Induced Pluripotent stem cells as a model system to study disease progression in myelodysplastic syndromes

BY

RUBA ALMAGHRABI

A thesis submitted to the University of Birmingham for the degree of

DOCTOR OF PHILOSOPHY

UNIVERSITY OF BIRMINGHAM

Institute of Cancer and Genomics Sciences

College of Medical and Dental Sciences

University of Birmingham

May 2021

Abstract

Myelodysplastic syndromes (MDS) are a heterogeneous group of age-associated hematopoietic diseases characterised by abnormal blood cell maturation and a high propensity for leukemic transformation. It is a clonal disease thought to originate in the haematopoietic stem cell (HSC). Therapeutic strategies in high-risk patients include demethylating agents and cytotoxic drugs, however, 50-60% of these patients do not respond to the treatment and progress to the worst stage. Therefore, there is an unmet clinical need to better understand the mechanisms leading to these blood disorders with the ultimate aim to facilitate the development of improved diagnostic and therapeutic strategies.

By making use of somatic reprogramming and CRISPR-Cas9 tools, our goal in this study was to generate an *in vitro* model of low-risk and high-risk MDS that could help to determine the molecular mechanisms leading to disease progression.

Peripheral blood mononuclear cells were collected from patient MDS27 when he was diagnosed with low-risk MDS; with cells at this stage harbouring *ASLX1*, *SRSF2* and *RUNX1* mutations. These cells were used to generate hiPSC using the non-integrated methods; sendai virus and episomal. Several clones were confirmed to harbour the same somatic mutations as the cells of origin. The pluripotent characteristics of the iPSC clones generated from patient MDS27 as well as from a hiPSC control line were confirmed. Furthermore, the differentiation potential of iPSC into hematopoietic progenitor cells indicated the ability of iPSC to differentiate to hematopoietic stem/progenitor cells (HSPCs) ($CD34^+$ $CD43^+$, $CD34^+$ $CD45^+$). In addition, HSPCs derived from MDS27-iPSC were able to form the different colony-forming unit (CFU) in methylcellulose semi-solid medium with less potential when compared to the hiPSC control.

Moreover, the study of the erythroid and myeloid lineage differentiation in liquid cultures indicated that HSPCs derived from MDS27-iPSC could differentiate to such lineages but with

aberrant morphology, validating our *in vitro* system. To generate a model for high-risk MDS, a mutation in *C/EBPα* causing disruption of the DNA binding domain was generated by CRISPR in the MDS27-iPSC cells, mimicking the additional mutation observed in MDS27 cells when the patient progressed to high-risk MDS. The differentiation of the *C/EBPα* mutant line (high-risk MDS-iPSC) showed a significant reduction in the myeloid and erythroid colony forming units (CFUs) with a block in granulocytic CFU formation. Furthermore, study of the myeloid lineage indicated that the high-risk MDS-iPSC had an impaired myeloid differentiation due to altered expression of key genes required for myeloid differentiation such as *PU.1*, *GATA2*, *LMO2* and *RUNX1*. The study of the erythroid lineage in high-risk MDS27 indicated that the four mutant genes induce erythroid differentiation, increasing the aberrant morphology. Our approach highlights the utility of human iPSCs to understand the molecular mechanisms leading to disease progression and their use as a platform for drug screening which will help to improve diagnostic and therapeutic strategies.

Acknowledgments

Neither words nor actions would express my sincere gratitude to my God and all those who support and help me throughout my PhD journey. First of all, I would like to thank my supervisor Dr Paloma García for her invaluable support, advice, encouragement, and warm hospitality during the past four years. I am appreciative for all the guidance when things were not going well and for all the ideas of how to push this project forward. I am eternally grateful for all the time and effort she puts into mentoring me, as I couldn't have done this work without her. I would also like to acknowledge the amazing members of the García group for support and discussions: Rachel, Nuria, Yara and Ola.

Whatever I am today is only because of my beloved Mama and Baba. I cannot express how grateful I am for their endless help and support. Without their wishes, prayers, endless love, and unlimited support, I would not complete this journey. I would also like to extend my special thanks to my sister; Rabab and brothers; Bander and Bader for their wishes and words to cheer me up and being supportive whenever I need them. A big appreciation goes to my husband (Abdulaziz) for his patience, encouraging me to complete this journey, supporting me spiritually throughout writing the thesis and my life in general. Thanks also to my beloved nephews (Firas & Omar) and nieces (Lamar, Mayan, Meral, Sulaf, Maya and Aleen) who provide unending innocent smiling and inspiration. I am also grateful to my friends I have met throughout this journey, for great memories, crazy journeys, lovely times, and their encouragements and helps to cheer me up whenever I have gotten depressed or homesick. Special thanks to Arwa, Fatama, Renad, Sarah and Mohannad.

Last but not least, I would acknowledge my beloved country, the Kingdom of Saudi Arabia. This work would not have been possible without the financial support.

Table of Contents

Chapter 1 : GENERAL INTRODUCTION 1

1.1 Haematopoiesis 2

1.1.1 Hematopoietic Stem Cells 3

1.1.2 Proposed models of the Hematopoietic system 4

1.1.3 Factors required for haematopoiesis 8

1.2 Erythropoiesis 10

1.2.1 Factors required for erythropoiesis regulation 12

1.3 Myeloid dysplastic syndromes 14

1.3.1 Classification of MDS 14

1.3.2 Chromosomal abnormalities in MDS 19

1.3.3 Recurrent somatic mutations in MDS 22

1.3.4 Evolution of MDS clones and leukemic transformation 28

1.3.5 Treatment of MDS 32

1.4 Human induced pluripotent stem cells 35

1.4.1 Somatic reprogramming 35

1.4.2 The original Yamanaka factors 38

1.4.3 Two-phase model of induced reprogramming 42

1.4.4 Models of somatic reprogramming 43

1.4.5 Limitations of human iPSC 45

1.4.6 The evolution of human iPSC 47

1.4.7 Advantages and applications of human iPSC 54

1.4.8 Use of iPSCs for the study of leukemic and pre-leukemic conditions 57

1.5 Genome editing and CRISPR-Cas9 60

1.5.1 Principle of CRISPR-Cas9 63

1.5.2 Applications of CRISPR-Cas9 65

1.5.3 Limitations and challenges of CRISPR-Cas9 66

1.6 Aim 69

Chapter 2 : Materials ans methods 71

2.1 Somatic reprogramming 72

2.1.1 Clinical history of patient MDS27 72
2.1.2 Patient samples preparation 72
2.1.3 Human iPSCs reprogramming using the integration-Free Sendai virus Vector 73
2.1.4 Human iPSCs generation by integration- Free episomal reprogramming 76
2.2 Human iPSCs culture methods 77
2.2.1 Matrigel 77
2.2.2 Thawing human iPSC on matrigel and passaging 78
2.2.3 Cryopreservation of iPSCs 79
2.3 Alkaline phosphatase staining (AP staining) 79
2.4 Cytogenetics 80
2.5 Immunofluorescence for pluripotent markers 81
2.5.1 Immunofluorescence for cytoplasmic marker TRA1-81 81
2.5.2 Immunofluorescence for the nuclear markers NANOG and SOX2 82
2.6 Human iPSC differentiation 83
2.6.1 Trilineage differentiation of human iPSC 83
2.6.2 Hematopoietic differentiation using Stemdiff protocol 84
2.6.3 Clonogenic progenitor assay 84
2.6.4 Erythroid differentiation 85
2.6.5 Myeloid differentiation 86
2.7 Flow Cytometric analysis of differentiation markers 86
P3.6.2.8.1 88
2.8 Cytospin and Kwik-Diff staining 88
2.9 Genomic DNA extraction from MDS-hiPSC clones 89
2.9.1 DNeasy blood and tissue kit 89
2.9.2 Genomic DNA extraction (Garcia's lab protocol) 89
2.10 Colony genotyping 90
2.11 RNA analysis for gene expression 91
2.11.1 RNA extraction 91
2.11.2 cDNA synthesis 92

2.11.3 Quantitative Polymerase chain reaction (qPCR, Taqman) 92

2.12 Using CRISPR-Cas9 to introduce C/EBPα mutation into MDS-hiPSC (C22) 93

2.12.1 Design of single guide RNA (sgRNA) 93

2.12.2 Annealing sgRNA oligoes 93

2.12.3 Cloning gRNA into pSpCas9 (BB)-2A-GFP (PX458) 94

2.13 Competent bacteria transformation 94

2.14 Plasmid DNA extraction (Miniprep & Maxiprep) 95

2.15 Nucleofection of MDS-hiPSC (C22) with PX458+ C/EBPα gRNA 96

2.16 PCR to amplify the C/EBPα target region 97

2.17 T7 endonuclease I (T7EI) assay to verify successful cutting and gene targeting .. 98

2.18 Agarose gel extraction 98

2.19 Western blotting for the C/EBPa protein 99

2.19.1 Protein extraction 99

2.19.2 Western blot 100

2.20 Statistical analysis 101

Chapter 3 : Generation and characterisation of human induced pluripotent stem cells (hiPSCs) from patient MDS27 102

3.1 Introduction 103

3.2 Results 105

3.2.1 Somatic reprogramming 105

3.2.2 Genotyping for hiPSCs generated from MDS27 107

3.2.3 Identifying the chromosomal stability after somatic reprogramming 112

3.2.4 Pluripotency characterisation of reprogrammed somatic cell lines 113

3.3 Discussion 123

3.3.1 The successful reprogramming of primary cells from MDS27 patient 123

3.3.2 Human iPSC lines from MDS27 patient harbour the same mutations as disease cells 125

3.3.3 Human iPSC lines from MDS27 patient exhibit normal karyotype 126

3.3.4 Human iPSC lines from MDS27 patient are pluripotent 127

Chapter 4 : Validation the disease phenotype of MDS27 129

4.1 Introduction 130

4.2 Results 131

4.2.1 Differentiation of hiPSC control and hiPSC from MDS27 patient to hematopoietic stem and progenitor cells (HSPCs) 131

4.2.2 Assessing the differentiation potential of HSPC from hiPSC control and MDS27-hiPSC in semi-solid medium 136

4.2.3 Studying the erythroid lineage of MDS27-hiPSC 140

4.3 Discussion 147

4.3.1 ASLX1, RUNX1 and SRSF2 mutations do not affect the formation of HSPCs from MDS27-iPSC clones 147

4.3.2 ASLX1, RUNX1 and SRSF2 impair the colony forming potential of MDS27 iPSC clones 149

4.3.3 ASLX1, RUNX1 and SRSF2 affect the quality of the erythroid cells derived from MDS27-C22 hiPSCs 150

Chapter 5 : Studying the contribution of C/EBPα mutation to the disease progression of patient MDS27 154

5.1 Introduction 155

5.2 Results 157

5.2.1 Engineering heterozygous C/EBPα mutation in MDS27-C22 hiPSC 157

5.2.2 Mutant clones display C/EBPα protein-truncating 164

5.2.3 Does the use of the CRISPR-Cas9 system affect the chromosomal stability and the pluripotency of hiPSC? 166

5.2.4 Differentiation of hiPSC control, MDS27-C22 and CRISPR-Cas9 clones to hematopoietic stem progenitor cells (HSPCs) 170

5.3 Differentiation potential of HSPC from hiPSC control, MDS27-C22 and CRISPR-Cas9 hiPSC clones in semi-solid medium 174

5.3.1 Investigating the role of ASLX1, RUNX1, SRSF2 and C/EBPα on the proliferation and self-renewal capacity of MDS27 HSPCs 182

5.3.2 Studying the myeloid differentiation potential of MDS27 hiPSC 185

5.3.3 Investigating the expression of C/EBPα target genes.....191
5.3.4 Studying the erythroid differentiation potential of MDS27 hiPSC containing the C/EBPα mutation.....194

5.4 Discussion 206

5.4.1 Successfully engineered heterozygous C/EBPα mutation, producing a high-risk MDS-iPSC line from low-risk MDS27-C22 line 206
5.4.2 Deletion in the DBD of one allele of C/EBPα lead to a truncated protein.....207
5.4.3 CRISPR-Cas9 does not affect the chromosomal stability and pluripotency of MDS27-C22 hiPSC 208
5.4.4 SRSF2, ASLX1, RUNX1 and C/EBPα mutations do not affect the differentiation of hiPSC into HSPCs 209
5.4.5 ASLX1, RUNX1, SRSF2 and C/EBPα mutations impair the colony-forming potential of myeloid lineage, induce erythroid lineage differentiation, and affect the self-renewal capacity of HSPCs 210
5.4.6 ASLX1, RUNX1, SRSF2 and C/EBPα mutations block the granulocytic differentiation 211
5.4.7 ASLX1, RUNX1, SRSF2 and C/EBPα mutations affect other TFs that have a role in regulating myeloid differentiation.....213
5.4.8 ASLX1, RUNX1, SRSF2 and C/EBPα mutations accelerate erythroid differentiation and generate cells with new dysplastic features 215

Chapter 6 : General discussion.....217

6.1 Implications of research findings 218

6.1.1 Human iPSC as a model system to study low-risk MDS and high-risk MDS 218
6.1.2 ASLX1, RUNX1 and SRSF2 mutations in Patient MDS27 at low-risk MDS lead to dysplasia in erythroid and myeloid cells 220
6.1.3 ASLX1, SRSF2, RUNX1 and C/EBPα mutations block myeloid differentiation and promote erythroid differentiation with an increase in the aberrant cells in high-risk MDS 223

6.2 Limitations of the model used in this study.....226

6.3 Further questions and future work.....227

6.3.1 Determine if the phenotype of high-risk MDS is not relative to the C/EBPα mutation only 227
6.3.2 Check the possible treatment options for the two stages of the disease 227
6.3.3 Xenotransplants using HSPCs from high-risk MDS stage 228

References 229

Supplementary 257

List of figures

Figure 1.1: Sites of haematopoiesis in the human embryo 3
Figure 1.2: Hematopoietic stem cells classical model 6
Figure 1.3: Alternative models of HSC 8
Figure 1.4: Erythropoiesis 12
Figure 1.5: Classification of MDS 16
Figure 1.6: Common driver mutations in MDS and AML 25
Figure 1.7: A model for clonal evolution during haematopoiesis 32
Figure 1.8: Current treatment strategy for MDS 35
Figure 1.9: Advances in the iPSC field 37
Figure 1.10: Generation of Human iPSC 37
Figure 1.11: Mechanistic insights into transcription factor mediated reprogramming 44
Figure 1.12: Methods for delivering reprogramming factors 54
Figure 1.13: Human iPSC applications 55
Figure 1.14: Development of CRISPR-Cas9 systems 63
Figure 1.15: CRISPR-Cas9 principle 64
Figure 3.1: Genotyping screening for MDS27 samples 106
Figure 3.2: Non-integrated method (Sendai virus) ables to generat hiPSC from patient MDS27 107
Figure 3.3: Genotyping screening shows that MDS27 clones harbour the same mutation as original sample 109
Figure 3.4: Different iPSC clones can be generated from PBMNCs of patient MDS27 using episomal method 111
Figure 3.5: Human iPSC clones generated from MDS27 have a normal number of chromosomes 113
Figure 3.6: Human iPSC morphology 114
Figure 3.7: Positive expression of alkaline phosphatase staining in all hiPSC lines 115
Figure 3.8: Positive expression of pluripotent markers by immunofluorescence staining 117
Figure 3.9: Trilineage differentiation of hiPSCs. 119
Figure 3.10: Positive expression of trilineage markers in all hiPSC lines 122
Figure 4.1: Human iPSCs differentiation to HSPCs 132
Figure 4.2: Human iPSCs differentiate to early HSPCs 134
Figure 4.3: Human iPSCs differentiate to late HSPCs 135
Figure 4.4: Differentiation potential of HSPCs in semi-solid medium 137
Figure 4.5: Type of CFUs 139
Figure 4.6: Morphology of CFUs 140
Figure 4.7: Erythroid differentiation of HSPCs 141
Figure 4.8: Assessment of erythroid differentiation of HSPCs 143
Figure 4.9: Morphological analysis of the erythroid cells 146
Figure 5.1: Utilizing CRISPR-Cas9 in hiPSC 158
Figure 5.2: C/EBPα sequence 159
Figure 5.3: Mismatch cleavage assay with T7EI 160

Figure 5.4: Sanger sequencing for positive clones 162
Figure 5.5: C/EBPα protein sequencing 163
Figure 5.6: C/EBPα protein expression in hiPSC clones and Kasumi-1 cells 165
Figure 5.7: Normal number of chromosoms in hiPSC clones after applying CRISPR-Cas9 167
Figure 5.8: Positive expression of pluripotent markers after applying CRISPR-Cas9 169
Figure 5.9: Human iPSCs differentiate to primitive HSPCs 172
Figure 5.10: Human iPSCs differentiate to definitive HSPCs 173
Figure 5.11: Further potential differentiation of HSPCs in semi-solid medium 175
Figure 5.12: Type of CFUs 177
Figure 5.13: Morphology of CFUs 178
Figure 5.14: Characterisation of CFUs 180
Figure 5.15: Morphology of CFUs 181
Figure 5.16: Self-renewal and re-plating capacity of CFUs 184
Figure 5.17: Myeloid differentiation of HSPCs 186
Figure 5.18: Assessment of myeloid differentiation 187
Figure 5.19: Morphological analysis of myeloid cells 189
Figure 5.20: Morphological analysis for aberrant myeloid cells 190
Figure 5.21: Gene expression in HSPCs and myeloid cells 193
Figure 5.22: Erythroid differentiation of HSPCs 197
Figure 5.23: Assessment of erythroid differentiation markers 198
Figure 5.24: Scoring erythroid cells 199
Figure 5.25: Erythroid cell morphology 200
Figure 5.26: DRAQ5 staining for nucleated erythrocytes 202
Figure 5.27: Assessment of enucleation rate 203
Figure 5.28: Morphological analysis of the erythroid cells 205
Figure 6.1: Low-risk MDS model 222
Figure 6.2: High-risk MDS model 225

List of tables

Table 1.1: The 2016 revision to the WHO classification of MDS 17
Table 1.2: Comparison between the 2008 and 2016 editions of the WHO classification of MDS 18
Table 1.3: Major mutant genes in MDS according to their pathways and functions 24
Table 1.4: Latest publication of successful generation of hiPSC from several haematological diseases 58
Table 2.1: Sendai virus reprogramming media 75
Table 2.2: Episomal reprogramming media 77
Table 2.3: Alkaline phosphatase staining buffers.....80
Table 2.4: Primary antibodies used in immunofluorescence.....82
Table 2.5: Secondary antibodies used in immunofluorescence.....83
Table 2.6: Antibodies used in flow Cytometry.....87
Table 2.7: Isotype antibodies used in flow cytometry.....88
Table 2.8: Taqman oligoes 93

List of abbreviations

Acute myeloid leukaemia (AML)

Adeno-associated viruses (AAVs)

Adult haemoglobin (Hb A)

Age- Related Macular Degeneration (ADM)

Alkaline phosphatase (AP)

Allogenic HSC transfusion (allo-HSCT)

Bone marrow (BM)

Bone morphogenetic proteins (BMPs)

Bovine serum albumin (BSA

Burst forming unit- erythroid (BFU-E)

CCAAT-enhancer-binding proteins Alfa and beta (C/EBP- α & β)

Chromosome 11q (del 11q)

chromosome 12p deletion (del 12p)

Chromosome 20q deletion (del 20q)

Chromosome 3 inversion (inv3)

Chromosome 3q deletion del(3q)

Chromosome 5q deletion del(5q)

Chromosome 7q deletion (del 7q)

Chromosome Y deletion (-Y)

Chronic myelomonocytic leukaemia (CMML)

Clonal haematopoiesis of indeterminate potential (CHIP)

Clustered regularly interspaced short palindromic repeats and CRISPR-associated protein 9 (CRISPR-Cas9)

Clustered Regularly Interspaced Short Palindromic Repeats-based gene activation (CRISPRa)

Colony forming unit erythroid (CFU-E)

Common lymphoid progenitors (CLPs)

Common myeloid progenitors (CMPs)

Common myeloid repopulating progenitors (CMRPs)

CRISPR-RNAs (crRNAs)

Daunorubicin and Cytarabine (ara-C) (DA)

Dimethylsulfoxide (DMSO)

DNA methyltransferase (DNMT)

DNA Methyltransferase 3 Alpha or beta (DNMT3A or B)

Dorsal aorta called aorta-gonad mesonephros (AGM)

Double-strand break (DSB)

Embryonic stem cells (ESCs)

Enhancer of zeste homolog 2 (EZH2)

Epithelial-mesenchymal transition (EMT)

EPO receptor (EPOR)

Erythropoietin (EPO)

Fanconi anaemia (FA)

Fetal haemoglobin (Hb F)

Fibroblast growth factor-4 (FGF-4)

Flavobacterium okeanokoites (FokI)

Fluorescence-activated cell sorting (FACS)

Fms-like tyrosine kinase-3 (FLT3)

French American British system (FAB)

Glycophorin A (GAP)

Graft versus host disease (GvHD).

Granulocyte-colony stimulating factor (G-CSF)

Granulocyte-macrophage colony-stimulating factor (GM-SCF)

Granulocyte-macrophage progenitors (GMPs)

H3 dimythylated at lysine 4 (H3K4me3)

Hematopoietic stem and progenitor cells (HSPCs)

Hematopoietic stem cell (HSC)

Histone deacetylase (HDAC)

Homologous recombination (HR)

Hypomethylating agents (HMAs)

Induced pluripotent stem cells (iPSCs)

Insulin-like growth factor-1(IGF-1)

Interleukin-3 (IL-3)

Interleukin-7 (IL-7)

Isocitrate Dehydrogenase-1(IDH1)

Isocitrate Dehydrogenase-2(IDH2)

Ki-ras2 Kirsten rat sarcoma viral oncogene (KRAS)

Kirsten rat sarcoma viral oncogene (KRAS)

Krüppel-like factor-1 or 4 (KLF-1or 4)

Leukocyte antigen (HLA)

LIM-only 2 (LMO-2)

Long-term (LT)

Macrophage-colony stimulating factor (M-CSF)

MDS with multilineage dysplasia (MDS-MLD)

MDS with single lineage dysplasia (MDS-SLD)

MDS-unclassifiable (MDS-U)

Megakaryocyte-erythrocyte progenitors (MEPs)

Megakaryocyte-erythrocyte repopulating progenitors (MERPs)

Mesenchymal stem cells (MSCs)

Mesenchymal-to-epithelial transition (MET)

Monosomy 7 (-7)

Mouse embryonic fibroblasts (MEFs)

Multipotent progenitors (MPPs)

Myelocytomatosis viral oncogene (c-Myc)

Myelodysplastic CMML (MDS-CMML)

Myeloid dysplastic syndrome (MDS)

Myeloproliferative CMML (MP-CMML)

Myeloproliferative leukemia virus oncogene (MPL)

Natural killer (NK)

Neuroblastoma RAS viral oncogene (NRAS)

Nonhomologous end joining (NHEJ)

Nucleophosmin-1 (NPM1)

OCT4, SOX2, KLF4, c-MYC) (OSKM)

Octamer-binding transcription factor 4 (OCT4)

Oestrogen-related receptor beta (ESRRb)

Peripheral blood (PB)

Peripheral blood mononuclear cells (PBMNC)

Phosphate buffered saline (PBS)

Protospacer-adjacent motif (PAM)

RA with an excess of the blast in transformation (RAEB-t)

RA with ring sideroblasts (RARS)

RA with the excess of blasts (RAEB)

Red blood cells (RBCs)

Refractory anaemia (RA)

Refractory cytopenia with unilineage dysplasia (RCUD),

Refractory cytopenias with multilineage dysplasia (RCMD)

Revised IPSS (IPSS-R)

Rho-associated coiled-coil protein kinase (ROCK1 and ROCK2)

Ribosomal protein S14 (RPS14)

Ring sideroblasts (RS)

Runt-related transcription factor-1 (RUNX-1)

Secondary AML (sAML)

Sendai viruses (SeV)

Serine/arginine-rich splicing factor 2 (SRSF2)

Sex-determining region Y box-2 (SOX2)

Short-term (ST)

Signal transducer and activator of transcription (STAT5)

Signalling lymphocyte activation molecule (SLAM)

Single nucleotide polymorphism (SNP)

Single-nucleotide variants (SNVs)

Splicing Factor 3b Subunit 1(SF3B1)

Splicing factors (SFs)

Stage-specific embryonic antigen 1 (SSEA1)

Staphylococcus aureus (*S. aureus*; SaCas9)

Stem cell factor (SCF)

Stem-like megakaryocyte committed progenitors (SL-MkPs)

Streptococcus pyogenes (*S. pyogenes*; SpCas9)

Tet methylcytosine dioxygenase 2 (TET2)

The International Prostate Symptom Score (IPSS)

Thrombopoietin (TPO)

Trans-encoded small (tracrRNA)

Transcription activator-like effectors (TALEs)

Transcription activators- like endonuclease (TALENs)

Transcription factors (TF)

Translocation 3q (t3q)

Trichostatin A (TSA)

Trisomy 19 (+19)

Trisomy 21 (+21)

Trisomy 8 (+8)

Tumour suppressor genes (TSG)

Valproic acid (VPA)

Vascular endothelial growth factor (VEGF)

von Willebrand factor (vWF)

Zinc Finger Nuclease (ZFNs).

Chapter 1 : GENERAL INTRODUCTION

1.1 Haematopoiesis

Haematopoiesis is the process of blood cell production that takes place within the hematopoietic system (Galloway and Zon, 2003). Haematopoiesis occurs under two different stages or waves at different places (Figure 1.1) the first wave of haematopoiesis called “primitive process” occurs during the initial gestation period, in the mouse embryo at day 7 and in the human embryo at 18-20 days. It begins in the extraembryonic yolk sac when the mesoderm differentiates into hemangioblast cells to generate erythrocytes and endothelial cells to supply the emerging yolk sac vasculature. Primitive haematopoiesis is limited, producing primitive erythroid, megakaryocyte and macrophage lineages that derive from primitive erythroid and myeloid progenitors (EMPs) (Tober et al., 2007, Palis et al., 1999). This primitive haematopoiesis is transient and replaced by a permanent wave of haematopoiesis called “definitive haematopoiesis”. The transition from primitive to definitive haematopoiesis or adult hematopoietic stem cells (HSC) occurs at various sites. The aorta-gonad-mesonephros (AGM) region is the main site of definitive haematopoiesis during mid-stage of gestation, soon after the cells migrate and colonise the placenta and foetal liver, the main sites for expansion. Then, around birth, the cells migrate to the bone marrow (BM), where HSCs gain the stem cell characteristics (such as self renewal) and the surface markers of adult HSC (Frame et al., 2013, Palis et al., 1999, Tober et al., 2007). EMPs are transported out of the yolk sac to the liver and differentiated into multiple types of blood cells, including megakaryocytes, enucleated erythrocytes and monocytes. Also, EMP can produce B and T lymphocytes (Boiers et al., 2013).

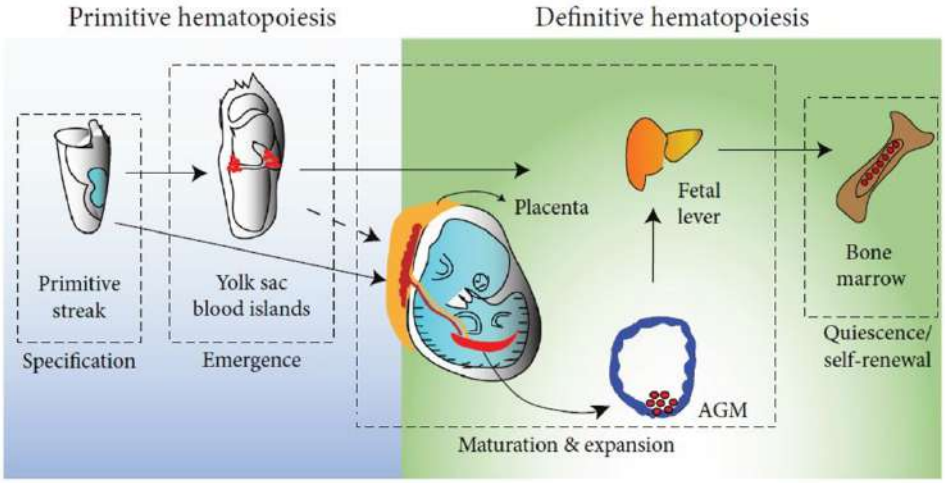


Figure 1.1: Sites of haematopoiesis in the human embryo

Schematic representing the sites of primitive and definitive HSCs during embryonic development. Taken from (Kumar et al., 2019).

1.1.1 Hematopoietic Stem Cells

The hematopoietic system is a hierarchical system with HSC located on the top of the haematopoiesis development hierarchy in the BM. HSCs possess the properties for both self-renewal and multipotency. Through their self-renewal capacity, they are able to divide giving rise to two identical daughter HSCs without differentiation, whilst by being multipotent they have the ability to differentiate into all functional blood cells that control immune function, homeostasis balance and response to microorganisms and inflammation. Till and McCulloch first proposed the significant clarification of the HSC concept in 1961. They observed that lethally irradiated mice transplanted with mouse BM cells developed colonies of hematopoietic cells in the spleen, and analysis of these colonies revealed that were originated from differentiated HSCs (Till and Mc, 1961, Becker et al., 1963).

Within the BM the hematopoietic niche helps to maintain the homeostasis of the hematopoietic system. The BM niche consists of stromal cells, macrophages, osteocytes, adipocytes, vascular endothelium, and the extracellular matrix. HSCs directly interact with BM endothelial cells via

cell adhesion molecules to support survival, proliferation, and differentiation of HSCs (Krebsbach et al., 1999, Zhang et al., 2003).

In addition, the components of the niche are critical regulators of HSC, and the genetic alteration of any major population within the niche has a direct impact on the hematopoietic system due to profound structural and biochemical changes. Consequently, genetic alterations can affect broadly the BM niche by remodelling the niche into an abnormal environment providing preferential support of malignant hematopoietic cells, protecting the cancer cells from therapy, inducing cell death, and driving disease progression (Coskun et al., 2014).

1.1.2 Proposed models of the Hematopoietic system

1.1.2.1 Classical model of haematopoiesis

The cellular function and lineage differentiation capacity of HSCs were determined primarily by Weissman's group in 1988 by using a combination of several markers to isolate mouse HSC and transplant them into lethal irradiated recipient mice, devoid of a functional endogenous hematopoietic system. This assay has long been the gold standard for defining the repopulation capacity of HSCs (Spangrude et al., 1988). Since then, a great effort has been made by several groups to identify surface marker combinations to further purify HSCs using fluorescence-activated cell sorting (FACS). To date, CD34+, Sca-1+, c-Kit+, the signalling lymphocyte activation molecule (SLAM) markers and others are still frequently used to purify HSCs and identify the different progenitor populations (Okada et al., 1992, Osawa et al., 1996, Oguro et al., 2013).

The immunophenotype-based tree-like hierarchy model was originally established by Weissman's group, and then by other groups to better interpret the relationship between HSC and the stepwise differentiation process (Morrison et al., 1997, Spangrude et al., 1988, Akashi

et al., 2000). In this classical model, HSCs are located at the top of the haematopoiesis hierarchy and it can be subdivided into two subpopulations according to their immuno-markers expression: $CD34^{-/low}$ c-Kit^{+} Sca-1^{+} Lin^{-} for long-term (LT)-HSCs and $CD34^{+}$ c-Kit^{+} Sca-1^{+} Lin^{-} for short-term (ST)-HSCs (Figure 1.2). LT-HSCs are a quiescent population in the BM and have long term reconstitution capacity (> 3∼4 months), whereas ST-HSCs only have a short-term reconstitution ability (mostly < 1 month). The differentiation begins from LT-HSCs into ST-HSCs and subsequently, ST-HSCs differentiate into multipotent progenitors (MPPs), which lose the ability to re-enter the HSC compartment and retain potency to differentiate into any mature lineage. Then, the cells at the MPP stage become committed to common myeloid progenitors (CMPs) and common lymphoid progenitors (CLPs). The first branch point at CMPs segregates to granulocyte-macrophage progenitors (GMPs) and megakaryocyte-erythrocyte progenitors (MEPs). Then, GMPs differentiate into mature granulocytes and monocytes and MEPs generate mature megakaryocytes and erythrocytes. On the other hand, CLPs further form T-cells, B-cells, natural killer (NK) and dendritic cells (Osawa et al., 1996, Yang et al., 2005, Metcalf, 2008, Seita and Weissman, 2010). The differentiation of theses lineages is controlled by many pathways, genes, and transcription factors (TF) such as Runt-related transcription factor1 (*RUNX1*) and *GATA2* (Figure 1.2).

With advances in single-cell technology and genetic mouse models, this classical model has been challenged over the past several years, and alternative models have evolved, that are discussed below.

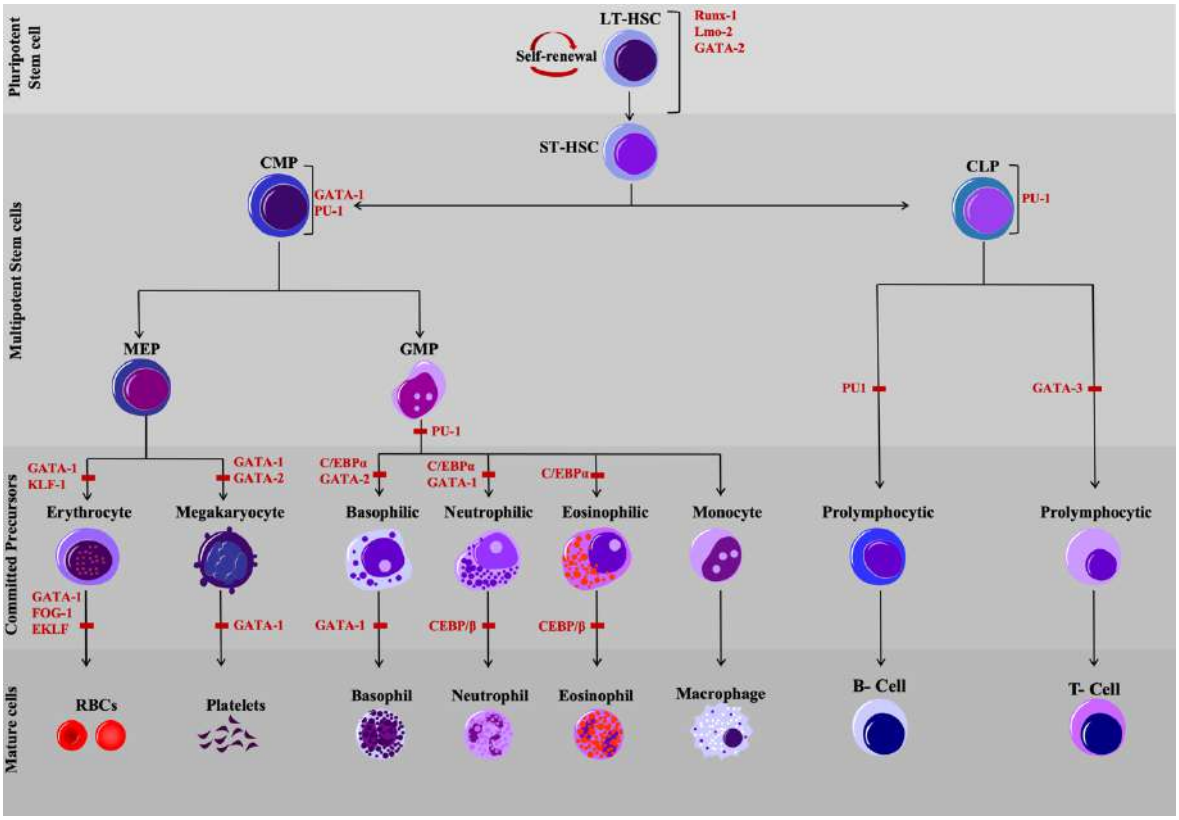


Figure 1.2: Hematopoietic stem cells classical model

Schematic representation of adult HSC differentiation into the specific blood lineages. Key genes known to control each specific lineage are named in red. LT-HSC (long-term hematopoietic stem cell), ST-HSC, (short-term hematopoietic stem cell) CMP (common myeloid progenitor), CLP (common lymphoid progenitor), MEP (megakaryocyte/erythroid progenitor), GMP (granulocyte/macrophage progenitor).

1.1.2.2 Alternative models for haematopoiesis

A series of recent studies have argued the origin of megakaryocytes for recent years. In 2005 Jacobsen's group identified that 25% of LT-HSCs express von Willebrand factor (vWF^{+}), a platelet specific gene (Figure 1.3, A). The vWF+ LT-HSCs myeloid lineage bias has the self-renewal capacity and ability to differentiate directly to MEP and GMP. Whilst, vWF^{-} lymphoid-biased HSCs could further differentiate into lymphoid- primed MPPs (LMPPs), which give rise to granulocytes, macrophages and lymphoid lineages but not megakaryocyte and erythrocyte lineages (Adolfsson et al., 2005). However, lineage-tracing studies challenge this view and suggest that LMPPs also differentiate into the megakaryocyte and erythrocyte lineage (Forsberg

et al., 2006, Boyer et al., 2011). Not only the group of Jacobsen, but also other studies have challenged the classical hierarchical model of haematopoiesis by using single-cell transplantation method. In 2013, Yamamoto *et al*., observed that megakaryocyte precursors are directly derived from HSCs (Figure 1.3, B) and HSCs branch to megakaryocyte repopulating progenitors (MkRPs), megakaryocyte-erythrocyte repopulating progenitors (MERPs) and common myeloid repopulating progenitors (CMRPs) (Yamamoto et al., 2013). In line with the previous study, another group reported that the HSC compartment contains stem-like megakaryocyte committed progenitors (SL-MkPs) and this cell population shares prevalent features with HSCs (Haas et al., 2015). Moreover, a recent study which tracked progenitors and mature lineage cells produced from a single transplanted HSC by Jacobsen's lab showed that a distinct class of HSCs adopts a fate towards LT and significant production of megakaryocytes/platelets without production of any other blood cell lineages (Carrelha et al., 2018). Collectively, these studies suggest that HSCs and the megakaryocyte are closer to one another in the hierarchy of HSC development and the megakaryocytes can arise independently of other lineages.

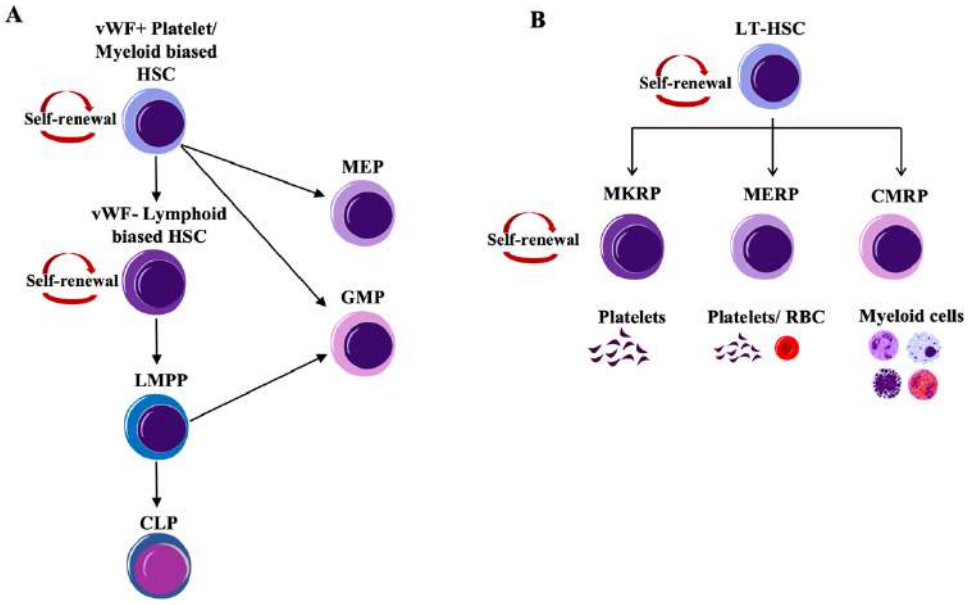


Figure 1.3: Alternative models of HSC

(A) Platelet-biased HSC model identified by Jacobsen's group (Adolfsson et al., 2005). The vWF+ platelet-biased HSCs sit at the apex of the hierarchy and can differentiate into all progenitors and mature cells. vWF− lymphoid-biased HSCs reside downstream of vWF+ HSCs. LMPPs cannot give rise to the megakaryocyte/erythrocyte lineage and MEPs are directly derived from HSCs. **(B)** the LT-HSC population branches directly to give CMRPs, MERPs, and MkRPs. (Yamamoto et al., 2013). vWF (von Willebrand factor), LMPPs (lymphoid- primed MPPs), MEP (megakaryocyte/erythroid progenitor), GMP (granulocyte/macrophage progenitor), CLP (common lymphoid progenitor), MkRPs (megakaryocyte repopulating progenitors), MkERPs (megakaryocyte-erythrocyte repopulating progenitors), and CMRPs (common myeloid repopulating progenitors).

1.1.3 Factors required for haematopoiesis

Haematopoiesis is a tightly regulated process which needs to maintain a balance between immature and mature functional cells, whilst at the same time being able to adapt and increase the production of mature cells in cases of need such as injury, stress, damage, inflammation, and infection. This regulation is achieved by a combination of intrinsic and extrinsic factors. For example, activation and silencing of TFs at different stages of haematopoiesis are important intrinsic factors that dictate the differentiation process (Figure 1.2). It has been identified by several studies that LIM-only 2 (*LMO-2*), *GATA2* and *RUNX1* are essential during the development of both primitive HSC in the yolk sac and definitive HSC in the foetal liver and adult BM (Robertson et al., 2000, Tsai et al., 1994, Yamada et al., 1998, Lie et al., 2018).

Additionally, other TFs are required later within the differentiation of individual blood lineages. For instance, *GATA1* and *PU-1* are known to enhance the differentiation of CMP toward the MEP and GMP (Rhodes et al., 2005). The overexpression of *GATA1* and other TF such as; Krüppel-like factor-1(*KLF-1*) and *P21*(cyclin-dependent kinase inhibitor 1) stimulate the differentiation of MEP to erythroid lineage (Papetti et al., 2010). Moreover, *PU-1* is an essential TF for the differentiation of GMP and CLP to granulocytes, macrophages and B-cells, whilst *PU-1* and *GATA3* stimulate T-cell differentiation (Rothenberg and Scripture-Adams, 2008). Finally, different CCAAT-enhancer-binding proteins alfa and beta (*C/EBP- α & β*) contribute with other TFs to regulate the differentiation of neutrophils and eosinophils (Avellino et al., 2016).

Furthermore, the extracellular matrix (collagen, proteoglycans, and adhesive proteins such as laminin and fibronectin) is an example of an extrinsic factor forming part of the BM niche. It has a role in the structure of the haematopoietic microenvironment and regulates haematopoiesis by containing the haematopoietic cells within the collagen-based haematopoietic microenvironments. The extracellular matrix also regulates the expansion, differentiation, and maturation of HSCs by controlling various interactions of binding proteins, signalling pathways, cytokines, and cells (Panoskaltsis et al., 2005, Charbord and In Zon, 2001). Besides, several cytokines also have a role in the regulation of the early haematopoietic cells such as vascular endothelial growth factor (VEGF), Bone morphogenetic proteins (BMPs), Thrombopoietin (TPO), Fibroblast Growth Factors (FGFs), Stem cell factor (SCF) and Interleukin-3 (IL-3). The combination of these cytokines is essential for survival, proliferation, and differentiation of HSC progenitor cells (Velazquez et al., 2002, Seita et al., 2007, Buza-Vidas et al., 2006, Cho et al., 2002, Donahue et al., 1988). However, other cytokines are key regulators of the differentiation of progenitor cells into the different committed cell lineages.

For example, Granulocyte-macrophage colony-stimulating factor (GM-CSF) is highly expressed on the CMP which promotes the differentiation to granulocyte and macrophage, whilst Interleukin-7 (IL-7) is produced by BM stromal cells and induces the differentiation of CLP. Granulocyte-colony stimulating factor (G-CSF) is a cytokine that drives the differentiation of granulocytes from primitive progenitor HSC whilst macrophage-colony stimulating factor (M-CSF) is a cytokine required to produce mature macrophages. Finally, the high expression of erythropoietin (EPO) stimulates the differentiation towards the erythroid lineage (Benbernou et al., 2000, Katsumoto et al., 2005, Donahue et al., 1988, Grover et al., 2014).

1.2 Erythropoiesis

Red blood cells (RBCs) are essential for delivering O_2 to the cells and tissue of an organism and their production occurs through the process of erythropoiesis. The production of mature RBC entails sequential stages of differentiation and maturation that can be defined by morphology, cell surface markers, and gene expression (Orkin, 2000). The first committed progenitor stage in the erythroid lineage is the burst-forming unit erythroid (BFU-E) (Figure 1.4). Then, BFU-E further differentiates into the colony-forming unit erythroid (CFU-E) and then to proerythroblasts, which are the earliest morphology recognizable as erythroid cells (Gregory and Eaves, 1977). The second phase involves the differentiation of the nucleated precursor from proerythroblast to basophilic, polychromatic and orthochromatic erythroblasts and each maturation stage can be recognised by their specific morphology with distinguishable cytoplasm and nucleus. A gradual accumulation of haemoglobin, progressive reduction in the cell size and nuclear condensation occurs during the second phase, which culminates in the expelling of the nuclei from orthochromatic erythroblasts to produce reticulocytes (Granick and

Levere, 1964). The last phase of erythropoiesis is the maturation of reticulocytes to mature RBCs that is characterised by the very small size and biconcave shape (Gifford et al., 2006). During erythroid differentiation, the hematopoietic markers CD34 and CD45 are gradually lost and the erythroid markers CD71 and CD235a (glycophorin A, GAP) are progressively increased (Figure 1.4) (Chen et al., 2009).

The primitive erythrocytes initially emerge in the yolk sac at 16–20 days of gestation; thereafter erythroid cells appear in the foetal liver at 23-30 days. The primitive erythroblasts are large in size, retain their nuclei and are produced in parallel with endothelial cells from a common precursor termed hemangioblast. Furthermore, these primitive erythroblasts primarily express embryonic globin genes which have a high affinity for O_2 (Palis et al., 1999, Kingsley et al., 2006). Unlike primitive, definitive haematopoiesis of erythrocytes starts from 10.5 weeks onwards in the BM and it is considered a lasting source of mature red cells (Baron et al., 2012). From the BM the erythrocytes enter the circulation as enucleated cells with expression of the foetal haemoglobin (Hb F/ $\alpha 2\gamma 2$) in the foetus which switches to adult haemoglobin (Hb A/ $\alpha 2\beta 2$) in adults (Cantu and Philipsen, 2014, Lu et al., 2008).

The termination of erythroid differentiation occurs in a niche known as erythroblast islands in the BM. The erythroblast islands support the maturation of erythroid cells in the presence of central macrophages (Chasis and Mohandas, 2008). The central macrophages provide the cellular interaction necessary for erythrocyte proliferation, differentiation and enucleation (Rhodes et al., 2008). Furthermore, the central macrophages are also required for supplying iron for heme synthesis during the erythroid development (Leimberg et al., 2008). Macrophages have a positive impact on erythropoiesis. For example, they secrete insulin-like growth factor-1(IGF-1) that enhances the proliferation and maturation of erythroid cells (Sawada et al., 1989). Interestingly, the mature RBCs remain in PB circulation for 120 days, when the RBCs reach

senescence or become damaged, the resident macrophages within the spleen are responsible for removing them from PB by engulfing them through a process called phagocytosis (Crosby, 1959).

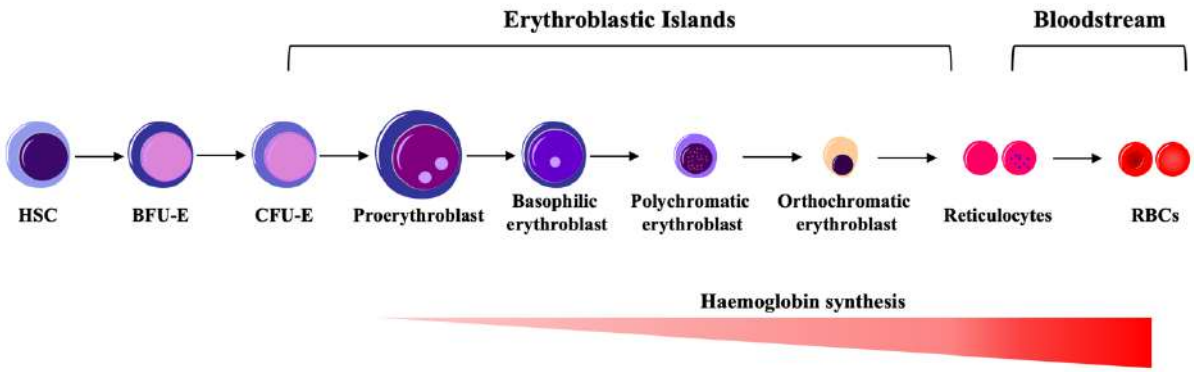


Figure 1.4: Erythropoiesis

Schematic representing the erythrocyte development from the HSC via several committed progenitor stages to mature erythrocyte. HSC (Hematopoietic stem cells), BFU-E (Burst forming unit erythroid), CFU-E (Colony forming unit erythroid), RBCs (Red blood cells).

1.2.1 Factors required for erythropoiesis regulation

Several factors are important for regulating the erythropoiesis process, the most essential factor being EPO. EPO is a glycoprotein hormone secreted mainly from oxygen sensing cells located in the kidney. Thus, during hypoxic stress such as bleeding, disease or senescence, high levels of EPO are secreted (Krantz, 1991). EPO receptor (EPOR) is expressed on the surface of erythroid cells from BFU-E stage onwards (Watowich, 2011, Krantz, 1991), and EPO-EPOR interactions promote the proliferation and differentiation of erythroid progenitors. The interaction between EPO and EPOR stimulates several signalling pathways that are required for erythropoiesis such as Janus kinase 2 (*JAK2*) and Signal transducer and activator of

transcription (*STAT5*). The activation of *JAK2* and *STAT5* promotes the expression of fundamental genes for erythroid progenitor proliferation and differentiation (Grebien et al., 2008). In addition, the STAT5 pathway is required to accelerate erythropoiesis during hypoxic stress (Yan et al., 2012).

Other cytokines have an essential impact on erythropoiesis, such as SCF and IL-3. SCF enhances the sensitivity of cultured erythroid progenitor cells to EPO and IL-3. Also, it has a role in activating several signalling pathways that are essential in erythropoiesis such as *PI3k* by binding to its tyrosine kinase receptor c-Kit (Dzierzak and Philipsen, 2013). IL-3, on the other hand, is important for HSC differentiation and proliferation, but together with SCF can induce the survival of erythroid cells during the early stages of differentiation (Robin et al., 2006).

In addition, other TFs have a crucial role in lineage restricted gene expression during erythroid differentiation. For example, *GATA1*, *GATA2* and *GATA3* negatively regulate the expression of EPO (Imagawa et al., 1997). The expression of *GATA1* is important for proliferation and terminal differentiation of erythroid cells by inhibiting the expression of other genes such as *c-MYB*, *c-MYC* and *c-KIT*; whilst *GATA1* and *GATA2* can stimulate the expression of genes required for erythroid maturation and ultimate ß-globin genes (Moriguchi and Yamamoto, 2014).

Other hormone receptors have been reported to be essential for erythropoiesis regulation such as insulin, glucocorticoids, dexamethasone and estradiol. Dexamethasone and estradiol enhance the proliferation of proerythroblasts and delay maturation in order to increase the number of progenitor cells (Migliaccio et al., 2010). On the other hand, glucocorticoids have a significant role in the development of BFU-E, but do not affect erythroid maturation (Flygare et al., 2011). Other minerals have a critical role in erythropoiesis such as iron, vitamin B12 and folate.

Transferrin -iron is required for Hb synthesis, whilst Vitamin B12 and folate are essential for DNA synthesis and cell division during the differentiation (Fillet and Beguin, 2001, Koury and Ponka, 2004).

Taken together, erythropoiesis is an essential process to provide the body with functional RBCs. However, several blood disorders such as myeloid dysplastic syndrome (MDS), myeloproliferative neoplasm, and chronic myelomonocytic leukaemia commonly exhibit anaemia. For example, in the different MDS stages, infective erythropoiesis is the result of abnormal growth of erythroid progenitors, abnormal erythroid differentiation and increased apoptosis of erythroid progenitors. As a consequence, patients with these type of blood disorders require regular blood transfusions because of impaired erythropoiesis.

1.3 Myeloid dysplastic syndromes

MDS is a group of haematological malignancies that present with abnormal growth and differentiation of HSC (Campo et al., 2011). They are characterised by blood cytopenias that occur because of somatic mutations leading to ineffective haematopoiesis and increase propensity to progress to AML (Corey et al., 2007). According to the statistics of cancer research in the United Kingdom (UK), the incidence of MDS is approximately 30 per 100,000 per year in people over 70 years old and only 1 in 5 people (20%) with MDS are younger than 50 years old (Aul et al., 2001).

1.3.1 Classification of MDS

The clinical heterogeneity of MDS has led to the development of different classifications and prognostic models to identify subsets of MDS patients with similar disease features, patterns of progression, molecular aetiology and the probability of response to standard therapies (Figure

1.5). French American British system (FAB) was the first classification, and it was established in 1982. FAB classification divided MDS into five subgroups based on the number of ring sideroblasts (RS), graduation of monocytosis and percentage of myeloblasts. The five subgroups are Refractory anaemia (RA), RA with ring sideroblasts (RARS), the chronic myelomonocytic leukaemia (CMML), RA with the excess of blasts (RAEB) and RA with an excess of the blast in transformation (RAEB-t) (Bennett et al., 1982).

The restrictions of FAB classification were considered by the WHO classification in 2001 which re-classified the RAEB-T into AML subgroups and placed the CMML into two subtypes: myelodysplastic CMML (MDS-CMML) and myeloproliferative CMML (MP-CMML), based on a white blood cell count cut off of 13×10^9/L because both have myelodysplastic and myeloproliferative features (Jaffe et al., 2001).

Subsequently, in 2008, WHO classification presented additional adjustments, classifying MDS into two main groups: (i) MDS without excess blast (<5% blast) which encompasses five subtypes: refractory cytopenia with unilineage dysplasia (RCUD), RARS, refractory cytopenias with multilineage dysplasia (RCMD), MDS with isolated chromosome 5q deletion *del(5q)* and MDS-unclassifiable (MDS-U); and (ii) MDS with an excess blast (>5% blast) which includes: RAEB-1 (<5% blast in PB and 5%-9% in BM) and RAEB-2, 5%-19% blast in PB and 10%-19% in BM (Vardiman et al., 2009, Swerdlow et al., 2008).

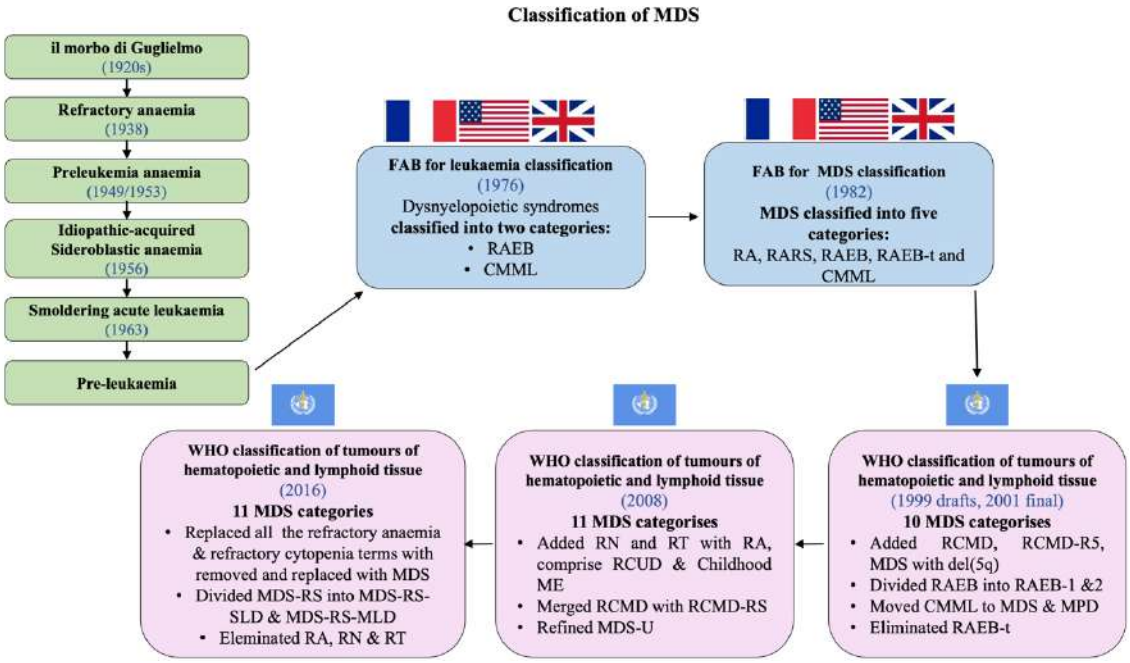


Figure 1.5: Classification of MDS

Schematic representing the evolution of classification of MDS, from the era when this syndrome was poorly characterised and collectively known as "preleukemia" and then, classified by different systems. FAB (French American-British), WHO (World health organization).

In 2016, the latest revision of the WHO classification provided an improvement in the cytopenia, morphological changes, percentage of the blast cells and genetic information in MDS diagnosis and classification as can be seen Table 1.1.

Table 1.1: The 2016 revision to the WHO classification of MDS

The table shows the latest revision of the MDS classification in 2016. WHO classifies MDS according to the number of linage dysplasia (1, 2 or 3 linages), different types of cell cytopenia (e.g. thrombocytopenia or anaemia), percentage of blast cells and type of chromosomal abnormalities.

Type	Dysplastic lineages	Cytopenia	Ring sideroblasts	Blasts	Cytogenetics
MDS-SLD	1	1 or 2	RS<15% (or <5%2)	PB <1% BM <5% No Auer rods	*del(5q)*
MDS-MLD	2 or 3	1-3	RS<15% (or<5%2)	PB <1% BM <5% No Auer rods	*del(5q)*
MDS-RS MDS-RS-SLD	1	1 or 2	RS≧15% (or ≧5%2)	PB <1% BM <5% No Auer rods	*del(5q)*
MDS-RS-MLD	2 or 3	1-3	RS≧15% (or ≧5%2)	PB <1% BM <5% No Auer rods	*del(5q)*
MDS with isolated del(5q)	1-3	1-2	None or any	PB <1% BM <5% No Auer rods	*del(5q)* alone or with 1 abnormality except -7 or del(7q)
MDS-EB MDS-EB-1	0-3	1-3	None or any	PB 2~4% BM 5~9%, no Auer rods	Any
MDS-EB-2	0-3	1-3	None or any	PB 5~19% BM10%~19% Auer	Any
MDS-U With 1% PB blast	1-3	1-3	None or any	PB=1%3, BM<5%, Auer rods	Any
with SLD and pancytopenia	1	3	None or any	PB <1% BM <5% No Auer rods	Any
Defining cytogenetic abnormality	0	1-3	<15%4	PB <1% BM <5% No Auer rods	MDS defining abnormality
RCC	1-3	1-3	None	PB <2% BM <5% No Auer rods	Any

However, MDS related to a single somatic mutation was not considered as diagnostic of MDS in the 2016 edition of the WHO classification (Hong and He, 2017, Arber et al., 2016). In a new edition, they focus on the diagnosis of MDS first and then classification of the disease, for that reason all the terms such as “refractory anaemia” and “refractory cytopenia” were removed and replaced with “myelodysplastic syndrome”. So, the new term of each subtype started with MDS and followed by the classification of the disease. For example, they classified the MDS according to the lineage dysplasia into MDS with single lineage dysplasia (MDS-SLD) and MDS with multilineage dysplasia (MDS-MLD) instead of RCUD and RCMD in 2008, respectively. The modification of other subtypes details is in the Table 1.2 (Hong and He, 2017, Arber et al., 2016).

Table 1.2: Comparison between the 2008 and 2016 editions of the WHO classification of MDS

WHO 2008	WHO 2016
Refractory cytopenia with unilineage dysplasia (RCUD) • Refractory anemia (RA) • Refractory neutropenia (RN) • Refractory thrombocytopenia (RT)	MDS with single lineage dysplasia (MDS-SLD)
Refractory anemia with ring sideroblasts (RARS)	MDS with ring sideroblasts (MDS-RS) • MDS-RS-SLD • MDS-RS-MLD
Refractory cytopenias with multilineage dysplasia	MDS with multilineage dysplasia (MDS-MLD)
Refractory anemia with excess blasts (RAEB) • RAEB-1 • RAEB-2	MDS with excess blasts (MDS-EB) • MDS-EB-1 • MDS-EB-2
MDS with isolated del(5q)	MDS with isolated del(5q)
MDS, unclassifiable (MDS-U)	MDS, unclassifiable (MDS-U)
Refractory cytopenia of childhood (provisional)	Refractory cytopenia of childhood (provisional)

The risk assessment of MDS patients or the risk of leukemic transformation needs to be determined correctly. For this reason, International Prognostic Scoring System (IPSS) published in 1997 and IPSS-R in 2012 provided the new MDS scoring systems, based on BM blast counts, the degree of cytopenia, and cytogenetic analysis (Greenberg et al., 1997, Greenberg et al., 2012). The IPSS system is still the most utilized system in the diagnosis of the MDS. IPSS has two main subtypes: a lower-risk group which encompasses patients who experience the symptoms of chronic anaemia with longer median survival; and, the higher risk group, which encompasses patients with high blast cell count, significant risk of transformation to AML and shorter median overall survival (Greenberg et al., 1997).

IPSS-R considers the depth of cytogenetic features, cytopenia and blast value (Greenberg *et al*., 2012). It is divided into five cytogenetic categories: (i) very good: *chromosome Y* deletion *(-Y)*, *chromosome 11q (del(11q); (ii)* good: normal, *chromosome 5q deletion* (*del(5q)*), *chromosome 20q deletion (del(20q)), chromosome 12p deletion (del(12p))*; (ii) intermediate: *trisomy 8* (*+8*), *chromosome 7q deletion* (*del* (*7q*))*; (iv)* poor: *chromosome 3 inversion (inv3);* and (v) complex karyotype: with 3 abnormalities (Greenberg et al., 2012). It is worth pointing out that the simplicity of the evaluation of the disease gives the IPSS- R system the strength to be the standard prognosis system.

1.3.2 Chromosomal abnormalities in MDS

Chromosomal abnormalities occur in approximately half of MDS patients. The highest frequency of cytogenetic abnormalities has been detected in patients with RAEB- 1 and RAEB-2 and the lowest frequency in patients with RARS with most frequent single cytogenetic abnormalities being *del(5q)*, *del(7q)* and *Trisomy 8 (+8)* (Tefferi and Vardiman, 2009). *Del(5q*) is one of the commonest chromosomal abnormalities in MDS which often has a

consistent clinical phenotype, termed the $5q^-$ syndrome which occurs in roughly 5% of MDS cases. This type of cytogenetic abnormality has a slow rate of progression to AML in comparison to other types of MDS (Tefferi and Vardiman, 2009, Sperling et al., 2017). In addition, *del(5q)* gives rise to the haploinsufficiency of a different number of genes, such as ribosomal protein S14 (*RPS14)*. Haploinsufficiency of the *RPS14* gene has been implicated in the pathogenesis of MDS with *del(5q)* by promoting activation of *p53* to block the differentiation and proliferation of the erythroid lineage (Germing et al., 2012). Furthermore, *micro RNAs*, *mir145* and *mir146a* are critically involved in MDS patients with *del(5q),* which lead to dysmegakaryopoiesis and thrombocytosis (Starczynowski et al., 2010). MDS patients with *del(5q)* show megakaryocytes with a specific distinctive morphology characterised by the presence of small, hypo-lobulated megakaryocytes and these patients show a favourable response to lenalidomide treatment and poor prognosis (Mohamedali and Mufti, 2009). Partial loss of *Del(7q)* or *monosomy 7(-7)* are associated with poor prognosis in MDS patients according to IPSS-R classification (Greenberg et al., 2012) (Greenberg et al., 2012). The role of chromosome 7 abnormalities in the pathogenesis of MDS is not yet well characterised and it was reported that the *del(7)* causes haploinsufficiency in several genes that have a role in different haematological malignancies, such as *EZH2*, *CYX1* and *MLL3* (Xie et al., 2014; Sperling et al., 2017). One study showed that the haploinsufficiency of *MLL3* as a result of *del(7)* in MDS patients leads to leukaemia by contributing to the *RAS* pathway and *TP53* (Chen et al., 2014). However, another study on 280 MDS patients with total deletion (*−7*) and partial deletion *del(7q)* concluded that *del(7q)* group had a more favourable outcome than the *-7* group. This was attributed to the fewer number of mutations in *del(7q)* patients compared to *-7* patients, indicating that the number and type of mutations are more relevant in terms of prognosis and possibly therapy response (Crisa et al., 2020).

Trisomy 8 (+8) is one of the most prevalent chromosome gains, occurring in around 5-7% of MDS patients (Saumell et al., 2012). These patients are classified under the intermediate cytogenetic risk group, according to IPSS- R (Greenberg et al., 2012). In contrast to the previous chromosome aberrations, the primary event of *+8* is not considered as a marker for MDS or any haematological malignancy in the absence of the morphological criteria (Tefferi and Vardiman, 2009). Notably, the occurrence of *+8* is mainly a secondary mutation in MDS patients, especially in MDS *del(5q)* treated with lenalidomide, and it is associated with poor prognosis (Fenaux et al., 2011). A study of 22 patients diagnosed with MDS with isolated *+8* suggested that isolated trisomy should be considered with enough evidence to diagnose MDS in normal and hypercellular bone marrow cases (Saumell et al., 2015).

Del(20q) is another chromosome abnormality that occurs in around 5-8% of MDS patients. The MDS patients with *del(20q)* have a low risk for AML transformation and a favourable prognosis. The patients with this mutation have thrombocytopenia, low blast cells and high reticulocyte counts (Bacher et al., 2009, Gupta et al., 2007, Braun et al., 2011). However, the *del(20q)* is not considered as a prominent marker for MDS in the absence of morphology evidence (Yin et al., 2010). In contrast, another study demonstrated that the *del(20q)* in de nova MDS patients could only be linked with no or minimal morphological dysplasia (Gupta et al., 2007). *Del(20q)* occurs in isolation in 2% of patients, having a median overall survival of only five years and even less when additional chromosomal abnormalities are present (Bacher et al., 2015). As a result of *del(20q)* one or more genes within this region are affected such as *MYBL2* (*B-Myb*), which has a role in cellular proliferation and is an essential protein involved in the maintenance of genome integrity (Aatola et al., 1992). Clarke and colleagues demonstrated a strong correlation between low *MYBL2* RNA expression and reduced expression of other genes related to DNA replication and checkpoint control pathways in CD34+ BM cells from RAEB2

MDS patients (Clarke et al., 2013). This observation was validated by an independent group showing that half of MDS patients display low levels of *MYBL2* regardless of *del(20q)*. These data have demonstrated that genes contained within del(20q) could have a more significant effect on MDS progression than previously anticipated (Heinrichs et al., 2013).

Other less frequent chromosomal abnormalities also have an important diagnostic value for MDS, even without clear morphologic evidence of myelodysplasia. They have been characterised for their importance on patient's prognosis according to IPSS-R such as *del(11q)* (very good prognosis), *isochromosome 17q* or *Trisomy 19 (+19)* (intermediate prognosis), or *inv 3* or *translocation 3q (t3q), chromosome 3q deletion del(3q)* (poor prognosis) (Arber et al., 2016, Hong and He, 2017).

1.3.3 Recurrent somatic mutations in MDS

Little was known about relevant gene mutations implicated in MDS pathogenesis until the early 2000s. In fact, only a handful of genes, including neuroblastoma RAS viral oncogene (*NRAS*), *TP53*, *ATRX* and *RUNX1*, were known to be mutated in MDS (Hirai et al., 1987, Imai et al., 2000, Sugimoto et al., 1993). The development of single nucleotide polymorphism (SNP) array-based genomic copy number analysis (or SNP array karyotyping) has enabled the comprehensive detection of genomic abnormalities and current mutations across the entire cancer genome (Nannya et al., 2005, Yamamoto et al., 2007). Until 2010, these techniques have identified additional gene mutations in MDS patients, including Nucleophosmin-1 (*NPM-1*), Tet methylcytosine dioxygenase 2 (*TET2*), Casitas B-lineage Lymphoma (*CBL*), enhancer of zeste homolog 2 (*EZH2*), Fms-like tyrosine kinase-3 (*FLT3*), Kirsten rat sarcoma viral oncogene (*KRAS*), myeloproliferative leukaemia virus oncogene (*MPL*), Additional sex combs-like (*ASXL1)*, Tyrosine-protein phosphatase non-receptor type 11(*PTPN11*), and *KIT* (Mardis et al., 2009, Yoshida et al., 2011). The revolution in sequencing technologies has introduced a

major development in the genetic analysis of different types of human cancer allowing the entire picture of genetic alterations in MDS in large cohorts of MDS patients to be confirmed. The power of next-generation sequencing has identified the mutations in DNA Methyltransferase 3 Alpha (*DNMT3A*) and Isocitrate Dehydrogenase-1(*IDH1*) based on whole-genome sequencing of AML patients, which turned out to be among the common mutations not only in AML but also in MDS (Mardis et al., 2009, Ley et al., 2010).

Therefore, the mutant genes in MDS are categorized into several discrete functional pathways, including RNA splicing factors (SF), DNA methylation, chromatin modifiers, transcription factors, signal transduction, DNA repair, tumour suppressor genes and cohesion molecules (Table 1.3) (Delhommeau et al., 2009, Haferlach et al., 2014, Papaemmanuil et al., 2013).

Table 1.3: Major mutant genes in MDS according to their pathways and functions

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

The number of mutations is different between MDS patients, depending on the disease classification. For example, a typical low-risk MDS patient harbours a median of two or three driver mutations, high-risk MDS (MDS-EB and MDS-MLD) and CMML patients tend to present higher numbers of mutations than lower-risk MDS patients (Makishima et al., 2017). The mutated genes identified through sequencing are most similar between MDS and primary AML, but the frequencies of these mutations are significantly variable between MDS and primary AML. For instance, the overrepresentation of mutations in receptor tyrosine kinase (*FLT3* and *KIT*), *RAS* pathway genes, *C/EBPα* and *IDH1*/ *IDH2* mutations are commonly founded in AML, while MDS patients have more prevalent mutations in SFs, epigenetic regulators, and chromosomal abnormalities (Figure 1.6). Other mutations in genes such as

RUNX1 and *TP53* are present in the germline and responsible for predisposition to AML and MDS (Makishima et al., 2017).

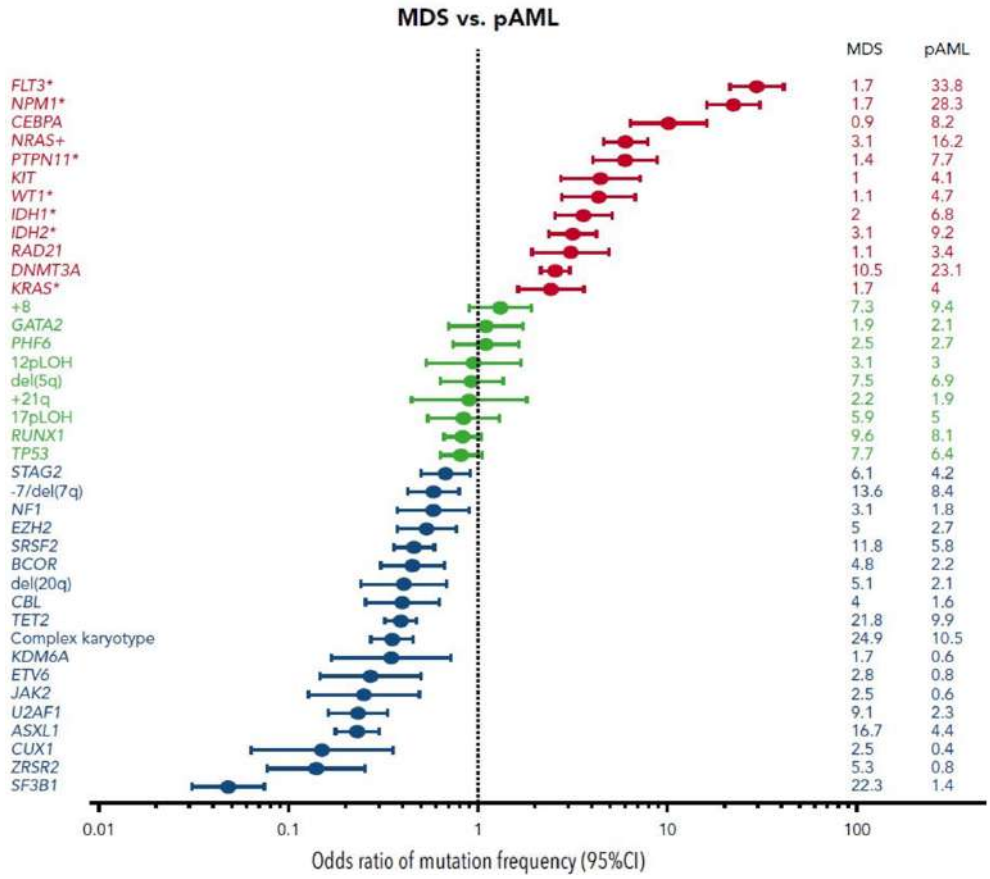


Figure 1.6: Common driver mutations in MDS and AML

The figure illustrates the odds ratios of 95% CIs of frequencies of major mutations and chromosomal abnormalities between MDS and pAML. The frequencies of these mutations are significantly different between MDS and pAML. For example, there is an overrepresentation of mutations in receptor tyrosine kinases (FLT3 and KIT) and RAS pathway genes, as well as CEBPA and IDH1/IDH2 mutations in AML, whilst mutations in splicing factors (SFs) and epigenetic regulators, as well as CNAs, are more prevalent in MDS. CIs (confidence intervals), pAML (Primary acute myeloid leukaemia). Taken from (Makishima et al., 2017).

1.3.3.1 Mutations in RNA splicing genes

RNA splicing factor is a basic cellular function found in all eukaryotes. This two-step process starts with the removal of intronic regions from the newly made pre-mRNA and the joining of exons to form a mature messenger (Wahl et al., 2009). The SF mutational analysis of a large

cohort of 582 cases of different subtypes of MDS and sAML and de novo AML revealed that SF mutations were a common feature of MDS and to a lesser extent in other myeloid neoplasms. Indeed, mutations in SF were detected in all subtypes of MDS at very high frequencies ranging from 45% to 85%, depending on subtype (Yoshida et al., 2011). Furthermore, Serine/arginine-rich splicing factor 2 (*SRSF2*) is one of the most frequently mutated SF genes in MDS (Yoshida et al., 2011). The mutation in *SRSF2* has been associated with worse survival outcomes in low-risk MDS patients and leukemic transformation (Zheng et al., 2017). *SRSF2* is found in many MDS patients and it is frequently mutated together with *TET2* and *ASXL1* mutations (Patel et al., 2017).

1.3.3.2 Mutations in Chromatin modifiers

Chromatin remodelling is a mechanism for modifying chromatin and allowing transcription signals to reach their destinations on the DNA strand (Kouzarides, 2007). *ASXL1* is one of the chromatin remodelling genes mutated in 11–21% of MDS patients and more frequent in high-risk cases of MDS (Lin et al., 2016, Yu et al., 2020). Considerably, *ASXL1* mutations coexist with other mutant genes such as *RUNX1* in MDS patients. The coexistence of *ASXL1* and *RUNX1* mutations have shown to play a role in rapidly progressing disease and driving a leukemic transformation of MDS (Inoue et al., 2015).

1.3.3.3 Mutations in DNA methylation genes

DNA methylation is an epigenetic process that consists of the addition of a methyl group to DNA, often changing gene expression (Moore et al., 2013). *TET2* is one of the highly mutated DNA methylation genes in MDS patients (approximately 50% of MDS cases) (Hussaini et al., 2018, Yu et al., 2020). Based on the known role of *TET2* in normal myeloid progenitor cells,

granulocytes and erythroid cells, mutations in *TET2* have been associated with disturbing erythroid differentiation (Qu et al., 2018).

1.3.3.4 Mutations in Transcription factors

Some hematopoietic transcription factors such as *RUNX1*, *GATA2* and *C/EBPα* are mutated in patients with inherited BM failure disorders and MDS patients and are sometimes associated with poor prognosis (Yu et al., 2020, Menssen and Walter, 2020). For example, *RUNX1* is a key gene for haematopoiesis, the mutation of which could be somatic or germline (Owen et al., 2008). Mutations in *RUNX1* are usually secondary events in MDS, although in a small number of patients they were identified as initiating events in clonal analyses (Papaemmanuil et al., 2013, Bellissimo and Speck, 2017). *GATA2* mutations are somatic mutations that occur in 1-2% of MDS patients (Haferlach et al., 2014, Owen et al., 2008). One study showed that *GATA2* deficiency is associated with haematological abnormalities and MDS, even when there is no evidence of morphologic abnormalities (dysplasia) (McReynolds et al., 2019). C*/EBPα* is another TF which has a role in controlling hematopoietic differentiation, in particular granulocytic differentiation (Wen et al., 2015). C*/EBPα* is highly mutated in AML and during the transformation from MDS to sAML. Although it cannot be used as a prognostic biomarker for MDS patients, it has been reported that *CEBP*α methylation is a common event in MDS (Wen et al., 2015, Yu et al., 2020).

1.3.3.5 Tumour Suppressor genes

Tumour suppressor genes (TSG) are important genes that regulate replication and cell division. They act as negative regulators of the cell cycle, normally acting to inhibit cell proliferation and tumour development by slowing down cell division, repairing DNA mistakes, and/or directing the cells to apoptosis (Cooper and Sunderland, 2000). *TP53* is the most frequently

TSG mutated, occurring in about 5% of MDS patients, but the acquisition of *TP53* mutations in more than 3% of MDS patients are therapy related. As a result of mutated *TP53*, the patients suffer from thrombocytopenia and high blast cell count (Wong et al., 2015).

1.3.3.6 Mutations in cohesion molecules

Cohesion molecules form a multimeric protein complex with a ring-like structure. They are recruited on chromatin in concert with cohesion associated molecules to prevent sister chromatids from premature separation during cell division (Kon et al., 2013). In MDS and other myeloid malignancies, approximately 10-15% of cases harbour mutations in cohesion molecules (Kon et al., 2013). *STAG2* is the most frequently mutated cohesion gene leading to modification of the chromatin structure making it accessible for several TFs, including *RUNX1*, *GATA2* and *ERG* (Noutsou et al., 2017). Furthermore, *STAG2* mutations coexist with mutations in *RUNX1*, *SRSF2*, *ASXL1*, *EZH2,* and are associated with poor prognosis for MDS patients (Haferlach et al., 2014).

Despite the increasing knowledge on the number of mutations observed in MDS and AML patients, all these mutations except one (*SF3B1*) are observed in healthy people during ageing and thus they do not constitute a diagnostic value in MDS. Moreover, how the different mutations and combination of these mutations contribute to the MDS phenotype or leukemic transformation remains unknown. Thus, understanding the role of different mutations will help to diagnose and predict the prognosis of MDS patients (Jaiswal et al., 2014).

1.3.4 Evolution of MDS clones and leukemic transformation

One of the main obstacles to understand MDS progression is the difficulty in tracking an entire history of clonal evolution of MDS back to its cell of origin. However, several attempts have

been made to understand how driver mutations shape MDS development and how these mutations cause the progression to the worse stages or sAML.

Several studies have shown that some of the commonly mutated genes and chromosomal abnormalities observed in MDS and AML patients can be detected in healthy elderly people who are asymptomatic, display normal blood counts and do not have dysplasia. Furthermore, these studies have shown that the age-related genetic changes occurring at the HSC level have an important impact on clonal haematopoiesis, also known as clonal haematopoiesis of indeterminate potential (CHIP) (Genovese et al., 2014, Jaiswal et al., 2014, Laurie et al., 2012).

Thus, CHIP is defined by evidence of somatic mutations or copy number variation in peripheral blood (VAF >2%) in individuals with normal blood count and morphology, and it usually has a low risk of leukemic transformation about 1-2% (Abelson et al., 2018). However, in other cases, patients can suffer from cytopenia and are likely to progress to different malignant phenotypes (Figure 1.7) (Genovese et al., 2014, Laurie et al., 2012, Loh et al., 2018, Steensma et al., 2015).

The first evidence of CHIP was reported in 1996 when Busque and colleagues found skewing of X-chromosome inactivation in blood cells in three different age groups: neonates, females 28 to 32 years old and, females aged 60 years. They found that the occurrence of this phenomenon increased with age (Busque et al., 1996). Then, several years later, the same group identified a *TET2* somatic mutation in some women with skewed X-chromosome inactivation, shedding new light on the existence of somatic mutations in individuals without haematological malignancies (Busque et al., 2012). Accordingly, many recent studies have identified genes frequently mutated in CHIP by sequencing a panel of genes commonly mutated in haematological malignancies. The three most common mutations in CHIP include *DNMT3A*, *TET2* and *ASLX1* along with other genes; *JAK2*, *SRSF2*, *SF3B1* and *TP53* which are mainly

related to leukemic transformation (Buscarlet et al., 2017, Grinfeld et al., 2018, Chen et al., 2019, Makishima et al., 2017).

The transition from CHIP to MDS likely involves a complex interplay between epigenetic alterations such as DNA methylation or chromatin modification within the HSCs and a dysfunctional bone marrow microenvironment. For instance, mutations in DNMT3A or TET2 lead to a hypercellular bone marrow as a result of an expansion of the mutant HSCs during ageing (Ley et al., 2010, Moran-Crusio et al., 2011, Zhang et al., 2016). This finding suggests that the genetic lesions in CHIP initiate the transformation to MDS, which leads to a proliferative advantage over normal HSCs and asymptomatic clonal expansion and eventually to overt disease. Several studies have used advanced methods to calculate the allele frequency of each gene in single-cell sequencing. They have reported that the mutations in SF (*SF3B1* and *SRSF2*), signalling pathways (*NOTCH2* and *KRAS*) and epigenetic modifiers (*TET2* and *DNMT3A*) tend to occur early in the development of MDS. Some of them are frequently detected in CHIP such as *TET2* and *DNMT3A* which indicates CHIP partially corresponds to the initiating mutations in MDS (Papaemmanuil et al., 2013, Makishima et al., 2017, Chen et al., 2019). However, the mutations in transcription factors such as *RUNX1* and *GATA2* are mainly detected as a late event of MDS (Chen et al., 2019, Papaemmanuil et al., 2013).

Importantly, one of the critical challenges in the clinical management of patients with MDS is controlling the disease progression to the worse stages or sAML. The progression of the disease is a result of the emergence of new clones that acquired additional mutations or other genetic abnormalities. Thus far, several studies have attempted to obtain direct insights into the pathogenesis of MDS and progression to sAML by utilising longitudinal, paired samples from different patients with MDS who had later progressed to sAML. These studies have reported a group of genes (*FLT3*, *NPM1*, *NRAS*, *PTPN11*, *WT1*, *IDH1*, *IDH2*) that are mutated in MDS

and sAML samples. These mutations are present in clones which dramatically increase in size during the progression to sAML (Chen et al., 2019, Makishima et al., 2017, Haferlach et al., 2014, Xie et al., 2014). On the other hand, other gene groups such as *TET2*, *TP53*, *EZH2*, *SETBP1* and *KRAS* reside commonly within the dominant clone with early mutations in MDS (low-risk) or CHIP and also have a significant role in leukemic transformation (Makishima et al., 2017, Chen et al., 2019).

Furthermore, mutations in several transcription factors such as *C/EBPα* also commonly occur during the progression to sAML by abrogating normal myeloid cell differentiation (Papaemmanuil et al., 2013, Wen et al., 2015). However, Chen and colleagues have reported that the mutations in Histone Deacetylase 4 (*HDAC4)*, *GLI1* and *RPL22* are observed in only small subclones of MDS stem cells. This study suggests that these mutations are an early event and not associated with disease progression (Chen et al., 2019).

In summary, the advances in genetic analysis provide insights into the clonal dynamics of the disease and allows for the use of subclone events as biomarkers of MDS progression and potentially for target therapy.

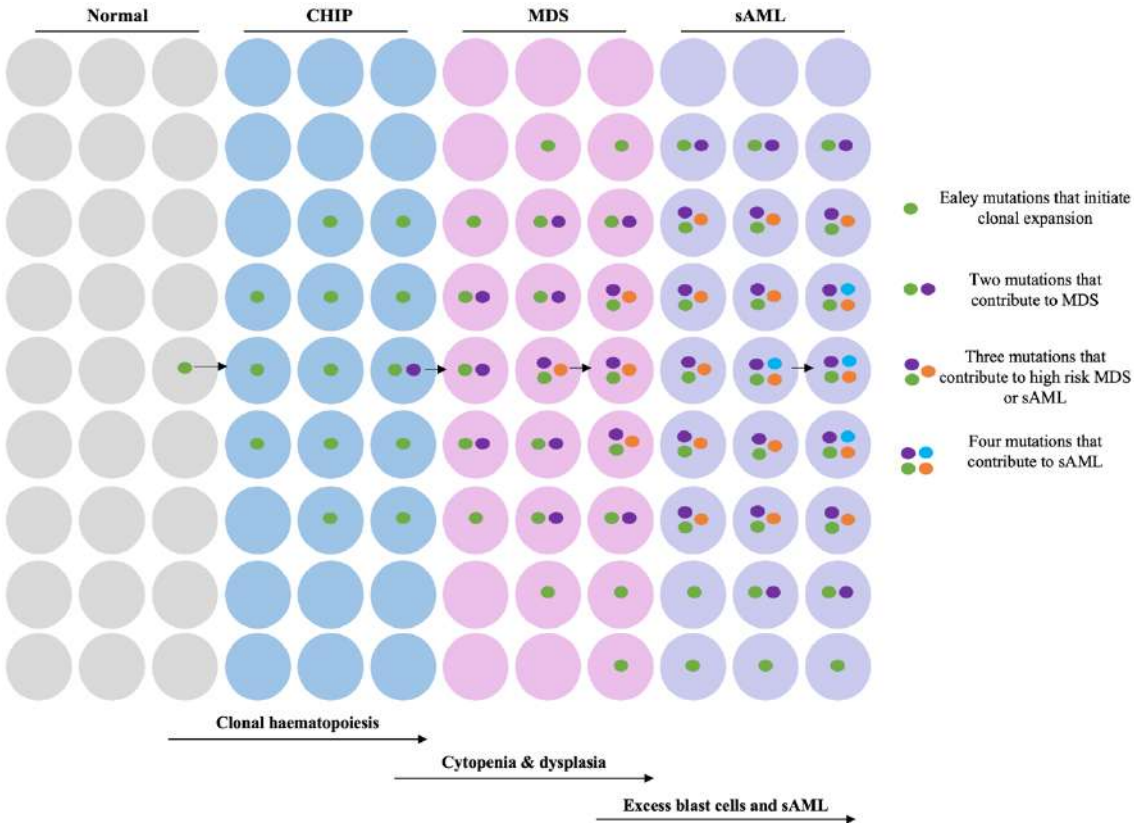


Figure 1.7: A model for clonal evolution during haematopoiesis
Schematic representing the development of polyclonal haematopoiesis from normal clone to CHIP and then in some cases to MDS or to sAML. CHIP (Clonal haematopoiesis of indeterminate potential), MDS (myelodysplastic syndrome), sAML (secondary acute myeloid leukaemia).

1.3.5 Treatment of MDS

Currently, there is no cure for MDS. The median age of MDS patients at diagnosis is over 65 years, and most patients poorly tolerate intensive therapeutic approaches such as chemotherapy, Hypomethylating agents (HMAs), and hematopoietic stem cell transplantation (HSCT). Thus, the main goals of treating patients with MDS are to minimize the disease-associated symptoms, the risk of disease progression and improving survival rates. The IPSS and IPSS-R progression systems help to select the appropriate treatments and give a dynamic estimation of disease prognosis during the treatment course. For example, patients with low and intermediate-1 risk, according to IPSS/IPSS-R, are asymptomatic at diagnosis with mild cytopenia. The treatment

of those patients at the early stage is not known to be beneficial in terms of preventing clonal evolution or death. Therefore, this group of MDS patients is followed regularly without any treatments until their cytopenia worsens or they become more symptomatic. The treatment is then chosen according to the type of cytopenia (Malcovati et al., 2013).

However, the treatment options for patients with high-risk MDS are divided into two main categories: intensive treatment and non-intensive treatment according to the National Health Service (NHS), UK (Figure 1.8), in which intensive treatment is the option for patients who are eligible for allo-HSCT and the combination of chemotherapy regimens (de Witte et al., 2017). A patient who can tolerate intensive treatment typically receives induction or consolidation of Daunorubicin and Cytarabine (ara-C) (DA). HSCT could be considered for those patients if complete remission is achieved, or the percentage of blast cells is reduced to <10%. HSCT is still the only curative option in patients who are fit for intensive treatment (de Witte et al., 2017). According to modelling analyses based on the Centre for International Blood and Marrow Transplant Research, HSCT should be considered for patients with intermediate-2 or high-risk MDS at the time of diagnosis. Further analysis has indicated that these patients with intermediate-2 or high-risk MDS are suitable for HSCT if they are up to the age 70 to 75 years, in good clinical condition and have no severe comorbidities; theses several factors should be taken into consideration because of their potential risk on patient outcome (de Witte et al., 2017, Platzbecker et al., 2012). Overall, it is challenging to determine the optimal moment to integrate HSCT into the therapeutic algorithm in many cases.

On the other hand, the non-intensive treatment option for high-risk MDS depends on the percentage of blast cells. HMAs such as azacitidine, are considered the first line of treatment in high-risk MDS patients who have 20-25% of blast cells (Malcovati et al., 2013). One trial showed that azacitidine has a significant survival benefit in high-risk MDS patients whilst a

higher rate of complete remission is detected in patients who undergo subsequent HSCT with azacitidine (Fenaux et al., 2009). However, several studies found that HMA responses are limited and not evident, especially in patients with mutations in epigenetic regulators (*TET2* and *DNMT3a*), complex cytogenetic abnormalities, high blasts percentage in the BM and high transfusion requirement. Furthermore, it was found that the median response duration was ~1 year, with only a few patients achieving long-lasting remissions; 50-60% of these patients did not respond to HMA and 30% of them developed sAML (Craddock et al., 2017, Bejar et al., 2014, Breccia et al., 2010, Jabbour et al., 2010). However, the Hydroxycarbamide or low dose of cytarabine is considered as a therapy for patients with 30% blast cells (already AML stage) (Craddock et al., 2017, Bejar et al., 2014, Breccia et al., 2010, Jabbour et al., 2010).

Taken together, a subset of MDS patients may benefit from the previous treatment approaches, but still, 40-60% of the high-risk MDS patients fail to respond to current therapies. Hence, to get a mechanistic insight into why such treatments fail in more than 50% of MDS patients and how disease progresses to AML, we need new model systems, allowing us to investigate other treatment options. As recently reported, the use of human induced pluripotent stem cells (hiPSC) could serve as a good system for MDS disease modelling (Kotini et al., 2015, Kotini et al., 2017). However, several animal models have been established to study MDS, such as mice, rats and zebrafish (Beachy and Aplan, 2010, Liu et al., 2017, Huang et al., 2017, Li et al., 2021). These models are very useful preclinical platforms for studying MDS, but they suffer from some limitations. For example, they have low transplantation efficiency, instability and safety concerns and produce phenotypes that are very different from the clinical patients (Li et al., 2021).

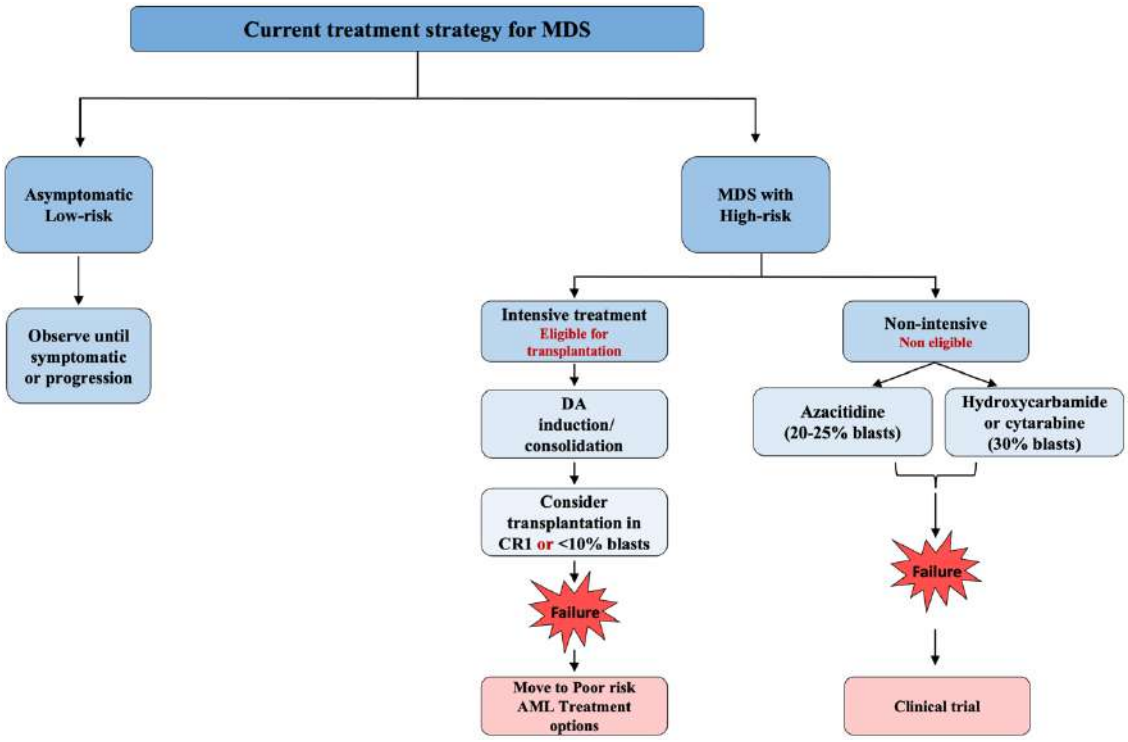


Figure 1.8: Current treatment strategy for MDS

Schematic representing the therapeutic decision tree for all subtypes of MDS. According to the National Health Service (NHS), UK, the treatment options of high-risk MDS depend on the eligibility of HSC transplantation. The treatment of high-risk can be classified as an intensive treatment, which is the combination of chemotherapy, and HSC transplantation and Non-intensive treatment which is only chemotherapy. Clinical trials should be considered for all patients that had treatment failure. CR1 (the first complete remission) DA (Daunorubicin).

1.4 Human induced pluripotent stem cells

1.4.1 Somatic reprogramming

The discovery of iPSCs resulted from the gain in knowledge and technologies of somatic reprogramming that have been developed over decades (Figure 1.9). The first successful somatic reprogramming was reported in 1962 by the group of John Gurdon; using the somatic cell nuclear transfer technique. The group of Gurdon successfully cloned a frog by transferring nuclei from adult somatic cells (intestine cells) of Xenopus into enucleated eggs, (Gurdon, 1962). Subsequent studies using somatic nuclear transfer led to the cloning of the first mammal,

Dolly the sheep (Wilmut et al., 1997). Then, isolation of mouse embryonic Stem Cells (Evans and Kaufman, 1981) and human embryonic Stem Cells hESCs (Thomson et al., 1998) provided the knowledge that resulted in the generation of the induced pluripotent stem cells (iPSCs) (Takashi, 2006 and 2007). First in mouse, and one year later in humans, the group of Yamanaka showed that by introducing four genes namely: Octamer-binding transcription factor 4 (*OCT4*), Sex-determining region Y box-2 (*SOX2*), Krüppel-like transcription factor 4 (*KLF4*) and myelocytomatosis viral oncogene (*c-MYC*) *(*known nowadays as OSKM or Yamanaka factors) into fibroblasts, these cells could revert to an embryonic stem cell (ESC)-like state, becoming pluripotent and capable of differentiating into any cell type in the body (Takahashi and Yamanaka, 2006, Takahashi et al., 2007) (Figure 1.10). The iPSC field has developed rapidly since its first discovery in 2006, with studies aiming to advance the iPSC methodology, the generation of iPSCs from patients with diseases and the use of iPSC with clinical applications (Figure 1.9).

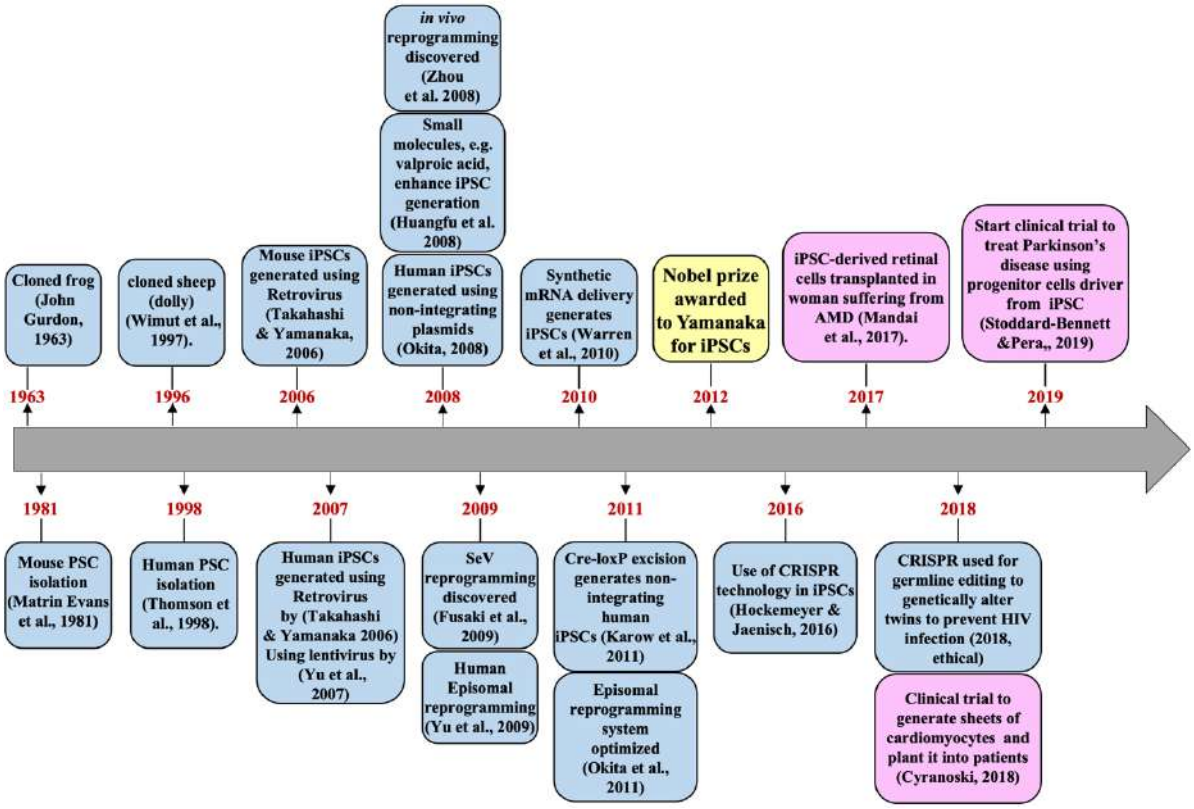


Figure 1.9: Advances in the iPSC field

Timeline representing the advances in the iPSC field since the ground-breaking discovery in 2006 by Yamanaka. Progress has been made on the methods of reprogramming, the cell types that have been reprogrammed and understanding the mechanisms of the reprogramming process. Several iPSCs clinical trials are ongoing in Japan and USA.

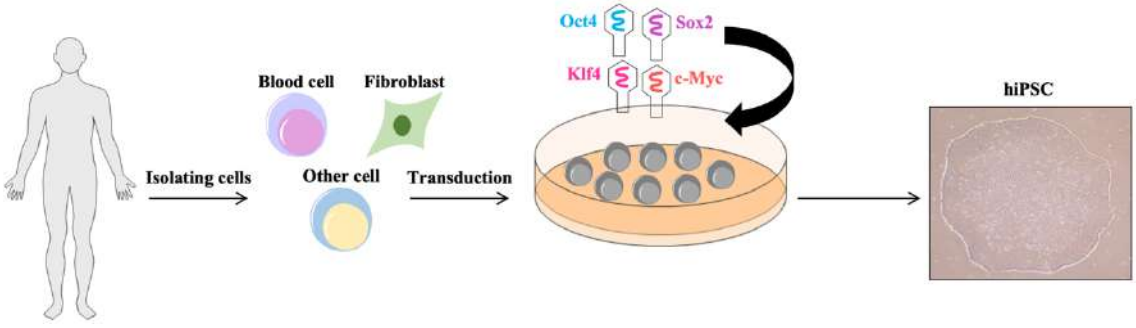


Figure 1.10: Generation of Human iPSC

Schematic representing the reprogramming of human somatic cells to iPSC. Somatic cells (blood cells, fibroblasts, etc) are isolated from a donor. The cells are transduced with four retroviruses containing the transcription factors OCT3/4, SOX2, KLF4, and c-MYC. After approximately a month, iPSCs are generated.

1.4.2 The original Yamanaka factors

1.4.2.1 Oct4

The POU domain transcription factor *OCT4* (also known as *OCT3* or *OCT3/4*) is one of TF in the somatic reprogramming cocktail. It was first described as a fundamental coordinating factor for pluripotency maintenance *in vitro* and *in vivo* (Okamoto et al., 1990, Scholer et al., 1990). *OCT4* is required for the development of the inner cell mass during embryonic development by inducing the secretion of fibroblast growth factor-4 (FGF-4) from the epiblast (Le Bin et al., 2014). Furthermore, it has been reported that *OCT4* acts during lineage specification to suppress the neural differentiation (Wang et al., 2012).

Recently, it was observed that the deletion of *OCT4* from epiblast cells affects the expression of lineage markers and impedes the epithelial to mesenchymal transition process by increasing the expression of E-cadherin; corroborating the requirement of *OCT4* during embryonic development and self-renewal capacity (Mulas et al., 2018).

OCT4 is considered the core pluripotency marker for iPSC generation. However, a recent study showed the successful iPSCs generation by omitting *OCT4* from the reprogramming cocktail. Generation of iPSCs under these conditions occurs through direct activation of *NANOG* by SKM factors and subsequent activation of endogenous *OCT4* and the rest of the pluripotency network by NANOG (Velychko et al., 2019). Moreover, according to RNA sequencing and ChIP sequencing data, somatic reprogramming using OSKM produces transient off-target gene upregulation driven by exogenous *OCT4* overexpression which cannot be detected in the cells reprogramed with the SKM cocktail (Velychko et al., 2019).

1.4.2.2 Sox2

SOX2 gene is a member of the SOX protein family in which Sry protein binds to specific DNA sequences through their high-mobility-group domain (Gubbay et al., 1990). To date, around 20 different *SOX* genes have been described in murine and human genomes which have been categorized into eight subgroups depending on the sequence identity and functional similarity (Wegner, 2010). Among all *SOX* genes, *SOX2* is likely the most studied due to its significant role in somatic reprogramming (Takahashi and Yamanaka, 2006, Takahashi et al., 2007). Indeed, *SOX2* is highly expressed in the inner cell mass at the early stage of embryonic development. Expression of *Sox2* at the 2-cell stage of murine embryos improves the development up to the blastocyst stage and its absence results in embryonic lethality (Pan and Schultz, 2011, Avilion et al., 2003). Many studies have reported that the deletion or overexpression of *SOX2* in ESCs results in their differentiation to ectoderm, mesoderm, endoderm and trophectoderm-like cells, demonstrating the essential role of *SOX2* in the maintenance of the stem cell phenotype in ESCs (Kopp et al., 2008, Boer et al., 2007, Adachi et al., 2010). Aside of its role in ESC, SOX2 also has a role in maintaining the stemness, proliferation and proper differentiation of mesenchymal stem cells (MSCs) (Yoon et al., 2014, Han et al., 2014). More recently, the importance of *Sox2* in adult stem cells has been revealed through the generation of a *Sox2* knockdown mouse model. *Sox2* knockdown mice suffer from premature ageing withe kyphosis, hair greying and reduced fat mass (Vilas et al., 2018).

During iPSC generation *OCT4* and *SOX2* cooperatively bind to the regions of different pluripotency genes such as *NANOG*, *FGF4* and *UTF* to induce their expression whilst inhibiting epithelial-mesenchymal transition (EMT) mediator genes such as *Snail* (Hoffding and Hyttel, 2015, Okumura-Nakanishi et al., 2005, Liu et al., 2018).

1.4.2.3 KLF4

KLFs belong to a relatively large family of Sp1-like transcription factors with over 20 members (Turner and Crossley, 1999). *KLF4* is one member of *KLF* family and is expressed in a wide range of tissues in mammals and plays a significant role in myriads of physiological processes such as proliferation, differentiation, development, transcriptional regulation, DNA repair, maintenance of normal tissue homeostasis and apoptosis (Feinberg et al., 2007, McConnell et al., 2007, Yoon et al., 2005, Yusuf et al., 2008). Furthermore, *KLF4* is known as a repressor of the cell cycle, blocking G1/S progression and mediating *p53*-dependent G1/S cell cycle arrest in response to DNA damage (Chen et al., 2001, Yoon et al., 2005). On the other hand, *KLF4* has a vital role as an activator and suppressor of essential genes that facilitate the reprogramming process. For example, *KLF4* plays a role in activating the transcription of E-cadherin, an adhesion protein important for mesenchymal-to-epithelial transition (MET) that occurs in cells undergoing somatic reprogramming towards iPSCs (LI et al., 2010). The global analysis of promoter occupancy by OSKM factors has identified that *Klf4* is an upstream regulator of a large feed-forward loop that contains these three TFs, *Oct4, Sox2* and *c-Myc*, as well as other common downstream factors including *Nanog* (Kim et al., 2008a, Zhang et al., 2010). Clearly, from previous findings, *KLF4* has a crucial role in somatic cell reprogramming and maintenance of ESC self-renewal.

1.4.2.4 c-MYC

MYC (myelocytomatosis viral oncogene) belongs to the *MYC* family of basic helix-loop-helix leucine zipper transcription factors together with, *MYCN* and *MYCL* (Blackwood and Eisenman, 1991). *c-myc* is known to be a proto-oncogene which causes leukaemia in animals such as birds and cats (Donner et al., 1982). It is central in the transcriptional network of many cells, and plays a role in the regulation of genes related to ESCs (Singh and Dalton, 2009). *c-*

MYC binds and regulates *OCT4* and *SOX2* genes and overexpression of *c-MYC* has been shown to increase transcript abundance of *OCT4* and *SOX2*, by accumulative occupancy at core promoters (Lin et al., 2012). Thus, *c-MYC* could play a role in switching on the endogenous *OCT4* and *SOX2* expression to establish the core pluripotency network.

Moreover, *c-MYC* regulates the genes involved in self-renewal and pluripotency in ESCs by transcriptionally activating *miRNAs* that block the differentiation and enhance cell division (Lin et al., 2009). Indeed, the regulation of self-renewal and pluripotency was corroborated by other studies where co-deletion of both *c-MYC* and *MYCN* disrupted the maintenance of ESCs and iPSCs and promoted their differentiation (Fagnocchi et al., 2016, Varlakhanova et al., 2010). In addition, it has been shown that removing *c-MYC* from the reprogramming cocktail affects reprogramming efficiency, in terms of the number of iPSC clones and the time required to reach the full reprogrammed state (Werning et al., 2008). Despite the possibility of generating iPSC in the absence of *c-MYC*, other studies have highlighted the importance of c-*MYC* to generate fully reprogrammed cells due to its ability to recruit histone acetyltransferases which is required to increase the expression of pluripotent genes at the early stage of reprogramming (Sridharan et al., 2013, Araki et al., 2011). Accordingly, different studies have investigated the role of histone deacetylases (HDAC) inhibitors such as the valproic acid (VPA) and trichostatin A (TSA), and found an improved somatic reprogramming efficiency of both human and murine fibroblasts in the absence of C-*MYC* (Araki et al., 2011, Huangfu et al., 2008, Duan et al., 2019). Taken together, these findings support the importance of *c-MYC* for efficient somatic reprogramming to iPSC.

However, one group has been able to generate iPSC from neural stem cells isolated from the adult murine brain using three different combinations of three factors *Oct4*, *Klf4* and *c-myc* (OKM); *Oct4*, *Klf4* and *Sox2* (OKS); and *Oct4*, *c-myc* and *Sox2*(OMS,) *Oct4* and *Klf4*

with the absence of Sox2 and *c-myc* and they have shown that three-factor iPSCs could be generated in the absence of *Sox2*, *Klf4* or *c-myc* in mouse neural stem cells, which endogenously express these three factors (Kim et al., 2008b).

1.4.3 Two-phase model of induced reprogramming

Different studies have shed light on how the ectopic expression of OSKM factors revert the somatic cells to the pluripotent state (Yamanaka, 2007, Hansson et al., 2012, Buganim et al., 2012). According to these studies, the somatic reprogramming process consists of two broad phases: **Early phase of reprogramming;** the exogenous *OCT4*/*SOX2* bind to the promoter of epithelial associate micro RNAs (*miRNA*) which leads to downregulation of the fibroblast genes (*SNAIL11*, *SNAIL2*, Zinc Finger E-Box Binding Homeobox 1 and 2 (*ZEB1 and 2)*) and upregulation of the epithelial genes (*Cadherin1*(*cdh1), epithelial cellular adhesion molecule (EpCAM)*) in order to activate the fundamental process, called MET. It is a critical step in the early reprogramming phase, and it can be recognised by morphological changes, increased proliferation, and the formation of cell clusters (Takaishi et al., 2016).

Several studies have used multiple surface and genomic markers to investigate the changes during the reprogramming process. These studies revealed that during the early phase of reprogramming OSKM factors induce the expression of markers such as alkaline phosphatase (AP), f-box 15 protein (Fbx15) and Stage-specific embryonic antigen 1 (SSEA1) (Brambrink et al., 2008, Hansson et al., 2012, Polo et al., 2012, Stadtfeld et al., 2008b).

Then, once the pluripotency network is activated, epigenetic remodelling takes place. Importantly, chromatin remodelling of the full array of pluripotent genes is needed to complete the early phase of the reprogramming. Chromatin remodelling includes the gradual unfolding of condensed heterochromatin to form an open euchromatin conformation and removes repressive H3K9me3 histone marks. This chromatin remodelling happens by the effect of *c-*

MYC, *Klf4* and histone modification enzymes (acetyltransferase and demethylases) (Soufi et al., 2012, Omole and Fakoya, 2018, Takahashi and Yamanaka, 2016).

Generally, in the early phase the completion of global chromatin remodelling is achieved, and the cells are ready to enter the late phase of reprogramming. **During the Late phase,** of the reprogramming most of pluripotency associated genes are gradually activated; some of the first markers to be detected are *Fbxo15*, *Sall4* and endogenous *Oct4*. Following, at the very end of this phase, *Sox2* or *Dppa4* can be detected as the cells can self-renew independently of the Yamanaka factors which is the main feature of late phase reprogramming (Stadtfeld et al., 2008b, Polo et al., 2012).

1.4.4 Models of somatic reprogramming

The molecular mechanisms of somatic reprogramming are still not fully investigated. Different models have evolved aiming to explain the mechanism of somatic reprogramming (Figure 1.11) (Smith et al., 2016, Takahashi and Yamanaka, 2016).

The "elite model" is the first model that assumes that progenitor cells within the heterogeneous population are able to induce the cells toward pluripotency, but the majority of differentiated cells never complete the reprogramming process (Takahashi and Yamanaka, 2016). However, this model has been ruled out by lineage tracking studies and clonal analysis, which have shown that differentiated cells such T and B lymphocytes are able to reprogram to iPSC (Stadtfeld et al., 2008a, Hanna et al., 2008).

Stochastic and deterministic models share the view that all the somatic cells can generate iPSC after being transduced with OSKM factors. In principle, the path to pluripotency could occur either via a stochastic or a deterministic mechanism: if the somatic reprogramming occurs through a stochastic process, the iPSCs appear at a different, random, and unpredictable times.

The stochastic model indicates that a reprogramming event may or may not be successfully achieved. However, the occurrence of somatic reprogramming in a deterministic manner leads to the iPSC appearing at a fixed and predictable time (Omole and Fakoya, 2018, Takahashi and Yamanaka, 2016).

Overall, successful reprogramming requires a precise and limited level of expression of the OSKM factors, and this balance of expression is fundamental for reprogramming.

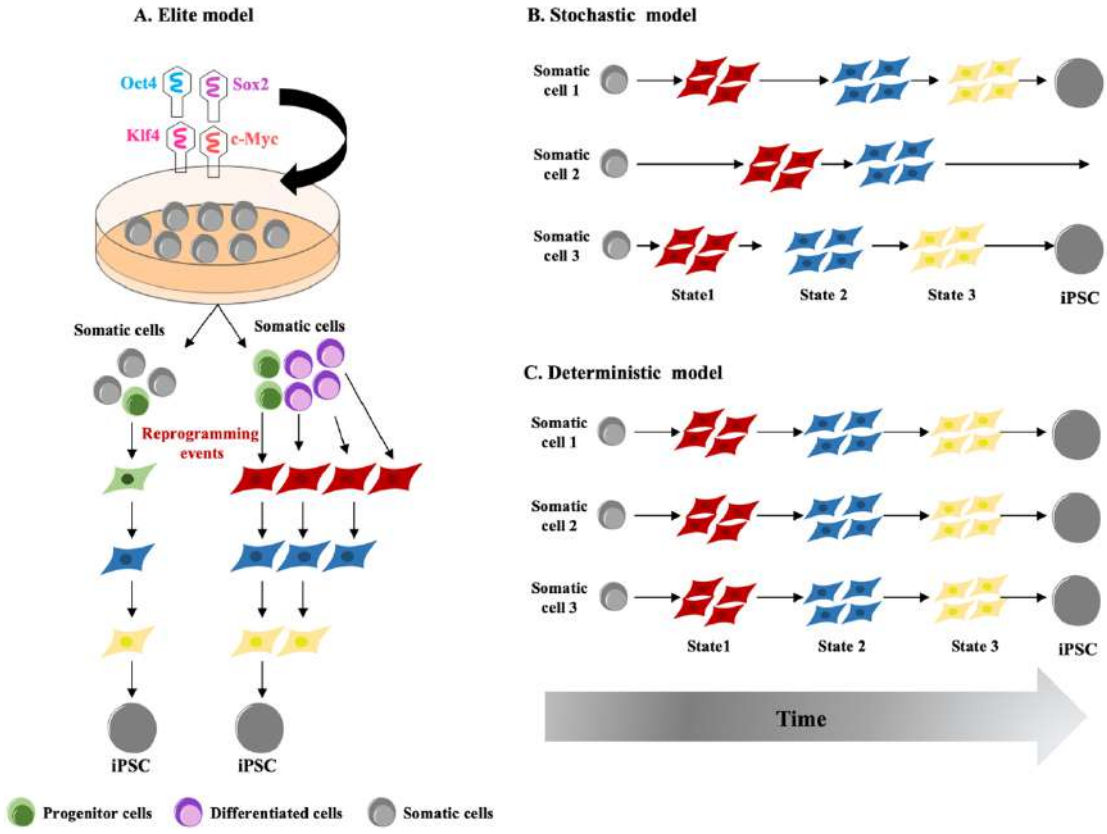


Figure 1.11: Mechanistic insights into transcription factor mediated reprogramming

Schematic representing the three models that aim to explain the mechanism of the reprogramming process. **(A)** The "elite model" assumes that only progenitor cells can generate iPSC, and the majority of differentiated cells never complete the reprogramming process. **(B)** The "stochastic model" is based on the ability of iPSCs to appear at different, random, and unpredictable times. **(C)** The "deterministic model" is based on the ability of iPSCs to appear at a fixed and predictable time.

1.4.5 Limitations of human iPSC

The work on somatic reprogramming by many researchers worldwide has led to the understanding of the limitations and improvements needed to bring iPSC closer to safe clinical applications. For example, despite improvements in efficiency from Yamanaka's original work in mouse and human cells (Takahashi et al., 2007, Takahashi and Yamanaka, 2006), the efficiency rate of the reprogramming is still incredibly low (often much less than 1%).

Furthermore, the Yamanaka factors, *OCT4*, *SOX2*, *KLF4* and *c-MYC,* have a role in the tumourigenicity of iPSCs (Okita et al., 2007). Particularly, the overexpression of *Oct4* is associated with murine epithelial cell dysplasia (Hochedlinger et al., 2005); serrated polyps and mucinous colon are detected with the high expression of *Sox2* (Park et al., 2008); *Klf4* and *c-MYC* are potent oncogenes, with overexpression of *KLF4* detected in breast cancer, and *c-MYC* overexpression associated with 70% of human cancers (Medvedev et al., 2010).

Another limitation about iPSCs generation is the delivery method of the OSKM factors. The original work of Yamanaka, implied the use of retrovirus to deliver OSKM factors, leading to a higher rate of genomic alterations (Takahashi et al., 2007). Indeed, the integration of retroviral DNA into the host cell genome could interrupt coding sequences or promoter elements affecting the transcription of many genes (Takahashi and Yamanaka, 2016).

Several studies have reported that the generation of hiPSCs leads to single-nucleotide variants (SNVs) and genome instability (Yoshihara et al., 2017, Bhutani et al., 2016). For example, one study has observed additional abnormalities (hyperploidy) in iPSC cell lines generated from a Fanconi anaemia (FA) patient (Yung et al., 2013). In addition, chromosomal abnormalities have

also been detected in iPSC lines making these cells useless for disease modelling, drugs screens and therapeutic applications (Taylor et al., 2014).

In addition, losing the features of malignancy, disease phenotype, gene expression, molecular changes, or drug response are other limitations of iPSC. This could be related to the resetting of the starting cell's epigenetic landscape to that of a pluripotent stem cell. For example, iPSC lines derived from myeloid malignancies behave as hematopoietic cells with normal morphology and function (Amabile et al., 2015). Moreover, other studies have demonstrated that differentiation of iPSC lines generated from patients to hematopoietic cells restores their malignant phenotypes (Chao et al., 2017, Kotini et al., 2017, Miyauchi et al., 2018).

Another issue related to iPSC generation is that some diseases have hurdles in reprogramming such as haematological malignancy. Presumably, this could be due to several reasons: Firstly, the genetic mutations and genomic instability associated with the disease almost certainly affect the reprogramming efficiency. Secondly, by natures cancers are often heterogeneous; the reprogramming process may occur more efficiently in normal cells over other cells with certain mutations and chromosomal aberrations. For example, one study has reported the failure to generate iPSC from different patients with FA, suggesting that mutations in genes related to DNA repair could affect the reprogramming process (Raya et al., 2009). Furthermore, another group aiming to generate hiPSC lines from AML patients, was able to generate only one clone that captured only the *ASLX1* mutation (Gomez Limia et al., 2017). In addition, another group reported the successful reprogramming of one AML patient sample out of 16 AML patients used (Lee et al., 2017a).

Despite all these drawbacks, much progress has been achieved in the past decade to address most of the previous limitations to improve the reprogramming technique and to use iPSCs safely in the clinic.

1.4.6 The evolution of human iPSC

1.4.6.1 Manipulation of transcription factors cocktail

Different strategies have been followed to overcome the drawbacks of somatic reprogramming technology in order to use iPSCs as effective research model systems, and ultimately translate this technology into clinical applications.

As mentioned previously, Yamanaka's cocktail of genes led to the generation of iPSCs with a low reprogramming efficiency and a high risk of tumourigenicity. Thus, several studies have focused on manipulating the TF cocktail by adding new TFs that can replace one of the original Yamanaka factors or collaborate with the original factors to improve efficiency. For example, one group demonstrated that *OCT4*, *SOX2*, *NANOG* and *LIN28* (OSNL) could reprogram human somatic cells to hiPSC. This finding showed that *NANOG* and *LIN28* can substitute the functional role of the oncogenic *KLF4* and *c-MYC* during somatic reprogramming (Yu et al., 2007). Furthermore, another study successfully converted mouse embryonic fibroblasts (MEFs) to iPSC by replacing *Klf4* with oestrogen-related receptor beta (*Esrrb*), which plays a significant role in recruiting the core pluripotency factors such as *Oct4*, *Sox2* and *Nanog* (Feng et al., 2009).

Cell cycle regulatory genes, such as *p53,* have been shown to have an essential role in somatic reprogramming. *p53* is a tumour suppressor protein that can negatively regulate the cell cycle by promoting senescence and inhibiting cell growth, thus having an inhibitory effect on iPSC generation (Hong et al., 2009, Kawamura et al., 2009, Marion et al., 2009). In 2009,

Yamanaka's group showed that *p53* downregulation increased the efficiency of somatic reprogramming. This work revealed a fivefold increase in iPSC colonies from *p53*+/- MEF, and significantly more colonies were obtained from *p53*-/-MEF compared with *p53*+/+ fibroblasts (Hong et al., 2009). Thereafter, several studies reported that inhibition of *p53* could significantly accelerate iPSC colony formation and promote the reprogramming efficiency with less susceptibility to differentiate (Marion et al., 2009, Kawamura et al., 2009, Brosh et al., 2013). Besides, the silencing of the genes that are accountable for *p53*-dependent cell cycle arrest and apoptosis, such as *p21*, *p27* and *Puma* can also promote the reprogramming of somatic cells into iPSC (Son et al., 2013, Lake et al., 2012, Zhan et al., 2019).

1.4.6.2 Epigenetic modifiers

The successful generation of hiPSC relys on epigenetic changes, from DNA methylation to histone modifications. Chromatin remodelling is a crucial step in the somatic reprogramming process, and thus researchers have examined the different chemical compounds that could simulate the changes in the epigenome occurring during reprogramming. As mentioned previously, HDAC inhibitors such as VPA and TSA can replace *c-MYC* activity by controlling global histone acetylation and enhancing the endogenous expression of *OCT4*, *NANOG* and *SOX2* in iPSC colonies (Duan et al., 2019).

Recently, Yoo et al., 2020, investigated the role of Porphyra 334 (P334) in promoting the generation of iPSC. P334 is a secondary metabolite found in diverse marine and terrestrial organisms which has several important effects on fibroblast proliferation, wound healing, and antioxidant activity. The use of P334, together with OSKM, significantly improved the generation of iPSC by regulating MET during the process of somatic reprogramming (Yoo et al., 2020).

Furthermore, incomplete DNA de-methylation could lead to low efficiency of reprogramming and trap cells in a pre-iPSC stage. Thus, several studies have used DNA methyltransferase (DNMT) inhibitors such as 5-azacytidine, G9a and CM272 to eliminate the epigenetic barriers on the road toward pluripotency (Rodriguez-Madoz et al., 2017, Mikkelsen et al., 2008).

In addition to HDAC inhibitors and DNMT inhibitors, other researchers have used ascorbic acid (vitamin C) to improve somatic reprogramming. Vitamin C is a co-factor that promotes hypomethylation, boosting the activity of histone demethylases and reducing cellular senescence by lowering the expression of *p53* and *p21* (Bagci and Fisher, 2013, Esteban et al., 2010).

During the last decade, the inhibition of two isoforms of the Rho-associated coiled-coil protein kinase (ROCK1 and ROCK2) has proven to be effective during the somatic reprogramming; enhancing the attachment, proliferation, and pluripotency maintenance of the dissociated iPSC (Croze et al., 2016, Gao et al., 2019). However, the prolonged treatment of the iPSC with ROCK inhibitor (Y-27632) could affect the morphology of hiPSC colony and increases the apoptosis and detachment of cells (Gao et al., 2019).

1.4.6.3 Improvement in the delivery methods

Finding other ways to deliver the Yamanaka factors into somatic cells in an effective and integration-free way has been the purpose of much research. Yamanaka used four different retroviral vectors to deliver each TF into mouse fibroblast because these vectors have the advantage of self-silencing towards the end of reprogramming (Takahashi and Yamanaka, 2006). Nonetheless, they held many downsides; for example, the silencing of the genes often occurs early in the reprogramming process, leading to partially reprogrammed cell lines that keep depending on exogenous factor expression and fail to activate endogenous expression of

pluripotent genes (Takahashi and Yamanaka, 2006, Mikkelsen et al., 2008). Also, the silenced transgenes could be reactivated later in iPSC lines which could interfere with their developmental potential and subsequently lead to tumour formation, as seen in chimeric animals (Takahashi and Yamanaka, 2006, Nakagawa et al., 2008, Okita et al., 2007). But the main limitation of using of a retroviral vector as a delivery system is the integration into the host genome, which leads to an increase in the chances of insertional mutagenesis.

Moreover, the infectivity of a retrovirus is restricted to dividing cells; limiting the range of cell types that can be reprogrammed (Okita et al., 2007, Takahashi and Yamanaka, 2006). Thus, much research has attempted to develop new methods to reduce the potentially harmful effects of reactivation of transgene expression and insertional mutagenesis. The advancements in reprogramming delivery methods can be grouped into two categories: integrative systems and non-integrative systems, the integrative systems being further subdivided into viral vectors (retrovirus and lentivirus) and non-viral vectors (PiggyBac). Similarly, the non-integrative system can be grouped into viral vectors (adenovirus and Sendai virus) and non-viral vectors (Episomal, DNA vectors, mRNA and proteins) (Figure 1.12).

Lentiviral vectors are another type of transgene integration method used to deliver the OSKM into any dividing and non-dividing cells with high expression levels. Unfortunately, like retroviral vectors, lentiviral vectors can integrate into the host genome, leading to a high risk of insertional mutagenesis. Also, their poor silencing once pluripotent iPSCs have been generated causes a differentiation block (Brambrink et al., 2008, Lois et al., 2002). To address these issues, researches have generated inducible lentiviral vectors in which expression of TFs can be controlled by doxycycline. Doxycycline inducible lentiviral vectors have given better monitoring of transgene expression, reducing the risk of inefficient silencing and transgene reactivation (Hockemeyer et al., 2008). In addition, using several lentiviral vectors to drive the

TFs has been replaced with the creation of polycistronic viral vectors in which polycistronic OSKM lentivirus flanked by loxP sites were developed, such as the STEMCCA lentiviral vector (Sommer et al., 2009). The cell of interest is infected with the lentivirus, and iPSC are generated. Then, by the addition of Cre recombinase, the OSKM genes are removed out of the genome. Thus, through the use of STEMCCA vectors, the expression of all OSKM factors will be driven by a single promoter and therefore enhance the reprogramming efficiency (Carey et al., 2009, Sommer et al., 2009). Nonetheless, a small thirteen base pairs insertion, corresponding to one loxP site, will remain within the genome of the cell.

A piggyBac transposon-based delivery system is a non-viral integrated method that has been applied by several studies to generate hiPSC. The PiggyBac system is known as a mobile genetic element (transposon) that in the presence of a transposase, can be integrated into chromosomal TTAA sites and subsequently is removed from the genome by the transposase. This process occurs without any genetic damage, thus avoiding the risk of insertional mutagenesis. However, the PiggyBac transposon delivery system is associated with low reprogramming efficiency (0.02– 0.05 %) compared with viral vectors (Woltjen et al., 2009). On the other hand, PiggyBac is an integrative method with a high risk of integration into the genome leading to insertional mutagenesis. Because the human genome has endogenous piggyback-like transposon elements, the transposase enzyme could be recognised and be mobilised, leading to undesired genomic changes (Brouwer et al., 2016).

Despite increased efficiency, integrative delivery methods are unreliable and unsafe approaches for therapeutic application or disease modelling. Thus, a great effort has been made to address all these major limitations by generating non-integrating methods.

Adenovirus is the first viral, non-integrated vector used to generate iPSC from mouse somatic cells in 2008 (Stadtfeld et al., 2008b) and later on to generate iPSC from human somatic cells

(Zhou and Freed, 2009). The use of adenovirus allows the generation of iPSC without the genomic integration characteristics of retroviruses or lentivirus. However, the reprogramming efficacy of adenovirus is low, only 0.001–0.0001% in mouse (Stadtfeld et al., 2008b) and 0.0002% in human cells (Zhou and Freed, 2009). Thus, Sendai viruses (SeV) are a good alternative for delivering of the Yamanaka factors into somatic cells. SeV is RNA virus that does not integrate into the nucleus, it can produce high amounts of protein, and it is diluted out of cells ~10 passages after the transduction. In addition, it can reprogram several types of cells with different efficiencies, for instance, 0.1% efficiency for blood cells and 1% for fibroblasts (Fusaki et al., 2009, Ban et al., 2011). So far, it is the most efficient and safest vector that has been used in producing hiPSC (Zhou and Zeng, 2013).

The use of Episomal vectors is a non-viral and non-integrating method that provides another alternative to the integration-free viruses. Episomes are extrachromosomal DNAs capable of replicating independently of the chromosomal DNA within a cell. The episomal vectors can drive the OSKM factors directly into somatic cells as a plasmid or minicircle DNA (Okita et al., 2010, Okita et al., 2011). In 2009, Thomson's group was the first laboratory that successfully derived hiPSC using non-integrating episomal vectors. This finding was important for two reasons: first, it overcomes one obstacle for the clinical application of hiPSCs, as the genome of the host cell will remain intact; and second, it demonstrates that somatic cells do not require genomic integration or the continued presence of exogenous reprogramming factors (Yu et al., 2009).

Additionally, a recent study comparing the SeV method with the episomal method in generating hiPSC from MDS patients revealed that episomal is the most effective method in modelling all subclones of MDS patients (Hsu et al., 2019). However, episomal reprogramming remains a concern for clinical application if using the shRNA as a silencer of *p53*. As mentioned

previously, reducing the activity of *p53* can be extremely effective for enhancing cell reprogramming, but at the same time, the defective cells could escape apoptosis and cause teratoma formation (Kawamura et al., 2009, Hong et al., 2009).

Yamanaka's factors can be directly delivered into somatic cells as recombinant proteins, but this method is extremely inefficient, and it also requires repeated transfection to preserve the intracellular protein level for reprogramming (Kim et al., 2009).

Other technological advances in the iPSC field have made possible the generation of hiPSC by using self-replicating RNA (srRNA), which is single-strand RNA with the coding sequences of the OSKM factors, has been used as an alternative RNA method for iPSC generation (Yoshioka et al., 2013). A recent study revealed that srRNA is efficient , less time consuming, faster, and bypasses the need for repeated transfections, making it an ideal method for clinical applications (Steinle et al., 2019).

Lastly, a novel Clustered Regularly Interspaced Short Palindromic Repeats-based gene activation (CRISPRa) has been used as a new reprogramming technique. CRISPRa can directly target the endogenous loci of the reprogramming factors and induce their expression (Weltner et al., 2018).

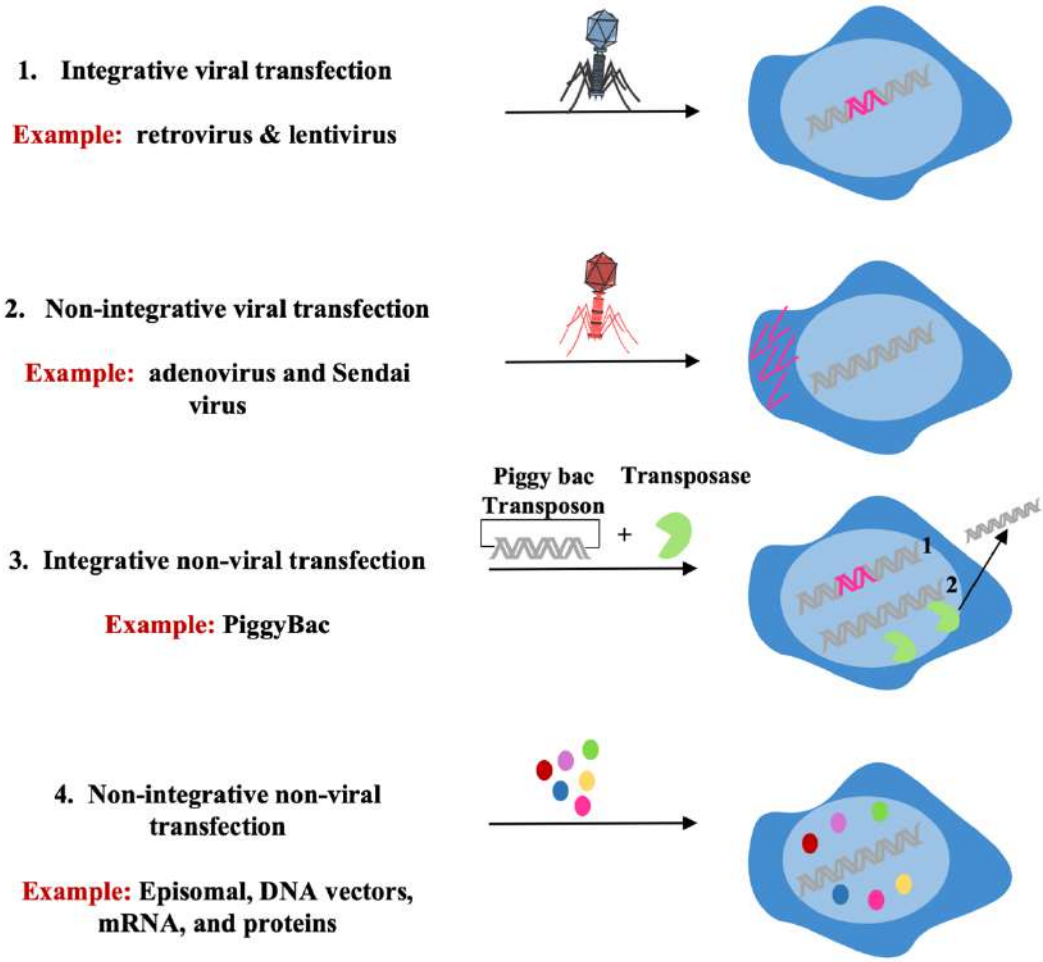


Figure 1.12: Methods for delivering reprogramming factors

Schematic representing the principle of each type of delivery method. Integrative viral methods cause the integration of exogenous DNA into the genome and contribute to teratoma formation. To avoid the integration, non-integrating vectors and non-integrating non-viral vectors are applied to improve the safety and efficacy of iPSCs.

1.4.7 Advantages and applications of human iPSC

The advantages of hiPSCs make them an invaluable tool to model human diseases; hold great therapeutic potential for drug development and drug screening, as well as a promising tool for regenerative medicine. Thus, iPSCs are rapidly becoming utilised in a wide range of medical and scientific fields (Figure 1.13).

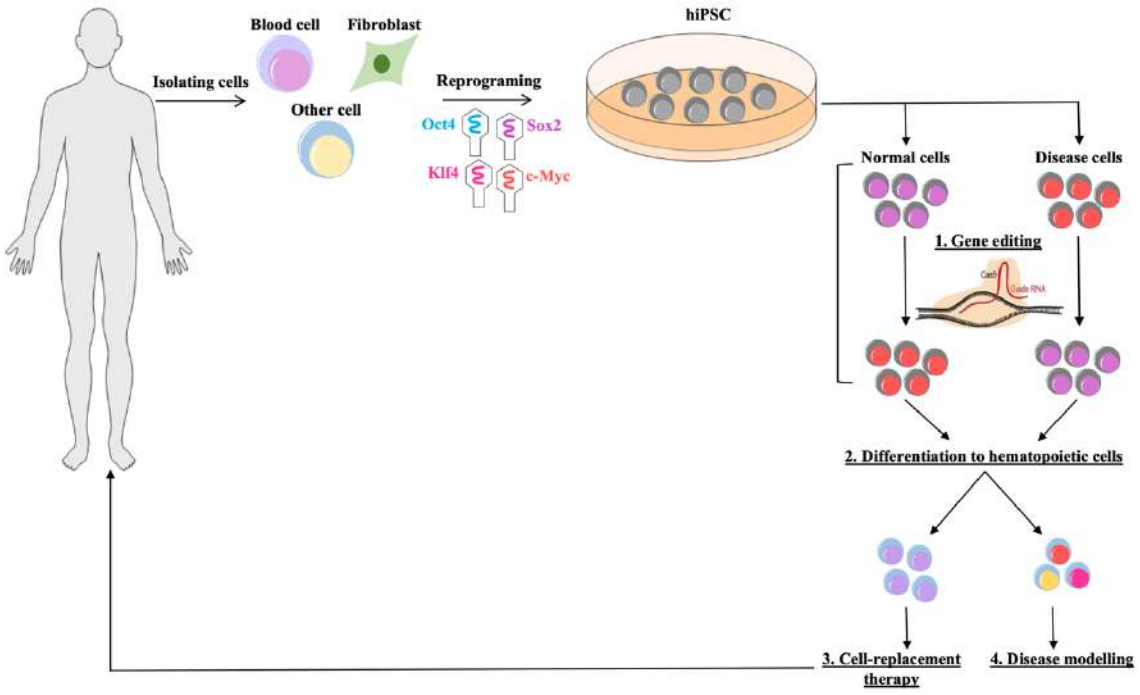


Figure 1.13: Human iPSC applications

Schematic representing the major applications of iPSC. **(1)** and **(2)** Isogenic clones or disease iPSCs can be generated through gene editing to either correct the mutation in disease iPSCs or to introduce it in normal iPSCs. **(3)** differentiation of iPSC into specific cells *in vitro* that can further be used for either cell-replacement therapy or **(4)** disease modelling applications.

One of the most important advantages of iPSC is that they provide an alternative approach to ESC which may eliminate ethical concerns because the establishment of iPSCs does not involve the use of early human embryos (Nakano-Okuno et al., 2014). Similar to ESC, they possess unlimited capacity to proliferate whilst maintaining their stem cell properties. They are also cryopreservable and thus, iPSCs meet any need in cell numbers for cellular, biochemical, molecular, and other downstream applications.

Human iPSCs like ESC, also differentiate to any cell type of the body and thus can be utilized to produce and replace a specific cell type that is not functional. For example, HSCs can be differentiated from iPSC through several protocols, such as the spin-embryoid body-based protocol, co-culture with stromal cell lines and monolayer cultures (Hansen et al., 2018, Bernecker et al., 2019).

Indeed, the generation of iPSC-derived HSC has a vital role in solving all the problems connected with human leukocyte antigen (HLA) compatibility by excluding complications caused by transplantation, such as graft versus host disease (GvHD). Human iPSC could be used to generate blood components to treat several diseases, such as severe anaemia. One study assessed the capacity of twenty-two hiPSC lines to produce the different components of blood. From these, it was found that fourteen hiPSC lines had high activity in hematopoietic CFU, especially for the erythroid lineage, whilst the remaining eight hiPSC lines had limited growth potential and decreased CFU activity (Arora and Daley, 2012).

Interestingly, the use of hiPSC in clinical trials has already started. In 2014 the first iPSCs clinical trial was launched at the RIKEN Institute in Japan to treat a patient with Age-Related Macular Degeneration (AMD). AMD affects the macula of the eye and leads to blurring of the central vision. The first transplantation of retinal pigment epithelium cells derived from iPSC in AMD patient showed an improvement in vision. However, the clinical trial was halted because of mutations generated during the reprogramming process of a second patient (Mandai et al., 2017, Kimbrel and Lanza, 2015). Subsequently, this study resumed in 2017 using HLA-matched allogeneic iPSCs to avoid the rejection or mutagenesis and to date five patients with AMD have been cured with hiPSC (Braganca et al., 2019).

Another major advantage of iPSC is the possibility of modelling many human inherited or somatic diseases such as microcephaly as a consequence of Zika virus (Tang et al., 2016), Alzheimer disease (Kang et al., 2016), type 1 diabetes mellitus (Hosokawa et al., 2017), Down syndrome (Huo et al., 2018), heart disease (Yla-Herttuala, 2018) and blood diseases (Papapetrou, 2019, Georgomanoli and Papapetrou, 2019, Spyrou and Papapetrou, 2020). In fact, hiPSCs provide a powerful tool for research in several fields in which iPSCs have more advantages over animal models. The animal models cannot perfectly mirror the right human

disease phenotype and the iPSC system is less expensive, and saves time compared with conventional animal systems (Omole and Fakoya, 2018).

1.4.8 Use of iPSCs for the study of leukemic and pre-leukemic conditions

Despite the successful generation of iPSC from disease cells that occur as a consequence of a specific point mutation, such as *ASLX1*, *SRSF2, SF3B1* and *EZH2,* iPSC generation from blood disorders and other diseases caused by a sequential accumulation of mutations have been very challenging. In the generation of iPSC for modelling of blood disorders, normal cells from the patient will be inadvertently co-isolated along with the disease cells. Thus, through the reprogramming process, disease iPSC and normal iPSC lines from the same patient will be generated (Ye et al., 2014, Chang et al., 2018, Hsu et al., 2019, Kotini et al., 2015, Kotini et al., 2017). In addition, somatic reprogramming could capture various disease progression stages, as cancers are often associated with the accumulation of mutations or lesions. Thus, somatic reprogramming can help to clarify the clonal composition or the clonal hierarchies of the starting cell population by identifying distinct clonal iPSCs with one, two, or more mutations (Ye et al., 2014, Chang et al., 2018, Hsu et al., 2019, Kotini et al., 2015, Kotini et al., 2017). Yet very few groups have successfully generated iPSC from myeloid leukemic cells Table 1.4.

Table 1.4: Latest publications of successful generation of hiPSC from several haematological diseases

Disease name	Somatic cell type	Genotype	Reprogramming method	Reference
Fanconi anaemia	PBMNCs	FANC genes mutation	SeV supplemented with P53 inhibition	(Qanash et al., 2018)
β-thalassemia	PBMNCs	deletions in the beta globin (*HBB*) gene	SeV	(Ou et al., 2016)
Hb-CS thalassemia	PBMNCs	Haemoglobin H-constant spring **α**-thalassaemia (- -/-α^{CS})	SeV	(Yingjun et al., 2019)
Aplastic anaemia	Fibroblasts	-	SeV	(Melguizo-Sanchis et al., 2018)
Juvenile myelomonocytic leukaemia (JMML)	PBMNCs & MBMNCs	*PTPN11,*E76K *CBL* mutations	SeV	(Tasian et al., 2019)
Chronic myeloid leukaemia (CML)	PBMNCs	t(9;22) (q34;q11)	Episomal	(Miyauchi et al., 2018)
Chronic myelomonocytic leukemia (CMML)	$CD34^+$	EZH2, NRAS & RUNX1 mutations	Episomal	(Taoka et al., 2018)
AML	$CD34^+$ cells	Translocation in 7q & 13q	SeV	(Yamasaki et al., 2019)
AML	PBMNCs	Del7q, KRAs, NRAs mutations	Cre-excisable OSKMS lentiviral vector	(Wesely et al., 2020)

The reprogramming of MDS cells into iPSCs has proven to be much more challenging and only two groups have been able to generate hiPSC from MDS patients (Papapetrou's group and Doulatov's group) (Ye et al., 2014, Chang et al., 2018, Hsu et al., 2019, Kotini et al., 2015, Kotini et al., 2017).

The first iPSCs generated from BM cells and PB cells of MDS patients was described by the group of Eirini Papapetrou in 2015 using a lentiviral vector (integrated method). This group reported the successful derivation of iPSC lines from two MDS patients with *del(7q)* and the normal iPSC lines from the same patients to study the contribution of *del(7q)* in MDS. The *del(7q)* iPSC recapitulated disease-associated phenotypes, including impaired hematopoietic differentiation. The phenotypes of this disease were easily represented in karyotypically normal cells by inducing homozygosity of defined segments, particularly the 20 Mb region (7q 31.3- q 31.6) on the long arm of chromosome 7 that affects normal differentiation. Using a phenotype rescue screening, candidate genes (*HIPK2*, *ATP6VOE2*, *LUC7L2* and *EZH2*) were shown to have a significant role in rescuing hematopoietic defects in *del(7q)* derived iPSCs. This work highlighted the utility of human iPSCs for functional studies aiming to determine key disease-associated genes located in large chromosomal deletions and for the discovery of haploinsufficient genes (Kotini et al., 2015).

In a subsequent study of the same group, several MDS-iPSC lines were generated with *del(7q)* and a mutation in the *SRSF2* gene. Furthermore, a series of normal iPSC lines harbouring *del(7q)*, the SF mutation *SRSF2 P95L*, either genetic lesions or none, were engineered using clustered regularly interspaced short palindromic repeats and CRISPR-associated protein 9 (CRISPR-Cas9). Interestingly, they investigated the specific phenotype and drug response of each genetic lesion of the hiPSC lines (Chang et al., 2018).

In 2017, the same group reported the generation of hiPSC lines from different MDS patients to capture the entire spectrum of myeloid transformation from the pre-leukaemia stage to low-grade MDS, high-grade MDS and sAML. In addition, they identified the specific disease phenotype and the severity of the disease of each hiPSC lines, using this information to construct a phenotypic framework of the clonal evolution of MDS to AML (Kotini et al., 2017).

Thereafter, in 2019 another group successfully generated different hiPSC cell lines from eight different MDS patients using non-integrated methods (SeV and episomal) in order to study the order of mutations contributing to the clonal evolution in MDS. The authors suggested that the *t(4;12)* abnormality is the initiating event of MDS clonality, followed by mutations in *SF3B1*, *EZH2* and *del(5q)* (Hsu et al., 2019).

Several iPSCs derived from MDS patients are being investigated for drug screening studies (Ye et al., 2014, Chang et al., 2018, Hsu et al., 2019, Kotini et al., 2015, Kotini et al., 2017). Interestingly, different clinical drug candidates have been derived from iPSC studies and are currently in clinical trials (Bright et al., 2015, McNeish et al., 2015, Mullard, 2015). For example, Papapetrou's group differentiated hiPSC lines (from pre-leukemic through low-risk MDS, high-risk MDS and sAML) to HSC to determine the effect of azacitidine at each stage of MDS. Interestingly, this study revealed that azacitidine promotes the hematopoietic differentiation in low-risk MDS iPSC lines whilst inducing cell growth inhibition in high-risk MDS iPSCs (Kotini et al., 2017).

In conclusion, as exemplified in the work of Papapetrou's group (Koitini, 2017, Cheg, 2018), the combination of genome editing tools such as CRISPR-Cas9 and iPSC has provided a remarkable advantage on disease modelling.

1.5 Genome editing and CRISPR-Cas9

Genome editing is broadly used in research to study gene function to repair or introduce specific mutations in the gene of interest. This technology is based on RNA sequences or proteins that bind to specific DNA sequences and recruit nucleases to introduce double-strand breaks (DSBs) in the target sequence. This DSB activates the DNA damage response aiming to repair the DNA lesion site. This could occur via homologous recombination (HR) or non-homologous end-

joining (NHEJ) processes. The combination of iPSCs and genome-editing tools has further improved the power of iPSCs in disease modelling and therapeutic potential. In the last decade, several genome editing approaches have been used to genetically target iPSCs such as Zinc Finger Nuclease (ZFNs), transcription activator-like effectors (TALEs) and Transcription activators- like endonuclease (TALENs) and Clustered regularly interspaced short palindromic repeats (CRISPR). These three gene editing methods differ in the type of repair and nucleases involved and are briefly explained below. ZFNs editing tool implies the use of zinc finger containing transcription factors that attach to the endonuclease domain of the bacterial Flavobacterium okeanokoites (FokI) restriction enzyme. Every zinc finger domain can recognise a 3 to 4-bp DNA sequence, and tandem domains can potentially bind to an extended nucleotide sequence for target cleavage (Cathomen and Keith Joung, 2008, Urnov et al., 2010). ZFNs are designed as a pair to cleave a specific site in the genome that can recognise two sequences flanking the site, one on the forward strand and the other on the reverse strand. So, the ZFNs binds on either side of the specific genome site, the pair of FokI domains dimerize and cleave the DNA at the site, generating a DNA DSB with 5′ overhangs (Urnov et al., 2010). Subsequently, the DNA DSB is repaired by either (a) NHEJ, which can take place during any phase of the cell cycle, but occasionally results in incorrect repair, or (b) HR, which typically occurs during late S phase or G2 phase when a sister chromatid is available to serve as a repair template.

TALEs tool is another class of FokI domain which has shed light on new possibilities for precise genome editing (Boch et al., 2009). TALE effector domains recognise single nucleotides, allowing many researchers to target a specific genome sequence of interest. However, the developments in genome editing tools have created a new type of editing site-specific nuclease called TALENs. The engineering of genome sites specifically occurs by fusing TALE domain

repeats to the FokI nuclease domain, which creates a DSB at the desired target site in the genome similar to ZFNs. Despite the successful genome editing achieved with these techniques, such as modifying the endogenous genes in hiPSC (Hockemeyer et al., 2011), these approaches have several potential disadvantages; highly costly, highly time-consuming, and they generate DSBs at off-target sites (Gupta and Musunuru, 2014, Li et al., 2020). More recently, the possibility of using CRISPR-Cas9 was presented as a unified and straightforward platform for genome editing (Gupta and Musunuru, 2014). CRISPRs are genome structures composed of 25-50 bp repeats separated by unique sequence spacers of similar length which are found in many prokaryotes, and were initially noticed by Dr Nakata's group in Escherichia coli in 1987 (Ishino et al., 1987).

In the early 2000s, the CRISPR-Cas system was shown to form part of the adaptive prokaryotic immune system after the discovery by Mojica, who found the majority of the sequences intercalated between the identical repeats were derived from invading phage and plasmid genomes (Mojica et al., 2005). However, this dramatic discovery by Mojica was grossly underappreciated at that time and was independently published in 2005 by three research groups (Mojica et al., 2005, Pourcel et al., 2005, Bolotin et al., 2005).

Since the discovery in 2005, scientists have spent almost a decade improving the system to be able to utilize it in a safe and efficient manner. These advances are summarized in the timeline in Figure 1.14. Finally, in 2013, the CRISPR-Cas9 system (Type II) was considered as the newest genome editing tool for the induction of site-specific DSBs and subsequent mutagenesis in plant, mouse, and human cells was completed with the ultimate goal of using this method in clinical trials (Cong et al., 2013).

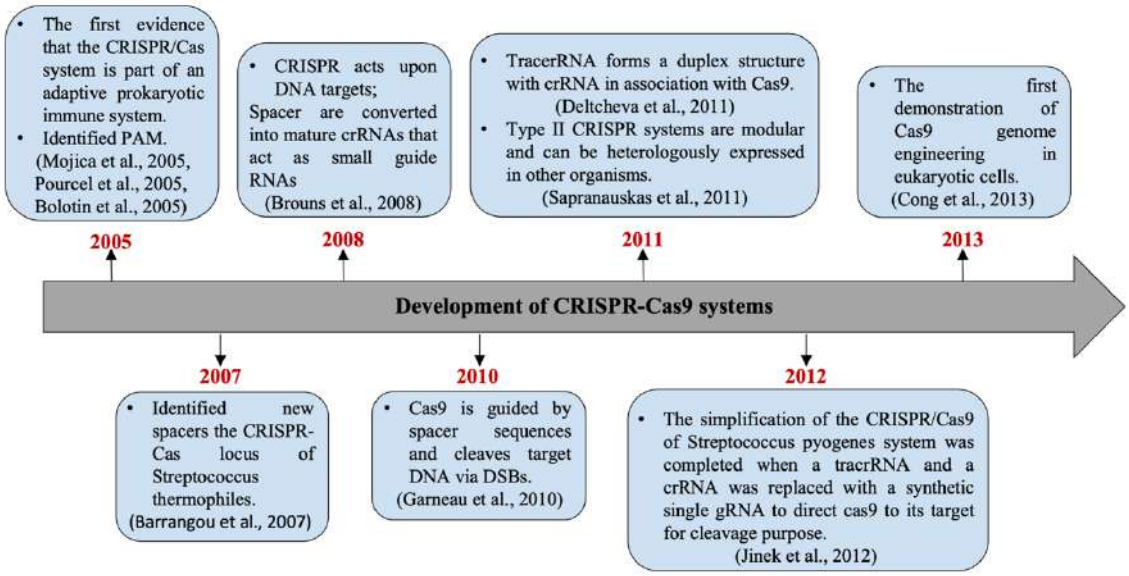


Figure 1.14: Development of CRISPR-Cas9 systems

1.5.1 Principle of CRISPR-Cas9

CRISPR-Cas9 has been utilised as a genomic modification tool in mammalian cells since 2013. Two components are essential in utilizing CRISPR-Cas9 as an editing tool: (i) the Cas9 nuclease, which is used to induce the DSB in the target DNA and (ii) the gRNA, a 20 nucleotide RNA sequence complementary to the DNA sequence to be edited. This gRNA helps to direct the Cas9 to the target region in the DNA. In addition, the protospacer adjacent motif (PAM) sequence (three nucleotides, NGG) is required in this system as only regions of the genomic DNA that contain PAMs can be identified and bound by the gRNA/Cas9 complex (Jinek et al., 2012).

Briefly, the gRNA directs the Cas9 to the target region in the genome where the DSB will occur, triggering the natural DNA repair processes of the cell such as NHEJ or HR in the presence of a donor DNA template (Adli, 2018). Two common mutations can be engineered using CRISPR-Cas9: (i) Frame-shifting or insertion/deletion, which are mediated by NHEJ and is frequently

used to generate gene knockouts and (ii) point mutations which are produced by HR, and is essential in single allele editing (Figure 1.15) (Doudna and Charpentier, 2014). Hence, the simplicity of CRISPR-Cas9 provides a valuable tool to study the effect of different gene expression alterations in various diseases such as cancer or diseases of the blood.

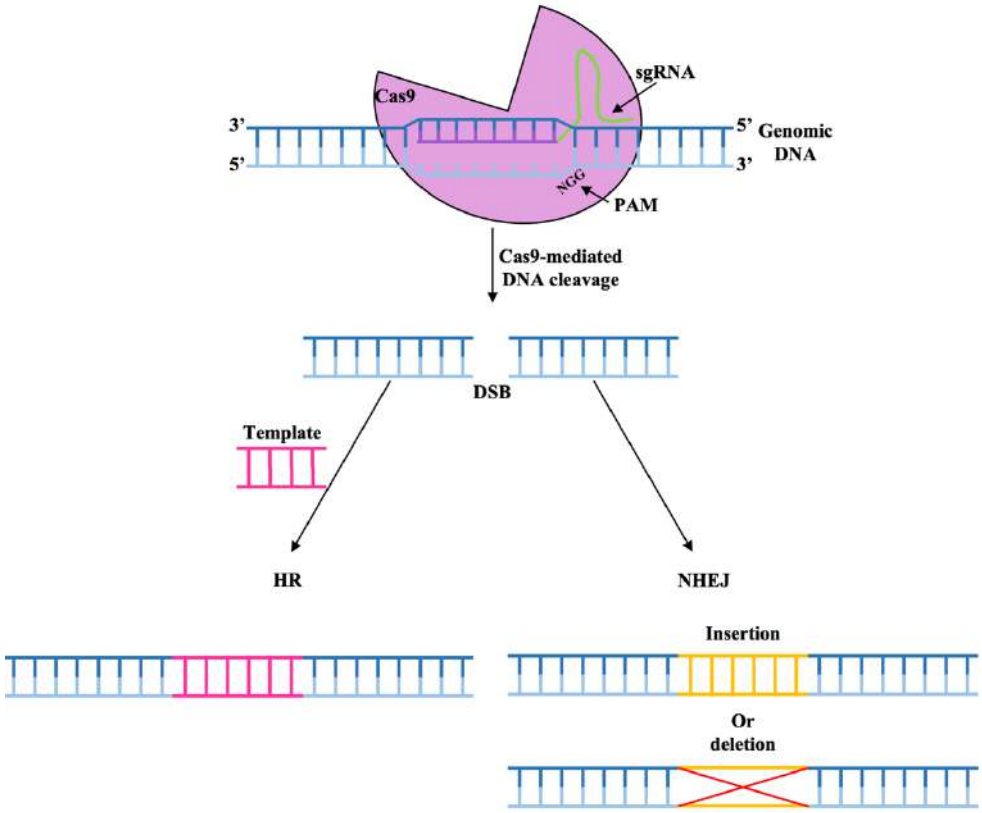


Figure 1.15: CRISPR-Cas9 principle

Schematic representing the complex of Cas9 and gRNA that binds to the target DNA close to the PAM site. Cas9 generates a DSB in the target site that could be repaired via NHEJ or HR. NHEJ repair usually results in insertions or deletions, or in frameshift that leads to knockout of the gene by interruption. HR repair occurs when a DNA donor with homology in the ends is provided, this DNA can be inserted to the target site for gene modification, introducing the nucleotides and leading to frameshifts or insertion of DNA.

1.5.2 Applications of CRISPR-Cas9

As mentioned earlier, the combination of both technologies, iPSC and CRISPR-Cas9, has a great potential for modelling many diseases and for therapeutic approaches, especially in the haematological field. This is exemplified in a study aiming at treating patients with thalassemia. Thalassemia is a haemoglobin disease caused by a mutation in the alpha-globin gene cluster of the haemoglobin (α- thalassemia) or beta-globin (β- thalassemia). Indeed, the generation of iPSC from β- thalassemia patients and correction of the β globin mutation of these cells by CRISPR-Cas9 technology has been reported (Song et al, 2015). The results illustrated that the differentiation of corrected iPSC was adequate and there was an increase in the percentage of HSC. More importantly, the transplantation of corrected iPSC into β- thalassemia patients restored β globin expression and reduced reactive oxygen species production compared with the control cell line (Song et al., 2015). Another achievement reported in the field of hemoglobinopathies is the use of CRISPR-Cas9 to insert a mutation to generate a benign genetic condition as an optional treatment for hereditary persistence of foetal haemoglobin (HPFH). The benign mutations reduce the shift of γ-globin to β-globin, causing a high level of HbF expression throughout life, which can relieve the clinical signs of β-thalassemia or sickle cell diseas (Traxler et al., 2016).

In the case of MDS, CRISPR-Cas9 technology has also been used to study the effect of specific genes in development of the disease phenotype. The study of Kotini et al., in 2015 described the hematopoietic phenotype in an MDS patient with *del(7q),* including blocking of hematopoietic differentiation. Interestingly, the inclusion of the deleted 7q region by CRISPR-Cas9 led to karyotypically normal cells, restoring the differentiation capacity (Kotini et al., 2015). Additionally, the combination of iPSC and CRISPR-Cas9 technology has also been useful to determine the contribution of the $SRSF2^{P95L}$ mutation to the MDS phenotype. By

engineering iPSC harbouring the *SRSF2*^{P95L} mutation and subsequent differentiation to the hematopoietic lineage, this work revealed that *SRSF2*^{P95L} does not affect the quantity of CD45+ cells (hematopoietic cells), but the quality of the cells (dysplastic features), indicating that *SRSF2* ^{P95L} mutation is important in driving overt dysplasia (Chang et al., 2018). Taken together, these studies highlight the power of the combination of CRISPR-Cas9 and iPSC technologies in modelling several diseases aiming to improve therapeutic targeting or to identify biomarkers of disease progression.

1.5.3 Limitations and challenges of CRISPR-Cas9

Despite the wide applications of CRISPR-Cas9 as a genome editing tool, it still faces many limitations that motivate many researchers to propose additional strategies to improve this system further. CRISPR-Cas9 faces the main challenge common with all gene-editing techniques, which is that they can miss their target region (off-targets). It has been reported that the Cas9 nuclease potentially has diverse tolerance to the mismatch between gRNA and targeted genomic DNA contributing to off-target effects; PAM mismatch could be another source of off-target effects (Fu et al., 2013). Therefore, tremendous efforts have been made by researchers to understand the incidence of off-target caused by CRISPR-Cas9.

Interestingly, whole-genome sequencing has identified that off-target sites mainly occur in mouse ESCs but at a very low frequency in human stem cells (Smith et al., 2014, Suzuki et al., 2014). However, several steps have been considered to improve the targeting specificity of the CRISPR-Cas9 tool; Firstly, well-designed gRNA is a very crucial step in CRISPR-Cas9, which helps to ensure high efficiency of genome cleavage and low incidence of off-targets. Several online tools are available for the proper designing of gRNA which also evaluate the incidence of off targets (Kim et al., 2016, Cradick et al., 2014, Montague et al., 2014, Naito et al., 2015).

Secondly, re-engineering the structure of Cas9 nuclease could help to improve the targeting specificity of CRISPR-Cas9; for instance, the use of mutant versions of Cas9 proteins that are able to introduce just a nick on the gRNA-bound strand without breaking the whole double-strand DNA (Ran et al., 2013). Among these mutant Cas9 proteins are Cas9 nickase (Cas9n), Streptococcus pyogenes Cas9 (SpCas9), SpCas9-high fidelity-1 (Kleinstiver et al., 2016), eSpCas9 1.1 (Slaymaker et al., 2016) and hyper-accurate Cas9 (Chen et al., 2017); these mutant Cas9 proteins are highly efficient and reduce the off-target frequency. Lastly, other studies have reported that manipulating the length of the gRNA sequence (decreasing or increasing) could increase the targeting specificity without sacrificing on-target cleavage efficiencies. For example, it was reported that truncated gRNA (less than 20 nucleotides) weakens the gRNA-DNA duplex stability at the off-target sites and thus significantly reduced the off-targets by 5000-fold (Fu et al., 2014). Additionally, it was shown that designing a hairpin structure to the 5′-end of the gRNA spacer region significantly reduced off-target incidence by an average of 55-fold (Kocak et al., 2019).

Editing efficiency is another challenging aspect of the gene-editing technique. CRISPR-Cas9 has been vastly applied in various cells and it has shown a high editing efficiency, especially in mouse cells. Nonetheless, it has been reported that the efficiency of CRISPR-Cas9 is really low in human stem cells such as HSC and MSC (Dang et al., 2015). To address this issue, several modifications have been applied to CRISPR-Cas9 to improve its clinical application. RNAs are generally sensitive and have a higher chance to be degraded within the human cells. Thus, the researchers focus on maintaining gRNA for a longer time prior to degradation by using chemical compositions (Houseley and Tollervey, 2009, Dang et al., 2015). For example, Hendel et al., 2015 have tried different chemical components to maintain the gRNA during the engineering of human primary T cells and $CD34^+$ cells. They have synthesised gRNA using 2′

O-thionocarbamate-protected nucleoside phosphonamidites and tried three different chemical modifications to evaluate their effects on efficacy (comprising 2′-O-methyl (M), 2′-O-methyl 3′phosphorothioate (MS), or 2′-O-methyl 3′thioPACE (MSP)) by incorporating at three terminal nucleotides for both ends 5' and 3' (Hendel et al., 2015). In addition, enhancing the efficiency of DSB repair methods leads to improved efficiency of genome-editing. Several chemical reagents have been identified to increase HR efficiency, such as SCR7 (Maruyama et al., 2015), RS-1 (Song et al., 2016), KU0060648 and NU7441 (Robert et al., 2015). Additionally, the use of a donor template in the form of ssDNA or gRNA-donor DNA fusion also enhances HR efficiency (Richardson et al., 2016, Lee et al., 2017b).

The efficient delivery of the CRISPR-Cas9 system into the cells has also limited the use of CRISPR-Cas9 technology. Depending on the delivery method, the genome editing system components can take various forms such as plasmids, RNAs, or protein. Most genome editing tools relay on viral vectors such as adeno-associated viruses (AAVs). However, these vectors cannot efficiently deliver gene sizes over ~4.5 kb. Validated Cas9 proteins are too large consisting of 1368 amino acids, and thus it is challenging to pack together with the gRNA-coding DNA into a single AAV vector (Moon et al., 2019). This problem has been addressed by lightweight Cas orthologs from archaea and bacteria such as NmCas9 (1082 aa), SaCas9 (1053 aa), CjCas9 (984 aa) and - ScCas9 (Ibraheim et al., 2018, Li et al., 2019, Kim et al., 2017). In addition, another study has successfully delivered the large SpCas9 into cells by splitting it into N-Cas9 (2–573 aa) and C-Cas9 (574–1368 aa) (Truong et al., 2015). Furthermore, a lentiviral vector has been efficiently utilized to deliver the Cas9 and gRNA into stem cells with low cellular toxicity (Toledo et al., 2015, Wang et al., 2014, Chen et al., 2013).

Finally, it is worth mentioning that the immunogenicity of CRISPR-Cas9 remains another limitation. Bacteria are the most common source of Cas9, such as *Staphylococcus aureus* (*S.*

aureus; SaCas9) and *Streptococcus pyogenes* (*S. pyogenes*; SpCas9); these bacteria are common human commensals that can also be pathogenic which lead to increased chance of inducing the immune response against those proteins. Indeed, pre-existing immunoglobulin G (IgG) antibodies against SaCas9 have been detected in the human blood samples from 79% of participants and 65% for antibodies against SpCas9 (Charlesworth et al., 2019). However, in cases of using Cas9 technology for *in vivo* gene therapy, the Cas9 immune response can potentially be overcome by using transient Cas9 expression and clearance of Cas9 protein before transplantation of the mutation-corrected cells back to patients (Crudele and Chamberlain, 2018).

1.6 Aim

MDS is a heterogeneous group of haematological diseases characterised by cytopenia and ineffective haematopoiesis. The occurrence of MDS is mainly in people over 65 years old, and 30% of these patients eventually progress to a more aggressive stage of AML. Patients with high-risk MDS may be treated with DNA demethylating agents such as azacytidine, but 50% to 60% of these patients do not respond to treatment. The reasons behind the treatment failure and progression to AML are not thoroughly investigated due to the insufficient number of cells that can be obtained from MDS patients and the lack of animal model systems. Therefore, there is an urgent need to develop new model systems that provide a sufficient number of cells for a better understanding of molecular mechanisms of these blood disorders with the aim to improve the diagnostic and therapeutic strategies of MDS.

The overall aim of this project is to develop an *in vitro* model system in which to study the molecular mechanism leading to MDS progression.

The aim will be met by achieving the following objectives:

1. Generation of hiPSC from PBMNCs from a low-risk MDS patient using non-integrated methods (SeV, Episomal).
2. Characterisation of hiPSC derived from MDS to determine the pluripotency by checking AP activity, pluripotent markers, and ability to differentiate to the three germline layers.
3. Re-differentiation of hiPSC back to hematopoietic cells to validate the disease phenotype by studying the proliferation, differentiating potential, and performing transcriptome studies to identify deregulated genes.
4. Determine the role of *C/EBPα* in a genetic background of *RUNX1*, *ASLX1* and *SRSF2* co-mutations in diseases progression to high-risk MDS by introducing a mutation in the *C/EBPα* gene using CRISPR-Cas9.

Chapter 2 : Materials ans methods

2.1 Somatic reprogramming

2.1.1 Clinical history of patient MDS27

MDS patient samples were obtained through Dr. Manoj Raghavan (Consultant Haematologist) via the Centre for Clinical Haematology at Queen Elizabeth hospital, Birmingham, UK. Ethical approval: **10/H1206/58**. MDS cells were used to generate hiPSC and were taken from patient MDS27. The patient MDS27 was a 65 year old male diagnosed with multi-lineage dysplasia and RA (low-risk MDS) in July 2013. Then, in April 2014 the patient progressed to RAEB2 (EB-2 under new revised nomenclature) which is the worse stage of MDS. In October 2014, the patient completed the first course of the Daunorubicin (DA) chemotherapy and showed slight haematological response, kidney injury, and sepsis. At this stage the patient was not suitable for further intensive treatment. Consequently, the patient relapsed, the amount of blast cells increased to 9-17% in March 2015 and a mutation in *C/EBPα* was detected. He started monthly azacitidine treatment in April 2015, but the blast cells increased to 15-30% with low haematological responded to treatment. In October 2016, the patient did not respond to chemotherapy, and he progressed to sAML dying soon after. The peripheral blood samples collected from MDS27 patient in 2013 and 2015 were used to generate hiPSC.

2.1.2 Patient samples preparation

Peripheral blood mononuclear cells (PBMNCs) were isolated from whole blood by standard Ficoll Paque™ PLUS (GE life Science, 17144003) density gradient centrifugation. Briefly, approximately 10 mL of heparinized, plasma-reduced blood was diluted with 10 ml RPMI medium, 1:1 dilution (Sigma, R8758). Then, 15 mL of ficoll was covered with a layer of 20 mL diluted blood (3:4 dilution). After 30 minutes of centrifugation (300g, room temperature, accel 3, break 2), the PBMNCs were collected and transferred to a new 15 ml tube. Then, the

PBMNCs were washed twice with RPMI medium and 10% (v/v) foetal bovine serum (FBS, Sigma, F7524). After two washing steps, the cells were counted and frozen at -80ºC in freezing medium containing 10% (v/v) dimethylsulfoxide (DMSO, Sigma, D2650) and 90% (v/v) FBS.

2.1.3 Human iPSCs reprogramming using the integration-Free Sendai virus Vector

Human iPSCs from MDS27 were generated from the PB sample collected in 2013 using media described in Table 2.1. MDS27 mononuclear cells were cultured in 1ml expansion medium in a 24-well plate (corning life science, 3524) for 3 days. Then, at day 3 the cells were centrifuged at 300g for 10 minutes and re-suspended in 1 ml of expansion medium and placed in a new well of a 24-well plate for 3 days. At day 6, the cells were counted and $25x10^4$ cells were resuspended in 300 µL of fresh expansion medium with 4µg/ml Polybrene (Sigma, TR-1003) and placed in a 15ml falcon tube for transduction, MOI 3 of Sendai Virus containing the four Yamanaka factors (generous gift from our collaborator Dr Pablo Menendez, Josep Carreras Leukaemia Research Institute, Barcelona, Spain) was added to the cells for three hours at 37ºC. Cells were washed with Phosphate buffered saline (PBS), cultured in 1ml of expansion medium in a 24-well plate and incubated at 37ºC for two days. Subsequently, at day 8, the transduced cells were centrifuged at 300g for 8 minutes and plated in a 6-well plate containing 90%-confluent of mitomycin C-treated MEFs and "day 8 medium" for three days at 37ºC, and every two days the medium was changed. At day 11, the medium was replaced with "day 11 medium" and incubated for four days at 37ºC and every two days the medium was changed with "day 11 medium". At day 16, the medium was changed to hiPSC medium and every day thereafter, the medium was changed to fresh "hiPSC medium". At day 25, emerging iPSC colonies were picked individually into a 6-well plate (corning life science, 3506) coated with Matrigel hESC-qualified Matrix (corning life science, 354277). The colonies were cultured in hiPSC medium

with 10 μM of Rock inhibitor (Y-27632, LKT laboratory, Y1000) and incubated at 37°C. From the first day after picking onwards, the colonies were cultured continuously using hiPSC medium without Rock inhibitor. The medium was changed daily until the cells could be stably passaged and ready for collection for subsequent identification. It is worth mentioning that the same protocol was used twice to generate hiPSC from the same patient after disease progression (MDS27 sample 2015 and 2016) but it failed to generate stable colonies. Furthermore, some optimisation was done to this protocol using 0.5 mM VPA (Sigma, PHR1061) to improve the efficiency of the reprogramming, as in (Kotini et al., 2015). From the 4^{th} day post transduction, 0.5 mM VPA was supplemented into the media for 1 week with a change of media every two days. Unfortunately, after VPA treatment, the cells from MDS27 2015 were not capable of generating any colonies.

Table 2.1: Sendai virus reprogramming media

Medium name	Component	Concentration	Company	Cat. number
Expansion medium	StemSpan H3000	Base	Stem cell technology	09850
	Vitamin C (VC)	50 µg/ml	Sigma	A92902
	SCF	50 ng/ml	R&D	255-SC-010
	IL-3	10 ng/ml	R&D	203-IL-010
	EPO	2 U/ml	Peprotech	100-64
	IGF-1	40 ng/ml	Peprotech	100-12
	Primocin	100 µg/ml	Invivogen	ant-pm-2
	Dexamethasone	1 µM/ml	Sigma	D2915
Day 8 medium	Iscove's Modified Dulbecco's Medium (IMDM)	Base	Thermofisher	12440053
	Heat-Inactivated FBS	10%	Sigma	12106C
	Non-essential aa	1%	Gibco	11140035
	2-mercaptoethanol	100 µM/ml	Sigma	M7522
	L-glutamine	1%	Gibco	25030024
	Dexamethasone	1 µM/ml	Sigma	D2915
	Primocin	100 µg/ml	Invivogen	ant-pm-2
	SCF	50 ng/ml	Peprotech	255-SC-010
	IL-3	10 ng/ml	Peprotech	203-IL-010
	EPO	2 U/ml	Peprotech	100-64
	IGF-1	40 ng/ml	Peprotech	100-12
	Basic fibroblast growth factor (bFGF)	10 ng/ml	Peprotech	100-18
	VC	50 µg/ml	Sigma	A92902
Day 11 medium	IMDM	Base	Sigma	12440053
	HIFBS	10%	Sigma	12106C
	bFGF	10 ng/ml	Peprotech	100-18
	VC	50 µg/ml	Sigma	A92902
Human iPSC medium	Knock out Dulbecco's Modified Eagle (DMEM)/ F12	Base	Gibco	10829018
	Knock out serum	20%	Gibco	10828028
	L-Glutamine	1x	Gibco	25030024
	Non-essential amino acid	1x	Gibco	11140035
	2-mercaptoethanol	100µM/ml	Sigma	M7522
	Primocin	100µg/ml	Invivogen	ant-pm-2
	bFGF	10ng/ml	Peprotech	100-18

2.1.4 Human iPSCs generation by integration- Free episomal reprogramming

The episomal reprogramming of MDS27 PBMNCs from the sample obtained in 2013 was performed as previously described (Hsu et al., 2019) using the episomal reprogramming plasmids pCXLE-OCT4+shP53 (27078), pCXLE-hOCT3/4 (27076), pCXLE-hSK (27078), pCXLE-hUL (27080) and pCXWB-EBNA1 (37624) obtained from Addgene. Briefly, PBMNCs were cultured in one well of a 24-well plate with 1ml of "PBMNCs expansion medium" for 3 days (Table 2.2). After 3 days, cells were collected by centrifugation and counted, $3x10^5$ were used for nucleofection. Two nucleofection reactions were performed +/- pCXLE-OCT4+shP53. Plasmids were used at 3 µg per reaction (<5% DNA volume). Condition one: 0.83 µg of pCXLE-OCT4+shP53 + 0.83 µg of pCXLE-hSK + 0.83 µg of pCXLE-hUL+ 0.5 µg of pCXWB-EBNA1. Condition two: 0.83 µg of pCXLE-hOCT3/4 + 0.83 µg of pCXLE-hSK + 0.83 µg of pCXLE-hUL+ 0.5 µg of pCXWB-EBNA1. The nucleofection was performed using the CD34+ cell nucleofector kit (Lonza, VPA-1003), program U-008, for the nucleofector 11/2b device (Lonza). Following nucleofection, the cells were cultured in one well of a 24-well plate with 1ml of "PBMNCs expansion medium" for 3 days. On day 3 after nucleofection, cells were plated onto 90% mitomycin C-treated MEFs in 6-well plate with 2ml of "PBMNCs expansion medium". On day 5 after nucleofection, an equal volume of hESC medium was added to the cells. On day 7 after nucleofection, the medium was replaced with hESC medium + 10 µM Y27632, and medium changed daily on day 8 onwards. On day 17 after nucleofection, the appearance of hiPSC colonies was apparent and the plate was maintained until day 32 to allow colony formation to occur. Individual colonies were picked and passaged for one passage into a single well of a 6-well plate coated with mitomycin C-treated MEFs and maintained in hESC medium without Y27632. Then, each hiPSC line was maintained on matrigel in StemFlex medium to use them for further experiments.

Table 2.2: Episomal reprogramming media

Medium	Component	Final concentration	Company	Cat. number
PBMNCs expansion medium	a-mem	Base	Merch	M8042
	HIFBS	10%	Sigma	12106C
	IL-3	10 ng/ml	Peprotech	200-03
	IL-6	10 ng/ml	Peprotech	200-06
	G-CSF	10 ng/ml	Peprotech	300-23
	GM-CSF	10 ng/ml	Peprotech	300-03
Human ESC medium	DMEM/F12	Base	Gibco	11320033
	Knock out Serum	20%	Gibco	10828028
	L-Glutamine	1 mM	Gibco	25030024
	Non-essential amino acids	1 mM	Gibco	11140035
	2-mercaptoethanol	0.1M	Sigma	M7522
	bFGF	10 ng/ml	Peprotech	100-18

2.2 Human iPSCs culture methods

Human iPSC BU3.10 was kindly provided by Dr. George Murphy from Boston University, Centre of Regenerative Medicine, Boston, USA. MDS-iPSC clones were generated from patient MDS27 as described in 2.1.3. Human iPSC lines were maintained in a 6-well plate coated with matrigel under a serum-free and feeder-independent protocol, according to the following steps:

2.2.1 Matrigel

Before aliquoting matrigel, everything was kept at -20°C to avoid clotting. To aliquot Corning matrige hESC-qualified Matrix, the vial was submerged in ice in a 4°C refrigerator overnight to thaw. Once thawed, the vial was swirled to ensure that matrigel was evenly mixed before aliquoting it into pre-chilled, sterile Eppendorf tubes using pre-chilled pipette tips. The aliquoting procedure was performed in a sterile biosafety cabinet, with changing of pipette tips every 5-7 tubes to prevent clotting of matrigel. The volume of aliquot was prepared according

to the dilution factor provided on the certificate of analysis. The dilution factor was calculated based on the protein concentration of each lot (0.5 mg/6-well plate and 1.0 mg/2x6-well plates).

To thaw an aliquot of matrigel and coat a plate with it, a 15 ml sterile conical tube containing 12.5 ml of cold DMEM/F-12 medium (Gibco, 11540446) was used to dilute one 1.0 mg matrigel aliquot taken directly from the -20°C freezer. The diluted mixture was pipetted gently up and down to dissolve the matrigel completely. Then, 1 ml of diluted matrigel was transferred to 1 well of a 6-well plate. This dilution was enough to cover two 6-well plates. Plates were allowed to set for 1hour at room temperature before use. Then, the solution was aspirated immediately and washed with Phosphate buffered saline (PBS, Sigma, D8537) before cells were passaged into the plate. If the plates were not used immediately, an additional 1 ml of hiPSC medium was added to each well before storing to prevent them from drying out. Plates were always wrapped in parafilm and stored at 2-8°C for 7-10 days or in incubator at 37°C for 5 days.

2.2.2 Thawing human iPSC on matrigel and passaging

A cryovial of iPSCs was carefully removed from the liquid nitrogen tank and thawed slowly at room temperature until only small ice crystals could be seen. Then, cells were gently pipetted into a sterile 15 ml conical tube containing StemFlex medium (Gibco, A3349401) supplemented with 100 µg/ml of primocin (Invivogen, ant-pm-2). Following, cells were pelleted by centrifugation at 150g for 5 minutes and resuspended very gently to avoid making single cells in a fresh 1 ml StemFlex medium. The cell solution was transferred to one coated well of a 6-well plate and the plate was gently shaken and left in incubator at 37°C, 5% CO_2 for 48 hours. Then, medium was changed daily, and the confluent cells were passaged at a 1:6 or 1:8 ratio every 6 to 7 days when the cells reached 80-90% confluency.

Cells were passaged using mechanical passaging techniques StemPro EZPassage (Invitrogen, 23181010) as described by the manufacturer's instruction, by cutting the colonies into small squares of around 100 μm in size.

2.2.3 Cryopreservation of iPSCs

Cells from an 80% confluent well of 6-well plate were harvested with StemPro EZPassage and centrifuged at 150g for 5 minutes. Cell chunks were gently resuspended in 1 ml of PSC Cryopreservation medium (Gibco, A2644601) and transferred into one pre-labelled cryovial. The cryovial was placed in Mr. Frosty and kept in -80°C freezer overnight. For long-term storage, cryovials were transferred to liquid Nitrogen.

2.3 Alkaline phosphatase staining (AP staining)

Human iPSCs were grown in 6-well plate until 70% confluent with the majority of colonies being large, compact, and exhibiting multi-layering at the centre. Then, the cells were fixed in 1ml/ well of 10% (v/v) neutral formalin buffer for 15 minutes on ice (Table 2.3). After the fixative was removed, the cells were rinsed once with cold distilled water (dH_2O), and then left in dH_2O for 15 minutes. The dH_2O was removed and the cells were stained with 1ml/well of AP staining for 30 minutes at room temperature. The staining solution was removed, and the cells were washed gently with dH_2O and then allowed to air dry. Red positively stained colonies were captured using a primo vert microscope (ZEISS) and Canon EOS 600D camera (Canon) to assess the pluripotency of generated hiPSC.

Table 2.3: Alkaline phosphatase staining buffers

Buffer	Contents
10% neutral formalin buffer	0.11M Na_2HPO_4 (VWR,102494C) 25.6mM of $NaH_2PO_4.H2O$ (VWR,125330) 4% of PFA (Thermo Scientific, 28906)
AP staining solution	0.1mM Naphthol AS-MX Phosphate (Sigma, 70485) 27.3mM N, N-Dimethylformamide (DMF, Sigma, 227056) 0.2M pH 8.3 of Tris-HCL (Sigma,10812846001) 0.7mM fast red violet LB salt (Sigma, F3381)

2.4 Cytogenetics

The cytogenetics protocol was kindly provided by Dr. Manar Samman and Mr. Khelad AlSaidi from King Fahd medical city, Riyadh, Saudi Arabia. Briefly, the day before harvesting metaphases, the cells were split 1:1 and plated in a 6-well plate coated with matrigel with 2 ml of stem flex medium for 24 hours in a 37 °C, 5% CO_2 incubator. Next day, 0.02 µg/ml of KaryoMAX Colcemid Solution (Thermofisher, cat no. 15212012) was added directly to the cells and placed at 37 °C for an additional 2 hours to allow cell cycle arrest at metaphase. Then, the cells were pelleted at 300g for 8 minutes and swelled by pre-warmed hypotonic lysis solution (0.075 M potassium chloride (KCl) for 30 minutes at 37°C. Cells were fixed with cold Carnoy's solution (3:1 (v/v) methanol/glacial acetic acid) at room temperature for 30 minutes, then washed twice in cold Carnoy's fixative.

Metaphase slides were made by resuspending the fixed cells in cold Carnoy's solution and dropping cells onto humidified, chilled glass slides. The mitotic index, quality of metaphase spread, presence of cytoplasm and overlaps were evaluated under a phase contrast microscope. Then, the slides were aged overnight on a heating block at 60°C.

Chromosomes during metaphase spreading were stained by Giemsa banding (G- banding) technique as follows; slides were immersed in 2.5 % (v/v) trypsin working solution at 37°C to digest the chromosomes, then 0.9 % (w/v) Sodium chloride (NaCl) solution at 37°C to stop the digestion. Finally, they were stained in Giemsa karyoMax (Gibco, 10092013) at 37°C and washed in H_2O at room temperature.

Chromosomes were visualized and counted under a light microscope (Leica DM6000, Leica Microsystems) at 100x using an oil immersion lens. Typically, 30 mitotic cells per cell line from three independent experiments were randomly selected and chromosomes were counted and imaged.

2.5 Immunofluorescence for pluripotent markers

2.5.1 Immunofluorescence for cytoplasmic marker TRA1-81

Several days prior to staining, the cells were passaged into 35-mm plates (Corning life science, 430165) containing a glass coverslip (18 mm X 18 mm, Thermo Scientific, 11798681) coated with matrigel and each plate receiving around 8 to 15 iPSC fragmented clones. For immunofluorescence staining, the cells were washed two times with StemFlex medium and then incubated with TRA1-81 mouse primary Ab (Table 2.4) for one hour at 37ºC. After three washes with StemFlex medium, the secondary Ab goat anti-mouse-IgM-Alexa, Fluor 488 (Table 2.5) was added to the cells and incubated for one hour at 37ºC. Then, the cells were washed with PBS cells and fixed with 2% (v/v) Formaldehyde methanol free (PFA, Thermo Scientific, 28906) for 10 minutes at room temperature. Finally, the cells were washed twice, mounted, observed, and imaged under a fluorescence microscope (Leica DM6000, Leica Microsystems).

2.5.2 Immunofluorescence for the nuclear markers NANOG and SOX2

Cells were fixed with 4% PFA (v/v) in PBS for 20 min at room temperature. Then, the cells were washed twice with PBS, and aldehyde groups quenched for 10 minutes with 50mM NH_4Cl (Sigma, 254134) in PBS at room temperature. Afterwards, the cells were permeabilized with 0.5% (v/v) Triton X-100 (Sigma, 11332481001) in PBS for 15 minutes at room temperature and blocked with PBS containing 1% (w/v) bovine serum albumin (BSA, Sigma, A1933) + 0.3% (v/v) Triton X-100 + 10% (v/v) FBS+ 1% (v/v) goat serum (Sigma, G9023) for 1hour at room temperature. Then, the cells were incubated with diluted primary mouse antibodies SOX2 and NANOG in blocking buffer for 1hour at room temperature (Table 2.4). After the primary antibody, the cells were washed for 30 minutes with PBS + 0.1% (v/v) Tween-20 (Sigma, P9416), and goat anti-mouse secondary antibody (Table 2.5) was applied to the cells for 1hour at room temperature. Finally, the cells were washed with PBS, mounted, observed, and imaged under a fluorescence microscope (Leica DM6000, Leica Microsystems).

Table 2.4: Primary antibodies used in immunofluorescence

Antibody	Immunoglobulin	Stock concentration (µg/ml)	Dilution Used	Clone	Supplier	Cat. number
TRA1-81	Mouse IgM	100	1:50	C1.261	Invitrogen	MA1-024
SOX2	Mouse IgG	100	1:100	245610	R&D	AF2018
NANOG	Goat IgG	100	1:100	ABZ92376	R&D	AF1997
OTX2	Goat IgG	100	1:100	246826	R&D	AF1979
SOX17	Goat IgG	100	1:100	245013	R&D	AF1924
Brachyury	Goat IgG	100	1:100	1161B	R&D	AF2085

Table 2.5: Secondary antibodies used in immunofluorescence

Antibody	Immunogen	Dilution Used	Supplier	Cat. number
Goat anti- mouse Alexa Flour 488	Mouse IgM	1:50	Invitrogen	A10684
Goat anti- mouse Alexa Fluor 633	Mouse IgG	1:200	Invitrogen	A21052
Donkey anti-goat Alexa Fluor 488	Goat IgG	1:200	Invitrogen	A32814

2.6 Human iPSC differentiation

2.6.1 Trilineage differentiation of human iPSC

Differentiation towards the three germ layers was assessed using the STEMdiff Trilineage differentiation kit (Stem cell technology, 05230) performed as described by the manufacturer. Briefly, 70% confluent hiPSC were detached from the plates into single cell suspension using the gentle cell dissociation reagent (Stem cell technology, 07174). 400,000 cells/well were seeded into matrigel-coated 12-well plates (Corning life science, 3513) to assess the differentiation to both the ectoderm and endoderm conditions and 100,000 cells/well to assess the mesoderm differentiation. At day 0, the cells to assess for ectoderm differentiation were treated with STEMdiff Trilineage Ecotderm medium + 10 µM of Y-27632 and the cells to assess for endoderm and mesoderm differentiation were treated with StemFlex + 10 µM of Y-27632 and incubated at 37^{o}C. On day 1, each type of cells was fed with the appropriate differentiation media, media refresh daily and cultures maintained in differentiation medium for one week. The capacity to differentiate into three germ layers was evaluated by immunofluorescence as described before in section 2.5.2 using specific antibodies (Table 2.4 and Table 2.5) for Ectoderm (OTX2), Endoderm (SOX17) and Mesoderm (Brachyury).

2.6.2 Hematopoietic differentiation using Stemdiff protocol

Stemdiff protocol (Stem cell technology, 05310) was followed to differentiate hiPSCs to hematopoietic progenitor cells as described by the manufacturer. To start the differentiation, between 16-20 (100-200 µm) colony fragments were plated into one well of a matrigel-coated 12-well plate for 24 hours at 37ºC. Next day, the plate was cleaned to either remove the fragments of colonies bigger than 200 µm or to remove the excess number of fragments if more than 20 fragmented clones were attached. Then, the cells were cultured with 1.5 ml/well of Medium A, with half of medium being changed at day 2. At day 3, the medium was replaced with 1.5 ml of medium B for 11 days. Every other day, 500 µL of medium was refreshed. At different time points (Day 7, 10, 12 and 14), the number of cells was counted, and the presence of hematopoietic progenitor cells assessed by flow cytometry as described in section (2.7) and by colony forming unit assay (2.6.3).

2.6.3 Clonogenic progenitor assay

Hematopoietic clonogenic assays were performed using the MethoCult protocol (Stem cell technology). MethoCult H4435 semisolid media (Stem cell technology, 04435) was thawed at 2°C to 8°C overnight, supplemented with 250 µL Penicillin-Streptomycin (Gibco, 15140122), contents mixed vigorously by shaking and media left to settle for 0.5-1 hour either at room temperature or at 4°C. Then, the medium was aliquoted and frozen as complete MethoCult medium. For clonogenic progenitor assay: the hematopoietic cells from day 12 of stemdiff differentiation were plated in 35-mm plastic dishes using 1.2 ml per dish of complete MethoCult H4435 medium at two different densities of cells (10,000 and 20,000 cells) and each of them by duplicate. Syringes and BD blunt fill needles 18G (BD, 30518) were used to accurately measure the desired volume due to the high viscosity of methylcellulose media. The dishes were placed on a 150 cm plate (Corning life science, 430599) and incubated at 37ºC, 5% CO_2

with a central plate without lid containing PBS to maintain hydration. Colonies were scored after 14 days of incubation based on the morphological criteria as erythroid colonies (CFU-E), granulocyte, erythrocyte, macrophage, megakaryocyte colonies (CFU-GEMM), granulocyte/macrophage colonies (CFU-GM), granulocyte colonies (CFU-G), and macrophage colonies (CFU-M).

2.6.4 Erythroid differentiation

Erythroid differentiation was carried out using a published protocol (Bayley et al., 2018) with some modifications. Firstly, hematopoietic differentiation for 10 days was carried out using the Stemdiff protocol. The single hematopoietic cells were plated at 1-2 x10^5/ml density in a 6-well plate and cultured for 18 or 20 days in IMDM containing 3% (w/v) Albumin serum (sigma, H4522), 3 mg/ml human plasma (Finished Product, QE Pharmacy) 10 µg/mL insulin (Sigma, 19278), 3 U/ml heparin (Sigma, H3149) and 500 µg/ml human transferrin (R&D, 2914-HT-001G). Hematopoietic cells were stimulated with the following cytokines (see table 2.2 for catalogue numbers):

Day 0 to Day 8: 100 ng/mL SCF, 5 ng/mL IL-3, and 3 U/mL EPO

Day 8 to Day 11: 100 ng/mL SCF and 3 U/mL EPO

Day 11 to Day 20: 3 U/mL EPO

The erythroid differentiation was monitored every 4 days by flow cytometry analysis of erythroid markers (CD71-APC, glycophorin A (CD235a-PE)) and DRAQ5- APCCy7 (staining for nucleated erythrocytes, ThermoFisher, 65088092) as described in section (2.7). Morphological analysis of erythrocytes was assessed after Kwik-Diff staining as described in section (2.8).

2.6.5 Myeloid differentiation

Myeloid differentiation was carried out using a published protocol (Hansen et al., 2018) with some modifications. After differentiating hiPSC to HSPC for 12 days using the Stemdiff protocol, the single hematopoietic cells were plated at 1 x10^5/ml density in a 12-well plate and cultured for 7 days in Stem line II supplemented with 1% Pen/strep and cytokines.

Hematopoietic cells were stimulated with the following cytokines (see table 2.2 for catalogue numbers):

Day0-day3: 10 ng/ml IL-3, 10 ng/ml GM-GSF, 30 ng/ml G-CSF, 50 ng/ml FLT-3 (300-19, Peprotech), 50 ng/ml hSCF

Day3-Day7: 30 ng/ml G-CSF

The myeloid differentiation was monitored on day 4 and day 7 by flow cytometry analysis of myeloid markers (CD14-APCCy7 and CD11b- PECy7) as described in section (2.7). Morphological analysis of myeloid cells was assessed after Kwik-Diff staining as described in section (2.8).

2.7 Flow Cytometric analysis of differentiation markers

Flow cytometric analyses were used for the assessment of the capability of hiPSC to differentiate to HSPCs. The hematopoietic cells or erythroid cells were collected at different time points during the Stemdiff protocol, erythroid differentiation, or myeloid differentiation protocols. The cells were centrifuged at 300g for 5 minutes and resuspended in 90 µL of 2% (v/v) FBS in PBS to perform surface staining followed by flow cytometry. Staining was performed on ice in 96 well plates (Costar, 3367). 10 µL of Fc block (eBioscience, 14916173, Clone AB468581) was added to the cells and incubated for 20 minutes on ice prior to the addition and one hour incubation of different antibody cocktails (Table 2.6). **For HSPCs**

marker: hemato-endothelial marker CD34, and Hematopoietic markers CD43 and CD45. **For erythroid lineage:** CD71 and CD235a. **For myeloid lineage:** CD33, CD11b and CD14. The isotype control cocktail is described in (Table 2.7). The cells were then washed with 100 µL of 2% (v/v) FBS in PBS and spun at 300g for 2 minutes. Then, the supernatant was discarded and 100 µL 2% (v/v) FBS in PBS was added to resuspend the cells. Finally, 300 µL of 2% (v/v) FBS in PBS solution was used to resuspend the cells in the tube for flow cytometry detection using a CyAn™ADP (Beckman coulter) flow cytometer. FlowJo v10.6.2 software was used to analyse the flow cytometry data.

Table 2.6: Antibodies used in flow Cytometry

Antibody	Fluorochrome	Stock concentration (mg/ml)	Dilution used	Clone	Supplier	Cat. number
CD34	PE	0.2	1:100	8G12	BD Pharmingen	550619
CD43	APC	0.2	1:100	1G10	BD Pharmingen	560198
CD45	FITC	0.2	1:100	HI30	eBioscience	11045942
CD33	FITC	0.2	1:100	HIM3-4	eBioscience	11033942
CD14	APC-CY7	0.2	1:100	61D3	eBioscience	47014942
CD71	APC	0.2	1:100	OKT-9	eBioscience	17071941
CD235a	PE	0.2	1:100	GA-R2	eBioscience	12-9987-80
CD11b	PE-CY7	0.2	1:100	ICRF44	eBioscience	15518356

Table 2.7: Isotype antibodies used in flow cytometry

Antibody	Fluorochrome	Stock concentration (mg/ml)	Dilution used	Clone	Supplier	Cat. number
Mouse-IgG1-Kappa	PE	0.2	1:100	MOPC-31C	BD Pharmingen	550617
Mouse IgG2b-Kappa, (Isotype control)	PE	0.2	1:100	Ebmg2b	eBioscience	12-4732-81
Mouse-IgG1-Kappa	APC	0.2	1:100	P3.6.2.8.1	eBioscience	17471482
Mouse-IgG1-Kappa	FITC	0.2	1:100	P3.6.2.8.1	eBioscience	11471481
Mouse-IgG1-Kappa	APC-CY7	0.2	1:100	P3.6.2.8.1	eBioscience	47471482
Mouse-IgG1-Kappa	PE-CY7	0.2	1:100	P3.6.2.8.1	eBioscience	11520627

2.8 Cytospin and Kwik-Diff staining

Erythrocyte morphology was checked by using Kwik-Diff staining (Thermofisher, 9990700). 100 µL of erythroid cells were loaded into cytospin slide chambers. After centrifugation at 200g for 5 minutes, the slides were air-dried at room temperature. Then, the air-dried slides were first dipped in Fixative Solution for 30 seconds, and immediately transferred to red solution (Eosin) for 30 seconds to stain the cytoplasm followed by being transferred to the blue Solution (Methylene Blue) for 20 seconds to stain the nucleus. The slides were rinsed in water and air dried. Finally, the slides were mounted, observed, and imaged under a bright field microscope (Leica DM6000, Leica Microsystems).

2.9 Genomic DNA extraction from MDS-hiPSC clones

2.9.1 DNeasy blood and tissue kit

For DNA extraction, the DNeasy blood and tissue kit manufacturers protocol was followed (Qiagen, 69504). Briefly, around $5x10^6$ hiPSC single cells were re-suspended in 200 µL PBS+ 20 µL of proteinase K to increase the efficiency of the extraction. Then, lysis buffer (AL) was added to the cells and incubated for 10 minutes at 56°C. To clean the sample, 100% of the Ethanol was added. The entire sample was transferred to a mini spin column to purify the DNA using different solutions. Finally, the DNA was eluted with 200 µL buffer AE. The DNA was quantified using a Nanodrop machine 2000c (Thermofisher).

2.9.2 Genomic DNA extraction (Garcia's lab protocol)

Each clone was seeded into 2 wells of a 24-well plate for 5 days until it reached 90% confluency. Then, 150 µL of cell suspension buffer (150 mM NaCl, 10 mM Ethylenediaminetetraacetic acid (EDTA) pH 7.4) was added to each well of each clone. This was followed by the addition of 150 µL of lysis buffer (20 mM NaCl, 10 mM EDTA pH 7.4, 20 mM Tris pH 7.4, 1% sodium dodecyl sulphate (SDS)) and 0.5 mg/ml proteinase K (Roche, 0311582800) was added just before use to give a final concentration of 0.5% (v/v) SDS, 10 mM EDTA, 10 mM Tris, 70 mM NaCl, 0.5 mg/ml proteinase K. The plate was covered with parafilm and incubated at 56°C in water bath overnight. Next day, 10 mg/ml of RNAseA (Roche, 10109142001) were added to each well after thorough mixing, the solution was incubated for 1 hour at room temperature. All the content of well was transferred to 1.5 ml Eppendorf tubes using cut pipette tips to avoid shredding of the genomic DNA. This was followed by addition of 1:1 equal volume of phenol: chloroform and the tubes were gently shaken and left on a rotating wheel for 1 hour. Then, the reaction mixture was centrifuged at

300g for 10 minutes and the aqueous phase was transferred to another Eppendorf tube. An equal volume of chloroform was added to the solution and the reaction was returned to the rotating wheel for 30 minutes to 1 hour. After the incubation, the reaction mixture was then centrifuged at 300g and supernatant was transferred to a new Eppendorf tube. Then, the volume was estimated, and 2 volumes of absolute Ethanol were added gently. The tubes were rocked slowly until the phases had mixed to form a DNA fibre. Finally, the DNA was purified by centrifugation at 1792g for 5 minutes at 4ºC. After centrifugation, the DNA (pellet) was washed with 70% Ethanol, dried and dissolved in 1x TE buffer. The tubes were left overnight at room temperature to allow the DNA to dissolve completely.

2.10 Colony genotyping

To corroborate that the colonies generated came from disease cells, a mutational screening was performed for MDS-hiPSC clones from both types of reprogramming using an array of the 40 most mutated genes in AML (*ABL1, BRAF, CBL, CSF3R, DNMT3A, FLT3, GATA2, HRAS, IDH1, IDH2, JAK2, KIT, KRAS, MPL, MYD88, NPM1, NRAS, PTPN11, SETBP1, SF3B1, SRSF2, U2AF1, WT1, ASXL1, BCOR, CALR, CEBPA, ETV6, EZH2, IKZF1, NF1, PHF6, PRPF8, RB1, RUNX1, SH2B3, STAG2, TET2, TP53, ZRSR2*). The genotyping was performed by Next generation sequencing and the results were directly analysed in excel and reviewed on the IGV software. The mutational screening of hiPSC clones was conducted through collaborations with Dr Eva Barragan, Hospital La Fe, Valencia, Spain.

2.11 RNA analysis for gene expression

2.11.1 RNA extraction

The RNA extraction was performed from HSPCs from day12 of stemdiff differentiation and myeloid cells from day4 and day7 of myeloid differentiation to examine the gene expression differences between the normal hiPSC, WT hiPSC (MDS27-C22), MDS27-C22+px458 vector and mutant clones (MDS27-C22.7 & MDS27-C22.20). Selected HSPCs or myeloid cells were washed with PBS and then the cell pellet was lysed by pipetting up and down in 250 µL Trizol reagent (15596026, Invitrogen) and stored at -80°C until ready for extraction. On the day of the extraction, the Trizol and cell solution was defrosted for 15 minutes at room temperature. 0.2 volumes of RNA-free chloroform (BP1145-1, Fisher Scientific) per sample was added to the mixture and shaken vigorously by hand for 15 seconds. Then, the samples were incubated for 10 minutes at room temperature and centrifuged at 14000g for 15 minutes at 4°C. After centrifugation, the aqueous phase was extracted into RNA-free Eppendorf tubes. 0.5 volumes of RNA-free isopropanol (P17490/15, Fisher Scientific) were added together with 1ul of RNA-free glycogen (10901393001, Roche). The samples were shaken gently and incubated at room temperature. After 10 minutes, the samples were centrifuged at 14000g for 10 minutes at 4°C. The supernatant was discarded, and the RNA pellet was washed with 250 µL of 70% RNA-free ethanol. Samples were centrifuged at 14000 rcf for 5 minutes at 4°C and the pellet was air dried for 5 minutes before resuspending in 14 µL RNase-free dH_2O. The concentration of RNA was measured using a Nanodrop 2000c (Thermo Scientific) spectrophotometer.

To remove the genomic DNA contamination, 4-5 µg of RNA were treated with DNAse 1 (1 U) (04716728001, Roche) with 1x DNAse 1 buffer. Sample volumes were made up to 10µl with water and incubated for 15 minutes at 37°C. Then the reaction of the enzyme was stopped by

incubating the RNA for 10 minutes at 75°C. Then the samples were cooled down on ice for 5 minutes.

2.11.2 cDNA synthesis

1-5 µg of RNA was mixed with 500 ng of oligo (dT) 15 primer (C1101, Promega) heated in heating blocks at 70°C for 5 minutes. Then, samples were cooled down on ice for 5 minutes. 1.25 µl of 10 mM dNTPs (10297018, Invitrogen), 5 µl of M-MLV reverse transcriptase 5X reaction buffer, 1 µL of 200 U/µl M-MLV (M1701, Promega), and 1 µL of RNase-out (10777019, Invitrogen) were added and mixed by gentle pipetting up and down. The mixture was then incubated at 40°C for 1 hour and stored in -20°C.

2.11.3 Quantitative Polymerase chain reaction (qPCR, Taqman)

Real time qPCR was used to analyse the cDNA produced in section 2.11.2. The reaction was performed in a MicroAmp optical 96-well reaction plate (430673, Life Technology). For each sample, 10 µL of Taqman universal master mix II (4426710, Thermo Fisher Scientific), 8 µL of RNA-free water and 1 µL Taqman oligo (Table 2.8) of the gene of interest were mixed. Then, 1 µL of cDNA was added to the mix in the plate. All reactions were done in triplicate. Then, the plate was sealed by a film and centrifuged at 3000g for 5 minutes at 4°C. The qPCR was carried out in a Stratagene Mx3005P (Agilent Technologies). The PCR conditions were 50°C for 2 minutes, 90°C for 10 minutes and then 40 cycles of 95ºC for 15 seconds and 60ºC for 1 minute. Ct values were calculated and generated by MxPro 3000 Stratagene software. Different gene relative expression values were calculated against GAPDH using the ΔΔCt mathematical model (Livak and Schmittgen, 2011).

Table 2.8: Taqman oligoes

Taqman Oligo	Supplier	Cat. number
C/EBPα	Thermofisher	Hs00269972_s1
RUNX1	Thermofisher	Hs01021970_m1
GATA2	Thermofisher	Hs00231119_m1
SPI1 (PU.1)	Thermofisher	Hs02786711_m1
LMO2	Thermofisher	Hs99999906_m1
GAPDH	Thermofisher	Hs02758991_g1

2.12 Using CRISPR-Cas9 to introduce C/EBPα mutation into MDS-hiPSC (C22)

pX458 plasmid (Addgene, 48138) was used to subclone a C/EBPα sgRNA into the BbsI restriction enzyme site as described in (Ran et al., 2013).

2.12.1 Design of single guide RNA (sgRNA)

A sgRNA targeting C/EBPα gene on chromosome 19 at position 13.11 was designed using an online tool (Trust Sanger Institute Editing database). Sequence surrounding the targeted region was submitted and sgRNA returned with score above 80 of complementary to the target was considered. Typically, sgRNA containing 18-20 bp of homology to the target sequence was used as a pair (Forward and Reverse) of complementary oligonucleotides. Nucleotide sequences "CACC" and "AAAC" were added to the 5' end of forward and reverse sgRNA. sgRNA was first phosphorylated and annealed in the following reaction, using the protocol provided by (Ran et al., 2013).

2.12.2 Annealing sgRNA oligoes

The forward and reverse sgRNA oligos needed to be annealed in order to use them as dsRNA fragments in the subsequent step. For this purpose, 100 µM of gRNA forward and 100 µM

gRNA Reverse were used with 10 U of T4 polynuclotide Kinase (PNK, Promega, M4101) in 1x T4 PNK buffer. Then, the mixture was transferred to a PCR machine at 37ºC for 30 minutes, followed by 95ºC for 5 minutes and the temperature was ramped down to 25ºC at 5ºC per minute. Then the annealed gRNA was diluted 1:200 to obtain a concentration of 0.1 µM and it was used for next step (cloning into the pX458 vector).

2.12.3 Cloning gRNA into pSpCas9 (BB)-2A-GFP (PX458)

This reaction consists of a simultaneous digestion of PX458 (Addgene, 48138) and ligation the gRNA into the vector in the same step. The reaction was performed by preparing a mixture of 100 ng of PX458 vector, 1x Tango buffer (ThermoFisher, By5), 10 mM DTT (Invitrogen, Y00147), 10 mM ATP (Invitrogen, 55082), 1 µL Fast digest BbsI (ThermoFisher, FD1014), 1500U T7 DNA Ligase (NEB, M0318S) and 0.2 annelid gRNA oligo. Following, the reaction was incubated in a thermal cycler at 37ºC for 5 minutes and 21ºC for 5 minutes and the same cycle was repeated 6 times. Once the digestion and ligation reaction were completed, the sample was processed to digest any residual linearized DNA. The Plasmid safer exonuclease kit (Lucigen, E3101K) was used to treat the ligation sample. To begin this process, the entire ligation sample from the previous step was treated with 1U of plasmedSafe exonuclease, 10mM ATP and 1x plasmid safer buffer. Then the whole reaction was incubated in a PCR machine using the following parameters: 37ºC for 30 minutes and 70ºC for 30 minutes. After the PCR run was finished, the DNA plasmid was ready for transformation.

2.13 Competent bacteria transformation

The transformation was performed by adding 4 µL of ligated plasmid to one shot Stbl3 chemically competent E.coli (life technology, C7373-03). Then, the mixture was transferred to 42ºC water bath for 30 seconds and transferred again to the ice for 5 minutes. Then, the bacteria

were plated on Lysogeny broth (LB) agar plate with 50 µg/ml ampicillin (Sigma, 10835242001) and the plates were incubated overnight at 37ºC. The next day, around 12 colonies were picked to grow in LB with 100 µg/ml ampicillin for miniprep.

2.14 Plasmid DNA extraction (Miniprep & Maxiprep)

QIAprep Spin Miniprep Kit (QIAGEN, 12123) was used according to the manufacturer to perform miniprep from bacteria colonies. In brief, 2 ml of bacterial culture was pelleted by centrifugation at 6800g for 3 minutes at room temperature. Then, the bacterial pellet was resuspended in P1 buffer and transferred to a microcentrifuge tube. P2 buffer was added to lyse the bacterial pellet and the tube was mixed gently by inverting it 4–6 times. After adding P3 buffer and mixing thoroughly, the solution was centrifuged for 10 minutes at 17,900g. Then, the supernatant was collected and transferred to a QIAprep 2.0 spin column for DNA purification. Following, the DNA pellet was washed with PE buffer and eluted with a EB buffer into a clean 1.5 ml microcentrifuge tube. The concentration of DNA was measured by nanodrop and stored at -20°. Then, the DNA of 12 colonies were sent for sequencing using the U6 promoter primer (5'TACGATACAAGGCTGTTAGA'3) to check for successful cloning. The sequencing was carried out by Source bioscience facility, Nottingham, UK.

The maxiprep for the plasmid obtained from a positive colony was performed using a plasmid plus maxi kit endotoxin-free (QIAGEN, 12362). A 3 ml bacterial pre-culture was inoculated in 400 ml of LB and incubated on a shaker at 250 rpm at 37°C overnight. On the next day the bacterial culture was transferred into a 500 ml tube and spun at 3000g at 4°C for 30 minutes. The supernatant was discarded, and the pellet was resuspended in 10 ml resuspension buffer (P1 buffer). 10 ml of lysis buffer (P2 buffer) was added to the bacterial suspension and the tube was inverted 3-4 times. It was then incubated for 5 minutes at room temperature. 10 ml of

neutralization buffer (P3 buffer) was added, and the tube was inverted 3-4 times. The bacterial lysate was immediately transferred into the QIAfilter Cartridge by pouring, not pipetting, and incubated for 10 minutes at room temperature to allow the precipitate float to the top of the solution. Once the precipitate has floated to the top, the plunger was inserted gently to pass the lysate through QIAfilter Maxi Cartridge into a new 50 ml tube. Next, the filtered lysate was collected in a 50 ml tube and 2.5 ml of ER buffer was added and incubated on ice for 30 minutes. Then, the lysate was loaded onto QIAGEN-tip column until the entire sample was passed through the QIAGEN-tip. The column was washed 3x with 30 ml QC buffer. After washing, the column was transferred to a new 50 ml tube and 15 ml QN buffer was added to the column to elute the DNA. Then, the DNA was precipitated by adding 10.5 ml isopropanol to the eluted DNA and the sample was mixed and centrifuged immediately at 15,000g for 30 minutes at 4°C. Then, the supernatant was carefully decanted, and the pellet was washed with 70% ethanol and re-dissolved with the appropriate amount of TE buffer. The DNA yield was determined by nanodrop and the sample was stored at -20°.

2.15 Nucleofection of MDS-hiPSC (C22) with PX458+ C/EBPα gRNA

For nucleofection, the P3 amaxa kit (Lonza, V4XP-3024) was used and the nucleofector solution P3 with supplement was prepared in the proportions described in the manufacturer's instructions (Lonza). To electroporate hiPSC clones, $1.5x10^6$ cells of MDS27 C22 were suspended in nucleofector reagent (solution P3 + supplement) 1 µg PX458+ C/EBPα gRNA and 1 µg of PX458 plasmid alone as control reaction. The mixture was transferred into a nucleocuvette and placed into the 4D nucleofection system (Lonza) with program DS-150. The transfection efficiency was assessed according to GFP expression, and the positive cells were sorted using BD FACSAria™ Fusion (BD bioscience) after gating the live cells based on a

negative control. Seven days post sorting, 24 clones were picked and expanded to extract genomic DNA as describe in (2.9) for assessing the gene editing efficiency.

2.16 PCR to amplify the C/EBPα target region

The target region of C/EBPα was amplified using Q5 High-Fidelity DNA polymerase (NEB, M0491). The thermocycling conditions of each sample were: 98 ºC for 30 seconds, 39x (98ºC for 10 seconds, 64ºC for 30 seconds, 72ºC for 30 seconds) and final extension at 72ºC for 2 minutes. Each reaction contained: 0.02 U/µL of Q5 High-Fidelity DNA polymerase, 200 µM dNTPs, 0.5 µM of each primer and 100 ng of DNA. The reaction was carried out in 25 µL and amplification was checked by gel electrophoresis.

Primers used for PCR

C/EBPα forward 5'GGCCTCTTCCCTTACCAGCC'3

C/EBPα reverse 5'CTGGTCAGCTCCAGCACCTT'3

The agarose gel electrophoresis was done by heating 1 % (w/v) agarose (Sigma, A9539, w/v) with 1x TAE buffer (40mM Tris pH7.6, 20mM acetic acid and 1 mM EDTA) and 0.5 µg/ml of Ethidium bromide (Sigma, E8751), using a microwave for 3 minutes on high power until dissolved. When the agarose solution had cooled down to about 40°C, ethidium bromide was added, then poured into a gel casting tray. The gel was solidified by keeping it at room temperature for about 20 minutes, then placed into a horizontal electrophoresis tank and covered with 1x TAE buffer. 6x loading buffer (Orange G) was added to each PCR product and the mixture loaded onto the gel, along with 5 µL of 100pb DNA ladder (NEB, N0551G). Gels were run at 100 V for about 1 hour. After 1 hour the gel was removed and placed into the ChemiDoc MP imaging system (Bio-Rad) to determine the size of plasmid according to the

DNA ladder. The expected size for the wild type band was 456 bp. Then, the positive clones were processed using a T7 endonuclease I assay.

2.17 T7 endonuclease I (T7EI) assay to verify successful cutting and gene targeting

PCR products were processed using a T7 endonuclease I (T7EI) assay using the Alt-R genome editing detection Kit (IDT, 1075932) according to the manufacturer's protocol. To estimate the editing efficiency of sgRNA, PCR products of the target region of C/EBPα from transfected cells were incubated with 1 U/μL of T7EI for 60 minutes at 37°C to recognise and cleave non-perfectly matched DNA created by gRNA. Two positive controls (Control A; homoduplex control, Control B; heteroduplex control) were used to monitor the function of T7EI assay. Fragments of DNA were analysed on a 2% agarose gel as described in 2.18. Then, the DNA of clones was extracted and sent for next generation sequencing to confirm the type of mutation.

2.18 Agarose gel extraction

The gel fragment was purified using a gel extraction kit (QIAGEN, 28704) according to the manufacturer's protocol. Briefly, the desired fragment of DNA was cut from an agarose gel with a sharp and clean scalpel under blue light. Then, buffer was added to the sample (3x the weight) and incubated at 50°C for 10 minutes to dissolve completely the gel slice. To precipitate the DNA, 1 volume of isopropanol was added to the sample and the entire mixture was transferred into a QIAquick column. Then, the DNA was washed with PE buffer. Finally, the DNA was eluted into a clean 1.5 ml microcentrifuge tube with EB buffer. The purified DNA concentration was measured using a Nanodrop 2000c (ThermoFisher). The DNA samples were sent for sequencing using the primers listed below. The sequencing was carried out by Source bioscience facility, Nottingham, UK.

C/EBPα forward 5'GGCCTCTTCCCTTACCAGCC'3

C/EBPα reverse 5'CTGGTCAGCTCCAGCACCTT'3

2.19 Western blotting for the C/EBPa protein

2.19.1 Protein extraction

TrypLE Express was used to harvest the iPSCs and centrifuged at 250 g at 4°C for 5 minutes. The pellets were resuspended in 1ml ice-cold PBS. Cells were counted and 1.5 x10^6 were centrifuged once more. The pellet was resuspended in 70 µL total protein lysis buffer (20mM Tris pH 7.4, 10mM EDTA, 100mM NaCl, 1% (v/v) Triton-100, 1mM Sodium Fluoride (NaF), 1mM β-glycerophosphate, 1mM Ethylene glycol tetra-acetic acid (EGTA), 5mM Polypropiolate sodium (PPNa), 1mM Phenylmethylsulfonyl fluoride (PMSF, Sigma, P7626), 1mM sodium orthovanadate (Na_3VO_4, Sigma, S6508), 1x halt protease inhibitor cocktail (ThermoFisher, 78430) on ice (50µl/1x10^6). The cell suspension for each sample was set aside on ice for 30 minutes before the centrifugation at 600 g for 10 minutes at 4°C. The supernatant from each sample was collected and transferred to an Eppendorf tube and 5 µL of the sample was set aside for determining protein concentration, Samples were snap frozen by putting the sample on dry ice and quickly adding 100% ethanol. The samples were then stored at -80 °C. To determine the protein concentration of each sample, a Bradford absorption assay was implemented. Firstly, a standard curve was generated using different concentrations of BSA (0, 1, 2, 4, 6, 8 and 10 mg/µl). The BSA samples were diluted in 800 µL H_2O plus 200 µL of Bradford reagent (Sigma, B6916) and the sample was mixed by inversion. 1 µL of each protein sample was also diluted in the same way. Nanodrop was used to measure the absorbance at 595 nm and a graph of absorbance vs concentration was plotted. Then, the absorbance measurement for each sample was used to determine the protein concentration in µg/µl.

2.19.2 Western blot

The purified proteins were prepared for gel electrophoresis by adding 4x sample buffer (40% (v/v) Glycerol, 240 mM Tris-HCl pH 6.8, 8% (w/v) SDS, 0.04% (w/v) bromophenol blue, 5% (v/v) β -mercaptoethanol) and heating at 95 ºC for 10 minutes. Protein lysate from hiPSC clones and protein lysate from positive control cells (Kasumi, and Kasumi+ siRNA-ETO RUNX1) and a protein ladder (Thermo fisher, 26634) were electrophoresed by SDS-PAGE electrophoresis on 10% acrylamide precast SDS-PAGE ready gels, in a vertical tank, filled with 1X SDS PAGE running buffer (25mM Tris, 190mM glycine, and 0.1% (w/v) SDS with adjusted pH of 8.3), at 100V for about 2 hours. Proteins separated by SDS-PAGE were transferred onto PDVF membrane (Millipore) using a semi-dry transfer system (Biorad transblot turbo) at 25V for 30 minutes. Then, the membrane was stained with 0.1% Ponceau S staining (Sigma, P7170) to visualize the successful protein transferred. Post-transfer, the membranes were blocked in 5% of Milk in TBST (50 mM Tris-Cl, pH 7.6; 150 mM NaCl plus 0.1% (v/v) Tween20) for 1 hour at room temperature. After blocking, the membranes were incubated with 1:1000 C/EBPa primary Ab (Rabbit, ab15043, Abcam) overnight at 4ºC on a roller. Next day, three TBST washes for 10 minutes each were applied to the membranes and then incubated with 1:5000 horseradish peroxidase (HRP) conjugated secondary antibody specific to the primary antibody (Amersham, NA9340) for 1 hour at room temperature on a roller. After further three washes in TBST, protein bands (42 and 30 kDa) were detected using the Super Signal west pico plus chemiluminescent substrate (Thermo Scientific, 34580). The chemiluminescence was developed onto photographic film by using the ChemiDoc MP imaging system (Bio-Rad).

2.20 Statistical analysis

The statistical analysis was performed using GraphPad Prism 8 (GraphPad Prism version 8.0 for Mac, GraphPad Software, San Diego California USA). All data are expressed as Mean ± Standard Error of the Mean (SEM). *P* values less than 0.05 were considered statistically significant and a star (*) was labelled in the figures. The following statistical analyses were used: Two-way ANOVA to analyse data describing two factors across multiple parametric groups. One-way ANOVA to analyse data describing one factor across multiple parametric.

Chapter 3 : Generation and characterisation of human induced pluripotent stem cells (hiPSCs) from patient MDS27

3.1 Introduction

MDSs are a group of heterogeneous clonal hematopoietic disorders which originate at the level of stem and progenitor cells (HSPCs). MDS is characterised by ineffective haematopoiesis, dysplastic changes, peripheral blood cytopenia and increased risk of transformation to AML (Greenberg et al., 1997).

The few treatment options are the main problem for MDS patients. The standard care of higher risk-MDS is treatment with hypomethylating agents such as azacytidine, but 50-60% of these patients do not respond to treatments, and 30% of MDS patients develop overt AML during the disease (Jabbour et al., 2010, Breccia et al., 2010).

The mechanisms of resistance to chemotherapy and transformation from a chronic MDS phase to a more aggressive phase are still poorly understood. This is mainly due to the limited number of cells obtained from MDS patients and the lack of mouse model systems that faithfully resemble human MDS. Thus, we urgently need innovative model systems which can provide sufficient cells to understand the molecular mechanisms behind the resistance to chemotherapy and progression to the worse stage of the disease. Generation of hiPSC from MDS patients could be a valuable model system to investigate the molecular mechanisms of MDS progression.

The generation of hiPSC from MDS patients has been reported by two groups. The first group led by Eirene Papapetrous reprogrammed two MDS patient samples harbouring *del(7q)* mutations using an integrating method (lentiviral vector). They found that *del(7q)* haploinsufficiency impaired the hematopoietic differentiation (Kotini et al., 2015). The other group generated different hiPSC lines from eight MDS patients to identify the stages of clonal evolution in MDS patients using non-integrating methods (SeV and episomal) (Hsu et al., 2019).

The aim of this chapter is twofold; Firstly, to generate hiPSC from patient MDS27 using non-integrated methods as a model system for MDS; and secondly, to characterise this model system by using different approaches to validate the pluripotency of hiPSC lines derived from patient MDS27.

3.2 Results

3.2.1 Somatic reprogramming

We derived hiPSC lines from PBMNCs of an MDS patient (MDS27) from a sample taken at the time of diagnosis in 2013. Patient MDS27 was diagnosed with multilineage dysplasia and RA (low-risk MDS) in July 2013. At this stage, he presented a normal karyotype and cells harboured three different mutations (*ASLX1*, *RUNX1*, *SRSF2*) as shown in Figure 3.1.

Two years later (2015) the patient relapsed, the number of blast cells increased to 9-17% and a tandem C-terminal duplication of the *C/EBPα* gene was detected (Figure 3.1). At this stage, the patient had progressed to high-risk MDS and was treated with demethylating chemotherapy agents but failed to respond. Patient MDS27 was selected for generating hiPSC because he had progressed from low-risk to high-risk, longitudinal samples were available and he had not responded to chemotherapy, eventually progressing to AML.

Sample	Gene	Locus	mutation type	VAF	Function	Coding	Protein
MDS27 sample in 2013	ASXL1	chr20:31023821	SNV	50.1	Missense	c.3306G>T	p.Glu1102Asp
	RUNX1	chr21:36206864	INDEL	15.11	Frameshift Insertion	c.648_649 ins TCC	p.Gly217fs
		chr21:36259324	SNV	51.15	Missense	c.167T>C	p.Leu56Ser
	SRSF2	chr17:74732935	SNV	30.29	Missense	c.284C>A	p.Pro95His
MDS27 Sample in 2015	ASXL1	chr20:31023821	SNV	48.54	Missense	c.3306G>T	p.Glu1102Asp
	RUNX1	chr21:36206864	INDEL	46.79	Frameshift Insertion	c.648_649 ins TCC	p.Gly217fs
		chr21:36259324	SNV	49.24	Missense	c.167T>C	p.Leu56Ser
	SRSF2	chr17:74732935	SNV	47.84	Missense	c.284C>A	p.Pro95His
	C/EBPα	chr19:33792551	INDEL	38.02	Frameshift Insertion	c.769_770 ins CTGG	p.Gly257fs

Figure 3.1: Genotyping screening for MDS27 samples

Genotyping was performed by the Hospital La Fe, Valencia, Spain on an MDS27 peripheral blood sample in 2013 and in 2015 when the patient progressed to an aggressive disease stage, using an array of the 40 most mutated genes in AML.

To create an *in vitro* model system that could help to understand the molecular mechanisms associated with treatment failure and disease progression, we generated hiPSC from MDS27 PBMNCs from a sample in 2013 using a non-integrating method: Sendai virus (SeV). The protocol for human iPSC induction is summarised in Figure 3.2 (A). After expansion and stabilisation in the expansion media, the cells were infected with SeV to deliver and express the Yamanaka factors. Two days after transduction, the cells were seeded into a 6-well plate containing inactivated MEFs and fed with media at different times. By 25 days after transduction, the culture of PBMNCs showed distinct types of colonies that were flat and resembled hESC cell colonies. Twenty-four colonies were picked, expanded in 6-well plates coated with matrigel and established as MDS27 iPSC cell lines (Figure 3.2, B).

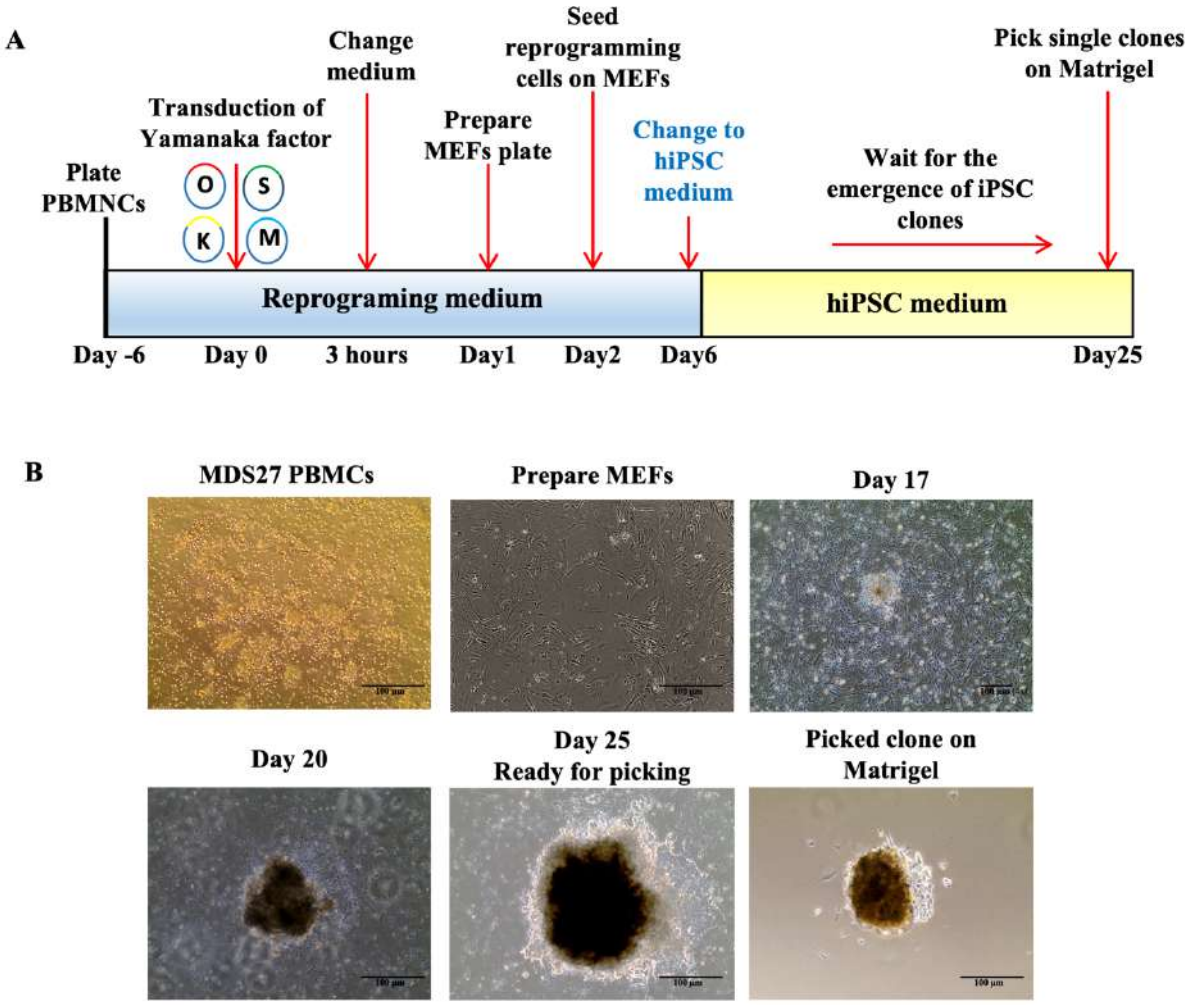


Figure 3.2: Non-integrating method (Sendai virus) enables to generate hiPSC from patient MDS27

(A) Schematic representation of the somatic reprogramming process using a non-integrating method: Sendai vector. PBMNCs were transduced with Sendai virus containing OSKM factors (*Oct4*, *Sox2*, *Klf4* and *c-Myc)* **(B)** Phase contrast microscope images showing PBMNCs reprogramming at day 17, 20, 25, after picking and expanding the clones. 10x magnification, 100 µm scale bar. PBMNCs (peripheral blood mononuclear cells), MEFs (mouse embryonic fibroblast).

3.2.2 Genotyping for hiPSCs generated from MDS27

The generated clones from the 2013 sample were genotyped by Next-generation sequencing using an array with the forty most mutated genes in AML at the Hospital La Fe, Spain. The purpose of this assay was to corroborate that the clones generated came from disease cells and not healthy cells. Moreover, this analysis would allow us to determine whether clones with

different combinations of mutations were generated and inform us whether additional mutations were produced during the reprogramming process. The genetic study was done through the analysis of the coding regions and flanking intronic regions of the genes included in the myeloid SMD panel. Genes included: *ABL1, BRAF, CBL, CSF3R, DNMT3A, FLT3, GATA2, HRAS, IDH1, IDH2, JAK2, KIT, KRAS, MPL, MYD88, NPM1, NRAS, PTPN11, SETBP1, SF3B1, SRSF2, U2AF1, WT1, ASXL1, BCOR, CALR, CEBPA, ETV6, EZH2, IKZF1, NF1, PHF6, PRPF8, RB1, RUNX1, SH2B3, STAG2, TET2, TP53, ZRSR2*.

The genotyping screening confirmed that 100% of the 24 hiPSC colonies examined were positive for the *ASLX1*, *RUNX1* and *SRSF2* mutations indicating that these colonies derived from the same MDS27 subclone. Unfortunately, this reprogramming failed to generate colonies with different combinations of these mutations, and the mutational screening confirmed that no isogenic healthy control clones from MDS27 were generated (Figure 3.3 and Supplement figure 1).

Sample	Gene	Locus	Mutation type	VAF	Function	Coding	Protein
MDS27 hiPSC clone1	ASXL1	chr20:31023821	SNV	50.90	Missense	c.3306G>T	p.Glu1102Asp
	RUNX1	chr21:36206864	INDEL	15.18	Frameshift Insertion	c.648_649 ins TCC	p.Gly217fs
		chr21:36259324	SNV	48.47	Missense	c.167T>C	p.Leu56Ser
	SRSF2	chr17:74732935	SNV	56.50	Missense	c.284C>A	p.Pro95His
MDS27 hiPSC clone8	ASXL1	chr20:31023821	SNV	51.70	Missense	c.3306G>T	p.Glu1102Asp
	RUNX1	chr21:36206864	INDEL	48.34	Frameshift Insertion	c.648_649 ins TCC	p.Gly217fs
		chr21:36259324	SNV	47.38	Missense	c.167T>C	p.Leu56Ser
	SRSF2	chr17:74732935	SNV	53.32	Missense	c.284C>A	p.Pro95His
MDS27 hiPSC clone11	ASXL1	chr20:31023821	SNV	51.17	Missense	c.3306G>T	p.Glu1102Asp
	RUNX1	chr21:36206864	INDEL	51.66	Frameshift Insertion	c.648_649 ins TCC	p.Gly217fs
		chr21:36259324	SNV	49.92	Missense	c.167T>C	p.Leu56Ser
	SRSF2	chr17:74732935	SNV	53.82	Missense	c.284C>A	p.Pro95His
MDS27 hiPSC clone22	ASXL1	chr20:31023821	SNV	49.64	Missense	c.3306G>T	p.Glu1102Asp
	RUNX1	chr21:36206864	INDEL	51.20	Frameshift Insertion	c.648_649 ins TCC	p.Gly217fs
		chr21:36259324	SNV	47.13	Missense	c.167T>C	p.Leu56Ser
	SRSF2	chr17:74732935	SNV	54.78	Missense	c.284C>A	p.Pro95His

Figure 3.3: Genotyping screening shows that MDS27 clones harbour the same mutations as the original sample

Genotyping was performed by the Hospital La Fe, Valencia, Spain for four selected clones (C1, C8, C11, C22) generated from MDS27 peripheral blood sample taken in 2013 by Sendai vector, using an array of the 40 most mutated genes in AML. All four selected clones harbour the same mutations as the original sample (2013).

The generation of healthy isogenic control iPSC clones was important for comparison purposes. Thus, we generated new clones by performing another round of somatic reprogramming from the same sample taken in 2013, using this time a different non-integrating method: episomal reprogramming. As shown in Figure 3.4 (A), PBMNCs were transfected with Yamanaka's factors plasmids by nucleofection. After three days of transfection, the cells were plated on MEFs and four days later the transfected cells plated in hESCs cell culture medium supplemented with the ROCK inhibitor Y27632, previously shown to enhance survival and clonogenicity of single dissociated hESCs cells. 28 days after transfection, twenty-four individual hiPSCs colonies were isolated from the culture and expanded by culturing on mitotically inactivated MEFs for the first passage. Then, the cells were maintained on matrigel and established as individual iPSC clones (Figure 3.4, B). As before, a mutational screening of the forty most common mutated genes in MDS was performed. Similar to SeV reprogramming, the clones generated from episomal reprogramming were all positive for *ASLX1*, *RUNX1* and *SRSF2* mutations (supplement figure 2).

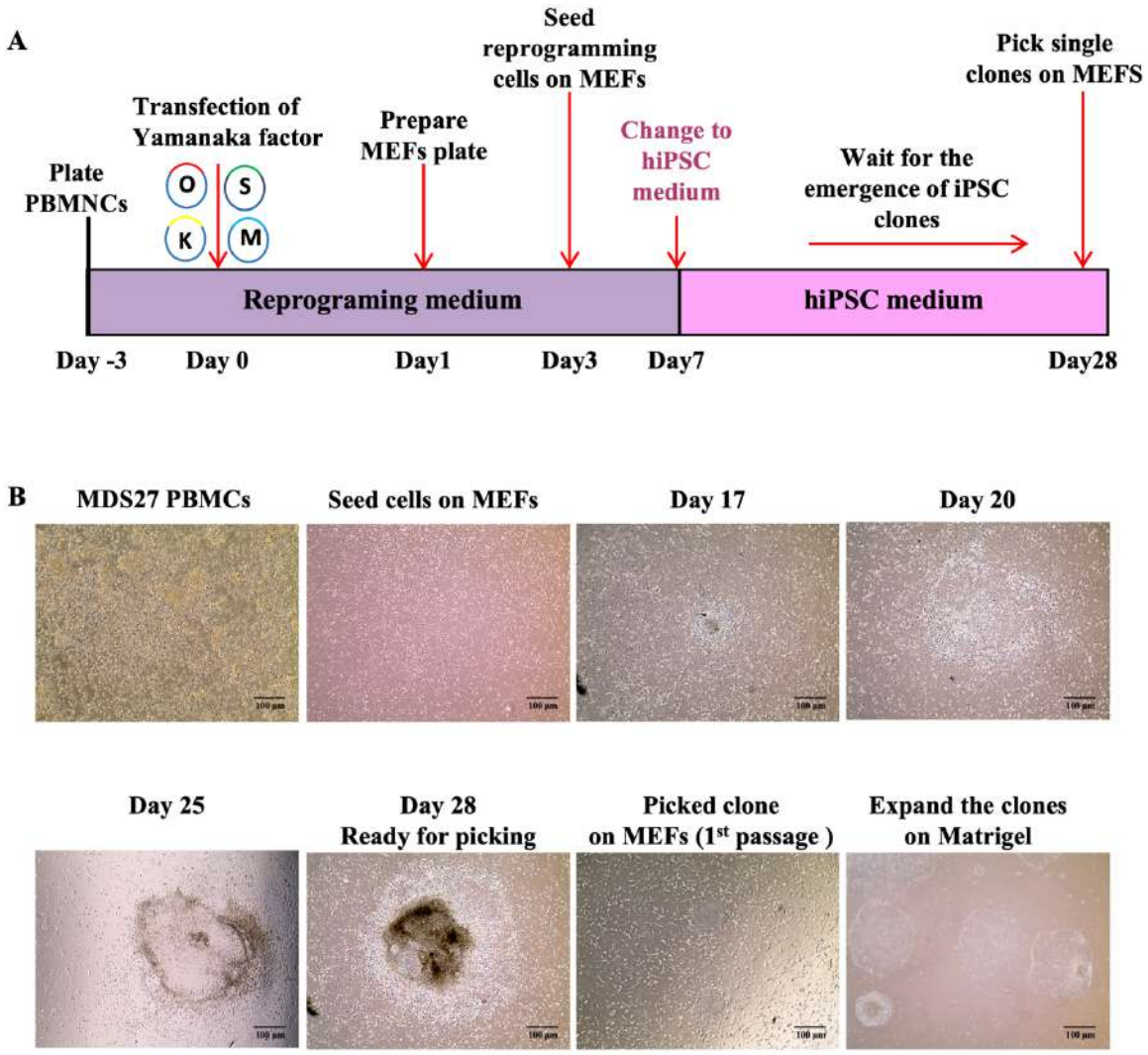


Figure 3.4: Different iPSC clones can be generated from PBMNCs of patient MDS27 using the episomal method

(A) Schematic representation of the somatic reprogramming process using the non-integrating method: Episomal reprogramming. PBMNCs were transfected with OSKM factor plasmids (Oct4, Sox2, Klf4, L-Myc and LIN-28) **(B)** Phase contrast microscope images showing PBMNCs reprogramming at day 17, 20, 25, after picking and expanding the clones. 4x magnification, 100 µm scale bar. PBMNCs (peripheral blood mononuclear cells), MEFs (mouse embryonic fibroblast).

Together these results showed that cells harbouring these three mutations seem to have a better predisposition for reprogramming towards a stem cell-like stage. Moreover, both methods had low efficiency in reprogramming the healthy cells of MDS27 as well as different sub clonal abnormalities. It is worth mentioning that several attempts to generate hiPSC from the same

patient after disease progression to high-risk and sAML (sample in 2015 and 2016) failed to generate stable colonies even with some optimisation to the protocol.

3.2.3 Identifying the chromosomal stability after somatic reprogramming

As mentioned above, an isogenic healthy iPSC clone from patient MDS27 was not obtained after two different methods of reprogramming. Subsequently, an iPSC cell line (BU3-10), generated from healthy PBMNC, was used as hiPSC control. Four different clones generated by SeV (C1, C8, C11, C22) were used for further hiPSC-MDS27 characterisation. Despite all clones harbouring the same mutations (*ASLX1*, *RUNX1*, *SRSF2*), the use of multiple clones would help to distinguish between effects due to the three mutations and effects due to additional mutations that could have occurred during the somatic reprogramming process.

Cells from the MDS27 patient taken on 2013, presented a normal karyotype. Thus, we first performed a karyotyping analysis to identify possible chromosomal numerical changes, as aneuploidy may have occurred during the reprogramming process. Four selected clones (C1, C8, C11 and C22) were cultured and treated with colcemid, a drug which depolymerises the mitotic spindle (Rieder and Palazzo, 1992, Taylor, 1965) to block cells at the metaphase stage. After two hours of treatment, chromosome preparation was performed, and the number of chromosomes counted. As human cells contain 46 chromosomes, metaphases displaying more or less than 46 chromosomes would represent aneuploidy.

All the MDS27-hiPSC lines (C1, C8, C22 and C22) showed a normal number of chromosomes (46, XY) as can be seen in Figure 3.5. This result proves that the PBMNCs from MDS27 did not suffer chromosome instability during their somatic reprogramming to hiPSC using the non-integrating SeV reprogramming approach.

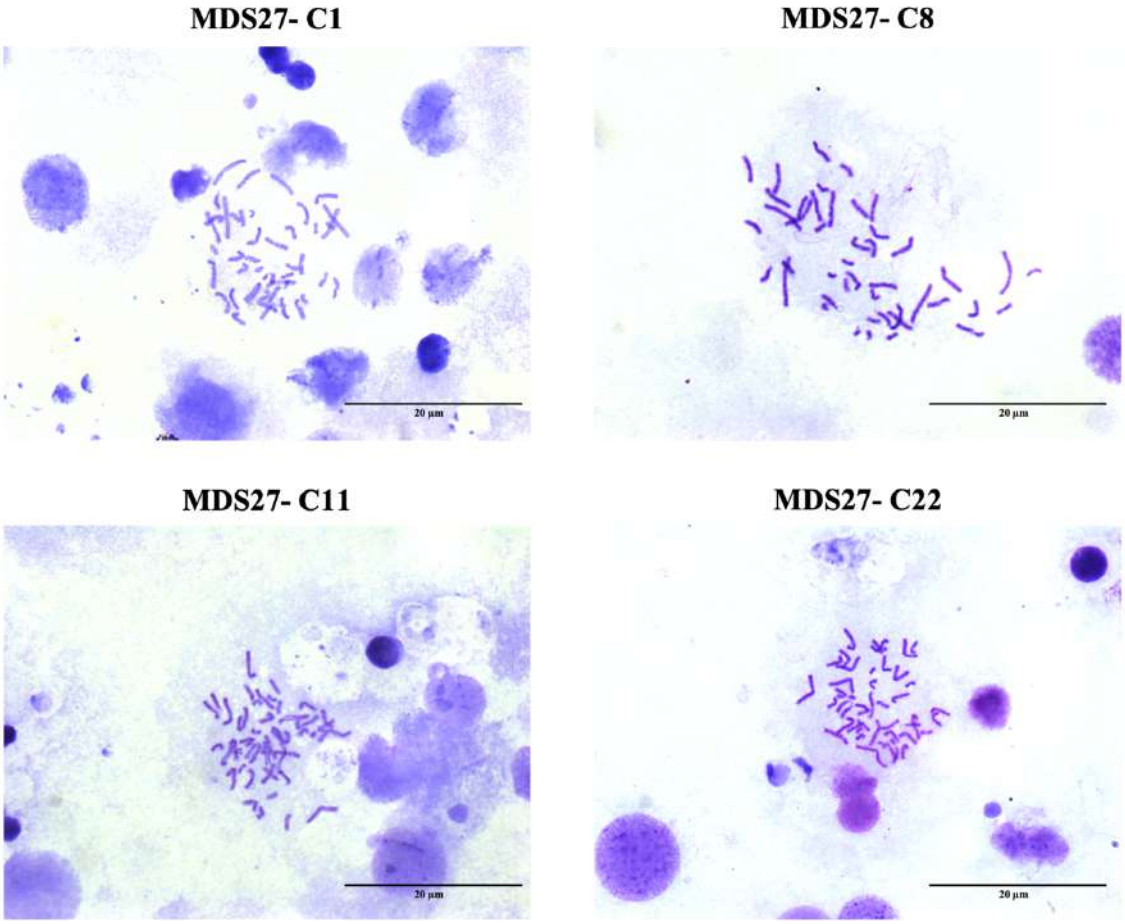


Figure 3.5: Human iPSC clones generated from MDS27 have a normal number of chromosomes

80 % confluent well of a 6-well plate for MDS27 clones were treated with 0.02 µg/ml colcemid for 2 hours, swollen with 0.075 M KCl, fixed with cold Carnoy's solution, dropped onto humidified, chilled glass slides, then stained with Giemsa staining and imaged at 100x magnification. (Data represent 3 independent experiments: >25 metaphases per sample per experiment), 20µm scale bar, 100x magnification, Leica DM6000 light microscope.

3.2.4 Pluripotency characterisation of reprogrammed somatic cell lines

After showing that the four clones displayed a normal karyotype, we next sought to characterise the pluripotency potential of the individual clones. The pluripotency characterisation of hiPSC was assessed based on (i) morphology, (ii) Alkaline phosphatase staining, (iii) positive expression of pluripotent protein markers and, (iv) their ability to differentiate to the three germline cells.

As shown in Figure 3.6, human iPCS colonies from the four selected clones presented the same morphology as the hiPSC control cell line (BU3.10): with the typical flat and tightly packed morphology, with large nuclei and scant cytoplasm (Takahashi et al., 2007).

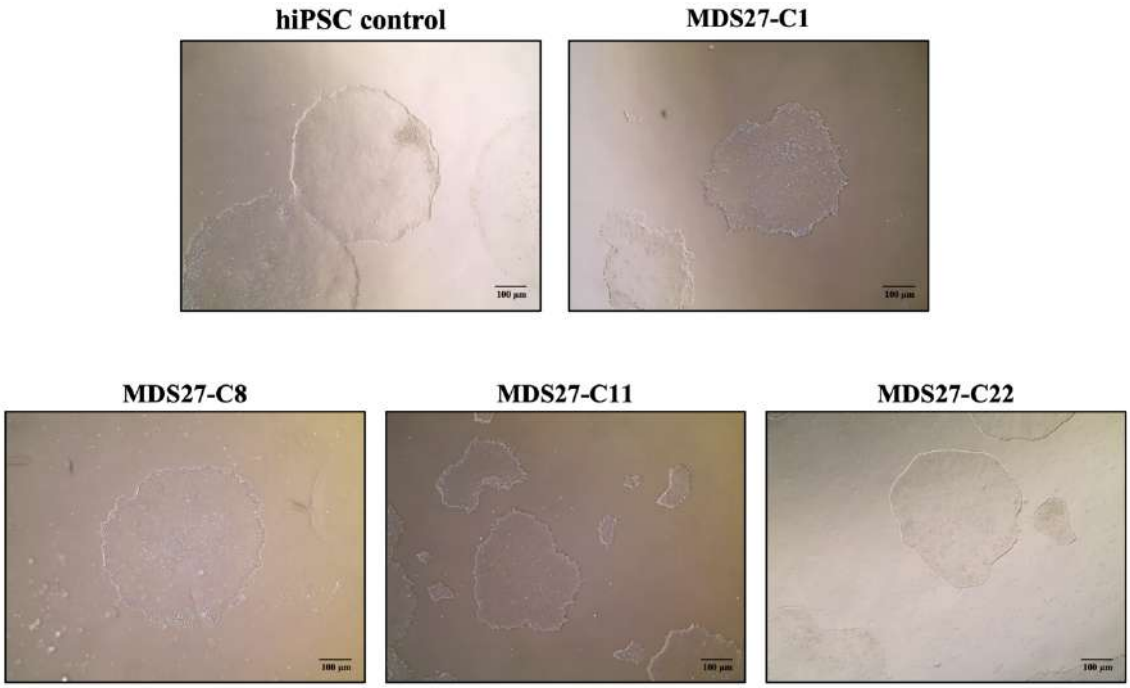


Figure 3.6: Human iPSC morphology

Human iPSC control (BU3-10) and MDS27 clones (C1, C2, C8 and C22) were grown in 6-well plates coated with matrigel and pictures were taken when the cells reached 80% confluency. 4x magnification, 100 µm, EVOS cell image system microscope (ThermoFisher Scientific).

Next, MDS27-hiPSC colonies (C1, C8, C11, C22) were cultured in a 6-well plate until they reached 80% confluency and stained for alkaline phosphatase activity. Alkaline phosphatase activity is a membrane-bound enzyme which is highly expressed by embryonic cells and downregulated in differentiated cells (Kim and Wyckoff, 1991). As expected, the hiPSC control showed positive AP staining as denoted by the pink-red cytoplasm. Similarly, colonies from the MDS27-hiPSC clones (C1, C8, C11, C22) showed positive staining for AP (Figure 3.7).

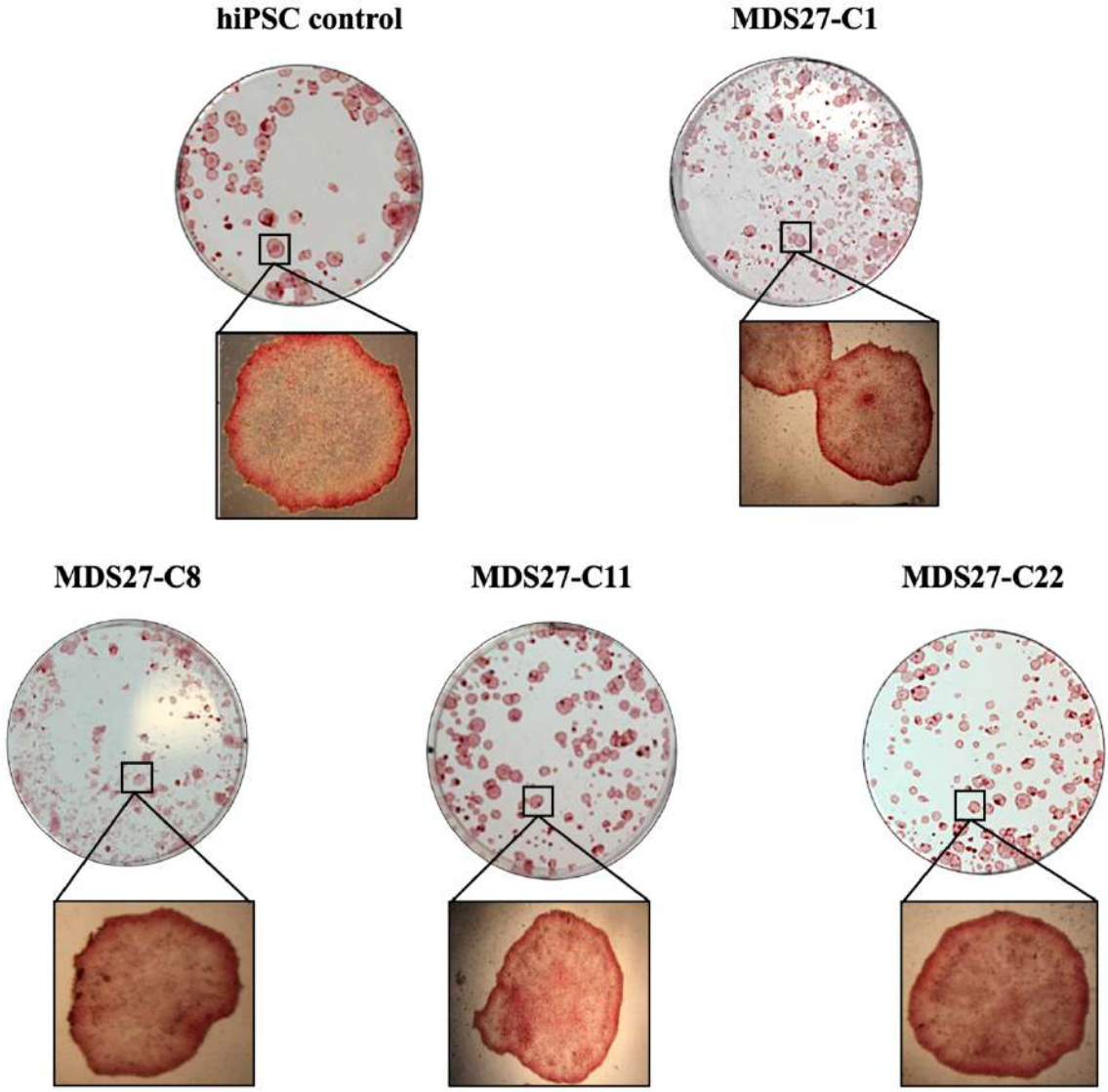
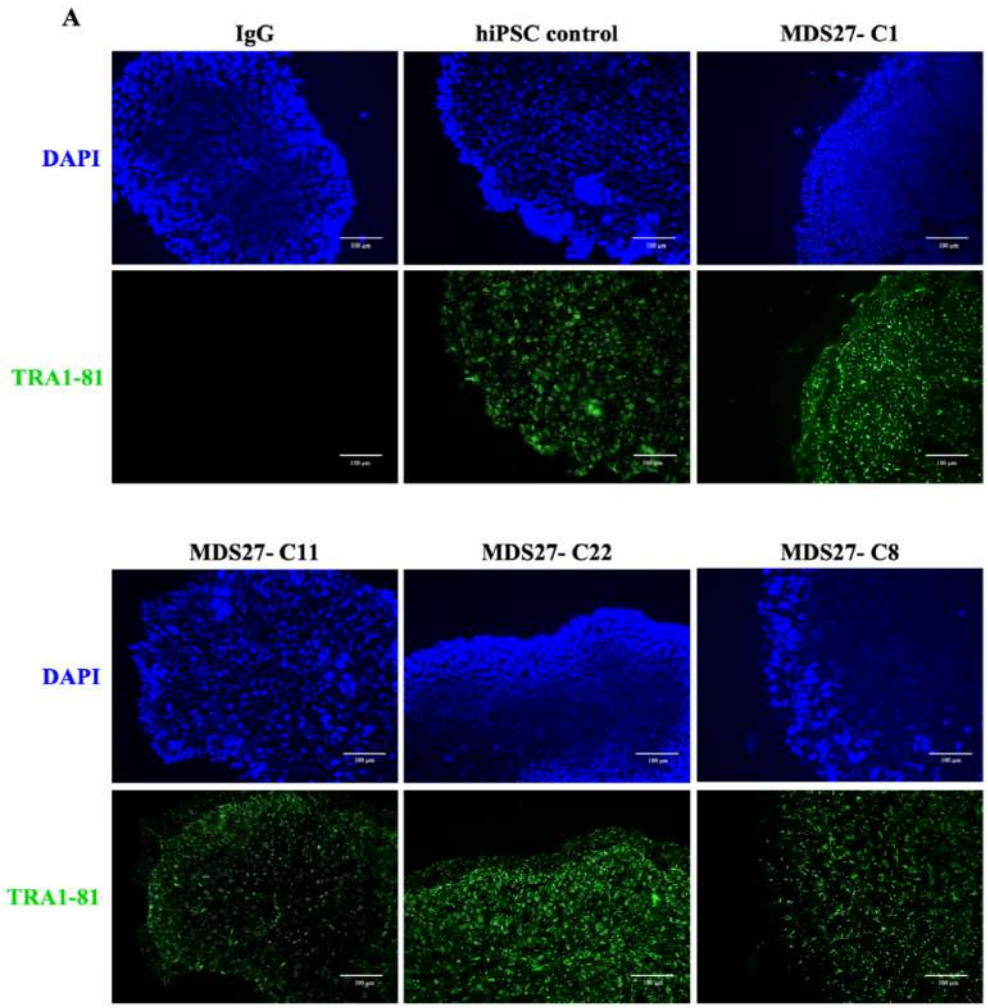


Figure 3.7: Positive expression of alkaline phosphatase staining in all hiPSC lines

Human iPSC control and early passage of MDS27 clones (C1, C2, C8 and C22) were grown in a 6-well plate coated with matrigel before staining for alkaline phosphatase activity for 30 minutes. Images show hiPSC clones with positive (red) alkaline phosphatase activity. 10x magnification, primo vert microscope (ZEISS). Data represent 3 independent experiments.

Moreover, we sought to determine the expression of pluripotent markers in the hiPSCs generated. MDS27-hiPSC C1, C8, C11 and C22 as well as the BU3.10 hiPSC control were grown on coverslips coated with matrigel, and then cells were stained with different types of antibodies, against the extracellular pluripotency marker TRA1-81 and the intracellular pluripotency markers NANOG and SOX2. TRA1-81 and NANOG were chosen as they are genes not contained within the reprogramming cocktail, and thus, their expression can only be due to successful reprogramming. All MDS27-hiPSC lines and the hiPSC control strongly expressed the pluripotent markers TRA1-81 (green) Figure 3.8 (A) , as well as NANOG (green) and SOX2 (Red) Figure 3.8 (B), further confirming that the iPSC generated from patient MDS27 were successfully stable iPSCs that maintain the expression of pluripotent markers.



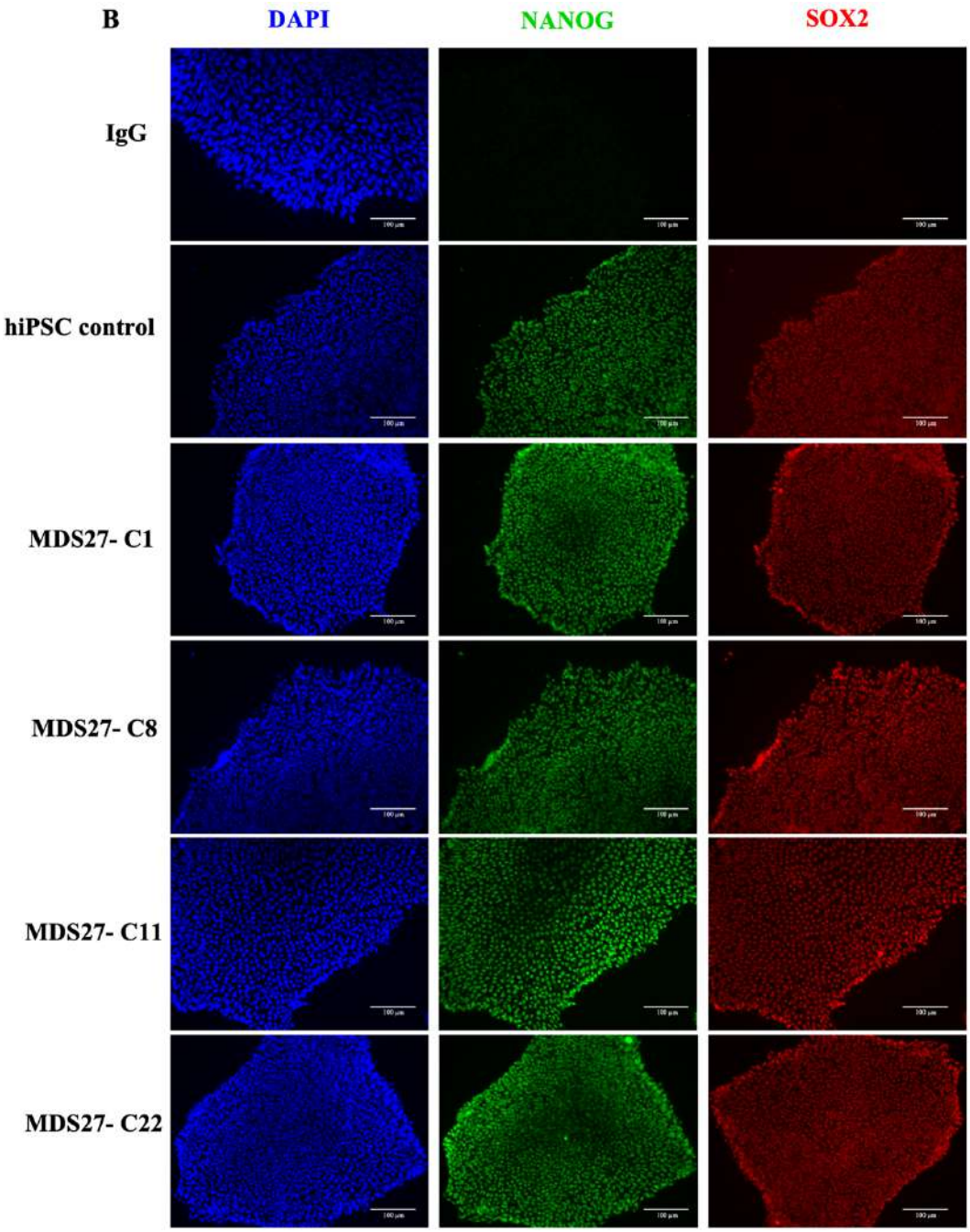


Figure 3.8: Positive expression of pluripotent markers by immunofluorescence staining

Human iPSCs control and MDS27 clones (C1, C8, C11 and C22) were grown on coverslips coated with matrigel. **(A)** Cells were stained with TRA1-81 (Green). The cells were stained with mouse Alexa Flour 488 IgM and fixed before being counterstained with DAPI (Blue). **(B)** Cells were fixed, permeabilised and then stained with Nanog (Green) and SOX2 (Red). The cells were then stained with anti-goat Alexa 488 and anti-mouse Alexa 633 antibodies before being counterstained with DAPI (Blue). 100 µm scale bar, 20x magnification, Leica DM6000. Data represent 4 independent experiments.

Lastly, the multipotent capability of MDS27-hiPSC lines was assessed by inducing their differentiation into the three germ layers (ectoderm, mesoderm, and endoderm), a requirement for them to be considered as a truly pluripotent stem cell. hiPSC lines (C1, C8, C11 and C22) as well as the hiPSC control line were differentiated to the three germ layers using the STEMdiff Trilineage differentiation kit (Stem cell technology) following the protocol described in Figure 3.9.

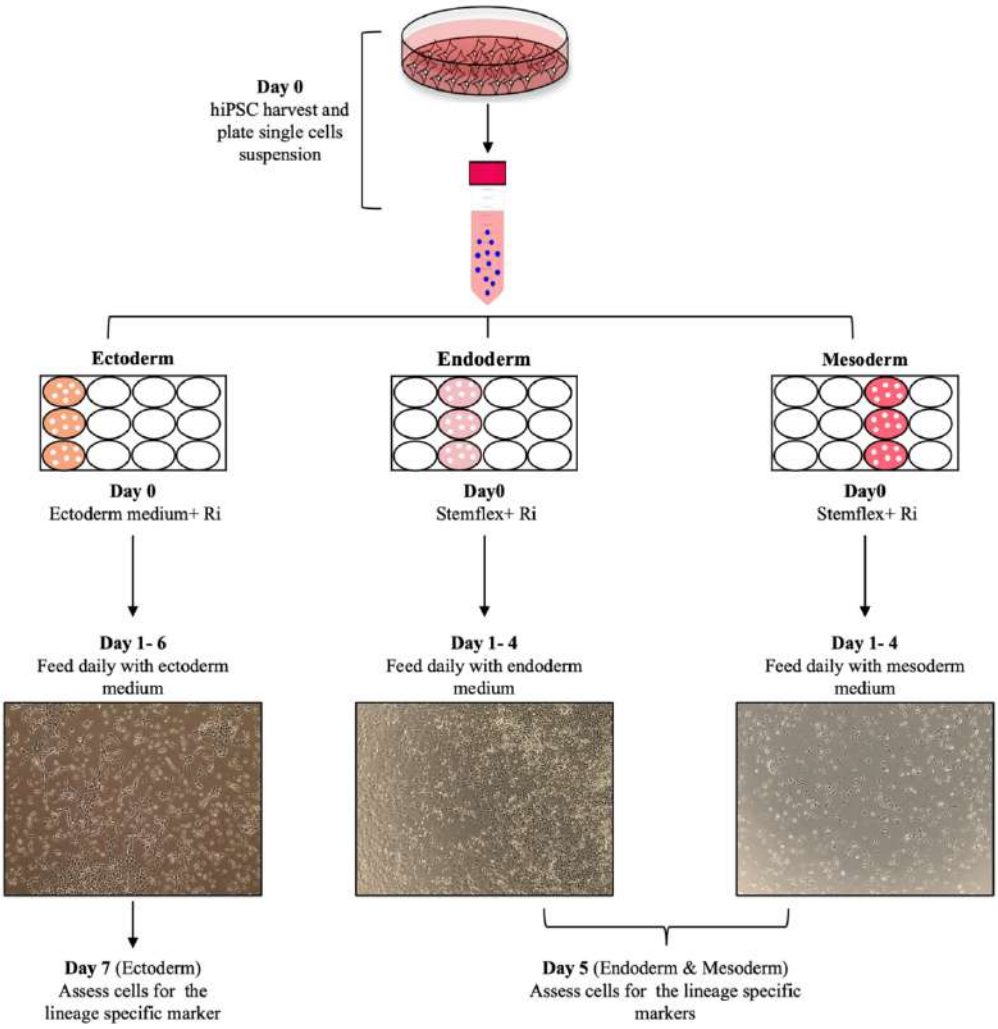


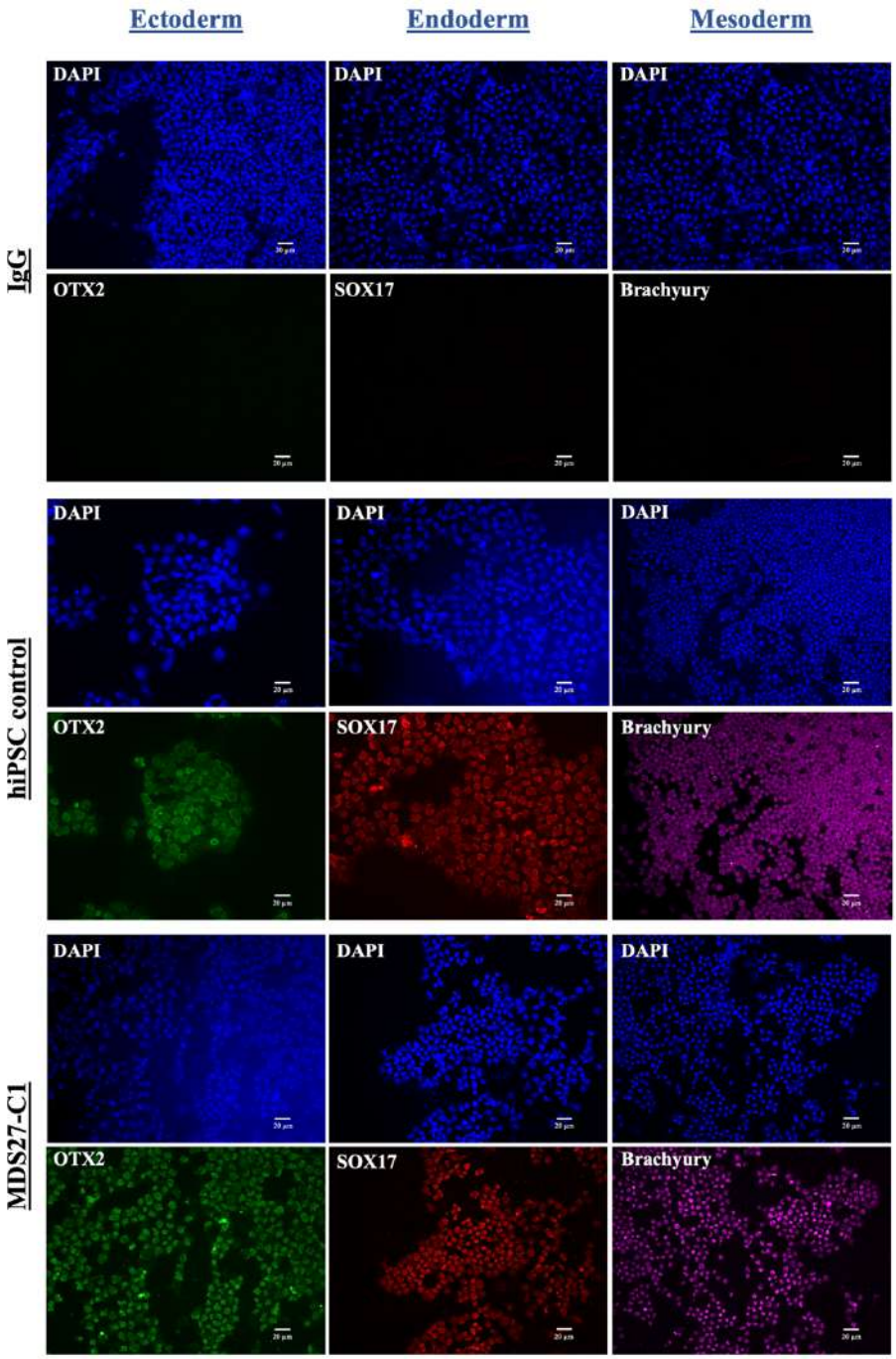
Figure 3.9: Trilineage differentiation of hiPSCs.

Schematic representation of the method of trilineage differentiation of hiPSCs lines using the STEMdiff Trilineage differentiation kit (Stem cell technology).

Then, the lineage commitment was evaluated by performing immunofluorescence staining and imaging for detection of OTX2 expression, a marker of ectoderm differentiation (Acampora et al., 1998); SOX17 expression, a marker of endoderm differentiation (D'Amour et al., 2005, Kanai-Azuma et al., 2002); and Brachyury expression, a marker for evaluating mesoderm differentiation (Lam et al., 2014). As shown in Figure 3.10, the hiPSC positive control and MDS27-iPSC clones all expressed OTX2, SOX17 and Brachyury when differentiated towards

ectoderm, endoderm, and mesoderm, respectively. Thus, this result indicates that the iPSC generated were capable of successfully differentiating towards the three germ layers.

Overall, we can infer that we successfully generated hiPSC clones from a patient with MDS (MDS27) with the same karyotype and mutations as the original sample and that these clones exhibited characteristics of pluripotent stem cells in all aspects: as these iPSCs are morphologically undifferentiated, are positive for AP, express the pluripotent markers and are able to differentiate into the three germ lineages.


Ectoderm
Endoderm
Mesoderm
IgG
hiPSC control
MDS27-C1
DAPI
OTX2
SOX17
Brachyury

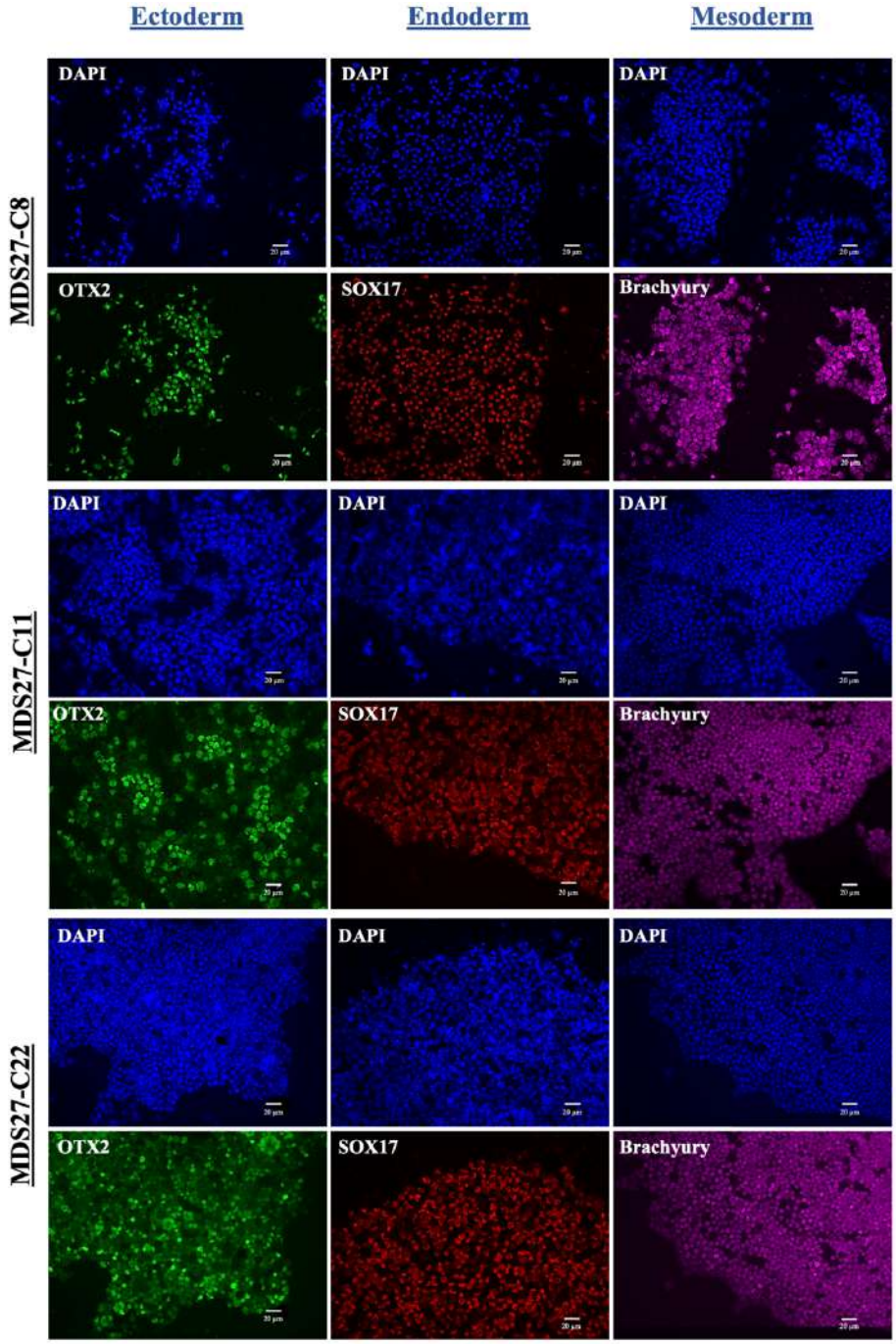


Figure 3.10: Positive expression of trilineage markers in all hiPSC lines

Cells of each germ line were cytospun at 300g for 8 minutes onto a glass slide. The cells were fixed, permeabilised and stained with OTX2 (Green), SOX17 (Read) and Brachyury (Purple). The cells were then stained with anti-goat Alexa 488 but the colour of SOX17 and Brachyury was modified to Red and Purple respectively using ImageJ software. 20µm scale bar, 20x magnification, Leica DM6000. Data represent 4 independent experiments.

3.3 Discussion

3.3.1 The successful reprogramming of primary cells from MDS27 patient

The generation of hiPSC from MDS patients enables unique opportunities to model and investigate MDS disease and the leukemic transformation of MDS. Patient MDS27 was chosen for presenting a normal karyotype, for having progressed to a more aggressive form of the disease and for failing to respond to drug treatment. Thus, making this patient sample the ideal subject to study the mechanism leading to progression and failure to treatment response.

In this study, non- integrating methods (SeV, Episomal) carrying Yamanaka's factors were used to infect/transfect the mononuclear cells from the PB of patient MDS27 and induce the formation of hiPSC. These two methods were considered in this study in order to have highly efficient reprogramming, free genome integration, and fewer genome rearrangements in the iPSC line.

Several studies have proven that SeV and Episomal methods are the most efficient non-integrating methods: SeV has the advantage of being a single RNA virus that does not enter the nucleus and is able to generate large amounts of protein to reprogram the somatic cells without the risk of exogenous integration. On the other hand, the episomal method is a mixture of plasmids that would allow for the generation of a footprint-free iPSC line (Kang et al., 2015, Hubscher et al., 2019, Schlaeger et al., 2015).

The formation of hiPSC was successful at the early stage of the disease (sample collected in 2013) but was not successful at the late stage (samples collected in 2015 and 2016). To overcome this problem, we enhanced the efficiency of the reprogramming by using 0.5 mM VPA at an early stage of reprogramming. VPA is widely used in reprogramming studies to

eliminate the original epigenetic memory of cells and plays an essential role in the upregulation of ESC specific genes (Kotini et al., 2015, Duan et al., 2019).

Unfortunately, even after incorporating VPA treatment in our reprogramming protocol, the cells from MDS27 from 2015 and 2016 were not capable of generating any colonies. The reason for that could be the difficulties to reprogram cells at an advanced stage of the disease, as at this time the patient had progressed to high-risk MDS and then to sAML. Indeed, less than a handful laboratories have been able to generate iPSC from patients with high-risk MDS and AML. For example, the group lead by Martin Bonamino, has faced difficulties to obtain iPSC from AML patients and were only able to generate one hiPSC clone with an *ASLX1* mutation (Gomez Limia et al., 2017). Also, another group tried to generate iPSC from 16 different AML patients, but only one AML patient out of 16 was able to give rise to hiPSC originating from transformed AML cells (Lee et al., 2017a). The reason for the failure to generate hiPSC from the late stage is not clear. One possible explanation is that it might be difficult to reprogram patient samples containing high percentage of blast cells harbouring mutations in epigenetic modulators, as epigenetic changes are key for the reprogramming process.

We could have attempted to generate iPSCs from MDS27 when they had progressed to high-risk MDS and sAML using different reprogramming methods, for example, episomal reprogramming or using alternative reprogramming factor cocktails, unfortunately after the several attempts, with/without VPA, we run out of cell samples after the disease progression.

3.3.2 Human iPSC lines from MDS27 patient harbour the same mutations as disease cells

Detailed genetic mutation analysis for the derived hiPSC lines allowed us to identify the efficiency of the reprogramming process. Our results showed that the clonal representation of cells in the hiPSC lines were skewed in favour of clones harbouring multi-mutations (*ASLX1, RUNX1, SRSF2*) over healthy clones. Contrary to expectations, both integrating-free methods used failed to capture any healthy isogenic clone or clones with different mutational composition. This finding is contrary to a previous study which reported that reprogramming using SeV is able mainly to capture the healthy cells from MDS during the reprogramming process whilst the episomal reprogramming has a higher efficiency of reprogramming disease subclones within MDS cells (Hsu et al., 2019). Furthermore, other studies have shown the possibility to capture various clones from normal stage to the full transformed status when samples from MDS patients were reprogrammed using an integrated method via a lentiviral vector (Kotini et al., 2015, Chang et al., 2018, Kotini et al., 2017).

A possible explanation for not being able to obtain an isogenic control in our model might be due to the long storage of the PBMCs (stored for more than 4 years at -80^{o}C, sample taken in 2013); the primary cells are extremely sensitive to changes in their environment, and thus it could be possible that the healthy cells did not survive for long at -80^{o}C. Another explanation could be the difference of starting material between our work and those from others. Most studies have been done using the pure population of hemato-endothelial cells (CD34+) (Kotini et al., 2015, Chang et al., 2018, Kotini et al., 2017), whilst our reprogramming studies were performed using PBMCs. It has been shown that using less differentiated cells (progenitors) could maximize reprogramming efficiency. This may be due to the progenitor cells having less

condensed chromatin in specific regions, which is more accessible to reprogramming factors (Eminli et al., 2009).

Furthermore, the reason for lacking other clonal abnormalities is not apparent, but it may be due to the peculiarities of the combinatorial mutations of MDS27: *SRSF2*, *RUNX1* and *ASLX1*. *SRSF2* is a splicing factor that is key for the regulation of gene expression. It has been shown that *SRSF2* mutations are present in 2–9% of MDS as an only event. *SRSF2* tends to be mutated early on during clonal hematopoiesis of indeterminate potential (CHIP) or during the progression to MDS in association with other mutations such as *ASLX1* and *RUNX1*. *RUNX1* is a transcription factor that has a critical role in hematopoietic development, and *RUNX1* mutation can be an early or a late event in MDS. However, ASLX1 is a chromatin-binding protein involved in the epigenetic regulation of hematopoietic genes. *ASLX1* mutation occurs mainly in MDS, and is among the earliest events in the process of AML transformation (Zheng et al., 2017, Papaemmanuil et al., 2013, Sperling et al., 2017, Fu and Maniatis, 1990, Jeromin et al., 2015, Hoischen et al., 2011).

Thus, each of these genetic mutations alone could have been present at a low frequency before the patient was diagnosed in 2013, and a second/third mutational hit could be responsible for the disease diagnosed in 2013. Thus, cells with single mutations could be at that point inexistent or too rare as to be detected directly by reprogramming.

3.3.3 Human iPSC lines from MDS27 patient exhibit normal karyotype

Despite the fact that somatic reprogramming has been reported to cause chromosomal instability (Taylor et al., 2014), the generation of hiPSC from MDS27 did not alter the number of the chromosomes, and each cell line had a normal karyotype. However, global karyotyping to determine changes in chromosome number is not enough to detect the genomic alteration

within each chromosome. Microarray comparative genomic hybridization analysis has a greater analytical sensitivity to detect the genomic imbalances, and it has been widely used to diagnose haematological malignancy in clinics (Peterson et al., 2015). Thus, the possibility of performing a microarray with MDS27-hiPSC lines would have been the ideal cytogenetic result.

3.3.4 Human iPSC lines from MDS27 patient are pluripotent

The pluripotency characterisation of hiPSC lines generated from patient MDS27 and the hiPSC control have proven that the hiPSCs generated were similar to hESCs in many aspects including, morphology, surface markers and ability to differentiate *in vitro* to the three germ layers.

Human iPSC colonies were typically tightly packed and flat, with large nuclei and scant cytoplasm (Takahashi et al., 2007). Each hiPSC line displayed high expression of AP staining, which is the traditional marker of pluripotency in both mouse and human ESCs (Brons et al., 2007, Tesar et al., 2007).

Besides, the exogenous expression of Yamanaka's factors activates the endogenous expression of pluripotent markers (TRA1-81 and NANOG) which indicated that MDS27 cells were efficiently reprogrammed.

These pluripotent genes have been shown to play a critical role in the maintenance of self-renewal and pluripotency of mouse and human ESCs. In particular, the transcription factors OCT4 and SOX2, collaborate to regulate the expression of other transcription factors such as NANOG, and the three together regulate a network of genes with a critical role in pluripotency and maintenance of the undifferentiated state of ESCs and iPSCs (Fong et al., 2008, Abujarour et al., 2013).

Moreover, hiPSC lines generated from MDS27 were able to differentiate to cells from the three germ layers: ectoderm, mesoderm, and endoderm *in vitro* as under the correct stimuli, they expressed specific markers for each germline.

Overall the characterisation of MDS27-hiPSCs has proven the feasibility of generating MDS-derived hiPSC, as they meet the gold standards to classify them as truly pluripotent: morphology of the iPCS colonies, expression of pluripotent protein markers and trilineage differentiation capacity (Hsu et al., 2019, Kotini et al., 2015, Takahashi et al., 2007, Castano et al., 2017). The successful generation of hiPSCs from patients with MDS places us at the same level as a handful of labs worldwide (Hsu et al., 2019, Kotini et al., 2015, Takahashi et al., 2007), providing us with a model system to explore the mechanisms leading to disease progression and drug resistance.

Chapter 4 : Validation the disease phenotype of MDS27

4.1 Introduction

In the previous chapter, we showed the successful generation of several hiPSC clones from patient MDS27 harbouring three somatic mutations (*ASLX1*, *RUNX1* and *SRSF2*) that were present in the original MDS27 patient sample (2013). The clones generated were confirmed to be pluripotent and able to differentiate towards the three germ layers but still it was imperative to determine whether they could be successfully differentiated into HSPCs, and whether after differentiation towards mature cells, they would phenocopy the disease phenotypes observed in the patient, such as anaemia.

These questions will be addressed in this fourth chapter by completing the following objectives:

- Differentiating hiPSC control and MDS27 hiPSC clones to HSPCs.
- Defining the differentiation potential of HSPCs obtained from control and MDS27 by performing colony assays.
- Assessing the erythroid lineage differentiation of HSPCs from control and MDS27 iPSCs in liquid culture by performing flow cytometry analysis and morphological characterisation.

4.2 Results

4.2.1 Differentiation of hiPSC control and hiPSC from MDS27 patient to hematopoietic stem and progenitor cells (HSPCs)

To assess the hematopoietic differentiation potential of hiPSC control and the MDS27-iPSC clones to HSPCs, we used the STEMdiff Hematopoietic protocol from STEM Cell Technology (Figure 4.1). When the hiPSC reached 80% confluency, around 16 to 20 fragments of hiPSC clones with a specific size of approximately 100 μm were seeded, which is the crucial step for both viability and successful differentiation. Medium A was used to feed the cells in the first stage of the differentiation to induce mesoderm differentiation. Then, in the next step, medium B was used to generate a transition wave of endothelial cells to hematopoietic cells differentiation. The successful differentiation could be observed by the morphological changes within the colony and the appearance of the single floating cells (HSPCs). The course of the differentiation was monitored using CD34 (a hemato-endothelial marker), CD43 (an early haematopoietic marker which persists in differentiating precursor cells) (Vodyanik et al., 2006)), and CD45 (the key marker of definitive hematopoietic cells) (Sturgeon et al., 2013). Expression of these markers was assessed by flow cytometric analysis at different days (Day7, day10, day12 and day 14) to determine the dynamics of haematopoietic differentiation. The gating strategy can be found in supplementary figure 3. Three different iPSC clones derived from the MDS27 patient sample (C8, C11 and C22) were investigated to control for possible effects during the reprogramming process.

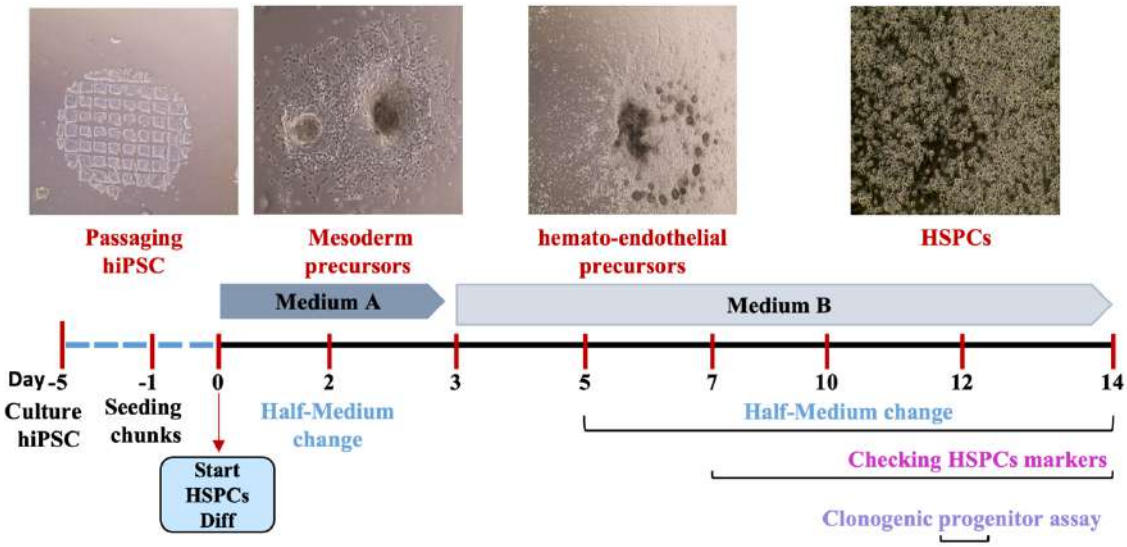


Figure 4.1: Human iPSCs differentiation to HSPCs

Schematic representation of the HSPC differentiation using hematopoietic STEMdiff protocol from stem cell technology. First, the specific size and number of hiPSC chunks were seeded before starting the differentiation. Then, hiPSCs were fed with medium A from day 0 to day 3 to induce the mesoderm precursors. Then, medium B was added on day 3 to induce the hemato-endothelial precursors and HSPCs. The HSPCs (floating cells) were collected on day 7, 10, 12 and 14 to check the hematopoietic markers. On day 12, HSPCs were collected for colony assay. HSPCs (hematopoietic stem and progenitor cells).

During the time course of the hematopoietic differentiation, both hiPSC control and MDS27 hiPSC clones behaved very similarly. The population of hematopoietic cells expressing $CD43^+$ was around 40% in all cases by day 7 and increased at day 12 and day 14. By day 14, typically, around 90% of the cells were $CD43^+$ (Figure 4.2, A).

According to the expression of CD45, a marker for mature hemopoiesis, the percentage of cells co-expressing $CD34^+$ and $CD45^+$ was quite low at day 7 (between 4-7%) and steadily increased during differentiation, with 25-35% of the cells co-expressing both markers by day 14. **(**Figure 4.3**,** A). The statistical analysis in Figure 4.2, (B) and Figure 4.3 (B).

showed no significant difference in the expression of hematopoietic markers between hiPSC control and MDS27 hiPSC clones. In addition, all hiPSC clones (C8, C11, C22) behaved similarly in the differentiation toward the HSPCs. The number of HSPCs after 14 days of the differentiation is presented in supplementary figure 4.

In summary, this result demonstrates that the HSPC differentiation protocol utilized successfully differentiated the hiPSC control and MDS27 hiPSC clones to both early and late hematopoietic cells. Also, these results indicate that the *ASLX1*, *RUNX1* and *SRSF2* mutations do not affect the generation of HSPCs ($CD34^+$ $CD43^+$ & $CD34^+$ $CD45^+$) of MDS27 hiPSC clones.

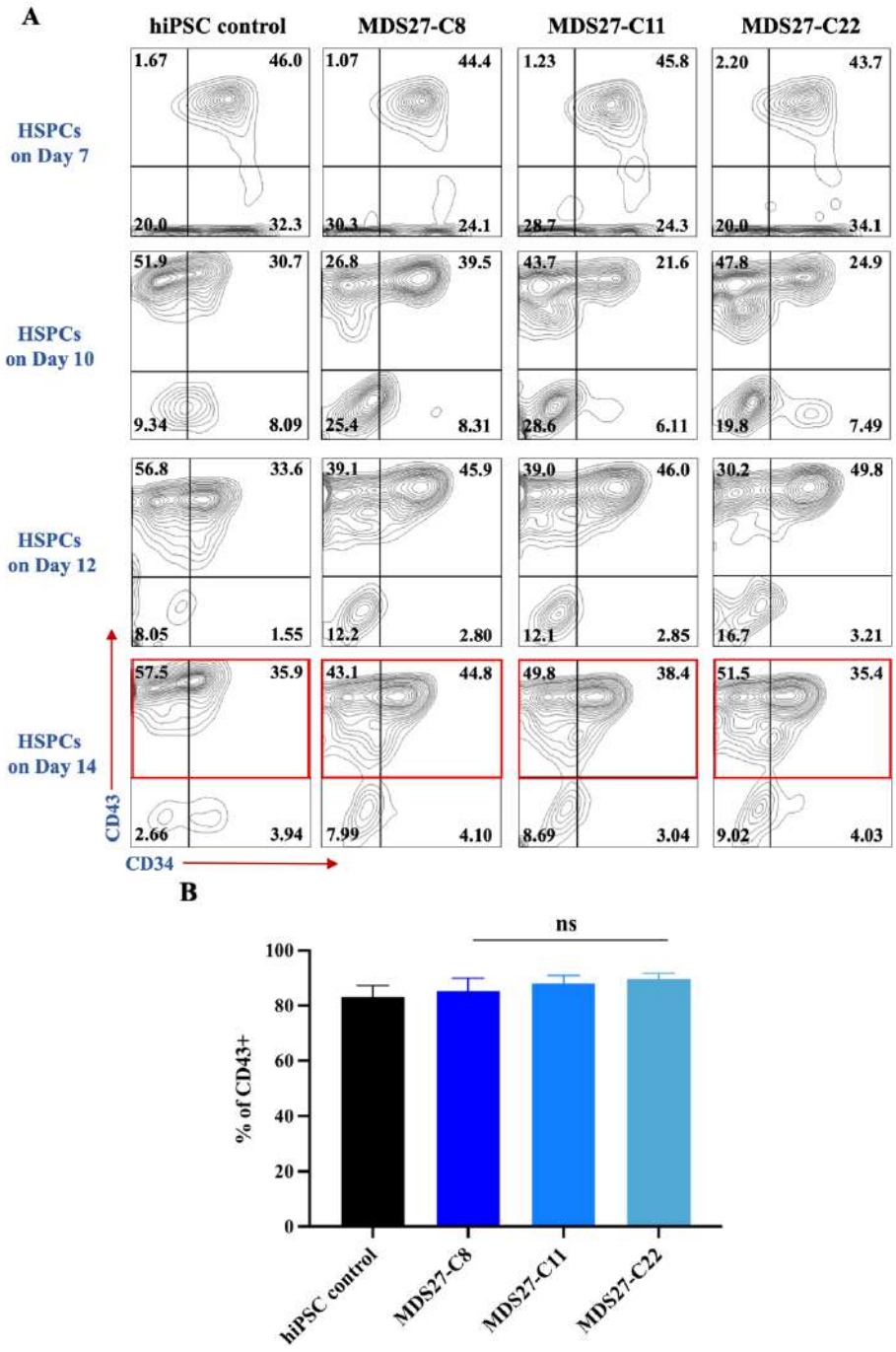


Figure 4.2: Human iPSCs differentiate to early HSPCs

(A) $CD34^+$ $CD43^+$ (early hematopoietic population) were monitored during the time course of differentiation on Day 7, 10, 12 and 14. The differentiating cells were analysed by flow cytometry and the cells were gated first based on FSC and SS of unstained cells and then analysed for hematopoietic markers based on the isotype control. **(B)** The figure shows the mean percentage of the hematopoietic markers $CD43^+$ on day 14. Statistical results are presented as mean ± standard error of the mean (SEM) and analysed using One-way ANOVA with multiple comparisons. p values were represented as ns for not significant. Data represent 4 independent experiments.

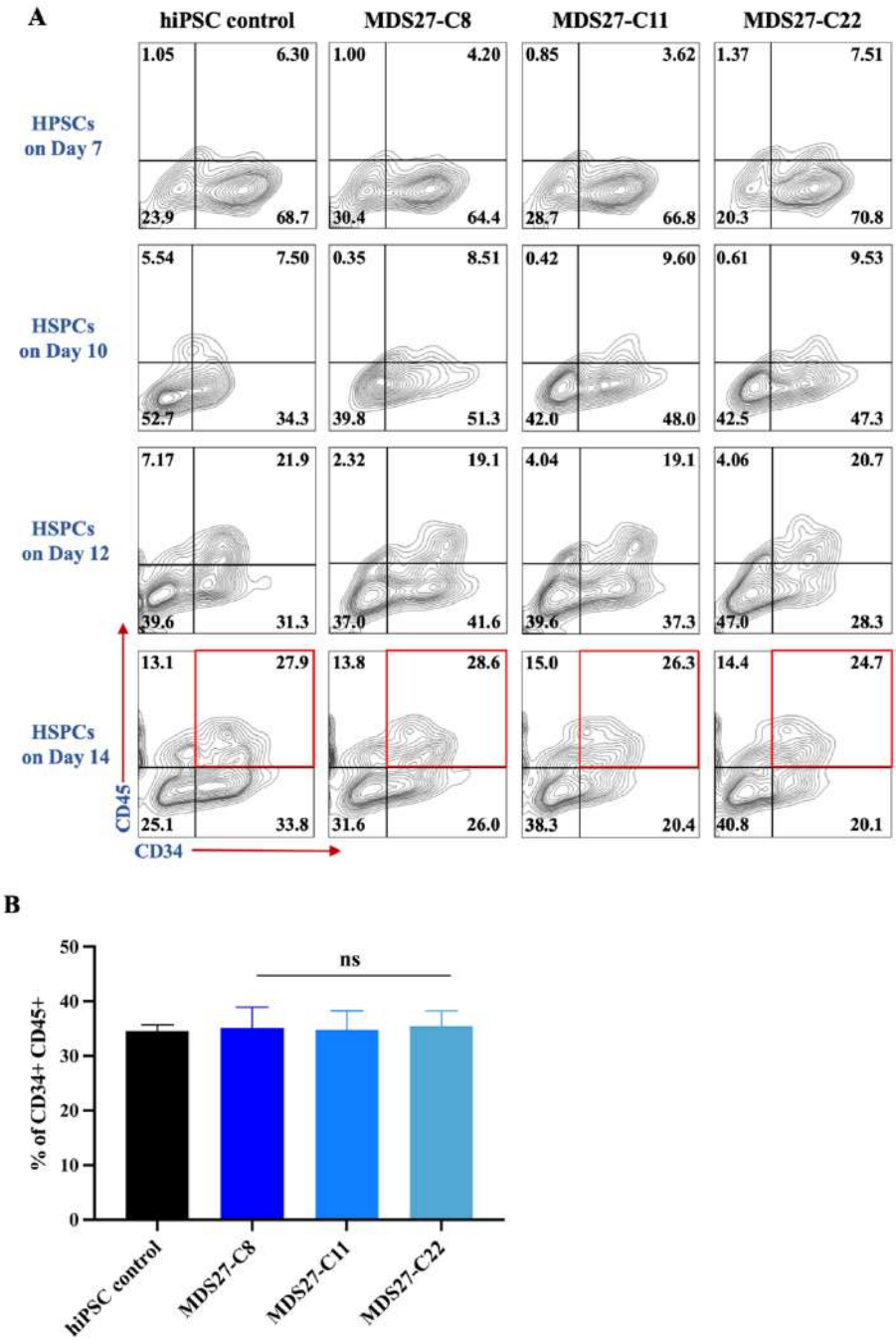


Figure 4.3: Human iPSCs differentiate to late HSPCs

(A) $CD34^+$ $CD45^+$ (late hematopoietic population) were monitored during the time course of differentiation on Day 10, 12 and 14. The differentiating cells were analysed by flow cytometry and the cells were gated first based on FSC and SS of unstained cells and then analysed for hematopoietic markers based on the isotype control. **(B)** Figure shows the mean percentage of hematopoietic markers ($CD34^+$ $CD45^+$) on day 14. Statistical results are presented as mean ± standard error of the mean (SEM) and analysed using One-way ANOVA with multiple comparisons. p values were represented a ns for not significant. Data represent 4 independent experiments.

4.2.2 Assessing the differentiation potential of HSPC from hiPSC control and MDS27-hiPSC in semi-solid medium

Once corroborated that the iPSC clones derived from the MDS27 patient were able to differentiate towards HSPCs, it was required to determine whether the mutations that these clones harbour (*SRSF2*, *ASXL1* and *RUNX1*) could affect the differentiation potential of the HSPCs. Thus, HSPCs from hiPSC control and MDS27 hiPSC clones were obtained at day 12 and plated in methylcellulose medium enriched with recombinant cytokines that induce the differentiation into mixed lineage (CFU-GEMM), committed erythroid (BFU-E), and myeloid lineage progenitors (CFU-G, CFU-M, CFU-GM) (Figure 4.4, A)**.** Two different HSPC cell numbers (10,000, and 20,000) were used to evaluate the hematopoietic differentiation potential and clonogenic capacity. Fourteen days after, the number of colonies were counted and scored based on their phenotypic characteristics. Quite consistently throughout the independent experiments, it was observed that the total number of colonies was lower in all the HSPCs derived from MDS27-hiPSC clones compared to the hiPSC control. This reduction in the number of myeloid and erythroid CFUs indicated an impaired hematopoietic colony-forming capacity of the MDS27-hiPSC clones (Figure 4.4, B).

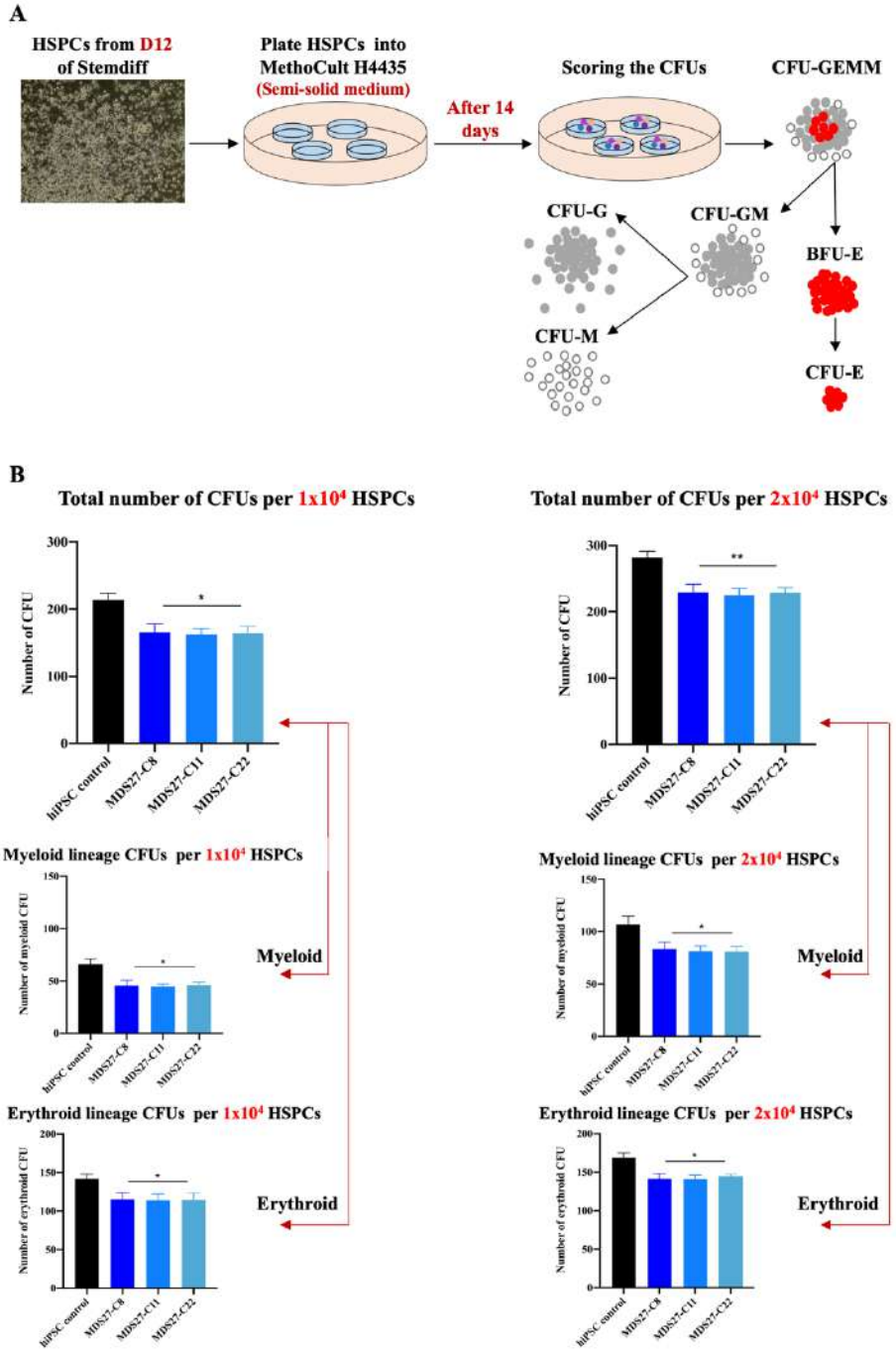


Figure 4.4: Differentiation potential of HSPCs in semi-solid medium

(A) Schematic representation of the haematopoietic potential of day 12 HSPCs (floating cells) was evaluated in colony forming assays. **(B)** After 14 days in semi solid medium, the total number of CFUs was scored and then from the total number of CFUs, the number of myeloid lineage CFUs and erythroid lineage CFUs were identified. Statistical results are presented as mean ± SEM and analysed using One-way ANOVA with multiple comparisons. p value for total CFU number * p 0.0216 (C8), 0.0148 (C11), 0.0192 (C22), ** p 0.0094 (C8), 0.0056 (C11), 0.0091 (C22). p Value for myeloid CFUs number * p 0.0138 (C8), 0.0102 (C11), 0.0160 (C22). p Value for Erythroid CFUs number * p 0.0116 (C8), 0.0103 (C11), 0.0244 (C22). Data represent 4 independent experiments.

In addition, all the clones derived from the MDS27 patient sample were able to generate all the types of CFUs (Figure 4.5, A), but the percentage of each type was lower in MDS27 clones compared with hiPSC control (Figure 4.5, B).

Despite the decrease in hematopoietic colony number in MDS27-hiPSC clones, no difference in the morphology and size of the colonies formed were observed in MDS27-hiPSC clones compared to hiPSC control, with a bias towards the formation of erythroid colonies versus myeloid colonies. Moreover, all the clones derived from the MDS27 patient sample behaved similarly (Figure 4.6).

Taken together, these results suggest that the HSPCs derived from the MDS27 patient (low-risk MDS) display a lower differentiation potential, which is in accordance with the data reported from MDS patients (Kotini et al., 2015, Kotini et al., 2017, Hsu et al., 2019).

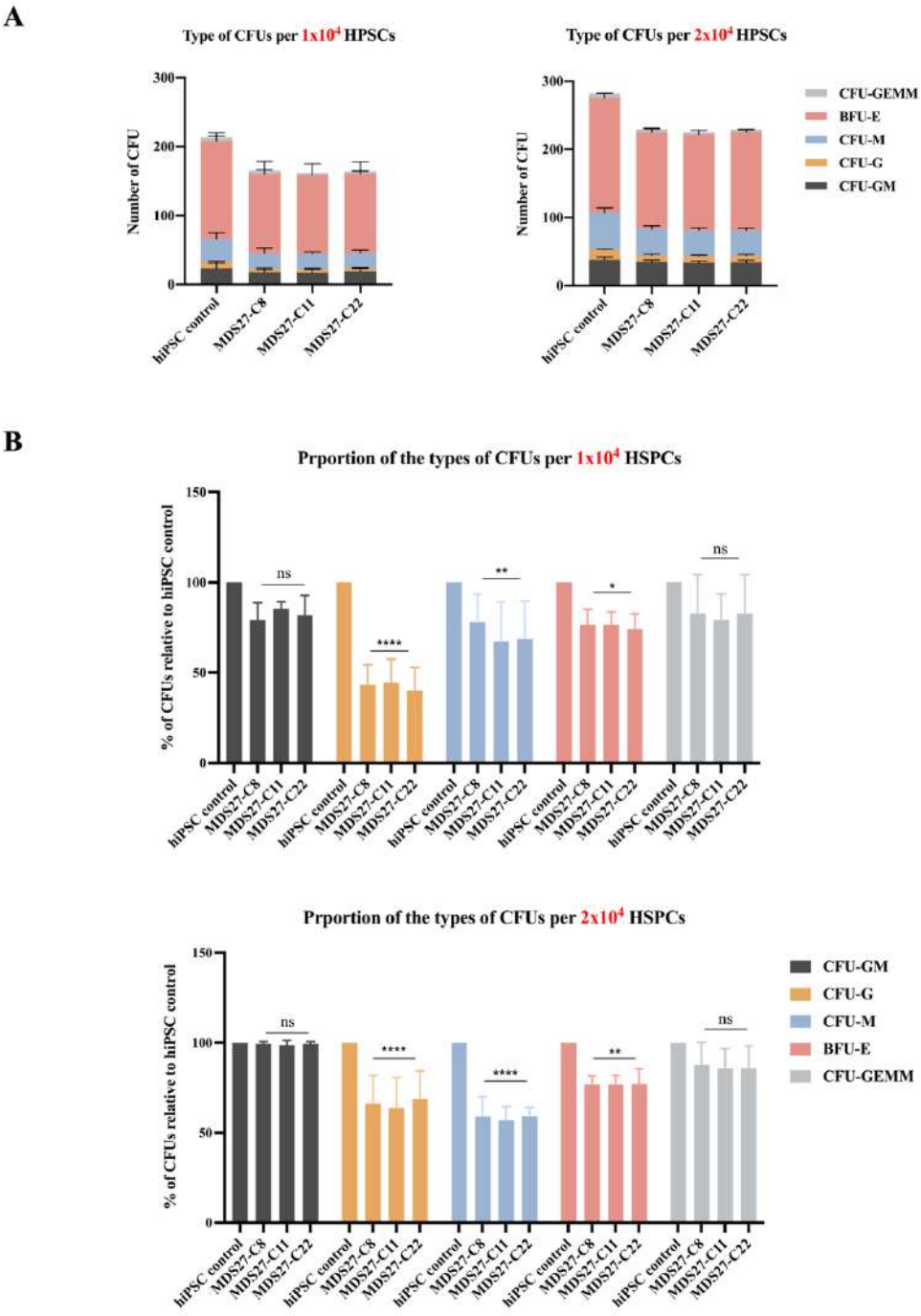


Figure 4.5: Type of CFUs

(A) Figure shows the mean number of the type of CFUs. A CFU-GEMM colony contains granulocytes, macrophages, erythrocytes, and megakaryocytes. A CFU-GM colony contains granulocytes and macrophages, a CFU-M colony contains only macrophages and BFU-E colony contains erythrocytes. Error bars represent the standard error of the mean (SEM). **(B)** Figure shows the proportion or relative percentage of each type of CFUs for $1x10^4$ and $2x10^4$ of HSPCs after 14 days in semisolid media. Statistical results are presented as mean ± SEM and analysed using Two-way ANOVA with multiple comparisons, *** $p < 0.0001$, ** p 0.002 and * p 0.04. Data represent 4 independent experiments.

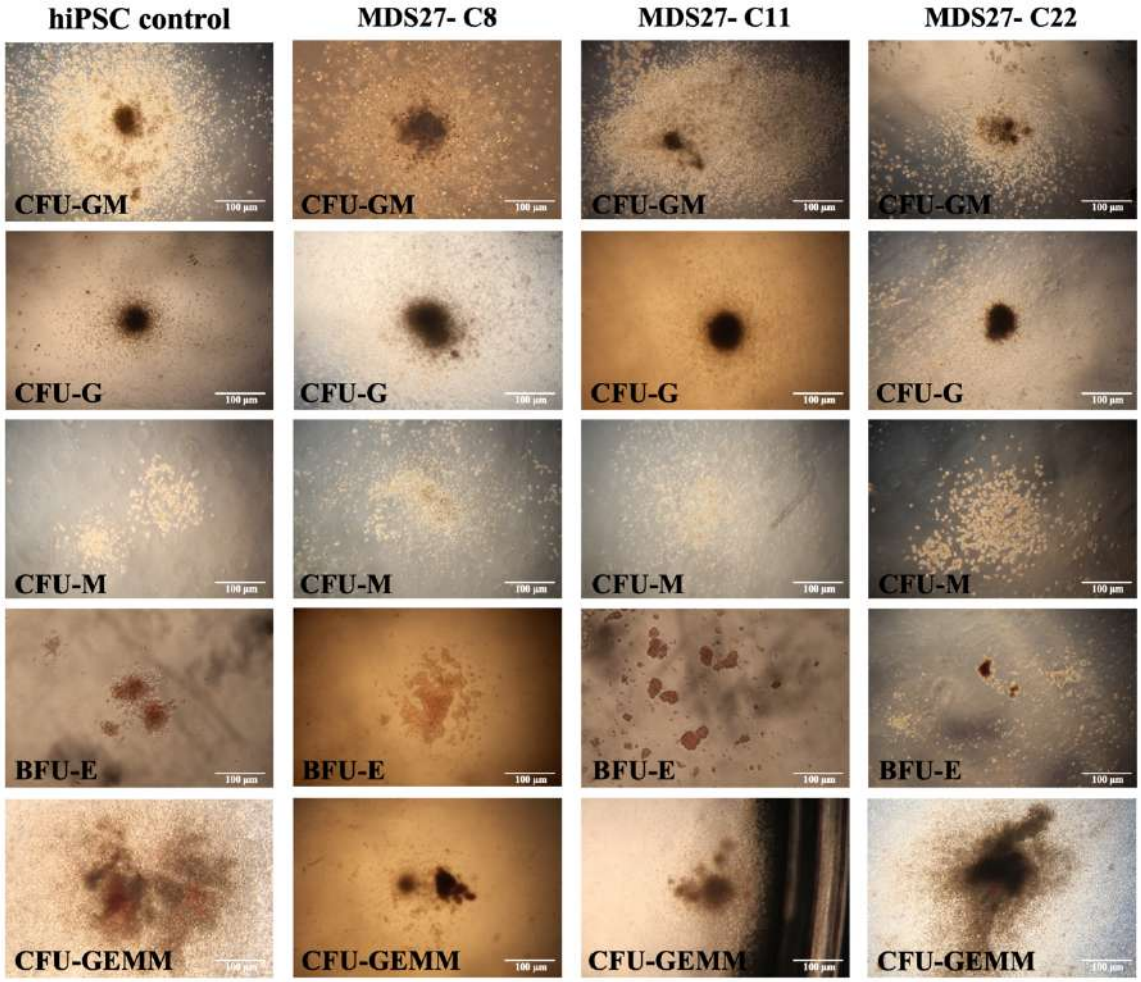


Figure 4.6: Morphology of CFUs

Phase contrast pictures show the morphology of CFUs for hiPSC control and MDS27 clones after 14 days in semi-solid medium. The pictures were taken by primo vert microscope (ZEISS) with Cannon camera at 4x magnification and size bar 100 µm scale bar . Data represent 4 independent experiments.

4.2.3 Studying the erythroid lineage of MDS27-hiPSC

According to the previous result, the MDS27-hiPSC clones display a deficiency in forming CFUs of myeloid lineage and erythroid lineage. Because anaemia is one of the common features of MDS patients, we sought to determine whether erythropoiesis might be affected in the clones derived from the low-risk MDS27 patient. Thus, we decided to evaluate the complete erythroid differentiation potential of HSPCs in liquid culture to investigate whether the decrease in

erythroid CFUs was due to the insufficient maturation of erythroid cells or dysplastic features that affects the quality of the cells.

We hypothesised that during the erythroid differentiation or maturation, the cells are blocked during maturation, which leads to a decrease in the number of the mature cells. For this purpose, HSPCs from hiPSC control and MDS27-hiPSC (C22) from day 10 were obtained and cultured in an established 3-phase erythropoiesis liquid culture over 18 days (Figure 4.7). The erythroid culture was monitored every 3-4 days, and the erythroid distribution over different maturation stages was assessed by the expression of CD71 (progenitor marker) and CD235a (Glycophorin A, mature marker). $CD71^{+}/CD235a^{-}$ defines erythrocyte progenitors, $CD71^{+}/CD235a^{+}$ defines erythroblast cells and $CD71^{low/-}$ $CD235a^{+}$ represents a more differentiated erythroid population or mature erythrocytes (Wangen et al., 2014). The gating strategy of erythroid cells is presented in supplementary figure 5.

Figure 4.7: Erythroid differentiation of HSPCs

Schematic representation of the experimental set up for the erythroid differentiation of HSPCs.

Depicted in Figure 4.8, (A)**,** the red arrows indicate the differentiation progression stages from progenitor erythroid cells ($CD71^+$) to a pure mature population ($CD235a^+$). After culturing HSPCs in erythroid media for 7 days, around 15% of cells were identified as $CD71^+$ (erythroid progenitor) and about 40% of cells co-expressed $CD71^+$ and $CD235^+$ (erythroblasts), while the expression of the mature marker ($CD235^+$) was still low at the early stage of the differentiation. This cellular distribution was very similar for both hiPSC control and MDS27-C22 hiPSC. Then, as differentiation continued, the percentage of erythroblasts decreased with a concomitant increase in the percentage of the mature cells. For example, on day 11 and day 14, the percentage of erythroblasts was around 20% and 10%, whilst the percentage of mature cells was approximately 35% and 50%, respectively, for both hiPSC control and MDS27-C22 hiPSCs. On the final day of erythroid differentiation (day 18), the erythroid lineage cells became more mature, with over 55% of the cells expressing $CD235a^+$ and only 5% at the erythroblast stage ($CD71^+$ $CD235a^+$) **(**Figure 4.8, A).

Curiously, analysis of these surface markers by flow cytometry did not detect any immunophenotypical differences in MDS27 clones (Figure 4.8, B)**.** As MDS patients, including patient MDS27, suffer from anaemia we decided to check the morphology of erythroid cells derived from hiPSC control and MDS27 clones to understand whether dyserythropoiesis was observed in our MDS27 iPSC clones or the clones generated did not recapitulate the disease phenotype.

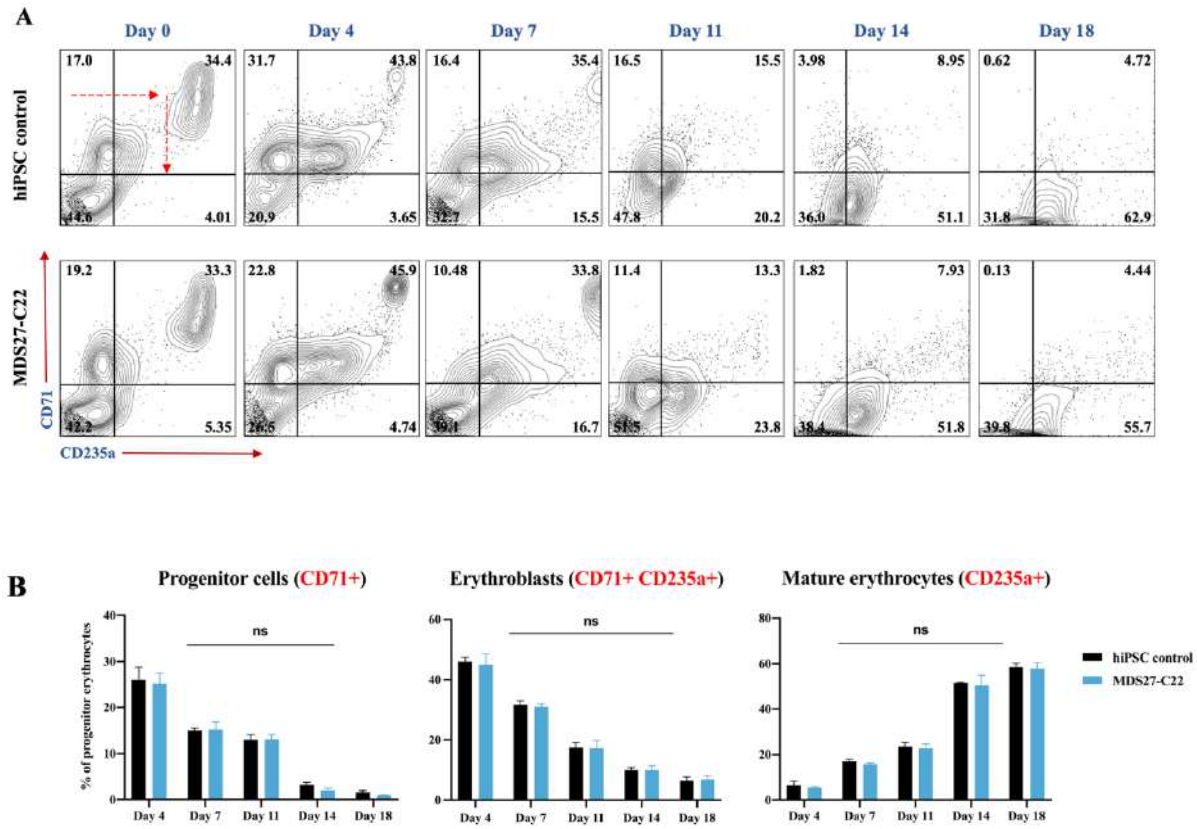


Figure 4.8: Assessment of erythroid differentiation of HSPCs

(A) CD71 and CD235a erythroid markers were monitored during the time course of differentiation on Day 0, 4, 7, 11, 14 and 18. The differentiating cells were analysed by flow cytometry and the cells were gated first based on FSC and SSC and then were analysed for erythroid markers based on the isotype control. The gating strategey is in supplemntry (figure 4) **(B)** Shows the mean percentage of erythrocyte ($CD71^+$), erythroblasts ($CD71^+$ $CD235a^+$) and mature erythrocytes ($CD235a^+$) during several time point of erythroid differentiation. Statistical results are presented as mean ± SEM and analysed using Two-way ANOVA with multiple comparisons. p values were represented a ns for not significant. Data represent 4 independent experiments.

Diff-quick staining was used to assess erythroid morphology at different time points along the differentiation process, on days 4, 7, 11, 14 and 18. Their distinguishable cytoplasm and nucleus can aid the recognition of changes in the morphology during the differentiation; a gradual accumulation of haemoglobin, progressive reduction in the cell size and nuclear condensation occurs during the second phase of the differentiation, which culminates with the expelling of the nuclei from orthochromatic erythroblasts to produce reticulocytes or mature cells (Granick and Levere, 1964, Gregory and Eaves, 1977).

The microscopic evaluation confirmed that hiPSC control and MDS27-C22 hiPSCs had given rise to a heterogeneous population of erythroid cells. All the stages of maturation appeared during the several days of the differentiation. The maturation stages are indicated according to the colour of the different arrows in the micrographs of Figure 4.9 (A): Proerythroblast (red arrow), Basophilic erythroblast (blue arrow), Polychromatic erythroblast (green arrow), Orthochromatic erythroblast (orange arrow) and Mature cells (yellow arrow). Nucleated erythroblasts were observed in the cultures from day 4 in both hiPSC control and MDS27-hiPSC clones. These cells were identified by their large size (10-15μm) and high nuclear to cytoplasmic ratio. A round nucleus occupies almost three-quarters of the cell that stains violet-blue and is surrounded by a narrow ring-shaped cytoplasm that stains light purple colour. Also, few enucleated erythroid cells (mature cells) were observed as early as day 7, according to their size (7-10μm) and pale pink colour. The enucleated cells increased towards day 18 in hiPSC control and MDS27-C22 hiPSC, parallel to a size reduction (7-8μm). However, many debris, dead cells and cells with damaged membranes were observed on the slides of hiPSC control and MDS27-C22 hiPSCs. This showed that the membrane stability was poor at the late stage of erythroid differentiation.

Despite the ability of MDS27-C22 hiPSCs to reach the mature stage, the morphological analysis revealed that the erythroid differentiation of MDS27-C22 hiPSCs was dysplastic as cells presented aberrant morphology (Figure 4.9, B). Examples of these atypical morphology are binucleated, multinucleated and nuclear bridges which are the common dysplastic features of MDS patients (Hasserjian et al., 2017). The aberrant morphology within 500 cells was scored, and the percentage of cells presenting aberrant morphology was calculated. However, few cells with hyper-vacuolated cytoplasm were scored as aberrant cells in hiPSC control. By calculating the percentage of the aberrant morphology of erythroid cells derived from MDS27-C22 hiPSCs

to the aberrant morphology of hiPSC control at the different time points (Day 4, 7, 11, 14 and 18), it was evident the increase in cells with aberrant morphology in the erythroid cultures from MDS27-C22 iPSCs compared to control hiPSC (Figure 4.9, C). These data indicate that *SRSF2*, *ASLX1* and *RUNX1* mutations promote an aberrant maturation of the erythroid cells in low-risk MDS, and importantly, that we could reproduce the disease phenotype *in vitro*.

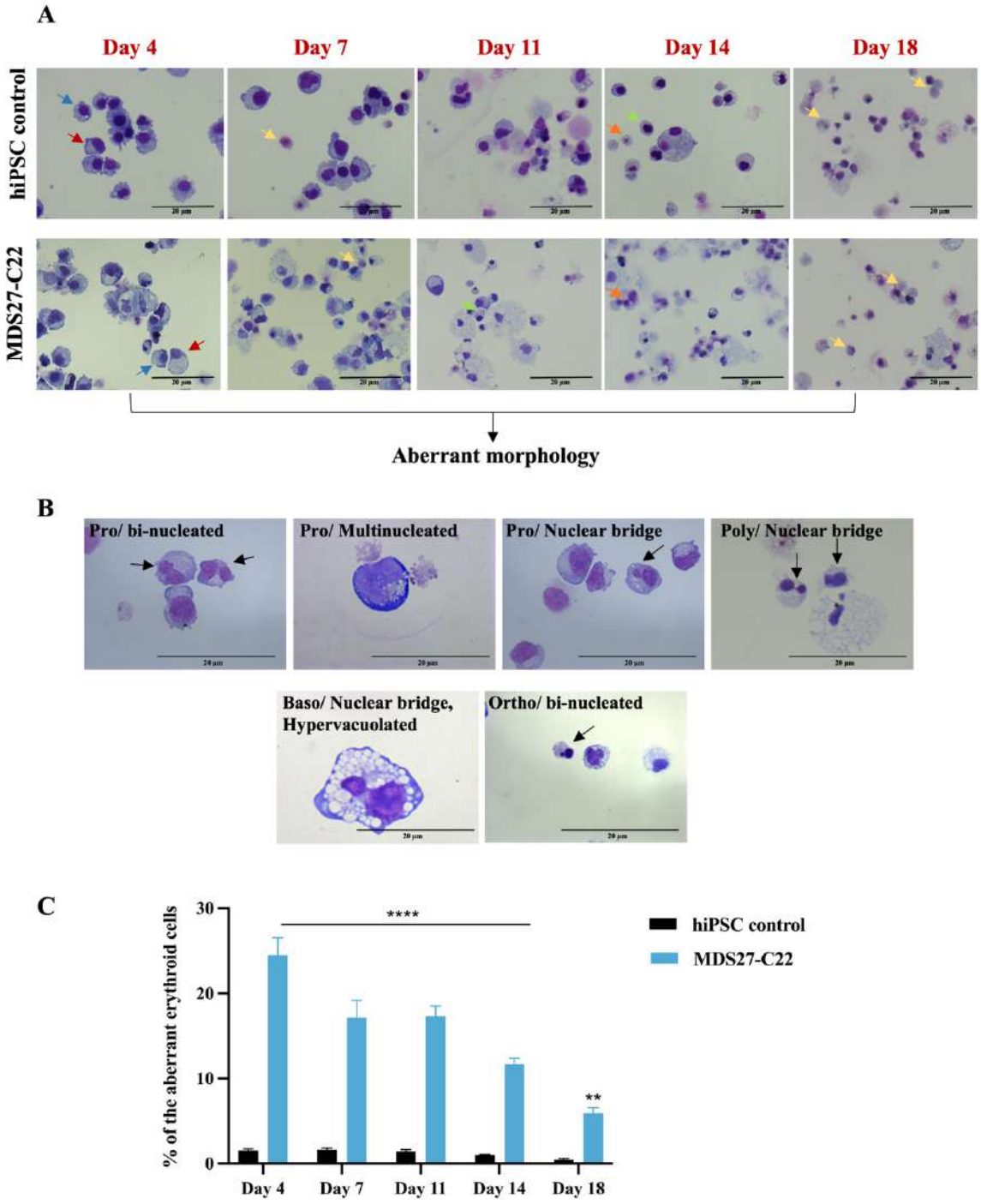


Figure 4.9: Morphological analysis of the erythroid cells

(A) Cytospin of erythroid cells shows the maturation stages during different days of the culture: Proerythroblast (red arrow), Basophilic erythroblast (blue arrow), Polychromatic erythroblast (green arrow), Orthochromatic erythroblast (orange arrow) & Mature cells (yellow arrow). **(B)** The common aberrant morphology (black arrow) in MDS27-C22. The pictures were taken Leica DM6000 at 40x and 100x magnification**,** 20 µm scale bar. **(C)** The percentage of the aberrant cells of MDS27-C22 to the aberrant morphology of hiPSC control. Results are presented as mean ± SEM and analysed using Two-way ANOVA with multiple comparisons. *** $p < 0.0001$ and ** p (0.0051 (day 18). The data represented 4 independent experiments.

4.3 Discussion

4.3.1 *ASLX1*, *RUNX1* and *SRSF2* mutations do not affect the formation of HSPCs from MDS27-iPSC clones

Despite the successful generation of iPSC from patient MDS27, it could have been possible the mutations that these clones harbour (*ASXL1, RUNX1* and *SRFS2*), could impede the generation of HSPC, and thus negating the possibility of having an *in vitro* system in which to understand the mechanisms of disease progression. Thus, iPSC clones from MDS27 were differentiated into HSPCs as an initial step to investigate the phenotype. Haematopoiesis is a convoluted process involving carefully balanced interactions among cytokines, different cell types, matrix factors and signalling pathways (Ditadi et al., 2017, Ivanovs et al., 2017). It involves two different waves; the primitive process is the first wave which begins the extraembryonic yolk sac during the initial gestation period in the human embryo (18-20 days) when the mesoderm differentiates into haemangioblasts cells to generate primitive erythroid, megakaryocyte and macrophage lineages that are derived from primitive EMPs (Tober et al., 2007, Palis et al., 1999). The second wave is called definitive haematopoiesis and produces the adult HSCs at various sites. For example, the AGM region is the main site of definitive haematopoiesis during the mid-stage of gestation, soon after the cells migrate and colonise the main sites for expansion; the placenta and foetal liver and finally the BM around birth where HSCs display the characteristic cell surface markers of adult HSC (Frame et al., 2013, Palis et al., 1999, Tober et al., 2007). EMPs are transported out of the yolk sac to the liver and differentiated into multiple types of blood cells, including megakaryocytes, enucleated erythrocytes, and monocytes. Also, EMP has produced B and T lymphocytes (Boiers et al., 2013). Unfortunately, several stablished protocols that aim to differentiate iPSCs to HSPCs do not produce mature HSPC (definitive) but recapitulate yolk sac hematopoiesis leading to foetal HSPCs (primitive)

(Hong et al., 2010, Chicha et al., 2011). Thus, a commercial protocol from stem cell technology was considered in this project to differentiate the hiPSC into HSPCs to limit the drawbacks of the previous protocols. The HSPCs differentiation was done for all MDS27 clones (1, 8, 11, 22), but then clone 1 was excluded because its result was highly variable compared with the other three clones.

Despite using a commercial protocol, it required a great deal of optimization, as depending on the number of colony fragments seeded and the size of the plate used, the percentage of hematopoietic cells obtained, the viability and the quality of the cells were affected. Nonetheless, once optimized, the percentage of definitive HSPCs was similar to other protocols used in the lab, such as the formation of HSPCs through embryoid body (EB) formation. This protocol successfully differentiated the hiPSC to early HSPC ($CD34^+$ $CD43^+$) and late HSPC ($CD34^+$ $CD45^+$) HSCs and importantly, we were able to obtain the same efficiency as other protocols already published for the differentiation of iPSCs towards HSPCs (Kotini et al., 2015, Kotini et al., 2017, Tursky et al., 2020, Hansen et al., 2018).

HSPC differentiation results indicated that MDS27-hiPSC clones were not impaired in their ability to generate HSPCs compared with hiPSC control. This finding is consistent with that of Chang et al., (2018), who found that *SRSF2* mutation in iPSC derived from a MDS patient did not affect the differentiation to HSPCs (Chang et al., 2018). Of note, our patient MDS27 presented a *SRSF2* mutation. These results also are in agreement with a study of Huang et al., (2017), who found that the co-existence of *Runx1* and *Srsf2* mutations in a murine model had not effect on the generation of HSPCs but they impaired the multipotent lineage haematopoiesis (Huang et al., 2017).

Contrary to our results, other studies have shown a significant reduction in the generation of HSPCs from hiPSC generated from MDS patients with *SF3B1*, *EZH2*, *de*

(5q) and *del(7q)* mutations (Hsu et al., 2019, Kotini et al., 2015, Kotini et al., 2017). These differences could be due to the differentiation protocol used or the type of mutation that these clones harboured.

In summary, the iPSC differentiation to HSPCs was successful and our data indicates that *ASLX1*, *SRSF2*, and *RUNX1* mutations do not have a role in affecting HSPC differentiation.

4.3.2 *ASLX1*, *RUNX1* and *SRSF2* impair the colony forming potential of MDS27 iPSC clones

The further differentiation potential of HSPCs was assessed by seeding the HSPCs in semisolid medium for 14 days. After this time, HSPCs from hiPSC control and MDS27-hiPSC clones gave rise to all the types of CFUs but the total number of colonies obtained was significantly reduced in MDS27 clones. The decrease in number of colonies was observed for both erythroid and myeloid lineages for all MDS27-hiPSC clones. This result implied that the reduction in the number of CFUs was because of the effect of these three mutations (*ASLX1, SRSF2* and *RUNX1*) on the multilineage differentiation capacity of HSPCs.

These results are in agreement with the clinical diagnosis of MDS27 patient, as this patient was diagnosed with refractory anaemia with multilineage dysplasia. Besides, the reduction in HSPC clonogenic capacity of the BM of MDS patients has been reported by different studies (Li et al., 2016, DeZern et al., 2013, Michalopoulou et al., 2004). Our findings are also consistent with those reported in the previous study using the double mutant *Srsf2* and *Runx1* murine model, as mouse hematopoietic cells from these animals displayed a reduced clonogenicity potential (Huang et al, 2017).

4.3.3 *ASLX1*, *RUNX1* and *SRSF2* affect the quality of the erythroid cells derived from MDS27-C22 hiPSCs

MDS is characterised by dysplastic or ineffective haematopoiesis, which leads to blood cytopenia. For example, anaemia is a common occurrence in MDS patients due to dysplastic cells in BM or a significant reduction in the number of reticulocytes or mature cells in the PB (Gupta et al., 2007). For that reason, it was of interest to investigate the differentiation potential towards the erythroid lineage of the iPSC clones generated from the MDS27 patient.

Our erythroid differentiation results indicated the successful differentiation of hiPSC control and MDS27-C22 hiPSCs into erythroid lineage cells. In this study, flow cytometry analysis was one of the methods used to monitor the erythroid differentiation. Surprisingly, our flow cytometry results did not detect any immunophenotypic changes during erythroid differentiation between hiPSC control and MDS27-C22 hiPSCs.

This inconsistency may be due to the use of an inappropriate hiPSC control (cell line) that has a different genetic background from MDS27. For example, it has been reported previously that control hiPSCs cell lines obtained from healthy volunteers behave differently, with lower differentiation potential than normal isogenic control iPSCs derived from the same MDS patient or aplastic anaemia patient. This indicates that changes in the genetic background could be the main source of the lack of significant differences in our results (Kotini et al., 2015, Melguizo-Sanchis et al., 2018).

Flow cytometry is a useful tool to characterise erythroid precursor maturation and to detect possible aberrant differentiation by looking at the expression of specific antigens (van de Loosdrecht et al., 2013). Despite this, flow cytometry suffers from some limitations that can affect the accuracy of the diagnosis of MDS. For instance, using flow cytometry to detect the

blast cells percentage based on the expression of $CD34^+$ is not ideal in MDS or AML because not every blast cell is positive for the CD34 marker; equally, not all cells will express an aberrant pattern of erythroid markers as for identifying dyserythropoiesis (Lewandowski et al., 2012). The aforementioned limitations may lead to a false diagnosis of MDS or count the percentage of blast cells in MDS patients (Lewandowski et al., 2012). For that reason, most cytogenetic laboratories rely on microscopic examination, which remains one of the key diagnostic procedures in haematology diseases. Indeed, for MDS diagnosis, morphological analysis is the main method to identify the dysplastic features and the transformation status (Greenberg et al., 1997); the microscope examination is always performed even when flow cytometry diagnosis of MDS is suspected (Goasguen et al., 2018).

A detection of at least 10% of aberrant cells in the BM is required for a diagnosis to be defined as dysplasia (Campo et al., 2011, Vardiman et al., 2009, Hasserjian et al., 2017). Our morphological results indicate that MDS27-C22 iPSCs have a high significant proportion of aberrant cells (12-23%) compared with the hiPSC control, and we could not detect this phenotype by flow cytometry. In addition, the type of aberrant morphology (dysplasia) that we observed in erythroid cells derived from MDS27-hiPSCs (internuclear bridge, binocularity, multinuclearity and hyper vacuolation) are common features present in MDS patients according to the WHO. In our data, the aberrant morphology was identified extensively in the nucleated erythrocytes or immature cells but not at the mature stage. A possible explanation of this result is the challenge to detect the dysplastic features in mature cells, as it needs specialised expertise, and this impediment has been reported previously (Mufti et al., 2008).

Moreover, according to our results, the erythroid differentiation of MDS27-C22 hiPSCs lead to two different erythrocyte populations, one presenting normal morphology and the other one with aberrant morphology. This result mimics the findings reported for the BM of MDS patients

that both normal and aberrant cells co-exist in the BM (Zini, 2017). Also, it could be possible that erythroblasts derived from MSD27-C22 hiPSCs reach a mature stage in which normal cells could complete the differentiation, but the aberrant cells could die during the differentiation. For that reason, analysis of apoptosis on the erythroid cultures were attempted, but the experiments were not successful. The main reason for this was due to the nature of the erythroid differentiation protocol: many cells with damaged membranes, fragile morphology, dead cells, and debris were observed in the culture with both hiPSC control and MDS27-C22 hiPSCs making the assessment of apoptosis very difficult. These observations were also reported by several studies in which they found that the generation of erythroid cells from hiPSC have a poor membrane, lower enucleation rate, and highly expressed foetal Hb (Fujita et al., 2016, Chang et al., 2011, Lapillonne et al., 2010, Lachmann et al., 2015).

Recently, many protocols have been reported which improve the limitation of erythroid differentiation from hiPSC (Hansen et al., 2018, Lopez-Yrigoyen et al., 2019, Bernecker et al., 2019, Tursky et al., 2020), and it would be useful for our study to attempt different protocols to generate high-quality erythroid cells.

It is worth mentioning that the erythroid differentiation was also performed for the other two clones (C8 and C11) and both clones gave similar results to MDS27-C22, indicating that the limitations encountered were not related to a specific clone.

Our results collectively are consistent with earlier findings that *ASLX1*, *SRSF2*, *TP53* and *RUNX1* mutations are associated with myeloid and erythroid dysplasia in MDS patients (Invernizzi et al., 2015). Therefore, these aberrant morphologies could affect the formation of mature and functional RBCs causing the anaemia observed in MDS27 patient. For example, the aberrant cells could affect the Hb synthesis, cell membrane stability, or O2 transport process

leading to production of non-functional mature cells or reduce the cells' survival rate (Zini, 2017).

The erythroid differentiation results are interesting, but they have some limitations. For instance, it would be interesting to analyse and determine the type of Hb at the protein level because as a result of dyserythropoiesis in MDS, most of the patients have a predominant foetal Hb that leads to increased O_2 binding affinity (Stomper et al., 2019). In addition, to develop a clear picture of the dysplasia found in MDS27 patient, we could have checked the expression of different genes associated with erythropoiesis, such as *GATA1* and *KLF1*.

To overcome this limitation and to clearly understand the mechanism behind this phenotype, we isolated two different cell populations from hiPSC control and MDS27-C22: (1) HSPCs from day 12 of HSPCs differentiation culture, and (2) Erythroid cells from day 5 of erythroid differentiation culture and performed single cell RNA sequencing (scRNA-seq) Single-cell RNA sequence results were not available at the time of writing this thesis.

Chapter 5 : Studying the contribution of *C/EBPα* mutation to the disease progression of patient MDS27

5.1 Introduction

In the previous chapter, we corroborated the disease phenotype of MDS27-hiPSCs generated by differentiating hiPSC control and MDS27-hiPSC clones to HSPCs. The results showed that *ASLX1*, *RUNX1* and *SRSF2* mutations did not affect the differentiation of HSPCs but reduced the clonogenic capacity of HSPCs in MDS27-hiPSC (low-risk MDS). According to the clinical diagnosis, the patient MDS27 had multilineage dysplasia and suffered from anaemia, encouraging us to study the erythroid lineage differentiation of MDS27-hiPSCs. The erythroid differentiation results demonstrated that *ASLX1*, *RUNX1* and *SRSF2* mutations were responsible for the dyserythropoiesis observed in MDS27-hiPSCs.

After successfully corroborating the MDS27 phenotype, we were keen to investigate the mechanisms behind the progression of MDS27 patient from low-risk to high-risk MDS. Previously, we have mentioned in chapter three that in 2015 the blast cells of patient MDS27 had increased, and he progressed to high-risk MDS, acquiring a heterozygous *C/EBPα* mutation. Several attempts to generate hiPSC from MDS27 after disease progression (sample in early 2015), failed and we were unable to maintain stable colonies even with some optimisation to the protocol. To obtain cells before and after disease progression with the same genetic background (harbouring mutations in *ASLX1*, *RUNX1* and *SRSF2)*, we decided to apply CRISPR-Cas9 to introduce the *C/EBPα* mutation into MDS27-C22 hiPSCs in order to understand the contribution of *C/EBPα* to the disease progression.

This aim will be addressed in this fifth chapter by completing these objectives:

- Designing gRNA to target the C-terminal region of *C/EBPα.*
- Generate the new lines from MDS27-C22 contains *C/EBPα* mutation.

- Characterising the pluripotency of the new lines to determine whether the CRISPR-Cas9 would affect pluripotency.

- Identifying the chromosomal stability of hiPSC clones after using CRISPR-Cas9.

- Checking the capability of the new clones to differentiate towards HSPC.

- Defining the differentiation potential of HSPCs from control and MDS27-iPSC clones harbouring *C/EBPα* mutation by performing colony assay.

- Assessing the myeloid and erythroid lineage differentiation of HSPCs from MDS27-iPSC clones harbouring *C/EBPα* mutation in liquid culture by performing flow cytometry analysis and morphological characterisation.

5.2 Results

5.2.1 Engineering heterozygous *C/EBPα* mutation in MDS27-C22 hiPSC

Cells from patient MDS27 harboured four mutations (*ASLX1, RUNX1, SRSF2* and *C/EBPα*) when they progressed to high-risk MDS. Thus, to determine the impact of the disruption of the *C/EBPα* DNA binding domain (DBD) on the MDS27 cells harbouring already three mutations (*ASLX1, RUNX1, SRSF*2), we employed CRISPR-Cas9 technology. For this purpose, a specific gRNA was designed using an online tool (Trust Sanger Institute Editing database) to direct Cas9 to the C-terminal region (DBD) of *C/EBPα* (the mutated region in patient MDS27). The gRNA oligonucleotide sequence was modified by the inclusion of CACC before the guide's forward complement and AAAC before the guide's reverse complement to be able to clone the gRNA into pSpCas9(BB)-2A-GFP (PX458) (addgene) at the BbsI restriction site. The PX458 vector contains the Cas9, GFP reporter marker, and the U6 promoter to regulate the expression of the gRNA. The gRNA primers were annealed and cloned into the PX458 vector. Then, the plasmid was transfected into MDS27-C22 iPSCs by amaxa nucleofection. Twenty-four hours post-nucleofection cells were dissociated, and the GFP population was sorted by FACS. All the GFP cells were replated at low density for isolation of single gene-mutated clones. After three weeks of the nucleofection, the gDNA was extracted from twenty-four individual clones, and a T7EI assay was performed to evaluate the gene editing efficiency of CRISPR-Cas9 in which T7 endonuclease can recognise and cleave non-perfectly matched DNA (figure 5.1 and Figure 5.2).

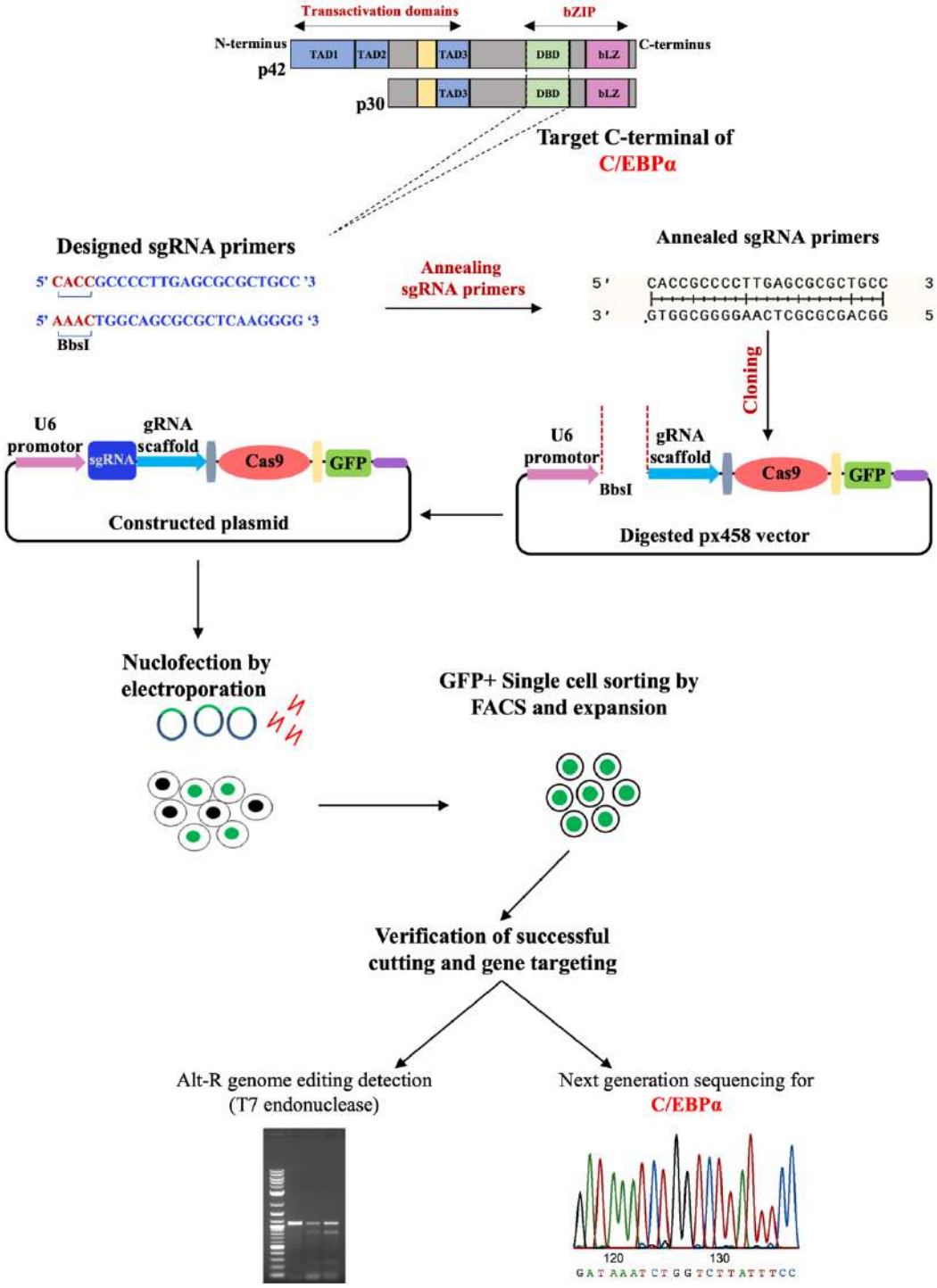


Figure 5.1: Utilizing CRISPR-Cas9 in hiPSC

Schematic representation of engineering C/EBPα mutation in MDS27-C22 using PX458 Cas9 plasmid. Guide RNA targeted DBD of C/EBPα were designed via an online tool (Trust Sanger Institute Editing database). Then, the sgRNA guide was cloned into an expression plasmid bearing both sgRNA scaffold backbone (BB) and Cas9, pSpCas9(BB) (PX458). The constructed plasmid was transfected into hiPSC MDS27-C22. Then, T7I assay and the next-generation sequencing were applied to identify the successful cutting and detect the type of mutation.DBD (DNA binding domain), TAD (Transcription activation domain) and bLZ (basic leucine zipper domain).

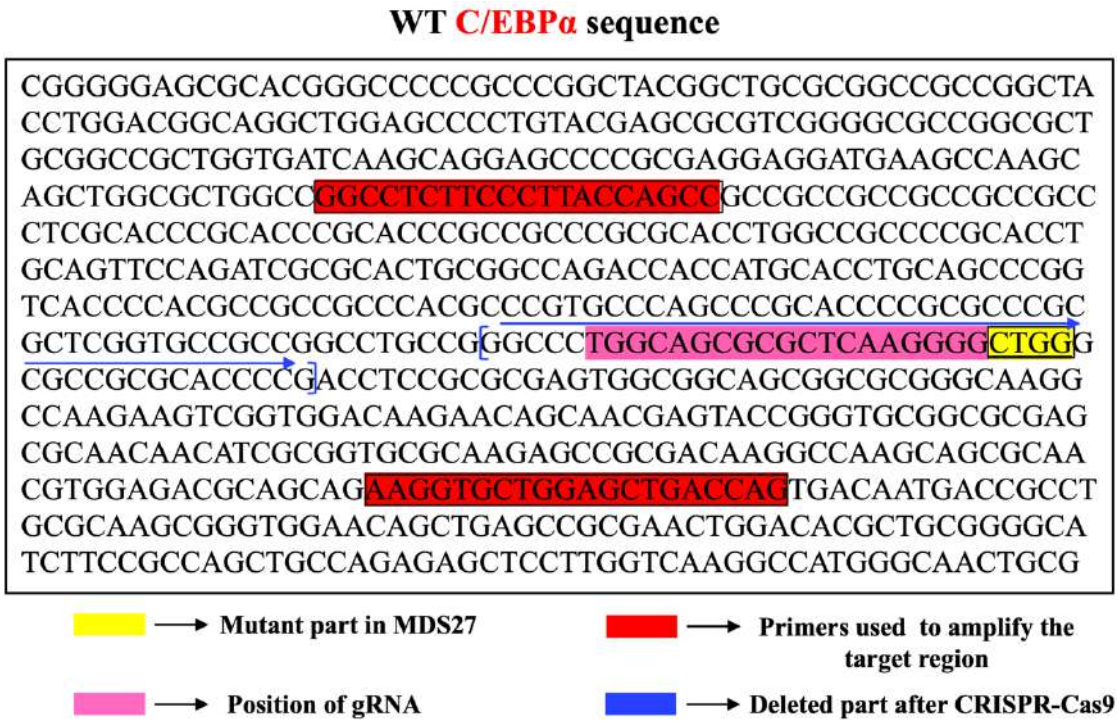


Figure 5.2: C/EBPα sequence

Snapshot of C/EBPα sequence centre in the mutated region. Position of gRNA (pink), primers used to amplify the targeted region (red) and the deleted nucleotides (blue) after CRISPR-Cas9 are highlighted.

We designed PCR primers (Figure 5.2) that amplify the *C/EBPα* target site for screening the positive clones, and the amplification was carried out using Q5 polymerase for MDS27-C22 (WT) and the twenty-four isolated clones. The agarose gel results showed that PCR amplification was successful in seven clones out of twenty-four (Figure 5.3, A). Then, the DNA of these seven clones and WT clone were incubated with T7EI enzyme and visualized using an agarose gel. Also, a homoduplexes PCR product control (A) (for homozygous mutation) and heteroduplexes PCR product control (A+B) (for heterozygous mutation) were also included on the agarose gel.

As shown in Figure 5.3, (B) three clones (C22.5, C22.7 and C22.20) out of 7 were found to contain a heterozygous mutation, showing the upper band of the expected size (456 bp) based

on the NCBI published sequence and a lower band with a smaller size, indicating that some kind of deletion had occurred in that amplified region as a consequence of CRISPR-Cas9. However, the (C22.2, C22.4, C22.13 and C22.23) had only one band which could mean that the CRISPR-Cas9 did not generate the mutation in the target region, or it could generate a small homozygous mutation.

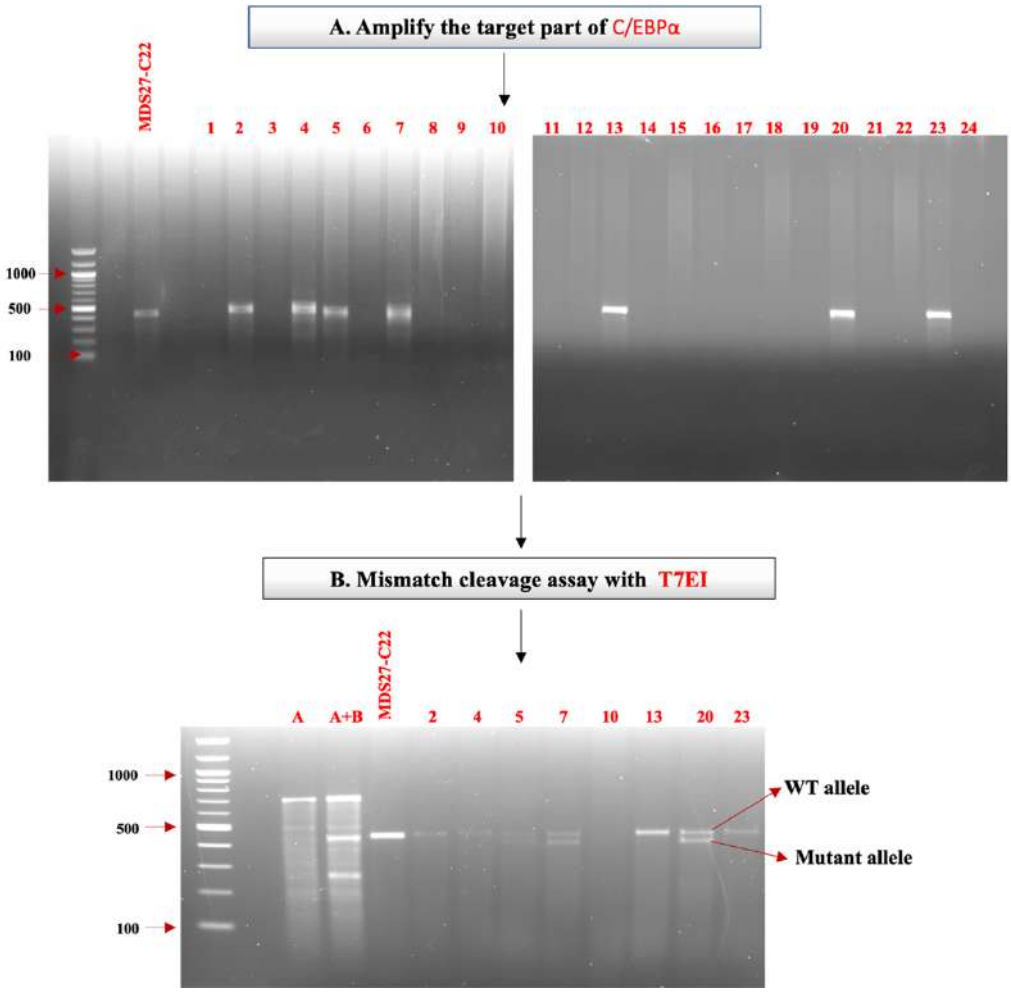


Figure 5.3: Mismatch cleavage assay with T7EI

(A) gDNA were extracted from MDS27-C22 (WT) and 24 clones after CRISPR editing. PCR was applied to amplify a DNA fragment flanking the target region (456 bp) using primers designed around the target region. The PCR result was visualised on an agarose gel. **(B)** The PCR products from step A were digested with T7EI. Digestion reactions were analysed on an agarose gel. Sample 1 contains Control A homoduplexes PCR products, while Sample 2 contains homoduplexes and heteroduplexes of Control A and B PCR products. Expected DNA fragment size is indicated as WT, and that showing a deletion indicated as mutant allele.

As mentioned earlier, the MDS27 patient had acquired a *C/EBPα* heterozygous mutation during disease progression that had led to disruption of the DBD of *C/EBPα.* Using the T7EI assay we could only determine which clones had suffered some kind of mutation in heterozygosity, and it was important to determine where the mutation had occurred and whether the mutation would disrupt the DBD of *C/EBPα.* Thus, mutant clones (MDS27- C22.5, MDS27-C22.7 and MDS27-C22.20) were subjected to sanger sequencing to detect possible insertion, deletion or mismatches introduced by CRISPR-Cas9 modification in the target area. Sanger sequencing confirmed that all three clones harboured the deletion in the target site in one allele of the *C/EBPα* gene (Figure 5.4). The sequencing data showed that MDS27-C22.5 contained multiple integrations and deletions whilst MDS27-C22.7 and MDS27-C22.20 had both a 43bp deletion in the desired region. Thus, MDS27-C22.5 was excluded from this study and MDS27-C22.7, and MDS27-C22.20 were expanded and used for further experiments. In addition, the protein sequence of C/EBPα indicated that deletion of the 43bp in the desired region resulted on a change in the open reading frame and premature termination of C/EBPα protein(Figure 5.5). Subsequently, this change on the protein sequence and truncation would alter the DNA binding domain of C/EBPα and thus its functionality.

In the further experiments, five selected clones were processed for each experiment: (i) hiPSC control, used as normal iPSC; (ii) MDS27-C22 used as the WT control with *SRSF2*, *ASLX1*, *RUNX1* mutations; (iii) MDS27-C22+PX458, as CRISPR-Cas9 control; and (iv) MDS27-C22.7 and MDS27-C22.20, the two clones that contain *SRSF2*, *ASLX1*, *RUNX1* and *C/EBPα* mutations.

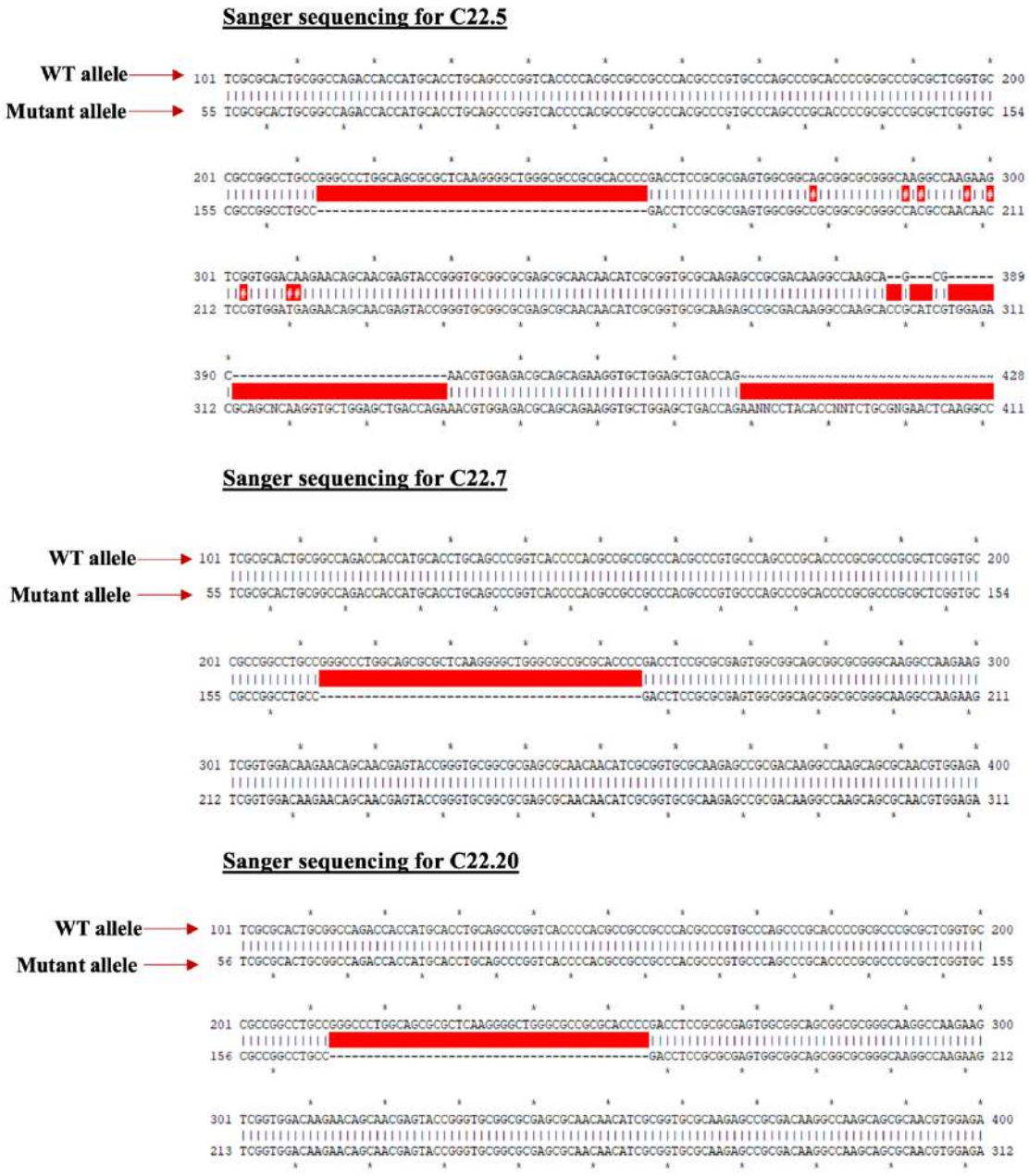


Figure 5.4: Sanger sequencing for positive clones

Three clones (MDS27-C22.5, MDS27-C22.7 and MDS27-C22.20) with heterozygous mutation were processed to sanger sequencing to identify the type of mutation. Sanger sequencing was performed by SourceBioscience with primers flanking the targeted region. Primer sequence can be found in materials and methods section 2.17. the sequence was aligned by using ApE-plasmid editor software.

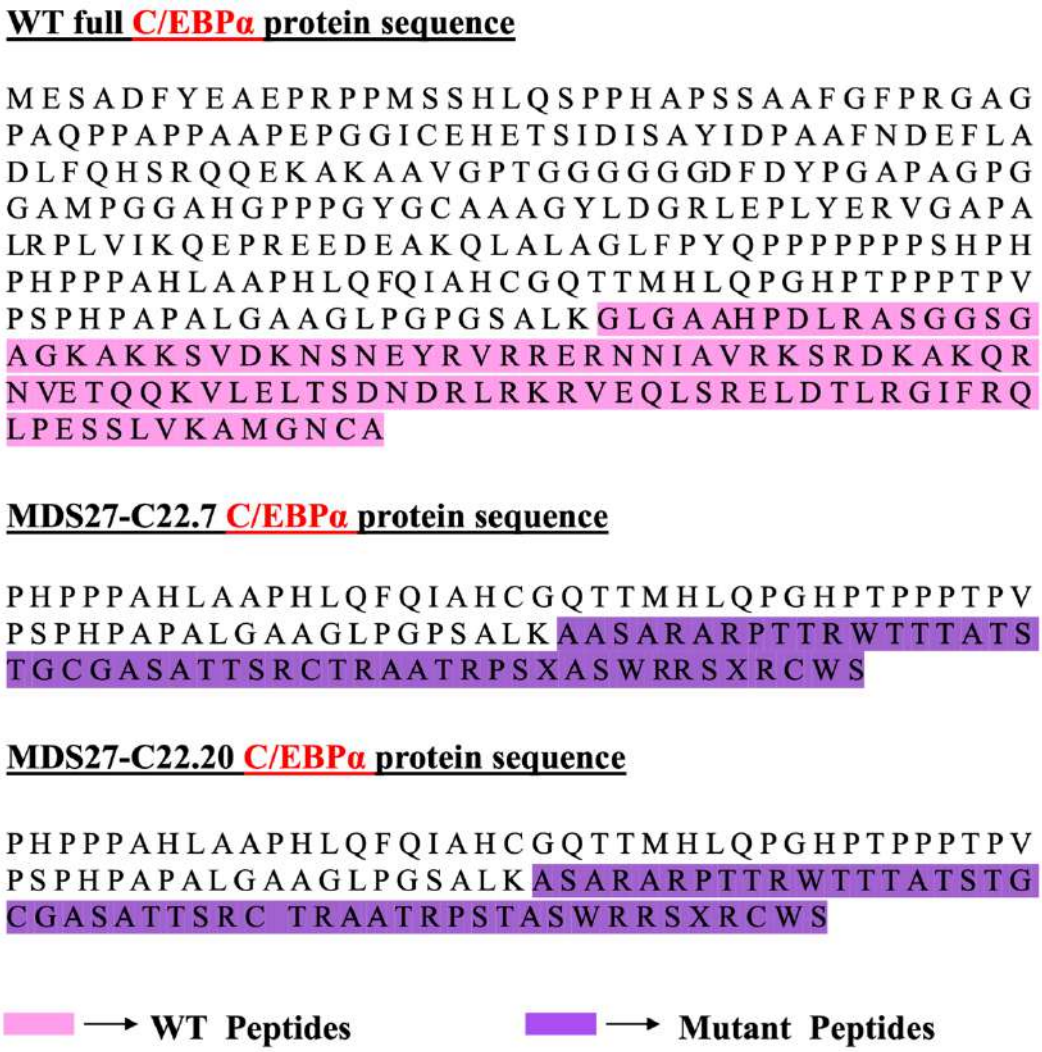


Figure 5.5: C/EBPα protein sequencing

The figure shows the protein sequence for C/EBPα protein in WT clones and mutant clones (MDS27-C22.7 and MDS27-C22.20), in which the deletion generated by CRISPR-Cas9 leads to a change in the open reading frame and premature termination. Protein sequencing was identified by Basic local alignment search tool (BLAST).

5.2.2 Mutant clones display a truncated C/EBP*α* protein and a change in the open reading frame (ORF)

It was important to determine whether the deletion generated by CRISPR could affect the protein levels or imbalance between the two isoforms of C/EBP*α*. Thus, C/EBPα protein expression from hiPSC control, MDS27-C22, MDS-C22+PX458, MDS27-C22.7 and MDS27-C22.20 was checked by western blotting and two positive controls were used: Kasumi-1 cells, which is a human AML immortal cell line with translocation (8;21) that expresses both isoforms of C/EBPα (Larizza et al., 2005), and Kasumi-1+ RUNX1-ETO siRNA that induces the expression of C/EBPα protein (generous gift from Bonifer‘s group (Ptasinska et al., 2019)).

C/EBPα protein expression was detectable in positive controls and hiPSC lines, and there was no remarkable difference in the expression of this protein between the mutant lines and WT lines (MDS27-C22, MDS-C22+PX458) (Figure 5.6, A and B). However, the WB result showed other bands with mutant clones that were not detected in the control cell lines. This result could indicate that a protein-truncation has occurred after CRISPR-Cas9, and truncation could affect the stability of the protein.

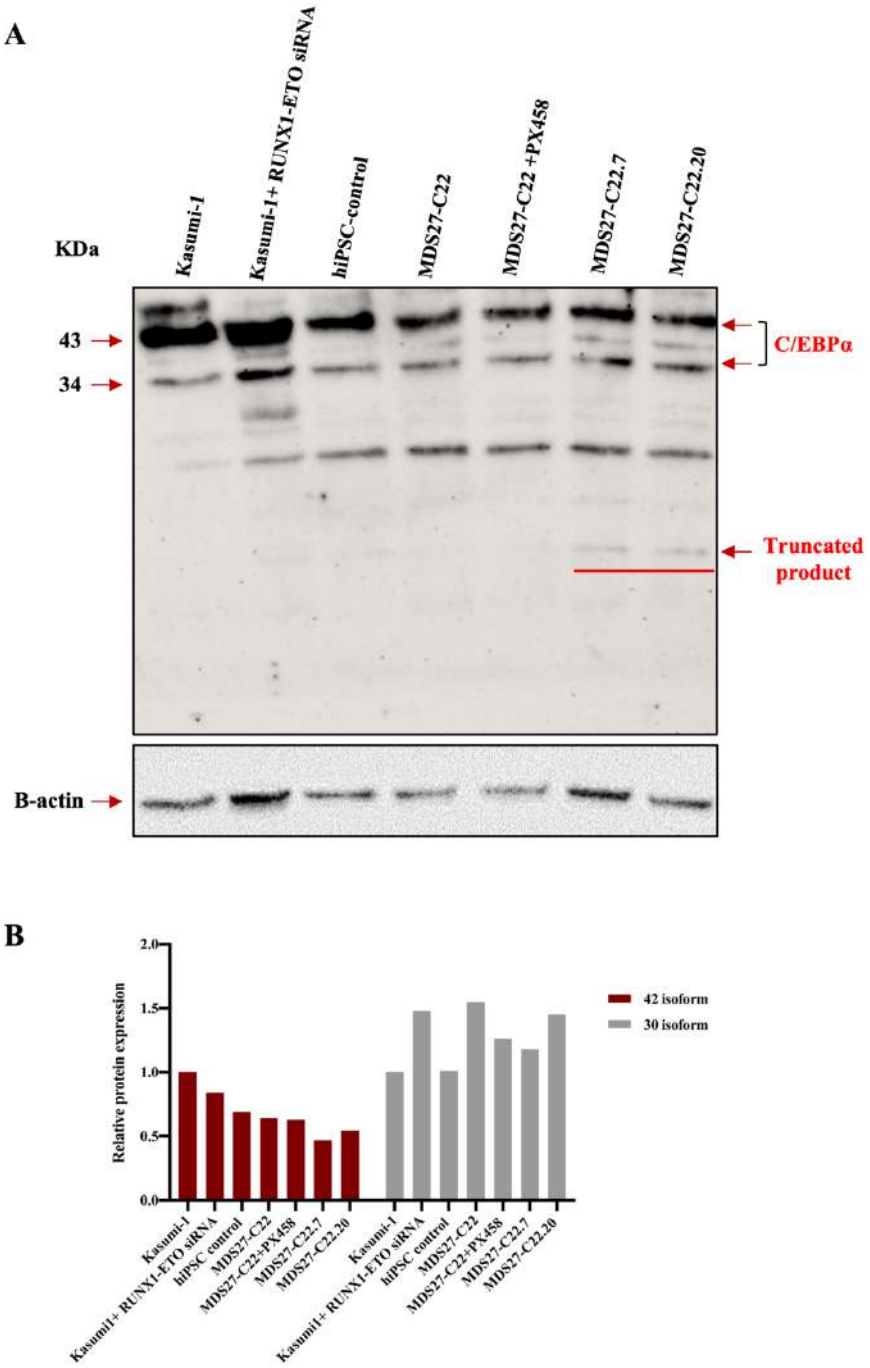


Figure 5.6: C/EBPα protein expression in hiPSC clones and Kasumi-1 cells

(A) Total protein was extracted from Kasumi-1, Kasumi-1+ RUNX1-ETO siRNA, hiPSC control, MDS27-C22, MDS27- C22+PX458, MDS27-C22.7 and MDS27-C22.20. 50 µg of total protein from each cell line was subjected to SDS-PAGE. Protein expression was detected by western blotting with antibodies against C/EBPα. The position of relevant molecular weight standards is shown on the left side of the blot, and arrows on the right indicate the specific bands for each antibody. **(B)** Quantification assay of C/EBPα protein expression normalized to b-actin and relative to Kasumi-1. Band intensity was calculated using ImageJ software. Data represent 2 independent experiments.

5.2.3 Does the use of the CRISPR-Cas9 system affect the chromosomal stability and the pluripotency of hiPSC?

We next sought to identify if CRISPR-Cas9 caused any alterations in the number of chromosomes. As described previously in 3.2.3, the three selected clones: MDS27-C22+PX458**,** MDS27-C22.7 and MDS27-C22.20 were cultured and treated with colcemid, a drug used to block cells at the metaphase stage (Rieder and Palazzo, 1992, Taylor, 1965). After two hours of treatment, chromosome preparation was conducted, and the number of chromosomes on the metaphase spreads counted. As human cells contain 46 chromosomes, metaphases displaying more, or less than 46 chromosomes would represent aneuploidy. Interestingly, All the CRISPR-Cas9 clones showed a normal number of chromosomes (46, XY), as can be seen in Figure 5.7.

This result indicates that the CRISPR-Cas9 system has not affected the chromosome stability of the hiPSC clones.

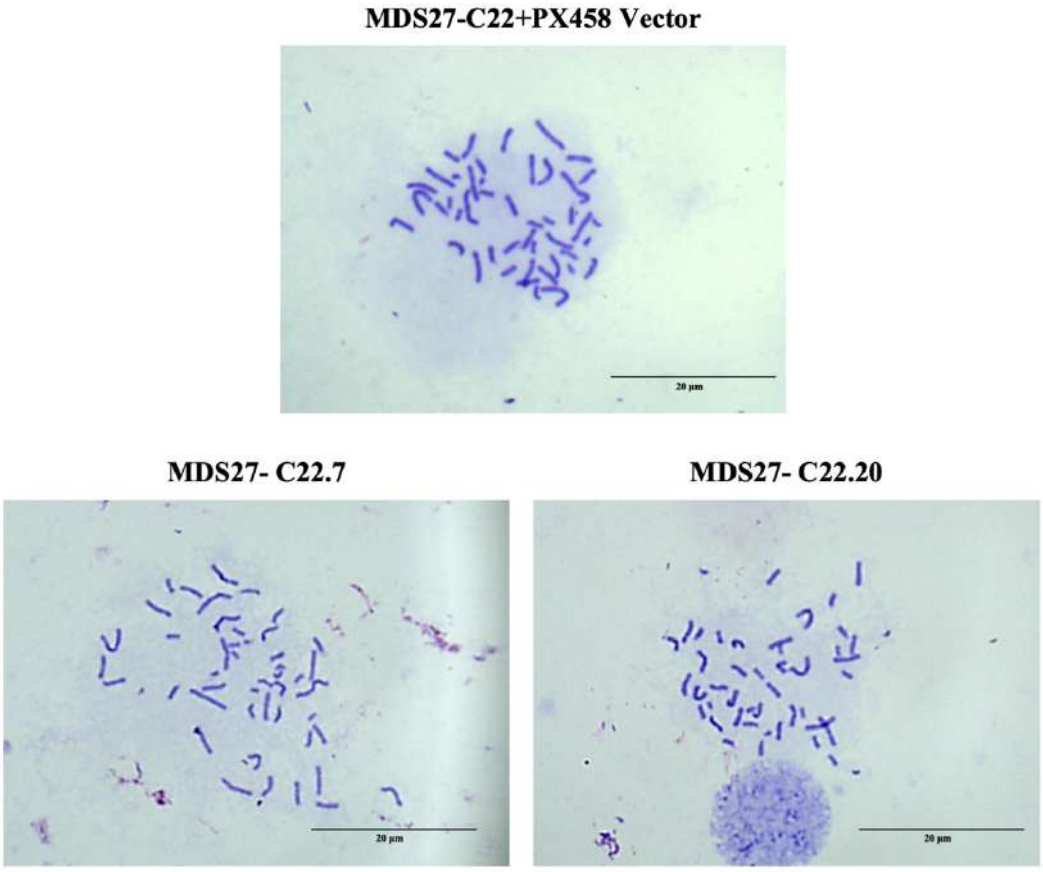
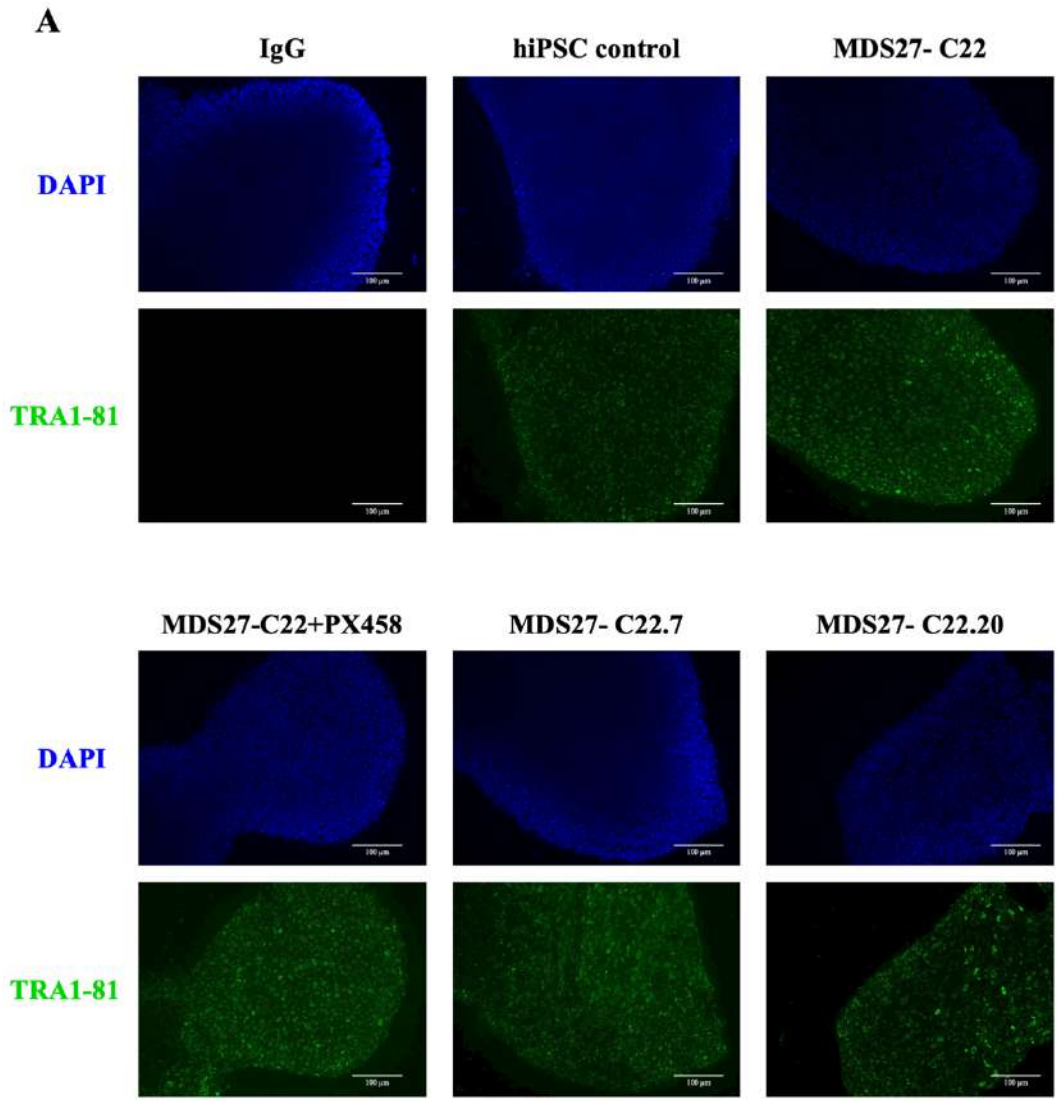


Figure 5.7: Normal number of chromosoms in hiPSC clones after applying CRISPR-Cas9.

An 80 % confluent well of a 6-well plate for MDS27- C22+PX458 Vector, MDS27-C22.7 and MDS27-C22.20 was treated with 0.02 µg/ml colcemid for 2 hours, swollen with 0.075 M KCl, fixed with cold Carnoy's solution, dropped onto humidified, chilled glass slides, then stained with Giemsa staining and imaged at 100x magnification. >25 metaphases per sample per experiment, 20µm scale bar, 100x magnification, Leica DM6000 light microscope. Data represent 3 independent experiments.

Then, we investigated if the CRISPR-Cas9 affected the pluripotency potential of the individual hiPSCs clones MDS27-C22+PX458**,** MDS27-C22.7 and MDS27-C22.20. The pluripotency characterisation of hiPSC was assessed based on the positive expression of pluripotent protein markers (TRA1-81, SOX2 and NANOG). The same methods in section 3.2.4 were used to stain the hiPSC clones with antibodies against the extracellular pluripotency marker TRA1-81 and the intracellular pluripotency markers NANOG and SOX2. Fortunately, CRISPR-Cas9 clones strongly expressed the pluripotent markers TRA1-81 (green, Figure 5.8, A) as well as NANOG (green) and SOX2 (red, Figure 5.8, B) when compared to the hiPSC control and MDS27-C22.

This is confirmation that the CRISPR-Cas9 system did not alter the expression of pluripotent markers, and all the generated clones after CRISPR-Cas9 maintained the expression of pluripotent markers.



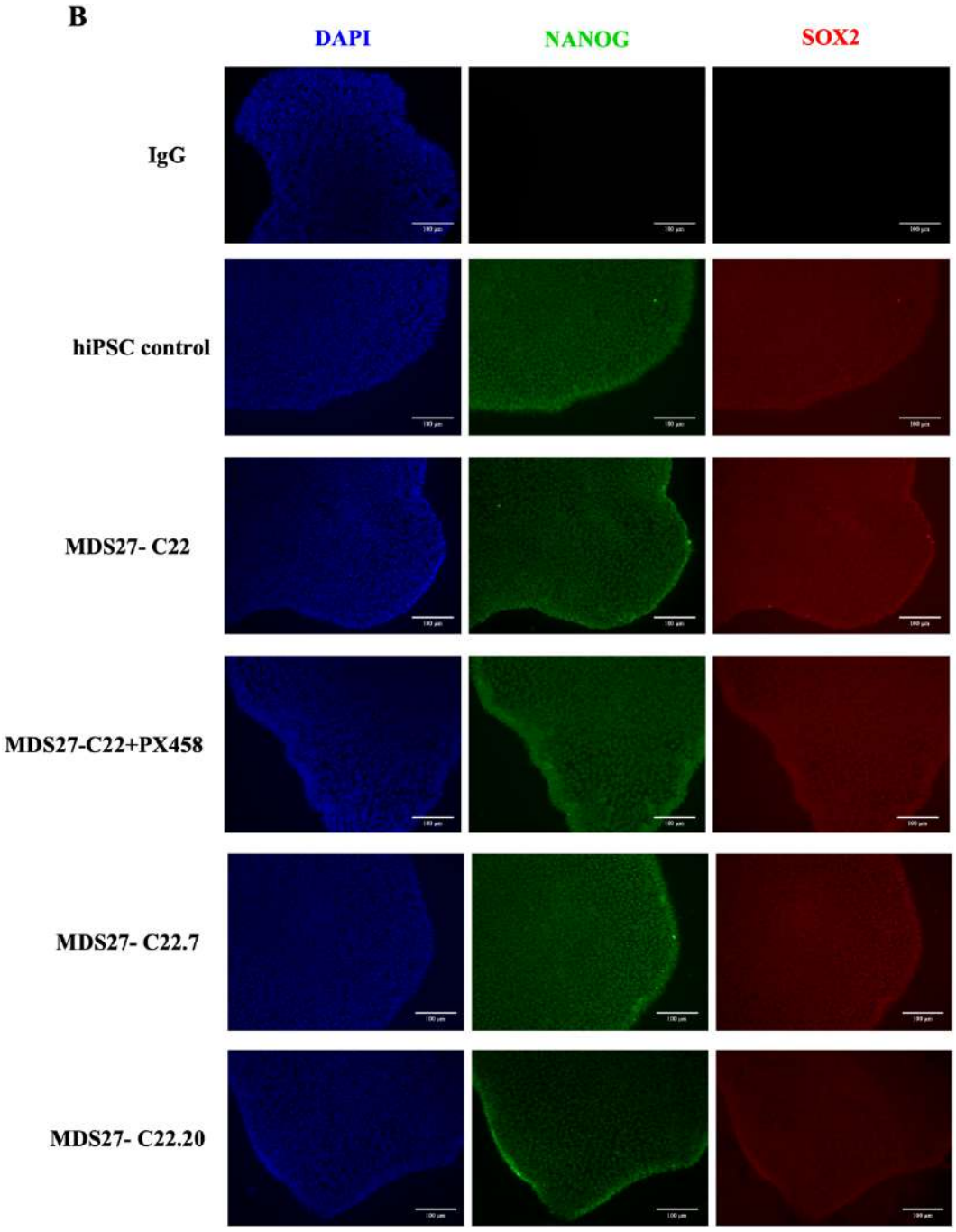


Figure 5.8: Positive expression of pluripotent markers after applying CRISPR-Cas9

Human iPSCs control, MDS27-C22, MDS27-C22+PX458, MDS27-C22.7 and MDS27-C22.20 were grown on coverslips coated with matrigel. **(A)** Cells were stained with TRA1-81 (Green). The cells were stained with mouse Alexa Flour 488 IgM and fixed before being counterstained with DAPI (Blue). **(B)** Cells were fixed, permeabilised and then stained with NANOG (Green) and SOX2 (Red). The cells were then stained with anti-goat Alexa 488 and anti-mouse Alexa 633 antibodies before being counterstained with DAPI (Blue). 100 µm scale bar, 20x magnification, Leica DM6000. Data represent 3 independent experiments.

5.2.4 Differentiation of hiPSC control, MDS27-C22 and CRISPR-Cas9 clones to hematopoietic stem progenitor cells (HSPCs)

Once we had corroborated that the use of CRISPR-Cas9 technology had not affected the chromosome stability nor the pluripotency potential of iPSCs, we were interested in defining whether the *C/EBPα* mutation generated in combination with the mutations in *ASLX1*, *RUNX1, SRSF2* could affect the differentiation potential of the iPSC to HSPCs. Thus, we differentiated the hiPSC control, MDS27-C22, MDS27-C22+PX458**,** MDS27-C22.7 and MDS27-C22.20 into HSPCs using the STEMdiff hematopoietic protocol from STEM Cell Technology as described in the previous chapter (4.2.1). The differentiation of each clone was monitored by assessing the expression of CD34 (a hemato-endothelial marker), CD43 (early haematopoietic marker which persists in differentiating precursor cells) and CD45 (the key marker of definitive hematopoietic cells) by flow cytometric analysis at different days (Day10, Day12 and Day 14) of haematopoietic differentiation. The cell number of HSPCs is presented in the supplementary figure 6.

The analysis showed that at day 10 of differentiation, 90% of the population were $CD43^+$ for hiPSC control, MDS27-C22, MDS27-C22+PX458**,** MDS27-C22.7 and MDS27-C22.20, and that this percentage did not seem to change by day 14 (Figure 5.9, A and B). In contrast, the percentage of cells co-expressing $CD34^+$ and $CD45^+$ appeared to increase from 14% at day 10 of differentiation to approximately 37% by day 14 for all hiPSC lines (Figure 5.10, A and B).

The comparison of hiPSC control, MDS27-C22 and MDS27-C22+PX458 with mutant clones indicated no statistically significant difference in the expression of hematopoietic markers at day 14, seeFigure 5.9, (B) and Figure 5.10, (B).

The comparison of MDS27-C22 (WT) with MDS27-C22+PX458 (CRISPR control) indicated that CRISPR control behaved similar to WT clone with respect to the capability to differente into HSPCs, and thus, the transient transfection of the empty PX458 vector did not affect the differentiation capacity of the cells (Figure 5.9, A and B) and (Figure 5.10, A and B).

In summary, this result demonstrated that the inclusion of the *C/EBPα* mutation in cells harbouring mutations in *ASLX1*, *RUNX1*, *SRSF2* did not affect the generation of HSPCs.

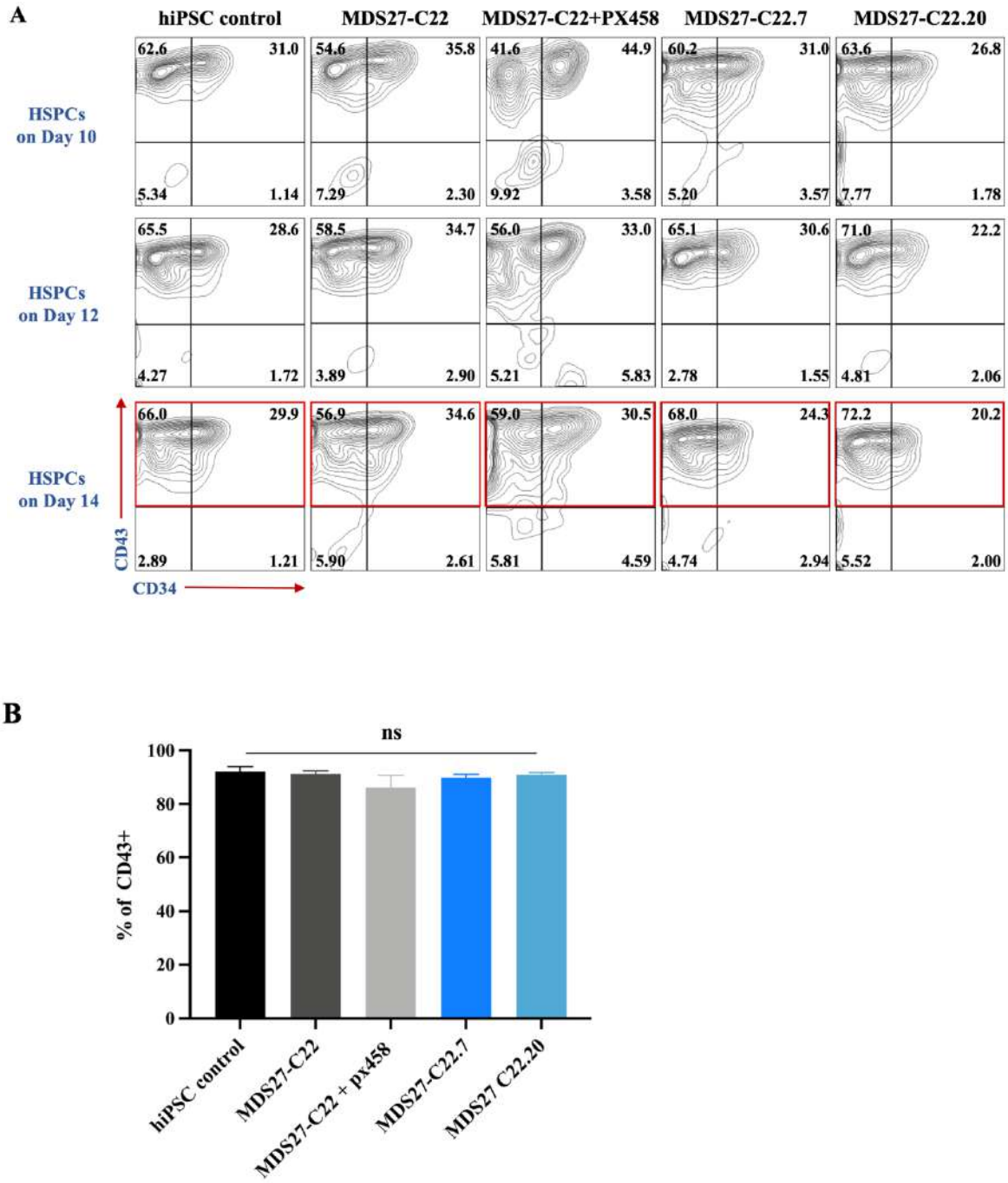


Figure 5.9: Human iPSCs differentiate to primitive HSPCs

(A) CD43^{+} (early hematopoietic population) were monitored during the differentiation time course on day 10, 12 and 14. The differentiating cells were analysed by flow cytometry; the cells were gated first based on FSC and SS of unstained cells and then analysed for hematopoietic markers based on the isotype control. **(B)** Figures show the mean percentage of hematopoietic markers (CD43^{+}) on day 14. Statistical results are presented as mean ± standard error of the mean (SEM) and analysed using One-way ANOVA with multiple comparisons. p values were represented as ns for not significant. Data represent 4 independent experiments.

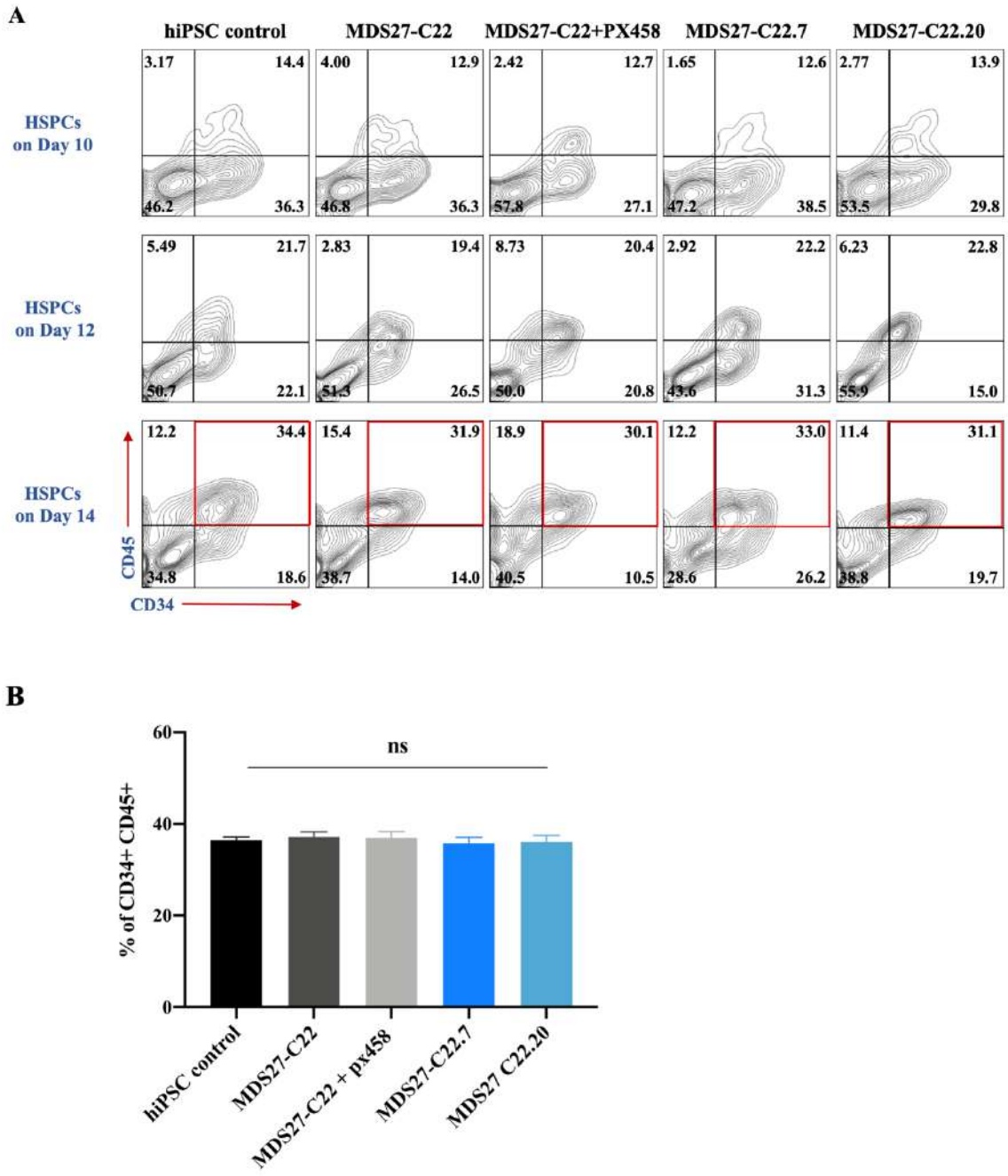


Figure 5.10: Human iPSCs differentiate to definitive HSPCs

(A) $CD34^+$ $CD45^+$ (definitive hematopoietic population) were monitored during the time course of differentiation Day 10, 12 and 14. The differentiating cells were analysed by flow cytometry and the cells were gated first based on FSC and SS of unstained cells and then analysed for hematopoietic markers based on the isotype control. **(B)** Figures show the mean percentage of hematopoietic markers ($CD34^+$ $CD45^+$) on day 14. Statistical results are presented as mean ± standard error of the mean (SEM) and analysed using One-way ANOVA with multiple comparisons. p values were represented a ns for not significant. Data represent 4 independent experiments.

5.3 Differentiation potential of HSPC from hiPSC control, MDS27-C22 and CRISPR-Cas9 hiPSC clones in semi-solid medium

The data obtained so far seemed to indicate that the inclusion of the *C/EBPα* mutation did not have any effect on the iPSC. Still, it was important to define whether the mutation in *C/EBPα* would affect or exacerbate the hematopoietic differentiation phenotype observed in cells containing the mutations in *ASXL1*, *SRSF2* and *RUNX1*, and thus serve as a good model system in which to study disease progression. Thus, the differentiation potential and clonogenic capacity of HSPCs derived from hiPSC control, MDS27-C22, MDS27-C22+PX458**,** MDS27-C22.7 and MDS27-C22.20 hiPSCs was assessed by plating $1x10^4$ cells in methylcellulose medium as described previously in 4.2.2 (Figure 5.11, A). Fourteen days after plating, types and number of colonies were scored. HSPCs from hiPSC control, MDS27-C22 and MDS27-C22 +PX458, gave rise to both myeloid and erythroid CFUs whilst the clones containing the *C/EBPα* mutation, MDS27-C22.7 and MDS27-C22.20, exhibited a significant reduction in their clonogenic capacity affecting mainly the myeloid lineage and differentiating to erythroid lineage (Figure 5.11, B).

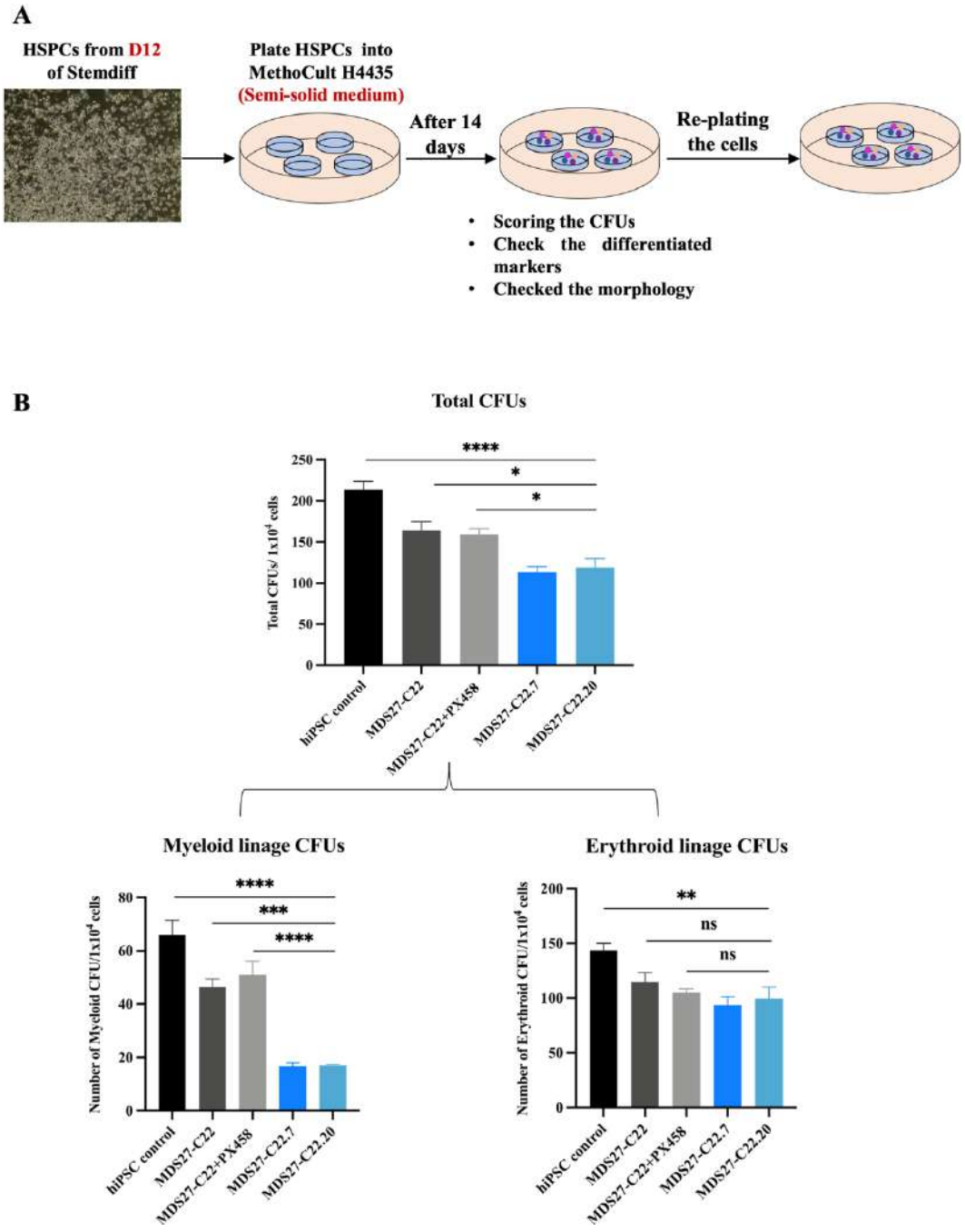


Figure 5.11: Further potential differentiation of HSPCs in semi-solid medium

(A) Schematic representation of the colony forming assay of HSPCs. **(B)** After 14 days in semi solid medium, the total number of CFUs was scored and then out from total number of CFUs, the number of myeloid lineage CFUs and erythroid lineage CFUs were identified. Statistical results are presented as mean ± SEM and analysed using One-way ANOVA with multiple comparisons. p values for total CFU number of mutant clones comparing with hiPSC control were **** <0.0001, comparing with MDS27-C22 were * 0.012 (C22.7) & 0.014 (C22.20), comparing with MDS27-C22+PX458 * 0.0113 (C22.7) & 0.025 (C22.20). p values for total myeloid lineage CFUs of mutant clones comparing with hiPSC control were **** <0.000, comparing with MDS27-C22 were *** 0.0001, comparing with MDS27-C22+PX458 were **** <0.0001. p values for total erythroid lineage CFUs of mutant clones comparing with hiPSC control were ** 0.0015 (C22.7) & 0.0041 (C22.20), comparing with MDS27-C22 & MDS27-C22+PX458 were ns. Data represent 4 independent experiments.

Interestingly, by scoring the type of CFUs, it was clear that mutant clones were not able to differentiate to CFU-G but were able mainly to form BFU-E (Figure 5.12, A). In addition, a high number of HSPCs of MDS27-C22.7 and MDS27-C22.20 were not able to form colonies in the methylcellulose and appeared as single cells. In addition, the percentage of each colony type was lower in MDS27 clones and mutant clones compared with hiPSC control (Figure 5.12, B).

Also, the size of the CFU-GEMM, CFU-GM and CFU-M was notably smaller in mutant clones than in control lines (hiPSC control, MDS27-C22 and MDS27-C22+PX458). In contrast, the size of BFU-E observed in *C/EBPα* mutant clones was larger than hiPSC control, MDS27-C22 and MDS27-C22+PX458 hiPSCs (Figure 5.13).

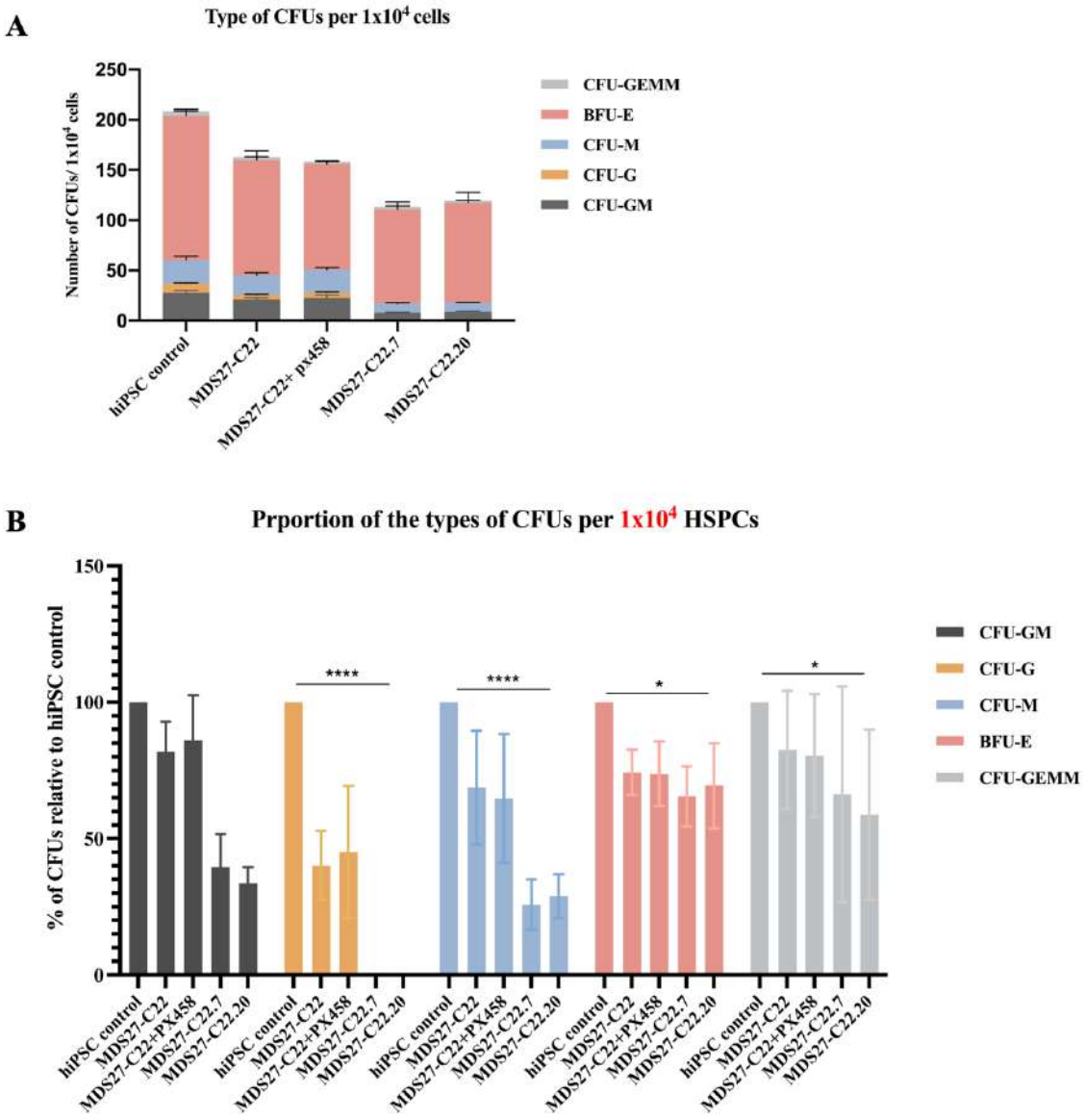


Figure 5.12: Type of CFUs

(A) Figure shows the mean number of the type of CFUs. A CFU-GEMM colony contains granulocytes, macrophages, erythrocytes and megakaryocytes. A CFU-GM colony contains granulocytes and macrophages, a CFU-M colony contains only macrophages and BFU-E colony contains erythrocytes. **(B)** The figure shows the proportion or relative percentage of each type of CFUs for $1x10^4$ HSPCs after 14 days in semisolid media. Statistical results are presented as mean ± SEM and analysed using Two-way ANOVA with multiple comparisons, *** $p < 0.0001$ and * p 0.08. Data represent 4 independent experiments.

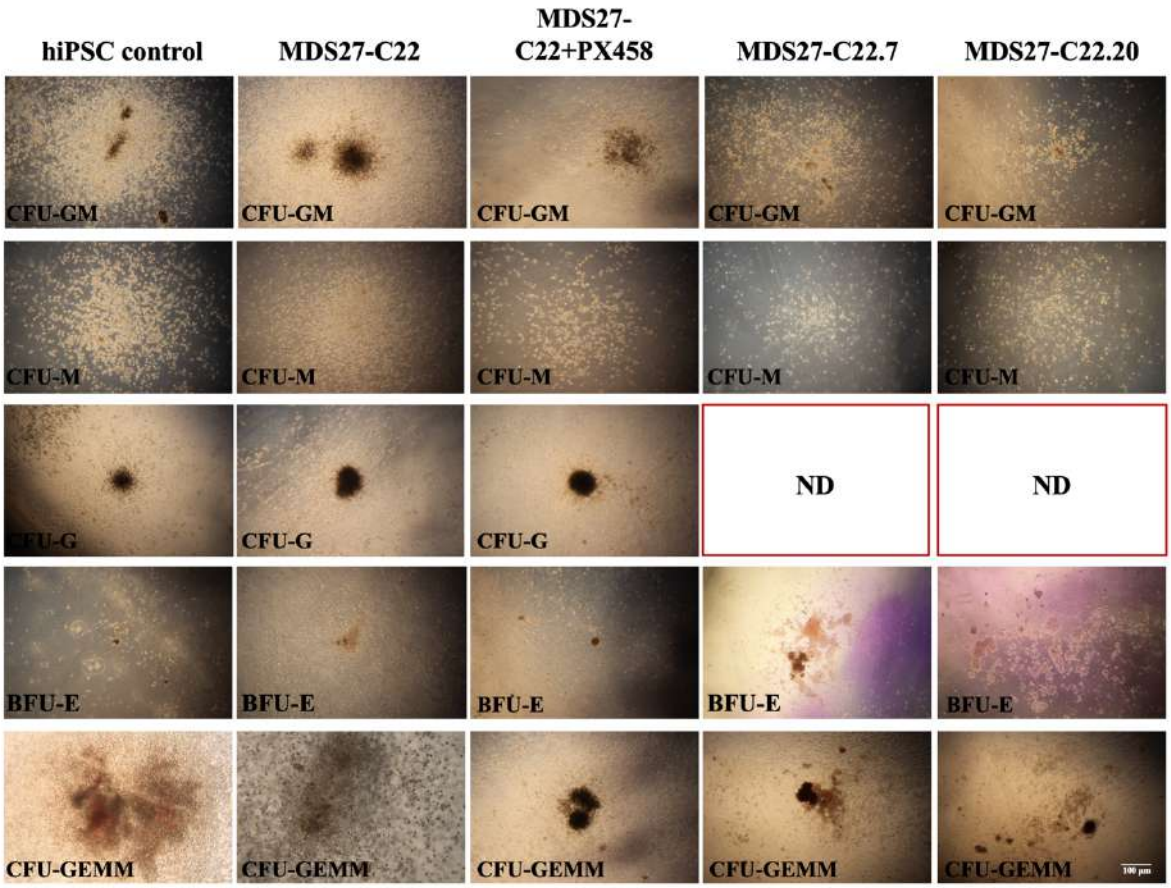


Figure 5.13: Morphology of CFUs

Phase contrast pictures show the morphology of CFUs for hiPSC control and MDS27 clones after 14 days in semi-solid medium. The pictures were taken by primo vert microscope (ZEISS) with Cannon camera at 4x magnification and size bar 100 µm scale bar . Data represent 4 independent experiments.

To corroborate that most of the CFUs in mutant clones were of the erythroid lineage, cells were collected from the methylcellulose media, washed several times with PBS, stained with myeloid markers (CD33, CD11b) and erythroid markers (CD71 and CD235a), and expression assessed by flow cytometry. This analysis revealed that *C/EBPα* mutant clones gave rise to a fewer percentage of myeloid cells composed mostly of immature cells $CD33^{+}CD11b^{+}$ (orange square) whilst the hiPSC control, MDS27-C22 and MDS27-C22+PX458 hiPSCs, were able to form mature myeloid $CD11b^{+}$ cell (green rectangle) (Figure 5.14, A&B).

Strikingly, the flow cytometry analysis confirmed that *C/EBPα* mutant clones, MDS27-C22.7 and MDS27-C22.20, could generate significantly erythroid colonies in methylcellulose culture

with approximately 45% of erythroblast cells ($CD71^{+}CD235a^{+}$, blue square) compared to 12.5% in control hiPSC, 14.4% in MDS27-C22 iPSCs and 18.5% MDS27-C22+PX458 iPSCs (Figure 5.14, A&B).

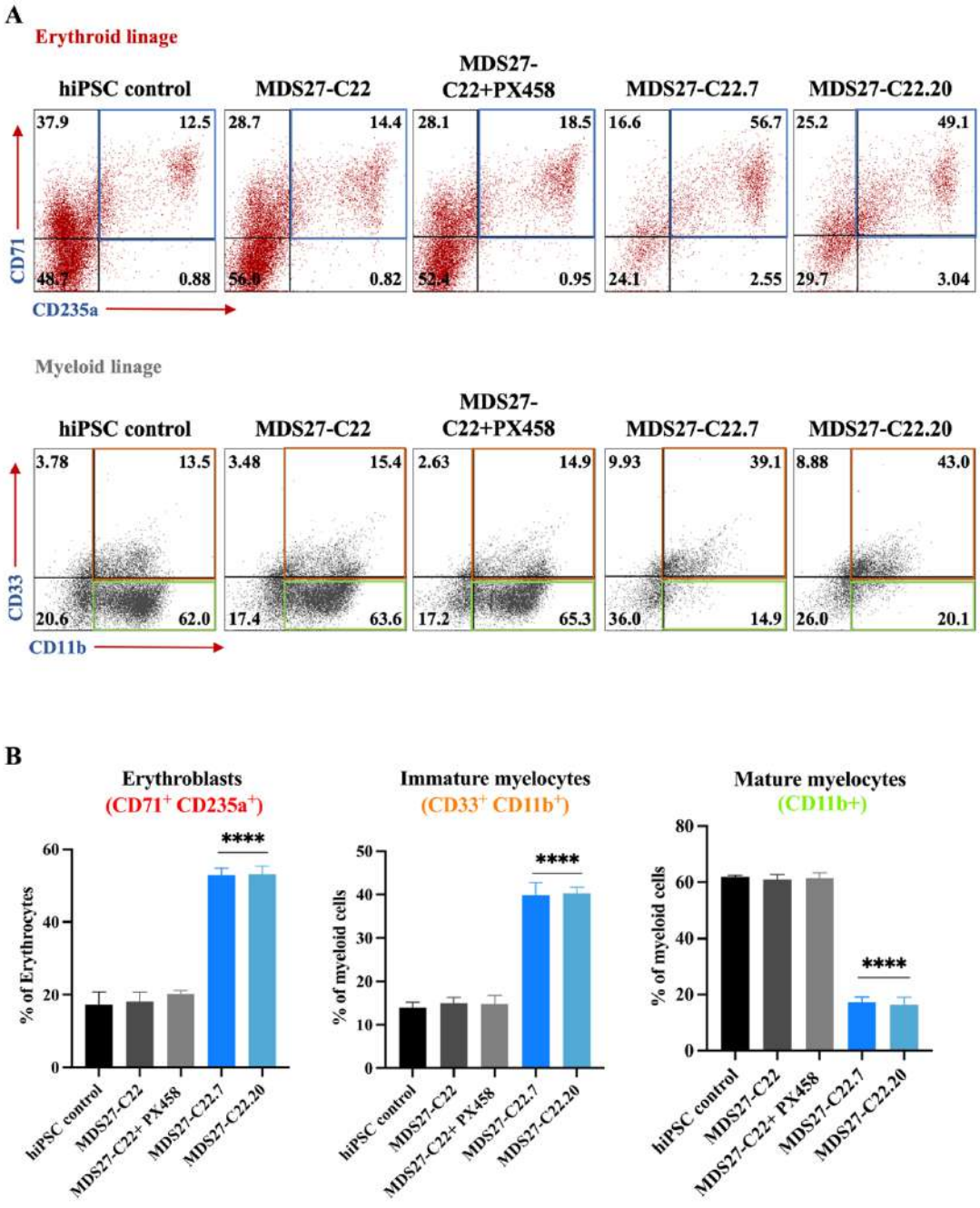


Figure 5.14: Characterisation of CFUs

(A) Erythroid markers (CD71 & CD235a) and myeloid markers (CD33 & CD11b) were monitored after 14 days from differentiation in methylcellulose medium. A blue square presents erythroblast cells ($CD71^+CD235a^+$), A orange square presents immature myeloid cells ($CD33^+$ $CD11b^+$), and a green rectangle presents mature myeloid cells ($CD11b^+$). The differentiating cells were analysed by flow cytometry, and the cells were gated first based on FSC and SS of unstained cells, and then the gated cells were analysed for hematopoietic markers based on the isotype control. The gating strategy shows in supplement figure 7. **(B)** Figures show the mean percentage of Erythroblasts markers ($CD71^+$ $CD235a^+$), immature myeloid markers ($CD33^+$ $CD11b^+$) and mature myeloid cells ($CD11b^+$). Statistical results are presented as mean ± standard error of the mean (SEM) and analysed using One-way ANOVA with multiple comparisons *** $p < 0.0001$. Data represent 3 independent experiments.

In summary, the morphological assessment of CFUs after 14 days in methylcellulose and the phenotypic assessment by flow cytometry showed that hiPSC control, MDS27-C22 hiPSCs and MDS27- C22+PX458 hiPSCs can differentiate and give rise to a mixture of myeloid and erythroid cells while the majority of cells in MDS27-C22.7 and MDS27-C22.20 hiPSCs were erythroblast cells (Figure 5.15).

Generally, these new hiPSC lines (MDS27-C22.7 and MDS27-C22.20) exhibited the characteristic of high-risk MDS in which the HSPCs have a blocked in differentiation into myeloid lineage and forced the HSPCs to differentiate mostly to erythroblast cells (Kotini et al., 2017).

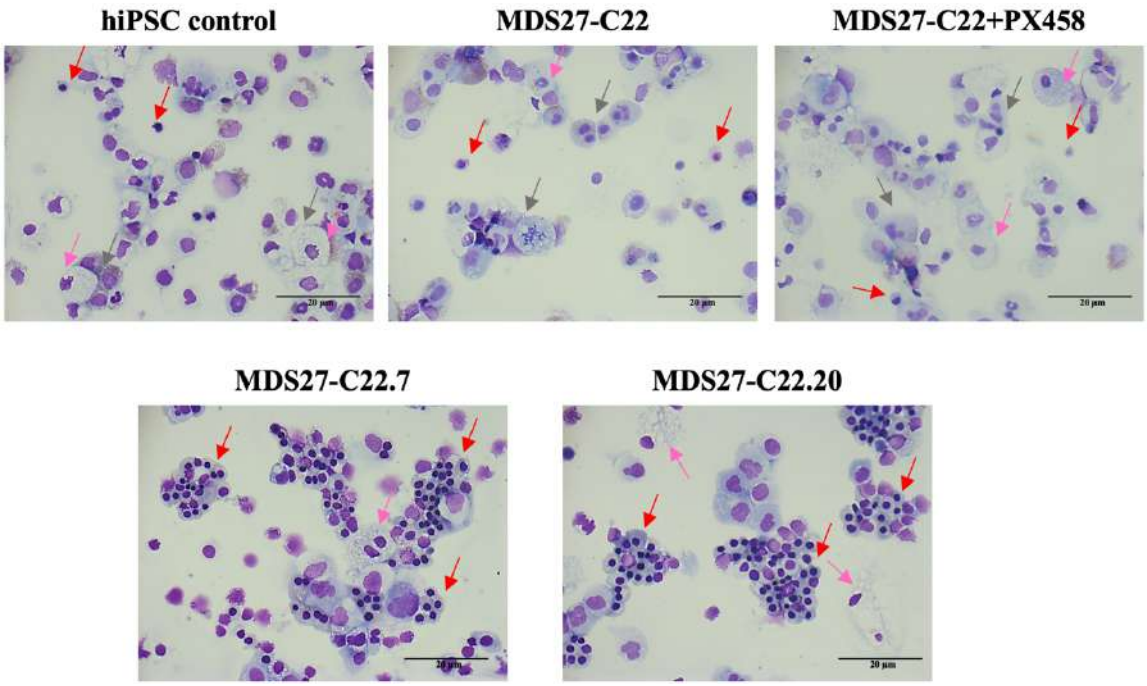


Figure 5.15: Morphology of CFUs
Cytospin of CFUs shows the erythroid cells (Red arrows) and granulocytes (Gray arrows) monocytes (Pink arrows). The pictures were taken Leica DM6000 at 40x, 20 µm scale bar. Data represent 3 independent experiments.

5.3.1 Investigating the role of *ASLX1, RUNX1, SRSF2* and *C/EBPα* on the proliferation and self-renewal capacity of MDS27 HSPCs

We further assessed the effect of *C/EBPα* mutation in combination with *ASLX1, RUNX1* and *SRSF2* mutations on self-renewal capacity by performing serial re-plating in methylcellulose. In the first plating, $1x10^4$ of HSPCs were plated in methylcellulose, and colonies were scored on the basis of their morphology after 14 days. Secondary re-plating was performed by collecting all the cells from the first plating, counting and seeding them in a new methylcellulose medium for an additional two weeks, after which colonies were scored and counted again. Remarkably, the *C/EBPα* mutation present in MDS27-C22.7 and MDS27-C22.20 hiPSCs resulted in the generation of BFU-E at a frequency of 264± and 243± colonies respectively in the second re-plating and around five colonies of CFU-GM (Figure 5.16, A).

In contrast, human iPSC control, MDS27-C22 iPSCs and MDS27-C22+PX458 hiPSCs, had very limited re-plating capacity with a frequency of 2 to 3 colonies of CFU-GM and no BFU-E colonies. In addition, when cells were collected from the plates and counted, a dramatic increase in the number of cells was observed in the *C/EBPα* mutant lines from first re-plating to second re-plating. This result was in contrast with the other iPSC cell lines (hiPSC control, MDS27-C22 and MDS27-C22+PX458) in which a decrease in the total number of cells was found from the first to the second re-plating. This data indicates that the *C/EBPα* mutation confer a greater proliferation capacity to HSPCs.

All the single cells from *C/EBPα* mutant clones were then plated again for the third re-plating for additional fourteen days. On day 14, the colonies derived from MDS27-C22.7 and MDS27-C22.20 iPSCs were scored, and it was observed that even though these cells were still able to

proliferate and give rise to colonies, the number of BFU-E had decreased in the third re-plating to 149 and 133 colonies, respectively. As before, cells were collected and although the number of cells had decreased from second to third re-plating, the cells were once more re-plated (fourth re-plating) to determine whether any cells could still proliferate and give rise to colonies. After 14 days, no CFUs had been formed and most of the cells had died (Figure 5.16, B).

Taken together, these results demonstrate that the *C/EBPα* mutation generated by CRISPR-Cas9 leads to functional consequences previously attributed to *C/EBPα* (i.e myeloid differentiation), validating our strategy. Our results also show that introducing the *C/EBPα* mutation to the previous mutations (*ASLX1, RUNX1, SRSF2*) induced the self-renewal capacity of the committed progenitor cells and blocks the myeloid differentiation, which is the main characteristic of high-risk MDS and sAML, indicating the successfully generation of high-risk lines by CRISPR-Cas9.

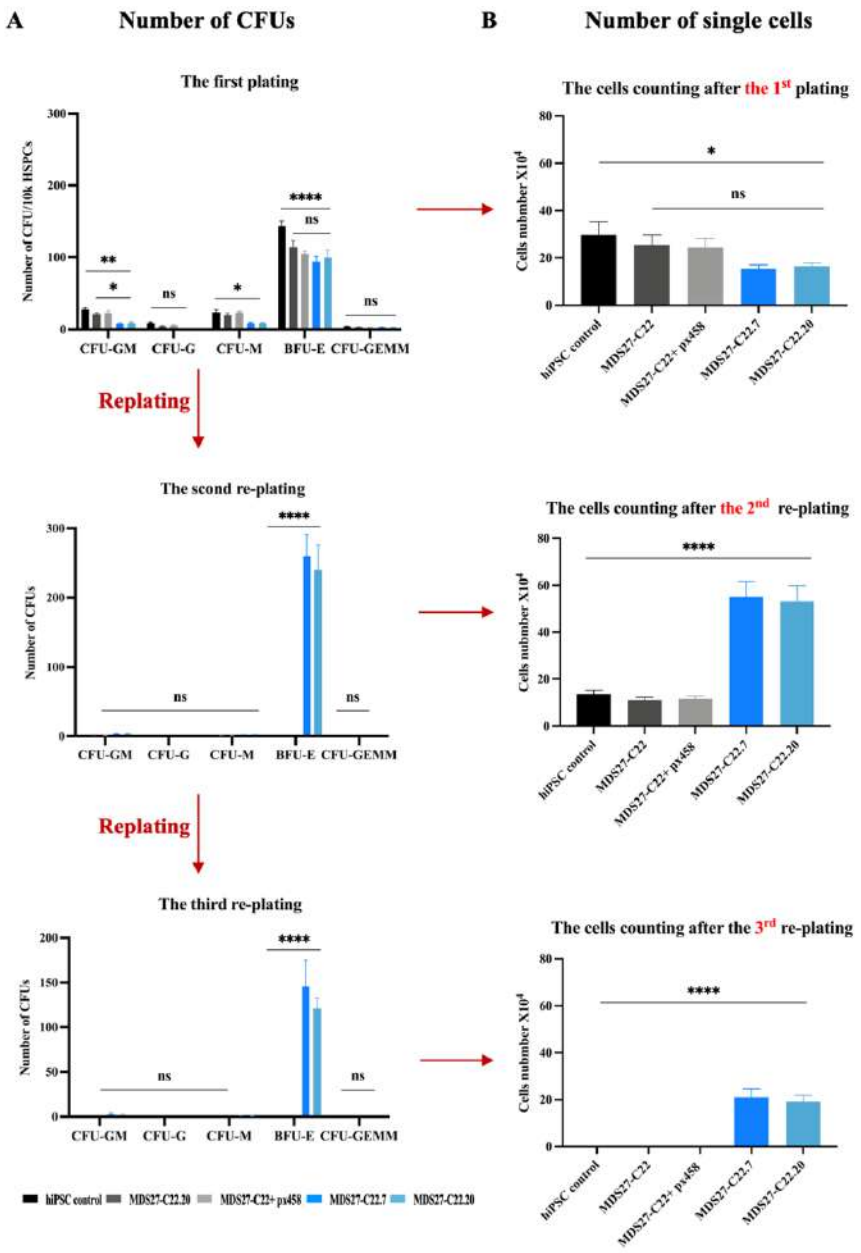


Figure 5.16: Self-renewal and re-plating capacity of CFUs

(A) CFUs obtained from WT clones and mutant clones were re-plated each 14 days for four times. The result indicates that mainly BFU-E was obtained at each generation. Statistical results are presented as mean ± SEM and analysed using Two-way ANOVA with multiple comparisons. p values for 1st plating comparing with hiPSC control (GM) were ** 0.0015 (C22.7) & 0.0024 (C22.20), (M) were * (0.0248), (E) were **** <0.0001* (mix) were ns, comparing with MDS27-C22 (GM) were * 0.0918 (C22.7) & 0.0954 (C22.20), (M) were * 0,0918 (C22.7 & C22.20), comparing with MDS27-C22+PX458 (GM) were * 0.0318 (C22.7) & 0.0454 (C22.20), (M) were * 0,0318 (C22.7 & C22.20). p-values for 2nd & 3rd re-plating comparing with hiPSC control, MDS27-C22 & MDS27-C22+PX458 were ns for all CFUs except (E) were **** <0.0001. **(B)** Shows the number of the single cells was obtained after each generation. Statistical results are presented as mean ± SEM and analysed using One-way ANOVA with multiple comparisons. p values for 1st plating comparing with hiPSC control were * 0,0435, comparing with MDS27-C22& MDS27-C22+PX458 were ns. p values for 2nd & 3rd re-plating comparing with hiPSC control, MDS27-C22& MDS27-C22+PX458 were **** <0.0001. Data represent 3 independent experiments.

5.3.2 Studying the myeloid differentiation potential of MDS27 hiPSC

As previously mentioned, *C/EBPα* mutation in combination with *ASLX1, RUNX1* and *SRSF2* mutations have affected the myeloid differentiation capacity of HSPCs in semisolid cultures (colony assay). Thus, to get a better understanding of how these mutations reduce the myeloid differentiation potential, we performed myeloid lineage differentiation in liquid culture. The myeloid differentiation was initiated by culturing the HSPCs in StemlineII medium with a cocktail of myeloid-induction cytokines (SCF, IL-3, GM-CSF, G-CSF, Flt3**)** for three days (Figure 5.17). Then, the cells were cultured for an additional four days, with G-CSF only, which constitutes the fundamental driving force for the generation, and differentiation of granulocytes (Welte et al., 1985).

The emergence of myeloid cells was assessed by flow cytometric analysis throughout the differentiation time course (Day 0, Day 4, and Day 7). The population of $CD11b^{+}CD14^{-}$ indicates the pure granulocytic cells and $CD11b^{-}$ $CD14^{+}$/$CD11b^{+}CD14^{+}$ represents monocytic cells. Furthermore, morphological analysis was performed to determine whether normal maturation of myeloid cells could be observed.

As shown in Figure 5.18 (A, B and C), by day 7 all the hiPSC lines without *C/EBPα* mutation were capable of producing a favourable percentage (around 40%) of the granulocytic population. In contrast, the *C/EBPα* mutant hiPSC lines (MD27-C22.7 and MDS27-C22.20) had a significant limited capacity to produce granulocytes and their differentiation seemed to be skewed to the production of monocytes.

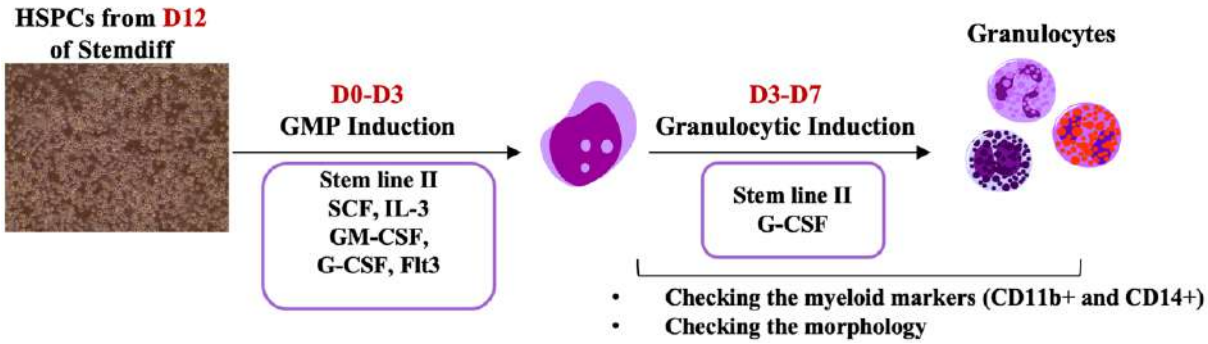


Figure 5.17: Myeloid differentiation of HSPCs

Schematic representation the myeloid differentiation of HSPCs.

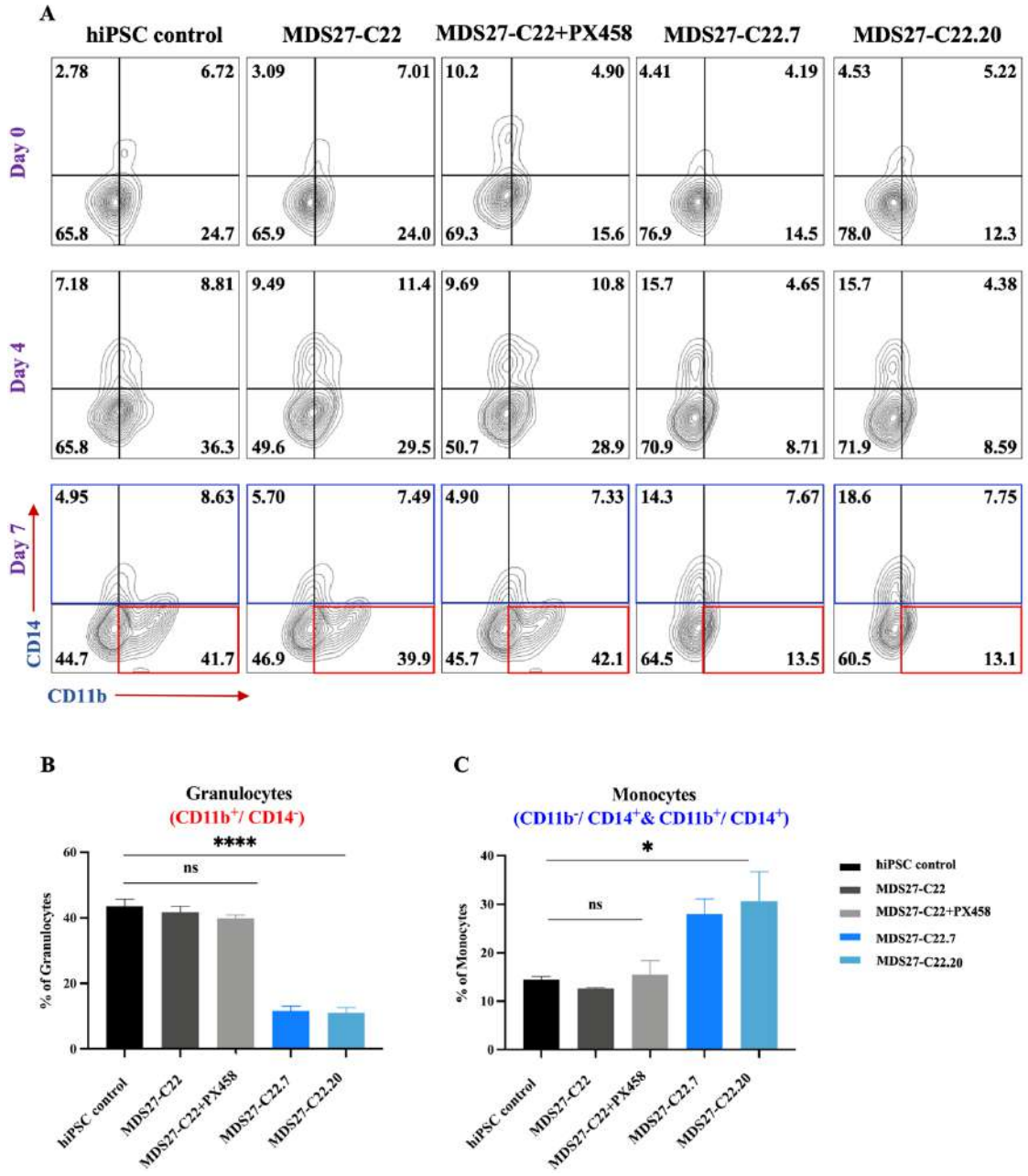


Figure 5.18: Assessment of myeloid differentiation

(A) $CD11b^+$ $CD14^-$ (granulocytes) and $CD11b^+$ $CD14^+$ (Monocytes) were monitored during the time course of differentiation Day 0, 4 and 7. The differentiating cells were analysed by flow cytometry and the cells were gated first based on FSC and SS of unstained cells and then the gated cells were analysed for myeloid markers based on the isotype control (supplement figure 8). **(B&C)** Figures show the mean percentage of $CD11b^+$ $CD14^-$ (granulocytes) and $CD11b^-$ $CD14^+$/$CD11b^+$ $CD14^+$ (Monocytes) on day 7. Statistical results are presented as mean ± standard error of the mean (SEM) and analysed using One-way ANOVA with multiple comparisons. p values for granulocytes of mutant clones comparing with hiPSC control, MDS27-C22& MDS27-C22+PX458 were **** <0.0001. p values for monocytes of mutant clones comparing with hiPSC control were * 0.0356 (C22.7) & 0.0118 (C22.20), MDS27-C22 were * 0.0169 (C22.7) & 0.0369 (C22.20), comparing with MDS27-C22+PX458 were * 0.0265 (C22.7) & 0.0187 (C22.20). Data represent 4 independent experiments.

Then, we sought to determine the morphology of the myeloid cells by light microscopy to get a better insight into proper maturation features such as granule content, morphology of the nucleus and cytoplasm/nucleus ratio. Interestingly, the committed eosinophil and neutrophil cells were identified within the myeloid population of hiPSC control, MDS27-C22 and MDS27-C22+PX458 hiPSCs, accompanied with appropriate segmented nucleus morphology and granules (Figure 5.19, grey arrows). However, the granulocytes were not detected in the myeloid cultures derived from MD27-C22.7 and MDS27-C22.20 hiPSCs, which proved the phenotypic disease detected by flow cytometry.

In addition, the morphological assessment of myeloid cells derived from MD27-C22.7 and MDS27-C22.20 hiPSCs showed that mutant clones differentiated into a heterogenous population that contained mainly erythrocytes (Figure 5.19, red arrows).

Interestingly, the morphological results of myeloid cells of WT lines (MDS27-C22 and MDS27-C22+PX458) revealed dysplastic morphology such as hypo-segmentations, ring nucleus, and hypo-granularity (Figure 5.20). These aberrant morphologies have been reported previously in MDS patients (Tohyama, 2018).

Overall, the *C/EBPα* mutation in the context of *ASLX1, RUNX1* and *SRSF2* mutations leads to an inhibition of the granulocytic differentiation and an increase in the erythroid lineage. However, the original mutations *ASLX1, RUNX1, SRSF2* are responsible of the dysplastic morphology observed in myeloid cells.

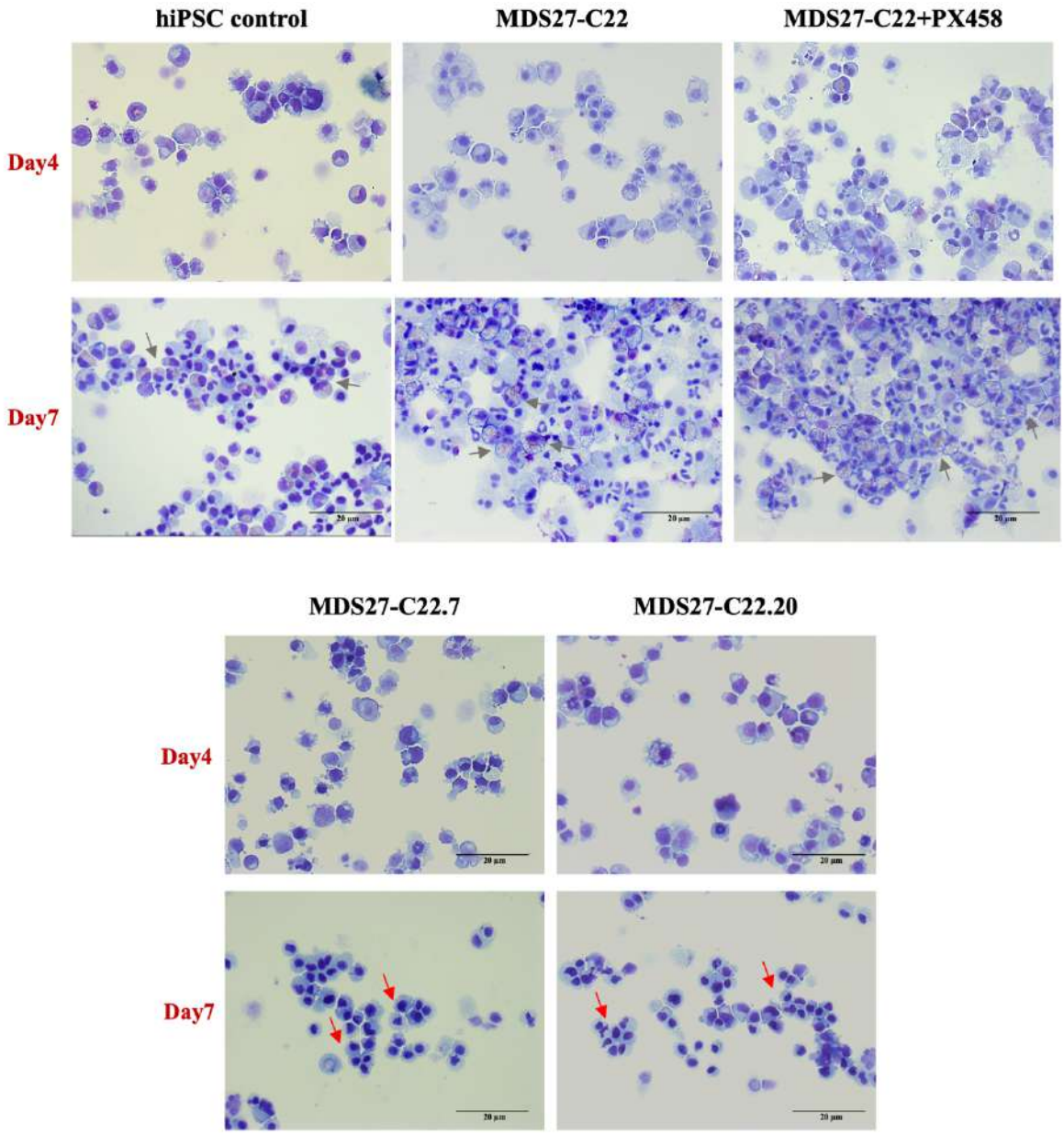


Figure 5.19: Morphological analysis of myeloid cells

Cytospin of myeloid cells indicates the presence of granulocytic cells (gray arrows) in hiPSC control, MDS27-C22 and MDS27-C22+PX458 and absence of granulocytes in the mutant clones and most of the population are erythroblasts (red arrows). The pictures were taken by Leica DM6000 microscope at 40x magnification, 20 µm scale bar. Data represent 4 independent experiments.

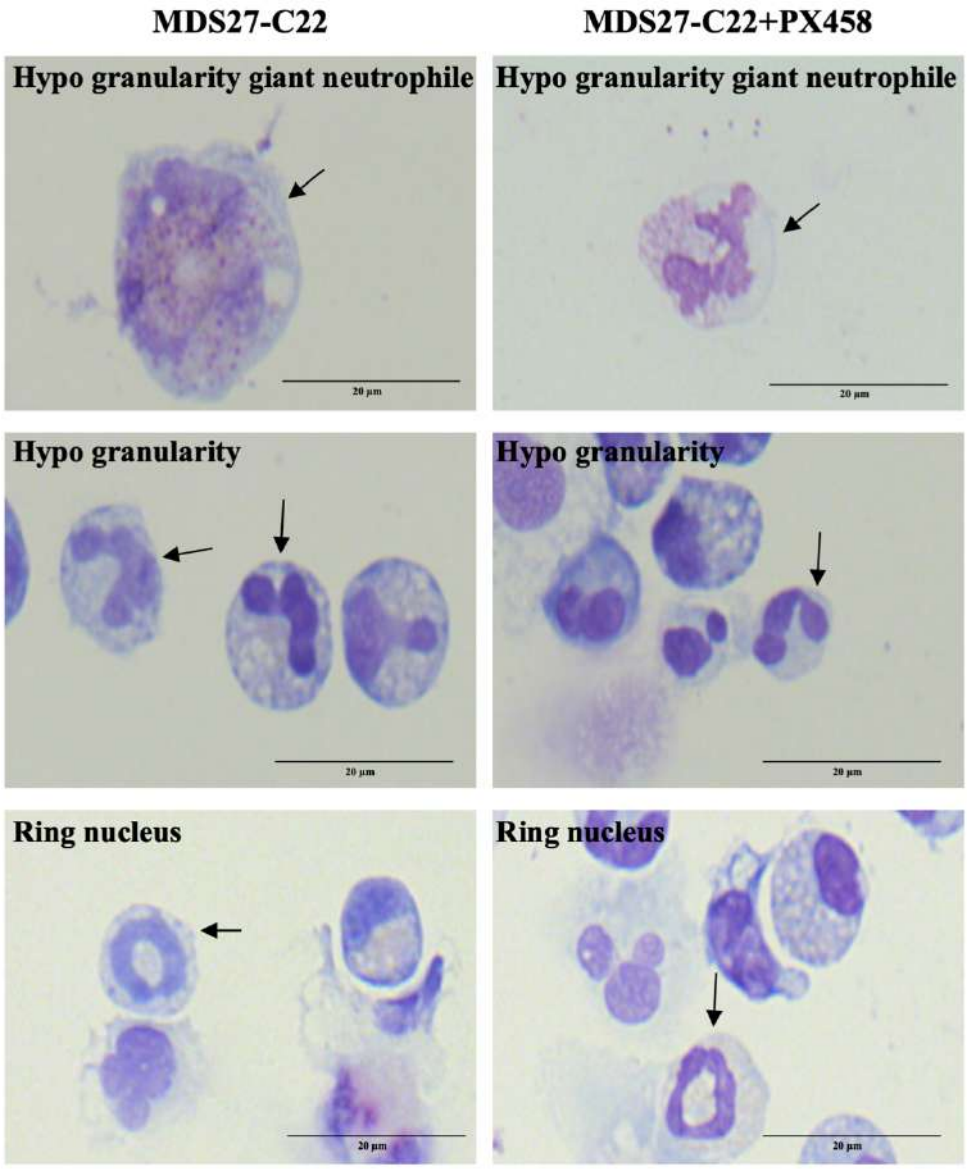


Figure 5.20: Morphological analysis for aberrant myeloid cells

The common aberrant morphology of myeloid cells (black arrows) in MDS27-C22 and MDS27-C22+PX458. The pictures were taken by Leica DM6000 microscope at 100x magnification, 20 µm scale bar. Data represent 4 independent experiments.

5.3.3 Investigating the expression of *C/EBPα* target genes

To further corroborate that the *C/EBPα* mutation in its DNA binding domain generated by CRISPR-Cas9 was indeed affecting the functionality of *C/EBPα*, we aimed to determine the expression levels of several genes essential for haematopoiesis and granulocytic differentiation which are known to be regulated by *C/EBPα*.

RNA was extracted from HSPCs at day 12 and myeloid cells from day 4 and day 7 derived from hiPSC control, MDS27-C22, MDS27-C22+PX458, MDS27-C22.7 and MDS27-C22.20 hiPSCs. Then, the RNA was treated with DNase to remove any gDNA contamination from RNA samples. cDNA was synthesised, and the expression of the genes checked by qRT-PCR (Figure 5.21, A).

As shown in Figure 5.21, (B) *C/EBPα* mRNA expression in the myeloid cells derived from MDS27-C22.7 and MDS27-C22.20 iPSC lines was reduced from day 4 and day 7 compared to the WT iPSC lines. In addition, all myeloid cells expressed a low level of *RUNX1* in all MDS27-hiPSC derived lines compared to hiPSC control. This could be the consequence of the two *RUNX1* mutations found in the MDS27 patient cells and derived iPSC clones, as one or both of these mutations could affect the expression of *RUNX1* mRNA.

PU.1 and *GATA2* are other TFs that were checked in myeloid and HSPCs derived from the hiPSC. The qRT-PCR result demonstrated the downregulation of the expression of *PU.1* and *GATA2* in myeloid cells derived from the *C/EBPα* mutant iPSC lines.

Lastly, *LMO2* was shown to be upregulated in all MDS27 WT iPSC lines and *C/EBPα* mutant iPSC lines by day 4 of myeloid differentiation compared to the iPSC control line. However, the expression of *LMO2* in the *C/EBPα* mutant lines showed a further increase by day 7. As *LMO2* has been shown to induce the erythroid differentiation (Hansson et al., 2007), the increase in

LMO2 expression in our cultures could explain the increase of erythroid cells during the myeloid differentiation.

Overall, the qRT-PCR data indicated that *C/EBPα* mutation leads to deregulation of a TF signature resulting in the inhibition of myeloid differentiation and induction of the transition to the erythroid lineage.

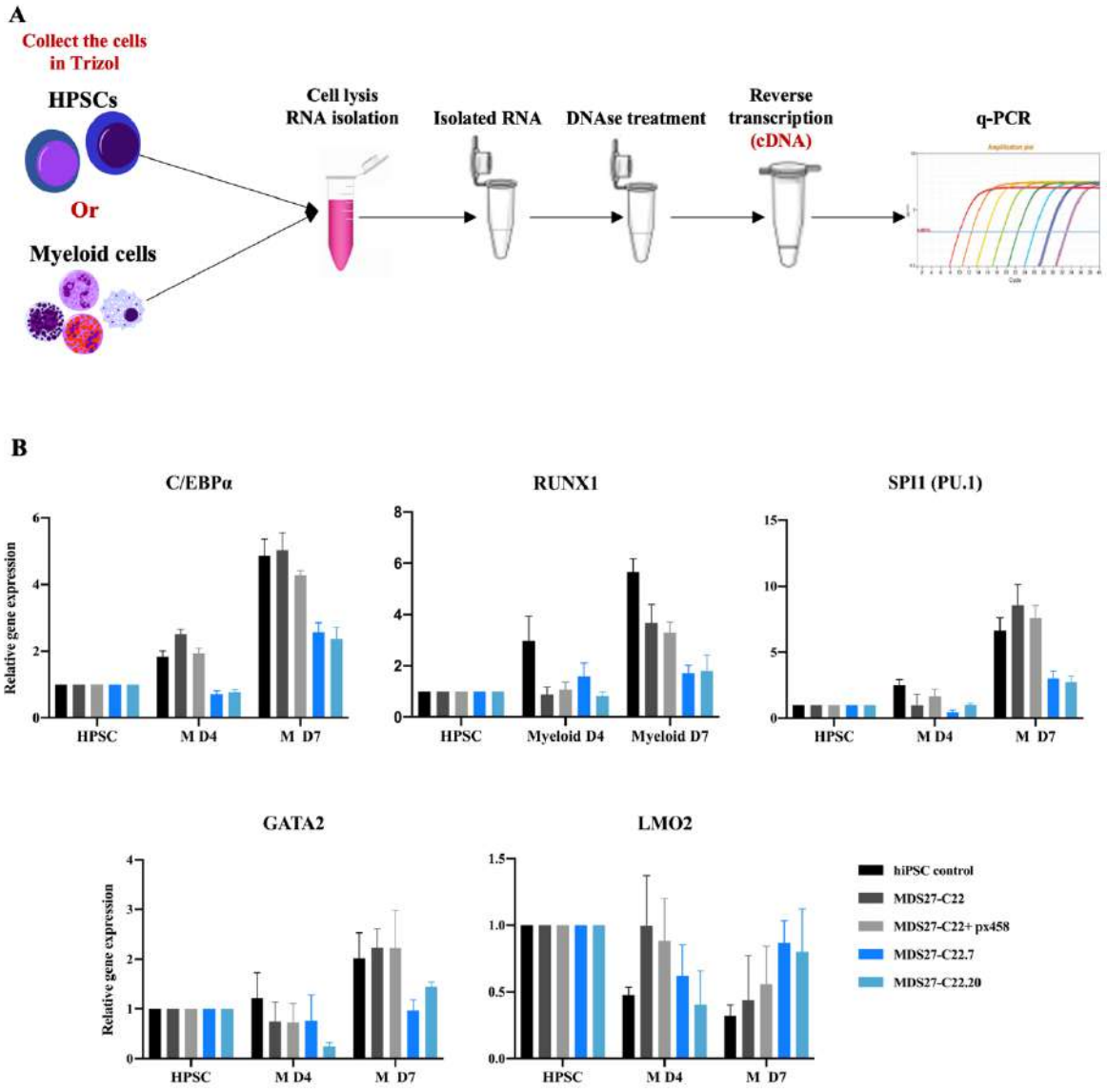


Figure 5.21: Gene expression in HSPCs and myeloid cells

(A) Schematic representation of the RNA extraction from HSPCs and myeloid cells. **(B)** Histogram representing the results of qRT-PCR of HSPCs and myeloid cells RNA from hiPSC control, MDS27-C22, MDS27-C22-PX458, MDS27-C22.7 and MDS27-C22.20 using primers designed to amplify *C/EBPα, RUNX1, SPI1 (PU.1), GATA2* and *LMO2*. The relative expression was calculated by normalising to *GAPDH* housekeeping gene, then expressed relatively to HSPCs expression. Data represent 4 independent experiments.

5.3.4 Studying the erythroid differentiation potential of MDS27 hiPSC containing the C/EBP*α* mutation

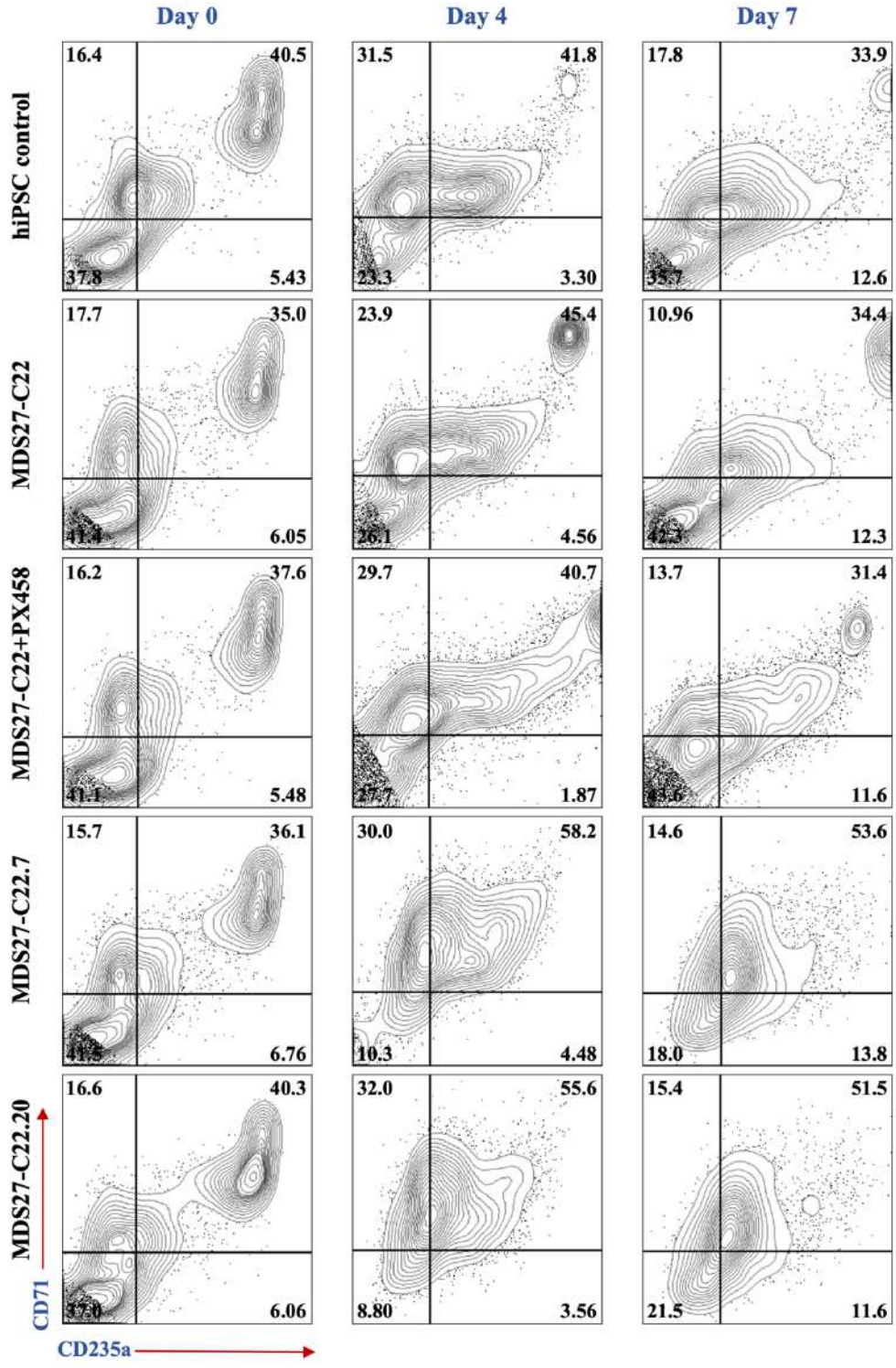
The colony assays results showed that the *C/EBPα* mutant MDS27-iPSC lines were proficient in the formation of erythroid colonies (Figure 5.11, B). Still, the formation of mature erythrocytes could be affected in the same way as granulocyte formation. Thus, we sought to determine whether mature erythrocytes could be formed in the mutant lines following the same erythroid differentiation protocol.

We determined the erythroid differentiation kinetics by measuring the erythroid markers CD71 and CD235a. The use of these two markers allows us to follow the progression of erythroid differentiation from erythroid progenitor cells (CD71+), to erythroblasts (CD71+ CD235a+) and mature erythrocytes (CD71- CD235a+). As shown in Figure 5.22, the kinetics of erythroid differentiation was similar in hiPSC control and MDS27-C22 and MDS27-C22+PX458 PSC lines, with similar percentages of erythrocyte progenitors, erythroblasts and mature erythrocytes at any given time point. In contrast, the *C/EBPα* mutant MDS27-hiPSC lines displayed a greater percentage of erythroblasts ($CD71^+$ $CD235^+$) at day 4 and day 7 of differentiation compared to any of the control cells.

At later timepoints, as expected, a decrease in the percentage of erythroblasts with a concomitant increase in mature cells was observed. However, the *C/EBPα* mutant MDS27-hiPSC lines still expressed a noticeable percentage of $CD71^+$ $CD235^+$ at the end of the differentiation. For example, on day 18, the cultures of hiPSC control and MDS27 WT lines contained mainly mature cells $CD235a^+$ >55% and 5% of erythroblasts ($CD71^+$ $CD235a^+$), whilst MDS27-C22.7 and MDS27-C22.20 cultures still contained 12% of erythroblasts by day 18. Surprisingly, despite the *C/EBPα* mutant MDS27-hiPSC lines displaying twice as many erythroblasts than control lines at day 18, the statistical analysis revealed this not to be

significantly different (Figure 5.23, A). Additionally, we determined the geometric mean (Mean Fluorescent Intensity (gMFI)) of CD71 expression on day 18 to check the intensity of CD71 expression on the terminal stages of the erythroid differentiation. As shown in Figure 5.23, (B) the intensity of CD71 expression was significantly higher in *C/EBPα* mutant MDS27-iPSC lines compared with hiPSC control, MDS27-C22 and MDS27-C22+PX458 hiPSCs.

Interestingly, this result indicates that the addition of *C/EBPα* mutation to cells harbouring *ASXL1, RUNX1 and SRSF2* mutations impaired the termination of the erythroid differentiation.


Day 0
Day 4
Day 7
hiPSC control
MDS27-C22
MDS27-C22+PX458
MDS27-C22.7
MDS27-C22.20
16.4
40.5
37.8
5.43
31.5
41.8
23.3
3.30
17.8
33.9
35.7
12.6
17.7
35.0
41.3
6.05
23.9
45.4
26.1
4.56
10.96
34.4
42.3
12.3
16.2
37.6
41.1
5.48
29.7
40.7
27.7
1.87
13.7
31.4
43.6
11.6
15.7
36.1
41.5
6.76
30.0
58.2
10.3
4.48
14.6
53.6
18.0
13.8
16.6
40.3
37.0
6.06
32.0
55.6
8.80
3.56
15.4
51.5
21.5
11.6
CD71
CD235a

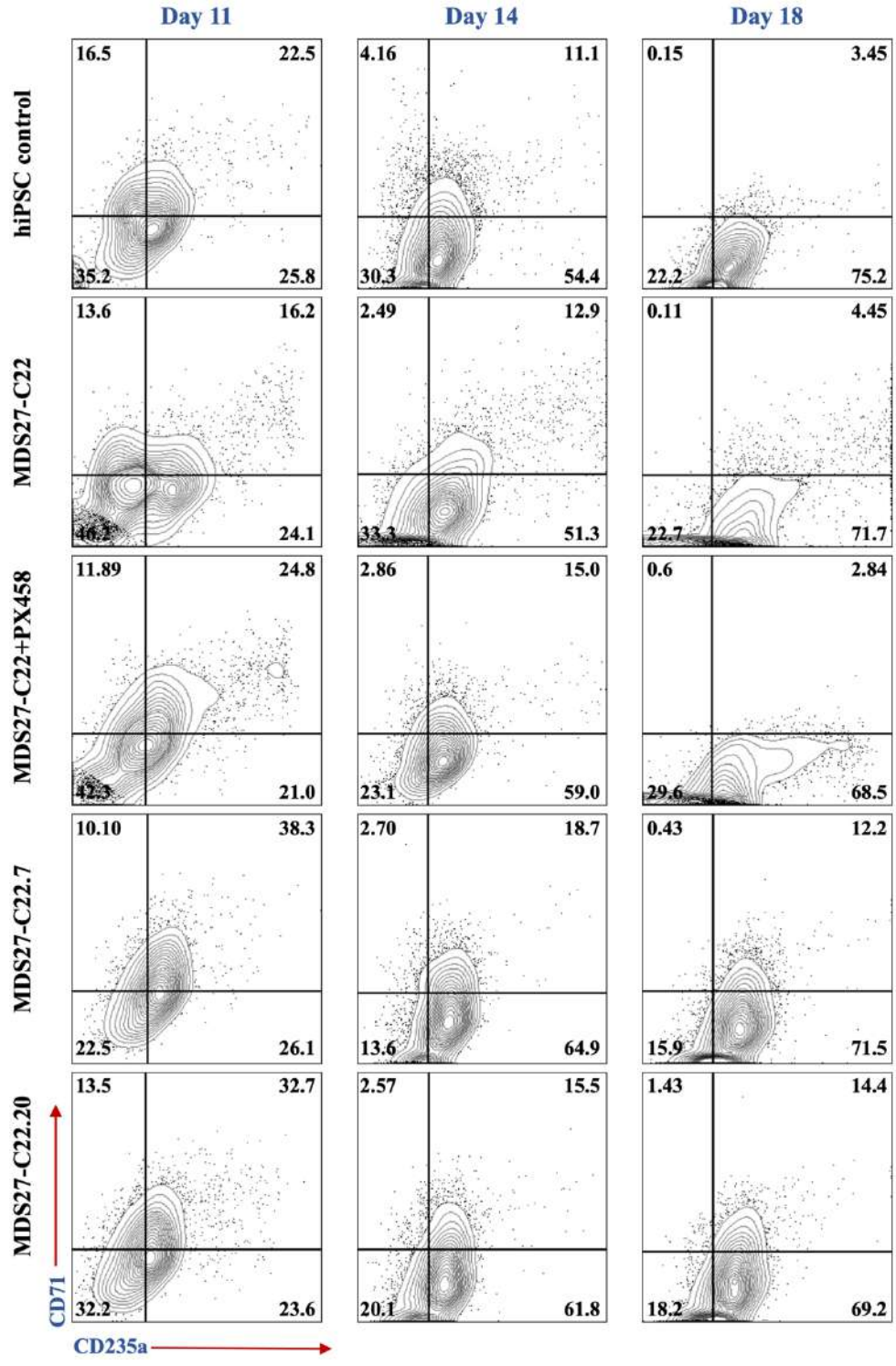


Figure 5.22: Erythroid differentiation of HSPCs

CD71 and CD235a erythroid markers were monitored during the time course of differentiation on Day 0, 4, 7, 11, 14 and 18. The differentiating cells were analysed by flow cytometry and the cells were gated first based on FSC and SSC Cells of unstained cells and then analysed for erythroid markers based on the isotype control. Data represent 4 independent experiments.

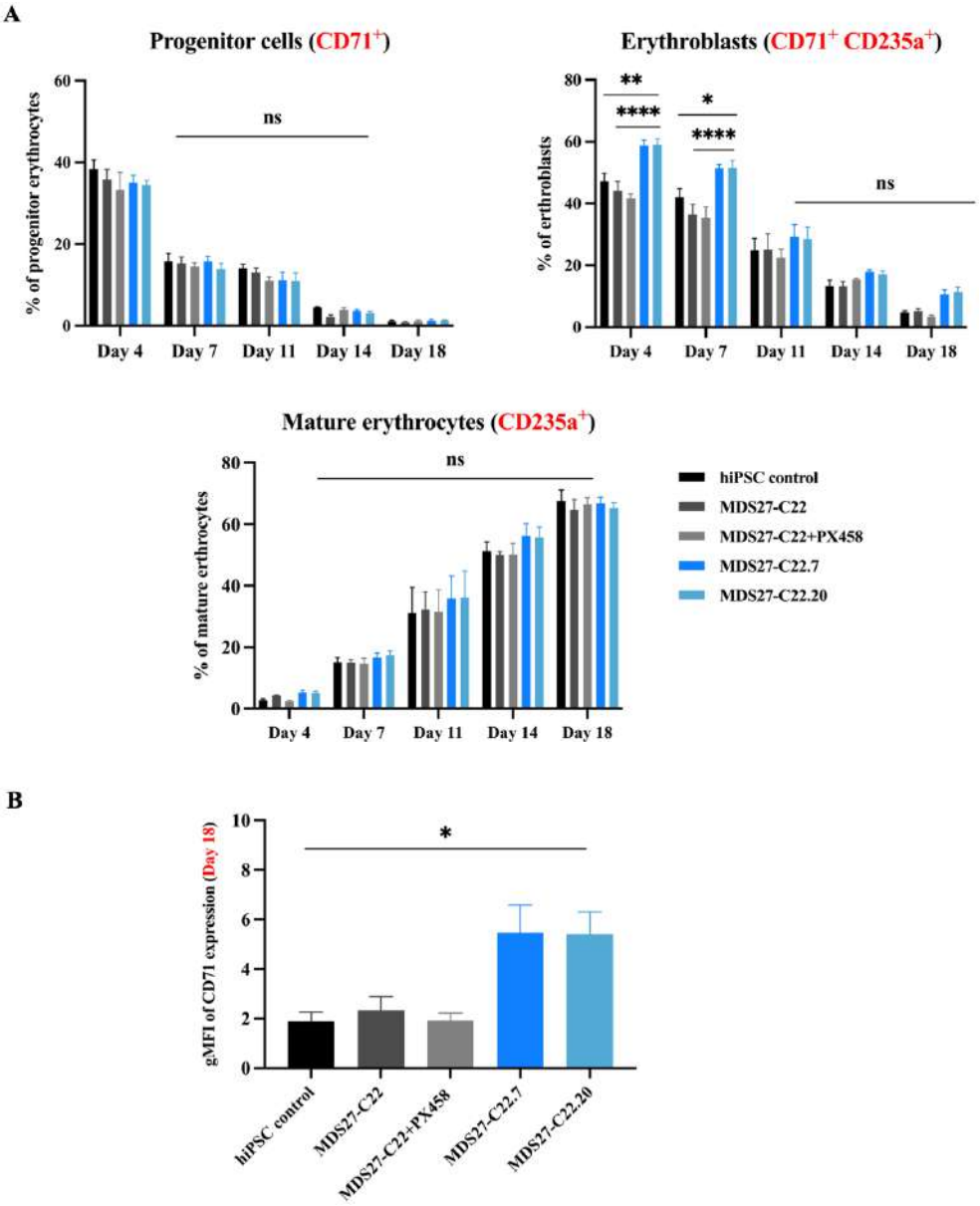


Figure 5.23: Assessment of erythroid differentiation markers

(A) The mean percentage of erythroid markers progenitor erythrocyte (CD71+), Erythroblasts (CD71+ CD235a+) and mature erythrocytes (CD71- CD235a+) during several time point of erythroid differentiation. Statistical results are presented as mean ± SEM and analysed using Two-way ANOVA with multiple comparisons. p-values for progenitor cells comparing with hiPSC control, MDS27-C22& MDS27-C22+PX458 were ns. p-values for erythroblasts comparing with hiPSc control on day 4 were ** 0.0054 (C22.7) & 0.0046 (C22.20), On day7 were * 0.0296 (C22.7) & 0.0264 (C22.20). p-values for erythroblasts comparing with MDS27-C22 and MDS27-C22+PX458 on day 4 & 7were **** <0.0001. p values for erythroblasts comparing with hiPSC control, MDS27-C22 and MDS27-C22+PX458 on day 11, 14 & 18 were ns. p values for mature cells comparing with hiPSC control, MDS27-C22 and MDS27-C22+PX458 were ns. **(B)** The geometric mean (gMFI) of each line at day 18 of erythroid differentiation was obtained for hiPSC control, MDS27-C22 and MDS27-C22+PX458 with MDS27-C22.7 & MDS27-C22.20 Statistical results are presented as mean ± SEM and analysed using One-way ANOVA with multiple comparisons. p values comparing with hiPSC control were * 0.0104 (C22.7) & 0.0117 (C22.20), comparing with MDS27-C22 were * 0.0243 (C22.7) & 0.0274 (C22.20), comparing with MDs27-C22+PX458 were * 0.0111 (C22.7) & 0.0126 (C22.20). Data represent 4 independent experiments.

To get a better understanding of the erythroid differentiation process, we performed a morphological analysis to identify the stages of erythroid maturation during the time course of the differentiation (Figure 5.24 and Figure 5.25).

Five hundred erythroid cells were scored from each line on different days. In contrast to the flow cytometry results, the morphological analysis showed significantly high percentages of mature cells in the *C/EBPα* mutant cell MDS27-iPSC lines in the early stage of differentiation (day 4 and day 7) compared with hiPSC control and MDS27-C22 and MDS27-C22+PX458 hiPSCs (Figure 5.24).

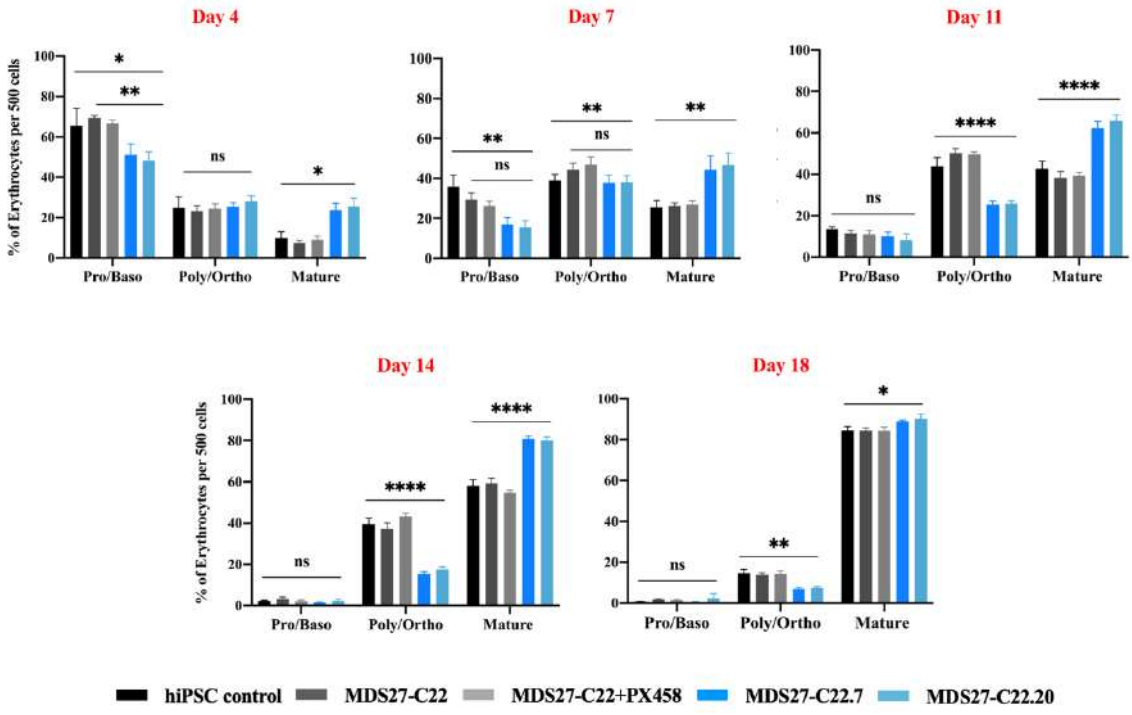


Figure 5.24: Scoring erythroid cells

Scoring 500 erythroid cells according to the stage of maturation and the percentage of each type of the cells was obtained. Results are presented as mean ± SEM and analysed using Two-way ANOVA with multiple comparisons. p value of each stage and day represents in Supplement figure 9. The data represented 4 independent experiments. Pro/Baso (Proerythrocyte/basophilic erythrocyte), Poly/Ortho (Polychromatic erythrocyte/ Orthochromatic erythrocyte).

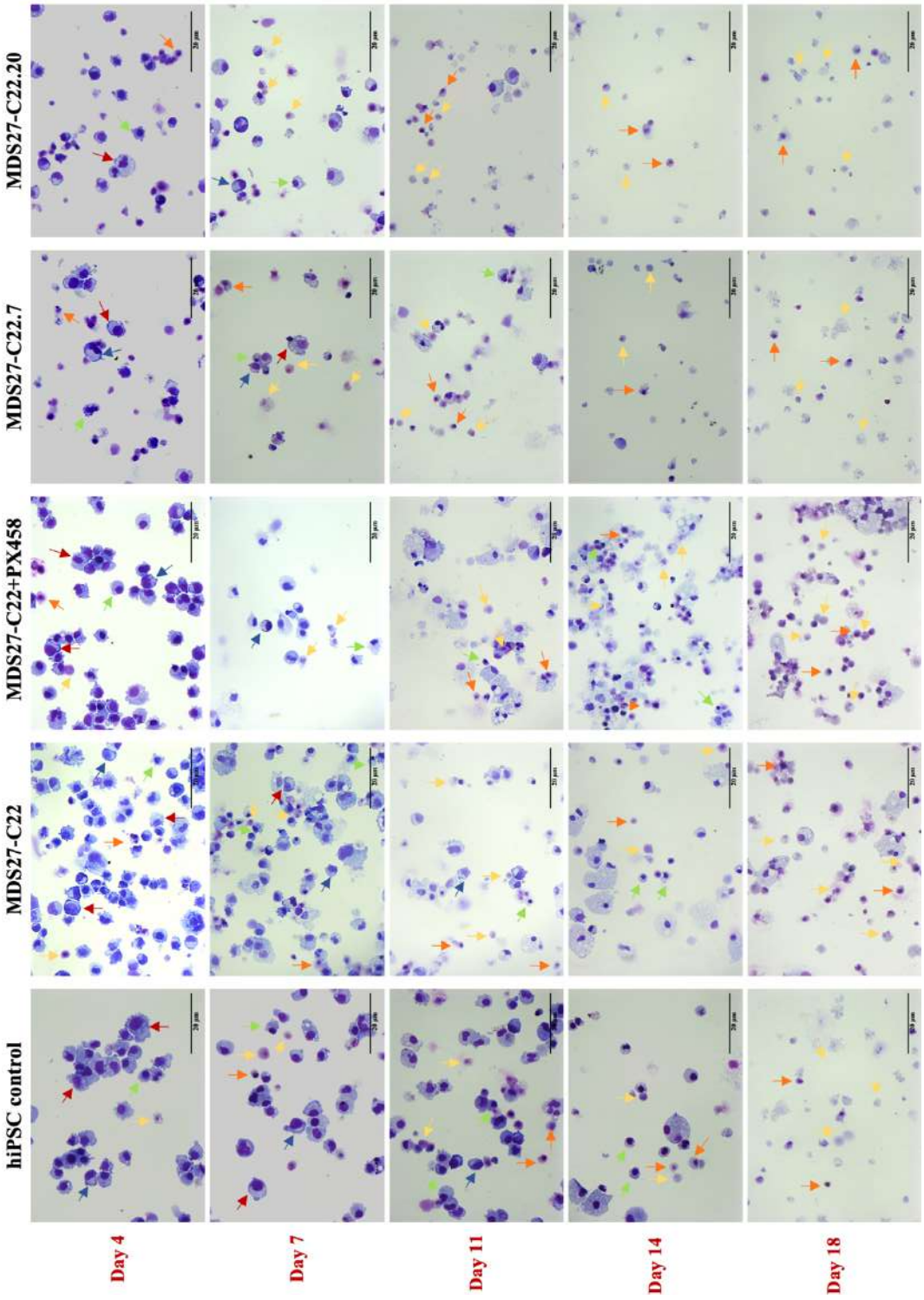


Figure 5.25: Erythroid cell morphology

Cytospin of erythroid cells shows the maturation stages during different days of the culture: Proerythroblast (red arrow), Basophilic erythroblast (blue arrow), Polychromatic erythroblast (green arrow), Orthochromatic erythroblast (orange arrow) & Mature cells (yellow arrow). The pictures were taken Leica DM6000 at 40x magnification, 20 µm scale bar. The data represented 4 independent experiments.

To confirm the morphological analysis result, we stained the erythroid cells derived from hiPSC control, MDS27-iPSCs lines (MDS27-C22, MDS27-C22+PX458) and *C/EBPα* mutant iPSC lines (MDS27-C22.7 and MDS27-C22.20) with DRAQ5 to evaluate the enucleation rate. DRAQ5 is an anthraquinone, a dye with high affinity for double-stranded DNA that can be used to distinguish nucleated and non-nucleated cells (Samsel and McCoy, 2015). The assessment of DRAQ5 staining was done by flow cytometry during the course of the differentiation (Day 0, 4, 7, 11, 14 and 18). Interestingly, our data showed that the enucleation rate was remarkably high in *C/EBPα* mutant iPSC lines compared to wild type iPSC lines, and that the differentiation towards the mature cells was very fast at the early days of differentiation (Figure 5.26). For example, on day 7 the percentage of nucleated cells in hiPSC clones and MDS27 WT iPSC lines was around 44%, while in *C/EBPα* mutant MDS27-iPSC lines decreased to 30%.

In addition, the statistical comparison of hiPSC control and MDS27 WT iPSC lines with *C/EBPα* mutant MDS27-iPSC lines indicated that the enucleation rate was significantly high in mutant clones during the early days of the differentiation (Figure 5.27).

This result agrees with the morphological assessment, showing that during erythroid differentiation the appearance of mature cells is faster in the *C/EBPα* mutant MDS27-iPSC lines.

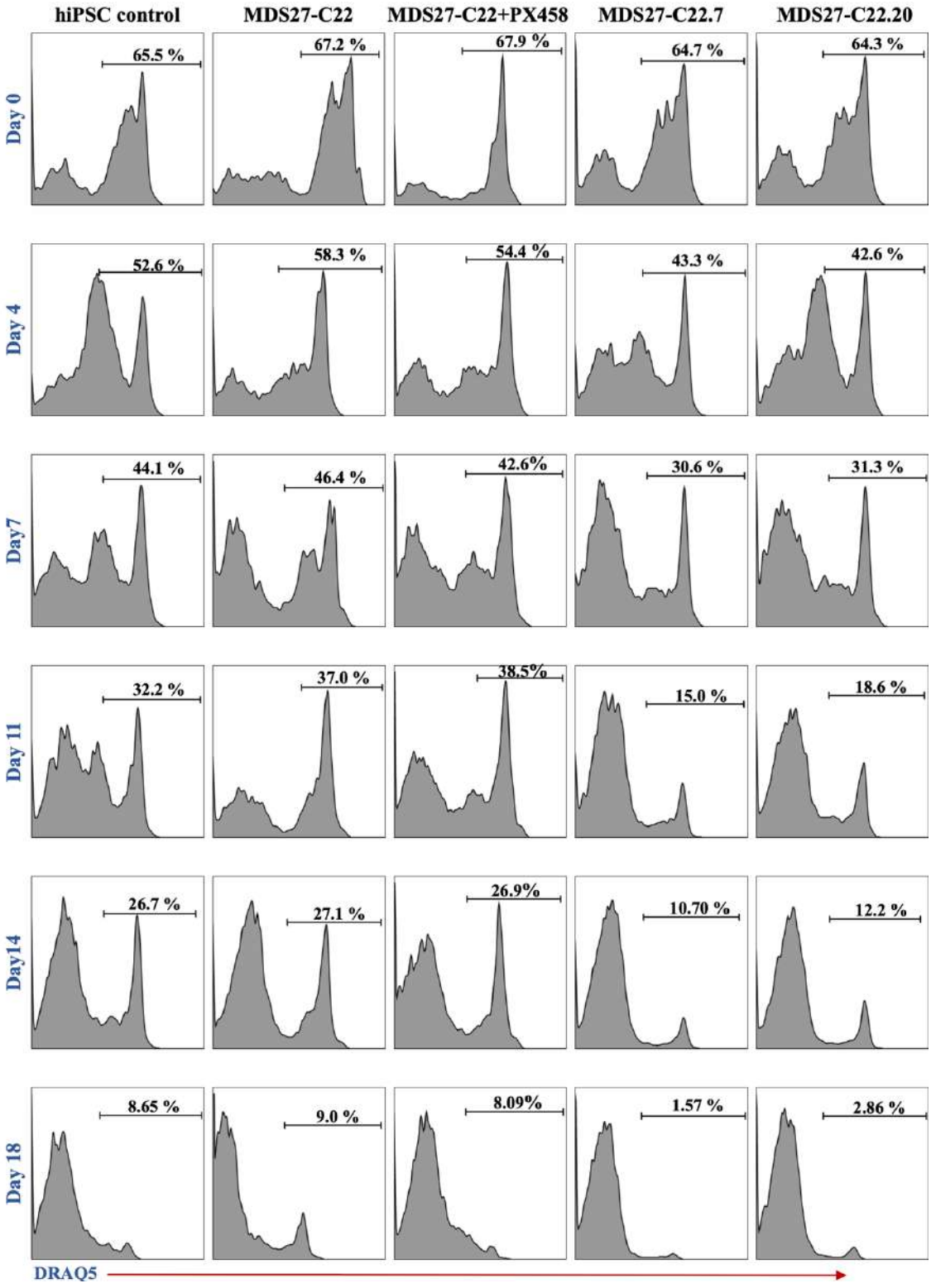


Figure 5.26: DRAQ5 staining for nucleated erythrocytes

Enucleation rates confirmed by DRAQ5 staining. Representative histogram figures show the percentage of nucleated erythrocytes. The DRAQ5 staining were analysed by flow cytometry and the cells were gated first based on FSC and SSC Cells of unstained cells and then the gated cells were analysed for DRAQ5.

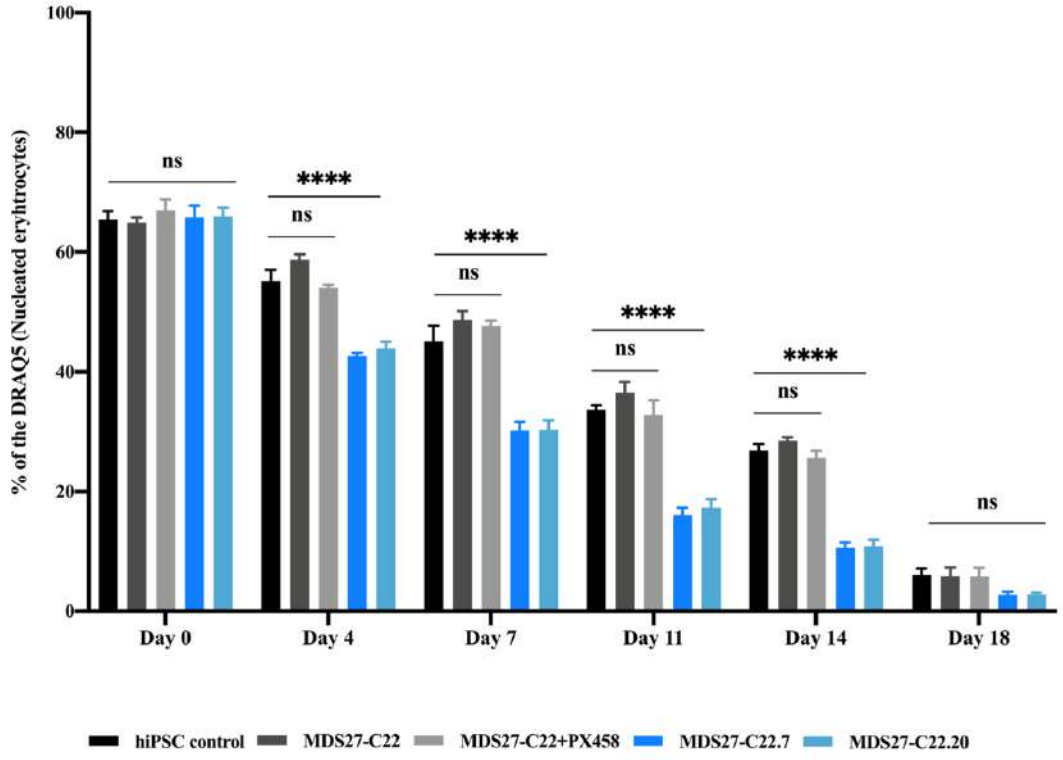


Figure 5.27: Assessment of enucleation rate

Enucleation rates confirmed by DRAQ5 staining during several days of erythroid differentiation. Results are presented as mean ± SEM and analysed using Two-way ANOVA with multiple comparisons. p values of day 4,7, 11 & 14 in comparing with hiPSC control, MDS27-C22 & MDS27-C22+PX458 were **** <0.0001and p-values were ns when compared hiPSC-control with MDS27-C22 and MDS27-C22+PX458. P values of day 0 & 18 in comparing with hiPSC control, MDS27-C22 & MDS27-C22+PX458 were ns. Data represent 4 independent experiments.

Our previous data in chapter 4 indicated the presence of dyserythropoiesis in cultures derived from MDS27-iPSC lines. To determine whether the addition of the *C/EBPα* mutation affected in any way the dysplastic phenotype in the erythroid lineage, a morphological assessment was performed for the *C/EBPα* mutant MDS27-iPSC lines (MDS27-C22.7 and MDS27-C22.20). The morphological analysis indicated that the erythroid cells derived from MDS27-C22.7 and MDS27-C22.20 iPSC lines harboured dysplastic morphology as observed for MDS27-C22 and MDS27-C22+PX458 iPSC lines. Interestingly, new aberrant morphologies such as giant erythroblast and mitotic erythroblast were detected in *C/EBPα* mutant MDS27-iPSC lines but not in MDS27-C22+PX458 iPSC, indicating that these new aberrant cells are associated with

the incorporation of the *C/EBPα* mutation to disease cells harbouring the *ASXL1, SRSF2* and *RUNX1* mutations (Figure 5.28, A).

Then, we scored the aberrant morphology within five hundred cells, and the percentage of the aberrant morphology was obtained. The percentage of aberrant morphology of *C/EBPα* mutant MDS27-iPSC lines to the aberrant morphology of hiPSC control was calculated at the different time points (Day 4, 7, 11, 14 and 18). As can be seen in Figure 5.28 (B), the percentage of aberrant morphology of MDS27-C22.7 and MDS27-C22.20 iPSC to control hiPSC was significantly high during the different days of the differentiation. However, the percentage of aberrant cells in *C/EBPα* mutant MDS27-hiPSC lines compared with MDS27-C22 and MDS27-C22+PX458 hiPSCs was slightly increased but not statistically significant.

Taken together, the four mutant genes (*SRSF2*, *ASLX1*, *RUNX1* and *C/EBPα*) have not affected the maturation of the erythroid cells. All lines have given rise to all stages of erythroid cells, but the erythroblast markers were expressed until the end of the differentiation and quantitative changes in the expression of CD71 and CD235a markers could be observed. Also, cells harbouring mutations in these four genes have accelerated the erythroid differentiation to the mature stage and new aberrant morphologies were detected.

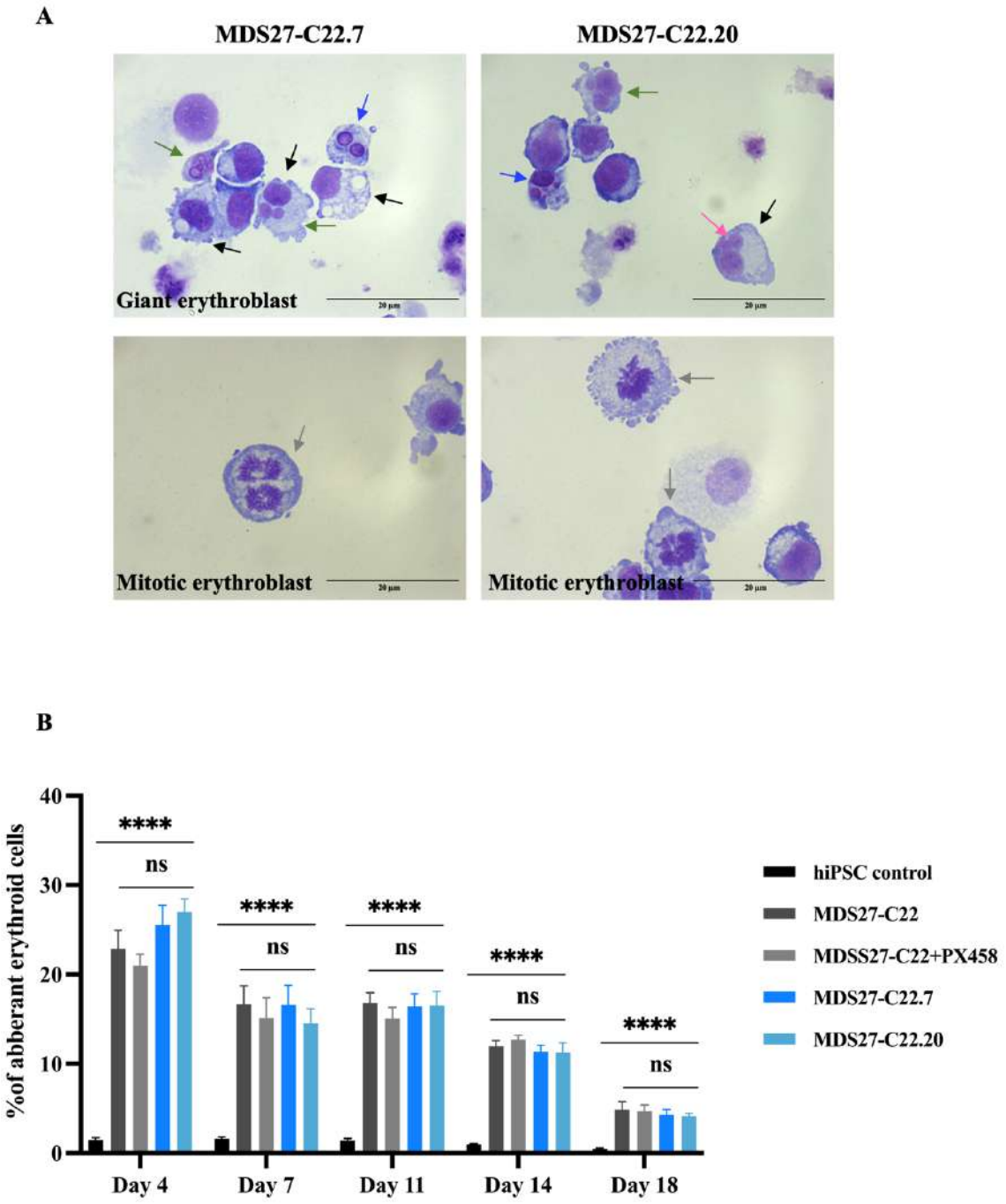


Figure 5.28: Morphological analysis of the erythroid cells

(A) A new aberrant morphology; Giant erythroblast (black arrow), Mitotic erythroblast (gray arrows) in MDS27-C22.7 and MDS27-C22.20 and the previous aberrant morphology as WT clone; Multinucleated erythroblast (green arrows), bi-nucleated erythroblast (blue arrows) and nuclear bridge (pink arrows). The pictures were taken Leica DM6000 at 100x magnification, 20 µm scale bar. **(B)** Bar grapg represents the percentage of the aberrant cells of MDS27 clones relative to the aberrant morphology of hiPSC control. Results are presented as mean ± SEM and analysed using Two-way ANOVA with multiple comparisons *** p >0.0001. The data represented 4 independent experiments.

5.4 Discussion

5.4.1 Successfully engineered heterozygous *C/EBPα* mutation, producing a high-risk MDS-iPSC line from low-risk MDS27-C22 line

In 2015, MDS27 patient had a heterozygous mutation in the C-terminal domain of *C/EBPα* in a region responsible for dimerization and DNA binding. C/EBPα, is one of the core hematopoietic TF; the mutation of this gene impairs myeloid differentiation and is associated with the progression from low-risk MDS to high-risk MDS (Sperling et al., 2017). It has been reported that the incidence of *C/EBPα* mutations in MDS patients is 5%-14% (Gombart et al., 2002, Snaddon et al., 2003). Furthermore, *C/EBPα* monoallelic mutation is observed in high-risk MDS and sAML after MDS progression, while biallelic mutations are common in de novo AML (Kato et al., 2011). However, a monoallelic point mutation in the C-terminal of *C/EBPα* alone is not able to generate leukaemia, but the incidence of a point mutation in combination with other mutations such as *ASLX1* and *RUNX1* is associated with leukemogenesis (Porse et al., 2005b). Thus, the co-existence of *SRSF2*, *ASLX1*, *RUNX1* and *C/EBPα* mutations in MDS27 has led to disease progression to high-risk MDS and eventually to sAML.

Implementing the CRISPR-Cas9 system has enabled a simple and affordable way to introduce the *C/EBPα* mutation into MDS27-C22 after several unsuccessful attempts to generate hiPSC clones from the blood sample taken from MDS27 patient in 2015. Our result demonstrates the successful targeting of the DNA binding domain of *C/EBPα* in MDS27-C22, engineering a heterozygous *C/EBPα* mutation.

5.4.2 Deletion in the DBD of one allele of C/EBPα lead to a truncated protein

C/EBPα protein comprises a DBD, a fork domain and a bZIP in its C-terminal (Lin et al., 1993). Basic ZIP is fundamental for homodimerization and heterodimerization with other members of the C/EBP family. The C-terminal region has a role in interaction with other transcription factors such as PU.1, c-JUN and RUNX1 (Lin et al., 1993, Ramji and Foka, 2002). Moreover, this intronless gene leads to two different protein isoforms: full-length C/EBPα protein (P42) and the shorter (P30) isoform that lacks the N-terminal domain. The N-terminal of C/EBPα is included in the full isoform only, and it consists of three transactivation domains (TA). TA can interact with the transcriptional machinery, cell regulators and chromatin re-modellers (Schwartz et al., 2003, Porse et al., 2006, Pedersen et al., 2001).

The effect of the different types of C/EBPα mutations on protein synthesis is not clearly investigated. Thus, we performed western blotting to check the protein level of C/EBPα after introducing the mutation in the C-terminal part. Interestingly, our result did not show a difference in the expression of the two isoforms of C/EBPα protein but other bands were detected in mutant clones which indicats that a a truncated protein could affect the stability of the function of the C/EBPα protein. Consistent with the literature, Leroy, et al, (2005) reported that insertions or deletions in specific regions of C/EBPα protein do not affect protein synthesis, but could inhibit the functionality of the C/EBPα protein (Leroy et al., 2005).

5.4.3 CRISPR-Cas9 does not affect the chromosomal stability and pluripotency of MDS27-C22 hiPSC

The CRISPR-Cas9 system is highly likely to cause unwanted chromosomal alteration. It has been reported that chromosomal abnormalities and large mutations at target loci are detected in clones of different cancer cell lines after using CRISPR-Cas9 (Rayner et al., 2019).

For this reason, we determined whether gross chromosome alterations (changes in the number of the chromosome) were observed in MDS27-C22+PX458, MDS27-C22.7 and MDS27-C22.20 iPSCs using a relatively straightforward karyotype method. Counting the number of the chromosomes demonstrated that none of the clones had suffered chromosomal instability after CRISPR-Cas9. We used our in-house facility to demonstrate the number of chromosomes, but other commercial services are readily available that can determine any alteration within chromosomes or any off-target mutations. For example, fluorescence *in situ* hybridization (FISH) is one of the cytogenetic techniques that is widely used for diagnostic applications, cytogenetic research and after using the CRISPR-Cas9 approach (Matsui et al., 2002, Peterson et al., 2015).

We also determined whether the CRISPR-Cas9 system could have affected the pluripotency of the selected iPSCs (MDS27-C22+PX458, MDS27-C22.7 and MDS27-C22.20). As expected, based on the morphology of the clones in culture, all selected hiPSC clones expressed the intercellular and extracellular pluripotent proteins analysed (TRA1-81, SOX2 and NANOG). Our work is directly in line with a previous study that reported that CRISPR-Cas9 mediated genome editing in iPSCs does not affect the pluripotency and differentiation potential of iPSC (Geng et al., 2020).

5.4.4 *SRSF2*, *ASLX1*, *RUNX1* and *C/EBPα* mutations do not affect the differentiation of hiPSC into HSPCs

After the successful generation of MDS27 clones with *C/EBPα* mutation, we compared the hematopoietic phenotype of MDS27-hiPSC before and after the *C/EBPα* mutation. Initially, the five selected clones were differentiated into HSPCs. For this, a commercial protocol from stem cell technology was followed in this project.

Interestingly, this protocol successfully differentiated all hiPSC clones to HSPCs, with both populations, earlly ($CD34^+$ $CD43^+$) and late ($CD34^+$ $CD45^+$) HSPCs, being detected.

HSPCs differentiation results indicated that the four mutant genes (*SRSF2*, *ASLX1*, *RUNX1* and *C/EBPα*) did not impair the ability of hiPSC to differentiate to HSPCs in comparison with hiPSC control or MDS27-C22 hiPSC. Despite not detecting any significant difference at the level of HSPC, our data revealed that C*/EBPα* mutation impaired the further differentiation of HSPCs to mature cells. This result is in agreement with published data showing that C/EBPα is required mainly during the differentiation of myeloid cells (Lin et al., 1993) Moreover, our results are consistent with data obtained in the study of Bereshchenko et al., (2009); using a mouse model, it was found that the C-terminal mutation of C*/EBPα* impaired only the myeloid differentiation by decreasing the myeloid genes expression and increasing erythroid-specific gene expression (Bereshchenko et al., 2009).

Overall, the differentiation to HSPCs was successful and we could conclude that *ASLX1, SRSF2,* and *RUNX1*, and *C/EBPα* mutations do not affect HSPC formation.

5.4.5 *ASLX1*, *RUNX1*, *SRSF2* and *C/EBPα* mutations impair the colony-forming potential of myeloid lineage, induce erythroid lineage differentiation, and affect the self-renewal capacity of HSPCs

To identify the impact of *ASLX1*, *RUNX1*, *SRSF2* and *C/EBPα* mutated genes in the differentiation potential of HSPCs we assessed the further differentiation of HSPCs to committed progenitor cells in semi-solid medium. Interestingly, the mutant HSPCs exhibited reduced clonogenicity capacity affecting the formation of myeloid CFUs, quite likely a reflection of the ineffective haematopoiesis observed in high-risk MDS and sAML patients.

Our model shows that introducing *C/EBPα* mutation in cells harbouring *ASLX1*, *RUNX1* and *SRSF2* mutations skewed the lineage differentiation of HSPCs, blocking the myeloid differentiation potential, specifically the granulocyte lineage, and increasing the erythroblast differentiation. Because granulocytic colonies could be formed from HSPCs harbouring *ASLX1*, *RUNX1* and *SRSF2* mutations, our data indicated that *C/EBPα* is required for the CMP to differentiate into the GMP. This result is in line with previous studies that found that the *C/EBPα* C-terminal mutation strongly affects the differentiation of myeloid cells into granulocytes leading to the accumulation of blast cells in the BM (Porse et al., 2005a).

In addition, it has been reported that HSPCs with *C/EBPα* mutation are hyperproliferative, fail to differentiate (especially to granulocytes), and show increased self-renewal potential *in vitro* (Koschmieder et al., 2009). To assess this observation in our model, we performed serial re-plating and evaluated the amount of CFUs in our *C/EBPα* mutant iPSC lines. Our data showed that the clones containing the four mutations were able to generate mainly erythroid colonies, in secondary and tertiary re-plating, indicative of the increase in self-renewal capacity of these cells. This data follows the previously published results obtained in mice where C-terminal

mutations of C/ebpα increase the self-renewal of the CFUs forming colonies after 6 rounds of re-plating (Kato et al., 2011, Heath et al., 2004). It is interesting to note that the mutation of *C/EBPα* leads to an enhanced competitive repopulating activity of HSCs, suggesting a normal role for *C/EBPα* to limit HSC self-renewal (Muller et al., 2004).

5.4.6 *ASLX1*, *RUNX1*, *SRSF2* and *C/EBPα* mutations block the granulocytic differentiation

Our model shows that introducing the C-terminal mutation of *C/EBPα* in cells harbouring other mutations (*ASLX1*, *RUNX1* and *SRSF2*) reduces the myeloid differentiation capacity and blocks the formation of granulocytes. Interestingly, the morphological assessment of the myeloid cells derived from MDS27-C22.7 and MDS27-C22.20 iPSCs revealed the presence of mainly erythroblast cells instead of granulocytes. In addition, MDS27-C22 and MDS27-C22+PX458 iPSC could differentiate to mature granulocytes but with aberrant morphology. These results indicate that *ASLX1*, *RUNX1* and *SRSF2* mutations drive the aberrant maturation of myeloid cells without blocking the differentiation; whilst the addition of *C/EBPα* mutation provokes myeloid differentiation blocking observed in MDS27-C22.7 and MDS27-C22.20 iPSC lines.

Our results agree with reports showing that *C/EBPα* has an important role in myeloid differentiation during the transition of GMP to granulocytes (Friedman, 2015, Avellino and Delwel, 2017). However, our observations regarding myeloid differentiation are contradictory with a study by Bapat *et al*., (2018), who report that *SRSF2* mutation in primary human $CD34^{+}$ causes abnormal differentiation by skewing GMP differentiation toward monocytes, and MEP differentiation toward megakaryocytes (Bapat et al., 2018). Others have shown that *ASXL1*/*SRSF2* co-mutated are frequently events in high-risk MDS and sAML, both mutations promoting GMP differentiation to monocytes and inducing leukemogenesis (Johnson et al.,

2019, Richardson et al., 2021). In addition, in high-risk MDS, *ASXL1* mutations frequently co-exist with *RUNX1* mutations reducing the myeloid differentiation and inducing leukemogenesis (Schnittger et al., 2013).

Our morphological results tie well with previous studies wherein the deletion of *ASXL1* in a mouse model and in human $CD34^+$ exhibit dysplastic morphology in the different lineages of HSCs (Jawhar et al., 2016, Metzeler et al., 2011). In line with previous studies, heterozygous *RUNX1* mutation in an iPSC model demonstrates that *RUNX1* only affects the differentiation of the megakaryocyte and the morphology of other lineages (Sakurai et al., 2014). However, *RUNX1* mutant-transduced cord blood $CD34^+$ cells continued for more than 100 days, which the cells displayed an immature granulocyte-macrophage progenitor-like (Gerritsen et al., 2019). In addition, *RUNX1* mutation only is not sufficient to induce leukaemia, and it is frequently found together with *ASXL1* to induce the transformation to sAML (Nagase et al., 2018). Interestingly, several studies have reported that mutations in *RUNX1* induce the *C/EBPα* mutation and consequently impair myeloid differentiation that may contribute to disease progression and leukemic transformation (Guo et al., 2012, Ptasinska et al., 2019).

Generally, according to our data, the co-existence of these three mutations (*ASXL1, SRSF2* and *RUNX1*) could have a role on inducing the *C/EBPα* mutation, contributing to disease progression and sAML. Thus, the occurrence of each mutation alone would not have a role in leukemogenesis, but the co-existence of them induces disease progression to the worse stage. To date, our study is the first study that has characterised the immune phenotype of *ASXL*1, *SRSF*2, *RUNX1*, and *C/EBPα* co-mutated in MDS progression to high-risk MDS. Also, it is the first project to study the effect of these co-mutated genes in a human model, which should have an important contribution in understanding disease progression and in treatment discovery.

Despite these promising results with myeloid differentiation, several questions need to be answered. For example, it would be imperative to introduce the *C/EBPα* mutation in normal hiPSC to prove that this phenotype is due to combining the four mutant genes and in this way, to get a complete picture ofthe role of *C/EBPα* mutation in disease progression to high-risk MDS.

5.4.7 *ASLX1*, *RUNX1*, *SRSF2* and *C/EBPα* mutations affect other TFs that have a role in regulating myeloid differentiation

In reviewing the literature, it has been shown that myeloid differentiation is directed at an early stage by *PU.1* and *RUNX1* and is further controlled by *C/EBPα* (Zhang et al., 2004, Back et al., 2005, Okuda et al., 1996). In addition, different ChIP-seq studies have demonstrated that *C/EBPα* binds the promoter region of *RUNX*, *GATA2*, *LMO2*, *PU.1* and other TFs (Collins et al., 2014, Wilson et al., 2010). Thus, it was important to demonstrate that the C/EBP*α* mutation generated by CRISPR would lead to the dysregulation of these transcription factors (*GATA2, LMO2, PU.1*) impacting on myeloid differentiation.

Interestingly, our data show that *RUNX1* expression is reduced in HSPCs derived from MDS27-C22 iPSC, and further reduction is observed after introducing the *C/EBPα* mutation. The reduction in the *RUNX1* expression could explain the block in myeloid differentiation in *C/EBPα* mutant lines because it has been reported that mutations in *RUNX1* strongly impaired myeloid commitment by reducing C/EBP*α* expression. In the same study, RNA seq data indicated that *RUNX1* mutation leads to the downregulation of genes associated with granulocytic differentiation (Gerritsen et al., 2019).

C/EBPα mutation also promoted a reduction in the *PU.1* expression in the *C/EBPα* mutant lines. These results go beyond previous reports, showing that *C/ebpα* knockdown in mice reduced the

expression of *PU.1* and markedly enhanced erythropoiesis (Ma et al., 2014). Several studies have confirmed that the reduction of C/EBPα expression strongly reduces the expression of *PU.1* which leads to the speculation that direct regulation of *GATA1* or a *GATA1* co-factor by *C/EBPα* might also restrict the CMP to MEP transition (Ma et al., 2014, Zhang et al., 2004, Cammenga et al., 2003).

Moreover, *GATA2* has been shown to have a role in myeloid differentiation: inducing the expression of *GATA2* in C/EBPα-expressing GMP leads to committing exclusively into the eosinophil lineage (Iwasaki et al., 2006). It is encouraging to compare our data of *GATA2* expression with previous findings that found that both *C/EBPα* and *GATA2,* have a role in controlling the proliferation and differentiation of myeloid progenitors, and mutations in both genes have been found to affect the granulocytic differentiation and induce the development of AML (Di Genua et al., 2020, Hahn et al., 2011) .

Also, *LMO2* has been reported to have a predominant role in erythropoiesis; with overexpression of *Lmo2* increasing the number of BFU-E at the expense of CFU-GM in the mouse BM (Hansson et al., 2007). This finding is consistent with the LMO2 expression data found in our mutant lines. It seems possible that the presence of erythroid cells in myeloid culture of *C/EBPα* mutant clones is due to the overexpression of LMO2.

Collectively, qRT-PCR results corroborate our hypothesis that the *C/EBPα* mutation generated by CRISPR-Cas9 affects the functionality of *C/EBPα*, resulting in the dysregulation of the expression of other TFs important for myeloid differentiation. In addition, dysregulation of these TFs has been associated with *C/EBPα* mutation, as the C-terminal region of C*/EBPα* interacts with *PU.1*, *GATA2*, *LMO2* and *RUNX1* (Lin et al., 1993, Ramji and Foka, 2002).

5.4.8 *ASLX1*, *RUNX1*, *SRSF2* and *C/EBPα* mutations accelerate erythroid differentiation and generate cells with new dysplastic features

In the previous chapter four, we reported that mutations in *ASLX1*, *RUNX1* and *SRSF2* affected the quality of erythroid cells. Moreover, during myeloid differentiation, the addition of the *C/EBPα* mutation to the other three mutations induced erythroid differentiation.

Indeed, an important observation of our erythroid differentiation data was that the mutant iPSC lines still contained a population of erythroblasts ($CD71^{+}$ $CD235^{+}$) at the end of erythroid differentiation (day 18). But the statistical analysis showed no significant difference between *C/EBPα* mutant lines, and control iPSC lines. However, the gMFI of CD71 expression was significantly higher in *C/EBPα* mutant iPSC lines in comparison to normal iPSC control and MDS27-iPSC lines. This result indicates that *C/EBPα* could slightly affect the maturation of erythroid cells, with a proportion of cells not able to fully mature.

Nonetheless, our DRAQ5 flow cytometry results and morphological assessment of the erythroid cells supported that erythroid differentiation in mutant lines goes faster than in the hiPSC control, MDS27-iPSC and MDS27-C22+PX458 iPSC lines. Acceleration of the erythroid differentiation could be attributed to *C/EBPα* mutation in which the *C/EBPα* mutation enhances erythropoiesis by altering some TFs important for erythroid differentiation such as GATA1 (Ma et al., 2014, Zhang et al., 2004, Cammenga et al., 2003). In addition, it has been reported that accelerated erythropoiesis in Thalassemia is associated with dyserythropoiesis which leads to increased apoptosis of erythrocytes (Centis et al., 2000). However, this outcome is contrary to the DRAQ5 flow cytometry results and morphological assessment, as these results suggest that introducing the *C/EBPα* mutation accelerates erythroid differentiation with mature erythroid cells detected at early stages of the differentiation. The possible explanation for this contradiction is that these mature cells are aberrant in their morphology, leading to co-

expression of $CD71^+$ and $CD235^+$. The morphological investigation supports this view, as cells with new aberrant morphology (giant erythroblast and mitotic erythroblast) were observed in the cultures derived from MD27-C22.7 and MD27-C22.20 hiPSC but not in the WT control or CRISPR control hiPSCs. It is important to mention that the new dysplastic morphology found in our erythroid cultures has been observed previously in the BM of high risk- MDS patient (Haria et al., 2014).

Moreover, not only new dysplastic features were observed but the percentage of the aberrant cells had increased in the *C/EBPα* mutant hiPSC lines. Therefore, it is possible to speculate that the increase in the proportion of cells that still express $CD71^+$ $CD235a^+$ by day 18 correspond to the aberrant cells found in the culture.

In summary, our results seem to indicate that at the stage of high-risk MDS, the four mutated genes have affected erythropoiesis by accelerating erythroid differentiation and increasing the proportion of aberrant cells, which could affect the functionality of RBCs and increase the severity of anaemia.

Chapter 6 : General discussion

6.1 Implications of research findings

6.1.1 Human iPSC as a model system to study low-risk MDS and high-risk MDS

MDS and other haematological diseases are devastating illnesses with a high economic and social burden. The researchers rely on model systems to understand the cellular and molecular mechanisms of the malignant state at the cellular, organ and organism level. These model systems include immortalised cell lines and genetically engineered mutant mice. However, the use of immortalised cell lines and mouse models have limitations in which they do not present all the stages and mutations of MDS or the effect of mutations could be different between mouse and human. Thus, the derivation of iPSCs from MDS patient cells provides a valuable tool in medical research for investigation of the disease, drug development and precision medicine. Thus, the main aim of this study was to develop a new *in vitro* system to study the molecular mechanisms leading to disease progression in MDS. Firstly, it was demonstrated that our lab is able to generate hiPSC from the PBMNCs of a MDS patient diagnosed with low-risk MDS harbouring *RUNX1*, *ASXL1* and *SRSF2* mutations. Also, we were able to apply CRISPR-Cas9 to generate a *C/EBPα* mutation in these hiPSC clones to create isogenic iPSC lines for high-risk MDS.

The use of these two experimental approaches have provided a remarkable advantage in this study in that the two stages of the disease were achieved. Furthermore, these models (low-risk & high-risk MDS) have beneficial implications for understanding disease evolution and helping to allocate efficient treatment options.

The generation of the low risk and high risk isogenic hiPSC lines from MDS27 was challenging but successful and has placed our lab within less than a handful of groups worldwide to generate iPSC from MDS patients. To date, only two groups have been able to generate hiPSC from

different MDS patients, and they are using this model to understand the disease phenotype and investigate the best treatment options (Kotini et al., 2015, Kotini et al., 2017, Hsu et al., 2019, Chang et al., 2018). Papapetrou's group were able to study different mutant genes and complex karyotypes such as *SRSF2*, *GATA2*, *PHF6* and *del(7q)* in the several stages of MDS, but most of them were not captured through the reprogramming process. Also, they applied CRISPR-Cas9 to have a larger collection of non-isogenic iPSC representing the normal, MDS and AML stages (Chang et al., 2018,Kotini et al., 2015, Kotini et al., 2017).

However, Doulatov's group generated iPSC from different MDS patients to study the clonal evolution during leukemic progression, but it is clear that not every patient will necessarily transition through each of the MDS stages (Hsu et al., 2019). Furthermore, MDS is a heterogeneous clonal disease. Several genetic mutations involving epigenetic regulators, chromatin modifiers, splicing factors, transcription factors, and signalling adaptors lead to MDS or disease progression. Thus, these previous models may not apply to all MDS patients, and the generation of several models will help enormously MDS studies.

In respect to the previous studies, our models are based on the non-integrating method to generate iPSC from MDS patients without introducing changes to the genome in which these iPSCs can be used safely for further clinical application. In addition, our models are the first models generated from the same MDS patient, thus both low risk and high risk iPSC clones came from the same genetic background, helping to identify the molecular mechanism of the disease progression more accurately. Based on our knowledge, our models are the first to study the co-mutations of *ASLX1, SRSF2, RUNX1* and *C/EBPα* in MDS patients.

Collectively, our isogenic MDS27-hiPSC lines serve as a good model system to study the molecular mechanisms leading to disease progression from low-risk to high risk MDS.

6.1.2 *ASLX1, RUNX1* and *SRSF2* mutations in Patient MDS27 at low-risk MDS lead to dysplasia in erythroid and myeloid cells

In chapters four and five, we were able to identify the disease phenotype of MDS27-hiPSC for the low-risk MDS stage. Low-risk MDS27-hiPSC harbouring *ASXL1*, *SRSF2* and *RUNX1* mutations were differentiated towards HSPCs. HSPCs derived from MDS27-iPSC displayed a reduced clonogenic capacity compared to healthy hiPSC (Figure 6.1).

Our findings are broadly consistent with the data reported from MDS patients that most of MDS patient have low number of CFUs progenitor cells (Chang et al., 2018, Kotini et al., 2015, Kotini et al., 2017, Michalopoulou et al., 2004, DeZern et al., 2013, Li et al., 2016).

However, the further investigation of erythroid and myeloid cells in liquid culture showed both lineages could differentiate to mature cells, but these mature cells were aberrant in their morphology; displaying nuclear bridges and multinuclear. Interestingly, these atypical aberrant morphologies are the common dysplastic features of MDS patients (Hasserjian et al., 2017, Tohyama, 2018). Furthermore, another group has shown that aberrant cells in MDS have a reduction in the survival rate (Zini, 2017). Based on these data, our results provide evidence that our MDS27-iPSCs model is a bona fide MDS model, showing erythroid and myeloid dysplasia similar to the anaemia and cytopenia of the MDS27 at the time of diagnosis.

This study has confirmed that *ASLX1*, *SRSF2* and *RUNX1* mutations play a role in the dysplastic features present in low-risk MDS. Thus, the model generated during this thesis could be used as a platform for the screening of drug therapies that could target these aberrant cells, improving differentiation and reducing disease progression. In line with this, a few preliminary experiments were done during the course of this thesis, specifically with Danazol.

Danazol is a synthetic steroid known to improve erythroid maturation in patients with Aplastic Anaemia (Townsley et al., 2016) and MDS (Colunga-Pedraza et al., 2018). In an attempt to determine the validity of our model system, the erythroid differentiation in the presence /absence of Danazol was studied in the lab. Our preliminary data seems to show a noticeable reduction in the percentage of aberrant cells, but these data need to be further confirmed. Recently, a new clinical trial has started for low and intermediate-risk MDS using repurposing drugs, among them Danazol, with our collaborator MD Manoj Raghavan as one of the clinical leads of the trial. Also, a new line of investigation, understanding how Danazol improves myeloid differentiation, is been followed in the lab using this model system and other iPSC that will be generated from responders and non-responders MDS patients from the trial.

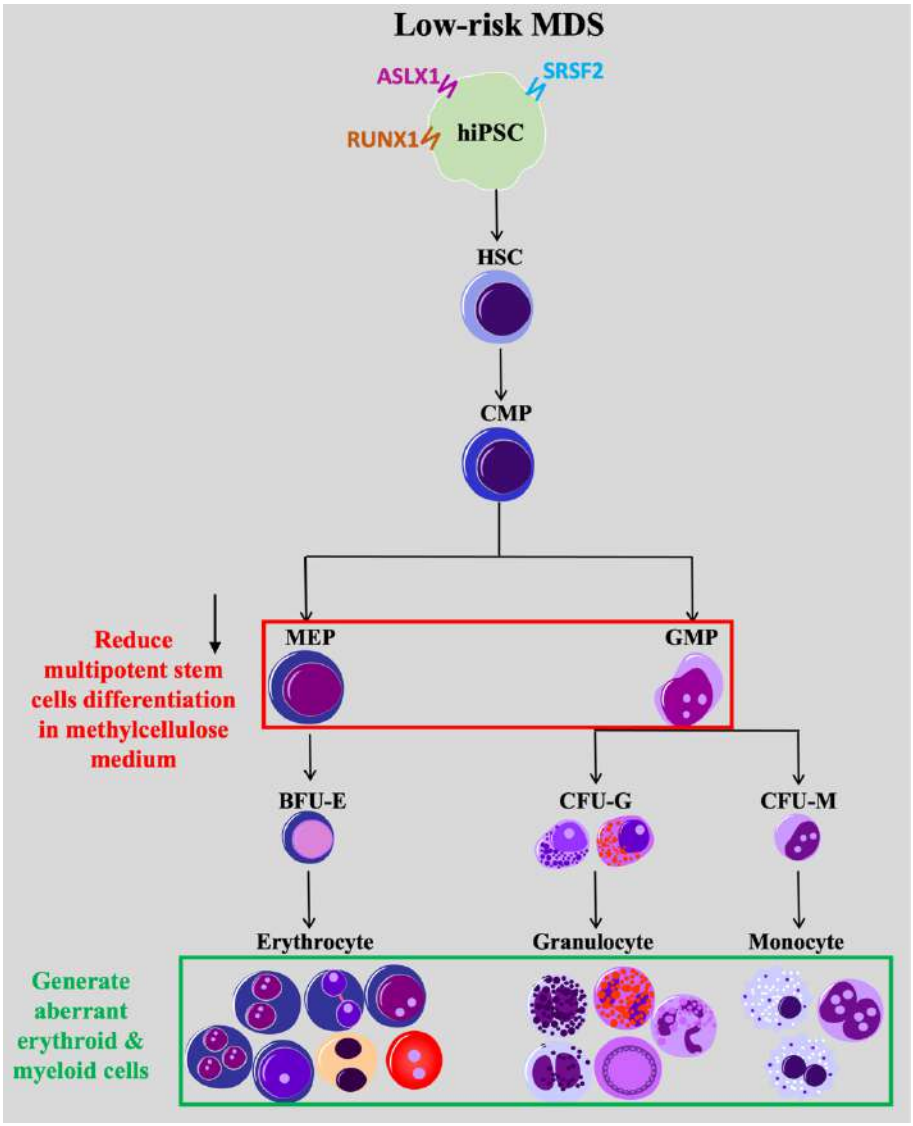


Figure 6.1: Low-risk MDS model

The low-risk MDS model generated from patient MDS27. The cells harboured *ASLX1*, *SRSF2* and *RUNX1* mutation. The phenotypical disease investigation shows that these three mutations reduce the clonogenic capacity and affect the quality of erythroid and myeloid cells.

.

6.1.3 *ASLX1, SRSF2, RUNX1* and *C/EBPα* mutations block myeloid differentiation and promote erythroid differentiation with an increase in the aberrant cells in high-risk MDS

In chapter five, we identified the disease phenotype of MDS27-hiPSC after disease progression to high-risk MDS. Our data shows that inclusion of *C/EBPα* to the MDS27-iPSC (harbouring the other mutant genes) did not affect the differentiation of hiPSC to HSPCs whilst it affected the multilineage differentiation of HSPCs. The colony assays revealed that the HSPCs from the high-risk model exhibited a reduced capacity in the formation of myeloid and erythroid CFUs and a block in the formation of granulocytic CFU. This result indicated the impaired hematopoietic colony-forming capacity of the high-risk MDS27 hiPSC clones in agreement with the data reported from MDS patients (Chang et al., 2018, Kotini et al., 2015, Kotini et al., 2017, Michalopoulou et al., 2004, DeZern et al., 2013, Li et al., 2016). Furthermore, it has been reported that an increase in the self-renewal of HSC in high-risk MDS is associated with an increase in the proliferation of disease cells that leads to accelerated disease progression and leukemic transformation (Walenda et al., 2014). Indeed, in our system, the inclusion of *C/EBPα* mutation led to an enhanced re-plating capacity of progenitors in methylcellulose medium, suggesting a normal role for *C/EBPα* to limit HSC self-renewal (Muller et al., 2004). Moreover, the investigation of myeloid differentiation in liquid culture showed a block in differentiation to mature granulocytes in *C/EBPα* mutant lines. This result supported our previous finding of using colony assays in that *C/EBPα* mutation in the context of *ASLX1*, *RUNX1* and *SRSF2* mutations led to an inhibition of the granulocytic differentiation and an increase in erythroid cells (Figure 6.2).

The functional consequences of *C/EBPα* mutation could be explained at a transcriptional level, as we found downregulation of *RUNX1, PU.1* and *GATA2* expression in *C/EBPα* mutant lines and upregulation of *LMO2* expression in these cells.

The qRT-PCR results support the work of others showing that mutation in *C/EBPα* affects the expression of *PU.1*, *LMO2* and *GATA2* (Yeamans et al., 2007, Cooper et al., 2015, Wilson et al., 2010). Subsequently, in our system, mutation of *C/EBPα* affected the expression of these genes leading to inhibition of myeloid differentiation and promoting differentiation towards the erythroid lineage.

How the cells from the patient harbouring three mutations *ASLX1*, *SRSF2* and *RUNX1* have evolved and acquired the *C/EBPα* mutation is presently unknown. One possible explanation is through genome instability generated by the *SRSF2* mutation, as mutations in splicing factors genes, including *SRSF2*, have been associated with increased replication stress in MDS (Pellagatti and Boultwood, 2020, Chen et al., 2018).

Collectively, our results are consistent with the RNA-seq results in patients with *C/EBPα* mutations in which this mutation leads to the upregulation of genes involved in erythroid differentiation including *GATA1*, (*ZFPM1* also known as *FOG1*), *HEMGN*, *EPOR*, *GFI1B*, *KLF1*, *ANK1*, *TFRC* (*CD71*), and genes encoding erythrocyte membrane proteins and haemoglobin chains (Marcucci et al., 2008).

It would be interesting to define the transcriptional changes in the two stages of the disease to clearly understand the mechanisms behind the aberrant morphology, the block in myeloid differentiation and the stimulation of erythroid differentiation (work in progress).

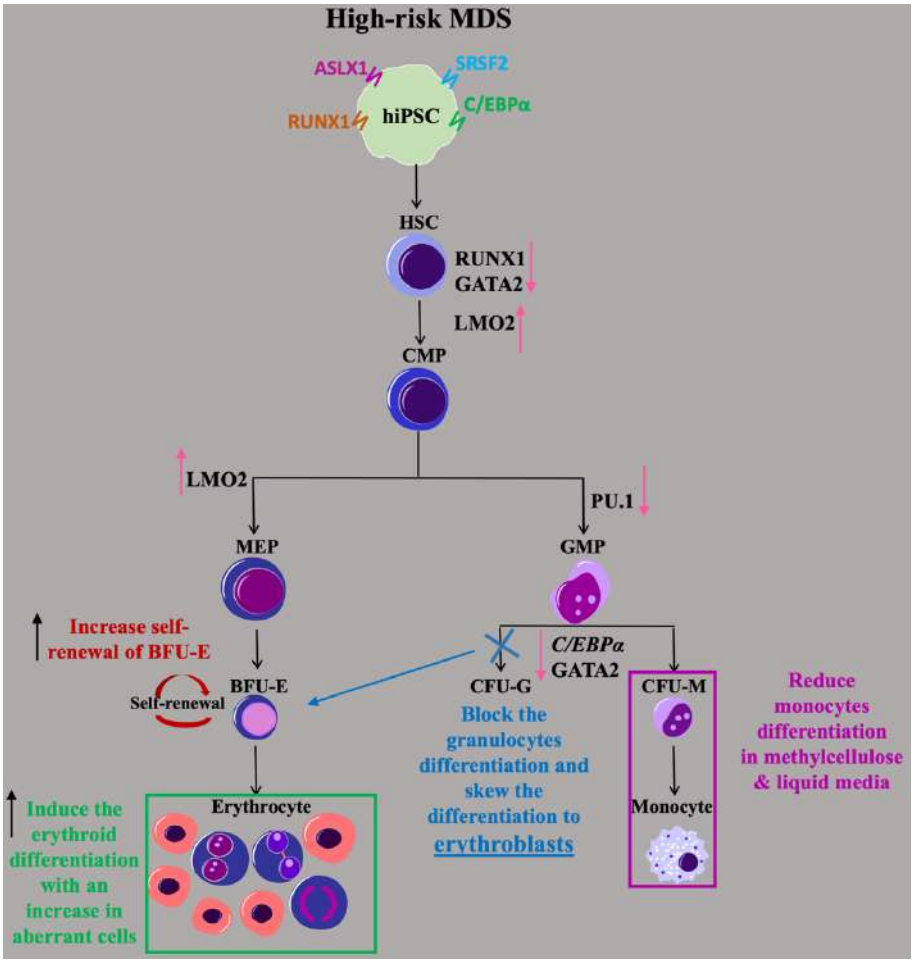


Figure 6.2: High-risk MDS model

Through CRISPR-Cas9 a *C/EBPα* mutation in the DNA binding domain was generated, to mimic the mutation acquired when MDS27 patient progressed to high-risk MDS. Introducing the *C/EBPα* to the clone harbouring *ASLX1, SRSF2* and *RUNX1* mutations reduced the clonogenic capacity, blocked myeloid differentiation and induced erythroid differentiation with an increased production of aberrant erythroid cells. Also, it increased the self-renewal capacity of BFU-E. *C/EBPα* mutation affected the expression of *PU.1, GATA2* and *LMO2* leading to a block in myeloid differentiation and an increase in erythroid cells.

In conclusion, these two models show the power of somatic reprogramming (hiPSC) as a new tool for investigating the molecular mechanisms of disease progression in MDS that could help in drug development and precision medicine. The exact mechanism of how the MDS27 cells progress from low-risk to high-risk MDS is still not fully elucidated, and further work is required. However, these models provide new opportunities to study HSPC populations in MDS which often cannot be easily obtained at adequate numbers from patient samples or propagated in patient-derived xenograft models. To our knowledge, these two models are the first models

to study the co-mutations of these TFs in MDS patient. In addition, this data can contribute to the understanding of the evaluation of the disease and represents a possible chance for the discovery of novel treatment of MDS.

6.2 Limitations of the model used in this study

Generation of iPSC from a patient is rather costly and time-consuming. It is estimated that the generation and expansion of an iPSC line and all the necessary tests to check its pluripotency and safety needs between 3 to 6 months for its total production. One of the main concerns about working with iPSC is that during the generation or working a long time with clones leads to genome instability. Over time, this would result in elevated mutation rates and chromosome aberrations. Therefore, it is not certain whether the phenotypes observed are related to the patient's mutations or whether they are due to the effect of the progressive acquisition of mutations. However, during this thesis, this possibility was reduced through limiting the number of passages of each line and doing regularly karyotype analysis. Lastly, despite the capability of iPSC to differentiate to any cell type, the differentiation protocols are lengthy, and thus the in depth study of the molecular mechanisms leading to disease progression could not be followed due to time constrains.

6.3 Further questions and future work

6.3.1 Determine if the phenotype of high-risk MDS is not relative to the *C/EBPα* mutation only

Further work should be undertaken to prove if the high-risk MDS phenotype is relative to the co-mutation of *ASLX1, SRSF2, RUNX1* and *C/EBPα*. To investigate this, we will apply the same gRNA of *C/EBPα* to hiPSC control (normal), to generate the same mutation as the mutant line. Then, the same experiments will be conducted to determine the contribution of *C/EBPα* single mutation to the hematopoietic differentiation.

6.3.2 Check the possible treatment options for the two stages of the disease

Future work should consider the potential effects of different treatment options such as Danazol, as mentioned previously. Also, it could be of interest to determine the effect of azacitidine treatment before and after the disease progression. Azacitidine is the treatment option in clinic for high-risk MDS patients, and MDS27 patient was treated with this drug once he progressed to high-risk MDS. Over time, the MDS27 patient became resistant to this treatment and he progressed to sAML, dying soon after. Thus, it would be extremely interesting to use this model system and determine whether cells resistant to azacitidine could be obtained to understand mechanisms of resistance to the chemotherapy. To do this, we will treat the cells with azacitidine in colony assays and determine whether clones emerged. The transcriptome analysis of those clones would then provide information about genes/pathways associated to this resistance.

6.3.3 Xenotransplants using HSPCs from high-risk MDS stage

Further work is certainly required to clearly determine the growth advantage of HSPCs after the disease progresses to high-risk MDS. To do this, we will differentiate the hiPSC to HSPCs and cells will be transplanted into immunocompromised mice such as NSG mice. This *in vivo* experiment would undoubtedly strengthen the *in vitro* results obtained through colony assays.

References

AATOLA, M., ARMSTRONG, E., TEERENHOVI, L. & BORGSTROM, G. H. 1992. Clinical significance of the del(20q) chromosome in hematologic disorders. *Cancer Genet Cytogenet,* 62**,** 75-80.

ABELSON, S., COLLORD, G., NG, S. W. K., WEISSBROD, O., MENDELSON COHEN, N., NIEMEYER, E., BARDA, N., ZUZARTE, P. C., HEISLER, L., SUNDARAVADANAM, Y., LUBEN, R., HAYAT, S., WANG, T. T., ZHAO, Z., CIRLAN, I., PUGH, T. J., SOAVE, D., NG, K., LATIMER, C., HARDY, C., RAINE, K., JONES, D., HOULT, D., BRITTEN, A., MCPHERSON, J. D., JOHANSSON, M., MBABAALI, F., EAGLES, J., MILLER, J. K., PASTERNACK, D., TIMMS, L., KRZYZANOWSKI, P., AWADALLA, P., COSTA, R., SEGAL, E., BRATMAN, S. V., BEER, P., BEHJATI, S., MARTINCORENA, I., WANG, J. C. Y., BOWLES, K. M., QUIROS, J. R., KARAKATSANI, A., LA VECCHIA, C., TRICHOPOULOU, A., SALAMANCA-FERNANDEZ, E., HUERTA, J. M., BARRICARTE, A., TRAVIS, R. C., TUMINO, R., MASALA, G., BOEING, H., PANICO, S., KAAKS, R., KRAMER, A., SIERI, S., RIBOLI, E., VINEIS, P., FOLL, M., MCKAY, J., POLIDORO, S., SALA, N., KHAW, K. T., VERMEULEN, R., CAMPBELL, P. J., PAPAEMMANUIL, E., MINDEN, M. D., TANAY, A., BALICER, R. D., WAREHAM, N. J., GERSTUNG, M., DICK, J. E., BRENNAN, P., VASSILIOU, G. S. & SHLUSH, L. I. 2018. Prediction of acute myeloid leukaemia risk in healthy individuals. *Nature,* 559**,** 400-404.

ABUJAROUR, R., VALAMEHR, B., ROBINSON, M., REZNER, B., VRANCEANU, F. & FLYNN, P. 2013. Optimized surface markers for the prospective isolation of high-quality hiPSCs using flow cytometry selection. *Sci Rep,* 3**,** 1179.

ACAMPORA, D., AVANTAGGIATO, V., TUORTO, F., BRIATA, P., CORTE, G. & SIMEONE, A. 1998. Visceral endoderm-restricted translation of Otx1 mediates recovery of Otx2 requirements for specification of anterior neural plate and normal gastrulation. *Development,* 125**,** 5091-104.

ADACHI, K., SUEMORI, H., YASUDA, S. Y., NAKATSUJI, N. & KAWASE, E. 2010. Role of SOX2 in maintaining pluripotency of human embryonic stem cells. *Genes Cells,* 15**,** 455-70.

ADLI, M. 2018. The CRISPR tool kit for genome editing and beyond. *Nat Commun,* 9**,** 1911.

ADOLFSSON, J., MANSSON, R., BUZA-VIDAS, N., HULTQUIST, A., LIUBA, K., JENSEN, C. T., BRYDER, D., YANG, L., BORGE, O. J., THOREN, L. A., ANDERSON, K., SITNICKA, E., SASAKI, Y., SIGVARDSSON, M. & JACOBSEN, S. E. 2005. Identification of Flt3+ lympho-myeloid stem cells lacking erythro-megakaryocytic potential a revised road map for adult blood lineage commitment. *Cell,* 121**,** 295-306.

AKASHI, K., TRAVER, D., MIYAMOTO, T. & WEISSMAN, I. L. 2000. A clonogenic common myeloid progenitor that gives rise to all myeloid lineages. *Nature,* 404**,** 193-197.

AMABILE, G., DI RUSCIO, A., MULLER, F., WELNER, R. S., YANG, H., EBRALIDZE, A. K., ZHANG, H., LEVANTINI, E., QI, L., MARTINELLI, G., BRUMMELKAMP, T., LE BEAU, M. M., FIGUEROA, M. E., BOCK, C. & TENEN, D. G. 2015. Dissecting the role of aberrant DNA methylation in human leukaemia. *Nat Commun,* 6**,** 7091.

ARAKI, R., HOKI, Y., UDA, M., NAKAMURA, M., JINCHO, Y., TAMURA, C., SUNAYAMA, M., ANDO, S., SUGIURA, M., YOSHIDA, M. A., KASAMA, Y. & ABE, M. 2011. Crucial role of c-Myc in the generation of induced pluripotent stem cells. *Stem Cells,* 29**,** 1362-70.

ARBER, D. A., ORAZI, A., HASSERJIAN, R., THIELE, J., BOROWITZ, M. J., LE BEAU, M. M., BLOOMFIELD, C. D., CAZZOLA, M. & VARDIMAN, J. W. 2016. The 2016 revision to the World Health Organization classification of myeloid neoplasms and acute leukemia. *Blood,* 127**,** 2391-405.

ARORA, N. & DALEY, G. Q. 2012. Pluripotent stem cells in research and treatment of hemoglobinopathies. *Cold Spring Harb Perspect Med,* 2**,** a011841.

AUL, C., GIAGOUNIDIS, A. & GERMING, U. 2001. Epidemiological features of myelodysplastic syndromes: results from regional cancer surveys and hospital-based statistics. *Int J Hematol,* 73**,** 405-410.

AVELLINO, R. & DELWEL, R. 2017. Expression and regulation of C/EBPalpha in normal myelopoiesis and in malignant transformation. *Blood,* 129**,** 2083-2091.

AVELLINO, R., HAVERMANS, M., ERPELINCK, C., SANDERS, M. A., HOOGENBOEZEM, R., VAN DE WERKEN, H. J., ROMBOUTS, E., VAN LOM, K., VAN STRIEN, P. M., GEBHARD, C., REHLI, M., PIMANDA, J., BECK, D., ERKELAND, S., KUIKEN, T., DE LOOPER, H., GROSCHEL, S., TOUW, I., BINDELS, E. & DELWEL, R. 2016. An autonomous CEBPA enhancer specific for myeloid-lineage priming and neutrophilic differentiation. *Blood,* 127**,** 2991-3003.

AVILION, A. A., NICOLIS, S. K., PEVNY, L. H., PEREZ, L., VIVIAN, N. & LOVELL-BADGE, R. 2003. Multipotent cell lineages in early mouse development depend on SOX2 function. *Genes Dev,* 17**,** 126-40.

BACHER, U., SCHANZ, J., BRAULKE, F. & HAASE, D. 2015. Rare cytogenetic abnormalities in myelodysplastic syndromes. *Mediterr J Hematol Infect Dis,* 7**,** e2015034.

BACHER, U., SCHNITTGER, S., KERN, W., WEISS, T., HAFERLACH, T. & HAFERLACH, C. 2009. Distribution of cytogenetic abnormalities in myelodysplastic syndromes, Philadelphia negative myeloproliferative neoplasms, and the overlap MDS/MPN category. *Ann Hematol,* 88**,** 1207-13.

BACK, J., ALLMAN, D., CHAN, S. & KASTNER, P. 2005. Visualizing PU.1 activity during hematopoiesis. *Exp Hematol,* 33**,** 395-402.

BAGCI, H. & FISHER, A. G. 2013. DNA demethylation in pluripotency and reprogramming: the role of tet proteins and cell division. *Cell Stem Cell,* 13**,** 265-9.

BAN, H., NISHISHITA, N., FUSAKI, N., TABATA, T., SAEKI, K., SHIKAMURA, M., TAKADA, N., INOUE, M., HASEGAWA, M., KAWAMATA, S. & NISHIKAWA, S. 2011. Efficient generation of transgene-free human induced pluripotent stem cells (iPSCs) by temperature-sensitive Sendai virus vectors. *Proc Natl Acad Sci U S A,* 108**,** 14234-9.

BAPAT, A., KEITA, N., MARTELLY, W., KANG, P., SEET, C., JACOBSEN, J. R., STOILOV, P., HU, C., CROOKS, G. M. & SHARMA, S. 2018. Myeloid Disease Mutations of Splicing Factor SRSF2 Cause G2-M Arrest and Skewed Differentiation of Human Hematopoietic Stem and Progenitor Cells. *Stem Cells,* 36**,** 1663-1675.

BARON, M. H., ISERN, J. & FRASER, S. T. 2012. The embryonic origins of erythropoiesis in mammals. *Blood,* 119**,** 4828-37.

BAYLEY, R., AHMED, F., GLEN, K., MCCALL, M., STACEY, A. & THOMAS, R. 2018. The productivity limit of manufacturing blood cell therapy in scalable stirred bioreactors. *J Tissue Eng Regen Med,* 12**,** e368-e378.

BEACHY, S. H. & APLAN, P. D. 2010. Mouse models of myelodysplastic syndromes. *Hematol Oncol Clin North Am,* 24**,** 361-75.

BECKER, A. J., MC, C. E. & TILL, J. E. 1963. Cytological demonstration of the clonal nature of spleen colonies derived from transplanted mouse marrow cells. *Nature,* 197**,** 452-4.

BEJAR, R., LORD, A., STEVENSON, K., BAR-NATAN, M., PEREZ-LADAGA, A., ZANEVELD, J., WANG, H., CAUGHEY, B., STOJANOV, P., GETZ, G., GARCIA-MANERO, G., KANTARJIAN, H., CHEN, R., STONE, R. M., NEUBERG, D., STEENSMA, D. P. & EBERT, B. L. 2014. TET2 mutations predict response to hypomethylating agents in myelodysplastic syndrome patients. *Blood,* 124**,** 2705-12.

BELLISSIMO, D. C. & SPECK, N. A. 2017. RUNX1 Mutations in Inherited and Sporadic Leukemia. *Front Cell Dev Biol,* 5**,** 111.

BENBERNOU, N., MUEGGE, K. & DURUM, S. K. 2000. Interleukin (IL)-7 induces rapid activation of Pyk2, which is bound to Janus kinase 1 and IL-7Ralpha. *J Biol Chem,* 275**,** 7060-5.

BENNETT, J. M., CATOVSKY, D., DANIEL, M. T., FLANDRIN, G., GALTON, D. A., GRALNICK, H. R. & SULTAN, C. 1982. Proposals for the classification of the myelodysplastic syndromes. *Br J Haematol,* 51**,** 189-99.

BERESHCHENKO, O., MANCINI, E., MOORE, S., BILBAO, D., MANSSON, R., LUC, S., GROVER, A., JACOBSEN, S. E., BRYDER, D. & NERLOV, C. 2009. Hematopoietic stem cell expansion precedes the generation of committed myeloid leukemia-initiating cells in C/EBPalpha mutant AML. *Cancer Cell,* 16**,** 390-400.

BERNECKER, C., ACKERMANN, M., LACHMANN, N., ROHRHOFER, L., ZAEHRES, H., ARAUZO-BRAVO, M. J., VAN DEN AKKER, E., SCHLENKE, P. & DORN, I. 2019. Enhanced Ex Vivo Generation of Erythroid Cells from Human Induced Pluripotent Stem Cells in a Simplified Cell Culture System with Low Cytokine Support. *Stem Cells Dev,* 28**,** 1540-1551.

BHUTANI, K., NAZOR, K. L., WILLIAMS, R., TRAN, H., DAI, H., DZAKULA, Z., CHO, E. H., PANG, A. W. C., RAO, M., CAO, H., SCHORK, N. J. & LORING, J. F. 2016. Whole-genome mutational burden analysis of three pluripotency induction methods. *Nat Commun,* 7**,** 10536.

BLACKWOOD, E. M. & EISENMAN, R. N. 1991. Max: a helix-loop-helix zipper protein that forms a sequence-specific DNA-binding complex with Myc. *Science,* 251**,** 1211-7.

BOCH, J., SCHOLZE, H., SCHORNACK, S., LANDGRAF, A., HAHN, S., KAY, S., LAHAYE, T., NICKSTADT, A. & BONAS, U. 2009. Breaking the code of DNA binding specificity of TAL-type III effectors. *Science,* 326**,** 1509-12.

BOER, B., KOPP, J., MALLANNA, S., DESLER, M., CHAKRAVARTHY, H., WILDER, P. J., BERNADT, C. & RIZZINO, A. 2007. Elevating the levels of Sox2 in embryonal carcinoma cells and embryonic stem cells inhibits the expression of Sox2:Oct-3/4 target genes. *Nucleic Acids Res,* 35**,** 1773-86.

BOIERS, C., CARRELHA, J., LUTTEROPP, M., LUC, S., GREEN, J. C., AZZONI, E., WOLL, P. S., MEAD, A. J., HULTQUIST, A., SWIERS, G., PERDIGUERO, E. G., MACAULAY, I. C., MELCHIORI, L., LUIS, T. C., KHARAZI, S., BOURIEZ-JONES, T., DENG, Q., PONTEN, A., ATKINSON, D., JENSEN, C. T., SITNICKA, E., GEISSMANN, F., GODIN, I., SANDBERG, R., DE BRUIJN, M. F. & JACOBSEN, S. E. 2013. Lymphomyeloid contribution of an immune-restricted progenitor emerging prior to definitive hematopoietic stem cells. *Cell Stem Cell,* 13**,** 535-48.

BOLOTIN, A., QUINQUIS, B., SOROKIN, A. & EHRLICH, S. D. 2005. Clustered regularly interspaced short palindrome repeats (CRISPRs) have spacers of extrachromosomal origin. *Microbiology,* 151**,** 2551-2561.

BOYER, S. W., SCHROEDER, A. V., BERDAN, S. & FORSBERG, E. C. 2011. All hematopoietic cells develop from hematopoietic stem cells through Flk2/Flt3-positive progenitor cells. *Cell stem cell,* 9**,** 64-73.

BRAGANCA, J., LOPES, J. A., MENDES-SILVA, L. & ALMEIDA SANTOS, J. M. 2019. Induced pluripotent stem cells, a giant leap for mankind therapeutic applications. *World J Stem Cells,* 11**,** 421-430.

BRAMBRINK, T., FOREMAN, R., WELSTEAD, G. G., LENGNER, C. J., WERNIG, M., SUH, H. & JAENISCH, R. 2008. Sequential expression of pluripotency markers during direct reprogramming of mouse somatic cells. *Cell Stem Cell,* 2**,** 151-9.

BRAUN, T., DE BOTTON, S., TAKSIN, A. L., PARK, S., BEYNE-RAUZY, O., COITEUX, V., SAPENA, R., LAZARETH, A., LEROUX, G., GUENDA, K., CASSINAT, B., FONTENAY, M., VEY, N., GUERCI, A., DREYFUS, F., BORDESSOULE, D., STAMATOULLAS, A., CASTAIGNE, S., TERRE, C., ECLACHE, V., FENAUX, P. & ADES, L. 2011. Characteristics and outcome of myelodysplastic syndromes (MDS) with isolated 20q deletion: a report on 62 cases. *Leuk Res,* 35**,** 863-7.

BRECCIA, M., LATAGLIATA, R., CANNELLA, L., CARMOSINO, I., SANTOPIETRO, M., LOGLISCI, G., FEDERICO, V. & ALIMENA, G. 2010. Refractory cytopenia with unilineage dysplasia: analysis of prognostic factors and survival in 126 patients. *Leuk Lymphoma,* 51**,** 783-8.

BRIGHT, J., HUSSAIN, S., DANG, V., WRIGHT, S., COOPER, B., BYUN, T., RAMOS, C., SINGH, A., PARRY, G., STAGLIANO, N. & GRISWOLD-PRENNER, I. 2015. Human secreted tau increases amyloid-beta production. *Neurobiol Aging,* 36**,** 693-709.

BRONS, I. G., SMITHERS, L. E., TROTTER, M. W., RUGG-GUNN, P., SUN, B., CHUVA DE SOUSA LOPES, S. M., HOWLETT, S. K., CLARKSON, A., AHRLUND-RICHTER, L., PEDERSEN, R. A. & VALLIER, L. 2007. Derivation of pluripotent epiblast stem cells from mammalian embryos. *Nature,* 448**,** 191-5.

BROSH, R., ASSIA-ALROY, Y., MOLCHADSKY, A., BORNSTEIN, C., DEKEL, E., MADAR, S., SHETZER, Y., RIVLIN, N., GOLDFINGER, N., SARIG, R. & ROTTER, V. 2013. p53 counteracts reprogramming by inhibiting mesenchymal-to-epithelial transition. *Cell Death Differ,* 20**,** 312-20.

BROUWER, M., ZHOU, H. & NADIF KASRI, N. 2016. Choices for Induction of Pluripotency: Recent Developments in Human Induced Pluripotent Stem Cell Reprogramming Strategies. *Stem Cell Rev Rep,* 12**,** 54-72.

BUGANIM, Y., FADDAH, D. A., CHENG, A. W., ITSKOVICH, E., MARKOULAKI, S., GANZ, K., KLEMM, S. L., VAN OUDENAARDEN, A. & JAENISCH, R. 2012. Single-cell expression analyses during cellular reprogramming reveal an early stochastic and a late hierarchic phase. *Cell,* 150**,** 1209-22.

BUSCARLET, M., PROVOST, S., ZADA, Y. F., BARHDADI, A., BOURGOIN, V., LEPINE, G., MOLLICA, L., SZUBER, N., DUBE, M. P. & BUSQUE, L. 2017. DNMT3A and TET2 dominate clonal hematopoiesis and demonstrate benign phenotypes and different genetic predispositions. *Blood,* 130**,** 753-762.

BUSQUE, L., MIO, R., MATTIOLI, J., BRAIS, E., BLAIS, N., LALONDE, Y., MARAGH, M. & GILLILAND, D. G. 1996. Nonrandom X-inactivation patterns in normal females: lyonization ratios vary with age. *Blood,* 88**,** 59-65.

BUSQUE, L., PATEL, J. P., FIGUEROA, M. E., VASANTHAKUMAR, A., PROVOST, S., HAMILOU, Z., MOLLICA, L., LI, J., VIALE, A., HEGUY, A., HASSIMI, M., SOCCI, N., BHATT, P. K., GONEN, M., MASON, C. E., MELNICK, A., GODLEY, L. A., BRENNAN, C. W., ABDEL-WAHAB, O. & LEVINE, R. L. 2012. Recurrent somatic TET2 mutations in normal elderly individuals with clonal hematopoiesis. *Nat Genet,* 44**,** 1179-81.

BUZA-VIDAS, N., ANTONCHUK, J., QIAN, H., MANSSON, R., LUC, S., ZANDI, S., ANDERSON, K., TAKAKI, S., NYGREN, J. M., JENSEN, C. T. & JACOBSEN, S. E. 2006. Cytokines regulate postnatal hematopoietic stem cell expansion: opposing roles of thrombopoietin and LNK. *Genes Dev,* 20**,** 2018-23.

CAMMENGA, J., MULLOY, J. C., BERGUIDO, F. J., MACGROGAN, D., VIALE, A. & NIMER, S. D. 2003. Induction of C/EBPalpha activity alters gene expression and differentiation of human CD34+ cells. *Blood,* 101**,** 2206-14.

CAMPO, E., SWERDLOW, S. H., HARRIS, N. L., PILERI, S., STEIN, H. & JAFFE, E. S. 2011. The 2008 WHO classification of lymphoid neoplasms and beyond: evolving concepts and practical applications. *Blood,* 117**,** 5019-32.

CANTU, I. & PHILIPSEN, S. 2014. adult hemoglobin expression in erythroid cells derived from cord blood and human induced pluripotent stem cells. *Haematologica,* 99**,** 1647-1649.

CAREY, B. W., MARKOULAKI, S., HANNA, J., SAHA, K., GAO, Q., MITALIPOVA, M. & JAENISCH, R. 2009. Reprogramming of murine and human somatic cells using a single polycistronic vector. *Proc Natl Acad Sci U S A,* 106**,** 157-62.

CARRELHA, J., MENG, Y., KETTYLE, L. M., LUIS, T. C., NORFO, R., ALCOLEA, V., BOUKARABILA, H., GRASSO, F., GAMBARDELLA, A., GROVER, A., HOGSTRAND, K., LORD, A. M., SANJUAN-PLA, A., WOLL, P. S., NERLOV, C. & JACOBSEN, S. E. W. 2018. Hierarchically related lineage-restricted fates of multipotent haematopoietic stem cells. *Nature,* 554**,** 106-111.

CASTANO, J., BUENO, C., JIMENEZ-DELGADO, S., ROCA-HO, H., FRAGA, M. F., FERNANDEZ, A. F., NAKANISHI, M., TORRES-RUIZ, R., RODRIGUEZ-PERALES, S. & MENENDEZ, P. 2017. Generation and characterization of a human iPSC cell line expressing inducible Cas9 in the "safe harbor" AAVS1 locus. *Stem Cell Res,* 21**,** 137-140.

CATHOMEN, T. & KEITH JOUNG, J. 2008. Zinc-finger Nucleases: The Next Generation Emerges. *Mol Ther,* 16**,** 1200-1207.

CENTIS, F., TABELLINI, L., LUCARELLI, G., BUFFI, O., TONUCCI, P., PERSINI, B., ANNIBALI, M., EMILIANI, R., ILIESCU, A., RAPA, S., ROSSI, R., MA, L., ANGELUCCI, E. & SCHRIER, S. L. 2000. The importance of erythroid expansion in determining the extent of apoptosis in erythroid precursors in patients with beta-thalassemia major. *Blood,* 96**,** 3624-9.

CHANG, C., KOTINI, A. G., OLSZEWSKA, M., GEORGOMANOLI, M., TERUYA-FELDSTEIN, J., SPERBER, H., SANCHEZ, R., DEVITA, R., MARTINS, T., ABDELWAHAB, O., BRADLEY, R. & PAPAPETROU, E. P. 2018. Dissecting the contributions of cooperating gene mutations to cancer phenotypes and drug responses with patient-derived iPSCs. *Stem cell reports,* 10**,** 1610-1624.

CHAO, M. P., GENTLES, A. J., CHATTERJEE, S., LAN, F., REINISCH, A., CORCES, M. R., XAVY, S., SHEN, J., HAAG, D., CHANDA, S., SINHA, R., MORGANTI, R. M., NISHIMURA, T., AMEEN, M., WU, H., WERNIG, M., WU, J. C. & MAJETI, R. 2017. Human AML-iPSCs Reacquire Leukemic Properties after Differentiation and Model Clonal Variation of Disease. *Cell Stem Cell,* 20**,** 329-344 e7.

CHARBORD, P. & IN ZON, L. 2001. *Microenvironmental cell populations essential for the support of hematopoietic stem cells,* New york, Hematopoiesis: a developmental approach. Oxford University Press.

CHARLESWORTH, C. T., DESHPANDE, P. S., DEVER, D. P., CAMARENA, J., LEMGART, V. T., CROMER, M. K., VAKULSKAS, C. A., COLLINGWOOD, M. A., ZHANG, L., BODE, N. M., BEHLKE, M. A., DEJENE, B., CIENIEWICZ, B., ROMANO, R., LESCH, B. J., GOMEZ-OSPINA, N., MANTRI, S., PAVEL-DINU, M., WEINBERG, K. I. & PORTEUS, M. H. 2019. Identification of preexisting adaptive immunity to Cas9 proteins in humans. *Nat Med,* 25**,** 249-254.

CHASIS, J. A. & MOHANDAS, N. 2008. Erythroblastic islands: niches for erythropoiesis. *Blood,* 112**,** 470-8.

CHEN, B., GILBERT, L. A., CIMINI, B. A., SCHNITZBAUER, J., ZHANG, W., LI, G. W., PARK, J., BLACKBURN, E. H., WEISSMAN, J. S., QI, L. S. & HUANG, B. 2013. Dynamic imaging of genomic loci in living human cells by an optimized CRISPR/Cas system. *Cell,* 155**,** 1479-91.

CHEN, J., KAO, Y. R., SUN, D., TODOROVA, T. I., REYNOLDS, D., NARAYANAGARI, S. R., MONTAGNA, C., WILL, B., VERMA, A. & STEIDL, U. 2019. Myelodysplastic syndrome progression to acute myeloid leukemia at the stem cell level. *Nat Med,* 25**,** 103-110.

CHEN, J. S., DAGDAS, Y. S., KLEINSTIVER, B. P., WELCH, M. M., SOUSA, A. A., HARRINGTON, L. B., STERNBERG, S. H., JOUNG, J. K., YILDIZ, A. & DOUDNA, J. A. 2017. Enhanced proofreading governs CRISPR-Cas9 targeting accuracy. *Nature,* 550**,** 407-410.

CHEN, K., LIU, J., HECK, S., CHASIS, J. A., AN, X. & MOHANDAS, N. 2009. Resolving the distinct stages in erythroid differentiation based on dynamic changes in membrane protein expression during erythropoiesis. *Proc Natl Acad Sci U S A,* 106**,** 17413-8.

CHEN, L., CHEN, J. Y., HUANG, Y. J., GU, Y., QIU, J., QIAN, H., SHAO, C., ZHANG, X., HU, J., LI, H., HE, S., ZHOU, Y., ABDEL-WAHAB, O., ZHANG, D. E. & FU, X. D. 2018. The Augmented R-Loop Is a Unifying Mechanism for Myelodysplastic Syndromes Induced by High-Risk Splicing Factor Mutations. *Mol Cell,* 69**,** 412-425 e6.

CHEN, X., JOHNS, D. C., GEIMAN, D. E., MARBAN, E., DANG, D. T., HAMLIN, G., SUN, R. & YANG, V. W. 2001. Kruppel-like factor 4 (gut-enriched Kruppel-like factor) inhibits cell proliferation by blocking G1/S progression of the cell cycle. *J Biol Chem,* 276**,** 30423-8.

CHICHA, L., FEKI, A., BONI, A., IRION, O., HOVATTA, O. & JACONI, M. 2011. Human pluripotent stem cells differentiated in fully defined medium generate hematopoietic CD34- and CD34+ progenitors with distinct characteristics. *PLoS One,* 6**,** e14733.

CHO, N. K., KEYES, L., JOHNSON, E., HELLER, J., RYNER, L., KARIM, F. & KRASNOW, M. A. 2002. Developmental control of blood cell migration by the Drosophila VEGF pathway. *Cell,* 108**,** 865-76.

CLARKE, M., DUMON, S., WARD, C., JAGER, R., FREEMAN, S., DAWOOD, B., SHERIFF, L., LORVELLEC, M., KRALOVICS, R., FRAMPTON, J. & GARCIA, P. 2013. MYBL2 haploinsufficiency increases susceptibility to age-related haematopoietic neoplasia. *Leukemia,* 27**,** 661-70.

COLLINS, C., WANG, J., MIAO, H., BRONSTEIN, J., NAWER, H., XU, T., FIGUEROA, M., MUNTEAN, A. G. & HESS, J. L. 2014. C/EBPalpha is an essential collaborator in Hoxa9/Meis1-mediated leukemogenesis. *Proc Natl Acad Sci U S A,* 111**,** 9899-904.

COLUNGA-PEDRAZA, P. R., COLUNGA-PEDRAZA, J. E., GARZA-LEDEZMA, M. A., JAIME-PEREZ, J. C., CANTU-RODRIGUEZ, O. G., GUTIERREZ-AGUIRRE, C. H., RENDON-RAMIREZ, E. J., LOPEZ-GARCIA, Y. K., LOZANO-MORALES, R. E., GOMEZ-DE LEON, A., SOTOMAYOR-DUQUE, G. & GOMEZ-ALMAGUER, D. 2018. Danazol as First-Line Therapy for Myelodysplastic Syndrome. *Clin Lymphoma Myeloma Leuk,* 18**,** e109-e113.

CONG, L., RAN, F. A., COX, D., LIN, S., BARRETTO, R., HABIB, N., HSU, P. D., WU, X., JIANG, W., MARRAFFINI, L. A. & ZHANG, F. 2013. Multiplex genome engineering using CRISPR/Cas systems. *Science,* 339**,** 819-23.

COOPER, G. M. & SUNDERLAND, M. A. 2000. *The Cell: A Molecular Approach,* USA.

COOPER, S., GUO, H. & FRIEDMAN, A. D. 2015. The +37 kb Cebpa Enhancer Is Critical for Cebpa Myeloid Gene Expression and Contains Functional Sites that Bind SCL, GATA2, C/EBPalpha, PU.1, and Additional Ets Factors. *PLoS One,* 10**,** e0126385.

COREY, S. J., MINDEN, M. D., BARBER, D. L., KANTARJIAN, H., WANG, J. C. & SCHIMMER, A. D. 2007. Myelodysplastic syndromes: the complexity of stem-cell diseases. *Nat Rev Cancer,* 7**,** 118-29.

COSKUN, S., CHAO, H., VASAVADA, H., HEYDARI, K., GONZALES, N., ZHOU, X., DE CROMBRUGGHE, B. & HIRSCHI, K. K. 2014. Development of the fetal bone marrow niche and regulation of HSC quiescence and homing ability by emerging osteolineage cells. *Cell Rep,* 9**,** 581-90.

CRADDOCK, C. F., HOULTON, A. E., QUEK, L. S., FERGUSON, P., GBANDI, E., ROBERTS, C., METZNER, M., GARCIA-MARTIN, N., KENNEDY, A., HAMBLIN, A., RAGHAVAN, M., NAGRA, S., DUDLEY, L., WHEATLEY, K., MCMULLIN, M. F., PILLAI, S. P., KELLY, R. J., SIDDIQUE, S., DENNIS, M., CAVENAGH, J. D. & VYAS, P. 2017. Outcome of Azacitidine Therapy in Acute Myeloid Leukemia Is not Improved by Concurrent Vorinostat Therapy but Is Predicted by a Diagnostic Molecular Signature. *Clin Cancer Res,* 23**,** 6430-6440.

CRADICK, T. J., QIU, P., LEE, C. M., FINE, E. J. & BAO, G. 2014. COSMID: A Web-based Tool for Identifying and Validating CRISPR/Cas Off-target Sites. *Mol Ther Nucleic Acids,* 3**,** e214.

CRISA, E., KULASEKARARAJ, A. G., ADEMA, V., SUCH, E., SCHANZ, J., HAASE, D., SHIRNESHAN, K., BEST, S., MIAN, S. A., KIZILORS, A., CERVERA, J., LEA, N., FERRERO, D., GERMING, U., HILDEBRANDT, B., MARTINEZ, A. B. V., SANTINI, V., SANZ, G. F., SOLE, F. & MUFTI, G. J. 2020. Impact of somatic mutations in myelodysplastic patients with isolated partial or total loss of chromosome 7. *Leukemia*.

CROSBY, W. H. 1959. Normal functions of the spleen relative to red blood cells: a review. *Blood,* 14**,** 399-408.

CROZE, R. H., THI, W. J. & CLEGG, D. O. 2016. ROCK Inhibition Promotes Attachment, Proliferation, and Wound Closure in Human Embryonic Stem Cell-Derived Retinal Pigmented Epithelium. *Transl Vis Sci Technol,* 5**,** 7.

CRUDELE, J. M. & CHAMBERLAIN, J. S. 2018. Cas9 immunity creates challenges for CRISPR gene editing therapies. *Nat Commun,* 9**,** 3497.

D'AMOUR, K. A., AGULNICK, A. D., ELIAZER, S., KELLY, O. G., KROON, E. & BAETGE, E. E. 2005. Efficient differentiation of human embryonic stem cells to definitive endoderm. *Nat Biotechnol,* 23**,** 1534-41.

DANG, Y., JIA, G., CHOI, J., MA, H., ANAYA, E., YE, C., SHANKAR, P. & WU, H. 2015. Optimizing sgRNA structure to improve CRISPR-Cas9 knockout efficiency. *Genome Biol,* 16**,** 280.

DE WITTE, T., BOWEN, D., ROBIN, M., MALCOVATI, L., NIEDERWIESER, D., YAKOUB-AGHA, I., MUFTI, G. J., FENAUX, P., SANZ, G., MARTINO, R., ALESSANDRINO, E. P., ONIDA, F., SYMEONIDIS, A., PASSWEG, J., KOBBE, G., GANSER, A., PLATZBECKER, U., FINKE, J., VAN GELDER, M., VAN DE LOOSDRECHT, A. A., LJUNGMAN, P., STAUDER, R., VOLIN, L., DEEG, H. J., CUTLER, C., SABER, W., CHAMPLIN, R., GIRALT, S., ANASETTI, C. & KROGER, N. 2017. Allogeneic hematopoietic stem cell transplantation for MDS and CMML: recommendations from an international expert panel. *Blood,* 129**,** 1753-1762.

DELHOMMEAU, F., DUPONT, S., DELLA VALLE, V., JAMES, C., TRANNOY, S., MASSE, A., KOSMIDER, O., LE COUEDIC, J. P., ROBERT, F., ALBERDI, A., LECLUSE, Y., PLO, I., DREYFUS, F. J., MARZAC, C., CASADEVALL, N., LACOMBE, C., ROMANA, S. P., DESSEN, P., SOULIER, J., VIGUIE, F., FONTENAY, M., VAINCHENKER, W. & BERNARD, O. A. 2009. Mutation in TET2 in myeloid cancers. *N Engl J Med,* 360**,** 2289-301.

DEZERN, A. E., PU, J., MCDEVITT, M. A., JONES, R. J. & BRODSKY, R. A. 2013. Burst-forming unit-erythroid assays to distinguish cellular bone marrow failure disorders. *Exp Hematol,* 41**,** 808-16.

DI GENUA, C., VALLETTA, S., BUONO, M., STOILOVA, B., SWEENEY, C., RODRIGUEZ-MEIRA, A., GROVER, A., DRISSEN, R., MENG, Y., BEVERIDGE, R., ABOUKHALIL, Z., KARAMITROS, D., BELDERBOS, M. E., BYSTRYKH, L., THONGJUEA, S., VYAS, P. & NERLOV, C. 2020. C/EBPalpha and GATA-2 Mutations Induce Bilineage Acute Erythroid Leukemia through Transformation of a Neomorphic Neutrophil-Erythroid Progenitor. *Cancer Cell,* 37**,** 690-704 e8.

DITADI, A., STURGEON, C. M. & KELLER, G. 2017. A view of human haematopoietic development from the Petri dish. *Nat Rev Mol Cell Biol,* 18**,** 56-67.

DONAHUE, R. E., SEEHRA, J., METZGER, M., LEFEBVRE, D., ROCK, B., CARBONE, S., NATHAN, D. G., GARNICK, M., SEHGAL, P. K., LASTON, D. & ET AL. 1988. Human IL-3 and GM-CSF act synergistically in stimulating hematopoiesis in primates. *Science,* 241**,** 1820-3.

DONNER, P., GREISER-WILKE, I. & MOELLING, K. 1982. Nuclear localization and DNA binding of the transforming gene product of avian myelocytomatosis virus. *Nature,* 296**,** 262-9.

DOUDNA, J. A. & CHARPENTIER, E. 2014. Genome editing. The new frontier of genome engineering with CRISPR-Cas9. *Science,* 346**,** 1258096.

DUAN, Q., LI, S., WEN, X., SUNNASSEE, G., CHEN, J., TAN, S. & GUO, Y. 2019. Valproic Acid Enhances Reprogramming Efficiency and Neuronal Differentiation on Small Molecules Staged-Induction Neural Stem Cells: Suggested Role of mTOR Signaling. *Front Neurosci,* 13**,** 867.

DZIERZAK, E. & PHILIPSEN, S. 2013. Erythropoiesis: development and differentiation. *Cold Spring Harb Perspect Med,* 3**,** a011601.

EMINLI, S., FOUDI, A., STADTFELD, M., MAHERALI, N., AHFELDT, T., MOSTOSLAVSKY, G., HOCK, H. & HOCHEDLINGER, K. 2009. Differentiation stage determines potential of hematopoietic cells for reprogramming into induced pluripotent stem cells. *Nat Genet,* 41**,** 968-76.

ESTEBAN, M. A., WANG, T., QIN, B., YANG, J., QIN, D., CAI, J., LI, W., WENG, Z., CHEN, J., NI, S., CHEN, K., LI, Y., LIU, X., XU, J., ZHANG, S., LI, F., HE, W., LABUDA, K., SONG, Y., PETERBAUER, A., WOLBANK, S., REDL, H., ZHONG, M., CAI, D., ZENG, L. & PEI, D. 2010. Vitamin C enhances the generation of mouse and human induced pluripotent stem cells. *Cell Stem Cell,* 6**,** 71-9.

EVANS, M. J. & KAUFMAN, M. H. 1981. Establishment in culture of pluripotential cells from mouse embryos. *Nature,* 292**,** 154-6.

FAGNOCCHI, L., CHERUBINI, A., HATSUDA, H., FASCIANI, A., MAZZOLENI, S., POLI, V., BERNO, V., ROSSI, R. L., REINBOLD, R., ENDELE, M., SCHROEDER, T., ROCCHIGIANI, M., SZKARLAT, Z., OLIVIERO, S., DALTON, S. & ZIPPO, A. 2016. A Myc-driven self-reinforcing regulatory network maintains mouse embryonic stem cell identity. *Nat Commun,* 7**,** 11903.

FEINBERG, M. W., WARA, A. K., CAO, Z., LEBEDEVA, M. A., ROSENBAUER, F., IWASAKI, H., HIRAI, H., KATZ, J. P., HASPEL, R. L., GRAY, S., AKASHI, K., SEGRE, J., KAESTNER, K. H., TENEN, D. G. & JAIN, M. K. 2007. The Kruppel-like factor KLF4 is a critical regulator of monocyte differentiation. *EMBO J,* 26**,** 4138-48.

FENAUX, P., GIAGOUNIDIS, A., SELLESLAG, D., BEYNE-RAUZY, O., MUFTI, G., MITTELMAN, M., MUUS, P., TE BOEKHORST, P., SANZ, G., DEL CANIZO, C., GUERCI-BRESLER, A., NILSSON, L., PLATZBECKER, U., LUBBERT, M., QUESNEL, B., CAZZOLA, M., GANSER, A., BOWEN, D., SCHLEGELBERGER, B., AUL, C., KNIGHT, R., FRANCIS, J., FU, T., HELLSTROM-LINDBERG, E. & GROUP, M. D. S. L. D. Q. S. 2011. A randomized phase 3 study of lenalidomide versus placebo in RBC transfusion-dependent patients with Low-/Intermediate-1-risk myelodysplastic syndromes with del5q. *Blood,* 118**,** 3765-76.

FENAUX, P., MUFTI, G. J., HELLSTROM-LINDBERG, E., SANTINI, V., FINELLI, C., GIAGOUNIDIS, A., SCHOCH, R., GATTERMANN, N., SANZ, G., LIST, A., GORE, S. D., SEYMOUR, J. F., BENNETT, J. M., BYRD, J., BACKSTROM, J., ZIMMERMAN, L., MCKENZIE, D., BEACH, C., SILVERMAN, L. R. & INTERNATIONAL VIDAZA HIGH-RISK, M. D. S. S. S. G. 2009. Efficacy of azacitidine compared with that of conventional care regimens in the treatment of higher-risk myelodysplastic syndromes: a randomised, open-label, phase III study. *Lancet Oncol,* 10**,** 223-32.

FENG, B., JIANG, J., KRAUS, P., NG, J. H., HENG, J. C., CHAN, Y. S., YAW, L. P., ZHANG, W., LOH, Y. H., HAN, J., VEGA, V. B., CACHEUX-RATABOUL, V., LIM, B., LUFKIN, T. & NG, H. H. 2009. Reprogramming of fibroblasts into induced pluripotent stem cells with orphan nuclear receptor Esrrb. *Nat Cell Biol,* 11**,** 197-203.

FILLET, G. & BEGUIN, Y. 2001. Monitoring of erythropoiesis by the serum transferrin receptor and erythropoietin. *Acta Clin Belg,* 56**,** 146-54.

FLYGARE, J., RAYON ESTRADA, V., SHIN, C., GUPTA, S. & LODISH, H. F. 2011. HIF1alpha synergizes with glucocorticoids to promote BFU-E progenitor self-renewal. *Blood,* 117**,** 3435-44.

FONG, H., HOHENSTEIN, K. A. & DONOVAN, P. J. 2008. Regulation of self-renewal and pluripotency by Sox2 in human embryonic stem cells. *Stem Cells,* 26**,** 1931-8.

FORSBERG, E. C., SERWOLD, T., KOGAN, S., WEISSMAN, I. L. & PASSEGUE, E. 2006. New evidence supporting megakaryocyte-erythrocyte potential of flk2/flt3+ multipotent hematopoietic progenitors. *Cell,* 126**,** 415-26.

FRAME, J. M., MCGRATH, K. E. & PALIS, J. 2013. Erythro-myeloid progenitors: "definitive" hematopoiesis in the conceptus prior to the emergence of hematopoietic stem cells. *Blood Cells Mol Dis,* 51**,** 220-5.

FRIEDMAN, A. D. 2015. C/EBPalpha in normal and malignant myelopoiesis. *Int J Hematol,* 101**,** 330-41.

FU, X. D. & MANIATIS, T. 1990. Factor required for mammalian spliceosome assembly is localized to discrete regions in the nucleus. *Nature,* 343**,** 437-41.

FU, Y., FODEN, J. A., KHAYTER, C., MAEDER, M. L., REYON, D., JOUNG, J. K. & SANDER, J. D. 2013. High-frequency off-target mutagenesis induced by CRISPR-Cas nucleases in human cells. *Nat Biotechnol,* 31**,** 822-6.

FU, Y., SANDER, J. D., REYON, D., CASCIO, V. M. & JOUNG, J. K. 2014. Improving CRISPR-Cas nuclease specificity using truncated guide RNAs. *Nat Biotechnol,* 32**,** 279-284.

FUSAKI, N., BAN, H., NISHIYAMA, A., SAEKI, K. & HASEGAWA, M. 2009. Efficient induction of transgene-free human pluripotent stem cells using a vector based on Sendai virus, an RNA virus that does not integrate into the host genome. *Proc Jpn Acad Ser B Phys Biol Sci,* 85**,** 348-62.

GALLOWAY, J. L. & ZON, L. I. 2003. Ontogeny of hematopoiesis: examining the emergence of hematopoietic cells in the vertebrate embryo. *Curr Top Dev Biol,* 53**,** 139-58.

GAO, L., NATH, S. C., JIAO, X., ZHOU, R., NISHIKAWA, S., KRAWETZ, R., LI, X. & RANCOURT, D. E. 2019. Post-Passage rock inhibition induces cytoskeletal aberrations and apoptosis in Human embryonic stem cells. *Stem Cell Res,* 41**,** 101641.

GENG, B. C., CHOI, K. H., WANG, S. Z., CHEN, P., PAN, X. D., DONG, N. G., KO, J. K. & ZHU, H. 2020. A simple, quick, and efficient CRISPR/Cas9 genome editing method for human induced pluripotent stem cells. *Acta Pharmacol Sin,* 41**,** 1427-1432.

GENOVESE, G., KAHLER, A. K., HANDSAKER, R. E., LINDBERG, J., ROSE, S. A., BAKHOUM, S. F., CHAMBERT, K., MICK, E., NEALE, B. M., FROMER, M., PURCELL, S. M., SVANTESSON, O., LANDEN, M., HOGLUND, M., LEHMANN, S., GABRIEL, S. B., MORAN, J. L., LANDER, E. S., SULLIVAN, P. F., SKLAR, P., GRONBERG, H., HULTMAN, C. M. & MCCARROLL, S. A. 2014. Clonal hematopoiesis and blood-cancer risk inferred from blood DNA sequence. *N Engl J Med,* 371**,** 2477-87.

GEORGOMANOLI, M. & PAPAPETROU, E. P. 2019. Modeling blood diseases with human induced pluripotent stem cells. *Dis Model Mech,* 12.

GERMING, U., LAUSEKER, M., HILDEBRANDT, B., SYMEONIDIS, A., CERMAK, J., FENAUX, P., KELAIDI, C., PFEILSTOCKER, M., NOSSLINGER, T., SEKERES, M., MACIEJEWSKI, J., HAASE, D., SCHANZ, J., SEYMOUR, J., KENEALY, M., WEIDE, R., LUBBERT, M., PLATZBECKER, U., VALENT, P., GOTZE, K., STAUDER, R., BLUM, S., KREUZER, K. A., SCHLENK, R., GANSER, A., HOFMANN, W. K., AUL, C., KRIEGER, O., KUNDGEN, A., HAAS, R., HASFORD, J. & GIAGOUNIDIS, A. 2012. Survival, prognostic factors and rates of leukemic transformation in 381 untreated patients with MDS and del(5q): a multicenter study. *Leukemia,* 26**,** 1286-92.

GERRITSEN, M., YI, G., TIJCHON, E., KUSTER, J., SCHURINGA, J. J., MARTENS, J. H. A. & VELLENGA, E. 2019. RUNX1 mutations enhance self-renewal and block granulocytic differentiation in human in vitro models and primary AMLs. *Blood Adv,* 3**,** 320-332.

GIFFORD, S. C., DERGANC, J., SHEVKOPLYAS, S. S., YOSHIDA, T. & BITENSKY, M. W. 2006. A detailed study of time-dependent changes in human red blood cells: from reticulocyte maturation to erythrocyte senescence. *Br J Haematol,* 135**,** 395-404.

GOMBART, A. F., HOFMANN, W. K., KAWANO, S., TAKEUCHI, S., KRUG, U., KWOK, S. H., LARSEN, R. J., ASOU, H., MILLER, C. W., HOELZER, D. & KOEFFLER, H. P. 2002. Mutations in the gene encoding the transcription factor CCAAT/enhancer binding protein alpha in myelodysplastic syndromes and acute myeloid leukemias. *Blood,* 99**,** 1332-40.

GOMEZ LIMIA, C. E., DEVALLE, S., REIS, M., SOCHACKI, J., CARNEIRO, M., MADEIRO DA COSTA, R., D'ANDREA, M., PADILHA, T., ZALCBERG, I. R., SOLZA, C., DAUMAS, A., REHEN, S., MONTE-MOR, B. & BONAMINO, M. H. 2017. Generation and characterization of a human induced pluripotent stem (iPS) cell line derived from an acute myeloid leukemia patient evolving from primary myelofibrosis carrying the CALR 52bp deletion and the ASXL1 p.R693X mutation. *Stem Cell Res,* 24**,** 16-20.

GRANICK, S. & LEVERE, R. D. 1964. Heme Synthesis in Erythroid Cells. *Prog Hematol,* 4**,** 1-47.

GREBIEN, F., KERENYI, M. A., KOVACIC, B., KOLBE, T., BECKER, V., DOLZNIG, H., PFEFFER, K., KLINGMULLER, U., MULLER, M., BEUG, H., MULLNER, E. W. & MORIGGL, R. 2008. Stat5 activation enables erythropoiesis in the absence of EpoR and Jak2. *Blood,* 111**,** 4511-22.

GREENBERG, P., COX, C., LEBEAU, M. M., FENAUX, P., MOREL, P., SANZ, G., SANZ, M., VALLESPI, T., HAMBLIN, T., OSCIER, D., OHYASHIKI, K., TOYAMA, K., AUL, C., MUFTI, G. & BENNETT, J. 1997. International scoring system for evaluating prognosis in myelodysplastic syndromes. *Blood,* 89**,** 2079-88.

GREENBERG, P. L., TUECHLER, H., SCHANZ, J., SANZ, G., GARCIA-MANERO, G., SOLE, F., BENNETT, J. M., BOWEN, D., FENAUX, P., DREYFUS, F., KANTARJIAN, H., KUENDGEN, A., LEVIS, A., MALCOVATI, L., CAZZOLA, M., CERMAK, J., FONATSCH, C., LE BEAU, M. M., SLOVAK, M. L., KRIEGER, O., LUEBBERT, M., MACIEJEWSKI, J., MAGALHAES, S. M., MIYAZAKI, Y., PFEILSTOCKER, M., SEKERES, M., SPERR, W. R., STAUDER, R., TAURO, S., VALENT, P., VALLESPI, T., VAN DE LOOSDRECHT, A. A., GERMING, U. & HAASE, D. 2012. Revised international prognostic scoring system for myelodysplastic syndromes. *Blood,* 120**,** 2454-65.

GREGORY, C. J. & EAVES, A. C. 1977. Human marrow cells capable of erythropoietic differentiation in vitro: definition of three erythroid colony responses. *Blood,* 49**,** 855-64.

GRINFELD, J., NANGALIA, J., BAXTER, E. J., WEDGE, D. C., ANGELOPOULOS, N., CANTRILL, R., GODFREY, A. L., PAPAEMMANUIL, E., GUNDEM, G., MACLEAN, C., COOK, J., O'NEIL, L., O'MEARA, S., TEAGUE, J. W., BUTLER, A. P., MASSIE, C. E., WILLIAMS, N., NICE, F. L., ANDERSEN, C. L., HASSELBALCH, H. C., GUGLIELMELLI, P., MCMULLIN, M. F., VANNUCCHI, A. M., HARRISON, C. N., GERSTUNG, M., GREEN, A. R. & CAMPBELL, P. J. 2018. Classification and Personalized Prognosis in Myeloproliferative Neoplasms. *N Engl J Med,* 379**,** 1416-1430.

GROVER, A., MANCINI, E., MOORE, S., MEAD, A. J., ATKINSON, D., RASMUSSEN, K. D., O'CARROLL, D., JACOBSEN, S. E. & NERLOV, C. 2014. Erythropoietin guides multipotent hematopoietic progenitor cells toward an erythroid fate. *J Exp Med,* 211**,** 181-8.

GUBBAY, J., COLLIGNON, J., KOOPMAN, P., CAPEL, B., ECONOMOU, A., MUNSTERBERG, A., VIVIAN, N., GOODFELLOW, P. & LOVELL-BADGE, R. 1990. A gene mapping to the sex-determining region of the mouse Y chromosome is a member of a novel family of embryonically expressed genes. *Nature,* 346**,** 245-50.

GUO, H., MA, O., SPECK, N. A. & FRIEDMAN, A. D. 2012. Runx1 deletion or dominant inhibition reduces Cebpa transcription via conserved promoter and distal enhancer sites to favor monopoiesis over granulopoiesis. *Blood,* 119**,** 4408-18.

GUPTA, R., SOUPIR, C. P., JOHARI, V. & HASSERJIAN, R. P. 2007. Myelodysplastic syndrome with isolated deletion of chromosome 20q: an indolent disease with minimal morphological dysplasia and frequent thrombocytopenic presentation. *Br J Haematol,* 139**,** 265-8.

GUPTA, R. M. & MUSUNURU, K. 2014. Expanding the genetic editing tool kit: ZFNs, TALENs, and CRISPR-Cas9. *J Clin Invest,* 124**,** 4154-61.

GURDON, J. B. 1962. The developmental capacity of nuclei taken from intestinal epithelium cells of feeding tadpoles. *J Embryol Exp Morphol,* 10**,** 622-40.

HAAS, S., HANSSON, J., KLIMMECK, D., LOEFFLER, D., VELTEN, L., UCKELMANN, H., WURZER, S., PRENDERGAST, A. M., SCHNELL, A., HEXEL, K., SANTARELLA-MELLWIG, R., BLASZKIEWICZ, S., KUCK, A., GEIGER, H., MILSOM, M. D., STEINMETZ, L. M., SCHROEDER, T., TRUMPP, A., KRIJGSVELD, J. & ESSERS, M. A. 2015. Inflammation-Induced Emergency Megakaryopoiesis Driven by Hematopoietic Stem Cell-like Megakaryocyte Progenitors. *Cell Stem Cell,* 17**,** 422-34.

HAFERLACH, T., NAGATA, Y., GROSSMANN, V., OKUNO, Y., BACHER, U., NAGAE, G., SCHNITTGER, S., SANADA, M., KON, A., ALPERMANN, T., YOSHIDA, K., ROLLER, A., NADARAJAH, N., SHIRAISHI, Y., SHIOZAWA, Y., CHIBA, K., TANAKA, H., KOEFFLER, H. P., KLEIN, H. U., DUGAS, M., ABURATANI, H., KOHLMANN, A., MIYANO, S., HAFERLACH, C., KERN, W. & OGAWA, S. 2014. Landscape of genetic lesions in 944 patients with myelodysplastic syndromes. *Leukemia,* 28**,** 241-7.

HAHN, C. N., CHONG, C. E., CARMICHAEL, C. L., WILKINS, E. J., BRAUTIGAN, P. J., LI, X. C., BABIC, M., LIN, M., CARMAGNAC, A., LEE, Y. K., KOK, C. H., GAGLIARDI, L., FRIEND, K. L., EKERT, P. G., BUTCHER, C. M., BROWN, A. L., LEWIS, I. D., TO, L. B., TIMMS, A. E., STOREK, J., MOORE, S., ALTREE, M., ESCHER, R., BARDY, P. G., SUTHERS, G. K., D'ANDREA, R. J., HORWITZ, M. S. & SCOTT, H. S. 2011. Heritable GATA2 mutations associated with familial myelodysplastic syndrome and acute myeloid leukemia. *Nat Genet,* 43**,** 1012-7.

HAN, S. M., HAN, S. H., COH, Y. R., JANG, G., CHAN RA, J., KANG, S. K., LEE, H. W. & YOUN, H. Y. 2014. Enhanced proliferation and differentiation of Oct4- and Sox2-overexpressing human adipose tissue mesenchymal stem cells. *Exp Mol Med,* 46**,** e101.

HANNA, J., MARKOULAKI, S., SCHORDERET, P., CAREY, B. W., BEARD, C., WERNIG, M., CREYGHTON, M. P., STEINE, E. J., CASSADY, J. P., FOREMAN, R., LENGNER, C. J., DAUSMAN, J. A. & JAENISCH, R. 2008. Direct reprogramming of terminally differentiated mature B lymphocytes to pluripotency. *Cell,* 133**,** 250-64.

HANSEN, M., VARGA, E., AARTS, C., WUST, T., KUIJPERS, T., VON LINDERN, M. & VAN DEN AKKER, E. 2018. Efficient production of erythroid, megakaryocytic and myeloid cells, using single cell-derived iPSC colony differentiation. *Stem Cell Res,* 29**,** 232-244.

HANSSON, A., ZETTERBLAD, J., VAN DUREN, C., AXELSON, H. & JONSSON, J. I. 2007. The Lim-only protein LMO2 acts as a positive regulator of erythroid differentiation. *Biochem Biophys Res Commun,* 364**,** 675-81.

HANSSON, J., RAFIEE, M. R., REILAND, S., POLO, J. M., GEHRING, J., OKAWA, S., HUBER, W., HOCHEDLINGER, K. & KRIJGSVELD, J. 2012. Highly coordinated proteome dynamics during reprogramming of somatic cells to pluripotency. *Cell Rep,* 2**,** 1579-92.

HARIA, P., MAJETHIA, N., MARGAM, S., GADGIL, N. & KALGUTKAR, A. 2014. Acute Myeloid Leukemia with Expanded Erythropoiesis; Diagnostic Dilemma. *international Journal of scientific study* 2.

HASSERJIAN, R. P., O'RAZI, A., BRUNNING, R. D., GERMING, U., LE BEAU, M. M., PORWIT, A., BAUMANN, I., HELLSTROM LINDBERG, E., LIST, A., CAZZOLA, M. & FOUCAR, K. 2017. Myelodysplastic syndromes/neoplasms, overview In: WHO Classifi- cation of Tumours of Haematopoietic and Lymphoid Tissues, revised, 4th edn. *Lyon, France: International Agency for Research on Cancer*.

HEATH, V., SUH, H. C., HOLMAN, M., RENN, K., GOOYA, J. M., PARKIN, S., KLARMANN, K. D., ORTIZ, M., JOHNSON, P. & KELLER, J. 2004. C/EBPalpha deficiency results in hyperproliferation of hematopoietic progenitor cells and disrupts macrophage development in vitro and in vivo. *Blood,* 104**,** 1639-47.

HEINRICHS, S., CONOVER, L. F., BUESO-RAMOS, C. E., KILPIVAARA, O., STEVENSON, K., NEUBERG, D., LOH, M. L., WU, W. S., RODIG, S. J., GARCIA-MANERO, G., KANTARJIAN, H. M. & LOOK, A. T. 2013. MYBL2 is a sub-haploinsufficient tumor suppressor gene in myeloid malignancy. *Elife,* 2**,** e00825.

HENDEL, A., BAK, R. O., CLARK, J. T., KENNEDY, A. B., RYAN, D. E., ROY, S., STEINFELD, I., LUNSTAD, B. D., KAISER, R. J., WILKENS, A. B., BACCHETTA, R., TSALENKO, A., DELLINGER, D., BRUHN, L. & PORTEUS, M. H. 2015. Chemically modified guide RNAs enhance CRISPR-Cas genome editing in human primary cells. *Nat Biotechnol,* 33**,** 985-989.

HIRAI, H., KOBAYASHI, Y., MANO, H., HAGIWARA, K., MARU, Y., OMINE, M., MIZOGUCHI, H., NISHIDA, J. & TAKAKU, F. 1987. A point mutation at codon 13 of the N-ras oncogene in myelodysplastic syndrome. *Nature,* 327**,** 430-2.

HOCHEDLINGER, K., YAMADA, Y., BEARD, C. & JAENISCH, R. 2005. Ectopic expression of Oct-4 blocks progenitor-cell differentiation and causes dysplasia in epithelial tissues. *Cell,* 121**,** 465-77.

HOCKEMEYER, D., SOLDNER, F., COOK, E. G., GAO, Q., MITALIPOVA, M. & JAENISCH, R. 2008. A drug-inducible system for direct reprogramming of human somatic cells to pluripotency. *Cell Stem Cell,* 3**,** 346-353.

HOCKEMEYER, D., WANG, H., KIANI, S., LAI, C. S., GAO, Q., CASSADY, J. P., COST, G. J., ZHANG, L., SANTIAGO, Y., MILLER, J. C., ZEITLER, B., CHERONE, J. M., MENG, X., HINKLEY, S. J., REBAR, E. J., GREGORY, P. D., URNOV, F. D. & JAENISCH, R. 2011. Genetic engineering of human pluripotent cells using TALE nucleases. *Nat Biotechnol,* 29**,** 731-4.

HOFFDING, M. K. & HYTTEL, P. 2015. Ultrastructural visualization of the Mesenchymal-to-Epithelial Transition during reprogramming of human fibroblasts to induced pluripotent stem cells. *Stem Cell Res,* 14**,** 39-53.

HOISCHEN, A., VAN BON, B. W., RODRIGUEZ-SANTIAGO, B., GILISSEN, C., VISSERS, L. E., DE VRIES, P., JANSSEN, I., VAN LIER, B., HASTINGS, R., SMITHSON, S. F., NEWBURY-ECOB, R., KJAERGAARD, S., GOODSHIP, J., MCGOWAN, R., BARTHOLDI, D., RAUCH, A., PEIPPO, M., COBBEN, J. M., WIECZOREK, D., GILLESSEN-KAESBACH, G., VELTMAN, J. A., BRUNNER, H. G. & DE VRIES, B. B. 2011. De novo nonsense mutations in ASXL1 cause Bohring-Opitz syndrome. *Nat Genet,* 43**,** 729-31.

HONG, H., TAKAHASHI, K., ICHISAKA, T., AOI, T., KANAGAWA, O., NAKAGAWA, M., OKITA, K. & YAMANAKA, S. 2009. Suppression of induced pluripotent stem cell generation by the p53-p21 pathway. *Nature,* 460**,** 1132-5.

HONG, M. & HE, G. 2017. The 2016 Revision to the World Health Organization Classification of Myelodysplastic Syndromes. *J Transl Int Med,* 5**,** 139-143.

HONG, S. H., WERBOWETSKI-OGILVIE, T., RAMOS-MEJIA, V., LEE, J. B. & BHATIA, M. 2010. Multiparameter comparisons of embryoid body differentiation toward human stem cell applications. *Stem Cell Res,* 5**,** 120-30.

HOSOKAWA, Y., TOYODA, T., FUKUI, K., BADEN, M. Y., FUNATO, M., KONDO, Y., SUDO, T., IWAHASHI, H., KISHIDA, M., OKADA, C., WATANABE, A., ASAKA, I., OSAFUNE, K., IMAGAWA, A. & SHIMOMURA, I. 2017. Insulin-producing cells derived from 'induced pluripotent stem cells' of patients with fulminant type 1 diabetes: Vulnerability to cytokine insults and increased expression of apoptosis-related genes. *J Diabetes Investig*.

HOUSELEY, J. & TOLLERVEY, D. 2009. The many pathways of RNA degradation. *Cell,* 136**,** 763-76.

HSU, J., REILLY, A., HAYES, B. J., CLOUGH, C. A., KONNICK, E. Q., TOROK-STORB, B., GULSUNER, S., WU, D., BECKER, P. S., KEEL, S. B., ABKOWITZ, J. L. & DOULATOV, S. 2019. Reprogramming identifies functionally distinct stages of clonal evolution in myelodysplastic syndromes. *Blood,* 134**,** 186-198.

HUANG, Y., XIONG, D., XU, Y., YU, J., XIAN, Y. & CAI, E. 2017. [Effect of the traditional Chinese medicine compound Yisui Lixue decoction on apoptosis of marrow cells in rats with myelodysplastic syndrome induced by dimethyl benzanthracene]. *Zhong Nan Da Xue Xue Bao Yi Xue Ban,* 42**,** 26-34.

HUANGFU, D., OSAFUNE, K., MAEHR, R., GUO, W., EIJKELENBOOM, A., CHEN, S., MUHLESTEIN, W. & MELTON, D. A. 2008. Induction of pluripotent stem cells from primary human fibroblasts with only Oct4 and Sox2. *Nat Biotechnol,* 26**,** 1269-75.

HUBSCHER, D., REBS, S., HAUPT, L., BORCHERT, T., GUESSOUM, C. I., TREU, F., KOHNE, S., MAUS, A., HAMBRECHT, M., SOSSALLA, S., DRESSEL, R., UY, A., JAKOB, M., HASENFUSS, G. & STRECKFUSS-BOMEKE, K. 2019. A High-Throughput Method as a Diagnostic Tool for HIV Detection in Patient-Specific Induced Pluripotent Stem Cells Generated by Different Reprogramming Methods. *Stem Cells Int,* 2019**,** 2181437.

HUO, H. Q., QU, Z. Y., YUAN, F., MA, L., YAO, L., XU, M., HU, Y., JI, J., BHATTACHARYYA, A., ZHANG, S. C. & LIU, Y. 2018. Modeling Down Syndrome with Patient iPSCs Reveals Cellular and Migration Deficits of GABAergic Neurons. *Stem Cell Reports,* 10**,** 1251-1266.

HUSSAINI, M. O., MIRZA, A. S., KOMROKJI, R., LANCET, J., PADRON, E. & SONG, J. 2018. Genetic Landscape of Acute Myeloid Leukemia Interrogated by Next-generation Sequencing: A Large Cancer Center Experience. *Cancer Genomics Proteomics,* 15**,** 121-126.

IBRAHEIM, R., SONG, C. Q., MIR, A., AMRANI, N., XUE, W. & SONTHEIMER, E. J. 2018. All-in-one adeno-associated virus delivery and genome editing by Neisseria meningitidis Cas9 in vivo. *Genome Biol,* 19**,** 137.

IMAGAWA, S., YAMAMOTO, M. & MIURA, Y. 1997. Negative regulation of the erythropoietin gene expression by the GATA transcription factors. *Blood,* 89**,** 1430-9.

IMAI, Y., KUROKAWA, M., IZUTSU, K., HANGAISHI, A., TAKEUCHI, K., MAKI, K., OGAWA, S., CHIBA, S., MITANI, K. & HIRAI, H. 2000. Mutations of the AML1 gene in myelodysplastic syndrome and their functional implications in leukemogenesis. *Blood,* 96**,** 3154-60.

INOUE, D., KITAURA, J., MATSUI, H., HOU, H. A., CHOU, W. C., NAGAMACHI, A., KAWABATA, K. C., TOGAMI, K., NAGASE, R., HORIKAWA, S., SAIKA, M., MICOL, J. B., HAYASHI, Y., HARADA, Y., HARADA, H., INABA, T., TIEN, H. F., ABDEL-WAHAB, O. & KITAMURA, T. 2015. SETBP1 mutations drive leukemic transformation in ASXL1-mutated MDS. *Leukemia,* 29**,** 847-57.

ISHINO, Y., SHINAGAWA, H., MAKINO, K., AMEMURA, M. & NAKATA, A. 1987. Nucleotide sequence of the iap gene, responsible for alkaline phosphatase isozyme conversion in Escherichia coli, and identification of the gene product. *J Bacteriol,* 169**,** 5429-33.

IVANOVS, A., RYBTSOV, S., NG, E. S., STANLEY, E. G., ELEFANTY, A. G. & MEDVINSKY, A. 2017. Human haematopoietic stem cell development: from the embryo to the dish. *Development,* 144**,** 2323-2337.

IWASAKI, H., MIZUNO, S., ARINOBU, Y., OZAWA, H., MORI, Y., SHIGEMATSU, H., TAKATSU, K., TENEN, D. G. & AKASHI, K. 2006. The order of expression of transcription factors directs hierarchical specification of hematopoietic lineages. *Genes Dev,* 20**,** 3010-21.

JABBOUR, E., GARCIA-MANERO, G., BATTY, N., SHAN, J., O'BRIEN, S., CORTES, J., RAVANDI, F., ISSA, J. P. & KANTARJIAN, H. 2010. Outcome of patients with myelodysplastic syndrome after failure of decitabine therapy. *Cancer,* 116**,** 3830-4.

JAFFE, E. S., HARRIS, N. L., STEIN, H. & VARDIMAN, J. W. 2001. World Health Organization Classification of Tumours: Pathology and Genetics of Tumours of Haematopoietic and Lymphoid Tissues. *ARC press Lyon*

JAISWAL, S., FONTANILLAS, P., FLANNICK, J., MANNING, A., GRAUMAN, P. V., MAR, B. G., LINDSLEY, R. C., MERMEL, C. H., BURTT, N., CHAVEZ, A., HIGGINS, J. M., MOLTCHANOV, V., KUO, F. C., KLUK, M. J., HENDERSON, B., KINNUNEN, L., KOISTINEN, H. A., LADENVALL, C., GETZ, G., CORREA, A., BANAHAN, B. F., GABRIEL, S., KATHIRESAN, S., STRINGHAM, H. M., MCCARTHY, M. I., BOEHNKE, M., TUOMILEHTO, J., HAIMAN, C., GROOP, L., ATZMON, G., WILSON, J. G., NEUBERG, D., ALTSHULER, D. & EBERT, B. L. 2014. Age-related clonal hematopoiesis associated with adverse outcomes. *N Engl J Med,* 371**,** 2488-98.

JAWHAR, M., SCHWAAB, J., SCHNITTGER, S., MEGGENDORFER, M., PFIRRMANN, M., SOTLAR, K., HORNY, H. P., METZGEROTH, G., KLUGER, S., NAUMANN, N., HAFERLACH, C., HAFERLACH, T., VALENT, P., HOFMANN, W. K., FABARIUS, A., CROSS, N. C. & REITER, A. 2016. Additional mutations in SRSF2, ASXL1 and/or RUNX1 identify a high-risk group of patients with KIT D816V(+) advanced systemic mastocytosis. *Leukemia,* 30**,** 136-43.

JEROMIN, S., HAFERLACH, T., WEISSMANN, S., MEGGENDORFER, M., EDER, C., NADARAJAH, N., ALPERMANN, T., KOHLMANN, A., KERN, W., HAFERLACH, C. & SCHNITTGER, S. 2015. Refractory anemia with ring sideroblasts and marked thrombocytosis cases harbor mutations in SF3B1 or other spliceosome genes accompanied by JAK2V617F and ASXL1 mutations. *Haematologica,* 100**,** e125-7.

JINEK, M., CHYLINSKI, K., FONFARA, I., HAUER, M., DOUDNA, J. A. & CHARPENTIER, E. 2012. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science,* 337**,** 816-21.

JOHNSON, S. M., RICHARDSON, D. R., GALEOTTI, J., ESPARZA, S., ZHU, A., FEDORIW, Y., WECK, K. E., FOSTER, M. C., COOMBS, C. C., ZEIDNER, J. F. & MONTGOMERY, N. D. 2019. Acute Myeloid Leukemia with Co-mutated ASXL1 and SRSF2 Exhibits Monocytic Differentiation and has a Mutational Profile Overlapping with Chronic Myelomonocytic Leukemia. *Hemasphere,* 3**,** e292.

KANAI-AZUMA, M., KANAI, Y., GAD, J. M., TAJIMA, Y., TAYA, C., KUROHMARU, M., SANAI, Y., YONEKAWA, H., YAZAKI, K., TAM, P. P. & HAYASHI, Y. 2002. Depletion of definitive gut endoderm in Sox17-null mutant mice. *Development,* 129**,** 2367-79.

KANG, J. M., YEON, B. K., CHO, S. J. & SUH, Y. H. 2016. Stem Cell Therapy for Alzheimer's Disease: A Review of Recent Clinical Trials. *J Alzheimers Dis,* 54**,** 879-889.

KANG, X., YU, Q., HUANG, Y., SONG, B., CHEN, Y., GAO, X., HE, W., SUN, X. & FAN, Y. 2015. Effects of Integrating and Non-Integrating Reprogramming Methods on Copy Number Variation and Genomic Stability of Human Induced Pluripotent Stem Cells. *PLoS One,* 10**,** e0131128.

KATO, N., KITAURA, J., DOKI, N., KOMENO, Y., WATANABE-OKOCHI, N., TOGAMI, K., NAKAHARA, F., OKI, T., ENOMOTO, Y., FUKUCHI, Y., NAKAJIMA, H., HARADA, Y., HARADA, H. & KITAMURA, T. 2011. Two types of C/EBPalpha mutations play distinct but collaborative roles in leukemogenesis: lessons from clinical data and BMT models. *Blood,* 117**,** 221-33.

KATSUMOTO, T. R., DUDA, J., KIM, A., WARDAK, Z., DRANOFF, G., CLAPP, D. W. & SHANNON, K. 2005. Granulocyte/macrophage colony-stimulating factor and accessory cells modulate radioprotection by purified hematopoietic cells. *J Exp Med,* 201**,** 853-8.

KAWAMURA, T., SUZUKI, J., WANG, Y. V., MENENDEZ, S., MORERA, L. B., RAYA, A., WAHL, G. M. & IZPISUA BELMONTE, J. C. 2009. Linking the p53 tumour suppressor pathway to somatic cell reprogramming. *Nature,* 460**,** 1140-4.

KIM, D., KIM, S., KIM, S., PARK, J. & KIM, J. S. 2016. Genome-wide target specificities of CRISPR-Cas9 nucleases revealed by multiplex Digenome-seq. *Genome Res,* 26**,** 406-15.

KIM, E., KOO, T., PARK, S. W., KIM, D., KIM, K., CHO, H. Y., SONG, D. W., LEE, K. J., JUNG, M. H., KIM, S., KIM, J. H., KIM, J. H. & KIM, J. S. 2017. In vivo genome editing with a small Cas9 orthologue derived from Campylobacter jejuni. *Nat Commun,* 8**,** 14500.

KIM, E. E. & WYCKOFF, H. W. 1991. Reaction mechanism of alkaline phosphatase based on crystal structures. Two-metal ion catalysis. *J Mol Biol,* 218**,** 449-64.

KIM, J., CHU, J., SHEN, X., WANG, J. & ORKIN, S. H. 2008a. An extended transcriptional network for pluripotency of embryonic stem cells. *Cell,* 132**,** 1049-61.

KIM, J. B., GREBER, B., ARAUZO-BRAVO, M. J., MEYER, J., PARK, K. I., ZAEHRES, H. & SCHOLER, H. R. 2009. Direct reprogramming of human neural stem cells by OCT4. *Nature,* 461**,** 649-3.

KIM, J. B., ZAEHRES, H., WU, G., GENTILE, L., KO, K., SEBASTIANO, V., ARAUZO-BRAVO, M. J., RUAU, D., HAN, D. W., ZENKE, M. & SCHOLER, H. R. 2008b. Pluripotent stem cells induced from adult neural stem cells by reprogramming with two factors. *Nature,* 454**,** 646-50.

KIMBREL, E. A. & LANZA, R. 2015. Current status of pluripotent stem cells: moving the first therapies to the clinic. *Nat Rev Drug Discov,* 14**,** 681-92.

KINGSLEY, P. D., MALIK, J., EMERSON, R. L., BUSHNELL, T. P., MCGRATH, K. E., BLOEDORN, L. A., BULGER, M. & PALIS, J. 2006. "Maturational" globin switching in primary primitive erythroid cells. *Blood,* 107**,** 1665-72.

KLEINSTIVER, B. P., PATTANAYAK, V., PREW, M. S., TSAI, S. Q., NGUYEN, N. T., ZHENG, Z. & JOUNG, J. K. 2016. High-fidelity CRISPR-Cas9 nucleases with no detectable genome-wide off-target effects. *Nature,* 529**,** 490-5.

KOCAK, D. D., JOSEPHS, E. A., BHANDARKAR, V., ADKAR, S. S., KWON, J. B. & GERSBACH, C. A. 2019. Increasing the specificity of CRISPR systems with engineered RNA secondary structures. *Nat Biotechnol,* 37**,** 657-666.

KON, A., SHIH, L. Y., MINAMINO, M., SANADA, M., SHIRAISHI, Y., NAGATA, Y., YOSHIDA, K., OKUNO, Y., BANDO, M., NAKATO, R., ISHIKAWA, S., SATO-OTSUBO, A., NAGAE, G., NISHIMOTO, A., HAFERLACH, C., NOWAK, D., SATO, Y., ALPERMANN, T., NAGASAKI, M., SHIMAMURA, T., TANAKA, H., CHIBA, K., YAMAMOTO, R., YAMAGUCHI, T., OTSU, M., OBARA, N., SAKATA-YANAGIMOTO, M., NAKAMAKI, T., ISHIYAMA, K., NOLTE, F., HOFMANN, W. K., MIYAWAKI, S., CHIBA, S., MORI, H., NAKAUCHI, H.,

KOEFFLER, H. P., ABURATANI, H., HAFERLACH, T., SHIRAHIGE, K., MIYANO, S. & OGAWA, S. 2013. Recurrent mutations in multiple components of the cohesin complex in myeloid neoplasms. *Nat Genet,* 45**,** 1232-7.

KOPP, J. L., ORMSBEE, B. D., DESLER, M. & RIZZINO, A. 2008. Small increases in the level of Sox2 trigger the differentiation of mouse embryonic stem cells. *Stem Cells,* 26**,** 903-11.

KOSCHMIEDER, S., HALMOS, B., LEVANTINI, E. & TENEN, D. G. 2009. Dysregulation of the C/EBPalpha differentiation pathway in human cancer. *J Clin Oncol,* 27**,** 619-28.

KOTINI, A. G., CHANG, C. J., BOUSSAAD, I., DELROW, J. J., DOLEZAL, E. K., NAGULAPALLY, A. B., PERNA, F., FISHBEIN, G. A., KLIMEK, V. M., HAWKINS, R. D., HUANGFU, D., MURRY, C. E., GRAUBERT, T., NIMER, S. D. & PAPAPETROU, E. P. 2015. Functional analysis of a chromosomal deletion associated with myelodysplastic syndromes using isogenic human induced pluripotent stem cells. *Nat Biotechnol,* 33**,** 646-55.

KOTINI, A. G., CHANG, C. J., CHOW, A., YUAN, H., HO, T. C., WANG, T., VORA, S., SOLOVYOV, A., HUSSER, C., OLSZEWSKA, M., TERUYA-FELDSTEIN, J., PERUMAL, D., KLIMEK, V. M., SPYRIDONIDIS, A., RAMPAL, R. K., SILVERMAN, L., REDDY, E. P., PAPAEMMANUIL, E., PAREKH, S., GREENBAUM, B. D., LESLIE, C. S., KHARAS, M. G. & PAPAPETROU, E. P. 2017. Stage-Specific Human Induced Pluripotent Stem Cells Map the Progression of Myeloid Transformation to Transplantable Leukemia. *Cell Stem Cell,* 20**,** 315-328 e7.

KOURY, M. J. & PONKA, P. 2004. New insights into erythropoiesis: the roles of folate, vitamin B12, and iron. *Annu Rev Nutr,* 24**,** 105-31.

KOUZARIDES, T. 2007. Chromatin modifications and their function. *Cell,* 128**,** 693-705.

KRANTZ, S. B. 1991. Erythropoietin. *Blood,* 77**,** 419-34.

KREBSBACH, P. H., KUZNETSOV, S. A., BIANCO, P. & ROBEY, P. G. 1999. Bone marrow stromal cells: characterization and clinical application. *Crit Rev Oral Biol Med,* 10**,** 165-81.

KUMAR, A., D'SOUZA, S. S. & THAKUR, A. S. 2019. Understanding the Journey of Human Hematopoietic Stem Cell Development. *Stem Cells Int,* 2019**,** 2141475.

LAKE, B. B., FINK, J., KLEMETSAUNE, L., FU, X., JEFFERS, J. R., ZAMBETTI, G. P. & XU, Y. 2012. Context-dependent enhancement of induced pluripotent stem cell reprogramming by silencing Puma. *Stem Cells,* 30**,** 888-97.

LAM, A. Q., FREEDMAN, B. S., MORIZANE, R., LEROU, P. H., VALERIUS, M. T. & BONVENTRE, J. V. 2014. Rapid and efficient differentiation of human pluripotent stem cells into intermediate mesoderm that forms tubules expressing kidney proximal tubular markers. *J Am Soc Nephrol,* 25**,** 1211-25.

LARIZZA, L., MAGNANI, I. & BEGHINI, A. 2005. The Kasumi-1 cell line: a t(8;21)-kit mutant model for acute myeloid leukemia. *Leuk Lymphoma,* 46**,** 247-55.

LAURIE, C. C., LAURIE, C. A., RICE, K., DOHENY, K. F., ZELNICK, L. R., MCHUGH, C. P., LING, H., HETRICK, K. N., PUGH, E. W., AMOS, C., WEI, Q., WANG, L. E., LEE, J. E., BARNES, K. C., HANSEL, N. N., MATHIAS, R., DALEY, D., BEATY, T. H., SCOTT, A. F., RUCZINSKI, I., SCHARPF, R. B., BIERUT, L. J., HARTZ, S. M., LANDI, M. T., FREEDMAN, N. D., GOLDIN, L. R., GINSBURG, D., LI, J., DESCH, K. C., STROM, S. S., BLOT, W. J., SIGNORELLO, L. B., INGLES, S. A., CHANOCK, S. J., BERNDT, S. I., LE MARCHAND, L., HENDERSON, B. E., MONROE, K. R., HEIT, J. A., DE ANDRADE, M., ARMASU, S. M., REGNIER, C., LOWE, W. L., HAYES, M. G., MARAZITA, M. L., FEINGOLD, E., MURRAY, J. C., MELBYE, M., FEENSTRA, B., KANG, J. H., WIGGS, J. L., JARVIK, G. P., MCDAVID, A. N., SESHAN, V. E., MIREL, D. B., CRENSHAW, A., SHAROPOVA, N., WISE, A., SHEN, J., CROSSLIN, D. R., LEVINE, D. M., ZHENG, X., UDREN, J. I., BENNETT, S., NELSON, S. C., GOGARTEN, S. M., CONOMOS, M. P., HEAGERTY, P., MANOLIO, T., PASQUALE, L. R., HAIMAN, C. A., CAPORASO, N. & WEIR, B. S. 2012. Detectable clonal mosaicism from birth to old age and its relationship to cancer. *Nat Genet,* 44**,** 642-50.

LE BIN, G. C., MUNOZ-DESCALZO, S., KUROWSKI, A., LEITCH, H., LOU, X., MANSFIELD, W., ETIENNE-DUMEAU, C., GRABOLE, N., MULAS, C., NIWA, H., HADJANTONAKIS, A. K. & NICHOLS, J. 2014. Oct4 is required for lineage priming in the developing inner cell mass of the mouse blastocyst. *Development,* 141**,** 1001-10.

LEE, J. H., SALCI, K. R., REID, J. C., ORLANDO, L., TANASIJEVIC, B., SHAPOVALOVA, Z. & BHATIA, M. 2017a. Brief Report: Human Acute Myeloid Leukemia Reprogramming to Pluripotency Is a Rare Event and Selects for Patient Hematopoietic Cells Devoid of Leukemic Mutations. *Stem Cells,* 35**,** 2095-2102.

LEE, K., MACKLEY, V. A., RAO, A., CHONG, A. T., DEWITT, M. A., CORN, J. E. & MURTHY, N. 2017b. Synthetically modified guide RNA and donor DNA are a versatile platform for CRISPR-Cas9 engineering. *Elife,* 6.

LEIMBERG, M. J., PRUS, E., KONIJN, A. M. & FIBACH, E. 2008. Macrophages function as a ferritin iron source for cultured human erythroid precursors. *J Cell Biochem,* 103**,** 1211-8.

LEROY, H., ROUMIER, C., HUYGHE, P., BIGGIO, V., FENAUX, P. & PREUDHOMME, C. 2005. CEBPA point mutations in hematological malignancies. *Leukemia,* 19**,** 329-34.

LEY, T. J., DING, L., WALTER, M. J., MCLELLAN, M. D., LAMPRECHT, T., LARSON, D. E., KANDOTH, C., PAYTON, J. E., BATY, J., WELCH, J., HARRIS, C. C., LICHTI, C. F., TOWNSEND, R. R., FULTON, R. S., DOOLING, D. J., KOBOLDT, D. C., SCHMIDT, H., ZHANG, Q., OSBORNE, J. R., LIN, L., O'LAUGHLIN, M., MCMICHAEL, J. F., DELEHAUNTY, K. D., MCGRATH, S. D., FULTON, L. A., MAGRINI, V. J., VICKERY, T. L., HUNDAL, J., COOK, L. L., CONYERS, J. J., SWIFT, G. W., REED, J. P., ALLDREDGE, P. A., WYLIE, T., WALKER, J., KALICKI, J., WATSON, M. A., HEATH, S., SHANNON, W. D., VARGHESE, N., NAGARAJAN, R., WESTERVELT, P., TOMASSON, M. H., LINK, D. C., GRAUBERT, T. A., DIPERSIO, J. F., MARDIS, E. R. & WILSON, R. K. 2010. DNMT3A mutations in acute myeloid leukemia. *N Engl J Med,* 363**,** 2424-33.

LI, A., LEE, C. M., HURLEY, A. E., JARRETT, K. E., DE GIORGI, M., LU, W., BALDERRAMA, K. S., DOERFLER, A. M., DESHMUKH, H., RAY, A., BAO, G. & LAGOR, W. R. 2019. A Self-Deleting AAV-CRISPR System for In Vivo Genome Editing. *Mol Ther Methods Clin Dev,* 12**,** 111-122.

LI, B., LIU, J., QU, S., GALE, R. P., SONG, Z., XING, R., LIU, J., REN, Y., XU, Z., QIN, T., ZHANG, Y., FANG, L., ZHANG, H., PAN, L., HU, N., CAI, W., ZHANG, P., HUANG, G. & XIAO, Z. 2016. Colony-forming unit cell (CFU-C) assays at diagnosis: CFU-G/M cluster predicts overall survival in myelodysplastic syndrome patients independently of IPSS-R. *Oncotarget,* 7**,** 68023-68032.

LI, H., YANG, Y., HONG, W., HUANG, M., WU, M. & ZHAO, X. 2020. Applications of genome editing technology in the targeted therapy of human diseases: mechanisms, advances and prospects. *Signal Transduct Target Ther,* 5**,** 1.

LI, R. H., LIANG, J. L., NI, S., ZHOU, T., QING, X. B., LI, H. P., HE, W. Z., CHEN, J. K., LI, F., ZHUANG, Q. A., QIN, B. M., XU, J. Y., LI, W., YANG, J. Y., GAN, Y., QIN, D. J., FENG, S. P., SONG, H., YANG, D. S., ZHANG, B. L., ZENG, L. W., LAI, L. X., & ESTEBAN, M. A. P., D. Q. 2010. A Mesenchymal-to-Epithelial Transition Initiates and Is Required for the Nuclear Reprogramming of Mouse Fibroblasts. *Cell Stem Cell,* 22**,** 3663-3673.

LI, W., LI, M., YANG, X., ZHANG, W., CAO, L. & GAO, R. 2021. Summary of animal models of myelodysplastic syndrome. *Animal Model Exp Med,* 4**,** 71-76.

LIE, A. L. M., MARINOPOULOU, E., LILLY, A. J., CHALLINOR, M., PATEL, R., LANCRIN, C., KOUSKOFF, V. & LACAUD, G. 2018. Regulation of RUNX1 dosage is crucial for efficient blood formation from hemogenic endothelium. *Development,* 145.

LIN, C. H., JACKSON, A. L., GUO, J., LINSLEY, P. S. & EISENMAN, R. N. 2009. Myc-regulated microRNAs attenuate embryonic stem cell differentiation. *EMBO J,* 28**,** 3157-70.

LIN, C. Y., LOVEN, J., RAHL, P. B., PARANAL, R. M., BURGE, C. B., BRADNER, J. E., LEE, T. I. & YOUNG, R. A. 2012. Transcriptional amplification in tumor cells with elevated c-Myc. *Cell,* 151**,** 56-67.

LIN, F. T., MACDOUGALD, O. A., DIEHL, A. M. & LANE, M. D. 1993. A 30-kDa alternative translation product of the CCAAT/enhancer binding protein alpha message: transcriptional activator lacking antimitotic activity. *Proc Natl Acad Sci U S A,* 90**,** 9606-10.

LIN, Y., ZHENG, Y., WANG, Z. C. & WANG, S. Y. 2016. Prognostic significance of ASXL1 mutations in myelodysplastic syndromes and chronic myelomonocytic leukemia: A meta-analysis. *Hematology,* 21**,** 454-61.

LIU, P., CHEN, M., LIU, Y., QI, L. S. & DING, S. 2018. CRISPR-Based Chromatin Remodeling of the Endogenous Oct4 or Sox2 Locus Enables Reprogramming to Pluripotency. *Cell Stem Cell,* 22**,** 252-261 e4.

LIU, W., WU, M., HUANG, Z., LIAN, J., CHEN, J., WANG, T., LEUNG, A. Y., LIAO, Y., ZHANG, Z., LIU, Q., YEN, K., LIN, S., ZON, L. I., WEN, Z., ZHANG, Y. & ZHANG, W. 2017. c-myb hyperactivity leads to myeloid and lymphoid malignancies in zebrafish. *Leukemia,* 31**,** 222-233.

LIVAK, K. J. & SCHMITTGEN, T. D. 2011. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods* 25**,** 402-408.

LOH, P. R., GENOVESE, G., HANDSAKER, R. E., FINUCANE, H. K., RESHEF, Y. A., PALAMARA, P. F., BIRMANN, B. M., TALKOWSKI, M. E., BAKHOUM, S. F., MCCARROLL, S. A. & PRICE, A. L. 2018. Insights into clonal haematopoiesis from 8,342 mosaic chromosomal alterations. *Nature,* 559**,** 350-355.

LOIS, C., HONG, E. J., PEASE, S., BROWN, E. J. & BALTIMORE, D. 2002. Germline transmission and tissue-specific expression of transgenes delivered by lentiviral vectors. *Science,* 295**,** 868-72.

LU, S. J., FENG, Q., PARK, J. S., VIDA, L., LEE, B. S., STRAUSBAUCH, M., WETTSTEIN, P. J., HONIG, G. R. & LANZA, R. 2008. Biologic properties and enucleation of red blood cells from human embryonic stem cells. *Blood,* 112**,** 4475-84.

MA, O., HONG, S., GUO, H., GHIAUR, G. & FRIEDMAN, A. D. 2014. Granulopoiesis requires increased C/EBPalpha compared to monopoiesis, correlated with elevated Cebpa in immature G-CSF receptor versus M-CSF receptor expressing cells. *PLoS One,* 9**,** e95784.

MAKISHIMA, H., YOSHIZATO, T., YOSHIDA, K., SEKERES, M. A., RADIVOYEVITCH, T., SUZUKI, H., PRZYCHODZEN, B., NAGATA, Y., MEGGENDORFER, M., SANADA, M., OKUNO, Y., HIRSCH, C., KUZMANOVIC, T., SATO, Y., SATO-OTSUBO, A., LAFRAMBOISE, T., HOSONO, N., SHIRAISHI, Y., CHIBA, K., HAFERLACH, C., KERN, W., TANAKA, H., SHIOZAWA, Y., GOMEZ-SEGUI, I., HUSSEINZADEH, H. D., THOTA, S., GUINTA, K. M., DIENES, B., NAKAMAKI, T., MIYAWAKI, S., SAUNTHARARAJAH, Y., CHIBA, S., MIYANO, S., SHIH, L. Y., HAFERLACH, T., OGAWA, S. & MACIEJEWSKI, J. P. 2017. Dynamics of clonal evolution in myelodysplastic syndromes. *Nat Genet,* 49**,** 204-212.

MALCOVATI, L., HELLSTROM-LINDBERG, E., BOWEN, D., ADES, L., CERMAK, J., DEL CANIZO, C., DELLA PORTA, M. G., FENAUX, P., GATTERMANN, N., GERMING, U., JANSEN, J. H., MITTELMAN, M., MUFTI, G., PLATZBECKER, U., SANZ, G. F., SELLESLAG, D., SKOV-HOLM, M., STAUDER, R., SYMEONIDIS, A., VAN DE LOOSDRECHT, A. A., DE WITTE, T., CAZZOLA, M. & EUROPEAN LEUKEMIA, N. 2013. Diagnosis and treatment of primary myelodysplastic syndromes in adults: recommendations from the European LeukemiaNet. *Blood,* 122**,** 2943-64.

MANDAI, M., KURIMOTO, Y. & TAKAHASHI, M. 2017. Autologous Induced Stem-Cell-Derived Retinal Cells for Macular Degeneration. *N Engl J Med,* 377**,** 792-793.

MARCUCCI, G., MAHARRY, K., RADMACHER, M. D., MROZEK, K., VUKOSAVLJEVIC, T., PASCHKA, P., WHITMAN, S. P., LANGER, C., BALDUS, C. D., LIU, C. G., RUPPERT, A. S., POWELL, B. L., CARROLL, A. J., CALIGIURI, M. A., KOLITZ, J. E., LARSON, R. A. & BLOOMFIELD, C. D. 2008. Prognostic significance of, and gene and microRNA expression signatures associated with, CEBPA mutations in cytogenetically normal acute myeloid leukemia with high-risk molecular features: a Cancer and Leukemia Group B Study. *J Clin Oncol,* 26**,** 5078-87.

MARDIS, E. R., DING, L., DOOLING, D. J., LARSON, D. E., MCLELLAN, M. D., CHEN, K., KOBOLDT, D. C., FULTON, R. S., DELEHAUNTY, K. D., MCGRATH, S. D., FULTON, L. A., LOCKE, D. P., MAGRINI, V. J., ABBOTT, R. M., VICKERY, T. L., REED, J. S., ROBINSON, J. S., WYLIE, T., SMITH, S. M., CARMICHAEL, L., ELDRED, J. M., HARRIS, C. C., WALKER, J., PECK, J. B., DU, F., DUKES, A. F., SANDERSON, G. E., BRUMMETT, A. M., CLARK, E., MCMICHAEL, J. F., MEYER, R. J., SCHINDLER, J. K., POHL, C. S., WALLIS, J. W., SHI, X., LIN, L., SCHMIDT, H., TANG, Y., HAIPEK, C., WIECHERT, M. E., IVY, J. V., KALICKI, J., ELLIOTT, G., RIES, R. E., PAYTON, J. E., WESTERVELT, P., TOMASSON, M. H., WATSON, M. A., BATY, J., HEATH, S., SHANNON, W. D., NAGARAJAN, R., LINK, D. C., WALTER, M. J., GRAUBERT, T. A., DIPERSIO, J. F., WILSON, R. K. & LEY, T. J. 2009. Recurring mutations found by sequencing an acute myeloid leukemia genome. *N Engl J Med,* 361**,** 1058-66.

MARION, R. M., STRATI, K., LI, H., MURGA, M., BLANCO, R., ORTEGA, S., FERNANDEZ-CAPETILLO, O., SERRANO, M. & BLASCO, M. A. 2009. A p53-mediated DNA damage response limits reprogramming to ensure iPS cell genomic integrity. *Nature,* 460**,** 1149-53.

MARUYAMA, T., DOUGAN, S. K., TRUTTMANN, M. C., BILATE, A. M., INGRAM, J. R. & PLOEGH, H. L. 2015. Increasing the efficiency of precise genome editing with CRISPR-Cas9 by inhibition of nonhomologous end joining. *Nat Biotechnol,* 33**,** 538-42.

MATSUI, S., SAIT, S., JONES, C. A., NOWAK, N. & GROSS, K. W. 2002. Rapid localization of transgenes in mouse chromosomes with a combined Spectral Karyotyping/FISH technique. *Mamm Genome,* 13**,** 680-5.

MCCONNELL, B. B., GHALEB, A. M., NANDAN, M. O. & YANG, V. W. 2007. The diverse functions of Kruppel-like factors 4 and 5 in epithelial biology and pathobiology. *Bioessays,* 29**,** 549-57.

MCNEISH, J., GARDNER, J. P., WAINGER, B. J., WOOLF, C. J. & EGGAN, K. 2015. From Dish to Bedside: Lessons Learned While Translating Findings from a Stem Cell Model of Disease to a Clinical Trial. *Cell Stem Cell,* 17**,** 8-10.

MCREYNOLDS, L. J., YANG, Y., YUEN WONG, H., TANG, J., ZHANG, Y., MULE, M. P., DAUB, J., PALMER, C., FORURAGHI, L., LIU, Q., ZHU, J., WANG, W., WEST, R. R., YOHE, M. E., HSU, A. P., HICKSTEIN, D. D., TOWNSLEY, D. M., HOLLAND, S. M., CALVO, K. R. & HOURIGAN, C. S. 2019. MDS-associated mutations in germline GATA2 mutated patients with hematologic manifestations. *Leuk Res,* 76**,** 70-75.

MEDVEDEV, S. P., SHEVCHENKO, A. I. & ZAKIAN, S. M. 2010. Induced Pluripotent Stem Cells: Problems and Advantages when Applying them in Regenerative Medicine. *Acta Naturae,* 2**,** 18-28.

MELGUIZO-SANCHIS, D., XU, Y., TAHEEM, D., YU, M., TILGNER, K., BARTA, T., GASSNER, K., ANYFANTIS, G., WAN, T., ELANGO, R., ALHARTHI, S., EL-HAROUNI, A. A., PRZYBORSKI, S., ADAM, S., SARETZKI, G., SAMARASINGHE, S., ARMSTRONG, L. & LAKO, M. 2018. iPSC modeling of severe aplastic anemia reveals impaired differentiation and telomere shortening in blood progenitors. *Cell Death Dis,* 9**,** 128.

MENSSEN, A. J. & WALTER, M. J. 2020. Genetics of Progression from MDS to Secondary Leukemia. *Blood*.

METCALF, D. 2008. Hematopoietic cytokines. *Blood,* 111**,** 485-91.

METZELER, K. H., BECKER, H., MAHARRY, K., RADMACHER, M. D., KOHLSCHMIDT, J., MROZEK, K., NICOLET, D., WHITMAN, S. P., WU, Y. Z., SCHWIND, S., POWELL, B. L., CARTER, T. H., WETZLER, M., MOORE, J. O., KOLITZ, J. E., BAER, M. R., CARROLL, A. J., LARSON, R. A., CALIGIURI, M. A., MARCUCCI, G. & BLOOMFIELD, C. D. 2011. ASXL1 mutations identify a high-risk subgroup of older patients with primary cytogenetically normal AML within the ELN Favorable genetic category. *Blood,* 118**,** 6920-9.

MICHALOPOULOU, S., MICHEVA, I., KOURAKLIS-SYMEONIDIS, A., KAKAGIANNI, T., SYMEONIDIS, A. & ZOUMBOS, N. C. 2004. Impaired clonogenic growth of myelodysplastic bone marrow progenitors in vitro is irrelevant to their apoptotic state. *Leuk Res,* 28**,** 805-12.

MIGLIACCIO, G., SANCHEZ, M., MASIELLO, F., TIRELLI, V., VARRICCHIO, L., WHITSETT, C. & MIGLIACCIO, A. R. 2010. Humanized culture medium for clinical expansion of human erythroblasts. *Cell Transplant,* 19**,** 453-69.

MIKKELSEN, T. S., HANNA, J., ZHANG, X., KU, M., WERNIG, M., SCHORDERET, P., BERNSTEIN, B. E., JAENISCH, R., LANDER, E. S. & MEISSNER, A. 2008. Dissecting direct reprogramming through integrative genomic analysis. *Nature,* 454**,** 49-55.

MIYAUCHI, M., KOYA, J., ARAI, S., YAMAZAKI, S., HONDA, A., KATAOKA, K., YOSHIMI, A., TAOKA, K., KUMANO, K. & KUROKAWA, M. 2018. ADAM8 Is an Antigen of Tyrosine Kinase Inhibitor-Resistant Chronic Myeloid Leukemia Cells Identified by Patient-Derived Induced Pluripotent Stem Cells. *Stem Cell Reports,* 10**,** 1115-1130.

MOHAMEDALI, A. & MUFTI, G. J. 2009. Van-den Berghe's 5q- syndrome in 2008. *Br J Haematol,* 144**,** 157-68.

MOJICA, F. J., DIEZ-VILLASENOR, C., GARCIA-MARTINEZ, J. & SORIA, E. 2005. Intervening sequences of regularly spaced prokaryotic repeats derive from foreign genetic elements. *J Mol Evol,* 60**,** 174-82.

MONTAGUE, T. G., CRUZ, J. M., GAGNON, J. A., CHURCH, G. M. & VALEN, E. 2014. CHOPCHOP: a CRISPR/Cas9 and TALEN web tool for genome editing. *Nucleic Acids Res,* 42**,** W401-7.

MOON, S. B., KIM, D. Y., KO, J. H. & KIM, Y. S. 2019. Recent advances in the CRISPR genome editing tool set. *Exp Mol Med,* 51**,** 1-11.

MOORE, L. D., LE, T. & FAN, G. 2013. DNA methylation and its basic function. *Neuropsychopharmacology,* 38**,** 23-38.

MORAN-CRUSIO, K., REAVIE, L., SHIH, A., ABDEL-WAHAB, O., NDIAYE-LOBRY, D., LOBRY, C., FIGUEROA, M. E., VASANTHAKUMAR, A., PATEL, J., ZHAO, X., PERNA, F., PANDEY, S., MADZO, J., SONG, C., DAI, Q., HE, C., IBRAHIM, S., BERAN, M., ZAVADIL, J., NIMER, S. D., MELNICK, A., GODLEY, L. A., AIFANTIS, I. & LEVINE, R. L. 2011. Tet2 loss leads to increased hematopoietic stem cell self-renewal and myeloid transformation. *Cancer Cell,* 20**,** 11-24.

MORIGUCHI, T. & YAMAMOTO, M. 2014. A regulatory network governing Gata1 and Gata2 gene transcription orchestrates erythroid lineage differentiation. *Int J Hematol,* 100**,** 417-24.

MORRISON, S. J., WANDYCZ, A. M., HEMMATI, H. D., WRIGHT, D. E. & WEISSMAN, I. L. 1997. Identification of a lineage of multipotent hematopoietic progenitors. *Development,* 124**,** 1929-39.

MULAS, C., CHIA, G., JONES, K. A., HODGSON, A. C., STIRPARO, G. G. & NICHOLS, J. 2018. Oct4 regulates the embryonic axis and coordinates exit from pluripotency and germ layer specification in the mouse embryo. *Development,* 145.

MULLARD, A. 2015. Stem-cell discovery platforms yield first clinical candidates. *Nat Rev Drug Discov,* 14**,** 589-91.

MULLER, C., CALKHOVEN, C. F., SHA, X. & LEUTZ, A. 2004. The CCAAT enhancer-binding protein alpha (C/EBPalpha) requires a SWI/SNF complex for proliferation arrest. *J Biol Chem,* 279**,** 7353-8.

NAGASE, R., INOUE, D., PASTORE, A., FUJINO, T., HOU, H. A., YAMASAKI, N., GOYAMA, S., SAIKA, M., KANAI, A., SERA, Y., HORIKAWA, S., OTA, Y., ASADA, S., HAYASHI, Y., KAWABATA, K. C., TAKEDA, R., TIEN, H. F., HONDA, H., ABDEL-WAHAB, O. & KITAMURA, T. 2018. Expression of mutant Asxl1 perturbs hematopoiesis and promotes susceptibility to leukemic transformation. *J Exp Med,* 215**,** 1729-1747.

NAITO, Y., HINO, K., BONO, H. & UI-TEI, K. 2015. CRISPRdirect: software for designing CRISPR/Cas guide RNA with reduced off-target sites. *Bioinformatics,* 31**,** 1120-3.

NAKAGAWA, M., KOYANAGI, M., TANABE, K., TAKAHASHI, K., ICHISAKA, T., AOI, T., OKITA, K., MOCHIDUKI, Y., TAKIZAWA, N. & YAMANAKA, S. 2008. Generation of induced pluripotent stem cells without Myc from mouse and human fibroblasts. *Nat Biotechnol,* 26**,** 101-6.

NAKANO-OKUNO, M., BORAH, B. R. & NAKANO, I. 2014. Ethics of iPSC-based clinical research for age-related macular degeneration: patient-centered risk-benefit analysis. *Stem Cell Rev Rep,* 10**,** 743-52.

NANNYA, Y., SANADA, M., NAKAZAKI, K., HOSOYA, N., WANG, L., HANGAISHI, A., KUROKAWA, M., CHIBA, S., BAILEY, D. K., KENNEDY, G. C. & OGAWA, S. 2005. A robust algorithm for copy number detection using high-density oligonucleotide single nucleotide polymorphism genotyping arrays. *Cancer Res,* 65**,** 6071-9.

NOUTSOU, M., LI, J., LING, J., WANG, Y. & CHEN, Y. 2017. The cohesin complex is necessary for epidermal progenitor cell function through maintenance of self-renewal genes. *Cell Rep.,* 20**,** 3005-3013.

OGURO, H., DING, L. & MORRISON, S. J. 2013. SLAM family markers resolve functionally distinct subpopulations of hematopoietic stem cells and multipotent progenitors. *Cell Stem Cell,* 13**,** 102-16.

OKADA, S., NAKAUCHI, H., NAGAYOSHI, K., NISHIKAWA, S., MIURA, Y. & SUDA, T. 1992. In vivo and in vitro stem cell function of c-kit- and Sca-1-positive murine hematopoietic cells. *Blood,* 80**,** 3044-50.

OKAMOTO, K., OKAZAWA, H., OKUDA, A., SAKAI, M., MURAMATSU, M. & HAMADA, H. 1990. A novel octamer binding transcription factor is differentially expressed in mouse embryonic cells. *Cell,* 60**,** 461-72.

OKITA, K., HONG, H., TAKAHASHI, K. & YAMANAKA, S. 2010. Generation of mouse-induced pluripotent stem cells with plasmid vectors. *Nat Protoc,* 5**,** 418-28.

OKITA, K., ICHISAKA, T. & YAMANAKA, S. 2007. Generation of germline-competent induced pluripotent stem cells. *Nature,* 448**,** 313-7.

OKITA, K., MATSUMURA, Y., SATO, Y., OKADA, A., MORIZANE, A., OKAMOTO, S., HONG, H., NAKAGAWA, M., TANABE, K., TEZUKA, K., SHIBATA, T., KUNISADA, T., TAKAHASHI, M., TAKAHASHI, J., SAJI, H. &

YAMANAKA, S. 2011. A more efficient method to generate integration-free human iPS cells. *Nat Methods,* 8**,** 409-12.

OKUDA, T., VAN DEURSEN, J., HIEBERT, S. W., GROSVELD, G. & DOWNING, J. R. 1996. AML1, the target of multiple chromosomal translocations in human leukemia, is essential for normal fetal liver hematopoiesis. *Cell,* 84**,** 321-30.

OKUMURA-NAKANISHI, S., SAITO, M., NIWA, H. & ISHIKAWA, F. 2005. Oct-3/4 and Sox2 regulate Oct-3/4 gene in embryonic stem cells. *J Biol Chem,* 280**,** 5307-17.

OMOLE, A. E. & FAKOYA, A. O. J. 2018. Ten years of progress and promise of induced pluripotent stem cells: historical origins, characteristics, mechanisms, limitations, and potential applications. *PeerJ,* 6**,** e4370.

ORKIN, S. H. 2000. Diversification of haematopoietic stem cells to specific lineages. *Nat Rev Genet,* 1**,** 57-64.

OSAWA, M., HANADA, K., HAMADA, H. & NAKAUCHI, H. 1996. Long-term lymphohematopoietic reconstitution by a single CD34-low/negative hematopoietic stem cell. *Science,* 273**,** 242-5.

OU, Z., NIU, X., HE, W., CHEN, Y., SONG, B., XIAN, Y., FAN, D., TANG, D. & SUN, X. 2016. The Combination of CRISPR/Cas9 and iPSC Technologies in the Gene Therapy of Human beta-thalassemia in Mice. *Sci Rep,* 6**,** 32463.

OWEN, C. J., TOZE, C. L., KOOCHIN, A., FORREST, D. L., SMITH, C. A., STEVENS, J. M., JACKSON, S. C., POON, M. C., SINCLAIR, G. D., LEBER, B., JOHNSON, P. R., MACHETA, A., YIN, J. A., BARNETT, M. J., LISTER, T. A. & FITZGIBBON, J. 2008. Five new pedigrees with inherited RUNX1 mutations causing familial platelet disorder with propensity to myeloid malignancy. *Blood,* 112**,** 4639-45.

PALIS, J., ROBERTSON, S., KENNEDY, M., WALL, C. & KELLER, G. 1999. Development of erythroid and myeloid progenitors in the yolk sac and embryo proper of the mouse. *Development,* 126**,** 5073-84.

PAN, H. & SCHULTZ, R. M. 2011. Sox2 modulates reprogramming of gene expression in two-cell mouse embryos. *Biol Reprod,* 85**,** 409-16.

PANOSKALTSIS, N., MANTALARIS, A. & WU, J. H. 2005. Engineering a mimicry of bone marrow tissue ex vivo. *J Biosci Bioeng,* 100**,** 28-35.

PAPAEMMANUIL, E., GERSTUNG, M., MALCOVATI, L., TAURO, S., GUNDEM, G., VAN LOO, P., YOON, C. J., ELLIS, P., WEDGE, D. C., PELLAGATTI, A., SHLIEN, A., GROVES, M. J., FORBES, S. A., RAINE, K., HINTON, J., MUDIE, L. J., MCLAREN, S., HARDY, C., LATIMER, C., DELLA PORTA, M. G., O'MEARA, S., AMBAGLIO, I., GALLI, A., BUTLER, A. P., WALLDIN, G., TEAGUE, J. W., QUEK, L., STERNBERG, A., GAMBACORTI-PASSERINI, C., CROSS, N. C., GREEN, A. R., BOULTWOOD, J., VYAS, P., HELLSTROM-LINDBERG, E., BOWEN, D., CAZZOLA, M., STRATTON, M. R., CAMPBELL, P. J. & CHRONIC MYELOID DISORDERS WORKING GROUP OF THE INTERNATIONAL CANCER GENOME, C. 2013. Clinical and biological implications of driver mutations in myelodysplastic syndromes. *Blood,* 122**,** 3616-27; quiz 3699.

PAPAPETROU, E. P. 2019. Modeling myeloid malignancies with patient-derived iPSCs. *Exp Hematol,* 71**,** 77-84.

PAPETTI, M., WONTAKAL, S. N., STOPKA, T. & SKOULTCHI, A. I. 2010. GATA-1 directly regulates p21 gene expression during erythroid differentiation. *Cell Cycle,* 9**,** 1972-80.

PARK, E. T., GUM, J. R., KAKAR, S., KWON, S. W., DENG, G. & KIM, Y. S. 2008. Aberrant expression of SOX2 upregulates MUC5AC gastric foveolar mucin in mucinous cancers of the colorectum and related lesions. *Int J Cancer,* 122**,** 1253-60.

PATEL, B. J., PRZYCHODZEN, B., THOTA, S., RADIVOYEVITCH, T., VISCONTE, V., KUZMANOVIC, T., CLEMENTE, M., HIRSCH, C., MORAWSKI, A., SOUAID, R., SAYGIN, C., NAZHA, A., DEMAREST, B., LAFRAMBOISE, T., SAKAGUCHI, H., KOJIMA, S., CARRAWAY, H. E., OGAWA, S., MAKISHIMA, H., SEKERES, M. A. & MACIEJEWSKI, J. P. 2017. Genomic determinants of chronic myelomonocytic leukemia. *Leukemia,* 31**,** 2815-2823.

PEDERSEN, T. A., KOWENZ-LEUTZ, E., LEUTZ, A. & NERLOV, C. 2001. Cooperation between C/EBPalpha TBP/TFIIB and SWI/SNF recruiting domains is required for adipocyte differentiation. *Genes Dev,* 15**,** 3208-16.

PELLAGATTI, A. & BOULTWOOD, J. 2020. Splicing factor mutant myelodysplastic syndromes: Recent advances. *Adv Biol Regul,* 75**,** 100655.

PETERSON, J. F., AGGARWAL, N., SMITH, C. A., GOLLIN, S. M., SURTI, U., RAJKOVIC, A., SWERDLOW, S. H. & YATSENKO, S. A. 2015. Integration of microarray analysis into the clinical diagnosis of hematological malignancies: How much can we improve cytogenetic testing? *Oncotarget,* 6**,** 18845-62.

PLATZBECKER, U., SCHETELIG, J., FINKE, J., TRENSCHEL, R., SCOTT, B. L., KOBBE, G., SCHAEFER-ECKART, K., BORNHAUSER, M., ITZYKSON, R., GERMING, U., BEELEN, D., EHNINGER, G., FENAUX, P., DEEG, H. J., ADES, L., GERMAN, M. D. S. S., COOPERATIVE TRANSPLANT STUDY, G., FRED HUTCHINSON CANCER RESEARCH, C. & GROUPE FRANCOPHONE DES, M. 2012. Allogeneic hematopoietic cell transplantation in patients age 60-70 years with de novo high-risk myelodysplastic syndrome or secondary acute myelogenous leukemia: comparison with patients lacking donors who received azacitidine. *Biol Blood Marrow Transplant,* 18**,** 1415-21.

POLO, J. M., ANDERSSEN, E., WALSH, R. M., SCHWARZ, B. A., NEFZGER, C. M., LIM, S. M., BORKENT, M., APOSTOLOU, E., ALAEI, S., CLOUTIER, J., BAR-NUR, O., CHELOUFI, S., STADTFELD, M., FIGUEROA, M. E., ROBINTON, D., NATESAN, S., MELNICK, A., ZHU, J., RAMASWAMY, S. & HOCHEDLINGER, K. 2012. A molecular roadmap of reprogramming somatic cells into iPS cells. *Cell,* 151**,** 1617-32.

PORSE, B., BRYDER, D., THEILGAARD-MONCH, K., HASEMANN, M. S., ANDERSEN, K., DAMGAARD, I., JABBOUR, S. & NERLOV, C. 2005a. Loss of C/EBPα cell cycle control increases myeloid progenitor proliferation and transforms the neutrophil granulocyte lineag. *JEM,* 202 (1)**,** 85-96.

PORSE, B. T., BRYDER, D., THEILGAARD-MONCH, K., HASEMANN, M. S., ANDERSON, K., DAMGAARD, I., JACOBSEN, S. E. & NERLOV, C. 2005b. Loss of C/EBP alpha cell cycle control increases myeloid progenitor proliferation and transforms the neutrophil granulocyte lineage. *J Exp Med,* 202**,** 85-96.

PORSE, B. T., PEDERSEN, T. A., HASEMANN, M. S., SCHUSTER, M. B., KIRSTETTER, P., LUEDDE, T., DAMGAARD, I., KURZ, E., SCHJERLING, C. K. & NERLOV, C. 2006. The proline-histidine-rich CDK2/CDK4 interaction region of C/EBPalpha is dispensable for C/EBPalpha-mediated growth regulation in vivo. *Mol Cell Biol,* 26**,** 1028-37.

POURCEL, C., SALVIGNOL, G. & VERGNAUD, G. 2005. CRISPR elements in Yersinia pestis acquire new repeats by preferential uptake of bacteriophage DNA, and provide additional tools for evolutionary studies. *Microbiology (Reading),* 151**,** 653-663.

PTASINSKA, A., PICKIN, A., ASSI, S. A., CHIN, P. S., AMES, L., AVELLINO, R., GROSCHEL, S., DELWEL, R., COCKERILL, P. N., OSBORNE, C. S. & BONIFER, C. 2019. RUNX1-ETO Depletion in t(8;21) AML Leads to C/EBPalpha- and AP-1-Mediated Alterations in Enhancer-Promoter Interaction. *Cell Rep,* 29**,** 2120.

QANASH, H., LINASK, K., BEERS, J., ZOU, J. & LAROCHELLE, A. 2018. Generation of Fanconi Anemia iPSC Clones By Addition of a Small Molecule Inhibitor of p53 during Reprogramming. *Blood,* 132.

QU, X., ZHANG, S., WANG, S., WANG, Y., LI, W., HUANG, Y., ZHAO, H., WU, X., AN, C., GUO, X., HALE, J., LI, J., HILLYER, C. D., MOHANDAS, N., LIU, J., YAZDANBAKHSH, K., VINCHI, F., CHEN, L., KANG, Q. & AN, X. 2018. TET2 deficiency leads to stem cell factor-dependent clonal expansion of dysfunctional erythroid progenitors. *Blood,* 132**,** 2406-2417.

RAMJI, D. P. & FOKA, P. 2002. CCAAT/enhancer-binding proteins: structure, function and regulation. *Biochem J,* 365**,** 561-75.

RAN, F. A., HSU, P. D., LIN, C. Y., GOOTENBERG, J. S., KONERMANN, S., TREVINO, A. E., SCOTT, D. A., INOUE, A., MATOBA, S., ZHANG, Y. & ZHANG, F. 2013. Double nicking by RNA-guided CRISPR Cas9 for enhanced genome editing specificity. *Cell,* 154**,** 1380-9.

RAYA, A., RODRIGUEZ-PIZA, I., GUENECHEA, G., VASSENA, R., NAVARRO, S., BARRERO, M. J., CONSIGLIO, A., CASTELLA, M., RIO, P., SLEEP, E., GONZALEZ, F., TISCORNIA, G., GARRETA, E., AASEN, T., VEIGA, A., VERMA, I. M., SURRALLES, J., BUEREN, J. & IZPISUA BELMONTE, J. C. 2009. Disease-corrected haematopoietic progenitors from Fanconi anaemia induced pluripotent stem cells. *Nature,* 460**,** 53-9.

RAYNER, E., DURIN, M. A., THOMAS, R., MORALLI, D., O'CATHAIL, S. M., TOMLINSON, I., GREEN, C. M. & LEWIS, A. 2019. CRISPR-Cas9 Causes Chromosomal Instability and Rearrangements in Cancer Cell Lines, Detectable by Cytogenetic Methods. *CRISPR J,* 2**,** 406-416.

RHODES, J., HAGEN, A., HSU, K., DENG, M., LIU, T. X., LOOK, A. T. & KANKI, J. P. 2005. Interplay of pu.1 and gata1 determines myelo-erythroid progenitor cell fate in zebrafish. *Dev Cell,* 8**,** 97-108.

RHODES, M. M., KOPSOMBUT, P., BONDURANT, M. C., PRICE, J. O. & KOURY, M. J. 2008. Adherence to macrophages in erythroblastic islands enhances erythroblast proliferation and increases erythrocyte production by a different mechanism than erythropoietin. *Blood,* 111**,** 1700-8.

RICHARDSON, C. D., RAY, G. J., DEWITT, M. A., CURIE, G. L. & CORN, J. E. 2016. Enhancing homology-directed genome editing by catalytically active and inactive CRISPR-Cas9 using asymmetric donor DNA. *Nat Biotechnol,* 34**,** 339-44.

RICHARDSON, D. R., SWOBODA, D. M., MOORE, D. T., JOHNSON, S. M., CHAN, O., GALEOTTI, J., ESPARZA, S., HUSSAINI, M. O., VAN DEVENTER, H., FOSTER, M. C., COOMBS, C. C., MONTGOMERY, N. D., SALLMAN, D. A. & ZEIDNER, J. F. 2021. Genomic characteristics and prognostic significance of co-mutated ASXL1/SRSF2 acute myeloid leukemia. *Am J Hematol,* 96**,** 462-470.

RIEDER, C. L. & PALAZZO, R. E. 1992. Colcemid and the mitotic cycle. *J Cell Sci,* 102 (Pt 3)**,** 387-92.

ROBERT, F., BARBEAU, M., ETHIER, S., DOSTIE, J. & PELLETIER, J. 2015. Pharmacological inhibition of DNA-PK stimulates Cas9-mediated genome editing. *Genome Med,* 7**,** 93.

ROBERTSON, S. M., KENNEDY, M., SHANNON, J. M. & KELLER, G. 2000. A transitional stage in the commitment of mesoderm to hematopoiesis requiring the transcription factor SCL/tal-1. *Development,* 127**,** 2447-59.

ROBIN, C., OTTERSBACH, K., DURAND, C., PEETERS, M., VANES, L., TYBULEWICZ, V. & DZIERZAK, E. 2006. An unexpected role for IL-3 in the embryonic development of hematopoietic stem cells. *Dev Cell,* 11**,** 171-80.

RODRIGUEZ-MADOZ, J. R., SAN JOSE-ENERIZ, E., RABAL, O., ZAPATA-LINARES, N., MIRANDA, E., RODRIGUEZ, S., PORCIUNCULA, A., VILAS-ZORNOZA, A., GARATE, L., SEGURA, V., GURUCEAGA, E., AGIRRE, X., OYARZABAL, J. & PROSPER, F. 2017. Reversible dual inhibitor against G9a and DNMT1 improves human iPSC derivation enhancing MET and facilitating transcription factor engagement to the genome. *PLoS One,* 12**,** e0190275.

ROTHENBERG, E. V. & SCRIPTURE-ADAMS, D. D. 2008. Competition and collaboration: GATA-3, PU.1, and Notch signaling in early T-cell fate determination. *Semin Immunol,* 20**,** 236-46.

SAKURAI, M., KUNIMOTO, H., WATANABE, N., FUKUCHI, Y., YUASA, S., YAMAZAKI, S., NISHIMURA, T., SADAHIRA, K., FUKUDA, K., OKANO, H., NAKAUCHI, H., MORITA, Y., MATSUMURA, I., KUDO, K., ITO, E., EBIHARA, Y., TSUJI, K., HARADA, Y., HARADA, H., OKAMOTO, S. & NAKAJIMA, H. 2014. Impaired hematopoietic differentiation of RUNX1-mutated induced pluripotent stem cells derived from FPD/AML patients. *Leukemia,* 28**,** 2344-54.

SAMSEL, L. & MCCOY, J. P., JR. 2015. Imaging flow cytometry for the study of erythroid cell biology and pathology. *J Immunol Methods,* 423**,** 52-9.

SAUMELL, S., FLORENSA, L., LUNO, E., SANZO, C., CANIZO, C., HERNANDEZ, J. M., CERVERA, J., GALLART, M. A., CARBONELL, F., COLLADO, R., ARENILLAS, L., PEDRO, C., BARGAY, J., NOMDEDEU, B., XICOY, B., VALLESPI, T., RAYA, J. M., BELLOCH, L., SANZ, G. F. & SOLE, F. 2012. Prognostic value of trisomy 8 as a single anomaly and the influence of additional cytogenetic aberrations in primary myelodysplastic syndromes. *Br J Haematol,* 159**,** 311-21.

SAUMELL, S., SOLE, F., ARENILLAS, L., MONTORO, J., VALCARCEL, D., PEDRO, C., SANZO, C., LUNO, E., GIMENEZ, T., ARNAN, M., POMARES, H., DE PAZ, R., ARRIZABALAGA, B., JEREZ, A., MARTINEZ, A. B., SANCHEZ-CASTRO, J., RODRIGUEZ-GAMBARTE, J. D., RAYA, J. M., RIOS, E., RODRIGUEZ-RIVERA, M., ESPINET, B. & FLORENSA, L. 2015. Trisomy 8, a Cytogenetic Abnormality in Myelodysplastic Syndromes, Is Constitutional or Not? *PLoS One,* 10**,** e0129375.

SAWADA, K., KRANTZ, S. B., DESSYPRIS, E. N., KOURY, S. T. & SAWYER, S. T. 1989. Human colony-forming units-erythroid do not require accessory cells, but do require direct interaction with insulin-like growth factor I and/or insulin for erythroid development. *J Clin Invest,* 83**,** 1701-9.

SCHLAEGER, T. M., DAHERON, L., BRICKLER, T. R., ENTWISLE, S., CHAN, K., CIANCI, A., DEVINE, A., ETTENGER, A., FITZGERALD, K., GODFREY, M., GUPTA, D., MCPHERSON, J., MALWADKAR, P., GUPTA, M., BELL, B., DOI, A., JUNG, N., LI, X., LYNES, M. S., BROOKES, E., CHERRY, A. B., DEMIRBAS, D., TSANKOV, A. M., ZON, L. I., RUBIN, L. L., FEINBERG, A. P., MEISSNER, A., COWAN, C. A. & DALEY, G. Q. 2015. A comparison of non-integrating reprogramming methods. *Nat Biotechnol,* 33**,** 58-63.

SCHNITTGER, S., EDER, C., JEROMIN, S., ALPERMANN, T., FASAN, A., GROSSMANN, V., KOHLMANN, A., ILLIG, T., KLOPP, N., WICHMANN, H. E., KREUZER, K. A., SCHMID, C., STAIB, P., PECENY, R., SCHMITZ, N., KERN, W., HAFERLACH, C. & HAFERLACH, T. 2013. ASXL1 exon 12 mutations are frequent in AML with intermediate risk karyotype and are independently associated with an adverse outcome. *Leukemia,* 27**,** 82-91.

SCHOLER, H. R., DRESSLER, G. R., BALLING, R., ROHDEWOHLD, H. & GRUSS, P. 1990. Oct-4: a germline-specific transcription factor mapping to the mouse t-complex. *EMBO J,* 9**,** 2185-95.

SCHWARTZ, C., BECK, K., MINK, S., SCHMOLKE, M., BUDDE, B., WENNING, D. & KLEMPNAUER, K. H. 2003. Recruitment of p300 by C/EBPbeta triggers phosphorylation of p300 and modulates coactivator activity. *EMBO J,* 22**,** 882-92.

SEITA, J., EMA, H., OOEHARA, J., YAMAZAKI, S., TADOKORO, Y., YAMASAKI, A., ETO, K., TAKAKI, S., TAKATSU, K. & NAKAUCHI, H. 2007. Lnk negatively regulates self-renewal of hematopoietic stem cells by modifying thrombopoietin-mediated signal transduction. *Proc Natl Acad Sci U S A,* 104**,** 2349-54.

SEITA, J. & WEISSMAN, I. L. 2010. Hematopoietic stem cell: self-renewal versus differentiation. *Wiley Interdiscip Rev Syst Biol Med,* 2**,** 640-53.

SINGH, A. M. & DALTON, S. 2009. The cell cycle and Myc intersect with mechanisms that regulate pluripotency and reprogramming. *Cell Stem Cell,* 5**,** 141-9.

SLAYMAKER, I. M., GAO, L., ZETSCHE, B., SCOTT, D. A., YAN, W. X. & ZHANG, F. 2016. Rationally engineered Cas9 nucleases with improved specificity. *Science,* 351**,** 84-8.

SMITH, C., GORE, A., YAN, W., ABALDE-ATRISTAIN, L., LI, Z., HE, C., WANG, Y., BRODSKY, R. A., ZHANG, K., CHENG, L. & YE, Z. 2014. Whole-genome sequencing analysis reveals high specificity of CRISPR/Cas9 and TALEN-based genome editing in human iPSCs. *Cell Stem Cell,* 15**,** 12-3.

SMITH, Z. D., SINDHU, C. & MEISSNER, A. 2016. Molecular features of cellular reprogramming and development. *Nat Rev Mol Cell Biol,* 17**,** 139-54.

SNADDON, J., SMITH, M. L., NEAT, M., CAMBAL-PARRALES, M., DIXON-MCIVER, A., ARCH, R., AMESS, J. A., ROHATINER, A. Z., LISTER, T. A. & FITZGIBBON, J. 2003. Mutations of CEBPA in acute myeloid leukemia FAB types M1 and M2. *Genes Chromosomes Cancer,* 37**,** 72-8.

SOMMER, C. A., STADTFELD, M., MURPHY, G. J., HOCHEDLINGER, K., KOTTON, D. N. & MOSTOSLAVSKY, G. 2009. Induced pluripotent stem cell generation using a single lentiviral stem cell cassette. *Stem Cells,* 27**,** 543-9.

SON, M. J., SON, M. Y., SEOL, B., KIM, M. J., YOO, C. H., HAN, M. K. & CHO, Y. S. 2013. Nicotinamide overcomes pluripotency deficits and reprogramming barriers. *Stem Cells,* 31**,** 1121-35.

SONG, B., FAN, Y., HE, W., ZHU, D., NIU, X., WANG, D., OU, Z., LUO, M. & SUN, X. 2015. Improved hematopoietic differentiation efficiency of gene-corrected beta-thalassemia induced pluripotent stem cells by CRISPR/Cas9 system. *Stem Cells Dev,* 24**,** 1053-65.

SONG, J., YANG, D., XU, J., ZHU, T., CHEN, Y. E. & ZHANG, J. 2016. RS-1 enhances CRISPR/Cas9- and TALEN-mediated knock-in efficiency. *Nat Commun,* 7**,** 10548.

SOUFI, A., DONAHUE, G. & ZARET, K. S. 2012. Facilitators and impediments of the pluripotency reprogramming factors' initial engagement with the genome. *Cell,* 151**,** 994-1004.

SPANGRUDE, G. J., HEIMFELD, S. & WEISSMAN, I. L. 1988. Purification and characterization of mouse hematopoietic stem cells. *Science,* 241**,** 58-62.

SPERLING, A. S., GIBSON, C. J. & EBERT, B. L. 2017. The genetics of myelodysplastic syndrome: from clonal haematopoiesis to secondary leukaemia. *Nat Rev Cancer,* 17**,** 5-19.

SPYROU, N. & PAPAPETROU, E. P. 2020. Studying leukemia stem cell properties and vulnerabilities with human iPSCs. *Stem Cell Res,* 50**,** 102117.

SRIDHARAN, R., GONZALES-COPE, M., CHRONIS, C., BONORA, G., MCKEE, R., HUANG, C., PATEL, S., LOPEZ, D., MISHRA, N., PELLEGRINI, M., CAREY, M., GARCIA, B. A. & PLATH, K. 2013. Proteomic and genomic approaches reveal critical functions of H3K9 methylation and heterochromatin protein-1gamma in reprogramming to pluripotency. *Nat Cell Biol,* 15**,** 872-82.

STADTFELD, M., BRENNAND, K. & HOCHEDLINGER, K. 2008a. Reprogramming of pancreatic beta cells into induced pluripotent stem cells. *Curr Biol,* 18**,** 890-4.

STADTFELD, M., MAHERALI, N., BREAULT, D. T. & HOCHEDLINGER, K. 2008b. Defining molecular cornerstones during fibroblast to iPS cell reprogramming in mouse. *Cell Stem Cell,* 2**,** 230-40.

STARCZYNOWSKI, D. T., KUCHENBAUER, F., ARGIROPOULOS, B., SUNG, S., MORIN, R., MURANYI, A., HIRST, M., HOGGE, D., MARRA, M., WELLS, R. A., BUCKSTEIN, R., LAM, W., HUMPHRIES, R. K. & KARSAN, A. 2010. Identification of miR-145 and miR-146a as mediators of the 5q- syndrome phenotype. *Nat Med,* 16**,** 49-58.

STEENSMA, D. P., BEJAR, R., JAISWAL, S., LINDSLEY, R. C., SEKERES, M. A., HASSERJIAN, R. P. & EBERT, B. L. 2015. Clonal hematopoiesis of indeterminate potential and its distinction from myelodysplastic syndromes. *Blood,* 126**,** 9-16.

STEINLE, H., WEBER, M., BEHRING, A., MAU-HOLZMANN, U., SCHLENSAK, C., WENDEL, H. P. & AVCI-ADALI, M. 2019. Generation of iPSCs by Nonintegrative RNA-Based Reprogramming Techniques: Benefits of Self-Replicating RNA versus Synthetic mRNA. *Stem Cells Int,* 2019**,** 7641767.

SUGIMOTO, K., HIRANO, N., TOYOSHIMA, H., CHIBA, S., MANO, H., TAKAKU, F., YAZAKI, Y. & HIRAI, H. 1993. Mutations of the p53 gene in myelodysplastic syndrome (MDS) and MDS-derived leukemia. *Blood,* 81**,** 3022-6.

SUZUKI, K., YU, C., QU, J., LI, M., YAO, X., YUAN, T., GOEBL, A., TANG, S., REN, R., AIZAWA, E., ZHANG, F., XU, X., SOLIGALLA, R. D., CHEN, F., KIM, J., KIM, N. Y., LIAO, H. K., BENNER, C., ESTEBAN, C. R., JIN, Y., LIU, G. H., LI, Y. & IZPISUA BELMONTE, J. C. 2014. Targeted gene correction minimally impacts whole-genome mutational load in human-disease-specific induced pluripotent stem cell clones. *Cell Stem Cell,* 15**,** 31-6.

SWERDLOW, S. H., CAMPO, E., HARRIS, N. L., JAFFE, E. S., PILERI, S. & STEIN, H. 2008. WHO Classification of Tumours of the Haematopoietic and Lymphoid Tissues. *Lyon, France: International Agency for Research on Cancer,* 1-439.

TAKAHASHI, K., TANABE, K., OHNUKI, M., NARITA, M., ICHISAKA, T., TOMODA, K. & YAMANAKA, S. 2007. Induction of pluripotent stem cells from adult human fibroblasts by defined factors. *Cell,* 131**,** 861-72.

TAKAHASHI, K. & YAMANAKA, S. 2006. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell,* 126**,** 663-76.

TAKAHASHI, K. & YAMANAKA, S. 2016. A decade of transcription factor-mediated reprogramming to pluripotency. *Nat Rev Mol Cell Biol,* 17**,** 183-93.

TAKAISHI, M., TARUTANI, M., TAKEDA, J. & SANO, S. 2016. Mesenchymal to Epithelial Transition Induced by Reprogramming Factors Attenuates the Malignancy of Cancer Cells. *PLoS One,* 11**,** e0156904.

TANG, H., HAMMACK, C., OGDEN, S. C., WEN, Z., QIAN, X., LI, Y., YAO, B., SHIN, J., ZHANG, F., LEE, E. M., CHRISTIAN, K. M., DIDIER, R. A., JIN, P., SONG, H. & MING, G. L. 2016. Zika Virus Infects Human Cortical Neural Progenitors and Attenuates Their Growth. *Cell Stem Cell,* 18**,** 587-90.

TAOKA, K., ARAI, S., KATAOKA, K., HOSOI, M., MIYAUCHI, M., YAMAZAKI, S., HONDA, A., AIXINJUELUO, W., KOBAYASHI, T., KUMANO, K., YOSHIMI, A., OTSU, M., NIWA, A., NAKAHATA, T., NAKAUCHI, H. &

KUROKAWA, M. 2018. Using patient-derived iPSCs to develop humanized mouse models for chronic myelomonocytic leukemia and therapeutic drug identification, including liposomal clodronate. *Sci Rep,* 8**,** 15855.

TASIAN, S. K., CASAS, J. A., POSOCCO, D., GANDRE-BABBE, S., GAGNE, A. L., LIANG, G., LOH, M. L., WEISS, M. J., FRENCH, D. L. & CHOU, S. T. 2019. Mutation-specific signaling profiles and kinase inhibitor sensitivities of juvenile myelomonocytic leukemia revealed by induced pluripotent stem cells. *Leukemia,* 33**,** 181-190.

TAYLOR, E. W. 1965. The Mechanism of Colchicine Inhibition of Mitosis. I. Kinetics of Inhibition and the Binding of H3-Colchicine. *J Cell Biol,* 25**,** SUPPL:145-60.

TAYLOR, T. H., GITLIN, S. A., PATRICK, J. L., CRAIN, J. L., WILSON, J. M. & GRIFFIN, D. K. 2014. The origin, mechanisms, incidence and clinical consequences of chromosomal mosaicism in humans. *Hum Reprod Update,* 20**,** 571-81.

TEFFERI, A. & VARDIMAN, J. W. 2009. Myelodysplastic syndromes. *N Engl J Med,* 361**,** 1872-85.

TESAR, P. J., CHENOWETH, J. G., BROOK, F. A., DAVIES, T. J., EVANS, E. P., MACK, D. L., GARDNER, R. L. & MCKAY, R. D. 2007. New cell lines from mouse epiblast share defining features with human embryonic stem cells. *Nature,* 448**,** 196-9.

THOMSON, J. A., ITSKOVITZ-ELDOR, J., SHAPIRO, S. S., WAKNITZ, M. A., SWIERGIEL, J. J., MARSHALL, V. S. & JONES, J. M. 1998. Embryonic stem cell lines derived from human blastocysts. *Science,* 282**,** 1145-7.

TILL, J. E. & MC, C. E. 1961. A direct measurement of the radiation sensitivity of normal mouse bone marrow cells. *Radiat Res,* 14**,** 213-22.

TOBER, J., KONISKI, A., MCGRATH, K. E., VEMISHETTI, R., EMERSON, R., DE MESY-BENTLEY, K. K., WAUGH, R. & PALIS, J. 2007. The megakaryocyte lineage originates from hemangioblast precursors and is an integral component both of primitive and of definitive hematopoiesis. *Blood,* 109**,** 1433-41.

TOHYAMA, K. 2018. Present status and perspective of laboratory hematology in Japan: On the standardization of blood cell morphology including myelodysplasia: On behalf of the Japanese Society for Laboratory Hematology. *Int J Lab Hematol,* 40 Suppl 1**,** 120-125.

TOLEDO, C. M., DING, Y., HOELLERBAUER, P., DAVIS, R. J., BASOM, R., GIRARD, E. J., LEE, E., CORRIN, P., HART, T., BOLOURI, H., DAVISON, J., ZHANG, Q., HARDCASTLE, J., ARONOW, B. J., PLAISIER, C. L., BALIGA, N. S., MOFFAT, J., LIN, Q., LI, X. N., NAM, D. H., LEE, J., POLLARD, S. M., ZHU, J., DELROW, J. J., CLURMAN, B. E., OLSON, J. M. & PADDISON, P. J. 2015. Genome-wide CRISPR-Cas9 Screens Reveal Loss of Redundancy between PKMYT1 and WEE1 in Glioblastoma Stem-like Cells. *Cell Rep,* 13**,** 2425-2439.

TOWNSLEY, D. M., DUMITRIU, B. & YOUNG, N. S. 2016. Danazol Treatment for Telomere Diseases. *N Engl J Med,* 375**,** 1095-6.

TRAXLER, E. A., YAO, Y., WANG, Y. D., WOODARD, K. J., KURITA, R., NAKAMURA, Y., HUGHES, J. R., HARDISON, R. C., BLOBEL, G. A., LI, C. & WEISS, M. J. 2016. A genome-editing strategy to treat beta-hemoglobinopathies that recapitulates a mutation associated with a benign genetic condition. *Nat Med,* 22**,** 987-90.

TRUONG, D. J., KUHNER, K., KUHN, R., WERFEL, S., ENGELHARDT, S., WURST, W. & ORTIZ, O. 2015. Development of an intein-mediated split-Cas9 system for gene therapy. *Nucleic Acids Res,* 43**,** 6450-8.

TSAI, F. Y., KELLER, G., KUO, F. C., WEISS, M., CHEN, J., ROSENBLATT, M., ALT, F. W. & ORKIN, S. H. 1994. An early haematopoietic defect in mice lacking the transcription factor GATA-2. *Nature,* 371**,** 221-6.

TURNER, J. & CROSSLEY, M. 1999. Mammalian Kruppel-like transcription factors: more than just a pretty finger. *Trends Biochem Sci,* 24**,** 236-40.

URNOV, F. D., REBAR, E. J., HOLMES, M. C., ZHANG, H. S. & GREGORY, P. D. 2010. Genome editing with engineered zinc finger nucleases. *Nat Rev Genet,* 11**,** 636-46.

VARDIMAN, J. W., THIELE, J., ARBER, D. A., BRUNNING, R. D., BOROWITZ, M. J., PORWIT, A., HARRIS, N. L., LE BEAU, M. M., HELLSTROM-LINDBERG, E., TEFFERI, A. & BLOOMFIELD, C. D. 2009. The 2008 revision of

the World Health Organization (WHO) classification of myeloid neoplasms and acute leukemia: rationale and important changes. *Blood,* 114**,** 937-51.

VARLAKHANOVA, N. V., COTTERMAN, R. F., DEVRIES, W. N., MORGAN, J., DONAHUE, L. R., MURRAY, S., KNOWLES, B. B. & KNOEPFLER, P. S. 2010. myc maintains embryonic stem cell pluripotency and self-renewal. *Differentiation,* 80**,** 9-19.

VELAZQUEZ, L., CHENG, A. M., FLEMING, H. E., FURLONGER, C., VESELY, S., BERNSTEIN, A., PAIGE, C. J. & PAWSON, T. 2002. Cytokine signaling and hematopoietic homeostasis are disrupted in Lnk-deficient mice. *J Exp Med,* 195**,** 1599-611.

VELYCHKO, S., ADACHI, K., KIM, K. P., HOU, Y., MACCARTHY, C. M., WU, G. & SCHOLER, H. R. 2019. Excluding Oct4 from Yamanaka Cocktail Unleashes the Developmental Potential of iPSCs. *Cell Stem Cell,* 25**,** 737-753 e4.

VILAS, J. M., CARNEIRO, C., DA SILVA-ALVAREZ, S., FERREIROS, A., GONZALEZ, P., GOMEZ, M., ORTEGA, S., SERRANO, M., GARCIA-CABALLERO, T., GONZALEZ-BARCIA, M., VIDAL, A. & COLLADO, M. 2018. Adult Sox2+ stem cell exhaustion in mice results in cellular senescence and premature aging. *Aging Cell,* 17**,** e12834.

WAHL, M. C., WILL, C. L. & LUHRMANN, R. 2009. The spliceosome: design principles of a dynamic RNP machine. *Cell,* 136**,** 701-18.

WALENDA, T., STIEHL, T., BRAUN, H., FROBEL, J., HO, A. D., SCHROEDER, T., GOECKE, T. W., RATH, B., GERMING, U., MARCINIAK-CZOCHRA, A. & WAGNER, W. 2014. Feedback signals in myelodysplastic syndromes: increased self-renewal of the malignant clone suppresses normal hematopoiesis. *PLoS Comput Biol,* 10**,** e1003599.

WANG, T., WEI, J. J., SABATINI, D. M. & LANDER, E. S. 2014. Genetic screens in human cells using the CRISPR-Cas9 system. *Science,* 343**,** 80-4.

WANG, Z., ORON, E., NELSON, B., RAZIS, S. & IVANOVA, N. 2012. Distinct lineage specification roles for NANOG, OCT4, and SOX2 in human embryonic stem cells. *Cell Stem Cell,* 10**,** 440-54.

WATOWICH, S. S. 2011. The erythropoietin receptor: molecular structure and hematopoietic signaling pathways. *J Investig Med,* 59**,** 1067-72.

WEGNER, M. 2010. All purpose Sox: The many roles of Sox proteins in gene expression. *Int J Biochem Cell Biol,* 42**,** 381-90.

WELTE, K., PLATZER, E., GABRILOVE, J. L., LU, L., LEVI, E., POLIVKA, A., MERTELSMANN, R. & MOORE, M. A. 1985. Purification to apparent homogeneity and biochemical characterization of human pluripotent hematopoietic colony-stimulating factor. *Haematol Blood Transfus,* 29**,** 398-401.

WELTNER, J., BALBOA, D., KATAYAMA, S., BESPALOV, M., KRJUTSKOV, K., JOUHILAHTI, E. M., TROKOVIC, R., KERE, J. & OTONKOSKI, T. 2018. Human pluripotent reprogramming with CRISPR activators. *Nat Commun,* 9**,** 2643.

WEN, X. M., HU, J. B., YANG, J., QIAN, W., YAO, D. M., DENG, Z. Q., ZHANG, Y. Y., ZHU, X. W., GUO, H., LIN, J. & QIAN, J. 2015. CEBPA methylation and mutation in myelodysplastic syndrome. *Med Oncol,* 32**,** 192.

WESELY, J., KOTINI, A. G., IZZO, F., LUO, H., YUAN, H., SUN, J., GEORGOMANOLI, M., ZVIRAN, A., DESLAURIERS, A. G., DUSAJ, N., NIMER, S. D., LESLIE, C., LANDAU, D. A., KHARAS, M. G. & PAPAPETROU, E. P. 2020. Acute Myeloid Leukemia iPSCs Reveal a Role for RUNX1 in the Maintenance of Human Leukemia Stem Cells. *Cell Rep,* 31**,** 107688.

WILMUT, I., SCHNIEKE, A. E., MCWHIR, J., KIND, A. J. & CAMPBELL, K. H. 1997. Viable offspring derived from fetal and adult mammalian cells. *Nature,* 385**,** 810-3.

WILSON, N. K., FOSTER, S. D., WANG, X., KNEZEVIC, K., SCHUTTE, J., KAIMAKIS, P., CHILARSKA, P. M., KINSTON, S., OUWEHAND, W. H., DZIERZAK, E., PIMANDA, J. E., DE BRUIJN, M. F. & GOTTGENS, B. 2010. Combinatorial transcriptional control in blood stem/progenitor cells: genome-wide analysis of ten major transcriptional regulators. *Cell Stem Cell,* 7**,** 532-44.

WOLTJEN, K., MICHAEL, I. P., MOHSENI, P., DESAI, R., MILEIKOVSKY, M., HAMALAINEN, R., COWLING, R., WANG, W., LIU, P., GERTSENSTEIN, M., KAJI, K., SUNG, H. K. & NAGY, A. 2009. piggyBac transposition reprograms fibroblasts to induced pluripotent stem cells. *Nature,* 458**,** 766-70.

WONG, T. N., RAMSINGH, G., YOUNG, A. L., MILLER, C. A., TOUMA, W., WELCH, J. S., LAMPRECHT, T. L., SHEN, D., HUNDAL, J., FULTON, R. S., HEATH, S., BATY, J. D., KLCO, J. M., DING, L., MARDIS, E. R., WESTERVELT, P., DIPERSIO, J. F., WALTER, M. J., GRAUBERT, T. A., LEY, T. J., DRULEY, T., LINK, D. C. & WILSON, R. K. 2015. Role of TP53 mutations in the origin and evolution of therapy-related acute myeloid leukaemia. *Nature,* 518**,** 552-555.

XIE, M., LU, C., WANG, J., MCLELLAN, M. D., JOHNSON, K. J., WENDL, M. C., MCMICHAEL, J. F., SCHMIDT, H. K., YELLAPANTULA, V., MILLER, C. A., OZENBERGER, B. A., WELCH, J. S., LINK, D. C., WALTER, M. J., MARDIS, E. R., DIPERSIO, J. F., CHEN, F., WILSON, R. K., LEY, T. J. & DING, L. 2014. Age-related mutations associated with clonal hematopoietic expansion and malignancies. *Nat Med,* 20**,** 1472-8.

YAMADA, Y., WARREN, A. J., DOBSON, C., FORSTER, A., PANNELL, R. & RABBITTS, T. H. 1998. The T cell leukemia LIM protein Lmo2 is necessary for adult mouse hematopoiesis. *Proc Natl Acad Sci U S A,* 95**,** 3890-5.

YAMAMOTO, G., NANNYA, Y., KATO, M., SANADA, M., LEVINE, R. L., KAWAMATA, N., HANGAISHI, A., KUROKAWA, M., CHIBA, S., GILLILAND, D. G., KOEFFLER, H. P. & OGAWA, S. 2007. Highly sensitive method for genomewide detection of allelic composition in nonpaired, primary tumor specimens by use of affymetrix single-nucleotide-polymorphism genotyping microarrays. *Am J Hum Genet,* 81**,** 114-26.

YAMAMOTO, R., MORITA, Y., OOEHARA, J., HAMANAKA, S., ONODERA, M., RUDOLPH, K. L., EMA, H. & NAKAUCHI, H. 2013. Clonal analysis unveils self-renewing lineage-restricted progenitors generated directly from hematopoietic stem cells. *Cell,* 154**,** 1112-1126.

YAMANAKA, S. 2007. Strategies and new developments in the generation of patient-specific pluripotent stem cells. *Cell Stem Cell,* 1**,** 39-49.

YAMASAKI, A. E., KING, N. E., MATSUI, H., JEPSEN, K. & PANOPOULOS, A. D. 2019. Two iPSC lines generated from the bone marrow of a relapsed/refractory AML patient display normal karyotypes and myeloid differentiation potential. *Stem Cell Res,* 41**,** 101587.

YAN, D., HUTCHISON, R. E. & MOHI, G. 2012. Critical requirement for Stat5 in a mouse model of polycythemia vera. *Blood,* 119**,** 3539-49.

YANG, L., BRYDER, D., ADOLFSSON, J., NYGREN, J., MANSSON, R., SIGVARDSSON, M. & JACOBSEN, S. E. 2005. Identification of Lin(-)Sca1(+)kit(+)CD34(+)Flt3- short-term hematopoietic stem cells capable of rapidly reconstituting and rescuing myeloablated transplant recipients. *Blood,* 105**,** 2717-23.

YE, Z., LIU, C. F., LANIKOVA, L., DOWEY, S. N., HE, C., HUANG, X., BRODSKY, R. A., SPIVAK, J. L., PRCHAL, J. T. & CHENG, L. 2014. Differential sensitivity to JAK inhibitory drugs by isogenic human erythroblasts and hematopoietic progenitors generated from patient-specific induced pluripotent stem cells. *Stem Cells,* 32**,** 269-78.

YEAMANS, C., WANG, D., PAZ-PRIEL, I., TORBETT, B. E., TENEN, D. G. & FRIEDMAN, A. D. 2007. C/EBPalpha binds and activates the PU.1 distal enhancer to induce monocyte lineage commitment. *Blood,* 110**,** 3136-42.

YIN, C. C., MEDEIROS, L. J. & BUESO-RAMOS, C. E. 2010. Recent advances in the diagnosis and classification of myeloid neoplasms--comments on the 2008 WHO classification. *Int J Lab Hematol,* 32**,** 461-76.

YINGJUN, X., YUHUAN, X., YUCHANG, C., DONGZHI, L., DING, W., BING, S., YI, Y., DIAN, L., YANTING, X., ZEYU, X., NENGQING, L., DIYU, C. & XIAOFANG, S. 2019. CRISPR/Cas9 gene correction of HbH-CS thalassemia-induced pluripotent stem cells. *Ann Hematol,* 98**,** 2661-2671.

YLA-HERTTUALA, S. 2018. iPSC-Derived Cardiomyocytes Taken to Rescue Infarcted Heart Muscle in Coronary Heart Disease Patients. *Mol Ther,* 26**,** 2077.

YOO, J., KIM, J., LEE, J. H., KIM, H., JANG, S. J., SEO, H. H., OH, S. T., HYEON, S. J., RYU, H., KIM, J. & MOH, S. H. 2020. Acceleration of somatic cell reprogramming into the induced pluripotent stem cell using a mycosporine-like amino acid, Porphyra 334. *Sci Rep,* 10**,** 3684.

YOON, D. S., CHOI, Y., JANG, Y., LEE, M., CHOI, W. J., KIM, S. H. & LEE, J. W. 2014. SIRT1 directly regulates SOX2 to maintain self-renewal and multipotency in bone marrow-derived mesenchymal stem cells. *Stem Cells,* 32**,** 3219-31.

YOON, H. S., GHALEB, A. M., NANDAN, M. O., HISAMUDDIN, I. M., DALTON, W. B. & YANG, V. W. 2005. Kruppel-like factor 4 prevents centrosome amplification following gamma-irradiation-induced DNA damage. *Oncogene,* 24**,** 4017-25.

YOSHIDA, K., SANADA, M., SHIRAISHI, Y., NOWAK, D., NAGATA, Y., YAMAMOTO, R., SATO, Y., SATO-OTSUBO, A., KON, A., NAGASAKI, M., CHALKIDIS, G., SUZUKI, Y., SHIOSAKA, M., KAWAHATA, R., YAMAGUCHI, T., OTSU, M., OBARA, N., SAKATA-YANAGIMOTO, M., ISHIYAMA, K., MORI, H., NOLTE, F., HOFMANN, W. K., MIYAWAKI, S., SUGANO, S., HAFERLACH, C., KOEFFLER, H. P., SHIH, L. Y., HAFERLACH, T., CHIBA, S., NAKAUCHI, H., MIYANO, S. & OGAWA, S. 2011. Frequent pathway mutations of splicing machinery in myelodysplasia. *Nature,* 478**,** 64-9.

YOSHIHARA, M., ARAKI, R., KASAMA, Y., SUNAYAMA, M., ABE, M., NISHIDA, K., KAWAJI, H., HAYASHIZAKI, Y. & MURAKAWA, Y. 2017. Hotspots of De Novo Point Mutations in Induced Pluripotent Stem Cells. *Cell Rep,* 21**,** 308-315.

YOSHIOKA, N., GROS, E., LI, H. R., KUMAR, S., DEACON, D. C., MARON, C., MUOTRI, A. R., CHI, N. C., FU, X. D., YU, B. D. & DOWDY, S. F. 2013. Efficient generation of human iPSCs by a synthetic self-replicative RNA. *Cell Stem Cell,* 13**,** 246-54.

YU, J., HU, K., SMUGA-OTTO, K., TIAN, S., STEWART, R., SLUKVIN, II & THOMSON, J. A. 2009. Human induced pluripotent stem cells free of vector and transgene sequences. *Science,* 324**,** 797-801.

YU, J., LI, Y., LI, T., LI, Y., XING, H., SUN, H., SUN, L., WAN, D., LIU, Y., XIE, X. & JIANG, Z. 2020. Gene mutational analysis by NGS and its clinical significance in patients with myelodysplastic syndrome and acute myeloid leukemia. *Exp Hematol Oncol,* 9**,** 2.

YU, J., VODYANIK, M. A., SMUGA-OTTO, K., ANTOSIEWICZ-BOURGET, J., FRANE, J. L., TIAN, S., NIE, J., JONSDOTTIR, G. A., RUOTTI, V., STEWART, R., SLUKVIN, II & THOMSON, J. A. 2007. Induced pluripotent stem cell lines derived from human somatic cells. *Science,* 318**,** 1917-20.

YUNG, S. K., TILGNER, K., LEDRAN, M. H., HABIBOLLAH, S., NEGANOVA, I., SINGHAPOL, C., SARETZKI, G., STOJKOVIC, M., ARMSTRONG, L., PRZYBORSKI, S. & LAKO, M. 2013. Brief report: human pluripotent stem cell models of fanconi anemia deficiency reveal an important role for fanconi anemia proteins in cellular reprogramming and survival of hematopoietic progenitors. *Stem Cells,* 31**,** 1022-9.

YUSUF, I., KHARAS, M. G., CHEN, J., PERALTA, R. Q., MARUNIAK, A., SAREEN, P., YANG, V. W., KAESTNER, K. H. & FRUMAN, D. A. 2008. KLF4 is a FOXO target gene that suppresses B cell proliferation. *Int Immunol,* 20**,** 671-81.

ZHAN, Z., SONG, L., ZHANG, W., GU, H., CHENG, H., ZHANG, Y., YANG, Y., JI, G., FENG, H., CHENG, T. & LI, Y. 2019. Absence of cyclin-dependent kinase inhibitor p27 or p18 increases efficiency of iPSC generation without induction of iPSC genomic instability. *Cell Death Dis,* 10**,** 271.

ZHANG, J., NIU, C., YE, L., HUANG, H., HE, X., TONG, W. G., ROSS, J., HAUG, J., JOHNSON, T., FENG, J. Q., HARRIS, S., WIEDEMANN, L. M., MISHINA, Y. & LI, L. 2003. Identification of the haematopoietic stem cell niche and control of the niche size. *Nature,* 425**,** 836-41.

ZHANG, P., ANDRIANAKOS, R., YANG, Y., LIU, C. & LU, W. 2010. Kruppel-like factor 4 (Klf4) prevents embryonic stem (ES) cell differentiation by regulating Nanog gene expression. *J Biol Chem,* 285**,** 9180-9.

ZHANG, P., IWASAKI-ARAI, J., IWASAKI, H., FENYUS, M. L., DAYARAM, T., OWENS, B. M., SHIGEMATSU, H., LEVANTINI, E., HUETTNER, C. S., LEKSTROM-HIMES, J. A., AKASHI, K. & TENEN, D. G. 2004. Enhancement of hematopoietic stem cell repopulating capacity and self-renewal in the absence of the transcription factor C/EBP alpha. *Immunity,* 21**,** 853-63.

ZHANG, X., SU, J., JEONG, M., KO, M., HUANG, Y., PARK, H. J., GUZMAN, A., LEI, Y., HUANG, Y. H., RAO, A., LI, W. & GOODELL, M. A. 2016. DNMT3A and TET2 compete and cooperate to repress lineage-specific transcription factors in hematopoietic stem cells. *Nat Genet,* 48**,** 1014-23.

ZHENG, X., ZHAN, Z., NAREN, D., LI, J., YAN, T. & GONG, Y. 2017. Prognostic value of SRSF2 mutations in patients with de novo myelodysplastic syndromes: A meta-analysis. *PLoS One,* 12**,** e0185053.

ZHOU, W. & FREED, C. R. 2009. Adenoviral gene delivery can reprogram human fibroblasts to induced pluripotent stem cells. *Stem Cells,* 27**,** 2667-74.

ZHOU, Y. Y. & ZENG, F. 2013. Integration-free methods for generating induced pluripotent stem cells. *Genomics Proteomics Bioinformatics,* 11**,** 284-7.

Supplementary

Sample	Gene	Locus	Mutation type	VAF	Function	Coding	Protein
MDS27 hiPSC clone 2	ASXL1	chr20:31023821	SNV	50.58	Missense	c.3306G>T	p.Glu1102Asp
	RUNX1	chr21:36206864	INDEL	51.40	Frameshift Insertion	c.648_649 ins TCC	p.Gly217fs
		chr21:36259324	SNV	49.30	Missense	c.167T>C	p.Leu56Ser
	SRSF2	chr17:74732935	SNV	56.28	Missense	c.284C>A	p.Pro95His
MDS27 hiPSC clone 3	ASXL1	chr20:31023821	SNV	53.00	Missense	c.3306G>T	p.Glu1102Asp
	RUNX1	chr21:36206864	INDEL	48.00	Frameshift Insertion	c.648_649 ins TCC	p.Gly217fs
		chr21:36259324	SNV	45.40	Missense	c.167T>C	p.Leu56Ser
	SRSF2	chr17:74732935	SNV	52.32	Missense	c.284C>A	p.Pro95His
MDS27 hiPSC clone 4	ASXL1	chr20:31023821	SNV	49.32	Missense	c.3306G>T	p.Glu1102Asp
	RUNX1	chr21:36206864	INDEL	55.43	Frameshift Insertion	c.648_649 ins TCC	p.Gly217fs
		chr21:36259324	SNV	65.75	Missense	c.167T>C	p.Leu56Ser
	SRSF2	chr17:74732935	SNV	56.02	Missense	c.284C>A	p.Pro95His
MDS27 hiPSC clone 9	ASXL1	chr20:31023821	SNV	56.00	Missense	c.3306G>T	p.Glu1102Asp
	RUNX1	chr21:36206864	INDEL	49.07	Frameshift Insertion	c.648_649 ins TCC	p.Gly217fs
		chr21:36259324	SNV	46.07	Missense	c.167T>C	p.Leu56Ser
	SRSF2	chr17:74732935	SNV	54.70	Missense	c.284C>A	p.Pro95His
MDS27 hiPSC clone 10	ASXL1	chr20:31023821	SNV	49.40	Missense	c.3306G>T	p.Glu1102Asp
	RUNX1	chr21:36206864	INDEL	52.56	Frameshift Insertion	c.648_649 ins TCC	p.Gly217fs
		chr21:36259324	SNV	47.90	Missense	c.167T>C	p.Leu56Ser
	SRSF2	chr17:74732935	SNV	53.99	Missense	c.284C>A	p.Pro95His
MDS27 hiPSC clone 13	ASXL1	chr20:31023821	SNV	51.33	Missense	c.3306G>T	p.Glu1102Asp
	RUNX1	chr21:36206864	INDEL	51.36	Frameshift Insertion	c.648_649 ins TCC	p.Gly217fs
		chr21:36259324	SNV	47.82	Missense	c.167T>C	p.Leu56Ser
	SRSF2	chr17:74732935	SNV	53.70	Missense	c.284C>A	p.Pro95His

Sample	Gene	Locus	Mutation type	VAF	Function	Coding	Protein
MDS27 hiPSC clone 14	ASXL1	chr20:31023821	SNV	50.85	Missense	c.3306G>T	p.Glu1102Asp
	RUNX1	chr21:36206864	INDEL	52.76	Frameshift Insertion	c.648_649 ins TCC	p.Gly217fs
		chr21:36259324	SNV	63.63	Missense	c.167T>C	p.Leu56Ser
	SRSF2	chr17:74732935	SNV	56.98	Missense	c.284C>A	p.Pro95His
MDS27 hiPSC clone 15	ASXL1	chr20:31023821	SNV	49.02	Missense	c.3306G>T	p.Glu1102Asp
	RUNX1	chr21:36206864	INDEL	51.96	Frameshift Insertion	c.648_649 ins TCC	p.Gly217fs
		chr21:36259324	SNV	65.90	Missense	c.167T>C	p.Leu56Ser
	SRSF2	chr17:74732935	SNV	56.85	Missense	c.284C>A	p.Pro95His
MDS27 hiPSC clone 16	ASXL1	chr20:31023821	SNV	51.68	Missense	c.3306G>T	p.Glu1102Asp
	RUNX1	chr21:36206864	INDEL	52.52	Frameshift Insertion	c.648_649 ins TCC	p.Gly217fs
		chr21:36259324	SNV	48.75	Missense	c.167T>C	p.Leu56Ser
	SRSF2	chr17:74732935	SNV	53.94	Missense	c.284C>A	p.Pro95His
MDS27 hiPSC clone 17	ASXL1	chr20:31023821	SNV	49.32	Missense	c.3306G>T	p.Glu1102Asp
	RUNX1	chr21:36206864	INDEL	51.81	Frameshift Insertion	c.648_649 ins TCC	p.Gly217fs
		chr21:36259324	SNV	49.10	Missense	c.167T>C	p.Leu56Ser
	SRSF2	chr17:74732935	SNV	51.65	Missense	c.284C>A	p.Pro95His
MDS27 hiPSC clone 20	ASXL1	chr20:31023821	SNV	51.05	Missense	c.3306G>T	p.Glu1102Asp
	RUNX1	chr21:36206864	INDEL	51.05	Frameshift Insertion	c.648_649 ins TCC	p.Gly217fs
		chr21:36259324	SNV	51.03	Missense	c.167T>C	p.Leu56Ser
	SRSF2	chr17:74732935	SNV	55.31	Missense	c.284C>A	p.Pro95His
MDS27 hiPSC clone 21	ASXL1	chr20:31023821	SNV	49.80	Missense	c.3306G>T	p.Glu1102Asp
	RUNX1	chr21:36206864	INDEL	53.17	Frameshift Insertion	c.648_649 ins TCC	p.Gly217fs
		chr21:36259324	SNV	51.03	Missense	c.167T>C	p.Leu56Ser
	SRSF2	chr17:74732935	SNV	56.90	Missense	c.284C>A	p.Pro95His

Sample	Gene	Locus	Mutation type	VAF	Function	Coding	Protein
MDS27 hiPSC clone 23	ASXL1	chr20:31023821	SNV	50.25	Missense	c.3306G>T	p.Glu1102Asp
	RUNX1	chr21:36206864	INDEL	52.11	Frameshift Insertion	c.648_649 ins TCC	p.Gly217fs
		chr21:36259324	SNV	49.40	Missense	c.167T>C	p.Leu56Ser
	SRSF2	chr17:74732935	SNV	52.92	Missense	c.284C>A	p.Pro95His
MDS27 hiPSC clone 24	ASXL1	chr20:31023821	SNV	51.85	Missense	c.3306G>T	p.Glu1102Asp
	RUNX1	chr21:36206864	INDEL	55.34	Frameshift Insertion	c.648_649 ins TCC	p.Gly217fs
		chr21:36259324	SNV	50.43	Missense	c.167T>C	p.Leu56Ser
	SRSF2	chr17:74732935	SNV	52.49	Missense	c.284C>A	p.Pro95His
MDS27 hiPSC clone 25	ASXL1	chr20:31023821	SNV	50.90	Missense	c.3306G>T	p.Glu1102Asp
	RUNX1	chr21:36206864	INDEL	52.45	Frameshift Insertion	c.648_649 ins TCC	p.Gly217fs
		chr21:36259324	SNV	50.25	Missense	c.167T>C	p.Leu56Ser
	SRSF2	chr17:74732935	SNV	54.75	Missense	c.284C>A	p.Pro95His
MDS27 hiPSC clone 26	ASXL1	chr20:31023821	SNV	49.32	Missense	c.3306G>T	p.Glu1102Asp
	RUNX1	chr21:36206864	INDEL	51.81	Frameshift Insertion	c.648_649 ins TCC	p.Gly217fs
		chr21:36259324	SNV	49.10	Missense	c.167T>C	p.Leu56Ser
	SRSF2	chr17:74732935	SNV	52.48	Missense	c.284C>A	p.Pro95His

Supplement figure 1: Genotyping screening

Genotyping was performed by the Hospital La Fe, Valencia, Spain for generated clones from MDS27 peripheral blood sample on 2013 by Sendai vector, using an array of the 40 most mutated genes in AML. All clones harbour the same mutations as the original sample (2013).

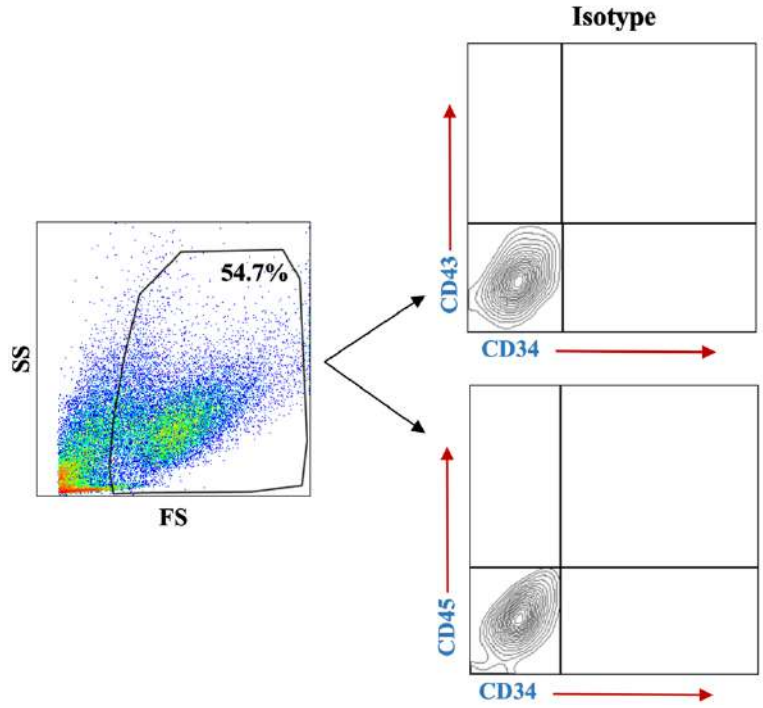
Sample	Gene	Locus	Mutation type	VAF	Function	Coding	Protein
MDS27 hiPSC clone 1	ASXL1	chr20:31023821	SNV	51.60	Missense	c.3306G>T	p.Glu1102Asp
	RUNX1	chr21:36206864	INDEL	49.80	Frameshift Insertion	c.648_649 ins TCC	p.Gly217fs
		chr21:36259324	SNV	63.63	Missense	c.167T>C	p.Leu56Ser
	SRSF2	chr17:74732935	SNV	56.83	Missense	c.284C>A	p.Pro95His
MDS27 hiPSC clone 2	ASXL1	chr20:31023821	SNV	49.47	Missense	c.3306G>T	p.Glu1102Asp
	RUNX1	chr21:36206864	INDEL	50.78	Frameshift Insertion	c.648_649 ins TCC	p.Gly217fs
		chr21:36259324	SNV	65.90	Missense	c.167T>C	p.Leu56Ser
	SRSF2	chr17:74732935	SNV	53.81	Missense	c.284C>A	p.Pro95His
MDS27 hiPSC clone 3	ASXL1	chr20:31023821	SNV	50.65	Missense	c.3306G>T	p.Glu1102Asp
	RUNX1	chr21:36206864	INDEL	50.43	Frameshift Insertion	c.648_649 ins TCC	p.Gly217fs
		chr21:36259324	SNV	48.75	Missense	c.167T>C	p.Leu56Ser
	SRSF2	chr17:74732935	SNV	55.09	Missense	c.284C>A	p.Pro95His
MDS27 hiPSC clone 4	ASXL1	chr20:31023821	SNV	50.70	Missense	c.3306G>T	p.Glu1102Asp
	RUNX1	chr21:36206864	INDEL	54.34	Frameshift Insertion	c.648_649 ins TCC	p.Gly217fs
		chr21:36259324	SNV	49.10	Missense	c.167T>C	p.Leu56Ser
	SRSF2	chr17:74732935	SNV	51.59	Missense	c.284C>A	p.Pro95His
MDS27 hiPSC clone 5	ASXL1	chr20:31023821	SNV	48.87	Missense	c.3306G>T	p.Glu1102Asp
	RUNX1	chr21:36206864	INDEL	51.30	Frameshift Insertion	c.648_649 ins TCC	p.Gly217fs
		chr21:36259324	SNV	51.03	Missense	c.167T>C	p.Leu56Ser
	SRSF2	chr17:74732935	SNV	52.96	Missense	c.284C>A	p.Pro95His
MDS27 hiPSC clone 6	ASXL1	chr20:31023821	SNV	49.80	Missense	c.3306G>T	p.Glu1102Asp
	RUNX1	chr21:36206864	INDEL	52.91	Frameshift Insertion	c.648_649 ins TCC	p.Gly217fs
		chr21:36259324	SNV	51.03	Missense	c.167T>C	p.Leu56Ser
	SRSF2	chr17:74732935	SNV	52.13	Missense	c.284C>A	p.Pro95His

Sample	Gene	Locus	Mutation type	VAF	Function	Coding	Protein
MDS27 hiPSC clone 7	ASXL1	chr20:31023821	SNV	50.25	Missense	c.3306G>T	p.Glu1102Asp
	RUNX1	chr21:36206864	INDEL	49.70	Frameshift Insertion	c.648_649 ins TCC	p.Gly217fs
		chr21:36259324	SNV	63.63	Missense	c.167T>C	p.Leu56Ser
	SRSF2	chr17:74732935	SNV	53.31	Missense	c.284C>A	p.Pro95His
MDS27 hiPSC clone 8	ASXL1	chr20:31023821	SNV	52.13	Missense	c.3306G>T	p.Glu1102Asp
	RUNX1	chr21:36206864	INDEL	50.45	Frameshift Insertion	c.648_649 ins TCC	p.Gly217fs
		chr21:36259324	SNV	65.90	Missense	c.167T>C	p.Leu56Ser
	SRSF2	chr17:74732935	SNV	53.94	Missense	c.284C>A	p.Pro95His
MDS27 hiPSC clone 11	ASXL1	chr20:31023821	SNV	48.50	Missense	c.3306G>T	p.Glu1102Asp
	RUNX1	chr21:36206864	INDEL	48.61	Frameshift Insertion	c.648_649 ins TCC	p.Gly217fs
		chr21:36259324	SNV	48.75	Missense	c.167T>C	p.Leu56Ser
	SRSF2	chr17:74732935	SNV	55.08	Missense	c.284C>A	p.Pro95His
MDS27 hiPSC clone 12	ASXL1	chr20:31023821	SNV	52.38	Missense	c.3306G>T	p.Glu1102Asp
	RUNX1	chr21:36206864	INDEL	51.37	Frameshift Insertion	c.648_649 ins TCC	p.Gly217fs
		chr21:36259324	SNV	49.10	Missense	c.167T>C	p.Leu56Ser
	SRSF2	chr17:74732935	SNV	54.30	Missense	c.284C>A	p.Pro95His
MDS27 hiPSC clone 13	ASXL1	chr20:31023821	SNV	51.30	Missense	c.3306G>T	p.Glu1102Asp
	RUNX1	chr21:36206864	INDEL	53.35	Frameshift Insertion	c.648_649 ins TCC	p.Gly217fs
		chr21:36259324	SNV	51.03	Missense	c.167T>C	p.Leu56Ser
	SRSF2	chr17:74732935	SNV	54.59	Missense	c.284C>A	p.Pro95His
MDS27 hiPSC clone 14	ASXL1	chr20:31023821	SNV	49.80	Missense	c.3306G>T	p.Glu1102Asp
	RUNX1	chr21:36206864	INDEL	51.05	Frameshift Insertion	c.648_649 ins TCC	p.Gly217fs
		chr21:36259324	SNV	51.03	Missense	c.167T>C	p.Leu56Ser
	SRSF2	chr17:74732935	SNV	57.19	Missense	c.284C>A	p.Pro95His

Sample	Gene	Locus	Mutation type	VAF	Function	Coding	Protein
MDS27 hiPSC clone 15	ASXL1	chr20:31023821	SNV	48.35	Missense	c.3306G>T	p.Glu1102Asp
	RUNX1	chr21:36206864	INDEL	51.17	Frameshift Insertion	c.648_649 ins TCC	p.Gly217fs
		chr21:36259324	SNV	63.63	Missense	c.167T>C	p.Leu56Ser
	SRSF2	chr17:74732935	SNV	53.33	Missense	c.284C>A	p.Pro95His
MDS27 hiPSC clone 16	ASXL1	chr20:31023821	SNV	53.60	Missense	c.3306G>T	p.Glu1102Asp
	RUNX1	chr21:36206864	INDEL	49.62	Frameshift Insertion	c.648_649 ins TCC	p.Gly217fs
		chr21:36259324	SNV	65.90	Missense	c.167T>C	p.Leu56Ser
	SRSF2	chr17:74732935	SNV	58.10	Missense	c.284C>A	p.Pro95His
MDS27 hiPSC clone 17	ASXL1	chr20:31023821	SNV	49.50	Missense	c.3306G>T	p.Glu1102Asp
	RUNX1	chr21:36206864	INDEL	46.58	Frameshift Insertion	c.648_649 ins TCC	p.Gly217fs
		chr21:36259324	SNV	48.75	Missense	c.167T>C	p.Leu56Ser
	SRSF2	chr17:74732935	SNV	56.92	Missense	c.284C>A	p.Pro95His
MDS27 hiPSC clone 18	ASXL1	chr20:31023821	SNV	48.55	Missense	c.3306G>T	p.Glu1102Asp
	RUNX1	chr21:36206864	INDEL	48.54	Frameshift Insertion	c.648_649 ins TCC	p.Gly217fs
		chr21:36259324	SNV	49.10	Missense	c.167T>C	p.Leu56Ser
	SRSF2	chr17:74732935	SNV	58.11	Missense	c.284C>A	p.Pro95His

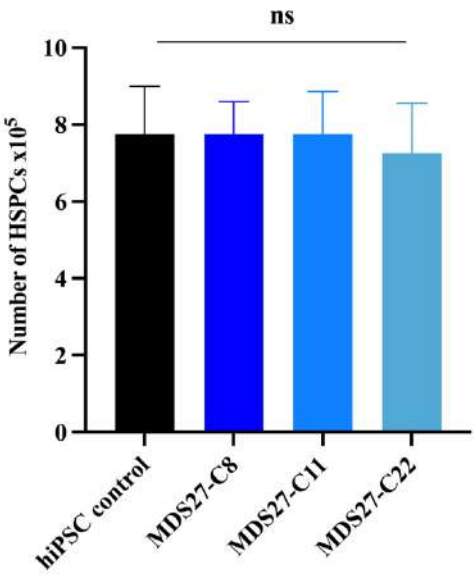
Supplement figure 2: Genotyping screening

Genotyping was performed by the Hospital La Fe, Valencia, Spain generated from MDS27 peripheral blood sample on 2013 by episomal method, using an array of the 40 most mutated genes in AML. All clones harbour the same mutations as the original sample (2013).



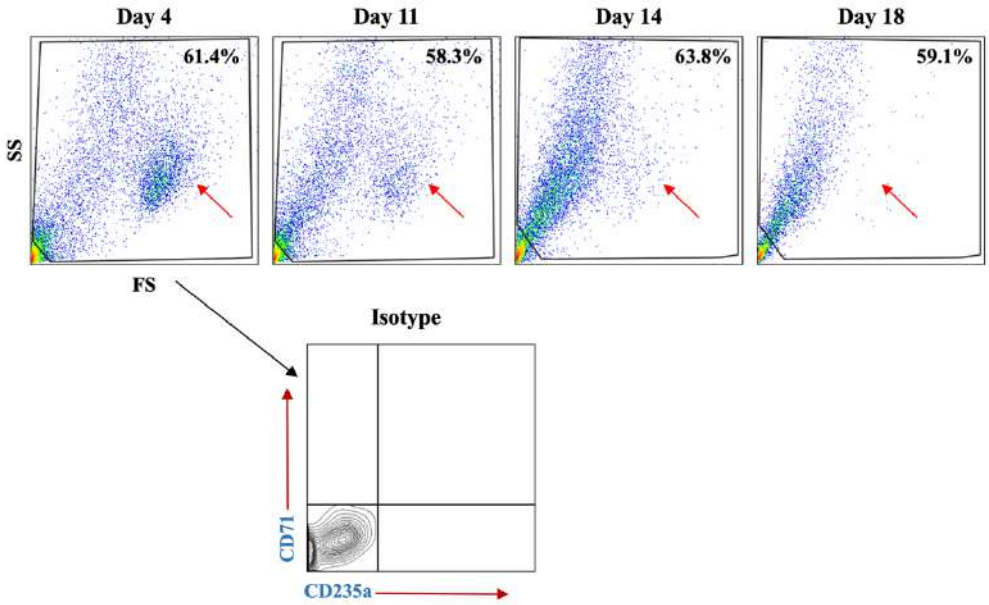
Supplement figure 3: Gating strategy of HSPCs

Gating strategy of HSPCs differentiation. The cells were gated first based on FSC and SSC Cells of unstained cells and then the live cells were analysed for HSPCs markers (CD34, CD43 & CD45) according to isotype control.



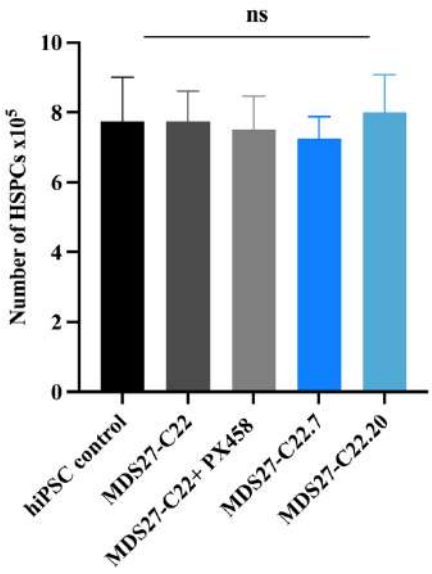
Supplement figure 4: Number of HSPCs

Number of HSPCs after 14 days in the differentiation culture and per one well of 12-well plate. Statistical results are presented as mean ± standard error of the mean (SEM) and analysed using One-way ANOVA with multiple comparisons. p values were represented a ns for not significant. Data represent 4 independent experiments.



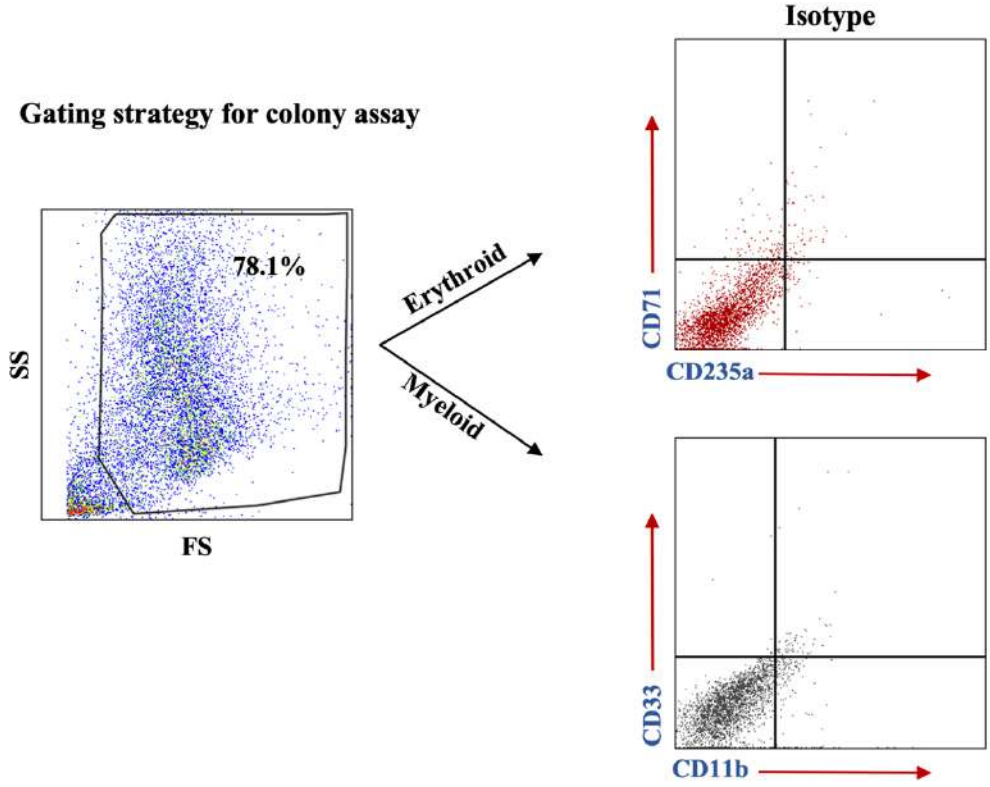
Supplement figure 5: Gating strategy of erythroid cells

Gating strategy of erythroid differentiation. The cells were gated first based on FSC and SS of unstained cells and then the live cells were analysed for erythroid markers (CD71 & CD235a) according to isotype control. The red arrows indicate the progenitor and erythroblast populations in which this population is lost during the maturation.



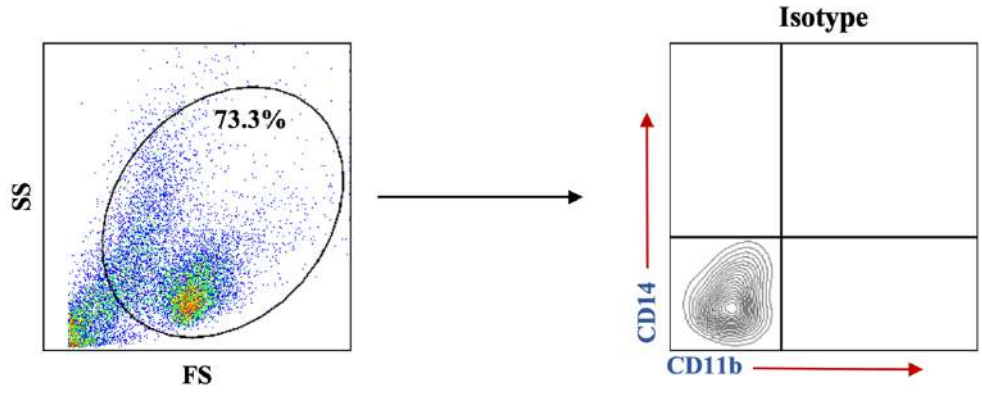
Supplement figure 6: Number of HSPCs

Number of HSPCs after 14 days in the differentiation culture and per one well of 12-well plate. Statistical results are presented as mean ± standard error of the mean (SEM) and analysed using One-way ANOVA with multiple comparisons. p values were represented a ns for not significant. Data represent 4 independent experiments.



Supplement figure 7: Gating strategy of CFUs

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



Supplement figure 8: Gating strategy of myeloid cells

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

1. Comparing with hIPSC control 2. Comparing with MDS27-C22 3. Comparing with MDS27-C22+PX458	MDS27-C22.7	MDS27-C22.20
P-Values of Pro/Baso on day 4	1. * 0.0448 2. ** 0.0070 3. ** 0.0098	1. * 0.0118 2. ** 0.0070 3. ** 0.0063
P-Values of Poly/Ortho on day 4	1. ns 2. ns 3. ns	1. ns 2. ns 3. ns
P-Values of mature on day 4	1. * 0.0499 2. * 0.0161 3. * 0.0337	1. * 0.0246 2. * 0.0991 3. * 0.0161
P-Values of Pro/Baso on day 7	1. ** 0.0069 2. ns 3. ns	1. ** 0.0034 2. ns 3. ns
P-Values of Poly/Ortho on day 7	1. ns 2. ns 3. ns	1. ns 2. ns 3. ns
P-Values of mature on day 7	1. ** 0.0076 2. ** 0.0010 3. ** 0.0013	1. ** 0.0021 2. ** 0.0031 3. ** 0.0044
P-Values of Pro/Baso on day 11&14	1. ns 2. ns 3. ns	1. ns 2. ns 3. ns
P-Values of Poly/Ortho& mature on day 11&14	1. ****<0.0001 2. ****<0.0001 3. ****<0.0001	1. ****<0.0001 2. ****<0.0001 3. ****<0.0001
P-Values of Pro/Baso & mature on day 18	1. ns 2. ns 3. ns	1. ns 2. ns 3. ns
P-Values of Poly/Ortho on day 18	1. ** 0.0010 2. ** 0.0012 3. ** 0.0098	1. ** 0.0016 2. ** 0.0055 3. ** 0.0030
P-Values of mature on day 18	1. * 0.0500 2. * 0.0498 3. * 0.0465	1. * 0.0173 2. * 0.0140 3. * 0.0131

Supplement figure 9: p values of erythroid cells

P values were obtained for different stages of erythrocyte differentiation during several time points. The results present in Figure 5.24