Investigation of haematopoietic cell growth and the impact of common mutations in myeloid malignancies



Nuffield Division of Clinical Laboratory Sciences,

Radcliffe Department of Medicine, University of Oxford

New College

Vickram S. Tittrea

Thesis Submitted for the Degree of Doctor of Philosophy

Supervisors: Prof. Jacqueline Boultwood and Prof. Suzanne Watt

Hilary Term, 2021

For my family: past, present, and future.

ABSTRACT

The myelodysplastic syndromes (MDS) are disorders of the haematopoietic stem cell and are characterised by ineffective haematopoiesis. The *TET2* and *ASXL1* genes have been implicated to be important in the pathology of MDS, but their function has not yet been characterised as co-mutations in erythroid and myeloid human cell lines. In this light using the CRISPR/Cas9 system, we engineered erythroid and myeloid models of myeloid malignancy in HUDEP2 and KBM5 cells respectively, with homozygous *TET2* or heterozygous *ASXL1* singular or co-mutations. Here, HUDEP2 and KBM5 cell models were robustly generated to express mutations in the *TET2* and/or *ASXL1* gene(s). Our cellular models demonstrate decreased mRNA expression of *TET2* in the homozygous TET2 cell variants and physiological mRNA expression of *ASXL1* in the heterozygous *ASXL1* cell variants. We found zero unintended off-targets of our *TET2* and *ASXL1* targeted sgRNAs in our HUDEP2 and KBM5 cell models. We next aimed to functionally characterise the effect of these mutations on key cellular function, such as proliferation, the cell cycle, apoptosis and differentiation. We demonstrate that cells exhibiting co-mutations of *TET2* and *ASXL1* gained a proliferative advantage. Studies on the cell cycle revealed that cells exhibiting co-mutations of *TET2* and *ASXL1* demonstrated an increase in cell cycle progression. We subsequently investigated apoptosis and differentiation in our cellular models. In our HUDEP2 cell model, the *TET2* mutants demonstrated a decrease in apoptosis and our *ASXL1* mutants demonstrated an increase in apoptosis. Next, we investigated differentiation; our HUDEP2 cell mutants demonstrated a block in erythroid differentiation. Finally, we expanded CD133+ cells using novel combinations of cytokines and small molecules, as a prelude to their gene editing.

PUBLICATIONS

MATA, M. F., HERNANDEZ, D., ROLOGI, E., GRANDOLFO, D., HASSAN, E., HUA, P., KALLMEIER, R., HIRANI, S., HEUTS, F., **TITTREA, V**., CHOO, Y., BARADEZ, M.-O., WATT, S. M. & TARUNINA, M. 2019. A modified CD34+ hematopoietic stem and progenitor cell isolation strategy from cryopreserved human umbilical cord blood. Transfusion, 59, 3560-3569.

RAHMANI, N. E., RAMACHANDRA, N., BHAGAT, T. D., GORDON, S., PRADHAN, K., RIVERA PENA, B., KAZEMI, M., GIRICZ, O., MAQBOOL, S., OLEA, R., DOLATSHAD, H., **TITTREA, V**., TATWAVEDI, D., SINGH, S., PELLAGATTI, A., BOULTWOOD, J. & VERMA, A. 2019. ASXL1 mutations are associated with widespread and distinct DNA methylation alterations. Blood, 134, 2989-2989.

ACKNOWLEDGEMENTS AND DECLARATION

Foremost, I would like to thank Prof. Jacqueline Boultwood and Prof. Suzanne Watt for their ongoing support and guidance throughout this DPhil project; without them this would not have been possible. I thank them for giving me the space to develop and implement my own ideas. I would also like to thank Dr. Andrea Pellagatti for his expertise in absolutely everything. He has been my go-to in time of crisis, and I thank him for his patience. I would also like to thank the postdoctoral team in both laboratories, whose world-class expertise at the bench has helped guide me through this profound journey. In particular, I would like to thank Dr. Hamid Dolatshad, Dr. Shalini Singh, Dr. Richard Armstrong, Dr. Violetta Steeples, Dr. Emma Harvey, Dr. Grace Yu and Dr. Alison Merryweather-Clarke. I would also like to thank the technical team for facilitating my research. In particular, I would like to thank Andrea Sanchi, Hoi Pat Tsang and Youyi Zhang. I would like to thank the dream team, my fellow DPhils, Dr. Dharamveer Tatwavedi, Juseong Lee and Tianyu Sun, for supporting me every step along the way, and for the countless hours we spent at the bench together. I would like to thank all University and College staff for helping in times of need. I would like to give huge thanks to all of my family and friends. A special mention to my best friend Zara Rachael Ryan, who has been my support network, and whose positivity knows no bounds. I would also like to thank the MRC and Bloodwise for funding this research. Finally, I would like to thank my Father for his laser-like wisdom and rigor, and my Mother for her love, understanding and support. I certify that I carried out more than 95% of the research for this thesis.

Contents

Chapter 1 1
General Introduction 1
1.1: Haematopoiesis 1
1.1.1: An introduction to haematopoietic stem cell biology 1
1.1.2: Ontogeny 1
1.1.3: Identifying the haematopoietic stem cell 6
1.1.4: The HSC lineage hierarchy 8
1.1.5: The HSC niche 17
1.1.6: Haematopoietic human growth factors 19
1.2: Clonal haematopoiesis of indeterminate potential 21
1.3: Myelodysplastic syndrome 25
1.4: Common mutations in MDS 29
1.5: *TET2* 30
1.6: *ASXL1* 32
1.7: The CRISPR Cas9 System 34
1.8: The aims and objective of this thesis 36
Chapter 2 37
Materials and Methods 37
2.1: Reagents/consumables and suppliers 37
2.2: Cells 43
2.2.1: Cell lines 43
2.2.2. Primary cells 45
2.3: Cell counting 51
2.4: Cryopreservation of cell lines 51
2.5: Flow cytometry analysis 52
2.6: Cell proliferation 52
2.7: Cell cycle analysis 52
2.8: Apoptosis assay 53
2.9: Analysis of erythroid and myeloid differentiation 54
2.10: Small scale plasmid purification 55
2.11: Large scale plasmid purification 56
2.12: Genomic DNA purification 56
2.13: DNA quantification and quality control 57
2.14: Polymerase chain reaction 57
2.15: Primers for PCR genotyping 58
2.16: RNA extraction 59

2.17: RNA quantification and quality control 60
2.18: RNA bead purification 60
2.19: Bioanalyzer 61
2.20: Dnase Treatment 62
2.21: Complementary DNA synthesis 62
2.22: Quantitative real-time polymerase chain reaction (qRT-PCR) 63
2.23: Gel electrophoresis of PCR products 64
2.24: Sanger Sequencing 65
2.25: Protein extraction 65
2.26: Protein quantification 65
2.27: Single guide RNA design 66
2.28: Cloning of sgRNA into PX458 CRISPR vector 67
2.29: Electroporation 69
2.30: Cell line expansion and single cell sorting 70
2.31: CRISPR/Cas9 off-target analysis 71
2.32: Data analyses 72
Chapter 3 73
Using CRISPR/Cas9 To Generate Models of Myeloid Malignancy 73
3.1: Introduction 73
3.2: Results 75
3.2.1: Generating *TET2* knockouts workflow 75
3.2.2: The *TET2* gene targeting strategy 76
3.2.3: Using the CRISPOR platform to design *TET2* targeted sgRNAs 77
3.2.4: *TET2* sgRNAs were successfully cloned into the PX458 plasmid 78
3.2.5: Validating the efficiency of the *TET2* sgRNA variants using ICE analysis 82
3.2.6: The *ASXL1* gene targeting strategy 85
3.2.7: Using the CRISPOR platform to design *ASXL1* targeted sgRNAs 86
3.2.8: Validating the efficiency of the *ASXL1* sgRNA variants using ICE analysis 87
3.2.9: Generating erythroid and myeloid models of myeloid malignancy 89
3.2.10 Messenger RNA expression of *TET2* and *ASXL1* in models of myeloid malignancy 93
3.2.11 Off-target analysis of *TET2* and *ASXL1* targeted sgRNAs 97
3.3: Discussion 99
Chapter 4 104
The Impact of *TET2* and *ASXL1* Mutations on Cellular Function 104
4.1: Introduction 104
4.2: Results 106
4.2.1: The effect of *TET2* and *ASXL1* mutations on the proliferation in HUDEP2 cells 106
4.2.2: The effect of *TET2* and *ASXL1* mutations on the proliferation in KBM5 cells 107

4.2.3: The effect of *TET2* and *ASXL1* mutations on the cell cycle in HUDEP2 cells 109
4.2.4: The effect of *TET2* and *ASXL1* mutations on the cell cycle in KBM5 cells 112
4.2.5: The effect of *TET2* and *ASXL1* mutations on the apoptosis in HUDEP2 cells 116
4.2.6: The effect of *TET2* and *ASXL1* mutations on the apoptosis in KBM5 cells 117
4.2.7: The effect of *TET2* and *ASXL1* mutations on erythroid differentiation in HUDEP2 cells.. 119
4.2.8: The effect of *TET2* and *ASXL1* mutations on myeloid differentiation in KBM5 cells..... 121
4.3: Discussion 122
Chapter 5 130
Human Haematopoietic Stem and Progenitor Cell Expansion 130
5.1: Introduction 130
5.2: Results 132
5.2.1: Isolation of CD133+ cells from fresh human umbilical cord blood 132
5.2.2: The selection of CD133+ cell donor samples isolated from fresh human UCB 134
5.2.3: The *ex vivo* expansion of UCB derived CD133+ HSPCs 137
5.2.4: The *ex vivo* expansion of UCB derived CD133+ HSPCs extended 147
Chapter 6 167
General Discussion 167
6.1: General discussion 167
6.2: Conclusions and future work 173
Bibliography 178

List of Figures

Figure 1: The organisation of physiological bone marrow. 3
Figure 2: Human aging increases maturing myeloid cell density in haematopoietic areas surrounding adipocytes in the bone marrow. 5
Figure 3: The classical haematopoietic hierarchy. In the classical model, LT-HSCs sit at the top of hierarchy. 10
Figure 4: A graphical representation of a continuum-based model of bone marrow haematopoiesis 14
Figure 5: The latest haematopoietic stem cell hierarchy model in mice. 16
Figure 6: There are two ways in which double stranded breaks, which have been induced by Cas9 may be repaired. 35
Figure 7: Representative gating strategy for the assessment of viability and purity of CD133+ cells isolated from fresh human umbilical cord blood. 47
Figure 8: Representative gating strategy for cell cycle analysis. 53
Figure 9: Representative gating strategy for apoptosis analysis. 54
Figure 10:The PX458 plasmid. 69
Figure 11: Experimental workflow for the design and engineering of sgRNAs targeted towards TET2. 75
Figure 12: The Human TET2 Gene (GRCh38.p13) Assembly. 76
Figure 13: Incorporation of downstream TET2 sgRNAs into the PX458 plasmid. 79
Figure 14: Incorporation of upstream TET2 sgRNAs into the PX458 plasmid. 79
Figure 15: Percentage of bulk sorted GFP+ HUDEP2 cells within live cells, for cells containing the PX458 plasmid expressing downstream TET2 sgRNA variants, n=1. Each plasmid containing a single sgRNA sequence was transfected into HUDEP2 cells, before bulk sorting for GFP. As can be observed, the percentage of bulk sorted HUDEP2 cells are approximately 99% GFP+ cells using Downstream A (black), Downstream B (dark grey)) and Downstream C (light grey). 80
Figure 16: Percentage of bulk sorted GFP+ HUDEP2 cells within live cells, for cells containing the PX458 plasmid expressing upstream TET2 sgRNA variants, n=1. Each plasmid containing a single sgRNA sequence was transfected into HUDEP2 cells, before bulk sorting for GFP. As can be observed, the percentage of bulk sorted HUDEP2 cells are approximately 99% GFP+ cells using Upstream A (black), Upstream B (dark grey)) and Upstream C (light grey). 81
Figure 17: ICE analysis of downstream TET2 sgRNA variants A, B and C, showing percentage efficiency as a product of ICE score and KO score. It can be observed, the most efficient downstream sgRNA was downstream B, with an ICE score of 93%. 83
Figure 18: Sanger sequence traces of downstream TET2 sgRNA variants. Nucleotides are represented by different colour peaks. As can be observed, the sequence collapse for downstream B is the most pronounced, correlating well with this sgRNA having the highest efficiency. 83
Figure 19: ICE analysis of upstream TET2 sgRNA variants A, B and C, showing percentage efficiency as a product of ICE score and KO score. It can be observed, the most efficient upstream sgRNA was upstream A, with an ICE score of 93%. 84
Figure 20: Sanger sequence traces of upstream TET2 sgRNA variants. Nucleotides are represented by different colour peaks. As can be observed, the sequence collapse for upstream A is the most pronounced, correlating well with this sgRNA having the highest efficiency. 84
Figure 21: The Structure of the ASXL1 gene. ASXL1 maps to the chromosome region 20q11 and comprises 12 exons. 85
Figure 22: The p.G646Wfs*12 ASXL1 Targeting Strategy. 86

Figure 23: The G710X ASXL1 Targeting Strategy. 86
Figure 24: ICE analysis of ASXL1 sgRNA variants A, B and C, showing percentage efficiency as a product of ICE score and KO score. It can be observed, that the most efficient ASXL1 sgRNA was ASXL1 A, with a knockout score of 41%. 88
Figure 25: Sanger sequence traces of ASXL1 A, B and C targeted sgRNA variant. Nucleotides are represented by different colour peaks. As can be observed, the sequence collapse for ASXL1 A is the most pronounced, correlating well with this sgRNA having the highest efficiency. 88
Figure 26: Experimental workflow for the engineering erythroid and myeloid models of myeloid malignancies in HUDEP2 and KBM5 cells. CRISPR/Cas9 cocktails were transfected into respective cell lines and subsequently single cell sorted into 96-well plates. Following DNA extraction, PCR analysis was conducted, and the samples were run on an agarose gel, in addition to Sanger sequencing to confirm the genotype. 90
Figure 27: The TET2 gene excision event. 90
Figure 28: Agarose gels depicting HUDEP2 and KBM5 genotypes in single cell colonies. 91
Figure 29: A Sanger sequencing trace showing a successful p.G646Wfs*12 ASXL1 heterozygous mutation knock-in. The incorporation can be depicted by the arrow. Nucleotides are represented by different colours. 92
Figure 30: Sanger sequencing trace showing successful heterozygous and homozygous G710X ASXL1 mutation correction. 92
Figure 31: TET2 mRNA expression in the HUDEP2 cell line variants. 94
Figure 32: ASXL1 mRNA expression in the HUDEP2 cell line variants. 95
Figure 33: TET2 mRNA expression in the KBM5 cell line variants. 96
Figure 34: ASXL1 mRNA expression in the KBM5 cell line variants. 97
Figure 35: The impact of TET2 and ASXL1 mutations on proliferation of HUDEP2 cells over a 4-day period, shown as fluorescence ($x10^4$). 107
Figure 36: The impact of TET2 and ASXL1 mutations on proliferation of KBM5 cells over a 4-day period, shown as fluorescence ($x10^4$). 109
Figure 37: The impact of TET2 and ASXL1 mutations on G0-G1 of the cell cycle in HUDEP2 cells. ... 110
Figure 38: The impact of TET2 and ASXL1 mutations on S-Phase of the cell cycle in HUDEP2 cells. . 111
Figure 39: The impact of TET2 and ASXL1 mutations on G2/M of the cell cycle in HUDEP2 cells. 112
Figure 40: The impact of TET2 and ASXL1 mutations on G0-G1 of the cell cycle in KBM5 cells. 114
Figure 41: The impact of TET2 and ASXL1 mutations on S-Phase of the cell cycle in KBM5 cells. 115
Figure 42: The impact of TET2 and ASXL1 mutations on G2/M of the cell cycle in KBM5 cells. 116
Figure 43: The impact of TET2 and ASXL1 mutations on apoptosis in HUDEP2 cells. 117
Figure 44: The impact of TET2 and ASXL1 mutations on apoptosis in KBM5 cells. 118
Figure 45: The impact of TET2 and ASXL1 mutations on HUDEP2 cell differentiation along the erythroid lineage as assessed by expression of CD36 (black histograms), CD71 (grey histograms) and CD235a (dark grey histograms) using FACS. 120
Figure 46: The impact of TET2 and ASXL1 mutations on myeloid differentiation markers in KBM5 cells as assessed by expression of CD11b (black histograms), CD14 (grey histograms) and CD15 (dark grey histograms) using FACS. 122
Figure 47: Representative gating strategy for the assessment of viability and purity of CD133+ cells isolated from fresh human umbilical cord blood. 133
Figure 48: Percentage viability and percentage purity of CD133+ cells isolated from fresh human UCB. 134

Figure 49: Representative gating strategy for TNC expansion of CD133+ cells used for donor selection. 135
Figure 50: Expansion and viability of human UCB CD133+ cells. 136
Figure 51: Representative gating strategy used to select HSPC subsets post expansion (from human UCB derived CD133+ cells) using small molecules and cytokines. CMP (Common Myeloid Progenitor), GMP (Granulocyte-Monocyte Progenitor), MEP (Megakaryocyte-Erythrocyte Progenitor), HSC (Haematopoietic Stem Cell), MPP (Multi-Potent Progenitor) and CLP (Common Lymphoid Progenitor). The data are represented as a dot plot. 139
Figure 52: Total nucleated cell (TNC) fold expansion from fresh human umbilical cord blood using novel combinations of small molecules at day 5 of day 0. Data are expressed as mean ± SD, n=3 biological replicates. Treatments are compared to basal, (**** P≤0.0001). A one-way ANOVA test was used for statistical analysis, coupled with the Dunnett's multiple comparison post hoc test. Treatment descriptions can be observed in table 33. 140
Figure 53: Lin-CD34+CD38-CD45RA-CD90+CD49f+ HSC (Haematopoietic Stem Cells) fold expansion from fresh human umbilical cord blood using novel combinations of small molecules at day 5 of day 0. Data are expressed as mean ± SD, n=3 biological replicates. Treatments are compared to basal, (ns=not significant). A one-way ANOVA test was used for statistical analysis, coupled with the Dunnett's multiple comparison post hoc test. Treatment descriptions can be observed in table 33. 141
Figure 54: Lin-CD34+CD38-CD45RA-CD90- MPP (Multi-Potent Progenitor) fold expansion from fresh human umbilical cord blood using novel combinations of small molecules at day 5 of day 0. Data are expressed as mean ± SD, n=3 biological replicates. Treatments are compared to basal, (*** P≤0.001). A one-way ANOVA test was used for statistical analysis, coupled with the Dunnett's multiple comparison post hoc test. Treatment descriptions can be observed in table 33. 142
Figure 55: Lin-CD34+CD38-CD45RA+CD90- CLP (Common Lymphoid Progenitor) fold expansion from fresh human umbilical cord blood using novel combinations of small molecules at day 5 of day 0. Data are expressed as mean ± SD, n=3 biological replicates. Treatments are compared to basal, (* P≤0.05). A one-way ANOVA test was used for statistical analysis, coupled with the Dunnett's multiple comparison post hoc test. Treatment descriptions can be observed in table 33. 143
Figure 56: Lin-CD34+CD38+CD123+CD45RA- CMP (Common Myeloid Progenitor) fold expansion from fresh human umbilical cord blood using novel combinations of small molecules at day 5 of day 0. Data are expressed as mean ± SD, n=3 biological replicates. Treatments are compared to basal, (ns=not significant). A one-way ANOVA test was used for statistical analysis, coupled with the Dunnett's multiple comparison post hoc test. Treatment descriptions can be observed in table 33. 144
Figure 57: Lin-CD34+CD38+CD123-CD45RA+ GMP (Granulocyte-Macrophage Progenitor) fold expansion from fresh human umbilical cord blood using novel combinations of small molecules at day 5 of day 0. Data are expressed as mean ± SD, n=3 biological replicates. Treatments are compared to basal, (**P≤0.01). A one-way ANOVA test was used for statistical analysis, coupled with the Dunnett's multiple comparison post hoc test. Treatment descriptions can be observed in table 33. 145
Figure 58: Lin-CD34+CD38+CD123-CD45RA- MEP (Megakaryocyte-Erythrocyte Progenitor) fold expansion from fresh human umbilical cord blood using novel combinations of small molecules at day 5 of day 0. Data are expressed as mean ± SD, n=3 biological replicates. Treatments are compared to basal, (ns=not significant). A one-way ANOVA test was used for statistical analysis, coupled with the Dunnett's multiple comparison post hoc test. Treatment descriptions can be observed in table 33. 146
Figure 59: Representative gating strategy used to select HSPC subsets post expansion (from human UCB derived CD133+ cells) using small molecules and cytokines. (CMP) Common Myeloid Progenitor,

GMP (Granulocyte-Macrophage Progenitor), (MEP) Megakaryocyte-Erythrocyte Progenitor, HSC (Haematopoietic Stem Cell), (MPP) Multi-Potent Progenitor and (CLP) Common Lymphoid Progenitor. The data are represented as a dot plot. 152

Figure 60: Total nucleated cell (TNC) fold expansion from fresh human umbilical cord blood using novel combinations of small molecules at day 5 of day 0. Data are expressed as mean ± SD, n=3 biological replicates. Treatments are compared to basal, (* P≤0.05), (*** P≤0.001), (**** P≤0.0001). A one-way ANOVA test was used for statistical analysis, coupled with the Dunnett's multiple comparison post hoc test. Treatment descriptions can be observed in table 34. 153

Figure 61: Lin-CD34+CD38-CD45RA-CD90+CD49f+ HSC (Haematopoietic Stem Cells) fold expansion from fresh human umbilical cord blood using novel combinations of small molecules at day 5 of day 0. Data are expressed as mean ± SD, n=3 biological replicates. Treatments are compared to basal, (ns=not significant). A one-way ANOVA test was used for statistical analysis, coupled with the Dunnett's multiple comparison post hoc test. Treatment descriptions can be observed in table 34. 154

Figure 62: Lin-CD34+CD38-CD45RA-CD90- MPP (Multi-Potent Progenitor) fold expansion from fresh human umbilical cord blood using novel combinations of small molecules at day 5 of day 0. Data are expressed as mean ± SD, n=3 biological replicates. Treatments are compared to basal, (ns=not significant). A one-way ANOVA test was used for statistical analysis, coupled with the Dunnett's multiple comparison post hoc test. Treatment descriptions can be observed in table 34..... 155

Figure 63: Lin-CD34+CD38-CD45RA+CD90- CLP (Common Lymphoid Progenitor) fold expansion from fresh human umbilical cord blood using novel combinations of small molecules at day 5 of day 0. Data are expressed as mean ± SD, n=3 biological replicates. Treatments are compared to basal, (ns=not significant), (**P≤0.01), (*** P≤0.001), (**** P≤0.0001). A one-way ANOVA test was used for statistical analysis, coupled with the Dunnett's multiple comparison post hoc test. Treatment descriptions can be observed in table 34. 156

Figure 64: Lin-CD34+CD38+CD123+CD45RA- CMP (Common Myeloid Progenitor) fold expansion from fresh human umbilical cord blood using novel combinations of small molecules at day 5 of day 0. Data are expressed as mean ± SD, n=3 biological replicates. Treatments are compared to basal, (ns=not significant). A one-way ANOVA test was used for statistical analysis, coupled with the Dunnett's multiple comparison post hoc test. Treatment descriptions can be observed in table 34. 157

Figure 65: Lin-CD34+CD38+CD123-CD45RA+ GMP (Granulocyte-Macrophage Progenitor) fold expansion from fresh human umbilical cord blood using novel combinations of small molecules at day 5 of day 0. Data are expressed as mean ± SD, n=3 biological replicates. Treatments are compared to basal, (ns=not significant). A one-way ANOVA test was used for statistical analysis, coupled with the Dunnett's multiple comparison post hoc test. Treatment descriptions can be observed in table 34. 158

Figure 66: Lin-CD34+CD38+CD123-CD45RA- MEP (Megakaryocyte-Erythrocyte Progenitor) fold expansion from fresh human umbilical cord blood using novel combinations of small molecules at day 5 of day 0. Data are expressed as mean ± SD, n=3 biological replicates. Treatments are compared to basal, (ns=not significant). A one-way ANOVA test was used for statistical analysis, coupled with the Dunnett's multiple comparison post hoc test. Treatment descriptions can be observed in table 34. 159

Figure 67: Cell surface biomarker analyses of the purified human UCB CD34+ cells..... 176

Figure 68: Primary and secondary engraftment in the bone marrow (BM) and spleen of NSG mice.. 177

List of Tables

Table 1: Reagents/consumables and supplier name. 42
Table 2: HUDEP2 cell media components, catalogue numbers, concentration and volumes used. ... 44
Table 3: KBM5 cell media components, catalogue numbers, concentrations and volumes used. 45
Table 4: CD133+ cell overnight media cocktail components, catalogue numbers, concentration and volumes used. 49
Table 5: Treatment combinations used for the ex vivo expansion of UCB derived CD133+ HSPCs..... 49
Table 6: Treatment combinations used for UCB derived CD133+ HSPC expansion. 50
Table 7: Antibody concentrations used for differentiation analysis sample staining. 55
Table 8: PCR reagents. 58
Table 9: PCR thermocycler conditions. 58
Table 10: PCR primer sequences and Tm temperatures for TET2 and ASXL1 primers. 59
Table 11: cDNA synthesis components. 63
Table 12: Thermocycler conditions for cDNA synthesis. 63
Table 13: qRT-PCR components. 64
Table 14: TaqMan Probes. 64
Table 15: TET2 sgRNA sequences. 66
Table 16: ASXL1 sgRNA sequences. 67
Table 17: Components for restriction enzyme-based digestion of plasmid. 68
Table 18: Components for plasmid ligation. 68
Table 19: Upstream A TET2 sgRNA off-target locus and primer sequences. 71
Table 20: Downstream B TET2 sgRNA off-target locus and primer sequences. 71
Table 21: ASXL1 p.G646Wfs*12 sgRNA off-targets locus and primer sequences. 72
Table 22: ASXL1 G710X sgRNA off-target locus and primer sequences 72
Table 23: Predicted specificity, efficiency and out of frame scores for TET2 sgRNA sequences. 78
Table 24: Predicted specificity and efficiency scores for ASXL1 sgRNA sequences. 87
Table 25: The HUDEP2 cell model variants engineered using CRISPR Cas9 technology. 93
Table 26: The KBM5 cell model variants engineered using CRISPR Cas9 technology. 93
Table 27: Upstream A TET2 sgRNA (TATATGTACATGCTGGTTGT) off-target analysis data. 98
Table 28: Downstream B TET2 sgRNA (TCTGGCATCCATCGCAAAGT) off-target analysis data. 98
Table 29: ASXL1 p.G646Wfs*12 sgRNA (GCGGCCACCACTGCCATCGG) off-targets analysis data. 98
Table 30: ASXL1 G710X sgRNA (TCTAAATGGGGAGCATACCC) off-targets analysis data 99
Table 31: CD biomarkers of HSPCs 131
Table 32: Factors that aid expansion of HSPCs ex vivo 131
Table 33: Treatment combinations used for the ex vivo expansion of UCB derived CD133+ HSPCs. 138
Table 34: Treatment combinations used for UCB derived CD133+ HSPC expansion. D=Day. Concentrations used were as follows: SCF (100ng/ml), TPO (20ng/ml), FLT3L (100ng/ml), Scriptaid (1µM), UM171 (35nM), SR1 (500nM), NAM (5mM), VPA (1mM), PEITC (10µM) and ZLLYFMK (10µM). 151

List of Abbreviations

AGAP2	ArfGAP With GTPase Domain, Ankyrin Repeat And PH Domain 2
AGM	Aorta-gonad-mesonephros
AML	Acute myeloid leukaemia
ANOVA	Analysis of variance
AQP10	Aquaporin 10
ARHGAP11A	Rho GTPase Activating Protein 11A
ASXL1	Additional Sex Combs Like Transcriptional Regulator 1
B2M	Beta-2-Microglobulin
BCA	Bicinchoninic Acid
BD	becton dickinson
BFU-E	Erythroid burst-forming units
BRCC3	BRCA1/BRCA2-Containing Complex Subunit 3
C4BPB	Complement Component 4 Binding Protein Beta
CAPN3	Calpain 3
CBL	Cbl Proto-Oncogene
CD	Cluster of differentiation
CDKN2A	Cyclin Dependent Kinase Inhibitor 2A
cDNA	Complementary DNA
CFU	Colony-forming unit
CFU-GEMM	Colony forming unit granulocyte erythrocyte macrophage megakaryocyte
CFU-GM	Colony forming unit granulocyte macrophage
CLP	Common lymphoid progenitors
CMML	Chronic myelomonocytic leukaemia
CMP	Common lymphoid progenitors

CRIPSR	Clustered regularly interspaced short palindromic repeats
CXCR4	CXC chemokine receptor type 4
DAPI	4’,6-diamidino-2-phenylindole
DEPC	Diethyl Pyrocarbonate
DLK-1	Protein delta homolog 1
DLL1	Delta-like protein 1
DMSO	Dimethyl Sulfoxide
DNA	Deoxyribonucleic acid
DNMT3A	DNA methyltransferase 3 alpha
DPBS	Dulbecco's phosphate-buffered saline
DYRK1B	Dual specificity tyrosine-phosphorylation-regulated kinase 1B
EBF1	EBF Transcription Factor 1
EDTA	Ethylenediaminetetraacetic acid
ETV6	ETS Variant Transcription Factor 6
EZH2	Enhancer of zeste homolog 2
FACS	Fluorescence-activated cell sorting
FAM53A	Family With Sequence Similarity 53 Member A
FBS	Fetal Bovine Serum
FITC	Fluorescein isothiocyanate
FLT-3L	FMS-like tyrosine kinase 3 ligand
FMO	Fluorescence Minus One
FSC	Forward scatter
GCSF	Granulocyte colony-stimulating factor
GFP	Green fluorescent protein
GMP	Granulocyte macrophage progenitors

GRK4	G protein-coupled receptor kinase 4
HDAC	Histone deacetylases
HDR	Homology-Directed Repair
HK2	Hexokinase 2
HOXA	Homeobox A
HOXD	Homeobox B
HPC	Haematopoietic progenitor cell
HPP	High proliferative potential
HSA	Human serum albumin
HSC	Haematopoietic stem cell
HSPC	Haematopoietic stem cells and progenitor cell
HUDEP2	Human Umbilical cord blood-Derived Erythroid Progenitor-2
ICE	Inference of CRISPR Edits
IDAX	CXXC Finger Protein 4
IDH2	Isocitrate Dehydrogenase (NADP(+)) 2
IL	Interleukins
IMDM	Iscove Modified Dulbecco Media
IMS	Industrial methylated spirit
INDEL	Insertion or deletion
iPSC	Induced pluripotent stem cells
IPSS-R	Revised International Prognostic Scoring System
JAK	Janus Kinase
KIT	KIT Proto-Oncogene, Receptor Tyrosine Kinase
KO	Knock out
KRAS	Kirsten rat sarcoma viral oncogene homolog

LB	Luria-Bertani
LCD	Liquid-crystal display
LSM	Lymphocyte separation medium
MDS	Myelodysplastic syndrome
MDS-U	Myelodysplastic syndrome, unclassified
MEP	Megakaryocyte-Erythrocyte Progenitor
MFI2	Melanotransferrin
MGDF	Megakaryocyte growth and development factor
MM	Master mix
MNC	Mononuclear cell
MPB	Mobilised peripheral blood
MPN	Myeloproliferative neoplasm
MPP	Multipotent progenitor cells
MYLK	Myosin Light Chain Kinase
NAM	Nicotinamide
NHEJ	Non-homologous end joining
NRAS	NRAS Proto-Oncogene, GTPase
OGT	O-linked β-N-acetylglucosamine transferase
PAM	Protospacer adjacent motif
PARP15	Poly(ADP-Ribose) Polymerase Family Member 15
PBS	Phosphate-buffered saline
PCR	Polymerase chain reaction
PDF	Portable Document Format
PEITC	Phenethyl isothiocyanate
PGE2	Prostaglandin E2

PGR	Progesterone receptor
PI	Propidium iodide
PIN4	Peptidylprolyl Cis/Trans Isomerase, NIMA-Interacting 4
PPP1R13L	Protein Phosphatase 1 Regulatory Subunit 13 Like
PRC2	polycomb repressive complex 2
PRKCE	Protein Kinase C Epsilon
PRPH2	Peripherin 2
QC	Quality Control
RA	Refractory anaemia
RAEB	Refractory anaemia with excess blasts
RARS	Refectory anaemia with ringed sideroblasts
RAS	Rat sarcoma
RCMD	Refractory cytopenia with multilineage dysplasia
RCMD-RS	Refractory cytopenia with multilineage dysplasia and ringed sideroblasts
RIPA	Radioimmunoprecipitation assay
RNA	Ribonucleic acid
RT	Reverse transcriptase
RTK	Receptor tyrosine kinase
RUNX1	RUNX Family Transcription Factor 1
sAML	Secondary acute myeloid leukaemia
SCF	Stem cell factor
SFEM	Serum-Free Expansion Medium
SIRT1	Sirtuin 1
SLC22A10	Solute Carrier Family 22 Member 10
SNX24	Sorting Nexin 24

SR1	StemRegenin 1
SRSF2	Serine And Arginine Rich Splicing Factor 2
SSC	Side-scatter
ssODN	Single-stranded oligodeoxynucleotide
STAT	Signal Transducer And Activator Of Transcription
SUZ12	SUZ12 Polycomb Repressive Complex 2 Subunit
TBE	Tris/Borate/EDTA
TET	Ten-eleven translocation
TET2	Ten-eleven translocation 2
TMEM27	Transmembrane Protein 27
TNC	Total nucleated cell
TPO	Thrombopoietin
UCB	Umbilical cord blood
UHRF2	Ubiquitin Like With PHD And Ring Finger Domains 2
UK	United kingdom
VPA	Valproic Acid
VprBP	Vpr (HIV-1) Binding Protein
WBC	White blood cell
WT	Wildtype
α-KG	Alpha ketoglutarate

Chapter 1
General Introduction

1.1: Haematopoiesis

1.1.1: An introduction to haematopoietic stem cell biology

The formation of blood cells from haematopoietic stem and/or progenitor cells is called haematopoiesis (Pinho and Frenette 2019). Haematopoiesis is critically important to survival as it generates a variety of circulating blood and tissue resident cells, the functions of which include the provision of oxygen to tissues, haemostasis in the event of bleeding and surveillance to prevent or modulate infections. Haematopoiesis is actively occurring in the body, but the actual process is modified throughout development. It can also be accelerated in the event of severe infection or bleeding, with skewing towards the generation of the haematopoietic lineages considered appropriate for the function required (Pinho and Frenette 2019). Remarkably, under normal homeostatic conditions in the adult human, between 1-5 hundred billion blood cells (the majority comprising platelets and erythrocytes) originate from haematopoietic precursors in the human bone marrow per day.

1.1.2: Ontogeny

Haematopoiesis occurs in various tissues depending on the stage of development and species studied (Ng and Alexander 2017). It has generally been divided into primitive and definitive haematopoiesis, with primitive haematopoiesis in amniotes commencing first in the blood islands of the extra-embryonic yolk sac. In mice, during the middle of gestation, definitive HSCs emerge from the aorta-gonad-mesonephros region (AGM) or more precisely from haemogenic endothelium of the ventral floor of the dorsal aorta (at approximately E10) (Cumano, Dieterlen-Lievre et al. 1996,

Medvinsky and Dzierzak 1996, Ng and Alexander 2017). Recent studies show that HSCs with a Sca1+ c-kit+ CD41+ phenotype emanate from central aortic cells in mice. It has also been shown that shear stress may be an important factor for this process to take place (Ng and Alexander 2017). The main growth factor pathways, which regulate early mouse haematopoiesis include the fibroblast growth factor pathway, the transforming growth factor, the bone morphogenic protein 4 pathway and the vascular endothelial growth factor pathway (Ng and Alexander 2017). The definitive HSCs seed the fetal liver in mice, and HSCs are detected at E9 of mouse development, with a substantial expansion of cells between E12-15 (Ivanovs, Rybtsov et al. 2017). Finally, on E18, the HSCs migrate to the fetal bone, which remains the main site of haematopoiesis in post-natal life. It was shown by Moore and colleagues that in *ex vivo* cultures, fetal liver derived stromal cells were able to support HSCs (Ivanovs, Rybtsov et al. 2017). These myelosupportive stromal cells are able to differentiate into osteoblasts, chondrocytes and adipocytes (Pittenger, Discher et al. 2019). More recently, in mice, the placenta has also been implicated as a key haematopoietic tissue throughout development (Wu, Zhang et al. 2018). Hormones, which mediate vascularisation are produced by the placenta, and since vasculogenesis and haematopoiesis are closely linked, the placenta is thought to have a role in regulating haematopoiesis (Wu, Zhang et al. 2018). Important factors which influence and promote haematopoiesis are produced by trophoblasts, which reside in the placenta, these factors also enter during maternal blood influx. It is well recognised that the placental microenvironment helps support the stemness of HSCs without affecting differentiation (Wu, Zhang et al. 2018). In the human, primitive haematopoiesis originates in the yolk sac blood islands at 2-3 weeks post-conception and prior to the development of the circulation (Baron, Isern et al. 2012, Popescu, Botting et al. 2019, Wittamer and Bertrand 2020). Definitive haematopoiesis is derived from HSCs that originate in the human from the AGM region 3-4 weeks post-conception and more specifically from the haemogenic endothelia of the ventral wall of

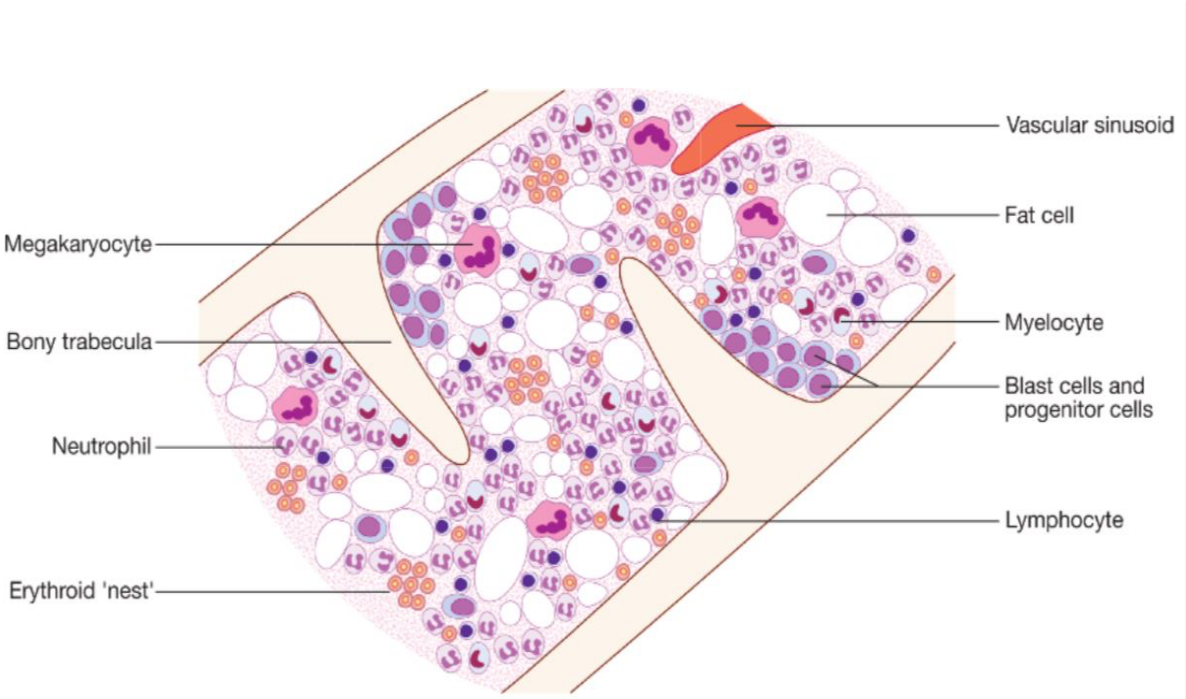


Figure 1: The organisation of physiological bone marrow.
Vascular sinusoid: Tubular space for the passage of blood. Fat cell: Primarily composed of adipose tissue and store energy as fat. Myelocyte: An early precursor of granulocytic series of white blood cells. Blast cell: Precursors to the mature, circulating blood cells. Progenitor cells: Descendants of stem cells that can differentiate to form one or more types of cells. Lymphocyte: A type of white blood cell in the immune system. Megakaryocyte: A large bone marrow cell responsible for the production of blood thrombocytes. Bony trabecula: Form of bone tissue organised into a network of interconnected rods surrounding the pores which comprise bone marrow. Neutrophil: The most abundant type of granulocyte and are important immune cells. Erythroid nest: A collection of Erythroid cells. Taken from: (Ralston, Penman et al. 2018). Copyright permission to reproduce this figure was obtained from Elsevier.

the dorsal aorta of the human embryo (Medvinsky and Dzierzak 1996, Tavian, Hallais et al. 1999, Popescu, Botting et al. 2019, Crosse, Gordon-Keylock et al. 2020). These HSCs migrate to and colonise the fetal liver, which then becomes the major haematopoietic organ in the developing human. Under normal human homeostasis, the spleen is colonised briefly by HSCs. The human bone marrow then becomes the major haematopoietic organ (with colonisation by HSCs commencing around 10.5 weeks post-conception and reaching a maximum around week 20), a function which continues throughout post-natal life (Mikkola and Orkin 2006, Galan-Diez and Kousteni 2017, Popescu, Botting et al. 2019, Li, Kong et al. 2020). Haematopoiesis is increasingly restricted throughout development postnatally, as the red marrow in children in replaced by yellow marrow, more commonly known as fat (Zhang, Zhang et al. 2019). This means that by adulthood, normal haematopoiesis only takes place in the pelvis, skull, ribs, proximal femora, sternum, vertebrae, upper humeri and clavicles (Zhang, Zhang et

al. 2019). However, it has been shown that, during the increased demand for blood cells, red marrow expands, and, in certain pathological circumstances, extra-medullary haematopoiesis can occur (Yang, Chen et al. 2020).

As illustrated in figure 1, the post-natal bone marrow contains both immature precursor cells including HSCs and their progeny at various stages of differentiation (Zhang, Zhang et al. 2019). Some of these cells are stored and then released when there is an increase in demand for them (Walkley, Shea et al. 2007). The surrounding connective tissue stroma, which is made up of macrophages, reticular cells, blood vessels, a variety of stromal cell subsets, nerve fibres and fat cells, interact closely with haematopoietic cells (Walkley, Shea et al. 2007). For example, during erythropoiesis, the normal bone marrow features a central macrophage, which serves as a source of iron and phagocytosis extruded nuclei for maturing erythroid cells. Platelets are produced and released into the vascular sinuses by polyploid megakaryocytes (Lucas 2017), although recent studies suggest that the process of platelet production from megakaryocytes can also occur in the lung (Yeung, Villacorta-Martin et al. 2020). Clusters of white cell precursors are in close proximity to the bone trabeculae; with the marrow spaces allowing maturing cells to migrate towards the vascular sinuses. The inter-trabecular spaces are populated with plasma cells, which are antibody secreting mature B cells, but represent less than 5% of the bone marrow population (Lucas 2017). Figure 2 demonstrates recent studies which show that conditions such as aging can alter haematopoietic cell distribution in an altered human bone marrow niche.

Multipotent HSCs are the main origin of adult blood cells, despite only comprising 0.01% of all bone marrow cells (Chitteti, Cheng et al. 2010). They can make more HSCs by a process called self-renewal and are able to differentiate to produce a lineage committed hierarchy of progenitor cells, although the classical model of haematopoiesis has been modified a number of times. Resulting differentiated erythromyeloid progeny are characterised according to the type of cell colony they form during a cell culture colony forming experiments (Chitteti, Cheng et al. 2010). There are multiple types of colony forming units (Frisch 2021). CFU-E denotes colony forming unit erythroid, CFU-G

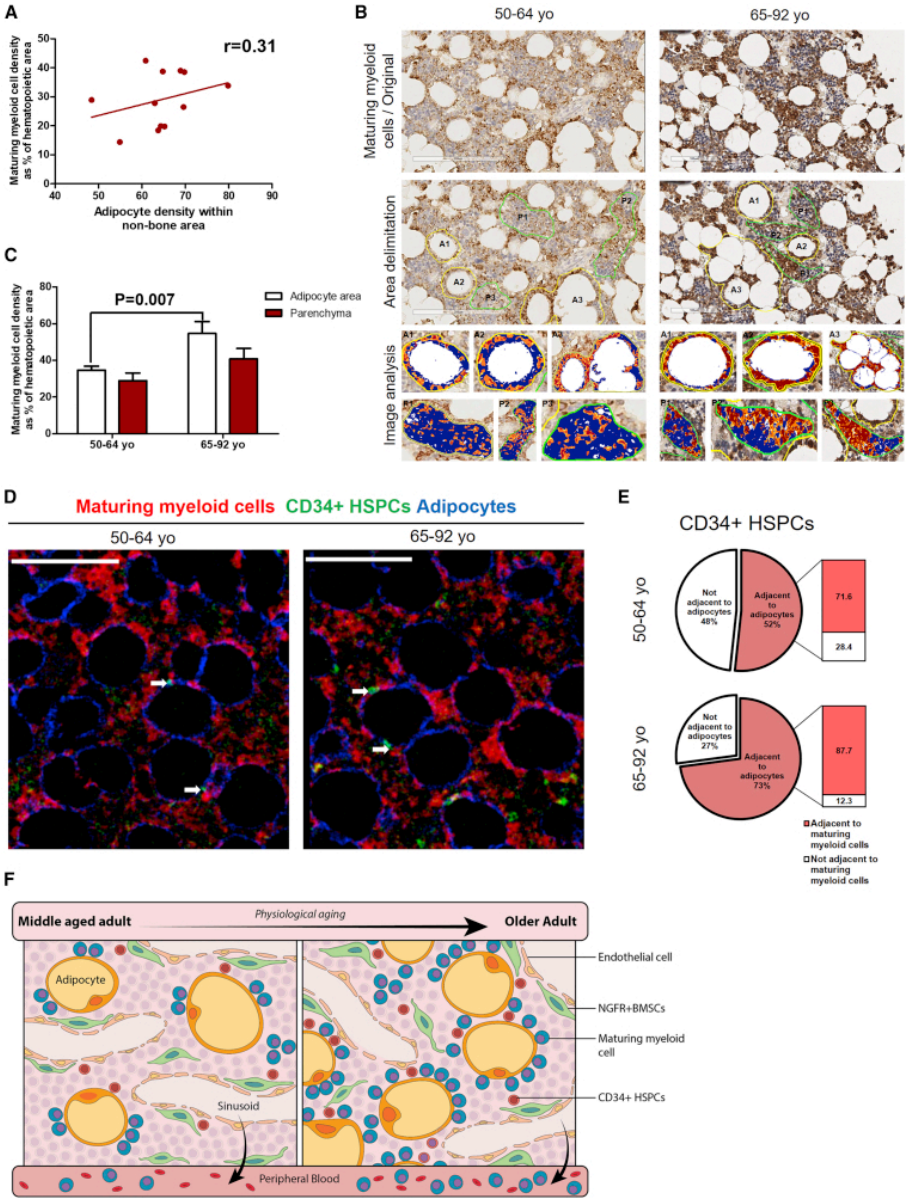


Figure 2: Human aging increases maturing myeloid cell density in haematopoietic areas surrounding adipocytes in the bone marrow. (BM) as assessed by (Aguilar-Navarro, Meza-León et al. 2020). (A) Weak correlation between adipocytes and maturing myeloid cell density. Pearson correlation test, r = 0.31, n = 11. (B) Representative IHC images of maturing myeloid cells (MPO) (original) in BM biopsies from haematologically healthy patients aged 50–64 and 65–92 years (upper panel). Adipocyte area (A) and parenchyma (P) delimitation for image analysis quantification (middle panel). Image analysis of A and P, showing positive (yellow to red) and negative (blue) areas (lower panel). Scale bar, 200µm. (C) Bar graph showing the density of maturing myeloid cells in adipocyte area and parenchyma of the 50–64-year (n = 7) and 65–92-year (n = 5) groups. Two-way ANOVA, p = 0.002, unpaired t test, adipocyte area, p = 0.007. (D) Representative IMC images showing an overlay of maturing myeloid cells (red channel), CD34+ HSPCs (green channel), and adipocytes (blue channel) in BM biopsies from haematologically healthy patients aged 50–64 years (n = 2) and 65–92 years (n = 3). Scale bar, 100µm. Adjacency between maturing myeloid cells, CD34+ HSPCs, and adipocytes are indicated by white arrows. (E) Diagram showing the percentage of CD34+ HSPCs adjacent to maturing myeloid cells out of percentage of CD34+ HSPCs adjacent to adipocytes in both age groups. Percentage of CD34+ HSPCs adjacent to adipocytes was calculated from IHC analysis and the percentage of CD34+ HSPCs adjacent to both adipocytes and maturing myeloid cells was calculated from IMC analysis. (F) Proposed model of physiological aging-related changes in human BM spatial organization. Increasing marrow adiposity correlates with an increased CD34+ HSPC and maturing myeloid density and adjacency with adipocytes, while contact with MSCs and perivascular localization remains unchanged across age groups. Taken from: (Aguilar-Navarro, Meza-León et al. 2020). Copyright permission to reproduce this figure was obtained from Elsevier.

denotes colony forming unit granulocyte, CFU-GM denotes colony forming unit granulocyte macrophage, BFU-E denotes burst forming unit erythroid, CFU-M denotes colony forming unit

macrophage, CFU-Meg denotes colony forming megakaryocyte and CFU-GEMM denotes colony forming unit granulocyte erythrocyte macrophage megakaryocyte (Frisch 2021). Below is a more in depth look into haematopoiesis, starting with the haematopoietic stem cell (HSC) concept and the discovery of haematopoietic stem cells.

1.1.3: Identifying the haematopoietic stem cell

When we first think of the word stem, we think of a plant or a tree. It is no coincidence that this is from where the term stem cell originates. When Ernst Haeckel first coined the term in 1868, he used it to represent the origin of a 'stem tree' (Stammbaum) which depicted the primordial unicellular organism as the source of all multicellular life. It was therefore no great leap to then assign the term 'stem cell' to the fertilised egg (Haeckel 1868).

Following on from Ernst Haeckel's initial use of the term, this stem cell concept was applied to normal and leukaemic haematopoiesis by histopathologists (Pappenheim 1896). This also created the notion that red and white blood cells originate from a common progenitor, which also act as a precursor of myeloid and lymphoid leukaemic cells (Pappenheim 1907). Returning to the tree analogy coined by Ernst Haeckel, stem cells have always represented the root of these 'stem trees', with ordered branches representing the steps by which multipotent stem cells form their progenies.

The first *in vivo* study used to elucidate stem cell function was based on the recovery of murine models from lethal irradiation using bone marrow transplantation (Jacobson, Simmons et al. 1951). Research into radiation sensitivity identified that, when transplanted into syngeneic irradiated murine recipients, donor adult bone marrow was able to protect these recipients from lethal irradiation, by regenerating the irradiation ablated haematopoietic system (Till and Mc 1961). This research constituted a pivotal turning point in our understanding of HSCs as cells in the bone marrow, being versatile enough to generate the complete blood cell system, although when this research took place, the specific cell had not yet been characterised or isolated. These early experiments successfully enabled donor derived colonies from multiple erythromyeloid haematopoietic lineages to be

observed macroscopically in the transplanted recipient's spleen (Till and Mc 1961). Despite not being HSCs, these spleen colony forming units (CFU-S) were still useful in characterising the progenitor cells that may play an integral part in haematopoietic reconstitution (Till and Mc 1961). Through the tracking of cytological abnormalities within distinct murine CFU-S colonies, Becker, McCulloch and Till subsequently were the first to report conclusive proof for the function of multipotent progenitor cells, *in vivo* (Becker, Mc et al. 1963).

The purification of transplantable human HSCs was facilitated by fluorescent activated cell sorting and the identification of the cell surface markers CD34 and CD133 on human HSCs and their immediate progeny (Civin, Strauss et al. 1984, Sutherland, Watt et al. 1988, Miraglia, Godfrey et al. 1997). Experimenting in mice, Spangrude Heimfeld and Weissman showed the usefulness of positive and negative selection strategies (Spangrude, Heimfeld et al. 1988) by highly purifying CFU-S from mouse bone marrow, although mistakenly identifying these cells as the *in vivo* equivalent of HSCs (Lord and Dexter 1988). Traditionally both self-renewal and multipotency have been the two primary criteria for defining HSCs. Serial transplant experiments, progressed significantly by John Dick's group (Dick 1991), are now often used to initially test the haematopoietic self-renewal capacity, with further studies using additional markers such as CD90 and CD49f allowing further purification of the human HSC (Notta, Doulatov et al. 2011). Barcoding studies and analyses in murine and non-human primate models as well as gene therapy studies in humans, together with single cell omics and newer imaging approaches, have enhanced our knowledge of both human and murine HSCs (Cordes, Wu et al. 2020, Ferrari, Jacob et al. 2020, Pei, Shang et al. 2020). On the other hand, the absence of extended self-renewal and a limited lineage differentiation capacity defines progenitor cells (Doulatov, Notta et al. 2010).

Progenitor cell populations which are downstream of HSCs were characterised in more detail around the year 2000, resulting in the haematopoietic differentiation or lineage tree model we see in many textbooks. As depicted in this model, lymphoid potential is segregated from myeloid, erythroid and megakaryocytic lineages. The branching on either side of the tree subsequently follows, where

the progenitor cells become multi, bi and finally uni-potent. By introducing additional cell surface markers, it appeared that there were several differences in the classical HSC differentiation tree model (Adolfsson, Månsson et al. 2005, Doulatov, Notta et al. 2010, Goardon, Marchi et al. 2011). For example, lymphoid and myeloid fates appeared to remain associated until later in the differentiation process and megakaryocytes appeared to branch earlier. The multipotent progenitor also seemed to divide up into specific subpopulations; furthermore since the pool of HSCs are molecularly heterogeneous, the story is more complex (Dykstra, Kent et al. 2007, Yamamoto, Morita et al. 2013, Pietras, Reynaud et al. 2015).

The murine system represents one of the most complex sections of the haemopoietic tree. Studies in this area are particularly advanced, and have demonstrated that it is virtually impossible to compartmentalise all these structures to a single fixed tree. We therefore need to remain creative when representing HSC differentiation graphically. New technologies are enabling us to adapt this classical view, and as a result it appears that there is a hierarchy of a heterogenous population, within which differentiation can take place gradually. These cell populations will be required to remain highly supple in order to meet the requirements of blood homeostasis. The HSC lineage hierarchy will be discussed further in section, '1.1.4: The HSC lineage hierarchy'.

1.1.4: The HSC lineage hierarchy

Accurately defining hierarchical lineage relationships and the regulatory pathways that control the huge generation of blood cells from rare human HSC and their progeny is essential not only for defining basic haematological mechanisms, but also for developing new, robust and high quality blood products to treat life threatening and chronic diseases. These approaches have benefitted from the most recent advances in single cell multi-omics, single cell quantitative RT-PCR and barcoding technologies combined with single cell fate mapping and transplant analyses (Hu, An et al. 2018). Studies on human haematopoietic lineage hierarchies have generally been preceded or informed by research using murine models, even though significant differences in haematopoietic development

exist between the two species, while biases may be introduced with single cell studies, particularly when HSPCs (haematopoietic stem and progenitor cells) are subjected to different microenvironmental influences such as stress-induced conditions, or assessed at different stages of ontology, as elegantly described recently (Popescu, Botting et al. 2019). Nevertheless, while these experimental approaches have enhanced our understanding of hierarchical relationships, they have questioned the classical model of haematopoiesis, most notably in terms of megakaryocyte and platelet lineage development from HSCs.

In the classical and balanced model of Haematopoiesis (Cheng, Zheng et al. 2020), rare self-renewing multipotent HSCs generate all haematopoietic lineages, gradually losing their self-renewing and long term (LT) serial transplantation capacities, and differentiating into increasingly lineage-restricted, but discrete, progenitor cell subpopulations, and eventually into functional mature blood cells. Classically, LT-HSCs were thought to differentiate into intermediate-term (IT)-HSCs and then short term (ST)-HSCs, before giving rise to multipotent progenitor cells (MPPs), which branch into two common, but distinct, progenitor subsets restricted to all myeloid (common myeloid progenitors or CMPs) or all lymphoid cell lineages (common lymphoid progenitors or CLPs) (Kondo 2010).

The classical HSC model was largely established by the Weissman group (Kondo, Weissman et al. 1997, Morrison, Wandycz et al. 1997, Akashi, Traver et al. 2000, Manz, Miyamoto et al. 2002). They suggest that depending on CD34 expression level, murine HSCs divide into two distinct categories, these are: CD34- long term HSCs and CD34+ short term HSCs. Both the human CD34- and the immature CD34+ HSPCs have long term serial repopulation abilities in immunodeficient mouse transplant experiments and can be identified as those cells expressing CD133 (Rocco, Liguori et al. 2012). Long term human HSCs have a more long term capacity of approximately 4-8 months with regards to reconstitution of haematopoiesis in murine models plus the ability to serially repopulate secondary and perhaps tertiary mice, and are quiescent in the bone marrow and rarer. On the other hand, short term human HSCs generally have a reconstitution capacity of less than 1 month to several months in murine models. Long term HSCs differentiate into short term HSCs, following which short term HSCs

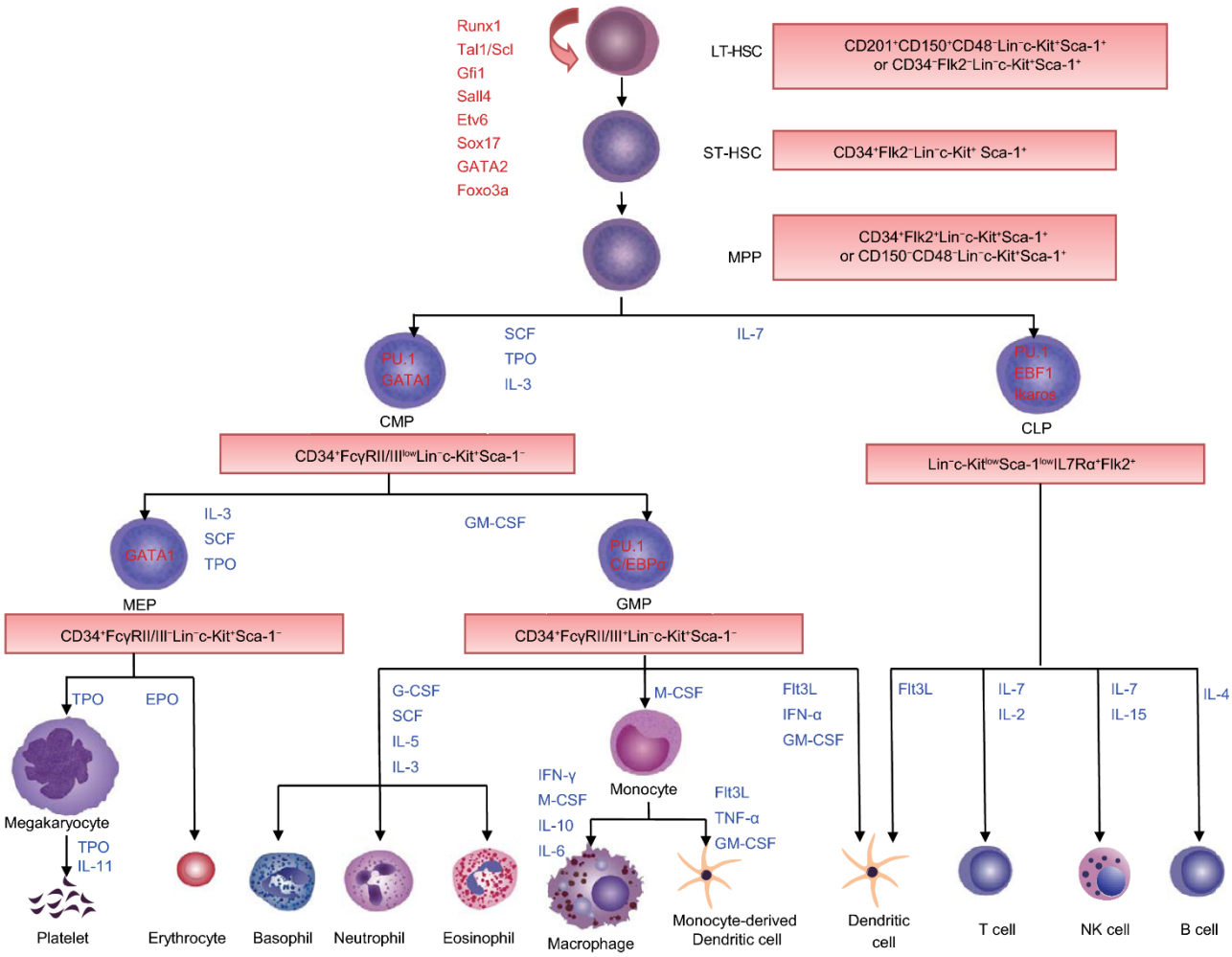


Figure 3: The classical haematopoietic hierarchy. In the classical model, LT-HSCs sit at the top of hierarchy. LT-HSCs differentiate into ST-HSCs, and subsequently to MPPs with reduced self-renewal ability. Downstream of MPPs, a strict separation between the myeloid (CMPs) and lymphoid (CLPs) branches is the first step in lineage commitment. CMPs can generate MEPs and GMPs. CLPs give rise to lymphocytes and dendritic cells. MEPs differentiate into megakaryocytes/platelets and erythrocytes. GMPs produce granulocytes, macrophages, and dendritic cells. Hematopoietic differentiation is controlled by extrinsic cytokines and intrinsic transcription factors. Taken from: (Cheng, Zheng et al. 2020). Copyright permission to reproduce this figure was obtained from Springer Nature.

differentiate into multipotent progenitors. As it stands, these multipotent progenitors do not appear to have a measurable self-renewal ability.

In the classical model, the multipotent progenitors segment into two distinct branches; Common myeloid progenitors (CMPs), which have myeloid, erythroid and megakaryocytic potential and common lymphoid progenitors (CLP) which have only lymphoid potential (Kondo, Weissman et al. 1997, Akashi, Traver et al. 2000, Zhang, Gao et al. 2018). The second segment at the CMP level splits into granulocyte macrophage progenitors (GMPs) and megakaryocyte erythrocyte progenitors (MEP) (Jamieson, Ailles et al. 2004, Zhang, Gao et al. 2018). CMP cells further differentiate and form T cells, B cells, NK Cells and dendritic cells (Zhang, Gao et al. 2018). Furthermore, GMPs further

differentiate into granulocytes and monocytes, whereas MEPs produce megakaryocytes and erythrocytes (Zhang, Gao et al. 2018). This process is carefully regulated by cytokines and key transcription factors as shown in figure 3.

HSCs crucially are able to become all types of differentiation blood cell (Jagannathan-Bogdan and Zon 2013). The myeloid and lymphoid output of HSCs differ, as described by the Müller-Sieburg and Eaves research groups who used limiting dilution analysis and single cell transplantation (Muller-Sieburg, Cho et al. 2004). Müller-Sieburg and colleagues conducted research on lineage biased HSCs. They were able to show that in mice, lineage biased HSCs give rise to all haematopoietic lineages, with varying ratios of myeloid and lymphoid cells (Muller-Sieburg, Cho et al. 2004). In addition, myeloid biased HSCs were shown to take significantly longer to repopulate, in serial transplantation experiments, compared to other types of HSCs. It was also observed that myeloid biased HSCs produced reduced levels of precursor cells for both T-lymphocyte and B-lymphocyte lineages, but produced physiological levels of myeloid precursors (Muller-Sieburg, Cho et al. 2004). They next went on to show using gene array analysis, that the expression levels of interleukin-7 were lower in the lymphoid progeny of the myeloid biased HSCs, suggesting interleukin 7 may play a role in HSC differentiation bias (Muller-Sieburg, Cho et al. 2004). In the study conducted by Brad Dykstra and colleagues, they transplanted 352 mice with mouse bone marrow derived $CD45^{mid}lin^{-}Rho^{-}SP$ cells, or the 4-day clonal progeny of these cells. They observed that the amount of circulating white blood cells in 93 of the 352 mice differed, up to the 6-month period after transplant (Dykstra, Kent et al. 2007). After analysing the clonal outputs at the 16 weeks timepoint, the data showed 4 distinct types of input cells, which they termed α, β, γ, and δ cells. They report that α cells remained as myeloid cells for many months, with little or no contributions to the lymphoid lineage.

These α cells also demonstrated profound self-renewal capability *in vivo*. By contrast the other three repopulating cell subtypes, β, γ, and δ cells were shown to have both myeloid and lymphoid differentiation potential (Dykstra, Kent et al. 2007). The subset of HSCs recently found by the Jacobson and Nakauchi groups tend to differentiate into megakaryocytes and platelets, giving these HSCs a so

called bias to this cell type (Sanjuan-Pla, Macaulay et al. 2013). The Jacobson group report a distinct subset of mouse HSCs, with a bias towards short and long term platelet reconstitution (Sanjuan-Pla, Macaulay et al. 2013). They suggest that thrombopoietin is a key regulator of the biased subsets of HSCs have towards platelets. This discovery shows that HSCs are able to be categorised into a cellular hierarchy, in with the platelet biased HSCs form the peak. It also shows that developmental platelet priming on both a molecular and functional level begins in a distinct population of HSCs (Sanjuan-Pla, Macaulay et al. 2013). Therefore, in order to create therapies to enhance platelet output, it is advantageous to identify this population of platelet primed HSCs.

Using a 5-lineage tracing system, the Nakauchi group identified myeloid biased progenitor cells in mice, in addition to the differentiation pathways which lead to these myeloid biased progenitors cells from HSCs (Yamamoto, Morita et al. 2013). After transplanting single Kusabira-Orange mouse haematopoietic stem and progenitor cells they examined reconstitution kinetics at intervals of 1-4 weeks for up to 20-24 weeks in primary and secondary recipients. The study involved examination of 287 haematopoietic transplants, which enabled the group to identify and classify HSCs into long term HSCs, intermediate term HSCs and short-term HSCs, these were based on reconstitution kinetics in both primary and secondary transplant recipients (Yamamoto, Morita et al. 2013). It was demonstrated that long term HSCs maintain their multilineage reconstitution potential in primary and secondary transplant recipients. By contrast, both intermediate term HSCs and short-term HSCs lose their multilineage reconstitution potential in both primary and secondary transplant recipients (Yamamoto, Morita et al. 2013). As haematopoietic heterogeneity can be enabled by serial transplantation, these studies indicate intrinsic programming, and that haematopoietic heterogeneity occurs randomly.

The Vyas group conducted a study to investigate the relationship between normal human HSPCs and cell populations with leukaemic stem cell (Goardon, Marchi et al. 2011). The group identified the existence of two HSPC immunophenotypic patterns in 80% of the CD34+ acute myeloid leukaemia cells. They demonstrate that both populations possess leukaemic stem cell properties but

are also ordered hierarchically (Goardon, Marchi et al. 2011). They used mRNA expression profiling of these two distinct immunophenotypic patterns, and showed that the leukaemic stem cells which are more mature resemble normal GMPs, whereas the leukaemic stem cells which are immature resemble a progenitor cell, similar to lymphoid primed multipotent progenitors (Goardon, Marchi et al. 2011).

In studies carried out by the Dick group in an investigation of the human umbilical cord blood derived HSC hierarchy, within the CD34+CD38- stem cell compartment, they identified multi lymphoid progenitors as a distinct population of Thy-1(CD90)-neg-lo-CD45RA+ cells (Doulatov, Notta et al. 2010). They observed that these cells gave rise to the complete set of lymphoid cell types. It was also demonstrated that these multi lymphoid progenitors also differentiate into monocytes, macrophages and dendritic cells, and this suggested that such myeloid lineages form early on in the lymphoid lineage specification (Doulatov, Notta et al. 2010). Therefore, the model of myeloid-lymphoid segregation followed in human haematopoiesis is not fixed, as demonstrated in their mouse model.

Interestingly, much research into investigating the HSC hierarchy has been done in non-human primates (Radtke, Adair et al. 2017). In a study conducted by Radtke and colleagues, the group investigated HSC engraftment kinetics, using an *in vivo* pigtailed macaque transplant model. They were able to conduct analysis on over 10^5 HSC clones in 5 five pigtailed macaques for 7.5 years after the initial transplant, in order to investigate the behaviour of the transplanted HSCs. They observed that fraction of CD34+CD45RA-CD90+ HSCs, made up all of the gene modified cells found in the bone marrow and peripheral blood after transplantation in their pigtailed macaques' *in vivo* model (Radtke, Adair et al. 2017). This indicated that contained in this cell phenotype was everything required for short and long term multilineage haematopoietic reconstitution (Radtke, Adair et al. 2017). Crucially, they showed that the transcriptomes, using RNA-Sequencing and the phenotype of the CD34+CD45RA-CD90+ HSCs, were extremely similar between nonhuman primates and humans (Radtke, Adair et al. 2017).

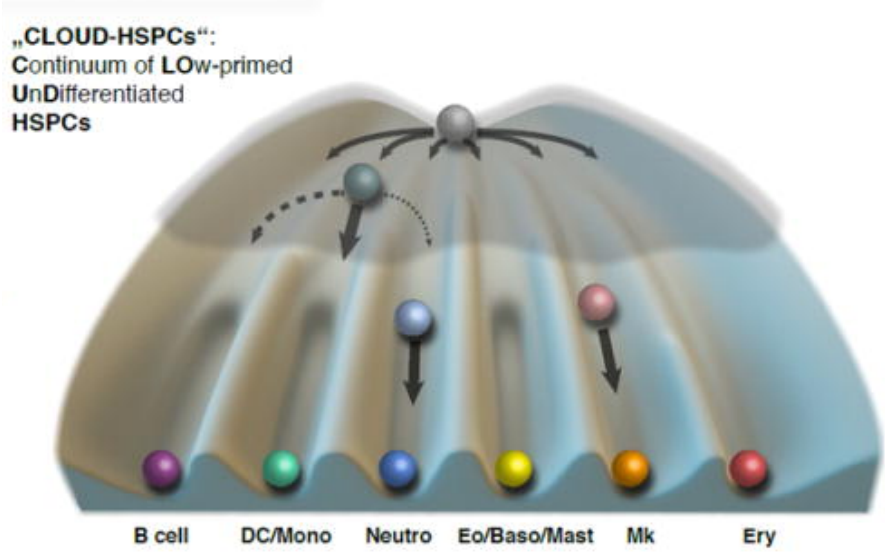


Figure 4: A graphical representation of a continuum-based model of bone marrow haematopoiesis
Graphical summary of a continuum-based model of bone marrow haematopoiesis. Some cell states and transitions are more likely than other due to the interactions of gene regulatory networks, this is represented by a lower elevation within a Waddington landscape. Early during lineage commitment, small barriers between lineages arise early, thereby creating lineage biases in HSCs. At the progenitor stage these barriers are already more pronounced, making the oligopotent stage less likely. T-cell and NK-cell development largely occurs outside the bone marrow. Taken from: (Velten, Haas et al. 2017). Copyright permission to reproduce this figure was obtained from Springer Nature.

Recent studies into HSC hierarchy oppose the classical stem cell lineage tree model (Velten, Haas et al. 2017). Studies conducted by Velten and colleagues dispute this stem cell lineage tree model, claiming that HSCs do not actually pass through discrete intermediate progenitor cells via binary decisions at branching points, instead claiming that a ‘Continuum of LOw-primed UnDifferentiated’ haematopoietic stem and progenitor cells, containing multipotent progenitors and multilymphoid progenitors which are transitory not discrete (see figure 4) (Velten, Haas et al. 2017). For Continuum of LOw-primed UnDifferentiated’ haematopoietic stem and progenitor cells, some transition in cell state and certain lineage combinations are more probable than others as they gradually gain transcriptomic lineage priming in various directions. The distinct lineages which come from the Continuum of LOw-primed UnDifferentiated haematopoietic stem and progenitor cells evolve sooner in the process than was predicted previously and again it is proposed that they do not pass through various discrete, stable progenitors (Velten, Haas et al. 2017). According to their data it seems that there is a multi-dimensional cellular and molecular process of steady state human haematopoiesis, which is identified by a continuous flow of differentiation and development of lineage trajectories, both of which take place independently (Velten, Haas et al. 2017). They propose that the Waddington’s landscape model helps show this in a more accurate way than the classic stem cell tree

model (Waddington 1957). The Waddington's landscape model depicts HSCs at the top and early barriers which separate individual cell lineages early on and increasingly deepen, which demonstrates how lineage biases are acquired; they are founded in the minor differences in gene expression of early fate, as soon as the barriers become impossible to overcome, cell type and commitment to a specific lineage are fixed (Waddington 1957, Velten, Haas et al. 2017).

The classical model gains more complexity in newer models of the HSC system. As observed in figure 5, which is the latest model of haematopoiesis in mice, which consist of MPP heterogeneity. The MPP population was further divided into MPP1, MPP2, MPP3 and MPP4 (Wilson, Laurenti et al. 2008, Pietras, Reynaud et al. 2015). They suggest that MPP1 is similar to intermediate term and short term HSCs. These cells have a multi lineage reconstitution capability of up to 4 months after transplant, by contrast, MPP2, MPP2 and MPP4 cell populations lack a self-renewal capability and only exhibit a short term myeloid reconstitution ability of less than 1 month (Wilson, Laurenti et al. 2008, Pietras, Reynaud et al. 2015). It is also thought that low level of T cells and B cells are produced by MPP2 and MPP3 sub populations, whereas low levels of myeloid cells are produced by MPP4 subpopulations (Wilson, Laurenti et al. 2008, Pietras, Reynaud et al. 2015). More importantly, MPP2 subpopulations produce a higher level of platelets compared to MPP3 and MPP4 subpopulations. The MPP2 subset are thought to be megakaryocyte biased, whereas the MPP3 subset is thought to be myeloid biased (Wilson, Laurenti et al. 2008, Pietras, Reynaud et al. 2015). Although HSCs are able to give rise to all three types of MPPs, the lymphoid biased MPP4 subset are functionally distinct from the MPP2 and MPP3 subset of cells (Wilson, Laurenti et al. 2008, Pietras, Reynaud et al. 2015). These data suggest that MPPs are a diverse subset of cells, with some subtypes exhibiting a distinct lineage biased towards certain cellular compartments.

In the classical murine HSC model, CD34 expression is used to distinguish CMP and MEP subpopulations, with CMPs being CD34+CD16/32- and MEPs being CD34-CD16/32- (Rashidi, Scott et al. 2014). Whether CMP or MEP subpopulations have distinct lineage potential was an interesting question asked by many research groups. Pronk and colleagues postulated the existence of novel

intermediate progenitors. Within the Lin-cKit+Sca1-, Pronk and colleagues defined the CD41+CD150+ cells as megakaryocyte progenitor cells (MkPs), which go on to form megakaryocytes. Within the CD41-CD150-CD16/32- cell population, they describe numerous distinct subpopulations referred to as: pre MegEs, pre GMs, pre CFU-Es and CFU-Es (Pronk, Rossi et al. 2007). Pre MegE cells are able to produce megakaryocyte and erythroid subsets. CMP differentiate into Pre GMs, which subsequently become GMs, by contrast the MEPs are sub divided into Pre CFU-Es, which intern become CFU-Es (Pronk, Rossi et al. 2007). Work conducted by the Amit lab and the Gottgens lab are both in line with

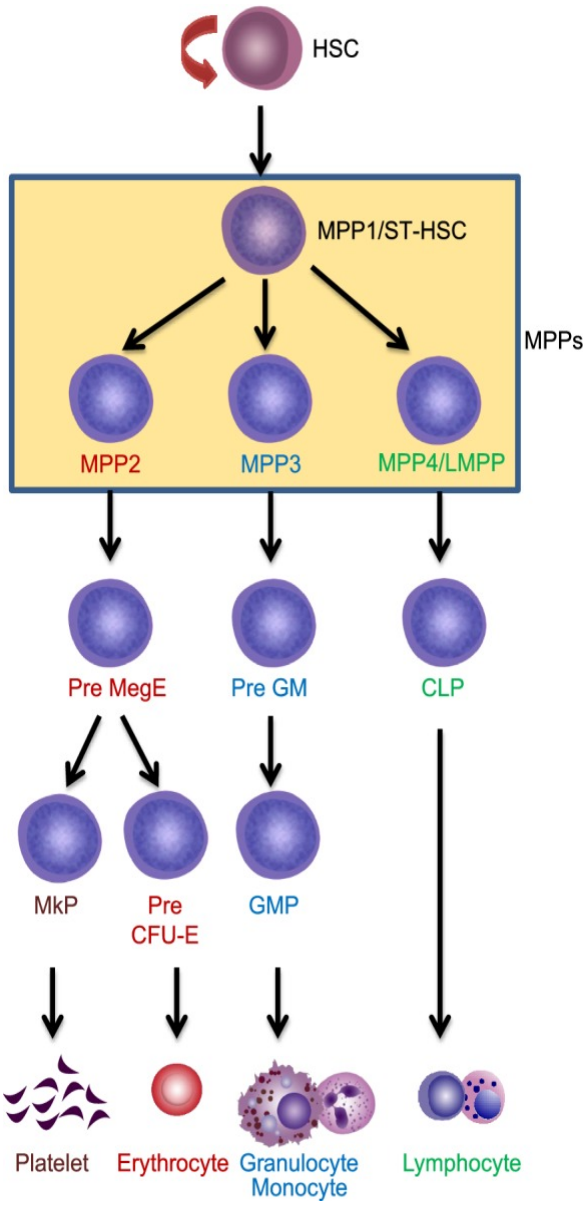


Figure 5: The latest haematopoietic stem cell hierarchy model in mice.
HSCs first differentiate into MPP1/ST-HSC, then give rise to MPP2, MPP3 and MPP4 (LMPP). MPP2 may generate pre MegE, and subsequently, pre MegE gives rise to platelets through MkP or produce erythrocytes through Pre CFU-E. MPP3 mostly give rise to granulocyte and monocyte lineages, and MPP4 (LMPP) mainly contribute to lymphocytes. Taken from: (Cheng, Zheng et al. 2020). Copyright permission to reproduce this figure was obtained from Springer Nature.

the heterogeneity of Lin-cKit+Sca1- progenitors described by Pronk and colleagues (Pronk, Rossi et al. 2007). The Amit laboratory, published a breakthrough paper identifying the transcriptomes of over 2600 mouse single Lin-cKit+Sca1- myeloid erythroid progenitors (Paul, Arkin et al. 2015). Similarly, the Gottgens lab reported on the transcriptomes of 1600 haematopoietic stem and progenitor cells (Nestorowa, Hamey et al. 2016). Taken together with the studies described above, it is clear that with the emergence of new technologies, more exciting descriptions of the HSC tree are being elucidated.

1.1.5: The HSC niche

The concept of a stem cell niche and with that a microenvironment was first coined by Schofield based on the relative difference between blood cells emanating from the spleen or bone marrow (Schofield 1978). In early 1978, Schofield proposed that, in the microenvironment, it is the association of the stem cell with other cells within its niche that determines the behaviour of the stem cell (Schofield 1978). Moreover, he suggested that stem cells respond to carefully orchestrated signals in their local environment in the form of growth factors and cellular signals (Schofield 1978, Scadden 2014). The concept of distinct sites having an impact on HSC differentiation was further clarified by Trentin (Trentin 1971). It was observed that the ratios of differentiating cells were different, for the bone marrow and the spleen, even though, both bone marrow and spleen give rise to several cell lineages (Trentin 1971). It was observed that erythropoiesis was more prevalent for the cells emanating from the spleen and granulopoiesis was more prevalent for cells emanating from the bone marrow. These findings suggest that interactions between HSCs and other cells types are required to propagate and influence differentiation of progenitor cells, in a local environment home to a collection of cells called the HSC microenvironment (Zhao, Lyu et al. 2018).

An example of how the stem cell niche is able to influence HSC biology is in the case of stress. During stress, the sympathetic nervous system is activated and with it a cascade of stress signals are initiated (Batsivari, Haltalli et al. 2020). During stress or after administration of granulocyte colony-stimulating factor (GCSF) or Plerixafor (a CXCR4 antagonist), HSCs leave the bone marrow and are

mobilised into peripheral blood (Batsivari, Haltalli et al. 2020). Furthermore, in a pathological state or during development, the stem cell niches freshly assemble. The stem cell niches are important in maintaining HSC homeostasis through regulating cell number, but the niches also respond to psychological cues such as stress (Batsivari, Haltalli et al. 2020).

Fetal liver cells are thought to be important in the stem cell niche. The amount and type of fetal liver cells which participate in the HSC niche are still not clear. Some studies suggested that DLK-1 may be an important protein in fetal liver cells to regulate haematopoiesis (Huang, Zhao et al. 2019). Although the various haematopoietic niches may consist of different cells and tissue, the key contribution to haematopoiesis homeostasis are shared, as demonstrated through engraftment of HSCs across different developmental barriers (Huang, Zhao et al. 2019). For example, adult bone marrow can receive an engraftment of fetal liver HSCs. The converse is also possible, with bone marrow derived HSCs being able to engraft into fetal liver cells (Huang, Zhao et al. 2019).

Within the bone marrow niche, the HSCs are joined by stromal cells. The extracellular matrix is produced by the stromal cells (Valletta, Thomas et al. 2020). The stromal cells consist of fibroblasts, endothelial cells, adipocytes, osteoblasts and myocytes. Haematopoietic cytokines which can promote, or block differentiation or proliferation of cells, are also produced by the stromal cells and can be enhanced by binding to the extracellular matrix (Gordon, Riley et al. 1987, Valletta, Thomas et al. 2020). Haematopoietic cells can interact with the local environment in several ways, this may be through cell to cell interactions, cell to cytokine interaction and finally cell to extracellular matrix interaction (Pinho and Frenette 2019). The cytokines may either be immobilised or secreted. Moreover, the cell surface receptors which interact with the HSCs are located on the stromal cells themselves or on the extracellular matrix (Pinho and Frenette 2019). These finely tuned interactions carefully regulate the normal haematopoietic process (Pinho and Frenette 2019).

It is well established that osteoblasts are key players in HSC homeostasis (Galan-Diez and Kousteni 2017). It was observed that HSCs reside near the osteoblastic niche and near the vascular niche (Galan-Diez and Kousteni 2017). It is hypothesised that the osteoblastic niche may provide

favourable characteristics for the HSCs to remain in a quiescent state. In contrast it is hypothesised that the vascular niche is important in providing differentiation and mobilisation signals to peripheral blood (Galan-Diez and Kousteni 2017). Key studies were conducted to identify important components of the HSC niche. An animal experimental study has identified the complex, micro-organisation of the various components of the bone marrow stem cell niche (Behrmann, Wellbrock et al. 2018). Since the HSC niche is a dynamic collection of cells and microstructures, *in vivo* imaging studies proved paramount in understand the microenvironment (Pinho and Frenette 2019). It must be said that in the current studies examining the HSC niches, conditioned mice have been invaluable, in that the mice are pre irradiated before transplantation of cells (Pinho and Frenette 2019). The HSCs are identified through quantifying localisation between the niche residual cells and the HSC (Pinho and Frenette 2019). The stem cell niche is important for the fate of HSCs, including the differentiation of HSCs.

1.1.6: Haematopoietic human growth factors

Key signalling pathways and molecules have been demonstrated to be important in *ex vivo* expansion of HSCs. Cytokines are a key class of cell signalling molecule essential in the immune system and inflammation (Khosrowshahi, Khoshfetrat et al. 2016). Recently, using experimental methods, numerous factors have been identified that stimulate expansion of HSCs *ex vivo*. Stem cell factor (SCF) is a key factor expressed in many cell types including stromal and mast cells (Lennartsson and Rönnstrand 2012). The cognate receptor for SCF is c-kit, a receptor tyrosine kinase (RTK). SCF exists in a transmembrane and soluble form as a product of alternative splicing (Lennartsson and Rönnstrand 2012). Research has demonstrated that defects in c-kit lead to a decreased repopulation of HSCs (Dykstra, Kent et al. 2007). A plethora of research studies have also shown that SCF suppresses apoptosis in HSCs (Hassan and Zander 1996). The majority of expansion studies using cytokines include SCF in their cocktail for the culturing of HSCs (Zhang and Lodish 2008, Khosrowshahi, Khoshfetrat et al. 2016). It has also been shown that the membrane bound SCF acts as an adhesion platform in the bone marrow microenvironment for HSCs (Zhang and Lodish 2008). Piacibello and colleagues have

validated that SCF is an important factor in HSC expansion and the regulation of apoptosis (Piacibello, Sanavio et al. 1999). They investigated the expansion potential of SCF in CD34+ cells. Following thawing, UCB CD34+ cells were cultured in the presence of FLT3L (FL) + megakaryocyte growth and development factor (MGDF) +/- SCF. They observed a doubling in fold expansion in the presence of SCF compared to FL+MGDF alone (Piacibello, Sanavio et al. 1999). More recently, He and colleagues also demonstrate that SCF has increased efficacy in the ability to expand HSCs *ex vivo* (He, Scott Swindle et al. 2017).

Another key cytokine used for the expansion of HSCs is thrombopoietin (TPO), which regulates production of platelets and megakaryocytes (Giannini and Peck-Radosavljevic 2015). TPO binds to the TPO receptor and signals via the JAK and STAT pathways to exert its effects. TPO is an important cytokine in platelet biogenesis (Giannini and Peck-Radosavljevic 2015). Recent evidence suggests that megakaryocytes and platelets regulate the amount of TPO that is made available to haematopoietic stem and progenitor cells (Lebois, Dowling et al. 2016). TPO binds to haematopoietic progenitors that express its cognate receptor, Mpl (Lebois, Dowling et al. 2016).

Fox and colleagues confirmed the importance of TPO in HSC expansion and radioprotection (Fox, Priestley et al. 2002). They transplanted normal mice marrow cells into lethally irradiated TPO+/+ or TPO-/- mice to investigate self-renewal and expansion. It was reported that 9 of 10 TPO+/+ mice survived post irradiation and transplant (Fox, Priestley et al. 2002). By contrast, only 2 of 10 TPO-/- mice survived. Next they investigated HSC expansion in TPO-/- mice and TPO+/+ control mice. They performed secondary transplants using donor cells from the primary transplant and observed that HSC expansion was up to 20 times greater in TPO+/+ mice compared with TPO-/-. They also report that this difference could be corrected by administration of physiological amounts of TPO to TPO-/- mice (Fox, Priestley et al. 2002). More recently, Khosrowshahi and colleagues also demonstrated that TPO has efficacy in the expansion of HSCs *ex vivo* (Khosrowshahi, Khoshfetrat et al. 2016). Taken together these findings suggest that TPO has an important role in HSC self-renewal and expansion.

FLT3 ligand is another important cytokine which has been demonstrated by numerous groups to facilitate HSC self-renewal (Zhang and Lodish 2008). McHenna and colleagues investigated the role FLT3L in *ex vivo* expansion of human CD34+ haematopoietic progenitor cells (McKenna, de Vries et al. 1995). Using a colony-forming assay they compared the effect of SCF and FLT3L on haematopoietic progenitor cell (HPC) production. They observed that FLT3L in the presence of interleukin 1-alpha, 3 and 6 promoted formation of granulocytic-monocytic (GM) and high proliferative potential (HPP) colonies. They also observed that FLT3L treatment alone, in isolated bone marrow CD34+ cell cultures, supported the maintenance and production of GM and HPP cell colonies (McKenna, de Vries et al. 1995). They also reported that combination of FLT3L with other factor such as interleukin 1-apha, 3 and 6 acted in synergy to expand progenitor cells *in vitro*.

The addition of FLT3L to a co-culture of irradiated bone marrow cell with CD34+ bone marrow cells enhanced progenitor and total cell production. Irradiated bone marrow cells were established following irradiation to bone marrow stromal cells using a ^{137}Cs irradiator (McKenna, de Vries et al. 1995). These findings show the important role for FLT3L in the regulation of haematopoietic progenitor cells. Anandasabapathy and colleagues conducted a study to investigate the ability of FLT3L to expand HSCs in healthy human volunteers (Anandasabapathy, Breton et al. 2015). They report that recombinant FLT3L (CDX-301) was able to expand monocytes in addition to haematopoietic stem and progenitor cells (Anandasabapathy, Breton et al. 2015). Previous work in our laboratory has also demonstrated that SCF, TPO and FLT3L when used together have a positive effect on haematopoietic stem and progenitor cell fold expansion (Tarunina, Hernandez et al. 2016). The small molecules used for expansion of cells in this thesis will be detailed in Chapter 5.

1.2: Clonal haematopoiesis of indeterminate potential

Studies involving genetic analysis of large populations revealed that somatic mutations in haematopoietic cells leading to clonal expansion (clonal haematopoiesis or CH) are commonly acquired during human aging (Genovese, Kähler et al. 2014, Jaiswal, Fontanillas et al. 2014, Xie, Lu et

al. 2014). As the human body ages, somatic mutations are inevitable (Gondek 2021). The expansion of clones can either be a consequence of increased fitness and/or due to a growth advantage of HSCs (Gondek 2021). It is well known that cancer related gene mutations often result in HSC expansion and clonal dominance. It is thought that Clonal Haematopoiesis of Indeterminate Potential (CHIP) is driven by HSC self-renewal and positive selection, over genetic drift (Gondek 2021). This is due to mutations in driver genes such as *DNMT3A* and *TET2* (Gondek 2021). Animal studies have shown that mutations which result in loss of *DNMT3A* augment HSC self-renewal, leading to expansion of mutated HSCs and clonal dominance (Challen and Goodell 2020). *TET2* is the second most mutated gene in CHIP (Ko, Huang et al. 2010). *In vivo* models exhibiting *TET2* loss demonstrate increased HSC self-renewal and increased myelomonocytic differentiation (Li, Cai et al. 2011). Due to aging, mutations in *TET2* and *ASXL1* are observed in approximately 2% of patients (Gondek 2021).

In addition to aging, CHIP is also caused by cell extrinsic factors (Gondek 2021). Such extrinsic factors include an immune attack observed in acquired aplastic anaemia (AA), in which CH is prevalent in up to 50% of patients. Interestingly 15% of patients with AA exhibit mutations in *DNMT3A* and *ASXL1* (Yoshizato, Dumitriu et al. 2015). Moreover, other extrinsic factors include genotoxic therapies such as chemotherapy, which increases the risk of myeloid neoplasms (Gondek 2021). It is thought that *TP53* and *PPM1D* protect the cell against DNA damage, and during genotoxic therapy *TP53* and *PPM1D* acquire mutations providing an inadequate response to DNA damage in addition to circumventing apoptosis (Jasek, Gondek et al. 2010, Bernard, Nannya et al. 2020). An example is complete loss of *TP53* which results in genetic defects and leukemic transformation. Inflammation is also thought to be a modifier of CH. It is thought that inflammatory cytokines such as IL-6 and IFN-gamma give mutated HSCs a selective growth advantage (Jasek, Gondek et al. 2010, Bernard, Nannya et al. 2020). CHIP can also be caused by an adaptation to intrinsic HSC defects underlying inherited bone marrow failure syndromes (iBMFs) (Gondek 2021). Several disorders may cause iBMFs such as: short telomere syndrome, impaired ribosome biogenesis, an increase in DNA damage and MIRAGE syndrome. In

iBMFs, through compensatory mechanisms, there may be an increase in HSC self-renewal (Gondek 2021).

A study was conducted by Genovese and colleagues in which peripheral-blood cell DNA of a cohort of 12,380 individuals was subject to Whole Exome Sequencing, to investigate the clinical importance of early mutations in cancer progression (Genovese, Kähler et al. 2014). The health data of patients was monitored, from between 2-7 years post DNA testing. Whilst it was found that most mutations were spread across the genome, four genes, namely *TET2*, *ASXL1*, *DNMT3A* and *PPM1D* demonstrated a high proportion of somatic mutations (Genovese, Kähler et al. 2014). *TET2*, *DNMT3A*, and *ASXL1* also demonstrate such mutations in haematological cancers, by contrast, *PPM1D* is implicated in non-haematological cancers, in the regulation of *TP53*. Genovese and colleagues found that the presence of CH was a strong risk indicator for subsequent haematological cancer (hazard ratio, 12.9) (Genovese, Kähler et al. 2014). They reported that approximately 42% of haematological cancers in their study arose in individuals who had clonality present when the DNA sample was collected, which was more than 6 months before cancer diagnosis (Genovese, Kähler et al. 2014).

A similar study was conducted by Jaiswal and colleagues in which they analysed DNA extracted from peripheral blood using Whole Exome Sequencing in 17,182 individuals not selected for an haematological phenotype (Jaiswal, Fontanillas et al. 2014). Somatic mutations were investigated in 160 genes, known to be recurrently mutated in haematological cancers. Haematological phenotypes, cardiovascular events and survival were also analysed in association with the presence of these mutations (Jaiswal, Fontanillas et al. 2014). Jaiswal and colleagues reported that in individuals under the age of 40, somatic mutations were rare, but rose with age (Jaiswal, Fontanillas et al. 2014). Clonal mutations were observed in 9.5% of individuals between the age of 70-79, in 11.7% of individuals between the age of 80-89 and in 18.4% of individuals aged between 90-108 (Jaiswal, Fontanillas et al. 2014). Most variants were found to occur in the *DNMT3A*, *TET2* and *ASXL1* genes. They also reported that the risk of haematological cancer, ischaemic stroke and coronary heart disease was increased, in

association with the presence of somatic mutations in these three genes (Jaiswal, Fontanillas et al. 2014).

In other studies, CH has been shown to be linked to an increased risk of myeloid or lymphoid neoplasia diagnosis (Steensma, Bejar et al. 2015). Although MDS are defined by the presence of cytopenias, dysplasia, and CH, most individuals with CH during aging will not go on to develop MDS (Steensma, Bejar et al. 2015). The acquisition of somatic mutations in the absence of cytopenia and dysplasia was therefore termed by Steensma as "clonal haematopoiesis of indeterminant potential" (CHIP), a benign state which is analogous to monoclonal gammopathy of undetermined significance and monoclonal B-cell lymphocytosis (Steensma, Bejar et al. 2015).

In CHIP, somatic mutations in HSCs and/or HSCPs are reported to contribute to the formation of subpopulations of blood cells which are genetically distinct (Genovese, Kähler et al. 2014). It has been proposed by Jaiswal and colleagues that the subpopulation is clonally derived from a single originator cell, and therefore contains the same mutations as the originator cell (Jaiswal and Libby 2020). As mentioned above, a clonal population of cells may be established when an HSC or HSPC acquires a competitive advantage by way of one or more acquired somatic mutation over cells without these 'acquired' somatic mutations (Bowman, Busque et al. 2018). The size of the clonal population may vary between individuals, but it is evident that the incidence of CH rises drastically with age (Jaiswal and Libby 2020). It has been observed that CH is found in less than 1% of individuals aged less than 50 but found in approximately 10% of individuals over the age of 65 (Genovese, Kähler et al. 2014). Having CH increases an individual's risk of developing blood cancer by 10-fold.

It is evident from a recent review by Hoermann et al. that the risk of progression of CHIP to myeloid neoplasms is only reviewed as modest (Hoermann, Greiner et al. 2020). However, in this review, CHIP represents a large risk factor for cardiovascular disease, such as: coronary artery disease, degenerative aortic valve stenosis and chronic heart failure, in addition to atherosclerosis (Hoermann, Greiner et al. 2020).

In another study conducted by Jaiswal and colleagues in an analysis of 4726 coronary heart disease patients and 3529 control patients, they confirmed increased prevalence of coronary heart disease in patients with CHIP, in which CHIP carriers were 1.9 times more likely to develop coronary heart disease compared to none carriers (Jaiswal, Natarajan et al. 2017). In addition, they report that the risk of myocardial infarction was greater for individuals with CHIP (hazard ratio 4.0) (Jaiswal, Natarajan et al. 2017). It was also observed that hypercholesterolemia-prone low-density lipoprotein-deficient *Ldlr*-/- mice had larger aortic lesions following engraftment of bone marrow from homozygous or heterozygous *TET2* knockout mice vs bone marrow from healthy controls (Jaiswal, Natarajan et al. 2017).

In a further study conducted by Mas-Peoiro and colleagues, they investigated the influence of CHIP on overall survival of patients after transcatheter aortic valve implantation (TAVI) in a cohort of 279 patients. In the first 8 months post-surgery, mutations in *TET2* were associated with increased mortality (hazard ratio 3.1) (Mas-Peiro, Hoffmann et al. 2020). Dorsheimer and colleagues examined 200 patients with chronic heart failure, post successful revascularization following myocardial infarction (Dorsheimer, Assmus et al. 2019). They found that in this cohort, CHIP correlated with significantly worse long-term survival and re-hospitalisations due to heart failure (Dorsheimer, Assmus et al. 2019).

A study was conducted by Bhattacharya and colleagues in which they examined the plasma genome sequence data from 78752 individuals to investigate the prevalence of CHIP (Bhattacharya, Zekavat et al. 2021). They subsequently assessed the association of CHIP with driver genes, *DNMT3A*, *TET2* and *ASXL1* with stroke. They report that CHIP is associated with an increase in the risk of small vessel and haemorrhagic stroke (Bhattacharya, Zekavat et al. 2021).

1.3: Myelodysplastic syndrome

The myelodysplastic syndromes (MDS) are disorders of the HSC and are characterised by ineffective haematopoiesis (Hellstrom-Lindberg, Tobiasson et al. 2020). Key hallmarks of MDS in patients are

anaemia, neutropenia, thrombocytopenia and dysplasia (Ogawa 2020). It is well documented that, per 100,000 individuals, approximately 7 people are affected by the disease, in the UK (Ogawa 2020). The typical age of onset is 70 years, with prognosis varying depending on the mutations and the cell types involved (Veryaskina, Titov et al. 2020). Approximately 30% of MDS patient cases transform to secondary acute myeloid leukaemia (sAML) (Hellstrom-Lindberg, Tobiasson et al. 2020). Patients with MDS most commonly succumb to progression to secondary AML, in addition to bone marrow failure (Ogawa 2020). MDS is diagnosed following a complete blood count and bone marrow test. Diagnosis is based on the occurrence of peripheral blood cytopenia, in addition to cell dysplasia in peripheral blood and the bone marrow (Veryaskina, Titov et al. 2020).

It was in 1982 that MDS was first classified, which originally was divided into subtypes, these were: chronic myelomonocytic leukaemia (CMML), refractory anaemia, refractory anaemia with ring sideroblasts, refractory anaemia with excess of blasts, refractory anaemia with an excess of blasts and refractory anaemia with excess of blasts in transformation (Bennett, Catovsky et al. 1982). During Refractory Anaemia, there is a maturity deficit in bone marrow cells, and they are unable to function properly (Awada, Thapa et al. 2020). Frequently, the cells die before leaving the bone marrow, or in some cases, once they reach the blood stream, they undergo dyserythropoiesis (Awada, Thapa et al. 2020).

The literature suggests that refractory anaemia accounts for 40% of all MDS cases, with diagnosis being more prevalent in the elderly population (Awada, Thapa et al. 2020). The presence of <5% of blasts in the bone marrow, blasts being absent in the peripheral blood, and iron treatment resistant anaemia are all hallmarks of refractory anaemia (Vazquez, Gonzalez et al. 2020). Anaemia is a classical sign of MDS; this is usually accompanied by tiredness and shortness of breath (Vazquez, Gonzalez et al. 2020). Neutropenia is also a sign of MDS, which increases patient susceptibility to infection (Meyer, Guffroy et al. 2020). Thrombocytopenia is also observed in patients with MDS, which increases the probability of bleeding and bruising (Song, Liu et al. 2020). Having said that, many patients with MDS are asymptomatic (Mohammad 2018). In addition to the probability of MDS

developing into AML, approximately 50% of deaths occur due to excessive bleeding or infection (Mohammad 2018). The world health organisation has updated the MDS classification into 7 subtypes: Refractory anaemia (RA), refractory anaemia with ringed sideroblasts (RARS), refractory cytopenia with multilineage dysplasia (RCMD), Refractory cytopenia with multilineage dysplasia and ringed sideroblasts (RCMD-RS), Refractory anaemia with excess blasts (RAEB), Myelodysplastic syndrome, unclassified (MDS-U) and MDS associated with isolated del(5q) (Mohammad 2018).

Anaemia is the primary indicator of refractory anaemia (Patnaik and Tefferi 2019). Although platelet and white blood cell numbers fall in the normal range, in the bone marrow, less than 5% of cells are blasts. It is highly unlikely for refractory anaemia to develop into acute myeloid leukaemia. In refractory anaemia with ringed sideroblasts, patients have anaemia, similar to refractory anaemia, but more than 10% of the red blood cells are sideroblasts (Patnaik and Tefferi 2019, Hall, Meti et al. 2020). In ring sideroblasts, phenotypically the iron in the cells presents itself as a ring in the centre of the cell (Bulaeva, Pellacani et al. 2020). The platelet and white blood cell count are usually normal levels. There is low risk of the disease developing into AML. In refractory cytopenia with multilineage dysplasia patients in the bone marrow, patients have less than 5% immature blood cells and less than 15% ringed sideroblasts (Yee, Papayannidis et al. 2020). There is a possibility that patients with refractory cytopenia with multilineage dysplasia may develop into AML (Yee, Papayannidis et al. 2020). Phenotypically, under the microscope, the other bone marrow cells also look abnormal. Refractory cytopenia with multilineage dysplasia and ringed sideroblasts is similar to refractory anaemia with ringed sideroblasts, where patients have in excess of 15% sideroblasts in addition to anaemia (Mannelli, Bencini et al. 2020). Moreover, additional cells of the bone marrow look irregular under the microscope (Mannelli, Bencini et al. 2020). Additionally, more than one blood cell counts are lower than normal.

Patients with Refractory cytopenia with multilineage dysplasia and ringed sideroblasts may also go on to develop acute myeloid leukaemia. Patients with refractory anaemia with excess blasts, have a reduction in most of their blood cell counts (Mannelli, Bencini et al. 2020). In the blood, there

are less than 5% blasts, by contrast in the bone marrow there are between 5% and 20% blasts. Acute myeloid leukaemia is diagnosed in a patient who possesses more than 20% blasts in the bone marrow (Ma, Lin et al. 2020). It is also observed in refractory anaemia with excess blasts that patients may have a lower white blood cell and platelet count. Approximately 40% of people diagnosed with refractory anaemia with excess blasts transform into acute myeloid leukaemia (Ma, Lin et al. 2020). In myelodysplastic syndrome unclassified, patients usually have a decreased amount of white and red blood cells, along with platelets. Moreover, in patients with myelodysplastic syndrome with del(5q), they suffer from anaemia but have fewer than 5% blast cells (Kirito 2020). The patients also have an abnormality with chromosome 5 (Kirito 2020).

The world health organisation has mixed classifications of myelodysplastic and myeloproliferative disorders, which lie outside of the 7 sub classifications of MDS described above. Examples of this are chronic myelomonocytic leukaemia and juvenile myelomonocytic leukaemia (Rumi, Barate et al. 2020). In both chronic myelomonocytic leukaemia and juvenile myelomonocytic leukaemia, the white blood cell count is much higher than normal (Barbui, Thiele et al. 2018). In these two sub disorders, the monocyte is the cell affected by a mutation. Chronic myelomonocytic leukaemia usually presents itself in people aged over the age of 65 (Barbui, Thiele et al. 2018). Juvenile myelomonocytic leukaemia is more prevalent in younger children under the age of 6 (Batai, Krizsan et al. 2020). Treatment of these disorders is, like MDS, using chemotherapy in the absence of, or in combination with a stem cell transplant (Batai, Krizsan et al. 2020).

The revised international prognostic scoring system (IPSS-R) is used to score prognostic risk in patients with MDS (Greenberg, Tuechler et al. 2012). The following are used to calculate a patient's IPSS-R: blood count at the time of diagnosis, the number of abnormal blasts in the bone marrow, and a bone marrow cytogenetic test at the time of diagnosis (Greenberg, Tuechler et al. 2012). The calculator is then totalled to give a final score, which is used to describe the expected risk of developing AML and expected survival. The risk is grouped into the following categories: very low, low, intermediate, high, and very high (Greenberg, Tuechler et al. 2012). Karyotype is one of the main

components of the IPSS-R scoring system. MDS is known to have several common karyotype abnormalities such as: Del(5q) (characterised by deletions of the long arm of chromosome 5), Monosomy 7 or del(7q) (characterised by abnormalities in chromosome 7), Trisomy 8 (defined as the presence of three copies of chromosome 8), Sex-chromosome abnormalities (–Y, –X), 3q abnormalities (Deletions, translocations and inversions), Trisomy 13 or Del(13q) (characterised by abnormalities in chromosome 13), Trisomy 21 or monosomy 21 (characterised by abnormalities in chromosome 21), Isochromosome 17q (abnormality on the long arm of chromosome 17) (Hale and Keane 2010, Zahid, Malik et al. 2017).

1.4: Common mutations in MDS

The major driver genes of MDS can be categorised into the pathways they affect, and if mutated, can contribute to the MDS phenotype (Ogawa 2019). For example, DNMTA and *TET2* are important regulators of DNA methylation. EZH2 and SUZ12 are important in chromatin regulation, whereas SF3B1 and SRSF2 are implicated in RNA splicing. RUNX1 and ETV6 are important in the role of transcription, whereas FLT3 and KIT are important cytokine receptors. NRAS and KRAS are important since they mediate RAS signalling whereas TP53 and CDKN2A are important in cell cycle checkpoint regulation. Finally, genes involved in DNA repair are also commonly muted in MDS, such as ATM and BRCC3 (Ogawa 2019).

A study conducted by Haferlach and colleagues, revealed the 6 most frequently mutated genes in 944 MDS patients to be, *TET2*, SF3B1, *ASXL1*, SRSF2, DNMT3A and RUNX1 (Haferlach, Nagata et al. 2014). *TET2* deletion in mice, lead to a reduction in 5-methylcytosine (5-mC) and an increase in 5-hydroxymethylcytosine (5-hmC) levels in bone marrow cells. It was shown that *TET2* -/- mice went on to develop myeloid malignancies (Li, Cai et al. 2011). Knock in mice, generated to express the SF3B1-K700E, the most common SF3B1 mutation in MDS, developed MDS like symptoms such as: impaired erythropoiesis and progressive anaemia without ringed sideroblasts (Mupo, Seiler et al. 2017). Moreover, it was shown by Wang and colleagues that mice exhibiting *ASXL1* -/- mutations, went

on to develop features of MDS, including dysplastic neutrophils and multilineage cytopenia (Wang, Li et al. 2014). It was shown by Kon and colleagues, that SRSF2 mutant transplanted mice, went on to develop an MDS like phenotype, demonstrated leukopenia and anaemia (Kon, Yamazaki et al. 2018). It was observed that MDS patients with a DNMT3A mutation had an higher platelet count, in addition, patients with ring sideroblasts also had high incidences of mutations in DNMT3A (Lin, Hou et al. 2018). A mild myeloproliferative phenotype was observed in RUNX1 deficient mice, in addition there was an increase in peripheral blood neutrophils and myeloid progenitor populations (Growney, Shigematsu et al. 2005).

1.5: *TET2*

The *TET2* gene is located on chromosome 4q24, and the *TET2* protein mediates hydroxy methylation of DNA; it achieves this by converting 5-mC to 5-hmC, which promotes DNA methylation (Jiang 2020). The way in which the *TET2* protein achieves this is through oxidation. It is well established that mutations within the *TET2* gene are closely linked to a decrease in levels of 5-hmC, this decrease in levels of 5-hmC are indicators of myeloid malignancies (Jiang 2020). The functional domain of *TET2* is located in the C-terminus of the protein; this consists of a cysteine rich region and a double stranded beta-helix fold domain (Cochran, Geier et al. 2020). The double stranded beta-helix fold domain contains vital metal binding residues for Fe (II)/α-KG. Oxygen, α-KG and Fe (II) are all required by TET2 to function (Cochran, Geier et al. 2020). Molecular oxygen is used as a substate in the oxidative decarboxylation of α-KG, generating Fe (IV)-oxo, and results in the conversion of 5-mC to 5-hmC (Cochran, Geier et al. 2020).

Since *TET2* was discovered, a number of studies have been focusing on the enzymatic effects of the protein, through its role in the methylation of DNA (Stremenova Spegarova, Lawless et al. 2020). Proteins which interact with *TET2* have been identified by tagging *TET2* with a purification tag, and using affinity purification and mass spectrometry, numerous binding partners have been identified, these include: transcription factors, micro RNAs and small molecules (Pan, Weeks et al. 2015). It is

well documented that O-linked β-N-acetylglucosamine transferase (OGT) is recruited to chromatin, to increase H3k4me3 levels, to regulate gene transcription (Pan, Weeks et al. 2015). Moreover, multiple studies have been conducted which demonstrate that *TET2*, at pre-mRNA level is negatively regulated by key microRNAs, these are: miR-7, miR-22, miR-26, miR-29, miR-101, and miR-125 (Cheng, Guo et al. 2013). It is also well established that vitamin C is able to increase the activity of *TET2* to suppress leukaemia (Li, Zhang et al. 2019). In addition, proteins such as WT1, VprBP, EBF1, IDAX and UHRF2 have been identified to co-localise with *TET2* to exert their biological function (Bowman and Levine 2017). It was also recently published, by (Yang, Sun et al. 2017) that sirtuin 1 (SIRT1) deacetylates *TET2* in MDS stem and progenitor cells (Yang, Sun et al. 2017). Furthermore, they demonstrated that SIRT1-deficient MDS CD34+ HSPCs showed enhanced cell growth and self-renewal (Yang, Sun et al. 2017). *TET2* is important in haematopoietic regulation of self-renewal, lineage commitment and terminal differentiation (Huang, Sun et al. 2020).

TET2 is expressed highly in HSCs (Huang, Sun et al. 2020). Loss of *TET2* has been shown to increase the percentage of progenitor cells (Figueroa, Abdel-Wahab et al. 2010), which suggest that loss of the *TET2* protein may regulate stem cell differentiation (Figueroa, Abdel-Wahab et al. 2010). It has also been shown that loss of *TET2* in CD34+ cells can also promote the expansion of monocytes, this finding suggests that *TET2* may play a role in differentiation (Cakouros, Hemming et al. 2019). *TET2* has also been implicated as being a tumour suppressor. Myeloid malignancies have been observed in mice exhibiting loss of the *TET2* gene (Cakouros, Hemming et al. 2019). Mutations, both homozygous and heterozygous, are frequent in myeloid malignancy (Ahn, Kim et al. 2015). The frequency of *TET2* mutations are ~25% in myelodysplastic syndrome (MDS), in chronic myelomonocytic leukaemia (CMML) ~45% and in acute myeloid leukaemia ~23%.

Mutations in *TET2* have been reported to cause a frame shift, generate stop codons, lead to in frame deletions and also cause amino acid substitutions (Ahn, Kim et al. 2015). The disease causing patterns in *TET2* mutations are attributed to the genotype, be it homozygous or heterozygous (Ahn, Kim et al. 2015). Furthermore, it was observed by (Solary, Bernard et al. 2014) that CMML had the

highest rate of *TET2* mutation alterations, reportedly 55%, which accounted for both homozygous and heterozygous mutations (Solary, Bernard et al. 2014). As revealed by genome sequencing of all coding exons of *TET2*, the homozygous mutations are associated with uniparental disomy (Jankowska, Szpurka et al. 2009), whereas the heterozygous mutations are associated with the lack of chromosomal lesions (Jankowska, Szpurka et al. 2009).

It has been shown that mice which exhibited a *TET2* heterozygous knockout, and showed an approximate 50% loss of gene expression of *TET2*, resulting in a less recurrent transformation when compared to a *TET2* homozygous knockout (Tripodi, Hoffman et al. 2010). Notably, in the *TET2* homozygous mutation, the level of 5-hmC level in DNA was far less than in the heterozygous knockout mice (Li, Cai et al. 2011). Within the first year of life, ~33% of mice with a homozygous *TET2* mutation went on to develop a lethal myeloid malignancy (Li, Cai et al. 2011). By comparison, ~8% of mice exhibiting a heterozygous *TET2* genotype developed a lethal myeloid malignancy in the first year of life (Li, Cai et al. 2011). Although in patients varying levels of *TET2* mutations were observed, the clinical presentations and pathological phenotype were the same. In addition, Kaplan Meier survival curves demonstrate that there are also no difference in overall survival between the homozygous and heterozygous *TET2* mutation cases (Jankowska, Szpurka et al. 2009). Moreover, patients homozygous for *TET2* showed a significantly lower event free survival and exhibited a higher relapse rate when compared to patients with heterozygous *TET2* mutations (Ahn, Kim et al. 2015). In mouse models, the *TET2* homozygous genotype demonstrates an increased susceptibility to malignancies, however in human, these distinctions are less clear (Ahn, Kim et al. 2015).

1.6: *ASXL1*

Additional sex combs-like 1 (*ASXL1*) is an epigenetic regulator and is commonly mutated in myeloid neoplasms, more frequently in myelodysplastic syndromes, myeloproliferative neoplasms, chronic myelogenous leukaemia and acute myeloid leukaemia (Micol and Abdel-Wahab 2016). Furthermore, mutations in *ASXL1* are associated with a poor prognosis (Micol and Abdel-Wahab 2016). Mutations

in *ASXL1* are also frequently found in clonal haematopoiesis of indeterminate potential, a precursor state for haematological neoplasms (Steensma, Bejar et al. 2015).

In order to understand the function of *ASXL1*, numerous *ASXL1* knockout studies have been studied in mice (Wang, Li et al. 2014). It has been demonstrated that a knockout of *ASXL1* leads to the development of myelodysplastic disease and myeloproliferative neoplasms (Abdel-Wahab, Gao et al. 2013). Moreover, *ASXL1* depletion resulting in aberrant haematopoiesis due to loss of PRC2 interaction. It was also demonstrated that loss of *ASXL1* in mice collaborated with NRASG12D, to enable myeloid leukaemogenesis (Abdel-Wahab, Adli et al. 2012).

Most mutations in the *ASXL1* gene are located on the last exon (exon 12), producing a protein truncated at the c-terminus, since it escapes nonsense mediated decay (Schnittger, Eder et al. 2013). A study was conducted which showed that overexpression of *ASXL1* in mouse transplant models was able to hinder myeloid differentiation and promote myelodysplastic syndrome (Inoue, Kitaura et al. 2013). There is increasing indication that mutated *ASXL1* is able to disrupt haematopoiesis and also promote myeloid transformation (Nagase, Inoue et al. 2018). It does this by affecting histone modifications leading to a gain of function (Nagase, Inoue et al. 2018). It was also shown by Valletta and colleagues that mice xenografted with *ASXL1* mutation corrected KBM5 cells, showed an increased survival compared to mice with uncorrected KBM5 xenografts (Valletta, Dolatshad et al. 2015). In addition, the *ASXL1* mutation corrected KBM5 cells showed a reduction in cell growth, compared to *ASXL1* uncorrected KBM5 cells (Valletta, Dolatshad et al. 2015).

Mutations of *ASXL1* are found in myeloid malignancies such as: myelodysplastic syndromes, acute myeloid leukaemia, and myeloproliferative neoplasms such as: chronic myelogenous leukaemia and chronic neutrophilic neoplasms (Lin, Zheng et al. 2016). Mutations of *ASXL1* are also observed in patients with chronic myelomonocytic leukaemia, in addition to juvenile myelomonocytic leukaemia (Perez, Kosmider et al. 2010). Furthermore, *ASXL1* mutations are observed in aplastic anemia, which commonly causes bone marrow failure (Lin, Wang et al. 2020).

Mutations in *ASXL1* are found in ~8% of patients with acute myeloid leukaemia, with mutations in *ASXL1* being more common in older patients (Wu, Song et al. 2016). *ASXL1* mutations which occur in patients with acute myeloid leukaemia, commonly co-exist with mutations in the genes for RUNX1 and IDH2 (Bera, Chiu et al. 2019). It is found that of patients with myelodysplastic syndrome, ~16% have mutations in *ASXL1*, these mutations in *ASXL1* are also associated with a poor outcome in patients with myelodysplastic syndromes (Lin, Zheng et al. 2016). The prevalence of *ASXL1* mutations are also apparent in patients who are classed as high-risk myelodysplastic syndrome (Lin, Zheng et al. 2016). In order to treat patients with high risk myelodysplastic syndrome, azacytidine or decitabine, which are DNA hypomethylating agents, are used (Papageorgiou, Kontos et al. 2018). Interestingly, it was recently shown that *TET2* mutations improve the response to these DNA hypomethylating agents (Bejar, Lord et al. 2014). Another study showed that patients who exhibit these mutations in *ASXL1* and are treated with DNA hypomethylating agents showed a shorter survival rate, compared to them not being used (Traina, Visconte et al. 2014).

1.7: The CRISPR Cas9 System

To determine the function of a gene, the gene can be interrupted by homologous recombination (Gerlai 2016) or by interfering with the messenger RNA of the gene (Agrawal, Dasaradhi et al. 2003). These can be achieved by conducting transfection in cells. Recently pioneering research has demonstrated the manipulation of genes in a broad spectrum of tissues and species (Principe, Narbutis et al. 2020). Genome editing technologies are proving invaluable in understanding disease (Iacomussi 2019). The ability to edit genomes is made possible due to specific nucleases, these highly efficient nucleases function by administering a double stranded break in the region of interest in the gene, which is later repaired by cellular mechanisms (Nishida and Kondo 2020). The repair mechanisms the cell employs to repair these double stranded breaks is by either non homologous end joining, or homology directed repair (Nishida and Kondo 2020). This may lead to INDELS (insertion(s)

or deletion(s)), in which these mutations may cause a disruption to the gene leading to aberrant function of the protein in the cell (Derry 2020).

The CRISPR Cas9 system consists of a guide RNA which serves to direct the Cas9 nuclease to the target site for the double stranded breaks, in the genome (Jinek, Chylinski et al. 2012) (see figure 6). The most common system used for genome editing is the type II CRISPR Cas9 system, which deploys the Cas9 protein, first discovered in Streptococcus pyogenes (Ledford 2020). In the normal endogenous system, crRNA co localised with tracrRNA to form a complex, which then in turn guides the Cas9 protein to the target site for double stranded cleavage (Scott, Urak et al. 2019). This specifically occurs, when the crRNA interacts with the target site for cleavage (Scott, Urak et al. 2019). There are two nuclease domains located in the multi-domain nuclease lobe of Cas9, the RuvC which cleaves the non-target DNA strand, and the HNH nuclease domain that cleaves the target strand of DNA. (Scott, Urak et al. 2019). Jinek and colleagues designed a chimerical guide RNA, which combines all of the essential Cas9 machinery, including the crRNA the tracrRNA and the Cas9 protein (Jinek, Chylinski et al. 2012). Since the development of this chimeric protein, a number of versions of the CRISPR Cas9 system have been invented, all of which recognise a nucleotide sequence of between 18 and 24 bases, in addition to a up to 4 nucleotides of the (PAM) protospacer adjacent motif site in the target DNA site (Kim, Hupperetz et al. 2020). Genome editing has been used to generate *in vivo* disease

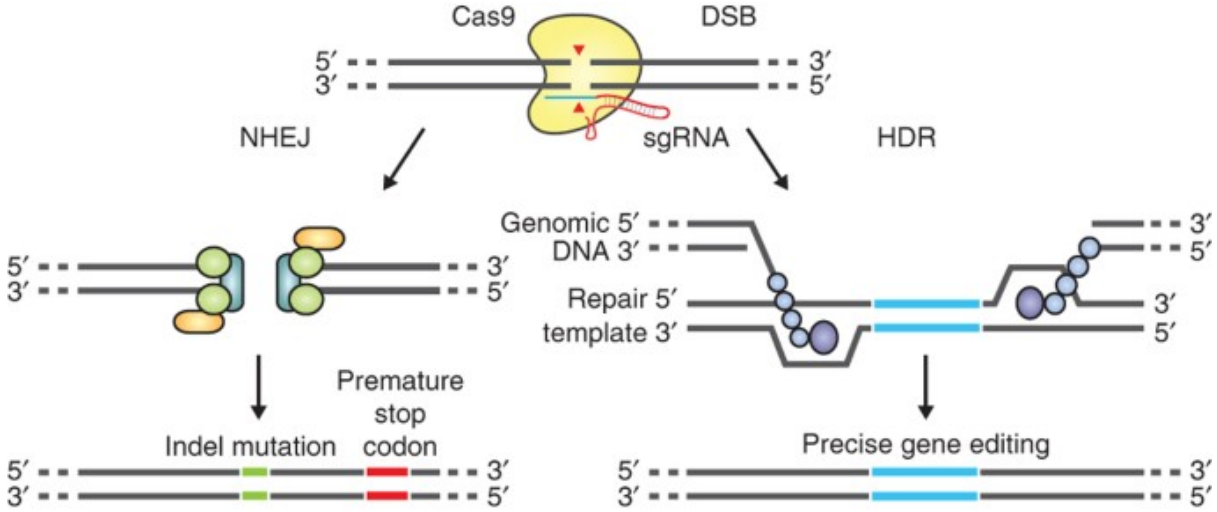


Figure 6: There are two ways in which double stranded breaks, which have been induced by Cas9 may be repaired. Firstly, the ends of a double stranded break can be processed and re-joined in the NHEJ pathway, resulting in random INDEL mutations. Alternatively, a plasmid or ssODN template can be used in conjunction with the HDR pathway, allowing highly specific gene editing. Taken from: (Ran, Hsu et al. 2013). Copyright permission to reproduce this figure was obtained from Springer Nature.

models and to study key genes in myeloid malignancy. Moreover, Torres and colleagues, were able to generate a robust cancer model in human cell lines and chromosomal translocation baring primary cells, resembling those in AML and Ewing's sarcoma (Torres, Martin et al. 2014). Chang and colleagues

used CRISPR/Cas9 in patient derived induced pluripotent stem cells (iPSCs) to investigate the contribution of the SRSF2 P95L mutation and the chromosome 7q deletion, in the development of myeloid malignancy (Chang, Kotini et al. 2018). In a study conducted by Cho and colleagues the CRISPR/Cas9 system was also used to disrupt the C4BPB gene in the K662 cell line (Cho, Kim et al. 2013).

Besides gene editing, CRISPR Cas9 technology has also been developed to regulate gene expression using dCas9, a mutated form of Cas9 which has no nuclease activity. The technique is called CRISPR interference, or CRISPRi (Larson, Gilbert et al. 2013). It was shown in E.coli that when coupled with an sgRNA of 20bp, dCas9 is able to repress any gene of interest with up to 99.9% repression (Qi, Larson et al. 2013). By contrast the activation of target genes has also been achieved using CRISPR activation or CRISPRa (Cheng, Wang et al. 2013). Considering, we saw the Nobel prize in chemistry awarded to Emmanuelle Charpentier and Jennifer A. Doudna for their ground-breaking work on the development of the CRIPSR Cas9 system, we should anticipate more expansive discoveries and uses for the CRIPSR Cas9 system (Uyhazi and Bennett 2020).

1.8: The aims and objective of this thesis

The objective of this thesis is to investigate the role of *TET2* and *ASXL1* in myeloid malignancy. There were three central aims:

1. To introduce mutations of *TET2* and *ASXL1*, both individually and in combination, in human cell lines using CRISPR/Cas9 technology.
2. To test the consequences of these cell mutations on key cellular function.
3. To optimise culture conditions for the expansion of CD133+ haematopoietic stem and progenitor cells as a prelude to gene editing of *TET2* and *ASXL1*.

Chapter 2
Materials and Methods

2.1: Reagents/consumables and suppliers

Reagent	Supplier
1 kb DNA Ladder	Sigma-Aldrich
1.5ml **Microcentrifuge tube(s)**	Thermo Fisher Scientific
1.7ml Microcentrifuge tube(s)	Sigma-Aldrich
10X NEBuffer 2.1	New England Biolabs
10X RT Buffer	Thermo Fisher Scientific
10X RT Random Primers	Thermo Fisher Scientific
12, 24 and 96 well culture plates	Sarstedt
15ml conical tube	Thermo Fisher Scientific
25X dNTP Mix (100mM)	Thermo Fisher Scientific
40 µm Nylon cell strainer	Sigma-Aldrich
50 ml Accuspin tubes	Sigma-Aldrich
6 well culture plate	Greiner
AGAP2 Primers	Invitrogen
Agar	Thermo Fisher Scientific
Agarose	Sigma-Aldrich
Agilent RNA 6000 Pico kit	Agilent
Amersham™ ECL™ Rainbow™ Marker - Full range	Sigma-Aldrich
Annexin V	Abcam
Anti-Beta Actin	Sigma-Aldrich

Anti-*TET2*	Sigma-Aldrich
AQP10 Primers	Invitrogen
ARHGAP11A Primers	Invitrogen
ASXL1 Primers	Invitrogen
ASXL1 qRT-PCR Probe	Applied biosystems
B2M qRT-PCR Probe	Applied biosystems
BbsI (NEB)	Thermo Fisher Scientific
BCA assay kit	Thermo Fisher Scientific
CAPN3 Primers	Invitrogen
CD11b-BV421	Biolegend
CD123- PE-Cy5	Biolegend
CD133 APC	Miltenyi
CD133 MicroBead Kit, human	Miltenyi Biotec
CD133/2-PE (293C3)	Miltenyi Biotec
CD14-APC-H7	BD
CD15-APC	Abcam
CD235a-APC	Invitrogen
CD34-AF700	BD
CD36-PE	Boilegend
CD38- PE-TxR	Caltag
CD45RA-APC-H7	BD
CD49f PE-Cy7	Biolegend
CD71-FITC	Invitrogen
CD90- PE	Biolegend
cDNA synthesis kit	Thermo Fisher Scientific

Cell Line NucleofectorTM Kit V	Lonza
Clear Polyolefin Star Seal	Starlab
CountBright Beads	Thermo Fisher Scientific
Cryopreservation vials	Thermo Fisher Scientific
DAPI	Sigma-Aldrich
DEPC-Treated Water	Invitrogen
Dexamethasone	Sigma-Aldrich
Dimethyl Sulphoxide	Sigma-Aldrich
DMSO	Sigma-Aldrich
DNase treatment	Thermo Fisher Scientific
Doxycycline	Sigma-Aldrich
DreamTaq	Thermo Fisher Scientific
Dulbecco's Phosphate-Buffered Saline	Thermo Fisher Scientific
DYRK1B Primers	Invitrogen
Erythropoietin	Thermo Fisher Scientific
ExoSAP-IT Express	Thermo Fisher Scientific
Expansion Supplement	Stem Cell Technologies
FAM53A Primers	Invitrogen
FcR Blocking Reagent, human	Miltenyi Biotec
Fetal Bovine Serum	Thermo Fisher Scientific
FLT-3L	R&D Systems
GRK4 Primers	Invitrogen
High-Capacity cDNA Reverse Transcription Kit	Thermo Fisher Scientific
HK2 Primers	Invitrogen
HOXD12 Primers	Invitrogen

Human Serum Albumin	Bio Products Laboratory Limited
Industrial methylated spirit	Thermo Fisher Scientific
Iscove's Modified Dulbecco's Medium	Thermo Fisher Scientific
LB broth	Sigma-Aldrich
LightCycler **TaqMan® Master Mix**	Roche
L**ymphocyte separation medium**	Lonza
MACS Buffer	Miltenyi Biotec
MFI2 Primers	Invitrogen
mIgG2b-PE isotype control	R&D Systems
MiniMACS Separator and Starting Kit	Miltenyi Biotec
Mr. **Frosty**	Sigma-Aldrich
MultiScribe Reverse Transcriptase	Thermo Fisher Scientific
MYLK Primers	Invitrogen
NAM	Sigma-Aldrich
Nitrocellulose/Filter Paper Sandwich	Thermo Fisher Scientific
NucleofectorTM Kit V	Lonza
NuPAGE™ 4 to 12%, Bis-Tris, 1.0 mm, Mini Protein Gel	Thermo Fisher Scientific
NuPAGE™ LDS Sample Buffer (4X)	Thermo Fisher Scientific
Oligonucleotides	Invitrogen
Orange DNA Loading Dye (6x)	Thermo Fisher Scientific
PARP15 Primers	Invitrogen
PEITC	Sigma-Aldrich
Penicillin-Streptomycin-Glutamine	Thermo Fisher Scientific
PGE2	Sigma-Aldrich

PGR Primers	Invitrogen
PIN4 Primers	Invitrogen
Plasmid digest solution	Sigma-Aldrich
PPP1R13L Primers	Invitrogen
PRKCE Primers	Invitrogen
Propidium iodide	Sigma-Aldrich
Protease and phosphatase inhibitor	Sigma-Aldrich
PRPH2 Primers	Invitrogen
pSpCas9(BB)-2A-GFP (PX458)	Addgene
PureLink™ Genomic DNA Mini Kit	Thermo Fisher Scientific
PureLink™ Genomic Lysis/Binding Buffer	Thermo Fisher Scientific
PX458 Plasmid	Addgene
QIAGEN Plasmid Plus Midi Kit	Qiagen
Qiagen QIAprep Spin Miniprep Kit	Qiagen
qRT-PCR culture plate	Thermo Fisher Scientific
qRT-PCR Kit	Roche
QuickExtract DNA Solution	Lucigen
Resazurin	Sigma-Aldrich
RIPA buffer	Sigma-Aldrich
RNAClean XP kit	Beckman Coulter
RNase Inhibitor	Thermo Fisher Scientific
RNaseZap™ RNase Decontamination Solution	Invitrogen
Roche MM (2X)	Roche
RS1	Stem Cell Technologies
S.O.C Media	Thermo Fisher Scientific

SCF	R&D Systems
Scriptaid	Sigma-Aldrich
SLC22A10 Primers	Invitrogen
SNX24 Primers	Invitrogen
SR1	Stem Cell Technologies
Stem Cell Factor	Stem Cell Technologies
StemSpan SFEM	Stem Cell Technologies
SYBR Safe DNA Gel Stain	Thermo Fisher Scientific
SYBR Safe DNA Gel Stain	Thermo Fisher Scientific
T4 DNA ligase	New England Biolabs
T4 DNA ligase buffer	New England Biolabs
Taqman Assay Probe	Roche
TBE buffer	Thermo Fisher Scientific
TET2 Primers	Invitrogen
TET2 qRT-PCR Probe	Applied biosystems
TMEM27 Primers	Invitrogen
TPO	R&D Systems
Tris/Borate/EDTA (TBE)	Thermo Fisher Scientific
Trypan Blue solution	Sigma-Aldrich
Tween 20	Thermo Fisher Scientific
UM171	Stem Cell Technologies
VPA	Sigma-Aldrich
Water	Ambion
ZLLYFMK	Abcam

Table 1: Reagents/consumables and supplier name.

2.2: Cells

2.2.1: Cell lines

2.2.1.1: HUDEP2 cells (plating and cell maintenance)

The HUDEP2 (Human Umbilical Cord Derived Erythroid Progenitor 2) cells are a human erythroid progenitor cell line, derived from umbilical cord CD34+ cells (Kurita 2013). HUDEP2 cells are suspension cells. The HUDEP2 cells were cultured in StemSpan SFEM medium (Stem Cell Technologies) with 0.005% (v/v) Stem Cell Factor (Stem Cell Technologies), 0.03% (v/v) Erythropoietin, 0.01% (v/v) Dexamethasone, 1% (v/v) Penicillin-Streptomycin-Glutamine and 0.04% (v/v) Doxycycline. All cell culture plates, (6-well Cat. No: 657185), (12-well Cat No: 83.3921500), (24-well, Cat No: 83.3922.500) and (96-well, Cat No: 83.3924.500), used in this thesis were purchased from Sarstedt Ltd. unless otherwise stated. HUDEP2 cells were cultured according to methods of Vinjamur (Vinjamur and Bauer 2018), with some modifications. The Biomat Class 2 microbial safety cabinet was subjected to ultraviolet light sterilisation for thirty minutes before each usage. Once powered on, the airflow was equilibrated for a further ten minutes. Once considered safe by the on-board LCD, the Biomat Class 2 microbial safety cabinet was sterilised using a 70% industrial methylated spirit (IMS). HUDEP2 cell media (see table 2) was pre-warmed to 37 degrees Celsius in a Grant SUB14 water bath. A cryopreservation vial containing 10^6 HUDEP2 cells was placed into the water bath (37 degrees Celsius) for 30 seconds to begin the thawing process, once thawed the cells were transferred to prewarmed HUDEP2 cell media and subsequently transferred to an appropriate culture dish. Cells were cultured in a Thermo Fisher Incubator at 37 degrees Celsius and 5% CO2 in air. HUDEP2 cell cultures were fed three times per week (½ media change). Once confluency was reached, a 1:10 split was conducted using freshly prepared pre-warmed (37 degrees Celsius) HUDEP2 cell media.

Reagent	Cat. No:	Concentration	Volume
StemSpan SFEM	09650	Stock	45ml
SCF	78062	100µl/ml	2.25µl
Erythropoietin	SRP6158	85µg/ml	13.5µl
Dexamethasone	D4902	10mg/ml	4.5µl
Penicillin-Streptomycin-Glutamine	10378016	50mg/ml	450µl
Doxycycline	10592-13-9	10mg/ml	18µl

Table 2: HUDEP2 cell media components, catalogue numbers, concentration and volumes used.

2.2.1.2: KBM5 cells (plating and cell maintenance)

KBM5 cells are a human chronic myeloid leukaemia cell line derived from cells obtained from a 67-year-old female (Beran, Pisa et al. 1993). These cells are suspension in nature. The KBM5 cells were cultured in Iscove's Modified Dulbecco's Medium (Thermo Fisher Scientific) with 10% (v/v) Fetal bovine serum and 1% (v/v) Penicillin-Streptomycin-Glutamine. KBM5 cells were cultured according to methods of Valletta and colleagues (Valletta, Dolatshad et al. 2015), with some modifications. KBM5 cell media (see table 3) was pre-warmed to 37 degrees Celsius in a Grant SUB14 water bath. Cells were cultured in a Themo Fisher Incubator at 37 degrees Celsius and 5% CO2. A cryopreservation vial containing 10^6 KBM5 cells was placed into the water bath (37 degrees Celsius) for 30 seconds to begin the thawing process, once thawed the cells were transferred to prewarmed KBM5 cell media and subsequently transferred to an appropriate culture dish. KBM5 cell cultures were fed three times per week. Cells were maintained with a ½ media change every other day. Once confluency was reached, a 1:10 split was conducted using freshly prepared prewarmed (37 degrees Celsius) KBM5 cell media.

Reagent	Cat. No:	Concentration	Volume
Iscove's Modified Dulbecco's Medium	12440053	Stock	50ml
Fetal Bovine Serum	16140071	Stock	5ml
Penicillin-Streptomycin-Glutamine	10378016	50mg/ml	500µl

Table 3: KBM5 cell media components, catalogue numbers, concentrations and volumes used.

2.2.2. Primary cells

2.2.2.1. Collection and isolation of human CD133+ cells

All CD133+ stem and progenitor cells used for these experiments were isolated using MACS® Separation. Cord blood units were diluted 1:1 if the white blood cell (WBC) count was $<7x10^6$ cells/ml and diluted 1:2 if the WBC count was $>7x10^6$ cells/ml in MNC wash buffer (DPBS) (Cat. No. 14190144) at room temperature and mixed by inversion. 30ml of diluted blood was layered onto Accuspin tubes (Cat. No. A2055-10EA) containing 17ml lymphocyte separation medium (LSM) (Cat. No. 17-829E). The Accuspin tubes were centrifuged at 800g with an acceleration setting of 9 and a deceleration setting of 0 for 35 minutes at 20 degrees Celsius in a Hettich Rotina 46R centrifuge. The upper layer was then aspirated leaving the mononuclear cell (MNC) layer undisturbed at the interphase.

Two MNC layers were transferred to a new 50ml tube and topped up to 50ml with MNC wash buffer. Subsequently, the suspension was centrifuged at 300g for 10 minutes with an acceleration setting of 9 and a deceleration setting of 9 at 4 degrees Celsius, followed by aspiration of the supernatant. 10ml MNC wash buffer was added to the first pellet of cells and was then transferred to the next tube. This process was continued to merge cells from two or three tubes into a single tube. The resulting tubes were topped up with MNC wash buffer to a final volume of 50ml. The tubes were then centrifuged at 200g for 15 minutes, at 4 degrees Celsius with an acceleration setting of 9, and a deceleration setting of 9. The supernatant was then aspirated. The pellets were pooled so that all the

cells were collected into a single tube. The tubes were topped up to 50ml with MNC wash buffer and the suspension was homogenised.

500µl sample of suspension was transferred into an Eppendorf tube and a counted on the Sysmex haematology analyser. The cell suspension was then centrifuged at 200g for 15 minutes at 4 degrees Celsius with an acceleration setting of 9, and a deceleration setting of 9. The supernatant was then aspirated. The pellet was next resuspended in 20ml MACS buffer pre chilled to 4 degrees Celsius, subsequently the cell suspension was filtered through a 40µm nylon cell strainer (Cat. No. CLS431751). The cell suspension was centrifuged at 200g for 10 minutes at 4 degrees Celsius with an acceleration setting of 9, and a deceleration setting of 9 and the supernatant was carefully aspirated.

Next CD133+ cells were isolated using MACS isolation beads. Cells were resuspended in a solution containing: 300µl MACS buffer (Cat. No. 130-091-221), 100µl FcR blocking reagent, human (Cat. No. 130-059-901) and 100µl CD133 microbeads (Cat. No. 130-097-049) per 10^8 total cells. The suspension was homogenised and incubated at 4 degrees Celsius for 30 minutes. Cells were washed with 20ml 4 degrees Celsius MACS buffer and centrifuged at 200g for 10 minutes. The cells were then resuspended in 5ml MACS buffer at 4 degrees Celsius. The separation was performed according to manufacturer instructions using the MiniMAC Separator and Starting Kit (Cat. No. 130-090-312).

An LS column was placed onto a MidiMACS separation magnet and was rinsed using a 3ml MACS buffer. The cell suspension was subsequently passed through a 30µm pre-separation filter and then through the LS attached to the MiniMACs separation magnet. The column was then repeatedly washed using a 4ml ice-cold MACS buffer, three times. The labelled cells were then flushed out using a 5ml ice-cold MACS buffer with a plunger (supplied with the kit). The washing procedure was repeated to increase purity. 10µl cell suspension was diluted 1:1 with 0.4% trypan blue solution (Cat. No. T8154) to assess viability.

Cells were then counted using a Neubauer haemocytometer and $2x10^4$ cells taken for assessment of purity using flow cytometry with Countbright beads (Cat. No. C36950) and DAPI (Cat. No. D9542). The concentration of DAPI used was 10µg/ml. The cell/bead mixture was allowed to

equilibrate to room temperature. The beads were then gently vortexed for 30 seconds. Immediately

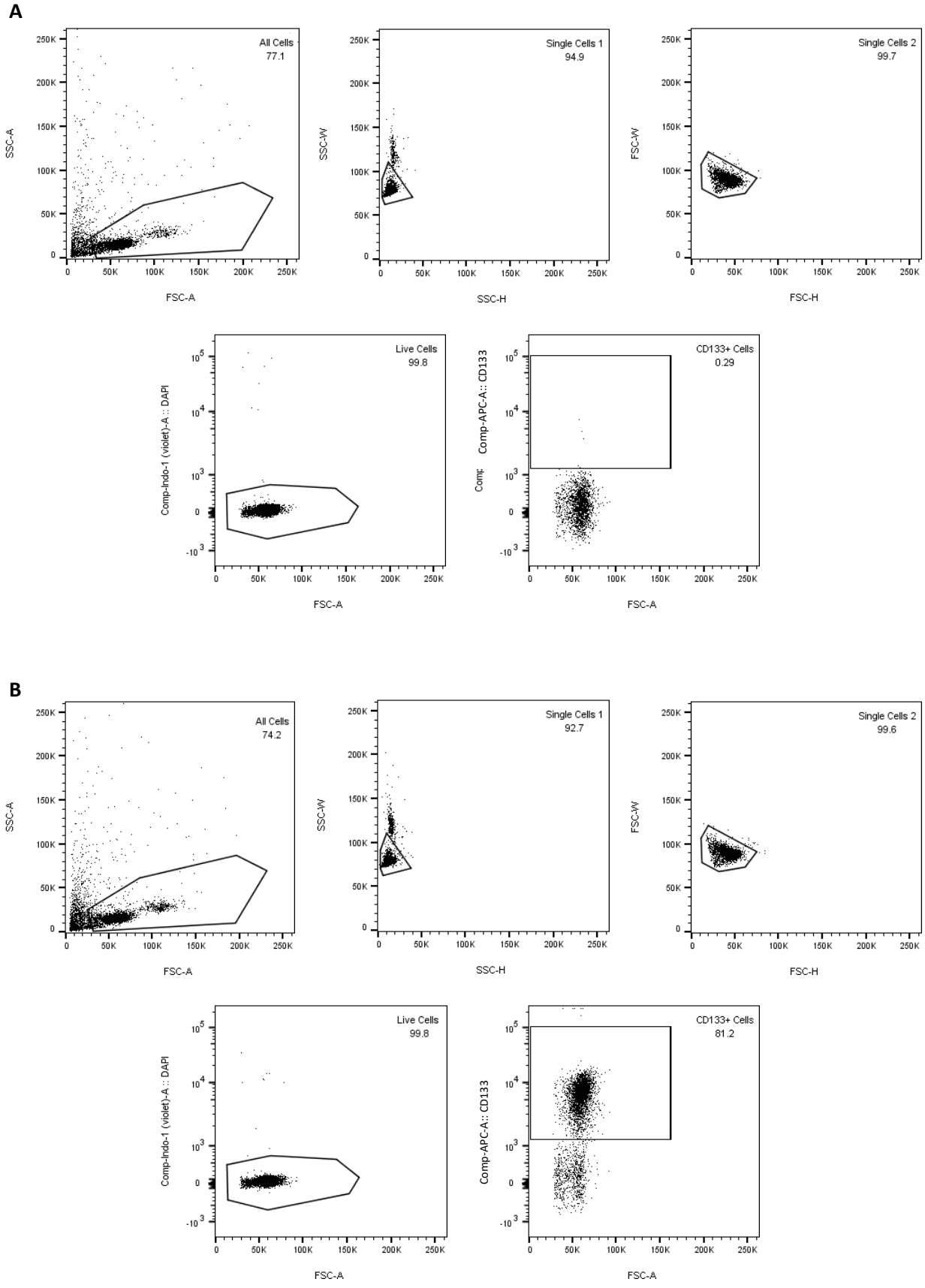


Figure 7: Representative gating strategy for the assessment of viability and purity of CD133+ cells isolated from fresh human umbilical cord blood. A) Gating workflow of the negative control. B) Gating workflow for CD133+ cells. The gating hierarchy is as follows: All cells > Single cells 1 > Single cells 2 > Live cells > CD133+ cells. The data are represented as dot plots.

after vortexing, the 50µl CountBright beads were added to a tube containing the previously stained samples. The samples were then run on a BD LSRII flow cytometer. Using the FACSDiva software, a forward scatter (FSC) and linear side scatter (SSC) plot was used to quantity the number of beads. The cells were also gated using the same metrics. This allowed us to calculate the absolute number of cells relative to the number of CountBright beads.

The remaining CD133+ cell suspension was centrifuged at 200g for 10 minutes at 4 degrees Celsius. Cells to be cryopreserved were resuspended in a cryopreservation solution consisting of 90% Fetal Bovine Serum and 10% Dimethyl Sulfoxide (DMSO) (Cat. No. 472301) at 1 x 10^5 cells/ml. Cells were then transferred to labelled cryopreservation vials (Cat. No. 4000198) and stored at -80 degrees Celsius overnight in a Mr. Frosty Freezing Container (Cat. No. C1562-1EA), containing isopropanol. They were then transferred within 24 hours to the vapour phase of liquid nitrogen.

Purity of the CD133+ cells was then assessed using flow cytometry, where 1x10^4 cells were stained with 10µl CD133/2-PE (Phycoerythrin) (293C3) (Cat. No. 130-113-186) and 10µl mIgG2b-PE isotype control (Cat. No. IC0041P) per 100µl MACS buffer and 10µl FcR blocking reagent for 30 minutes at 4 degrees Celsius. Data were obtained using flow cytometry on the BD LSRII. Following doublet and dead cell exclusion, cells were gated using PE CD133 and FSC-A, the gating strategy can be observed in figure 7.

2.2.2.2: CD133+ cells (plating and cell expansion)

CD133+ cell vials were thawed at 37 degrees Celsius using a water bath. The cells were then transferred to a 15ml conical tube (Cat. No. 11507411). Subsequently, the cells were washed using 10ml Iscove's Modified Dulbecco's Medium (IMDM) (Life Technologies) supplemented with 0.5% human serum albumin (HSA) (Zenalb20). The samples were then centrifuged for 7 minutes at 200g at 22 degrees Celsius. Following centrifugation, the supernatant was discarded, and the cell pellet was resuspended in an overnight media cocktail (see table 4), consisting of 5ml serum-free StemSpan SFEM (Cat. No. 09650), along with the relevant cytokine cocktail for overnight culture (5µl (100µg/ml) SCF,

TPO 5µl (20µg/ml) and FLT-3L (5µl (100µg/ml), at 37 degrees Celsius and 5% CO2 in air. Cells were subsequently cultured for according to the expansion protocols shown in tables 5 and 6, over 5 days. Cluster of differentiation (CD) marker specific antibodies (CD34-AF700, Cat. No. 560518), (CD38-PE-TXR, Cat. No. CYT-38F2), (CD45RA-APC-H7, Cat. No. 561212), (CD90-PE, Cat. No. 328109), (CD49F-PE-CY7, Cat No. 313621), (CD123-PERCP-CY5.5, Cat No. 306015) and (CD133-APC, Cat No. 130-090-851) were used to identify HSPCs.

Reagent	Cat. No:	Concentration	Volume
StemSpan SFEM	09650	Stock	5ml
SCF	255-SC-050	100ng/ml	5µl
TPO	288-TP-005	20ng/ml	5µl
FLT-3L	308-FK-025	100ng/ml	5µl

Table 4: CD133+ cell overnight media cocktail components, catalogue numbers, concentration and volumes used.

Treatment	SCF	TPO	FLT3L	Scriptaid	ZLLYFMK	PEITC
Basal	D-1, D0, D2	D-1, D0, D2	D-1, D0, D2	-	-	-
(1) Scriptaid	D-1, D0, D2	D-1, D0, D2	D-1, D0, D2	D0, D2	-	-
(2) Scriptaid + ZLLYFMK 1	D-1, D0, D2	D-1, D0, D2	D-1, D0, D2	D0, D2	D0	-
(3) Scriptaid + ZLLYFMK 2	D-1, D0, D2	D-1, D0, D2	D-1, D0, D2	D0, D2	D0, D2	-
(4) Scriptaid + ZLLYFMK 3	D-1, D0, D2	D-1, D0, D2	D-1, D0, D2	D0, D2	D2	-
(5) Scriptaid + PEITC 1	D-1, D0, D2	D-1, D0, D2	D-1, D0, D2	D0, D2	-	D0
(6) Scriptaid + PEITC 2	D-1, D0, D2	D-1, D0, D2	D-1, D0, D2	D0, D2	-	D0, D2

Table 5: Treatment combinations used for the ex vivo expansion of UCB derived CD133+ HSPCs.
D=Day. Concentrations used were as follows: SCF (100ng/ml), TPO (20ng/ml), FLT3L (100ng/ml), Scriptaid (1µM), ZLLYFMK (10µM) and PEITC (10µM).

Treatment	SCF	TPO	FLT3L	Scriptaid	UM171	SR1	NAM	VPA	PEITC	ZLLYFMK
Basal	D-1, D0, D2	D-1, D0, D2	D-1, D0, D2	-	-	-	-	-	-	-
(1) Scriptaid	D-1, D0, D2	D-1, D0, D2	D-1, D0, D2	D-1, D0	-	-	-	-	-	-
(2) UM171	D-1, D0, D2	D-1, D0, D2	D-1, D0, D2	-	D0, D2	-	-	-	-	-
(3) SR1	D-1, D0, D2	D-1, D0, D2	D-1, D0, D2	-	-	D0, D2	-	-	-	-
(4) UM171 + SR1 1	D-1, D0, D2	D-1, D0, D2	D-1, D0, D2	-	D0, D2	D0, D2	-	-	-	-
(5) UM171 + SR1 2	D-1, D0, D2	D-1, D0, D2	D-1, D0, D2	-	D0	D2	-	-	-	-
(6) UM171 + SR1 3	D-1, D0, D2	D-1, D0, D2	D-1, D0, D2	-	D2	D0	-	-	-	-
(7) NAM (Nicotinamide)	D-1, D0, D2	D-1, D0, D2	D-1, D0, D2	-	-	-	D0, D2	-	-	-
(8) VPA (Valproic Acid)	D-1, D0, D2	D-1, D0, D2	D-1, D0, D2	-	-	-	-	D0, D2	-	-
(9) Scriptaid + ZLLYFMK + PEITC	D-1, D0, D2	D-1, D0, D2	D-1, D0, D2	D0	-	-	-	-	D2	D0
(10) Scriptaid + PEITC	D-1, D0	D-1, D0	D-1, D0	D0	-	-	-	-	D2	-
(11) ZLLYFMK + PEITC	D-1, D0	D-1, D0	D-1, D0	-	-	-	-	-	D2	D0

Table 6: Treatment combinations used for UCB derived CD133+ HSPC expansion.
D=Day. Concentrations used were as follows: SCF (100ng/ml), TPO (20ng/ml), FLT3L (100ng/ml), Scriptaid (1µM), UM171 (35nM), SR1 (500nM), NAM (5mM), VPA (1mM), PEITC (10µM) and ZLLYFMK (10µM).

2.3: Cell counting

To achieve the correct plating density, the number of cells in the cell suspension were counted using the trypan blue exclusion method following methods by (Strober, 2001). Forty microlitres of cell suspension was transferred to a 1.5ml microcentrifuge tube (Cat No. 2150N), to which 40µl Trypan Blue solution (0.4 v/v %) (Cat. No. T8154) was added. The 1:1 mixture was gently vortexed, using a Scientific Industries Genie 2 vortex. Ten microlitres of the cell suspension and Trypan Blue solution (0.4%) mixture was added to a Neubauer Improved Haemocytometer. To calculate viable cells/ml the following method of calculation was performed. The average cell count from each of the sets of 16 squares was taken. This number was multiplied by 10^4 which was then multiplied by the dilution factor of 2 to achieve cells/ml. Once a viable cell count was established, a calculation was performed to determine total cells in suspension, by multiplying cells/ml by the total volume of cell suspension. Subsequently, the cell media to cell suspension ratio required to achieve the required cell plating density was calculated. The cells were then diluted using the correct volume of media and subsequently plated into the appropriate cell culture plate.

2.4: Cryopreservation of cell lines

All cell manipulations were conducted under sterile conditions and were the same method for HUDEP2 and KBM5. One million cells were counted using trypan blue exclusion and centrifuged at 300g for 5 minutes in a tabletop Beckman Coulter centrifuge. Following on, the cell media were removed, and the cells were reconstituted in cryopreservation solution, consisting of 90% FBS (Cat. No: 16140071) and 10% DMSO (Cat No: 67-68-5). The reconstituted cells were transferred to suitable cryopreservation vials (Cat No: V3135-500EA) and stored at -80 degrees Celsius in a Nalgene Mr. Frosty. After 24 hours, the cell vials were transferred to an RS Series liquid nitrogen storage unit (vapour phase) for long term storage.

2.5: Flow cytometry analysis

Flow cytometry was performed using the BD LSRII and the data was acquired in conjunction with FACS DIVA 9 software. Once acquired, the data were analysed using Flowjo 10 software. Cells were stained with DAPI to account for dead cells. Unstained, fluorescence minus one (FMO) and single colour controls were used to positively and negatively gate live cell populations for each fluorochrome.

2.6: Cell proliferation

Cells were counted using trypan blue exclusion. 50,000 cells were plated into 4 different round bottom black 96-well plates. The cells were cultured for a total of 4 days, every day 20µl media was added to each well. On the day of reading, cells were stained with Resazurin (Cat No: 62758-13-8), according to the manufacturer's protocol. The fluorescence, which is proportional to live cells, was then read using a Molecular Devices SpectraMax i3X at 562nm.

2.7: Cell cycle analysis

Cells were counted using trypan blue exclusion. On hundred thousand cells were transferred into a labelled flow cytometry tube and centrifuged at 300g for 5 minutes in a tabletop Beckman Coulter centrifuge. The cell media was removed, and the flow tube was topped up with PBS (Cat No: 14190-094). The cells were then centrifuged at 1600 rpm for 5 minutes. The wash was repeated twice in total. The cell pellet and PBS solution were gently vortexed, and 200µl ice-cold PBS was added. The tube was topped up with PBS and centrifuged at 1600 rpm for 5 minutes. This wash was once again repeated. A propidium iodide (PI) (Cat No: P4864**)** solution was subsequently prepared in water, according to the manufacturer's protocol. Ten microlitres PI (1mg/ml) was mixed with 250µl PBS and 2.5µl RNase (10mg/ml). The solution was gently vortexed and covered with aluminium foil. The sample tubes were then centrifuged at 1600 rpm for 5 minutes and the supernatant was discarded. Subsequently, the pellets were resuspended in the PI staining solution and covered with foil. The

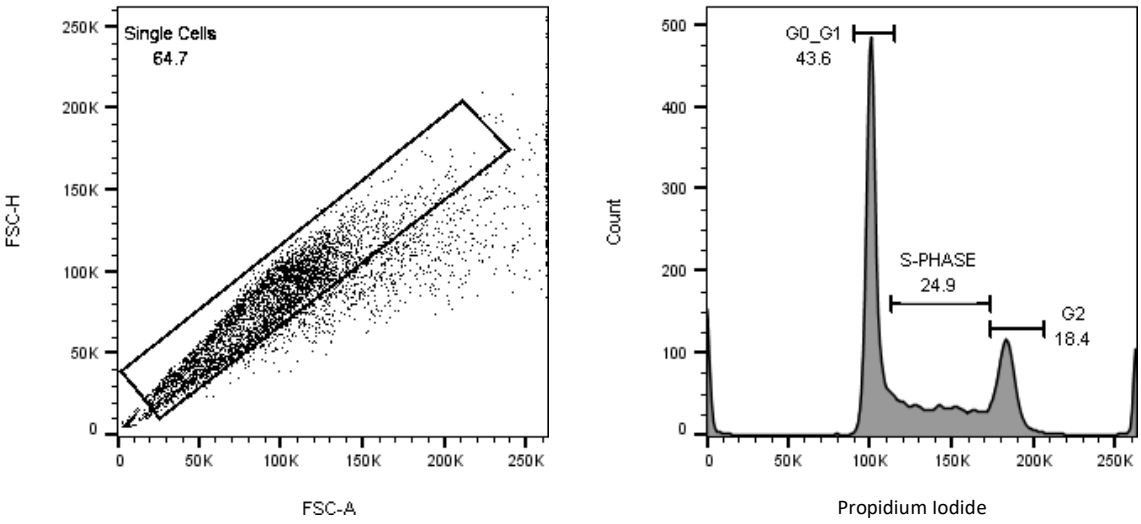


Figure 8: Representative gating strategy for cell cycle analysis.
The single cell data are represented as a dot plot, whereas a histogram was used to select distinct phases of the cell cycle.

samples were then incubated at room temperature for 60 minutes. Following the incubation, the flow tubes were read using an LSRII Flow Machine and subsequently analysed using FACS DIVA software. The gating strategy used can be observed in figure 8.

2.8: Apoptosis assay

Cells were counted using trypan blue exclusion. 100,000 cells were transferred into a labelled flow cytometry tube and centrifuged at 300g for 5 minutes. The cell media were removed, and the flow tubes were topped up with 500µl DPBS (Cat No: 14190-094) and centrifuged at 1600 rpm for 5 minutes. This wash was conducted a total of 2 times. After the final wash, the DPBS was removed, and the cells were resuspended in 100µl Binding Buffer (Cat. No. K182302). Five microlitres Annexin V (0.25µg/ml) (Cat No: ab14082) was added to the Annexin V positive control and samples to be analysed. These tubes were then incubated in the dark on ice for 30 minutes. One hundred microlitres binding buffer was added to each sample. 200µl DPBS (containing 1/10,000 DAPI, Cat. No. 10236276001) was added to all samples along with a PI-positive control (containing heat-shocked cells). The cell populations were then quantified using the BD LSRII in conjunction with FACS DIVA software. The gating strategy used can be observed in figure 9.

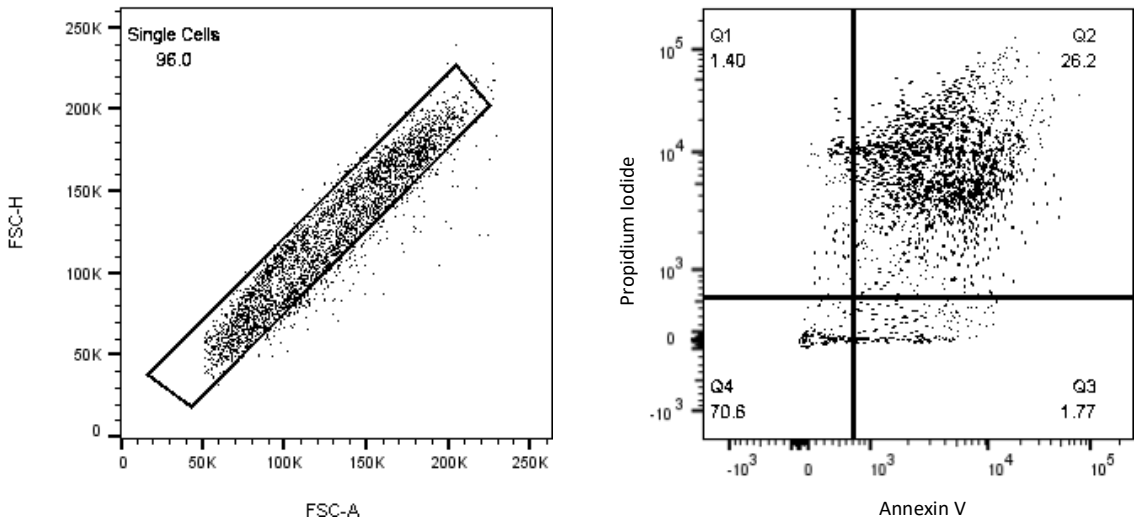


Figure 9: Representative gating strategy for apoptosis analysis.
The data are represented as a dot plot. A quadrant was used to gate Annexin V positive cells.

2.9: Analysis of erythroid and myeloid differentiation

For differentiation marker analysis, cells were counted using trypan blue exclusion. 100,000 cells were transferred into a flow tube and centrifuged at 1600 rpm for 5 minutes. The cells were subsequently washed twice with ice-cold PBS and following the final wash, the cells were resuspended in 200µl PBS. Antibodies was added to each sample tube according to manufacturer instructions. For HUDEP2 cells, the following directly conjugated antibodies were used: CD36 (PE) (Cat. No. 336205), CD71 (FITC) (Cat No. 11-0719-42), and CD235a (APC) (Cat No. 17-9987-42). By contrast, for KBM5 cells, the following antibodies were used: CD11b (BV421) (Cat No. 101235), CD14 (APC-H7) (Cat. No. 561384) and CD15 (ACP) Cat No. ab91151. The cells were stained for a total of 60 minutes in the dark on ice. For antibody dilutions, see table 7. Following the incubation, the tubes were topped up with PBS and centrifuged at 1600 rpm for 5 minutes. This wash was then repeated, and the cells were resuspended in 200µl PBS. Before flow cytometry, DAPI (Cat No: 10236276001) (1/10,000) was added to each sample to aid the exclusion of dead cells. Data were acquired using the BD LSRII in conjunction with FACS DIVA.

Antibody	Conjugate	Concentration
CD36	PE	200µg/ml
CD71	FITC	25µg/ml
CD235a	APC	3µg/ml
CD11b	BV421	100µg/ml
CD14	APC-H7	20µg/ml
CD15	APC	100µg/ml

Table 7: Antibody concentrations used for differentiation analysis sample staining.

2.10: Small scale plasmid purification

Small scale plasmid purification was conducted using the Qiagen QIAprep Spin Miniprep Kit (Cat. No. 27106). Mini preparation was conducted according to the manufacturer's instructions. Competent bacterial E. coli cells were centrifuged at 6800 x g for 3 minutes. The bacterial E. coli pellet was resuspended in 250µl P1 Buffer (supplied with kit), these were then transferred to a microcentrifuge tube. 250µl P2 buffer (supplied with kit) was then added to the preparation and gently vortexed for 5 seconds. 350µl N3 buffer (supplied with kit) was subsequently added to the mixture and inverted 6 times to ensure a homogenous solution preparation. This preparation was then centrifuged for 10 minutes at 17,900 x g in a Beckton Cutler tabletop centrifuge. Following pellet formation, the 800µl supernatant was added to the QIAprep 2.0 column included with the kit, and centrifuged for 60 seconds. The QIAprep 2.0 column was subsequently washed using 750µl PE buffer (supplied with kit) and centrifuged for a further 60 seconds. The flow-through was then discarded and the QIAprep 2.0 column was centrifuged for a further 1 minute. The QIAprep 2.0 column was placed in a clean 1.7ml microcentrifuge tube (Cat. No. CLS3620-500EA) followed by the addition of 50µl EB elution buffer (supplied with kit). Finally, the preparation was left at room temperature for 1 minute, then centrifuged for 1 minute at 13000 rpm. The preparation was stored at -80 degrees Celsius until required.

2.11: Large scale plasmid purification

Large scale plasmid purification was conducted using the QIAGEN Plasmid Plus Midi Kit (Cat. No. 12943). Midi preparation was conducted according to the manufacturer's conditions. The bacterial culture (30ml) was centrifuged at 6000g for 15 minutes at 4 degrees Celsius. The bacterial pellet was resuspended in a 4ml P1 buffer (supplied with kit). 4ml P2 buffer (supplied with kit) was subsequently added. The preparation was then gently mixed by inverting the tube 6 times and incubated at room temperature for 5 minutes. 4ml of P3 buffer (supplied with kit) was added and mixed by inverted 6 times. The preparation was subsequently incubated on ice for 15 minutes. The preparation as then centrifuged for 30 minutes at 20,000g. During this centrifugation period, 4ml QBT Buffer (supplied with kit) was added to a fresh QIAGEN-tip and allowed to empty via gravitational flow. Next, the preparation supernatant was added to the equilibrated QIAGEN-tip. The QIAGEN-tip containing the preparations washed twice using 10ml QC Buffer (supplied with kit). Gravitation pull was used to filter the solution through the QIAGEN-tip tube before elution. DNA was subsequently eluted using 5ml QF Buffer (supplied with kit) into a clean collection tube. DNA was next precipitated using 3.5ml isopropanol and centrifuged at 15,000g for 30 minutes. The supernatant was next carefully decanted. The pellet was finally air-dried for 10 minutes and reconstituted in an appropriate volume of TE Buffer (supplied with kit). This was stored at -80 degrees Celsius until further use.

2.12: Genomic DNA purification

Genomic DNA from the cell lines (HUDEP2 or KBM5) were purified using the PureLink™ Genomic DNA Mini Kit (Cat No: K182001), according to manufacturer's instructions. To prepare cell lysates, cell suspensions (3×10^6) were centrifuged at 300g for 5 minutes using a Beckman Cutler tabletop centrifuge. The growth medium was removed, and the cell pellet was resuspended in 200µl DPBS (Cat. No. 14190-094). Following on, 20µl Proteinase K (supplied with kit) was added along with 20µl RNase A (supplied with kit) and incubated at room temperature. Subsequently, 200µl PureLink® Genomic Lysis Buffer (supplied with kit) was added and gently mixed. The preparation was then incubated at

55 degrees Celsius for 10 minutes using a Beckton Cutler incubator. 200µl 100% ethanol was added to the lysate and vortexed for 5 seconds. The homogenous solution was next added to a PureLink® Spin Column (supplied with kit) and placed in a collection tube and then centrifuged at 10,000g for 1 minute at room temperature. The PureLink® Spin Column was then placed in a fresh collection tube. 500µl wash buffer 1 (supplied with kit) was added to the PureLink® Spin Column and centrifuged at room temperature at 10,000g for 1 minute. 500µl Wash Buffer 2 (supplied with kit) was next added to the column, and the column was centrifuged at 17,000g for 3 minutes. Finally, before elution, the spin column was placed into a sterile 1.7ml microcentrifuge collection tube. 30µl Elution Buffer (supplied with kit) was subsequently added to the spin column and was incubated at room temperature for 1 minute. Finally, the spin column was centrifuged for 2 minutes at room temperature. Extracted DNA was stored at -20 degrees Celsius until required.

2.13: DNA quantification and quality control

Eluted DNA quantity and quality was determined using the NanoDrop ND-1000 Spectrophotometer, coupled with the NanoDrop 1000 software, version 3.8.1. Purity was determined using the 260/280 absorbance ratio. A nucleic acid sample with a 260/280 ratio of ~1.8 was considered pure DNA.

2.14: Polymerase chain reaction

PCR was conducted using the Applied Biosystems Veriti 96 Fast Thermal Cycler. Primers were reconstituted in Invitrogen DEPC-Treated Water (Cat. No. 4387937) to 100µM, according to manufacturer's instructions. Primers were thawed at room temperature along with DreamTaq (Thermo Scientific: Cat No: K9012). 1.5ml microcentrifuge tubes were clearly labelled according to each condition. A PCR master mix was prepared according to table 8. The master mix was added to corresponding wells of a 96-well PCR plate before thermocycling. The thermocycling conditions can also be observed in table 9.

Reagent	Concentration	Volume
DreamTaq	5 U/µl	15µl
Primer(s)	10µM	1.5µl
DNA	100ng	1µl
DMSO	100%	0.5µl
Water	100%	10.5µl

Table 8: PCR reagents.

Initial Denaturation	Denaturation	Annealing	Extension	Final Extension	Cycles
95°C	95°C	59°C	72°C	72°C	-
3 Minutes	30 Seconds	30 Seconds	1 Minute	10 Minutes	35

Table 9: PCR thermocycler conditions.

2.15: Primers for PCR genotyping

PCR primers were designed using the primer blast online platform (www.ncbi.nlm.nih.gov/primer-blast). A 1000bp sequence of interest was input using the search box within the blast software. This sequence was carefully selected based on the region of the gene(s) we were interested in. This automatically generated sets of forward and reverse primer sequences. Primer sequences can be observed in table 10.

Gene Name	Sequence	Tm Temp.
***TET2* 1 Forward**	AGGGAAGTAGGAAGAAAAAGGGGT	59
***TET2* 1 Reverse**	TACTCATTTGCAACACTAGGTGAA	59
***TET2* 2 Forward**	TGGGTCTGGATTGACTATTCTTATT	59
***TET2* 2 Reverse**	ACCACACTTTACCAGTCTTTCTG	59
***ASXL1* G710X Forward**	AGAGGCAGCAGCAGTGGTGA	52
***ASXL1* G710X Reverse**	GAGGCTGCTCCACTAATCTCTCA	52
***ASXL1* G646 Forward**	AGGTCAGATCACCCAGTCAGTT	52
***ASXL1* G646 Reverse**	TAGCCCATCTGTGAGTCCAACTGT	52

Table 10: PCR primer sequences and Tm temperatures for TET2 and ASXL1 primers.

2.16: RNA extraction

Surfaces and experimental apparatus were thoroughly cleaned using Invitrogen™ RNaseZap™ Rnase Decontamination Solution (Cat. No. AM9780). RNA was extracted using the PureLink™ RNA Mini Kit (Cat No: 12183018A), according to manufactures instructions with slight modifications. Suspension cells were centrifuged at 300g for 5 minutes to pellet cells. The supernatant was subsequently discarded. The pellet was transferred to a 15ml falcon tube, to which 600µl Lysis Buffer with 1% v/v 2-mercaptoethanol (Cat No: M6250-100ML) (Per 5 x 10^6 cells) was added. The preparation was then vortexed until the cell pellet was lysed, and the preparation was transferred to another sterile 15ml falcon tube. The preparation was centrifuged for 2 minutes at 12,000g. One volume of 70% ethanol (supplied with kit) was then added. The preparation was vortexed gently until a precipitate was no longer visible. The solution was then transferred to a spin cartridge suspended inside a fresh collection tube. The cartridge containing the solution was then centrifuged for 15 seconds at 12,000g, at room temperature. The flow-through was discarded and a new collection tube was used. This process was repeated until the entire sample had been processed. Seven hundred microlitres Wash Buffer 1 (supplied with kit) was added to the spin cartridge and was then centrifuged at for 15 seconds at

12,000g. Once again the flow-through was discarded and the spin tube was placed into a new collection tube. Five hundred microlitres Wash Buffer 2 (supplied with kit) was subsequently added to the spin tube and centrifuged at for 15 seconds at 12,000g. This wash was conducted twice, with the flow-through discarded after every wash. The spin cartridge was then centrifuged for 2 minutes at 12,000g. The collection tube was discarded, and the spin tube was placed inside an RNA sample recovery tube. Thirty microlitres Rnase free water (supplied with kit) was added to the centre of the spin cartridge and it was incubated at room temperature for 1 minute. Subsequently, the spin cartridge containing the elution solution (Rnase free water) was centrifuged for 2 minutes at 12,000g to elute the RNA into the recovery tube.

2.17: RNA quantification and quality control

RNA was quantified, and quality controlled using the NanoDrop ND-1000 Spectrophotometer, coupled with the NanoDrop 1000 software version 3.8.1. A nucleic acid 260/280 absorbance ratio of 2 indicated pure RNA.

2.18: RNA bead purification

RNA bead purification was conducted using the RNAClean XP kit (Cat. No. A63987). The RNA Clean XP bottle was mixed using a rocker at room temperature for 30 minutes before use. RNAClean XP and extracted RNA were added to an appropriate round-bottom 96-well plate using the correct ratios recommended by the protocol. For example, for a sample of 20µl, 36µl of RNAClean XP was added. The preparation was mixed 10 times using a 200µl pipette. Following on, the round bottom 96-well plate was placed in a 96 super Magner Plate for 10 minutes, which separated the beads from the solution. The clear solution was carefully removed. Two hundred microlitres 70% ethanol (wash solution) was carefully dispensed into each well of the round bottom 96-well plate, and the plate was incubated at room temperature for 30 seconds. The ethanol was then discarded. The wash was repeated a total of three times. After the final ethanol volume was removed, the plate was left to air

dry for 10 minutes. Forty microlitres Rnase-free water was added to each well of the reaction plate, each well was mixed 10 times using a 200µl pipette. The RNA was contained within a clear solution. This was transferred to clearly labelled microcentrifuge tubes and stored at -80 degrees Celsius for later use.

2.19: Bioanalyzer

RNA was extracted and bead purified according to the methods described previously. Bioanalysis was conducted using the Agilent RNA 6000 Pico kit (Cat No: G2938-90046) according to the manufacturer's instructions. Initially, all the reagents were equilibrated at room temperature. The RNA 6000 Pico dye concentrate was then vortexed for 10 seconds and gently centrifuged. One microlitre RNA 6000 Pico dye concentrate was then mixed with 65µl filtered gel. The filtered gel was prepared by centrifuging 550µl RNA Pico gel matrix for 10 minutes at 1500g using the provided spin filter. The mixture was then vigorously vortexed and centrifuged at 13000g for ten minutes. The gel-dye mixture was then equilibrated at room temperature for a further 30 minutes. A brand-new RNA chip was placed on the priming station and 9µl gel-dye mixture was loaded into the appropriate well. The plunger was pressed down to promote gel dispersion for 30 seconds and then released. Following on, 9µl gel-dye mixture was added to the 2 remaining gel wells. Then 9µl RNA 6000 Pico conditioning solution was added to the appropriate well. Next, 5µl RNA Pico marker was added to all sample wells, in addition to the well representing the ladder. One microlitre RNA 6000 Pico ladder was added to the appropriate well. One microlitre RNA sample was then added to appropriate sample wells. Subsequently, using the IKA MS3 Vortex, the samples were mixed for 60 seconds at 2400 rpm. Next, the RNA chip was loaded into the Agilent 2100 Bioanalyzer Instrument, coupled with the Agilent 2100 Expert software package. The data were then exported as a PDF for analysis.

2.20: Dnase Treatment

Dnase treatment (Cat No: EN0521, Thermo Scientific™) of RNA samples was conducted according to manufacturer's instructions. 1µg RNA, 1µl 10x reaction buffer with MgCl2, 1µl Dnase I was added to a 500µl microcentrifuge tube. The tube was subsequently topped up to 10µl with RNAse free water. The tube was then incubated at 37 degrees Celsius for 15 minutes. 1µl of 25mM EDTA (per µg of RNA) was added (supplied with kit). The mixture was heated for 10 minutes at 65 degrees Celsius. The samples were then stored at -80 degrees Celsius for further use.

2.21: Complementary DNA synthesis

cDNA was synthesised according to the manufacturer's instructions using the High-Capacity cDNA Reverse Transcription Kit (Cat. No. 4368814). An RT master mix was prepared according to table 11. 10µl 2x RT master mix was added to well(s) of a 96-well plate, corresponding to each sample. 10µl (1µg total RNA) or 10µl water (negative control) was added to corresponding wells and gently mixed. The 96-well plates were sealing using Clear Polyolefin Star Seal (Cat No: E2796-9793) and centrifuged gently. The plate was then placed into the Veriti thermocycler and run using the program shown in table 12.

Reagent	Volume
10x RT Buffer	2µl
25x dNTP Mix (100mM)	0.8µl
10x RT Random Primers	2µl
MultiScribe Reverse Transcriptase	1µl
Rnase Inhibitor	1µl
Nuclease-free H_2O	3.2µl
Total per reaction	10µl

Table 11: cDNA synthesis components.

Setting	Step 1	Step 2	Step 3	Step 4
Temperature	25°C	37°C	85°C	4°C
Time	10 minutes	120 minutes	5 minutes	∞

Table 12: Thermocycler conditions for cDNA synthesis.

2.22: Quantitative real-time polymerase chain reaction (qRT-PCR)

qRT-PCR was conducted according to the manufacturer's instructions using the LightCycler® TaqMan® Master Mix (Cat. No. 04535286001). A master mix was prepared according to table 13. 12.5µl master mix was added to each well of a white bottom qRT-PCR culture plate (Cat. No. AB0900W). The qRT-PCR probes used are detailed in table 14. β2-Microglobulin (B2M) was used as a reference gene. The plate was run using the Roche Light cycler 96. Once complete, the were data analysed using the Lightcycler 96 software package.

Component	Volume/Reaction
Roche MM (2x)	6.25µl
cDNA (25ng)	Vol. for 25ng
Taqman Assay Probe	1.25µl (10x)
H_2O	Top up to 12.5µl
Total	12.5µl

Table 13: qRT-PCR components.

Gene Name	Probe ID
TET2	HS00758658658_M1
ASXL1	Hs00898213_M1
B2M	HS99999907_M1

Table 14: TaqMan Probes.

2.23: Gel electrophoresis of PCR products

Gel electrophoresis of PCR products was conducted to differentiate samples according to fragment size. A 1% w/v agarose gel was prepared by mixing 1g agarose (Cat. No: A5304-100G) in 100ml Tris/Borate/EDTA (TBE) (Cat. No. B52). The preparation was then heated for 2 minutes using a microwave at 600 Watts. The preparation was then mixed and heated for a further 30 seconds. The solution was cooled at room temperature for 10 minutes. Subsequently, 1µl SYBR Safe DNA Gel Stain (Cat. No. S33102) was added to the preparation. Next, the electrophoresis gel chamber was assembled with the appropriate combs, and the agarose preparation was poured into the chamber. Once the preparation was cooled the combs were removed. Next the chamber was placed into a Consort EV231. The Consort EV231 was then appropriately topped up with TBE. Once loaded with 5µl 1 kb DNA Ladder (Cat No: Do428) and 10µl PCR Product with 1:5 Orange DNA Loading Dye (6x) (Cat No: R0631), the machine was run using the following settings: Voltage: 120 Volts, Current: 100 Amps, Power: 80 Watts and Time: 60 minutes. Images were acquired using a Gene Flash and printed using the Syngene UP895-MD.

2.24: Sanger Sequencing

PCR products were run on a 1% w/v agarose gel to confirm the presence of a band(s) of the expected size. Following on, 10µl PCR product was run with 4µl ExoSAP-IT Express (Cat. No. 75001.1EA), according to the manufacturer's instructions. The mixture was gently vortexed and incubated at 37 degrees Celsius for 4 minutes and then at 80 degrees Celsius for 1 minute using the Applied Biosystems Veriti 96 Fast Thermal Cycler. Once ExoSAP-IT treated, the diluted PCR products (100ng/µl), along with primers (3.2pmol/µl) were sent to Source Bioscience (Oxford, UK) for Sanger sequencing.

2.25: Protein extraction

Three million cells were pelleted by centrifugation at 300g for 5 minutes. The cell media were then discarded and washed three times with chilled DPBS. Cell lysis solution was prepared by mixing 2ml RIPA buffer (Cat No: R0278) with 20µl protease and phosphatase inhibitor (Cat No: PPC1010). The DPBS from the final wash was discarded and the pellet was resuspended in 300µl lysis solution. The cell lysis solution was incubated at room temperature. Subsequently, the supernatant (containing protein) was transferred to a labelled 1.5ml microcentrifuge tube. The protein lysate was stored at -20 degrees Celsius until required or taken directly for protein quantitation.

2.26: Protein quantification

Proteins were quantified using a BCA assay (Cat No: 23227) according to the manufacturer's instructions, with some modifications. Bovine serum albumin standards were prepared using concentrations ranging from 25µg/ml to 2000µg/ml. A working BCA reagent was prepared by mixing 5ml Reagent A (supplied with kit) with 100µl Reagent 2 (supplied with kit). One hundred microlitres working BCA reagent was added to labelled wells of a flat bottom 96-well plate (Cat No: 83.3924), in duplicate. One row corresponded to the standards and one corresponding to the main samples. Ten microlitres of samples and 10µl of standards were added to corresponding wells. The plate was then

gently agitated using the Thermo Fisher Scientific Titer Plate Shaker, at speed 3 for 30 seconds. Subsequently, the plate was incubated at 37 degrees Celsius for 30 minutes and then loaded onto a Molecular Devices SpectraMax i3X. The onboard software was used to quantify total protein amount using absorbance at 532nm.

2.27: Single guide RNA design

Single Guide RNAs (sgRNAs) were designed using the CRISPOR online design platform (www.crispor.tefor.net). Target sequences were put into the software. Subsequently, sgRNA sequences were selected based on predicted efficiency, predicted specificity, and off-target activity. sgRNA sequences for *TET2* and *ASXL1* are shown in table 15 and table 16 respectively. To facilitate sgRNA incorporation into the PX458 plasmid (Cat. No. 48138), the following adaptor sequences were used: forward (CACC_G_G_n20) and reverse (AAAC_n20) (Ran, et al. 2013).

sgRNA Name	Guide Sequence + PAM
Downstream A	TTGACACTTTATTATCAGGT GGG
Downstream B	TCTGGCATCCATCGCAAAGT GGG
Downstream C	AGGACTAATAGTCTGCTAGC TGG
Upstream A	TATATGTACATGCTGGTTGT TGG
Upstream B	TTAAAATCCAAGTAATAGGT AGG
Upstream C	CTGGAAGCCTACCTATTACT TGG

Table 15: TET2 sgRNA sequences.

sgRNA Name	Guide Sequence + PAM
ASXL1 A	GCGGCCACCACTGCCARCGG AGG
ASXL1 B	CATTAAAGCCCGTGCTCTGC AGG
ASXL1 C	CATCGGAGGGGGGGGTGGCC CGG

Table 16: ASXL1 sgRNA sequences.

2.28: Cloning of sgRNA into PX458 CRISPR vector

The components detailed in table 17 were sequentially added to a labelled microcentrifuge tube and heated to 37 degrees Celsius for 60 minutes. The sgRNA forward and reverse primers were thawed at room temperature (100µM). 10µl 2x annealing buffer was added to the mixture. 5µl forward and reverse sgRNAs were added to corresponding wells, containing the annealing buffer. The PCR plate was covered using Starseal Sealing Tape. The PCR plate was run in the thermocycler using the ramp setting 95 degrees Celsius for 5 minutes and 25 degrees for 0.15 seconds. Next, the annealed sgRNAs were diluted using nuclease-free water (1:8). The ligation components in table 18 were mixed in labelled microcentrifuge tubes. The ligation samples were heated at 37 degrees Celsius for 60 minutes using a heat block. During this hour, competent cells were thawed on ice. Once the hour had elapsed, 5µl post ligation solution was added to the competent cells. This preparation was then mixed using a gentle flicking technique and subsequently incubated on ice for 20 minutes. The samples were then heat shocked for 1 minute at 42 degrees Celsius, and then cooled on ice for 5 minutes immediately after. 150µl S.O.C Media (Cat No: 15544034) was added and mixed using a pipette. This preparation was then shaken for 60 minutes at 37 degrees Celsius. Finally, the cell solution mixture was transferred onto LB-(Ampicillin) plates corresponding to each sample. The ampicillin plates were prepared by mixing 12g of LB Broth (Cat. No. L3022) with 600ml DDH2O, this was followed by the addition of 9g agar (Cat: No. LP0011). The solution was then autoclaved. A flat-ended mixer was used to equally distribute the solution once poured into a culture dish. The lid was replaced on the petri dish and they

were subsequently wrapped using parafilm. The plates were incubated overnight at 37 degrees Celsius. The PX458 plasmid schematic can be observed in figure 10.

Name	Volume/Concentration
PX458 Plasmid	1µg
10X NEBuffer 2.1	2µl
ddH2O	16µl
BbsI (NEB) (10 U/µl)	1µl
Total	20µl

Table 17: Components for restriction enzyme-based digestion of plasmid.

Name	Volume
Plasmid digest solution	20µl
Diluted oligonucleotide (1:8)	1µl
T4 DNA ligase buffer	2.5µl
T4 DNA ligase	1.5µl
Total	25µl

Table 18: Components for plasmid ligation.



Figure 10:The PX458 plasmid.
The U6 promoter encodes short RNAs and regulates the synthesis of sgRNA. The CBh promoter modulates Cas9 expression that is followed by a T2A self-cleaving peptide that releases a GFP reporter protein. The bovine growth hormone polyadenylation signal (bGH-PolyA) is used after the GFP gene (Nerys-Junior et al. 2018). Taken from: www.addgene.org.

2.29: Electroporation

Electroporation for both HUDEP2 and KBM5 cells was conducted using the Lonza 2b electroporator using the protocol (U-008). Following trypan blue exclusion, at least $1x10^6$ cells at 80% confluency were transferred to a labelled microcentrifuge tube. Following centrifugation at 300g for 5 minutes the cell media was removed. Subsequently, the cells were resuspended in the nucleofection solution from Cell Line NucleofectorTM Kit V (VVCA-1003). For a synthetic sgRNA preparation, the nucleofection solution consisted of 100µl nucleofector solution, 5µl Cas9 protein (20µg), synthetic sgRNA (1200pmol). The synthetic sgRNA preparation was used to knockout *TET2* in HUDEP2 cells since this showed the highest CRISPR efficiency in preliminary experiments. By contrast for a plasmid preparation, the nucleofection solution consisted of a PX458 plasmid (5µg), mutant ssODN (500pmol), RS1 (7.5µM/ml) and 100µl nucleofection solution. The plasmid-based method was used because it allowed bulk sorting on positive cells using the integrated GFP tag within the PX458 plasmid. These

respective solutions were then used to resuspend the cell pellet. The nucleofection cell solution was then transferred to an appropriate cuvette and placed into the nucleofector well and electroporated using the correct program (U-008) for both HUDEP2 and KBM5 cells. Once complete the cuvette was removed from the electroporation machine. Using a supplied pipette, the cell solution was removed from the cuvette into a 6 well culture plate containing 2ml prewarmed media. The plate was then placed into a Thermo Scientific HELA Cell 200 incubator at 37 degrees Celsius and 5% CO2.

2.30: Cell line expansion and single cell sorting

Following electroporation, the electroporated cuvette was removed from the electroporator. Subsequently, the contents (electroporated HUDEP2 or KBM5 cells) of the cuvette were removed using a plastic pipette supplied with the Cell Line NucleofectorTM Kit V (VVCA-1003) and placed into a 6-well suspension well plate contained prewarmed cell media (HUDEP2 or KBM5 media). The cuvette was washed using 1ml media to remove any residual cells. The cells were then transferred to a Themofisher Incubator at 37 degrees Celsius and 5% CO2 and expanded for 7 days. Every other day a 50% media change was conducted. Next, the cells were single-cell sorted into round-bottom 96-well plates using the BD FACSAriaIII, which were gated for live cells (DAPI). Before single-cell sorting, cells were pelleted and resuspended in either HUDEP2 or KBM5 media containing DAPI. Single cells were then expanded over 14 days, with the addition of 50µl to each well (using respective cell culture media) until a maximal volume of 200µl was reached. Once the cells reached more than 10,000/well approximately 40µl was used to extract DNA using the QuickExtract™ DNA Extraction Solution (Cat No: QE09050). Each sample was taken through the PCR pipeline and was subsequently genotyped. Positive colonies were expanded further using 24-well and 12-well culture plates. Once confluent, these cells were run through the quality control pipeline (qRT-PCR, off-target, etc) and were functionally assessed.

2.31: CRISPR/Cas9 off-target analysis

To investigate potential off-target cleavage in our CRISPR/Cas9 system, Sanger sequencing was used to analyse top predicted exonic changes predicted by CRISPOR (www.crispor.tefor.net) platform, which were used to design *TET2* and *ASXL1* sgRNAs. The off-target gene names, loci and primer sequences are displayed in tables 19, 20, 21 and 22.

Gene Name	Locus	Forward Primer	Reverse Primer
SLC22A10	chr11:63311147-63311169	AAGCCTTCATGAGCTGATTGGT	AGCCAGCCACATCCTAATTCTG
PGR	chr11:101031343-101031365	TGTGACTTCATGAGCCCTCC	CAACGGGGGTCTATGACCAG
ARHGAP11A	chr15:32638899-32638921	GCCAGTAGGTGTACCTTGTGTT	GCACTTTGAATTTCATGAGAGCCT
SNX24	chr5:123008869-123008891	CCTGGGTTTACTCAAGGGTGT	TTTAGTTCTTGCTTGTTTCCCAC

Table 19: Upstream A TET2 sgRNA off-target locus and primer sequences.

Gene Name	Locus	Forward Primer	Reverse Primer
AGAP2	chr12:57727510-57727532	GGGGAATTCAATCTGCGTGG	ACAGTAGCTGCTCGTACTTGG
PARP15	chr3:122635832-122635854	AGATTGGGCATGGTTCTACACA	TGACTAATCCTGCACGTCCC
TMEM27	chrX:15644995-15645017	GTCACCAACCGGTGATGGAG	TGGCTGATTGCACCTCAACA
MFI2	chr3:197024374-197024396	ACGAGAACGCTATGCACACA	GAGTACATCGCAGCCCATCA
PRKCE	chr2:46001449-46001471	ACCGTCGATGTGAGACCAAC	TGCTAGCCTGAAAAGGGCTC

Table 20: Downstream B TET2 sgRNA off-target locus and primer sequences.

Gene Name	Locus	Forward Primer	Reverse Primer
FAM53A	chr4:1640216-1640238	TCTGTGTTCCCCACGAAAGG	ATCGCAAACTCGAAACCCCT
DYRK1B	chr19:39833447-39833469	ACCAGCGACCAATGGTACAG	CCTTGGAGCTTCCCCCAAAA
MYLK	chr3:123752359-123752381	TTGCATAGTGGATTCTGGACC	AGGCCTTGGGGTAACTGAGA
PPP1R13L	chr19:45392194-45392216	ATGCCCTGGAAGTTGTGGAG	TACTTACCATGCTCTGCGGC
GRK4	chr4:2963902-2963924	CTAAGCCGTTAGCGCCGAG	CTTTCAGCAGCAGCGAGTTG

*Table 21: ASXL1 p.G646Wfs*12 sgRNA off-targets locus and primer sequences.*

Gene Name	Locus	Forward Primer	Reverse Primer
HK2	chr2:+75118439	GCTCATCACTGCTGTGGCCT	GGAATCCACAGGGCAAATGC
CAPN3	chr15:−42703613	CACCTGATAATCTCCAGTCTGCTCC	TTCAGGCATACATGGTGAGCTG
AQP10	chr1:−154295508	TGAGATGGAGTTTCACTCTGTTGCC	CGAGGAACAGGACCAAAGTAGTGTC
PRPH2	chr6:+42665654	CTACAGACGTCGCTGGATGGTG	CCCCTGCCTTGAAATTCTCTTG
PIN4	chrX:+71406302	GTGGCCCAAGATAATTCTTCCA	TGCGCCCAGCTACAATTTTG

Table 22: ASXL1 G710X sgRNA off-target locus and primer sequences
Taken from: (Valletta, Dolatshad et al. 2015). Copyright permission to reproduce these data was obtained from Springer Nature.

2.32: Data analyses

All data analyses were carried out using Microsoft Excel 2019, GraphPad Prism 8, and FlowJo 10. Data are expressed as Mean ± SD (standard deviation). One-way ANOVA analysis was conducted assuming Gaussian distribution. This was followed by Post hoc analysis using Dunnett's and Tukey's multiple comparison tests, which allowed 2 or more groups to be analysed. Each experimental condition was repeated multiple times to increase the robustness of the data. P values < 0.05 were considered statistically significant.

Chapter 3
Using CRISPR/Cas9 To Generate Models of Myeloid Malignancy

3.1: Introduction

Ten Eleven Translocation 2 (*TET2*) is a tumour suppressor protein, which regulates myelopoiesis (Li, Cai et al. 2011). It converts 5-methylcytosine (5-mC) to 5-hydroxymethylcytosine (5-hmC) (Chiba 2017). It is commonly mutated in the Myelodysplastic Syndromes (MDS), with mutations distributed throughout the coding region of the gene and considered to be loss-of-function (Itzykson, Kosmider et al. 2011).

The Additional Sex Combs Like Transcriptional Regulator 1 gene, codes for the polycomb group protein, *ASXL1*. *ASXL1* is an epigenetic modifier and tumour suppressor (Gelsi-Boyer, Trouplin et al. 2009). Somatic mutations of the *ASXL1* gene are common in MDS, with one of the most common mutation being p.G646Wfs*12 (Roche-Lestienne, Alice et al. 2010). Mutations of *ASXL1* are also drivers of secondary AML (Kakosaiou, Panitsas et al. 2018). *TET2* and *ASXL1* mutations can co-occur in myeloid malignancy (Rocquain, Carbuccia et al. 2010). Therefore, the aim of this chapter was to knock out *TET2* and to introduce specific point mutations in *ASXL1* in human erythroid and myeloid leukaemia cell lines.

To achieve this, we decided to genetically engineer cell lines using CRISPR/Cas9 technology (Jinek, Chylinski et al. 2012). Cas9 is a site-specific RNA guided endonuclease that introduces a double-stranded DNA break at intended target sites in the genome (Jinek, Chylinski et al. 2012). The Cas9 protein is guided to its intended target site by a 20-nucleotide single guide RNA (sgRNA) (Dang, Jia et

al. 2015). When the double-stranded DNA break occurs, cellular machinery is deployed to repair the breaks (Dang, Jia et al. 2015). The repair can occur in one of two ways. In the absence of an oligonucleotide template, the cell uses (NHEJ) non-homologous end joining to repair the breaks, which can often result in insertions or deletions (INDELS) (Burma, Chen et al. 2006). On the other hand, if an oligonucleotide template is provided, the cell uses homology-directed repair (HDR) to repair the break (Weber, Wefers et al. 2015).

The cell lines that we chose to study are HUDEP2 and KBM5, since the main cell lineages affected in MDS are cells of the erythroid and myeloid lineages (Malcovati, Della Porta et al. 2005, Della Porta, Malcovati et al. 2006). HUDEP2 (Human Umbilical Cord Derived Erythroid Progenitor 2) is an immortalized CD34+ haematopoietic stem cell-derived erythroid precursor cell line that can differentiate into erythroid cells (Vinjamur and Bauer 2018). HUDEP2 cells can be efficiently genetically manipulated and expanded (Vinjamur and Bauer 2018), making them a good choice for use with CRISPR/Cas9 based modelling. In HUDEP2 cells, we decided to use CRISPR/Cas9 technology to introduce the *ASXL1* p.G646Wfs*12 mutation, individually or in combination with a *TET2* knockout.

KBM5 cells are derived from a 67-year female with chronic myeloid leukaemia in blast phase (Beran, Pisa et al. 1993). KBM5 cells have an homozygous *ASXL1* G710X mutation, and our group has reported the effects of CRISPR/Cas9-mediated correction of this mutation in KBM5 cells in a previous study (Valletta, Dolatshad et al. 2015). In KBM5 cells, we decided to knock out the *TET2* gene in the *ASXL1* G710X heterozygous mutant variant, and in the *ASXL1* G710X homozygous corrected KBM5 variant (wildtype). The *ASXL1* mutations, be they p.G646Wfs*12 in the case of HUDEP2 cells, or G710X in the case of the KBM5 cells, both lie in the functional region of the gene, which suggests that these mutations may have functional, and therefore clinical importance (Valletta, Dolatshad et al. 2015, Rotunno, Mannarelli et al. 2019). This hypothesis will be tested in the next chapter.

3.2: Results

3.2.1: Generating *TET2* knockouts workflow

The initial stage of the project was to select regions on the *TET2* gene we wanted to target, which would give us the highest probability of a highly efficient knockout. Subsequently, we used the CRISPOR platform to design multiple sgRNAs, up and downstream of the intended target site. Following on, the sgRNA were cloned into the PX458 plasmid (which contains a GFP selection marker) and sequenced for confirmation of incorporation. Next, the PX458 plasmid was transfected into HUDEP2 cells. Subsequently, the HUDEP2 cells were sorted for GFP since it serves as a positive selection marker for the PX458 plasmid. DNA was extracted from these GFP+ HUDEP2 cells, and following PCR, the samples were Sanger sequenced and put through the ICE software pipeline (www.ice.synthego.com), which estimates sgRNA targeting efficiency and specificity. A cartoon depicting the *TET2* knockout workflow can be observed in figure 11.

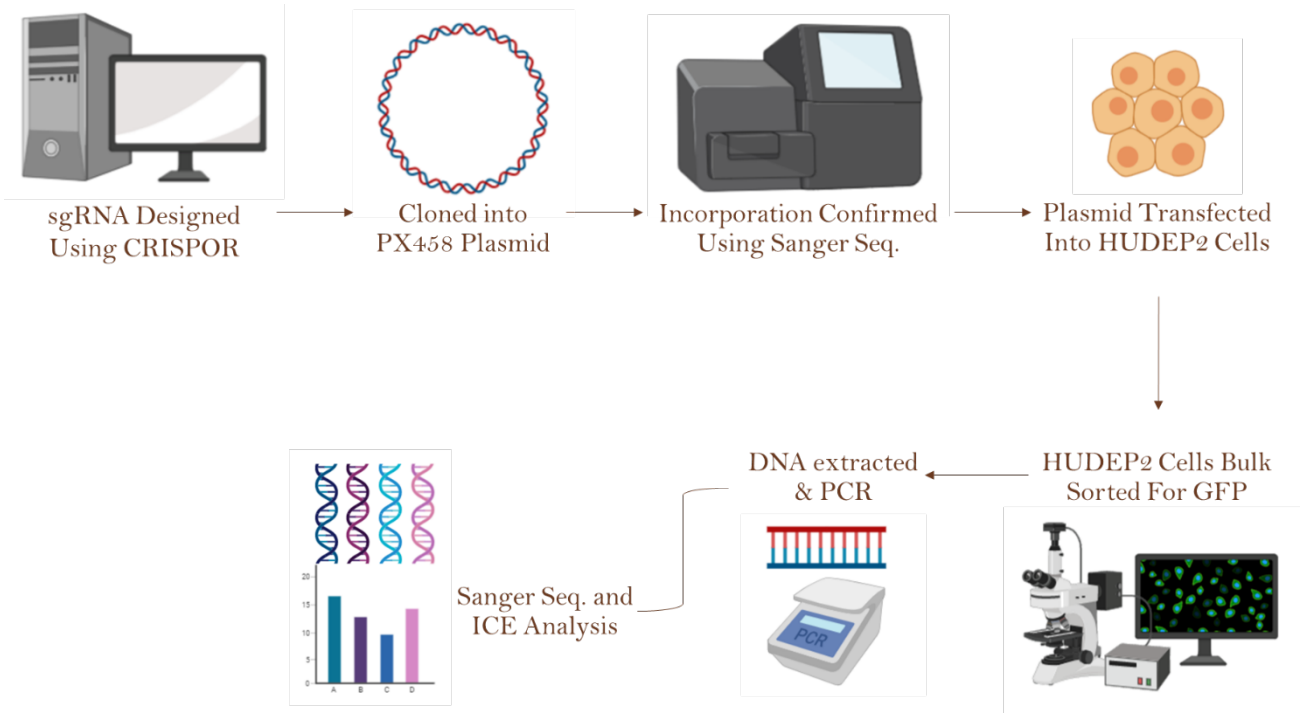


Figure 11: Experimental workflow for the design and engineering of sgRNAs targeted towards TET2.
The CRIPSPOR platform was used to design multiple sgRNAs. Subsequently, the sgRNAs were cloned into the PX458 plasmid, which was next sequenced to verify incorporation of sequences. The PX458 plasmid was then transfected into HUDEP2 cells, which were then sorted for GFP. Following DNA extraction and PCR, the samples were Sanger sequenced and used in the ICE pipeline, a software that estimates targeting efficiency and specificity.

3.2.2: The *TET2* gene targeting strategy

Previous pilot studies in our laboratory aimed to knock out the *TET2* gene by introducing a single base deletion. Unfortunately, this proved ineffective due to the presence of an alternate initiation codon downstream of the cut site, the indels introduced using CRISPR did not cause premature termination of translation and normal protein levels were detected through a western blot and therefore a more robust targeting strategy was required. We hypothesised that a larger, more important region of the *TET2* gene needed to be targeted to achieve a frameshift, and with it a complete knockout. Exon 4 is largely conserved across the majority of *TET2* isoforms, and it lies in the function region of the *TET2* protein (Yamazaki, Taby et al. 2012). We decided to employ a targeting strategy that would cause a double-stranded cut downstream and upstream of exon 4, leading to excision of the exon and frameshift and resulting in the generation of a premature stop codon, because of the premature termination, the transcript would be expected to be subject to non-sense mediated decay in the cell, thereby recapitulating loss of function mutations in MDS patients. The *TET2* gene assembly is shown in figure 12.

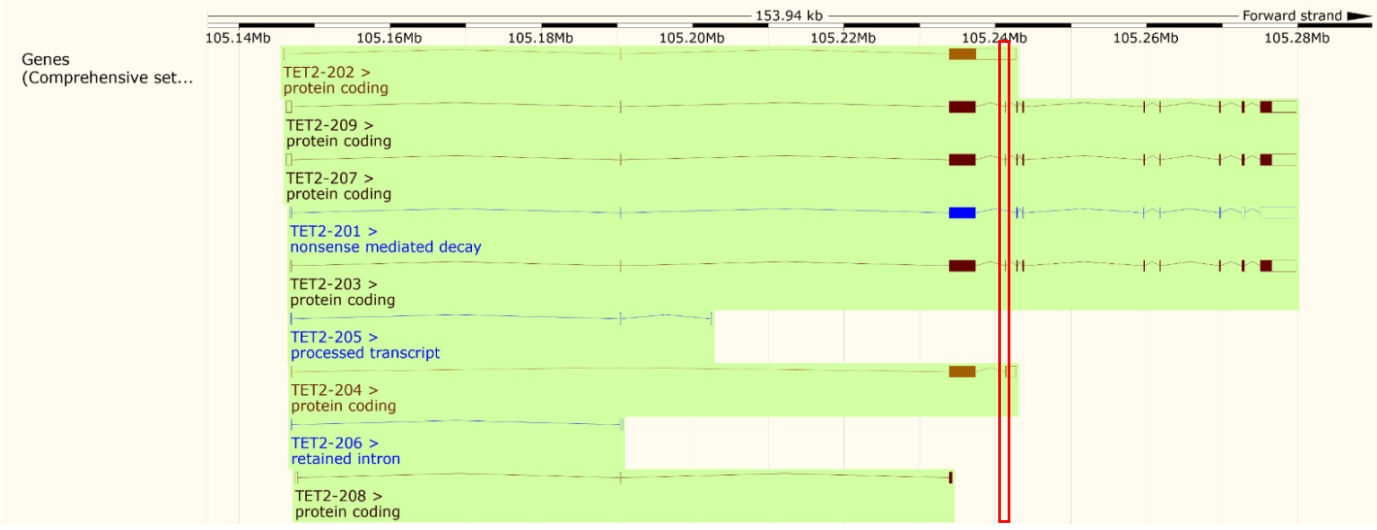


Figure 12: The Human TET2 Gene (GRCh38.p13) Assembly.
This figure has been adopted from the Ensembl genome browser showing various isoforms of TET2 expressed in the cell. The isoforms include protein coding transcripts and transcripts that undergo nonsense mediated decay. The thin horizontal line in each transcript represents intronic regions in the gene and the filled bars represent coding exons. As this figure shows, a majority of all protein coding transcripts include exon 4 (red box) and therefore, by excising exon 4 using CRISPR, we introduce a frameshift downstream leading to premature termination of all protein coding isoforms.

3.2.3: Using the CRISPOR platform to design *TET2* targeted sgRNAs

Since our *TET2* knockout strategy involved flanking exon 4 with a pair of sgRNAs, we needed to design an upstream set of sgRNAs and a downstream set of sgRNAs. Using the CRISPOR platform pipeline (Concordet and Haeussler 2018), coupled with Benchling software (Pellegrini 2016), we input a sequence of 500 base pairs, upstream and downstream of exon 4 into the CRISPOR online platform. This software-generated potential sgRNA sequences, which were ranked using a specificity and efficiency score. The higher the specificity score, the lower the predicted effects of the off-targets on the genome. Specificity scores range from 0-100 and measure the uniqueness of a guide in the genome. It is recommended that sgRNAs with values above 50 are used. As seen in table 23, the specificity scores were above 50 for all selected upstream and downstream sgRNAs. The higher the predicted efficiency score, the more likely the system is to cleave at the intended cut site. Similarly, in table 23, the efficiency scores are all above 50, but seeing as this is predicted efficiency, physiologically these values may be different. The final parameter, we used to shortlist sgRNA sequences was the out of frame score. The out of frame score (0-100) predicts the percentage of clones that will carry the out of frame deletions, this is based on the microhomology flanking the intended target site. As can be observed in table 23, the out of frame scores were all above 50, in line with the recommended guidelines, when selecting a sgRNA sequence.

	Guide Sequence + PAM	Specificity	Predicted Efficiency	Out of Frame Score
Downstream A	TTGACACTTTATTATCAGGT GGG	75	57	64
Downstream B	TCTGGCATCCATCGCAAAGT GGG	90	58	56
Downstream C	AGGACTAATAGTCTGCTAGC TGG	79	64	80
Upstream A	TATATGTACATGCTGGTTGT TGG	73	53	70
Upstream B	TTAAAATCCAAGTAATAGGT AGG	62	51	74
Upstream C	CTGGAAGCCTACCTATTACT TGG	90	34	77

Table 23: Predicted specificity, efficiency and out of frame scores for TET2 sgRNA sequences.
These were derived from the online sgRNA predictive CRISPOR platform, and served as the basis for initial sgRNA selection. Specificity: A score of the predicted effects of the off-targets on the genome. Predicted efficiency: A score of how likely the more likely the system is to cleave at the intended cut site. Out of frame score: The percentage of clones that may carry out of frame deletions.

3.2.4: *TET2* sgRNAs were successfully cloned into the PX458 plasmid

The *TET2* targeted sgRNAs were cloned individually into the PX458 plasmid and transformed into competent bacterial cells. Transformed colonies were picked and mini preparation was conducted. Subsequently, the presence of *TET2* sgRNA was verified using Sanger sequencing. Maxi preparation was carried out on single positive clones.

Sanger sequencing was conducted to confirm the incorporation of each *TET2* sgRNA into the plasmid. It can be observed in figure 13 that the downstream *TET2* sgRNAs were successfully incorporated into the PX458 plasmid. Similarly, in figure 14, the upstream *TET2* sgRNA sequences were also incorporated into the PX458 plasmid. Each plasmid containing a single sgRNA sequence was subsequently transfected into HUDEP2 cells. Then, the HUDEP2 cells were bulk sorted for GFP and further expanded. Figure 15 confirms that 99% (Downstream A), 99% (Downstream B) and 98.9% (Downstream C) of live cells in the bulk sorted HUDEP2 population were GFP+. Similarly, it can be observed in figure 16 that 99.1% (Upstream A), 99.2% (Upstream B) and 99% (Upstream C) of the live cells in the bulk sorted HUDEP2 cell population were GFP+. Following on, genomic DNA was extracted

from bulk sorted GFP+ HUDEP2 cells, PCR was conducted, and Sanger sequencing analysis was conducted.

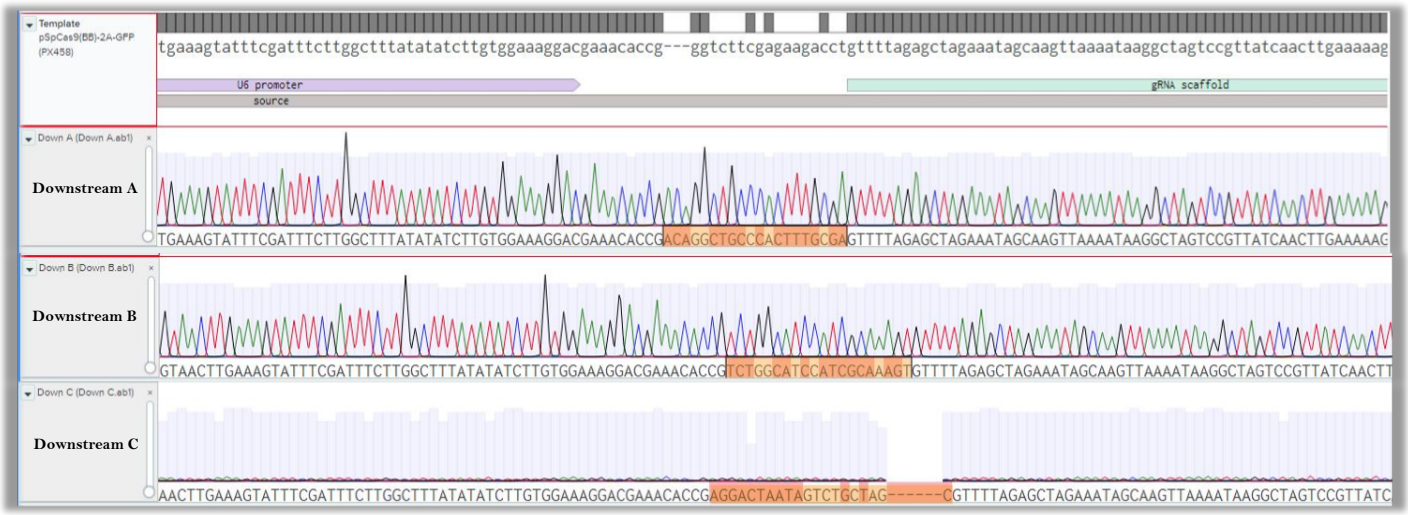


Figure 13: Incorporation of downstream TET2 sgRNAs into the PX458 plasmid.
A Sanger sequencing trace showing successful integration of sgRNA into the PX458 Cas9 backbone. The regions highlighted in orange show the 20-nucleotide sequence of 3 different sgRNAs namely downstream A, downstream B and downstream C cloned into the plasmid.

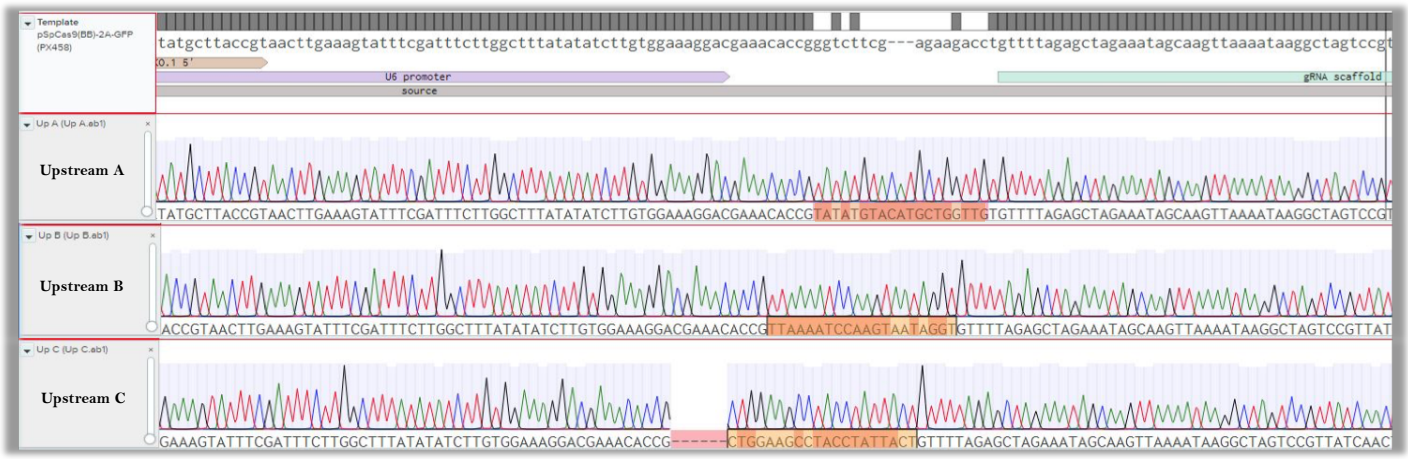


Figure 14: Incorporation of upstream TET2 sgRNAs into the PX458 plasmid.
A Sanger sequencing trace showing successful integration of sgRNA into the PX458 Cas9 backbone. The regions highlighted in orange show the 20-nucleotide sequence of 3 different sgRNAs namely upstream A, upstream B and upstream C cloned into the plasmid.

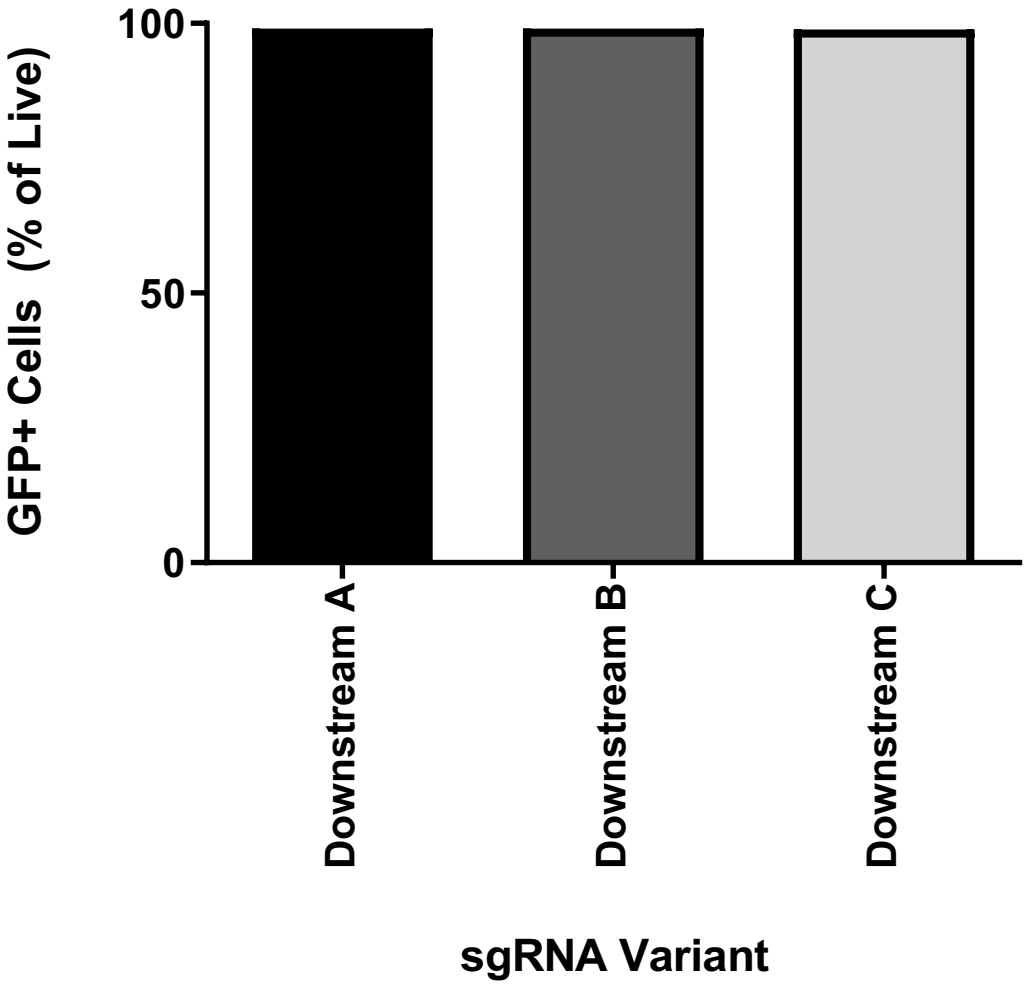


Figure 15: Percentage of bulk sorted GFP+ HUDEP2 cells within live cells, for cells containing the PX458 plasmid expressing downstream TET2 sgRNA variants, n=1. Each plasmid containing a single sgRNA sequence was transfected into HUDEP2 cells, before bulk sorting for GFP. As can be observed, the percentage of bulk sorted HUDEP2 cells are approximately 99% GFP+ cells using Downstream A (black), Downstream B (dark grey)) and Downstream C (light grey).

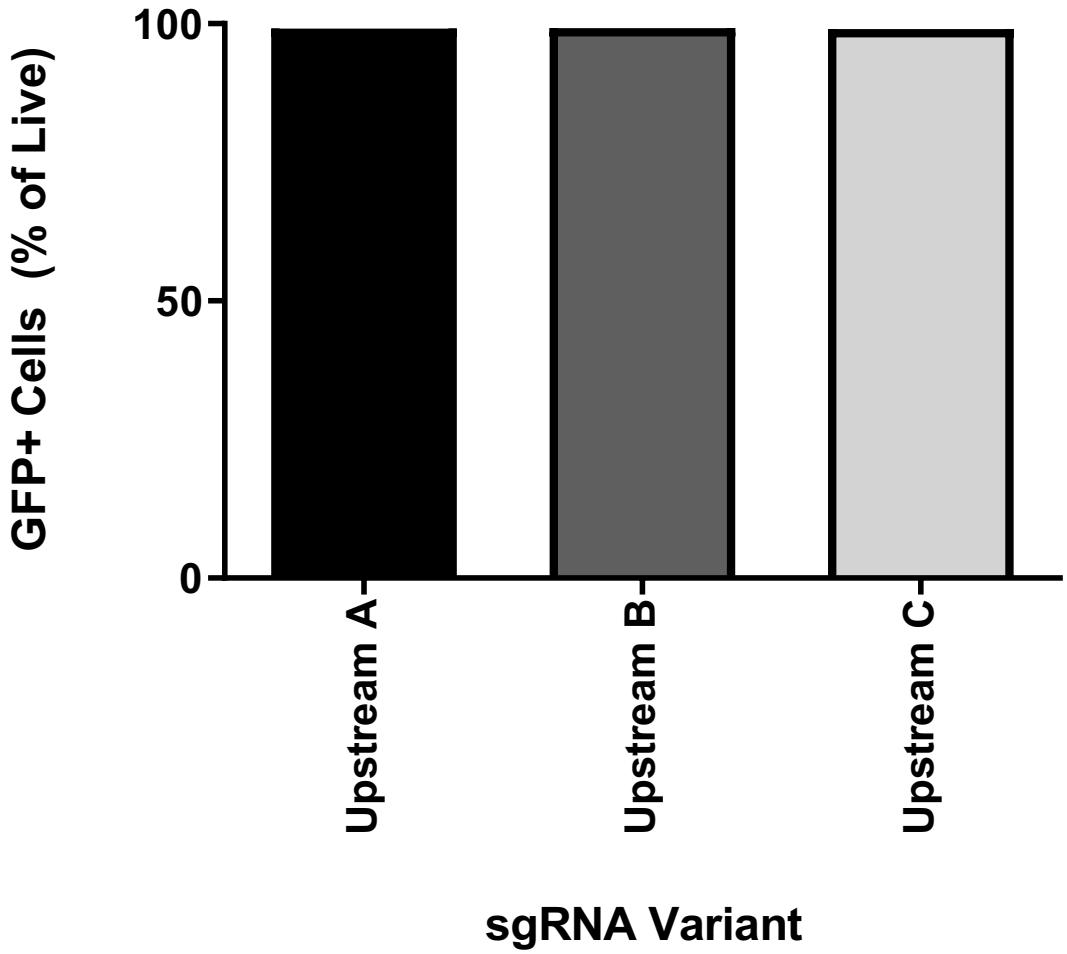


Figure 16: Percentage of bulk sorted GFP+ HUDEP2 cells within live cells, for cells containing the PX458 plasmid expressing upstream TET2 sgRNA variants, n=1. Each plasmid containing a single sgRNA sequence was transfected into HUDEP2 cells, before bulk sorting for GFP. As can be observed, the percentage of bulk sorted HUDEP2 cells are approximately 99% GFP+ cells using Upstream A (black), Upstream B (dark grey)) and Upstream C (light grey).

3.2.5: Validating the efficiency of the *TET2* sgRNA variants using ICE analysis

DNA extracted from bulk sorted GFP+ HUDEP2 cell populations were subjected to Sanger sequencing and results were then decomposed by using an online web tool i.e., ICE (Inference of CRISPR Edits) to rapidly assess genome editing by CRISPR-Cas9 at a target locus. Based on the quantitative sequence trace data from two standard capillary sequencing reactions, the ICE software quantifies the editing efficacy, and identifies the predominant types of insertions and deletions (INDELS) in the DNA of a targeted cell pool. Frequency of indels caused by NHEJ for each sgRNA correlates to sgRNA efficiency at the genomic site. The ICE software generates an efficiency score and a KO score. The KO score represents the number of cells that have incurred frameshift mutation or a greater than 21 base pair indel. As shown in figure 17, the Downstream B sgRNA had the highest ICE value (93) and KO-score (81) respectively and was the most efficient out of all the downstream sgRNAs tested. The next most efficient sgRNA was Downstream A with an ICE score of 42 and a KO score of 33. Finally, the least efficient sgRNA was Downstream C, with an ICE score of 0 and a KO score of 0 (see figure 17). The cutting efficiency of each downstream *TET2* sgRNA variant also correlates well with the observed results of the Sanger sequencing traces shown in figure 18. The decrease in the peak of signal after the PAM site can be observed in the traces, representing the more efficient sgRNA variants.

The most efficient upstream *TET2* sgRNA was Upstream A, with an ICE score of 93 and a KO score of 91. The second most efficient sgRNA was Upstream B with an ICE score of 84 and a KO score of 69. Finally, the least efficient upstream sgRNA variant was upstream C with an ICE score of 2 and a KO score of 2 (see figure 19). The cutting efficiency of the upstream *TET2* sgRNA variants also correlated well with the Sanger sequencing traces shown in figure 20, with a decrease in trace signal after the PAM site visible, which represent the more efficient sgRNAs.



SUCCESSFULLY ANALYZED SAMPLES

SAMPLE	GUIDE TARGET	PAM SEQUENCE	ICE	R^2	KO-SCORE	
Downstream A	ACAGGCTGCCCACTTTGCGA	TGG	42	0.43	33	VIEW DETAILS
Downstream B	TCTGGCATCCATCGCAAAGT	GGG	93	0.94	81	VIEW DETAILS
Downstream C	AGGACTAATAGTCTGCTAGC	TGG	0	1	0	VIEW DETAILS

Figure 17: ICE analysis of downstream TET2 sgRNA variants A, B and C, showing percentage efficiency as a product of ICE score and KO score. It can be observed, the most efficient downstream sgRNA was downstream B, with an ICE score of 93%.

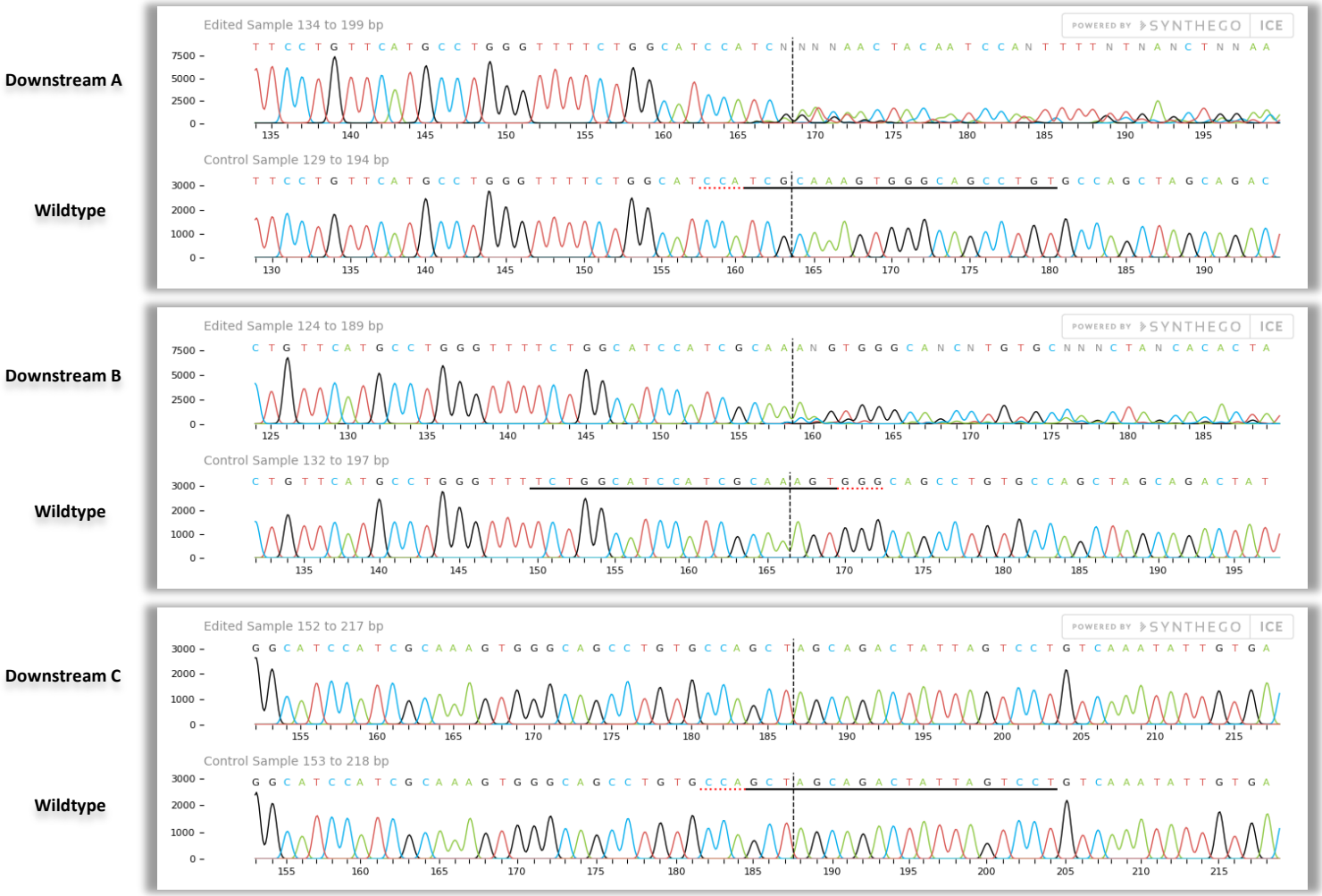


Figure 18: Sanger sequence traces of downstream TET2 sgRNA variants. Nucleotides are represented by different colour peaks. As can be observed, the sequence collapse for downstream B is the most pronounced, correlating well with this sgRNA having the highest efficiency.

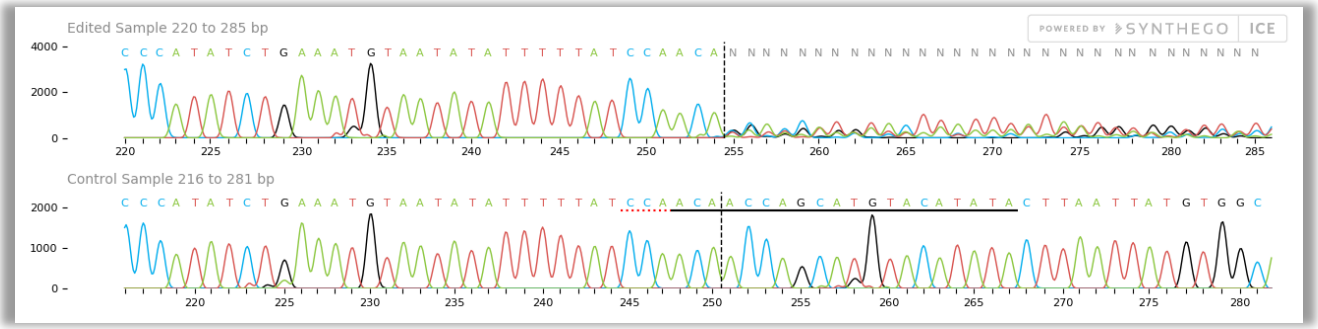


SUCCESSFULLY ANALYZED SAMPLES

SAMPLE	GUIDE TARGET	PAM SEQUENCE	ICE	R^2	KO-SCORE	
Upstream A	TATATGTACATGCTGGTTGT	TGG	93	0.94	91	VIEW DETAILS
Upstream B	TTAAAATCCAAGTAATAGGT	AGG	84	0.87	69	VIEW DETAILS
Upstream C	CTGGAAGCCTACCTATTACT	TGG	2	0.99	2	VIEW DETAILS

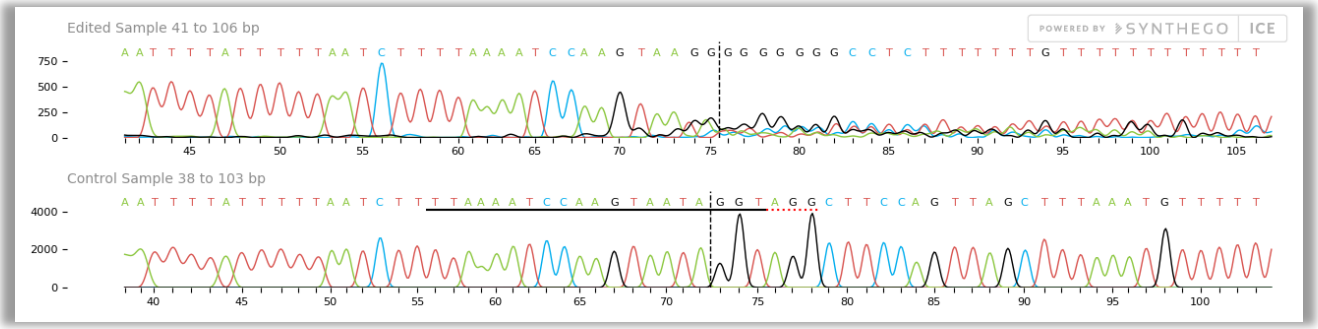
Figure 19: ICE analysis of upstream TET2 sgRNA variants A, B and C, showing percentage efficiency as a product of ICE score and KO score. It can be observed, the most efficient upstream sgRNA was upstream A, with an ICE score of 93%.

Upstream A



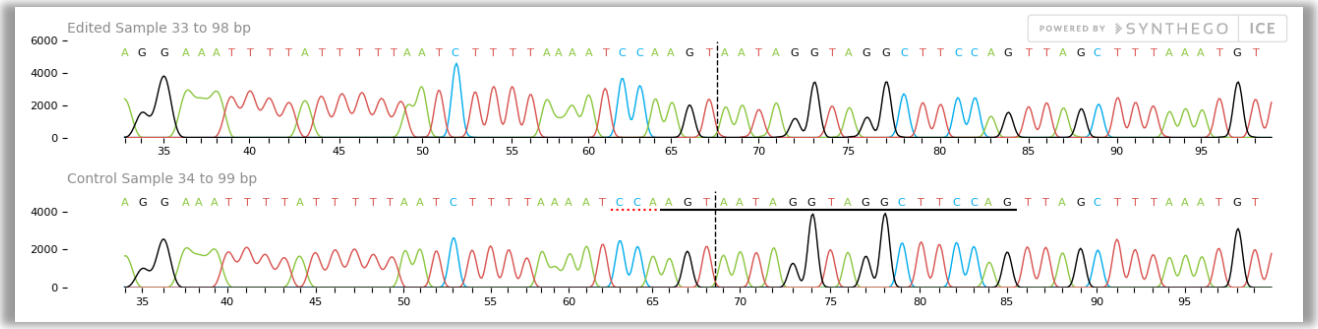
Wildtype

Upstream B



Wildtype

Upstream C



Wildtype

Figure 20: Sanger sequence traces of upstream TET2 sgRNA variants. Nucleotides are represented by different colour peaks. As can be observed, the sequence collapse for upstream A is the most pronounced, correlating well with this sgRNA having the highest efficiency.

3.2.6: The *ASXL1* gene targeting strategy

In MDS, *ASXL1* mutations are commonly located within exon 12 (Schnittger, Eder et al. 2013). It is highly conserved and helps *ASXL1* bind to chromatin, propagating its role as a chromatin modifier (Binder, Gaspar Maia et al. 2019). Common MDS mutations of *ASXL1* are p.G646Wfs*12 (Schnittger, Eder et al. 2011) and G710X (Valletta, Dolatshad et al. 2015) (see figure 21). We decided to use HDR to introduce the p.G646Wfs*12 mutation into HUDEP2 cells. We designed *ASXL1* targeting sgRNAs and a 127bp mutant single-stranded oligonucleotide (91 and 36 nucleotides upstream and downstream, respectively, of the intended cut site), which incorporated the intended seed changes, conducted according to methods by Richardson and colleagues (Richardson, Ray et al. 2016).

The targeting strategy for the p.G646Wfs*12 *ASXL1* mutant can be observed in figure 22. The targeting strategy for the G710X *ASXL1* mutant (which was previously engineered in our laboratory) can be observed in figure 23 (Valletta, Dolatshad et al. 2015). The seed changes are used to introduce a restriction enzyme cutting site for easier screening of oligonucleotide integrated clones. Silent changes to the PAM are usually made to prevent re-cutting at the site of the double stranded break. A silent change was not compatible with the PAM site and therefore we introduced silent changes in the seed region of the sgRNA (8-12 nucleotides upstream of the sgRNA) to decrease affinity of the sgRNA at the intended locus, once the double stranded break has been repaired using the template, in order to prevent re-cutting.

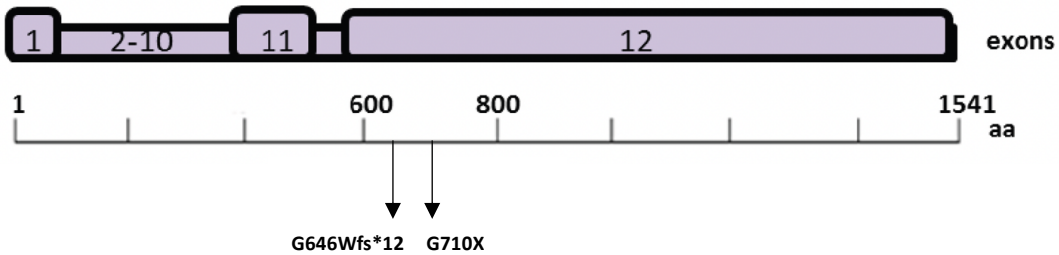


*Figure 21: The Structure of the ASXL1 gene. ASXL1 maps to the chromosome region 20q11 and comprises 12 exons. The location of the p.G646Wfs*12 and G710X ASXL1 Mutation Sites are shown by the black arrows and are located in exon 12.* Taken from*: (Valletta, Dolatshad et al. 2015). Copyright permission to reproduce this figure was obtained from Springer Nature.*

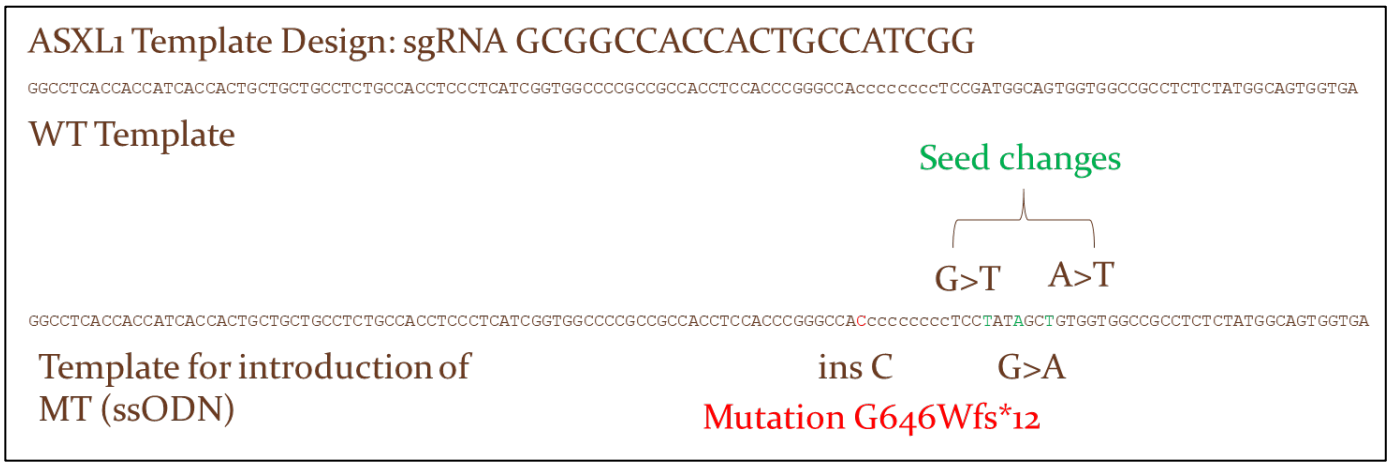


*Figure 22: The p.G646Wfs*12 ASXL1 Targeting Strategy.*
*The bottom sequence shows the asymmetric ssODN repair template used for HDR with the nucleotide change (C>A) associated with mutation G646Wfs*12, most commonly seen in patients with myeloid malignancies, and the silent nucleotide changes (highlighted in green) in the seed region, region immediately downstream of the PAM site. The silent changes are to prevent recutting by Cas9 at the PAM site once the changes have been incorporated.*

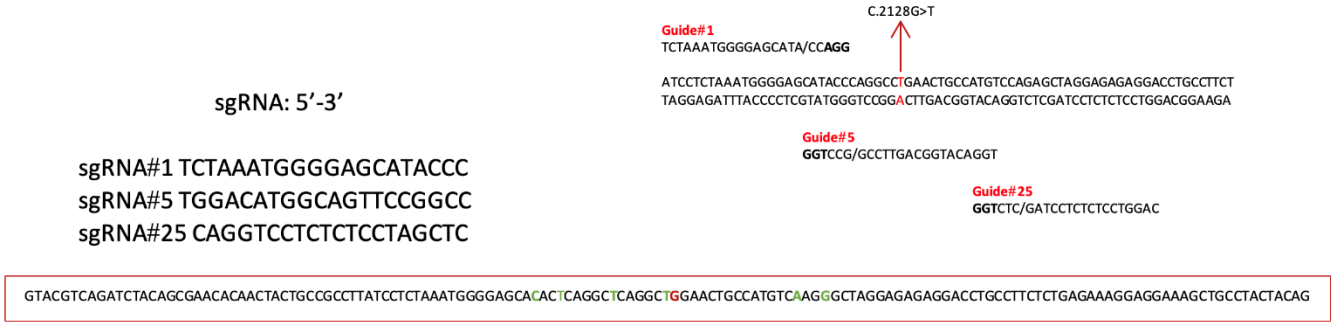


Figure 23: The G710X ASXL1 Targeting Strategy.
The three sgRNA sequences, namely #1, #5 and #25, shown at the top, were initially chosen to target the intended cut site. After testing them for cutting efficiency, only one of them was used to introduce mutations. Shown below is the sequence for the asymmetric donor template used for HDR with the intended mutation change (G>T; highlighted in red).

3.2.7: Using the CRISPOR platform to design *ASXL1* targeted sgRNAs

We designed *ASXL1* sgRNAs targeted towards the intended cut site G646 using the CRISPOR platform pipeline, coupled with Benchling software. Using the CRISPOR online platform, we inserted a sequence of 100 base pairs upstream and downstream of the intended cut site. The CRISPOR platform generated potential sgRNA sequences, were ranked using specificity (off-target score) and efficiency (on-target score). In addition, the distance of the intended cut site from the mutation insertion site (ideally fewer than 15 nucleotides) was an important consideration.

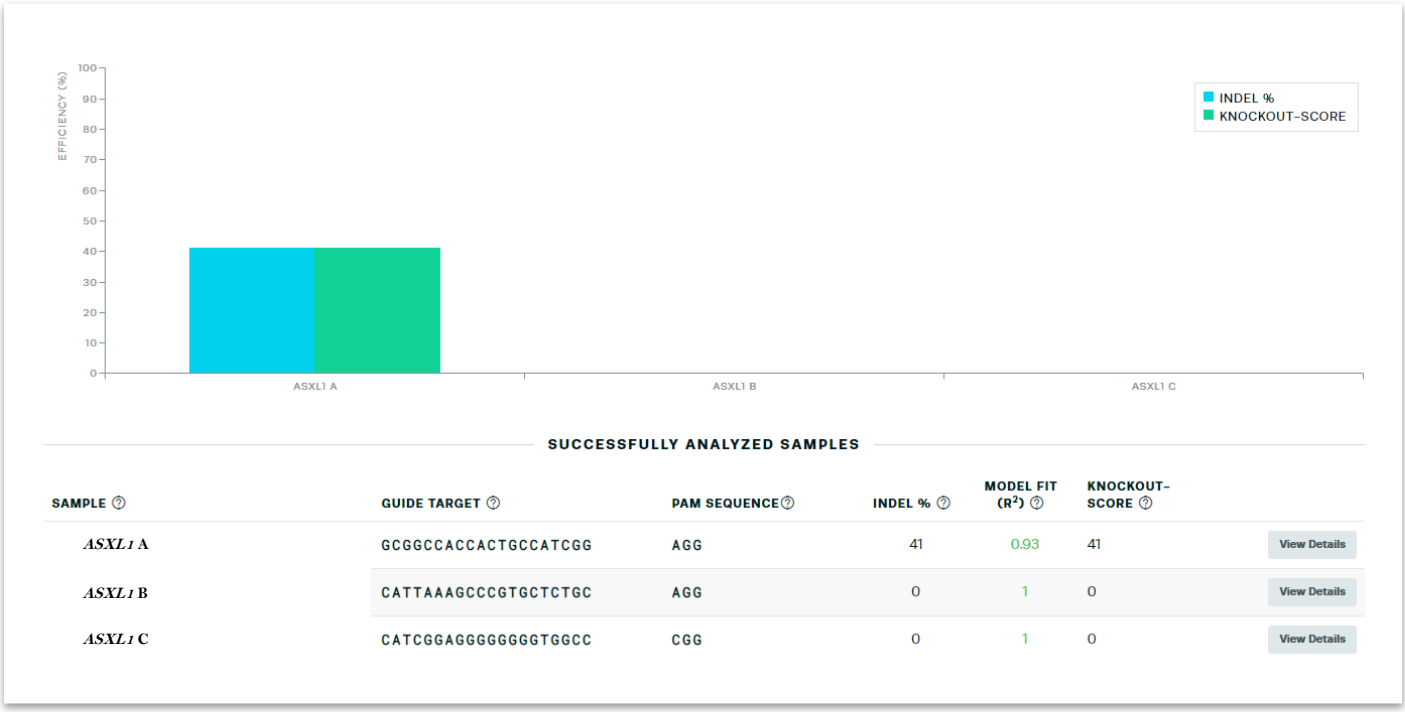
The higher the specificity score, the lower the effects of the off-targets on the genome. Specificity scores range from 0-100 and measure the uniqueness of a guide in the genome. As seen in table 24, the specificity scores for the *ASXL1* targeted sgRNAs are 79 (*ASXL1* A), 87 (*ASXL1* B) and 54 (*ASXL1* C). The second parameter we used to select the sgRNAs is predicted efficiency (above 50 is optimal). The predicted efficiencies for the *ASXL1* targeted sgRNAs were 54 (*ASXL1* A), 45 (*ASXL1* B) and 31 (*ASXL1* C).

sgRNA Name	Guide Sequence + PAM	Specificity	Predicted Efficiency
***ASXL1* A**	GCGGCCACCACTGCCARCGG AGG	79	54
***ASXL1* B**	CATTAAAGCCCGTGCTCTGC AGG	87	45
***ASXL1* C**	CATCGGAGGGGGGGGTGGCC CGG	54	31

Table 24: Predicted specificity and efficiency scores for ASXL1 sgRNA sequences.
These were derived from the online sgRNA predictive CRISPR platform and served as the basis for initial sgRNA selection.

3.2.8: Validating the efficiency of the *ASXL1* sgRNA variants using ICE analysis

Following PCR and Sanger sequencing of the exon 12 region of the *ASXL1* gene, we used the Inference of CRISPR Edits (ICE) software tool (Hsiau, Maures et al. 2018) to generate an INDEL % and knockout score. We observed that the *ASXL1* A sgRNA variant had the highest INDEL % value (41) and KO-score (41) respectively and was the most efficient out of all the *ASXL1* targeted sgRNAs tested. *ASXL1* B and *ASXL1* C variants showed an INDEL % value and a knockout score of 0 (see figure 24). The cutting efficiency of the *ASXL1* A sgRNA variant also correlates well with the sequence collapse, and decrease in peak intensity after the PAM site, as can be observed in the Sanger sequencing trace shown in figure 25.



SUCCESSFULLY ANALYZED SAMPLES

SAMPLE	GUIDE TARGET	PAM SEQUENCE	INDEL %	MODEL FIT (R^2)	KNOCKOUT-SCORE	
ASXL1 A	GCGGCCACCACTGCCATCGG	AGG	41	0.93	41	View Details
ASXL1 B	CATTAAAGCCCGTGCTCTGC	AGG	0	1	0	View Details
ASXL1 C	CATCGGAGGGGGGGGTGGCC	CGG	0	1	0	View Details

Figure 24: ICE analysis of ASXL1 sgRNA variants A, B and C, showing percentage efficiency as a product of ICE score and KO score. It can be observed, that the most efficient ASXL1 sgRNA was ASXL1 A, with a knockout score of 41%.

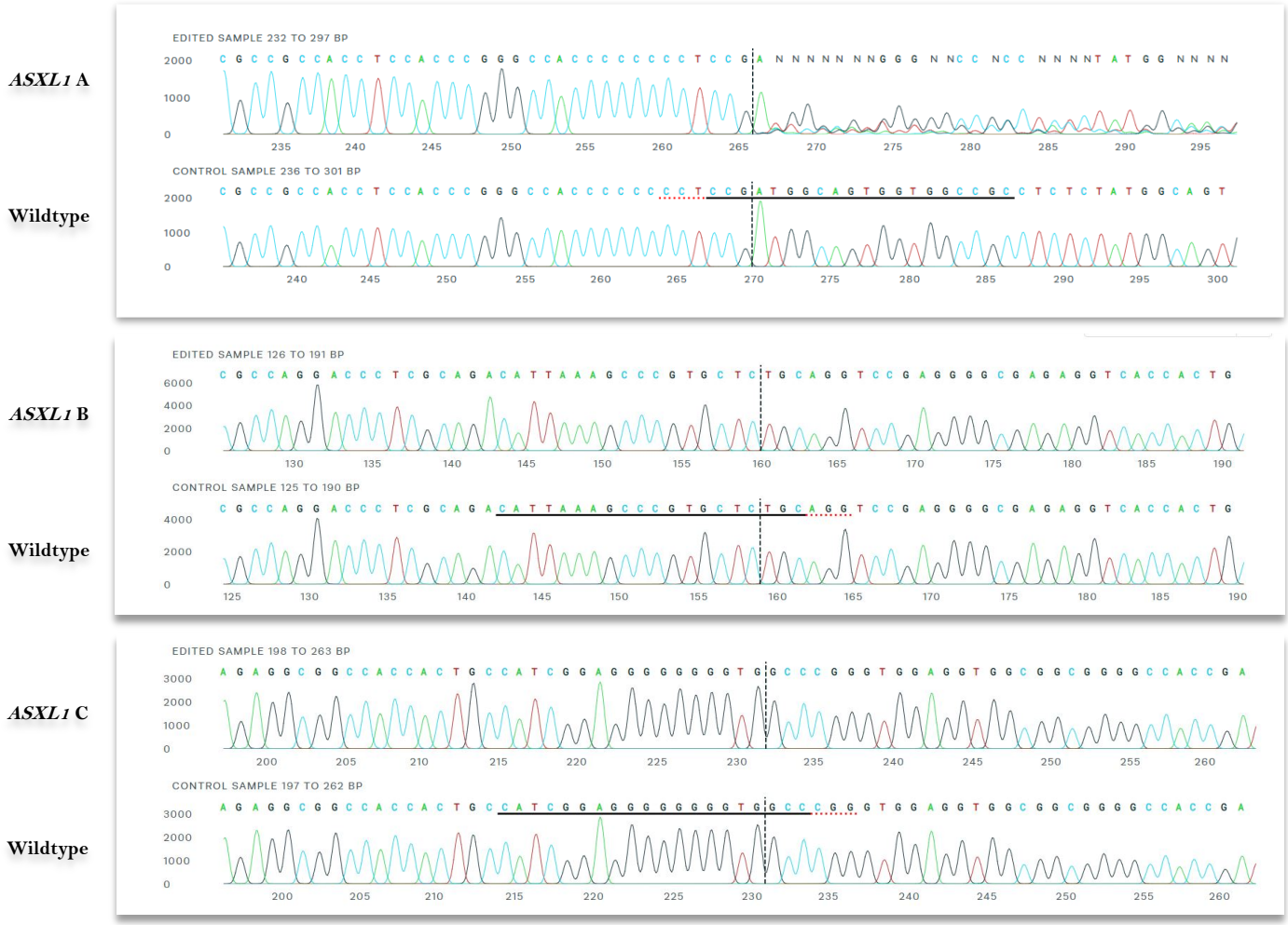


Figure 25: Sanger sequence traces of ASXL1 A, B and C targeted sgRNA variant. Nucleotides are represented by different colour peaks. As can be observed, the sequence collapse for ASXL1 A is the most pronounced, correlating well with this sgRNA having the highest efficiency.

3.2.9: Generating erythroid and myeloid models of myeloid malignancy

The most efficient sgRNAs targeting *TET2* and *ASXL1* were used to attempt to knock out *TET2* and introduce *ASXL1* mutation in HUDEP2 and KBM5 cells The *TET2* targeted sgRNAs we decided to take forward were Downstream B and Upstream A, which both had an ICE score of 93 respectively.

The most efficient *ASXL1* targeted sgRNA variant was *ASXL1* A, which had an ICE score of 41, and was therefore taken forward. Since the KBM5 cells already contained an *ASXL1* G710X mutation (Valletta, Dolatshad et al. 2015), the p.G646Wfs*12 was only to be introduced into HUDEP2 cells. *TET2*, however, was to be knocked out in both HUDEP2 and KBM5 cells. Since the *TET2* targeted sgRNA efficiency scores were above the recommended threshold, we decided to use synthetic sgRNA, whereas for the *ASXL1* targeted sgRNA we decided to use a plasmid-based system as the predicted efficiency was lower. Using a plasmid-based system allowed for GFP sorting of the *ASXL1* transfected cell population, leading to a higher probability of isolated cells containing the CRISPR/Cas9 components.

Following transfection of our CRISPR/Cas9 cocktail into the respective cell line, the cultures were single-cell sorted into 96-well plates. Once individual colonies reached 50% confluency, DNA was extracted, and PCR analysis was conducted. The samples were then run using an agarose gel, along with Sanger sequencing to confirm the genotype. Following on, mRNA expression levels were investigated using qRT-PCR. Finally, an off-target analysis was conducted. The workflow can be observed in figure 26.

Pure *TET2* -/- single-cell HUDEP2 and KBM5 colonies can be observed on the agarose gels, identified by the 500bp bands, in figure 28. The *TET2* gene excision event can be observed in figure 27. Moreover, the Sanger sequencing trace in figure 29 shows successful introduction of *ASXL1* p.G646Wfs*12 mutation into HUDEP2 cells. Similarly, the Sanger sequence traces in figure 30 show successful heterozygous and homozygous G710X *ASXL1* mutation correction, in KBM5 cells (conducted previously in our laboratory) (Valletta, Dolatshad et al. 2015). The model variants of

HUDEP2 and KBM5 cells we genetically engineered using CRISPR/Cas9, to KO *TET2* and/or express *ASXL1* p.G646Wfs*12 or *ASXL1* G710X mutations, are summarised in table 25 and 26 respectively.

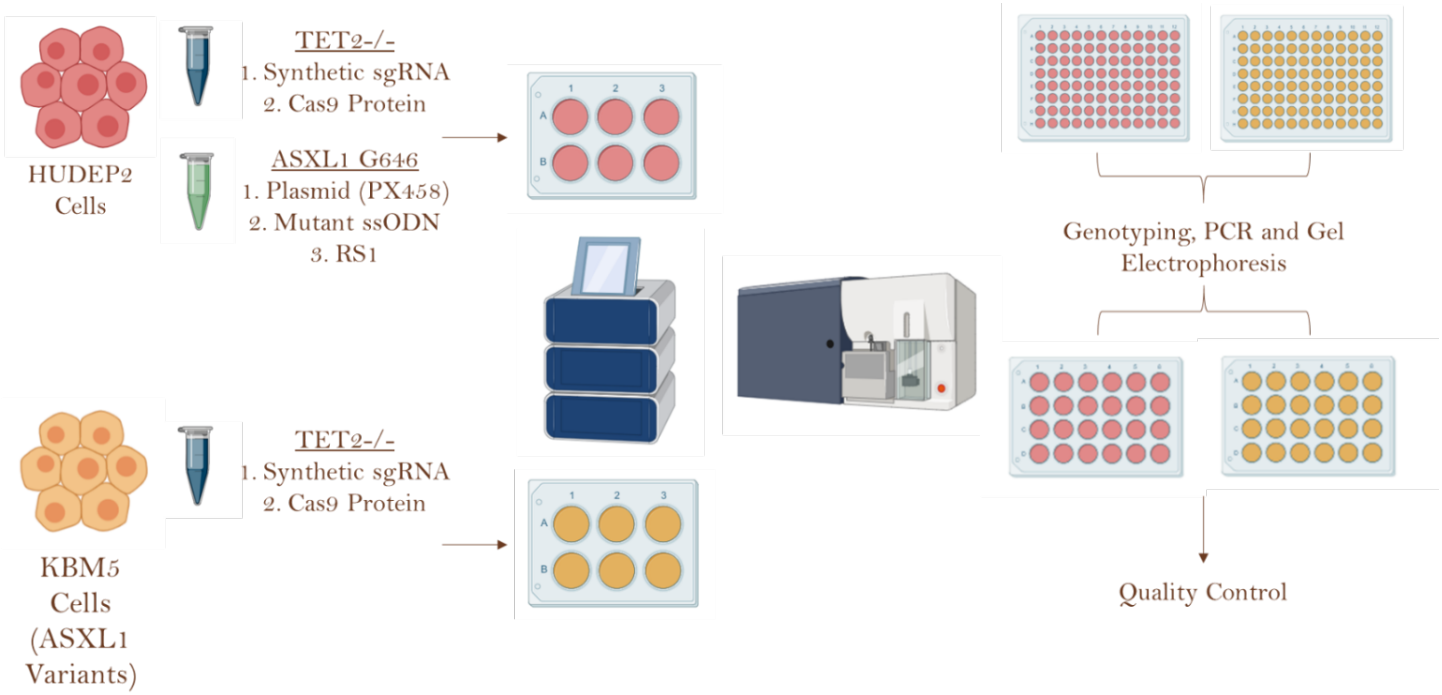


Figure 26: Experimental workflow for the engineering erythroid and myeloid models of myeloid malignancies in HUDEP2 and KBM5 cells. CRISPR/Cas9 cocktails were transfected into respective cell lines and subsequently single cell sorted into 96-well plates. Following DNA extraction, PCR analysis was conducted, and the samples were run on an agarose gel, in addition to Sanger sequencing to confirm the genotype.

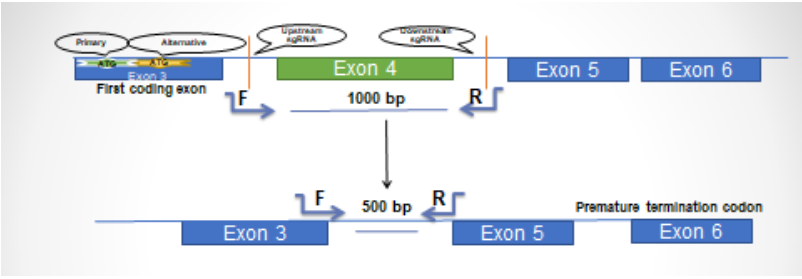


Figure 27: The TET2 gene excision event.
We decided to employ a targeting strategy that would cause a double-stranded cut downstream and upstream of exon 4, leading to excision of the exon and a frameshift, resulting in the generation of a premature stop codon, and because of the premature termination, the transcript would be expected to be subject to non-sense mediated decay in the cell, thereby recapitulating TET2 loss of function mutations in MDS patients.

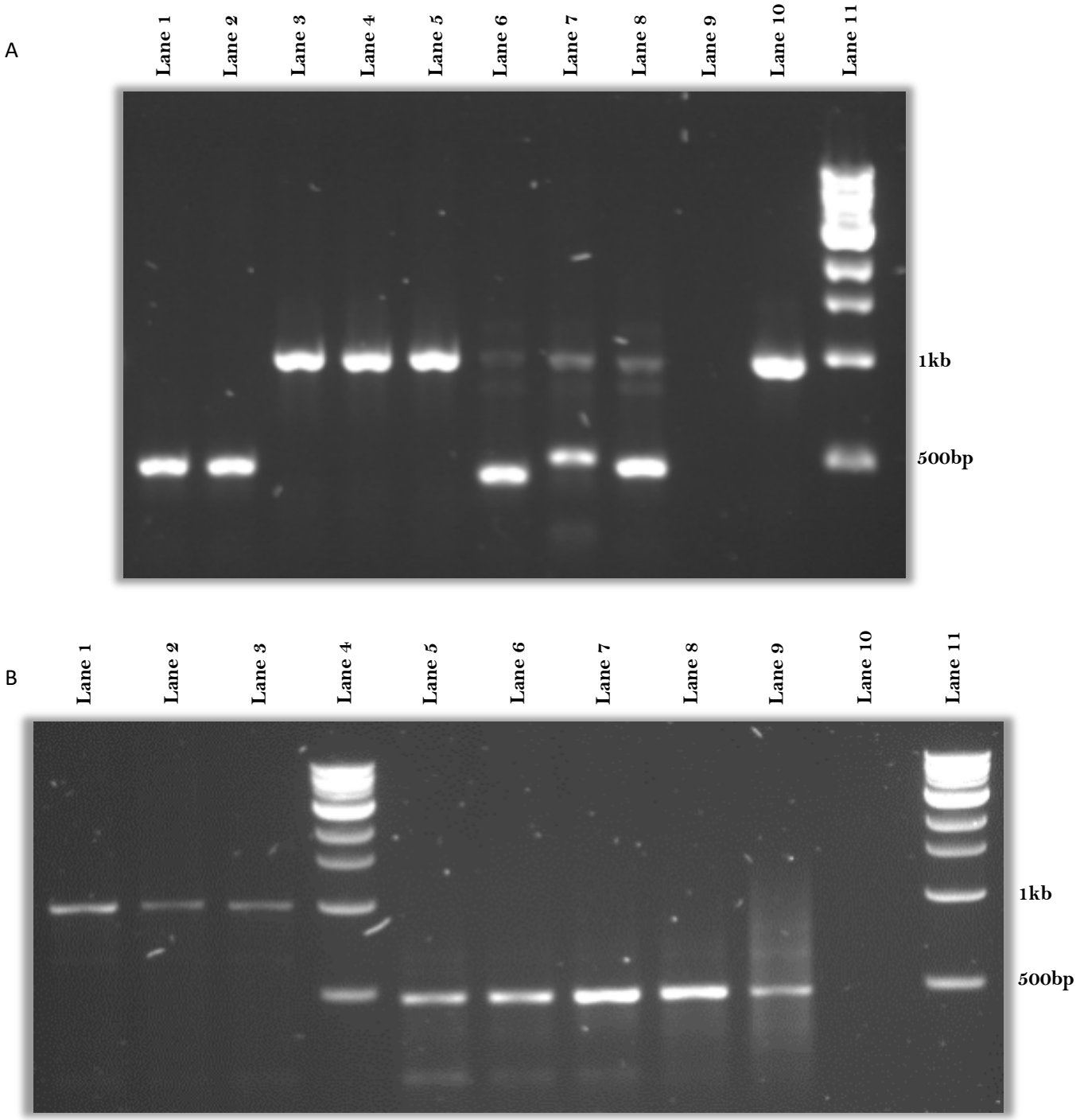


Figure 28: Agarose gels depicting HUDEP2 and KBM5 genotypes in single cell colonies.
A) HUDEP2 Cells: Lanes 1-2: TET-/-, Lanes 3-5: Wildtype, Lanes 6-8: TET2 -/+, Lane 9: Negative control, Lane 10: Positive control and Lane 11: 1kB ladder. The 1kb band is the product of two primers designed flanking exon 4. The 500bp band represents the product after excision of exon 4. B) KBM5 Cells: Lanes 1: Positive control, Lanes 2-3: Wildtype, Lane 4: 1kb ladder, Lanes 5-9: TET-/-, Lane 10: Negative control and Lane 11: 1kb ladder. The 1kb band is the product of two primers designed flanking exon 4. The 500bp band represents the product after excision of exon 4.

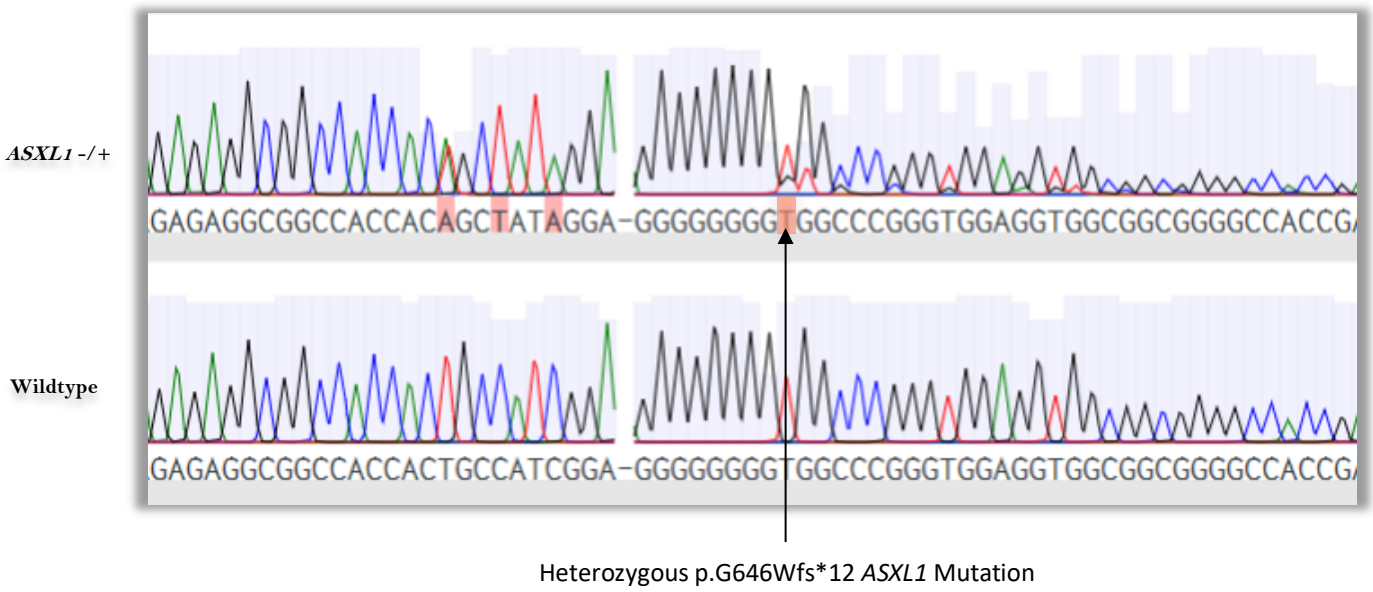


*Figure 29: A Sanger sequencing trace showing a successful p.G646Wfs*12 ASXL1 heterozygous mutation knock-in. The incorporation can be depicted by the arrow. Nucleotides are represented by different colours.*

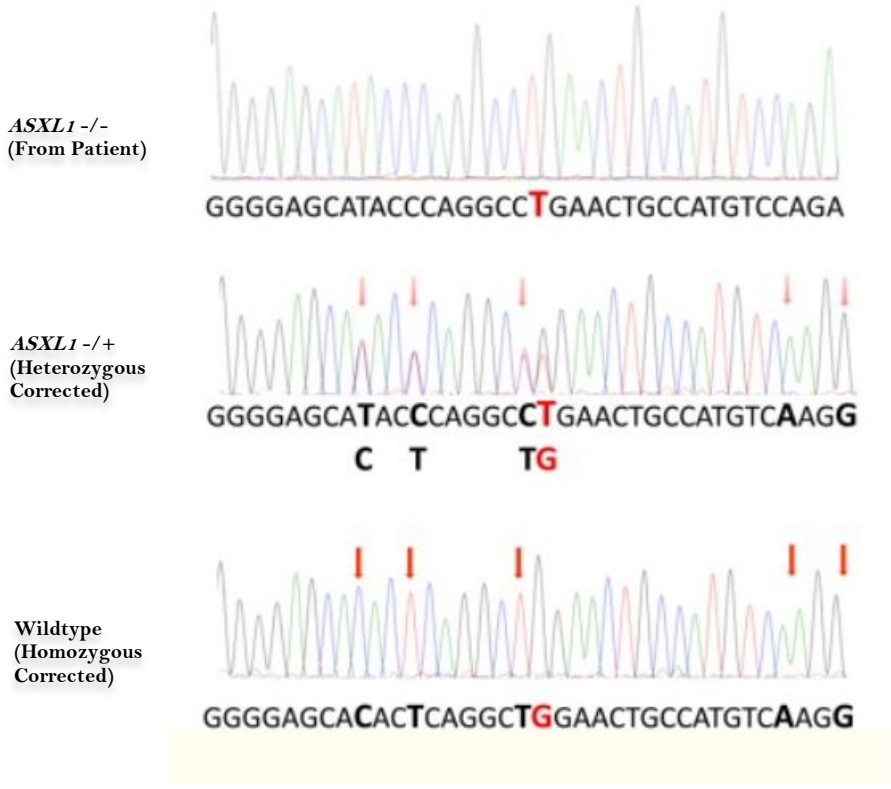


Figure 30: Sanger sequencing trace showing successful heterozygous and homozygous G710X ASXL1 mutation correction. Nucleotides are represented by different colours. These KBM5 mutant corrected cell lines were previously produced by colleagues in our laboratory (Valletta, Dolatshad et al. 2015). Copyright permission to reproduce this figure was obtained from Springer Nature.

HUDEP2 Cell Model	*TET2* Status	*ASXL1* p.G646Wfs*12 Status
Wildtype	Wildtype	Wildtype
ASXL1 -/+	Wildtype	Heterozygous
TET2 -/-	Homozygous	Wildtype
TET2 -/- *ASXL1* -/+	Homozygous	Heterozygous

Table 25: The HUDEP2 cell model variants engineered using CRISPR Cas9 technology.

KBM5 Cell Model	*TET2* Status	*ASXL1* G710X Status
Wildtype	Wildtype	Wildtype
ASXL1 -/+	Wildtype	Heterozygous
TET2 -/-	Homozygous	Wildtype
TET2 -/- *ASXL1* -/+	Homozygous	Heterozygous

Table 26: The KBM5 cell model variants engineered using CRISPR Cas9 technology.

3.2.10 Messenger RNA expression of *TET2* and *ASXL1* in models of myeloid malignancy

Following introduction of mutations most commonly associated with myeloid malignancy (*TET2* and *ASXL1*) in HUDEP2 and KBM5 cell lines, we decided to measure mRNA expression of *TET2* and *ASXL1* in each cell line model variant as a direct measure of gene expression, and with it, a direct read of mutation status.

TET2 mRNA expression fold change in WT (1.00 ± 0.00) and *ASXL1* -/+ (0.96 ± 0.03) HUDEP2 cell variants was significantly higher than in *TET2* -/- (0 ± 0) and *TET2* -/- *ASXL1* -/+ (0 ± 0) HUDEP2 cell variants (**** P≤0.0001). Furthermore, there was no significant difference in *TET2* mRNA expression fold change between WT (1.00 ± 0.00) and *ASXL1* -/+ (0.96 ± 0.03) HUDEP2 cell variants (P>0.05) (see figure 31). Moreover, there was no significant difference in *ASXL1* mRNA expression fold change

between WT (1.00 ± 0.00), *ASXL1* -/- (0.96 ± 0.03), *TET2* -/- (0.98 ± 0.01) and *TET2* -/- *ASXL1* -/+ (0.96 ± 0.02) HUDEP2 cell variants (P>0.05) (see figure 32).

In KBM5 cell models, *TET2* mRNA expression fold change in WT (1.00 ± 0.00) and *ASXL1* -/- (0.98 ± 0.02) cell variants was significantly higher than in *TET2* -/- (0 ± 0) and *TET2* -/- *ASXL1* -/+ (0 ± 0) cell variants (**** P≤0.0001). Furthermore, in the KBM5 cell models, there was no significant difference in *TET2* mRNA expression fold change between WT (1.00 ± 0.00) and *ASXL1* -/+ (0.98 ± 0.02) cell variants (P>0.05) (see figure 33). Additionally, in KBM5 cell models, there was no significant difference in *ASXL1* expression fold change between WT (1.00 ± 0.00), *ASXL1* -/- (0.97 ± 0.02), *TET2* -/- (0.97 ± 0.03) and *TET2* -/- *ASXL1* -/+ (0.96 ± 0.03) cell variants (P>0.05) (see figure 34).

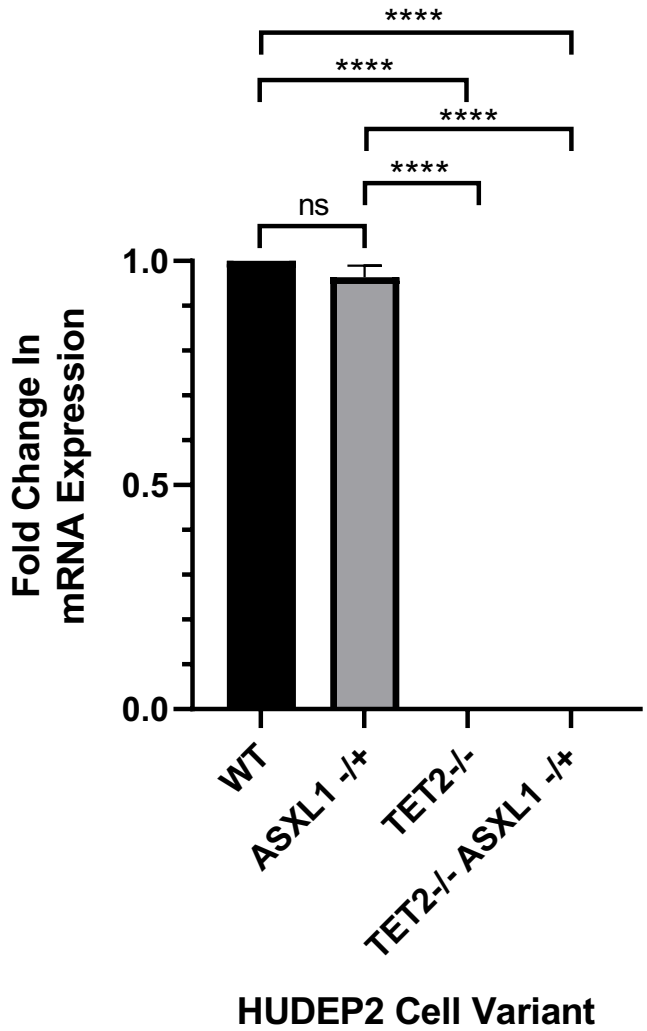


Figure 31: TET2 mRNA expression in the HUDEP2 cell line variants.
*Data are expressed as mean ± SD, n=3 biological replicates, (**** P≤0.0001), (ns = not significant, P>0.05). A one-way ANOVA test was used for statistical analysis, coupled with the Tukey's multiple comparison post hoc test. WT (wildtype), ASXL1 -/+ (heterozygous ASXL1 knockout), TET2 -/- (homozygous TET2 knockout), and TET2-/- ASXL1 -/+ (heterozygous ASXL1 knockout + homozygous TET2 knockout double mutant).*

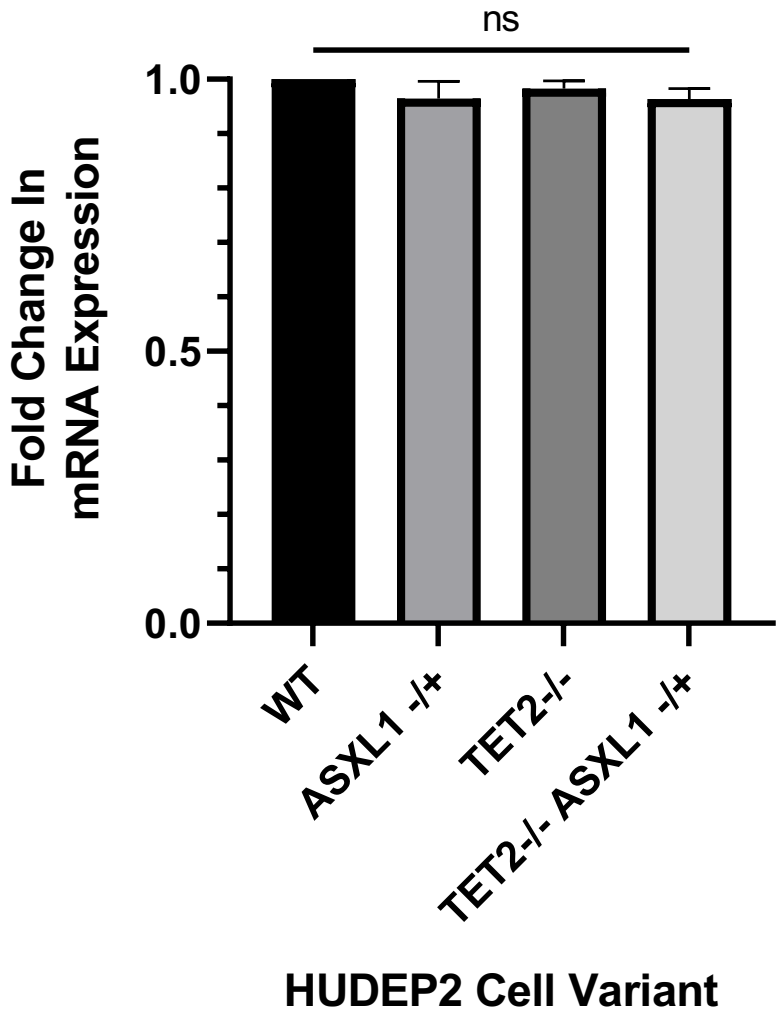


Figure 32: ASXL1 mRNA expression in the HUDEP2 cell line variants.
Data are expressed as mean ± SD, n=3 biological replicates, (ns = not significant, P>0.05). A one-way ANOVA test was used for statistical analysis, coupled with the Tukey's multiple comparison post hoc test. WT (wildtype), ASXL1 -/+ (heterozygous ASXL1 knockout), TET2 -/- (homozygous TET2 knockout), and TET2-/- ASXL1 -/+ (heterozygous ASXL1 knockout + homozygous TET2 knockout double mutant).

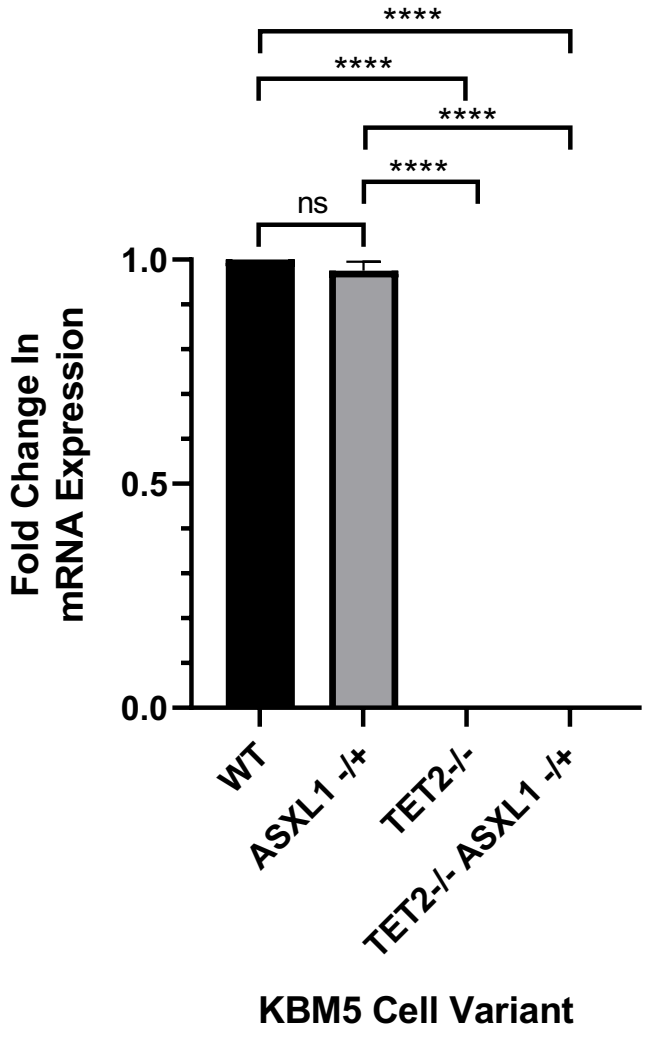


Figure 33: TET2 mRNA expression in the KBM5 cell line variants.
*Data are expressed as mean ± SD, n=3 biological replicates, (**** P≤0.0001), (ns = not significant, P>0.05). A one-way ANOVA test was used for statistical analysis, coupled with the Tukey's multiple comparison post hoc test. WT (wildtype), ASXL1 -/+ (heterozygous ASXL1 knockout), TET2 -/- (homozygous TET2 knockout), and TET2-/- ASXL1 -/+ (heterozygous ASXL1 knockout + homozygous TET2 knockout double mutant).*

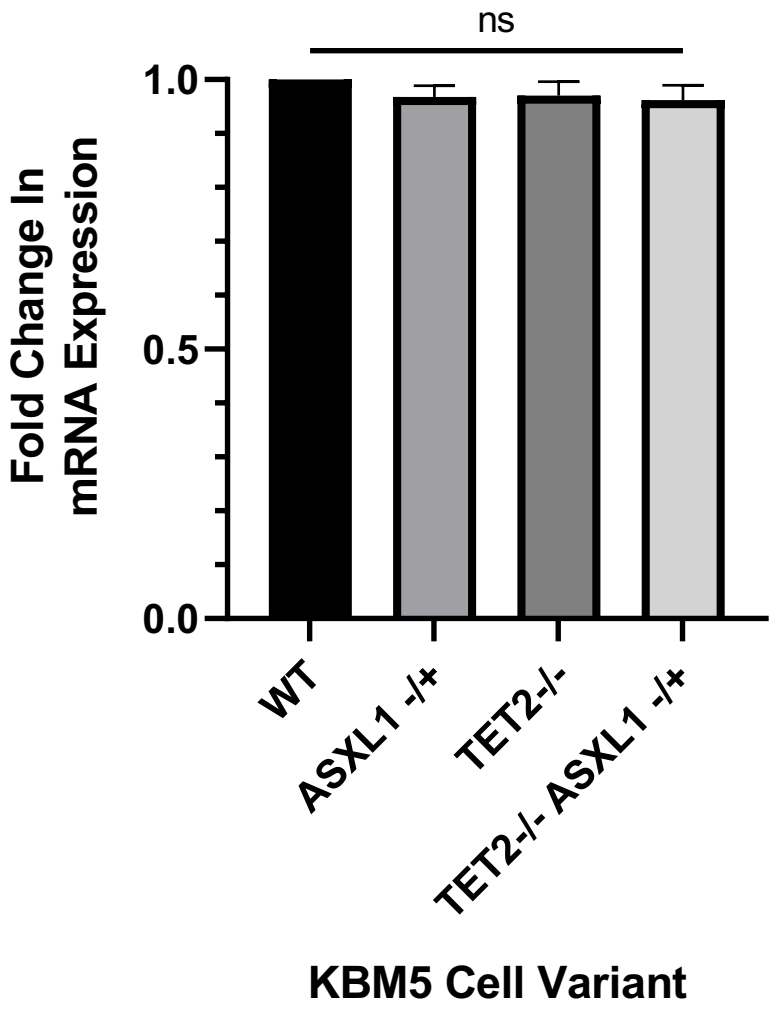


Figure 34: ASXL1 mRNA expression in the KBM5 cell line variants.
Data are expressed as mean ± SD, n=3 biological replicates, (ns = not significant, P>0.05). A one-way ANOVA test was used for statistical analysis, coupled with the Tukey's multiple comparison post hoc test. WT (wildtype), ASXL1 -/+ (heterozygous ASXL1 knockout), TET2 -/- (homozygous TET2 knockout), and TET2-/- ASXL1 -/+ (heterozygous ASXL1 knockout + homozygous TET2 knockout double mutant).

3.2.11 Off-target analysis of *TET2* and *ASXL1* targeted sgRNAs

In order to investigate potential off-target cleavage in our CRISPR/Cas9 system, using Sanger sequencing we analysed the top predicted exonic changes at off-target sites predicted by the CRISPOR platform we used to design each unique *TET2* or *ASXL1* targeted sgRNA. The off-target site for each sgRNA is predicted and refers to the potential predicted nonspecific and unintended mutations. No unintended sequence alterations were detected in all HUDEP2 and KBM5 cell clones screened, which expressed mutations in *TET2* and/or *ASXL1*. This suggests a high specificity of the CRISPR/Cas9 system

used in this thesis. The off-targets were analysed, and the gene candidates can be observed in tables 27, 28, 29 and 30.

Gene Name	Locus	Forward Primer	Reverse Primer
SLC22A10	chr11:63311147-63311169	AAGCCTTCATGAGCTGATTGGT	AGCCAGCCACATCCTAATTCTG
PGR	chr11:101031343-101031365	TGTGACTTCATGAGCCCTCC	CAACGGGGGTCTATGACCAG
ARHGAP11A	chr15:32638899-32638921	GCCAGTAGGTGTACCTTGTGTT	GCACTTTGAATTTCATGAGAGCCT
SNX24	chr5:123008869-123008891	CCTGGGTTTACTCAAGGGTGT	TTTAGTTCTTGCTTGTTTCCCAC

Table 27: Upstream A TET2 sgRNA (TATATGTACATGCTGGTTGT) off-target analysis data.

Gene Name	Locus	Forward Primer	Reverse Primer
AGAP2	chr12:57727510-57727532	GGGGAATTCAATCTGCGTGG	ACAGTAGCTGCTCGTACTTGG
PARP15	chr3:122635832-122635854	AGATTGGGCATGGTTCTACACA	TGACTAATCCTGCACGTCCC
TMEM27	chrX:15644995-15645017	GTCACCAACCGGTGATGGAG	TGGCTGATTGCACCTCAACA
MFI2	chr3:197024374-197024396	ACGAGAACGCTATGCACACA	GAGTACATCGCAGCCCATCA
PRKCE	chr2:46001449-46001471	ACCGTCGATGTGAGACCAAC	TGCTAGCCTGAAAAGGGCTC

Table 28: Downstream B TET2 sgRNA (TCTGGCATCCATCGCAAAGT) off-target analysis data.

Gene Name	Locus	Forward Primer	Reverse Primer
FAM53A	chr4:1640216-1640238	TCTGTGTTCCCCACGAAAGG	ATCGCAAACTCGAAACCCCT
DYRK1B	chr19:39833447-39833469	ACCAGCGACCAATGGTACAG	CCTTGGAGCTTCCCCCAAAA
MYLK	chr3:123752359-123752381	TTGCATAGTGGATTCTGGACC	AGGCCTTGGGGTAACTGAGA
PPP1R13L	chr19:45392194-45392216	ATGCCCTGGAAGTTGTGGAG	TACTTACCATGCTCTGCGGC
GRK4	chr4:2963902-2963924	CTAAGCCGTTAGCGCCGAG	CTTTCAGCAGCAGCGAGTTG

*Table 29: ASXL1 p.G646Wfs*12 sgRNA (GCGGCCACCACTGCCATCGG) off-targets analysis data.*

Gene Name	Locus	Forward Primer	Reverse Primer
HK2	chr2:+75118439	GCTCATCACTGCTGTGGCCT	GGAATCCACAGGGCAAATGC
CAPN3	chr15:−42703613	CACCTGATAATCTCCAGTCTGCTCC	TTCAGGCATACATGGTGAGCTG
AQP10	chr1:−154295508	TGAGATGGAGTTTCACTCTGTTGCC	CGAGGAACAGGACCAAAGTAGTGTC
PRPH2	chr6:+42665654	CTACAGACGTCGCTGGATGGTG	CCCCTGCCTTGAAATTCTCTTG
PIN4	chrX:+71406302	GTGGCCCAAGATAATTCTTCCA	TGCGCCCAGCTACAATTTTG

Table 30: ASXL1 G710X sgRNA (TCTAAATGGGGAGCATACCC) off-targets analysis data
Taken from: (Valletta, Dolatshad et al. 2015). Copyright permission to reproduce these data was obtained from Springer Nature.

3.3: Discussion

The aim of this study was to introduce common MDS-associated *TET2* and *ASXL1* gene mutations individually and in combination, in erythroid (HUDEP2 cells) and myeloid (KBM5 cells) cell lines. In HUDEP2 cells we aimed to KO the *TET2* gene, a tumour suppressor, which regulates myelopoiesis (Langemeijer, Kuiper et al. 2009). *TET2* is commonly mutated in MDS, and *TET2* mutations in these disorders are usually loss-of-function (Langemeijer, Kuiper et al. 2009). We effectively achieved a *TET2* KO in HUDEP2 cells, with several colonies expanded from single cell populations. In order to KO the *TET2* gene, we decided to target a large isoform conserved region of the *TET2* gene located in exon 4. Since the majority of protein coding transcripts of TET2 include exon 4, we hypothesised that by excising exon 4 using CRISPR, we could introduce a frameshift which may introduce a premature termination codon in all protein coding isoforms (see figure 12). The chosen strategy was to induce two distinct cuts using the CRISPR Cas9 system, one upstream of exon 4 and one downstream of exon 4. The use of the CRISPR Cas9 system to KO TET2 has been previously reported in the literature. In order to assess the function of *TET2* in breast cancer, Wang and colleagues designed a strategy with Exon 3 as the target for their CRISPR Cas9 system (Wang, Ozark et al. 2018). Furthermore, CRISPR Cas9

was used by Zhang and colleagues to KO *TET2* in CNE1 cells to investigate its role in nasopharyngeal carcinoma (Zhang, Yang et al. 2020).

The CRISPOR platform was used to design potential upstream and downstream TET2 targeted sgRNA sequences around the exon 4 region. Once the predictive software was run, the most favourable upstream and downstream TET2 targeted sgRNA sequences were selected (see table 23). We then decided to clone the top 3 downstream and top 3 upstream TET2 targeted sgRNAs into the PX458 plasmid construct. The PX458 plasmid is a recently engineered plasmid construct, which is used to deliver a CRISPR/Cas9 payload to target cells (Ran, Hsu et al. 2013, Dara, Dianatpour et al. 2019). The advantage of using the PX458 system is that it already contains Cas9 and GFP, aiding in cell selection and sorting. Initially we cloned the upstream and downstream *TET2* targeted sgRNAs into the PX458 plasmid. We used Sanger sequencing to confirm incorporation of the *TET2* targeted sgRNA sequences, which demonstrated successful incorporation for both upstream and downstream *TET2* targeted sgRNA (see figures 13 and 14). These constructs were then transfected into HUDEP2 cells as a prerequisite to select the most efficient pair of upstream and downstream *TET2* targeted sgRNAs to be taken forward to subsequent experiments. The afore-mentioned PX458 plasmid was used, since it expressed GFP, allowing the HUDEP2 cells containing the sgRNA sequence to be sorted according to the GFP tag. Following transfection, the percentage of bulk sorted live GFP+ cells was ~100% for all upstream and downstream *TET2* targeted sgRNAs tested (see figures 15 and 16). Following on, we used the ICE software to elucidate actual efficiency of each upstream and downstream *TET2* targeted sgRNA sequence (see figures 17 and 19). We then deduced that the upstream A and downstream B *TET2* targeted sgRNA variants demonstrated the highest efficiency and were selected to be taken forward to further experiments.

Next using CRISPR Cas9 technology, we aimed to introduce a heterozygous *ASXL1* p.G646Wfs*12 mutation in HUDEP2 cells, since this is one of the most common *ASXL1* mutation in MDS (Abdel-Wahab, Gao et al. 2013). We successfully introduced the *ASXL1* p.G646Wfs*12 mutation in HUDEP2 cells, in several colonies expanded from several single cell populations. Various research

groups have used the CRISPR Cas9 technology to introduce mutations in *ASXL1*. In order to study cell cycle progression and proliferation in U937 cells, Wu and colleagues used CRISPR Cas9 technology to generate heterozygous and homozygous *ASXL1* mutations (Wu, Zhao et al. 2018). They observed a defect in monocyte/macrophage differentiation in the ASXL1 mutated cells. Following transcriptome analysis, it was revealed that the ASXL1 mutation altered myeloid differentiation through a cytochrome B-245 β chain and C-type lectin domain family 5, member A mediated mechanism (Wu, Zhao et al. 2018).

Furthermore, in order to study malignant myeloid disease, Fang and colleagues used a Zebra fish model in which they knocked out *ASXL1* using CRISPR Cas9 technology (Fang, Xu et al. 2021). They were able to generate a C-terminal truncated ASXL1 mutant in zebrafish which mimics human myeloid malignancy (Fang, Xu et al. 2021).

In order to design *ASXL1* G646 targeted sgRNAs we selected sequences close to the G646 site, and used the CRISPOR online platform to design *ASXL1* G646 sgRNA constructs. We then selected the 3 best hits, based on predicted efficiency (see table 24). Following on, the *ASXL1* constructs were introduced into HUDEP2 cells, sorted for GFP and Sanger sequenced. ICE analysis revealed one active guide around *ASXL1* and with a much lower efficiency of 40% (see figure 24). The mutation and DNA break site is a G rich region and due of the high binding energy, we expected the low cutting at this site. Hence, to increase our chances of finding a clone with HDR, we attempted to edit *ASXL1* using the PX458 plasmid since it carries a selection marker in the form of GFP, and cells can be sorted based on plasmid uptake.

Heterozygous corrected *ASXL1* G710X KBM5 cells (*ASXL1* -/+) and homozygous corrected *ASXL1* G710X KBM5 cells (wildtype) were previously engineered in our laboratory (Valletta, Dolatshad et al. 2015). Using the optimal sgRNA constructs for both *TET2* and *ASXL1*, we successfully engineered erythroid and myeloid models of myeloid malignancy, in HUDEP2 cells and KBM5 cells respectively, which contained individual or co-mutations of *TET2* and *ASXL1*. An overview of the HUDEP2 cell model

variants can be observed in table 25, by contrast the KBM5 cell model variants can be observed in table 26.

We used qRT-PCR to quantify *TET2* and *ASXL1* gene expression, in both HUDEP2 and KBM5 cells for all mutation variants. *TET2* expression fold change was significantly reduced in *TET2* -/- and *TET2* -/- *ASXL1* -/+ mutant variants in both HUDEP2 (see figure 31) and KBM5 (see figure 33) cell types compared to wildtype cells. The expression fold change of *ASXL1* in all mutant variants was comparable to the expression of *ASXL1* in wildtype cells in both HUDEP2 (see figure 32) and KBM5 (see figure 34) cell variants, with no significant difference in expression fold change observed. For *TET2* these findings were expected since the *TET2* mutation was homozygous, and the KO creates a frameshift leading to emergence of a stop codon and nonsense-mediated decay. We would have expected the *ASXL1* heterozygous mutations to cause a reduction in mRNA expression, since one allele was mutated, but his was not observed in our clones possibly due to ASXL1 escaping nonsense mediated decay. It has been reported that heterozygous mutations of *ASXL1* in exon 12 may be gain of function, even if *ASXL1* mRNA expression in the mutant cell variants are comparable to wildtype (Yang, Kurtenbach et al. 2018).

Following on, we conducted off-target analysis to investigate whether the sgRNAs target any other genes, which may share some sequence homology with the *TET2* or *ASXL1* targeted sgRNAs, in HUDEP2 and KBM5 cells. Off-target effects can result in the introduction of mutations at sites other than the intended on-target sites. If there is a mutation at a target site other than the original locus, then any cellular phenotypes linked to the intended mutation may be confounded by the off-target effect (Zhang, Tee et al. 2015). We did not find sequence alterations in any of the off-target sites examined for the sgRNAs targeting *TET2* and *ASXL1* in our HUDEP2 and KBM5 cell models, suggesting that our CRISPR/Cas9-based gene editing was highly specific for the genes of interest. A limitation to using web-based algorisms to predict off-targets is that the algorithm assumes the off-target sequence is closely related to the intended on-target sequence, which may exclude off-target sequences with less sequence homology. The use of whole genome sequencing would allow a comprehensive

evaluation of off-target effects (Zhang, Tee et al. 2015), however the higher cost of this approach precluded its use in our study.

In this chapter we described the engineering of novel models of myeloid malignancy using CRISPR Cas9. These models may be valuable for the study of the mechanisms underlying disease progression and provide a platform that could be used for drug screening to identify new therapeutic agents. The cell line models described in this chapter were then taken forward to functional experiments, described in the next chapter.

Chapter 4
The Impact of *TET2* and *ASXL1* Mutations on Cellular Function

4.1: Introduction

It is recognised that *TET2* has an important role in basic cellular functions such as proliferation (Wang, Xiao et al. 2015), the cell cycle (Cimmino, Dolgalev et al. 2017), apoptosis (Cai, Kotzin et al. 2018) and differentiation (Scopim-Ribeiro, Machado-Neto et al. 2017).

It was reported by Moran-Crusio and colleagues, that in mice, TET2 loss leads to myeloproliferation along with splenomegaly, monocytosis, and extramedullary haematopoiesis (Moran-Crusio, Reavie et al. 2011). A study was conducted by Wang et al. (Wang, Miao et al. 2020), where they examined the relationship between *TET2* mutation status, and the prevalence of myelodysplastic syndrome. Wang and colleague (Wang, Miao et al. 2020) constructed 214 *TET2* -/- transgenic mice. It was found that 93% of the *TET2* -/- transgenic mice went on to develop myeloid malignancies (Wang, Miao et al. 2020). A study conducted by Muto and colleagues (Muto, Sashida et al. 2013) demonstrated that *TET2* depletion in mice caused development of MDS/MPN, which had myeloproliferative characteristics; these included increased apoptosis, dysplasia in myeloid cells and enhanced HSC repopulating capacity.

It is also recognised that *ASXL1* has an important role in key cellular functions such as: proliferation (Valletta, Dolatshad et al. 2015), cell cycle (Asada, Goyama et al. 2018), apoptosis (Wang, Li et al. 2014) and differentiation (Inoue, Kitaura et al. 2013). The *ASXL1* mutations are primarily nonsense and frameshift mutations, located in exon 12 before the c-terminus homeo finger region

(Carbuccia, Murati et al. 2009). As a result, most *ASXL1* mutations observed in myeloid malignancies in fact generate stable but truncated proteins, giving rise to functional disparities in the cell (Inoue, Kitaura et al. 2013). A study conducted by Abdel-Wahab and colleagues, demonstrated that ASXL1 associates with the Polycomb repressive complex 2 (PRC2), and that ASXL1 loss collaborates with NRASG12D, to promote myeloid leukemogenesis, *in vivo* (Abdel-Wahab, Adli et al. 2012).

It was shown by Wang and colleagues (Wang, Li et al. 2014) that *ASXL1* (+/-) mice went on to develop features of MDS including dysplastic neutrophils and multi lineage cytopenia; furthermore they had a medium survival rate of up to 42 days. *ASXL1* +/- and *ASXL1* -/- cells also demonstrated an increase in apoptosis. These findings are consistent with the presentation of MDS in humans (Mangaonkar, Gangat et al. 2018).

Gjini and colleagues generated ASXL1 heterozygous and null zebra fish to investigate myeloid cell transformation. They demonstrated that loss of ASXL1 leads to apoptosis in newly formed HSCs (Gjini, Jing et al. 2019). Moreover, approximately half of the ASXL1 heterozygous zebra fish went on to develop myeloproliferative neoplasms by 5 months of age. They also showed that heterozygous ASXL1 combined with heterozygous loss of TET2 lead to a more pronounced myeloproliferative neoplasm phenotype. Furthermore, combination of heterozygous ASXL1 with complete loss of TET2 led to AML (Gjini, Jing et al. 2019).

The effects of TET2 and ASXL1 have been evaluated individually but not in combination in mice or in erythroid and myeloid cell lines. We therefore decided to study the effect of *TET2* and *ASXL1* mutations, individually and in combination, on proliferation, the cell cycle, apoptosis and differentiation in the HUDEP2 and KBM5 cell lines.

4.2: Results

4.2.1: The effect of *TET2* and *ASXL1* mutations on the proliferation in HUDEP2 cells

Firstly, we decided to test the functional consequence of *TET2* and *ASXL1* mutations on proliferation in HUDEP2 cells. Proliferation was measured using the resazurin assay, coupled with the SpectraMax i3x Microplate Reader, in which fluorescence is proportional to cell density.

On day 1, although not significant, there was a mild increase in fluorescence in *ASXL1* -/+ (89.61 ± 6.38), *TET2* -/- (62.75 ± 1.60) and *TET2* -/- *ASXL1* -/+ (86.22 ± 0.61) HUDEP2 cell variants, compared to wildtype (55.57 ± 1.17) ($P>0.05$). There was no significant difference between mutant groups, although the HUDEP2 *ASXL1* -/+ (89.61 ± 6.38) cell variant showed the highest fluorescence ($P>0.05$).

On day 2, there was a significant increase in fluorescence in *ASXL1* -/+ (137.38 ± 2.82) ($P\leq0.05$), *TET2* -/- (154.74 ± 6.66) ($P\leq0.01$) and *TET2* -/- *ASXL1* -/+ (187.14 ± 9.47) ($P\leq0.001$) HUDEP2 cell variants compared to wildtype (73.46 ± 1.33). There was no significant difference between mutant groups, although the *TET2* -/- *ASXL1* -/+ (187.14 ± 9.47) HUDEP2 cell variant showed the highest fluorescence ($P>0.05$).

On day 3, there was a significant increase in fluorescence in *ASXL1* -/+ (256.98 ± 6.01), *TET2* -/- (270.03 ± 6.35) and *TET2* -/- *ASXL1* -/+ (292.93 ± 16.79) HUDEP2 cell variants compared to wildtype (84.68 ± 6.78) ($P\leq0.0001$). There was no significant difference between mutant groups, although the *TET2* -/- *ASXL1* -/+ (292.93 ± 16.79) HUDEP2 cell variant showed the highest fluorescence ($P>0.05$).

On day 4, there was a significant increase in fluorescence in *ASXL1* -/+ (475.99 ± 51.05), *TET2* -/- (568.85 ± 70.30) and *TET2* -/- *ASXL1* -/+ (604.59 ± 63.40) HUDEP2 cell variants compared to wildtype (301.65 ± 67.00) ($P\leq0.0001$). In addition, within the mutant group, there was a significant increase in fluorescence in *TET2* -/- (568.85 ± 70.30) ($P\leq0.05$) and *TET2* -/- *ASXL1* -/+ (604.59 ± 63.40) ($P\leq0.001$) HUDEP2 cell variants, compared to the *ASXL1* -/+ (475.99 ± 51.05) HUDEP2 cell variant. The highest fluorescence was observed in the *TET2* -/- *ASXL1* -/+ (604.59 ± 63.40) HUDEP2 cell variant (see figure 35).

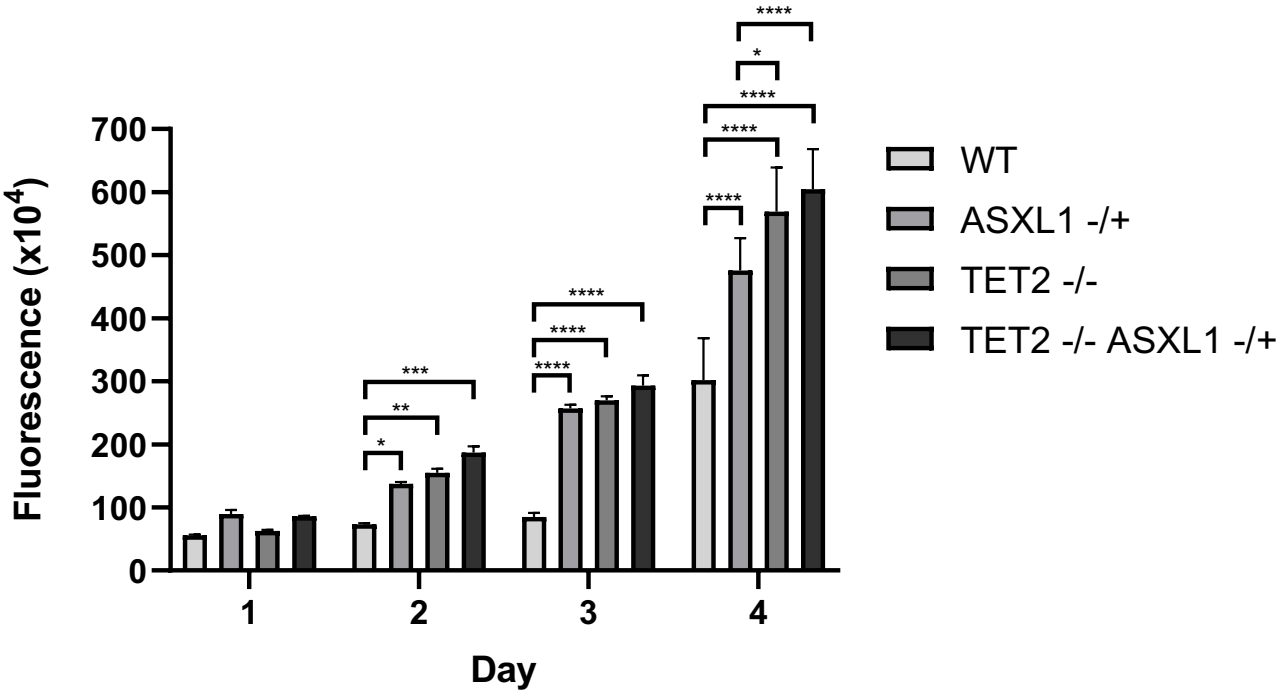


Figure 35: The impact of TET2 and ASXL1 mutations on proliferation of HUDEP2 cells over a 4-day period, shown as fluorescence (x10⁴). Wildtype cells (WT) are compared with ASXL1-/-, TET2-/- and TET2-/- ASXL1-/+ cell lines. Data are expressed as mean ± SD, n=3 biological replicates, (P≤0.05), (**P≤0.01), (*** P≤0.001), (**** P≤0.0001). A two-way ANOVA test was used for statistical analysis, coupled with the Tukey's multiple comparison post hoc test.*

4.2.2: The effect of *TET2* and *ASXL1* mutations on the proliferation in KBM5 cells

Next, we decided to test the functional consequence of *TET2* and *ASXL1* mutations on proliferation in our KBM5 cell models. There was a significant increase in fluorescence in the *TET2* -/- *ASXL1* -/+ (534.46 ± 13.03) KBM5 cell variant, compared to wildtype (200.82 ± 31.72) (P≤0.05). Although not significant, the *ASXL1* -/+ (316.45 ± 31.03) and *TET2* -/- (320.22 ± 3.34) KBM5 cell variants showed a mild increase in fluorescence compared to wildtype (200.82 ± 31.72) (P>0.05). There was no significant difference between all mutant groups, although the *TET2* -/- *ASXL1* -/+ (534.46 ± 13.03) KBM5 cell variant showed the highest fluorescence (P>0.05).

On day 2, there was a significant increase in fluorescence in *TET2* -/- *ASXL1* -/+ (1277.23 ± 13.87) KBM5 cell variants compared to wildtype (654.14 ± 7.91) (P≤0.0001). Although not significant, the *ASXL1* -/+ (751.84 ± 9.85) and *TET2* -/- (911.94 ± 16.49) KBM5 cell variants showed a mild increase in fluorescence compared to wildtype (654.14 ± 7.91) (P>0.05). In addition, within the mutant group, there was a significant increase in fluorescence in *TET2* -/- *ASXL1* -/+ (1277.23 ± 13.87) compared with

ASXL1 -/+ (751.84 ± 9.85) (P≤0.0001) and *TET2* -/- (911.94 ± 16.49) (P≤0.01) KBM5 cell variants. The highest fluorescence was observed in the *TET2* -/- *ASXL1* -/+ (1277.23 ± 13.87) KBM5 cell variant.

On day 3, although not significant, the fluorescence in the *ASXL1* -/+ (1256.68 ± 34.72) and *TET2* -/- (1269.40 ± 58.52) KBM5 variant was comparable to wildtype (1234.29 ± 34.40) (P>0.05). In addition, although not significant, the *TET2* -/- *ASXL1* -/+ (1497.42 ± 22.31) KBM5 cell variant showed a mild increase in fluorescence compared to wildtype (1234.29 ± 34.40) (P>0.05). There was no significant difference between all mutant groups, although the *TET2* -/- *ASXL1* -/+ (1497.42 ± 22.31) KBM5 cell variant showed the highest fluorescence (P>0.05).

On day 4, there was a significant increase in fluorescence in *ASXL1* -/+ (2369.87 ± 272.70), *TET2* -/- (2392.73 ± 81.96) and *TET2* -/- *ASXL1* -/+ (2840.28 ± 327.33) KBM5 cell variants compared to wildtype (1754.55 ± 230.95) (P≤0.0001). Between mutant groups, there was a significant increase in fluorescence in the *TET2* -/- *ASXL1* -/+ (2840.28 ± 327.33) KBM5 cell variant compared to *ASXL1* -/+ (2369.87 ± 272.70) and *TET2* -/- (2392.73 ± 81.96) KBM5 cell variants (P≤0.001). There was no significant difference in the level of fluorescence between *ASXL1* -/+ (2369.87 ± 272.70) and *TET2* -/- (2392.73 ± 81.96) KBM5 cell variants (P>0.05). The highest fluorescence was observed in the *TET2* -/- *ASXL1* -/+ (2840.28 ± 327.33) KBM5 cell variant (see figure 36).

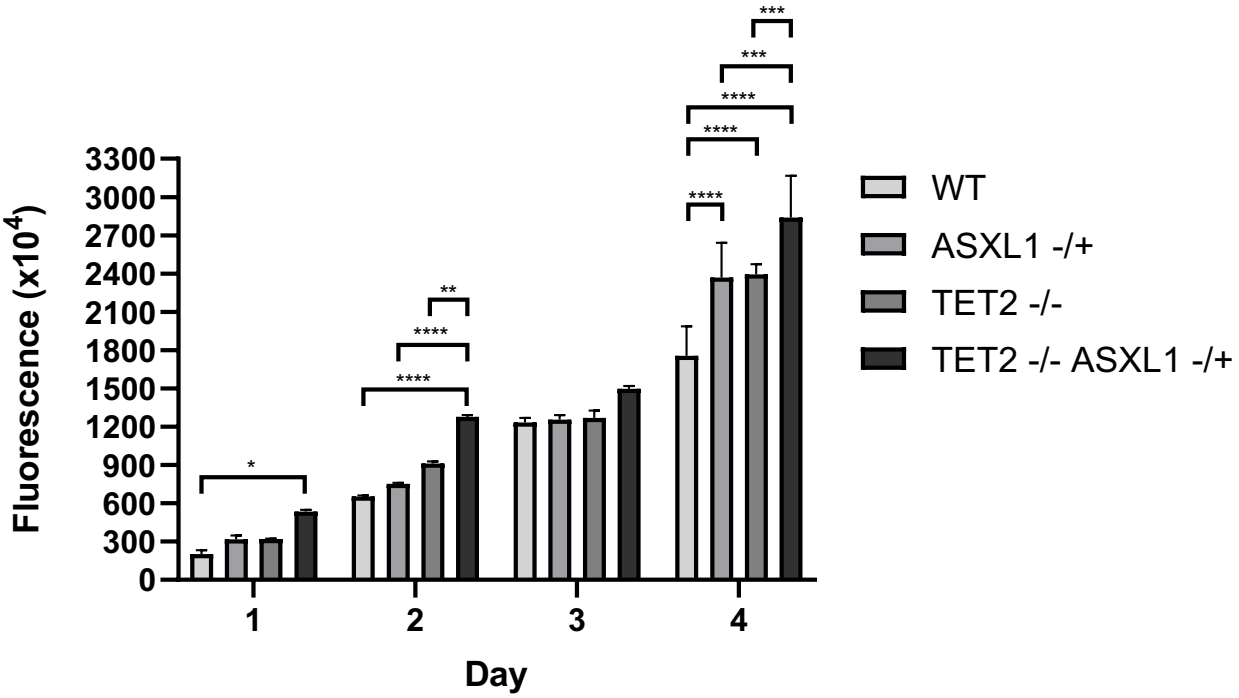


Figure 36: The impact of TET2 and ASXL1 mutations on proliferation of KBM5 cells over a 4-day period, shown as fluorescence (x10⁴). Wildtype cells (WT) are compared with ASXL1-/-, TET2-/- and TET2-/- ASXL1-/+ cell lines. Data are expressed as mean ± SD, n=3 biological replicates, (P≤0.05), (**P≤0.01), (*** P≤0.001), (**** P≤0.0001). A two-way ANOVA test was used for statistical analysis, coupled with the Tukey's multiple comparison post hoc test.*

4.2.3: The effect of *TET2* and *ASXL1* mutations on the cell cycle in HUDEP2 cells

We decided to assess the impact of *TET2* and *ASXL1* mutations on the cell cycle in HUDEP2 cell variants. There was a significant decrease in the percentage of the *TET2* -/- *ASXL1* -/+ (36.47 ± 0.42%) HUDEP2 cell variant in G0-G1, compared to wildtype (40.47 ± 1.22%) (P≤0.01). Although not significant, there was also a decrease in the percentage of *ASXL1* -/+ (38.90 ± 0.99%) and *TET2* -/- (37.85 ± 0.64%) HUDEP2 cell variants in G0-G1, compared to wildtype (40.47 ± 1.22%) (P>0.05). There was no significant difference between all mutant groups, although the *TET2* -/- *ASXL1* -/+ (36.47 ± 0.42%) HUDEP2 cell variant had the fewest percentage of cells in G0-G1 (P>0.05) (see figure 37).

There was a significant increase in the percentage of the *TET2* -/- *ASXL1* -/+ (24.80 ± 0.26%) HUDEP2 cell variant in S-Phase, compared to wildtype (20.73 ± 0.15%) (P≤0.05). Although not significant, there was also an increase in the percentage of *ASXL1* -/+ (22.20 ± 0.36%) and *TET2* -/- (23.35 ± 0.35%) HUDEP2 cell variants in S-Phase, compared to wildtype (20.73 ± 0.15%) (P>0.05). There

was no significant difference between all mutant groups, although the *TET2* -/- *ASXL1* -/+ (24.80 ± 0.26%) HUDEP2 cell variant had the highest percentage of cells in S-Phase (P>0.05) (see figure 38).

There was a significant increase in the percentage of *TET2* -/- (19.0 ± 0.14%) (P≤0.05) and *TET2* -/- *ASXL1* -/+ (20.33 ± 0.06%) (P≤0.01) HUDEP2 cell variants in G2/M, compared to wildtype (16.77 ± 0.57%). Although not significant, there was also an increase in the percentage of the *ASXL1* -/+ (18.27 ± 0.15%) HUDEP2 cell variant in G2/M, compared to wildtype (16.77 ± 0.57%) (P>0.05). There was no significant difference between all mutant groups, although the *TET2* -/- *ASXL1* -/+ (20.33 ± 0.06%) HUDEP2 cell variant had the highest percentage of cells in G2/M (P>0.05) (see figure 39).

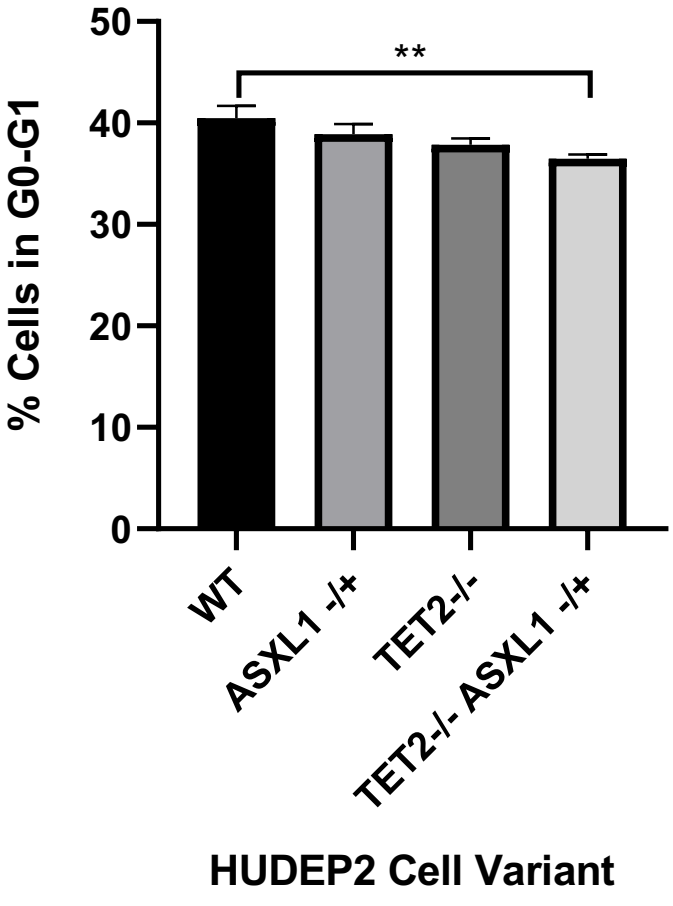


Figure 37: The impact of TET2 and ASXL1 mutations on G0-G1 of the cell cycle in HUDEP2 cells.
*Wildtype cells (WT) are compared with ASXL1-/-, TET2-/- and TET2-/- ASXL1-/+ cell lines. Data are expressed as mean ± SD, n=3 biological replicates, (**P≤0.01). A one-way ANOVA test was used for statistical analysis, coupled with the Tukey's multiple comparison post hoc test.*

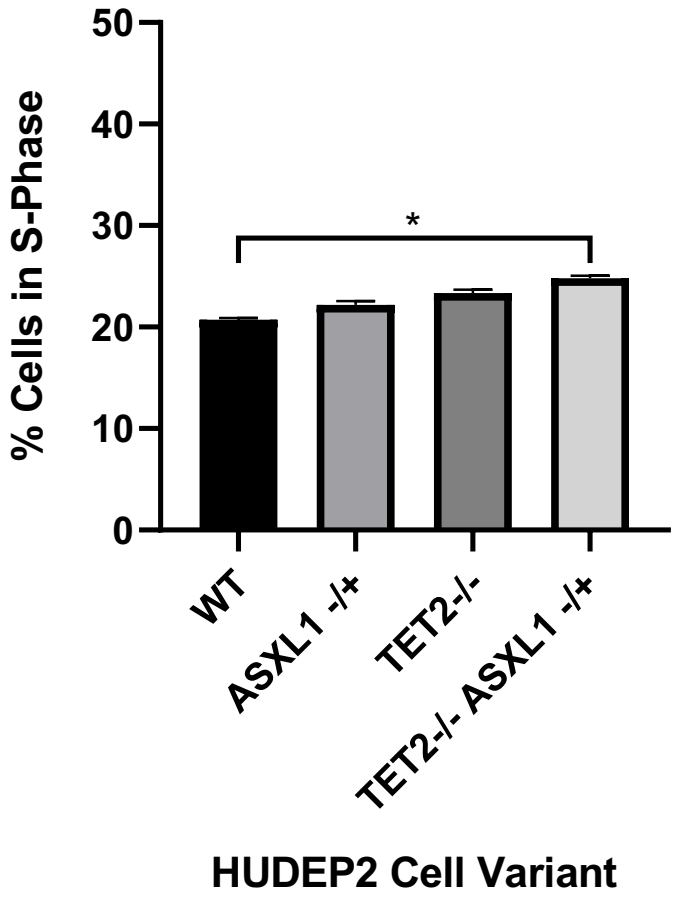


Figure 38: The impact of TET2 and ASXL1 mutations on S-Phase of the cell cycle in HUDEP2 cells.
Wildtype cells (WT) are compared with ASXL1-/-, TET2-/- and TET2-/- ASXL1-/+ cell lines. Data are expressed as mean ± SD, n=3 biological replicates, (P≤0.05). A one-way ANOVA test was used for statistical analysis, coupled with the Tukey's multiple comparison post hoc test.*

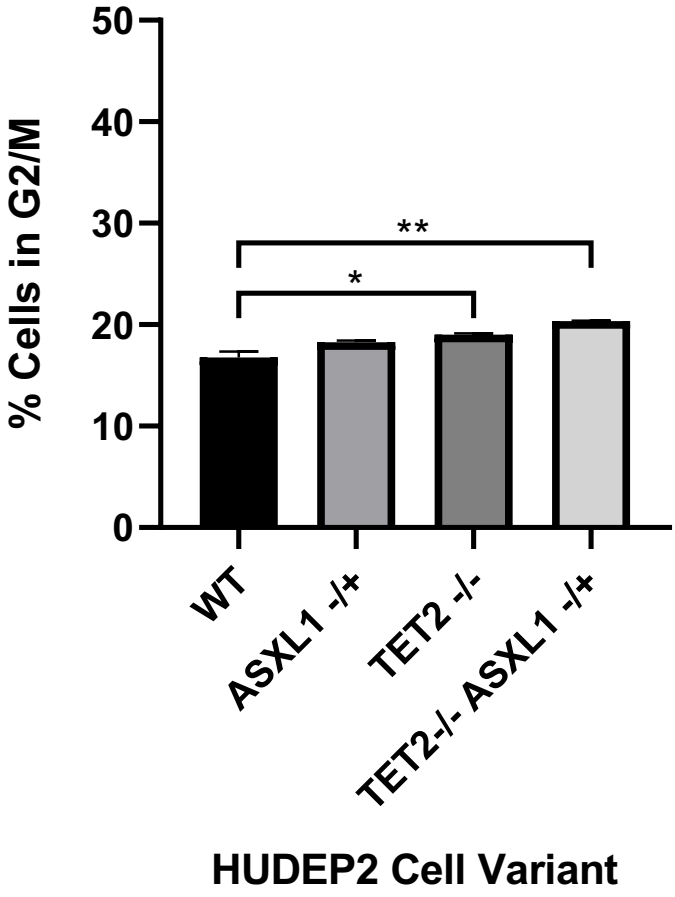


Figure 39: The impact of TET2 and ASXL1 mutations on G2/M of the cell cycle in HUDEP2 cells.
Wildtype cells (WT) are compared with ASXL1-/-, TET2-/- and TET2-/- ASXL1-/+ cell lines. Data are expressed as mean ± SD, n=3 biological replicates, (P≤0.05), (** P≤0.01). A one-way ANOVA test was used for statistical analysis, coupled with the Tukey's multiple comparison post hoc test.*

4.2.4: The effect of *TET2* and *ASXL1* mutations on the cell cycle in KBM5 cells

In order to assess the functional consequence of *TET2* and *ASXL1* mutations, we decided to assess the impact of these mutations on the cell cycle in KBM5 cell variants. There was a significant decrease in the percentage of *TET2* -/- (33.8 ± 3.4%) (P≤0.001) and *TET2* -/- *ASXL1* -/+ (42.8 ± 1.8%) (P≤0.05) KBM5 cell variants in G0-G1, compared to wildtype (49.3 ± 1.8%). Although not significant, there was a mild decrease in the percentage of *ASXL1* -/+ (48.4 ± 1.4%) KBM5 cell variants in G0-G1, compared to wildtype (49.3 ± 1.8%) (P>0.05). Between mutant groups, there was a significant decrease in the percentage of the *TET2* -/- (33.8 ± 3.4%) KBM5 cell variant in G0-G1, compared to *ASXL1* -/+ (48.4 ±

1.4%) (P≤0.001) and *TET2* -/- *ASXL1* -/+ (42.8 ± 1.8%) (P≤0.01) KBM5 cell variants. The *TET2* -/- (33.8 ± 3.4%) KBM5 cell variant had the fewest percentage of cells in G0-G1 (see figure 40).

There was no significant difference between the percentage of cells in S-Phase between all KBM5 cells variants (P>0.05). Although not significant, there was a mild increase in the percentage of *ASXL1* -/+ (22.0 ± 1.4%) and *TET2* -/- *ASXL1* -/+ (21.5 ± 1.7%) KBM5 cell variants in S-Phase, compared to wildtype (19.4 ± 2.2%) (P>0.05). Moreover, although not significant, there was a mild decrease in the percentage of the *TET2* -/- (19.0 ± 1.4%) KBM5 cell variant in S-Phase, compared to wildtype (19.4 ± 2.2%) (P>0.05). There was no significant difference between all mutant groups, although the *ASXL1* -/+ (22.0 ± 1.4%) KBM5 cell variant had the highest percentage of cells in S-Phase (P>0.05) (see figure 41).

There was a significant increase in the percentage of *TET2* -/- (21.8 ± 0.8%) (P≤0.05) and *TET2* -/- *ASXL1* -/+ (25.0 ± 1.2%) (P≤0.01) KBM5 cell variants in G2/M, compared to wildtype (18.1 ± 1.5%). Although not significant, there was also an increase in the percentage of the *ASXL1* -/+ (19.9 ± 0.6%) KBM5 cell variant in G2/M, compared to wildtype (18.1 ± 1.5%) (P>0.05). Between all mutant groups, we observed a significant increase in the percentage of the *TET2* -/- *ASXL1* -/+ (25.0 ± 1.2%) KBM5 cell variant in G2/M, compared to *ASXL1* -/+ (19.9 ± 0.6%) (P≤0.01) and *TET2* -/- (21.8 ± 0.8%) (P≤0.05) KBM5 cell variants. Although not significant, we observed an increase in the percentage of the *TET2* -/- (21.8 ± 0.8%) KBM5 cell variant in G2/M, compared to the *ASXL1* -/+ (19.9 ± 0.6%) KBM5 cell variant (P>0.05). The *TET2* -/- *ASXL1* -/+ (20.33 ± 0.06%) KBM5 cell variant had the highest percentage of cells in G2/M (see figure 42).

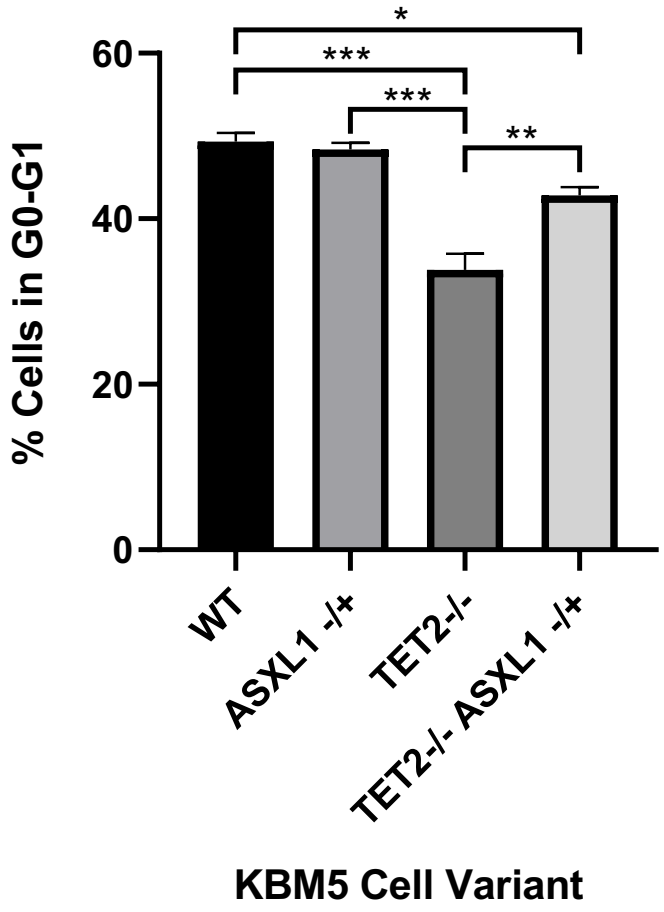


Figure 40: The impact of TET2 and ASXL1 mutations on G0-G1 of the cell cycle in KBM5 cells.
Wildtype cells (WT) are compared with ASXL1-/-, TET2-/- and TET2-/- ASXL1-/+ cell lines. Data are expressed as mean ± SD, n=3 biological replicates, (P≤0.05), (**P≤0.01), (*** P≤0.001). A one-way ANOVA test was used for statistical analysis, coupled with the Tukey's multiple comparison post hoc test.*

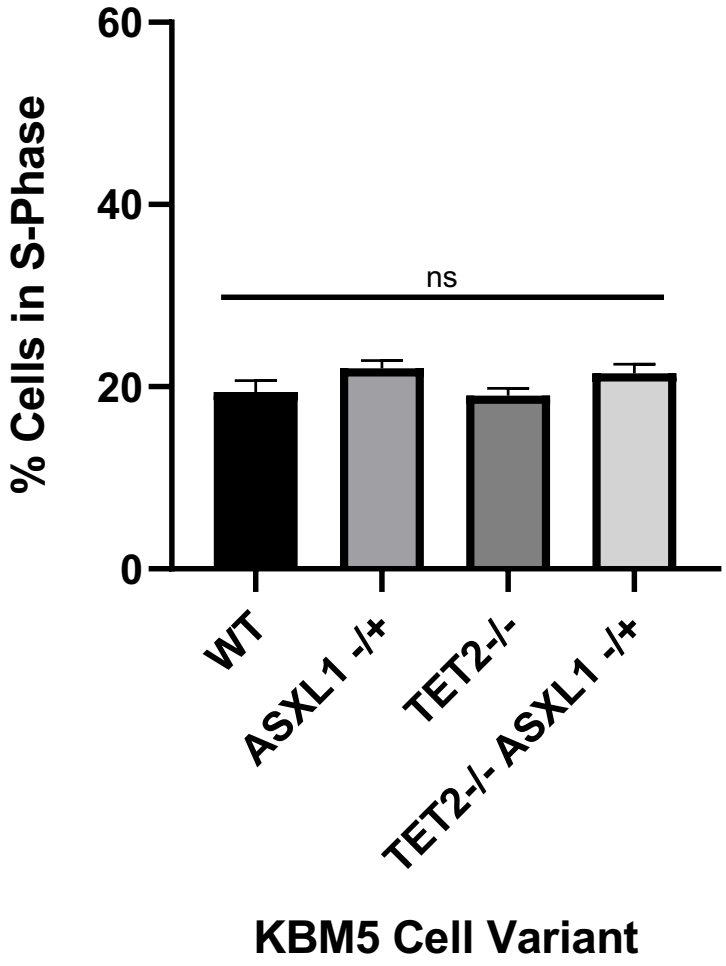


Figure 41: The impact of TET2 and ASXL1 mutations on S-Phase of the cell cycle in KBM5 cells.
Wildtype cells (WT) are compared with ASXL1-/-, TET2-/- and TET2-/- ASXL1-/+ cell lines. Data are expressed as mean ± SD, n=3 biological replicates, (ns=not significant). A one-way ANOVA test was used for statistical analysis, coupled with the Tukey's multiple comparison post hoc test.

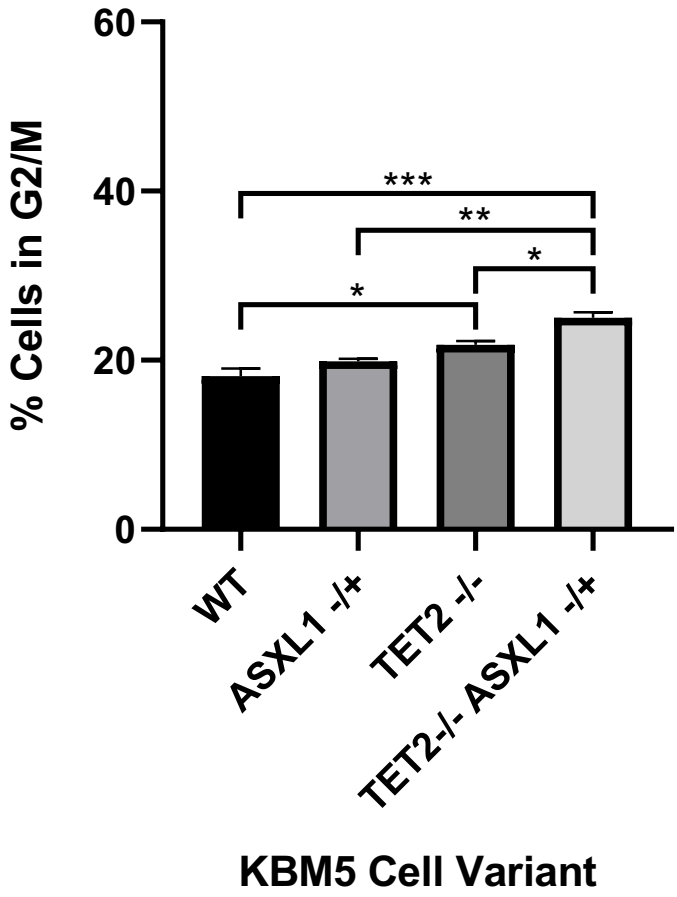


Figure 42: The impact of TET2 and ASXL1 mutations on G2/M of the cell cycle in KBM5 cells.
Wildtype cells (WT) are compared with ASXL1-/-, TET2-/- and TET2-/- ASXL1-/+ cell lines. Data are expressed as mean ± SD, n=3 biological replicates, (P≤0.05), (**P≤0.01), (*** P≤0.001). A one-way ANOVA test was used for statistical analysis, coupled with the Tukey's multiple comparison post hoc test.*

4.2.5: The effect of *TET2* and *ASXL1* mutations on the apoptosis in HUDEP2 cells

The ability of a cell to evade apoptosis is a well characterised hallmark of cancer pathophysiology (Hanahan and Weinberg 2011). In this light, we decided to investigate the functional consequence of mutations in *TET2* and *ASXL1* in HUDEP2 cells on apoptosis. We used Annexin V as a marker to quantify apoptosis. The values are expressed as percentage of live cells.

There was a significant increase in the percentage of apoptotic cells in *ASXL1* -/+ (10.69 ± 1.01%) and *TET2* -/- *ASXL1* -/+ (11.30 ± 0.71%) HUDEP2 cell variants compared to wildtype (7.11 ± 0.33%) (P≤0.01). There was a significant decrease in the percentage of apoptotic cells in the *TET2* -/- (4.36 ± 0.13%) HUDEP2 cell variant, compared to wildtype (7.11 ± 0.33%) (P≤0.05). Between mutant groups, we observed a significant increase in the percentage of apoptotic cells in *ASXL1* -/+ (10.69 ± 1.01%) and *TET2* -/- *ASXL1* -/+ (11.30 ± 0.71%) HUDEP2 cell variants, compared to the *TET2* -/- (4.36 ±

0.13%) HUDEP2 cell variant (P≤0.001). Although there was no significant difference in the percentage of apoptotic cells between *ASXL1* -/+ (10.69 ± 1.01%) and *TET2* -/- *ASXL1* -/+ (11.30 ± 0.71%) HUDEP2 cell variants, the *TET2* -/- *ASXL1* -/+ (11.30 ± 0.71%) HUDEP2 cell variant demonstrated a slight increase (P>0.05). The *TET2* -/- *ASXL1* -/+ (11.30 ± 0.71%) HUDEP2 cell variant demonstrated the highest percentage of apoptotic cells (see figure 43).

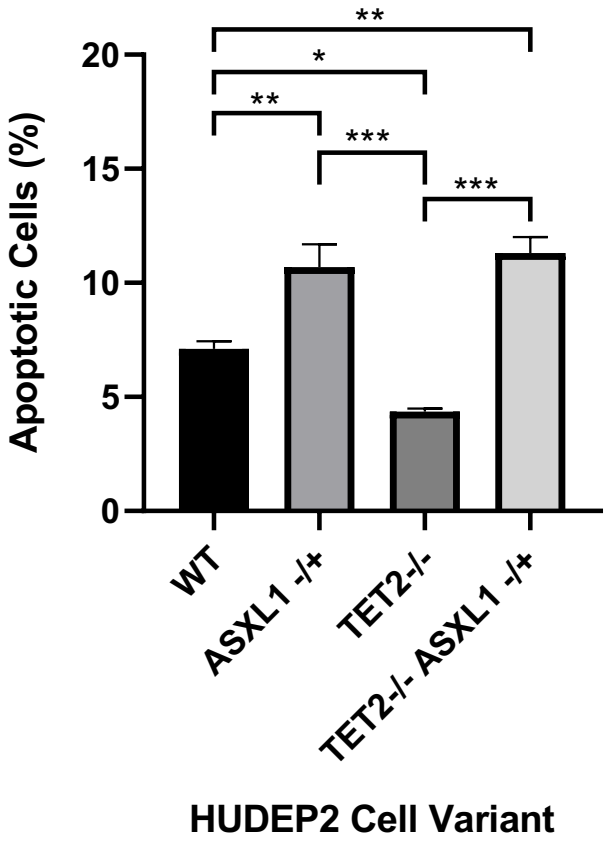


Figure 43: The impact of TET2 and ASXL1 mutations on apoptosis in HUDEP2 cells.
Wildtype cells (WT) are compared with ASXL1-/-, TET2-/- and TET2-/- ASXL1-/+ cell lines. Data are expressed as mean ± SD, n=3 biological replicates, (P≤0.05), (**P≤0.01), (*** P≤0.001). A one-way ANOVA test was used for statistical analysis, coupled with the Tukey's multiple comparison post hoc test.*

4.2.6: The effect of *TET2* and *ASXL1* mutations on the apoptosis in KBM5 cells

We also decided to investigate the functional consequence of mutations in *TET2* and *ASXL1* in our KBM5 cell model, on apoptosis. The values are expressed as percentage of live cells. There was a significant increase in the percentage of apoptotic cells in the *TET2* -/- (24.95 ± 1.77%) KBM5 cell variant compared to wildtype (4.67 ± 0.87%) (P≤0.0001). By contrast, there was a significant decrease

in the percentage of apoptotic cells in the *ASXL1* -/+ (2.28 ± 0.21%) KBM5 cell variant, compared to wildtype (4.67 ± 0.87%) (P≤0.05). Moreover, although not significant, there was a decrease in in the percentage of apoptotic cells in the *TET2* -/- *ASXL1* -/+ (3.65 ± 0.38%) KBM5 cell variant, compared to wildtype (4.67 ± 0.87%) (P>0.05). Between mutant groups, we observed a significant increase in the percentage of apoptotic cells in the *TET2* -/- (24.95 ± 1.77%) KBM5 cell variant, compared to the *ASXL1* -/+ (2.28 ± 0.21%) and *TET2* -/- *ASXL1* -/+ (3.65 ± 0.38%) KBM5 cell variants (P≤0.0001). Although there was no significant difference in the percentage of apoptotic cells between *ASXL1* -/+ (2.28 ± 0.21%) and *TET2* -/- *ASXL1* -/+ (3.65 ± 0.38%) KBM5 cell variants, the *TET2* -/- *ASXL1* -/+ (3.65 ± 0.38%) KBM5 cell variant demonstrated an increase in the percentage of apoptotic cells (P>0.05). The *TET2* -/- (24.95 ± 1.77%) KBM5 cell variant demonstrated the highest percentage of apoptotic cells (see figure 44).

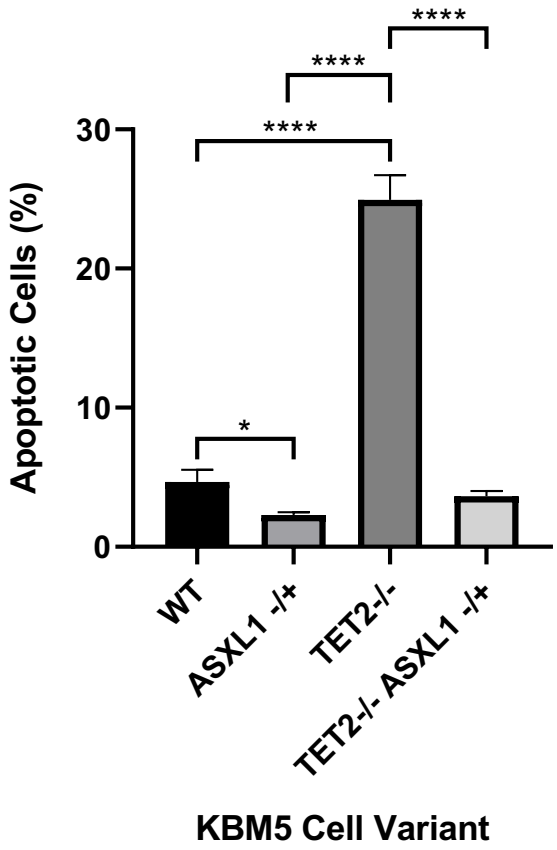


Figure 44: The impact of TET2 and ASXL1 mutations on apoptosis in KBM5 cells.
Wildtype cells (WT) are compared with ASXL1-/-, TET2-/- and TET2-/- ASXL1-/+ cell lines. Data are expressed as mean ± SD, n=3 biological replicates, (P≤0.05), (**** P≤0.0001). A one-way ANOVA test was used for statistical analysis, coupled with the Tukey's multiple comparison post hoc test.*

4.2.7: The effect of *TET2* and *ASXL1* mutations on erythroid differentiation in HUDEP2 cells

A block of differentiation is a characteristic of myeloid malignancy (Tenen 2014). Therefore, we decided to investigate the effect mutations in *TET2* and *ASXL1* have on HUDEP2 cell differentiation. The erythroid markers we investigated were CD36, CD71 and CD235a. The values are expressed as percentage of live cells.

The percentages of cells which were CD36+, were significantly lower in the *ASXL1* -/+ (10.48 ± 2.93%) HUDEP2 cell variant, compared to wildtype (25.49 ± 1.76%) (P≤0.01). By contrast, the percentages of CD36+ cells were significantly higher in the *TET2* -/- cell variant (60.84 ± 8.44%) HUDEP2 cell variant, compared to wildtype (25.49 ± 1.76%) (P≤0.0001). Furthermore, the was no significant difference in the percentage of cell which were CD36+ between the *TET2* -/- *ASXL1* -/+ (25.97 ± 0.91%) cell variant and wildtype (25.49 ± 1.76%) (P>0.05). Within the mutant group, the percentage of CD36+ cells were significantly higher in the *TET2* -/- (60.84 ± 8.44%) HUDEP2 cell variant, compared to *ASXL1* -/+ (10.48 ± 2.93%) and *TET2* -/- *ASXL1* -/+ (25.97 ± 0.91%) HUDEP2 cell variants (P≤0.0001). The percentages of CD36+ cells, were significantly higher in the *TET2* -/- *ASXL1* -/+ (25.97 ± 0.91%) HUDEP2 cell variant, compared with the *ASXL1* -/+ (10.48 ± 2.93%) HUDEP2 cell variant (P≤0.01).

The percentages of CD71+ cells were significantly higher in *ASXL1* -/+ (52.80 ± 0.42%) (P≤0.01) and *TET2* -/- (66.86 ± 7.19%) (P≤0.0001) HUDEP2 cell variants, compared to wildtype (37.06 ± 4.80%). Moreover, although not significant, the percentages of CD71+ cells were higher in the *TET2* -/- *ASXL1* -/+ (42.55 ± 5.23%) HUDEP2 cell variant, compared to wildtype (37.06 ± 4.80%) (P>0.05). Within the mutant group, the percentage of CD71+ cells, were significantly higher in the *TET2* -/- (66.86 ± 7.19%) HUDEP2 cell variant, compared to *ASXL1* -/+ (52.80 ± 0.42%) (P≤0.05) and *TET2* -/- *ASXL1* -/+ (42.55 ± 5.23%) (P≤0.0001) HUDEP2 cell variants. Although not significant, the percentages of CD71+ cells, were higher in the *ASXL1* -/+ (52.80 ± 0.42%) HUDEP2 cell variant, compared with the *TET2* -/- *ASXL1* -/+ (42.55 ± 5.23%) HUDEP2 cell variant (P>0.05).

The percentages of CD235a+ cells, although not significant, were higher in *ASXL1* -/+ (9.15 ± 0.69%) and *TET2* -/- *ASXL1* -/+ (5.98 ± 3.80%) HUDEP2 cell variants, compared to wildtype (2.22 ±

0.98%) (P>0.05). Moreover, although not significant, the percentages of CD235a+ cells were lower in the *TET2* -/- (0.14 ± 0.08%) HUDEP2 cell variant, compared to wildtype (2.22 ± 0.98%) (P>0.05). Within the mutant group, although not significant, the percentage of CD235a+ cells were higher in the *ASXL1* -/+ (9.15 ± 0.69%) HUDEP2 cell variant, compared to the *TET2* -/- (0.14 ± 0.08%) and *TET2* -/- *ASXL1* -/+ (5.98 ± 3.80%) HUDEP2 cell variants (P>0.05). Furthermore, although not significant, the percentage of CD235a+ cells were high in the *TET2* -/- *ASXL1* -/+ (5.98 ± 3.80%) HUDEP2 cell variant, compared to the *TET2* -/- (0.14 ± 0.08%) HUDEP2 cell variant (P>0.05) (see figure 45).

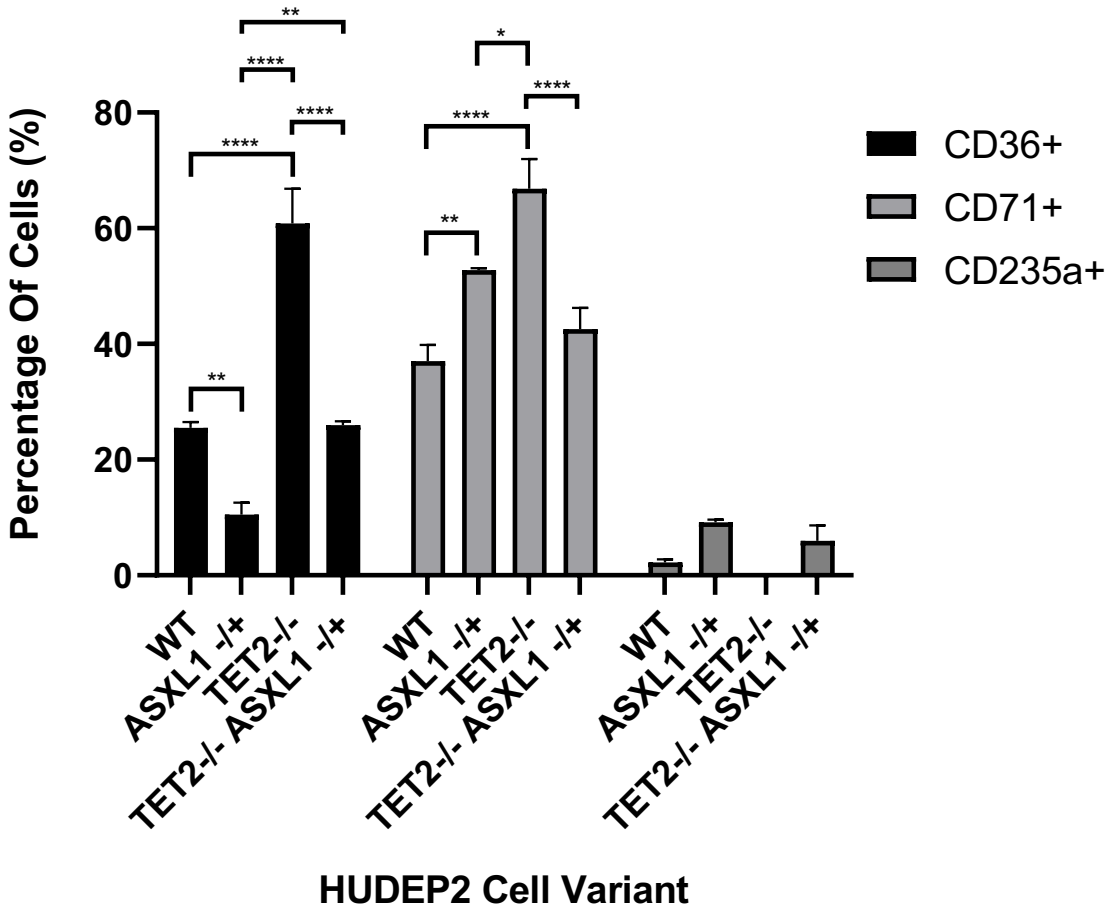


Figure 45: The impact of TET2 and ASXL1 mutations on HUDEP2 cell differentiation along the erythroid lineage as assessed by expression of CD36 (black histograms), CD71 (grey histograms) and CD235a (dark grey histograms) using FACS. Wildtype cells (WT) are compared with ASXL1-/-, TET2-/- and TET2-/- ASXL1-/+ cell lines. Data are expressed as mean ± SD, n=3 biological replicates, (P≤0.05), (**P≤0.01), (**** P≤0.0001). A two-way ANOVA test was used for statistical analysis, coupled with the Tukey's multiple comparison post hoc test.*

4.2.8: The effect of *TET2* and *ASXL1* mutations on myeloid differentiation in KBM5 cells

We decided to investigate the effect mutations in *TET2* and *ASXL1* have on KBM5 cell differentiation. The myeloid markers we investigated were CD11b, CD14 and CD15. The values are expressed as percentage of live cells.

The percentages of CD11b+ cells were significantly higher in *ASXL1* -/+ (24.48 ± 3.32%) and *TET2* -/- (25.68 ± 3.08%) KBM5 cell variants, compared to wildtype (13.03 ± 0.06%) (P≤0.0001). Moreover, although not significant, the percentages of CD11b+ cells were higher in the *TET2* -/- *ASXL1* -/+ (16.30 ± 3.17%) KBM5 cell variant, compared to wildtype (13.03 ± 0.06%) (P>0.05). Within the mutant group, the percentage of CD11b+ cells, were significantly higher in *ASXL1* -/+ (24.48 ± 3.32%) and *TET2* -/- (25.68 ± 3.08%) KBM5 cell variants, compared to the *TET2* -/- *ASXL1* -/+ (16.30 ± 3.17%) KBM5 cell variant (P≤0.001). Although not significant, the percentages of CD11b+ cells, were higher in the *TET2* -/- (25.68 ± 3.08%) KBM5 cell variant, compared to the *ASXL1* -/+ (24.48 ± 3.32%) KBM5 cell variant (P>0.05).

The percentages of CD14+ cells although not significant, were lower in *ASXL1* -/+ (0.02 ± 0.01%), *TET2* -/- (0.75 ± 0.26%) and *TET2* -/- *ASXL1* -/+ (0.11 ± 0.07%) KBM5 cell variants, compared to wildtype (1.28 ± 0.40%) (P>0.05). By contrast, in mutant group, although not significant, the percentages of CD14+ cells was comparable in *ASXL1* -/+ (0.02 ± 0.01%) and *TET2* -/- *ASXL1* -/+ (0.11 ± 0.07%) KBM5 cell variants (P>0.05). In the mutant group, the *TET2* -/- (0.75 ± 0.26%) KBM5 cell variant, had the highest percentage of CD14+ cells.

The percentages of CD15+ cells were significantly higher in *ASXL1* -/+ (16.67 ± 2.98%) and *TET2* -/- (17.53 ± 3.04%) KBM5 cell variants, compared to wildtype (6.33 ± 0.23%) (P≤0.0001). Moreover, although not significant, the percentages of CD15+ cells were higher in the *TET2* -/- *ASXL1* -/+ (10.16 ± 2.64%) KBM5 cell variant, compared to wildtype (6.33 ± 0.23%) (P>0.05). Within the mutant group, the percentage of CD15+ cells were significantly higher in *ASXL1* -/+ (16.67 ± 2.98%) and *TET2* -/- (17.53 ± 3.04%) KBM5 cell variants, compared to the *TET2* -/- *ASXL1* -/+ (10.16 ± 2.64%) KBM5 cell variant (P≤0.01). Although not significant, the percentages of CD15+ cells, were higher in the *TET2* -/- (17.53

± 3.04%) KBM5 cell variant, compared to the *ASXL1* -/+ (16.67 ± 2.98%) KBM5 cell variant (P>0.05) (see figure 46).

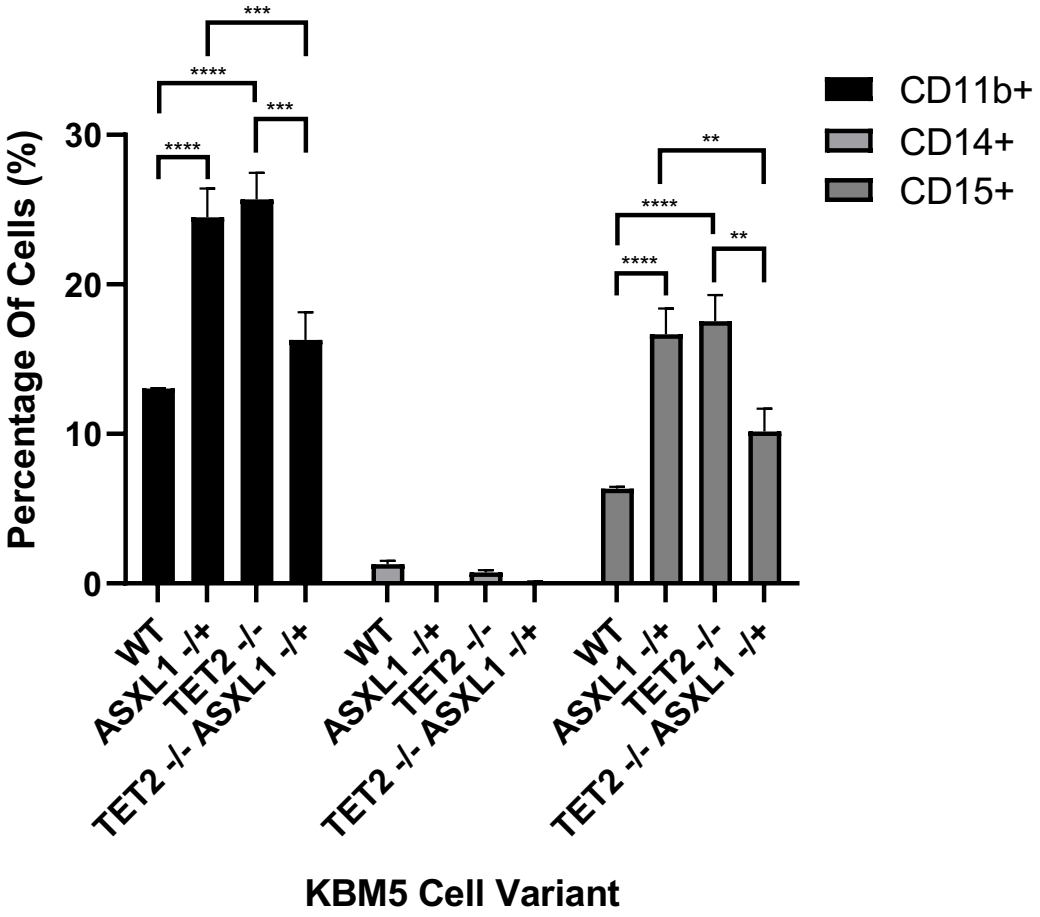


*Figure 46: The impact of TET2 and ASXL1 mutations on myeloid differentiation markers in KBM5 cells as assessed by expression of CD11b (black histograms), CD14 (grey histograms) and CD15 (dark grey histograms) using FACS. Wildtype cells (WT) are compared with ASXL1-/-, TET2-/- and TET2-/- ASXL1-/+ cell lines. Data are expressed as mean ± SD, n=3 biological replicates, (** P≤0.01), (***P≤0.001), (**** P≤0.0001). A two-way ANOVA test was used for statistical analysis, coupled with the Tukey's multiple comparison post hoc test.*

4.3: Discussion

In this study, we elucidated the impact of common MDS mutations on basic cellular functions. We modelled the erythroid disease using HUDEP2 cells, in which we introduced an *ASXL1* p.G646Wfs*12 mutation, in addition to knocking out *TET2*. To study the myeloid disease, we used the KBM5 cell line, which had a G710X mutation (previously produced in our laboratory) (Valletta, Dolatshad et al. 2015), and in which we also knocked out *TET2*. This allowed us to study the impact of *TET2* and *ASXL1* gene

mutations on cellular function singularly and as co-mutations. We decided to investigate the impact of these mutations on key cellular functions such as proliferation, the cell cycle, apoptosis and lineage differentiation.

Firstly, we investigated the effect *ASXL1* and/or *TET2* mutations had on proliferation in both HUDEP2 and KBM5 cell lines. In HUDEP2 cells, we observed that our mutant cell line variants, had a higher proliferation compared to wildtype. The most profound increase in proliferation after 4 days, was observed in the co-mutated *TET2* -/- *ASXL1* -/+ HUDEP2 cell variant. Comparably, in KBM5 cells, we observed that our mutant cell line variants, had a higher proliferation compared to wildtype. The most profound increase in proliferation after 4 days, was observed in the co-mutated *TET2* -/- *ASXL1* -/+ KBM5 cell variant.

The effects of mutations in *TET2* and *ASXL1* on cell proliferation have been widely published in the literature. *TET2* has been shown to suppress proliferation in leukaemic cells (Wang, Xiao et al. 2015). Moreover, it has been shown in osteoblasts, that *TET2* loss increases proliferation (Li, Zhou et al. 2018). It was reported by Montagner and colleagues that *TET2* loss increased mast cell proliferation, and that in the absence of *TET2*, mast cells showed altered 5hmC deposition and disrupted gene expression (Montagner, Leoni et al. 2016). Rajan and colleagues report that, in the presence of external stimuli such as infection, loss of *TET2* leads to an increase in proliferation of HSCs and as a result accumulation of immature cells prone to becoming leukaemogenic. Although this phenomenon was observed in Zebra fish, it also extends to human HSCs (Rajan, Collett et al. 2021). In a study conducted by Li and colleagues, it was demonstrated that deletion of *TET2* in mice increased the proliferative capacity of HSCs. In addition, they suggested that the increased HSC pool and proliferation of HSCs and HSPCs in *TET2* -/- mice may lead to myeloid malignancy (Li, Cai et al. 2011).

By contrast, it was demonstrated by Hye and colleagues that deletion of *ASXL1* led to cellular senescence in embryonic cells, suggesting an important role for *ASXL1* in proliferation (Youn, Kim et al. 2017). Moreover, it was demonstrated by Valletta and colleagues that *ASXL1* homozygous mutation corrected cells demonstrated a significant growth suppression compared to uncorrected *ASXL1*

mutant KBM5 cells (Valletta, Dolatshad et al. 2015). The data described within this thesis showing that *TET2* and *ASXL1* mutations lead to increased proliferation are in line with the literature. We also observed that in co-mutated cell lines, harbouring *TET2* and *ASXL1* mutations, the proliferative advantage was more profound, compared to cells with either mutation. This is an important finding since not only does it confirm that *TET2* and *ASXL1* are important regulators of cell proliferation, but also suggests that *TET2* and *ASXL1* have an additive effect on proliferation. Our data are in line with the literature, in which Abdel-Wahab and colleagues report an MDS phenotype and as a consequence a hastened death for mice exhibiting a compound *ASXL1*/*TET2* deletion compared to single-gene KO mice (Abdel-Wahab, Gao et al. 2013). In addition, they also report that mice with simultaneous loss of *ASXL1* and *TET2* develop a larger increase in HSCPs and decrease in the number of circulating mature cells compared with single gene KO mice (Abdel-Wahab, Gao et al. 2013). Loss of *TET2* had a more prominent effect on cell proliferation, since proliferation was greater in cell lines harbouring *TET2* mutations than in cell lines with *ASXL1* mutations. It would also be interesting to conduct this study over a longer period to investigate if this effect would be sustained.

We next decided to investigate the effect *ASXL1* and/or *TET2* mutations had on cell cycle progression, in both HUDEP2 and KBM5 cell lines. In HUDEP2 cells, we observed that our mutant cell line variants, showed a decrease in the percentage of cells in G0-G1, compared to wildtype. The most profound decrease in the percentage of cells in G0-G1, was observed in the co-mutated *TET2* -/- *ASXL1* -/+ HUDEP2 cell variant. By contrast, in HUDEP2 cells, we observed that our mutant cell line variants, showed an increase in the percentage of cells in S-Phase, compared to wildtype. The most profound increase in the percentage of cells in S-Phase, was observed in the co-mutated *TET2* -/- *ASXL1* -/+ HUDEP2 cell variant. Finally, for HUDEP2 cells, we observed that our mutant cell line variants, showed an increase in the percentage of cells in G2/M, compared to wildtype. The most profound increase in the percentage of cells in G2/M, was observed in the co-mutated *TET2* -/- *ASXL1* -/+ HUDEP2 cell variant.

In KBM5 cells, we observed that our mutant cell line variants, showed a decrease in the percentage of cells in G0-G1, compared to wildtype. The most profound decrease in the percentage of cells in G0-G1, was observed in the *TET2* -/- cell variant. By contrast, in KBM5 cells, we observed that our *ASXL1* -/+ and co-mutated *TET2* -/- *ASXL1* -/+ KBM5 cell line variants, showed a slight increase in the percentage of cells in S-Phase, whereas our *TET2* -/- KBM5 cell line variant, showed a slight decrease in the percentage of cells in S-Phase, compared to wildtype. Finally, for KBM5 cells, we observed that our mutant cell line variants, showed an increase in the percentage of cells in G2/M, compared to wildtype. The most profound increase in the percentage of cells in G2/M, was observed in the co-mutated *TET2* -/- *ASXL1* -/+ KBM5 cell variant.

It has been shown by various research groups that *TET2* and *ASXL1* are important in cell cycle regulation. In a study conducted by Liu and colleagues, they demonstrated that a *TET2* knockdown in HaCaT cells, significantly decreased the proportion of cells observed in G0-G1, and significantly increased the proportion of cell observed in S-Phase and G2/M, compared to wildtype (Liu, Wang et al. 2020). It was also demonstrated by Cimmino and colleagues that restoration of *TET2* using vitamin C, in an inducible cKit+ *TET2* knockdown cell line, delayed cell cycle progression (Cimmino, Dolgalev et al. 2017). Competitive transplant experiments were conducted by Chen and colleagues, in which they purified lineage^{low}Kit^{high}Sca1^{+} (LSK) cells from *TET2* null mice. Following staining and flow cytometry of key cell cycle markers, they observed a significant increase of cells in S/G2M of the cell cycle, compared to WT (Chen, Schneider et al. 2015).

It was shown by Asada and colleagues that a mutation in *ASXL1* leading to expression of a C-terminally truncated form of *ASXL1*, increased the proportion of S-Phase and G2/M cells in human RUNX1-ETO cells (Asada, Goyama et al. 2018). In Human umbilical cord blood (UCB) cells, Takeda and colleagues investigated the effect of *ASXL1* in myeloid leukemogenesis. It was shown that cells expressing *ASXL1* and *HHEX* knockouts, showed an increased proportion of cells in S/G2/M-phase (Takeda, Asada et al. 2020).

The data presented in this thesis is in line with the literature, in that we observed a lower percentage of G0-G1 cells, in all mutant HUDEP2 and KBM5 cell variants, compared to wildtype. In addition, we also observed a higher percentage of S-Phase cells, in all mutant HUDEP2 and KBM5 cell variants, except for the *TET2* -/- KBM5 cell variant, compared to wildtype.

We also observed a higher percentage of G2/M cells in all mutant HUDEP2 and KBM5 cell variants, compared to wildtype. We observed that our co-mutated *TET2* -/- *ASXL1* -/+ HUDEP2 or KBM5 cell variants, showed the most profound increase in the percentage of cells in the latter part of the cell cycle, compared to individual *ASXL1* +/- or *TET2* -/- mutations; this is expected since both *TET2* and *ASXL1* mutations drive cell cycle progression.

It is well established that evading apoptosis is a hallmark of cancer (Hanahan and Weinberg 2011). In this light we next decided to investigate the effect *ASXL1* and/or *TET2* mutations have on apoptosis, in both HUDEP2 and KBM5 cell lines.

In HUDEP2 cells, we observed an increase in the percentage of apoptotic cells in *ASXL1* -/+ and *TET2* -/- *ASXL1* -/+ HUDEP2 cell variants, compared to wildtype. By contrast we saw a decrease in the percentage of apoptotic cells in the TET -/- HUDEP2 cell variant, compared to wildtype. In KBM5 cell, we observed the opposite effect. We observed a decrease in the percentage of apoptotic cells in *ASXL1* -/+ and *TET2* -/- *ASXL1* -/+ KBM5 cell variants, compared to wildtype. By contrast we saw an increase in the percentage of apoptotic cells in the TET -/- KBM5 cell variant, compared to wildtype.

The effect of *TET2* and *ASXL1* mutations have on apoptosis is well characterised in the literature. Deng and colleagues observed that *TET2* overexpression induced apoptosis in gastric cancer cells (Deng, Wang et al. 2016). Furthermore, Cai and colleagues demonstrated that in *TET2* loss reduced apoptosis in HSCs, and that this may be due to hyperactivation of IL-6/Shp2/Stat3/*Morrbid* (Cai, Kotzin et al. 2018). Abegunde and colleagues conducted a study to investigate the effect *TET2* has on tumour necrosis factor alpha (TNFα) induced apoptosis. They reported that *TET2*-/- murine bone marrow (BM) cells had a lower apoptotic index compared with WT BM cells*, in vitro* (Abegunde, Buckstein et al. 2018).

It was demonstrated by Wang and colleagues that loss of *ASXL1* led to an increase in apoptosis, in mice (Wang, Li et al. 2014). A study was conducted by Abdel-Wahab and colleagues, in which conditional deletion of *ASXL1* was associated with a profound increase in apoptosis, *in vivo* (Abdel-Wahab, Gao et al. 2013). Furthermore, in a study conducted by Hilgendorf and colleagues loss of *ASXL1* in UCB CD34+ cells caused a marked increase in apoptosis (Hilgendorf, Folkerts et al. 2016).

Our HUDEP2 cell data are in line with the literature, in that we observed that a mutation in *ASXL1* in these HUDEP2 cells led to an increase in the percentage of apoptotic cells, compared to control. Moreover, we observed that a *TET2* KO in HUDEP2 cells led to a decrease in the percentage of apoptotic cells compared to control. Interestingly, the co-mutated *TET2* -/- *ASXL1* -/+ HUDEP2 cell variant showed an increase in the percentage of apoptotic cells compared to control. This may be because the *ASXL1* mutation is more enabling than the *TET2* mutation which is suppressive, with regards to apoptosis, which may counterbalance the pro-apoptotic effect of ASXL1. Our KBM5 data showed the opposite result. This may be due to differences in the apoptosis regulatory mechanisms between the two cell lines.

It is well established that a differentiation block is a characteristic of myeloid neoplasms (Tenen 2014). In this light we next decided to investigate the effect *ASXL1* and/or *TET2* mutations have on erythroid and myeloid differentiation markers in HUDEP2 and KBM5 cell lines respectively.

For HUDEP2 cells, which are erythroid in nature, we decided to look at CD36, CD235a and CD71. CD71 and CD36 are early markers (Marsee, Pinkus et al. 2010, Ng and Wood 2014), whereas CD235a (Glycophorin A) is a marker of more mature erythroid cells (Marsee, Pinkus et al. 2010). In the HUDEP2 cell variants, all mutant cell line variants had a higher percentage of CD71+ cells compared to wildtype, with the *TET2* -/- being the most pronounced. CD71 expression also positively correlated with cell cycle progression in HUDEP2 cell mutant variants. In HUDEP2 cells, the majority of the cells that were CD36+ belonged to the *TET2* -/- mutant cell line. Expression of CD36 in the *TET2* -/- *ASXL1* -/+ co-mutant cell line and wildtype HUDEP2 cells were comparable. However, we observed that the *ASXL1* -/+ variant had a fewer number of CD36+ cells compared with wildtype cells. In HUDEP2 cells,

the number of CD235a+ cells were very low for all cell variants, with the highest being the *ASXL1* -/+ mutant variant.

We then repeated these differentiation experiments in the KBM5 mutant cell variants, using the myeloid markers CD11b, CD14 and CD15. In KBM5 cells, the mutant cell lines had a higher percentage of CD11b+ and CD15+ cells compared to wildtype cells. We also observed that the expression of CD14 was very low in all cell variants, including wildtype.

The role of TET2 and ASXL1 in erythroid and myeloid differentiation are well characterised in the literature. It was demonstrated by Scopim and colleagues that *TET2* expression increased during erythroid differentiation of primary haematopoietic progenitor cells (Scopim-Ribeiro, Machado-Neto et al. 2017). Qu and colleagues demonstrated that *TET2* deficiency in CD34+ cells lead to impaired differentiation of erythroid cells (Qu, Zhang et al. 2018). Hilgendorf and colleagues conducted a study to determine the role of ASXL1 in HSCs and HSPCs in UCB CD34+ cells. They report that in the presence of an *ASXL1* knockdown, there was an increase in erythroid lineage progenitor cells (Hilgendorf, Folkerts et al. 2016). Moreover, it was demonstrated by Shi and colleagues that knockdown of *ASXL1* impairs erythroid differentiation, in umbilical cord blood CD34+ cells (Shi, Yamamoto et al. 2016). It was shown by Ko and colleagues that *TET2* deficiency delayed HSC differentiation significantly (Ko et al., 2011).

Furthermore, it was demonstrated by Inoue and colleagues that myeloid differentiation was inhibited in the presence of a truncating *ASXL1* mutation, *in vivo*, which induced an MDS like disease (Inoue, Kitaura et al. 2013). It was also reported by Valletta and colleagues, that increased myeloid differentiation was observed in *ASXL1* mutation-corrected cells, compared to KBM5 mutant cells (Valletta, Dolatshad et al. 2015).

Although our HUDEP2 cell erythroid differentiation data are in line with the literature, our KBM5 cell myeloid differentiation data are not. One reason why we observed a different result to what was expected in comparison to the literature is that our KBM5 cell models were expanded from a single

clone, and these data may have been more representative of the findings by others if we used a bulk KBM5 cell population. Another reason for the difference in data may be due to cell lines being immortalised and therefore do not differentiate as well as primary cells. It would be interesting to conduct this experiment *in vivo*, since immortalised cell lines are a closed system, and not clinically viable. In addition, a larger panel of differentiation markers could be used to increase the resolution of results.

Chapter 5
Human Haematopoietic Stem and Progenitor Cell Expansion

5.1: Introduction

In previous Chapters, we successfully generated and functionally characterised erythroid and myeloid models of myeloid malignancy in HUDEP2 and KBM5 cells respectively. In this chapter, as a prelude to gene editing of human umbilical cord blood (UCB) derived CD133+ haematopoietic stem and progenitor cells (HSPCs), we conducted experiments to develop expansion methodologies for these cells. Our aim was to generate more robust and clinically relevant models of myeloid malignancy. Since it would be advantageous to increase the number of edited HSCs and since cycling/*in vitro* culture is a requirement for some genome editing technologies, in the present study the ability of small molecules and cytokines to expand human CD133+ cells *ex vivo* were investigated. Cluster of differentiation (CD) markers were used to identify Haematopoietic Stem and Progenitor Cells that have been reported to phenotypically classify the CD133+ cells (see table 31). Previous research, at the commencement of the studies for this thesis and involving our laboratory, has identified differential and sequential cytokine and small molecule treatments to expand human UCB CD133+ derived HSPCs *ex vivo* (Gullo, van der Garde et al. 2015). Furthermore, numerous compounds, that have been reported in the literature to have efficacy in the expansion of HSCs *ex vivo* are summarised in table 32.

Biomarker	References
CD34	(Matsuoka, Ebihara et al. 2001)
CD38	(Dürig, Naschar et al. 2002)
CD45RA	(Kersten, Valkering et al. 2016)
CD90	(Lechman, Gentner et al. 2012)
CD49F	(Huntsman, Bat et al. 2015)
CD123	(Testa, Pelosi et al. 2014)
CD133	(Toren, Bielorai et al. 2005)

Table 31: CD biomarkers of HSPCs

Molecule	Mechanism of Action	References
SCF	Increases HSC Survival	(Sasaki, Mizuochi et al. 2010)
TPO	Supports HSC Quiescence, Platelet and Megakaryocyte Production	(Fox, Priestley et al. 2002)
FLT3L	Stimulates Differentiation of HSCs	(McKenna, de Vries et al. 1995)
Scriptaid	HDAC Inhibition	(Hua, Kronsteiner et al. 2019)
ZLLYFMK	Calpain Inhibition	(Kale and Limaye 2010)
UM171	HSC Self Renewal and Inhibits Erythroid and Megakaryocyte Differentiation	(Fares, Chagraoui et al. 2014)
SR1	Aryl Hydrocarbon Receptor Antagonist	(Wagner Jr, Brunstein et al. 2016)
VPA (Valproic Acid)	HDAC Inhibition	(Bug, Gül et al. 2005)
NAM (Nicotinamide)	SIRT1 Inhibitor	(Peled, Shoham et al. 2012)
PEITC	Notch Activator	(Chan and Miskimins 2012)

Table 32: Factors that aid expansion of HSPCs ex vivo

5.2: Results

5.2.1: Isolation of CD133+ cells from fresh human umbilical cord blood

Initially, CD133+ cells were isolated from human UCB in order to establish a bank of CD133+ cells to be used for HSPC expansion studies. Following doublet exclusion, live and CD133+ cells were gated (see figure 47). CD133+ cell numbers and purity were quantified using flow cytometry. These data show that, for 12 UCB donations, the cell viability was 96.01 ± 2.47% and the CD133+ cell purity was 90.67 ± 5.55% (see figure 48). The cells were subsequently cryopreserved.

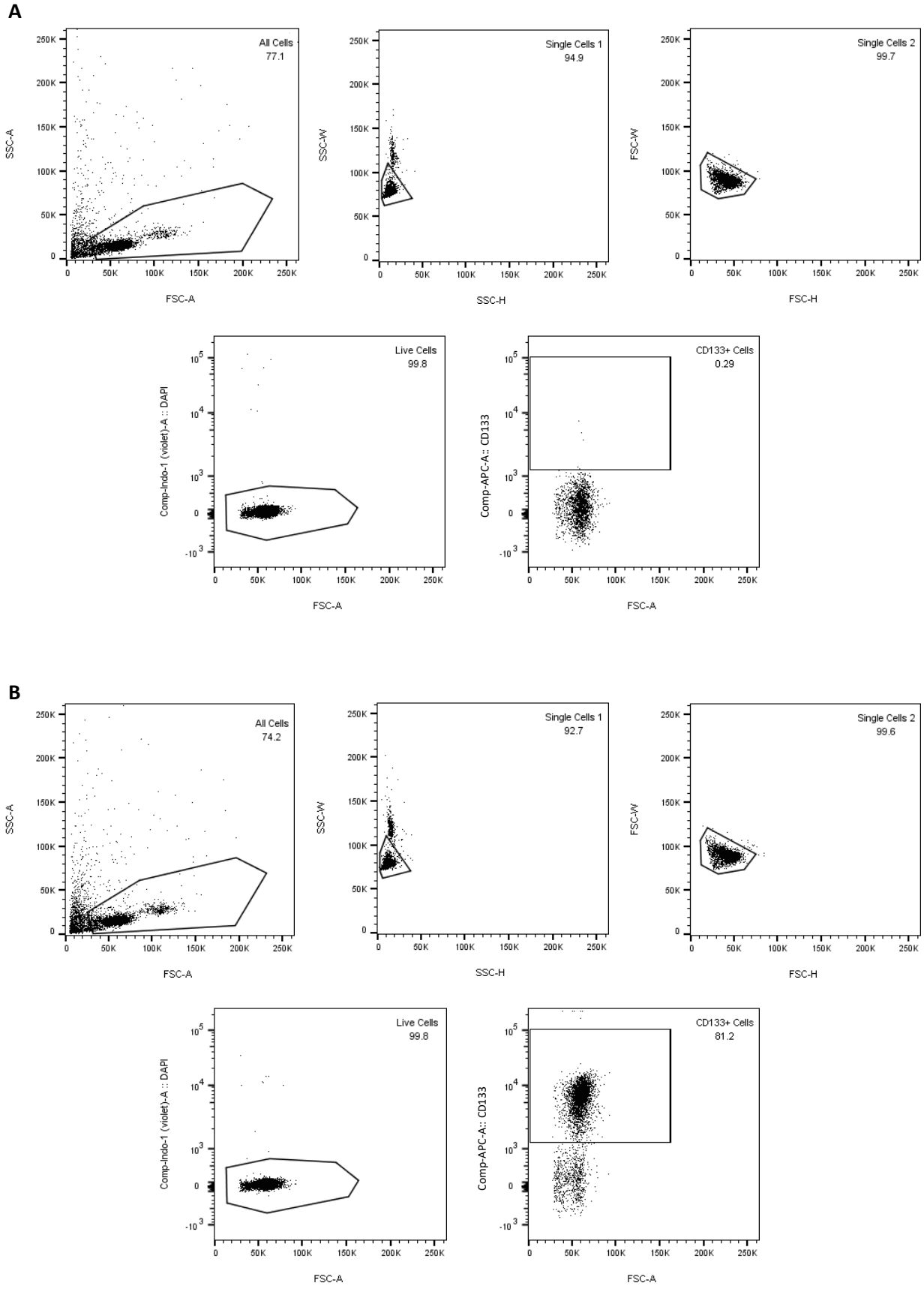
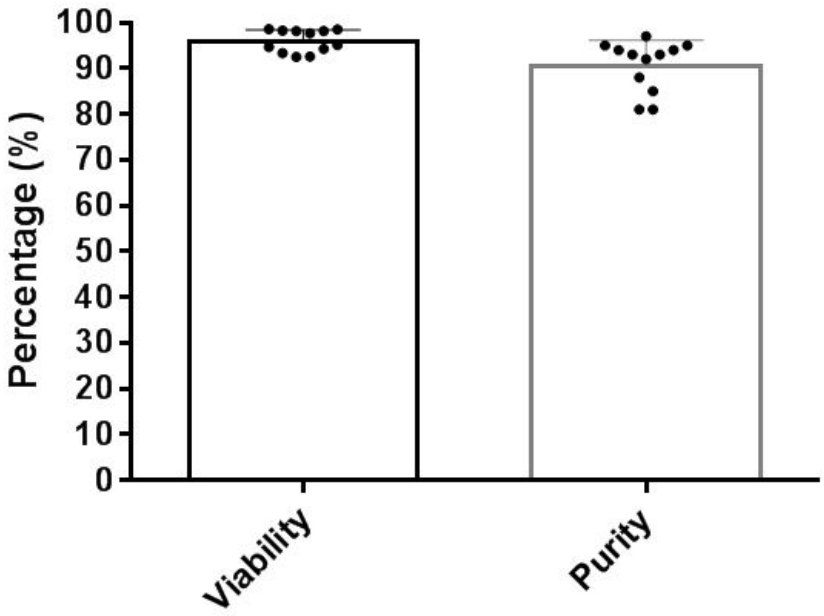


Figure 47: Representative gating strategy for the assessment of viability and purity of CD133+ cells isolated from fresh human umbilical cord blood. A) Gating workflow of the negative control. B) Gating workflow for CD133+ cells. The gating hierarchy is as follows: All cells > Single cells 1 > Single cells 2 > Live cells > CD133+ cells. The data are represented as a dot plot.

A



B

	Viability	Purity
n=12 Donors, Mean ± SD	96.01 ± 2.47%	90.67 ± 5.55%

Figure 48: Percentage viability and percentage purity of CD133+ cells isolated from fresh human UCB. A) Percentage viability and purity of cells. B) Numerical data showing viability and purity of cells. Data are expressed as mean ± SD, n=12.

5.2.2: The selection of CD133+ cell donor samples isolated from fresh human UCB

The *ex vivo* expansion of cryopreserved donor samples was tested by culturing cells in liquid culture and this served as a selection platform for donor samples. Expansion was observed in all donor samples using SCF, TPO and FLT3L, which were previously optimised by studies conducted in our laboratory (Gullo, van der Garde et al. 2015). Flow analysis was conducted on Day 5 for each donor sample (see figure 49 for gating strategy). Fold increase ranged from 1.87 ± 0.32 to 18.64 ± 3.47 (see figure 50). The donor samples which expanded the most were taken forward for further small molecule expansion experiments. These were donor samples: 610 (13.28 ± 2.33), 757 (10.13 ± 1.69), 875 (14.86 ± 1.73), 461 (18.64 ± 3.47), 453 (3.93 ± 0.63) and 456 (18.47 ± 4.27).

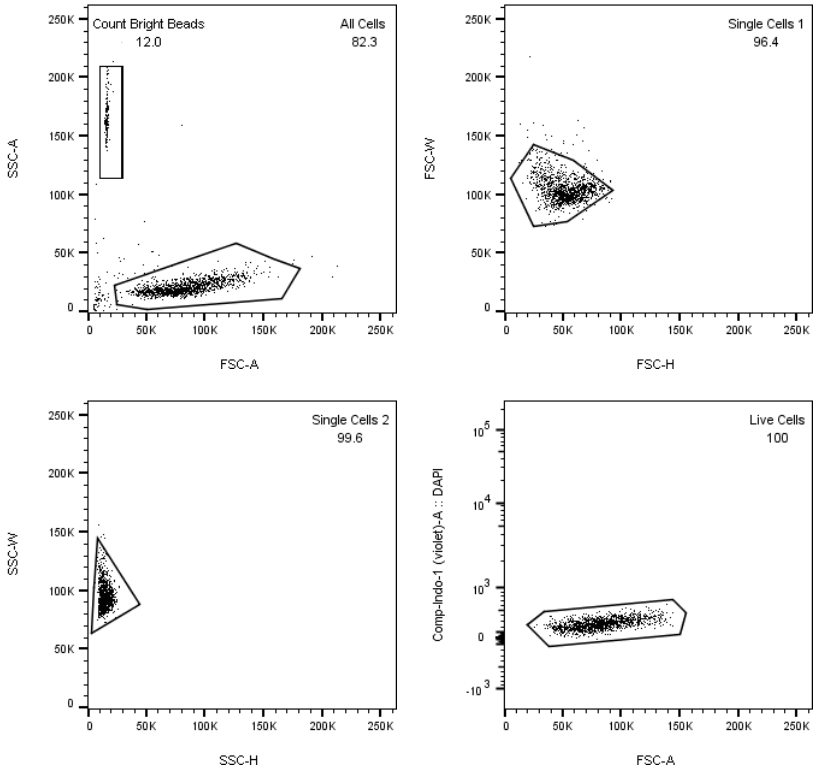
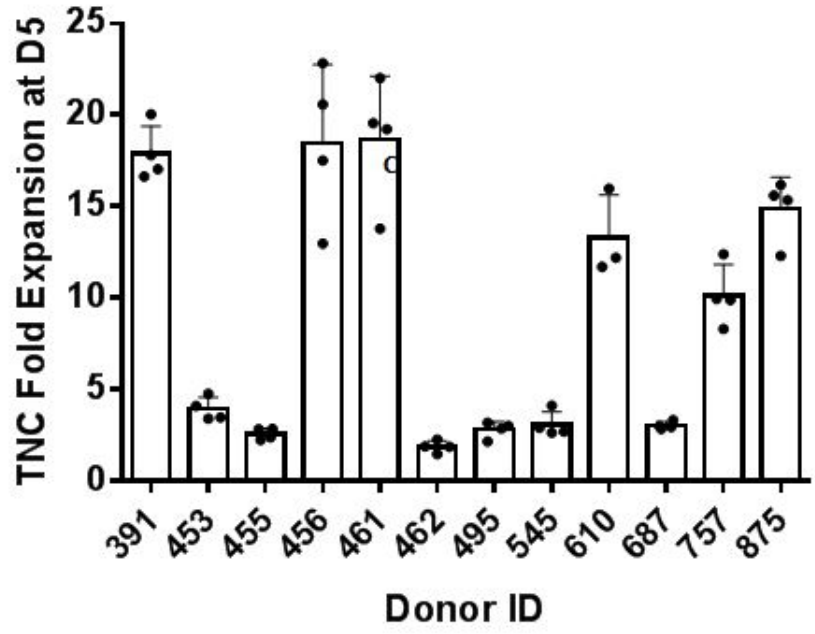


Figure 49: Representative gating strategy for TNC expansion of CD133+ cells used for donor selection. The gating hierarchy is as follows: All cells > Single cells 1 > Single cells 2 > Live cells. The data are represented as a dot plot.

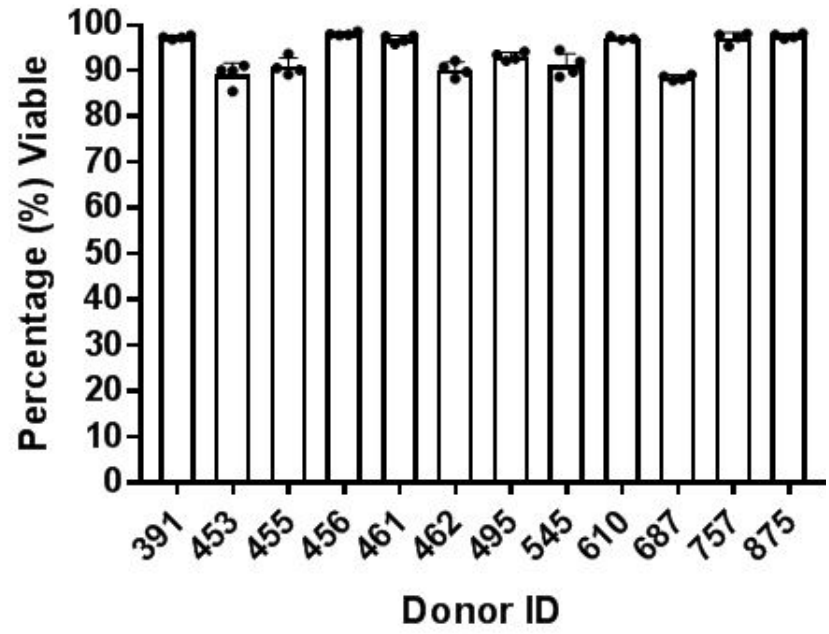
A



B

Donor ID	Mean	SD
391	17.88	1.51
453	3.93	0.63
455	2.57	0.31
456	18.47	4.27
461	18.64	3.47
462	1.87	0.32
495	2.80	0.45
545	3.09	0.70
610	13.28	2.33
687	3.03	0.20
757	10.13	1.69
875	14.86	1.73

C



D

Donor ID	Mean	SD
391	97.35	0.37
453	89.18	2.50
455	90.93	1.94
456	98.08	0.36
461	96.93	0.83
462	90.30	1.66
495	93.15	0.88
545	91.25	2.57
610	97.10	0.36
687	88.53	0.59
757	97.18	1.23
875	97.63	0.53

Figure 50: Expansion and viability of human UCB CD133+ cells.
A) Fold expansion at D5 (day 5) of D0 (day 0). B) Mean and standard deviation for fold expansion. C) Percentage viability D) Mean and standard deviation for percentage viability. Data are expressed as mean ± SD, n=3-4. A one-way ANOVA test was used for statistical analysis, coupled with the Tukey's multiple comparison post hoc test.

5.2.3: The *ex vivo* expansion of UCB derived CD133+ HSPCs

Following the previous donor recovery and expansion study, we selected the 6 best donor expanders (610, 757, 875, 461, 453 and 456) to use in the presence of novel combinations of cytokines and small molecules in order to facilitate expansion of haematopoietic stem and progenitor cells (HSPCs). A breakdown of combinations can be observed in table 33. The gating strategy used to select HSPCs post expansion is presented in figure 51.

All combination treatments demonstrated a statistically significant lower fold expansion of total nucleated cells (between 7.89 ± 0.97 and 11.95 ± 1.25 fold) compared to basal (26.53 ± 5.63 fold) ($P \leq 0.0001$). The treatment which exhibited the highest decrease in fold change of total nucleated cells was treatment 3, with a fold change of 7.89 ± 0.97 fold (see figure 52).

By contrast, although not statistically significant, all treatments showed an increase in Lin-CD34+CD38-CD45RA-CD90+CD49f+ HSC (phenotypically defined Haematopoietic Stem Cells) fold expansion (between 93.33 ± 87.01 and 109.21 ± 123.55 fold) compared to basal (11.13 ± 10.49 fold) ($P > 0.05$). The treatment which exhibited the highest increase in fold change of Lin-CD34+CD38-CD45RA-CD90+CD49f+ HSCs was treatment 2, with a fold change of 109.21 ± 123.55 fold (see figure 53).

Fold expansion of the Lin-CD34+CD38-CD45RA-CD90- MPP (Multi-Potent Progenitor) cells in all treated cells was statistically significantly lower (between 1.00 ± 0.88 and 2.26 ± 1.31 fold) than in the cells grown on basal media (8.73 ± 3.31 fold) ($P \leq 0.001$). The treatment which exhibited the highest decrease in fold change of Lin-CD34+CD38-CD45RA-CD90- MPP was treatment 2, with a fold change of 109.21 ± 123.55 fold (see figure 54).

The Lin-CD34+CD38-CD45RA+CD90- CLP (Common Lymphoid Progenitor) cell population showed a statistically significant lower expansion following all combination treatments (between 3.91 ± 2.64 and 12.57 ± 9.94 fold), when compared to basal (105.24 ± 98.60 fold) ($P \leq 0.05$). The treatment which exhibited the highest decrease in fold change of Lin-CD34+CD38-CD45RA+CD90- CLP was treatment 3, with a fold change of 3.91 ± 2.64 fold (see figure 55).

Although not statistically significant, all combination treatments demonstrated a decrease in fold expansion of the Lin-CD34+CD38+CD123+CD45RA- CMP (Common Myeloid Progenitor) cell populations (between 2.45 ± 0.78 and 3.31 ± 0.94 fold), compared with basal (3.85 ± 0.70 fold) ($P>0.05$). The treatment which exhibited the highest decrease in fold change of Lin-CD34+CD38+CD123+CD45RA- CMP was treatment 3, with a fold change of 2.45 ± 0.78 fold (see figure 56).

For the Lin-CD34+CD38+CD123-CD45RA+ GMP (Granulocyte-Macrophage Progenitor) cell population, all treatments resulted in a statistically significant decrease in fold change (between 2.61 ± 0.10 and 5.48 ± 1.34 fold), compared with basal (16.22 ± 8.19 fold) ($P\leq0.01$). The treatment which exhibited the highest decrease in fold change of Lin-CD34+CD38+CD123-CD45RA+ GMP was treatment 3, with a fold change of 2.61 ± 0.10 fold (see figure 57).

Finally, all treatment combinations showed a non-significant increase in Lin-CD34+CD38+CD123-CD45RA- MEP (Megakaryocyte-Erythrocyte Progenitor) numbers (between 11.22 ± 6.11 and 14.10 ± 8.22 fold) compared with basal (9.72 ± 4.00 fold) ($P>0.05$). The treatment which exhibited the highest increase in fold change of Lin-CD34+CD38+CD123-CD45RA- MEP was treatment 6, with a fold change of 14.10 ± 8.22 fold (see figure 58).

Treatment	SCF	TPO	FLT3L	Scriptaid	ZLLYFMK	PEITC
Basal	D-1, D0, D2	D-1, D0, D2	D-1, D0, D2	-	-	-
(1) Scriptaid	D-1, D0, D2	D-1, D0, D2	D-1, D0, D2	D0, D2	-	-
(2) Scriptaid + ZLLYFMK 1	D-1, D0, D2	D-1, D0, D2	D-1, D0, D2	D0, D2	D0	-
(3) Scriptaid + ZLLYFMK 2	D-1, D0, D2	D-1, D0, D2	D-1, D0, D2	D0, D2	D0, D2	-
(4) Scriptaid + ZLLYFMK 3	D-1, D0, D2	D-1, D0, D2	D-1, D0, D2	D0, D2	D2	-
(5) Scriptaid + PEITC 1	D-1, D0, D2	D-1, D0, D2	D-1, D0, D2	D0, D2	-	D0
(6) Scriptaid + PEITC 2	D-1, D0, D2	D-1, D0, D2	D-1, D0, D2	D0, D2	-	D0, D2

Table 33: Treatment combinations used for the ex vivo expansion of UCB derived CD133+ HSPCs.
D=Day. Concentrations used were as follows: SCF (100ng/ml), TPO (20ng/ml), FLT3L (100ng/ml), Scriptaid (1µM), ZLLYFMK (10µM) and PEITC (10µM).

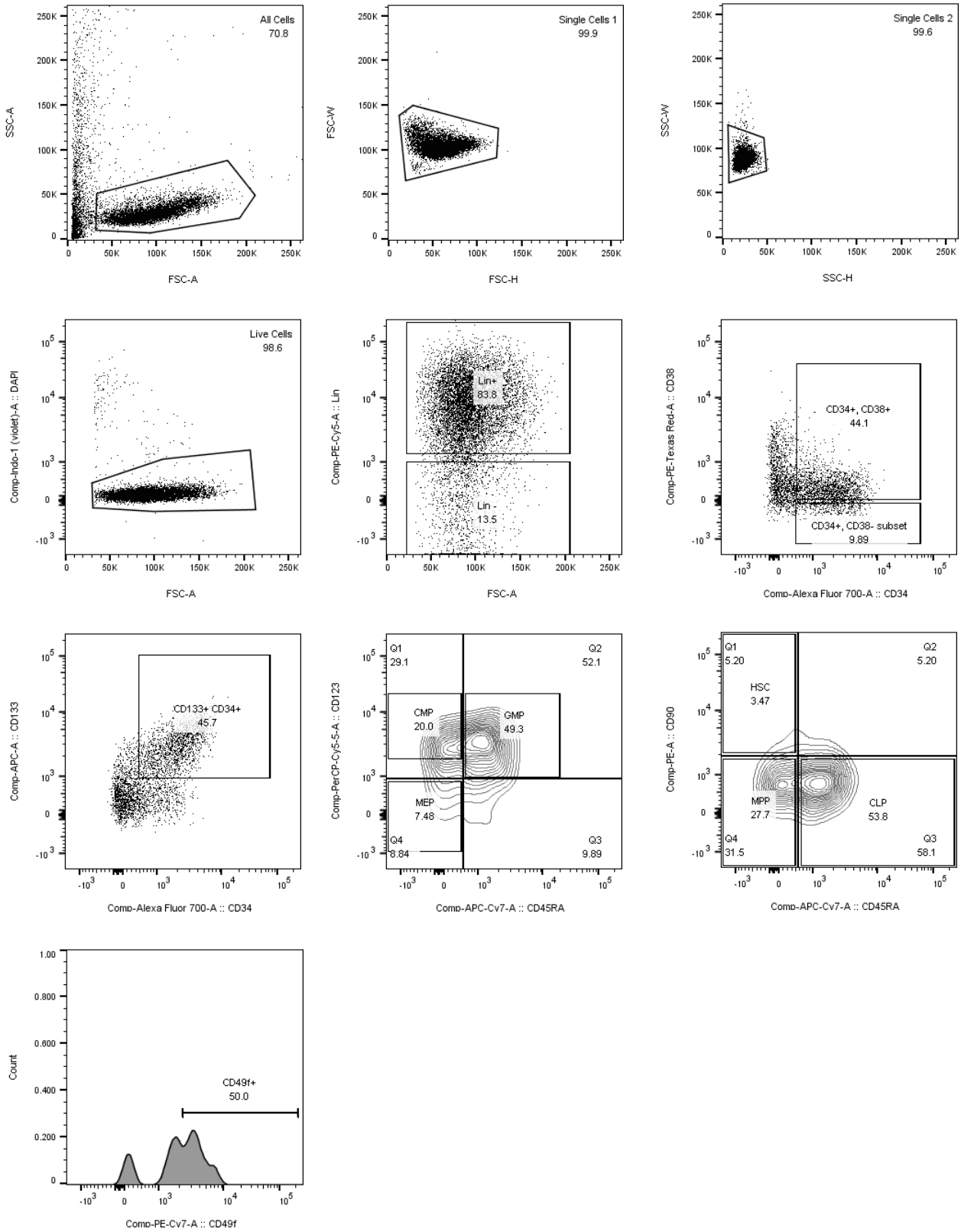
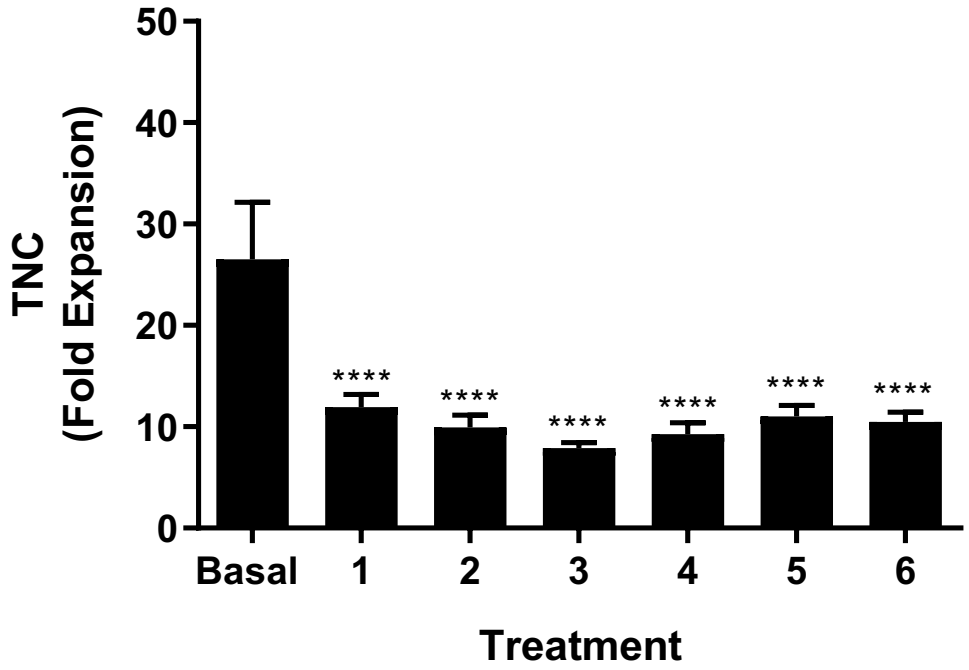
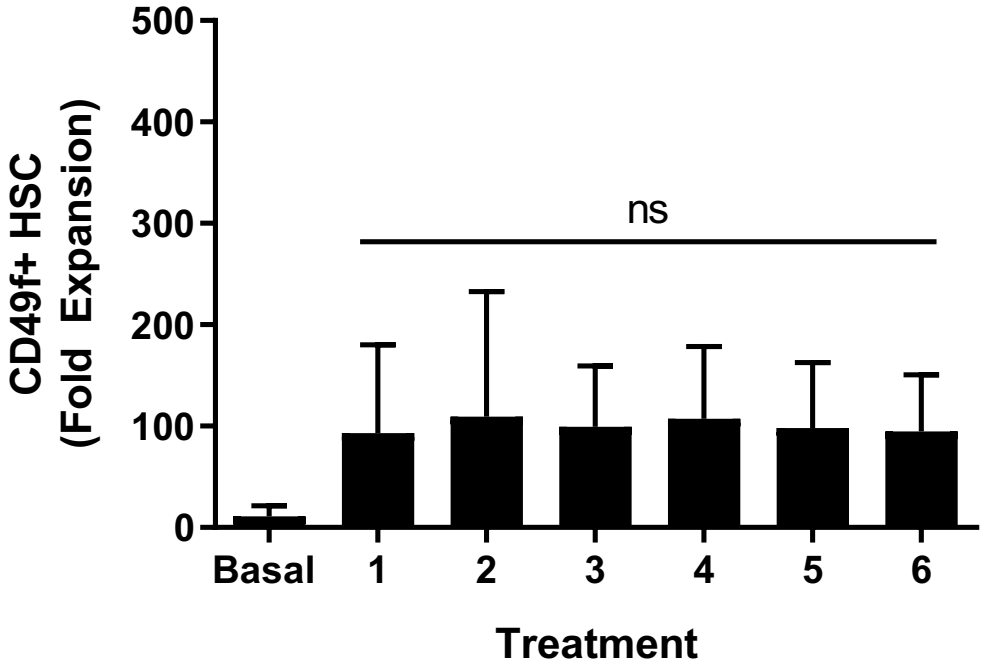


Figure 51: Representative gating strategy used to select HSPC subsets post expansion (from human UCB derived CD133+ cells) using small molecules and cytokines. *CMP (Common Myeloid Progenitor), GMP (Granulocyte-Monocyte Progenitor), MEP (Megakaryocyte-Erythrocyte Progenitor), HSC (Haematopoietic Stem Cell), MPP (Multi-Potent Progenitor) and CLP (Common Lymphoid Progenitor). The data are represented as a dot plot.*



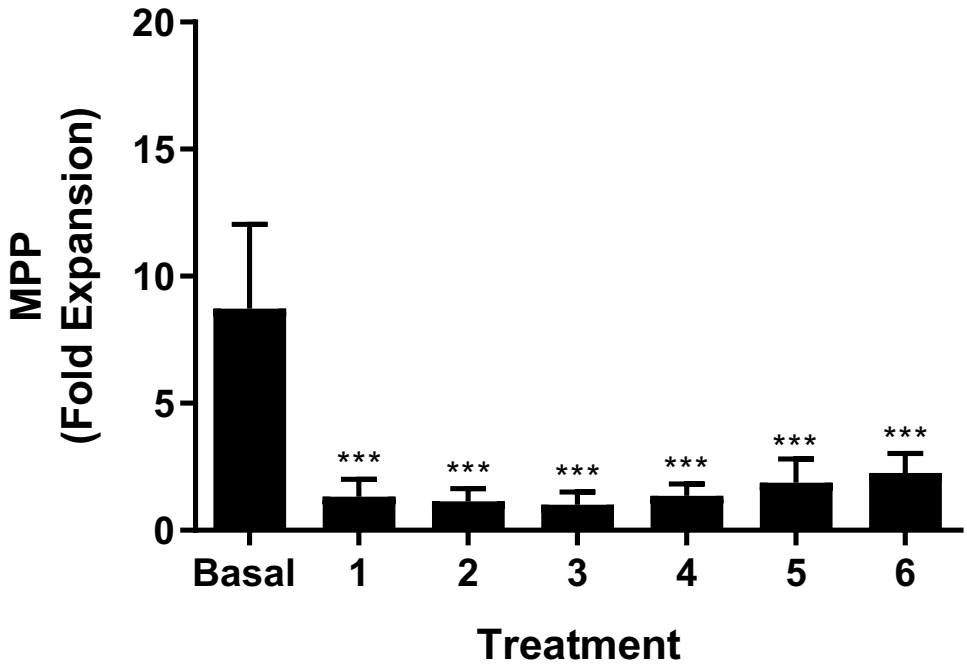
Treatment	Number
Basal	-
Scriptaid	1
Scriptaid + ZLLYFMK 1	2
Scriptaid + ZLLYFMK 2	3
Scriptaid + ZLLYFMK 3	4
Scriptaid + PEITC 1	5
Scriptaid + PEITC 2	6

*Figure 52: Total nucleated cell (TNC) fold expansion from fresh human umbilical cord blood using novel combinations of small molecules at day 5 of day 0. Data are expressed as mean ± SD, n=3 biological replicates. Treatments are compared to basal, (**** P≤0.0001). A one-way ANOVA test was used for statistical analysis, coupled with the Dunnett's multiple comparison post hoc test. Treatment descriptions can be observed in table 33.*



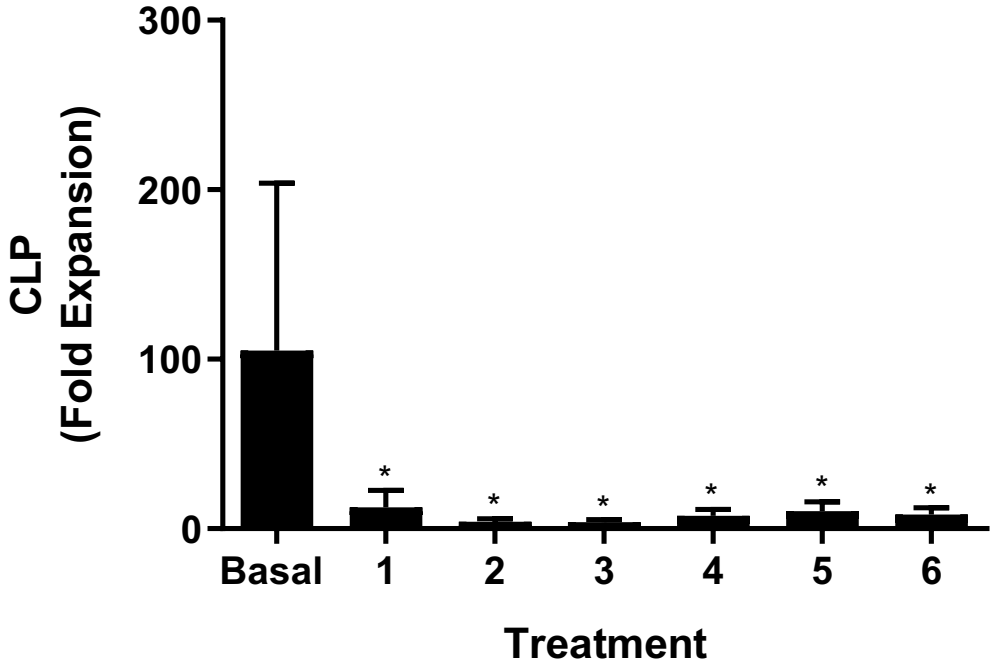
Treatment	Number
Basal	-
Scriptaid	1
Scriptaid + ZLLYFMK 1	2
Scriptaid + ZLLYFMK 2	3
Scriptaid + ZLLYFMK 3	4
Scriptaid + PEITC 1	5
Scriptaid + PEITC 2	6

Figure 53: Lin-CD34+CD38-CD45RA-CD90+CD49f+ HSC (Haematopoietic Stem Cells) fold expansion from fresh human umbilical cord blood using novel combinations of small molecules at day 5 of day 0. Data are expressed as mean ± SD, n=3 biological replicates. Treatments are compared to basal, (ns=not significant). A one-way ANOVA test was used for statistical analysis, coupled with the Dunnett's multiple comparison post hoc test. Treatment descriptions can be observed in table 33.



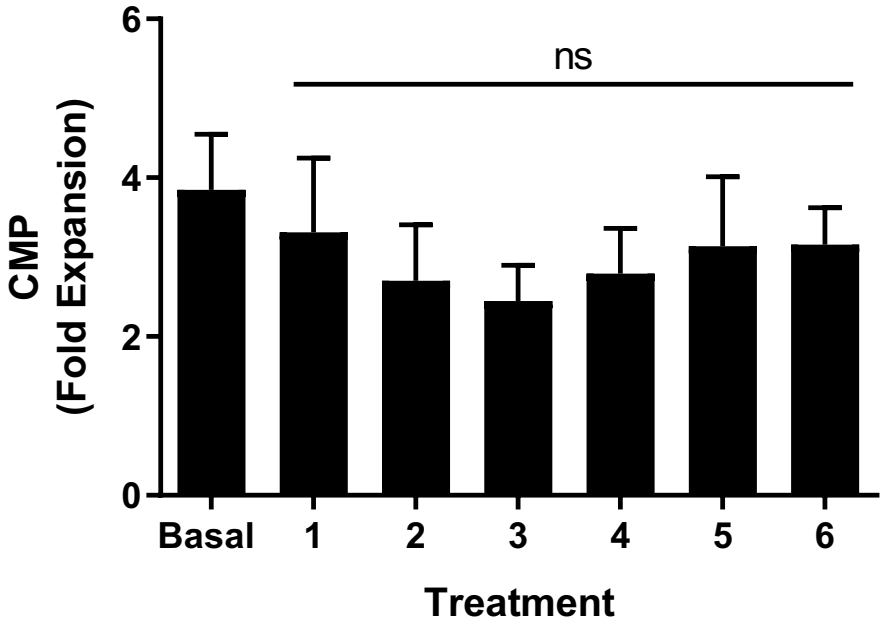
Treatment	Number
Basal	-
Scriptaid	1
Scriptaid + ZLLYFMK 1	2
Scriptaid + ZLLYFMK 2	3
Scriptaid + ZLLYFMK 3	4
Scriptaid + PEITC 1	5
Scriptaid + PEITC 2	6

*Figure 54: Lin-CD34+CD38-CD45RA-CD90- MPP (Multi-Potent Progenitor) fold expansion from fresh human umbilical cord blood using novel combinations of small molecules at day 5 of day 0. Data are expressed as mean ± SD, n=3 biological replicates. Treatments are compared to basal, (*** P≤0.001). A one-way ANOVA test was used for statistical analysis, coupled with the Dunnett's multiple comparison post hoc test. Treatment descriptions can be observed in table 33.*



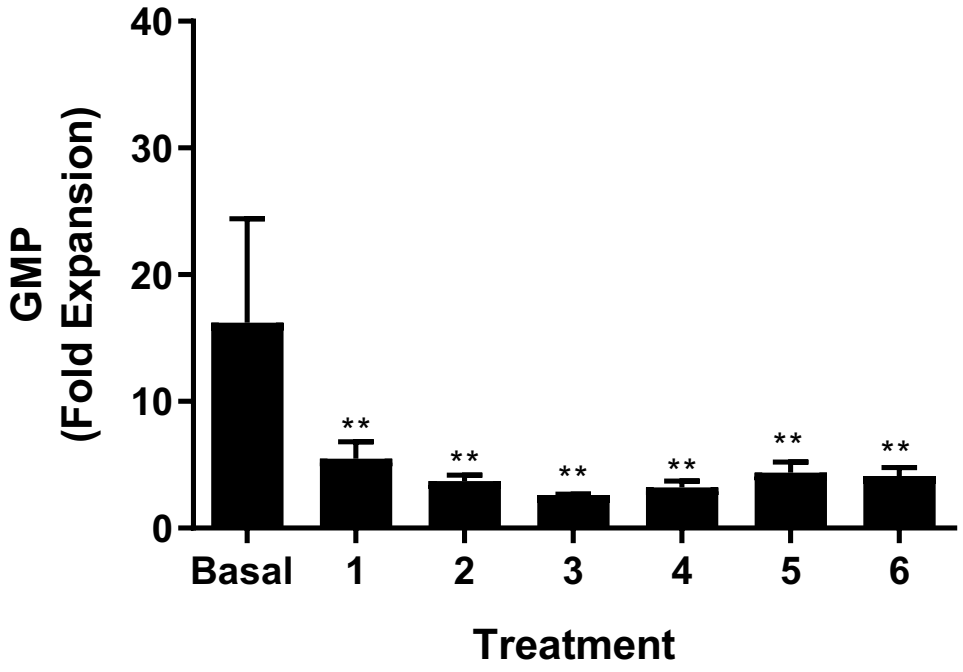
Treatment	Number
Basal	-
Scriptaid	1
Scriptaid + ZLLYFMK 1	2
Scriptaid + ZLLYFMK 2	3
Scriptaid + ZLLYFMK 3	4
Scriptaid + PEITC 1	5
Scriptaid + PEITC 2	6

Figure 55: Lin-CD34+CD38-CD45RA+CD90- CLP (Common Lymphoid Progenitor) fold expansion from fresh human umbilical cord blood using novel combinations of small molecules at day 5 of day 0. Data are expressed as mean ± SD, n=3 biological replicates. Treatments are compared to basal, (P≤0.05). A one-way ANOVA test was used for statistical analysis, coupled with the Dunnett's multiple comparison post hoc test. Treatment descriptions can be observed in table 33.*



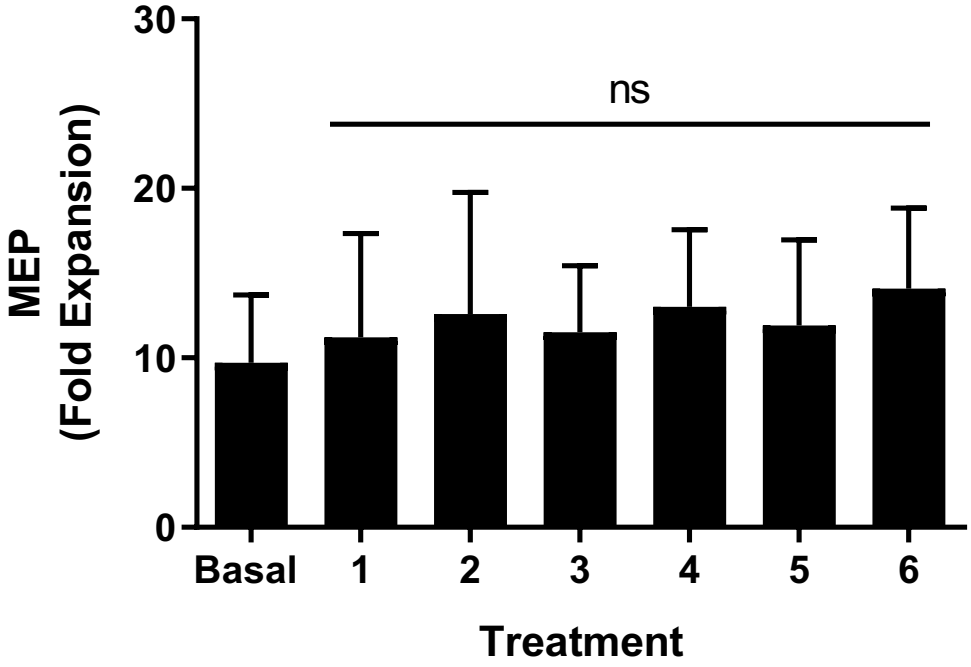
Treatment	Number
Basal	-
Scriptaid	1
Scriptaid + ZLLYFMK 1	2
Scriptaid + ZLLYFMK 2	3
Scriptaid + ZLLYFMK 3	4
Scriptaid + PEITC 1	5
Scriptaid + PEITC 2	6

Figure 56: Lin-CD34+CD38+CD123+CD45RA- CMP (Common Myeloid Progenitor) fold expansion from fresh human umbilical cord blood using novel combinations of small molecules at day 5 of day 0. Data are expressed as mean ± SD, n=3 biological replicates. Treatments are compared to basal, (ns=not significant). A one-way ANOVA test was used for statistical analysis, coupled with the Dunnett's multiple comparison post hoc test. Treatment descriptions can be observed in table 33.



Treatment	Number
Basal	-
Scriptaid	1
Scriptaid + ZLLYFMK 1	2
Scriptaid + ZLLYFMK 2	3
Scriptaid + ZLLYFMK 3	4
Scriptaid + PEITC 1	5
Scriptaid + PEITC 2	6

*Figure 57: Lin-CD34+CD38+CD123-CD45RA+ GMP (Granulocyte-Macrophage Progenitor) fold expansion from fresh human umbilical cord blood using novel combinations of small molecules at day 5 of day 0. Data are expressed as mean ± SD, n=3 biological replicates. Treatments are compared to basal, (**P≤0.01). A one-way ANOVA test was used for statistical analysis, coupled with the Dunnett's multiple comparison post hoc test. Treatment descriptions can be observed in table 33.*



Treatment	Numb er
Basal	-
Scriptaid	1
Scriptaid + ZLLYFMK 1	2
Scriptaid + ZLLYFMK 2	3
Scriptaid + ZLLYFMK 3	4
Scriptaid + PEITC 1	5
Scriptaid + PEITC 2	6

Figure 58: Lin-CD34+CD38+CD123-CD45RA- MEP (Megakaryocyte-Erythrocyte Progenitor) fold expansion from fresh human umbilical cord blood using novel combinations of small molecules at day 5 of day 0. Data are expressed as mean ± SD, n=3 biological replicates. Treatments are compared to basal, (ns=not significant). A one-way ANOVA test was used for statistical analysis, coupled with the Dunnett's multiple comparison post hoc test. Treatment descriptions can be observed in table 33.

5.2.4: The *ex vivo* expansion of UCB derived CD133+ HSPCs extended

The previous HSPC expansion screen described in section 5.2.3, was extended, but this time using novel combinations of additional molecules including UM171 (Fares, Chagraoui et al. 2017), SR1 (Wagner Jr, Brunstein et al. 2016), NAM (Nicotinamide) (Peled, Shoham et al. 2012) and VPA (Valproic Acid) (Bug, Gül et al. 2005). A breakdown of these treatment combinations can be observed in table 34. The gating strategy used to select HSPCs post expansion can be observed in figure 59.

All treatment combinations showed a statistically significant lower total nucleated cell (TNC) fold change (between 7.66 ± 2.45 and 28.71 ± 5.15 fold) compared to basal (38.18 ± 3.85 fold) ($P \leq 0.05$), ($P \leq 0.001$) and ($P \leq 0.0001$). The treatment which showed the highest significant fold change in TNC compared to basal (38.18 ± 3.85 fold) was treatment 8, with a fold change of 7.66 ± 2.45 ($P \leq 0.0001$) (see figure 60).

Although combination treatments showed no statistically significant change in Lin-CD34+CD38-CD45RA-CD90+CD49f+ HSC expansion compared to basal, some showed an increase, and some showed a decrease in expansion. Treatments 1-4 and 6-10 over 5 days *in vitro* culture showed an increase in Lin-CD34+CD38-CD45RA-CD90+CD49f+ HSC expansion (between 20.83 ± 28.68 and 187.08 ± 224.06 fold) compared to basal (20.67 ± 33.18 fold) ($P > 0.05$). The treatment which showed the highest increase in fold change of Lin-CD34+CD38-CD45RA-CD90+CD49f+ HSCs was treatment 1, with a fold change of 187.08 ± 224.06 fold compared to basal (20.67 ± 33.18 fold) ($P > 0.05$). By contrast, treatments 5 and 11 exhibited a lower fold change of CD34+CD38-CD45RA-CD90+CD49f+ HSCs (between 1.36 ± 1.90 and 20.25 ± 21.39), compared to basal (20.67 ± 33.18 fold) ($P > 0.05$). The treatment which exhibited the highest decrease in fold change of Lin-CD34+CD38-CD45RA-CD90+CD49f+ HSCs was treatment 11, with a fold change of 1.36 ± 1.90 fold compared to basal ($P > 0.05$) (see figure 61).

When compared to basal (5.77 ± 2.53 fold), although not significant, treatment 7 showed an increase of Lin-CD34+CD38-CD45RA-CD90- MPP (Multi-Potent Progenitor) fold expansion, with a fold change of 8.02 ± 2.26 ($P > 0.05$). All other treatments exhibited a decrease in expansion of Lin-

CD34+CD38-CD45RA-CD90- MPPs (between 0.09 ± 0.11 and 4.59 ± 1.59 fold), compared to basal (5.77 ± 2.53 fold) (P>0.05). Of the treatments which showed a decreased in expansion of Lin-CD34+CD38-CD45RA-CD90- MPP, treatment 1 (0.48 ± 0.42 fold) and treatment 8 (0.09 ± 0.11 fold) demonstrated a statistically significant decrease in fold change compared to basal (5.77 ± 2.53 fold) (P≤0.001). The treatment which showed the highest decrease in fold change of Lin-CD34+CD38-CD45RA-CD90- MPPs compared to Basel was treatment 8, with a fold change decrease of 0.09 ± 0.11 fold (P≤0.001) (see figure 62).

For Lin-CD34+CD38-CD45RA+CD90- CLP (Common Lymphoid Progenitor) cell populations, when compared to basal (69.28 ± 15.80 fold), treatment 7 showed a none significant increase of Lin-CD34+CD38-CD45RA+CD90- CLP fold expansion, with a fold change of 70.81 ± 10.26 (P>0.05). All other treatments exhibited a decrease in fold expansion of Lin-CD34+CD38-CD45RA+CD90- CLPs (between 0.56 ± 0.64 and 57.79 ± 16.86 fold), compared to basal (69.28 ± 15.80 fold) (P≤0.01), (P≤0.001) and (P≤0.0001). Of the treatments which showed a decreased in expansion of Lin-CD34+CD38-CD45RA+CD90- CLPs, treatment 1 (3.36 ± 2.95 fold) (P≤0.0001), treatment 2 (20.42 ± 5.87 fold) (P≤0.0001), treatment 4 (13.16 ± 3.13 fold) (P≤0.0001), treatment 5 (27.23 ± 8.79 fold) (P≤0.001), treatment 6 (30.28 ± 13.08 fold) (P≤0.001), treatment 8 (0.56 ± 0.64 fold) (P≤0.0001), treatment 9 (8.44 ± 7.12 fold) (P≤0.0001), treatment 10 (14.00 ± 2.76 fold) (P≤0.0001) and treatment 11 (33.04 ± 14.09 fold) (P≤0.01) demonstrated a statistically significant decrease in fold change compared to basal (69.28 ± 15.80 fold). The treatment which showed the highest decrease in fold change of Lin-CD34+CD38-CD45RA+CD90- CLPs compared to basal was treatment 8, with a fold change decrease of 0.56 ± 0.64 fold (P≤0.0001) (see figure 63).

Although combination treatments showed no statistically significant change in Lin-CD34+CD38-CD45RA-CD90- CMP expansion compared to basal, some showed an increase, and some showed a decrease in expansion. Treatments 2-6 showed an increase in Lin-CD34+CD38-CD45RA-CD90- CMP expansion (between 1.52 ± 1.63 and 5.29 ± 7.33 fold) compared to basal (0.51 ± 0.56 fold) (P>0.05). The treatment which showed the highest increase in fold change of Lin-CD34+CD38-CD45RA-

CD90- CMP was treatment 4, with a fold change of 5.29 ± 7.33 fold, compared to basal (0.51 ± 0.56 fold) (P>0.05). By contrast, treatment(s) 1 and 7-11, exhibited a lower fold change of Lin-CD34+CD38-CD45RA-CD90- CMP (between 0 ± 0 and 0.28 ± 0.34 fold), compared to basal (0.51 ± 0.56 fold) (P>0.05). The treatment which exhibited the highest decrease in fold change of Lin-CD34+CD38-CD45RA-CD90- CMP was treatment 9, with a fold change of 0 ± 0 fold, compared to basal (0.51 ± 0.56 fold) (P>0.05) (see figure 64).

For the Lin-CD34+CD38-CD45RA-CD90- GMP cell sub type, although some treatments demonstrated an increase or decrease in fold expansion compared to basal, none of these were statistically significant changes. Treatments 3, 5 and 6 showed an increase in Lin-CD34+CD38-CD45RA-CD90- GMP expansion (between 8.39 ± 6.23 and 15.05 ± 9.43 fold) compared to basal (7.14 ± 4.48 fold) (P>0.05). The treatment when compared to basal, which showed the highest increase in fold change of Lin-CD34+CD38-CD45RA-CD90- GMP was treatment 3, with a fold change of 15.05 ± 9.43 fold (P>0.05). By contrast, treatment(s) 1, 2, 4 and 7-11 exhibited a lower fold change of Lin-CD34+CD38-CD45RA-CD90- GMP (between 0.66 ± 0.44 and 6.68 ± 2.81 fold), compared to basal (7.14 ± 4.48 fold) (P>0.05). The treatment which exhibited the highest decrease in fold change of Lin-CD34+CD38-CD45RA-CD90- GMP, compared to basal was treatment 8, with a fold change of 0.66 ± 0.44 fold (P>0.05) (see figure 65).

Although combination treatments showed no statistically significant change in Lin-CD34+CD38+CD123-CD45RA- MEP expansion compared to basal, some showed an increase, and some showed a decrease in expansion. Treatment(s) 1 and 7-9 showed an increase in Lin-CD34+CD38+CD123-CD45RA- MEP expansion (between 8.76 ± 5.28 and 10.34 ± 2.91 fold) compared to basal (7.11 ± 2.89 fold) (P>0.05). The treatment which showed the highest increase in fold change of Lin-CD34+CD38+CD123-CD45RA- MEP was treatment 7, with a fold change of 10.34 ± 2.91 fold, compared to basal (7.11 ± 2.89 fold) (P>0.05). By contrast, treatment(s) 2-6, 10 and 11 exhibited a lower fold change of Lin-CD34+CD38+CD123-CD45RA- MEP (between 3.24 ± 1.06 and 7.07 ± 3.06 fold), compared to basal (7.11 ± 2.89 fold) (P>0.05). The treatment, compared to basal, which exhibited the

highest decrease in fold change of Lin CD34+CD38+CD123-CD45RA- MEP was treatment 11, with a fold change of 3.24 ± 1.06 fold ($P>0.05$) (see figure 66).

Treatment	SCF	TPO	FLT3L	Scriptaid	UM171	SR1	NAM	VPA	PEITC	ZLLYFMK
Basal	D-1, D0, D2	D-1, D0, D2	D-1, D0, D2	-	-	-	-	-	-	-
(1) Scriptaid	D-1, D0, D2	D-1, D0, D2	D-1, D0, D2	D-1, D0	-	-	-	-	-	-
(2) UM171	D-1, D0, D2	D-1, D0, D2	D-1, D0, D2	-	D0, D2	-	-	-	-	-
(3) SR1	D-1, D0, D2	D-1, D0, D2	D-1, D0, D2	-	-	D0, D2	-	-	-	-
(4) UM171 + SR1 1	D-1, D0, D2	D-1, D0, D2	D-1, D0, D2	-	D0, D2	D0, D2	-	-	-	-
(5) UM171 + SR1 2	D-1, D0, D2	D-1, D0, D2	D-1, D0, D2	-	D0	D2	-	-	-	-
(6) UM171 + SR1 3	D-1, D0, D2	D-1, D0, D2	D-1, D0, D2	-	D2	D0	-	-	-	-
(7) NAM (Nicotinamide)	D-1, D0, D2	D-1, D0, D2	D-1, D0, D2	-	-	-	D0, D2	-	-	-
(8) VPA (Valproic Acid)	D-1, D0, D2	D-1, D0, D2	D-1, D0, D2	-	-	-	-	D0, D2	-	-
(9) Scriptaid + ZLLYFMK + PEITC	D-1, D0, D2	D-1, D0, D2	D-1, D0, D2	D0	-	-	-	-	D2	D0
(10) Scriptaid + PEITC	D-1, D0	D-1, D0	D-1, D0	D0	-	-	-	-	D2	-
(11) ZLLYFMK + PEITC	D-1, D0	D-1, D0	D-1, D0	-	-	-	-	-	D2	D0

Table 34: Treatment combinations used for UCB derived CD133+ HSPC expansion. D=Day. Concentrations used were as follows: SCF (100ng/ml), TPO (20ng/ml), FLT3L (100ng/ml), Scriptaid (1µM), UM171 (35nM), SR1 (500nM), NAM (5mM), VPA (1mM), PEITC (10µM) and ZLLYFMK (10µM).

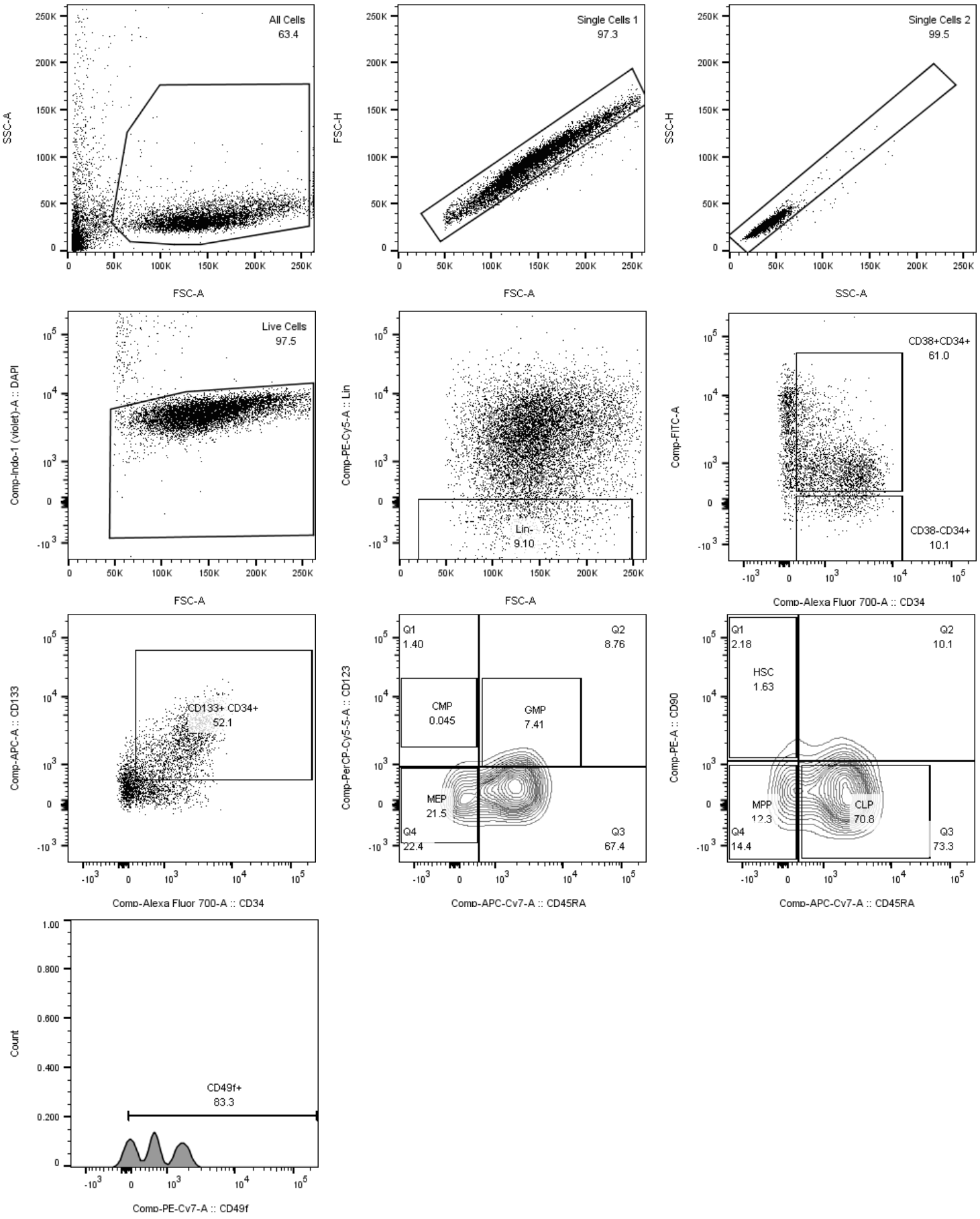
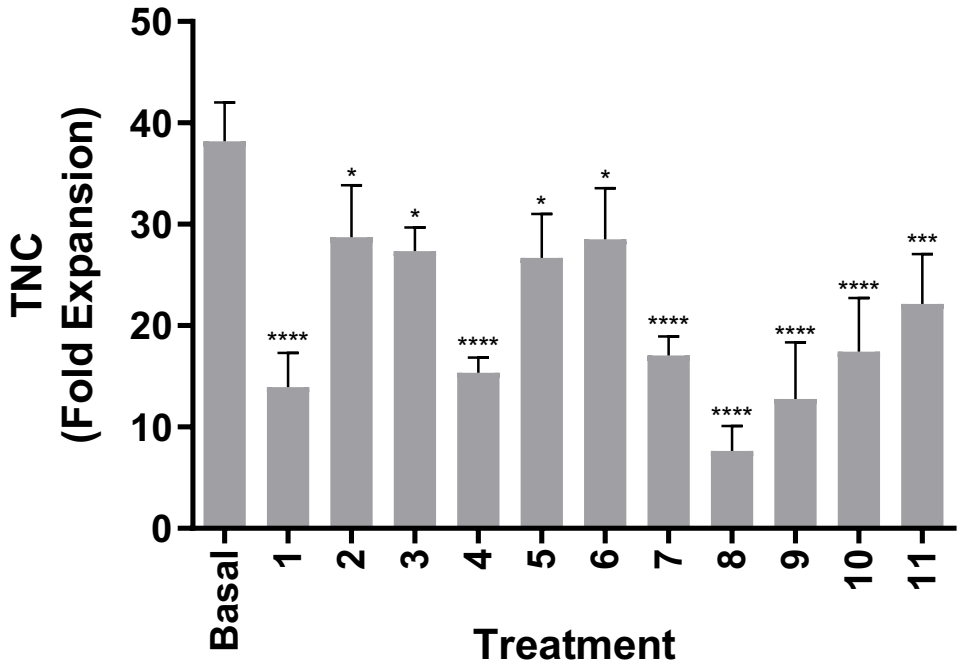
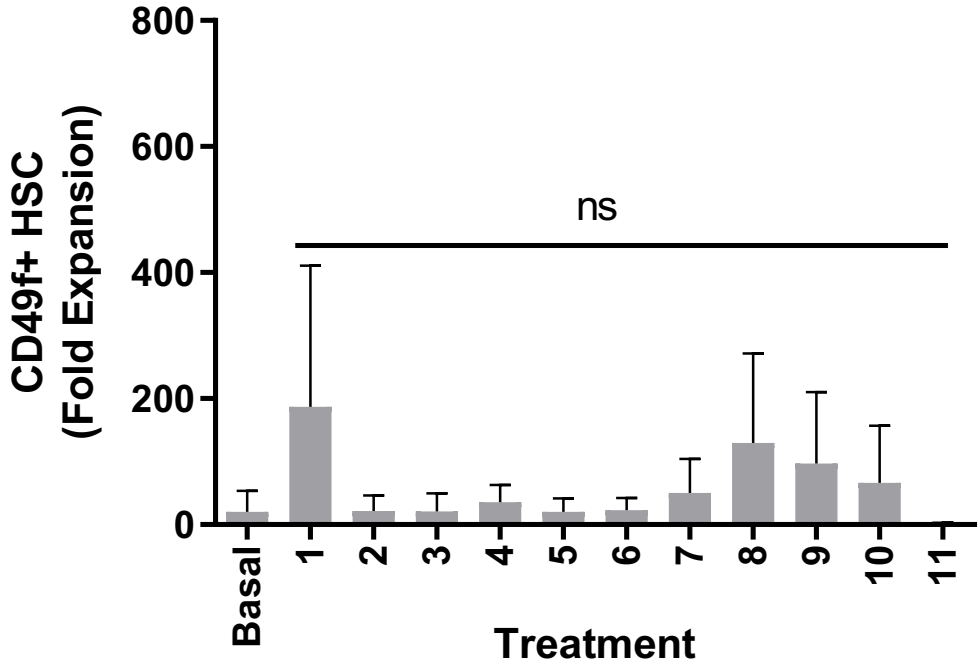


Figure 59: Representative gating strategy used to select HSPC subsets post expansion (from human UCB derived CD133+ cells) using small molecules and cytokines. (CMP) Common Myeloid Progenitor, GMP (Granulocyte-Macrophage Progenitor), (MEP) Megakaryocyte-Erythrocyte Progenitor, HSC (Haematopoietic Stem Cell), (MPP) Multi-Potent Progenitor and (CLP) Common Lymphoid Progenitor. The data are represented as a dot plot.



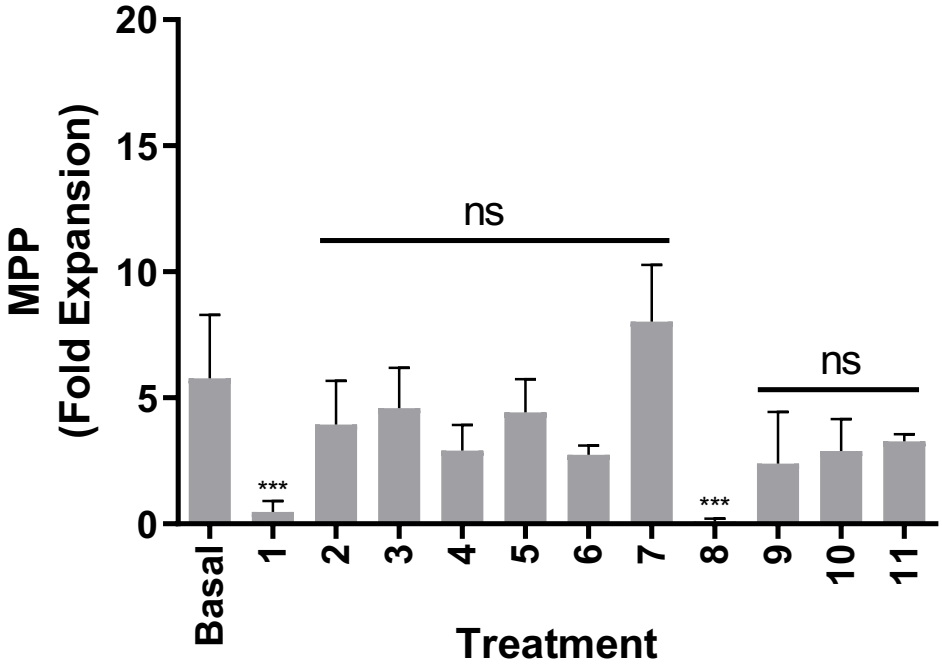
Treatment	Number
Basal	-
Scriptaid	1
UM171	2
SR1	3
UM171 + SR1 1	4
UM171 + SR1 2	5
UM171 + SR1 3	6
NAM (Nicotinamide)	7
VPA (Valproic Acid)	8
Scriptaid + ZLLYFMK + PEITC	9
Scriptaid + PEITC	10
ZLLYFMK + PEITC	11

Figure 60: Total nucleated cell (TNC) fold expansion from fresh human umbilical cord blood using novel combinations of small molecules at day 5 of day 0. *Data are expressed as mean ± SD, n=3 biological replicates. Treatments are compared to basal, (* P≤0.05), (*** P≤0.001), (**** P≤0.0001). A one-way ANOVA test was used for statistical analysis, coupled with the Dunnett's multiple comparison post hoc test. Treatment descriptions can be observed in table 34.*



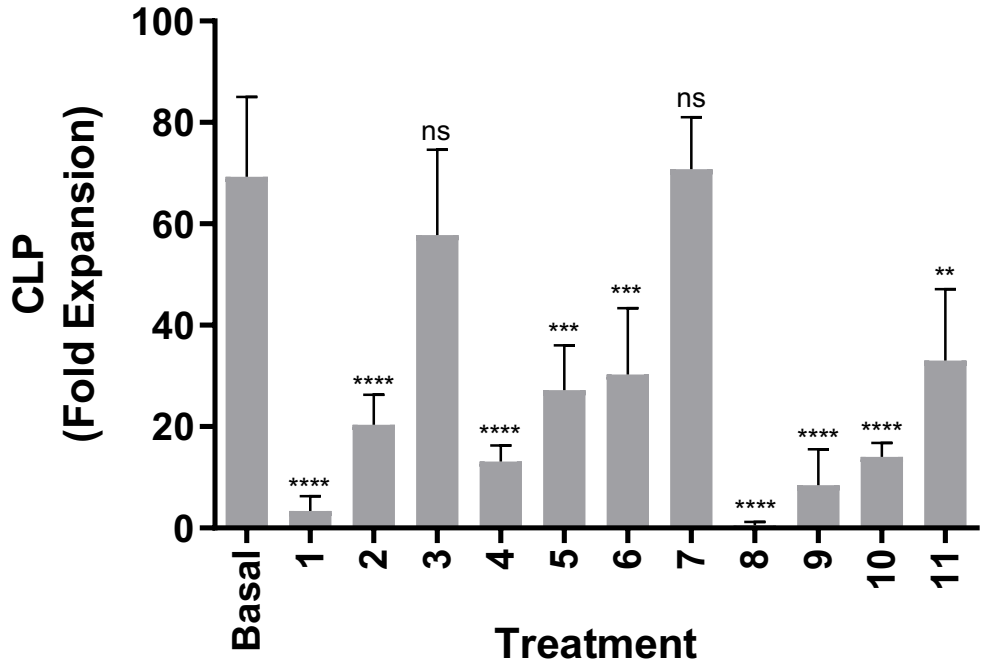
Treatment	Number
Basal	-
Scriptaid	1
UM171	2
SR1	3
UM171 + SR1 1	4
UM171 + SR1 2	5
UM171 + SR1 3	6
NAM (Nicotinamide)	7
VPA (Valproic Acid)	8
Scriptaid + ZLLYFMK + PEITC	9
Scriptaid + PEITC	10
ZLLYFMK + PEITC	11

Figure 61: Lin-CD34+CD38-CD45RA-CD90+CD49f+ HSC (Haematopoietic Stem Cells) fold expansion from fresh human umbilical cord blood using novel combinations of small molecules at day 5 of day 0. Data are expressed as mean ± SD, n=3 biological replicates. Treatments are compared to basal, (ns=not significant). A one-way ANOVA test was used for statistical analysis, coupled with the Dunnett's multiple comparison post hoc test. Treatment descriptions can be observed in table 34.



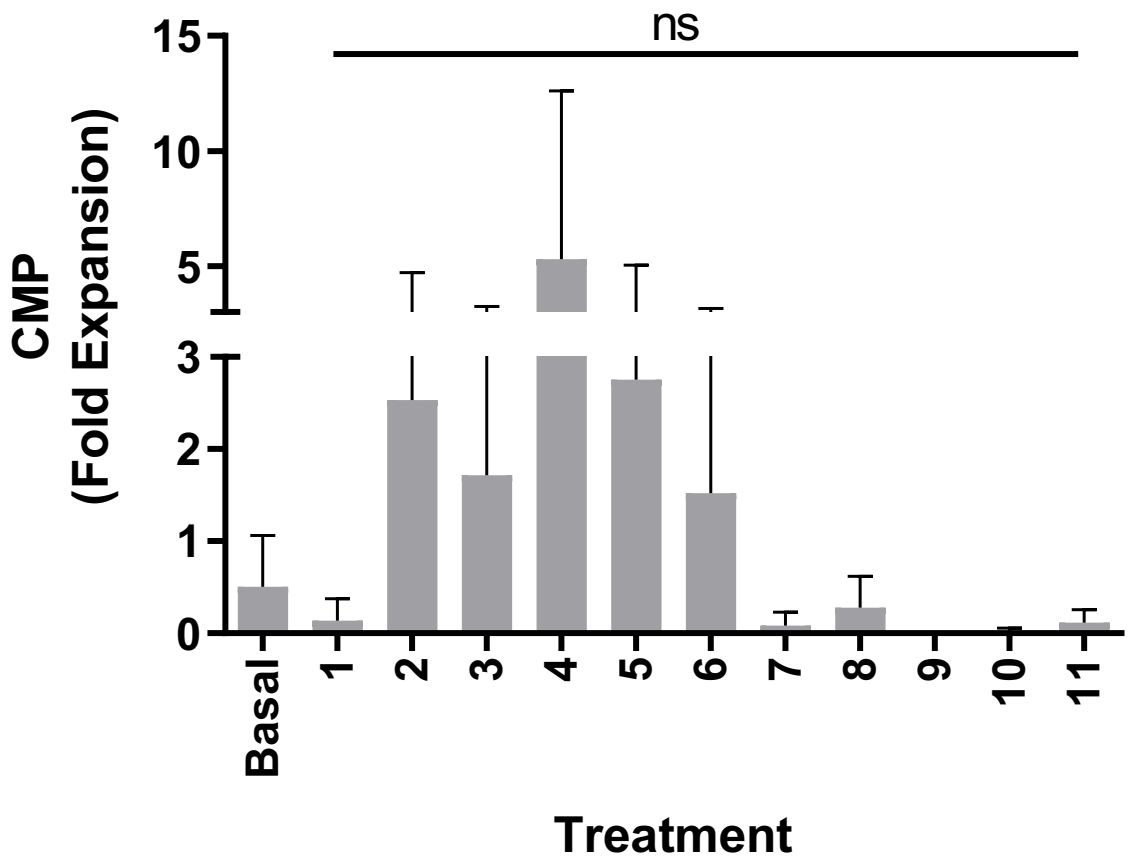
Treatment	Number
Basal	-
Scriptaid	1
UM171	2
SR1	3
UM171 + SR1 1	4
UM171 + SR1 2	5
UM171 + SR1 3	6
NAM (Nicotinamide)	7
VPA (Valproic Acid)	8
Scriptaid + ZLLYFMK + PEITC	9
Scriptaid + PEITC	10
ZLLYFMK + PEITC	11

Figure 62: Lin-CD34+CD38-CD45RA-CD90- MPP (Multi-Potent Progenitor) fold expansion from fresh human umbilical cord blood using novel combinations of small molecules at day 5 of day 0. Data are expressed as mean ± SD, n=3 biological replicates. Treatments are compared to basal, (ns=not significant). A one-way ANOVA test was used for statistical analysis, coupled with the Dunnett's multiple comparison post hoc test. Treatment descriptions can be observed in table 34.



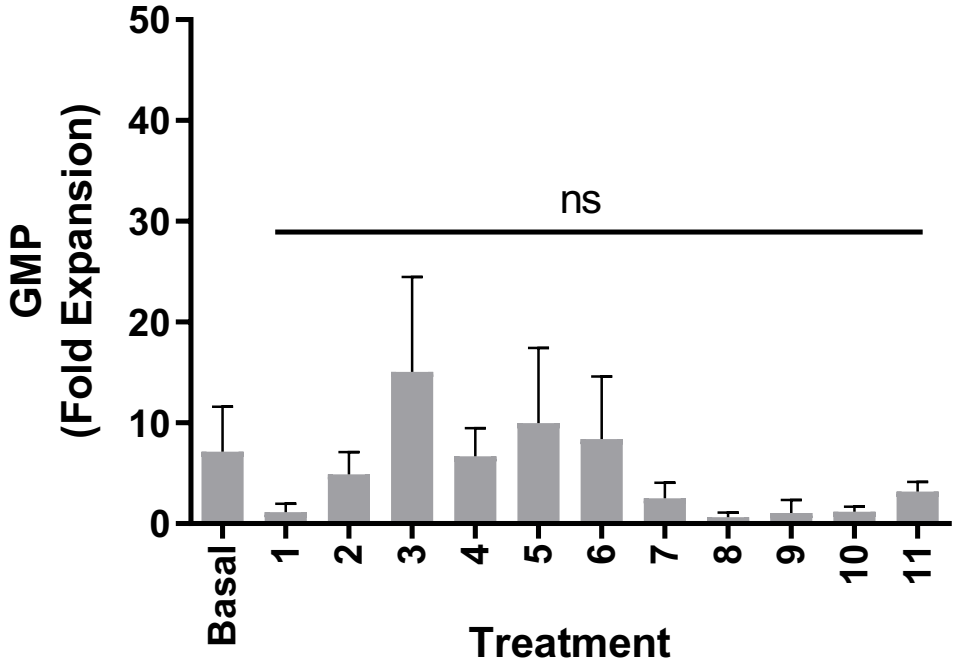
Treatment	Number
Basal	-
Scriptaid	1
UM171	2
SR1	3
UM171 + SR1 1	4
UM171 + SR1 2	5
UM171 + SR1 3	6
NAM (Nicotinamide)	7
VPA (Valproic Acid)	8
Scriptaid + ZLLYFMK + PEITC	9
Scriptaid + PEITC	10
ZLLYFMK + PEITC	11

*Figure 63: Lin-CD34+CD38-CD45RA+CD90- CLP (Common Lymphoid Progenitor) fold expansion from fresh human umbilical cord blood using novel combinations of small molecules at day 5 of day 0. Data are expressed as mean ± SD, n=3 biological replicates. Treatments are compared to basal, (ns=not significant), (**P≤0.01), (*** P≤0.001), (**** P≤0.0001). A one-way ANOVA test was used for statistical analysis, coupled with the Dunnett's multiple comparison post hoc test. Treatment descriptions can be observed in table 34.*



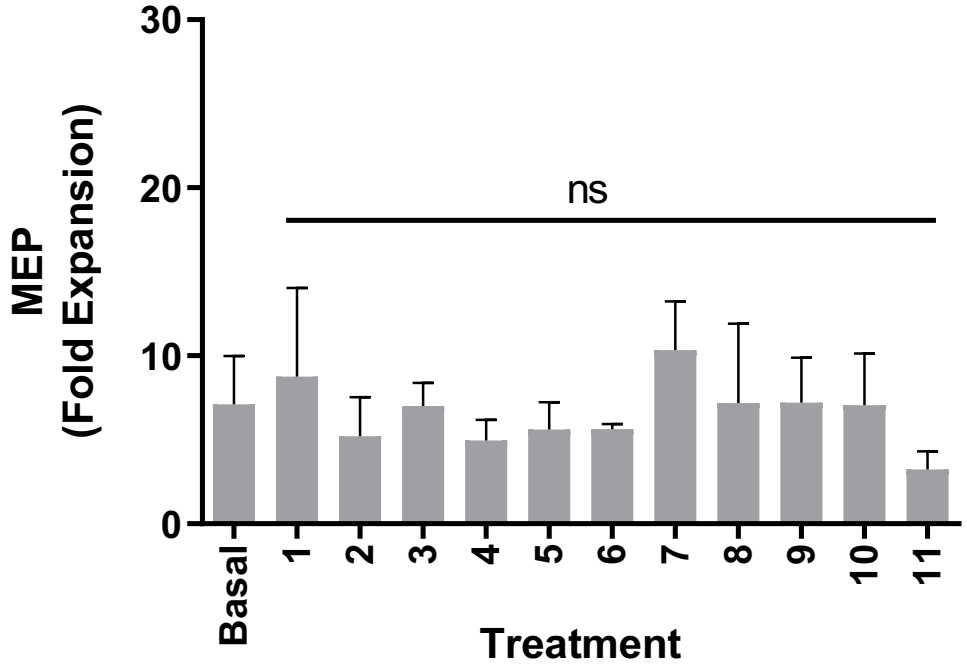
Treatment	Number
Basal	-
Scriptaid	1
UM171	2
SR1	3
UM171 + SR1 1	4
UM171 + SR1 2	5
UM171 + SR1 3	6
NAM (Nicotinamide)	7
VPA (Valproic Acid)	8
Scriptaid + ZLLYFMK + PEITC	9
Scriptaid + PEITC	10
ZLLYFMK + PEITC	11

Figure 64: Lin-CD34+CD38+CD123+CD45RA- CMP (Common Myeloid Progenitor) fold expansion from fresh human umbilical cord blood using novel combinations of small molecules at day 5 of day 0. *Data are expressed as mean ± SD, n=3 biological replicates. Treatments are compared to basal, (ns=not significant). A one-way ANOVA test was used for statistical analysis, coupled with the Dunnett's multiple comparison post hoc test. Treatment descriptions can be observed in table 34.*



Treatment	Number
Basal	-
Scriptaid	1
UM171	2
SR1	3
UM171 + SR1 1	4
UM171 + SR1 2	5
UM171 + SR1 3	6
NAM (Nicotinamide)	7
VPA (Valproic Acid)	8
Scriptaid + ZLLYFMK + PEITC	9
Scriptaid + PEITC	10
ZLLYFMK + PEITC	11

Figure 65: Lin-CD34+CD38+CD123-CD45RA+ GMP (Granulocyte-Macrophage Progenitor) fold expansion from fresh human umbilical cord blood using novel combinations of small molecules at day 5 of day 0. Data are expressed as mean ± SD, n=3 biological replicates. Treatments are compared to basal, (ns=not significant). A one-way ANOVA test was used for statistical analysis, coupled with the Dunnett's multiple comparison post hoc test. Treatment descriptions can be observed in table 34.



Treatment	Number
Basal	-
Scriptaid	1
UM171	2
SR1	3
UM171 + SR1 1	4
UM171 + SR1 2	5
UM171 + SR1 3	6
NAM (Nicotinamide)	7
VPA (Valproic Acid)	8
Scriptaid + ZLLYFMK + PEITC	9
Scriptaid + PEITC	10
ZLLYFMK + PEITC	11

Figure 66: Lin-CD34+CD38+CD123-CD45RA- MEP (Megakaryocyte-Erythrocyte Progenitor) fold expansion from fresh human umbilical cord blood using novel combinations of small molecules at day 5 of day 0. Data are expressed as mean ± SD, n=3 biological replicates. Treatments are compared to basal, (ns=not significant). A one-way ANOVA test was used for statistical analysis, coupled with the Dunnett's multiple comparison post hoc test. Treatment descriptions can be observed in table 34.

5.3: Discussion

A main, but future, aim of this project is to conduct genome engineering on primary HSCs, in order to model myeloid malignancy. To facilitate this, the main objective of the current study was to attempt to find optimal combinations of small molecules and cytokines to expand UCB derived HSPCs *ex vivo* with the aim of increasing HSC numbers. These cells could then be used as a model for genome engineering approaches. To this aim, CD133+ cells were isolated from fresh human UCB, and their expansion was investigated after exposure to different novel combinations of small molecules and cytokines, which have been shown to individually influence proliferation, survival and differentiation of these cells previously.

The UCB derived CD133+ cell donor samples to be tested were selected based on their viability and purity. A basal cocktail of SCF, TPO and FLT3L was used to initially test the expansion of the 12 donor samples selected. This combination of cytokines was selected because previous work in our laboratory demonstrated their increased efficacy in the *ex vivo* expansion of CD133+ cells. All 12 donor CD133+ cell samples demonstrated high viability. Although all donor samples were viable (greater than 89%), their expansion following treatment with SCF, TPO and FLT3L varied. This variability is expected from expansion studies and has previously been reported (Gullo, van der Garde et al. 2015).

The 6 best expanders, which were combined into 3 donor pools, were used for further *ex vivo* expansion experiments. Novel combinations of small molecules and cytokines were used to expand human UCB derived HSPCs. In the initial screen the molecules used were SCF, TPO, FLT3L, Scriptaid, ZLLYFMK, and PEITC.

In the initial combination *ex vivo* expansion, although the majority of combined treatments were able to alter the number of TNCs and HSPCs, between treatment groups containing Scriptaid there was no statistically significant change. Since each combination treatment included Scriptaid on Day 0 and Day 2 of culture, this histone deacetylase (HDAC) inhibitor appeared to have the major influence on the expansion of the phenotypically defined Lin-CD34+CD38-CD45RA-CD90+CD49f+ HSCs. This finding is also now well documented in the literature (Hua, Kronsteiner et al. 2019).

Chaurasia and colleagues tested the ability of several HDAC inhibitors to expand human UCB HSCs *ex vivo* (Chaurasia, Gajzer et al. 2014). They treated human UCB derived CD34+ cells over a 7-day period using SCF, TPO, FLT3L and IL-3 in combination with trichostatin A, suberoylanilide hydroxamic acid, VPA, Scriptaid, CAY10433, CAY10398 or CAY10603 HDAC inhibitors (Chaurasia, Gajzer et al. 2014). They demonstrated that Scriptaid and VPA have the highest effect on CD34+CD90+ HSC cell expansion in the presence of cytokines, compared to cytokines alone. They also showed that VPA in the presence of cytokines for 7 days expanded long term repopulating cells when transplanted into NSG mice by 36 fold compared with primary UCB derived CD34+ cells (Chaurasia, Gajzer et al. 2014). In addition, it was shown by Bug and colleagues that in HSCs, VPA increases both proliferation and self-renewal of HSCs (Bug, Gül et al. 2005).

Our data demonstrated that phenotypically defined Lin-CD34+CD38-CD45RA-CD90+CD49f+ HSC numbers increased by 93.33 ± 87.01 to 109.21 ± 123.55 fold compared to basal whenever Scriptaid was present. In the study by Chaurasia and colleagues the phenotypic markers differed from our experiment; they consisted of CD34, CD90, CD184, CD117, CD49f, and CD45RA, moreover the treatment combination used by Chaurasia and colleagues included IL-3, whereas our study consisted of SCF, TPO and FLT3L (Gajzer, Ross et al. 2016). Furthermore, our treatment regimens were conducted on different days, and therefore may have resulted in a different result. We have demonstrated expansion of phenotypically defined Lin-CD34+CD38-CD45RA-CD90+CD49f+ HSC using HDAC inhibition, which is in line with the literature, and these data suggest that HDAC inhibition may be a viable strategy to expand CD133+ cells prior to genome editing.

We conducted a further screen to compare compounds used in the initial UCB derived CD133+ HSPC *ex vivo* expansion combination screen with a variety of other small molecules described in the literature as having the efficacy to expand human UCB derived HSPCs *ex vivo*. These included combinations of SCF, TPO, FLT3L, Scriptaid, UM171, SR1, NAM, VPA, ZLLYFMK and PEITC. Combination treatments of these molecules were able to alter the number of phenotypically defined HSPCs *ex vivo.* Like the initial screen, we observed that Scriptaid was able to increase phenotypically defined Lin-

CD34+CD38-CD45RA-CD90+CD49f+ HSCs numbers, as also observed by Chaurasia and colleagues (Gajzer, Ross et al. 2016). However, the variability found in these particular experiments was such that our results failed to reach statistical significance.

We further observed that combination treatments of Scriptaid + ZLLYFMK + PEITC and Scriptaid + PEITC also increase absolute Lin-CD34+CD38-CD45RA-CD90+CD49f+ HSCs numbers compared to basal. Having said that, the largest effect was observed with Scriptaid alone. With the exception of MEPs, in the presence of Scriptaid, alone or in combination, all HSPC subsets showed a smaller fold change increase compared to basel.

In addition to Scriptaid, we decided to test the expansion efficacy of VPA, another HDAC inhibitor. In agreement with Chaurasia and colleagues (Chaurasia, Gajzer et al. 2014), we showed that VPA treatment increases fold expansion of Lin-CD34+CD38-CD45RA-CD90+CD49f+ HSCs compared to basal. Again however, there was substantial variability between the UCB expansions. Mahmud and colleagues also showed that VPA has efficacy in the expansion of human UCB derived CD34+CD90+ cells (Mahmud, Petro et al. 2014). Similar to Scriptaid, with the exception of MEPs, all HSPC subsets showed a smaller fold change increase, in the presence of VPA, compared to basal. A study was conducted by Papa and Colleagues (Papa, Zimran et al. 2018) in which they demonstrated that *ex vivo* treatment of UCB derived CD34+ cells with VPA increased the number of HSCs in culture. A similar finding was reported in a study conducted by Zimran and colleagues (Zimran, Papa et al. 2020), where they treated cells isolated from bone marrow and peripheral blood with VPA to demonstrate an expansion of putative CD34+CD45RA-CD90+CD49f+ cells.

It has been reported that priming HSPCs *in vitro* with chromatin modifying agents such as VPA and Scriptaid together with cytokines under serum-free-conditions significantly enhances HSC numbers following transplantation into NSG mice (Papa, Zimran et al. 2018). Further studies from our laboratory (Hua, Kronsteiner et al. 2019) assessed the effects on human CD133+ HSPCs over 5 days of culture with Scriptaid and cytokines, SCF, TPO and FL, and detected an absolute increase in Lin-CD34+CD38-CD45RA-CD90+CD49f+ putative HSC numbers, coinciding with a decrease in the Lin-CD38-

CD34+CD45RA-CD90- subset. These studies showed that most of the CD90- cells had become reprogrammed to CD90+ HSPCs, without other significant differences in their transcriptomic profile, which suggests that Scriptaid maintains stem cell gene expression programmes despite expansion in HSC numbers. Likewise, Papa and colleagues have also reported that VPA rapidly reprogrammed the CD90- HSPCs to CD90+ cells, while maintaining transcriptomic and mitochondrial profiles that resemble those in the primary HSCs (Papa, Zimran et al. 2018). Like our initial expansion screen, we have shown expansion of phenotypically defined Lin-CD34+CD38-CD45RA-CD90+CD49f+ HSC using HDAC inhibition, specifically Scriptaid and VPA. Our findings further confirm that using HDAC inhibition to expand CD133+ cells prior to genome editing, may be a viable strategy. One method to explore HDAC inhibition further would be to use different classes of HDAC inhibitors in the same cocktail, to see if there is synergy, since specific subtypes of HDACs might be more important to target.

We also tested combinations of ZLLYFMK, a calpain inhibitor, since it has been reported to be important in HSC expansion (Kale and Limaye 2010). ZLLYFMK was tested in combination with, SCF, TPO, FLT3L, as well as Scriptaid and PEITC. The best combination within this group to expand Lin-CD34+CD38-CD45RA-CD90+CD49f+ HSCs was Scriptaid + ZLLYFMK + PEITC. Furthermore, with the exception of MEP, we observed a decrease in TNC or HPC numbers following treatment with the combination of ZLLYFMK with Scriptaid and PEITC in the presence of SCF, TPO and FLT3L compared to basal treatment alone.

The ability of calpain inhibition to increase HSC numbers is well documented in the literature (Zhang and Gao 2016). Sangeetha and colleagues investigated the importance of calpain inhibition in the expansion of human UCB HSPCs (Kale and Limaye 2010). They reported that ZLLYFMK treatment of human UCB derived CD34+ cells increased total cell yield by up to 70 fold following 10 days in culture. They also observed that following double immunostaining of CD34 and CD33, 70% of the cells that expanded in the presence of ZLLYFMK were myeloid cells. The increase in expansion may be a consequence of a decrease in apoptosis, since they also report that expression of Annexin V and caspase 3, key pro apoptotic mediators, were significantly reduced compared to untreated cells. These

data suggest that inhibiting calpain may have efficacy in the expansion of HSCs, since inhibiting calpain may reduce the level of apoptotic and necrotic cell death. Our findings align with the literature in that we also showed an increase of Lin-CD34+CD38-CD45RA-CD90+CD49f+ in the presence of calpain inhibition. However, Scriptaid treatment in isolation was able to expand Lin-CD34+CD38-CD45RA-CD90+CD49f+ more so than the Scriptaid + ZLLYFMK + PEITC cocktail. This may be explained due to high levels of endogenous calpastatin, which may occupy calpain binding sites, causing ZLLYFMK to compete for binding. This could be experimentally determined by conducting dose response experiments with ZLLYFMK in isolation, in the presence of calpastatin depletion using siRNA (Liu, Schneider et al. 2011). Similar studies could be conducted using PEITC in isolation and will be discussed later in the chapter.

We also decided to test UM171 and SR1, individually and in combination. UM171 was identified by Fares and colleagues following a high throughput screen of a compound library. Following 7-day culture of mobilised peripheral blood (MPB) derived CD34+CD45RA- cells in the presence of UM171, they report a 100-fold increase in CD34+CD38-CD90+CD49f+ cells compared to control groups (Fares, Chagraoui et al. 2014). We observed a slight increase in Lin-CD34+CD38-CD45RA-CD90+CD49f+ UCB HSC numbers following treatment with UM171 over 5 days of culture. This could in part reflect a differential effect of UM171 on MPB versus UCB, and/or the high donor variability we observed in our studies. With the exception of CMPs, the expansion of HPCs was lower in our UM171 treated samples when compared to basal.

Furthermore, we decided to test expansion of HSCs in the presence of SR1, an Aryl hydrocarbon receptor inhibitor identified by Boitano and colleagues (Boitano, Wang et al. 2010). We observed that SR1 treated cells showed only a small increase in Lin-CD34+CD38-CD45RA-CD90+CD49f+ HSC numbers. A study by Wagner and colleagues (Wagner Jr, Brunstein et al. 2016) demonstrated that CD34+CD133+CD90+ HSC expansion for 15 days in the presence of SR1 increased by 330 fold compared to control groups. Similar to UM171, with the exception of CMPs, the expansion of HPCs were lower in SR1 treated samples when compared to basal.

Combinations of UM171 and SR1 exhibited variable results of Lin-CD34+CD38-CD45RA-CD90+CD49f+ HSC numbers compared to basal. The best combination of UM171 SR1 was treatment 4, which exhibited a 1.7 fold increase in expansion compared to basal over 5 days of culture. Fares and colleagues reported that SR1 enhanced UM171 mediated expansion of CD34+CD38-CD90+CD49f+ cells (Ngom, Imren et al. 2018). With the exception of CMPs, the expansion of HPCs were lower in all combinations of UM171 and SR1 treated samples when compared to basal. Lower expansion of Lin-CD34+CD38-CD45RA-CD90+CD49f+ HSCs was observed in the present study in the presence of UM171 or SR1, compared to the studies already discussed. This may have been due to concentrations used and the length, method and frequency of treatment. In the study conducted by Fares and colleagues (Ngom, Imren et al. 2018), they used 35nM UM171 or 750nM SR1 (500nM SR1 for combined). In the study conducted by Wagner and colleagues (Wagner Jr, Brunstein et al. 2016), they used 750nm SR1 in the presence of 50ng/ml SCF, TPO, FLT3L and IL-6. By contrast, we used 500nM SR1 (since Fares and colleagues (Ngom, Imren et al. 2018) used it for their combination screen), 100ng/ml SCF, 20ng/ml TPO and 100ng/ml FLT3L. Fares and colleagues (Ngom, Imren et al. 2018) used a batch fed cell culture system spanning 7 days and Wagner and colleagues (Wagner Jr, Brunstein et al. 2016) treated cells for 15 days, whereas we conducted treatments on Day 0, 2 and 4 of a 5-day culture period. It would be of interest to recreate the experiments conducted by Fare and colleagues (Fares, Chagraoui et al. 2014) and by Boitano and Colleagues (Boitano, Wang et al. 2010), but in the presence of Scriptaid, since in our study it showed the greatest expansion of Lin-CD34+CD38-CD45RA-CD90+CD49f+ HSCs.

We also decided to include NAM, a potent SIRT1 inhibitor, in our expansion screen since it has been implicated to be important in expansion of HSPCs (Peled, Shoham et al. 2012). We observed that NAM treatment increased Lin-CD34+CD38-CD45RA-CD90+CD49f+ HSC expansion compared to basal, although not significantly given the expansion variability of the UCB samples used. With the exception of CMP and GMP, the expansion of HPCs were higher in NAM treated samples when compared to basal. Peled and colleagues reported that, in the presence of NAM together with SCF, TPO and FLT3L, there was an increase in CD34+CD38- cells derived from human UCB, whereas there were fewer

numbers of differentiated cells compared to control conditions (Peled, Shoham et al. 2012). The phenomenon was also described by Horwitz and colleagues (Horwitz, Chao et al. 2014). Since NAM is a broad spectrum SIRT inhibitor, and inhibits SIRT1, SIRT2, SIRT3, SIRT5 and SIRT6 it would be interesting to use a subtype specific SIRT inhibitor, to see if targeting specific SIRT subtypes leads to a more pronounced Lin-CD34+CD38-CD45RA-CD90+CD49f+ HSC expansion (Hu, Jing et al. 2014). It would also be interesting to expand Lin-CD34+CD38-CD45RA-CD90+CD49f+ HSC cells in the presence of Scriptaid and NAM, since Scriptaid showed the most profound expansion of Lin-CD34+CD38-CD45RA-CD90+CD49f+ HSC.

Finally, in the presence of combinations containing PEITC, a Notch activator, Lin-CD34+CD38-CD45RA-CD90+CD49f+ HSC expansion was variable. Studies conducted by Csaszar and colleagues demonstrated that DLL1, a Notch Activator, significantly increased CD34+CD90+ cells over 16-day culture period. They also reported that DLL1 treatment reduced total cell number (Csaszar, Wang et al. 2014). One possible explanation for the low variable response to treatments containing PEICT was that the concentration used was too low. Further experiments could be aimed at using varied concentrations of PEICT along with other Notch inhibitors such as DLL1, in isolation and in combination.

Chapter 6
General Discussion

6.1: General discussion

The objective of this thesis was to investigate the role of *TET2* and *ASXL1* in myeloid malignancies, and to this objective there were three central aims: 1) To generate cell lines using the CRISPR/Cas9 system expressing mutations in *TET2* and *ASXL1*, both individually and as co mutations; 2) To test the consequences of these mutations on key cellular function; 3) To expand CD133+ haematopoietic stem and progenitor cells as a prelude to gene editing of *TET2* and *ASXL1*. To these aims, we successfully engineered HUDEP2 and KBM5 cell lines expressing mutations in *TET2* and/or *ASXL1*, as individual or co mutations. The cellular consequences of these mutations were successfully investigated, and finally we successfully expanded CD133+ using novel combinations of cytokines and small molecules, as a prelude to gene editing.

We used the CRISPR/Cas9 system (Qi et al., 2013) to engineer cell lines with mutations in *TET2* and/or *ASXL1*. The CRISPR/Cas9 system was used previously in our laboratory, which was used to correct mutations in *ASXL1* with high specificity and with a low incidence of off targets, so for direct comparative analysis and congruence in our cellular models we decided to use CRISPR/Cas9 for this study (Valletta, Dolatshad et al. 2015).

Initially we successfully achieved a *TET2* KO in HUDEP2 cell lines with zero off-targets identified. HUDEP2 cells were selected to represent the erythroid component of myeloid malignancies and the use of which has been well documented (Vinjamur and Bauer 2018). Moreover, KBM5 cells pre-engineered in our laboratory (Valletta, Dolatshad et al. 2015) containing an existing *ASXL1* mutation were also mutated to express a *TET2* KO. KBM5 cells were selected to represent the myeloid component of myeloid malignancies.

Furthermore, in the HUDEP2 with the *TET2* KO, using CRISPR/Cas9, more specifically using HDR, we also introduced an *ASXL1* mutation. The *TET2* KO efficiency we achieved was 8% in HUDEP2 cells and 5% in KBM5 cells. The actual efficiency was extremely low compared to the predicted efficiency. Whereas the *ASXL1* mutation efficiency in HUDEP2 cells was 5%. One reason for this is that the predictive software used to design the sgRNA does not account for the behaviours of a biological system. In addition, this software does not account for cellular death, which may be why efficiency appears to be low. The CRISPR/Cas9 system is known to induce off-target effects (Wienert, Wyman et al. 2019). However, no sequence alterations were detected in the predicted off target genes examined in all the *TET2* and *ASXL1* mutant clones engineered for this thesis.

Ito and colleagues demonstrated that TET2 loss leads to myeloid and lymphoid disorders in mice (Ito, Lee et al. 2019). By contrast, it was shown by Abdel-Wahab and colleagues that mutations in ASXL1 promoted myeloid leukemogenesis (Abdel-Wahab, Adli et al. 2012).

In order to investigate the effect these mutations have on cellular function, we subsequently conducted functional studies. For both our HUDPE2 cell and KBM5 cell models, we observed an increase in proliferation for the mutant variants of each cell line, with the co-mutation exhibiting an increase in proliferation in each cell line and suggests that the co mutations, TET2 and ASXL1, may be having an additive effect. These findings are in line with the literature (Wang, Xiao et al. 2015, Youn, Kim et al. 2017, Li, Zhou et al. 2018). Moreover, it has also been shown that TET2 is recruited by WT1 to mediate proliferation of cells in AML (Wang, Xiao et al. 2015).

So, it may be a possibility, if TET2 is acting through other factors, to target both TET2 and secondary factors this gene may interact with, preventing crosstalk (Wang, Xiao et al. 2015).

Next we investigated the cell cycle, particularly in our HUDEP2 cell model, we observed that WT cells were observed more in early parts of the cell cycle, with mutant cells being more prevalent in latter parts of the cell cycle, suggesting the mutations have a pro proliferative effect. Interestingly, in our HUDEP2 cell models, we also observed the co-mutations had a more pronounced effect, with

an increase in cell cycle progression. These data suggest the effects on *TET2* and *ASXL1* may be additive.

Similarly, in our KBM5 cell model, the mutant cell variants showed less prevalence in the early stages of the cell cycle. However, the mutations increased the number of cells in later parts of the cell cycle, which also suggests that the cells are cycling more and that the mutations may have a pro proliferative effect. As with HUDEP2 cells, our KBM5 model demonstrated an additive effect of *TET2* and *ASXL1* in cell cycle progression. These data are also in line with our proliferation studies described above. Therapy to treat myeloid malignancy could be targeted at both genes, a so-called combination therapy approach. Mahfoudhi and colleagues report that *TET2* activity slows down S-Phase cell cycle progression in a P53 dependent manner (Mahfoudhi, Talhaoui et al. 2016). It was reported by Fujino and colleagues, that AXSL1 regulates the cell cycle and cell proliferation through the Akt/mTOR pathway (Fujino, Goyama et al. 2020). These reports show that *TET2* and *ASXL1* have an indirect effect on cell cycle progression, suggesting that targeting multiple mechanisms may achieve a more profound effect.

The ability to evade apoptosis is a well-known hallmark of cancer (Hanahan and Weinberg 2011). We observed the percentage of apoptotic cells was less in our HUDEP2 *TET2* KO mutant variant compared to WT. By contrast in our HUDEP2 cell model, the *ASXL1* and co-mutation variants showed a higher percentage of apoptotic cells. In our KBM5 cell model, the percentage of apoptotic cells was higher in the *TET2* KO mutant variant, compared to WT. By contrast, the *ASXL1* and co-mutation cell variants showed a lower percentage of apoptotic cells. In our HUDEP2 cell models, our findings are in line with the literature, in which *TET2* loss leads to a decrease in apoptosis, whereas the *ASXL1* mutation leads to an increase in apoptosis compared to WT (Wang, Li et al. 2014, Deng, Wang et al. 2016, Cai, Kotzin et al. 2018).

In our KBM5 cell model, the opposite was observed, in which *ASXL1* mutants showed lower percentages of apoptotic cells compared to WT, whereas the *TET2* mutant showed an increase in apoptosis. This may have been due to other drivers of apoptosis in the cell, and their expression levels.

In addition, KBM5 cells were showing more mature markers compared to WT in our differentiation studies, which suggest that our KBM5 cell culture populations may have consisted of mainly mature cells, having a compensatory apoptotic mechanism. A further investigatory approach could be to explore the effect of *TET2* and *ASXL1* mutations on early and late apoptosis, as this may provide a more specific insight into the result observed in KBM5 cells. Moreover, it would be interested to expand these experiments to other myeloid cell lines, in order to determine the consistency of the effects of TET2 and ASXL1 mutations on apoptosis.

It is well established that differentiation block is a key hallmark of myeloid neoplasms (Tenen 2014). We observed that early erythroid markers showed higher expression in most mutant variants of our HUDEP2 cell model. Moreover, the expression of more mature erythroid markers was low in all mutant variants of our HUDEP2 cell model. By contrast, in our KBM5 model, expression of early myeloid markers was lower than expression of mature myeloid markers in our mutant variants. It has been shown that *TET2* absence is able to delay differentiation in HSCs (Ko, Bandukwala et al. 2011). It has also been shown that mutations in *ASXL1* leading to a truncation inhibits myeloid differentiation (Inoue, Kitaura et al. 2013). Although our HUDEP2 model data are in line with the literature, the reasons why our KBM5 model data are showing opposing results may be attributed to our KBM5 models being expanded from a single clone. Another consideration might be to use a larger panel of differentiation markers, in order to obtain a more comprehensive picture of the differentiation stage.

In the final part of this thesis, we examined the ability of novel combinations of cytokines and small molecules to expand CD133+ cells. In our initial screen we used Scriptaid, a potent HDAC inhibitor (Hua, Kronsteiner et al. 2019), in combination ZLLYFMK a Calpain inhibitor (Kale and Limaye 2010) and PEITC, which is a Notch activator and known anticancer agent (Chan and Miskimins 2012). Scriptaid alone or in combination with ZLLYFMK or PEITC showed an increased percentage of Lin-CD34+CD38-CD45RA-CD90+CD49f+ phenotypically defined HSCs. It was also observed that Scriptaid alone or in combination showed a decrease in the percentage of early and late progenitor cells compared to Basal treatment (cytokines but no small molecule inhibitors). This suggests that Scriptaid

along with ZLLYFMK and PEITC are acting to maintain the CD133+ cells as HSCs thereby blocking erythroid and myeloid differentiation.

We continued our expansion study by including combinations of known HSC expansion agents. These were: UM171 (Fares, Chagraoui et al. 2017), SR1 (Wagner Jr, Brunstein et al. 2016), VPA (HDAC inhibitor) (Bug, Gül et al. 2005) and NAM (SIRT1 inhibitor) (Peled, Shoham et al. 2012). Treatments containing HDAC inhibitors had the most profound expansion of Lin-CD34+CD38-CD45RA-CD90+CD49f+ phenotypically defined HSC numbers, and this was subsequently shown by Hua et al. to involve the 'reprogramming of CD90- MMP to CD90+ 'phenotypically defined' HSCs (Hua, Kronsteiner et al. 2019). Lower expansion of early and late progenitor cells was observed in the presence of treatments containing HDAC inhibitors, except for MEPs. It was demonstrated by Papa and colleagues that treatment with VPA, a potent HDAC inhibitor was able to increase HSC numbers *ex vivo* at least in terms of the ability of the expanded cells to serially repopulate immunodeficient mice (Papa, Zimran et al. 2018). Moreover, it was demonstrated by Bug and colleagues that VPA prevented differentiation of HSCs in a dose dependant manner (Bug, Gul et al. 2004). In addition, Hua and colleagues showed that treatment with Scriptaid increased the number of Lin−CD34+CD38−CD45RA−CD90+CD49f+ HSPCs, *ex vivo* (Hua, Kronsteiner et al. 2019). Therefore, our expansion data for HDAC inhibitors is in line with further studies now established in the literature.

Treatments containing combinations of UM171 and SR1 over a 5 day period had little or no effect on Lin-CD34+CD38-CD45RA-CD90+CD49f+ HSC expansion, but did however show an increase in the number of early and late progenitor cells. It was demonstrated by Fares and colleagues that UM171 enabled a profound *in vivo* expansion of UCB cells (Fares, Chagraoui et al. 2014). Cohen and colleagues conducted a clinical trial to investigate the feasibility of using UM171 to expand UCB CD34+ cells over a 7 day period, prior to transplantation into patients with haematological malignancies (Cohen, Roy et al. 2020). They demonstrated that 96% of UCB units were expanded using UM171, with no graft failure following transplant (median follow up was 18 months) (Cohen, Roy et al. 2020).

Boitano and colleagues showed that SR1 promotes the *ex vivo* expansion of CD34+ cells (Boitano, Wang et al. 2010). A clinical trial was conducted by Wagner and colleagues to investigate the effectiveness of SR1 to expand UCB CD34+ cells prior to transplant (Wagner, Brunstein et al. 2015). They observed that SR1 was able to expand UCB CD34+ cells by 328-fold, in addition SR1 expanded UCB CD34+ cells drastically increased early haematopoietic recovery (Wagner, Brunstein et al. 2015).

One explanation why we saw the opposite effects to these studies (Fares, Chagraoui et al. 2014) (Boitano, Wang et al. 2010) may have been due to the method of treatment. Fares and colleagues used a UM171 batch feed system over a 12 day period, whereas we used a 5 day protocol, in which the cells were treated with fresh factor on day(s) 0 and 2 only. Similarly, Boitano and colleagues (Boitano, Wang et al. 2010) treated their culture for 21 days with SR1, whereas we treated our CD133+ cells day(s) 0 and 2 with SR1, and ran our experiment for 5 days.

SIRT1 inhibition using NAM was also able to increase Lin-CD34+CD38-CD45RA-CD90+CD49f+ HSC fold change. In addition, NAM treatment was able to increase the fold expansion of all early and late progenitor cells except CMPs. Peled and colleagues demonstrated that treatment of NAM in UCB CD34+ increased expansion and inhibited differentiation (Peled, Shoham et al. 2012). A clinical trial was conducted by Horwitz and colleagues to investigate the safety and efficacy of NAM to expand UCB CD34+ cells prior to transplant. They observed that NAM was able to expand UCB CD34+ cells safely prior to transplant (Horwitz, Wease et al. 2019).

In the presence of a calpain inhibitor (ZLLYFMK) and Notch activator (PEITC) there was very little expansion of Lin-CD34+CD38-CD45RA-CD90+CD49f+ phenotypically defined HSCs. By contrast, expansion of early and late progenitor cells, with the exception of CMPs, was high. Kale and colleagues showed that the use of ZLLYFMK, a potent calpain inhibitor, improved expansion of CD34+ cells (Kale and Limaye 2010). It was shown by Satyan and colleagues that PEITC treatment inhibited differentiation through a c-Myc dependent manner (Satyan, Swamy et al. 2006). In addition a Notch mediated expansion of HSPCs was demonstrated by Delaney and colleagues (Delaney, Heimfeld et al. 2010). Furthermore, Notch ligand was used to expand UCB CD34+ cells prior to transplant into human

subjects in a phase 1 trial, showing enhanced rates of myeloid engraftment with a shortened time to neutrophil recovery (Delaney, Heimfeld et al. 2010). One explanation for not seeing HSC expansion observed by these researchers (Kale and Limaye 2010) (Delaney, Heimfeld et al. 2010) may have also been due to a difference in culture protocol. Our CD133+ cell expansion in the presence of ZLLYFMK and PEITC ran over a 5 day protocol whereas (Kale and Limaye 2010) used a 10 day expansion protocol and (Delaney, Heimfeld et al. 2010) used a 21 day protocol. From these studies, it was evident that there are multiple methods to expand HSCs. With complex mechanisms of action governing HSC homeostasis, more combinations of molecules will need to be investigated in the expansion of HSCs.

6.2: Conclusions and future work

In this study we engineered *in vitro* HUDEP2 and KBM5 cell models of *TET2* and/or *ASXL1* cells representing erythroid and myeloid cells in myeloid malignancies. We further confirmed the role of *TET2* and *ASXL1* in myeloid malignancies by conducting functional studies and finally we investigated the ability of novel combinations of cytokines and small molecules to expand CD133+ cells. We think these data will provide new important insights into how we can design new scientific strategies to help understand myeloid malignancy. However, further research was hindered by the SARS-CoV-2 pandemic in 2020.

Future work could be aimed at completing the HSPC expansion studies in human CD133+ cells, and subsequently introducing the *TET2* and *ASXL1* mutations described, alone or together, into these primary cells for functional assessment. However, we have now moved away from the use of CD133+ cells to CD34+ cells as we had access to cryopreserved cord blood and bone marrow for these studies and we have developed an improved GMP compatible selection strategy for CD34+ cells from cryopreserved cord blood units (Mata, Hernandez et al. 2019), since we could not achieve consistent results using CD133+ selection beads with the cryopreserved cells. Results of these CD34+ cell isolation studies from cryopreserved umbilical cord blood (UCB) units are shown in figure 67 and the serial transplantation of these cells into NSG mice is shown in figure 68 and have been published (Mata,

Hernandez et al. 2019). This approach would next be applied to the purification of CD34+ cells from cryopreserved adult human bone marrow.

CRISPR gene editing during human HSPC culture *ex vivo* would optimally achieve high genome engineering efficiencies and retain high or optimal yields of transplantable engineered human HSCs while minimising off-target effects. Here, we would compare two approaches for *ex vivo* cell culture, while editing *ASXL1* and/or *TET2* genes. The first would test UM171, which has been reported to enhance human cord blood and adult bone marrow lentiviral-mediated transduction and enhance yields of long-term lympho-myeloid repopulating activity in immunodeficient mice (Ngom, Imren et al. 2018).

The second approach would be based on further studies by some of our collaborators (Hua, Hester et al. 2020), who have recently demonstrated that long term UCB HSC expansion, as assessed both phenotypically and in an *in vivo* immunodeficient mouse serial transplant model, can be promoted by the addition of another chromatin modifying factor, the bromodomain and extra-terminal motif (BET) inhibitor CPI203, to our cytokine cocktail of SCF, TPO and FLT3L, which minimises HSC loss, over a short period of time, namely a 5 day culture period. These approaches may provide an improved HSC expansion/culture strategy for our further studies. While UCB is more easily obtained than human bone marrow, it would be clinically relevant to introduce the *TET2* and *ASXL1* mutations described in this thesis into cultured/expanded BM derived CD34+ cells.

The next step would be to assess off-target effects on the human HSPC engineered samples, and to characterise these primary human cell models to identify expression changes in key genes, which may be important for cellular function when *TET2* or *ASXL1* genes are altered. Where possible such techniques as Sanger sequencing, qRT-PCR, flow cytometry and western blot analysis would first be used. Following on, as with the cell line models described in this thesis, proliferation, cell cycle, apoptosis and differentiation would be investigated in these primary human cell models *ex vivo*. It would be important to assess the effects of these mutations on individual and defined CD34+ HSPC subsets identified by flow cytometry and assessed by single cell omics (e.g. RNA-seq) coupled where

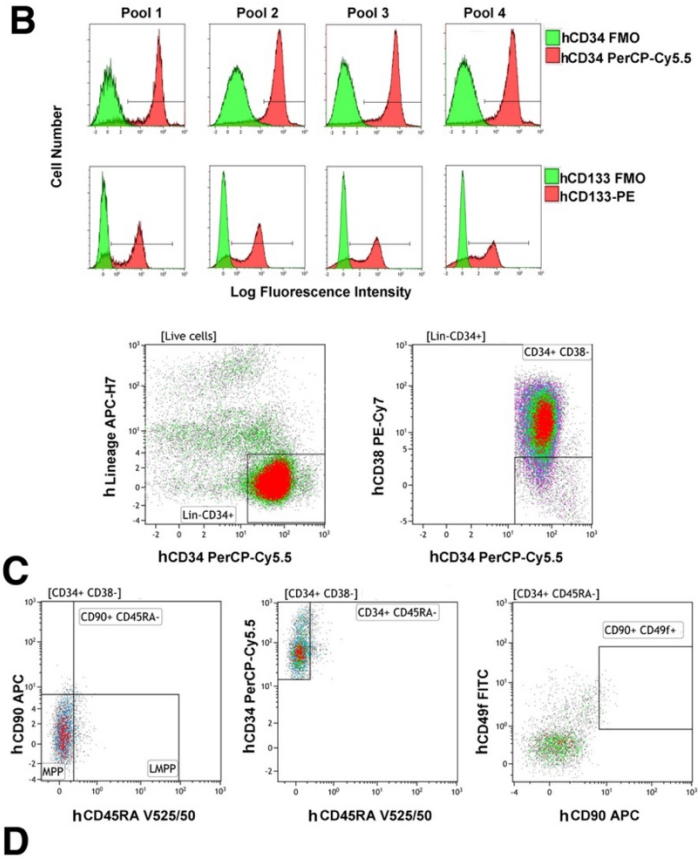
possible with functional *ex vivo* analysis. Quantifying levels of 5-mC and 5-hmC (methylation changes) in *TET2* modified versus unmodified human HSPCs, and investigating HOXA gene family expression in *ASXL1* mutant versus unmodified human HSPCs would be of particular interest since they are important in MDS (Fisher, Pineault et al. 2010).

In addition, we would use mouse xenograft studies to investigate the effects of the *TET2* or ASXL1 mutations on human cells *in vivo*. The NSG (NOD scid gamma mouse) mice would be injected with *TET2* and *ASXL1* mutant primary human cells, alone and in combination, to investigate their differential effects on human neonatal (UCB) versus adult BM HSC survival and self-renewal, differentiation bias or clonal haematopoiesis, and their ability to contribute to serial transplantation, as well as disease progression and survival using Kaplan-Meier analysis. It would also be interesting to look at additional driver mutations of MDS such as: *RUNX1*, *CBL*, *SRSF2*, *RAS*, *IDH2*, and to introduce these into our primary human cell models (Sperling, Gibson et al. 2017). Finally, while such studies are possible in cell lines, we would need to assess if it would be possible to use CRISPR activation (Cheng, Wang et al. 2013) to rescue our mutant primary human cells in order to see if the functional effects observed *in vitro* and *in vivo* in surrogate immunodeficient murine models are reversible, should expression of *TET2* and/or *ASXL1* be restored.

A

UCB ID	Pre-freeze TNC ($x10^6$)	Pre-freeze CD34+ ($x10^6$)	Post-thaw TNC ($x10^6$)	Post-thaw viability (%)	Pre-column TNC ($x10^6$)	Pre-column CD34+ (%)	Post column TNC ($x10^6$)	Yield (%)	Post column CD34+ (%)	Post column CD34+/ CD133+ (%)
S1	12.01	4.03	4.89	54.8	3.03	0.4	1.8	0.6	60	57
S2	10.0	4	4.47	54.9	3.08	0.55	1.92	0.62	81	79
S3	17.22	4.07	4.77	47.5	4	0.5	1.39	0.34	74	71.5
Pool 1 CB 1	12	5.3	4.6	95.6						
Pool 1 1 CB 2	14.7	5.2	5.21	96.4						
Pool 1 CB 3	12.5	4.6	3.49	95.6						
Pool 1 CB 4	10.6	5.1	4.96	94.4						
Pool 1				79.1	14.9	n/a	12.4	0.83	97.7	95.2
Pool 2 CB 1	16.6	10.02	5.93	97.2						
Pool 2 CB 2	12	16.1	7.28	96.1						
Pool 2 CB 3	13.5	6.5	6.13	97.9						
Pool 2 CB 4	12	5.6	7.08	95.6						
Pool 2				85.3	20.9	n/a	24.4	1.17	76.3	71.3
Pool 3 CB1	12.4	4.08	5.37	97.2						
Pool 3 CB2	11	3.45	3.71	96.1						
Pool 3 CB3	18.3	9.5	7.89	97.9						
Pool 3 CB4	19.4	18.05	7.24	95.6						
Pool 3				n/a	31.1	n/a	34.5	1.11	96.8	90.8
Pool 4 CB1	13.1	5.93	5.48	97.2						
Pool 4 CB2	17	6.6	7.05	96.1						
Pool 4 CB3	11.5	3.3	8.74	97.9						
Pool 4 CB4	13.5	3.8	6.01	95.6						
Pool 4				n/a	22.0	n/a	16.1	0.73	96.2	85.5

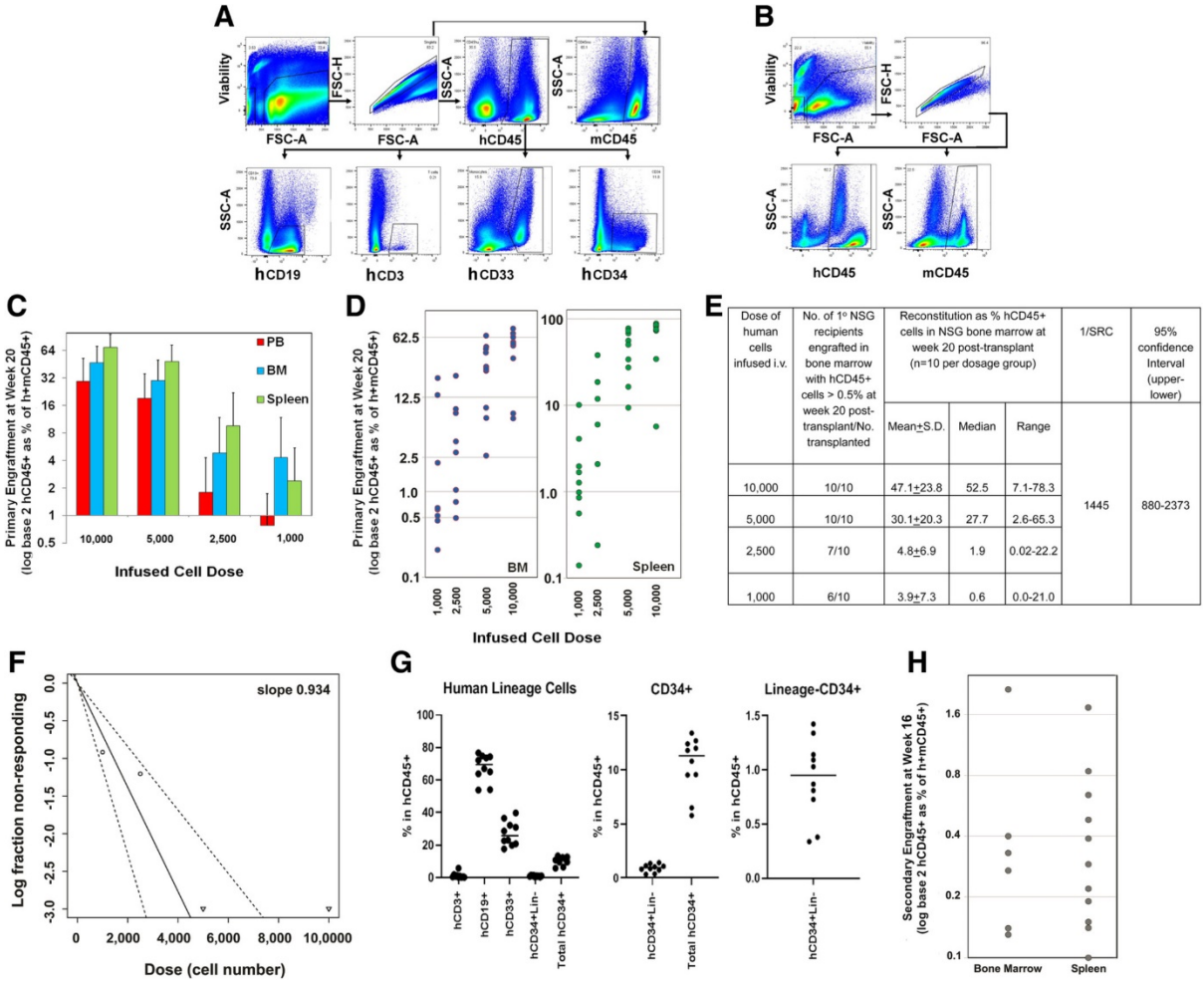
S = single UCB units; Pool = each of 4 UCB units used to create the different cell pools labeled 1-4.



D

Percentage of Enriched hCD34 Cells (Mean±S.D.; n=4)				
%hCD34+ Cells	% hCD133+ Cells	% MPP	%LMPP	%HSC
CD34+	CD133+	Lin-CD34+CD38-CD90-CD45RA-	Lin-CD34+CD38-CD90-CD45RA+	Lin-CD34+CD38-CD90+CD45RA-CD49f+
87.8±2.4	77.6±2.1	6.6±4.7	0.7±0.1	0.2±0.1

Figure 67: Cell surface biomarker analyses of the purified human UCB CD34+ cells.
(A) Individual data on UCB units that were used alone (S) or pooled (Pool 1-4) for CD34+ cell isolation on the Whole Blood Column. (B-D): Pooled UCB units (four units per pool) were thawed in the Plasmatherm, labelled with CD34 magnetic beads and isolated using Whole Blood MACS Columns (n = 4 independent pools). (B) Flow cytometric histograms showing the purity of the viable cells eluted for each cell pool. In the upper and lower histogram panels, the nucleated cells were labelled with hCD34 and hCD133 antibodies respectively (red) and analysed against FMO controls (green). (C) Flow cytometric dot plots showing subsets of viable hCD34+ cells labelled with hCD38, hCD90, hCD45RA and hCD49f to define phenotypes of HSC and early hematopoietic progenitors. (D) Summary of the quantitative analyses of the four purified pools of UCB showing the mean ± standard deviation for hCD34 purity, the percentage of hCD34+ cells expressing CD133, the percentage of MPPs defined as Lin-CD34 + CD38-CD90-CD45RA- cells, the percentage of LMPPs defined as Lin-CD34 + CD38-CD90-CD45RA+ cells, and the percentage of HSCs defined as Lin-CD34 + CD38-CD90 + CD45RA-CD49f + cells (Mata, Hernandez et al. 2019).



Dose of human cells infused i.v.	No. of 1° NSG recipients engrafted in bone marrow with hCD45+ cells > 0.5% at week 20 post-transplant/No. transplanted	Reconstitution as % hCD45+ cells in NSG bone marrow at week 20 post-transplant (n=10 per dosage group)			1/SRC	95% confidence Interval (upper-lower)
		Mean±S.D.	Median	Range		
10,000	10/10	47.1±23.8	52.5	7.1-78.3	1445	880-2373
5,000	10/10	30.1±20.3	27.7	2.6-65.3		
2,500	7/10	4.8±6.9	1.9	0.02-22.2		
1,000	6/10	3.9±7.3	0.6	0.0-21.0		

Figure 68: Primary and secondary engraftment in the bone marrow (BM) and spleen of NSG mice.
After exclusion of mCD45+ cells, (A) flow cytometric plots and gating of hCD45+ leukocytes in the BM of primary NSG recipients at Week 20 after transplant and (B) flow cytometric plots and gating of hCD45+ leukocytes in the spleen of primary NSG recipients at Week 20 after transplant. (C) hCD45+ hematologic reconstitution of the BM (blue) and spleen (green) compared to the peripheral blood (red) at Week 20 after primary transplants of NSG mice receiving 10,000, 5000, 2500, and 1000 enriched hCD34+ cells. Values are means ± standard deviation for n = 10 mice per cell dose injected, except for mice receiving 1000 enriched hCD34+ cells, where 9 of the 10 mice survived to Week 20 (D) Engraftment of greater than 0.1% of hCD45+ cells into the BM and spleen of each primary NSG recipient at Week 20 after transplant according to primary enriched hCD34+ cell dose. (E) Quantitative analyses of hCD45+ cell engraftment in the bone marrow primary NSG recipients at Week 20 after transplant at different cell doses infused. The mean ± standard deviation, median, and range of engraftment of all NSG mice (n = 40) are shown. The number of mice per group engrafting hCD45+ cells is shown with a more stringent BM engraftment cut off of greater than 0.5% hCD45+ as used by Csaszar E, Kirouac DC, Yu M, et al. (Csaszar, Kirouac et al. 2012) and the average Scid repopulating cell (SRC) content with 95% confidence intervals of the primary graft calculated as shown in E and F. (G) The lineage distribution of the hCD45+ cells in Week 20 bone marrow of NSG mice receiving 10,000 cells per mouse intravenously shows the percentage of hCD45+ cells expressing CD3+ T and CD19+ B lymphoid biomarkers, hCD33+ myeloid biomarkers and with hCD34+ progenitor characteristics. (H) Half the primary BM from each NSG mouse receiving initial hCD34+ enriched cell dosages of 10,000 and 5000 cells (n = 10 each dose) was injected into individual secondary NSG recipients and the engraftment of hCD45+ leukocytes in the BM and spleen analysed 16 weeks later (Mata, Hernandez et al. 2019).

Bibliography

Abdel-Wahab, O., et al. (2012). "ASXL1 mutations promote myeloid transformation through loss of PRC2-mediated gene repression." Cancer Cell **22**(2): 180-193.

Abdel-Wahab, O., et al. (2013). "Deletion of Asxl1 results in myelodysplasia and severe developmental defects in vivoConditional deletion of Asxl1 results in MDS." The Journal of experimental medicine **210**(12): 2641-2659.

Abegunde, S. O., et al. (2018). "An inflammatory environment containing TNFα favors Tet2-mutant clonal hematopoiesis." Experimental hematology **59**: 60-65.

Adolfsson, J., et al. (2005). "Identification of Flt3+ lympho-myeloid stem cells lacking erythro-megakaryocytic potential a revised road map for adult blood lineage commitment." Cell **121**(2): 295-306.

Agrawal, N., et al. (2003). "RNA interference: biology, mechanism, and applications." Microbiol Mol Biol Rev **67**(4): 657-685.

Aguilar-Navarro, A. G., et al. (2020). "Human aging alters the spatial organization between CD34+ hematopoietic cells and adipocytes in bone marrow." Stem Cell Reports **15**(2): 317-325.

Ahn, J. S., et al. (2015). "Adverse prognostic effect of homozygous TET2 mutation on the relapse risk of acute myeloid leukemia in patients of normal karyotype." Haematologica **100**(9): e351-353.

Akashi, K., et al. (2000). "A clonogenic common myeloid progenitor that gives rise to all myeloid lineages." Nature **404**(6774): 193-197.

Anandasabapathy, N., et al. (2015). "Efficacy and safety of CDX-301, recombinant human Flt3L, at expanding dendritic cells and hematopoietic stem cells in healthy human volunteers." Bone marrow transplantation **50**(7): 924-930.

Asada, S., et al. (2018). "Mutant ASXL1 cooperates with BAP1 to promote myeloid leukaemogenesis." Nat Commun **9**(1): 2733.

Awada, H., et al. (2020). "The Genomics of Myelodysplastic Syndromes: Origins of Disease Evolution, Biological Pathways, and Prognostic Implications." Cells **9**(11).

Barbui, T., et al. (2018). "The 2016 WHO classification and diagnostic criteria for myeloproliferative neoplasms: document summary and in-depth discussion." Blood Cancer J **8**(2): 15.

Baron, M. H., et al. (2012). "The embryonic origins of erythropoiesis in mammals." Blood, The Journal of the American Society of Hematology **119**(21): 4828-4837.

Batai, B., et al. (2020). "Juvenile myelomonocytic leukaemia presentation after preceding juvenile xanthogranuloma harbouring an identical somatic PTPN11 mutation." Pediatr Blood Cancer **67**(9): e28368.

Batsivari, A., et al. (2020). "Dynamic responses of the haematopoietic stem cell niche to diverse stresses." Nature cell biology **22**(1): 7-17.

Becker, A. J., et al. (1963). "Cytological demonstration of the clonal nature of spleen colonies derived from transplanted mouse marrow cells." Nature **197**: 452-454.

Behrmann, L., et al. (2018). "Acute Myeloid Leukemia and the Bone Marrow Niche-Take a Closer Look." Front Oncol **8**: 444.

Bejar, R., et al. (2014). "TET2 mutations predict response to hypomethylating agents in myelodysplastic syndrome patients." Blood **124**(17): 2705-2712.

Bennett, J. M., et al. (1982). "Proposals for the classification of the myelodysplastic syndromes." Br J Haematol **51**(2): 189-199.

Bera, R., et al. (2019). "RUNX1 mutations promote leukemogenesis of myeloid malignancies in ASXL1-mutated leukemia." J Hematol Oncol **12**(1): 104.

Beran, M., et al. (1993). "Biological properties and growth in SCID mice of a new myelogenous leukemia cell line (KBM-5) derived from chronic myelogenous leukemia cells in the blastic phase." Cancer research **53**(15): 3603-3610.

Bernard, E., et al. (2020). "Implications of TP53 allelic state for genome stability, clinical presentation and outcomes in myelodysplastic syndromes." Nature medicine **26**(10): 1549-1556.

Bhattacharya, R., et al. (2021). "Clonal Hematopoiesis Is Associated With Higher Risk of Stroke." Stroke: STROKEAHA. 121.037388.

Binder, M., et al. (2019). "Epigenomic Determinants of Transcriptional Activity in ASXL1-Mutant Chronic Myelomonocytic Leukemia." Blood **134**(Supplement_1): 2987-2987.

Boitano, A. E., et al. (2010). "Aryl hydrocarbon receptor antagonists promote the expansion of human hematopoietic stem cells." Science **329**(5997): 1345-1348.

Bowman, R. L., et al. (2018). "Clonal hematopoiesis and evolution to hematopoietic malignancies." Cell stem cell **22**(2): 157-170.

Bowman, R. L. and R. L. Levine (2017). "TET2 in Normal and Malignant Hematopoiesis." Cold Spring Harb Perspect Med **7**(8).

Bug, G., et al. (2004). Valproic Acid Accelerates Cell Cycle Progression of Hematopoietic Stem Cells Via the Activation of GSK3beta-Dependent Signaling Pathways, American Society of Hematology.

Bug, G., et al. (2005). "Valproic acid stimulates proliferation and self-renewal of hematopoietic stem cells." Cancer research **65**(7): 2537-2541.

Bulaeva, E., et al. (2020). "MYC-induced human acute myeloid leukemia requires a continuing IL-3/GM-CSF costimulus." Blood **136**(24): 2764-2773.

Burma, S., et al. (2006). "Role of non-homologous end joining (NHEJ) in maintaining genomic integrity." DNA repair **5**(9-10): 1042-1048.

Cai, Z., et al. (2018). "Inhibition of inflammatory signaling in Tet2 mutant preleukemic cells mitigates stress-induced abnormalities and clonal hematopoiesis." Cell stem cell **23**(6): 833-849. e835.

Cakouros, D., et al. (2019). "Specific functions of TET1 and TET2 in regulating mesenchymal cell lineage determination." Epigenetics Chromatin **12**(1): 3.

Carbuccia, N., et al. (2009). "Mutations of ASXL1 gene in myeloproliferative neoplasms." Leukemia **23**(11): 2183-2186.

Challen, G. A. and M. A. Goodell (2020). "Clonal hematopoiesis: mechanisms driving dominance of stem cell clones." Blood **136**(14): 1590-1598.

Chan, D. K. and W. K. Miskimins (2012). "Metformin and phenethyl isothiocyanate combined treatment in vitro is cytotoxic to ovarian cancer cultures." Journal of ovarian research **5**(1): 1-9.

Chang, C. J., et al. (2018). "Dissecting the Contributions of Cooperating Gene Mutations to Cancer Phenotypes and Drug Responses with Patient-Derived iPSCs." Stem Cell Reports **10**(5): 1610-1624.

Chaurasia, P., et al. (2014). "Epigenetic reprogramming induces the expansion of cord blood stem cells." The Journal of clinical investigation **124**(6): 2378-2395.

Chen, E., et al. (2015). "Distinct effects of concomitant Jak2V617F expression and Tet2 loss in mice promote disease progression in myeloproliferative neoplasms." Blood, The Journal of the American Society of Hematology **125**(2): 327-335.

Cheng, A. W., et al. (2013). "Multiplexed activation of endogenous genes by CRISPR-on, an RNA-guided transcriptional activator system." Cell research **23**(10): 1163-1171.

Cheng, H., et al. (2020). "New paradigms on hematopoietic stem cell differentiation." Protein Cell **11**(1): 34-44.

Cheng, J., et al. (2013). "An extensive network of TET2-targeting MicroRNAs regulates malignant hematopoiesis." Cell Rep **5**(2): 471-481.

Chiba, S. (2017). "Dysregulation of TET2 in hematologic malignancies." International journal of hematology **105**(1): 17-22.

Chitteti, B. R., et al. (2010). "Impact of interactions of cellular components of the bone marrow microenvironment on hematopoietic stem and progenitor cell function." Blood, The Journal of the American Society of Hematology **115**(16): 3239-3248.

Cho, S. W., et al. (2013). "Targeted genome engineering in human cells with the Cas9 RNA-guided endonuclease." Nature biotechnology **31**(3): 230-232.

Cimmino, L., et al. (2017). "Restoration of TET2 Function Blocks Aberrant Self-Renewal and Leukemia Progression." Cell **170**(6): 1079-1095 e1020.

Civin, C. I., et al. (1984). "Antigenic analysis of hematopoiesis. III. A hematopoietic progenitor cell surface antigen defined by a monoclonal antibody raised against KG-1a cells." J Immunol **133**(1): 157-165.

Cochran, J. N., et al. (2020). "Non-coding and Loss-of-Function Coding Variants in TET2 are Associated with Multiple Neurodegenerative Diseases." Am J Hum Genet **106**(5): 632-645.

Cohen, S., et al. (2020). "Hematopoietic stem cell transplantation using single UM171-expanded cord blood: a single-arm, phase 1–2 safety and feasibility study." The Lancet Haematology **7**(2): e134-e145.

Concordet, J.-P. and M. Haeussler (2018). "CRISPOR: intuitive guide selection for CRISPR/Cas9 genome editing experiments and screens." Nucleic acids research **46**(W1): W242-W245.

Cordes, S., et al. (2020). "Clonal tracking of haematopoietic cells: insights and clinical implications." Br J Haematol.

Crosse, E. I., et al. (2020). "Multi-layered Spatial Transcriptomics Identify Secretory Factors Promoting Human Hematopoietic Stem Cell Development." Cell stem cell **27**(5): 822-839. e828.

Csaszar, E., et al. (2012). "Rapid expansion of human hematopoietic stem cells by automated control of inhibitory feedback signaling." Cell stem cell **10**(2): 218-229.

Csaszar, E., et al. (2014). "Blood stem cell fate regulation by Delta-1–mediated rewiring of IL-6 paracrine signaling." Blood, The Journal of the American Society of Hematology **123**(5): 650-658.

Cumano, A., et al. (1996). "Lymphoid potential, probed before circulation in mouse, is restricted to caudal intraembryonic splanchnopleura." Cell **86**(6): 907-916.

Dang, Y., et al. (2015). "Optimizing sgRNA structure to improve CRISPR-Cas9 knockout efficiency." Genome biology **16**(1): 1-10.

Dara, M., et al. (2019). "Synthesis a New Viral Base Vector Carrying Single Guide RNA (sgRNA) and Green Florescent Protein (GFP)." Journal of Sciences, Islamic Republic of Iran **30**(3): 205-209.

Delaney, C., et al. (2010). "Notch-mediated expansion of human cord blood progenitor cells capable of rapid myeloid reconstitution." Nat Med **16**(2): 232-236.

Della Porta, M. G., et al. (2006). "Flow cytometry evaluation of erythroid dysplasia in patients with myelodysplastic syndrome." Leukemia **20**(4): 549-555.

Deng, W., et al. (2016). "TET2 regulates LncRNA-ANRIL expression and inhibits the growth of human gastric cancer cells." IUBMB Life **68**(5): 355-364.

Derry, W. B. (2020). "CRISPR: development of a technology and its applications." FEBS J.

Dick, J. (1991). "Immune-deficient mice as models of normal and leukemic human hematopoiesis." Cancer cells (Cold Spring Harbor, NY: 1989) **3**(2): 39-48.

Dorsheimer, L., et al. (2019). "Association of mutations contributing to clonal hematopoiesis with prognosis in chronic ischemic heart failure." JAMA cardiology **4**(1): 25-33.

Doulatov, S., et al. (2010). "Revised map of the human progenitor hierarchy shows the origin of macrophages and dendritic cells in early lymphoid development." Nat Immunol **11**(7): 585-593.

Dürig, J., et al. (2002). "CD38 expression is an important prognostic marker in chronic lymphocytic leukaemia." Leukemia **16**(1): 30-35.

Dykstra, B., et al. (2007). "Long-Term Propagation of Distinct Hematopoietic Differentiation Programs In Vivo." Cell stem cell **1**(2): 218-229.

Fang, X., et al. (2021). "Asxl1 C-terminal mutation perturbs neutrophil differentiation in zebrafish." Leukemia **35**(8): 2299-2310.

Fares, I., et al. (2014). "Pyrimidoindole derivatives are agonists of human hematopoietic stem cell self-renewal." Science **345**(6203): 1509-1512.

Fares, I., et al. (2014). "Cord blood expansion. Pyrimidoindole derivatives are agonists of human hematopoietic stem cell self-renewal." Science **345**(6203): 1509-1512.

Fares, I., et al. (2017). "EPCR expression marks UM171-expanded CD34+ cord blood stem cells." Blood **129**(25): 3344-3351.

Ferrari, S., et al. (2020). "Efficient gene editing of human long-term hematopoietic stem cells validated by clonal tracking." Nat Biotechnol **38**(11): 1298-1308.

Figueroa, M. E., et al. (2010). "Leukemic IDH1 and IDH2 mutations result in a hypermethylation phenotype, disrupt TET2 function, and impair hematopoietic differentiation." Cancer Cell **18**(6): 553-567.

Fisher, C. L., et al. (2010). "Loss-of-function Additional sex combs like 1 mutations disrupt hematopoiesis but do not cause severe myelodysplasia or leukemia." Blood, The Journal of the American Society of Hematology **115**(1): 38-46.

Fox, N., et al. (2002). "Thrombopoietin expands hematopoietic stem cells after transplantation." The Journal of clinical investigation **110**(3): 389-394.

Frisch, B. J. (2021). "Hematopoietic Stem Cell Cultures and Assays." Methods Mol Biol **2230**: 467-477.

Fujino, T., et al. (2020). "3011–MUTANT ASXL1 INDUCES EXPANSION OF HEMATOPOIETIC STEM CELLS THROUGH ACTIVATION OF AKT/MTOR PATHWAY." Experimental hematology **88**: S41.

Gajzer, D., et al. (2016). "Epigenetic and molecular signatures of cord blood CD 34+ cells treated with histone deacetylase inhibitors." Vox Sanguinis **110**(1): 79-89.

Galan-Diez, M. and S. Kousteni (2017). "The osteoblastic niche in hematopoiesis and hematological myeloid malignancies." Curr Mol Biol Rep **3**(2): 53-62.

Gelsi-Boyer, V., et al. (2009). "Mutations of polycomb-associated gene ASXL1 in myelodysplastic syndromes and chronic myelomonocytic leukaemia." British journal of haematology **145**(6): 788-800.

Genovese, G., et al. (2014). "Clonal hematopoiesis and blood-cancer risk inferred from blood DNA sequence." New England Journal of Medicine **371**(26): 2477-2487.

Gerlai, R. (2016). "Gene Targeting Using Homologous Recombination in Embryonic Stem Cells: The Future for Behavior Genetics?" Frontiers in Genetics **7**(43).

Giannini, E. G. and M. Peck-Radosavljevic (2015). Platelet dysfunction: status of thrombopoietin in thrombocytopenia associated with chronic liver failure. Seminars in thrombosis and hemostasis, Thieme Medical Publishers.

Gjini, E., et al. (2019). "Disruption of asxl1 results in myeloproliferative neoplasms in zebrafish." Disease models & mechanisms **12**(5).

Goardon, N., et al. (2011). "Coexistence of LMPP-like and GMP-like leukemia stem cells in acute myeloid leukemia." Cancer Cell **19**(1): 138-152.

Gondek, L. P. (2021). "CHIP: is clonal hematopoiesis a surrogate for aging and other disease?" Hematology **2021**(1): 384-389.

Gordon, M. Y., et al. (1987). "Compartmentalization of a haematopoietic growth factor (GM-CSF) by glycosaminoglycans in the bone marrow microenvironment." Nature **326**(6111): 403-405.

Greenberg, P. L., et al. (2012). "Revised International Prognostic Scoring System for Myelodysplastic Syndromes." Blood **120**(12): 2454-2465.

Growney, J. D., et al. (2005). "Loss of Runx1 perturbs adult hematopoiesis and is associated with a myeloproliferative phenotype." Blood **106**(2): 494-504.

Gullo, F., et al. (2015). "Computational modeling of the expansion of human cord blood CD133+ hematopoietic stem/progenitor cells with different cytokine combinations." Bioinformatics **31**(15): 2514-2522.

Haeckel, E. (1868). Natürliche Schöpfungsgeschichte. Gemeinverständliche wissenschaftliche Vorträge über die Entwickelungslehre im Allgemeinen und diejenige von Darwin, Goethe und Lamarck im Besonderen... Mit Tafeln, Holzschnitten, etc.

Haferlach, T., et al. (2014). "Landscape of genetic lesions in 944 patients with myelodysplastic syndromes." Leukemia **28**(2): 241-247.

Hale, N. E. and J. F. Keane, Jr. (2010). "Piecing together a picture of trisomy 8 mosaicism syndrome." J Am Osteopath Assoc **110**(1): 21-23.

Hall, M., et al. (2020). "Refractory Autoimmune Hemolytic Anemia and Pure Red Cell Aplasia Secondary to Immunotherapy Requiring Prolonged Immunosuppression." JCO Oncol Pract **16**(10): 699-700.

Hanahan, D. and R. A. Weinberg (2011). "Hallmarks of cancer: the next generation." Cell **144**(5): 646-674.

Hassan, H. and A. Zander (1996). Stem cell factor as a survival and growth factor in human normal and malignant hematopoiesis. Molecular Biology of Hematopoiesis 5, Springer**:** 549-558.

He, Q., et al. (2017). "Enhanced Hematopoietic Stem Cell Self-Renewal-Promoting Ability of Clonal Primary Mesenchymal Stromal/Stem cells Versus Their Osteogenic Progeny." Stem Cells **35**(2): 473-484.

Hellstrom-Lindberg, E., et al. (2020). "Myelodysplastic syndromes: moving towards personalized management." Haematologica **105**(7): 1765-1779.

Hilgendorf, S., et al. (2016). "Loss of ASXL1 triggers an apoptotic response in human hematopoietic stem and progenitor cells." Experimental hematology **44**(12): 1188-1196. e1186.

Hoermann, G., et al. (2020). "Clonal Hematopoiesis of Indeterminate Potential: A Multidisciplinary Challenge in Personalized Hematology." Journal of Personalized Medicine **10**(3): 94.

Horwitz, M. E., et al. (2014). "Umbilical cord blood expansion with nicotinamide provides long-term multilineage engraftment." The Journal of clinical investigation **124**(7): 3121-3128.

Horwitz, M. E., et al. (2019). "Phase I/II study of stem-cell transplantation using a single cord blood unit expanded ex vivo with nicotinamide." Journal of Clinical Oncology **37**(5): 367.

Hsiau, T., et al. (2018). "Inference of CRISPR edits from sanger trace data." BioRxiv: 251082.

Hu, J., et al. (2014). "Sirtuin inhibitors as anticancer agents." Future medicinal chemistry **6**(8): 945-966.

Hu, Y., et al. (2018). "Single cell multi-omics technology: methodology and application." Frontiers in cell and developmental biology **6**: 28.

Hua, P., et al. (2020). "The BET inhibitor CPI203 promotes ex vivo expansion of cord blood long-term repopulating HSCs and megakaryocytes." Blood **136**(21): 2410-2415.

Hua, P., et al. (2019). "Single-cell assessment of transcriptome alterations induced by Scriptaid in early differentiated human haematopoietic progenitors during ex vivo expansion." Scientific reports **9**(1): 1-9.

Huang, F., et al. (2020). "HDAC4 inhibition disrupts TET2 function in high-risk MDS and AML." Aging (Albany NY) **12**(17): 16759-16774.

Huang, J., et al. (2019). "Distinct roles of Dlk1 isoforms in bi-potential differentiation of hepatic stem cells." Stem Cell Res Ther **10**(1): 31.

Huntsman, H. D., et al. (2015). "Human hematopoietic stem cells from mobilized peripheral blood can be purified based on CD49f integrin expression." Blood, The Journal of the American Society of Hematology **126**(13): 1631-1633.

Iacomussi, S. (2019). "Regulating genome editing technologies: a comparison of expert recommendations in the U.K. and in the U.S.A." Emerg Top Life Sci **3**(6): 701-705.

Inoue, D., et al. (2013). "Myelodysplastic syndromes are induced by histone methylation-altering ASXL1 mutations." J Clin Invest **123**(11): 4627-4640.

Ito, K., et al. (2019). "Non-catalytic roles of Tet2 are essential to regulate hematopoietic stem and progenitor cell homeostasis." Cell reports **28**(10): 2480-2490. e2484.

Itzykson, R., et al. (2011). "Impact of TET2 mutations on response rate to azacitidine in myelodysplastic syndromes and low blast count acute myeloid leukemias." Leukemia **25**(7): 1147-1152.

Ivanovs, A., et al. (2017). "Human haematopoietic stem cell development: from the embryo to the dish." Development **144**(13): 2323-2337.

Jacobson, L. O., et al. (1951). "Recovery from radiation injury." Science **113**(2940): 510-511.

Jagannathan-Bogdan, M. and L. I. Zon (2013). "Hematopoiesis." Development **140**(12): 2463-2467.

Jaiswal, S., et al. (2014). "Age-related clonal hematopoiesis associated with adverse outcomes." New England Journal of Medicine **371**(26): 2488-2498.

Jaiswal, S. and P. Libby (2020). "Clonal haematopoiesis: connecting ageing and inflammation in cardiovascular disease." Nature Reviews Cardiology **17**(3): 137-144.

Jaiswal, S., et al. (2017). "Clonal Hematopoiesis and Risk of Atherosclerotic Cardiovascular Disease." N Engl J Med **377**(2): 111-121.

Jamieson, C. H., et al. (2004). "Granulocyte–macrophage progenitors as candidate leukemic stem cells in blast-crisis CML." New England Journal of Medicine **351**(7): 657-667.

Jankowska, A. M., et al. (2009). "Loss of heterozygosity 4q24 and TET2 mutations associated with myelodysplastic/myeloproliferative neoplasms." Blood **113**(25): 6403-6410.

Jasek, M., et al. (2010). "TP53 mutations in myeloid malignancies are either homozygous or hemizygous due to copy number-neutral loss of heterozygosity or deletion of 17p." Leukemia **24**(1): 216-219.

Jiang, S. (2020). "Tet2 at the interface between cancer and immunity." Commun Biol **3**(1): 667.

Jinek, M., et al. (2012). "A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity." Science **337**(6096): 816-821.

Kakosaiou, K., et al. (2018). "ASXL1 mutations in AML are associated with specific clinical and cytogenetic characteristics." Leukemia & lymphoma **59**(10): 2439-2446.

Kale, V. P. and L. S. Limaye (2010). "Expansion of cord blood CD34+ cells in presence of zVADfmk and zLLYfmk improved their in vitro functionality and in vivo engraftment in NOD/SCID mouse." PLoS One **5**(8): e12221.

Kersten, B., et al. (2016). "CD 45 RA, a specific marker for leukaemia stem cell sub-populations in acute myeloid leukaemia." British journal of haematology **173**(2): 219-235.

Khosrowshahi, Y. B., et al. (2016). "Ex vivo expansion of hematopoietic stem cells in a proliferation chamber with external stirred conditioning tank: Sequential optimization of growth factors." Engineering in Life Sciences **16**(3): 254-262.

Kim, S., et al. (2020). "Genome editing of immune cells using CRISPR/Cas9." BMB Rep.

Kirito, K. (2020). "Myeloid neoplasm with isolated del(5q) and the MPLW515L mutation fulfills the WHO diagnostic criteria for ET." Int J Hematol **112**(2): 238-242.

Ko, M., et al. (2011). "Ten-Eleven-Translocation 2 (TET2) negatively regulates homeostasis and differentiation of hematopoietic stem cells in mice." Proceedings of the National Academy of Sciences **108**(35): 14566-14571.

Ko, M., et al. (2010). "Impaired hydroxylation of 5-methylcytosine in myeloid cancers with mutant TET2." Nature **468**(7325): 839-843.

Kon, A., et al. (2018). "Physiological Srsf2 P95H expression causes impaired hematopoietic stem cell functions and aberrant RNA splicing in mice." Blood **131**(6): 621-635.

Kondo, M. (2010). "Lymphoid and myeloid lineage commitment in multipotent hematopoietic progenitors." Immunological reviews **238**(1): 37-46.

Kondo, M., et al. (1997). "Identification of clonogenic common lymphoid progenitors in mouse bone marrow." Cell **91**(5): 661-672.

Langemeijer, S. M., et al. (2009). "Acquired mutations in TET2 are common in myelodysplastic syndromes." Nature genetics **41**(7): 838-842.

Larson, M. H., et al. (2013). "CRISPR interference (CRISPRi) for sequence-specific control of gene expression." Nature protocols **8**(11): 2180-2196.

Lebois, M., et al. (2016). "Regulation of platelet lifespan in the presence and absence of thrombopoietin signaling." Journal of Thrombosis and Haemostasis **14**(9): 1882-1887.

Lechman, E. R., et al. (2012). "Attenuation of miR-126 activity expands HSC in vivo without exhaustion." Cell stem cell **11**(6): 799-811.

Ledford, H. (2020). "CRISPR gene therapy shows promise against blood diseases." Nature.

Lennartsson, J. and L. Rönnstrand (2012). "Stem cell factor receptor/c-Kit: from basic science to clinical implications." Physiological reviews **92**(4): 1619-1649.

Li, J., et al. (2019). "Association of Leukemia Target Genes Tet2, Bcl2, and Slc23a2 in Vitamin C Pathways." Cancer Genomics Proteomics **16**(5): 333-344.

Li, R., et al. (2018). "TET2 Loss Dysregulates the Behavior of Bone Marrow Mesenchymal Stromal Cells and Accelerates Tet2(-/-)-Driven Myeloid Malignancy Progression." Stem Cell Reports **10**(1): 166-179.

Li, Y., et al. (2020). "Single-Cell Analysis of Neonatal HSC Ontogeny Reveals Gradual and Uncoordinated Transcriptional Reprogramming that Begins before Birth." Cell stem cell **27**(5): 732-747. e737.

Li, Z., et al. (2011). "Deletion of Tet2 in mice leads to dysregulated hematopoietic stem cells and subsequent development of myeloid malignancies." Blood, The Journal of the American Society of Hematology **118**(17): 4509-4518.

Lin, M. E., et al. (2018). "Dynamics of DNMT3A mutation and prognostic relevance in patients with primary myelodysplastic syndrome." Clin Epigenetics **10**: 42.

Lin, Y., et al. (2020). "Clinical characteristics and prognostic study of adult acute myeloid leukemia patients with ASXL1 mutations." Hematology **25**(1): 446-456.

Lin, Y., et al. (2016). "Prognostic significance of ASXL1 mutations in myelodysplastic syndromes and chronic myelomonocytic leukemia: A meta-analysis." Hematology **21**(8): 454-461.

Liu, T., et al. (2011). "Calpastatin is regulated by protein never in mitosis gene A interacting-1 (PIN1) in endothelial cells." Biochemical and biophysical research communications **414**(3): 581-586.

Liu, X., et al. (2020). "TET2 is involved in DNA hydroxymethylation, cell proliferation and inflammatory response in keratinocytes." Molecular medicine reports **21**(4): 1941-1949.

Lord, B. I. and T. M. Dexter (1988). "Purification of haemopoietic stem cells--the end of the road?" Immunol Today **9**(12): 376-377.

Lucas, D. (2017). The bone marrow microenvironment for hematopoietic stem cells. Stem Cell Microenvironments and Beyond, Springer**:** 5-18.

Ma, T. T., et al. (2020). "Development and validation of a prognostic model for adult patients with acute myeloid leukaemia." EBioMedicine **62**: 103126.

Mahfoudhi, E., et al. (2016). "TET2-mediated 5-hydroxymethylcytosine induces genetic instability and mutagenesis." DNA repair **43**: 78-88.

Mahmud, N., et al. (2014). "Differential effects of epigenetic modifiers on the expansion and maintenance of human cord blood stem/progenitor cells." Biology of Blood and Marrow Transplantation **20**(4): 480-489.

Malcovati, L., et al. (2005). "Flow cytometry evaluation of erythroid and myeloid dysplasia in patients with myelodysplastic syndrome." Leukemia **19**(5): 776-783.

Mangaonkar, A. A., et al. (2018). "Prognostic impact of ASXL1 mutations in patients with myelodysplastic syndromes and multilineage dysplasia with or without ring sideroblasts." Leuk Res **71**: 60-62.

Mannelli, F., et al. (2020). "Multilineage Dysplasia as Assessed by Immunophenotype in Acute Myeloid Leukemia: A Prognostic Tool in a Genetically Undefined Category." Cancers (Basel) **12**(11).

Manz, M. G., et al. (2002). "Prospective isolation of human clonogenic common myeloid progenitors." Proc Natl Acad Sci U S A **99**(18): 11872-11877.

Marsee, D. K., et al. (2010). "CD71 (transferrin receptor): an effective marker for erythroid precursors in bone marrow biopsy specimens." Am J Clin Pathol **134**(3): 429-435.

Mas-Peiro, S., et al. (2020). "Clonal haematopoiesis in patients with degenerative aortic valve stenosis undergoing transcatheter aortic valve implantation." European heart journal **41**(8): 933-939.

Mata, M. F., et al. (2019). "A modified CD34+ hematopoietic stem and progenitor cell isolation strategy from cryopreserved human umbilical cord blood." Transfusion **59**(12): 3560-3569.

Matsuoka, S., et al. (2001). "CD34 expression on long-term repopulating hematopoietic stem cells changes during developmental stages." Blood, The Journal of the American Society of Hematology **97**(2): 419-425.

McKenna, H. J., et al. (1995). "Effect of flt3 ligand on the ex vivo expansion of human CD34+ hematopoietic progenitor cells."

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

Meyer, A., et al. (2020). "Systemic lupus erythematosus and neutropaenia: a hallmark of haematological manifestations." Lupus Sci Med **7**(1).

Micol, J. B. and O. Abdel-Wahab (2016). "The Role of Additional Sex Combs-Like Proteins in Cancer." Cold Spring Harb Perspect Med **6**(10).

Mikkola, H. K. and S. H. Orkin (2006). "The journey of developing hematopoietic stem cells." Development **133**(19): 3733-3744.

Miraglia, S., et al. (1997). "A novel five-transmembrane hematopoietic stem cell antigen: isolation, characterization, and molecular cloning." Blood **90**(12): 5013-5021.

Mohammad, A. A. (2018). "Myelodysplastic syndrome from theoretical review to clinical application view." Oncol Rev **12**(2): 397.

Montagner, S., et al. (2016). "TET2 regulates mast cell differentiation and proliferation through catalytic and non-catalytic activities." Cell reports **15**(7): 1566-1579.

Moran-Crusio, K., et al. (2011). "Tet2 loss leads to increased hematopoietic stem cell self-renewal and myeloid transformation." Cancer Cell **20**(1): 11-24.

Morrison, S. J., et al. (1997). "Identification of a lineage of multipotent hematopoietic progenitors." Development **124**(10): 1929-1939.

Muller-Sieburg, C. E., et al. (2004). "Myeloid-biased hematopoietic stem cells have extensive self-renewal capacity but generate diminished lymphoid progeny with impaired IL-7 responsiveness." Blood **103**(11): 4111-4118.

Mupo, A., et al. (2017). "Hemopoietic-specific Sf3b1-K700E knock-in mice display the splicing defect seen in human MDS but develop anemia without ring sideroblasts." Leukemia **31**(3): 720-727.

Muto, T., et al. (2013). "Concurrent loss of Ezh2 and Tet2 cooperates in the pathogenesis of myelodysplastic disorders." J Exp Med **210**(12): 2627-2639.

Nagase, R., et al. (2018). "Expression of mutant Asxl1 perturbs hematopoiesis and promotes susceptibility to leukemic transformation." J Exp Med **215**(6): 1729-1747.

Nestorowa, S., et al. (2016). "A single-cell resolution map of mouse hematopoietic stem and progenitor cell differentiation." Blood **128**(8): e20-31.

Ng, A. P. and W. S. Alexander (2017). "Haematopoietic stem cells: past, present and future." Cell Death Discov **3**: 17002.

Ng, D. P. and B. Wood (2014). "Unsupervised Discovery of Early Markers of Erythroid Maturation in Human Donor Marrow." Blood **124**(21): 4304-4304.

Ngom, M., et al. (2018). "UM171 enhances lentiviral gene transfer and recovery of primitive human hematopoietic cells." Molecular Therapy-Methods & Clinical Development **10**: 156-164.

Nishida, K. and A. Kondo (2020). "CRISPR-derived genome editing technologies for metabolic engineering." Metab Eng.

Notta, F., et al. (2011). "Isolation of single human hematopoietic stem cells capable of long-term multilineage engraftment." Science **333**(6039): 218-221.

Ogawa, S. (2019). "Genetics of MDS." Blood **133**(10): 1049-1059.

Ogawa, S. (2020). "Genetic basis of myelodysplastic syndromes." Proc Jpn Acad Ser B Phys Biol Sci **96**(3): 107-121.

Pan, F., et al. (2015). "The TET2 interactors and their links to hematological malignancies." IUBMB Life **67**(6): 438-445.

Papa, L., et al. (2018). "Ex vivo human HSC expansion requires coordination of cellular reprogramming with mitochondrial remodeling and p53 activation." Blood advances **2**(20): 2766-2779.

Papageorgiou, S. G., et al. (2018). "The outcome of patients with high-risk MDS achieving stable disease after treatment with 5-azacytidine: A retrospective analysis of the Hellenic (Greek) MDS Study Group." Hematol Oncol **36**(4): 693-700.

Pappenheim, A. (1896). "Ueber entwickelung und ausbildung der erythroblasten." Archiv für pathologische Anatomie und Physiologie und für klinische Medicin **145**(3): 587-643.

Pappenheim, A. (1907). "Zwei Fälle akuter grosslymphozytärer Leukämie." Fol. Haematol **4**: 301-308.

Patnaik, M. M. and A. Tefferi (2019). "Refractory anemia with ring sideroblasts (RARS) and RARS with thrombocytosis: "2019 Update on Diagnosis, Risk-stratification, and Management"." Am J Hematol **94**(4): 475-488.

Paul, F., et al. (2015). "Transcriptional Heterogeneity and Lineage Commitment in Myeloid Progenitors." Cell **163**(7): 1663-1677.

Pei, W., et al. (2020). "Resolving Fates and Single-Cell Transcriptomes of Hematopoietic Stem Cell Clones by PolyloxExpress Barcoding." Cell stem cell **27**(3): 383-395 e388.

Peled, T., et al. (2012). "Nicotinamide, a SIRT1 inhibitor, inhibits differentiation and facilitates expansion of hematopoietic progenitor cells with enhanced bone marrow homing and engraftment." Experimental hematology **40**(4): 342-355. e341.

Pellegrini, R. (2016). Edit single bases with Benchling.

Perez, B., et al. (2010). "Genetic typing of CBL, ASXL1, RUNX1, TET2 and JAK2 in juvenile myelomonocytic leukaemia reveals a genetic profile distinct from chronic myelomonocytic leukaemia." Br J Haematol **151**(5): 460-468.

Piacibello, W., et al. (1999). "Engraftment in nonobese diabetic severe combined immunodeficient mice of human CD34+ cord blood cells after ex vivo expansion: evidence for the amplification and self-renewal of repopulating stem cells." Blood, The Journal of the American Society of Hematology **93**(11): 3736-3749.

Pietras, E. M., et al. (2015). "Functionally Distinct Subsets of Lineage-Biased Multipotent Progenitors Control Blood Production in Normal and Regenerative Conditions." Cell stem cell **17**(1): 35-46.

Pinho, S. and P. S. Frenette (2019). "Haematopoietic stem cell activity and interactions with the niche." Nat Rev Mol Cell Biol **20**(5): 303-320.

Pittenger, M. F., et al. (2019). "Mesenchymal stem cell perspective: cell biology to clinical progress." NPJ Regen Med **4**: 22.

Popescu, D.-M., et al. (2019). "Decoding human fetal liver haematopoiesis." Nature **574**(7778): 365-371.

Principe, D. R., et al. (2020). "Frequency and prognostic value of mutations associated with the homologous recombination DNA repair pathway in a large pan cancer cohort." Sci Rep **10**(1): 20223.

Pronk, C. J., et al. (2007). "Elucidation of the phenotypic, functional, and molecular topography of a myeloerythroid progenitor cell hierarchy." Cell stem cell **1**(4): 428-442.

Qi, L. S., et al. (2013). "Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression." Cell **152**(5): 1173-1183.

Qu, X., et al. (2018). "TET2 deficiency leads to stem cell factor–dependent clonal expansion of dysfunctional erythroid progenitors." Blood, The Journal of the American Society of Hematology **132**(22): 2406-2417.

Radtke, S., et al. (2017). "A distinct hematopoietic stem cell population for rapid multilineage engraftment in nonhuman primates." Science translational medicine **9**(414): eaan1145.

Rajan, V., et al. (2021). "Stress hematopoiesis induces a proliferative advantage in TET2 deficiency." Leukemia: 1-12.

Ralston, S. H., et al. (2018). Davidson's Principles and Practice of Medicine E-Book, Elsevier Health Sciences.

Ran, F. A., et al. (2013). "Genome engineering using the CRISPR-Cas9 system." Nature protocols **8**(11): 2281-2308.

Rashidi, N. M., et al. (2014). "In vivo time-lapse imaging shows diverse niche engagement by quiescent and naturally activated hematopoietic stem cells." Blood, The Journal of the American Society of Hematology **124**(1): 79-83.

Richardson, C. D., et al. (2016). "Enhancing homology-directed genome editing by catalytically active and inactive CRISPR-Cas9 using asymmetric donor DNA." Nat Biotechnol **34**(3): 339-344.

Rocco, A., et al. (2012). "CD133 and CD44 cell surface markers do not identify cancer stem cells in primary human gastric tumors." Journal of cellular physiology **227**(6): 2686-2693.

Roche-Lestienne, C., et al. (2010). Mutations of TET2, IDH1, IDH2 and ASXL1 In Chronic Myeloid Leukemia, American Society of Hematology.

Rocquain, J., et al. (2010). "Combined mutations of asxl1, cbl, flt3, idh1, idh2, jak2, kras, npm1, nras, runx1, tet2 and wt1 genes in myelodysplastic syndromes and acute myeloid leukemias." BMC cancer **10**(1): 1-7.

Rotunno, G., et al. (2019). "Spectrum of ASXL1 mutations in primary myelofibrosis: prognostic impact of the ASXL1 p. G646Wfs* 12 mutation." Blood, The Journal of the American Society of Hematology **133**(26): 2802-2808.

Rumi, E., et al. (2020). "Myeloproliferative and lymphoproliferative disorders: State of the art." Hematol Oncol **38**(2): 121-128.

Sanjuan-Pla, A., et al. (2013). "Platelet-biased stem cells reside at the apex of the haematopoietic stem-cell hierarchy." Nature **502**(7470): 232-236.

Sasaki, T., et al. (2010). "Regulation of hematopoietic cell clusters in the placental niche through SCF/Kit signaling in embryonic mouse." Development **137**(23): 3941-3952.

Satyan, K. S., et al. (2006). "Phenethyl isothiocyanate (PEITC) inhibits growth of ovarian cancer cells by inducing apoptosis: role of caspase and MAPK activation." Gynecol Oncol **103**(1): 261-270.

Scadden, D. T. (2014). "Nice neighborhood: emerging concepts of the stem cell niche." Cell **157**(1): 41-50.

Schnittger, S., et al. (2011). Somatic mutations and inborn variants in exon 12 of ASXL1 in different myeloid neoplasms, American Society of Hematology.

Schnittger, S., et al. (2013). "ASXL1 exon 12 mutations are frequent in AML with intermediate risk karyotype and are independently associated with an adverse outcome." Leukemia **27**(1): 82-91.

Schofield, R. (1978). "The relationship between the spleen colony-forming cell and the haemopoietic stem cell." Blood Cells **4**(1-2): 7-25.

Scopim-Ribeiro, R., et al. (2017). "TET2 is upregulated during erythroid differentiation of CD34+ cells from healthy donors and myelodysplastic syndrome patients." Applied Cancer Research **37**(1): 38.

Scott, T., et al. (2019). "Improved Cas9 activity by specific modifications of the tracrRNA." Sci Rep **9**(1): 16104.

Shi, H., et al. (2016). "ASXL1 plays an important role in erythropoiesis." Sci Rep **6**: 28789.

Solary, E., et al. (2014). "The Ten-Eleven Translocation-2 (TET2) gene in hematopoiesis and hematopoietic diseases." Leukemia **28**(3): 485-496.

Song, J. C., et al. (2020). "Expert consensus on the diagnosis and treatment of thrombocytopenia in adult critical care patients in China." Mil Med Res **7**(1): 15.

Spangrude, G. J., et al. (1988). "Purification and characterization of mouse hematopoietic stem cells." Science **241**(4861): 58-62.

Sperling, A. S., et al. (2017). "The genetics of myelodysplastic syndrome: from clonal haematopoiesis to secondary leukaemia." Nature Reviews Cancer **17**(1): 5-19.

Steensma, D. P., et al. (2015). "Clonal hematopoiesis of indeterminate potential and its distinction from myelodysplastic syndromes." Blood **126**(1): 9-16.

Stremenova Spegarova, J., et al. (2020). "Germline TET2 loss of function causes childhood immunodeficiency and lymphoma." Blood **136**(9): 1055-1066.

Sutherland, D. R., et al. (1988). "Structural and partial amino acid sequence analysis of the human hemopoietic progenitor cell antigen CD34." Leukemia **2**(12): 793-803.

Takeda, R., et al. (2020). "HHEX promotes myeloid transformation in cooperation with mutant ASXL1." Blood **136**(14): 1670-1684.

Tarunina, M., et al. (2016). "A Novel High-Throughput Screening Platform Reveals an Optimized Cytokine Formulation for Human Hematopoietic Progenitor Cell Expansion." Stem cells and development **25**(22): 1709-1720.

Tavian, M., et al. (1999). "Emergence of intraembryonic hematopoietic precursors in the pre-liver human embryo." Development **126**(4): 793-803.

Tenen, D. G. (2014). "Myeloid differentiation and the leukemia-initiating cell." Leuk Suppl **3**(Suppl 1): S25-26.

Testa, U., et al. (2014). "CD 123 is a membrane biomarker and a therapeutic target in hematologic malignancies." Biomarker research **2**(1): 1-11.

Till, J. E. and C. E. Mc (1961). "A direct measurement of the radiation sensitivity of normal mouse bone marrow cells." Radiat Res **14**: 213-222.

Toren, A., et al. (2005). "CD133-positive hematopoietic stem cell “stemness” genes contain many genes mutated or abnormally expressed in leukemia." Stem Cells **23**(8): 1142-1153.

Torres, R., et al. (2014). "Engineering human tumour-associated chromosomal translocations with the RNA-guided CRISPR–Cas9 system." Nature communications **5**(1): 1-8.

Traina, F., et al. (2014). "Impact of molecular mutations on treatment response to DNMT inhibitors in myelodysplasia and related neoplasms." Leukemia **28**(1): 78-87.

Trentin, J. J. (1971). "Determination of bone marrow stem cell differentiation by stromal hemopoietic inductive microenvironments (HIM)." Am J Pathol **65**(3): 621-628.

Tripodi, J., et al. (2010). "Frequency of heterozygous TET2 deletions in myeloproliferative neoplasms." Cancer Manag Res **2**: 219-223.

Uyhazi, K. E. and J. Bennett (2020). "A CRISPR view of the 2020 Nobel Prize in Chemistry." The Journal of clinical investigation.

Valletta, S., et al. (2015). "ASXL1 mutation correction by CRISPR/Cas9 restores gene function in leukemia cells and increases survival in mouse xenografts." Oncotarget **6**(42): 44061.

Valletta, S., et al. (2020). "Micro-environmental sensing by bone marrow stroma identifies IL-6 and TGFbeta1 as regulators of hematopoietic ageing." Nat Commun **11**(1): 4075.

Vazquez, C., et al. (2020). "Bevacizumab for treating Hereditary Hemorrhagic Telangiectasia patients with severe hepatic involvement or refractory anemia." PLoS One **15**(2): e0228486.

Velten, L., et al. (2017). "Human haematopoietic stem cell lineage commitment is a continuous process." Nature cell biology **19**(4): 271-281.

Veryaskina, Y. A., et al. (2020). "Prognostic Markers of Myelodysplastic Syndromes." Medicina (Kaunas) **56**(8).

Vinjamur, D. S. and D. E. Bauer (2018). "Growing and Genetically Manipulating Human Umbilical Cord Blood-Derived Erythroid Progenitor (HUDEP) Cell Lines." Methods Mol Biol **1698**: 275-284.

Vinjamur, D. S. and D. E. Bauer (2018). Growing and genetically manipulating human umbilical cord blood-derived erythroid progenitor (HUDEP) cell lines. Erythropoiesis, Springer**:** 275-284.

Waddington, C. (1957). The Strategy of the Genes, George Allen & Unwin.

Wagner, J. E., et al. (2015). "Acceleration of umbilical cord blood (UCB) stem engraftment: Results of a phase I clinical trial with stemregenin-1 (SR1) expansion culture." Biology of Blood and Marrow Transplantation **21**(2): S48-S49.

Wagner Jr, J. E., et al. (2016). "Phase I/II trial of StemRegenin-1 expanded umbilical cord blood hematopoietic stem cells supports testing as a stand-alone graft." Cell stem cell **18**(1): 144-155.

Walkley, C. R., et al. (2007). "Rb regulates interactions between hematopoietic stem cells and their bone marrow microenvironment." Cell **129**(6): 1081-1095.

Wang, J., et al. (2014). "Loss of Asxl1 leads to myelodysplastic syndrome–like disease in mice." Blood, The Journal of the American Society of Hematology **123**(4): 541-553.

Wang, J., et al. (2014). "Loss of Asxl1 leads to myelodysplastic syndrome-like disease in mice." Blood **123**(4): 541-553.

Wang, J., et al. (2020). "Characteristics of myeloid sarcoma in mice and patients with TET2 deficiency." Oncol Lett **19**(6): 3789-3798.

Wang, L., et al. (2018). "TET2 coactivates gene expression through demethylation of enhancers." Science advances **4**(11): eaau6986.

Wang, Y., et al. (2015). "WT1 recruits TET2 to regulate its target gene expression and suppress leukemia cell proliferation." Molecular cell **57**(4): 662-673.

Weber, T., et al. (2015). "Increasing the efficiency of homology-directed repair for CRISPR-Cas9-induced precise gene editing in mammalian cells." Nature biotechnology **33**(5): 543-548.

Wienert, B., et al. (2019). "Unbiased detection of CRISPR off-targets in vivo using DISCOVER-Seq." Science **364**(6437): 286-289.

Wilson, A., et al. (2008). "Hematopoietic stem cells reversibly switch from dormancy to self-renewal during homeostasis and repair." Cell **135**(6): 1118-1129.

Wittamer, V. and J. Y. Bertrand (2020). "Yolk sac hematopoiesis: does it contribute to the adult hematopoietic system?" Cellular and Molecular Life Sciences: CMLS.

Wu, L., et al. (2016). "Genetic landscape of recurrent ASXL1, U2AF1, SF3B1, SRSF2, and EZH2 mutations in 304 Chinese patients with myelodysplastic syndromes." Tumour Biol **37**(4): 4633-4640.

Wu, M., et al. (2018). "Comparison of the Biological Characteristics of Mesenchymal Stem Cells Derived from the Human Placenta and Umbilical Cord." Sci Rep **8**(1): 5014.

Wu, Z.-J., et al. (2018). "CRISPR/Cas9-mediated ASXL1 mutations in U937 cells disrupt myeloid differentiation." International journal of oncology **52**(4): 1209-1223.

Xie, M., et al. (2014). "Age-related mutations associated with clonal hematopoietic expansion and malignancies." Nature medicine **20**(12): 1472-1478.

Yamamoto, R., et al. (2013). "Clonal analysis unveils self-renewing lineage-restricted progenitors generated directly from hematopoietic stem cells." Cell **154**(5): 1112-1126.

Yamazaki, J., et al. (2012). "Effects of TET2 mutations on DNA methylation in chronic myelomonocytic leukemia." Epigenetics **7**(2): 201-207.

Yang, H., et al. (2018). "Gain of function of ASXL1 truncating protein in the pathogenesis of myeloid malignancies." Blood **131**(3): 328-341.

Yang, J., et al. (2017). "Activation of Sirt1/FXR Signaling Pathway Attenuates Triptolide-Induced Hepatotoxicity in Rats." Front Pharmacol **8**: 260.

Yang, X., et al. (2020). "The mechanisms of pathological extramedullary hematopoiesis in diseases." Cell Mol Life Sci **77**(14): 2723-2738.

Yee, K., et al. (2020). "Murine double minute 2 inhibition alone or with cytarabine in acute myeloid leukemia: Results from an idasanutlin phase 1/1b study small star, filled." Leuk Res **100**: 106489.

Yeung, A. K., et al. (2020). "Lung megakaryocytes display distinct transcriptional and phenotypic properties." Blood Adv **4**(24): 6204-6217.

Yoshizato, T., et al. (2015). "Somatic mutations and clonal hematopoiesis in aplastic anemia." New England Journal of Medicine **373**(1): 35-47.

Youn, H. S., et al. (2017). "Asxl1 deficiency in embryonic fibroblasts leads to cellular senescence via impairment of the AKT-E2F pathway and Ezh2 inactivation." Sci Rep **7**(1): 5198.

Zahid, M. F., et al. (2017). "Cytogenetic Abnormalities in Myelodysplastic Syndromes: An Overview." Int J Hematol Oncol Stem Cell Res **11**(3): 231-239.

Zhang, C. C. and H. F. Lodish (2008). "Cytokines regulating hematopoietic stem cell function." Current opinion in hematology **15**(4): 307.

Zhang, P., et al. (2019). "The physical microenvironment of hematopoietic stem cells and its emerging roles in engineering applications." Stem Cell Res Ther **10**(1): 327.

Zhang, X., et al. (2020). "TET2 suppresses nasopharyngeal carcinoma progression by inhibiting glycolysis metabolism." Cancer Cell International **20**(1): 1-14.

Zhang, X. H., et al. (2015). "Off-target Effects in CRISPR/Cas9-mediated Genome Engineering." Mol Ther Nucleic Acids **4**: e264.

Zhang, Y., et al. (2018). "Hematopoietic hierarchy–an updated roadmap." Trends in cell biology **28**(12): 976-986.

Zhang, Y. and Y. Gao (2016). "Novel chemical attempts at ex vivo hematopoietic stem cell expansion." International journal of hematology **103**(5): 519-529.

Zhao, H. Y., et al. (2018). "An unbalanced monocyte macrophage polarization in the bone marrow microenvironment of patients with poor graft function after allogeneic haematopoietic stem cell transplantation." Br J Haematol **182**(5): 679-692.

Zimran, E., et al. (2020). "Expansion and preservation of the functional activity of adult hematopoietic stem cells cultured ex vivo with a histone deacetylase inhibitor." Stem cells translational medicine **9**(4): 531-542.