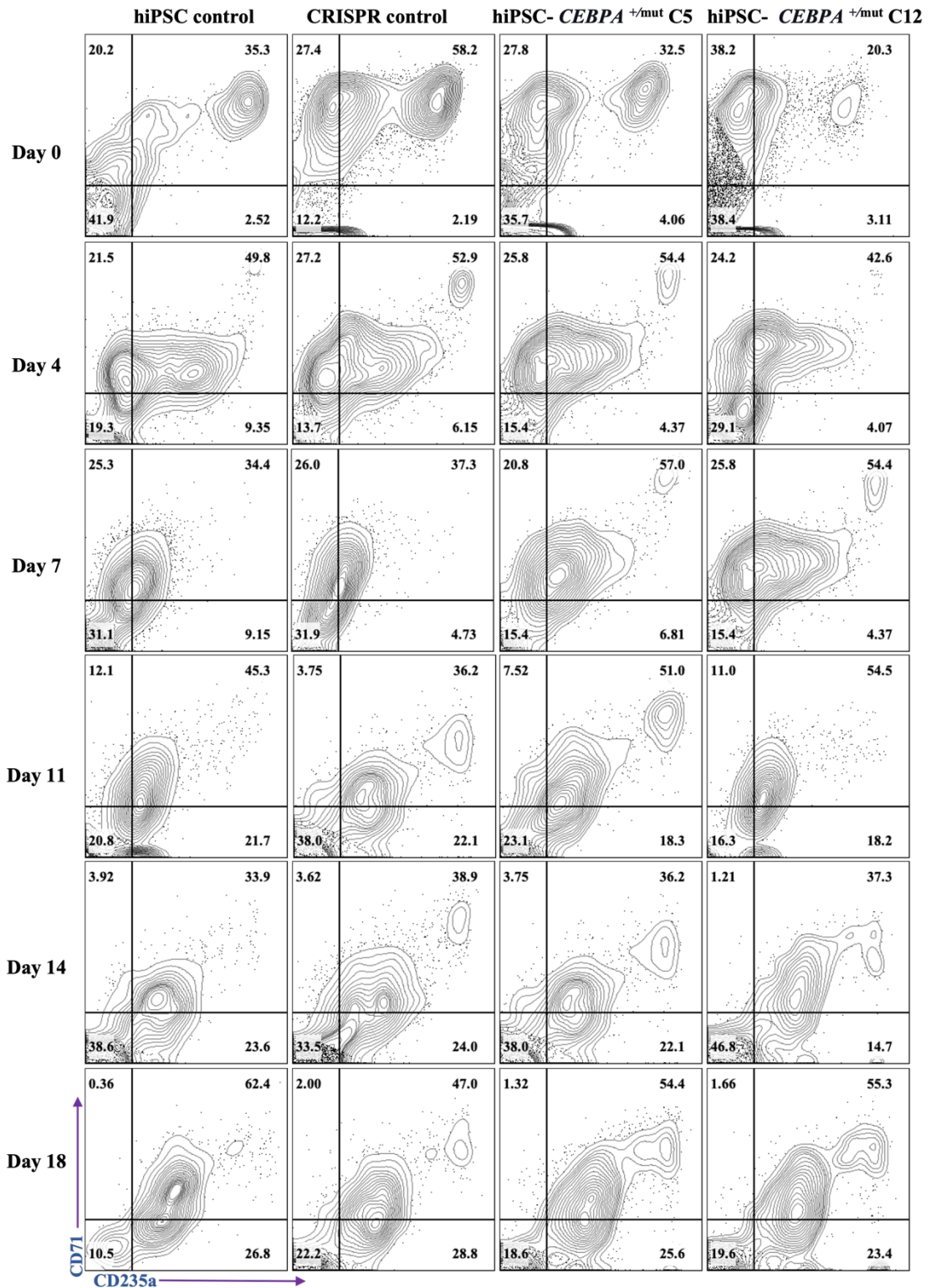
HSPCs from hiPSC control and CRISPR control and hiPSC – *CEBPA*<sup>+/-mut</sup> clones (C5 and C12) were harvested on day 10 of STEMdiff and cultured for 18 days in a three-phase erythropoiesis liquid culture (Figure 4.15). The erythroid culture was checked on days 0, 4, 7, 11, 14 and 18, and the distribution of erythroid cells throughout different maturation stages was determined by the expression of progenitor marker (CD71) and erythroid mature marker glycophorin A (CD235a). We were able to track the course of erythroid differentiation from progenitor cells (CD71+) to erythroblasts (CD71+ / CD235a+) and mature erythrocytes (CD71- / CD235a+) by using these markers (Macrì et al. 2015). The erythroid cell gating method is shown in Supplemental Figure 4.





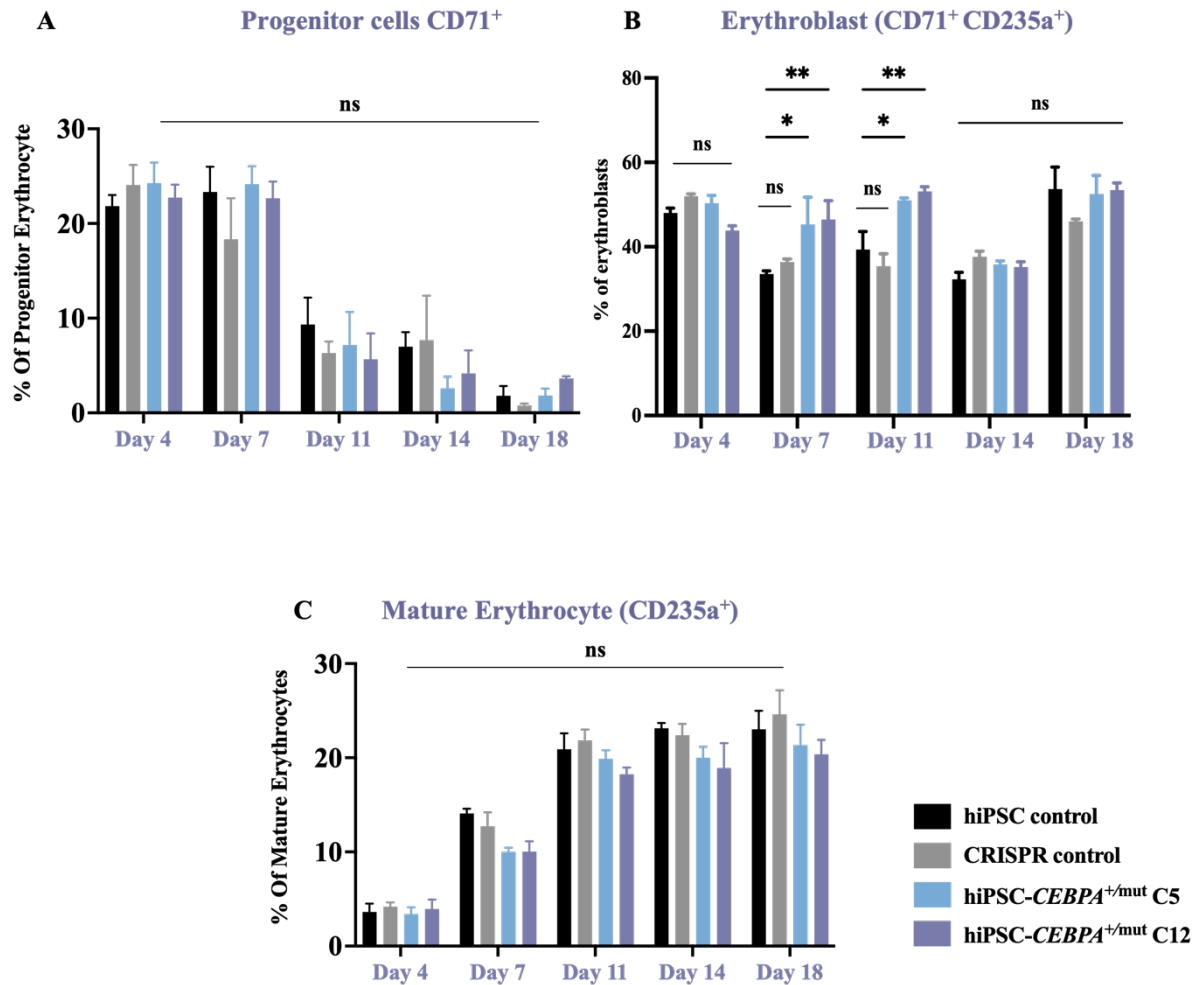
**Figure 4.15 Schematic representation method and timeline of erythroid differentiation.**

During erythroid differentiation, comparable proportions of erythrocyte progenitors, erythroblasts and mature erythrocytes were observed in all of the hiPSCs on the day of the differentiation (Figure 4.16). CD71<sup>+</sup> was measured in the hiPSC-control, CRISPR control and hiPSC-*CEBPA*<sup>+/<sup>mut</sup></sup> clones after culturing in an erythroid medium for four days. Our findings revealed that approximately 21% and 27% were CD71<sup>+</sup> in the hiPSC-control and the CRISPR control, respectively, and 24–25% in the hiPSC-*CEBPA*<sup>+/<sup>mut</sup></sup> clones, all of which exhibited approximately 40–50% co-expression of CD71<sup>+</sup> and CD235<sup>+</sup> (erythroblasts). In contrast, the expression of CD235<sup>+</sup> alone was low on day 4 of the differentiation. We also observed that the erythroid cells obtained from hiPSC-*CEBPA*<sup>+/<sup>mut</sup></sup> exhibited the same percentage range of CD71<sup>+</sup> as those from the hiPSC control and the CRISPR control clones, which was 20–25% on day 7. However, the mutant clones showed a high percentage of erythroblast (CD71<sup>+</sup> / CD235<sup>+</sup>) differentiation compared to the parental and CRISPR control hiPSCs, and this pattern was sustained until day 11. At later time points, the percentage of CD71<sup>+</sup> cells was reduced, concomitant with an increase in CD235<sup>+</sup> cells. However, at the end of the differentiation phase, all the hiPSC lines exhibited a significant proportion of cells expressing the (CD71<sup>+</sup> / CD235<sup>+</sup>) phenotype. On day 18, the control cultures, as well as the hiPSC – *CEBPA*<sup>+/<sup>mut</sup></sup> clones, contained mainly mature CD235a<sup>+</sup> cells (>20%). Statistical analysis shown this percentage to be similar among the clones (Figure 4.17).



**Figure 4.16 Analysis of erythroid differentiation of HSPCs derived from hiPSC-*CEBPA*<sup>+/*mut*</sup>.**

Contour plots showing the assessment of erythroid progenitor marker CD71 and erythroid mature marker CD235a for hiPSC-control, CRISPR control and hiPSC-*CEBPA*<sup>+/*mut*</sup> cells over multiple days of differentiation. The contour plots are representative examples of three independent experiments.



**Figure 4.17 Analysis of erythroid differentiation derived from hiPSC- *CEBPA*<sup>+/mut</sup>.**

(A) Bar graph shows mean percentage of erythrocyte progenitors (CD71<sup>+</sup>). The *p*-values for progenitor cells from hiPSC- *CEBPA*<sup>+/mut</sup> compared with the hiPSC control and hiPSC CRISPR control were ns. (B) Bar graph shows mean percentage of erythroblasts (CD71<sup>+</sup> CD235a<sup>+</sup>). The *p*-values for erythroblasts of hiPSC- *CEBPA*<sup>+/mut</sup> clones compared with the hiPSC control on day 4 were ns, and on day 7 were \* 0.0147 (C5) and \*\* 0.0077 (C12). The *p*-values for erythroblasts compared on day 11 were \* 0.0421(C5) and \*\* 0.0033 (C12). The *p*-values for erythroblasts of hiPSC- *CEBPA*<sup>+/mut</sup> clones compared with hiPSC control and hiPSC CRISPR control clones on days 14 and 18 were ns. (C) Bar graph shows mean percentage of mature erythrocytes (CD235a<sup>+</sup>) at different times of erythroid differentiation. The *p*-values for mature cells compared with hiPSC control and hiPSC CRISPR control clones were ns. Results are depicted as mean ± SEM and analysed using two-way ANOVA with multiple comparisons. ns=not significant. Data represent 3 independent experiments.

We then conducted a morphological analysis at different stages of erythroid maturation to obtain a better insight into the erythroid differentiation process and to investigate whether the *CEBPA*<sup>+/-mut</sup> promoted a dysplastic phenotype in the erythroid lineage. Microscopy evaluation revealed that hiPSC control and hiPSC – *CEBPA*<sup>+/-mut</sup> clones generated a diversity of erythroid cells, with different phases of maturation at days 4 and 7 of differentiation. The maturation phases of these erythroid cells are showed in Figure 4.18.

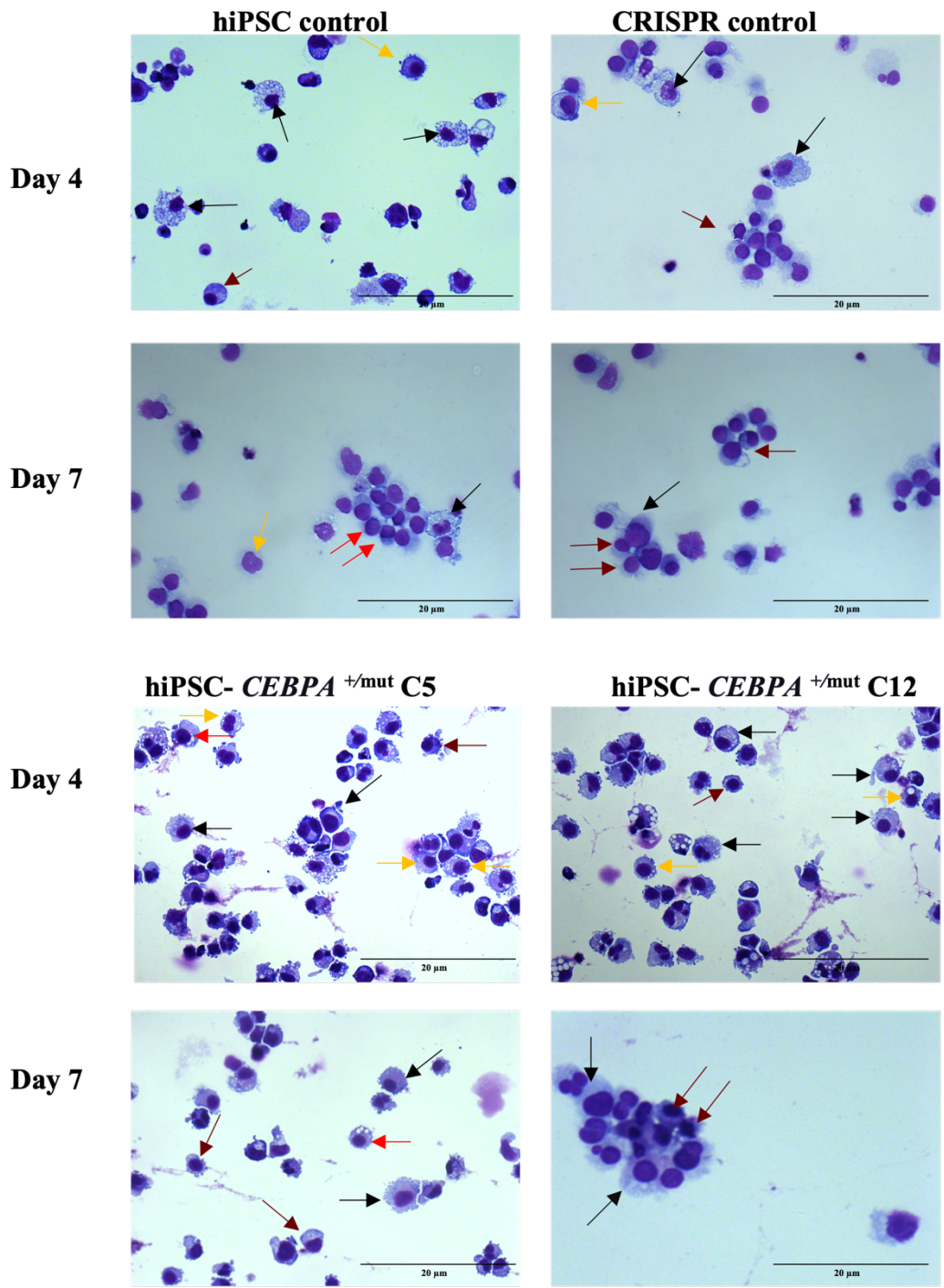
These cells were identified according to the features of each stage, including size, nuclei and cytoplasm: nucleus ratios at days 4 and 7. All these features were clearly reduced during the maturation of erythrocytes. According to the American Society of Hematology image bank, a proerythroblast is defined as a large cell with agranular cytoplasm and a spherical nucleus, a basophilic erythroblast is characterised by more condensed chromatin, a polychromatic erythroblast contains a nucleus that occupies a relatively small part of the cell, and an orthochromatic erythroblast has a small, round, dense nucleus and a slightly less basophilic cytoplasm than that of a mature red cell.

The identification of cells during the advanced stages of differentiation was challenging due to the presence of dead cells, significant debris and cells exhibiting impaired membranes on the slides of all hiPSC clones.

Interestingly, we determined during the cell morphology analysis that the erythroid cells generated from the hiPSC – *CEBPA*<sup>+/-mut</sup> clones (C5 and C12) exhibited a dysplastic appearance. As shown in Figure 4.19A, these abnormal cells included bi-nucleated erythroblasts (blue arrows), giant erythroblasts (black arrows), mitotic erythroblasts (pink arrows) and multinucleated erythroblasts (red arrows). The percentage of aberrant morphology was determined by scoring 500 of these aberrant cells on days 4 and 7 and compared for the hiPSC – *CEBPA*<sup>+/-mut</sup> clones relative to the hiPSC control clones.

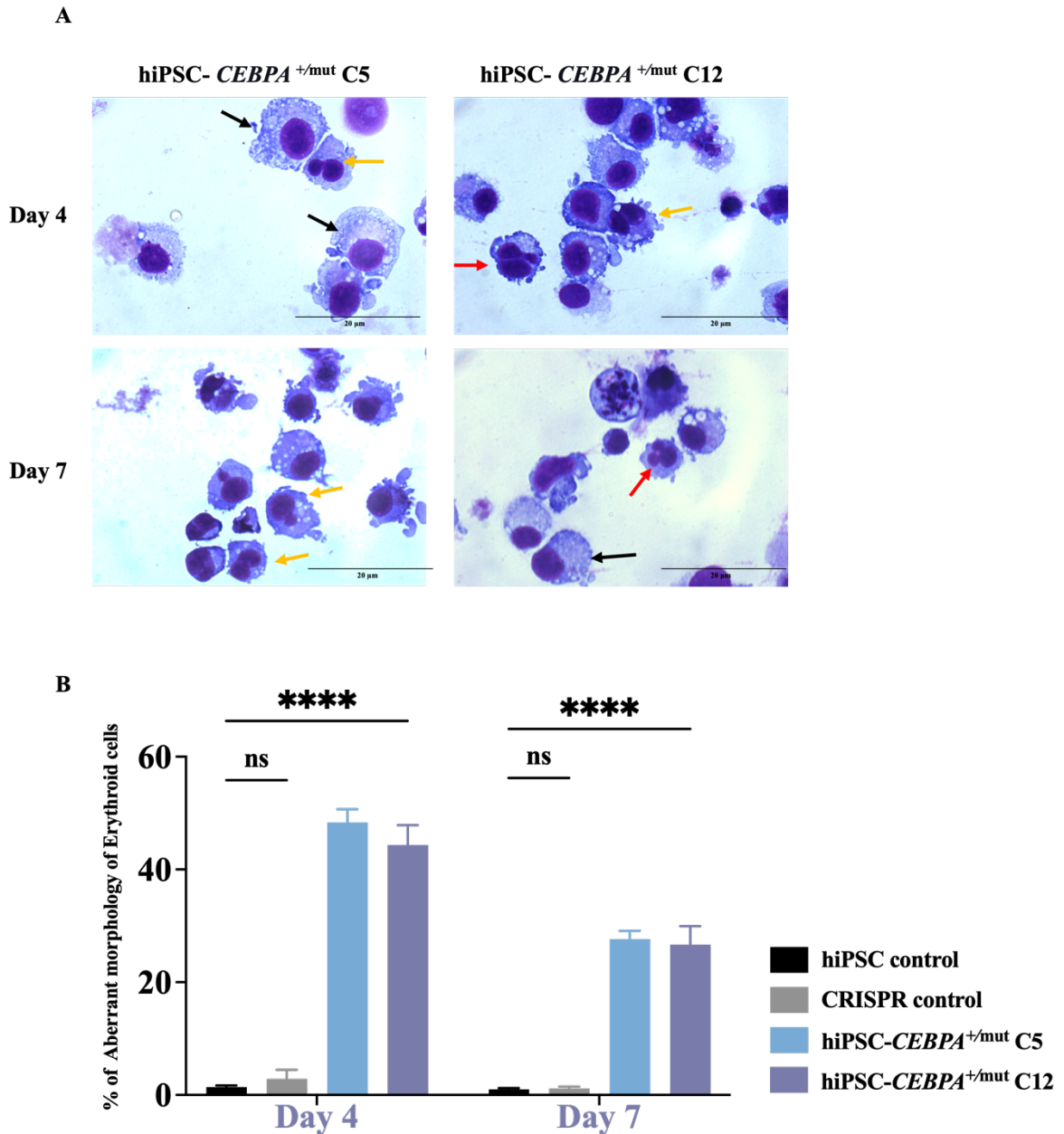
The hiPSC – *CEBPA*<sup>+mut</sup> C5 clones exhibited abnormal morphology in approximately 50% of the generated erythroid cells at day 4 of differentiation, while the percentage of these aberrant cells was around 40% for the hiPSC – *CEBPA*<sup>+mut</sup> C12 clones. By day 7, both clones showed reductions in the proportion of erythroid cells with abnormal morphology, with approximately 30% observed in each clone. This reduction was due to the immature erythroid cells undergoing a series of changes during the later days of differentiation, including changes in cell size, reduction in nucleus size, and enucleation, as well as cells undergoing apoptosis (Hattangadi et al. 2011). (Fig. 4.19 B).

Together, these findings reveal that the monoallelic mutation of *CEBPA* disrupting the bZIP domain affected the generation of normal erythroid cells. All HSPCs derived from hiPSC – *CEBPA*<sup>+mut</sup> clones had the ability to generate erythroid cells that showed normal expression of erythroblast markers until the end of differentiation; however, the erythroid cells produced by clones with monoallelic mutations in *CEBPA* exhibited distinct aberrant morphologies.



**Figure 4.18 Erythroid cell morphology.**

Erythroid cells collected at day 4 and day 7 of differentiation culture from hiPSCs control, CRISPR control and hiPSCs - *CEBPA*<sup>+/-mut</sup>: Proerythroblast (**black arrows**), basophilic erythroblast (**yellow arrows**), polychromatic erythroblast (**red arrows**) and orthochromatic erythroblast (**brown arrows**), Images were captured with a Leica DM6000 microscope at 40× magnification. Scale bar=20 μm, representative images of n=3.



**Figure 4.19 Morphology analysis of aberrant erythroid cells.**

(A) hiPSCs - *CEBPA*<sup>+mut</sup> C5 and C12 clones generate erythroid cells with aberrant morphology and bi-nucleated erythroblasts (**yellow arrows**), giant erythroblasts (**black arrows**), and multinucleated erythroblasts (**red arrows**). The images were captured with a Leica DM6000 microscope at 40 $\times$ . Scale bar=20  $\mu$ m, representative images of n=3. (B) Bar graph shows the percentage of aberrant cells generated by the hiPSCs - *CEBPA*<sup>+mut</sup> clones relative to aberrant cells of hiPSC control. Results are presented as mean  $\pm$  SEM and analysed using two-way ANOVA with multiple comparisons. \*\*\*\*  $p$ -value <0.0001. The data represent three independent experiments.

## 4.3 Discussion

### 4.3.1 *CEBPA*<sup>+mut</sup> have no impact on hiPSC differentiation into HSPCs

The phenotype of MDS and the risk of transforming into AML are influenced by several factors, including genomic instability, microenvironmental aberrations and somatic mutations. The higher incidence of MDS in elderly patients emphasises the increased risk of acquired mutations in HSCs and HSPCs (Welch et al. 2012; Kontandreopoulou et al. 2022).

The successful generation of hiPSC clones carrying the *CEBPA* monoallelic mutation which disrupts the bZIP domain led us to investigate whether this mutation could cause an alteration in the haematopoietic phenotype. Differentiation of HSPCs *in vitro* can provide a genetically tractable system for analysing and studying the physiology and pathology of human tissue development and differentiation. The four hiPSC lines (WT hiPSC, CRISPR control and the two hiPSC- *CEBPA*<sup>+mut</sup> clones) were differentiated into HSPCs to establish a successful and efficient *in vitro* HSPC differentiation method from iPSCs.

Haematopoiesis is well known to occur in two successive waves. The first wave, called ‘primitive haematopoiesis’ can generate primitive erythroblasts megakaryocytes and macrophages, but no T or B lymphocytes are generated in this wave (Baron et al. 2012; Sturgeon et al. 2014; McGrath et al. 2015b). The second wave of haematopoiesis is called ‘definitive haematopoiesis’ and occurs in the dorsal aorta of the aorta-gonad-mesonephros area of the foetus, in the foetal liver, and in BM (Medvinsky and Dzierzak 1996; Barminko et al. 2016). This second wave generates long-term repopulating HSCs with the capacity to form myeloid cells, T lymphocytes and enucleated erythrocytes expressing adult-type haemoglobin (Bertrand et al. 2010).

This knowledge has been vital for the development of sophisticated protocols that can generate HSPCs from pluripotent stem cells with a distinct and definitive phenotype. Nevertheless, some established techniques that promote the differentiation of iPSCs into HSPCs fail to create



mature HSPCs (definitive) but instead recapitulate yolk sac haematopoiesis, resulting in the formation of foetal HSPCs (primitive). A recent protocol described by Calvanese et al. (2022) validated the production of definitive HSPCs from iPSCs; however, this protocol was published during the finalising of the research and writing of this thesis.

We utilised a standard stem cell protocol to model the earliest and latest stages of HSPC development and to characterise potential target cells for differentiation into myeloid and erythroid lineages. This commercial methodology was already optimised in our lab. For example, previous work performed in the lab determined that the size of the plate used, and number of colony chunk seeded would affect the proportion of haematopoietic cells obtained and their quality. Interestingly, using this method, all hiPSC clones studied in this thesis were successfully differentiated into HSPCs and were phenotypically identified as early/primitive ( $CD34^+ CD43^+$ ) and late/definitive ( $CD34^+ CD45^+$ ) HSPC populations. The proportion of cells expressing these surface markers increased steadily from day 7 to day 14.

CD34 is expressed by numerous cell types, including endothelium and mesenchymal stem cells (Traktuev et al. 2008). Therefore,  $CD34^+$  expression alone does not prove the presence of haematopoietic stem cells. CD43 is widely expressed in haematopoietic lineage cells, including haematopoietic stem cells (Vodyanik et al. 2006; Ditadi et al. 2015; Dou et al. 2016; Daniel et al. 2016). In this study, a high percentage of  $CD43^+$  cells were present on day 7 and the proportion dramatically increased by day 14 of differentiation, in agreement with the work of the Dorn group, who reported the appearance of  $CD43^+$  cells as the first sign of haematopoietic differentiation in iPSCs and that these  $CD43^+$  cells are capable of erythropoiesis differentiation *in vitro* (Kessel et al. 2017). CD45 is a unique haematopoietic marker (Trowbridge and Thomas 1994; Ogata et al. 2005; Ngo et al. 2006). Once refined, the percentage of late HSPCs was comparable to those obtained with other methods we used in the lab, such as generating HSPC from embryoid bodies. Notably, we obtained the same differentiation efficiency of iPSCs into

HSPCs as previously described by other groups (Hansen et al., 2018; Flippe et al., 2020; Tursky et al., 2020). The HSPC differentiation results demonstrated that the monoallelic mutation in *CEBPA* disrupting the bZIP domain did not inhibit the capacity of hiPSCs to differentiate into HSPCs when compared to hiPSC control or hiPSC CRISPR control cells. Although we could not find any statistically significant differences in the formation of HSPCs, previous work performed in zebrafish revealed that *cebpa*<sup>mut/mut</sup> had a dramatic effect on myelopoiesis occurring from the early stages of differentiation to definitive haematopoiesis (Hockings et al. 2018).

By contrast, our findings are consistent with previously published results using mouse models, in which *CEBPA* knockout in mice (*Cebpa*<sup>Δ/Δ</sup>) had no effect on either the LT-HSC or ST-HSC phenotype (Higa et al. 2021). This discrepancy in outcomes between zebrafish and mice could have several explanations, including genetic deviation, physiological variations, and developmental differences. Zebrafish and mice have diverged evolutionarily, leading to changes in their genomes and gene expression profiles. Although they may share functional similarities between certain genes, they also have differences in gene structure, regulation and function, which may result in differences in biological processes and outcomes when studying specific genes.

#### **4.3.2 *CEBPA*<sup>+mut</sup> diminishes the colony-forming potential of the myeloid lineage and increases erythroid lineage differentiation, but does not impact the self-renewal ability of HSPCs**

We determined the effect of *CEBPA*<sup>+mut</sup> on the differentiation capability of HSPCs by evaluating the differentiation of HSPCs into mature haematopoietic cells when cultured on a semi-solid medium. Interestingly, the HSPCs derived from hiPSC- *CEBPA*<sup>+mut</sup> displayed a reduced potential for forming colonies, which dramatically affected the generation of myeloid CFUs, presumably a reflection of the impaired haematopoiesis associated with *CEBPA*

functions (Zhang et al. 2004; Wang et al. 2006). Our model demonstrates that the generation of the *CEBPA* monoallelic mutation disrupts the differentiation lineage programme of HSPCs by reducing the potential for myeloid differentiation, especially in the granulocyte lineage, and increases the potential for erythroid differentiation.

The Keller group has reported that the foetal liver of a mouse model that lacks the ability to produce the C/EBP $\alpha$  protein showed a marked reduction in the populations of GMPs and CFU-M, along with an increase in megakaryocyte/erythroid progenitors (Heath et al. 2004). Another group also showed that the inducible ablation of C/EBP $\alpha$  in the bone marrow caused a decrease in neutrophils and monocytes and eradicated eosinophils in peripheral blood (Zhang et al. 2004).

In our study, granulocytic colonies were only produced from HSPCs derived from the WT and CRISPR controls; therefore, C/EBP $\alpha$  appears to be necessary for CMP cells to differentiate into GMP cells. This agrees with previous studies performed in both human and mouse experimental model systems. For example, biallelic mutation of *cebpa* in mouse model shows profound developmental effects during the transition of myeloid cells into granulocytes, resulting in an accumulation of blast cells in the bone marrow. In addition, the complete loss of C/EBP $\alpha$  leads to a block upstream of GMP differentiation (Zhang et al. 2004; Porse et al. 2005; Hasemann et al. 2014; Avellino and Delwel 2017). Previous work has also indicated the essential role of C/EBP $\alpha$  in controlling the quiescence of HSCs, with C/EBP $\alpha$  deficient HSCs showing increased competitive repopulation activity in mouse transplant studies (Zhang et al. 2004). A role for C/EBP $\alpha$  in HSC self-renewal and repopulating activity has also been proposed (Zhang et al. 2004). It has an important function in regulating cell proliferation and differentiation (Johansen et al., 2001; Porse et al., 2001; Castilla, 2008).

We evaluated whether *CEBPA*<sup>+/*mut*</sup> affected the self-renewal capacity of HSPCs in our model by conducting serial re-plating of the hiPSC- *CEBPA*<sup>+/*mut*</sup> lines and counting the number of

CFUs. The *CEBPA*<sup>+mut</sup> did not noticeably affect the self-renewal of HSPCs. One possible explanation is that our model has only the monoallelic mutation. HSPCs derived from mice carrying *C/EBPα*<sup>-/-</sup> are reportedly hyperproliferative and fail to differentiate (Zhang et al. 2004). One reason could be that *C/EBPα*<sup>-/-</sup> HSCs have elevated levels of the Bmi-1, which promotes self-renewal (Iwama et al. 2004). Other studies using the same model also indicated that CFUs continued to expand in number up to a fourth re-plating (Higa et al. 2021).

#### **4.3.3 *CEBPA*<sup>+mut</sup> affects granulocyte differentiation**

*C/EBPα* appears to enhance myeloid differentiation during the transition from CMPs to GMPs (Pundhir et al. 2018). This gene was most strongly expressed in CD11b<sup>+</sup> mature myeloid cells, which are nearly exclusively neutrophils, with rare monocytes, in the bone marrow. This morphology indicates that *C/EBPα* is produced initially at the time of myeloid precursor commitment and persists in mature myeloid cells but not in lymphoid cells (Traver et al. 2001; Miyamoto et al. 2002). Our findings show that clones harbouring the *CEBPA*<sup>+mut</sup> mutation displayed a reduction in myeloid differentiation and a lack of normal granulocyte cells. These findings align with previous studies highlighting *C/EBPα* as a critical regulator of granulopoiesis, as its expression enables haematopoietic progenitors to differentiate (Grossmann et al. 2012). Nerlov's group also showed that mouse models harbouring a mutation in the C-terminal domain are incapable of promoting myeloid differentiation (Bereshchenko et al. 2009). Given the profound phenotype observed by blocking granulocytic differentiation, we could speculate that the monoallelic mutation of *CEBPA* disrupting the bZIP domain may act as a dominant negative mutation. This possibility is supported by a previously published study showing that the mutation-disrupted function of *C/EBPα* reduces myeloid priming of HSCs and simultaneously impedes commitment to the myeloid lineage (Bereshchenko et al. 2009). Moreover, these data tie in well with previous studies showing a stronger inhibition of neutrophil differentiation in response to G-CSF in a mouse model expressing a monoallelic C-

terminal mutation than an N-terminal mutation, indicating that the C-terminal mutations are responsible for this inhibition (Kato et al. 2011) The previous work also showed that the differentiation block arose because of the suppressive activity of C/EBP $\alpha$  and not because of the reduced expression of the G-CSF-R protein downstream of C/EBP $\alpha$  (Kato et al. 2011).

Morphological analysis of granulocyte cultures from our mutant clones also revealed dysplastic cells. Surprisingly, a morphological analysis of myeloid cells obtained from the hiPSC- *CEBPA*<sup>+mut</sup> C5 and C12 clones revealed that approximately 50% of the aberrant myeloid cells were neutrophils with either hypo-granularity or pseudo-Pelger abnormality. Hypo-granularity is considered a particular dysplastic feature (Goasguen et al. 2014; Invernizzi et al. 2015), as the International Working Group on Morphology of MDS proposal for improving the criteria for dysgranulopoiesis defines neutrophils as dysplastic if they show at least a 2/3 reduction in granule content or the pseudo-Pelger abnormality of the nucleus. Both of these features were observed in both of our hiPSC-*CEBPA*<sup>+mut</sup> C5 and C12 clones. Although granulocyte colonies were not observed in the colony-forming experiment, granulocytes with abnormal morphology were detected in myeloid culture. One possible explanation for this discrepancy is that we scored colony formation after 14 days following the 12th day of HSPC, whereas myeloid liquid culture cells were evaluated on day 7. These aberrant morphologies may trigger cells to undergo apoptosis, so those cells may no longer have been present at the time of scoring.

Overall, these results are consistent with previously published results indicating that C/EBP $\alpha$  is required most specifically during the process of myeloid cell development (Pundhir et al. 2018). Our findings demonstrate the indispensable function of C/EBP $\alpha$  in neutrophil differentiation and maturation, as well as its susceptibility to dysregulation by oncogenic processes that control haematopoiesis stability. Generally, according to our data, we have

shown that the disruption of *CEBPA* in one allele behaves as a dominant negative regulator, as illustrated by the substantial impacts on myeloid cell development.

#### **4.3.4 *CEBPA*<sup>+/-mut</sup> affects the expression of TFs that regulate myeloid development**

Generating mature and functional blood cells from multipotent HSPCs demands the precise regulation of gene expression at each developmental phase. Specific transcription factors (TFs) are crucial to this process because they effect the expression of their target genes in a lineage-specific manner. Crucial transcriptional regulators have been discovered that play a significant role in establishing and sustain gene expression patterns during haematopoietic development. Several studies indicate that *C/EBPα* may function as a pioneer TF that operates in conjunction with other TFs, such as PU.1, RUNX1 and GATA2, to initiate the myeloid gene expression programme during the earliest stages of haematopoiesis (van Oevelen et al. 2015; Prange et al. 2014). Therefore, determining whether a monoallelic mutation of *CEBPA* would result in the dysregulation of these TFs (PU.1, GATA2, RUNX1) and impair myeloid development was an important focus of the present research.

*C/EBPα* has been shown to prime chromatin at myeloid-specific sites in collaboration with PU.1 and RUNX1 (Koschmieder et al. 2009; Hasemann et al. 2014; Avellino and Delwel 2017).

Our data revealed that *RUNX1* expression decreased in HSPCs obtained from our mutant clones. *C/EBPα* and RUNX1 interact in a complex and dynamic fashion to regulate haematopoietic differentiation. Decreases in *CEBPA* expression as a result of *RUNX1* deletion inhibits granulopoiesis and is the primary cause of the resultant myeloproliferative phenotype (Guo et al. 2012). This deregulation of *RUNX1* impairs myeloid commitment by reducing *CEBPA* expression, which could explain the blocking of myeloid differentiation in *CEBPA*<sup>+/-mut</sup> lines (Ptasinska et al. 2012). PU.1 plays an essential role in haematopoietic differentiation by interacting with *C/EBPα* through its C-terminal bZIP domain (Kato et al. 2011). The mutation

of *CEBPA* enhanced the reduced of *PU.1* expression in our hiPSC- *CEBPA*<sup>+mut</sup> clones, and this finding is supported by previous reports highlighting that both PU.1 and C/EBP $\alpha$  control granulocyte-monocyte differentiation (Pundhir et al. 2018). Additionally, ChIP-seq of C/EBP $\alpha$  data obtained from previous study revealed an unexpected role in modulating the PU.1 function, while the absence of C/EBP $\alpha$  caused an early GM lineage differentiation block upstream of the CMP/pre-GMP transition (Zhao et al. 2022).

GATA2 has also been reported to be a critical haematopoietic transcription factor important for the formation and lineage output of HSPCs, mainly by myelopoiesis (Gao et al. 2013; de Pater et al. 2013). Cooperative interactions between GATA2 and C/EBP $\alpha$  have been demonstrated to play a vital role in eosinophil commitment (Nerlov 2007; Iwasaki et al. 2006; Koschmieder et al. 2009). Our data on *GATA2* expression are in agreement with the findings of previous studies that reported an involvement of both C/EBP $\alpha$  and GATA2 in the regulation of myeloid progenitor proliferation and differentiation and that mutations in both genes can cause individuals to develop AML (Taskesen et al. 2011; Di Genua et al. 2020).

By contrast, the downregulation of myeloid genes in hiPSC- *CEBPA*<sup>+mut</sup> clones revealed the enrichment of several erythroid-specific genes, such as *EPO-R* and *TFRC*. Our finding is in agreement with a previous study by the Nerlov group, who found an impartially uniform effect of C/EBP $\alpha$  mutation across the LT-HSC compartment, meaning that the loss of C/EBP $\alpha$  function could lead to changes in lineage priming in addition to changes at the single-cell level (Bereshchenko et al. 2009). Despite the fact that normal C/EBP $\alpha$  role is important for eliciting normal myeloid differentiation (Tenen, 2003), previous studies have shown that haematopoietic cells that are impaired in C/EBP $\alpha$  function are shifted towards erythroid differentiation (Mannelli et al., 2017).

The increase in the expression of erythroid genes aligns with the work of others that showed increasing levels of *GATA-1* expression in haematopoietic precursor cells from

*C/EBPα*<sup>-/-</sup> foetal livers (Wagner et al., 2006). GATA1 is tightly controlled, and its levels increase during erythropoiesis (Iwasaki et al. 2003). An assessment of *GATA-1* in our hiPSC-*CEBPA*<sup>+mut</sup> clones could help to confirm this increase in erythroid markers. Another possibility that could explain the increased expression of the erythroid genes is the reduced expression of *PU.1*, as PU.1 binds GATA-1 to prevent erythropoiesis (Zhang et al. 1999). Reducing *PU.1* RNA in response to *CEBPA*<sup>+mut</sup> would increase the potential to induce erythropoiesis.

In conclusion, our findings demonstrate that adequate *CEBPA* levels are required to initiate the myeloid programme critical for producing GMPs (Avellino and Delwel 2017), and that C-terminal or dysregulated expression mutations might be a key to the irregular granulopoiesis state. More techniques, such as ChIP-Seq, will help address this issue. The use of this technology, in combination with *in vitro* and *in vivo* reperch studies, should provide a helpful platform for delineating the function of *CEBPA* in HSPCs.

#### **4.3.5 *CEBPA*<sup>+mut</sup> does not block erythroid differentiation but promotes dyserythropoiesis**

We demonstrated that heterozygous mutations that disrupt the DBD in *CEBPA* influence lineage commitment of HSPCs. The qPCR results showed reduced myeloid gene expression and elevated expression of genes involved in erythroid differentiation. Consequently, our findings reveal the possibility of a relationship between the expression of TFs involved in HSC programs and their lineage potential. Flow cytometry analysis used to detect erythroid differentiation revealed the interesting finding that no immunophenotypic differences existed between the mutants and the hiPSC controls during erythroid development. Nevertheless, the possibility of evaluating the relevant markers for addressing the abnormal erythroid phenotypes, such as CD34, CD117 and CD105, could improve the assessment accuracy (Westers et al. 2017).

The use of flow cytometry alone is often inadequate to achieve a conclusive diagnosis. Our findings indicate that morphological analysis should also be integrated. Comparing the



morphology of the hiPSC- *CEBPA*<sup>+mut</sup> cells to that of the hiPSC control cells revealed the morphological features of nucleated erythrocytes or immature cells. Based on the difficulty of finding dysplastic characteristics in mature cells, which requires specialised knowledge and has been documented previously (Mufti et al. 2008), we decided to score the aberrant morphology no later than day 7 of differentiation. This revealed dysplastic erythropoiesis in the mutant clones, with approximately 40–50% of the cells on day 4 and 30% on day 7, displaying abnormalities. Consequently, we hypothesise that the increase in the percentage of cells that express CD71<sup>+</sup> CD235a<sup>+</sup> corresponds to the presence of abnormal cells in the culture. Overall, *CEBPA*<sup>+mut</sup> could impact erythroid function by producing dysplasia that affects erythropoiesis. However, further research is needed to determine how *CEBPA*<sup>+mut</sup> affects the erythroid lineage. This may involve analysing the effects of these mutations on the gene expression patterns (determined by RNA sequencing), protein-protein interactions, and transcriptional regulatory networks operating in erythroid cells. Live cell imaging techniques, such as confocal or time-lapse microscopy, could also help to image live cells and enable the visualisation of dynamic events occurring inside living cells. Tracking and measuring changes in cell morphology and other cellular behaviours related to aberrant morphology would be feasible by collecting time-lapse images of cells during differentiation.

## **Chapter 5: Investigating the role of MYBL2 in haematological disorders**

## 5.1 Introduction

A parallel body of work carried out during this research was related to investigating the effect of low levels of *MYBL2* on MDS. *MYBL2* is a gene located in the long arm of chromosome 20, in a region known to be deleted in 5–10% of myeloid malignancies, including AML and MDS (Brunning 2003; Okada et al. 2012). The presence of del(20q) as a secondary genetic mutation combined with other abnormalities is associated with a poor prognosis, leading to increased risk of AML (Odenike et al. 2011). Thus, one or more genes contained within del(20q) should act as a tumour suppressor gene, the deletion of which promotes disease progression. In del(20q), the minimum common deleted region contains nine genes, five of which are expressed in HSCs, and one of the genes is the transcription factor *MYBL2* (Clarke et al. 2013). According to previous studies, including work from our lab, 50–60% of patients with MDS with a high risk of progressing to AML express significantly lower *MYBL2* levels than healthy controls, regardless of del(20q) cytogenetic abnormality (Clarke et al. 2013; Heinrichs et al. 2013). Mouse models expressing low levels of *MYBL2* display MDS-like disease during ageing (Clarke et al., 2013; Heinrichs et al., 2013). Moreover, patients with MDS showing low *MYBL2* levels exhibit a defective DNA repair signature (Bayley et al., 2018), and changes in the DNA damage response in HSCs have been associated with the appearance of numerous blood diseases (Mohrin et al. 2010).

All the evidence presented above supports the involvement of *MYBL2* as a tumour suppressor gene within del(20q). Nevertheless, whether the sole disruption of *MYBL2* in HSCs will result in deficient haematopoiesis and MDS-like characteristics is outstanding. We hypothesise that reducing the level of this tumour suppressor could affect erythroid differentiation. The overarching aim of this chapter was to determine the connection between *MYBL2* deficiency and lineage dysplasia present in patients with MDS. To do so, we decided to perform our studies in two systems: (i) HSPCs derived from hiPSCs harbouring *MYBL2* haploinsufficiency

(generated by CRISPR-Cas9)—to study their possible differentiation towards myeloid and erythroid lineages; and (ii) immortalised myeloid cell lines in which *MYBL2* was downregulated by shRNA to study their differentiation to the erythroid lineage following induction by cytarabine (Ara-C).

## 5.2 Result:

### 5.2.1 Generation of stable hiPSC with low MYBL2 expression

We were keen to alter MYBL2 levels to understand and explore its role in MDS and to assess the connection between *MYBL2* haploinsufficiency and lineage dysplasia that arises in patients with MDS. CRISPR-Cas9 technology was utilised in the same way as explained in Chapter 3 to generate stable cells with *MYBL2* indels, which would result in a reduction in MYBL2 protein levels. The sgRNA was designed using online tools and selected to target Exon 5 of *MYBL2*. Exon 5 was chosen for two reasons: first, it is the longest exon and forms part of the DBD, and second, previous work conducted in our lab revealed that altering DBD in a mouse model produced the MDS phenotype.



To enable cloning of the gRNA into pSpCas9(BB)-2A-GFP (PX458) (Addgene) at the BbsI restriction site, the gRNA oligonucleotide sequence was altered by adding CACC before the forward complement of the guide and AAAC before the reverse complement of the guide.

(Figure 5.1) (Ran et al. 2013).

Then, sgRNA primers were annealed and cloned into the PX458 vector (Addgene), as described in the chapter 3 (3.2.1). The clonal selection was performed after transfection into the healthy iPSC line BU3.10. Eighteen iPSC clones were individually picked and expanded, and genomic DNA was extracted from five to evaluate the gene-editing efficiency of CRISPR-Cas9. Genomic DNA from BU3-10 iPSCs was also extracted as a control validation of the modified region by CRISPR, which was conducted by digesting the DNA with the T7EI endonuclease and amplifying the *MYBL2* genomic DNA fragment (420 bp) by PCR using primers flanking the area of interest (see section 2.16). The agarose gel electrophoresis data demonstrated successful PCR amplification of a 420 bp DNA fragment in the five selected clones (Figure 5.2A). After the T7EI endonuclease reaction, no changes in the size of the DNA fragment (420 bp) were observed, indicating that none of the clones exhibited a heterozygous

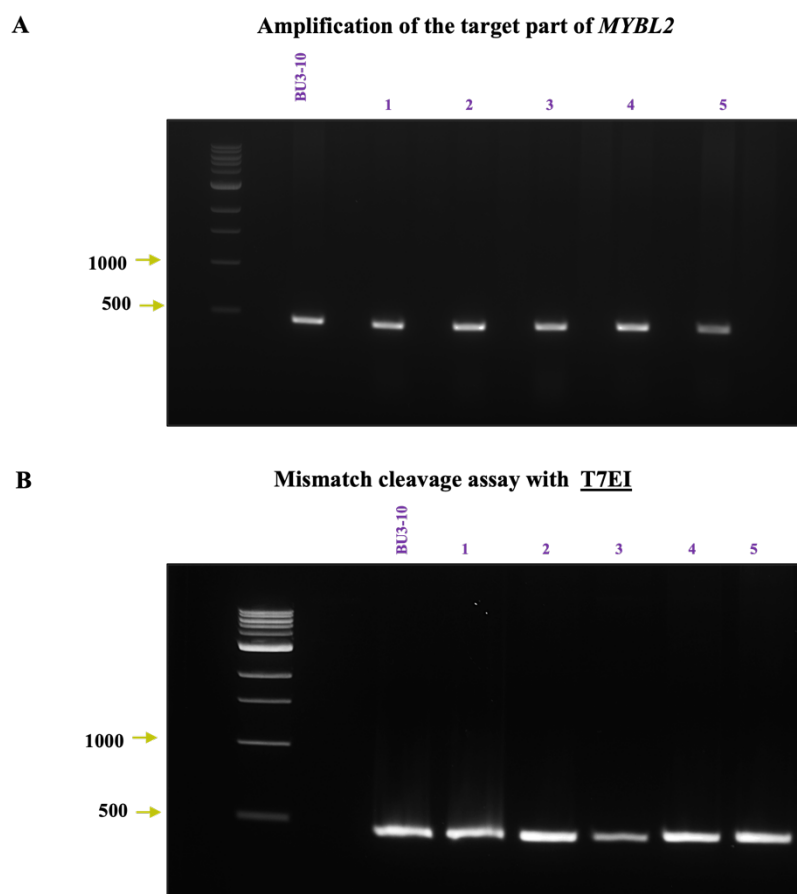
mutation (Figure 5.2B). Our data indicated that despite the clonal selection based on the expression of the GFP reporter gene, the isolated clones did not contain indels generated by CRISPR-Cas9 in the desired region.

```
TCAGGTGGATGTGAAGGGCTATGGTGGGGGCGAGAGAGGGGCCGGTGCAAGTGGC
TACCCAGAGACTTTTCTCTAAGGTCATCGAGCTGGTTAAGAAGTATGGCACAAAGC
AGTGGACACTGATTGCCAAGCACCTGAAGGGCCGGCTGGGGAAGCAGTGCCGTGA
ACGCTGGCACAACCACCTCAACCCTGAGGTGAAGAAGTCTTGCTGGACCGAGGAG
GAGGACCGCATCATCTGCCAGGCCACAAGGTGCTGGGCAACCGCTGGGCCGAGA
TCGCAAGATGTTGCCAGGGAGGTAAGCTGTCTTCTTGGGGGTTGGGACAGGTTCC
CGGGAGGCCAGGCCCGTGTTTCTGATGGAGGAGGGTTCCTGGGTGGGTATTGTGGT
CCTGTGCTCTTGTTCTGTAAACCCAGCCTCGGCTCCAGGGCTCCAGAGGCATGC
ATAGTCAGGAGGAAGGGTG
```

 Primers to amplify target region  
 gRNA position

**Figure 5.1 MYBL2 sequence.**

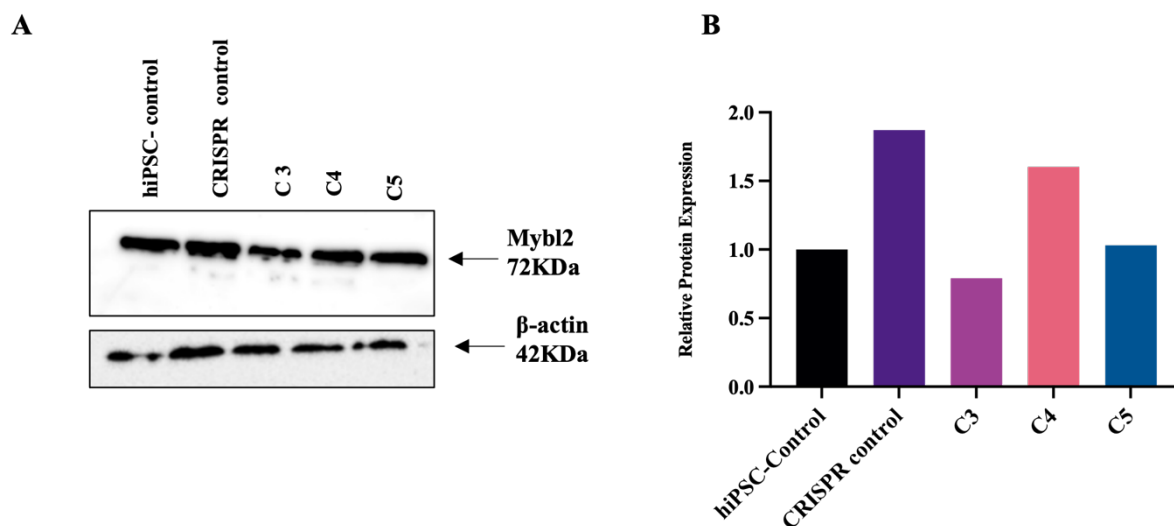
Snapshot of *MYBL2* sequence. The position of gRNA (green) and primers used to amplify the targeted region (yellow) are highlighted.



**Figure 5.2 T7E1 mismatch cleavage assay.**

**A)** Following CRISPR-Cas9 editing, gDNA was extracted from hiPSC (BU310), and *MYBL2* CRISPR clones were selected. PCR was performed to amplify a DNA fragment flanking the target region (420 bp) using primers designed around the targeted area. Agarose gel electrophoresis was used to visualise the PCR results. **(B)** T7E1 was used to digest the PCR products from step A, and digestion reactions were detected using agarose gel. The size of the WT allele (420 bp) and the mutant allele band were undetected.

The T7E assay has been reported to be particularly inefficient when minor changes occur (Sentmanat et al. 2018). Consequently, the DNA from selected clones were subjected to Sanger sequencing in order to identify any possible of point mutation that may have been introduced by the CRISPR-Cas9 modification in the targeted area. The results obtained from Sanger sequencing demonstrated none of each of these clones are exhibited deletions in the *MYBL2* gene in the target area (Data not shown). Thus, despite the lack of indels identified by T7E assay, or Sanger sequencing, we decided to determine whether, in any of the clones generated, the levels of MYBL2 protein were affected as a consequence of CRISPR by comparing the clones with the parental BU3.10 control line. Western blotting analysis showed that the isolated clones displayed similar MYBL2 protein levels (Figure 5.3A, B), corroborating that our CRISPR-Cas9 approach failed. Thus, we decided to change the experimental system to gain some insight into the possible role of MYBL2 in erythropoiesis.



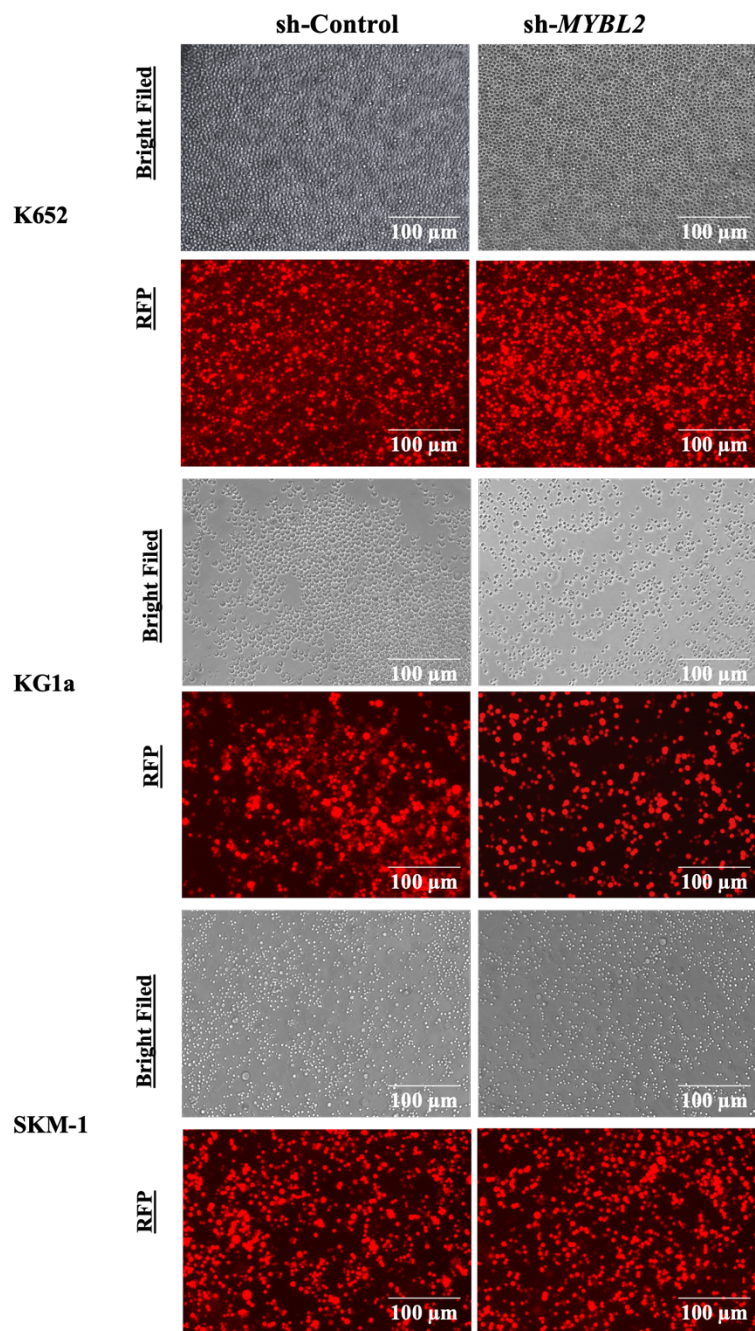
**Figure 5.3 MYBL2 protein expression detection in hiPSC**

(A) Western blot analysis showing the expression of MYBL2 protein in BU310, CRISPR control (BU310+PX458), and selected clones. A total of 50 µg of protein from the cell line was subjected to SDS-PAGE, followed by detection using antibodies against MYBL2. The right side of the blot represents the relevant molecular weight standards for the protein. (B) A quantification assay of MYBL2 protein expression was performed by normalising to β-actin. Identifying the band intensity in ImageJ software.



### 5.2.2 Downregulation of MYBL2 by shRNA in leukaemia cell line

Given the lack of success in generating iPSC lines expressing lower levels of *MYBL2* by CRISPR-Cas9, we decided to change the approach and utilise a lentivirus containing doxycycline-inducible *MYBL2* shRNA (pTRIPZ-sh*MYBL2*). Transduction of healthy iPSCs with pTRIPZ lentivirus was unsuccessful; either the cells were challenging to transduce, or the cells transduced died, as no positive colonies were obtained after two attempts. Thus, to gain some insight into the function of low *MYBL2* in haematopoiesis in a human cell line setting, we decided to use the pTRIPZ-sh*MYBL2* lentivirus (and pTRIPZ shControl) and transduce human cell lines representing different myeloid malignancies: K562 cells (chronic myeloid leukaemia), KG1a cells (AML), and SKM-1 cells (MDS/AML). This lentivirus (pTRIPZ) also contains a doxycycline-inducible reporter gene red fluorescent protein (RFP), which helps determine the extent of induction. K562, KG1a, and SKM-1 leukaemia cell lines transduced with pTRIPZ-sh*MYBL2* and pTRIPZ-shControl were selected with puromycin (1 µg/ml). After selection, the addition of doxycycline (dox) for 24 h confirmed that the infection had been successful, as the cells expressed the reporter protein RFP (Figure 5.4).



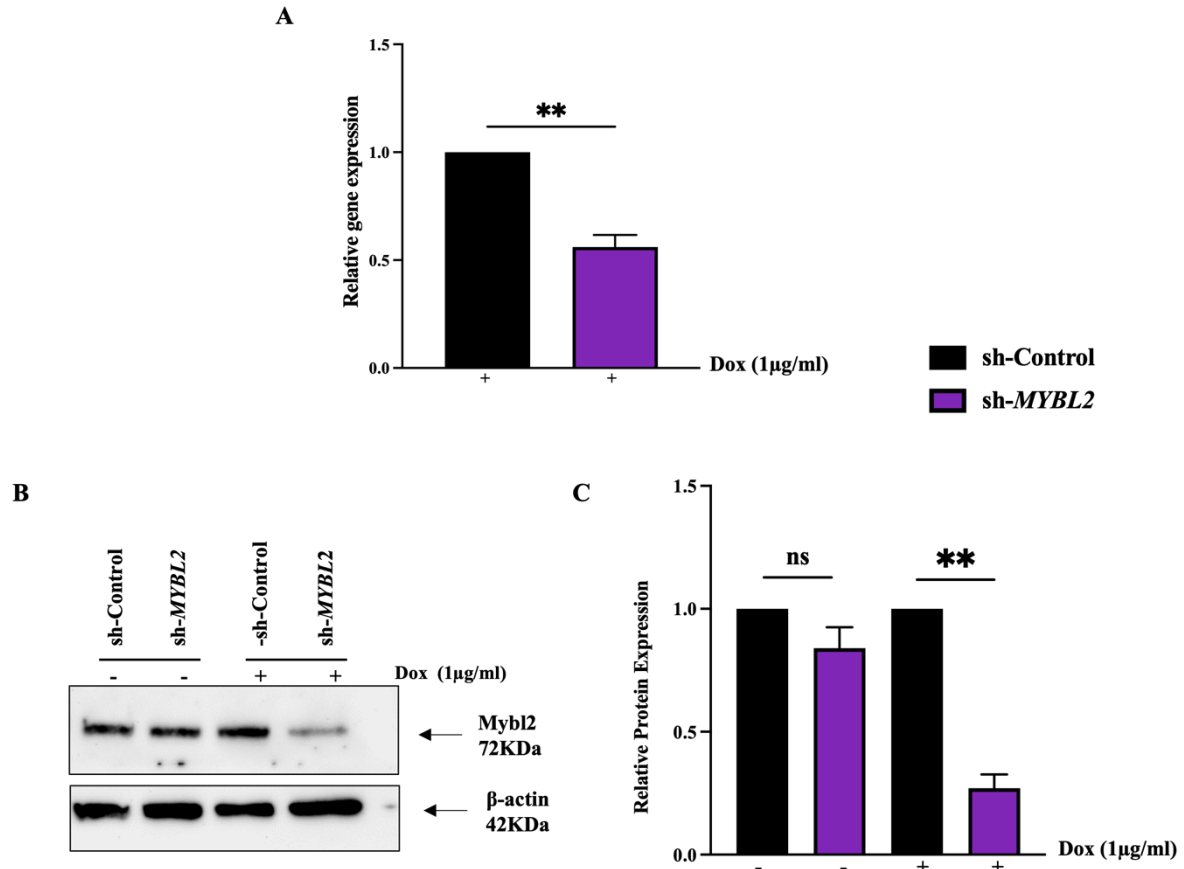
**Figure 5.4 Knockdown of MYBL2 in K562, KG1a, and SKM-1.**

The fluorescence of the cells at 3 days following transfection with the shRNA-Control or sh-MYBL2 in the presence of 1 $\mu$ g/ml Doxycycline for 24 h pictures captured was assessed using a fluorescence microscope. Scale bar: 100  $\mu$ m.

We performed quantitative real-time PCR (qPCR) and western blotting to evaluate the knockdown efficiency. Cells transduced with pTRIPZ-sh*MYBL2* and pTRIPZ-shControl were cultured in the presence or absence of 1 µg/ml dox for 3 days. *MYBL2* mRNA expression in K562 sh-*MYBL2* lines was around 60% reduced compared to the K562-Control cells (Figure 5.5A). Western blotting was performed to determine whether this reduction in mRNA levels translated into a reduction at the protein level. *MYBL2* protein expression was reduced by 35% to 40% compared to the controls (Figure 5.5B, C).

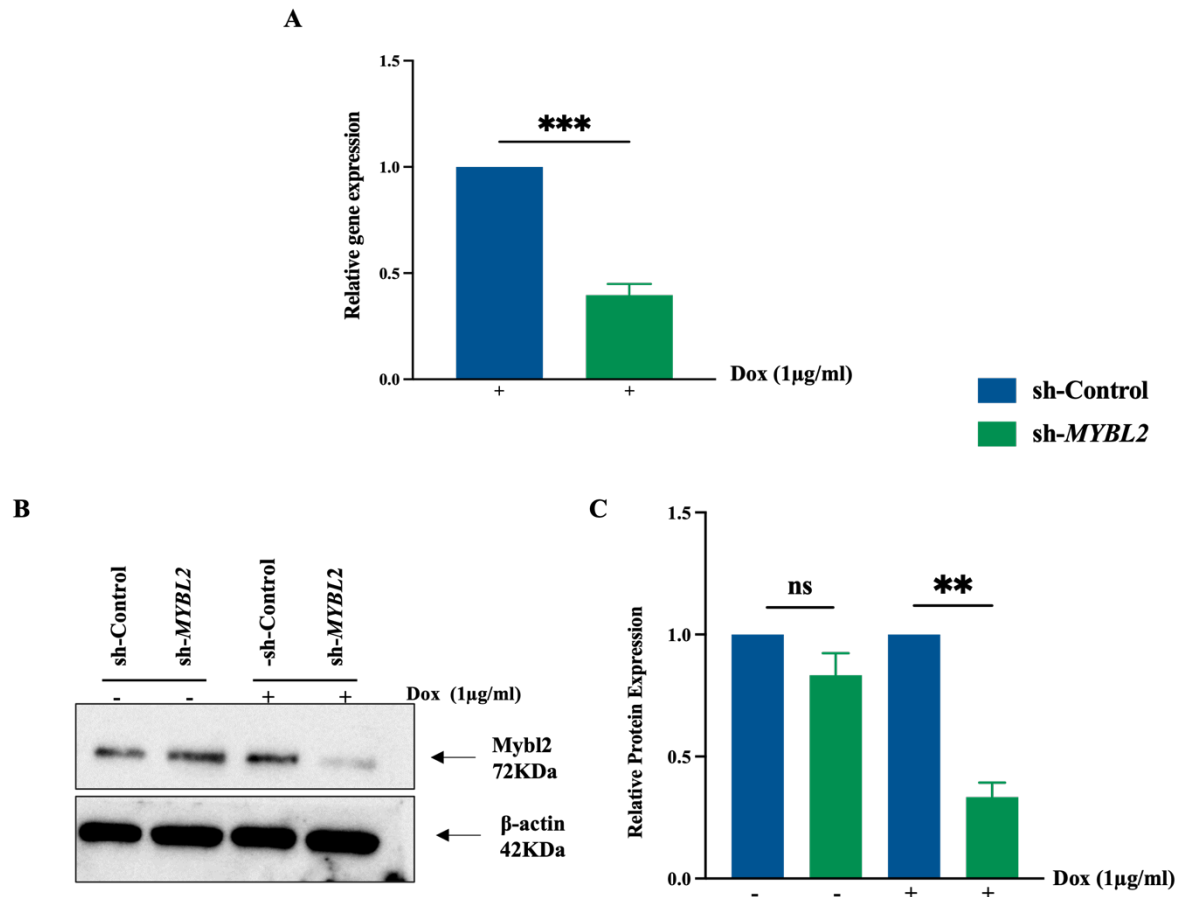
Similar to K562, *MYBL2* mRNA in KG1a cells and SKM-1 cells transduced with pTRIPZ-sh*MYBL2* displayed approximately a 50% decrease (Figure 5.6A) and a 40% decrease (Figure 5.7A), respectively, upon dox treatment. *MYBL2* protein levels in KG1a and SKM-1 cells transduced with pTRIPZ-sh*MYBL2* were also analysed after 3 days of dox treatment. A 50% reduction in *MYBL2* protein expression in KG1a cells and SKM-1 was observed compared to the corresponding controls (Figure 5.6B, C) (Figure 5.7B, C).

In conclusion, our results confirmed the generation of three different human cell lines where *MYBL2* downregulation could be induced for further functional studies.



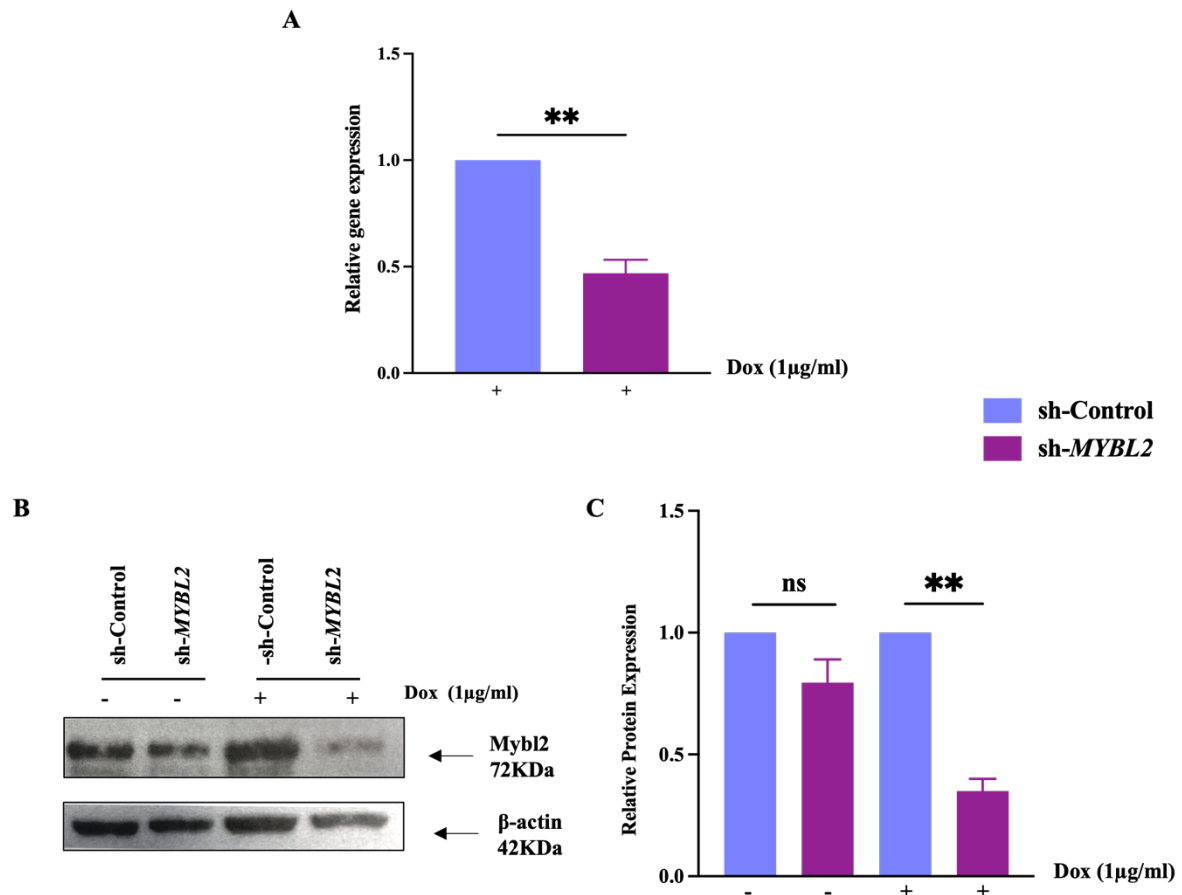
**Figure 5.5 Knockdown of MYBL2 in K562.**

(A) Histogram showing the relative transcript level of *MYBL2* compared to the control. The relative expression was calculated by normalising to the GAPDH housekeeping gene and then expressed relative to sh-Control. The results are presented as the mean  $\pm$  SEM and analysed using an unpaired t-test.  $**p < 0.01$ . The data represent three independent experiments. (B) Western blot analysis showing the expression of MYBL2 protein. Lane 1 and 2 are sh-Control and sh-MYBL2, both untreated. Lanes 3 and 4 are sh-Control and sh-MYBL2 cells, respectively, treated with 1  $\mu$ g/ml dox for 72 h. The right side of the blot indicates the relevant protein molecular weight standards. (C) A quantification assay of MYBL2 protein expression was performed by normalising to  $\beta$ -actin and sh-MYBL2 relative to sh-Control and sh-MYBL2 + 1  $\mu$ g dox relative to sh-Control + 1  $\mu$ g dox, and band intensities were identified using ImageJ software. The statistical results are presented as the mean  $\pm$  standard error of the mean (SEM) and analysed using an unpaired t-test.  $P$ -value = ns (no significant)  $**p < 0.01$ . The data represent three independent experiments.



### Figure 5.6 Knockdown of MYBL2 in KG1a.

(A) Histogram showing the relative transcript level of *MYBL2* compared to the control. The relative expression was calculated by normalising to the *GAPDH* housekeeping gene and then expressed relative sh -Control. The results are presented as the mean  $\pm$  SEM and analysed using an unpaired t-test.  $***p < 0.001$ . The data represented three independent experiments. (B) Western blot analysis showing the expression of MYBL2 protein. Lanes 1 and 2 are sh-Control and sh-MYBL2, both untreated. Lanes 3 and 4 are sh-Control and sh-MYBL2 cells, respectively, treated with 1  $\mu$ g/ml dox for 72 h. The right side of the blot indicates the relevant protein molecular weight standards. (C) A quantification assay of MYBL2 protein expression was performed by normalising to  $\beta$ -actin and sh-MYBL2 relative to sh-Control and sh-MYBL2 + 1  $\mu$ g dox relative sh-Control + 1  $\mu$ g dox. Band intensities were identified using ImageJ software. The statistical results are presented as the mean  $\pm$  standard error of the mean (SEM) and analysed using an unpaired t-test. *P*-value ns (no significant)  $**p < 0.01$ . The data represent three independent experiments.

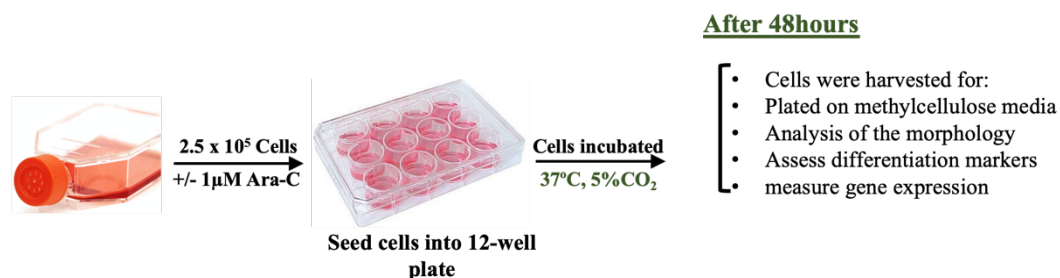


### Figure 5.7 Knockdown of MYBL2 in SKM-1.

(A) Histogram showing the relative transcript level of *MYBL2* compared to the control. The relative expression was calculated by normalising to the *GAPDH* housekeeping gene and then expressed relative to sh-Control. The results are presented as the mean  $\pm$  SEM and analysed using an unpaired t-test.  $***p < 0.001$ . The data represented three independent experiments. (B) Western blot analysis showing the expression of MYBL2 protein. Lanes 1 and 2 are sh-Control and sh-MYBL2, both untreated. Lanes 3 and 4 are sh-Control and sh-MYBL2 cells, respectively, treated with 1  $\mu$ g/ml dox for 72 h. The right side of the blot indicates the relevant protein molecular weight standards. (C) A quantification assay of MYBL2 protein expression was performed by normalising to  $\beta$ -actin and sh-MYBL2 relative to sh-Control and sh-MYBL2 + 1  $\mu$ g dox relative to sh-Control + 1  $\mu$ g dox. Band intensities were identified using ImageJ software. The statistical results are presented as the mean  $\pm$  standard error of the mean (SEM) and analysed using an unpaired t-test. *P*-value ns (no significant)  $**p < 0.01$ . The data represent 3 independent experiments.

### 5.2.3 Studying the effect of Ara-C on cell proliferation

After the successful generation of cell lines with modified MYBL2 protein levels, we were keen to determine whether a lower level of MYBL2 could affect erythroid differentiation by using differentiation-inducing agents that are safe for human use at effective concentrations. Ara-C is a therapeutic agent for myeloid leukaemia treatment and is known to induce erythroid differentiation in patients with myeloid malignancies (Hiddemann, 1991). Moreover, Ara-C has been shown to block proliferation and induce myeloid or erythroid differentiation of K562, KG1a, and SKM-1 cells (Gañán-Gómez et al. 2014). Thus, we decided to culture our cells in the presence and absence of Ara-C and determine (i) whether we could reproduce the work of others (Gañán-Gómez et al. 2014) and (ii) whether MYBL2 downregulation would affect the proliferation and differentiation capacity of the cells. Hence, the cells from each condition were grown exponentially, and 1  $\mu\text{g}/\text{ml}$  doxycycline was added for 72 h. We then cultured  $2.5 \times 10^5$  cells from each cell line in the presence or absence of 1  $\mu\text{M}$  Ara-C for 48 h, as a previous study reported that this Ara-C concentration has no effect on cell viability (Gañán-Gómez et al. 2014). The treated cells were harvested for RNA extraction or cultured methylcellulose semi-solid media to assess proliferation, or for flow cytometry to determine differentiation capacity (Figure 5.8).

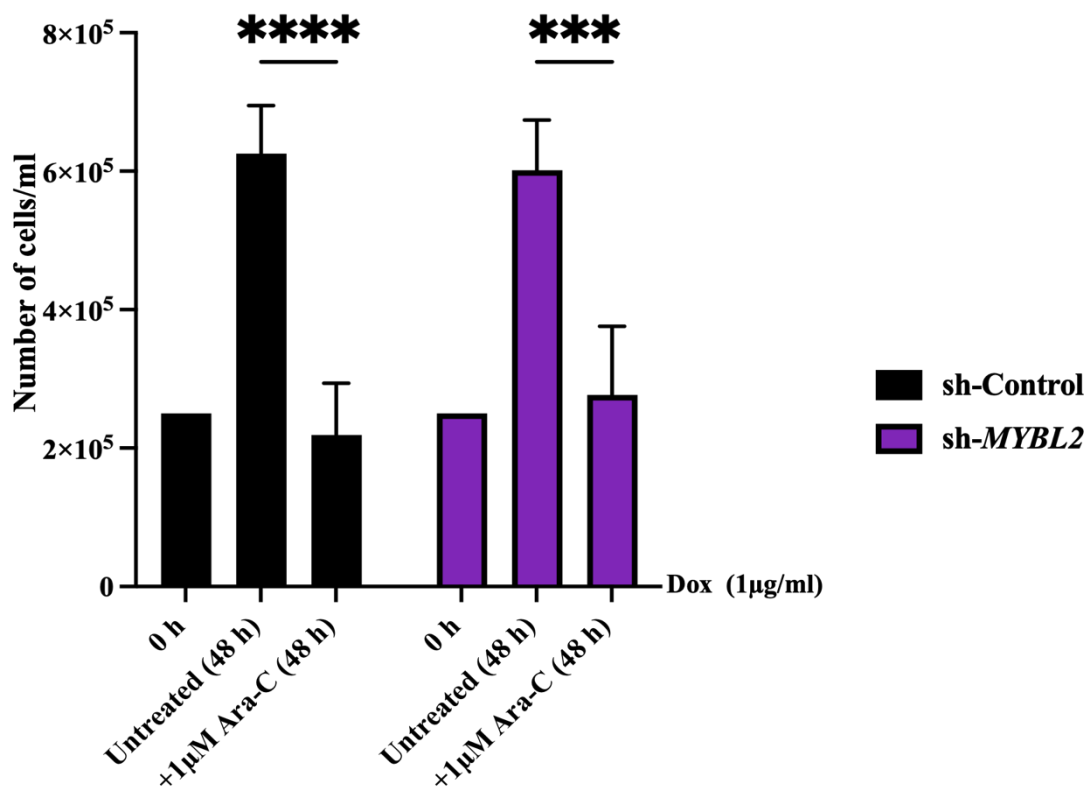


**Figure 5.8 Schematic representations of inducing K563, KG1a, and SKM-1 cells with Ara-C.**

Cells were cultured in the presence or absence of 1  $\mu\text{M}$  Ara-C for 48 h before harvesting to assess the effect of Ara-C on erythroid and myeloid differentiation.

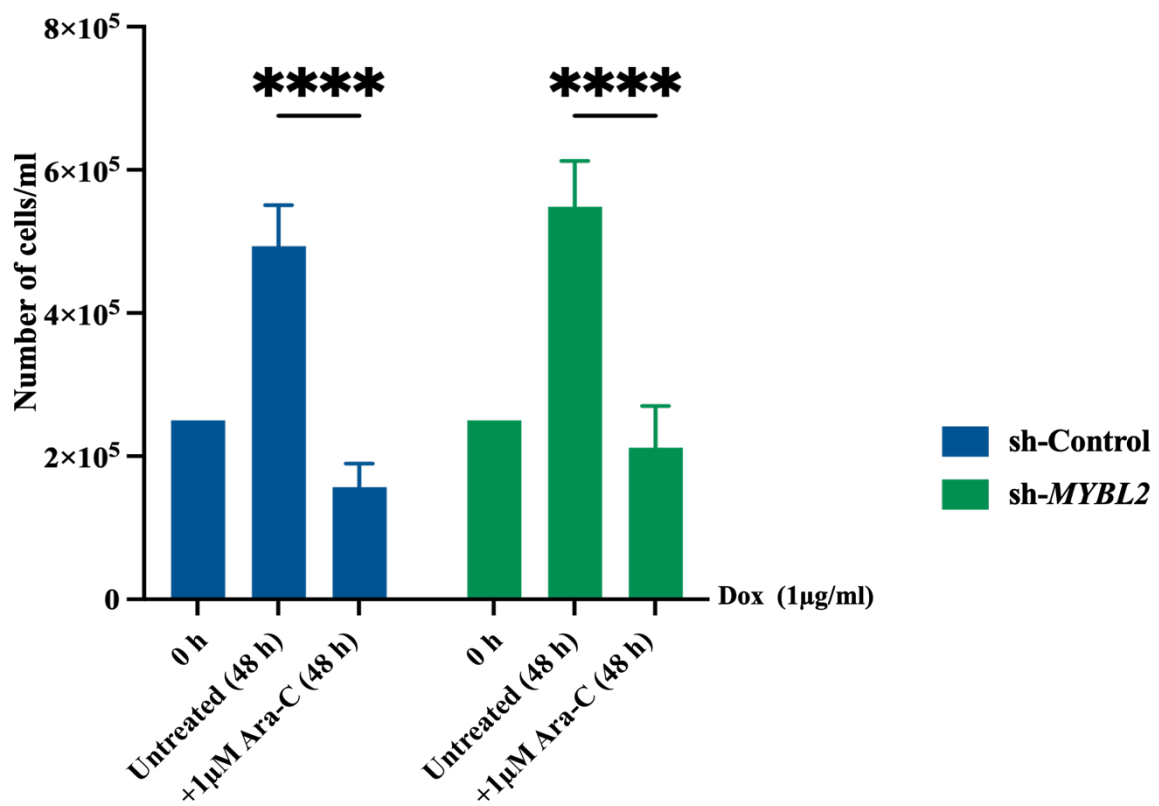
Since the scope of this chapter includes investigating the effect of Ara-C as a chemotherapeutic drug on leukaemia cell lines, the effective concentrations of Ara-C treatment on cell numbers were monitored using trypan blue exclusion. The sensitivity of these cell lines to treatments with 1  $\mu$ M Ara-C was assessed after 48 h. Untreated cells showed increased cell numbers after 48 h in culture, corresponding to proliferating cells. By contrast, cells cultured in the presence of Ara-C did not proliferate, with cell numbers similar to those at the start of the experiment (Time 0). Additionally, no changes were observed in the K562 cells with sh-Control or sh-*MYBL2* (Figure 5.9). Consistent with the results obtained for K562, KG1a cells and SKM-1 cells behaved very similarly in the presence of Ara-C, with cells cultured in the presence of Ara-C showing a block in proliferation and with no changes between the control and sh-*MYBL2* cell lines. As we did not observe an increase in dead cells with trypan blue manual counting, BrdU assays can be further used to validate the blocking of proliferation (Figure 5.10 & Figure 5.11).





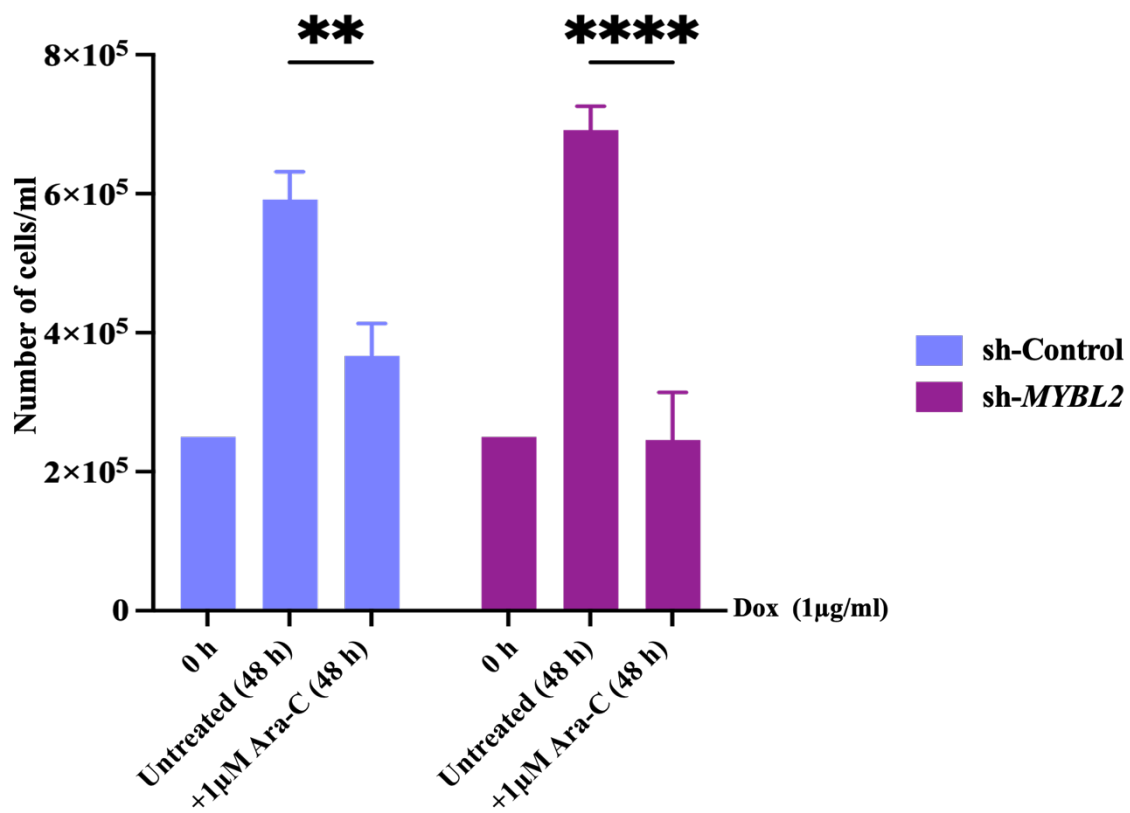
**Figure 5.9 Ara-C treatment affects the proliferation of K652 cells.**

The cell density of K562 sh-Control and sh-MYBL2. Cells were treated with 1 µM Ara-C for 48 h in the presence of 1 µg/ml dox. The data are presented as the mean ± standard error of the mean (SEM) and analysed using two-way ANOVA with multiple comparisons, \*\*\* $p < 0.001$  and \*\*\*\* $p < 0.0001$ . The data represent three independent experiments.



**Figure 5.10 Ara-C treatment affects the proliferation of KG1a cells.**

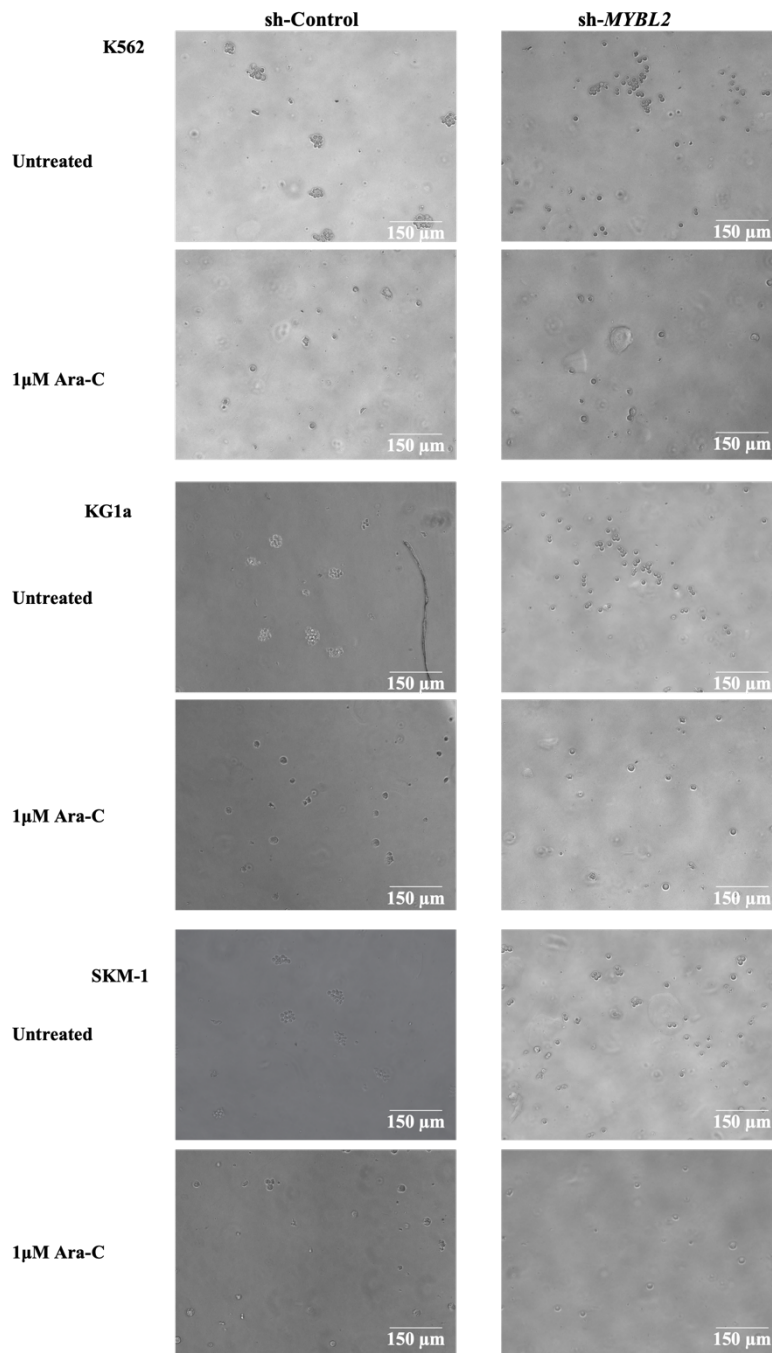
The cell density of KG1a sh-Control and sh-MYBL2. Cells were treated with 1 µM Ara-C for 48 h in the presence of 1 µg/ml dox. The data are presented as the mean ± standard error of the mean (SEM) and analysed using two-way ANOVA with multiple comparisons, \*\*\*\* $p < 0.0001$ . The data represent three independent experiments.



**Figure 5.11 Ara-C treatment affects the proliferation of SKM-1 cells.**

The cell density of SKM-1 sh-Control and sh-MYBL2. Cells were treated with 1 µM Ara-C for 48 h in the presence of 1 µg/ml dox. The data are presented as the mean ± standard error of the mean (SEM) and analysed using a two-way ANOVA with multiple comparisons, \*\* $p = 0.0037$  and \*\*\*\* $p < 0.0001$ . The data represent three independent experiments.

Next, we sought to investigate the effect of lowering MYBL2 on cell proliferation, accompanied by the Ara-C-induced cytotoxicity in these cell lines. Therefore, 500 cells from each condition were seeded on methylcellulose media for 3 days with 1  $\mu\text{g/ml}$  doxycycline. The media type and time of adding the drug to the cells were chosen following previously published literature (Gañán-Gómez et al. 2014). As depicted in Figure 5.12, low levels of MYBL2 did not affect cell proliferation, although the cells were not as tightly compact as the control cells. By contrast, 1  $\mu\text{M}$  Ara-C concentration reduced cell proliferation in sh-Control and sh-MYBL2. Overall, our findings suggest that low levels of MYBL2 did not affect cell numbers or cell proliferation, whereas 1  $\mu\text{M}$  Ara-C significantly reduced cell proliferation.



**Figure 5.12 Ara-C affects cell proliferation.**

Phase contrast of colonies formed in cells K562, KG1a, and SKM-1 harbouring downregulation of MYBL2 in the presence of 1 μg/ml dox and the presence or absence of 1 μM Ara-C. Cells were plated in methylcellulose semi-solid media for 3 days. Scale bar: 150 μm; 10× magnification, representative images of n=3

#### 5.2.4 Ara-C induces erythroid differentiation

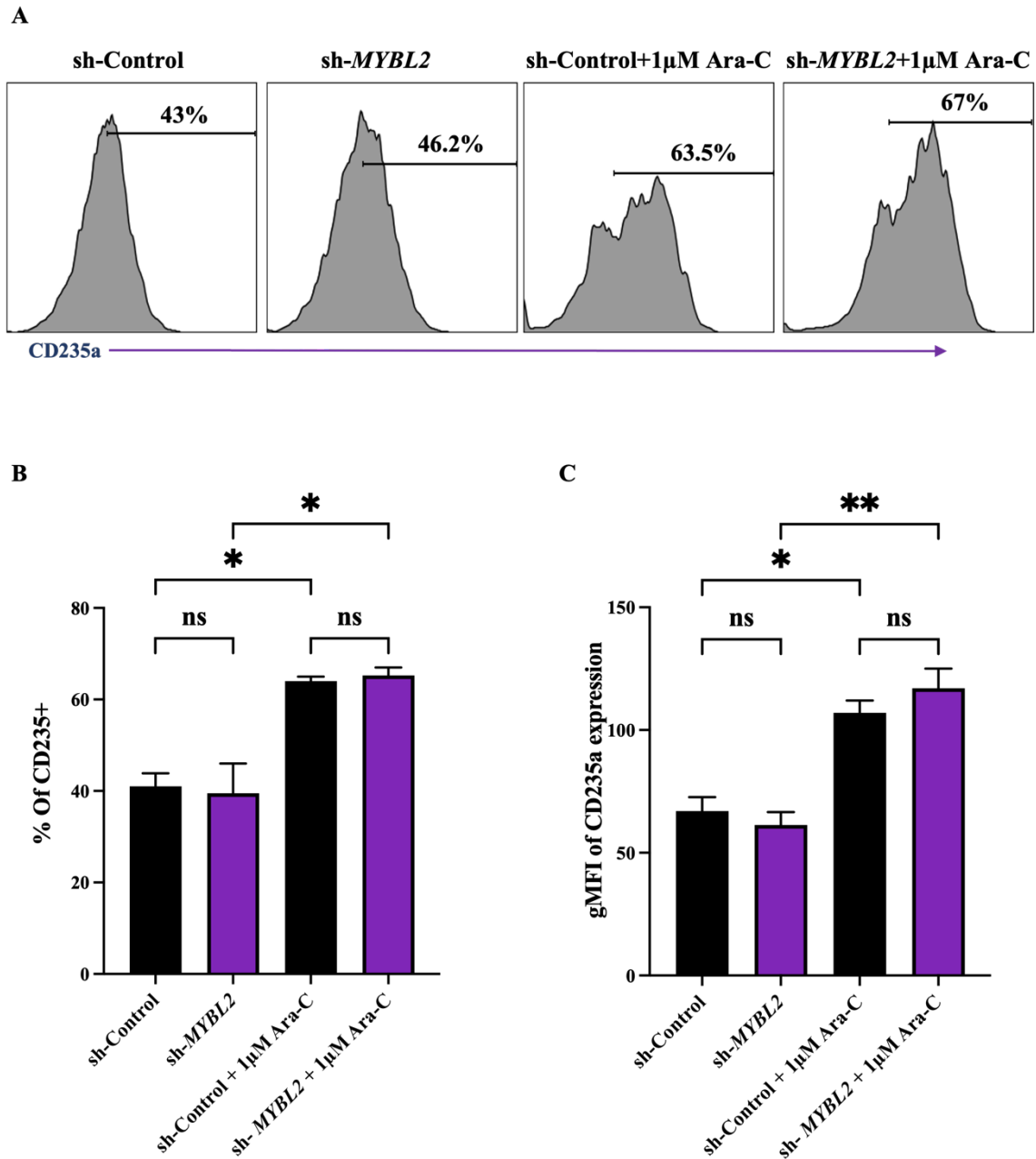
As mentioned above, Ara-C treatment has been reported to block proliferation and induce erythroid differentiation of leukaemia cells (Tomic et al. 2022). Our data also suggest that Ara-C treatment induces a block in proliferation, regardless of MYBL2 protein levels. Thus, we next sought to determine whether Ara-C could trigger cell differentiation. For this purpose, the proportion of cells differentiating into erythroid lineages was analysed. Cells treated with 1  $\mu$ M Ara-C were harvested after 48 h and stained with erythroid markers (CD235a<sup>+</sup>) only due to time constrain in the lab and technical issues. The expression of CD235a<sup>+</sup> was evaluated by flow cytometry. The phenotypic analysis revealed that a significant proportion of K562 sh-Control and K562 sh-*MYBL2* express the mature erythroid marker CD235a<sup>+</sup> cells at 43% and 46%, respectively, and that this expression increases upon Ara-C treatment in both K562 sh-Control and K562 sh-*MYBL2* (63.5% and 67%, respectively) (Figure 5.13A, B).

Additionally, the geometric fluorescent intensity (gMFI) of CD235<sup>+</sup> expression was determined to assess the intensity of CD235<sup>+</sup> expression at the terminal stages of erythroid differentiation. As displayed in Figure 5.13C, the intensity of CD235 expression was significantly higher in treated cells compared to untreated sh-Control and sh-*MYBL2*, whereas no difference was shown between sh-Control and sh-*MYBL2*.

The flow cytometry analysis further indicated that treatment of KG1a sh-Control and sh-*MYBL2* with Ara-C induced erythroid differentiation, as observed by the higher percentage of erythroid cells, 66% and 73%, compared to approximately 20% in both KG1a sh-Control and sh-*MYBL2*, respectively (Figure 5.14 A, B). The geometric mean gMFI of CD235 expression was significantly higher in treated cells compared with untreated sh-Control and sh-*MYBL2*. By contrast, no difference was shown between sh-Control and sh-*MYBL2* (Figure 5.14 C). Furthermore, SKM-1 cells treated with Ara-C did not respond the same as K562 and KG1a, showing limited potential to differentiate towards erythroid lineages, with only 2.5% and

1.45% CD235a<sup>+</sup> in sh-Control and sh-*MYBL2*, respectively. Nonetheless, the same effect was found in both sh-Control and sh-*MYBL2* cells (Figure 5.15).

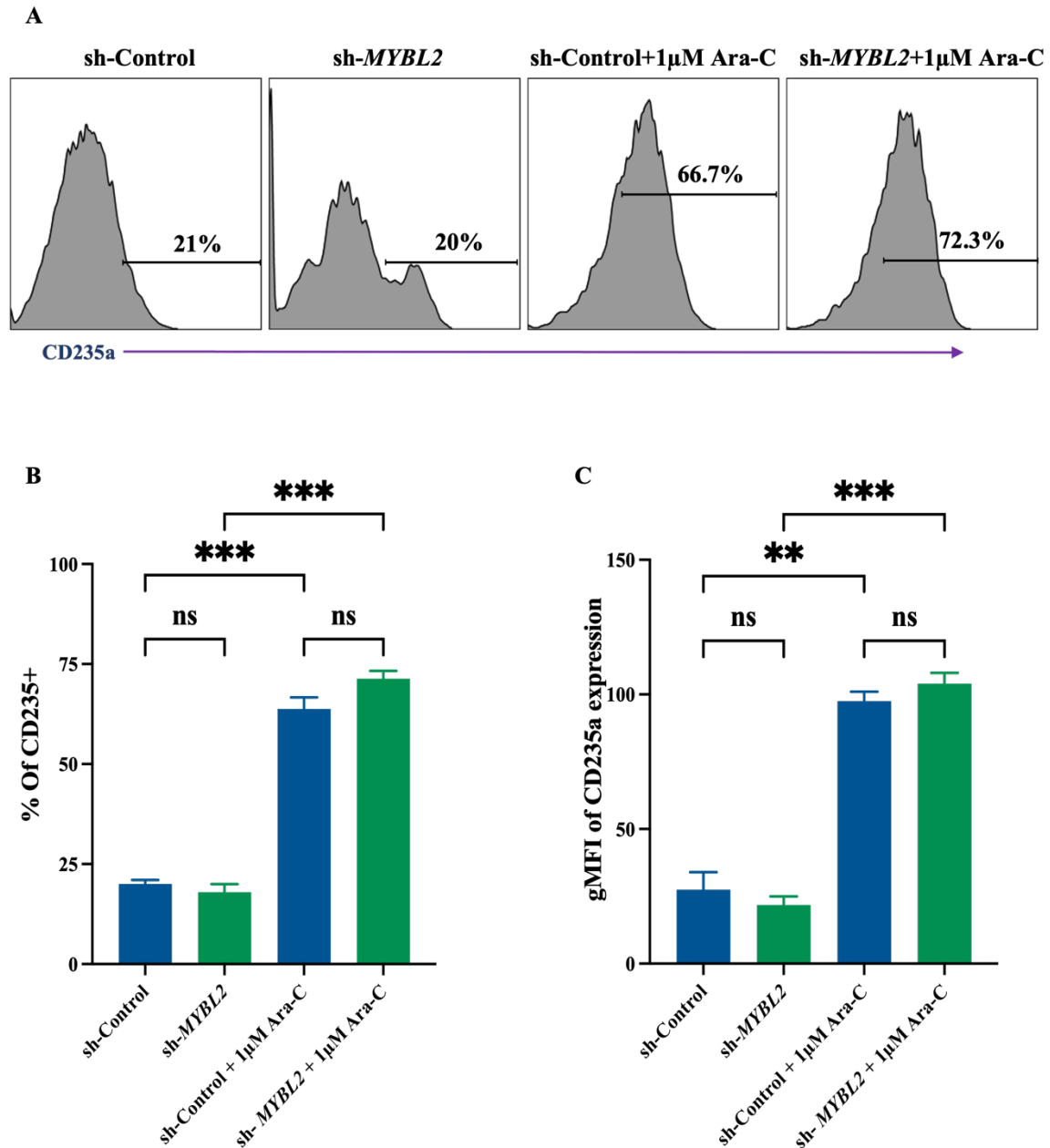
These results indicate that 1  $\mu$ M Ara-C induces erythroid differentiation in K562 and KG1a. In contrast, it did not influence the erythroid differentiation in SKM-1. One explanation for that could be related to the duration used to evaluate the drug's effect on cells, which may not have been sufficient to observe the impact of the drug on cell differentiation. It would be ideal to evaluate the drug's effect after 96 h. The reduction of MYBL2 did not affect erythroid differentiation, nor did it enhance or reduce the capability of the cells to differentiate. one possibility is that the knockdown of MYBL2 may not have been entirely effective and may have left behind sufficient amounts of functional MYBL2 proteins to sustain the differentiation.



**Figure 5.13 Ara-C induces erythroid differentiation in K562 cells.**

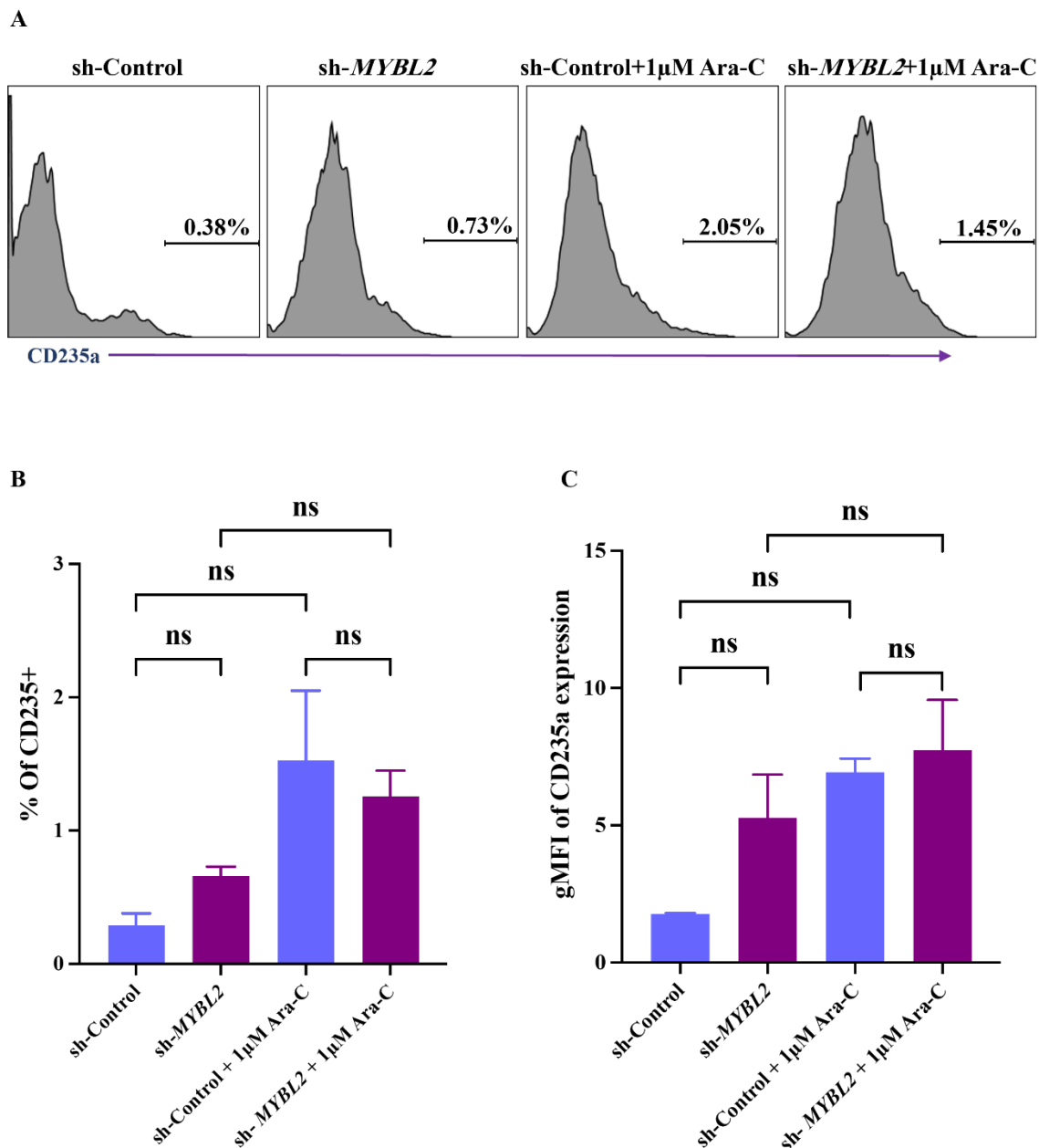
(A) CD235a<sup>+</sup> erythroid marker was evaluated in K562 sh-Control and sh-MYBL2 after being treated with 1 µM Ara-C for 48 h in the presence of 1 µg/ml dox. The cells were analysed by flow cytometry. (B) Bar graph representing the mean percentage of CD235<sup>+</sup> for cells treated or not for 48 h with 1 µM Ara-C. (C) Bar graph illustrating the geometric mean (gMFI). The statistical results are presented as the mean ± SEM and analysed using one-way ANOVA with multiple comparisons. The data are presented as the mean ± standard error of the mean (SEM), ns = not significant, \* $p < 0.05$  \*\* $p < 0.01$ . The data represent three independent experiments.





**Figure 5.14 Ara-C induced erythroid differentiation in KG1a cells.**

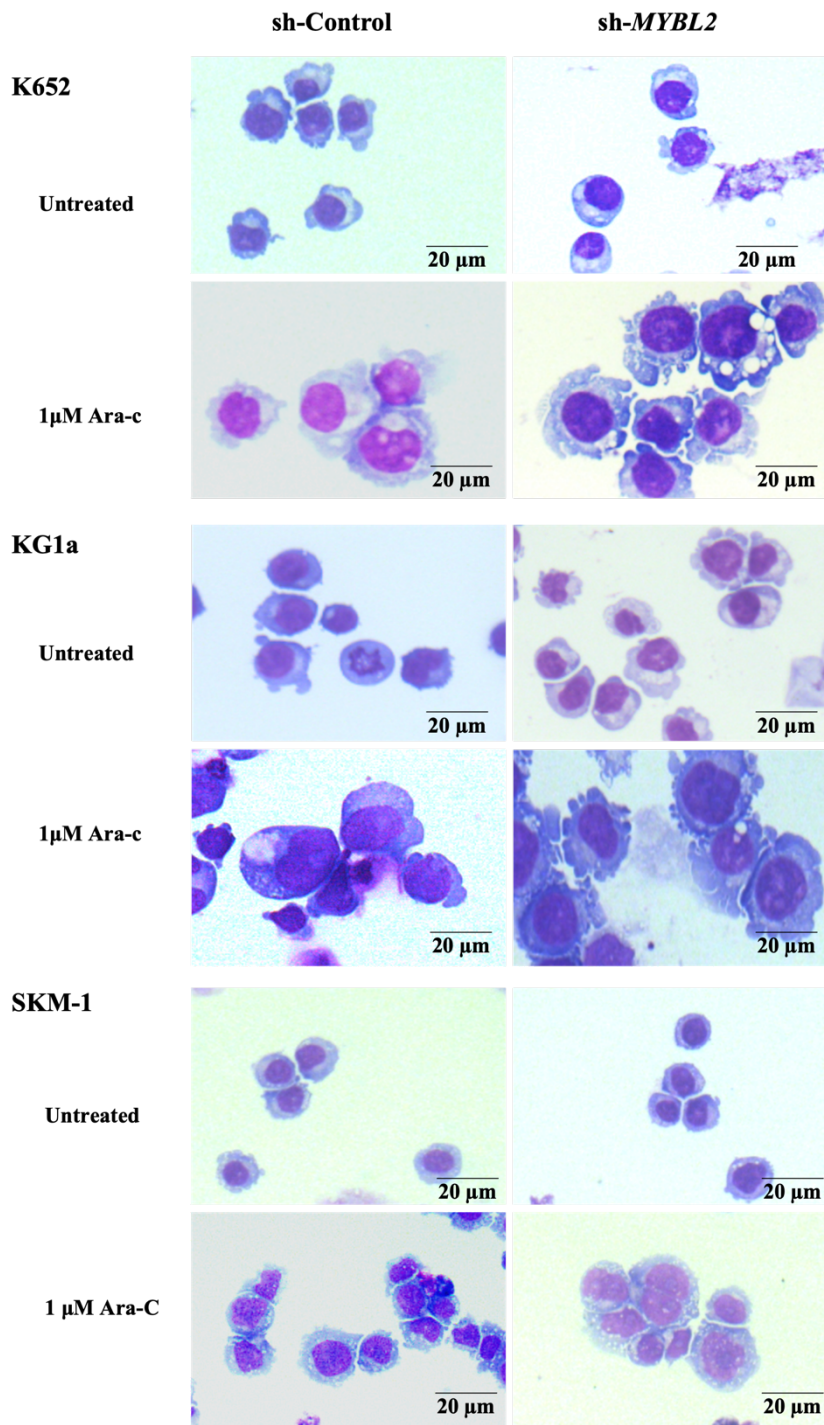
(A) CD235<sup>+</sup> erythroid marker was evaluated in KG1a sh-Control and sh-MYBL2 after being treated with 1 µM Ara-C for 48 h in the presence of 1 µg/ml dox. The cells were analysed using flow cytometry. (B) Bar graph representing the mean percentage of CD235<sup>+</sup> for cells treated or not for 48 h with 1 µM Ara-C. (C) Bar graph representing the geometric mean (gMFI). The statistical results are presented as the mean ± SEM and analysed using one-way ANOVA with multiple comparisons. The data are presented as the mean ± standard error of the mean (SEM), ns = not significant, \*\**p* < 0.01, \*\*\**p* < 0.001. The data represent three independent experiments.



**Figure 5.15 Ara-C does not induce erythroid differentiation in SKM-1 cells.**

(A) CD235a<sup>+</sup> erythroid marker was evaluated in SKM-1 sh-Control and sh-MYBL2 after being treated with 1 µM Ara-C for 48 h in the presence of 1 µg/ml dox. The cells were analysed using flow cytometry. (B) Bar graph representing the mean percentage of CD235<sup>+</sup> for cells treated or not for 48 h with 1 µM Ara-C. (C) Bar graph representing the geometric mean (gMFI). The statistical results evaluated using one-way ANOVA with multiple comparisons and presented as the mean ± SEM and.) ns = not significant. The data represent three independent experiments.

In addition, the cell morphology was also analysed to assess the drug's effect on the cells. Cells treated with Ara-C for 48h following published literature (Gañán-Gómez et al. 2014) then, stained with diff-quick dye and observed under a microscope. Compared with the control group, cells subjected to Ara-C treatment presented morphologic features of differentiation, indicating a more mature phenotype, increased cell size, coarse chromatin and reduced nucleus-cytoplasm ratio (Figure 5.16). Additionally, morphological changes were apparent for cells with reduced MYBL2 levels, with more giant cells in both KG1a and SKM-1 cells. These results reveal that Ara-C treatment promoted the differentiation of the cells that could be measured by the surface expression of mature antigens and also by phenotypic changes after morphological examination.



**Figure 5.16 Morphology of Ara-C treated cell line.**

Cytospin of K562, KG1a, and SKM-1 harboured low levels of MYBL2 cells after 48 h of Ara-C treatment in the presence of 1 μg/ml dox. Pictures were captured using a Leica DM6000 at 100× magnification, scale bar: 20 μm, representative images of n=3.

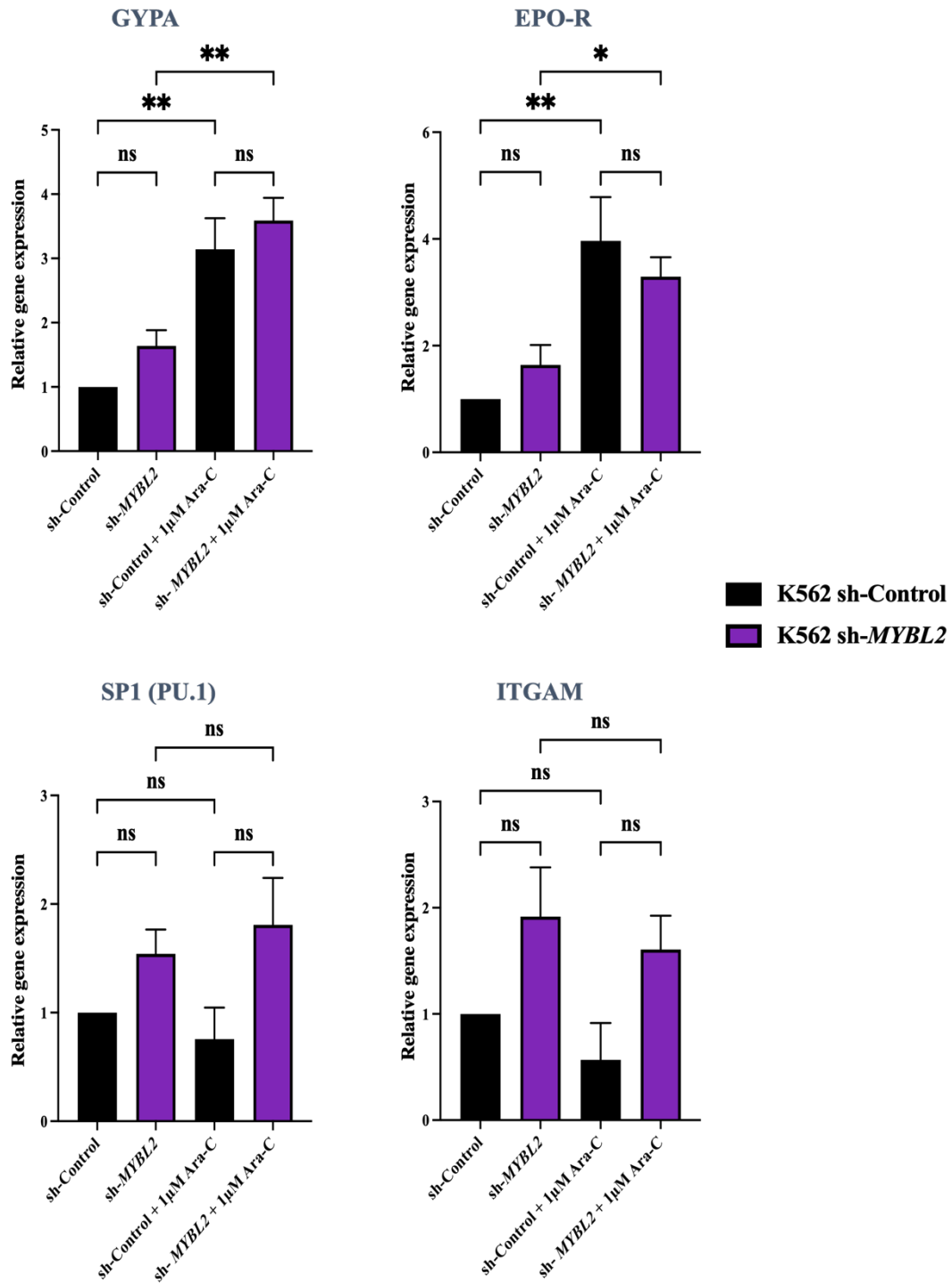
Lastly, to further corroborate the effect of Ara-C on erythroid and myeloid differentiation, the common erythroid markers: glycophorin A (*GYP A*) (Andersson et al. 1979), erythropoietin receptor (*EPO-R*) (Youssoufian et al. 1993) and the myeloid differentiation markers: integrin  $\gamma$ -M (*ITGAM*) (Pahl et al. 1992) and *PUI* (*SP11*) (Lloberas et al. 1999) were analysed by qPCR.

K562 sh-Control and sh-*MYBL2* show that upon Ara-C treatment, the erythropoietin receptor glycophorin A (*GYP A*, CD235a) and *EPO-R* mRNA levels were significantly increased. In contrast, no statistically significant effect was observed in the expression levels of myeloid markers *ITGAM* and *PUI* in either cell. Also, no statistically significant difference was observed between sh-Control and sh-*MYBL2* treated or untreated with Ara-C (Figure 5.17).

In KG1a cells, similar to K562, no difference was observed in erythroid markers, *EPO-R* and *GYP A* between sh-Control and sh-*MYBL2*. However, adding Ara-C statistically increased the expression of these markers in both cells. Interestingly, the expression of the myeloid marker *ITGAM* was observed to be significantly reduced in sh-*MYBL2* compared to sh-Control; however, the treatment of Ara-C for 48 hours showed an increase in this myeloid marker, suggesting that Ara-C has the potential to stimulate myeloid differentiation, whilst no significant alteration was detected in *PUI* (figure 5.18).

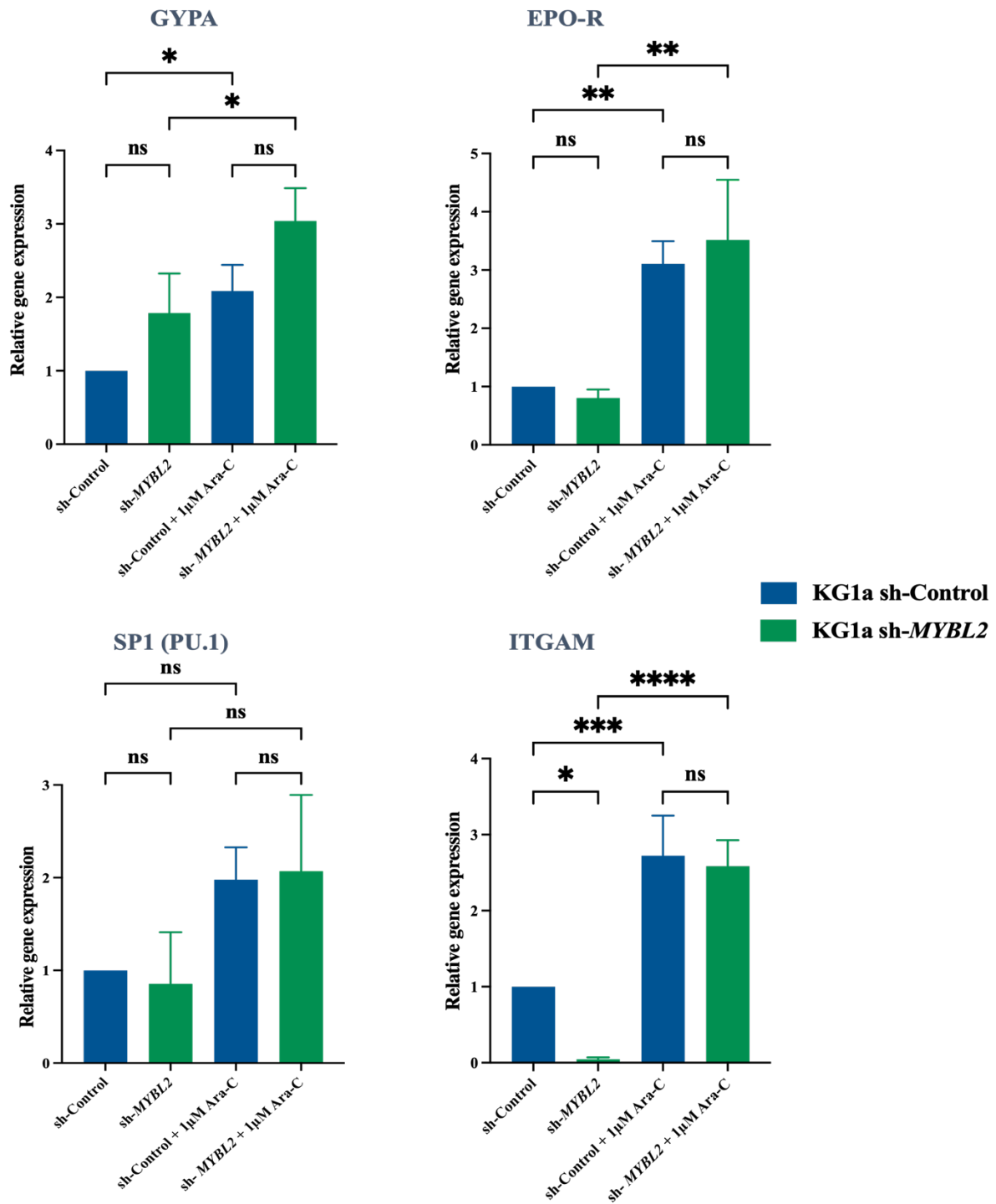
Additionally, an increase in the expression of *GYP A* and *EPO-R* was noted in SKM-1 Sh-Control and sh-*MYBL2* cell lines. Furthermore, *ITGAM* exhibited upregulation, whereas Ara-C treatment did not elicit any significant alteration in the level of *PUI*. No statistically significant difference was seen between sh-Control and sh-*MYBL2* treated or untreated with Ara-C, as illustrated in Figure 5.19.

Altogether, these results indicate that Ara-C induces erythroid differentiation in these cell lines and that *MYBL2* protein levels seem not to affect the erythroid differentiation potential of the cells.



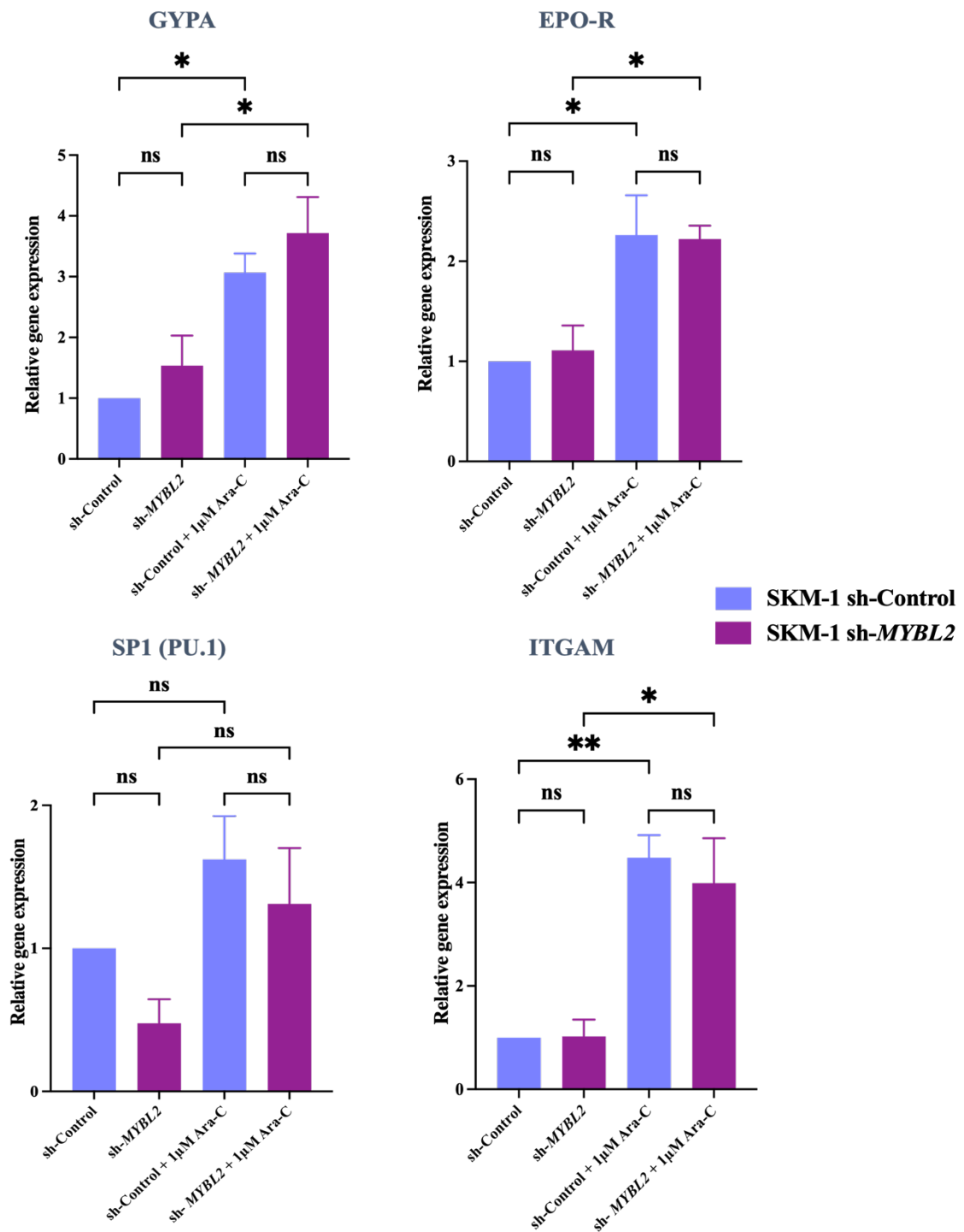
**Figure 5.17 Erythroid and myeloid marker expression in Ara-C-treated K562 cells.**

Graphs showing qRT-PCR of different markers in K562 sh-Control & K562 sh-MYBL2. RNA was extracted from untreated cells and cells treated with 1 µM Ara-C after 48 h in the presence of 1 µg/ml dox and retrotranscribed to cDNA. qRT-PCR was performed using primers designed to amplify *GYPA*, *EPO-R*, *PU.1*, and integrin γ-M (*ITGAM*). The relative expression was calculated by normalising to the *GAPDH* housekeeping gene and then expressed relative to sh-Control expression. The data are presented as the mean ± standard error of the mean (SEM) and analysed using one-way ANOVA with multiple comparisons. ns = not significant, \* $p < 0.05$  and \*\* $p < 0.01$ . The data represent three independent experiments.



**Figure 5.18 Erythroid and myeloid marker expression in Ara-C-treated KG1a cells.**

Graphs showing qRT-PCR of different markers in KG1a sh-Control & KG1a sh-MYBL2. RNA was extracted from untreated cells and cells treated with 1 µM Ara-C after 48 h in the presence of 1 µg/ml dox and retrotranscribed to cDNA. qRT-PCR was performed using primers designed to amplify *GYPA*, *EPO-R*, *PU.1*, and integrin  $\gamma$ -M (*ITGAM*). The relative expression was calculated by normalising to the *GAPDH* housekeeping gene and then expressed relative to sh-Control expression. The data are presented as the mean  $\pm$  standard error of the mean (SEM) and analysed one-way ANOVA with multiple comparisons. ns = not significant, \* $p < 0.05$  and \*\* $p < 0.01$ , \*\*\* $p < 0.001$ , \*\*\*\* $p < 0.0001$ . The data represent three independent experiments.



**Figure 5.19 Erythroid and myeloid marker expression in Ara-C-treated SKM-1 cells.**

Graphs showing qRT-PCR of different markers in SKM-1 Sh-Control & SKM-1 sh-MYBL2. RNA was extracted from untreated cells and cells treated with 1  $\mu$ M Ara-C after 48 h in the presence of 1  $\mu$ g/ml dox and retrotranscribed to cDNA. qRT-PCR was performed using primers designed to amplify *GYPA*, *EPO-R*, *PU.1*, and integrin  $\gamma$ -M (*ITGAM*). The relative expression was calculated by normalising to the *GAPDH* housekeeping gene and then expressed relative to sh-Control expression. The data are presented as the mean  $\pm$  standard error of the mean (SEM) and analysed one-way ANOVA with multiple comparisons. ns = not significant, \* $p < 0.05$  and \*\* $p < 0.01$ . The data represent 3 independent experiments.



## **5.3 Discussion:**

### **5.3.1 Engineering human hiPSCs with low *MYBL2* levels by CRISPR-Cas9**

This chapter aims to investigate and define the role of *MYBL2* in haematological diseases, primarily MDS, using two model systems: hiPSCs and leukaemia/MDS cell lines. Our group and others previously reported that 50% of high-risk (worse prognosis) patients with MDS display lower levels of *MYBL2* (Clarke et al. 2013; Heinrichs et al. 2013) and that lower *MYBL2* expression can be observed even in patients with normal karyotype and not associated with the interstitial deletion of chromosome 20q, where *MYBL2* is located (Clarke et al. 2013; Heinrichs et al. 2013). Interestingly, in an MDS case, heterozygous missense mutations in *MYBL2* result in the loss of gene function, similar to the effect of monoallelic gene deletion, indicating that the *MYBL2* transcription factor is targeted by a variety of molecular events in MDS (Heinrichs et al. 2013). Using mouse model systems, our group and Thomas Look's group showed that lowering the levels of *MYBL2* led to MDS-like disease in mice (Heinrichs et al., 2013; Clarke et al., 2013). Nevertheless, whether the sole disturbance of *MYBL2* can drive MDS in a human setting is unknown.

The purpose of the present chapter was to investigate *MYBL2* role in MDS to study the connection between *MYBL2* deficiency and the lineage dysplasia characteristic of patients with MDS; the CRISPR-Cas9 system was implemented as a simple tool for introducing 'indels' into the *MYBL2* locus in hiPSCs and generating hiPSC clones with low expression of functional *MYBL2*. A "one in all" system, a plasmid (pX458) containing Cas9, a small guidance RNA under the U6 promoter, and the reporter gene (GFP), was chosen to facilitate the integration of all the components into the same cell (Ran et al. 2013). We aimed to target exon 5 of *MYBL2*, engineering a heterozygous mutation. Successful subcloning of guide RNAs, transfection of the plasmids into the BU3.10 cells, around 52% of GFP+ cells were sorted, growth, and

isolation of individual clones after sorting (individual iPSCs), and amplification of a 420 bp region flanking the targeted CRISPR area were achieved.

Unfortunately, our attempts to validate the effectiveness of CRISPR-Cas9 in targeting the *MYBL2* gene using the T7 endonuclease I cleavage assay were unsuccessful. Despite our careful design of specific guide RNAs and rigorous optimisation of the CRISPR-Cas9 system, we could not detect any significant cleavage of *MYBL2* DNA fragments by T7 endonuclease I. This failure to validate CRISPR-Cas9 targeting in *MYBL2* may be attributed to different factors, including the known function of MYBL2 in maintaining pluripotency (Tarasov et al. 2008; Zhan et al. 2012). The pluripotency of iPSCs may be compromised by reducing MYBL2 expression, which results in affecting self-renewal and proliferation capacity. This may result in the inability to preserve the intended iPSC properties, thereby limiting the intended gene manipulation.

Moreover, our attempt to transduce iPSCs with TRIPZ to knockdown MYBL2 did not achieve the intended results, and we were unable to successfully transduce iPSCs. Other groups have reported facing difficulty in transducing iPSCs with different lentiviruses (Alex Thomson, personal communication). Owing to COVID-imposed restrictions, it was vital to maximise lab time, as CRISPR validation has been shown to be time-intensive. We decided to use human cell lines to uncover the role of *MYBL2* in erythroid and myeloid differentiation in heterogeneous myeloid malignancies.

### **5.3.2 Use of human cell lines to study the role of MYBL2 in erythroid and myeloid differentiation**

As mentioned above, due to COVID restrictions, after repeated attempts to validate the success of the CRISPR system, we decided to utilise immortalised human cell lines as a different approach to determine the involvement of *MYBL2* in erythroid/myeloid differentiation. We decided to use three distinct cell lines that represent different myeloid malignancies: (i) K562

cells, derived from a female patient with chronic myeloid leukaemia (Andersson et al., 1979b); (ii) KG1a cells, taken from a male AML patient (Koeffler and Golde 1978); and (iii) SKM-1 cells, a male MDS/AML cell line (Nakagawa et al. 2008). Previous studies have shown that Ara-C treatment could stimulate myeloid differentiation in these cell lines (Winter et al. 1985; Ossenkoppele et al. 1991; Wang et al. 2016).

We utilised an shRNA approach to knock down *MYBL2* expression in leukaemia cells to understand the functional role of *MYBL2* in leukaemia and to explore its potential as a therapeutic target. Fortunately, lentiviral-mediated transduction successfully reduced *MYBL2* expression. This technique has been widely used to knock down gene expression for functional studies and potential therapeutic applications (Brummelkamp et al. 2002). It provides an easy way to introduce shRNA into dividing or non-dividing cells and is often associated with lower cytotoxicity than adenovirus-mediated transduction (Taxman et al. 2010). Further, an appropriate control was used to validate the assays. This control has an identical lentiviral backbone expressing a similar reporter protein (RFP) (Taxman et al. 2010).

The doxycycline-inducible shRNA system is a tetracycline-regulated gene expression system that allows for the controlled silencing of a target gene in response to doxycycline administration (Gossen and Bujard 1992). Our system utilises a vector approach comprising a Tet repressor-expressing vector and a Tet-responsive shRNA expression vector, allowing for reversible and tunable knockdown of the target gene. Adding doxycycline relieves repression and induces shRNA expression, leading to downregulation of the target gene. (Szulc and Aebischer 2008).

In summary, our study demonstrates that the doxycycline-inducible shRNA system provides a powerful tool for the controlled and reversible downregulation of *MYBL2* expression in leukaemia cell lines, highlighting its functional consequences and therapeutic implications.

### 5.3.3 Ara-C affects cell proliferation

*MYBL2* is essential for cell proliferation and is involved in controlling the G1/S and G2/M transitions of the cell cycle, which are necessary checkpoints for cell division (Musa et al. 2017). Cytarabine (also called cytosine arabinose or Ara-C) is an antimetabolite chemotherapy drug used for treating AML for over 40 years (Lowenberg et al. 1999). Aside from AML, Ara-C is used to treat a variety of other haematological cancers, including acute lymphoblastic leukaemia, chronic myelocytic leukaemia, erythroleukemia, and mantle-cell lymphoma (Lamba 2009). Therefore, the current study aims to determine if Ara-C could enhance differentiation in leukaemia cells after the downregulation of *MYBL2*.

The first conclusion obtained from the present study on this cell line is that the downregulation level of *MYBL2* does not impair cell numbers. Reducing the expression of *MYBL2* in these cells (K562, KG1a, and SKM-1) could not impede their ability to divide and grow. *MYBL2* regulates genes that are critical for cell cycle progression and are involved in regulating the G1/S and G2/M transitions of the cell cycle; therefore, downregulation of this gene leads to cell cycle arrest and disrupts the progression of the cell cycle, resulting in cell cycle arrest and, ultimately, a decrease in leukaemia cell proliferation (Okada 2002; Santilli et al. 2005). One possibility could be that a low level of *MYBL2* did not affect cell numbers because of its low knockdown efficiency. Another reason could be that the reduction in *MYBL2* may be compensated for by the increased expression of other genes that also promote cell cycle progression and proliferation. For example, other cyclin or CDK genes may increase their expression to compensate for reduced *MYBL2*. Leukaemia cells are known for their plasticity. This means that they can adapt to changes, including the loss or reduction of specific genes, and find alternative ways to proliferate. Techniques such as BrdU could confirm proliferation capacity.

Moreover, we investigated the effect of Ara-C treatment on cell viability. Our results revealed that K562, KG1a, and SKM-1 were responsive to Ara-C treatment, with 1  $\mu$ M Ara-C significantly reducing cell number after 48 h. The results of our work on the two AML cell lines and SKM-1 align with prior findings indicating that dosages of up to 100 nM Ara-C induce differentiation of leukaemia cells *in vitro*, whereas greater (500  $\mu$ M) concentrations increase cell death (Wang et al. 2000; Chen et al. 2017). Furthermore, these findings are consistent with studies that have discovered that Ara-C inhibits the proliferation of leukaemia cell lines. Ara-C therapy inhibited cell proliferation in K562 cells (Chen et al., 2011). The observed decreased proliferation based on cell counter and colony assays might be due to cells undergoing slow proliferation. Although these results could reflect a lack of proliferation but not apoptosis, our manual count suggests that apoptosis did not happen because no increase in dead cells (trypan blue-positive cells) was detected during the counting of the cells. Ara-C treatment has been previously reported to diminish cell numbers and trigger cell death in KG1a and SKM-1 (Gañán-Gómez et al. 2014). Our results are also consistent with a recent study indicating that Ara-C reduces cell numbers and stimulates differentiation in leukaemia cells (Tomic et al. 2022). Indeed, several studies previously reported that Ara-C could induce apoptosis in AML cells by blocking DNA and RNA synthesis, enhanced by a potent differentiation stimulus while remaining non-toxic to untreated AML cells (Wang et al. 2016). Apoptosis assays, such as Annexin V, would have helped clarify whether the lack of proliferation observed in our cells after Ara-C treatment was due to increased differentiation or apoptosis. Unfortunately, time constraints and technical issues prevented the assessment of apoptosis in these cells.

In summary, our results indicate that Ara-C effectively inhibited the proliferation of K562, KG1a, and SKM-1, regardless of MYBL2 levels. The primary mechanism through which Ara-C exerts these effects is possibly by inhibiting DNA synthesis, which induces cell cycle

arrest. However, this needs to be confirmed by conducting cell cycle, apoptosis, and proliferation assays.

#### **5.3.4 Ara-C induces differentiation of leukaemia cells**

Recent studies have suggested that Ara-C may promote differentiation in leukaemia cell lines. Our work is in line with this previous work (Tomic et al. 2022) based on two different sets of results: (i) Ara-C treatment-induced morphological alterations and changes in surface marker expression consistent with erythroid differentiation in all three leukaemia cell lines and (ii) qPCR results that further supported the differentiation-inducing effect of Ara-C, with upregulation of genes associated with lineage-specific differentiation.

The ability of cells to differentiate in response to the drug was assessed by evaluating CD235a. Our findings revealed an increase in two cell lines, K562 and KG1a, while the differentiation was limited in SKM-1. This increase was confirmed by measuring the lineage-specific differentiation for erythroid, showing that the upregulation of genes associated with EPO-R, and glycophorin A (*GYPA*) mRNA levels were increased upon treatment with Ara-C. (Gaán-Gómez et al., 2014). Nonetheless, changes in these markers did not seem affected by reducing the level of MYBL2. However, assessing multiple markers for erythroid and myeloid differentiation is recommended to provide more robust and reliable results.

Another finding was that treatment of cells with Ara-C resulted in increased expression of myeloid differentiation markers, which confirmed studies that discovered that Ara-C treatment induces the expression of CD11b in KG1a cells, and upregulation of *ITGAM* in SKM-1. This may support evidence that a low dose of Ara-C is associated with a good response in patients with MDS (Visani et al. 2004; Wang et al. 2016; Tomic et al. 2022) .

The molecular mechanisms underlying Ara-C-induced differentiation may involve activating erythroid transcription factors. Our work shows that erythropoietin receptor (*EPO-R*) and glycophorin A (*GYPA*, CD235a+) mRNA levels increased following Ara-C treatment.

Moreover, our data revealed that reducing MYBL2 in K562, KG1a, and SKM-1 did not affect the expression of these factors, the increase in which would have indicated erythroid progenitor differentiation. GYPA is a sialoglycoprotein of the erythrocyte membrane expressed during erythroid differentiation from the proerythroblast stage (Fajtova et al. 2013; Li et al. 2014). An increase in both *EPO-R* and *GYPA* after Ara-C treatment has been reported previously (Gañán-Gómez et al. 2014). Ara-C treatment promotes heme and globin synthesis (Zhang et al. 2007). Benzidine staining on the leukaemia cells to detect the pseudo peroxidase activity of haemoglobin following Ara-C treatment would have helped to confirm the drug's effectiveness in inducing erythroid differentiation.

Another finding obtained was that Ara-C resulted in increased expression of myeloid differentiation markers CD11b (Wang et al. 2016).

Our results align with a recent study that discovered that Ara-C treatment of KG1a cells induces the expression of CD11b in these cells (Tomic et al. 2022). In the case of MDS/AML cell line SKM-1, upregulation of *ITGAM* was observed. This could be consistent with evidence that a low dose of Ara-C is associated with a good response in patients with MDS (Visani et al. 2004).

## **Chapter 6: General discussion**



## 6.1 iPSCs as a model system for studying haematological disorders

The difficulty of reproducing the clinical characteristics of MDS in preclinical models is one of the obstacles to advances in MDS therapy. The reason behind these is related to many aspects. For example, MDS cell lines derived from patients before disease progression (*in vitro* models) are still limited. The lack of MDS cell numbers, difficulty in cell proliferation in culture, and challenged genetic engineering of these primary MDS cells in the lab constrain the direct examination of MDS patient samples (Stanchina et al., 2021).

iPSCs provide a robust system for studying the molecular mechanisms underlying haematological disorders and devising potential therapeutic interventions (Takahashi et al. 2007). These cells are derived from human cells and can therefore be used to model human diseases more precisely than animal models. Certain features of human diseases cannot be recapitulated in mice or other animal models due to differences in species (Lancaster and Knoblich, 2014). Thus, human genetic variations in diseases need to be investigated in cells that can be derived from patients with the disease.

Mutations in the transcription factor C/EBP $\alpha$  are relatively common in AML and MDS (van Waalwijk van Doorn-Khosrovani et al. 2003; Pabst and Mueller 2009; Heyes et al. 2021). *CEBPA* has a carboxyterminal moiety with a bZIP domain that binds to DNA and other transcription factors (Sierra and Nomdedeu 2021). Around 30% of *CEBPA* mutations in patients with AML and MDS are located in the bZIP domain (Wakita et al. 2022). In Chapter 3, we aimed to generate iPSCs harbouring a specific *CEBPA* mutation that would disrupt the C-terminal region using CRISPR-Cas9 and assess the effects of this mutation on the differentiation and function of haematopoietic cells. We isolated two clones that differed in their deletion length and successfully confirmed that these iPSC clones maintained their pluripotency after genetic manipulation by CRISPR based on various aspects. Conducting

these experiments was crucial for validating disease modelling, giving consistency and reliability to the experimental results (Soldner et al. 2009).

Overall, our *CEBPA*<sup>+mut</sup> iPSCs can complement conventional animal models and primary patient samples, providing a valuable resource for elucidating the role of this transcription factor in the pathogenesis of haematological diseases (Robinton and Daley 2012).

## **6.2 *CEBPA*<sup>+mut</sup> block granulocyte differentiation and lead to myeloid and erythroid**

### **dysplasia**

In Chapter 4, we analysed the haematopoietic disease phenotype of our cells after introducing a mutation disrupting *CEBPA* bZIP domain. Our data showed that hiPSC-*CEBPA*<sup>+mut</sup> clones were able to differentiate into HSPCs, although they impaired the differentiation of HSPCs into myeloid lineages, mainly granulocytes. The results of the colony assays showed that HSPCs derived from hiPSC-*CEBPA*<sup>+mut</sup> had a diminished myeloid CFU capacity and impeded the emergence of granulocytic CFU. The reduced capacity to form haematopoietic colonies in hiPSC-*CEBPA*<sup>+mut</sup> clones emphasises the importance of *C/EBPα* as a differentiation mediator in myeloid lineages. This finding highlights its essential function as a TF for granulocytic differentiation (Wouters et al. 2007; Pabst and Mueller 2009; Avellino and Delwel 2017). Indeed, in our system, the *CEBPA*<sup>+mut</sup> mutation led to the limited capacity of progenitors in methylcellulose medium since *CEBPA* has also been shown to regulate several myeloid-specific genes. It interacts specifically with functional regions proximal to the promoters of the G-CSF receptor and GM-CSF receptors to regulate the expression of these receptors, which are needed for early myeloid differentiation (Smith et al. 1996; Ai and Udalova 2020). *CEBPA* plays an important role in normal granulopoiesis. It binds to regulatory sequences of genes required for myeloid differentiation, for example, *GFI-1*, *IL-6R*, *KLF1*, and *CSF3R* (Sierra and Nomdedeu 2021).

Moreover, our findings indicate that hiPSC- *CEBPA*<sup>+mut</sup> had impaired differentiation into myeloid cells in liquid culture. This is evidenced by (i) decreased expression of lineage-specific markers of myeloid differentiation, and (ii) morphological abnormalities, such as pseudo-Pelger-Huët anomaly, seen in many conditions, including MDS (Wang et al. 2011). This result confirmed our previous finding obtained from colony assays in that *CEBPA*<sup>+mut</sup> led to an inhibition of granulocytic differentiation. These findings support previous research linking *CEBPA* mutations to haematological disorders (Koschmieder et al. 2009) and provide additional evidence for the essential role of *CEBPA* in the development and function of myeloid cells. (Zhang et al. 2004).

The functional effects of mutations in *CEBPA* may be elucidated at the transcriptional level, as the expression of *GATA2*, *RUNX1*, and *PU.1* was downregulated in *CEBPA* mutant cells. The downregulation of myeloid genes in hiPSC- *CEBPA*<sup>+mut</sup> revealed the enrichment of erythroid-specific genes, such as *EPO-R* and *TFRC*. qRT-PCR results corroborate previous research findings indicating that *CEBPA* mutation impaired the expression of *GATA2* and *PU.1* (Fasan et al. 2013; Pundhir et al. 2018). PU.1 and C/EBP $\alpha$  regulate granulocytic-monocytic differentiation, and ChIP-seq studies showed that C/EBP $\alpha$  plays an unexpected role in modulating PU.1 function (Zhao et al. 2022). Cooperating interactions between *CEBPA* and *GATA2* are indispensable for eosinophil commitment (Nerlov 2007; Koschmieder et al. 2009). Overall, the *CEBPA*<sup>+mut</sup> altered the expression of key genes regulating myeloid differentiation, thereby inhibiting myeloid differentiation, and promoting erythroid differentiation. Our findings align with the RNA-seq findings observed in patients with *CEBPA* mutations. These mutations have been shown to cause an increase in the expression of genes involved in erythroid differentiation, such as *TFRC* (CD71), *EPOR*, *HEMG*, *KLF1*, *GATA1* and genes encoding haemoglobin chains and erythrocyte membrane proteins (Marcucci et al. 2008).

Furthermore, our investigation of erythroid differentiation in liquid culture showed that hiPSC-*CEBPA*<sup>+mut</sup> might differentiate into mature cells. Nevertheless, these cells displaying abnormal morphology, such as multinuclear erythroid cells. However, as mentioned in Chapter 4, the action of the mechanism of *CEBPA*<sup>+mut</sup> required specific analysis. To decipher the direct regulatory function of *CEBPA*, we employed Cut and Run technology. Regrettably, our attempts were unsuccessful due to the high background in our sequencing, indicating over digestion of the chromatin and the impossibility to assign peaks bound by *CEBPA*. Ideally, conducting ChIP-seq would have allowed us to understand how this binding is influenced by this mutation. This technique's insights can deepen our understanding of the role of *CEBPA*, shedding light on its functional complexities during erythroid differentiation.

Overall, understanding the effect of *CEBPA*<sup>+mut</sup> on erythroid and myeloid cell differentiation and function can provide valuable insights into the pathogenesis of haematological disorders associated with this mutation (Friedman 2007). Our findings may contribute to the development of novel therapeutic strategies to target dysregulated molecular pathways in patients with *CEBPA* mutated haematological diseases, such as AML and MDS (Wouters and Delwel 2016). It would be helpful to clarify the disease's transcriptional changes to comprehend the mechanisms underlying the block of myeloid differentiation and the presence of abnormal morphology in myeloid and erythroid cells.

### **6.3 The impact of MYBL2 on cell proliferation and differentiation in leukaemia cells was negligible, whereas Ara-C demonstrated a significant effect**

In Chapter 5, we conducted parallel work using iPSCs to assess the role of MYBL2 in erythroid differentiation. Our goal was to elucidate the role of MYBL2 in MDS/leukaemia using iPSCs, as we can potentially control their differentiation process and guide them towards specific lineages (Laurent et al. 2011). This flexibility allows iPSCs to faithfully recapitulate the genetic and phenotypic characteristics of the disease of interest. However, we did not achieve the

intended results. The main reasons for the lack of success were: (i) work restrictions imposed due to COVID-19, (ii) incapability to validate CRISPR in iPSCs, and (iii) failure to transduce iPSCs with lentivirus vectors to inducible reduced MYBL2 levels. Therefore, we generated AML cell lines (K562, KG1a, and SKM-1) for transduction-targeted reduction of MYBL2 levels using a lentivirus containing a dox-inducible MYBL2shRNA and induced the differentiation of these lines with Ara-C, as previously reported (Gañán-Gómez et al. 2014; Tomic et al. 2022). We determined that reducing MYBL2 expression in these cells (K562, KG1a, and SKM-1) did not affect their ability to divide and grow. The reduction of MYBL2 did not affect the response of the cells to Ara-C, as the cells behaved similarly to dox-treated control cells. Surprisingly, the reduction of MYBL2 in the cells did not affect erythroid differentiation, as previously, MYBL2 downregulation has been associated with blocking erythroid differentiation and causing anaemia in mouse models. For example, Heinrichs et al. (2013) found that knocking down MYBL2 resulted in clonal dominance of hematopoietic progenitors and the development of a myeloproliferative or myelodysplastic disorder. Another study indicated that inactivation of MYBL2 in a mouse model leads to depletion of the HSC pool, resulting in substantial reductions in mature erythroid, and myeloid cells (Baker et al. 2014). Moreover, pervious study showed that mice expressing half the normal levels of MYBL2 develop anaemia (Clarke et al. 2013). Although the level of MYBL2 reduction in our study was 50%, which is approximately the same as the reduction in previously published studies, it is quite possible that the discrepancy between our study and the published data regarding the role of MYBL2 in haematopoietic differentiation is due to the more efficient knockdown of MYBL2 in their studies. Another possibility is the experimental system used in our studies: the genetic drift and clonal selection are among the limitations of the leukaemia cell lines (Gillet et al. 2013). Additionally, over time, the characteristics of a cell line may diverge from those of the original patient sample, thereby diminishing their utility as a model

for researching leukaemia (Ben-David et al. 2019). The leukaemia cell lines used in this study possess distinct genetic and molecular characteristics for each subtype and are heterogeneous in their differentiation potential. The heterogeneous character of leukaemia cell differentiation poses a significant obstacle. This is because leukaemia comprises a wide variety of subtypes, each exhibiting unique patterns of differentiation potential and therapeutic agent response (Song et al. 2021). Consequently, it becomes difficult to establish a standardised framework for tracking differentiation based on morphology. The multifactorial nature of differentiation, influenced by genetic and epigenetic factors, interprets the observed morphological changes and their correlations with drug-induced differentiation as subjective and requires expert validation.

#### **6.4 Limitations of this study**

All of the experiments were conducted comparing two clones of *CEBPA*<sup>+mut</sup> generated by CRISPR-Cas9 with isogenic controls: WT hiPSCs, and CRISPR control. Nonetheless, it would be helpful to use an additional iPSC line and generate *CEBPA*<sup>+mut</sup> and CRISPR control isogenic lines to validate the results. Despite this possible limitation, our results mirrored those described in the literature using different experimental systems, providing confidence in the validity of our findings.

The major limitation of this study is the *in vitro* differentiation of erythroid and myeloid cells. *In vitro* differentiation offers several advantages but still includes drawbacks associated with this approach. The microenvironment or niche plays a crucial role in regulating hematopoietic cell differentiation. It provides essential cues, such as, extracellular matrix components, cell–cell interactions and growth factor, influencing HSPCs fate decisions and erythroid and myeloid cells' behaviour. This could lead to altered cell behaviour and functionality (P. Zhang et al. 2019). Differentiation methods *in vitro*, including ours, often fail to recapitulate the complex signals present in the bone marrow microenvironment. Regarding evaluating the role

of MYBL2 in myeloid and erythroid dysplasia, the main limitation of our study was the use of a leukaemia cell line that as mentioned in the general discussion is not the perfect model for studying differentiation. This is due to (i) the limited differentiation potential, and (ii) the differentiation has to be promoted by the use of chemicals such as Ara-C instead of specific cytokines known to drive the differentiation of specific lineages.

## **6.5 Final conclusion and future work**

The role of *CEBPA*<sup>+/-mut</sup> in HSPC biology require more investigation. This is primarily masked by the heterogeneity of available animal models, which may not fully recapitulate human disease (Lancaster and Knoblich 2014), making it challenging to clearly define the function of the gene within an HSPC population. One method that could overcome this issue is single-cell RNA sequencing. Combining this technology with *in vitro* and *in vivo* experiments provides a platform for defining the role of *CEBPA* in human HSPCs in health and disease. Further work using human bone marrow organoids would provide an alternative to overcome the main limitation of this study (the lack of microenvironment cues). Recently, a vascularised human bone marrow–organoid using iPSCs has been generated (Khan et al. 2023). These bone marrow organoids contain haematopoietic elements and support active endogenous haematopoiesis and the growth and survival of haematopoietic cells (Khan et al. 2023). Additional investigations, such as xenograft experiments, will be performed to elucidate the impact of monoallelic mutation of *CEBPA*. To accomplish this, immunocompromised mice will undergo transplantation with HSPCs derived from hiPSC-*CEBPA*<sup>+/-mut</sup>. Studying *CEBPA*<sup>+/-mut</sup> using these model systems may provide potentially powerful and support our *in vitro* findings obtained through phenotyping assays.

With respect to evaluating the role of MYBL2 in myeloid and erythroid dysplasia, neonatal primary human dermal fibroblasts (HDFn) may be infected with pTRIPZ sh*MYBL2* to reduce MYBL2 levels and then reprogramme the transduced HDFn into iPSCs. This strategy allows

us to control the downregulation of MYBL2 by doxycycline treatment in the same way as performed in our cell line studies. Also, it represents an advantage as MYBL2 levels would remain intact during the reprogramming process and maintenance of iPSC and could be manipulated at the level of the HSPCs to differentiate them into erythroid and myeloid lineages.



## REFERENCES

- Abdallah, N.A., Prakash, C.S. and McHughen, A.G. (2015). Genome editing for crop improvement: Challenges and opportunities. *GM Crops & Food*, 6(4), pp.183–205.
- Abelson, S. et al. (2018). Prediction of acute myeloid leukaemia risk in healthy individuals. *Nature*, 559(7714), pp.400–404.
- Abujarour, R., Valamehr, B., Robinson, M., Rezner, B., Vranceanu, F. and Flynn, P. (2013). Optimized Surface Markers for the Prospective Isolation of High-Quality hiPSCs using Flow Cytometry Selection. *Scientific Reports*, 3(1), p.1179.
- Adès, L., Itzykson, R. and Fenaux, P. (2014). Myelodysplastic syndromes. *The Lancet*, 383(9936), pp.2239–2252.
- Adolfsson, J., Månsson, R., Buza-Vidas, N., Hultquist, A., Liuba, K., Jensen, C.T., Bryder, D., Yang, L., Borge, O.-J., Thoren, L.A.M., Anderson, K., Sitnicka, E., Sasaki, Y., Sigvardsson, M. and Jacobsen, S.E.W. (2005). Identification of Flt3<sup>+</sup> Lympho-Myeloid Stem Cells Lacking Erythro-Megakaryocytic Potential. *Cell*, 121(2), pp.295–306.
- Ahn, J.-S. et al. (2016). Normal karyotype acute myeloid leukemia patients with CEBPA double mutation have a favorable prognosis but no survival benefit from allogeneic stem cell transplant. *Annals of Hematology*, 95(2), pp.301–310.
- Ai, Z. and Udalova, I.A. (2020). Transcriptional regulation of neutrophil differentiation and function during inflammation. *Journal of Leukocyte Biology*, 107(3), pp.419–430.
- Akashi, K., Traver, D., Miyamoto, T. and Weissman, I.L. (2000). A clonogenic common myeloid progenitor that gives rise to all myeloid lineages. *Nature*, 404(6774), pp.193–197.
- Alagoz, M. and Kherad, N. (2020). Advance genome editing technologies in the treatment of human diseases: CRISPR therapy (Review). *International Journal of Molecular Medicine*, 46(2), pp.521–534.
- Andersson, L.C., Gahmberg, C.G., Teerenhovi, L. and Vuopio, P. (1979). Glycophorin a as a cell surface marker of early erythroid differentiation in acute leukemia. *International Journal of Cancer*, 24(6), pp.717–720.
- Antoniou, P., Miccio, A. and Brusson, M. (2021). Base and Prime Editing Technologies for Blood Disorders. *Frontiers in Genome Editing*, 3.
- Arber, D.A., Orazi, A., Hasserjian, R., Thiele, J., Borowitz, M.J., Le Beau, M.M., Bloomfield, C.D., Cazzola, M. and Vardiman, J.W. (2016). The 2016 revision to the World Health Organization classification of myeloid neoplasms and acute leukemia. *Blood*, 127(20), pp.2391–2405.
- Armes, H., Rio-Machin, A., Krizsán, S., Bödör, C., Kaya, F., Bewicke-Copley, F., Alnajar, J., Walne, A., Péterffy, B., Tummala, H., Rouault-Pierre, K., Dokal, I., Vulliamy, T. and Fitzgibbon, J. (2022). Acquired somatic variants in inherited myeloid malignancies. *Leukemia*, 36(5), pp.1377–1381.

- Asou, H., Gombart, A.F., Takeuchi, S., Tanaka, H., Tanioka, M., Matsui, H., Kimura, A., Inaba, T. and Koefler, H.P. (2003). Establishment of the acute myeloid leukemia cell line Kasumi-6 from a patient with a dominant-negative mutation in the DNA-binding region of the C/EBP $\alpha$  gene. *Genes, Chromosomes and Cancer*, 36(2), pp.167–174.
- Avellino, R. and Delwel, R. (2017). Expression and regulation of C/EBP $\alpha$  in normal myelopoiesis and in malignant transformation. *Blood*, 129(15), pp.2083–2091.
- Bacher, U., Haferlach, T., Schnittger, S., Zenger, M., Meggendorfer, M., Jeromin, S., Roller, A., Grossmann, V., Krauth, M.-T., Alpermann, T., Kern, W. and Haferlach, C. (2014). Investigation of 305 patients with myelodysplastic syndromes and 20q deletion for associated cytogenetic and molecular genetic lesions and their prognostic impact. *British Journal of Haematology*, 164(6), pp.822–833.
- Baker, S.J., Ma'ayan, A., Lieu, Y.K., John, P., Reddy, M.V.R., Chen, E.Y., Duan, Q., Snoeck, H.-W. and Reddy, E.P. (2014). B- myb is an essential regulator of hematopoietic stem cell and myeloid progenitor cell development. *Proceedings of the National Academy of Sciences*, 111(8), pp.3122–3127.
- Baliou, S., Adamaki, M., Kyriakopoulos, A., Spandidos, D., Panagiotidis, M., Christodoulou, I. and Zoumpourlis, V. (2018). CRISPR therapeutic tools for complex genetic disorders and cancer (Review). *International Journal of Oncology*, 6 June 2018.
- Barak, M., Fedorova, V., Pospisilova, V., Raska, J., Vochyanova, S., Sedmik, J., Hribkova, H., Klimova, H., Vanova, T. and Bohaciakova, D. (2022). Human iPSC-Derived Neural Models for Studying Alzheimer's Disease: from Neural Stem Cells to Cerebral Organoids. *Stem Cell Reviews and Reports*, 18(2), pp.792–820.
- Barbany, G., Arthur, C., Liedén, A., Nordenskjöld, M., Rosenquist, R., Tesi, B., Wallander, K. and Tham, E. (2019). Cell-free tumour <sc>DNA</sc> testing for early detection of cancer – a potential future tool. *Journal of Internal Medicine*, 8 April 2019, p.joim.12897.
- Barminko, J., Reinholt, B. and Baron, M.H. (2016). Development and differentiation of the erythroid lineage in mammals. *Developmental & Comparative Immunology*, 58, pp.18–29.
- Baron, M.H., Isern, J. and Fraser, S.T. (2012). The embryonic origins of erythropoiesis in mammals. *Blood*, 119(21), pp.4828–4837.
- Barzi, A. and Sekeres, M.A. (2010). Myelodysplastic syndromes: A practical approach to diagnosis and treatment. *Cleveland Clinic Journal of Medicine*, 77(1), pp.37–44.
- Bayley, R., Blakemore, D., Cancian, L., Dumon, S., Volpe, G., Ward, C., Almaghrabi, R., Gujar, J., Reeve, N., Raghavan, M., Higgs, M.R., Stewart, G.S., Petermann, E. and García, P. (2018). MYBL2 Supports DNA Double Strand Break Repair in Hematopoietic Stem Cells. *Cancer Research*, 78(20), pp.5767–5779.
- BECKER, A.J., McCULLOCH, E.A. and TILL, J.E. (1963). Cytological Demonstration of the Clonal Nature of Spleen Colonies Derived from Transplanted Mouse Marrow Cells. *Nature*, 197(4866), pp.452–454.
- Bejar, R., Stevenson, K., Abdel-Wahab, O., Galili, N., Nilsson, B., Garcia-Manero, G., Kantarjian, H., Raza, A., Levine, R.L., Neuberg, D. and Ebert, B.L. (2011). Clinical Effect of

Point Mutations in Myelodysplastic Syndromes. *New England Journal of Medicine*, 364(26), pp.2496–2506.

Bejar, R., Stevenson, K.E., Caughey, B.A., Abdel-Wahab, O., Steensma, D.P., Galili, N., Raza, A., Kantarjian, H., Levine, R.L., Neuberg, D., Garcia-Manero, G. and Ebert, B.L. (2012). Validation of a Prognostic Model and the Impact of Mutations in Patients With Lower-Risk Myelodysplastic Syndromes. *Journal of Clinical Oncology*, 30(27), pp.3376–3382.

Bench, A.J., Nacheva, E.P., Hood, T.L., Holden, J.L., French, L., Swanton, S., Champion, K.M., Li, J., Whittaker, P., Stavrides, G., Hunt, A.R., Huntly, B.J., Campbell, L.J., Bentley, D.R., Deloukas, P. and Green, A.R. (2000a). Chromosome 20 deletions in myeloid malignancies: reduction of the common deleted region, generation of a PAC/BAC contig and identification of candidate genes. *Oncogene*, 19(34), pp.3902–3913.

Bench, A.J., Nacheva, E.P., Hood, T.L., Holden, J.L., French, L., Swanton, S., Champion, K.M., Li, J., Whittaker, P., Stavrides, G., Hunt, A.R., Huntly, B.J., Campbell, L.J., Bentley, D.R., Deloukas, P. and Green, A.R. (2000b). Chromosome 20 deletions in myeloid malignancies: reduction of the common deleted region, generation of a PAC/BAC contig and identification of candidate genes. *Oncogene*, 19(34), pp.3902–3913.

Ben-David, U., Beroukhi, R. and Golub, T.R. (2019). Genomic evolution of cancer models: perils and opportunities. *Nature Reviews Cancer*, 19(2), pp.97–109.

Bennett, J.M., Catovsky, D., Daniel, M.-T., Flandrin, G., Galton, D.A.G., Gralnick, H.R. and Sultan, C. (1976). Proposals for the Classification of the Acute Leukaemias French-American-British (FAB) Co-operative Group. *British Journal of Haematology*, 33(4), pp.451–458.

Benveniste, P., Frelin, C., Janmohamed, S., Barbara, M., Herrington, R., Hyam, D. and Iscove, N.N. (2010). Intermediate-Term Hematopoietic Stem Cells with Extended but Time-Limited Reconstitution Potential. *Cell Stem Cell*, 6(1), pp.48–58.

Bereshchenko, O., Mancini, E., Moore, S., Bilbao, D., Månsson, R., Luc, S., Grover, A., Jacobsen, S.E.W., Bryder, D. and Nerlov, C. (2009). Hematopoietic Stem Cell Expansion Precedes the Generation of Committed Myeloid Leukemia-Initiating Cells in C/EBP $\alpha$  Mutant AML. *Cancer Cell*, 16(5), pp.390–400.

Bertrand, J.Y., Chi, N.C., Santoso, B., Teng, S., Stainier, D.Y.R. and Traver, D. (2010). Haematopoietic stem cells derive directly from aortic endothelium during development. *Nature*, 464(7285), pp.108–111.

Bhatia, S., Pilquill, C., Roth-Albin, I. and Draper, J.S. (2013). Demarcation of Stable Subpopulations within the Pluripotent hESC Compartment. *PLoS ONE*, 8(2), p.e57276.

Bies, J., Hoffman, B., Amanullah, A., Giese, T. and Wolff, L. (1996). B-Myb prevents growth arrest associated with terminal differentiation of monocytic cells. *Oncogene*, 12(2), pp.355–63.

Blakemore, D., Vilaplana-Lopera, N., Almaghrabi, R., Gonzalez, E., Moya, M., Ward, C., Murphy, G., Gambus, A., Petermann, E., Stewart, G.S. and García, P. (2021). MYBL2 and ATM suppress replication stress in pluripotent stem cells. *EMBO reports*, 22(5).

- Böiers, C. et al. (2013). Lymphomyeloid Contribution of an Immune-Restricted Progenitor Emerging Prior to Definitive Hematopoietic Stem Cells. *Cell Stem Cell*, 13(5), pp.535–548.
- Boisset, J.-C., van Cappellen, W., Andrieu-Soler, C., Galjart, N., Dzierzak, E. and Robin, C. (2010). In vivo imaging of haematopoietic cells emerging from the mouse aortic endothelium. *Nature*, 464(7285), pp.116–120.
- Boyer, L.A., Lee, T.I., Cole, M.F., Johnstone, S.E., Levine, S.S., Zucker, J.P., Guenther, M.G., Kumar, R.M., Murray, H.L., Jenner, R.G., Gifford, D.K., Melton, D.A., Jaenisch, R. and Young, R.A. (2005). Core Transcriptional Regulatory Circuitry in Human Embryonic Stem Cells. *Cell*, 122(6), pp.947–956.
- Brafman, D.A., Moya, N., Allen-Soltero, S., Fellner, T., Robinson, M., McMillen, Z.L., Gaasterland, T. and Willert, K. (2013). Analysis of SOX2-Expressing Cell Populations Derived from Human Pluripotent Stem Cells. *Stem Cell Reports*, 1(5), pp.464–478.
- Braun, T. et al. (2011). Characteristics and outcome of myelodysplastic syndromes (MDS) with isolated 20q deletion: A report on 62 cases. *Leukemia Research*, 35(7), pp.863–867.
- Briegel, K., Lim, K.C., Plank, C., Beug, H., Engel, J.D. and Zenke, M. (1993). Ectopic expression of a conditional GATA-2/estrogen receptor chimera arrests erythroid differentiation in a hormone-dependent manner. *Genes & Development*, 7(6), pp.1097–1109.
- Broudy, V.C. (1997). Stem Cell Factor and Hematopoiesis. *Blood*, 90(4), pp.1345–1364.
- de Bruijn, M. and Dzierzak, E. (2017). Runx transcription factors in the development and function of the definitive hematopoietic system. *Blood*, 129(15), pp.2061–2069.
- Brummelkamp, T.R., Bernards, R. and Agami, R. (2002). A System for Stable Expression of Short Interfering RNAs in Mammalian Cells. *Science*, 296(5567), pp.550–553.
- Brunner, E., Yagi, R., Debrunner, M., Beck-Schneider, D., Burger, A., Escher, E., Mosimann, C., Hausmann, G. and Basler, K. (2019). CRISPR-induced double-strand breaks trigger recombination between homologous chromosome arms. *Life Science Alliance*, 2(3), p.e201800267.
- Brunning, R.D. (2003). MDS—new classification, new problem?. *Leukemia Research*, 27(7), pp.567–569.
- Calkhoven, C.F., Müller, C. and Leutz, A. (2000). Translational control of C/EBP $\alpha$  and C/EBP $\beta$  isoform expression. *Genes & Development*, 14(15), pp.1920–1932.
- Calvi, L.M. and Link, D.C. (2015). The hematopoietic stem cell niche in homeostasis and disease. *Blood*, 126(22), pp.2443–2451.
- Campagna, A. et al. (2022). Myelodysplastic Syndromes with Isolated 20q Deletion: A New Clinical–Biological Entity?. *Journal of Clinical Medicine*, 11(9), p.2596.
- Carey, B.W., Markoulaki, S., Hanna, J., Saha, K., Gao, Q., Mitalipova, M. and Jaenisch, R. (2009). Reprogramming of murine and human somatic cells using a single polycistronic vector. *Proceedings of the National Academy of Sciences*, 106(1), pp.157–162.

- Carr, M.I. and Jones, S.N. (2016). Regulation of the Mdm2-p53 signaling axis in the DNA damage response and tumorigenesis. *Translational Cancer Research*, 5(6), pp.707–724.
- Carrelha, J., Meng, Y., Kettyle, L.M., Luis, T.C., Norfo, R., Alcolea, V., Boukarabila, H., Grasso, F., Gambardella, A., Grover, A., Högstrand, K., Lord, A.M., Sanjuan-Pla, A., Woll, P.S., Nerlov, C. and Jacobsen, S.E.W. (2018). Hierarchically related lineage-restricted fates of multipotent haematopoietic stem cells. *Nature*, 554(7690), pp.106–111.
- Castaño, J., Bueno, C., Jiménez-Delgado, S., Roca-Ho, H., Fraga, M.F., Fernandez, A.F., Nakanishi, M., Torres-Ruiz, R., Rodríguez-Perales, S. and Menéndez, P. (2017). Generation and characterization of a human iPSC cell line expressing inducible Cas9 in the “safe harbor” AAVS1 locus. *Stem Cell Research*, 21, pp.137–140.
- Castilla, L.H. (2008). C/EBP $\alpha$  in Leukemogenesis: A Matter of Being in the Right Place with the Right Signals. *Cancer Cell*, 13(4), pp.289–291.
- Chakraborty, A. and Twardyabc, D.J. (1998). Stat3 and G-CSF-Induced Myeloid Differentiation. *Leukemia & Lymphoma*, 30(5–6), pp.433–442.
- Challen, G.A., Sun, D., Jeong, M., Luo, M., Jelinek, J., Berg, J.S., Bock, C., Vasanthakumar, A., Gu, H., Xi, Y., Liang, S., Lu, Y., Darlington, G.J., Meissner, A., Issa, J.-P.J., Godley, L.A., Li, W. and Goodell, M.A. (2012). Dnmt3a is essential for hematopoietic stem cell differentiation. *Nature Genetics*, 44(1), pp.23–31.
- Chang, C.-J., Kotini, A.G., Olszewska, M., Georgomanoli, M., Teruya-Feldstein, J., Sperber, H., Sanchez, R., DeVita, R., Martins, T.J., Abdel-Wahab, O., Bradley, R.K. and Papapetrou, E.P. (2018). Dissecting the Contributions of Cooperating Gene Mutations to Cancer Phenotypes and Drug Responses with Patient-Derived iPSCs. *Stem Cell Reports*, 10(5), pp.1610–1624.
- Chen, J., Kao, Y.-R., Sun, D., Todorova, T.I., Reynolds, D., Narayanagari, S.-R., Montagna, C., Will, B., Verma, A. and Steidl, U. (2019). Myelodysplastic syndrome progression to acute myeloid leukemia at the stem cell level. *Nature Medicine*, 25(1), pp.103–110.
- Chen, L., Guo, P., Zhang, Y., Li, X., Jia, P., Tong, J. and Li, J. (2017). Autophagy is an important event for low-dose cytarabine treatment in acute myeloid leukemia cells. *Leukemia Research*, 60, pp.44–52.
- Chen, M., Wu, Y., Zhang, H., Li, S., Zhou, J. and Shen, J. (2020). The Roles of Embryonic Transcription Factor BRACHYURY in Tumorigenesis and Progression. *Frontiers in Oncology*, 10.
- Chen, P., Aimiwu, J., Xie, Z., Wei, X., Liu, S., Klisovic, R., Marcucci, G. and Chan, K.K. (2011). Biochemical Modulation of Aracytidine (Ara-C) Effects by GTI-2040, a Ribonucleotide Reductase Inhibitor, in K562 Human Leukemia Cells. *The AAPS Journal*, 13(1), pp.131–140.
- Christiansen, D.H., Andersen, M.K. and 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(5), pp.1474–1481.

- Churpek, J.E. and Bresnick, E.H. (2019). Transcription factor mutations as a cause of familial myeloid neoplasms. *Journal of Clinical Investigation*, 129(2), pp.476–488.
- Clarke, M., Dumon, S., Ward, C., Jäger, R., Freeman, S., Dawood, B., Sherif, L., Lorvellec, M., Kralovics, R., Frampton, J. and García, P. (2013). MYBL2 haploinsufficiency increases susceptibility to age-related haematopoietic neoplasia. *Leukemia*, 27(3), pp.661–670.
- Cogle, C.R., Kurtin, S.E., Bentley, T.G.K., Broder, M.S., Chang, E., Megaffin, S., Fruchtman, S., Petrone, M.E. and Mukherjee, S. (2017). The Incidence and Health Care Resource Burden of the Myelodysplastic Syndromes in Patients in Whom First-Line Hypomethylating Agents Fail. *The Oncologist*, 22(4), pp.379–385.
- Corces-Zimmerman, M.R. and Majeti, R. (2014). Pre-leukemic evolution of hematopoietic stem cells: the importance of early mutations in leukemogenesis. *Leukemia*, 28(12), pp.2276–2282.
- Cordeiro Gomes, A., Hara, T., Lim, V.Y., Herndler-Brandstetter, D., Nevius, E., Sugiyama, T., Tani-ichi, S., Schlenner, S., Richie, E., Rodewald, H.-R., Flavell, R.A., Nagasawa, T., Ikuta, K. and Pereira, J.P. (2016). Hematopoietic Stem Cell Niches Produce Lineage-Instructive Signals to Control Multipotent Progenitor Differentiation. *Immunity*, 45(6), pp.1219–1231.
- Crispino, J.D. (2005). GATA1 in normal and malignant hematopoiesis. *Seminars in Cell & Developmental Biology*, 16(1), pp.137–147.
- D'Amour, K.A., Agulnick, A.D., Eliazer, S., Kelly, O.G., Kroon, E. and Baetge, E.E. (2005). Efficient differentiation of human embryonic stem cells to definitive endoderm. *Nature Biotechnology*, 23(12), pp.1534–1541.
- Daniel, M.G., Pereira, C.-F., Lemischka, I.R. and Moore, K.A. (2016). Making a Hematopoietic Stem Cell. *Trends in Cell Biology*, 26(3), pp.202–214.
- Daniel-Moreno, A., Lamsfus-Calle, A., Raju, J., Antony, J.S., Handgretinger, R. and Mezger, M. (2019). CRISPR/Cas9-modified hematopoietic stem cells—present and future perspectives for stem cell transplantation. *Bone Marrow Transplantation*, 54(12), pp.1940–1950.
- Delhommeau, F. et al. (2009). Mutation in *TET2* in Myeloid Cancers. *New England Journal of Medicine*, 360(22), pp.2289–2301.
- Desai, P., Mencia-Trinchant, N., Savenkov, O., Simon, M.S., Cheang, G., Lee, S., Samuel, M., Ritchie, E.K., Guzman, M.L., Ballman, K. V., Roboz, G.J. and Hassane, D.C. (2018). Somatic mutations precede acute myeloid leukemia years before diagnosis. *Nature Medicine*, 24(7), pp.1015–1023.
- Dick, J.E. (2003). Self-renewal writ in blood. *Nature*, 423(6937), pp.231–232.
- Dickinson, R.E. et al. (2014). The evolution of cellular deficiency in GATA2 mutation. *Blood*, 123(6), pp.863–874.
- Diesch, J., Zwick, A., Garz, A.-K., Palau, A., Buschbeck, M. and Götze, K.S. (2016). A clinical-molecular update on azanucleoside-based therapy for the treatment of hematologic cancers. *Clinical Epigenetics*, 8(1), p.71.

DiNardo, C.D. and Cortes, J.E. (2016). Mutations in AML: prognostic and therapeutic implications. *Hematology*, 2016(1), pp.348–355.

Ditadi, A., Sturgeon, C.M., Tober, J., Awong, G., Kennedy, M., Yzaguirre, A.D., Azzola, L., Ng, E.S., Stanley, E.G., French, D.L., Cheng, X., Gadue, P., Speck, N.A., Elefanty, A.G. and Keller, G. (2015). Human definitive haemogenic endothelium and arterial vascular endothelium represent distinct lineages. *Nature Cell Biology*, 17(5), pp.580–591.

Dou, D.R., Calvanese, V., Sierra, M.I., Nguyen, A.T., Minasian, A., Saarikoski, P., Sasidharan, R., Ramirez, C.M., Zack, J.A., Crooks, G.M., Galic, Z. and Mikkola, H.K.A. (2016). Medial HOXA genes demarcate haematopoietic stem cell fate during human development. *Nature Cell Biology*, 18(6), pp.595–606.

Drevon, L. et al. (2018). Myelodysplastic syndrome (MDS) with isolated trisomy 8: a type of MDS frequently associated with myeloproliferative features? A report by the Groupe Francophone des Myélodysplasies. *British Journal of Haematology*, 182(6), pp.843–850.

Dufour, A., Schneider, F., Metzeler, K.H., Hoster, E., Schneider, S., Zellmeier, E., Benthaus, T., Sauerland, M.-C., Berdel, W.E., Büchner, T., Wörmann, B., Braess, J., Hiddemann, W., Bohlander, S.K. and Spiekermann, K. (2010). Acute Myeloid Leukemia With Biallelic *CEBPA* Gene Mutations and Normal Karyotype Represents a Distinct Genetic Entity Associated With a Favorable Clinical Outcome. *Journal of Clinical Oncology*, 28(4), pp.570–577.

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

Dzierzak, E. and Philipsen, S. (2013). Erythropoiesis: Development and Differentiation. *Cold Spring Harbor Perspectives in Medicine*, 3(4), pp.a011601–a011601.

Elsaid, R., Soares-da-Silva, F., Peixoto, M., Amiri, D., Mackowski, N., Pereira, P., Bandeira, A. and Cumano, A. (2020). Hematopoiesis: A Layered Organization Across Chordate Species. *Frontiers in Cell and Developmental Biology*, 8.

Engel, I. and Murre, C. (1999). Transcription factors in hematopoiesis. *Current Opinion in Genetics & Development*, 9(5), pp.575–579.

Ernst, T., Chase, A.J., Score, J., Hidalgo-Curtis, C.E., Bryant, C., Jones, A. V, Waghorn, K., Zoi, K., Ross, F.M., Reiter, A., Hochhaus, A., Drexler, H.G., Duncombe, A., Cervantes, F., Oscier, D., Boultonwood, J., Grand, F.H. and Cross, N.C.P. (2010). Inactivating mutations of the histone methyltransferase gene *EZH2* in myeloid disorders. *Nature Genetics*, 42(8), pp.722–726.

Essers, M.A.G., Offner, S., Blanco-Bose, W.E., Waibler, Z., Kalinke, U., Duchosal, M.A. and Trumpp, A. (2009). IFN $\alpha$  activates dormant haematopoietic stem cells in vivo. *Nature*, 458(7240), pp.904–908.

Fajtova, M., Kovarikova, A., Svec, P., Kankuri, E. and Sedlak, J. (2013). Immunophenotypic profile of nucleated erythroid progenitors during maturation in regenerating bone marrow. *Leukemia & Lymphoma*, 54(11), pp.2523–2530.

Fasan, A., Eder, C., Haferlach, C., Grossmann, V., Kohlmann, A., Dicker, F., Kern, W., Haferlach, T. and Schnittger, S. (2013). GATA2 mutations are frequent in intermediate-risk

karyotype AML with biallelic CEBPA mutations and are associated with favorable prognosis. *Leukemia*, 27(2), pp.482–485.

Fenaux, P., Platzbecker, U. and Adeva, L. (2020). How we manage adults with myelodysplastic syndrome. *British Journal of Haematology*, 189(6), pp.1016–1027.

Feng, Y., Zhang, S. and Huang, X. (2014). A robust TALENs system for highly efficient mammalian genome editing. *Scientific Reports*, 4(1), p.3632.

Feusier, J.E., Arunachalam, S., Tashi, T., Baker, M.J., VanSant-Webb, C., Ferdig, A., Welm, B.E., Rodriguez-Flores, J.L., Ours, C., Jorde, L.B., Prechal, J.T. and Mason, C.C. (2021). Large-scale Identification of Clonal Hematopoiesis and Mutations Recurrent in Blood Cancers. *Blood Cancer Discovery*, 2(3), pp.226–237.

Flippe, L., Gaignerie, A., Sérazin, C., Baron, O., Saulquin, X., Themeli, M., Guillonnet, C. and David, L. (2020). Rapid and Reproducible Differentiation of Hematopoietic and T Cell Progenitors From Pluripotent Stem Cells. *Frontiers in Cell and Developmental Biology*, 8.

Fong, H., Hohenstein, K.A. and Donovan, P.J. (2008). Regulation of Self-Renewal and Pluripotency by Sox2 in Human Embryonic Stem Cells. *Stem Cells*, 26(8), pp.1931–1938.

Friedman, Alan D. (2015). C/EBP $\alpha$  in normal and malignant myelopoiesis. *International Journal of Hematology*, 101(4), pp.330–341.

Friedman, Alan D. (2015). C/EBP $\alpha$  in normal and malignant myelopoiesis. *Int J Hematol*, 101, pp.330–341.

Friedman, A.D. (2007). Transcriptional control of granulocyte and monocyte development. *Oncogene*, 26(47), pp.6816–6828.

Friedman, A.D. and McKnight, S.L. (1990). Identification of two polypeptide segments of CCAAT/enhancer-binding protein required for transcriptional activation of the serum albumin gene. *Genes & Development*, 4(8), pp.1416–1426.

Fujiwara, Y., Chang, A.N., Williams, A.M. and Orkin, S.H. (2004). Functional overlap of GATA-1 and GATA-2 in primitive hematopoietic development. *Blood*, 103(2), pp.583–585.

Gambari, R., Finotti, A., Breda, L., Lederer, C., Bianchi, N., Zuccato, C., Klenathous, M. and Rivella, S. (2015). Recent trends in the gene therapy of  $\beta$ -thalassemia. *Journal of Blood Medicine*, February 2015, p.69.

Gammill, L.S. and Sive, H. (1997). Identification of otx2 target genes and restrictions in ectodermal competence during *Xenopus* cement gland formation. *Development*, 124(2), pp.471–481.

Gammill, L.S. and Sive, H. (2001). otx2 Expression in the Ectoderm Activates Anterior Neural Determination and Is Required for *Xenopus* Cement Gland Formation. *Developmental Biology*, 240(1), pp.223–236.

Gañán-Gómez, I., Wei, Y., Yang, H., Pierce, S., Bueso-Ramos, C., Calin, G., Boyano-Adánez, M. del C. and García-Manero, G. (2014). Overexpression of miR-125a in Myelodysplastic



Syndrome CD34+ Cells Modulates NF- $\kappa$ B Activation and Enhances Erythroid Differentiation Arrest. *PLoS ONE*, 9(4), p.e93404.

Gao, X., Johnson, K.D., Chang, Y.-I., Boyer, M.E., Dewey, C.N., Zhang, J. and Bresnick, E.H. (2013). Gata2 cis-element is required for hematopoietic stem cell generation in the mammalian embryo. *Journal of Experimental Medicine*, 210(13), pp.2833–2842.

Garcia-Alegria, E., Menegatti, S., Fadlullah, M.Z.H., Menendez, P., Lacaud, G. and Kouskoff, V. (2018). Early Human Hemogenic Endothelium Generates Primitive and Definitive Hematopoiesis In Vitro. *Stem Cell Reports*, 11(5), pp.1061–1074.

Garcia-Manero, G. (2008). Demethylating agents in myeloid malignancies. *Current Opinion in Oncology*, 20(6), pp.705–710.

Di Genua, C., Valletta, S., Buono, M., Stoilova, B., Sweeney, C., Rodriguez-Meira, A., Grover, A., Drissen, R., Meng, Y., Beveridge, R., Aboukhalil, Z., Karamitros, D., Belderbos, M.E., Bystrykh, L., Thongjuea, S., Vyas, P. and Nerlov, C. (2020). C/EBP $\alpha$  and GATA-2 Mutations Induce Bilineage Acute Erythroid Leukemia through Transformation of a Neomorphic Neutrophil-Erythroid Progenitor. *Cancer Cell*, 37(5), pp.690-704.e8.

Germing, U., Kobbe, G., Haas, R. and Gattermann, N. (2013). Myelodysplastic Syndromes. *Deutsches Ärzteblatt international*, 15 November 2013.

Gewin, V. (2015). Medicine: Expanding possibilities. *Nature*, 528(7580), pp.S10–S11.

Ghosh, D., Kumar, A. and Sinha, N. (2021). Targeted genome editing. In: *Advances in Animal Genomics*. Elsevier, pp.75–89.

Giagounidis, A.A.N., Germing, U., Strupp, C., Hildebrandt, B., Heinsch, M. and Aul, C. (2005). Prognosis of patients with del(5q) MDS and complex karyotype and the possible role of lenalidomide in this patient subgroup. *Annals of Hematology*, 84(9), pp.569–571.

Gillet, J.-P., Varma, S. and Gottesman, M.M. (2013). The Clinical Relevance of Cancer Cell Lines. *JNCI Journal of the National Cancer Institute*, 105(7), pp.452–458.

Goasguen, J.E., Bennett, J.M., Bain, B.J., Brunning, R., Vallespi, M.-T., Tomonaga, M., Zini, G. and Renault, A. (2014). Proposal for refining the definition of dysgranulopoiesis in acute myeloid leukemia and myelodysplastic syndromes. *Leukemia Research*, 38(4), pp.447–453.

Göhring, G. et al. (2010). Complex karyotype newly defined: the strongest prognostic factor in advanced childhood myelodysplastic syndrome. *Blood*, 116(19), pp.3766–3769.

Göllner, S. et al. (2017). Loss of the histone methyltransferase EZH2 induces resistance to multiple drugs in acute myeloid leukemia. *Nature Medicine*, 23(1), pp.69–78.

Gombart, A.F., Hofmann, W.-K., Kawano, S., Takeuchi, S., Krug, U., Kwok, S.H., Larsen, R.J., Asou, H., Miller, C.W., Hoelzer, D. and Koefler, H.P. (2002). Mutations in the gene encoding the transcription factor CCAAT/enhancer binding protein  $\alpha$  in myelodysplastic syndromes and acute myeloid leukemias. *Blood*, 99(4), pp.1332–1340.

- Gossen, M. and Bujard, H. (1992). Tight control of gene expression in mammalian cells by tetracycline-responsive promoters. *Proceedings of the National Academy of Sciences*, 89(12), pp.5547–5551.
- Grass, J.A., Boyer, M.E., Pal, S., Wu, J., Weiss, M.J. and Bresnick, E.H. (2003). GATA-1-dependent transcriptional repression of *GATA-2* via disruption of positive autoregulation and domain-wide chromatin remodeling. *Proceedings of the National Academy of Sciences*, 100(15), pp.8811–8816.
- Green, C.L., Koo, K.K., Hills, R.K., Burnett, A.K., Linch, D.C. and Gale, R.E. (2010). Prognostic Significance of *CEBPA* Mutations in a Large Cohort of Younger Adult Patients With Acute Myeloid Leukemia: Impact of Double *CEBPA* Mutations and the Interaction With *FLT3* and *NPM1* Mutations. *Journal of Clinical Oncology*, 28(16), pp.2739–2747.
- Greenberg, P., Cox, C., LeBeau, M.M., Fenaux, P., Morel, P., Sanz, G., Sanz, M., Vallespi, T., Hamblin, T., Oscier, D., Ohyashiki, K., Toyama, K., Aul, C., Mufti, G. and Bennett, J. (1997). International Scoring System for Evaluating Prognosis in Myelodysplastic Syndromes. *Blood*, 89(6), pp.2079–2088.
- Greenberg, P.L. et al. (2012). Revised International Prognostic Scoring System for Myelodysplastic Syndromes. *Blood*, 120(12), pp.2454–2465.
- Grinfeld, J. et al. (2018). Classification and Personalized Prognosis in Myeloproliferative Neoplasms. *New England Journal of Medicine*, 379(15), pp.1416–1430.
- Gritz, E. and Hirschi, K.K. (2016). Specification and function of hemogenic endothelium during embryogenesis. *Cellular and Molecular Life Sciences*, 73(8), pp.1547–1567.
- Grossmann, V., Bacher, U., Kohlmann, A., Butschalowski, K., Roller, A., Jeromin, S., Dicker, F., Kern, W., Schnittger, S., Haferlach, T. and Haferlach, C. (2012). Expression of CEBPA is reduced in RUNX1-mutated acute myeloid leukemia. *Blood Cancer Journal*, 2(8), pp.e86–e86.
- Grossmann, V., Schnittger, S., Schindela, S., Klein, H.-U., Eder, C., Dugas, M., Kern, W., Haferlach, T., Haferlach, C. and Kohlmann, A. (2011). Strategy for Robust Detection of Insertions, Deletions, and Point Mutations in CEBPA, a GC-Rich Content Gene, Using 454 Next-Generation Deep-Sequencing Technology. *The Journal of Molecular Diagnostics*, 13(2), pp.129–136.
- Guan, J. et al. (2022). Chemical reprogramming of human somatic cells to pluripotent stem cells. *Nature*, 605(7909), pp.325–331.
- Guo, H., Ma, O., Speck, N.A. and Friedman, A.D. (2012). Runx1 deletion or dominant inhibition reduces Cebpa transcription via conserved promoter and distal enhancer sites to favor monopoiesis over granulopoiesis. *Blood*, 119(19), pp.4408–4418.
- Gutiérrez, L., Nikolic, T., van Dijk, T.B., Hammad, H., Vos, N., Willart, M., Grosveld, F., Philipson, S. and Lambrecht, B.N. (2007). Gata1 regulates dendritic-cell development and survival. *Blood*, 110(6), pp.1933–1941.
- Haase, D. et al. (2007). New insights into the prognostic impact of the karyotype in MDS and correlation with subtypes: evidence from a core dataset of 2124 patients. *Blood*, 110(13), pp.4385–4395.

- Haase, D. et al. (2019). TP53 mutation status divides myelodysplastic syndromes with complex karyotypes into distinct prognostic subgroups. *Leukemia*, 33(7), pp.1747–1758.
- Hansen, M., Varga, E., Aarts, C., Wust, T., Kuijpers, T., von Lindern, M. and van den Akker, E. (2018). Efficient production of erythroid, megakaryocytic and myeloid cells, using single cell-derived iPSC colony differentiation. *Stem Cell Research*, 29, pp.232–244.
- Harada, H., Harada, Y., Niimi, H., Kyo, T., Kimura, A. and 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(6), pp.2316–2324.
- Hasemann, M.S., Lauridsen, F.K.B., Waage, J., Jakobsen, J.S., Frank, A.-K., Schuster, M.B., Rapin, N., Bagger, F.O., Hoppe, P.S., Schroeder, T. and Porse, B.T. (2014). C/EBP $\alpha$  Is Required for Long-Term Self-Renewal and Lineage Priming of Hematopoietic Stem Cells and for the Maintenance of Epigenetic Configurations in Multipotent Progenitors. *PLoS Genetics*, 10(1), p.e1004079.
- Hattangadi, S.M., Wong, P., Zhang, L., Flygare, J. and Lodish, H.F. (2011). From stem cell to red cell: regulation of erythropoiesis at multiple levels by multiple proteins, RNAs, and chromatin modifications. *Blood*, 118(24), pp.6258–6268.
- Hauke, R.J. and Tarantolo, S.R. (2000). Hematopoietic Growth Factors. *Laboratory Medicine*, 31(11), pp.613–615.
- Heath, V., Suh, H.C., Holman, M., Renn, K., Gooya, J.M., Parkin, S., Klarmann, K.D., Ortiz, M., Johnson, P. and Keller, J. (2004). C/EBP $\alpha$  deficiency results in hyperproliferation of hematopoietic progenitor cells and disrupts macrophage development in vitro and in vivo. *Blood*, 104(6), pp.1639–1647.
- Heinrichs, S., Conover, L.F., Bueso-Ramos, C.E., Kilpivaara, O., Stevenson, K., Neuberg, D., Loh, M.L., Wu, W.-S., Rodig, S.J., Garcia-Manero, G., Kantarjian, H.M. and Look, A.T. (2013). MYBL2 is a sub-haploinsufficient tumor suppressor gene in myeloid malignancy. *eLife*, 2.
- Heyes, E., Schmidt, L., Manhart, G., Eder, T., Proietti, L. and Grebien, F. (2021). Identification of gene targets of mutant C/EBP $\alpha$  reveals a critical role for MSI2 in CEBPA-mutated AML. *Leukemia*, 35(9), pp.2526–2538.
- Hiddemann, W. (1991). Cytosine arabinoside in the treatment of acute myeloid leukemia: The role and place of high-dose regimens. *Annals of Hematology*, 62(4), pp.119–128.
- Higa, K.C., Goodspeed, A., Chavez, J.S., De Dominici, M., Danis, E., Zaberezhnyy, V., Rabe, J.L., Tenen, D.G., Pietras, E.M. and DeGregori, J. (2021). Chronic interleukin-1 exposure triggers selection for *Cebpa* -knockout multipotent hematopoietic progenitors. *Journal of Experimental Medicine*, 218(6).
- Hirai, H., Kobayashi, Y., Mano, H., Hagiwara, K., Maru, Y., Omine, M., Mizoguchi, H., Nishida, J. and Takaku, F. (1987). A point mutation at codon 13 of the N-ras oncogene in myelodysplastic syndrome. *Nature*, 327(6121), pp.430–432.
- Hirsch, P. et al. (2016). Genetic hierarchy and temporal variegation in the clonal history of acute myeloid leukaemia. *Nature Communications*, 7(1), p.12475.

- Hoban, M.D., Orkin, S.H. and Bauer, D.E. (2016). Genetic treatment of a molecular disorder: gene therapy approaches to sickle cell disease. *Blood*, 127(7), pp.839–848.
- Hockings, C., Deaner, V., Hoade, Y., Dace, P., Lubin, A. and Payne, E. (2018). A Zebrafish Model of Cooperating C and N Terminal CEBPA Mutations Reveals Defects in Early Myelopoiesis and HSPCs Leading to Leukaemogenesis. *Blood*, 132(Supplement 1), pp.1343–1343.
- Hoeffel, G. et al. (2015). C-Myb<sup>+</sup> Erythro-Myeloid Progenitor-Derived Fetal Monocytes Give Rise to Adult Tissue-Resident Macrophages. *Immunity*, 42(4), pp.665–678.
- Hofmann, W.-K., Lübbert, M., Hoelzer, D. and Phillip Koeffler, H. (2004). Myelodysplastic syndromes. *The Hematology Journal*, 5(1), pp.1–8.
- Hong, M. and He, G. (2017). The 2016 revision to the World Health Organization classification of myelodysplastic syndromes. *Journal of Translational Internal Medicine*, 5(3), pp.139–143.
- Hou, Z., Zhang, Y., Propson, N.E., Howden, S.E., Chu, L.-F., Sontheimer, E.J. and Thomson, J.A. (2013). Efficient genome engineering in human pluripotent stem cells using Cas9 from *Neisseria meningitidis*. *Proceedings of the National Academy of Sciences*, 110(39), pp.15644–15649.
- Hsu, J., Reilly, A., Hayes, B.J., Clough, C.A., Konnick, E.Q., Torok-Storb, B., Gulsuner, S., Wu, D., Becker, P.S., Keel, S.B., Abkowitz, J.L. and Doulatov, S. (2019). Reprogramming identifies functionally distinct stages of clonal evolution in myelodysplastic syndromes. *Blood*, 134(2), pp.186–198.
- Huang, X., Wang, Y., Yan, W., Smith, C., Ye, Z., Wang, J., Gao, Y., Mendelsohn, L. and Cheng, L. (2015). Production of Gene-Corrected Adult Beta Globin Protein in Human Erythrocytes Differentiated from Patient iPSCs After Genome Editing of the Sickle Point Mutation. *Stem Cells*, 33(5), pp.1470–1479.
- Hughes, A.E.O. et al. (2014). Clonal Architecture of Secondary Acute Myeloid Leukemia Defined by Single-Cell Sequencing. *PLoS Genetics*, 10(7), p.e1004462.
- Ikuta, K. and Weissman, I.L. (1992). Evidence that hematopoietic stem cells express mouse c-kit but do not depend on steel factor for their generation. *Proceedings of the National Academy of Sciences*, 89(4), pp.1502–1506.
- Invernizzi, R., Quaglia, F. and Della Porta, M.G. (2015). IMPORTANCE OF CLASSICAL MORPHOLOGY IN THE DIAGNOSIS OF MYELODYSPLASTIC SYNDROME. *Mediterranean Journal of Hematology and Infectious Diseases*, 7, p.e2015035.
- Ishino, Y., Shinagawa, H., Makino, K., Amemura, M. and Nakata, A. (1987). Nucleotide sequence of the iap gene, responsible for alkaline phosphatase isozyme conversion in *Escherichia coli*, and identification of the gene product. *Journal of Bacteriology*, 169(12), pp.5429–5433.
- Iwama, A., Oguro, H., Negishi, M., Kato, Y., Morita, Y., Tsukui, H., Ema, H., Kamijo, T., Katoh-Fukui, Y., Koseki, H., van Lohuizen, M. and Nakauchi, H. (2004). Enhanced Self-Renewal of Hematopoietic Stem Cells Mediated by the Polycomb Gene Product Bmi-1. *Immunity*, 21(6), pp.843–851.

- Iwasaki, H., Mizuno, S., Wells, R.A., Cantor, A.B., Watanabe, S. and Akashi, K. (2003). GATA-1 Converts Lymphoid and Myelomonocytic Progenitors into the Megakaryocyte/Erythrocyte Lineages. *Immunity*, 19(3), pp.451–462.
- Iwasaki, H., Mizuno, S., Arinobu, Y., Ozawa, H., Mori, Y., Shigematsu, H., Takatsu, K., Tenen, D.G. and Akashi, K. (2006). The order of expression of transcription factors directs hierarchical specification of hematopoietic lineages. *Genes & Development*, 20(21), pp.3010–3021.
- Jagannathan-Bogdan, M. and Zon, L.I. (2013). Hematopoiesis. *Development*, 140(12), pp.2463–2467.
- Jaiswal, S. et al. (2014). Age-Related Clonal Hematopoiesis Associated with Adverse Outcomes. *New England Journal of Medicine*, 371(26), pp.2488–2498.
- Johansen, L.M., Iwama, A., Lodie, T.A., Sasaki, K., Felsher, D.W., Golub, T.R. and Tenen, D.G. (2001). c-Myc Is a Critical Target for C/EBP $\alpha$  in Granulopoiesis. *Molecular and Cellular Biology*, 21(11), pp.3789–3806.
- Jung, M., Cordes, S., Zou, J., Yu, S.J., Guitart, X., Hong, S.G., Dang, V., Kang, E., Donaires, F.S., Hassan, S.A., Albitar, M., Hsu, A.P., Holland, S.M., Hickstein, D.D., Townsley, D., Dunbar, C.E. and Winkler, T. (2018). GATA2 deficiency and human hematopoietic development modeled using induced pluripotent stem cells. *Blood Advances*, 2(23), pp.3553–3565.
- Kanai-Azuma, M., Kanai, Y., Gad, J.M., Tajima, Y., Taya, C., Kurohmaru, M., Sanai, Y., Yonekawa, H., Yazaki, K., Tam, P.P.L. and Hayashi, Y. (2002). Depletion of definitive gut endoderm in *Sox17* -null mutant mice. *Development*, 129(10), pp.2367–2379.
- Kantarjian, H., O'Brien, S., Ravandi, F., Borthakur, G., Faderl, S., Bueso-Ramos, C., Abruzzo, L., Pierce, S., Shan, J., Issa, J.-P. and Garcia-Manero, G. (2009). The heterogeneous prognosis of patients with myelodysplastic syndrome and chromosome 5 abnormalities. *Cancer*, 115(22), pp.5202–5209.
- Kantor, A., McClements, M. and MacLaren, R. (2020). CRISPR-Cas9 DNA Base-Editing and Prime-Editing. *International Journal of Molecular Sciences*, 21(17), p.6240.
- Kato, N., Kitaura, J., Doki, N., Komeno, Y., Watanabe-Okochi, N., Togami, K., Nakahara, F., Oki, T., Enomoto, Y., Fukuchi, Y., Nakajima, H., Harada, Y., Harada, H. and Kitamura, T. (2011). Two types of C/EBP $\alpha$  mutations play distinct but collaborative roles in leukemogenesis: lessons from clinical data and BMT models. *Blood*, 117(1), pp.221–233.
- Katsumura, K.R. and Bresnick, E.H. (2017). The GATA factor revolution in hematology. *Blood*, 129(15), pp.2092–2102.
- Keeshan, K., Santilli, G., Corradini, F., Perrotti, D. and Calabretta, B. (2003). Transcription activation function of C/EBP $\alpha$  is required for induction of granulocytic differentiation. *Blood*, 102(4), pp.1267–1275.
- Kessel, K.U., Bluemke, A., Schöler, H.R., Zaehres, H., Schlenke, P. and Dorn, I. (2017). Emergence of CD43-Expressing Hematopoietic Progenitors from Human Induced Pluripotent Stem Cells. *Transfusion Medicine and Hemotherapy*, 44(3), pp.143–150.

- Ketley, N.J. and Newland, A.C. (1997). Haemopoietic growth factors. *Postgraduate Medical Journal*, 73(858), pp.215–221.
- Khalil, A.M. (2020). The genome editing revolution: review. *Journal of Genetic Engineering and Biotechnology*, 18(1), p.68.
- Khan, A.O. et al. (2023). Human Bone Marrow Organoids for Disease Modeling, Discovery, and Validation of Therapeutic Targets in Hematologic Malignancies. *Cancer Discovery*, 13(2), pp.364–385.
- Kiel, M.J., Yilmaz, Ö.H., Iwashita, T., Yilmaz, O.H., Terhorst, C. and Morrison, S.J. (2005). SLAM Family Receptors Distinguish Hematopoietic Stem and Progenitor Cells and Reveal Endothelial Niches for Stem Cells. *Cell*, 121(7), pp.1109–1121.
- Killick, S.B. et al. (2021). British Society for Haematology guidelines for the management of adult myelodysplastic syndromes. *British Journal of Haematology*, 194(2), pp.267–281.
- King, A. and Shenoy, S. (2014). Evidence-based focused review of the status of hematopoietic stem cell transplantation as treatment of sickle cell disease and thalassemia. *Blood*, 123(20), pp.3089–3094.
- Klaver-Flores, S., Zittersteijn, H.A., Canté-Barrett, K., Lankester, A., Hoeben, R.C., Gonçalves, M.A.F. V., Pike-Overzet, K. and Staal, F.J.T. (2021). Genomic Engineering in Human Hematopoietic Stem Cells: Hype or Hope?. *Frontiers in Genome Editing*, 2.
- Knott, G.J. and Doudna, J.A. (2018). CRISPR-Cas guides the future of genetic engineering. *Science*, 361(6405), pp.866–869.
- Koeffler, H.P. and Golde, D.W. (1978). Acute Myelogenous Leukemia: A Human Cell Line Responsive to Colony-Stimulating Activity. *Science*, 200(4346), pp.1153–1154.
- Kondo, M., Weissman, I.L. and Akashi, K. (1997). Identification of Clonogenic Common Lymphoid Progenitors in Mouse Bone Marrow. *Cell*, 91(5), pp.661–672.
- Kontandreopoulou, C.-N., Kalopisis, K., Viniou, N.-A. and Diamantopoulos, P. (2022). The genetics of myelodysplastic syndromes and the opportunities for tailored treatments. *Frontiers in Oncology*, 12.
- Koschmieder, S., Halmos, B., Levantini, E. and Tenen, D.G. (2009). Dysregulation of the C/EBP $\alpha$  Differentiation Pathway in Human Cancer. *Journal of Clinical Oncology*, 27(4), pp.619–628.
- Kotini, A.G. et al. (2017). Stage-Specific Human Induced Pluripotent Stem Cells Map the Progression of Myeloid Transformation to Transplantable Leukemia. *Cell Stem Cell*, 20(3), pp.315–328.e7.
- Koury, S., Koury, M., Bondurant, M., Caro, J. and Graber, S. (1989). Quantitation of erythropoietin-producing cells in kidneys of mice by in situ hybridization: correlation with hematocrit, renal erythropoietin mRNA, and serum erythropoietin concentration. *Blood*, 74(2), pp.645–651.

- Lam, A.Q., Freedman, B.S., Morizane, R., Lerou, P.H., Valerius, M.T. and Bonventre, J. V. (2014). Rapid and Efficient Differentiation of Human Pluripotent Stem Cells into Intermediate Mesoderm That Forms Tubules Expressing Kidney Proximal Tubular Markers. *Journal of the American Society of Nephrology*, 25(6), pp.1211–1225.
- Lamba, J.K. (2009). Genetic factors influencing cytarabine therapy. *Pharmacogenomics*, 10(10), pp.1657–74.
- Lancaster, M.A. and Knoblich, J.A. (2014). Generation of cerebral organoids from human pluripotent stem cells. *Nature Protocols*, 9(10), pp.2329–2340.
- Lancrin, C., Sroczynska, P., Stephenson, C., Allen, T., Kouskoff, V. and Lacaud, G. (2009). The haemangioblast generates haematopoietic cells through a haemogenic endothelium stage. *Nature*, 457(7231), pp.892–895.
- Larizza, L., Magnani, I. and Beghini, A. (2005). The Kasumi-1 cell line: a t(8;21)-kit mutant model for acute myeloid leukemia. *Leukemia & Lymphoma*, 46(2), pp.247–255.
- Latham, K.E., Litvin, J., Orth, J.M., Patel, B., Mettus, R. and Reddy, E.P. (1996). Temporal patterns of A-myb and B-myb gene expression during testis development. *Oncogene*, 13(6), pp.1161–8.
- Laurent, L.C. et al. (2011). Dynamic Changes in the Copy Number of Pluripotency and Cell Proliferation Genes in Human ESCs and iPSCs during Reprogramming and Time in Culture. *Cell Stem Cell*, 8(1), pp.106–118.
- Leonard, M., Brice, M., Engel, J. and Papayannopoulou, T. (1993). Dynamics of GATA transcription factor expression during erythroid differentiation. *Blood*, 82(4), pp.1071–1079.
- Leroy, H., Roumier, C., Huyghe, P., Biggio, V., Fenaux, P. and Preudhomme, C. (2005). CEBPA point mutations in hematological malignancies. *Leukemia*, 19(3), pp.329–334.
- Li, H., Yang, Y., Hong, W., Huang, M., Wu, M. and Zhao, X. (2020). Applications of genome editing technology in the targeted therapy of human diseases: mechanisms, advances and prospects. *Signal Transduction and Targeted Therapy*, 5(1), p.1.
- Li, J., Hale, J., Bhagia, P., Xue, F., Chen, L., Jaffray, J., Yan, H., Lane, J., Gallagher, P.G., Mohandas, N., Liu, J. and An, X. (2014). Isolation and transcriptome analyses of human erythroid progenitors: BFU-E and CFU-E. *Blood*, 124(24), pp.3636–3645.
- Li, L., Yi, H., Liu, Z., Long, P., Pan, T., Huang, Y., Li, Y., Li, Q. and Ma, Y. (2022). Genetic correction of concurrent  $\alpha$ - and  $\beta$ -thalassemia patient-derived pluripotent stem cells by the CRISPR-Cas9 technology. *Stem Cell Research & Therapy*, 13(1), p.102.
- Li, Q., Bohin, N., Wen, T., Ng, V., Magee, J., Chen, S.-C., Shannon, K. and Morrison, S.J. (2013). Oncogenic Nras has bimodal effects on stem cells that sustainably increase competitiveness. *Nature*, 504(7478), pp.143–147.
- Liang, H.-B. et al. (2017). MYBL2 is a Potential Prognostic Marker that Promotes Cell Proliferation in Gallbladder Cancer. *Cellular Physiology and Biochemistry*, 41(5), pp.2117–2131.

- List, A., Ebert, B.L. and Fenaux, P. (2018). A decade of progress in myelodysplastic syndrome with chromosome 5q deletion. *Leukemia*, 32(7), pp.1493–1499.
- Livak, K.J. and Schmittgen, T.D. (2001). Analysis of Relative Gene Expression Data Using Real-Time Quantitative PCR and the  $2^{-\Delta\Delta CT}$  Method. *Methods*, 25(4), pp.402–408.
- Lloberas, J., Soler, C. and Celada, A. (1999). The key role of PU.1/SPI-1 in B cells, myeloid cells and macrophages. *Immunology Today*, 20(4), pp.184–189.
- Lorvellec, M., Dumon, S., Maya-Mendoza, A., Jackson, D., Frampton, J. and Garcia, P. (2010). B-Myb is Critical for Proper DNA Duplication During an Unperturbed S Phase in Mouse Embryonic Stem Cells. *Stem Cells*, 28(10), pp.1751–1759.
- De Los Angeles, A., Loh, Y.-H., Tesar, P.J. and Daley, G.Q. (2012). Accessing naïve human pluripotency. *Current Opinion in Genetics & Development*, 22(3), pp.272–282.
- Lowenberg, B., Downing, J.R. and Burnett, A. (1999). Acute Myeloid Leukemia. *New England Journal of Medicine*, 341(14), pp.1051–1062.
- Lu, R., Neff, N.F., Quake, S.R. and Weissman, I.L. (2011). Tracking single hematopoietic stem cells in vivo using high-throughput sequencing in conjunction with viral genetic barcoding. *Nature Biotechnology*, 29(10), pp.928–933.
- Luo, W., Li, S., Peng, B., Ye, Y., Deng, X. and Yao, K. (2013). Embryonic Stem Cells Markers SOX2, OCT4 and Nanog Expression and Their Correlations with Epithelial-Mesenchymal Transition in Nasopharyngeal Carcinoma. *PLoS ONE*, 8(2), p.e56324.
- Ma, O., Hong, S., Guo, H., Ghiaur, G. and Friedman, A.D. (2014). Granulopoiesis Requires Increased C/EBP $\alpha$  Compared to Monopoiesis, Correlated with Elevated Cebpa in Immature G-CSF Receptor versus M-CSF Receptor Expressing Cells. *PLoS ONE*, 9(4), p.e95784.
- Macri, S., Pavesi, E., Crescitelli, R., Aspesi, A., Vizziello, C., Botto, C., Corti, P., Quarello, P., Notari, P., Ramenghi, U., Ellis, S.R. and Dianzani, I. (2015). Immunophenotypic Profiling of Erythroid Progenitor-Derived Extracellular Vesicles in Diamond-Blackfan Anaemia: A New Diagnostic Strategy. *PLOS ONE*, 10(9), p.e0138200.
- Makishima, H. et al. (2017). Dynamics of clonal evolution in myelodysplastic syndromes. *Nature Genetics*, 49(2), pp.204–212.
- Malcovati, L. et al. (2017). Clinical significance of somatic mutation in unexplained blood cytopenia. *Blood*, 129(25), pp.3371–3378.
- Malcovati, L. et al. (2014). Driver somatic mutations identify distinct disease entities within myeloid neoplasms with myelodysplasia. *Blood*, 124(9), pp.1513–1521.
- Malcovati, L., Germing, U., Kuendgen, A., Della Porta, M.G., Pascutto, C., Invernizzi, R., Giagounidis, A., Hildebrandt, B., Bernasconi, P., Knipp, S., Strupp, C., Lazzarino, M., Aul, C. and Cazzola, M. (2007). Time-Dependent Prognostic Scoring System for Predicting Survival and Leukemic Evolution in Myelodysplastic Syndromes. *Journal of Clinical Oncology*, 25(23), pp.3503–3510.



- Mannelli, F. et al. (2017). *CEBPA* –double-mutated acute myeloid leukemia displays a unique phenotypic profile: a reliable screening method and insight into biological features. *Haematologica*, 102(3), pp.529–540.
- Manz, M.G., Miyamoto, T., Akashi, K. and Weissman, I.L. (2002). Prospective isolation of human clonogenic common myeloid progenitors. *Proceedings of the National Academy of Sciences*, 99(18), pp.11872–11877.
- Marcucci, G., Maharry, K., Radmacher, M.D., Mrózek, K., Vukosavljevic, T., Paschka, P., Whitman, S.P., Langer, C., Baldus, C.D., Liu, C.-G., Ruppert, A.S., Powell, B.L., Carroll, A.J., Caligiuri, M.A., Kolitz, J.E., Larson, R.A. and Bloomfield, C.D. (2008). Prognostic Significance of, and Gene and MicroRNA Expression Signatures Associated With, *CEBPA* Mutations in Cytogenetically Normal Acute Myeloid Leukemia With High-Risk Molecular Features: A Cancer and Leukemia Group B Study. *Journal of Clinical Oncology*, 26(31), pp.5078–5087.
- Marechal, A. and Zou, L. (2013). DNA Damage Sensing by the ATM and ATR Kinases. *Cold Spring Harbor Perspectives in Biology*, 5(9), pp.a012716–a012716.
- Martin, D.I.K., Zon, L.I., Mutter, G. and Orkin, S.H. (1990). Expression of an erythroid transcription factor in megakaryocytic and mast cell lineages. *Nature*, 344(6265), pp.444–447.
- Martinez, I. and DiMaio, D. (2011). B-Myb, Cancer, Senescence, and MicroRNAs. *Cancer Research*, 71(16), pp.5370–5373.
- Martinez-Høyer, S. and Karsan, A. (2020). Mechanisms of lenalidomide sensitivity and resistance. *Experimental Hematology*, 91, pp.22–31.
- Martins, T.M. da M., de Paula, A.C.C., Gomes, D.A. and Goes, A.M. (2014). Alkaline Phosphatase Expression/Activity and Multilineage Differentiation Potential are the Differences Between Fibroblasts and Orbital Fat-Derived Stem Cells – A Study in Animal Serum-Free Culture Conditions. *Stem Cell Reviews and Reports*, 10(5), pp.697–711.
- Maruyama, H., Ishitsuka, Y., Fujisawa, Y., Furuta, J., Sekido, M. and Kawachi, Y. (2014). B-Myb enhances proliferation and suppresses differentiation of keratinocytes in three-dimensional cell culture. *Archives of Dermatological Research*, 306(4), pp.375–384.
- De Masi, C., Spitalieri, P., Murdocca, M., Novelli, G. and Sangiuolo, F. (2020). Application of CRISPR/Cas9 to human-induced pluripotent stem cells: from gene editing to drug discovery. *Human Genomics*, 14(1), p.25.
- Mason, C.C. et al. (2016). Age-related mutations and chronic myelomonocytic leukemia. *Leukemia*, 30(4), pp.906–913.
- Mawaribuchi, S., Aiki, Y., Ikeda, N. and Ito, Y. (2019). mRNA and miRNA expression profiles in an ectoderm-biased substate of human pluripotent stem cells. *Scientific Reports*, 9(1), p.11910.
- McGowan, K.A., Pang, W.W., Bhardwaj, R., Perez, M.G., Pluvinau, J. V., Glader, B.E., Malek, R., Mendrysa, S.M., Weissman, I.L., Park, C.Y. and Barsh, G.S. (2011). Reduced ribosomal protein gene dosage and p53 activation in low-risk myelodysplastic syndrome. *Blood*, 118(13), pp.3622–3633.

- McGrath, K.E., Frame, J.M., Fegan, K.H., Bowen, J.R., Conway, S.J., Catherman, S.C., Kingsley, P.D., Koniski, A.D. and Palis, J. (2015a). Distinct Sources of Hematopoietic Progenitors Emerge before HSCs and Provide Functional Blood Cells in the Mammalian Embryo. *Cell Reports*, 11(12), pp.1892–1904.
- McGrath, K.E., Frame, J.M., Fegan, K.H., Bowen, J.R., Conway, S.J., Catherman, S.C., Kingsley, P.D., Koniski, A.D. and Palis, J. (2015b). Distinct Sources of Hematopoietic Progenitors Emerge before HSCs and Provide Functional Blood Cells in the Mammalian Embryo. *Cell Reports*, 11(12), pp.1892–1904.
- McKinney-Freeman, S.L., Naveiras, O., Yates, F., Loewer, S., Philitas, M., Curran, M., Park, P.J. and Daley, G.Q. (2009). Surface antigen phenotypes of hematopoietic stem cells from embryos and murine embryonic stem cells. *Blood*, 114(2), pp.268–278.
- Medvinsky, A. and Dzierzak, E. (1996). Definitive Hematopoiesis Is Autonomously Initiated by the AGM Region. *Cell*, 86(6), pp.897–906.
- Mendoza, H., Podoltsev, N.A. and Siddon, A.J. (2021). Laboratory evaluation and prognostication among adults and children with *CEBPA* -mutant acute myeloid leukemia. *International Journal of Laboratory Hematology*, 43(S1), pp.86–95.
- Menssen, A.J. and Walter, M.J. (2020). Genetics of progression from MDS to secondary leukemia. *Blood*, 136(1), pp.50–60.
- Miyamoto, T., Iwasaki, H., Reizis, B., Ye, M., Graf, T., Weissman, I.L. and Akashi, K. (2002). Myeloid or Lymphoid Promiscuity as a Critical Step in Hematopoietic Lineage Commitment. *Developmental Cell*, 3(1), pp.137–147.
- Mohrin, M., Bourke, E., Alexander, D., Warr, M.R., Barry-Holson, K., Le Beau, M.M., Morrison, C.G. and Passegué, E. (2010). Hematopoietic Stem Cell Quiescence Promotes Error-Prone DNA Repair and Mutagenesis. *Cell Stem Cell*, 7(2), pp.174–185.
- Mollgard, L., Saft, L., Treppendahl, M.B., Dybedal, I., Norgaard, J.M., Astermark, J., Ejerblad, E., Garelius, H., Dufva, I.H., Jansson, M., Jadersten, M., Kjeldsen, L., Linder, O., Nilsson, L., Vestergaard, H., Porwit, A., Gronbaek, K. and Lindberg, E.H. (2011). Clinical effect of increasing doses of lenalidomide in high-risk myelodysplastic syndrome and acute myeloid leukemia with chromosome 5 abnormalities. *Haematologica*, 96(7), pp.963–971.
- Moradi, S., Mahdizadeh, H., Šarić, T., Kim, J., Harati, J., Shahsavarani, H., Greber, B. and Moore, J.B. (2019). Research and therapy with induced pluripotent stem cells (iPSCs): social, legal, and ethical considerations. *Stem Cell Research & Therapy*, 10(1), p.341.
- Moran-Crusio, K. et al. (2011). Tet2 Loss Leads to Increased Hematopoietic Stem Cell Self-Renewal and Myeloid Transformation. *Cancer Cell*, 20(1), pp.11–24.
- Mortensen, A.H., Schade, V., Lamonerie, T. and Camper, S.A. (2015). Deletion of OTX2 in neural ectoderm delays anterior pituitary development. *Human Molecular Genetics*, 24(4), pp.939–953.
- Mucenski, M.L., McLain, K., Kier, A.B., Swerdlow, S.H., Schreiner, C.M., Miller, T.A., Pietryga, D.W., Scott, W.J. and Potter, S.S. (1991). A functional c-myc gene is required for normal murine fetal hepatic hematopoiesis. *Cell*, 65(4), pp.677–689.

- Mufti, G.J., Bennett, J.M., Goasguen, J., Bain, B.J., Baumann, I., Brunning, R., Cazzola, M., Fenaux, P., Germing, U., Hellstrom-Lindberg, E., Jinnai, I., Manabe, A., Matsuda, A., Niemeyer, C.M., Sanz, G., Tomonaga, M., Vallespi, T. and Yoshimi, A. (2008). Diagnosis and classification of myelodysplastic syndrome: International Working Group on Morphology of myelodysplastic syndrome (IWGM-MDS) consensus proposals for the definition and enumeration of myeloblasts and ring sideroblasts. *Haematologica*, 93(11), pp.1712–1717.
- Müller, C., Calkhoven, C.F., Sha, X. and Leutz, A. (2004). The CCAAT Enhancer-binding Protein  $\alpha$  (C/EBP $\alpha$ ) Requires a SWI/SNF Complex for Proliferation Arrest. *Journal of Biological Chemistry*, 279(8), pp.7353–7358.
- Musa, J., Aynaud, M.-M., Mirabeau, O., Delattre, O. and Grünewald, T.G. (2017). MYBL2 (B-Myb): a central regulator of cell proliferation, cell survival and differentiation involved in tumorigenesis. *Cell Death & Disease*, 8(6), pp.e2895–e2895.
- Nakagawa, T., Matozaki, S., Murayama, T., Nishimura, R., Tsutsumi, M., Kawaguchi, R., Yokoyama, Y., Hikiji, K., Isobe, T. and Chihara, K. (2008). Establishment of a leukaemic cell line from a patient with acquisition of chromosomal abnormalities during disease progression in myelodysplastic syndrome. *British Journal of Haematology*, 85(3), pp.469–476.
- Nakajima, H. (2021). Molecular Pathogenesis and Treatment of Myelodysplastic Syndromes. *Internal Medicine*, 60(1), pp.15–23.
- Nazha, A., Al-Issa, K., Hamilton, B.K., Radivoyevitch, T., Gerds, A.T., Mukherjee, S., Adema, V., Zarzour, A., Abuhadra, N., Patel, B.J., Hirsch, C.M., Advani, A., Przychodzen, B., Carraway, H.E., Maciejewski, J.P. and Sekeres, M.A. (2017). Adding molecular data to prognostic models can improve predictive power in treated patients with myelodysplastic syndromes. *Leukemia*, 31(12), pp.2848–2850.
- Nazha, A., Sekeres, M.A., Bejar, R., Rauh, M.J., Othus, M., Komrokji, R.S., Barnard, J., Hilton, C.B., Kerr, C.M., Steensma, D.P., DeZern, A., Roboz, G., Garcia-Manero, G., Erba, H., Ebert, B.L. and Maciejewski, J.P. (2019). Genomic Biomarkers to Predict Resistance to Hypomethylating Agents in Patients With Myelodysplastic Syndromes Using Artificial Intelligence. *JCO Precision Oncology*, December 2019, pp.1–11.
- Neo, W.H., Lie-A-Ling, M., Fadlullah, M.Z.H. and Lacaud, G. (2021). Contributions of Embryonic HSC-Independent Hematopoiesis to Organogenesis and the Adult Hematopoietic System. *Frontiers in Cell and Developmental Biology*, 9.
- Nerlov, C. (2004a). C/EBP $\alpha$  mutations in acute myeloid leukaemias. *Nature Reviews Cancer*, 4(5), pp.394–400.
- Nerlov, C. (2004b). C/EBP $\alpha$  mutations in acute myeloid leukaemias. *Nature Reviews Cancer*, 4(5), pp.394–400.
- Nerlov, C. (2007). The C/EBP family of transcription factors: a paradigm for interaction between gene expression and proliferation control. *Trends in Cell Biology*, 17(7), pp.318–324.
- Ness, S.A. (2003). Myb protein specificity: evidence of a context-specific transcription factor code. *Blood Cells, Molecules, and Diseases*, 31(2), pp.192–200.

- Ngo, N., Patel, K., Isaacson, P.G. and Naresh, K.N. (2006). Leucocyte common antigen (CD45) and CD5 positivity in an 'undifferentiated' carcinoma: a potential diagnostic pitfall. *Journal of Clinical Pathology*, 60(8), pp.936–938.
- Nielsen, H. V, Christensen, J.P., Andersson, E.C., Marker, O. and Thomsen, A.R. (1994). Expression of type 3 complement receptor on activated CD8<sup>+</sup> T cells facilitates homing to inflammatory sites. *The Journal of Immunology*, 153(5), pp.2021–2028.
- Nitsche, A., Junghahn, I., Thulke, S., Aumann, J., Radonić, A., Fichtner, I. and Siegert, W. (2003). Interleukin-3 Promotes Proliferation and Differentiation of Human Hematopoietic Stem Cells but Reduces Their Repopulation Potential in NOD/SCID Mice. *STEM CELLS*, 21(2), pp.236–244.
- Odenike, O., Anastasi, J. and Le Beau, M.M. (2011). Myelodysplastic Syndromes. *Clinics in Laboratory Medicine*, 31(4), pp.763–784.
- Ogata, K., Satoh, C., Tachibana, M., Hyodo, H., Tamura, H., Dan, K., Kimura, T., Sonoda, Y. and Tsuji, T. (2005). Identification and Hematopoietic Potential of CD45<sup>-</sup> Clonal Cells with Very Immature Phenotype (CD45<sup>-</sup> CD34<sup>-</sup> CD38<sup>-</sup> Lin<sup>-</sup>) in Patients with Myelodysplastic Syndromes. *STEM CELLS*, 23(5), pp.619–630.
- Ogawa, S. (2019). Genetics of MDS. *Blood*, 133(10), pp.1049–1059.
- Oguro, H., Ding, L. and Morrison, S.J. (2013). SLAM Family Markers Resolve Functionally Distinct Subpopulations of Hematopoietic Stem Cells and Multipotent Progenitors. *Cell Stem Cell*, 13(1), pp.102–116.
- Okada, M., Suto, Y., Hirai, M., Shiseki, M., Usami, A., Okajima, K., Teramura, M., Mori, N. and Motoji, T. (2012). Microarray CGH analyses of chromosomal 20q deletions in patients with hematopoietic malignancies. *Cancer Genetics*, 205(1–2), pp.18–24.
- Okada, M. (2002). Myb controls G2/M progression by inducing cyclin B expression in the Drosophila eye imaginal disc. *The EMBO Journal*, 21(4), pp.675–684.
- Okita, K., Hong, H., Takahashi, K. and Yamanaka, S. (2010). Generation of mouse-induced pluripotent stem cells with plasmid vectors. *Nature Protocols*, 5(3), pp.418–428.
- Okita, K., Ichisaka, T. and Yamanaka, S. (2007). Generation of germline-competent induced pluripotent stem cells. *Nature*, 448(7151), pp.313–317.
- Orkin, S.H. (2000). Diversification of haematopoietic stem cells to specific lineages. *Nature Reviews Genetics*, 1(1), pp.57–64.
- Osato, M., Asou, N., Abdalla, E., Hoshino, K., Yamasaki, H., Okubo, T., Suzushima, H., Takatsuki, K., Kanno, T., Shigesada, K. and Ito, Y. (1999). Biallelic and Heterozygous Point Mutations in the Runt Domain of the AML1/PEBP2 $\square$ B Gene Associated With Myeloblastic Leukemias. *Blood*, 93(6), pp.1817–1824.
- Ossenkoppele, G.J., Wijermans, P.W., Nauta, J.J.P., Huijgens, P.C. and Langenhuijsen, M.M.A.C. (1991). Low-dose cytarabine for acute myeloid leukaemia and myelodysplastic syndromes: in vivo and in vitro cytotoxicity. *European Journal of Cancer and Clinical Oncology*, 27(7), pp.842–845.

- Ottersbach, K. (2019). Endothelial-to-haematopoietic transition: an update on the process of making blood. *Biochemical Society Transactions*, 47(2), pp.591–601.
- Pabst, T., Mueller, B.U., Zhang, P., Radomska, H.S., Narravula, S., Schnittger, S., Behre, G., Hiddemann, W. and Tenen, D.G. (2001). Dominant-negative mutations of CEBPA, encoding CCAAT/enhancer binding protein- $\alpha$  (C/EBP $\alpha$ ), in acute myeloid leukemia. *Nature Genetics*, 27(3), pp.263–270.
- Pabst, T. and Mueller, B.U. (2009). Complexity of CEBPA Dysregulation in Human Acute Myeloid Leukemia. *Clinical Cancer Research*, 15(17), pp.5303–5307.
- Pahl, H., Rosmarin, A. and Tenen, D. (1992). Characterization of the myeloid-specific CD11b promoter. *Blood*, 79(4), pp.865–870.
- Palomo, L. and Solé, F. (2020). SF3B1: the lord of the rings in MDS. *Blood*, 136(2), pp.149–151.
- Pan, G. and Thomson, J.A. (2007). Nanog and transcriptional networks in embryonic stem cell pluripotency. *Cell Research*, 17(1), pp.42–49.
- Papaemmanuil, E. et al. (2013). Clinical and biological implications of driver mutations in myelodysplastic syndromes. *Blood*, 122(22), pp.3616–3627.
- Papaemmanuil, E. et al. (2016). Genomic Classification and Prognosis in Acute Myeloid Leukemia. *New England Journal of Medicine*, 374(23), pp.2209–2221.
- Papapetrou, E.P. and Schambach, A. (2016). Gene Insertion Into Genomic Safe Harbors for Human Gene Therapy. *Molecular Therapy*, 24(4), pp.678–684.
- Papathanasiou, S., Markoulaki, S., Blaine, L.J., Leibowitz, M.L., Zhang, C.-Z., Jaenisch, R. and Pellman, D. (2021). Whole chromosome loss and genomic instability in mouse embryos after CRISPR-Cas9 genome editing. *Nature Communications*, 12(1), p.5855.
- Papetti, M. and Augenlicht, L.H. (2011). MYBL2, a link between proliferation and differentiation in maturing colon epithelial cells. *Journal of Cellular Physiology*, 226(3), pp.785–791.
- Pardanani, A.D., Levine, R.L., Lasho, T., Pikman, Y., Mesa, R.A., Wadleigh, M., Steensma, D.P., Elliott, M.A., Wolanskyj, A.P., Hogan, W.J., McClure, R.F., Litzow, M.R., Gilliland, D.G. and Tefferi, A. (2006). MPL515 mutations in myeloproliferative and other myeloid disorders: a study of 1182 patients. *Blood*, 108(10), pp.3472–3476.
- de Pater, E., Kaimakis, P., Vink, C.S., Yokomizo, T., Yamada-Inagawa, T., van der Linden, R., Kartalaei, P.S., Camper, S.A., Speck, N. and Dzierzak, E. (2013). Gata2 is required for HSC generation and survival. *Journal of Experimental Medicine*, 210(13), pp.2843–2850.
- Paul, P., Malakar, A.K. and Chakraborty, S. (2019). The significance of gene mutations across eight major cancer types. *Mutation Research/Reviews in Mutation Research*, 781, pp.88–99.
- Pellagatti, A. and Boultonwood, J. (2015). The molecular pathogenesis of the myelodysplastic syndromes. *European Journal of Haematology*, 95(1), pp.3–15.

Perna, F., Gurvich, N., Hoya-Arias, R., Abdel-Wahab, O., Levine, R.L., Asai, T., Voza, F., Menendez, S., Wang, L., Liu, F., Zhao, X. and Nimer, S.D. (2010). Depletion of L3MBTL1 promotes the erythroid differentiation of human hematopoietic progenitor cells: possible role in 20q- polycythemia vera. *Blood*, 116(15), pp.2812–2821.

Pfeilstöcker, M. et al. (2016). Time-dependent changes in mortality and transformation risk in MDS. *Blood*, 128(7), pp.902–910.

Pinkel, D. and Albertson, D.G. (2005). Array comparative genomic hybridization and its applications in cancer. *Nature Genetics*, 37(S6), pp.S11–S17.

Platzbecker, U., Kubasch, A.S., Homer-Bouthiette, C. and Prebet, T. (2021). Current challenges and unmet medical needs in myelodysplastic syndromes. *Leukemia*, 35(8), pp.2182–2198.

Platzbecker, U. (2019). Treatment of MDS. *Blood*, 133(10), pp.1096–1107.

Porcher, C., Chagraoui, H. and Kristiansen, M.S. (2017). SCL/TAL1: a multifaceted regulator from blood development to disease. *Blood*, 129(15), pp.2051–2060.

Porse, B.T., Pedersen, T.Å., Xu, X., Lindberg, B., Wewer, U.M., Friis-Hansen, L. and Nerlov, C. (2001). E2F Repression by C/EBP $\alpha$  Is Required for Adipogenesis and Granulopoiesis In Vivo. *Cell*, 107(2), pp.247–258.

Porse, B.T., Bryder, D., Theilgaard-Mönch, K., Hasemann, M.S., Anderson, K., Damgaard, I., Jacobsen, S.E.W. and Nerlov, C. (2005). Loss of C/EBP $\alpha$  cell cycle control increases myeloid progenitor proliferation and transforms the neutrophil granulocyte lineage. *Journal of Experimental Medicine*, 202(1), pp.85–96.

Della Porta, M.G. et al. (2016). Clinical Effects of Driver Somatic Mutations on the Outcomes of Patients With Myelodysplastic Syndromes Treated With Allogeneic Hematopoietic Stem-Cell Transplantation. *Journal of Clinical Oncology*, 34(30), pp.3627–3637.

Della Porta, M.G. et al. (2015). Validation of WHO classification-based Prognostic Scoring System (WPSS) for myelodysplastic syndromes and comparison with the revised International Prognostic Scoring System (IPSS-R). A study of the International Working Group for Prognosis in Myelodysplasia (IWG-PM). *Leukemia*, 29(7), pp.1502–1513.

Prange, K.H.M., Singh, A.A. and Martens, J.H.A. (2014). The genome-wide molecular signature of transcription factors in leukemia. *Experimental Hematology*, 42(8), pp.637–650.

Preudhomme, C., Sagot, C., Boissel, N., Cayuela, J.-M., Tigaud, I., de Botton, S., Thomas, X., Raffoux, E., Lamandin, C., Castaigne, S., Fenaux, P. and Dombret, H. (2002). Favorable prognostic significance of CEBPA mutations in patients with de novo acute myeloid leukemia: a study from the Acute Leukemia French Association (ALFA). *Blood*, 100(8), pp.2717–2723.

Ptasinska, A. 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(8), pp.1829–1841.

Pundhir, S., Bratt Lauridsen, F.K., Schuster, M.B., Jakobsen, J.S., Ge, Y., Schoof, E.M., Rapin, N., Waage, J., Hasemann, M.S. and Porse, B.T. (2018). Enhancer and Transcription Factor

Dynamics during Myeloid Differentiation Reveal an Early Differentiation Block in Cebpa null Progenitors. *Cell Reports*, 23(9), pp.2744–2757.

Radomska, H.S., Huettner, C.S., Zhang, P., Cheng, T., Scadden, D.T. and Tenen, D.G. (1998). CCAAT/Enhancer Binding Protein  $\alpha$  Is a Regulatory Switch Sufficient for Induction of Granulocytic Development from Bipotential Myeloid Progenitors. *Molecular and Cellular Biology*, 18(7), pp.4301–4314.

Ran, F.A., Hsu, P.D., Lin, C.-Y., Gootenberg, J.S., Konermann, S., Trevino, A.E., Scott, D.A., Inoue, A., Matoba, S., Zhang, Y. and Zhang, F. (2013). Double Nicking by RNA-Guided CRISPR Cas9 for Enhanced Genome Editing Specificity. *Cell*, 154(6), pp.1380–1389.

Raschell, G., Negroni, A., Sala, A., Pucci, S., Romeo, A. and Calabretta, B. (1995). Requirement of B- myb Function for Survival and Differentiative Potential of Human Neuroblastoma Cells. *Journal of Biological Chemistry*, 270(15), pp.8540–8545.

Ravindran, A., He, R., Ketterling, R.P., Jawad, M.D., Chen, D., Oliveira, J.L., Nguyen, P.L., Viswanatha, D.S., Reichard, K.K., Hoyer, J.D., Go, R.S. and Shi, M. (2020). The significance of genetic mutations and their prognostic impact on patients with incidental finding of isolated del(20q) in bone marrow without morphologic evidence of a myeloid neoplasm. *Blood Cancer Journal*, 10(1), p.7.

Rayner, E., Durin, M.-A., Thomas, R., Moralli, D., O’Cathail, S.M., Tomlinson, I., Green, C.M. and Lewis, A. (2019). CRISPR-Cas9 Causes Chromosomal Instability and Rearrangements in Cancer Cell Lines, Detectable by Cytogenetic Methods. *The CRISPR Journal*, 2(6), pp.406–416.

Ren, F., Wang, L., Shen, X., Xiao, X., Liu, Z., Wei, P., Wang, Y., Qi, P., Shen, C., Sheng, W. and Du, X. (2015). MYBL2 is an independent prognostic marker that has tumor-promoting functions in colorectal cancer. *American journal of cancer research*, 5(4), pp.1542–52.

Renneville, A., Mialou, V., Philippe, N., Kagialis-Girard, S., Biggio, V., Zobot, M.-T., Thomas, X., Bertrand, Y. and Preudhomme, C. (2009). Another pedigree with familial acute myeloid leukemia and germline CEBPA mutation. *Leukemia*, 23(4), pp.804–806.

Le Rhun, A., Escalera-Maurer, A., Bratovič, M. and Charpentier, E. (2019). CRISPR-Cas in *Streptococcus pyogenes*. *RNA Biology*, 16(4), pp.380–389.

Ria, R., Moschetta, M., Reale, A., Mangialardi, G., Castrovilli, A., Vacca, A. and Dammacco, F. (2009). Managing myelodysplastic symptoms in elderly patients. *Clinical interventions in aging*, 4, pp.413–23.

Ribeil, J.-A., Arlet, J.-B., Dussiot, M., Cruz Moura, I., Courtois, G. and Hermine, O. (2013). Ineffective Erythropoiesis in  $\beta$  -Thalassemia. *The Scientific World Journal*, 2013, pp.1–11.

Rieger, M.A. and Schroeder, T. (2012). Hematopoiesis. *Cold Spring Harbor Perspectives in Biology*, 4(12), pp.a008250–a008250.

Robb, L. (2007). Cytokine receptors and hematopoietic differentiation. *Oncogene*, 26(47), pp.6715–6723.

- Robinton, D.A. and Daley, G.Q. (2012). The promise of induced pluripotent stem cells in research and therapy. *Nature*, 481(7381), pp.295–305.
- Rodda, D.J., Chew, J.-L., Lim, L.-H., Loh, Y.-H., Wang, B., Ng, H.-H. and Robson, P. (2005). Transcriptional Regulation of Nanog by OCT4 and SOX2. *Journal of Biological Chemistry*, 280(26), pp.24731–24737.
- Romito, A. and Cobellis, G. (2016). Pluripotent Stem Cells: Current Understanding and Future Directions. *Stem Cells International*, 2016, pp.1–20.
- Rossetti, M., Gregori, S. and Roncarolo, M.G. (2010). Granulocyte-colony stimulating factor drives the in vitro differentiation of human dendritic cells that induce anergy in naïve T cells. *European Journal of Immunology*, 40(11), pp.3097–3106.
- Roussel, M., Saule, S., Lagrou, C., Rommens, C., Beug, H., Graf, T. and Stehelin, D. (1979). Three new types of viral oncogene of cellular origin specific for haematopoietic cell transformation. *Nature*, 281(5731), pp.452–455.
- Sackett, D.L. and Varma, J.K. (1993). Molecular mechanism of colchicine action: Induced local unfolding of  $\beta$ -tubulin. *Biochemistry*, 32(49), pp.13560–13565.
- Sala, A. (2005). B-MYB, a transcription factor implicated in regulating cell cycle, apoptosis and cancer. *European Journal of Cancer*, 41(16), pp.2479–2484.
- Sanada, M. et al. (2009). Gain-of-function of mutated C-CBL tumour suppressor in myeloid neoplasms. *Nature*, 460(7257), pp.904–908.
- Sanchez, J.F. (2011). Treatment of myelodysplastic syndromes in elderly patients. *Advances in Therapy*, 28(S2), pp.1–9.
- Sanjuan-Pla, A. et al. (2013). Platelet-biased stem cells reside at the apex of the haematopoietic stem-cell hierarchy. *Nature*, 502(7470), pp.232–236.
- Santilli, G., Schwab, R., Watson, R., Ebert, C., Aronow, B.J. and Sala, A. (2005). Temperature-dependent Modification and Activation of B-MYB. *Journal of Biological Chemistry*, 280(16), pp.15628–15634.
- Savona, M.R. et al. (2020). Clinical Efficacy and Safety of Oral Decitabine/Cedazuridine in 133 Patients with Myelodysplastic Syndromes (MDS) and Chronic Myelomonocytic Leukemia (CMML). *Blood*, 136(Supplement 1), pp.37–38.
- Saygin, C. and Carraway, H.E. (2021). Current and emerging strategies for management of myelodysplastic syndromes. *Blood Reviews*, 48, p.100791.
- Scalzulli, E., Pepe, S., Colafigli, G. and Breccia, M. (2021). Therapeutic strategies in low and high-risk MDS: What does the future have to offer?. *Blood Reviews*, 45, p.100689.
- Schanz, J., Steidl, C., Fonatsch, C., Pfeilstöcker, M., Nösslinger, T., Tuechler, H., Valent, P., Hildebrandt, B., Giagounidis, A., Aul, C., Lübbert, M., Stauder, R., Krieger, O., Garcia-Manero, G., Kantarjian, H., Germing, U., Haase, D. and Estey, E. (2011). Coalesced Multicentric Analysis of 2,351 Patients With Myelodysplastic Syndromes Indicates an



Underestimation of Poor-Risk Cytogenetics of Myelodysplastic Syndromes in the International Prognostic Scoring System. *Journal of Clinical Oncology*, 29(15), pp.1963–1970.

Schanz, J. et al. (2012). New Comprehensive Cytogenetic Scoring System for Primary Myelodysplastic Syndromes (MDS) and Oligoblastic Acute Myeloid Leukemia After MDS Derived From an International Database Merge. *Journal of Clinical Oncology*, 30(8), pp.820–829.

Scharenberg, C., Giai, V., Pellagatti, A., Saft, L., Dimitriou, M., Jansson, M., Jädersten, M., Grandien, A., Douagi, I., Neuberg, D.S., LeBlanc, K., Boultonwood, J., Karimi, M., Jacobsen, S.E.W., Woll, P.S. and Hellström-Lindberg, E. (2017). Progression in patients with low- and intermediate-1-risk del(5q) myelodysplastic syndromes is predicted by a limited subset of mutations. *Haematologica*, 102(3), pp.498–508.

Schmidt, L., Heyes, E. and Grebien, F. (2020). Gain-of-Function Effects of N-Terminal *CEBPA* Mutations in Acute Myeloid Leukemia. *BioEssays*, 42(2), p.1900178.

Schopperle, W.M. and DeWolf, W.C. (2007). The TRA-1-60 and TRA-1-81 Human Pluripotent Stem Cell Markers Are Expressed on Podocalyxin in Embryonal Carcinoma. *Stem Cells*, 25(3), pp.723–730.

Schroeder, I.S., Sulzbacher, S., Nolden, T., Fuchs, J., Czarnota, J., Meisterfeld, R., Himmelbauer, H. and Wobus, A.M. (2012). Induction and Selection of Sox17-Expressing Endoderm Cells Generated from Murine Embryonic Stem Cells. *Cells Tissues Organs*, 195(6), pp.507–523.

Sebaa, A., Ades, L., Baran-Marzack, F., Mozziconacci, M.-J., Penther, D., Dobbstein, S., Stamatoullas, A., Récher, C., Prebet, T., Moulessehou, S., Fenaux, P. and Eclache, V. (2012). Incidence of 17p deletions and *TP53* mutation in myelodysplastic syndrome and acute myeloid leukemia with 5q deletion. *Genes, Chromosomes and Cancer*, 51(12), pp.1086–1092.

Seita, J. and Weissman, I.L. (2010). Hematopoietic stem cell: self-renewal versus differentiation. *WIREs Systems Biology and Medicine*, 2(6), pp.640–653.

Sekeres, M.A. and Taylor, J. (2022). Diagnosis and Treatment of Myelodysplastic Syndromes. *JAMA*, 328(9), p.872.

Sekine, K., Tsuzuki, S., Yasui, R., Kobayashi, T., Ikeda, K., Hamada, Y., Kanai, E., Camp, J.G., Treutlein, B., Ueno, Y., Okamoto, S. and Taniguchi, H. (2020). Robust detection of undifferentiated iPSC among differentiated cells. *Scientific Reports*, 10(1), p.10293.

Sentmanat, M.F., Peters, S.T., Florian, C.P., Connelly, J.P. and Pruett-Miller, S.M. (2018). A Survey of Validation Strategies for CRISPR-Cas9 Editing. *Scientific Reports*, 8(1), p.888.

Serin, I., Eren, R. and Dogu, M.H. (2020). Lenalidomide Plus Decitabine Treatment in a Myelodysplastic Syndrome Patient With Deletion 5q and Excess Blasts. *Journal of Hematology*, 9(1–2), pp.33–36.

Shih, L.-Y., Huang, C.-F., Lin, T.-L., Wu, J.-H., Wang, P.-N., Dunn, P., Kuo, M.-C. and Tang, T.-C. (2005). Heterogeneous Patterns of *CEBPA* Mutation Status in the Progression of Myelodysplastic Syndrome and Chronic Myelomonocytic Leukemia to Acute Myelogenous Leukemia. *Clinical Cancer Research*, 11(5), pp.1821–1826.

- Shin, J.Y., Hu, W., Naramura, M. and Park, C.Y. (2014). High c-Kit expression identifies hematopoietic stem cells with impaired self-renewal and megakaryocytic bias. *Journal of Experimental Medicine*, 211(2), pp.217–231.
- Sierra, J. and Nomdedeu, J.F. (2021). *CEBPA* bZip mutations: just a single shot. *Blood*, 138(13), pp.1091–1092.
- Silverman, L.R., Demakos, E.P., Peterson, B.L., Kornblith, A.B., Holland, J.C., Odchimar-Reissig, R., Stone, R.M., Nelson, D., Powell, B.L., DeCastro, C.M., Ellerton, J., Larson, R.A., Schiffer, C.A. and Holland, J.F. (2002). Randomized Controlled Trial of Azacitidine in Patients With the Myelodysplastic Syndrome: A Study of the Cancer and Leukemia Group B. *Journal of Clinical Oncology*, 20(10), pp.2429–2440.
- Singh, V.K., Kalsan, M., Kumar, N., Saini, A. and Chandra, R. (2015). Induced pluripotent stem cells: applications in regenerative medicine, disease modeling, and drug discovery. *Frontiers in Cell and Developmental Biology*, 3.
- Siva, N., Gupta, S., Gupta, A., Shukla, J.N., Malik, B. and Shukla, N. (2021). Genome-editing approaches and applications: a brief review on CRISPR technology and its role in cancer. *3 Biotech*, 11(3), p.146.
- Sloand, E.M., Pfannes, L., Chen, G., Shah, S., Solomou, E.E., Barrett, J. and Young, N.S. (2007). CD34 cells from patients with trisomy 8 myelodysplastic syndrome (MDS) express early apoptotic markers but avoid programmed cell death by up-regulation of antiapoptotic proteins. *Blood*, 109(6), pp.2399–2405.
- Smink, J.J. and Leutz, A. (2010). Rapamycin and the transcription factor *C/EBPβ* as a switch in osteoclast differentiation: implications for lytic bone diseases. *Journal of Molecular Medicine*, 88(3), pp.227–233.
- Smith, L., Hohaus, S., Gonzalez, D., Dziennis, S. and Tenen, D. (1996). PU.1 (Spi-1) and C/EBP alpha regulate the granulocyte colony-stimulating factor receptor promoter in myeloid cells. *Blood*, 88(4), pp.1234–1247.
- Soldner, F., Hockemeyer, D., Beard, C., Gao, Q., Bell, G.W., Cook, E.G., Hargus, G., Blak, A., Cooper, O., Mitalipova, M., Isacson, O. and Jaenisch, R. (2009). Parkinson's Disease Patient-Derived Induced Pluripotent Stem Cells Free of Viral Reprogramming Factors. *Cell*, 136(5), pp.964–977.
- Solé et al. (2000). Incidence, characterization and prognostic significance of chromosomal abnormalities in 640 patients with primary myelodysplastic syndromes. *British Journal of Haematology*, 108(2), pp.346–356.
- Sommer, A.G., Rozelle, S.S., Sullivan, S., Mills, J.A., Park, S.-M., Smith, B.W., Iyer, A.M., French, D.L., Kotton, D.N., Gadue, P., Murphy, G.J. and Mostoslavsky, G. (2012). Generation of Human Induced Pluripotent Stem Cells from Peripheral Blood Using the STEMCCA Lentiviral Vector. *Journal of Visualized Experiments*, 31 October 2012.
- Sommer, C.A., Sommer, A.G., Longmire, T.A., Christodoulou, C., Thomas, D.D., Gostissa, M., Alt, F.W., Murphy, G.J., Kotton, D.N. and Mostoslavsky, G. (2010). Excision of Reprogramming Transgenes Improves the Differentiation Potential of iPS Cells Generated with a Single Excisable Vector. *Stem Cells*, 28(1), pp.64–74.

- Sommer, C.A., Stadtfeld, M., Murphy, G.J., Hochedlinger, K., Kotton, D.N. and Mostoslavsky, G. (2009). Induced Pluripotent Stem Cell Generation Using a Single Lentiviral Stem Cell Cassette. *Stem Cells*, 27(3), pp.543–549.
- Song, B., Fan, Y., He, W., Zhu, D., Niu, X., Wang, D., Ou, Z., Luo, M. and Sun, X. (2015). Improved Hematopoietic Differentiation Efficiency of Gene-Corrected Beta-Thalassemia Induced Pluripotent Stem Cells by CRISPR/Cas9 System. *Stem Cells and Development*, 24(9), pp.1053–1065.
- Song, J., Du, L., Liu, P., Wang, F., Zhang, B., Xie, Y., Lu, J., Jin, Y., Zhou, Y., Lv, G., Zhang, J., Chen, S., Chen, Z., Sun, X., Zhang, Y. and Huang, Q. (2021). Intra-heterogeneity in transcription and chemoresistant property of leukemia-initiating cells in murine *Setd2*<sup>-/-</sup> acute myeloid leukemia. *Cancer Communications*, 41(9), pp.867–888.
- Sood, R., Kamikubo, Y. and Liu, P. (2017). Role of RUNX1 in hematological malignancies. *Blood*, 129(15), pp.2070–2082.
- Sperling, A.S., Gibson, C.J. and Ebert, B.L. (2017). The genetics of myelodysplastic syndrome: from clonal haematopoiesis to secondary leukaemia. *Nature Reviews Cancer*, 17(1), pp.5–19.
- Stanchina, M., Chaudhry, S., Karr, M. and Taylor, J. (2021). Current State and Challenges in Development of Targeted Therapies in Myelodysplastic Syndromes (MDS). *Hemato*, 2(2), pp.217–236.
- Steensma, D.P., Higgs, D.R., Fisher, C.A. and Gibbons, R.J. (2004). Acquired somatic ATRX mutations in myelodysplastic syndrome associated with  $\alpha$  thalassemia (ATMDS) convey a more severe hematologic phenotype than germline ATRX mutations. *Blood*, 103(6), pp.2019–2026.
- Steensma, D.P., Bejar, R., Jaiswal, S., Lindsley, R.C., Sekeres, M.A., Hasserjian, R.P. and Ebert, B.L. (2015). Clonal hematopoiesis of indeterminate potential and its distinction from myelodysplastic syndromes. *Blood*, 126(1), pp.9–16.
- Steensma, D.P. et al. (2021). Imetelstat Achieves Meaningful and Durable Transfusion Independence in High Transfusion–Burden Patients With Lower-Risk Myelodysplastic Syndromes in a Phase II Study. *Journal of Clinical Oncology*, 39(1), pp.48–56.
- Steensma, D.P. (2018). Myelodysplastic syndromes current treatment algorithm 2018. *Blood Cancer Journal*, 8(5), p.47.
- Stefanska, M., Batta, K., Patel, R., Florkowska, M., Kouskoff, V. and Lacaud, G. (2017). Primitive erythrocytes are generated from hemogenic endothelial cells. *Scientific Reports*, 7(1), p.6401.
- Štefková, K., Procházková, J. and Pacherník, J. (2015). Alkaline Phosphatase in Stem Cells. *Stem Cells International*, 2015, pp.1–11.
- Stein, E.M., Fathi, A.T., DiNardo, C.D., Pollyea, D.A., Roboz, G.J., Collins, R., Sekeres, M.A., Stone, R.M., Attar, E.C., Frattini, M.G., Tosolini, A., Xu, Q., See, W.L., MacBeth, K.J., de Botton, S., Tallman, M.S. and Kantarjian, H.M. (2020). Enasidenib in patients with mutant IDH2 myelodysplastic syndromes: a phase 1 subgroup analysis of the multicentre, AG221-C-001 trial. *The Lancet Haematology*, 7(4), pp.e309–e319.

Stomper, J., Meier, R., Ma, T., Pfeifer, D., Ihorst, G., Blagitko-Dorfs, N., Greve, G., Zimmer, D., Platzbecker, U., Hagemeijer, A., Schmitt-Graeff, I. and Lübbert, M. (2021). Integrative study of EZH2 mutational status, copy number, protein expression and H3K27 trimethylation in AML/MDS patients. *Clinical Epigenetics*, 13(1), p.77.

Stoyka, W.W., Frankel, D.Z. and Kay, J.C. (1978). The linear relation of cerebral blood flow to arterial oxygen saturation in hypoxic hypoxia induced with nitrous oxide or nitrogen. *Canadian Anaesthetists' Society journal*, 25(6), pp.474–8.

Sturgeon, C.M., Ditadi, A., Awong, G., Kennedy, M. and Keller, G. (2014). Wnt signaling controls the specification of definitive and primitive hematopoiesis from human pluripotent stem cells. *Nature Biotechnology*, 32(6), pp.554–561.

Su, L., Tan, Y., Lin, H., Liu, X., Yu, L., Yang, YanPing, Liu, S., Bai, O., Yang, Yan, Jin, F., Sun, J., Liu, C., Liu, Q., Gao, S. and Li, W. (2018). Mutational spectrum of acute myeloid leukemia patients with double *CEBPA* mutations based on next-generation sequencing and its prognostic significance. *Oncotarget*, 9(38), pp.24970–24979.

Sugimoto, K., Hirano, N., Toyoshima, H., Chiba, S., Mano, H., Takaku, F., Yazaki, Y. and Hirai, H. (1993). Mutations of the p53 gene in myelodysplastic syndrome (MDS) and MDS-derived leukemia. *Blood*, 81(11), pp.3022–3026.

Sun, N. and Zhao, H. (2014). Seamless correction of the sickle cell disease mutation of the *HBB* gene in human induced pluripotent stem cells using TALENs. *Biotechnology and Bioengineering*, 111(5), pp.1048–1053.

Swain, N., Thakur, M., Pathak, J. and Swain, B. (2020). SOX2, OCT4 and NANOG: The core embryonic stem cell pluripotency regulators in oral carcinogenesis. *Journal of Oral and Maxillofacial Pathology*, 24(2), p.368.

Syed, K., Naguib, S., Liu, Z.-J., Cimmino, L. and Yang, F.-C. (2020). Novel combinations to improve hematopoiesis in myelodysplastic syndrome. *Stem Cell Research & Therapy*, 11(1), p.132.

Szulc, J. and Aebischer, P. (2008). Conditional Gene Expression and Knockdown Using Lentivirus Vectors Encoding shRNA. In: *Gene Therapy Protocols*. Totowa, NJ: Humana Press, pp.291–300.

Takahashi, K., Tanabe, K., Ohnuki, M., Narita, M., Ichisaka, T., Tomoda, K. and Yamanaka, S. (2007). Induction of Pluripotent Stem Cells from Adult Human Fibroblasts by Defined Factors. *Cell*, 131(5), pp.861–872.

Takahashi, K. and Yamanaka, S. (2006). Induction of Pluripotent Stem Cells from Mouse Embryonic and Adult Fibroblast Cultures by Defined Factors. *Cell*, 126(4), pp.663–676.

Tanaka, Y., Patestos, N.P., Maekawa, T. and Ishii, S. (1999). B-myb Is Required for Inner Cell Mass Formation at an Early Stage of Development. *Journal of Biological Chemistry*, 274(40), pp.28067–28070.

Tang, G., Zhang, L., Fu, B., Hu, J., Lu, X., Hu, S., Patel, A., Goswami, M., Khoury, J.D., Garcia-Manero, G., Medeiros, L.J. and Wang, S.A. (2014). Cytogenetic risk stratification of

417 patients with chronic myelomonocytic leukemia from a single institution. *American Journal of Hematology*, 89(8), pp.813–818.

Tarasov, K. V., Tarasova, Y.S., Tam, W.L., Riordon, D.R., Elliott, S.T., Kania, G., Li, J., Yamanaka, S., Crider, D.G., Testa, G., Li, R.A., Lim, B., Stewart, C.L., Liu, Y., Van Eyk, J.E., Wersto, R.P., Wobus, A.M. and Boheler, K.R. (2008). B-MYB Is Essential for Normal Cell Cycle Progression and Chromosomal Stability of Embryonic Stem Cells. *PLoS ONE*, 3(6), p.e2478.

Taskesen, E., Bullinger, L., Corbacioglu, A., Sanders, M.A., Erpelinck, C.A.J., Wouters, B.J., van der Poel-van de Luytgaarde, S.C., Damm, F., Krauter, J., Ganser, A., Schlenk, R.F., Löwenberg, B., Delwel, R., Döhner, H., Valk, P.J.M. and Döhner, K. (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(8), pp.2469–2475.

Taube, F. et al. (2022). *CEBPA* mutations in 4708 patients with acute myeloid leukemia: differential impact of bZIP and TAD mutations on outcome. *Blood*, 139(1), pp.87–103.

Tawana, K. et al. (2015). Disease evolution and outcomes in familial AML with germline CEBPA mutations. *Blood*, 126(10), pp.1214–1223.

Taxman, D.J., Moore, C.B., Guthrie, E.H. and Huang, M.T.-H. (2010). Short Hairpin RNA (shRNA): Design, Delivery, and Assessment of Gene Knockdown. In: pp.139–156.

Tefferi, A. and Vardiman, J.W. (2009). Myelodysplastic Syndromes. *New England Journal of Medicine*, 361(19), pp.1872–1885.

Tenen, D.G. (2003). Disruption of differentiation in human cancer: AML shows the way. *Nature Reviews Cancer*, 3(2), pp.89–101.

Tie, R., Li, H., Cai, S., Liang, Z., Shan, W., Wang, B., Tan, Y., Zheng, W. and Huang, H. (2019). Interleukin-6 signaling regulates hematopoietic stem cell emergence. *Experimental & Molecular Medicine*, 51(10), pp.1–12.

Tien, F.-M. et al. (2018). Concomitant *WT1* mutations predict poor prognosis in acute myeloid leukemia patients with double mutant *CEBPA*. *Haematologica*, 103(11), pp.e510–e513.

TILL, J.E. and McCULLOCH, E.A. (1961). A direct measurement of the radiation sensitivity of normal mouse bone marrow cells. *Radiation research*, 14, pp.213–22.

Tomic, B., Smoljo, T., Lalic, H., Dembitz, V., Batinic, J., Batinic, D., Bedalov, A. and Visnjic, D. (2022). Cytarabine-induced differentiation of AML cells depends on Chk1 activation and shares the mechanism with inhibitors of DHODH and pyrimidine synthesis. *Scientific Reports*, 12(1), p.11344.

Toscani, A., Mettus, R. V., Coupland, R., Simpkins, H., Litvin, J., Orth, J., Hatton, K.S. and Reddy, E.P. (1997). Arrest of spermatogenesis and defective breast development in mice lacking A-myb. *Nature*, 386(6626), pp.713–717.

Traktuev, D.O., Merfeld-Clauss, S., Li, J., Kolonin, M., Arap, W., Pasqualini, R., Johnstone, B.H. and March, K.L. (2008). A Population of Multipotent CD34-Positive Adipose Stromal

Cells Share Pericyte and Mesenchymal Surface Markers, Reside in a Periendothelial Location, and Stabilize Endothelial Networks. *Circulation Research*, 102(1), pp.77–85.

Trauth, K., Mutschler, B., Jenkins, N.A., Gilbert, D.J., Copeland, N.G. and Klempnauer, K.H. (1994). Mouse A-myb encodes a trans-activator and is expressed in mitotically active cells of the developing central nervous system, adult testis and B lymphocytes. *The EMBO Journal*, 13(24), pp.5994–6005.

Traver, D., Miyamoto, T., Christensen, J., Iwasaki-Arai, J., Akashi, K. and Weissman, I.L. (2001). Fetal liver myelopoiesis occurs through distinct, prospectively isolatable progenitor subsets. *Blood*, 98(3), pp.627–635.

Trowbridge, I.S. and Thomas, M.L. (1994). CD45: An Emerging Role as a Protein Tyrosine Phosphatase Required for Lymphocyte Activation and Development. *Annual Review of Immunology*, 12(1), pp.85–116.

Trusler, O., Huang, Z., Goodwin, J. and Laslett, A.L. (2018). Cell surface markers for the identification and study of human naive pluripotent stem cells. *Stem Cell Research*, 26, pp.36–43.

Tsai, F.-Y. and Orkin, S.H. (1997). Transcription Factor GATA-2 Is Required for Proliferation/Survival of Early Hematopoietic Cells and Mast Cell Formation, But Not for Erythroid and Myeloid Terminal Differentiation. *Blood*, 89(10), pp.3636–3643.

Tsiftoglou, A.S., Vizirianakis, I.S. and Strouboulis, J. (2009). Erythropoiesis: Model systems, molecular regulators, and developmental programs. *IUBMB Life*, 61(8), pp.800–830.

Tursky, M.L., Loi, T.H., Artuz, C.M., Alateeq, S., Wolvetang, E.J., Tao, H., Ma, D.D. and Molloy, T.J. (2020). Direct Comparison of Four Hematopoietic Differentiation Methods from Human Induced Pluripotent Stem Cells. *Stem Cell Reports*, 15(3), pp.735–748.

Valent, P. et al. (2017). Proposed Terminology and Classification of Pre-Malignant Neoplastic Conditions: A Consensus Proposal. *EBioMedicine*, 26, pp.17–24.

van Oevelen, C., Collombet, S., Vicent, G., Hoogenkamp, M., Lepoivre, C., Badeaux, A., Bussmann, L., Sardina, J.L., Thieffry, D., Beato, M., Shi, Y., Bonifer, C. and Graf, T. (2015). C/EBP $\alpha$  Activates Pre-existing and De Novo Macrophage Enhancers during Induced Pre-B Cell Transdifferentiation and Myelopoiesis. *Stem Cell Reports*, 5(2), pp.232–247.

Vardiman, J.W., Thiele, J., Arber, D.A., Brunning, R.D., Borowitz, M.J., Porwit, A., Harris, N.L., Le Beau, M.M., Hellström-Lindberg, E., Tefferi, A. and Bloomfield, C.D. (2009). The 2008 revision of the World Health Organization (WHO) classification of myeloid neoplasms and acute leukemia: rationale and important changes. *Blood*, 114(5), pp.937–951.

Vink, C.S., Mariani, S.A. and Dzierzak, E. (2022). Embryonic Origins of the Hematopoietic System: Hierarchies and Heterogeneity. *HemaSphere*, 6(6), p.e737.

Visani, G., Malagola, M., Piccaluga, P.P. and Isidori, A. (2004). Low Dose Ara-C for Myelodysplastic Syndromes: is it Still a Current Therapy?. *Leukemia & Lymphoma*, 45(8), pp.1531–1538.

- Vodyanik, M.A., Thomson, J.A. and Slukvin, I.I. (2006). Leukosialin (CD43) defines hematopoietic progenitors in human embryonic stem cell differentiation cultures. *Blood*, 108(6), pp.2095–2105.
- Volpe, G., Clarke, M., Garcia, P., Walton, D.S., Vegiopoulos, A., Del Pozzo, W., O’Neill, L.P., Frampton, J. and Dumon, S. (2015). Regulation of the Flt3 Gene in Haematopoietic Stem and Early Progenitor Cells. *PLOS ONE*, 10(9), p.e0138257.
- Volpe, V.O., Garcia-Manero, G. and Komrokji, R.S. (2022). SOHO State of the Art & Next Questions: Myelodysplastic Syndromes: A New Decade. *Clinical Lymphoma Myeloma and Leukemia*, 22(1), pp.1–16.
- van Waalwijk van Doorn-Khosrovani, S.B., Erpelinck, C., Meijer, J., van Oosterhoud, S., van Putten, W.L.J., Valk, P.J.M., Berna Beverloo, H., Tenen, D.G., Löwenberg, B. and Delwel, R. (2003). Biallelic mutations in the CEBPA gene and low CEBPA expression levels as prognostic markers in intermediate-risk AML. *The Hematology Journal*, 4(1), pp.31–40.
- Wagner, K., Zhang, P., Rosenbauer, F., Drescher, B., Kobayashi, S., Radomska, H.S., Kutok, J.L., Gilliland, D.G., Krauter, J. and Tenen, D.G. (2006). Absence of the transcription factor CCAAT enhancer binding protein  $\alpha$  results in loss of myeloid identity in bcr/abl-induced malignancy. *Proceedings of the National Academy of Sciences*, 103(16), pp.6338–6343.
- Wakao, S., Kitada, M., Kuroda, Y., Ogura, F., Murakami, T., Niwa, A. and Dezawa, M. (2012). Morphologic and Gene Expression Criteria for Identifying Human Induced Pluripotent Stem Cells. *PLoS ONE*, 7(12), p.e48677.
- Wakita, S. et al. (2022). Prognostic impact of CEBPA bZIP domain mutation in acute myeloid leukemia. *Blood Advances*, 6(1), pp.238–247.
- Wang, D., D’Costa, J., Civin, C.I. and Friedman, A.D. (2006). C/EBP $\alpha$  directs monocytic commitment of primary myeloid progenitors. *Blood*, 108(4), pp.1223–1229.
- Wang, E., Boswell, E., Siddiqi, I., Lu, C.M., Sebastian, S., Rehder, C. and Huang, Q. (2011). Pseudo-Pelger-Huët Anomaly Induced by Medications. *American Journal of Clinical Pathology*, 135(2), pp.291–303.
- Wang, X., Harrison, J.S. and Studzinski, G.P. (2016). Enhancement of arabinocytosine (AraC) toxicity to AML cells by a differentiation agent combination. *The Journal of Steroid Biochemistry and Molecular Biology*, 164, pp.72–78.
- Wang, Z., Wang, S., Fisher, P.B., Dent, P. and Grant, S. (2000). Evidence of a functional role for the cyclin-dependent kinase inhibitor p21CIP1 in leukemic cell (U937) differentiation induced by low concentrations of 1- $\beta$ -D-Arabinofuranosylcytosine. *Differentiation*, 66(1), pp.1–13.
- Ward, C., Volpe, G., Cauchy, P., Ptasinska, A., Almaghrabi, R., Blakemore, D., Nafria, M., Kestner, D., Frampton, J., Murphy, G., Buganim, Y., Kaji, K. and García, P. (2018). Fine-Tuning Mybl2 Is Required for Proper Mesenchymal-to-Epithelial Transition during Somatic Reprogramming. *Cell Reports*, 24(6), pp.1496-1511.e8.
- Wattanapanitch, M. (2019). Recent Updates on Induced Pluripotent Stem Cells in Hematological Disorders. *Stem Cells International*, 2019, pp.1–15.

- Weiss, M.J., Keller, G. and Orkin, S.H. (1994). Novel insights into erythroid development revealed through in vitro differentiation of GATA-1 embryonic stem cells. *Genes & Development*, 8(10), pp.1184–1197.
- Welch, J.S. et al. (2012). The Origin and Evolution of Mutations in Acute Myeloid Leukemia. *Cell*, 150(2), pp.264–278.
- Wen, X., Hu, J., Yang, J., Qian, W., Yao, D., Deng, Z., Zhang, Y., Zhu, X., Guo, H., Lin, J. and Qian, J. (2015). CEBPA methylation and mutation in myelodysplastic syndrome. *Medical Oncology*, 32(7), p.192.
- Westers, T.M. et al. (2017). Immunophenotypic analysis of erythroid dysplasia in myelodysplastic syndromes. A report from the IMDSFlow working group. *Haematologica*, 102(2), pp.308–319.
- Whitfield, M.L., George, L.K., Grant, G.D. and Perou, C.M. (2006). Common markers of proliferation. *Nature Reviews Cancer*, 6(2), pp.99–106.
- Winter, J.N., Variakojis, D., Gaynor, E.R., Larson, R.A. and Miller, K.B. (1985). Low-dose cytosine arabinoside (Ara-C) therapy in the myelodysplastic syndromes and acute leukemia. *Cancer*, 56(3), pp.443–449.
- Wouters, B.J., Jordà, M.A., Keeshan, K., Louwers, I., Erpelinck-Verschueren, C.A.J., Tielemans, D., Langerak, A.W., He, Y., Yashiro-Ohtani, Y., Zhang, P., Hetherington, C.J., Verhaak, R.G.W., Valk, P.J.M., Löwenberg, B., Tenen, D.G., Pear, W.S. and Delwel, R. (2007). Distinct gene expression profiles of acute myeloid/T-lymphoid leukemia with silenced CEBPA and mutations in NOTCH1. *Blood*, 110(10), pp.3706–3714.
- Wouters, B.J., Löwenberg, B., Erpelinck-Verschueren, C.A.J., van Putten, W.L.J., Valk, P.J.M. and 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(13), pp.3088–3091.
- Wouters, B.J. and Delwel, R. (2016). Epigenetics and approaches to targeted epigenetic therapy in acute myeloid leukemia. *Blood*, 127(1), pp.42–52.
- Xie, F., Ye, L., Chang, J.C., Beyer, A.I., Wang, J., Muench, M.O. and Kan, Y.W. (2014). Seamless gene correction of  $\beta$ -thalassemia mutations in patient-specific iPSCs using CRISPR/Cas9 and *piggyBac*. *Genome Research*, 24(9), pp.1526–1533.
- Xue, H., Wu, S., Papadeas, S.T., Spusta, S., Swistowska, A.M., MacArthur, C.C., Mattson, M.P., Maragakis, N.J., Capecchi, M.R., Rao, M.S., Zeng, X. and Liu, Y. (2009). A Targeted Neuroglial Reporter Line Generated by Homologous Recombination in Human Embryonic Stem Cells. *Stem Cells*, 27(8), pp.1836–1846.
- Yamamoto, R., Morita, Y., Ooehara, J., Hamanaka, S., Onodera, M., Rudolph, K.L., Ema, H. and Nakauchi, H. (2013). Clonal Analysis Unveils Self-Renewing Lineage-Restricted Progenitors Generated Directly from Hematopoietic Stem Cells. *Cell*, 154(5), pp.1112–1126.
- Yamanaka, S. (2012). Induced Pluripotent Stem Cells: Past, Present, and Future. *Cell Stem Cell*, 10(6), pp.678–684.



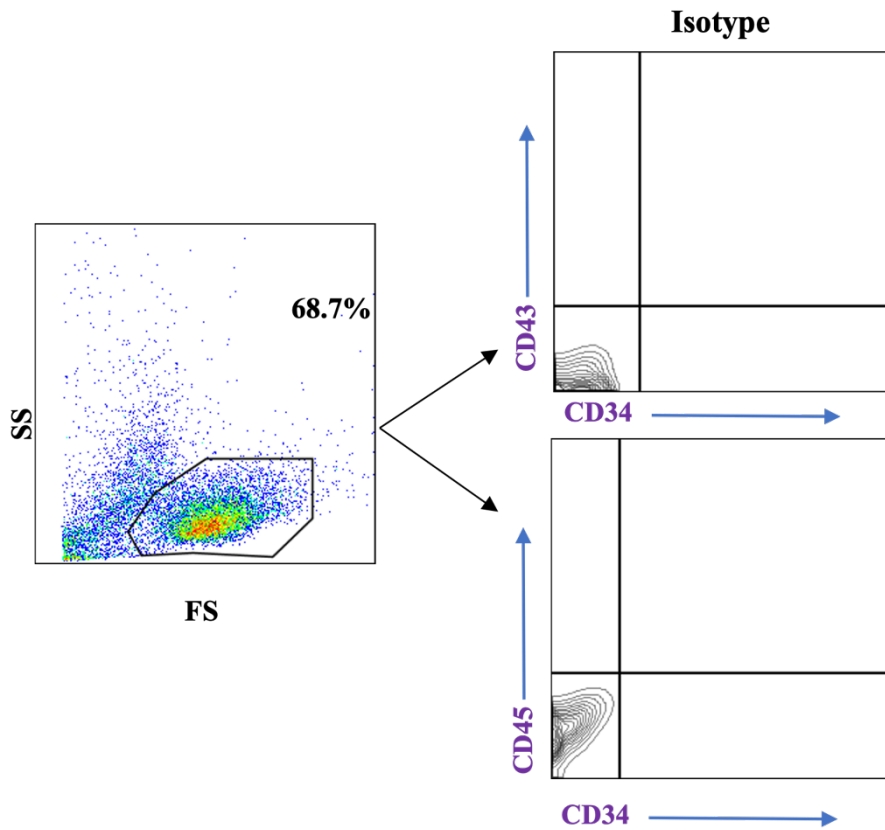
- Yang, Y., Zhang, X., Yi, L., Hou, Z., Chen, J., Kou, X., Zhao, Y., Wang, H., Sun, X.-F., Jiang, C., Wang, Y. and Gao, S. (2016). Naïve Induced Pluripotent Stem Cells Generated From  $\beta$ -Thalassemia Fibroblasts Allow Efficient Gene Correction With CRISPR/Cas9. *Stem Cells Translational Medicine*, 5(1), pp.8–19.
- Ye, M., Zhang, H., Amabile, G., Yang, H., Staber, P.B., Zhang, P., Levantini, E., Alberich-Jordà, M., Zhang, J., Kawasaki, A. and Tenen, D.G. (2013). C/EBP $\alpha$  controls acquisition and maintenance of adult haematopoietic stem cell quiescence. *Nature Cell Biology*, 15(4), pp.385–394.
- Yin, C.C., Peng, J., Li, Y., Shamanna, R.K., Muzzafar, T., DiNardo, C., Khoury, J.D., Li, S., Medeiros, L.J., Wang, S.A. and Tang, G. (2015). Clinical significance of newly emerged isolated del(20q) in patients following cytotoxic therapies. *Modern Pathology*, 28(8), pp.1014–1022.
- Ymer, S., Tucker, W.Q.J., Sanderson, C.J., Hapel, A.J., Campbell, H.D. and Young, I.G. (1985). Constitutive synthesis of interleukin-3 by leukaemia cell line WEHI-3B is due to retroviral insertion near the gene. *Nature*, 317(6034), pp.255–258.
- Yokomizo, T., Ideue, T., Morino-Koga, S., Tham, C.Y., Sato, T., Takeda, N., Kubota, Y., Kurokawa, M., Komatsu, N., Ogawa, M., Araki, K., Osato, M. and Suda, T. (2022). Independent origins of fetal liver haematopoietic stem and progenitor cells. *Nature*, 609(7928), pp.779–784.
- Yoshida, K. et al. (2011). Frequent pathway mutations of splicing machinery in myelodysplasia. *Nature*, 478(7367), pp.64–69.
- You, L., Tong, R., Li, M., Liu, Y., Xue, J. and Lu, Y. (2019). Advancements and Obstacles of CRISPR-Cas9 Technology in Translational Research. *Molecular Therapy - Methods & Clinical Development*, 13, pp.359–370.
- Yousoufian, H., Longmore, G., Neumann, D., Yoshimura, A. and Lodish, H. (1993). Structure, function, and activation of the erythropoietin receptor. *Blood*, 81(9), pp.2223–2236.
- Yu, J., Li, Y., Zhang, D., Wan, D. and Jiang, Z. (2020). Clinical implications of recurrent gene mutations in acute myeloid leukemia. *Experimental Hematology & Oncology*, 9(1), p.4.
- Yu, J., Li, Yingmei, Li, T., Li, Yafei, Xing, H., Sun, H., Sun, L., Wan, D., Liu, Y., Xie, X. and Jiang, Z. (2020). Gene mutational analysis by NGS and its clinical significance in patients with myelodysplastic syndrome and acute myeloid leukemia. *Experimental Hematology & Oncology*, 9(1), p.2.
- Yu, J., Hu, K., Smuga-Otto, K., Tian, S., Stewart, R., Slukvin, I.I. and Thomson, J.A. (2009). Human Induced Pluripotent Stem Cells Free of Vector and Transgene Sequences. *Science*, 324(5928), pp.797–801.
- Zhan, M., Riordon, D.R., Yan, B., Tarasova, Y.S., Bruweleit, S., Tarasov, K. V., Li, R.A., Wersto, R.P. and Boheler, K.R. (2012). The B-MYB Transcriptional Network Guides Cell Cycle Progression and Fate Decisions to Sustain Self-Renewal and the Identity of Pluripotent Stem Cells. *PLoS ONE*, 7(8), p.e42350.

- Zhang, D., Zhang, Z., Unver, T. and Zhang, B. (2021). CRISPR/Cas: A powerful tool for gene function study and crop improvement. *Journal of Advanced Research*, 29, pp.207–221.
- Zhang, D., Cho, E. and Wong, J. (2007). A critical role for the co-repressor N-CoR in erythroid differentiation and heme synthesis. *Cell Research*, 17(9), pp.804–814.
- Zhang, H. and McCarty, N. (2016). CRISPR-Cas9 technology and its application in haematological disorders. *British Journal of Haematology*, 175(2), pp.208–225.
- Zhang, H.-X., Zhang, Y. and Yin, H. (2019). Genome Editing with mRNA Encoding ZFN, TALEN, and Cas9. *Molecular Therapy*, 27(4), pp.735–746.
- Zhang, P., Iwasaki-Arai, J., Iwasaki, H., Fenyus, M.L., Dayaram, T., Owens, B.M., Shigematsu, H., Levantini, E., Huettner, C.S., Lekstrom-Himes, J.A., Akashi, K. and Tenen, D.G. (2004). Enhancement of Hematopoietic Stem Cell Repopulating Capacity and Self-Renewal in the Absence of the Transcription Factor C/EBP $\alpha$ . *Immunity*, 21(6), pp.853–863.
- Zhang, P., Behre, G., Pan, J., Iwama, A., Wara-aswapati, N., Radomska, H.S., Auron, P.E., Tenen, D.G. and Sun, Z. (1999). Negative cross-talk between hematopoietic regulators: GATA proteins repress PU.1. *Proceedings of the National Academy of Sciences*, 96(15), pp.8705–8710.
- Zhang, P., Zhang, C., Li, J., Han, J., Liu, X. and Yang, H. (2019). The physical microenvironment of hematopoietic stem cells and its emerging roles in engineering applications. *Stem Cell Research & Therapy*, 10(1), p.327.
- Zhang, P., Iwama, A., Datta, M.W., Darlington, G.J., Link, D.C. and Tenen, D.G. (1998). Upregulation of Interleukin 6 and Granulocyte Colony-Stimulating Factor Receptors by Transcription Factor CCAAT Enhancer Binding Protein  $\alpha$  (C/EBP $\alpha$ ) Is Critical for Granulopoiesis. *Journal of Experimental Medicine*, 188(6), pp.1173–1184.
- Zhang, Y., Gao, S., Xia, J. and Liu, F. (2018). Hematopoietic Hierarchy – An Updated Roadmap. *Trends in Cell Biology*, 28(12), pp.976–986.
- Zhao, W., Ji, X., Zhang, F., Li, L. and Ma, L. (2012). Embryonic Stem Cell Markers. *Molecules*, 17(6), pp.6196–6236.
- Zhao, X., Li, J., Liu, Z. and Powers, S. (2021). Combinatorial CRISPR/Cas9 Screening Reveals Epistatic Networks of Interacting Tumor Suppressor Genes and Therapeutic Targets in Human Breast Cancer. *Cancer Research*, 81(24), pp.6090–6105.
- Zhao, X., Bartholdy, B., Yamamoto, Y., Evans, E.K., Alberich-Jordà, M., Staber, P.B., Benoukraf, T., Zhang, P., Zhang, J., Trinh, B.Q., Crispino, J.D., Hoang, T., Bassal, M.A. and Tenen, D.G. (2022). PU.1-c-Jun interaction is crucial for PU.1 function in myeloid development. *Communications Biology*, 5(1), p.961.
- Zheng, Y.L. (2016). Some Ethical Concerns About Human Induced Pluripotent Stem Cells. *Science and Engineering Ethics*, 22(5), pp.1277–1284.
- Zhu, J. and Emerson, S.G. (2002). Hematopoietic cytokines, transcription factors and lineage commitment. *Oncogene*, 21(21), pp.3295–3313.

Zou, J., Mali, P., Huang, X., Dowey, S.N. and Cheng, L. (2011). Site-specific gene correction of a point mutation in human iPS cells derived from an adult patient with sickle cell disease. *Blood*, 118(17), pp.4599–4608.

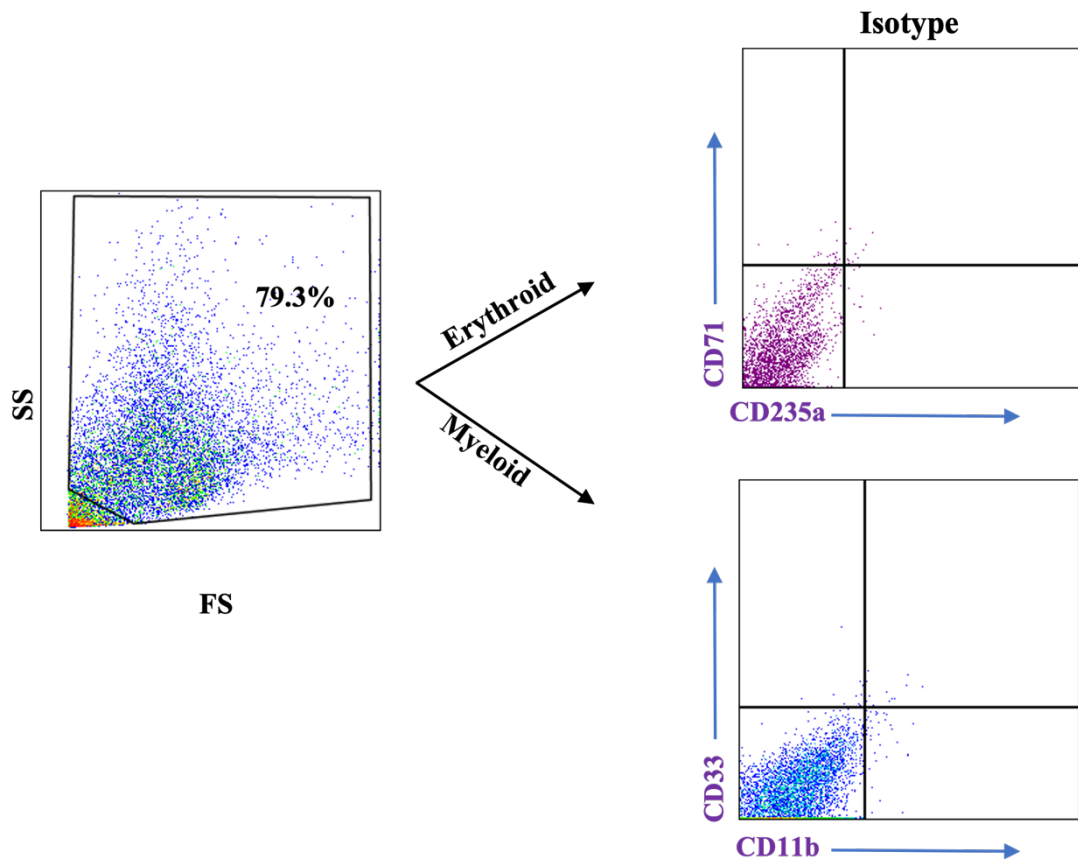
Zovein, A.C., Hofmann, J.J., Lynch, M., French, W.J., Turlo, K.A., Yang, Y., Becker, M.S., Zanetta, L., Dejana, E., Gasson, J.C., Tallquist, M.D. and Iruela-Arispe, M.L. (2008). Fate Tracing Reveals the Endothelial Origin of Hematopoietic Stem Cells. *Cell Stem Cell*, 3(6), pp.625–636.

## SUPPLEMENTARY



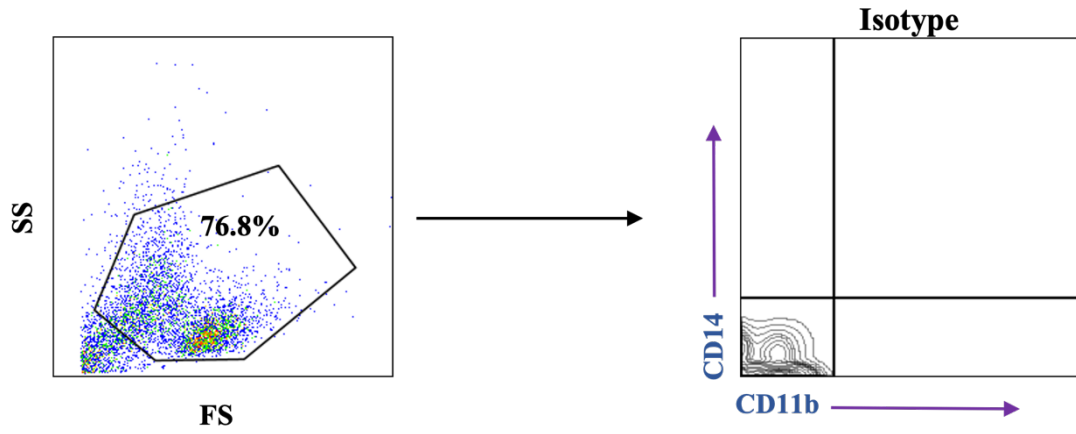
### Supplementary figure 1: Gating strategy for HSPCs.

Gating strategy of HSPCs Differentiation. The cells were gated according to FSC and SSC. Live cells and unstained cells were analysed for HSPCs markers (CD34, CD43, and CD45) based on an isotype control.



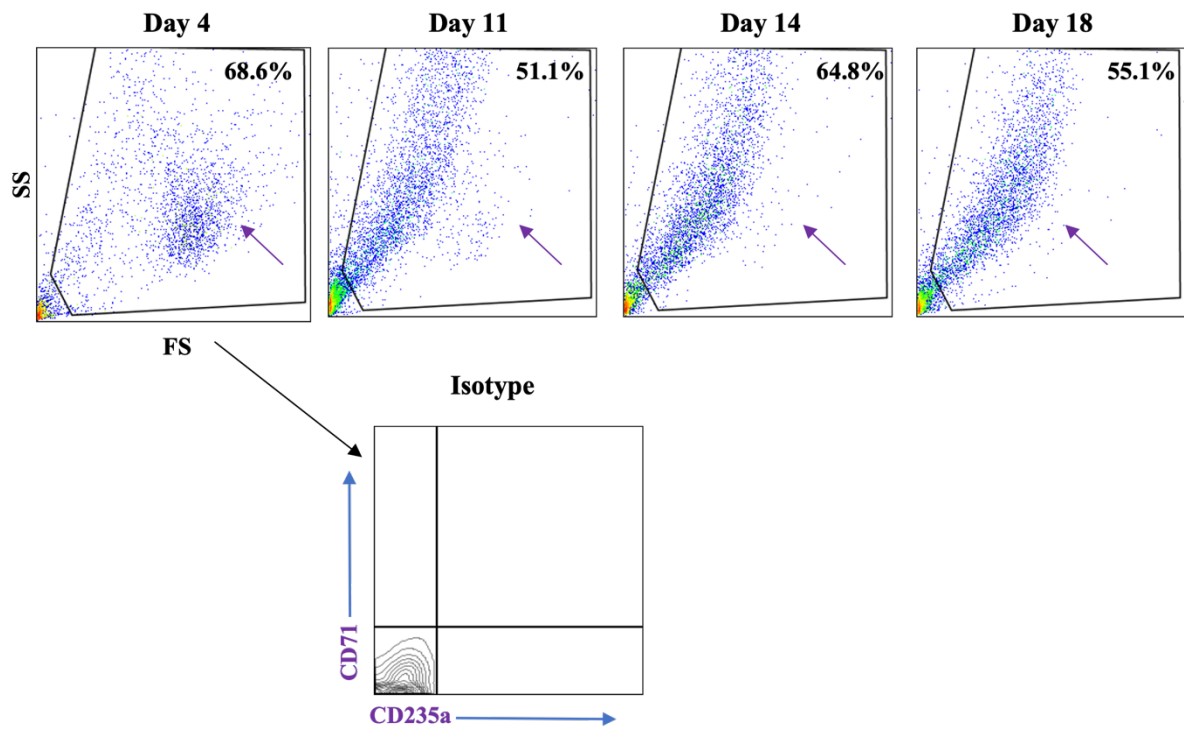
**Supplementary figure 2: Gating strategy for colony assay**

Gating strategy of colony assay differentiation. The cells were gated first based on FSC and SSC Cells of unstained cells and then the live cells were analysed for erythroid and myeloid markers according to isotype control.



### Supplementary figure 3: Gating strategy for Myeloid cells

Gating strategy of myeloid differentiation. The cells were gated first based on FSC and SSC Cells of unstained cells, and then the live cells were analysed for CD11b<sup>+</sup> CD14<sup>-</sup> (granulocytes) and CD11b<sup>+</sup> CD14<sup>+</sup> (Monocytes) according to isotype control.



#### Supplementary figure 4: Gating strategy for Erythroid cells

Strategy for strategy erythroid differentiation. First, the cells were gated according to the FSC and SS of unstained cells, and then the living cells were analysed for erythroid markers (CD71 and CD235a) using an isotype control. The purple arrows represent the populations of progenitor and erythroblast cells that are lost during maturation.