

University of Nottingham
UK | CHINA | MALAYSIA

Understanding the role of HOXA9 in haematopoiesis

Emma Margaret Collins MSc, BSc

School of Medicine

Biodiscovery Institute

The University of Nottingham

Thesis submitted to the University of Nottingham for the Degree of Doctor of Philosophy

June 2022

Acknowledgements

Firstly, I would like to thank my principal supervisor Dr Alexander Thompson for his support, guidance, and encouragement over the past five years of this PhD. I would also like to thank my secondary supervisor Professor Chris Denning for his expertise and advice throughout this PhD and the technical and administrative staff at the Biodiscovery Institute who have provided me with invaluable help and support.

I would like to thank Dr Monica Nafria for her advice in optimisation of haematopoietic differentiation. I thank Dr David Onion, Dr Nicola Croxall and Dr Martin Grundy for their help and advice with flow cytometry experiments and Dr Nigel Mongan for his help with RNA sequencing data analysis. I am especially grateful to Dr Alexander Kondrashov and Dr Ashley Ramsawhook for being so generous with their time, expertise and for teaching me everything I now know about molecular biology. I would also like to thank Dr Aishah Nasir for her assistance with fluorescence imaging, for her support and friendship.

A special thanks to the BCASC group and my surrogate lab, the Cardio group. Especially, Dr Kate Butcher, Lara Dolan, Mays Abuhantash and Ben Johnson. I have truly enjoyed working with you, thank you for your encouragement and kind words during the stressful times of the last few years.

I would not have made it through this PhD without Paulina Durzcak, Maria Eleftheriou, Sara Bodbin, Kate Vo, Peggy Lo, Analy Serna Valverde, Paige Lowe-Hallett and Kavita Raniga. From the extended coffee breaks and pub trips, I have so many happy memories of this experience. You have been there for me every step of this PhD and meeting each of you has made it all worthwhile. I will be forever grateful for your support and friendship.

I would like to thank Evelyn Shiel, Emmet Leyden, Edu Vico Oton and Sam Wandi. Thank you for always checking-in on me, I am truly lucky to have you as friends.

Finally, I would like to thank my family, especially my parents, Marie and Tony, who have supported me in all my endeavours. Thank you for your constant support, love and for always believing in me.

Declaration

I hereby declare that this thesis has been composed by myself and has not been submitted for any other degree previously. This project was supervised by Dr Alexander Thompson and Professor Chris Denning. Acknowledgements of specific procedures not performed by myself are stated; otherwise, the work described is my own.

Abstract

Embryonic haematopoiesis can be modeled *in vitro* using human pluripotent stem cells (hPSCs). These models aim to recapitulate key developmental stages of *in vivo* haematopoiesis and have been demonstrated to closely replicate several phenotypes of extra- and intra-embryonic haematopoiesis. However, generation of *bona fide* haematopoietic stem cells which possess multilineage long-term reconstitution ability remains a challenge of hPSC models. In the past decade, numerous RNA sequencing studies have demonstrated human cord blood, foetal liver and bone marrow-derived haematopoietic progenitors possess a unique *HOXA* expression profile; a signature which is absent in hPSC-derived progenitors. Notably, *HOXA9* has been identified as one of the most differentially expressed genes in a comparison of foetal liver- and hPSC-derived progenitors. During hPSC-based haematopoietic differentiation, *HOXA9* expression has been described in mesoderm and arterial endothelium derivatives which subsequently generate progenitors with erythro-myeloid and lymphoid potential. However, the precise function of HOXA9 during haematopoietic development remains unknown. To assess the function of HOXA9 at distinct stages of haematopoietic development, an inducible HOXA9 overexpression system was generated using CRISPR/Cas9 gene editing. In parallel, an hPSC haematopoietic differentiation system was optimised. Characterisation of haematopoietic differentiation using flow cytometry and qPCR revealed differentiating cells were first directed to a mesoderm state before acquiring haemato-vascular potential and giving rise to floating haematopoietic progenitors. Notably, two distinct populations of progenitors were generated; an early erythroid-biased population and a later population with robust erythro-myeloid differentiation potential and a lymphoid gene signature, as assessed by colony forming unit assays and RNA sequencing analysis. The inducible expression system stimulated robust HOXA9 overexpression during early haematopoietic differentiation, however, this effect was lost as differentiation progressed, with preliminary data suggesting methylation within the TRE promoter and AAVS1 were responsible. Therefore, the inducible system was used to stimulate transient overexpression of HOXA9 solely during mesoderm patterning. Transcriptome analysis at this timepoint suggested overexpression of HOXA9 promoted specification of all mesoderm sub-types and did not act to promote haemogenic mesoderm specifically. Analysis of haematopoietic progenitors derived from HOXA9-induced and uninduced cultures revealed similar immunophenotypes and functional activities, suggesting transient HOXA9 overexpression during early haematopoietic differentiation does not influence haematopoietic development or cell fate.

Table of Contents

Chapter 1: Introduction 1

1.1 HOX genes 1

1.1.1 HOX genes in development 1

1.1.2 HOX specificity 3

1.2 Embryonic haematopoiesis 4

1.2.1 Primitive haematopoietic wave 5

1.2.2 Definitive haematopoietic waves 6

1.2.3 HSC expansion 7

1.2.4 Endothelial origin of definitive haematopoiesis 8

1.3 Regulation of embryonic haematopoiesis 9

1.4 HOX genes in haematopoiesis 13

1.4.1 HOX in HSC maintenance and differentiation 13

1.4.2 HOX in malignant haematopoiesis 15

1.4.3 HOX in HSC specification and emergence 16

1.5 Hypothesis and Aims 20

Chapter 2: Materials and Methods 22

2.1 Cell Culture 22

2.1.1 Cell lines 22

2.1.2 Human induced pluripotent stem cell maintenance and passage 22

2.1.3 Cryopreservation 23

2.1.4 Karyotyping 23

2.1.5 Nucleofection of live cells 24

2.1.6 Cardiac differentiation and dissociation 24

2.1.7 Haematopoietic differentiation using STEMdiff™ Hematopoietic Kit 25

2.1.8 Spin Embryoid Body haematopoietic differentiation 25

2.1.9 Colony forming unit assay 27

2.2 Flow Cytometry 27

2.2.1 Staining of cells for flow cytometry 27
2.2.2 Flow cytometry analysis 29
2.3 Molecular Biology 29
2.3.1 RNA extraction 29
2.3.2 cDNA synthesis 29
2.3.3 Quantitative real-time Polymerase Chain Reaction (qPCR) 29
2.3.4 Immunostaining 31
2.3.5 Genomic DNA extraction 32
2.3.6 Methylated DNA immunoprecipitation (MeDIP) 32
2.4 Generation of plasmids for CRISPR/Cas9 transfections 35
2.4.1 Primer design 35
2.4.2 Polymerase chain reaction (PCR) 35
2.4.3 Fusion PCR 38
2.4.4 Restriction digestion 39
2.4.5 Dephosphorylation 40
2.4.6 Ligation 40
2.4.7 Agarose gel electrophoresis 40
2.4.8 Bacterial transformation 41
2.4.9 DNA mini- and midi-prep 41
2.4.10 Sanger Sequencing 41
2.5 Microscopy 41
2.6 RNA-sequencing 42
2.7 Statistical analysis 42
Chapter 3: Engineering an inducible HOXA9-mScarlet overexpression system in hiPSCs 43
3.1 Introduction 43
3.1.1 CRISPR/Cas9 genome editing 43
3.1.2 Genomic safe-harbour sites 46
3.1.3 Inducible gene expression systems 48

3.2 Aims 49
3.3 Results 50
3.3.1 AAVS1 targeting strategy 50
3.3.2 Building of the HOXA9-mScarlet donor template plasmid 50
3.3.3 Transient transfection of pAAVS1-TRE-HOXA9-mScarlet 60
3.3.4 Survival analysis 61
3.3.5 Targeting of the AAVS1 locus 62
3.3.6 Inducible HOXA9 clones are functional and specific for DOX 66
3.3.7 Inducible HOXA9 clones retain their pluripotency following DOX treatment 70
3.3.8 Loss of DOX sensitivity following mesoderm specification in iA9 clone differentiation to haematopoietic cells 71
3.4 Discussion 74
Chapter 4: Cardiac differentiation proof of concept study 79
4.2 Introduction 79
4.2.1 Cardiomyocyte generation *in vivo* and *in vitro* 79
4.2.2 DNA methylation 80
4.3 Aims 82
4.4 Results 83
4.4.1 Optimisation of DOX treatment conditions 83
4.4.2 Loss of inducible HOXA9-mScarlet expression during cardiac differentiation of iA9 clones 88
4.4.3 DNA methylation identified within the AAVS1 locus and TRE promoter during cardiac differentiation 93
4.4.4 *HOXA9*-mScarlet expression can be maintained by continuous DOX treatment 97
4.5 Discussion 99
Chapter 5: Optimisation of haematopoietic differentiation of hiPSCs 107
5.1 Introduction 107
5.1.1 Human PSC-based haematopoietic differentiation 107
5.1.2 Directed differentiation 112

5.1.3 Transcription factor-driven differentiation 113
5.1.4 Forward reprogramming 114
5.2 Aims 114
5.3 Results 115
5.3.1 Method selection and optimisation 115
5.3.2 STEMdiff™ Hematopoietic kit produces $CD34^{+}CD38^{-}$ progenitors at varying efficiencies 117
5.3.3 Spin EB differentiation shows poor mesoderm differentiation 119
5.3.4 Spin EB differentiation of feeder- and Matrigel-cultured hPSCs do not produce suspension cells 120
5.3.5 Cross titration of BMP4 and Activin A identified conditions which produce $CD34^{+}CD43/45^{+}$ haematopoietic progenitors 126
5.3.6 SB/CHIR treatment is essential to specify haematopoietic progenitors 131
5.3.7 Spin EB differentiation yields two functionally different haematopoietic progenitor populations 139
5.4 Discussion 147
Chapter 6: Transient HOXA9 overexpression during mesoderm patterning 155
6.1 Introduction 155
6.2 Aims 156
6.3 Results 156
6.3.1 Optimisation of DOX treatment conditions in iPSC-derived mesendoderm 156
6.3.2 DOX inducibility is not maintained throughout spin EB differentiation 159
6.3.3 Transient overexpression of HOXA9 during mesoderm patterning 163
6.3.4 Transcriptome analysis of mesoderm populations 172
6.3.5 Transcriptome analysis of haematopoietic populations 182
6.4 Discussion 188
6.4.1 Untreated cultures have impaired mesoderm differentiation compared with SB/CHIR-treated cultures 189
6.4.2 Transient HOXA9 overexpression does not specifically promote haemogenic mesoderm formation 190

6.4.3 Haematopoietic progenitors generated on d18 possess a gene signature and functional activity similar to extra-embryonic EMPs 191

6.4.4 Early transient overexpression of HOXA9 produced minor changes in the transcriptome profile of d18 haematopoietic progenitors 193

Chapter 7: General Discussion 195

List of Figures

Figure 1.1: Genomic organisation and expression patterns of HOX genes 3
Figure 1.2: HOX gene and protein structure 4
Figure 1.3: Mouse embryonic haematopoiesis 6
Figure 1.4: EHT in the mouse AGM 9
Figure 1.5: Proposed role for HOXA in a hPSC model of haematopoiesis 19
Figure 3.1: Schematic of AAVS1 locus targeting strategy 51
Figure 3.2: Inducible HOXA9 donor plasmid design 52
Figure 3.3: Building the TRE-HOXA9-T2A fragment .....
Figure 3.4: Building of the BsiWI-mScarlet-pA fragment 55
Figure 3.5: Building of the final TRE-HOXA9-T2A-mScarlet-pA fragment 56
Figure 3.6: Amplification of TRE-HOXA9-mScarlet 57
Figure 3.7: Restriction enzyme digestion and ligation to create pAAVS1-TRE-HOXA9-mScarlet 58
Figure 3.8: Map of CRISPR donor template plasmid, pAAVS1-TRE-HOXA9-mScarlet 59
Figure 3.9: Transient transfection of inducible HOXA9 donor plasmid 61
Figure 3.10: Identification of optimal G418 concentration for selection of targeted clones 62
Figure 3.11: CRISPR/Cas9n workflow 63
Figure 3.12: Genotyping PCR and sequencing identifies correctly targeted clones 65
Figure 3.13: All clones are heterozygous 66
Figure 3.14: All clones upregulate *HOXA9* and mScarlet in the presence of DOX .....
Figure 3.15: Quantification of mScarlet images using Columbus™ 68
Figure 3.16: HOXA9 and mScarlet proteins induced expressed at similar levels following DOX treatment 70
Figure 3.17: iA9 clone 54 iPSCs retain pluripotency in response to DOX 71
Figure 3.18: STEMdiff™ Hematopoietic kit differentiation overview 72
Figure 3.19: DOX mediated *HOXA9*/mScarlet induction is not maintained throughout haematopoietic differentiation 73
Figure 4.1: mScarlet and 2A staining in DOX-treated iA9 clone 54 iPSCs 84
Figure 4.2: Quantification of fluorescence staining relative to nuclear staining using Columbus™ 85
Figure 4.3: DOX concentration and number of treatments impact the level of mScarlet and 2A expression 86
Figure 4.4: mScarlet intensity increases with successive DOX treatments 87
Figure 4.5: iA9 clones can be differentiated to cardiomyocytes 88

Figure 4.6: DOX at 2 µg/mL DOX reduces total cell number on day 2 of cardiac differentiation 89
Figure 4.7: The ability of DOX to induce HOXA9 decreases as cardiac differentiation proceeds 91
Figure 4.8: mScarlet expression reduces as cardiac differentiation progresses 92
Figure 4.9: Day 2 of cardiac differentiation corresponds to mesendoderm formation 93
Figure 4.10: Representation of regions examined for DNA methylation within the TRE promoter and AAVS1 locus 95
Figure 4.11: iA9 clone 5 is enriched for methylation on d2 and 6 of cardiac differentiation 96
Figure 4.12: Inducible *HOXA9*/mScarlet expression can be maintained by continuous DOX treatment 98
Figure 4.13: Tet responsive elements 104
Figure 5.1: Optimisation of aggregate seeding conditions for STEMdiff™ Hematopoietic differentiation 116
Figure 5.2: Haematopoietic differentiation using STEMdiff™ Hematopoietic kit produces CD34+CD38- haematopoietic progenitors at varying efficiencies 118
Figure 5.3: Spin EB haematopoietic differentiation shows poor mesoderm specification 119
Figure 5.4: RPAT and S/R-H9 cell lines can produce NCAM+KDR+ mesoderm 123
Figure 5.5: RPAT and S/R-H9 cells differ in CD90 and CD34 expression on day 7 of differentiation 124
Figure 5.6: Differentiation of both RPAT and S/R-H9 cell lines failed to produce suspension cells 125
Figure 5.7: Concentration of BMP4 and Activin A influences NCAM+KDR+ mesoderm formation 127
Figure 5.8: CD43 expression in d7 EBs varied with BMP4 and Activin A concentration 128
Figure 5.9: All BMP4/Activin A conditions produce suspension cells 129
Figure 5.10: Flow cytometry analysis and CFU assay demonstrate 5 ng/mL BMP4/7.5 ng/mL Activin A yields cells with the greatest haematopoietic potential 130
Figure 5.11: SB/CHIR treatment on day 2 of differentiation enhances EB size 132
Figure 5.12: Addition of SB/CHIR on day 2 of differentiation is essential for specification of KDR+ mesoderm 133
Figure 5.13: Day 2 SB/CHIR treatment enhances the formation of haemato-vascular mesoderm .....
Figure 5.14: Optimisation of CFU assay conditions .....

Figure 5.15: SB/CHIR-treated cultures generate two populations of CD34+CD43+CD90+CD38- haematopoietic progenitors 138
Figure 5.16: Addition of SB/CHIR 38h or 40h post-seeding does not affect the efficiency of mesoderm or haemato-vascular cell formation..... 141
Figure 5.17: HOXA signature upregulated by d7 of differentiation..... 143
Figure 5.18: Haematopoietic progenitors with different colony forming abilities are produced with addition of SB/CHIR at 38h or 40h of differentiation 144
Figure 5.19: qPCR reveals gradual acquisition of key haematopoietic transcription factors as differentiation progresses 146
Figure 6.1: Optimisation of DOX treatment conditions in iPSC-derived mesendoderm 158
Figure 6.2: Inducible *HOXA9/mScarlet* expression is not retained after d3 of spin EB haematopoietic differentiation 160
Figure 6.3: Spin EB differentiation of iA9 iPSCs yields haematopoietic progenitors .. 162
Figure 6.4: Experimental plan to establish the function of HOXA9 in mesoderm patterning and haematopoietic differentiation. 164
Figure 6.5: SB/CHIR treatment is necessary to form NCAM+KDR+ mesendoderm .. 165
Figure 6.6: Formation of haemato-vascular cells is enhanced in cultures treated with SB/CHIR 166
Figure 6.7: DOX induction at 40h post-seeding produces robust HOXA9 expression on d4 of differentiation 168
Figure 6.8: Transient HOXA9 overexpression at 40h post-seeding does not affect the formation or functionality of HSPCs in SB/CHIR treated cultures. 169
Figure 6.9: Similar gene expression profile observed in d14 and d18 SC+ and SC+DOX+ suspension cells 170
Figure 6.10: Relationship between d4, d14 and d18 cells under the different treatment conditions used..... 174
Figure 6.11: Gene expression changes induced in d4 mesoderm 174
Figure 6.12: Pathway analysis of differentially expressed genes in d4 SC- cultures versus d4 SC+ and SC+DOX+ cultures..... 176
Figure 6.13: Pathway analysis of differentially expressed genes in d4 SC-DOX+ cultures versus d4 SC+ and SC+DOX+ cultures 177
Figure 6.14: Pathway analysis of differentially expressed genes in +SB/CHIR and -SB/CHIR groups treated +/- DOX 178
Figure 6.15: GSEA and KEGG pathway analysis of d14 and d18 suspension cells .. 181
Figure 6.16: Gene expression of key haematopoietic transcription factors and surface markers in d14 and d18 suspension cells 183

Figure 6.17: Expression of lineage-specific genes in d14 and d18 suspension cells . 185
Figure 6.18: Pathways associated with downregulated genes in d18SC+ cells versus d18 SC+DOX+ cells..... 188
Supplementary Figure 1: Optimisation of genotyping PCR primers..... 243
Supplementary Figure 2: Unsupervised hierarchal clustering of d4, d14 and d18 samples 244

List of Tables

Table 2.1: Number of EBs required for downstream assays 26
Table 2.2: Number of day 14/18 cells required for downstream assays 27
Table 2.3: Conjugated antibodies used for flow cytometry 28
Table 2.4: Conjugated isotype antibodies used for flow cytomtery 28
Table 2.5: qPCR primers 31
Table 2.6: Primary and secondary antibodies used for immunostaining 32
Table 2.7: MeDIP qPCR primers 34
Table 2.8: PCR thermocycling conditions 35
Table 2.9: PCR reaction set-up 36
Table 2.10: Genotyping PCR primers 36
Table 2.11: Primer pairs used for PCR and Fusion PCR 37
Table 2.12: Annealing template reaction set-up 38
Table 2.13: Annealing PCR conditions 38
Table 2.14: Extension of annealing template reaction set-up 39
Table 2.15: PCR amplification conditions of fusion product 39
Table 4.1: Summary of Tet-On inducible expression systems which demonstrate robust transgene expression in different cell type 102
Table 5.1: Summary of key hPSC-derived haematopoietic differentiation protocols .. 111
Table 5.2: Comparison of spin EB differentiation procedures 121
Table 6.1: DEGs in d14 and d18 cultures treated +/-DOX at 40h 186
Supplementary Table 1: Number of DEGs in d4 samples comparisons under indicated thresholds 245
Supplementary Table 2: List of DEGs in d4 SC+DOX+ versus d4 SC+ cultures and d4 SC-DOX+ versus d4 SC- cultures 245
Supplementary Table 3: Number of DEGs in d14 and d18 sample comparisons under indicated thresholds 246

Abbreviations

Abbreviation	Description
AAVS1	Adeno-Associated Virus Integration Site 1
AGM	Aorta-gonad-mesonephros
ALL	Acute lymphoid leukaemia
AML	Acute myeloid leukaemia
APEL	Albumin Polyvinylalcohol Essential Lipids
B2M	Beta-2-Microglobulin
BMP4	Bone Morphogenetic Protein 4
C/EBPb	CCAAT Enhancer Binding Protein Beta
CCR5	C-C Motif Chemokine Receptor 5
CDH5	Cadherin 5
CDX	Caudal Type Homeobox 1
CFC	Colony forming cell
CFU	Colony forming unit
CRISPR	Clustered regularly interspaced short palindromic repeats
CS	Carnegie stage
CXCR4	C-X-C chemokine receptor type 4
CYLBL	Citramalyl-CoA Lyase
DAPI	4′,6-diamidino-2-phenylindole
DC	Dendritic cell
DLL4	Delta Like Canonical Notch Ligand 4
DOX	Doxycycline
dpc	days post conception
DSB	Double stand break
EDTA	Ethylenediaminetetraacetic acid
EFNB2	Ephrin B2
EMP	Erythro-myeloid progenitor
EOMES	Eomesodermin
EPO	Erythropoietin
ERG	ETS Transcription Factor ERG
ESAM	Endothelial Cell Adhesion Molecule
ETV2	ETS Variant Transcription Factor 2
FCER1A	Fc Epsilon Receptor Ia
FGF2	(Fibroblast Growth Factor 2
FLI1	friend leukaemia transcription factor 1
FLT3	Fms Related Receptor Tyrosine Kinase 3
FLT-ITD	FLT internal tandem duplication
FOXF1	Forkhead Box F1
GATA	GATA-binding factor
GBX	Gastrulation Brain Homeobo
GCDR	Gentle cell dissociation agent
GFI1	Growth factor independence 1
GFI1B	Growth factor independent 1B
GM-CSF	Granulocyte macrophage-colony stimulating factor
GO	Gene Ontology
GP9	Glycoprotein IX Platelet

gRNA	guide RNA
GSC	Goosecoid Homeobox
GSK	Glycogen synthase kinas
HBSS	Hanks' Balanced Salt Solution
HDR	Homology-Directed Repair
HES1	Hes Family BHLH Transcription Factor 1
HEY2	Hairy/enhancer-of-split related with YRPW motif protein 2
HLA	Human leukocyte antigen
HSC	Haematopoietic stem cell
IGF2	insulin-like growth factor 2
IL3	Interleukin 3
IL6	Interleukin 6
ITGA	Integrin Subunit Alpha
KDR	kinase insert domain receptor
KEGG	Kyoto Encyclopedia of Genes and Genomes
KIT	KIT Proto-Oncogene, Receptor Tyrosine Kinase
LCOR	Ligand Dependent Nuclear Receptor Corepressor
LDB1	LIM Domain Binding 1
LMO2	Lim domain only 2
LSD1	Lysine-specific demethylase 1
LSK	Lin− Sca-1+ c-Kit+
MEIS	myeloid ectropic insertion site
MESP1	Mesoderm Posterior BHLH Transcription Factor 1
MIXL1	Mix Paired-Like Homeobox
MLL	Mixed lineage leukaemia
MYB	MYB Proto-Oncogene
NEA1	nuclear paraspeckle assembly transcription factor
NFE2	Nuclear Factor, Erythroid 2
NHEJ	Non-homologous end joining
NPM1	Nucleophosmin
NSG	NOD scid gamma
NUP98	Nucleoporin 98
PECAM1	Platelet endothelial cell adhesion molecule-1
PAM	protospacer adjacent motif
PBS	phosphate buffered saline
PBX	pre-B-cell leukaemia
PF4	Platelet factor 4
PP1A	Protein phosphatase 1
pSp	Para-aortic splanchnopleura
PTPRC	Protein Tyrosine Phosphatase Receptor Type C
RI	Rock inhibitor
RORA	RAR Related Orphan Receptor A
rtTA	reverse tetracycline-controlled transactivator
RUNX1	Runt-related transcription factor 1
SCF	Stem cell factor
SHOX2	Short Stature Homeobox
SHS	Safe harbour site
SOX17	SRY-Box Transcription Factor 17

SOX4	SRY-Box Transcription Factor 4)
SP1	Specificity Protein 1
SPI1	Spi-1 Proto-Oncogene
SPN	Sialophorin
TAL1	TAL BHLH transcription factor
TALE	Three amino acid extension
TALEN	Transcription activator-like effector nucleases
TBX3	T-Box Transcription Factor 3
TBXT	T-Box Transcription Factor T
TEK	TEK Receptor Tyrosine Kinase
Tet	Tetracycline
tetO	tet Operon
TPO	Thyroid peroxidase
TRE	Tet responsive element
tTA	tetracycline-controlled transactivator
UBE2C	Ubiquitin Conjugating Enzyme E2
UBE2S	Ubiquitin Conjugating Enzyme E2 S
VEGF	Vascular endothelial growth factor
ZFN	Zinc finger nucleases

Chapter 1: Introduction

1.1 HOX genes

1.1.1 HOX genes in development

The homeobox gene family encodes the second largest family of transcription factors in the human genome, with approx. 300 genes identified (Holland et al., 2007; Luo et al., 2019). Homeobox genes are often involved in developmental processes and are characterised by the presence of a 183-nucleotide Homeobox sequence which encodes a DNA-binding homeodomain (Holland et al., 2007). The Hox gene family are a subgroup of homeobox genes which are evolutionarily highly conserved and present in the genome of all plants, fungi and animals. In bilateral animals, all Hox genes function in patterning of the anterior-posterior (AP) axis during early embryogenesis. This role was first described in *Drosophila melanogaster,* where mutation of *bithorax* and *antennapedia* genes resulted in transformation of one part of the body onto another. The *bithorax* and *antennapedia* genes were shown to possess a homologous DNA sequence, termed 'Homeobox', and were subsequently found to be part of two clusters, comprising of 3 – 5 Homeobox-containing genes, collectively referred to as the homeotic complex (*HOM-C*). Hox genes represent the vertebrate equivalent of the bithorax/antennapedia cluster. While, *Drosophila* possess a single cluster of dispersed genes, vertebrates usually possess four gene clusters (Grier et al., 2005a; Lappin et al., 2006; Singh and Krumlauf, 2022). Vertebrate and invertebrate Hox gene share a common ancestor, with the four Hox clusters typically found in vertebrate genomes proposed to have arisen during two rounds of duplication from their common ancestor during evolution (Garcia-Fernàndez, 2005).

The mammalian Hox cluster comprises of 39 Hox genes organsied into four clusters on four distinct chromosomes, *Hoxa* (7p15), *Hoxb* (17q21), *Hoxc* (12q13) and *Hoxd* (2q31) (Rice and Licht, 2007). Each Hox cluster consists of 9 – 13 genes and each individual gene can be aligned with different clusters in up to 13 paralog groups, or with genes in the *Drosophilla HOM-C* cluster (Figure 1.1), based on their position in the cluster and sequence similarity (Kessel and Gruss, 1990; Mann et al., 2009). The order of Hox genes in a cluster corresponds to their expression pattern along the AP axis during early development; a distinct feature of Hox genes known as collinearity. Hox genes exhibit spatial, temporal and functional collinearity. Spatial collinearity describes the region in which Hox genes are expressed along the AP axis. 3' Hox genes are expressed at the

anterior end of the AP axis, while 5' Hox genes are expressed more posteriorly. Temporal collinearity refers to the order in which Hox genes are expressed; 3' Hox genes are expressed first and 5' Hox genes are expressed sequentially later (Durston, 2012; Grier et al., 2005a; Kmita and Duboule, 2003; Lappin et al., 2006). For example, in specification of the axial skeleton, Hox3 and Hox4 have critical roles in specifying the cervical vertebrae in the neck, Hox5 – 9 regulate ribcage formation, while Hox11 mediate sacrum and tail formation (Mallo et al., 2010; Wellik, 2007). Posterior Hox genes have well-described roles in limb development; whereby formation of the humerous, radius and ulna, and the digits are regulated by Hox9 – 10, Hox11 and Hox13 genes, respectively (Pineault and Wellik, 2014). As was observed with mutation of bithorax/antennapedia genes in *Drosophilia*, mutation of mammalian Hox genes can result in body malformations, such as synpolydactylyl, the fusion or formation of extra digits, caused by mutations in *HOXD13* (Muragaki et al., 1996).

Hox genes also possess tight self-regulation controls, a feature referred to as functional collinearity or posterior prevalence (Durston, 2012). In this process, the function of more posterior Hox genes is dominant over that of anterior Hox genes. For example, during axial skeletal patterning, Hox6 genes induce rib formation. This activity is inhibited by Hox10 to prevent the formation of extra ribs and permit the formation of the sacrum (Mallo et al., 2010).

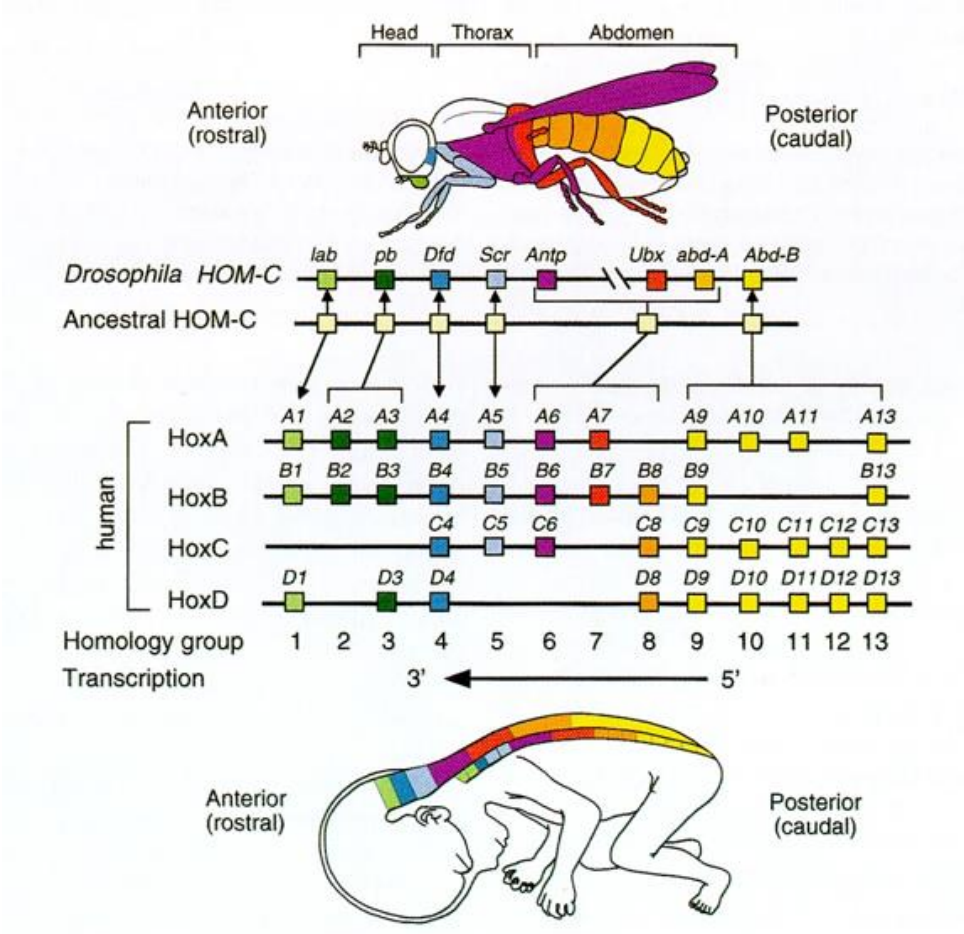


Figure 1.1: Genomic organisation and expression patterns of HOX genes

Schematic depiction of the genomic organisation of the *Drosophila melanogaster* homeotic complex (*HOM-C*) and the four human *HOX* clusters (*HOXA – D*). Homology between human and drosophila is indicated by the colour of individual HOX genes (boxes). Illustrations of *drosophila* and the human foetus highlight spatial and temporal collinearity in the patterning of the body axis. The colours of different anatomical regions correspond to the regulatory HOX gene involved, with 3' HOX genes expressed more anteriorly than 5' genes. Image from Mark et al. (1997).

1.1.2 HOX specificity

All mammalian Hox genes comprise of two exons and a single intron. The Homeobox is always contained within the second exon and encodes the 60 amino acid homeodomain (Grier et al., 2005b). The homeodomain region is highly similar amongst Hox proteins and directly interacts with AT-rich DNA sequences through a helix-turn-helix DNA binding motif (Figure 1.2). Due to their structural similarities, Hox proteins have poor DNA binding selectivity as monomers (Mann et al., 2009; Singh and Krumlauf, 2022). Thus, Hox

proteins rely on interactions with co-factors in order to mediate their specific functional activities *in vivo*. The PBC class of TALE homeodomain-containing transcription factors are well-known Hox co-factors. PBC proteins includes Pbx, Meis and Prep proteins. Pbx proteins (Pbx1 – 4) and Meis (Meis1, 2) represent the vertebrate orthologs of the *Drosophilia* PBC proteins Extradenticle (Exd) and Homothorax (Hth), respectively (Mann et al., 2009; Merabet and Mann, 2016). Hox proteins interact with PBC proteins through a conserved 6 amino acid hexapeptide motif. Interaction with PBC proteins results in alteration of DNA binding preferances and affinities of Hox proteins, thus facilitating target binding specificity (Dard et al., 2018; Singh and Krumlauf, 2022; Slattery et al., 2011).

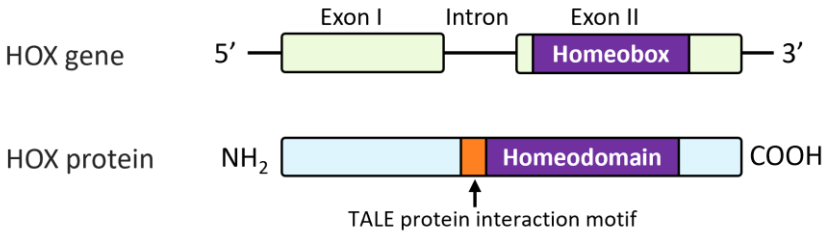


Figure 1.2: HOX gene and protein structure

HOX genes comprise of one intron and two exons. Exon II contains the Homeobox (~ 180 nucleotides), which encodes a 60 amino acid DNA-binding homeodomain. A TALE (three-amino-acid-loop-extension) motif lies upstream of the homeodomain and permits interaction of HOX proteins with co-factors MEIS and PBX. Image adapted from Grier et al. (2005).

1.2 Embryonic haematopoiesis

The haematopoietic system emerges in the early embryo in a series of waves, with each successive wave producing haematopoietic progenitors with increasing complexity. Studies of human embryonic haematopoiesis is limited due to tissue availability; however, embryonic haematopoiesis has been extensively studied in the mouse, and shares several features with the human process (Frame et al., 2013; Ivanovs et al., 2014). The first haematopoietic progenitors are detected in the extra-embryonic yolk sac on embryonic day 7.0 (E7.0) in mice (Palis et al., 1999) and day 16 - 18.5 post-conception (16 - 18.5 dpc) in the human embryo (Ivanovs et al., 2017; Tavian et al., 1999). These progenitors have limited potential and produce predominately primitive erythrocytes (Palis and Yoder, 2001; Palis et al., 1999), hence the first wave is known as primitive

hematopoiesis. The second and third wave of haematopoiesis produce multipotent progenitors in a process described as definitive haematopoiesis (see Figure 1.3 for overview). The two definitive waves differ in the haematopoietic potential and location of progenitors generated: extra-embryonic yolk sac-derived erythro-myeloid progenitors (EMPs) are produced in the first definitive wave, while intra-embryonic-derived haematopoietic stem cells (HSCs) are generated in the second wave. Although EMPs have lympho-myeloid potential, only HSCs can generate all blood lineages. The functional activity of HSCs is assessed by their ability to self-renew and reconstitute the entire blood system following transplantation into irradiated mice (Zovein et al., 2008). HSCs are specified in the dorsal aorta on E10.5 and 27 dpc in mouse (Palis et al., 1999) and human embryos (Tavian et al., 2001), respectively. HSCs subsequently expand and colonise other tissues, including the bone marrow, which remains the primary niche for HSC self-renewal and differentiation throughout the life-span of the organism (Ivanovs et al., 2017).

1.2.1 Primitive haematopoietic wave

The first wave of blood progenitor formation occurs in the blood islands of the yolk sac and functions to provide the developing embryo with oxygen prior to the onset of blood circulation (Frame et al., 2013). At E7.0 – E7.25, primitive erythrocyte, megakaryocyte and macrophage progenitors are derived from mesodermal cells in the proximal yolk sac of the mouse embryo (Palis et al., 1999; Tober et al., 2007). The progenitors expand until E8.25 and are no longer detected in the yolk sac by E9.0 (Palis and Yoder, 2001; Palis et al., 1999). Primitive erythrocytes are the predominant lineage generated and are distinguished from their foetal and adult counterparts by their large size, nucleus and expression of embryonic (βH1) globins (Kaut s et al., 2016). In the human embryo, haematopoiesis begins during the third week of development at 16 - 18.5 dpc (also referred to as Carnegie stage (CS) 7 – 8) with specification of primitive erythrocytes and small numbers of macrophage and megakaryocyte progenitors in the yolk sac (Ivanovs et al., 2017; Tavian et al., 1999).

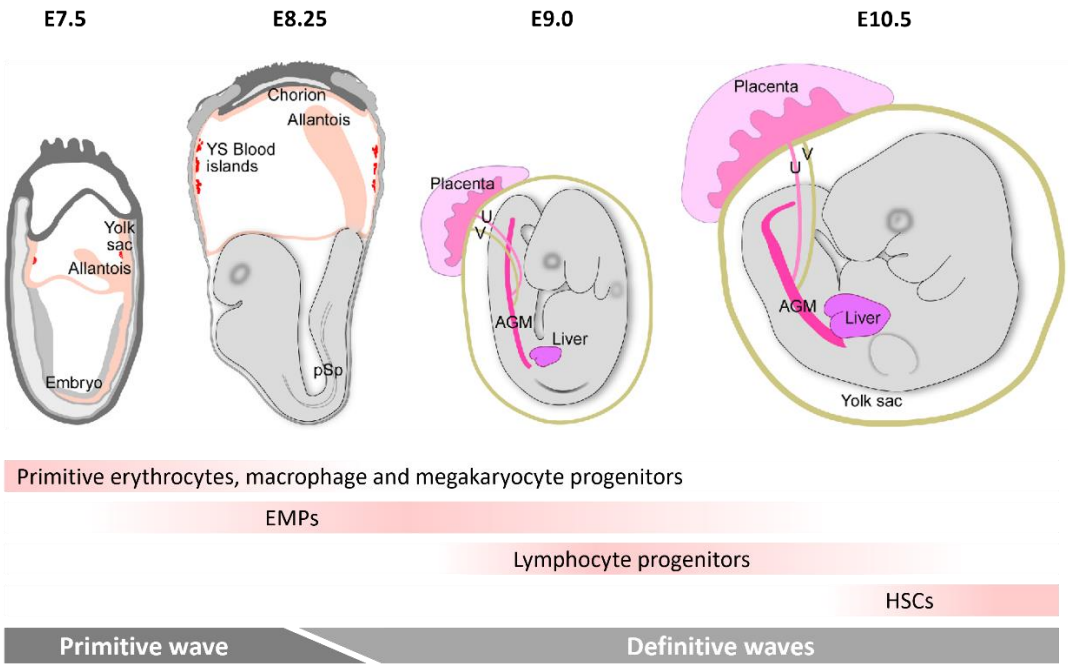


Figure 1.3: Mouse embryonic haematopoiesis

Overview of the times and sites of mouse haematopoietic development in the early embryo. Embryonic haematopoiesis takes place in three waves of haematopoietic activity. The first, primitive, wave takes place between embryonic day (E) 7.0 and E9.0 in the blood islands of the yolk sac (ys) and leads to generation of primitive erythrocytes, macrophage, and megakaryocyte progenitors. The second and third wave of haematopoiesis, known as definitive haematopoiesis, produce erythroid-myeloid progenitors (EMPs) and haematopoietic stem cells (HSCs), respectively. EMPs are first produced in the yolk sac blood islands at E8.25, before migration to the embryo proper at the onset of blood circulation (E8.5). Between E9.0 – 9.5 B and T lymphocyte progenitors are specified and detected in the yolk sac and para-aortic splanchnopleura (pSp). Finally, the third wave begins at E10.5 in the embryo proper with the specification of HSCs in the aorta-gonad-mesonephros (AGM) and umbilical (U) and vitelline (V) arteries. Image adapted from Dzierzak and Speck, 2008.

1.2.2 Definitive haematopoietic waves

The first wave of definitive haematopoiesis overlaps briefly with the primitive wave. Definitive erythrocyte, neutrophil, mast cell and granulocyte-macrophage progenitors (collectively termed EMPs) initially emerge in the mouse yolk sac at E8.25 and expand to the bloodstream and developing foetal liver at the onset of blood circulation at E8.5 (Palis et al., 1999). Unlike primitive erythrocytes, definitive erythroid progenitors express adult (βmajor) globins and undergo enucleation (Kauts et al., 2016). Macrophages are

produced in both waves of definitive haematopoiesis; however, yolk sac-derived macrophages can persist in the brain, liver, epidermis and lung into adulthood and are not specified through a monocyte precursor (Gomez Perdiguero et al., 2015; Schulz et al., 2012). By E9.5, B and T lymphocyte progenitors are detected in the yolk sac and para-aortic splanchnopleura, precursor to the aorta-gonad-mesonephros (AGM) region, and prior to HSC formation (Yoshimoto et al., 2011, 2012). Yolk sac-derived B cell progenitors are restricted to innate B1 and Marginal Zone B cells, with generation of adaptive B2 cell progenitors dependent on HSCs (Lacaud and Kouskoff, 2017; Yoshimoto et al., 2011). In the developing human embryo, yolk sac-derived EMPs colonise the foetal liver at 24 – 25 dpc, after the onset of blood formation (Tavian 1999).

The first HSCs are produced in the embryo proper on E10.5 in the dorsal aorta of the mouse AGM and umbilical and vitelline arteries, which connect the dorsal aorta to the placenta and yolk sac, respectively (De Bruijn et al., 2000; Müller et al., 1994). By E11.5, HSCs are detected in the placenta (Gekas et al., 2005) and in the yolk sac and foetal liver by E12.0 (Kumaravelu et al., 2002). At E16.5, HSCs migrate to the bone marrow, their primary niche throughout adulthood (Lacaud and Kouskoff, 2017). In the human embryo, HSCs first appear at 27 dpc (CS13) in the floor of the ventral wall of the dorsal aorta (Ivanovs et al., 2014). Although HSCs in the mouse embryo are identified in the placenta shortly after their first appearance in the AGM, CD34+ haematopoietic cells are not evident in the human placenta until gestation week 6 and transplantable haematopoietic progenitors are not identified until week 9 (Robin et al., 2009). In the foetal liver, CD34+ cells are evident shortly after HSC specification in the AGM from gestation week 4 (CS13) (Tavian 1999). However, haematopoietic progenitors with *in vivo* reconstitution ability do not appear in the foetal liver until after week 5 - 6 (CS17) (Ivanovs et al., 2011, 2017). Migration of HSCs to the bone marrow signals the end of embryonic haematopoiesis at approx. week 8 of development (CS23) (Ivanovs et al., 2017).

1.2.3 HSC expansion

The AGM is a temporary site for definitive haematopoiesis. In the mouse AGM, low numbers of HSCs are produced, reaching maximum levels by E11.0 and no longer detected by E12.0 (Kumaravelu et al., 2002). Reduction in HSC activity in the AGM coincides with expansion to the yolk sac and foetal liver (Ema and Nakauchi, 2000; Kumaravelu et al., 2002). The foetal liver is a major site of HSC expansion, with the number of HSCs in the foetal liver increasing 50-fold within 24h of their initial colonisation

(Kumaravelu et al., 2002). Similarly, haematopoietic activity in the human AGM is diminished by CS17 and produces low numbers of HSCs (Ivanovs et al., 2011; Tavian et al., 2001). HSC expansion takes place outside the AGM, predominately in the foetal liver, which remains an important site for HSC expansion and differentiation throughout *in utero* development (Ivanovs et al., 2017).

1.2.4 Endothelial origin of definitive haematopoiesis

Haematopoietic and endothelial lineages develop in close proximity. Both are specified from $KDR^{+}CD34^{-}$ mesodermal precursors (Cortes et al., 1999) and share several surface markers. Haematopoietic cells have been shown in several lineage tracing and imaging studies to develop directly from endothelial cells. In the adult mouse bone marrow, $CD34^{+}CD45^{+}$ blood cells were found to be derived from an endothelial precursor expressing VE-cadherin/CD144 (Oberlin et al., 2010). The majority of adult blood was found to be derived from VE-cadherin precursors (Oberlin 2010, Chen 2009). In zebrafish, multipotent haematopoietic progenitors were observed to be specified in an endothelial-to-haematopoietic transition (EHT), whereby cells bud off from the KDR+ endothelium lining the aortic floor and enter into suspension, acquiring haematopoietic surface markers, such as CD41 and CD45, in the process (Bertrand et al., 2010; Kissa and Herbomel, 2010). In the mouse AGM, fate mapping demonstrated VE-cadherin lineages in the AGM have multilineage haematopoietic potential (Zovein et al., 2008). Similar imaging studies in the E10.5 mouse dorsal aorta also demonstrated HSCs emerge from CD31+ endothelial cells during EHT (see Figure 1.4 for illustration) (Boisset et al., 2010) Furthermore, human AGM-derived HSCs are characterised as $CD34^{+}CD144^{+}CD45^{+}$ cells which positively express endothelial surface markers cKit, CD90 and Endoglin, while lack expression of mature HSC markers CD38 and CD45RA (Ivanovs et al., 2014), highlighting their endothelial origin.

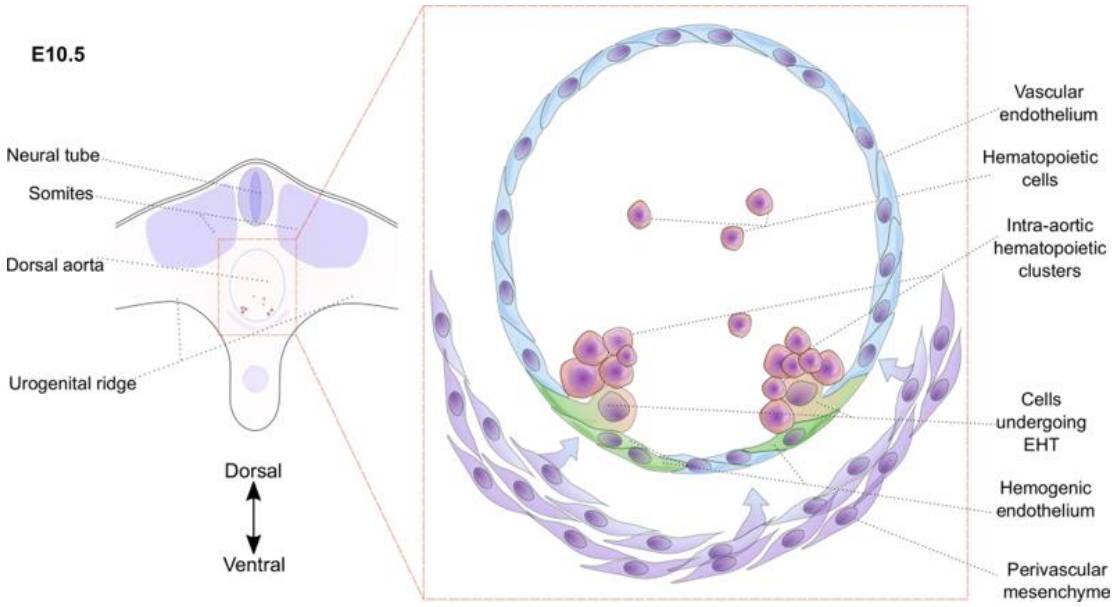


Figure 1.4: EHT in the mouse AGM

Schematic of a cross-section through the mouse E10.5 AGM, with the dorsal aorta magnified (red box) to depict cell types involved in EHT. Haemogenic endothelial cells lining the ventral wall of the dorsal aorta undergo EHT, whereby cells lose their endothelial identity and bud-off into the aortic lumen as haematopoietic cells. Image from Lange et al. (2021).

1.3 Regulation of embryonic haematopoiesis

Haematopoietic lineage specification is regulated by a network of transcription factors, whose coordinated activity drive expression of haematopoietic-specific genes, while repressing alternative lineage fates (Goode et al., 2016). Lineage tracing studies in the mouse embryo and *in vitro* mouse or human pluripotent stem cell (PSC) models of embryonic haematopoiesis have provided insight into the mechanisms by which transcription factors direct haematopoietic commitment and differentiation.

Haematopoietic and endothelial systems develop in close association and share a similar regulatory network during early embryogenesis. At the top of this regulatory network is ETS-factor family transcription factor FLI1. Mice lacking Fli1 die by E12.5 from defects in blood vessel formation and impaired haematopoiesis (Spyropoulos et al., 2000). Fli1 is an upstream regulator of other key haematopoietic regulators, including Tal1/Scl, Lmo2, Flk1/Kdr, Runx1, and Gata2 (Liu et al., 2008; Nottingham et al., 2007). TAL1/SCL and LMO2 are pivotal regulators of early-stage haematopoietic development. In the

mouse embryo, Tal1 is first expressed at the onset of haematopoiesis at E7.5 in both embryonic and extra-embryonic mesoderm (Porcher et al., 1996). Mice lacking Tal1 do not develop a blood system and die between E8.5 and E10.5 (Shivdasanl et al., 1995). Lmo2$^{-/-}$ mice present with a similar phenotype; knockout mice do not undergo yolk sac-derived erythropoiesis and die by E10.5 (Warren et al., 1994). LMO2 is a central component in several DNA-binding transcription factor complexes. In erythroid cells, Lmo2 acts as a bridging molecule in a complex consisting of Tal1, E2A, Gata1 and Ldb1 (Mikkola et al., 2003; Wadman et al., 1997). Lmo2 also forms a complex with Tal1 and Ldb1 in mESC-derived HE. Both Tal1$^{-/-}$ and Lmo2$^{-/-}$ mESCs have reduced ability to generate HE from Flk1+ mesoderm and do not produce CD41+ or CD45+ progenitors (D'Souza et al., 2005; Lancrin et al., 2009; Stanulović et al., 2017). Thus, TAL1 and LMO2 are essential regulators of HE specification from mesoderm. Furthermore, Tal1 is necessary to initiate Lmo2 expression, while Lmo2 is essential to position the TAL1/LMO2/LDB1 complex to key haematopoietic regulator elements at the HE stage of development (Stanulović et al., 2017). Tal1 also functions to repress cardiac fate during mesoderm development (Van Handel et al., 2012). Tal1 was found to bind enhancers in Flk1+ mesoderm, mediating suppression of cardiac transcription factors, such as Gata4, Gata6 and Tbx3 while activating key haematopoietic transcription factors, such as Runx1, Gata1 and cMyb (Org et al., 2015).

Other important regulators of early haematopoiesis include SP1 and ETV2. SP1 activates important signaling pathways and regulators of early haematopoiesis; including BMP and WNT signaling pathways, and CDX and HOX genes (Gilmour et al., 2014). ETV2 is an ETS transcription factor and, like FLI1, regulates both haematopoietic and endothelial lineage specification. During haematopoietic development, ETV2 is only required briefly at the mesoderm stage, where it is a direct regulator of TAL1. Etv2$^{-/-}$ mESCs are unable to generate CD41+ haematopoietic progenitors. This activity can be fully restored by expression of TAL1 yet only partly rescued by expression of either GATA2 or FLI1 (Wareing et al., 2012). Thus, mesoderm differentiation to HE is tightly regulated by TAL1 and LMO2.

RUNX1 and GATA2 are critical regulators of EHT and specification of the definitive program. Lineage tracing studies in the developing mouse embryo have identified Runx1 expression at key sites of definitive haematopoiesis. Runx1 is initially expressed in extra-embryonic mesoderm of the presumptive blood islands at E7.5 and is no longer detected in the yolk sac following the onset of blood circulation at E8.5; thereafter appearing in

endothelial cells lining the vitelline artery and floor of the dorsal aorta and in circulating blood cells (North et al., 1999; Samokhvalov et al., 2007). By E10.5, Runx1 expression is observed in major sites of definitive haematopoiesis: the AGM, vitelline and umbilical arteries, and the foetal liver (North et al., 1999). Runx1$^{-/-}$ embryos die by E12.5 and show major defects in definitive haematopoiesis, while primitive haematopoiesis is unaffected (Cai et al., 2000; Okuda et al., 1996). Runx1 is required for HSC emergence from the HE. Conditional deletion of Runx1 in VE-Cadherin-expressing cells does not impair HE formation but prevents HSC specification (Chen et al., 2009; Swiers et al., 2013). Similar deletion of Runx1 in Vav+ haematopoietic cells in the mouse foetal liver does not impair haematopoietic cell function, demonstrating RUNX1 is critically required for EHT and specification of HSCs but not thereafter (Chen et al., 2009). RUNX1 is also required for specification of both EMP- and HSC-derived definitive progenitors (Tober et al., 2013). Additionally, Runx1 is responsible for reorganisation of the transcriptional network during EHT. During early haematopoietic development, regulatory transcription factors FLI1, TAL1 and C/EBPb initially bind haematopoietic genes, such as Notch1, Pu1/Spi1, Nfe2 and Erg. Following expression of Runx1 in the HE, the pattern of FLI1 and SCL transcription factor binding is rearranged towards that of committed haematopoietic cells (Lichtinger et al., 2012). The expression of several lineage committed genes are primed in the HE, including PU.1 (Lichtinger et al., 2012). PU.1 is essential for myelopoiesis and orchestrates the differentiation of macrophages, granulocytes and lymphocytes (Gupta et al., 2009). Although PU.1 is expressed only after haematopoietic commitment, Runx1 mediates chromatin reorganisation at PU.1 during HE, acting to promote an open chromatin conformation (Hoogenkamp et al., 2009).

Similar to RUNX1, GATA2 expression in the mouse embryo parallels the emergence of definitive haematopoietic cells. Gata2 is first expressed at E7.5 in extra-embryonic mesoderm and subsequently at E10.5 in endothelial cells lining the dorsal aorta and in the foetal liver at E12.5 (Minegishi et al., 1999). Mouse embryos lacking Gata2 die by E10.5 and show similar defects in definitive haematopoiesis as observed in Runx1$^{-/-}$ embryo. However, Gata2$^{-/-}$ embryos also present with severe anaemia, indicative of impaired primitive haematopoiesis (Tsai et al., 1994). GATA2 is essential for EHT and HSC survival. Deletion of Gata2 in VE-cadherin-expressing endothelial cells in the mouse embryo hampers EHT and HSC specification. Additionally, deletion of Gata2 in Vav+ haematopoietic cells in the foetal liver impairs their functional activity in the CFU assay (de Pater et al., 2013). Some haematopoietic progenitors can be produced

independent of GATA2. All mouse AGM-derived HSCs express Gata2, however, prior to HSC emergence, the yolk sac and AGM can generate both Gata2-expressing and non-expressing EMPs (Kaimakis et al., 2016). Therefore, while GATA2 is implicated at all stages of development, it is only essential for intra-embryonic definitive haematopoiesis and HSC survival.

Downregulation of the endothelial program is an essential step in EHT and requires transcriptional repressors GFI1 and GFI1B (Thambyrajah et al., 2016; Wilson et al., 2010). Gfi1/Gfi1b double knock-out mice develop a normal vasculature, yet $Gfi1^{KO}$/$Gfi1b^{KO}$ cells in the yolk sac are unable to migrate to the embryo proper, retaining an endothelial morphology and high expression of endothelial markers Tie2 and Flk1 (Lancrin et al., 2012). During EHT, Gfi1 and Gfi1b recruit histone demethylase LSD1 to endothelial regulator genes, inducing their repression and consequently haematopoietic development (Thambyrajah et al., 2016). GFI1 and GFI1B are regulated by RUNX1 and GATA2. Both Gfi1 and Gfi1b are downstream targets of Runx1 and in mESCs lacking Runx1, Gfi1 or Gfi1b expression can induce downregulation of the endothelial program in HE cells, stimulating EHT and specification of CD41+ haematopoietic progenitors (Lancrin et al., 2012). However, progenitors lack CFU activity, demonstrating while GFI1/GFI1B can direct EHT, RUNX1 is still required to generate progenitors. Gata2 mediates the activity of Gfi1b, while Gfi1 can directly repress Gata2 in mouse primary blood cells (Moignard et al., 2013). More recently, Gata2 was shown to directly regulate Gfib expression in mouse E11.0 AGM suggesting, like Runx1, Gata2 also functions through Gfi1b to suppress endothelial identity during EHT (Koyunlar et al., 2022).

HSCs specifically develop from the major arterial regions of the mouse embryo, that is the dorsal aorta, vitelline and umbilical vessels (De Bruijn et al., 2000). As such proper specification of arterial endothelium is also essential for haematopoiesis. SOX17 is a key regulator of arterial HE development and interacts with Notch signaling pathways to govern arterial fate (Clarke et al., 2013; Irion et al., 2007; Lizama et al., 2015). Knockdown of *Sox17* in early mouse embryos prevents the development of major arteries, alters vascular remodeling, and upregulates venous endothelial markers, while downregulating arterial genes (Corada et al., 2013). Downregulation of SOX17 is also an important process for HSC specification. Continuous SOX17 expression in hPSC-derived HE results in a differential block, and cells retain their haemato-endothelial phenotype (Nakajima-takagi et al., 2013). Sox17 was also shown to specifically prevent EHT by repressing Runx1 and Gata2 (Lizama et al., 2015). Intra-embryonic definitive

haematopoiesis is dependent on Notch signaling (Hadland et al., 2004; Uenishi et al., 2018). In hiPSC-based haematopoietic differentiations, overexpression of SOX17 during HE development enriches the Notch-dependent $DLL4^{+}CXCR4^{+}$ arterial-HE population and upregulates HOXA expression within this population (Jung et al., 2021). Like Notch dependance, upregulation of HOXA genes is a defining feature of intra-embryonic definitive haematopoiesis and HOXA genes are important regulators of this process.

1.4 HOX genes in haematopoiesis

HOX genes have well-established roles in HSC self-renewal, differentiation, and lineage commitment. Regulatory roles for genes of the HOXA, HOXB and HOXC clusters have all been described in haematopoiesis (Alharbi et al., 2013; Bhatlekar et al., 2018). HOX genes, in particular HOXA, are expressed in human primitive haematopoietic stem and progenitor cells and are downregulated as differentiation and lineage commitment proceeds (Pineault et al., 2002; Sauvageau et al., 1994) Despite the structural similarities shared amongst the four HOX clusters, HOXD genes are not expressed in haematopoiesis (Alharbi et al., 2013). In further support of their central role in haematopoiesis, dysregulated HOX expression is a feature of leukaemia and a prognosis marker for some AMLs (Alharbi et al., 2013). In a less well understood process, HOX genes also function in the specification and emergence of HSCs during embryonic haematopoiesis. Human PSC-based *in vitro* models of embryonic haematopoiesis have implicated HOX genes as essential factors in derivation of HSCs.

1.4.1 HOX in HSC maintenance and differentiation

HOX gain- and loss-of-function studies in mouse and human model systems have established clear roles for HOX genes in HSC self-renewal, maintenance, and lineage commitment. Taking this approach, roles for several HOXC genes in erythroid differentiation were identified. HOXC3, HOXC4 and HOXC8 were shown to promote the expansion of early erythroid progenitors (Daga et al., 2000; Takeshita et al., 1993), while Hoxc8 was also implicated in granulocyte-macrophage differentiation *in vitro* (Shimamoto et al., 1999). However, genes of the HOXA and HOXB clusters are the most prevalent in haematopoiesis.

HOXA genes are predominately expressed in myeloid lineages (Alharbi et al., 2013). Ectopic expression of Hoxa5 promotes mouse HSCs to re-enter the cell cycle and undergo myeloid differentiation (Yang et al., 2015). Similarly, HOXA5 overexpression in

human cord blood-derived CD34+ cells promotes myeloid and suppresses erythroid differentiation (Crooks et al., 1999; Fuller et al., 1999). HOXA10 has also been observed to enhance myeloid differentiation when overexpressed in cord blood-derived CD34+ cells (Buske et al., 2001). HOXA10 has dose-dependent effects on haematopoiesis. In a tetracycline-inducible HOXA10 expression mouse model, induction of high HOXA10 expression in the HSC-enriched $Lin^{-}Sca1^{+}cKit^{+}$ (LSK) population blocked erythropoiesis and impaired lymphoid differentiation, while low HOXA10 promoted expansion of the LSK population and conferred engraftment ability *in vivo* (Magnusson et al., 2007a).

HOXA9 is the most highly expressed HOX gene in the HSC compartment and has the most profound effect on hematopoiesis in both overexpression and knock-out studies (Alharbi et al., 2013). In mouse HSCs, Hoxa9 overexpression suppresses B lymphopoiesis and enhances myelopoiesis and HSC repopulation ability in transplant assays (Thorsteinsdottir et al., 2002). Conversely, $HOXA9^{-/-}$ mouse chimeras have reduced numbers of granulocytes, B and T lymphocytes and committed myeloid and erythroid progenitors in their peripheral blood (Lawrence et al., 1997). $Hoxa9^{-/-}$ bone marrow cells exhibit reduced repopulation activity in competitive transplantation assays, a defect which is not observed with $Hoxa10^{-/-}$ cells (Lawrence et al., 2005) despite HOXA10's reported critical role in HSC engraftment (Magnusson et al., 2007a). The haematopoietic defects observed in HOXA9 knock-out studies are particularly striking given, due to the functional redundancy amongst HOX genes, HOX gene knockout in general has a minimal effect on haematopoiesis. $Hoxa^{+/-}$ mice survive into adulthood and possess no major haematopoietic defect despite an approx. 50% reduction in Hoxa expression in bone marrow-derived cells. However, a mild defect in maturation of B cell progenitors and bone marrow reconstitution is observed in older mice (Lebert-Ghali et al., 2010). Conditional homozygous deletion of Hoxa in adult mouse HSPCs similarly reduced B cell progenitors and $Hoxa^{-/-}$ HSPCs completely lacked repopulation activity in secondary transplantation assays; a defect which could be recovered by overexpression of Hoxa9 (Lebert-Ghali et al., 2016, 2017). These loss-of-function studies demonstrate the HOXA cluster is indispensable for early B cell development but also that HOXA9 has a critical role in conferring repopulation activities to HSPCs.

HOXB genes have also been implicated to have an important role in HSC self-renewal and maintenance. Overexpression of HOXB6 in mouse bone marrow promotes expansion of HSCs and myeloid precursors both *in vitro* and *in vivo*, at the expense of erythropoiesis and lymphopoiesis (Fischbach et al., 2005). Similarly, HOXB3

overexpression in adult mouse bone marrow blocks T lymphopoiesis in the thymus, impairs B cell development and promotes myelopoiesis (Sauvageau et al., 1997). Interestingly, Hoxb3$^{-/-}$ mutant mice are normal after birth, only developing impairments in B lymphopoiesis after 3 months of ages, suggesting compensatory mechanism by other Hox genes prevent more severe defects in younger mice (Ko et al., 2007). Overexpression of HOXB4 in both adult mouse bone marrow and human cord blood-derived progenitors enhances the expansion of primitive haematopoietic progenitors, without negatively impacting blood lineage development, as is observed with HOXB6, HOXA10 and HOXA9 overexpression (Antonchuk et al., 2001; Buske et al., 2002). Furthermore, in mESCs differentiation to haematopoietic cells, overexpression of Hoxb4 during the haematopoietic specification stage results in progenitors which possessed long term multilineage reconstitution abilities in primary and secondary transplant assays (Kyba et al., 2002). Despite the pivotal role of HOXB4 in HSC self-renewal proposed in overexpression studies, Hoxb4$^{-/-}$ and Hoxb1-b9$^{-/-}$ foetal liver cells retain their repopulation and reconstitution potential following transplantation (Bijl et al., 2006). While bone marrow cells derived from compound Hoxa9/b3/b4 null mice exhibited a repopulation defect in a competitive transplantation assay, Hoxa9-/- mutants possessed a similar defect (Magnusson et al., 2007b), suggesting Hoxa9 is responsible for the functional impairment observed.

1.4.2 HOX in malignant haematopoiesis

HOX genes are frequently dysregulated in leukaemia. Aberrant HOXA expression is associated with MLL translocations, in particular, high HOXA expression is a characteristic of all MLL-rearranged leukaemias (Armstrong et al., 2002). HOXB genes are upregulated in AML patient samples with FLT-ITD mutations (Lindblad et al., 2015), while AML patients with NPM1c mutations exhibit upregulation of genes from both HOXA and HOXB clusters (Alcalay et al., 2005). HOXA and HOXB gene expression patterns, chromatin accessibility and methylation profiles are strikingly similar between CD34+ cells from healthy donors and AML patients and HOXA and HOXB genes have been hypothesized to be regulated in a similar manner in normal and malignant haematopoietic cells (Spencer et al., 2015).

In culture, mouse bone marrow cells can be immortalised by overexpression of HOXA genes, which, when overexpressed together with HOX co-factor MEIS1, can drive leukaemic transformation upon transplantation to immunodeficient mice (Bach et al., 2010). Several HOX genes can also directly induce leukaemia through formation of

fusion proteins with nucleoporin gene NUP98, including HOXA9 (Nakamura et al., 1996), HOXA11 (Suzuki et al., 2002), HOXC13 (Panagopoulos et al., 2003) and HOXD13 (Raza-Egilmez et al., 1998). Balanced HOX expression is essential for normal haematopoietic development. Overexpression of HOXB3, HOXB6, HOXA9 or HOXA10 results in dysregulated haematopoiesis and leukaemia within 3 – 10 months (Fischbach et al., 2005; Sauvageau et al., 1997; Thorsteinsdottir et al., 2002). Consistent with its central role in HSC self-renewal, HOXA9 is a particularly potent oncogene. HOXA9 overexpression is associated with poor clinical outcome and prognosis in AML and ALL patients (Golub et al., 1999). While Hoxa9 alone cannot induce leukaemia in mice, co—expression with Meis1 is sufficient to drive leukaemic transformation (Kroon et al., 1998).

1.4.3 HOX in HSC specification and emergence

Human and mouse PSC haematopoietic differentiation systems aim to recapitulate *in vivo* embryonic haematopoiesis and reproduce the stages and cell types generated during initial HSC specification in the AGM. Human cord blood-, foetal liver- and bone marrow-derived HSCs possess multilineage reconstitution potential, however, generation of PSC-derived HSCs with these capabilities remains a challenge. While haematopoietic progenitors with multilineage reconstitution abilities have been generated from hPSCs, genetic modification of differentiating cells or programming with transcription factors was required (Sugimura et al., 2017; Tan et al., 2018). Human PSC-derived progenitors generated without genetic modification fail to achieve multilineage reconstitution upon mouse bone marrow transplantation. In the past decade numerous RNA-sequencing studies of *in vivo*-derived human HSCs have determined HOXA upregulation to be a key distinguishing feature of long-term repopulating HSCs (Dou et al., 2016; Ivanovs et al., 2017; Ng et al., 2016; Sugimura et al., 2017). Lack of engraftment in hPSC-derived progenitors has been attributed to deficiencies in HOXA signaling (Ivanovs et al., 2017). Upregulation of *HOXA* during hematopoietic differentiation has become synonymous with definitive potential and is used to distinguish between intra- and extra-embryonic pathways (Bruveris et al., 2020; Dege et al., 2020; Dou et al., 2016; Ng et al., 2016).

The precise mechanism and cell type in which HOXA genes influence haematopoiesis remains unknown. Study of HOX function specifically during embryonic haematopoiesis *in vivo* is challenging due to the essential roles HOX play in patterning of the body plan. Deletion of the Hoxa cluster in mice results in early embryonic lethality (Lebert-Ghali et al., 2010). However, Hoxa deletion has been studied specifically in mesoderm-derived

lineages using a conditional deletion approach. The haematopoietic compartment of the resultant mice was reduced in size, and mice also suffered skeletal, cardiac and urogenital defects (Di-Poï et al., 2010). Differentiation of foetal liver cells isolated from mutant embryos had reduced reconstitution abilities in transplantation assays and HSC differentiation was partially blocked (Di-Poï et al., 2010). Conversely, foetal liver cells isolated from Hoxb1-b9$^{-/-}$ mice show no defects in reconstitution in transplantation assays (Bijl et al., 2006). Compensation from other HOX genes was likely responsible for the lack of a defect in Hoxb1-b9$^{-/-}$ foetal liver cells, however, the reduced reconstitution ability of foetal liver cells observed following Hoxa deletion in mesoderm lineages (Di-Poï et al., 2010) suggests HOXA genes have a prominent role in HSC specification, which cannot be overcome by compensation.

Mouse ESC differentiation studies have demonstrated a role for Hoxb4 in enhancement of haematopoietic progenitor formation. In particular, Hoxb4 overexpression was found to support the formation of the haematopoietic niche (Jackson et al., 2012) and specification of haemato-endothelial precursors (Teichweyde et al., 2018). Ectopic expression of Hoxb4 during mESC haematopoietic differentiation was sufficient to endow haematopoietic progenitors with multilineage repopulation abilities (Kyba et al., 2002). However, this effect was not translated to the human system. Although overexpression of HOXB4 in hESC differentiation enhanced the production of CD34^{+}CD43^{+} progenitors, further differentiation to erythroid cells resulted in immature cells which retained their embryonic globin profile (Jackson et al., 2016). Similarly, ectopic expression of HOXB4 in hESC-derived haematopoietic progenitors enhanced progenitor cell expansion *in vitro* but was unable to confer reconstitution ability following transplantation of HOXB4-transduced progenitors to immunodeficient mice (Wang et al., 2005).

HOXA genes have been shown to influence haematopoietic progenitor formation at mesoderm, HE and EHT stages of hPSC-based differentiation. Manipulation of Wnt and Activin signaling pathways during mesoderm patterning in hESCs upregulates CDX genes resulting in sustained *HOXA* expression and generation of SOX17^{+}RUNX1c^{+} endothelium reminiscent of human AGM (Ng et al., 2016). The Wnt/CDX/HOXA signaling axis plays a crucial role in patterning of definitive haemogenic mesoderm, where enforced expression of CDX4 in differentiating hPSCs specifically promotes *HOXA* upregulation and formation of KDR^{+}CD235a^{-} mesoderm and the definitive lineage (Creamer et al., 2017). In hPSC-derived HE, high *HOXA* expression is associated with specification of intra-embryonic haematopoiesis, while low or no *HOXA* expression is

characteristic of extra-embryonic haematopoiesis (Dege et al., 2020). Similarly, high *HOXA* expression was found preferentially in $DLL4^{+}CXCR4^{+}$ arterial-HE, an effect mediated by SOX17 interaction with the CDX2 promoter (Jung et al., 2021). Retinoic acid signaling is a positive regulator of HSC development (Chanda et al., 2013) and activation of retinoic acid signaling in hPSC-derived haemato-vascular cells resulted in upregulation of transcriptional programs associated with definitive HE as well as medial HOXA genes (*HOXA5*, *HOXA6*, *HOXA9* and *HOXA10*) (Dou et al., 2016). While most HOXA genes promote HSC generation, HOXA3 was found to be a negative regulator of haematopoiesis. In mESC differentiation, Hoxa3 overexpression directed the bipotent haemato-endothelial population towards an endothelial fate (Iacovino et al., 2011). Hoxa3 acts in part by suppressing Runx1 and Notch signaling, thus preventing EHT (Iacovino et al., 2011; Sanghez et al., 2017).

Given the prominent roles for HOXA genes in specification of intra-extraembryonic haematopoiesis, a model in which HOXA expression is at the apex of haematopoietic fate divergence has been proposed (Figure 1.5). Extra-embryonic and intra-embryonic haematopoietic fates are proposed to diverge during mesoderm patterning in hPSC models of haematopoiesis. Addition of Wnt agonist CHIR and activin antagonist SB during mesoderm patterning direct cells towards an intra-embryonic haematopoietic fate which is characterised by upregulation of HOXA genes. Conversely, absence of SB/CHIR and lack of *HOXA* is associated with extra-embryonic haematopoietic fates (Dege et al., 2020; Ivanovs et al., 2017).

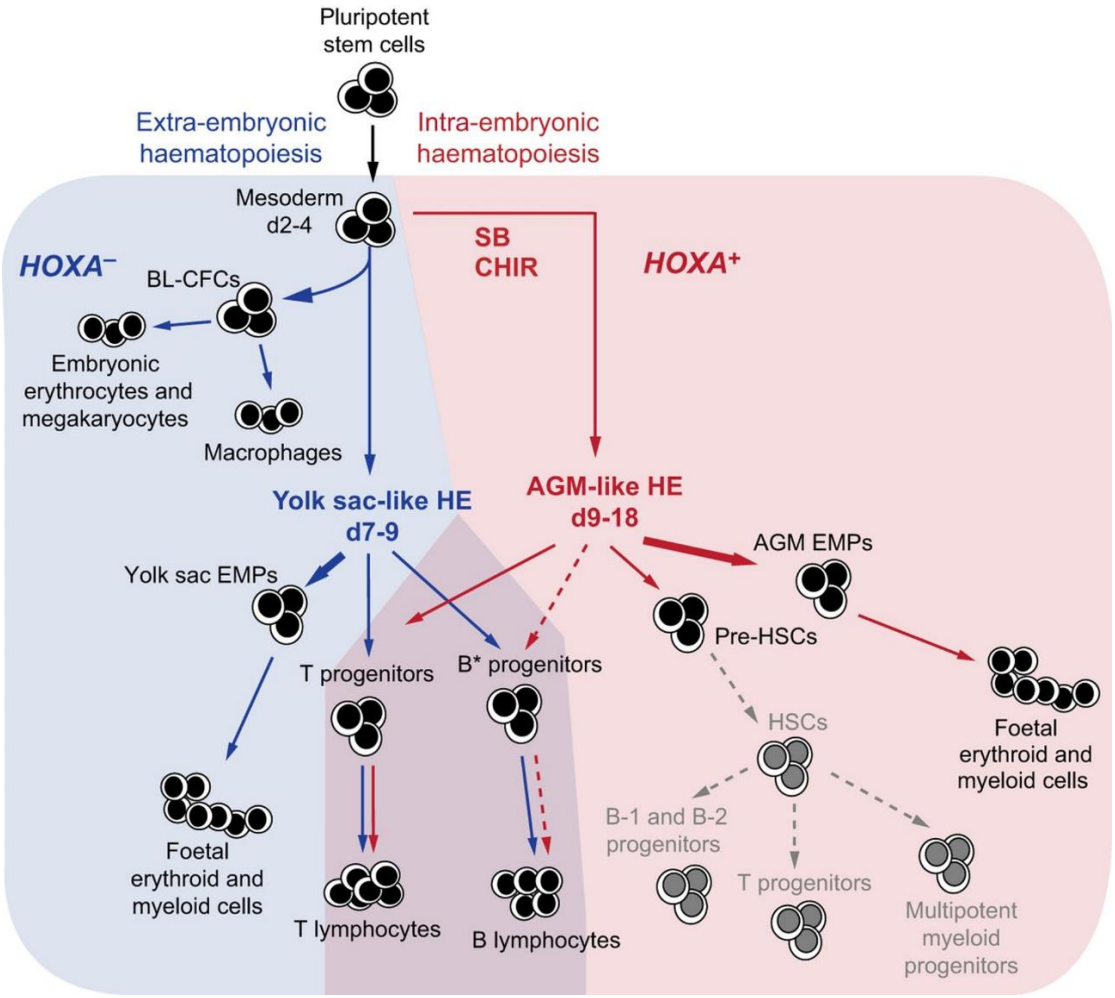


Figure 1.5: Proposed role for HOXA in a hPSC model of haematopoiesis

Extra- and intra-embryonic haematopoietic programs diverge during mesoderm patterning and can be distinguished by HOXA expression. Wnt agonists (CHIR) and activin antagonists (SB) directs cells to AGM-like HE in intra-embryonic haematopoiesis, which is characterised by upregulation of HOXA genes. In the absence of SB/CHIR, HOXA are not upregulated and primitive haematopoiesis and yolk sac-like HE is generated, recapitulating extra-embryonic haematopoiesis. Image from Ivanovs et al. (2017).

Several HOXA genes also appear to have a role in specification of hPSC-derived haematopoietic progenitors with multilineage differentiation potential. Transfection of hPSC-derived HE with seven transcription factors (*ERG, HOXA5, HOXA9, HOXA10, LCOR, RUNX1* and *SPII*) followed by transplantation to mouse bone marrow conferred long-term multilineage engraftment in primary and secondary recipients. Furthermore, multilineage engraftment was compromised if either *ERG, HOXA5, HOXA9, HOXA10* or

RUNX1 were excluded (Sugimura et al., 2017). ERG is a downstream target of HOXA9, while HOXA9 interacts with enhancers in SPII (Huang et al., 2012) suggesting functional co-operation between HOXA9 and other HSC-specific factors plays a role in specification of HSCs with multilineage engraftment potential. In another study, overexpression of five transcription factors (*HOXA9, ERG, RORA, SOX4* and *MYB)* in hPSC-derived $CD34^{+}CD45^{+}$ haematopoietic progenitors facilitated short-term engraftment upon transplantation to mice. In particular, *in vitro* CFU assays demonstrated HOXA9 and ERG were essential to confer self-renewal and multilineage potential to the progenitor cells (Doulatov et al., 2013).

1.5 Hypothesis and Aims

HOXA expression is associated with mesoderm, HE and haematopoietic progenitors of the definitive intra-embryonic haematopoietic program. Amongst the HOXA genes, HOXA9 is consistently upregulated in these cell types, suggesting a prominent role for HOXA9 in specification of intra-embryonic haematopoiesis. Furthermore, HOXA9 was identified as the most differentially expressed gene in a comparison of cord blood-derived CD34+ cells and hESC-derived CD34+ cells (Wang et al., 2005). When HOXA9 expression is tracked during hESC differentiation to haematopoietic progenitors, HOXA9 expression follows the pattern of haematopoietic commitment (Ramos-Mejía et al., 2014). Continuous overexpression of HOXA9 throughout haematopoietic differentiation, either starting in hESCs (Ramos-Mejía et al., 2014) or during mesoderm or HE formation (Zeng et al., 2021) promotes haematopoiesis but it is unclear if HOXA9 specifically acts to promotes intra- or extra-embryonic haematopoiesis. We hypothesise HOXA9 functions specifically in promotion of intra-embryonic haematopoiesis.

As discussed, HOXA expression has been observed at several stages of haematopoietic development and is proposed to mark the intra-embryonic haematopoietic program. However, it remains unknown as to why HOXA upregulation achieved during hPSC-directed haematopoietic differentiation is unable to be sustained in hPSC-derived HSPCs. The aim of this thesis was to specifically study the function of HOXA9 at different stages of haematopoietic development in an hPSC model of haematopoiesis. In doing so, we aimed to generate an inducible HOXA9 overexpression system, in which the impact of HOXA9 overexpression on haematopoietic development and haematopoietic progenitor formation could be studied at distinct stages of differentiation. Specifically, we aimed to:

1. Utilise CRISPR/Cas9 technology and Tet-On expression systems to generate an inducible HOXA9 expression system in hiPSC that will facilitate upregulation of HOXA9 at distinct stages of differentiation.
2. Model embryonic haematopoiesis through generation of an *in vitro* haematopoietic differentiation system
3. Investigate the impact transient HOXA9 overexpression during mesoderm patterning, HE specification or EHT has on haematopoietic lineage development and haematopoietic progenitor formation and function.

Chapter 2: Materials and Methods

2.1 Cell Culture

All cells were maintained at 37°C in a humidified atmosphere with 5% CO_2.

2.1.1 Cell lines

Cell lines used in this thesis included RPAT, a human iPSC line (Hammad et al., 2016), and genetically modified human ESC lines, the inducible-Shox2a-NanoLuc reporter H9 line and the dual reporter *SOX17*$^{mCherry/w}$ and *RUNX1c*$^{GFP/w}$ H9 line (Ng et al., 2016). An inducible HOXA9-mScarlet reporter cell line was generated as part of this thesis in RPAT iPSCs.

2.1.2 Human induced pluripotent stem cell maintenance and passage

Human iPSCs were cultured and maintained in Essential 8 (E8) media (Life Technologies A1517001) on Matrigel-coated tissue culture plastic. Coating was performed by diluting Matrigel (Corning #354234) 1:100 in DMEM and adding approx. 0.2mL/cm^2 per T25/well-plate. Matrigel was allowed to polymerise at room temperature for 45 min or overnight at 4°C. Human iPSCs were passaged every 72 – 96h using either TrypLE, EDTA or Gentle Cell Dissociation Reagent (GCDR).

2.1.2.1 TrypLE passaging

For single cell passaging and setting up cardiac differentiations, TrypLE (Thermo Fisher Scientific #12563029) was used. Cells were washed with PBS and dissociated by adding 2 mL TrypLE per 25 cm^2 T25. Cells were incubated for 3 – 5 min at 37°C, TrypLE was aspirated, and the flask was bashed 4 – 5 times to detach the hiPSCs. Cells were counted manually using a haemocytometer and seeded at 20,000 cells/cm^2 in E8 with the addition of 10 µM Y-27632 Rock Inhibitor (RI, Tocris #1254).

2.1.2.2 EDTA passaging

For gentle dissociation and setting up spin embryoid body differentiations, EDTA (Sigma-Aldrich #03690) was used. Cells were dissociated as described for TrypLE; except for incubation times and seeding. Cells were incubated for 5 – 7 mins in 0.5 mM EDTA at 37°C and seeded in ratios (usually 1:10 – 1:15 for passaging from T25 to T25).

2.1.2.3 GCDR passaging

For clump passaging and setting up haematopoietic differentiations using the STEMdiff™ Hematopoietic Kit, GCDR (STEMCELL Technologies #07174) was used. Cells were routinely maintained in 6-well plates and dissociated once 70 – 80% confluent. Cells were washed with PBS and incubated in 1 mL GCDR for 7 – 10 mins at room temperature. GCDR was aspirated and 1 mL E8 was added to the well. A cell scraper (Sigma CLS3008) was used to lift hiPSC colonies from the plate surface. Colonies were pipetted up-and-down 1 – 3 times depending on the size of the aggregates dissociated. Colonies were seeded in ratios (1:8 – 1:12).

2.1.3 Cryopreservation

Approx. 1 x 10^6 cells were centrifuged at 160 *x g* for 4 mins and resuspended in 0.5 mL FBS (Fisher #11550-356) containing 10% DMSO (Sigma #C1646). Cells were transferred to cryovials and placed in Mr Frosty freezing containers and stored at -80^{o}C for 24 hours. Cryovials were transferred to liquid nitrogen for long-term storage.

2.1.4 Karyotyping

Human iPSC cultures in log phase growth were treated with 100 ng/mL Karyomax Colecmid (Life Technologies #25212-012) and incubated at 37^{o}C for 1 hour. The media and Karyomax were collected in a 50 mL falcon tube, and cells were dissociated by incubation with 1 mL TrypLE for 3 – 5 mins. TrypLE was inactivated by addition of 4mL E8 and dissociated cells were transferred to the 50 mL tube. The cell suspension was centrifuged at 160 *x g* for 4 mins. The supernatant was aspirated, and the cell pellet was resuspended in 0.3 mL E8 medium. 8 mL of fresh 0.6% hypotonic solution (300 mg Sodium Citrate (Fisher #BP327-500) dissolved in 50 mL distilled water) was added dropwise to the cell suspension whilst vortexing. Cells were incubated for 20 min at room temperature to swell. The cell suspension was centrifuged at 300 *x g* for 5 min and the majority of the supernatant was discarded leaving approx. 0.5 mL supernatant, which was used to resuspend the pellet with a P1000 pipette. 8 mL of fresh 5:1 methanol (Fisher #10284580): glacial acetic acid (Fisher #10304980) fix solution was added dropwise to the cell suspension whilst vortexing. The cell suspension was centrifuged at 300 *x g* for 5 min and the supernatant was discarded. The addition of the fix solution was repeated, and the final cell pellet was resuspended in 0.5 mL of fix solution. Slide preparation and G-Banding was performed by Nigel Smith at the Department of Cytogenetics, City Hospital, Nottingham University NHS Trust.

2.1.5 Nucleofection of live cells

Transfection of hiPSCs was achieved using the Amaxa™ 4D-Nucleofector™. One million hiPSCs were resuspended in 100 µL P3 buffer (Lonza V4XP-3024) in the presence of 1 – 3 µg DNA and transferred to a Nucleocuvette. The Nucleocuvette was placed the Amaxa™ nucleofector and program CA 137 (optimised for hiPSC transfection) was run. Cells were incubated for 5 min at 37°C before being flushed with 2 mL E8 + 10 µM RI and added to a single well of a Matrigel-coated 6-well plate.

2.1.6 Cardiac differentiation and dissociation

Cardiac differentiation of hiPSCs was performed using an in-house monolayer protocol (Mosqueira et al., 2018). On day -3, hiPSCs were dissociated with TrypLE and seeded in a Matrigel-coated 12-well plate at 0.15 x 10^6 cells/well in E8 media + 10 µM RI. Media was changed to fresh E8 on day -2 and -1. Approx. 16 hours prior to day 0, a pre-conditioning step was performed by addition of a Matrigel overlay (1:100 dilution of Matrigel in StemPro™-34 Serum-Free Medium (SP34, Gibco #10639011)) supplemented with 1 ng/mL BMP4 (R&D Systems #314-BP). On day 0, media was changed to SP34 supplemented with 10 ng/mL BMP4 and 8 ng/mL Activin A (Life Technologies PHC9561). After 48 hours (day 2), media was changed to RPMI supplemented with B27 minus insulin (Life Technologies #A1895601),10 µM KY02111 (R&D Systems #4731) and 10 µM XAV939 (R&D Systems #3748). On day 4, 10 µM KY02111 and 10 µM XAV939 were added with RPMI supplemented with B27 plus insulin (B27+Ins, Life Technologies #0080085-SA). Thereafter, media was changed every 2 days with RPMI supplemented with B27+Ins until day 15 of differentiation, where hiPSC-derived cardiomyocytes (hiPSC-CMs) were dissociated.

To dissociate hiPSC-derived CMs, cells were washed twice in Ca^{2+}-free Hank's Balanced Salt Solution (HBSS, Life Technologies #14175095) and incubated for 3 – 4 hours at 37°C with 200 U/mL Collagenase II (Life Technologies #17101015) in HBSS, supplemented with 1 mM HEPES pH 7.4 (Sigma #H4034), 10 µM Y-27632 and 30 µM N-Benzyl-p-toluenesulfonamide (BTS, TCI Chemicals #B3082). Dissociated cardiomyocytes were collected and washed with RPMI supplemented with 24 µg/mL Deoxyribonuclease II (DNAse II, Sigma #D8764) and centrifuged at 100 *x g* for 15 mins. Cells were resuspended in RPMI and passed through a 100 µm cell strainer (VWR #89508-840) to remove large clumps. Cell counts were taken using an automated CEDEX counter and centrifuged again as described above. Cells were resuspended in

RPMI supplemented with B27+Ins and 10 µM RI and seeded at 30,000 cells/well in a Matrigel-coated 96-well plate for immunostaining.

2.1.7 Haematopoietic differentiation using STEMdiff™ Hematopoietic Kit

Differentiation was performed using STEMdiff™ Hematopoietic Kit according to the manufacturer's instructions (STEMCELL Technologies #05310). Briefly, hiPSCs were dissociated as cell aggregates with GCDR for 7 – 10 minutes at room temperature, followed by scraping. Cell aggregates (50 – 200 µm in diameter) were seeded at different ratio densities in E8 media on Matrigel-coated 12-well plates. The following day, only wells which contained 16 – 40 colonies ›50 µm in diameter were selected to continue with differentiation and media was changed to Media A (day 0 of differentiation). On day 2, a half Media A change was performed. On day 3, media was changed to Media B and half Media B changes were performed on days 5, 7 and 10. On day 12 suspension cells were harvested for downstream analysis.

2.1.8 Spin Embryoid Body haematopoietic differentiation

Haematopoietic differentiation of hiPSCs was performed using the spin embryoid body (EB) method (Nafria et al., 2020a; Ng et al., 2008, 2016). Three T25s of hiPSCs were used to set-up 10 low attachment 96-well round-bottom plates (ThermoFisher #268200), at 60 EBs/plate. On day -1 of differentiation, three 80 – 90% confluent T25s were passaged 1:1 using EDTA into fresh Matrigel-coated T25s. On day 0, prior to initiation of differentiation, media was changed to fresh E8. After at least 2 hours, hiPSCs were dissociated using EDTA and resuspended in STEMdiff™ APEL™2 medium (APEL2, STEMCELL Technologies #05275) supplemented with 5 ng/mL BMP4, 7.5 ng/mL Activin A, 25 ng/mL vascular endothelial growth factor (VEGF, PeproTech #100-20), 25 ng/mL stem cell factor (SCF, PeproTech #300-07), 10 ng/mL fibroblast growth factor 2 (FGF2, PeproTech #100-18B) and 10 µM RI (Stage I media). The cell suspension was added at 80 µL/well to the 60 inner wells of the 96-well plates, while 80 µL of sterile ddH_2O was added to the 36 outer wells. EBs were formed by centrifugation at 400 *x g* for 2 min at 4ºC. At approx. 40 hours post-seeding, 20 µL of APEL2 supplemented with 4 µM SB431542 (SB, Selleckchem #S1067), 3 µM CHIR99021 (CHIR) and/or 1 µg/mL doxycycline (DOX) was added to EBs where indicated. On day 4 of differentiation, 80 µL of Stage I media removed from each well and 100 µL of APEL2 supplemented with 5 ng/mL BMP4, 50 ng/mL VEGF, 50 ng/mL SCF, 10 ng/mL FGF2 and 30 ng/mL insulin-like growth factor 2 (IGF2, PeproTech #100-12) (Stage II media) was added per well. Mesoderm formation was assessed on day 4 via flow cytometry, qPCR and RNA-seq.

Several EBs (see Table 2.1 for number of EBs required for each assay) were collected and washed in PBS before incubation in 1 mL Accutase (Sigma #A6964) for 20 min at 37°C to dissociate the EBs. Accutase was neutralised with 3 mL APEL2, and the cell suspension was further diluted up to 10 mL in PBS. Dissociated cells were centrifuged at 300 *x g* for 5 min and pellets were frozen at -80°C for qPCR and RNA-seq or stained for flow cytometry analysis (see 2.2.1).

Table 2.1: Number of EBs required for downstream assays

Downstream assay	Number of EBs required	
	Day 4	Day 7
Flow cytometry	30	30
qPCR	20	10
RNA-seq	30	N/A

On day 7 of differentiation, EBs were transferred to adherent culture. Matrigel-coated 6-well plates were washed in PBS and approx. 20 EBs were added to each well. Plates were incubated for at least 24 hours undisturbed to allow EBs to attach to the surface. Haemato-vascular mesoderm formation was assessed on day 7. EBs were harvested and dissociated as described for day 4, however, due to their larger size, day 7 EBs were dissociated in Accutase for 1.25 – 1.5 hours at 37°C (see Table 2.1 for number of EBs required). On day 9 of differentiation, Stage II media was removed and 2 mL/well of APEL2 supplemented with 50 ng/mL VEGF, 50 ng/mL SCF, 50 ng/mL interleukin 3 (IL30, PeproTech, #200-03), 25 ng/mL interleukin 6 (IL6, PeproTech #200-06), 25 ng/mL thrombopoietin (TPO, PeproTech #300-18), 25 ng/mL fms-like tyrosine kinase 3 (FLT3, PeproTech #300-19), 20 ng/mL IGF2 and 10 ng/mL FGF2 (Stage III media) were added. On day 11, each well was topped up with 1 mL Stage III media. On day 14, suspension cells were harvested and counted using a haemocytometer. Suspension cells were pelleted and taken for downstream analysis to assess their hematopoietic potential (see Table 2.2 for number of cells required for each assay).

Table 2.2: Number of day 14/18 cells required for downstream assays

Downstream assay	Number of suspension cells required
Flow cytometry	minimum 0.1 x 10^6 cells/sample
Colony forming unit assay	0.135 x 10^6 cells/triplicate
qPCR	minimum 0.3 x 10^6 cells/sample
RNA-seq	0.6 x 10^6 cells

Adherent EB cultures continue to produce suspension cells, therefore fresh Stage III media was added to cultures on day 14 and topped up again on day 16. On day 18 of differentiation, newly produced suspension cells were harvested, counted, and analysed via downstream assays as described above for day 14 suspension cells.

2.1.9 Colony forming unit assay

Haematopoietic potential of suspension cells from day 14 and 18 cultures were assess by culture in methylcellulose. A 100 mL bottle of MethoCult™ Serum-free H4646 (STEMCELL Technologies #04646) was thawed overnight at 4°C, shaken to ensure homogeneity, and dispended into bijou tubes at 1.5 mL/tube to set-up triplicate assays. Pelleted suspension cells harvested on day 14 or day 18 (see Table 2.2) were resuspended in 100 µL APEL2 followed by addition to the MethoCult™ bijou tube and vortexed to ensure adequate mixing. The suspension cell-methylcellulose mix was allowed to sit at room temperature for bubbles to dissipate before 400 µL/well was dispensed into non-treated 24-well plates (ThermoFisher #144530) using a 18G x 38 mm blunt-end needle (SLS #305180) and 1 mL syringe, giving a final density of 3 x 10^3 cells/well. Only the 8 inner wells of the 24-well plate were used, while the remaining wells and plate chamber were filled with sterile ddH_2O to maintain humidity. Plates were scored for colony formation after 14 days.

2.2 Flow Cytometry

2.2.1 Staining of cells for flow cytometry

Flow cytometric analysis of suspension cells harvested from haematopoietic differentiations using the STEMdiff™ Hematopoietic Kit were analysed using the BD FACSCanto II cytometer. Suspension cells were washed in PBS and fixed in 4% PFA for 15 min at room temperature. Fixed cells were stained with CD34-PerCp (BD Biosciences #345803, 1:250) and CD38-PE (Santa-Cruz #SC-7325, 1:250) for 30 min at

room temperature and then washed twice with PBS supplemented with 1% FBS (1% PBA) and analysed.

Table 2.3: Conjugated antibodies used for flow cytometry

Antibody	Fluorochrome	Manufacturer	Catalog #	Dilution
CD38	PE	Santa-Cruz	SC-7325	1:20
CD34	PeCy7	Thermofisher	25-0349-42	1:20
CD34	FITC	BD Bioscience	555821	1:5
NCAM	PE	Biolegend	362508	1:20
NCAM	BV421	Biolegend	318328	1:20
KDR	AF647	BD Bioscience	560495	1:10
CD43	APC	Biolegend	343206	1:40
CD90	BV421	Biolegend	328122	1:20
CD45	PE	BD Bioscience	55548	1:5
CD38	PerCp-Cy5.5	Biolegend	303522	1:20
CD38	PE	Biolegend	356604	1:20

Table 2.4: Conjugated isotype antibodies used for flow cytomtery

Isotype	Fluorochrome	Manufacturer	Catalog #	Dilution
Mouse IgG1, k	PeCy7	Thermofisher	25-4717-80	1:800
Mouse IgG1, k	FITC	BD Bioscience	555748	1:20
Mouse IgG1, k	PE	Biolegend	400113	1:50
Mouse IgG1, k	BV421	Biolegend	400157	1:67
Mouse IgG1, k	AF647	BD Bioscience	560495	1:80
Mouse IgG1, k	APC	Biolegend	400119	1:20
Mouse IgG1, k	BV421	Biolegend	400157	1:20
Mouse IgG1, k	PE	BD Bioscience	555749	1:10
Mouse IgG1, k	PerCp-Cy5.5	Biolegend	400149	1:10
Mouse IgG1, k	PE	Biolegend	400113	1:40

Flow cytometric analysis of cells harvested on day 4, 7, 14 and 18 using the spin EB haematopoietic differentiation method were analysed using a Beckman Coulter MoFlo Astrios Cell Sorter or Sony ID7000 Spectral Flow Analyser. Cell pellets were resuspended in 50 µL PBS and stained with 1:2000 Zombie NIR viability stain (Biolegend #423105) for 10 min at room temperature followed by staining with conjugated antibodies

for 20 min at 4ºC (Table 2.3 for list of antibodies). Isotype, single-stain and unstained controls were prepared by resuspending pelleted cells in 50 µL PBS followed by staining for 20 min at 4ºC with a conjugated isotype antibody (Table 2.4), a single antibody (Table 2.3) or no antibodies (unstained). Cell suspensions were then washed twice in 1% PBA and resuspended in 500 µL PBS for flow cytometric analysis.

2.2.2 Flow cytometry analysis

Flow cytometric data was analysed using Beckman Coulter Kaluza Analysis Software.

2.3 Molecular Biology

2.3.1 RNA extraction

RNA extraction was performed using NucleoSpin® RNA Isolation kit (Machery-Nagel #740955.50) according to the manufacturer's instructions. RNA was eluted in 30 µL of RNase-free water and quantified using NanoDrop spectrophotometer (ThermoFisher)

For RNA-seq, RNA was extracted using TRI Reagent (Sigma Aldrich #T9424) and Direct-Zol RNA miniprep (Zymo #R2050). RNA was eluted in 20 µL of RNase-free water and quantified using QuBit fluorometer (Invitrogen).

2.3.2 cDNA synthesis

Total RNA was converted to cDNA in a 20 µL reaction. RNA was first incubated with 10 mM dNTPs (NEB #N0447L) and 0.3 µg/µL random primers (Promega #C1181) at 65ºC for 5 min followed by 1 min on ice. 5X First-Strand Buffer, 0.1 M DTT and Superscript III reverse transcriptase (Invitrogen #18080044) were then added to the reaction and incubated initially at room temperature for 5 minutes and then at 50ºC for 3 hours. Heat inactivation of reverse transcriptase was performed for 15 minutes at 70ºC. The reaction was subsequently diluted in nuclease-free water and cDNA was stored at -20ºC.

2.3.3 Quantitative real-time Polymerase Chain Reaction (qPCR)

Gene expression analysis was measured via qPCR on the Applied Biosystems 7500 Fast Real-time PCR system. All qPCR reactions were performed in triplicates. All data was analysed using the comparative CT method (Schmittgen and Livak, 2008). Data is expressed relative to undifferentiated iPSCs unless other wise indicated.

6.1.3.1 TaqMan Assay

HOXA9 expression during cardiac differentiation and in hiPSCs was measured via TaqMan qPCR using a previously described HOXA9 primer-probe set (Dickson et al., 2009) and endogenous control primer-probe sets from Applied Biosystems, 18S rRNA (4319413E) or B2M (4326319E). Each 15 µL qPCR reaction contained 7.5 µL TaqMan Universal Master Mix II (Applied Biosystems, 4440038), 1.5 µL primer-probe and 2 µL cDNA (10 ng/µL). Cycling conditions were as follows: 50^{o}C for 2 mins, 95^{o}C for 10 mins, and 40 cycles of 95^{o}C for 15 seconds and 60^{o}C for 1 min.

6.1.3.2 SYBR Green Assay

Gene expression during haematopoietic differentiation was measured using SYBR Green qPCR (see Table 2.5 for list of primers used). Each 10 µL qPCR reaction contained 5 µL GoTaq qPCR master mix (Promega #A6001), 0.25 µL 10 µM forward and reverse primer (Sigma Aldrich) and 2 µL cDNA. Cycling conditions included an initial denaturation at 95^{o}C for 5 min followed by 40 cycles at 95^{o}C for 15s, 60^{o}C for 30s and 72^{o}C for 30s. Following amplification, a melt-curve step was performed at 95^{o}C for 15s, 60^{o}C for 1 min, 95^{o}C for 30s and 60^{o}C for 15s. Gene expression was normalised to an endogenous control, protein phosphatase 1A (PP1A).

Table 2.5: qPCR primers

Gene	Primer (5' – 3')	
	Forward	**Reverse**
OCT4	AGTGAGAGGCAACCTGGAGA	ACACTCGGACCACATCCTTC
NANOG	CATGAGTGTGGATCCAGCTTG	CCTGAATAAGCAGATCCATGG
SOX2	TGGACAGTTACGCGCACAT	CGAGTAGGACATGCTGTAGGT
HOXA3	GATGGAGCCATGGGAAGATTAC	CATAGGGAGGAGGCTGAGATATAG
HOXA5	TCCCATCGCTTCCCTACCT	GCTTTGGAACAGCCTACAGCTT
HOXA7	AGCTTGGAAATTCTGCTCACTTCT	TCTGATGTCATGGCCAAATTTG
HOXA9	GCCGGCCTTATGGCATTAA	CAGGGACAAAGTGTGAGTGTCAA
HOXA10	AAAGCCTCGCCGGAGAA	CTCCAGTGTCTGGTGCTTCGT
CDX1	GAGAAGGAGTTTCATTACAGCCG	GTTCACTTTGCGCTCCTTTGCC
CDX2	CCGAACAGGGACTTGTTTAGAG	AGGTTGGCTCTGGCATTTATAG
CDX4	CGAGAAGACTGGAGCGTGTA	CTGTAGTCGGTCGAGCAGAA
MIXL1	GTGCTCCTGGAACTGAAACG	GCTAGAGTCAGAGATGCCCC
HAND1	AAAGGCTCAGGACCCAAGAA	GCGCAGAGTCTTGATCTTGG
mScarlet	GAAGGGCGACATTAAGCACG	AACTTGCGGTCGACGTTGTA
GATA2	TCAGCCACTCCGGACACAT	GTCCTCGACGTCCATCTGTT
RUNX1c	CCTTCGTACCCACAGTGCTT	CAACGCCTCGCTCATCTT
SPI1	GACACGGATCTATACCAACGCC	CCGTGAAGTTGTTCTCGGCGAA
TAL1	GACACAGTGCAAGCTGGAAGAC	AGTCAGGCTCTTGATCCTCACC
PP1A	TCTTTCACTTTGCCAAACACC	CATCCTAAAGCATACGGGTCC

2.3.4 Immunostaining

Samples were fixed in 4% PFA for 15 mins at room temperature and washed twice in PBS at room temperature. Samples were incubated for 30 mins at room temperature in a 10% blocking solution, comprising of 10% FBS and 0.1% Triton X-100 (Sigma #T8787) in PBS. Blocking solution was removed and samples were incubated overnight with the primary antibody diluted in a 1% wash solution (1% blocking solution in PBS) 4ºC. Samples were washed 3 times in 1% wash solution before incubation in secondary antibody diluted in 1% wash solution for 1 hour at room temperature. Primary and secondary antibodies used are listed in Table 2.6. Samples were washed a further 3 times in 1% wash solution and incubated with 0.5 µg/mL DAPI (Sigma-Aldrich #D9542)

for 5 mins at room temperature for nuclear staining. Samples were washed a final two more times in 1% wash solution and stored at 4°C before imaging within 24 hours.

Table 2.6: Primary and secondary antibodies used for immunostaining

Antibody	Type	Manufacturer	Dilution
Anti-2A rabbit polyclonal	Primary	Millipore #ABS31	1:1000
Anti-α-actinin mouse monoclonal	Primary	Sigma-Aldrich #A7811	1:800
Anti-OCT4 mouse monoclonal	Primary	Santa-Cruz #SC-5279	1:200
Anti-NANOG goat polyclonal	Primary	R&D Systems #AF1997	1:200
Goat anti-mouse AF488	Secondary	Invitrogen #A11029	1:400
Donkey anti-rabbit AF647	Secondary	Invitrogen #A31573	1:400

2.3.5 Genomic DNA extraction

Genomic DNA was extracting using the DNeasy Blood & Tissue kit (Qiagen #69504) according to the manufacturer's instructions. DNA was eluted 50 µL Nuclease-free water and quantified using Nanodrop spectrophotometry.

2.3.6 Methylated DNA immunoprecipitation (MeDIP)

The methylation status of the AAVS1 locus and TRE promoter during iA9 clone hiPSC differentiation to cardiomyocytes was assessed via MeDIP. Samples were collected on day 0, 2 and 6 of differentiation. Large quantities of gDNA are required for MeDIP, therefore hiPSCs were seeded in T25s at 1 x 10^6 cells/T25 and differentiated as described in Section 2.1.6.

6.1.3.3 Genomic DNA extraction

Genomic DNA was extracted using phenol/chloroform. Cell pellets were resuspended in 1 mL lysis buffer (20 mM Tris pH 8.0, 4 mM EDTA, 20 mM NaCl, 1% SDS) followed by incubation with 3.6 µg/mL RNase A (Qiagen #19101) at 37°C for 1 hour. Proteinase K (20 mg/mL, Promega #V3021) was added and samples were incubated at 55°C overnight. Extraction was performed twice with one volume phenol:chloroform:isoamyl alcohol (25:24:1, Sigma #P2069) and DNA was precipitated with 1/10 5M NaCl diluted in one volume isopropanol. Samples were incubated for 1 hour at -20°C and centrifuged at 14,000 *x g* for 10 mins to pellet DNA. The DNA pellets were washed with 70% ethanol and centrifuged again under the same conditions. The supernatant was removed, and

the pellets were allowed to dry at room temperature to remove residual ethanol before being resuspended in 50 µL Nuclease-free water. DNA was quantified using the Nanodrop spectrophotometer.

6.1.3.4 Genomic DNA sonication

DNA was fragmented by sonication. DNA samples (20 – 60 µg) were diluted in 400 µL nuclease-free water and sonicated using Covaris S220 ultrasonicator at a frequency of 550 KHz. DNA was sonicated for three 15 min cycles of 2 min on/1 min off to generate fragments ~400 bp in length. Aliquots of each sample were run after each 15 min cycle on an agarose gel to assess fragment length and ensure equal sonication.

6.1.3.5 Immunoprecipitation of methylated DNA

DNA was denatured for 10 min at 95ºC, followed by incubation on ice for 10 min. 51 µL of 10X IP buffer (100 mM Na-Phosphate pH 7.1, 1.4 M NaCl, 0.5% Triton X-100) and 7 µL of 5-methylcytosine (5-mC) mouse monoclonal antibody (Diagenode #C15200081) were added to the samples and incubated at 4ºC for 2 hours whilst spinning. Sheep anti-mouse Dynabeads M-289 (Invitrogen #11201D) were used to precipitate 5mC fragments. The beads were resuspended in 1X IP buffer and 20 µL was added to each sample, followed by a further 2 hour incubation at 4ºC whilst spinning. The samples were placed in a magnetic rack to remove the supernatant and washed twice with 1X IP buffer. Proteinase K (25 mg/mL, Promega #MC5005) was added to the samples which were then incubated shaking at 55ºC overnight. Samples were then column purified using QIAquick PCR purification kit (Qiagen #28104) according to the manufacturer's instructions. Immunoprecipitated DNA was quantified using Qubit™ dsDNA HS assay kit (Thermofisher #28104) according to the manufacturer's instructions. MeDIP controls included sonicated gDNA which did not undergo immunoprecipitation (Input control) and sonicated gDNA which underwent immunoprecipitation in the absence of an antibody (no antibody control).

6.1.3.6 QPCR of methylated regions

Methylation status of samples was determined by SYBR Green qPCR for methylated regions using the Applied Biosystems 7500 Fast Real-time PCR system. Each qPCR reaction contained 5 µL GoTaq qPCR master mix (Promega #A6001), 0.25 µL 10 µM forward and reverse primer (Sigma Aldrich) and 4 µg gDNA immunoprecipitated with 5-mC antibody (MeDIP fraction), 4 µg no antibody control gDNA or 400 ng Input control gDNA. Cycling conditions included an initial denaturation at 95ºC for 5 min followed by

50 cycles at 95°C for 15s, 60°C for 30s and 72°C for 30s. Following amplification, a melt-curve step was performed at 95°C for 15s, 60°C for 1 min, 95°C for 30s and 60°C for 15s. Data was analysed using the comparative CT method. The MeDIP fraction and no antibody control for each timepoint was normalised relative to their respective Input control. Enrichment in the MeDIP fraction was subsequently determined by comparison with the sample's corresponding no antibody control. Fold changes >1 indicate enrichment for methylation in the sample. As the PCR templates comprise of fragmented DNA of unknown length, several different primer pairs targeting the TRE promoter and the AAVS1 locus were designed. Primer pairs were designed so products contained 2 – 6 GCs and were approx. 150 – 200 bp long. A primer pair was deemed to be effective if it produced a product during amplification of Input controls and a single peak during the melt curve cycle. The functional primers used are listed in Table 2.7. Their location within the TRE promoter and AAVS1 locus and their design are also indicated.

Table 2.7: MeDIP qPCR primers

Primer	Sequence (5' – 3')	Location	Design
UBE2B F	CTCAGGGGTGGATTGTTGAC	Negative control	Mohn *et al.*, 2009
UBE2B R	TGTGGATTCAAAGACCACGA		
TRE #1 F	GTATGTCGAGGTAGGCGTGT	End of TRE promoter	SnapGene
TRE #1 R	GGCGATCTGACGGTTCACTA		
TRE #2 F	TCACGAGGCCCTTTCGT	Start of TRE promoter	SnapGene
TRE #2 R	TAAACTCGACATCGTTCTC		
Pre-EcoRI F	TCCCCAATTTATATTGTTCCT	Downstream of EcoRI site in AAVS1 locus	SnapGene
Pre-EcoRI R	TTCTCACAAAGGGAGTTTTC		
EcoRI F	CCCTTTGTGAGAATGGTGCG	EcoRI site at 2756 bp in AAVS1 locus	Ordovas *et al.*, 2015
EcoRI R	GCTTGCCAAGGACTCAAACC		
Post-EcoRI F	GGTGACTGCTTGGCAAACTC	Upstream of EcoRI site in AAVS1 locus	SnapGene
Post-EcoRI R	TGTTTTCCTAGGACTGAGGGT		
Post-exon I F	AGGTCCACCCTCTGCTGCG	After PPP1R12C exon I	SnapGene
Post-exon I R	TCGAAGGGGAATGGTAAGGG		

2.4 Generation of plasmids for CRISPR/Cas9 transfections

2.4.1 Primer design

All primers were designed using SnapGene software. Melting temperature of primers were calculating using the Oligonucleotides Property Calculator provided by Northwestern University.

2.4.2 Polymerase chain reaction (PCR)

PCR was performed using Phusion® High-Fidelity DNA Polymerase (NEB #M0530) or Taq DNA Polymerase (NEB #M0273). Thermocycling conditions and standard reaction mixes are indicated in Table 2.8 and Table 2.9, respectively. Extension time was calculated at 1 kb per minute. Primer pairs used in generation of pAAVS1-TRE-HOXA9-mScarlet are listed in Table 2.11. Primers used for genotyping PCRs are listed in Table 2.10.

Table 2.8: PCR thermocycling conditions

Step	Temperature		Time	Cycles
	Phusion	Taq		
Initial Denaturation	98°C	95°C	5 min	1
Denaturation	98°C	95°C	30s	25 - 30
Annealing	56 - 74°C		30s	
Extension	72°C		30s – 2 min 30s	
Final Extension	72°C		5 min	1

Table 2.9: PCR reaction set-up

Component	Final Concentration	
	Phusion	**Taq**
5X Phusion HF Reaction Buffer	1X	-
10X Standard *Taq* Reaction Buffer	-	1X
10 mM dNTPs	200 µM	200 µM
10 µM Forward Primer	0.5 µM	0.5 µM
10 µM Reverse Primer	0.5 µM	0.5 µM
Template	<250 ng	<1000 ng
Polymerase	1 unit/50 µL	1.25 units/50 µL
Nuclease-free water	Up to 50 µL	Up to 50 µL

Table 2.10: Genotyping PCR primers

Primer name	Sequence (5' – 3')
5-int F	TCCGGACCACTTTGAGCTCTACTGGC
5-int R	TGCGGCCTGTCCAATACCTCCCG
3-int F	GCCTCGGTGCACATGCTTTACATG
3-int R	CAAGCAGTCACCCCACAGTTGGAGGAG
AAVS1 F	TCTCCTGAACCTGAGCCAGCTC
AAVS1 R	AGCAGTCACCCCACAGTTGGA

Table 2.11: Primer pairs used for PCR and Fusion PCR

Template	Primer	Sequence (5' – 3')	Reaction
pAAVS1-TRE-Shox2a-Nluc	P1	CCCCACA**ACTAGT**ACGTCTCTCAAGGA	PCR
	P2	TCGTAGGGGTAAGCCATGGAGGCCTTAAGGGCGATCTGACGGTTCACTA	
pMSCVhygro-HA-hHOXA9	P3	TAGTGAACCGTCAGATCGCCCTTAAGGCCTCCATGGCTTACCCCTACGA	PCR
	P4	CACGTCGCCACAGGTCAGCAGAGAGCCCCGGCCCTCACCCTCGTCTTTTGCTCGGT	
TRE-HOXA9-AflII + AflII-HOXA9-T2A	P1	CCCCACA**ACTAGT**ACGTCTCTCAAGGA	Fusion PCR
	P4	CACGTCGCCACAGGTCAGCAGAGAGCCCCGGCCCTCACCCTCGTCTTTTGCTCGGT	
TRE-HOXA9-AflII + AflII-HOXA9-T2A	P1	CCCCACA**ACTAGT**ACGTCTCTCAAGGA	Fusion PCR
	P5	ACGGGTCGTACGGGGGCCGGGGTTCTCCTCCACGTCGCCACAGGTCAG	
pmScarlet_H_C1	P6	GCAGGTCGTACGATGGTGAGCAAGGGCGAGGCA	PCR
	P7	GCACAGCTTAAGCTACTTGTACAGCTCGTCCATGC	
pAAVS1-TRE-Shox2a-Nluc	P8	CAAGTAGCTTAAGCTGTGCCTTCTAGTTGCCAG	PCR
	P9	TTAATA**ACTAGT**CAATAATCAATGTCGGCC	
BsiWI-mScarlet-AflII + AflII-pA	P6	GCAGGTCGTACGATGGTGAGCAAGGGCGAGGCA	Fusion PCR
	P9	TTAATA**ACTAGT**CAATAATCAATGTCGGCC	

2.4.3 Fusion PCR

Fusion PCR was performed using two PCR amplified fragments with complementary overhangs. Fusion PCR was carried out in three steps using Phusion® High-Fidelity DNA Polymerase and PCR products generated using PCR primers in Table 2.11 as templates. In step 1, templates were denatured and realligned at corresponding overhangs. Annealing template reaction set-up and PCR conditions are shown in Table 2.12 and Table 2.13, respectively.

Table 2.12: Annealing template reaction set-up

Component	Final Concentration
5X Phusion GC Reaction Buffer	1X
Fragment 1	8 ng/µL
Fragment 2	8 ng/µL
Nuclease-free water	Up to 15 µL

Table 2.13: Annealing PCR conditions

Step	Temperature	Time
Initial Denaturation	94°C	5 min
Annealing Cycle	94°C	30 s
	75°C	5 min
	52°C	5 min
	45°C	5 min

In step 2, Phusion polymerase and dNTPs were added to the annealing reaction and incubated for 2 mins at 72°C (see Table 2.14 for reaction set-up).

Table 2.14: Extension of annealing template reaction set-up

Component	Final Concentration
5X Phusion GC Reaction Buffer	1X
10mM dNTPs	200 µM
Annealing template	From step 1
Phusion polymerase	1 unit/50 µL reaction
Nuclease-free water	Up to 25 µL

In step 3, forward and reverse primers (see Table 2.11 for primer pairs used) were added and the annealing template was amplified. Two different annealing temperature were used. Loop I annealing temperature was calculated based on the primer sequence which binds to the template (i.e. the primer which does not contain the overhang sequence). Loop II annealing temperature was calculate based on the complete primer sequence. PCR conditions for amplification are shown in Table 2.15.

Table 2.15: PCR amplification conditions of fusion product

Step	Temperature	Time
Initial Denaturation	98°C	2 min
Loop I (5 cycles)	98°C	15 s
	Primer dependent	30 s
	72°C	1 kb/min
Loop II (25 cycles)	98°C	15 s
	Primer dependent	30 s
	72°C	1 kb/min
Final extension	72°C	5 min

2.4.4 Restriction digestion

Restriction digests were performed using 100 – 500 ng of PCR products or 1 – 2 µg of plasmid DNA in a 50 µL reaction containing 1 µL of restriction enzyme and the appropriate reaction buffer. Reactions were incubated overnight at 37°C, unless otherwise indicated. All restriction enzymes were purchased from NEB. BsiWI (#R0553S) was used to digest TRE-HOXA9-T2A and mScarlet-pA to generate the TRE-HOXA9-T2A-mScarlet-pA fragment. FspI (#R0135S) and SpeI-HF (#R3133S) were used

to isolate TRE-HOXA9-T2A-mScarlet-pA from the pMiniT vector. SpeI-HF was used to generate compatible overhangs for insertion of the TRE-HOXA9-T2A-mScarlet-pA into the backbone vector from pAAVS1-TRE-Shox2a-Nluc. SalI (#R0183S), SpeI-HF, BsiWI and AflII (R0520S) were used in diagnostic digests to confirm the generation of pAAVS1-TRE-HOXA9-mScarlet. For diagnostic digests, restriction enzymes were incubated for 1h at 37°C.

2.4.5 Dephosphorylation

To prevent self-ligation of pAAVS1-TRE-Shox2a-Nluc following digestion with SpeI-HF, the linearised backbone vector was dephosphorylated with Antarctic Phosphatase (NEB #M0289). Dephosphorylation removes the 5' phosphate group from overhangs, preventing DNA ligase from forming a phosphodiester bond with the 3' hydroxyl group. A 20 µL reaction containing 1 pmol of the linearised backbone vector, 10X Antarctic Phosphatase Reaction Buffer and 5 units of Antarctic Phosphatase was prepared and incubated at 37°C. After 30 mins, an additional 5 units of Antarctic Phosphatase was added, and the reaction was incubated for a further 30 mins at 37°C. To remove the phosphatase, backbone DNA was column-purified using QIAquick PCR Purification Kit according to the manufacturer's instructions (Qiagen #28104).

2.4.6 Ligation

Ligation of fragments with compatible overhangs was performed overnight at 4°C using 1 µL T4 DNA ligase (NEB #M0202) in a 20 µL reaction with T4 DNA ligase reaction buffer. Ligation of fusion PCR products following BsiWI digestion was performed at a molar ratio of 1:1. Generation of pAAVS1-TRE-HOXA9-mScarlet was formed by ligation at ratios of 1:3 or 1:5 vector:insert. T4 DNA ligase was heat inactivated at 65°C for 10 mins.

2.4.7 Agarose gel electrophoresis

Agarose gel electrophoresis was performed to detect the products of PCR and restriction digestion reactions. DNA products were run on 0.8 – 1.5% agarose gels containing 10 mg/mL ethidium bromide (Invitrogen #15585011) at 80V for 1 – 1.5 hours. DNA samples were loaded with 6X loading dye (NEB #B57021S) diluted to 1X, alongside 5 µg DNA ladders for reference (100 bp or 1 kb, NEB #N3231S or #N3232S). Gels were visualised using LAS-4000 Fujifilm Luminescent Image Analyser. Desired bands were extracted under UV light and column purified using QIAquick Gel extraction Kit according to the

manufacturer's instructions (Qiagen #28704). DNA was eluted in nuclease-free water and quantified using the Nanodrop spectrophotometer.

2.4.8 Bacterial transformation

Plasmids were transformed into NEB 5-alpha Competent *E. coli* cells (NEB #C2987I) according to the manufacturer's instructions. Briefly, 5-alpha cells were thawed on ice followed by addition of 2 – 5 µL of plasmid DNA (1 pg – 100 ng) and incubated on ice for 30 minutes. Cells were heat shocked for 30 seconds at 42°C, then incubated on ice for a further 5 minutes. 950 µL of SOC media was added and the mixture was shaken at 225 rpm for 1 hour at 37°C. 100 µL of the mixture was spread on LB agar plates containing 100 µg/mL Ampicillin and incubated overnight at 37°C. Bacterial colonies were cultured in 5 mL LB Broth with 100 µg/mL Ampicillin for 14 – 16 hours at 37°C shaking at 225 rpm. DNA was extracted from subcultures by miniprep and screened using diagnostic digests (Section 2.4.4) and plasmid sequencing.

2.4.9 DNA mini- and midi-prep

DNA mini-preps were performed using Wizard® *Plus* SV Minipreps DNA Purification System (Promega #A1340), while midi-preps were performed using NucleoBond® Xtra Midi (Takara Bio #740410.50). Midi-preps were carried out when large quantities of pure plasmid DNA was required, such as for CRISPR/Cas9 transfections. In both cases, bacteria subcultures were harvested by centrifugation at 5,000 *x g* for 5 minutes and plasmid DNA was isolated and purified according to the manufacturer's instructions. DNA was eluted in nuclease-free water and quantified using the Nanodrop spectrophotometer.

2.4.10 Sanger Sequencing

Samples were sequenced using Sanger Sequencing services provided by Source Biosciences or Eurofins. Sequencing results were analysed using SnapGene.

2.5 Microscopy

Live and fixed fluorescence microscopy was performed using the Operetta® high content image analysis system (PerkinElmer). At least three technical replicates were performed for each experiment, unless otherwise indicated. A technical replicate represents a single well of a well plate in an experiment, with 10 – 20 fields of few imaged per well. Biological replicates of independent experiments were performed where indicated. Columbus™

analysis software (PerkinElmer) was used to quantify fluorescence expression. Quantification results are shown as the mean ± standard deviation (SD) of technical replicates. Mean and SD were calculated using GraphPad Prism.

2.6 RNA-sequencing

RNA-sequencing was performed on iA9 haematopoietic differentiation samples. Total RNA was extracted from two independent biological replicates using TRI Reagent and Direct-Zol RNA miniprep kit (see Section 2.3.1). RNA samples were submitted to Novogene for quality control checks, library preparation and sequencing using the Illumina platform. Read alignment was performed using Hisat2 and reads were mapped to the GRCh38/hg 38 Ensembl human genome. Unsupervised and supervised hierarchical clustering was performed with sample RPKM values using Cluster 3.0 and expression heat maps were generated using TreeView software. Differential gene expression analysis was performed using DESeq2. P values were adjusted using the Benjamini and Hochberg's approach and raw read count was normalised. KEGG and gene ontology analysis was performed using WebGestalt (http://www.webgestalt.org/). Volcano plots were made using VolcaNoseR (https://huygens.science.uva.nl/VolcaNoseR/).

2.7 Statistical analysis

All data is presented as mean ± SD unless otherwise stated. Mean and SD are determined based on three technical replicates for qPCR and CFU assay analysis, and at least 3 technical replicates for fluorescent image quantification using Columbus™. For flow cytometry experiments, mean and SD were determined based on biological replicates, unless otherwise indicated. Statistical analysis was performed using GraphPad Prism and only on samples with three biological replicates (n = 3), unless otherwise stated.

Chapter 3: Engineering an inducible HOXA9-mScarlet overexpression system in hiPSCs

3.1 Introduction

3.1.1 CRISPR/Cas9 genome editing

Precise, robust, and efficient modification of the human genome has become readily attainable with the advent of designer programmable nucleases. The development of zinc-finger nuclease (ZFN), transcription activator-like effector nuclease (TALEN), and clustered regularly interspaced short palindromic repeat (CRISPR)/Cas technologies have revolutionised gene editing and its applications. All three technologies act in a similar fashion; each introduces target-specific double-stranded breaks (DSBs) in DNA and harness endogenous repair machinery to permanently modify the target sequence (Kim and Kim, 2014). However, all three technologies also differ in their design, efficiency, and specificity; factors which much be taken into consideration when deciding which technology to use for gene editing.

ZFNs are customised hybrid restriction enzymes consisting of a DNA-binding zinc finger protein fused to the DNA cleavage domain of FokI endonuclease (Kim et al., 1996). Zinc finger proteins are composed of 3 – 6 zinc fingers, each binding 3 bp DNA sequences, creating DNA target sequences 9 – 18 bp in length (Kim and Kim, 2014; Urnov et al., 2005). FokI must dimerise in order to form a functional nuclease, and as a result two ZFNs are required for effective DNA cleavage, each targeting adjacent DNA sites separated by 6 bps (Bibikova et al., 2001). Advances in ZFN design have improved ZFN activity, in particular, modification of the FokI domain has greatly reduced FokI homodimer formation responsible for high off-target cleavage and cytotoxicity (Miller et al., 2007; Szczepek et al., 2007). However, despite these improvements, ZFN cleavage efficiency is variable and cleavage is often greatest in G-rich regions, restricting possible target sites (Kim and Kim, 2014).

Similar to ZFNs, TALENs are comprised of a DNA-binding domain, in this case a transcription activator-like effector (TALE) protein, fused to the FokI cleavage domain (Boch et al., 2009). TALEs are composed of 15.5 – 19.5 repeats, with each repeat containing ~34 amino acids. Target recognition is mediated through two amino acids at positions 12 and 13, which recognise one nucleotide. Thus, one repeat recognises one nucleotide, making the design of customised TALENs relatively straightforward (Holkers

et al., 2013). Like ZFNs, two TALENs are required to dimerise to form a functional FokI nuclease. This has the advantage of increasing the target sequence length, thereby minimising the potential for off targets. TALENs also have a high cleavage efficiency but they are limited by their highly repetitive structure and large size (~3 kb per TALEN) which creates challenges in cell delivery (Holkers et al., 2013; Miller et al., 2011).

The CRISPR/Cas system is an adaptive immune defence system native to bacteria and archaea. CRISPR systems are classified into at least five sub-types (types I – V) and share several features; all are regulated by Cas proteins and consist of a CRISPR locus comprising of a series of repeat-spacer-repeat sequences (Makarova et al., 2015). Repeat sequences are conserved within a CRISPR locus, while spacer sequences are unique and acquired from invading plasmid or viral DNA (Makarova et al., 2011; Wiedenheft et al., 2012). CRISPR sub-types differ in their method of transcribing and processing the CRISPR locus into single repeat-spacer RNAs, termed CRISPR-associated RNAs (crRNAs), which together with Cas proteins are responsible for silencing of invading sequences (Jore et al., 2019). Among the CRISPR sub-types, only Type II systems utilise and express specialised *trans*-activating CRISPR-RNAs (tracrRNAs) for this process. TracrRNA forms a complex with repeats in the CRISPR locus and recruits RNase III and Cas9 proteins to process the crRNA-tracrRNA duplex into mature crRNA (Deltcheva et al., 2011; Jinek et al., 2012).

The type II CRISPR system of *Streptococcus pyogenes* has been adapted to create a genome engineering tool in eukaryotes. Through fusion of mature crRNA to tracrRNA, a chimeric single guide RNA (gRNA) was shown to be sufficient to direct Cas9 protein to cleave a specific target site *in vitro* (Jinek et al., 2012) and subsequently in mammalian cells (Cong et al., 2013; Mali et al., 2013). Site-specific targeting is mediated through a 20 nt sequence within the gRNA, which binds to its complementary sequence in the target DNA (termed 'protospacer'), and a short protospacer adjacent motif (PAM). As such, provided target DNA directly follows a 5' NGG PAM site, the 20 nt sequence may be altered to direct Cas9 to any target site in the genome (Ran et al., 2013a). As a 5' NGG sequence is reported to occur every 8 – 12 bp, this requirement places minor restrictions on potential Cas9 targets (Cong et al., 2013; Hsu et al., 2013). Once directed to a target, Cas9 protein induces a DSB 3 bp upstream of the PAM site through its two endonuclease domains, HNH and RuvC, which are responsible for cleaving complementary and non-complementary DNA strands, respectively (Jinek et al., 2012; Sapranauskas et al., 2011). The induced DSBs stimulate innate DNA repair by either

non-homologous end-joining (NHEJ) or homology-directed repair (HDR). During NHEJ, insertion/deletion (indel) mutations are introduced into target DNA. In coding regions, this can lead to frameshifts and premature stop codons causing gene knock-out. In contrast, HDR facilitates precise gene editing. Plasmid or single-stranded DNA repair templates bearing regions homologous to the target site stimulate the HDR pathway and enable incorporation of the repair 'donor' template (Ran et al., 2013a). The CRISPR/Cas9 system has many advantages over other nucleases; the system is easily customised, has few target restrictions and is relatively small in size (Cas9 gene is ~4 Kb) for ease of delivery (Cho et al., 2013). However, the CRISPR/Cas9 system lacks the stringent target specificity of systems such as ZFNs and TALENs, and is limited by the potential for non-specific Cas9 cleavage (Ran et al., 2013a).

Cas9 specificity and activity varies with target sites and the number and position of mismatches between the 20 nt guide sequence and its complementary DNA (Cho et al., 2014; Fu et al., 2013). In general, mismatches are better tolerated at the PAM distal region of the guide sequence than the proximal region. However, Cas9 has been shown to induce off-target cleavage events when mismatches occur within the PAM proximal region (Fu et al., 2013; Hsu et al., 2013; Pattanayak et al., 2014). Efforts to minimise off-target events focus on modifying gRNA design and improving Cas9 specificity. Online guide design tools, such as those provided by the Broad Institute, can aid in design of optimal guides for a given target site and predict potential for off-targets. GC content of guide sequences must also be considered during design. Guides with high GC have been reported to have a high tolerance of mismatches, possibly due to stabilisation of gRNA-DNA binding (Fu et al., 2013).

Cas9 targeting specificity can be enhanced by use of a paired Cas9 nickase (Cas9n) approach. Mutating either the RuvC (D10A mutation) or HNH (H840A mutation) nuclease domain of Cas9 generates Cas9n, which cleaves one DNA strand to produce 5' or 3' overhangs (Ran et al., 2013b). Therefore, introduction of DSBs requires two gRNAs binding in close proximity to adjacent target strands, increasing the specificity of targeting. Paired Cas9n has been demonstrated in multiple human cell lines to yield fewer off-target events than wild-type Cas9, while also stimulating NHEJ- and HDR-mediated gene editing to a similar efficiency (Cong *et al.*, 2013; Mali *et al.*, 2013; Cho *et al.*, 2014). As such, Cas9n represents a robust approach for gene editing with minimal off-target effects.

3.1.2 Genomic safe-harbour sites

Gene editing applications often require stable integration of transgenes into the host genome. In applications where the integration site is not pre-determined (for example, modification of endogenous genes), genomic safe-harbour sites (SHS) are frequently targeted (Pellenz et al., 2019). SHS are purported to permit stable transgene expression without adversely affecting target or flanking gene expression and function (Sadelain et al., 2012). As such, SHS represent ideal locations for transgene integration. To aid SHS identification, a list of criteria was established which potential sites must adhere to. For safety purposes, sites must be located at least 50 kb from coding and non-coding genes and 300 kb from cancer-related and microRNA genes. Potential sites must be unique and accessible to aid with targeting efficiency; located within open chromatin; and exclude copy number variable regions. In addition, sites must be located outside transcriptional units and ultra-conserved regions to avoid disruption of endogenous genes (Pellenz et al., 2019). To date, no *bona fide* SHS has been identified which meets all criteria. However, several presumptive sites exist which have been routinely targeted to varying degrees of success (Papapetrou and Schambach, 2016).

The adeno-associated virus site 1 (AAVS1) and chemokine (C-C motif) receptor 5 (*CCR5*) loci are the most frequently targeted sites for transgene insertion (Papapetrou and Schambach, 2016). The human ortholog of the mouse Rosa26 locus and citrate lyase beta-like (*CLYBL*) gene have also been shown to be amenable to transgene insertion, although they are less well studied than the AAVS1 and *CCR5* loci. The AAVS1 locus is located within the first intron of *PPP1R12C* on chromosome 19 and is a natural integration site for the AAV virus (Kotin et al., 1992). Disruption of *PPP1R12C* is not associated with any known pathology, making it a desirable site for transgene insertion (Oceguera-Yanez et al., 2016). CCR5 on chromosome 3 is a co-receptor for HIV-1 and was discovered as a potential SHS following the identification of HIV-1 resistant individuals who possess a natural homozygous deletion in *CCR5* (Liu et al., 1996). While *CCR5* disruption was initially associated with no advert phenotype, increased sensitivity to West Nile virus infection was later discovered (Glass et al., 2006). The mouse *Rosa26* locus permits robust, ubiquitous transgene expression and is widely targeted in mouse transgenesis studies (Papapetrou and Schambach, 2016). Search for regions with high sequence homology to mouse *Rosa26* identified the human counterpart (*ROSA26*) on chromosome 3. Human *ROSA26* has been demonstrated to support transgene expression in human ESCs and following differentiation to cells of the three germ layers (Irion et al., 2007), and also to mature cell types of the cardiac, neural and hepatic

lineages (Bertero et al., 2016). However, the *ROSA26* locus is infrequently used and more studies are required to confirm the potential broad utility, efficiency and safety of human *ROSA26* as a SHS (Sadelain et al., 2012). Another potential site, intron 2 of *CLYBL*, has been investigated using TALENs and demonstrated to support 10 times greater transgene expression than the AAVS1 locus in hiPSCs (Cerbini et al., 2015). Despite this promising finding, further studies into the suitability of *CYLBL* as a SHS have not been performed.

Stable transgene integration at both the AAVS1 and *CCR5* sites has been demonstrated following ZFN-mediated targeting in several human cell lines, including iPSCs and ESCs (Hockemeyer et al., 2009; Lombardo et al., 2007). However, in a comparative study of both sites, the AAVS1 locus was shown to permit higher transgene expression than *CCR5* in lymphoid, hepatic and iPS cell lines. Investigation of the epigenetic features of each target site by chromatin immunoprecipitation (ChIP) analysis in lymphoblasts, iPSCs and HSCs revealed a repressive chromatin structure surrounding the CCR5 site, while the AAVS1 target site possessed an open chromatin conformation (Lombardo et al., 2011; Van Rensburg et al., 2013). The presence of an insulator element in the promoter of *PPP1R12C* likely supports the open chromatin conformation of the AAVS1 locus, improving accessibility and facilitating high transgene expression (Ogata et al., 2003). In contrast to *PPP1R12C* which is ubiquitously expressed, CCR5 is mainly expressed on T cell, macrophages, microglia and dendritic cells. Consistent with this expression pattern, CCR5 was found to be poorly transcribed in iPSCs while PPP1R12C was actively transcribed (Van Rensburg et al., 2013). Recently, CRISPR/Cas9-mediated editing of CCR5 in cord blood-derived HSCs was shown to permit robust transgene expression, which was maintained after *in vivo* transplantation (Gomez-Ospina et al., 2019). Therefore, *CCR5* may be most suited for haematopoietic lineage-specific gene editing applications.

A major advantage of the AAVS1 locus is its broad applications; high efficient transgene expression has been shown in numerous cell types, including following hPSC-based differentiation to cells of mesoderm, endoderm and ectoderm lineage (Bertero et al., 2016). However, DNA methylation within the locus has also been reported, resulting in reduced transgene expression in hESCs and hESC-derived hepatocytes (Ordovas et al., 2015). Despite this, the majority of studies demonstrate the locus supports robust transgene expression with limited silencing and adverse effects. Thus, while the AAVS1 site does not meet all SHS criteria, it remains the preferred site for transgene integration.

3.1.3 Inducible gene expression systems

Inducible gene expression systems are useful tools to study the role a given gene may play in a developmental process or phenotype. Inducible systems, such as tetracycline (Tet)-On/Off and tamoxifen-ERT2, permit spatial and temporal control of gene expression and can be modified to either induce ectopic gene expression or gene knockdown. The Tet-On/Off system utilises a bacterial mechanism of resistance to the antibiotic tetracycline to control gene expression. In bacteria, tetracycline repressor proteins (TetR) bind tet operons (tetO) to prevent gene expression. However, in the presence of tetracycline, TetR undergoes a conformational change and is no longer able to bind tetO resulting in expression of genes, such as those involved in drug efflux (Das et al., 2016). Fusion of TetR with a VP16 activator domain produces a tetracycline-controlled transcriptional activator (tTA) which induces expression of genes under the control of a promoter containing tetO sequences, while supressing gene expression in the presence of tetracycline (Gossen and Bujard, 1992). Modification of the tTA structure creates a reverse tTA (rtTA), whereby tetracycline, or its derivative doxycycline (DOX), binds to rtTA to activate gene expression (Gossen et al., 1995).

Since their generation, further optimisation of the tTA and rtTA configuration has improved the sensitivity and activity of the Tet-On/Off system (Kallunki et al., 2019). In early versions of the system, high basal activity of the transactivator in the uninduced state hindered broad application of the system (Roney et al., 2016). The transactivator thus underwent several modifications with the aim of reducing background activity and increasing sensitivity to DOX, while maintaining the high level of induction already achievable; resulting in the commonly used rtTA-M2 variant (Urlinger et al., 2000). In addition, a series of tetracycline response element (TRE) promoters, which comprised of a minimal promoter upstream of seven tetO repeats, were generated. These promoters had substantially reduced background promoter activity and enabled dose-dependent stimulation of the system with tetracycline or DOX in a reversible manner (Loew et al., 2010).

The Tet-On/Off system has been used together with CRISPR/Cas9 to mediate tightly controlled temporal gene editing. Features of the Tet-Off system have been harnessed to generate an inducible gRNA system in human immortalised cell lines (Aubrey et al., 2015). Similarly, the Tet-On system has been utilised to create an inducible siRNA system which has been demonstrated to mediate siRNA expression during differentiation to mesoderm, endoderm and ectoderm lineages (Bertero et al., 2016). An inducible Cas9

system (iCRISPR) has also been generated via knock-in of a TRE promoter-driven Cas9 and rtTA into the AAVS1 locus, facilitating generation of single (Qian et al., 2014) and multiple gene knock-out systems in hPSCs (González et al., 2014).

The tamoxifen-ERT2 system can be used to create inducible gene overexpression systems. Cloning of the gene of interest in-frame with ERT2 will generate a fusion protein which remains localised in the cytoplasm, dependent on tamoxifen for translocation to the nucleus and DNA binding. HOXB4-ERT2 (Jackson et al., 2016) and KLF1-ERT2 (Yang et al., 2017a) systems have been generated in hPSCs by random integration and CRISPR/Cas9-mediated targeting of the AAVS1 locus to achieve inducible HOXB4/KLF1 expression during haematopoietic differentiation. The tamoxifen-ERT2 system has also been used with Cre/LoxP and FRT/Flpe recombination systems to mediate tightly controlled gene knockout (Chen et al., 2015). Both recombination systems are highly sensitive, with minimal Cre/Flpe activity required to excise the gene flanked by LoxP/FRT sites. As such, unintentional gene excision can occur if Cre/Flpe are not tightly controlled (Chen et al., 2015). Generation of Cre/Flpe-ERT2 fusion proteins dependent on tamoxifen for nuclear translocation enables tight control over recombinase activity and has been used in inducible gene knockout (Chen et al., 2015) and lineage tracing systems (Chen et al., 2018).

3.2 Aims

HOXA9 expression during mesoderm patterning (Ng et al., 2016), haemogenic endothelium specification (Doulatov et al., 2013; Sugimura et al., 2017) and EHT (Dou et al., 2016) has been implicated in definitive haematopoiesis and HSC generation. In this chapter we aim to utilise CRISPR/Cas9 technology and the Tet-On gene expression systems to create an inducible HOXA9 hiPSC line which will permit controlled, temporal study of HOXA9 expression during haematopoietic differentiation. In order to achieve stable integration into the genome with minimal adverse effects to endogenous genes, we targeted the inducible HOXA9 expression system to the AAVS1 safe harbour locus. Furthermore, we tagged HOXA9 with mScarlet fluorescent protein to facilitate detection of HOXA9 expression in live cells.

3.3 Results

3.3.1 AAVS1 targeting strategy

The AAVS1 locus, located in intron 1 of *PPP1R12C,* was chosen as the site for insertion of the inducible HOXA9 expression system (see Figure 3.1 for over view of strategy). A dual gRNA/Cas9n targeting approach was adopted using gRNAs kindly provided by Dr Jamie Bhagwan and which had previously been demonstrated to direct highly precise and efficient editing in the AAVS1 locus (Bhagwan et al., 2020a). An inducible HOXA9 donor plasmid utilising the Tet-On system was generated. HOXA9 cDNA was cloned from pMSCVhygro-HA-hHOXA9 (provided by Dr Guy Sauvageau) and linked to red fluorescent protein mScarlet (Bindels et al., 2016) via self-cleaving peptide T2A. HOXA9-mScarlet expression was placed under the control of the TRE promoter, while the constitutively expressed CAG promoter was used to drive the expression of the rtTA domain and a G418 selection marker. This cassette was then flanked on the left and right with 600 bp and 800 bp of homology to the AAVS1 locus, respectively (see Figure 3.1 for overview of targeting strategy). The plasmid backbone and these core components were all derived from a parental vector plasmid, pAAVS1-TRE-Shox2a-Nluc, kindly donated by Dr Mine Bakur.

3.3.2 Building of the HOXA9-mScarlet donor template plasmid

Direct replacement of the Shox2a-Nluc fragment in pAAVS1-TRE-Shox2a-Nluc with HOXA9-T2A-mScarlet was not possible due to the lack of flanking restriction enzyme sites. Two SpeI sites, one located upstream of the TRE promoter and the other in the poly A (pA) tail following Nanoluc (Nluc), were instead used to excise the Shox2a-Nluc fragment. Therefore, the HOXA9-T2A-mScarlet fragment was generated with the TRE promoter and pA tail fragments incorporated at the 5’ and 3’ ends, respectively (Figure 3.2). The entire segment (TRE-HOXA9-T2A-mScarlet-pA) was produced sequentially, with each individual fragment first PCR amplified with addition of complementary overhanging regions followed by Fusion PCR to join the fragments. Fusion PCR requires an overlapping region (minimum 20 bp) between two fragments for them to be joined together. In a Fusion PCR reaction, the overlapping regions are first annealed at an appropriate temperature and then extended upon addition of the polymerase and primers. The overhangs also permit insertion of additional elements. A Kozak sequence was added before the start of HOXA9 to ensure efficient translation. AflII sites were incorporated 5’ of HOXA9 and 3’ of mScarlet to facilitate easy excision of HOXA9-T2A-mScarlet fragment for future use of the construct in Tet-On expression systems.

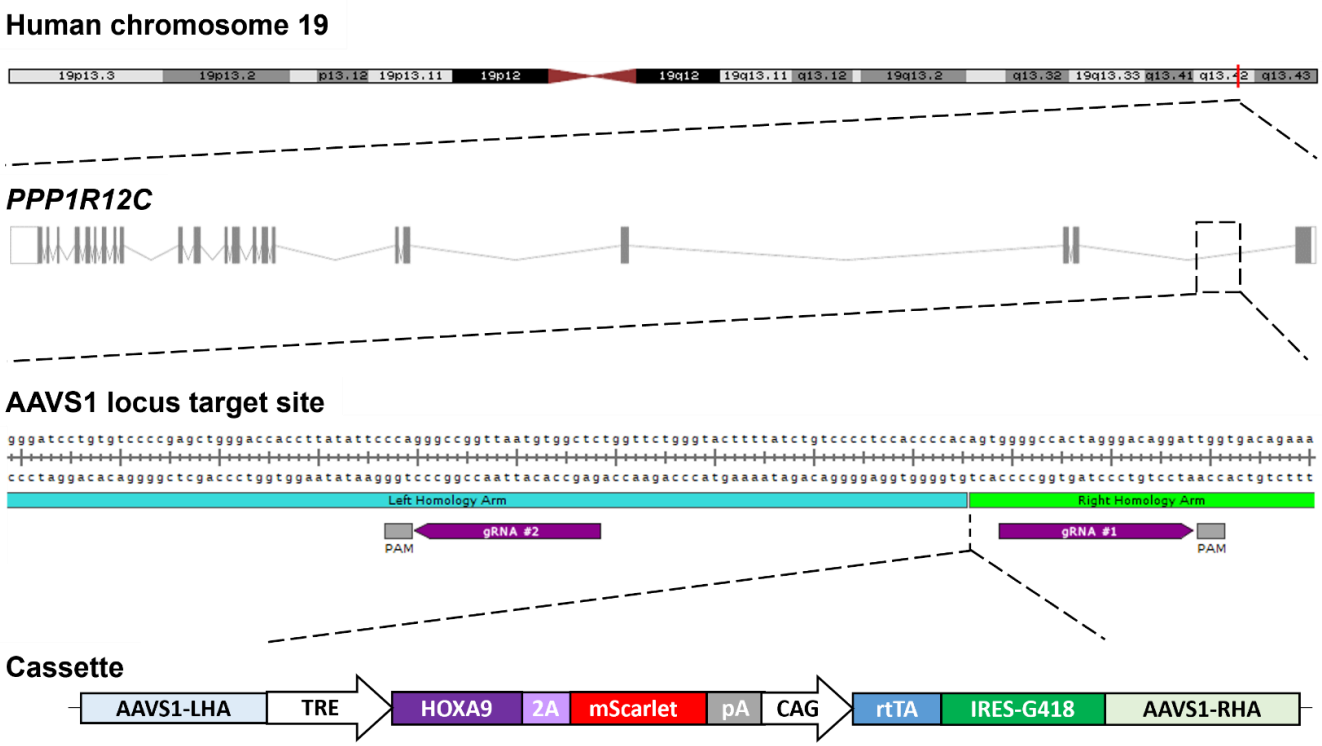


Figure 3.1: Schematic of AAVS1 locus targeting strategy

CRISPR/Cas9n was used to knock-in an inducible HOXA9-mScarlet expression cassette into intron 1 of *PPP1R12C*, the AAVS1 locus. The location of the two gRNAs and their respective PAM sites within the AAVS1 locus are indicated. The cassette comprises of a TRE promoter driving expression of HOXA9 and mScarlet, joined via the T2A peptide, and a CAG promoter driving expression of the rtTA domain and G418 selection marker, linked via an IRES element. The cassette is flanked by arms homologous to the AAVS1 locus (AAVS1-LHA and –RNA).

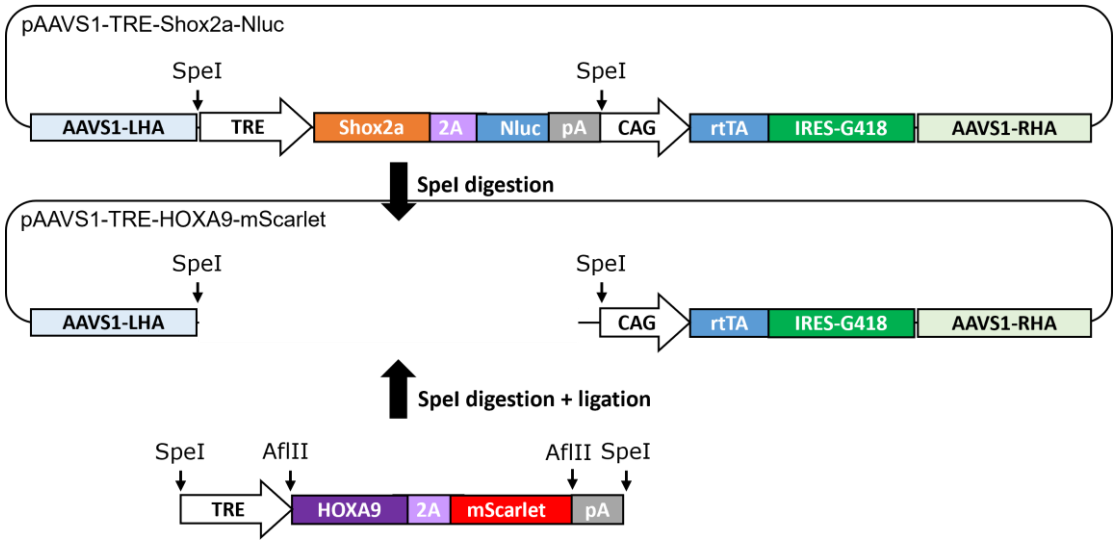


Figure 3.2: Inducible HOXA9 donor plasmid design

Illustration of pAAVS1-TRE-Shox2a-Nluc parental plasmid with location of SpeI sites. Digestion with SpeI facilitates excision of TRE-Shox2a-T2A-Nluc-pA fragment. A TRE-HOXA9-T2A-mScarlet-pA fragment generated by PCR cloning and flanked by SpeI sites is subsequently inserted into the linearised vector backbone. AflII sites incorporated during cloning of the fragment enable excision of HOXA9-T2A-mScarlet.

Briefly, the TRE promoter, including the SpeI site, was PCR amplified from pAAVS1-TRE-Shox2a-Nluc with a 29 nt overhang (encompassing an AflII site, Kozak sequence and part of HOXA9) added at the 3' end and yielding a 675 bp product (Figure 3.3A). HOXA9 was similarly amplified from pMSCVhygro-HA-hHOXA9 with addition of an equivalent 29 nt 5' overhang and 36 nt of the T2A peptide at the 3' end, generating a 932 bp product (Figure 3.3B). The complementary overhangs between the TRE and HOXA9 PCR products enabled union of the two fragments by Fusion PCR, creating a 1558 bp product (Figure 3.3C).

The initial molecular cloning plan involved fusion of TRE-HOXA9-T2A to mScarlet via the 54 bp T2A peptide using the same overlapping nucleotide strategy. The intention was to add a 36 nt T2A overhang to the 3' end of HOXA9 and 5' end of mScarlet, creating an overlap of 18 bp and facilitating fusion of the two fragments by Fusion PCR. However, despite numerous attempts, addition of the 36 nt overhang to mScarlet was not possible. A restriction enzyme digestion approach was thus adopted to join the two fragments

together (Figure 3.5A). Thus, the remaining 18 nt of the T2A peptide and a BsiWI site were added to the TRE-HOXA9-T2A fragment, generating the 1588bp TRE-HOXA9-T2A BsiWI fragment (Figure 3.3C). Correct insertion and union of all components was confirmed by Sanger Sequencing (Figure 3.3D).

The mScarlet fragment was PCR amplified from pmScarlet_H_C1 (Addgene #85043) with a BsiWI and AflII site added at the 5' and 3' end respectively, creating a 723 bp product (Figure 3.4A). To create a complementary overhang, AflII was added to the 5' end of pA tail amplified from pAAVS1-TRE-Shox2a-Nluc (Figure 3.4B). The two PCR products were then joined together by Fusion PCR generating the 990 bp BsiWI-mScarlet-pA fragment (Figure 3.4C), with correct insertion verified by Sanger Sequencing (Figure 3.4D).

TRE-HOXA9-T2A-BsiWI and BsiWI-mScarlet-pA fragments were digested with BsiWI (Figure 3.5B) and ligated to form the 2560 bp TRE-HOXA9-T2A-mScarlet-pA segment (TRE-HOXA9-mScarlet) (Figure 3.5C). Ligation was confirmed by Sanger Sequencing (Figure 3.5D).

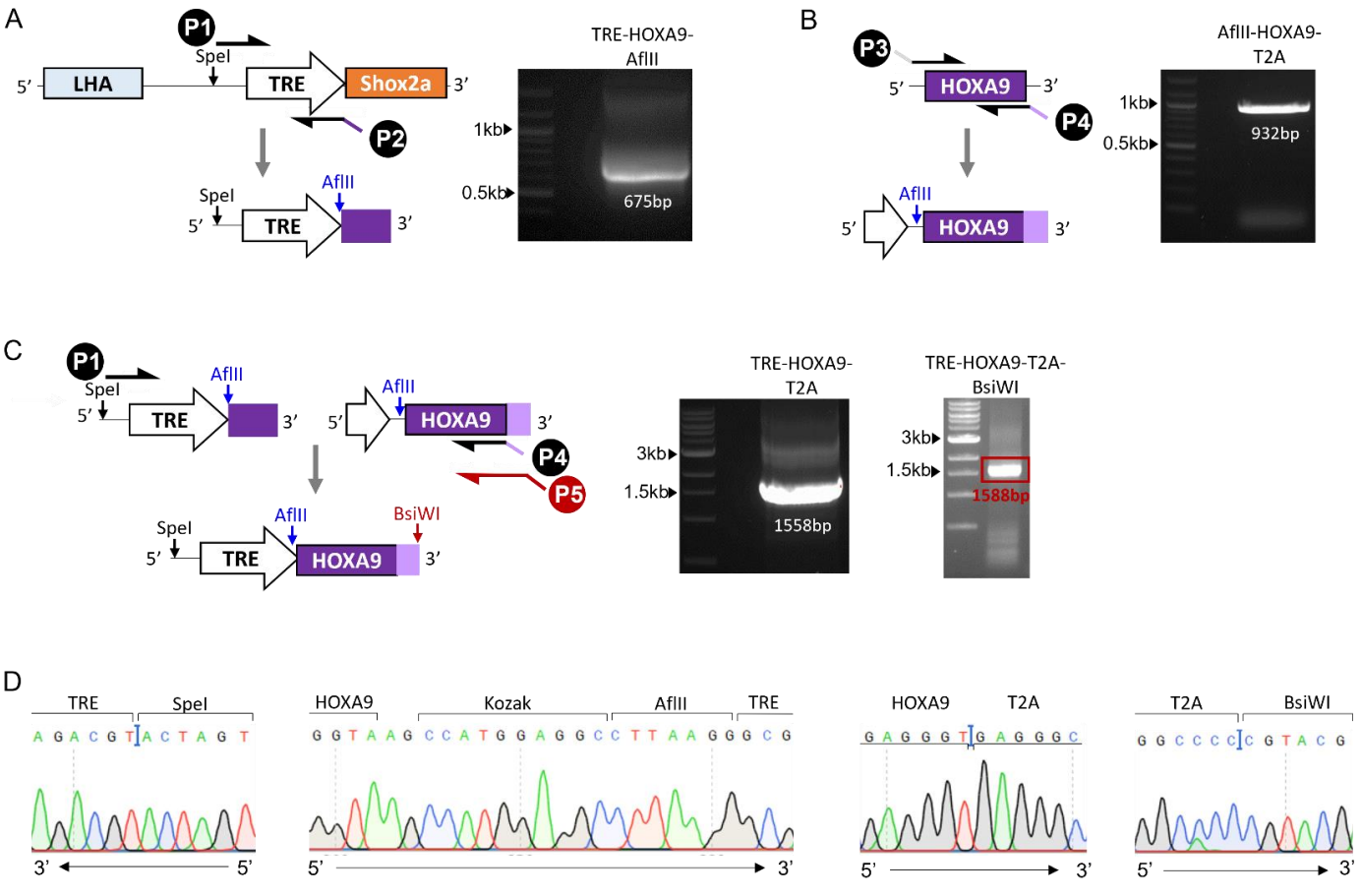


Figure 3.3: Building the TRE-HOXA9-T2A fragment

Schematic illustration of overhang addition to TRE and HOXA9 fragments by PCR, with expected product shown in agarose gel images. Primers used are indicated by black circles (see Table 2.11). (A) PCR of the TRE promoter with 3' addition of an AflII site and part of HOXA9. (B) PCR of HOXA9 with 5' addition of an AflII site and part of TRE and 3' addition of 36 nucleotides of the T2A peptide. (C) Fusion PCR of the TRE and HOXA9 PCR products resulting in the 1558 bp TRE-HOXA9 fragment. Addition of the remaining T2A nucleotides and a BsiWI to TRE-HOXA9 by PCR creating the 1588 bp TRE-HOXA9-T2A fragment. (D) Chromatograms demonstrate fusion of overlapping sequences and addition of restriction enzyme sites.

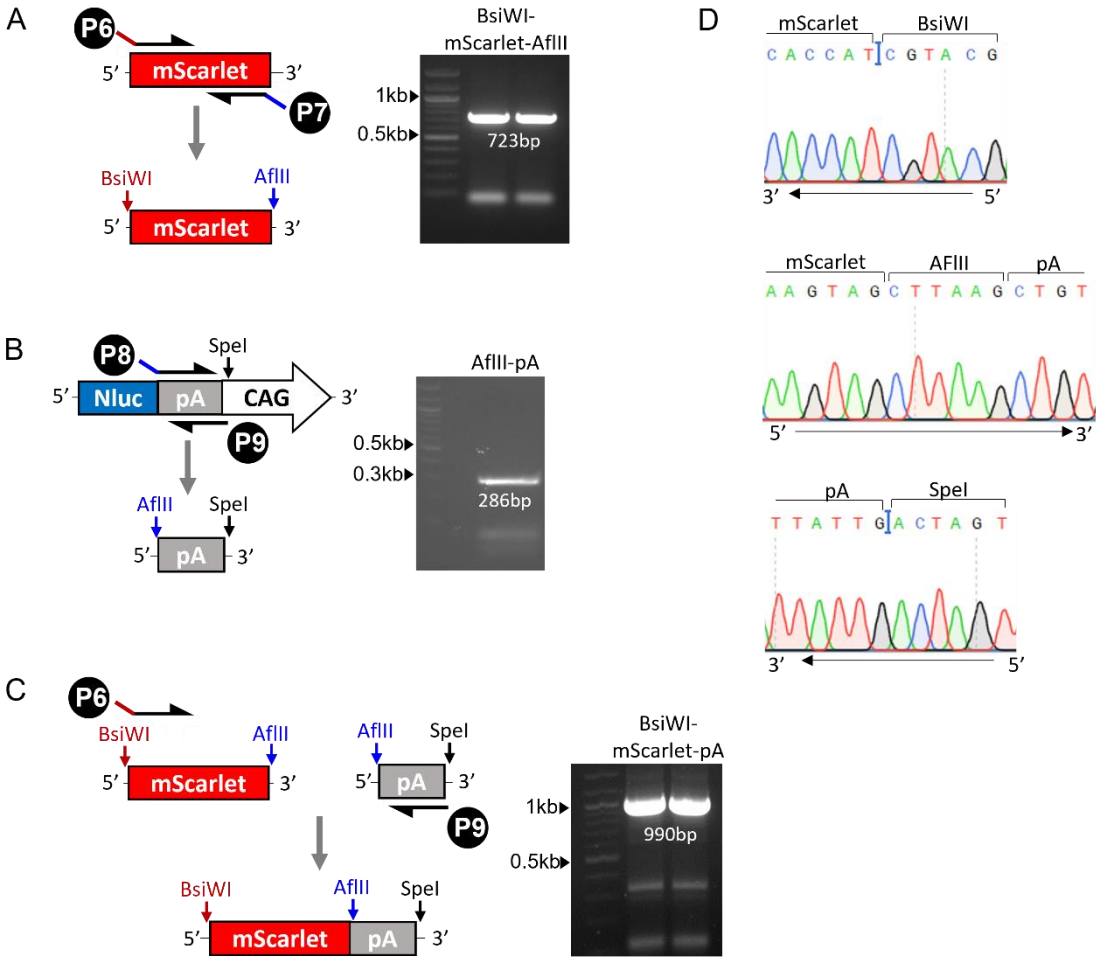


Figure 3.4: Building of the BsiWI-mScarlet-pA fragment

Schematic illustration of overhang addition to mScarlet and pA fragments by PCR, with expected product shown in agarose gel images. Primers used are indicated by black circles (see Table 2.11). (A) PCR of mScarlet with addition of BsiWI at the 5' end and AflII at the 3' end. (B) PCR of pA with addition of an AflII site at the 5' end. (C) Fusion PCR of the mScarlet and pA PCR products resulting in the 990 bp BsiWI-mScarlet-pA fragment. (D) Chromatograms demonstrate fusion of overlapping sequences and addition of the two restriction enzyme sites.

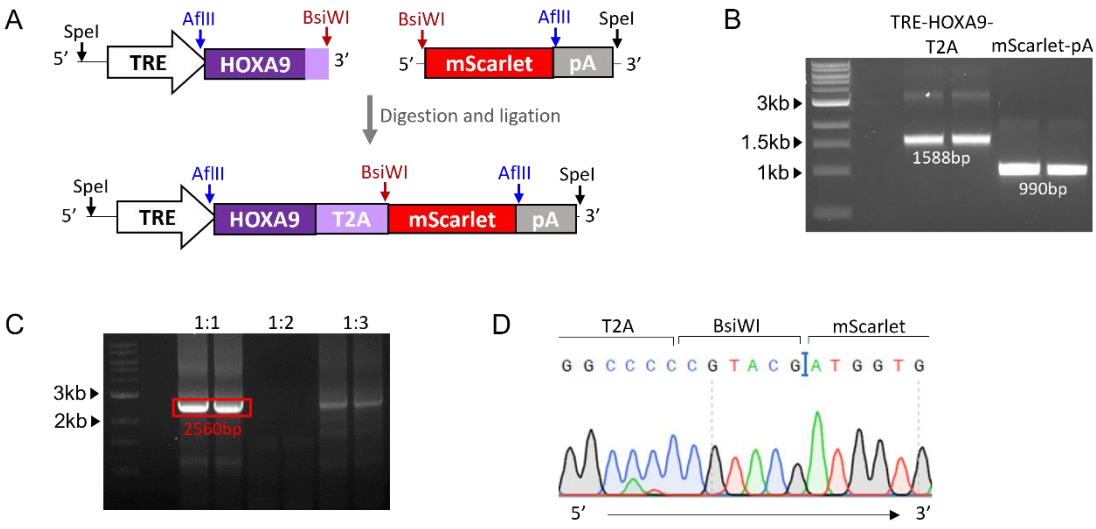


Figure 3.5: Building of the final TRE-HOXA9-T2A-mScarlet-pA fragment

(A) Schematic illustrating generation of the TRE-HOXA9-T2A-mScarlet-pA fragment by restriction enzyme digestion and ligation. (B) Digestion of TRE-HOXA9-T2A-BsiWI and BsiWI-mScarlet-pA fragments with BsiWI and purification on an agarose gel. (C) Ligation of the digestion products in a 1:1 molar ratio resulting in the final product (red box). (D) Chromatogram confirms ligation of TRE-HOXA9-T2A and mScarlet-pA fragments.

Digestion and ligation of the TRE-HOXA9-mScarlet segment into the backbone plasmid was hindered by insufficient quantity of the segment following purification. Attempts to repeat the BsiWI digestion and PCR of the whole segment also yielded low product. Therefore, to generate multiple copies and sufficient quantity of TRE-HOXA9-mScarlet, the segment was ligated into the blunt-end ligation vector pMiniT 2.0 (Figure 3.6A) and amplified in 5-alpha bacteria. Ligation was confirmed in an AflII diagnostic digest which demonstrated 5/6 purified plasmid constructs had the predicted bands 1636 bp and 3513 bp (Figure 3.6B). As TRE-HOXA9-mScarlet (2560 bp) is approximately the same size as pMiniT (2588 bp), one purified plasmid (construct 3) was taken forward and digested with SpeI and FspI to excise the TRE-HOXA9-mScarlet segment. FspI is unique to pMiniT and double digestion yields three products, with a band at 2560 bp corresponding to the TRE-HOXA9-mScarlet segment, which was subsequently purified (Figure 3.6C).

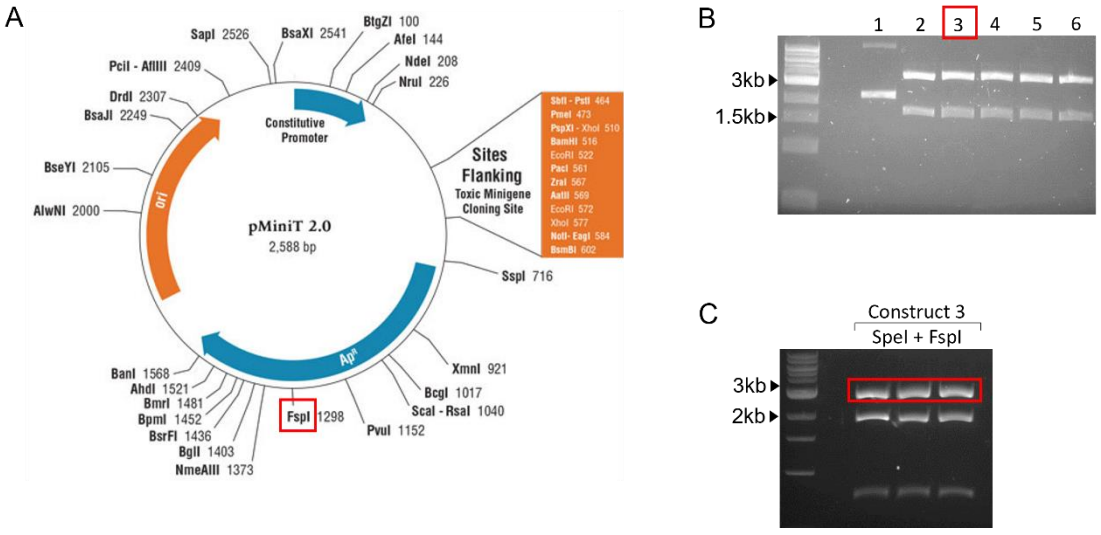


Figure 3.6: Amplification of TRE-HOXA9-mScarlet

(A) Map of pMiniT blunt-end ligation vector. Red box indicates location of FspI site. (B) AflII digestion of plasmids (1 – 6) isolated from bacteria transformed with pMiniT-TRE-HOXA9-mScarlet ligation reaction. (C) Digestion of plasmid 3 (red box in B) with SpeI and FspI. Red box indicates the TRE-HOXA9-mScarlet band.

The parental pAAVS1-TRE-SHOX2A-Nluc vector was digested with SpeI to isolate the 8899 bp backbone plasmid, which was dephosphorylated to prevent self-ligation (Figure 3.7A). The TRE-HOXA9-mScarlet fragment isolated from the pMiniT vector and the linearised backbone were then ligated together at different molar ratios and transformed into 5-alpha bacteria. Bacterial colonies were screened via colony PCR for correct insertion using primers which amplify across the 5' SpeI site. Several colonies were identified to have the expected 1258 bp product (Figure 3.7B). One colony was taken forward (colony 3), and the plasmid DNA was purified and subsequently assessed via diagnostic restriction enzyme digest and Sanger sequencing. Insertion of TRE-HOXA9-mScarlet in the correct orientation produces a second SalI site and yields two products, 3188 bp and 8253 bp, following digestion. SpeI digestion similarly produces two products, 2542 bp and 8899 bp. BsiWI is unique to the plasmid and thus digestion results in linearisation (see Figure 3.8 for plasmid map and restriction enzyme sites). The diagnostic digest produced the predicted bands (Figure 3.7C) while PCR of the junctions joining TRE-HOXA9-mScarlet to the backbone plasmid and subsequent sequencing (Figure 3.7D) further confirmed generation of the inducible HOXA9 donor plasmid pAAVS1-TRE-HOXA9-mScarlet.

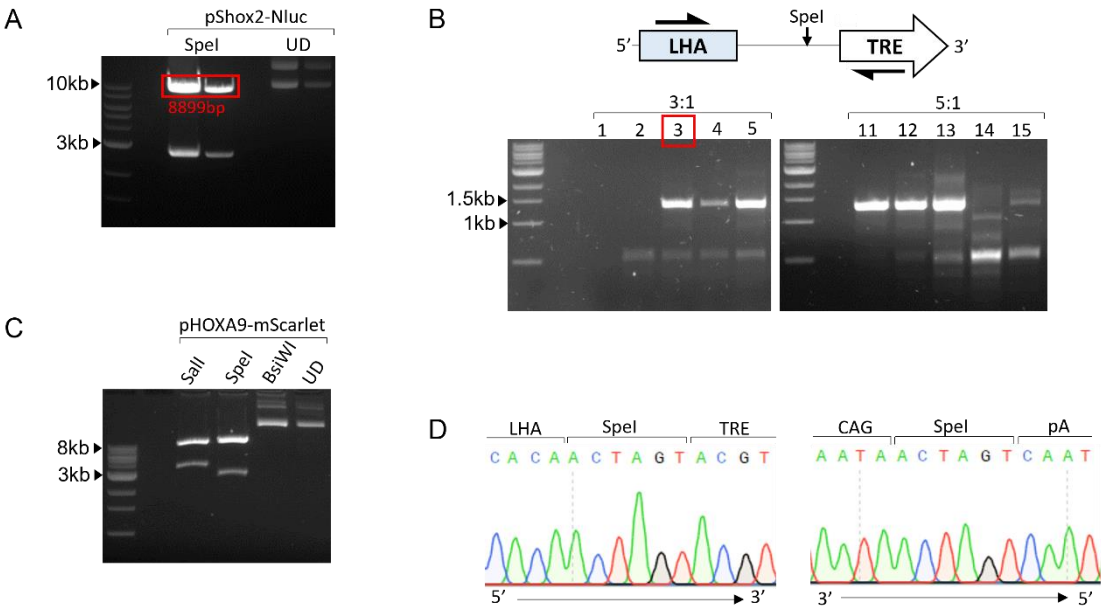


Figure 3.7: Restriction enzyme digestion and ligation to create pAAVS1-TRE-HOXA9-mScarlet

(A) Digestion of pAAVS1-TRE-SHOX2A-Nluc (pSHOX2A-Nluc) with SpeI. Red box indicates backbone vector. (B) Colony PCR of bacteria transformed with TRE-HOXA9-mScarlet and backbone vector ligation reaction at 3:1 or 5:1 molar ratios. Several colonies produce the expected PCR product at 1258 bp. Schematic indicates approximate location of colony PCR primers. (C) Diagnostic digest of pAAVS1-TRE-HOXA9-mScarlet (pHOXA9-mScarlet) isolated from bacterial colony 3 (red box in B) with SalI, SpeI and BsiWI. (D) Chromatogram indicates ligation of the left homology arm (LHA) to the TRE promoter and the pA tail to the CAG promoter. UD; undigested plasmid.

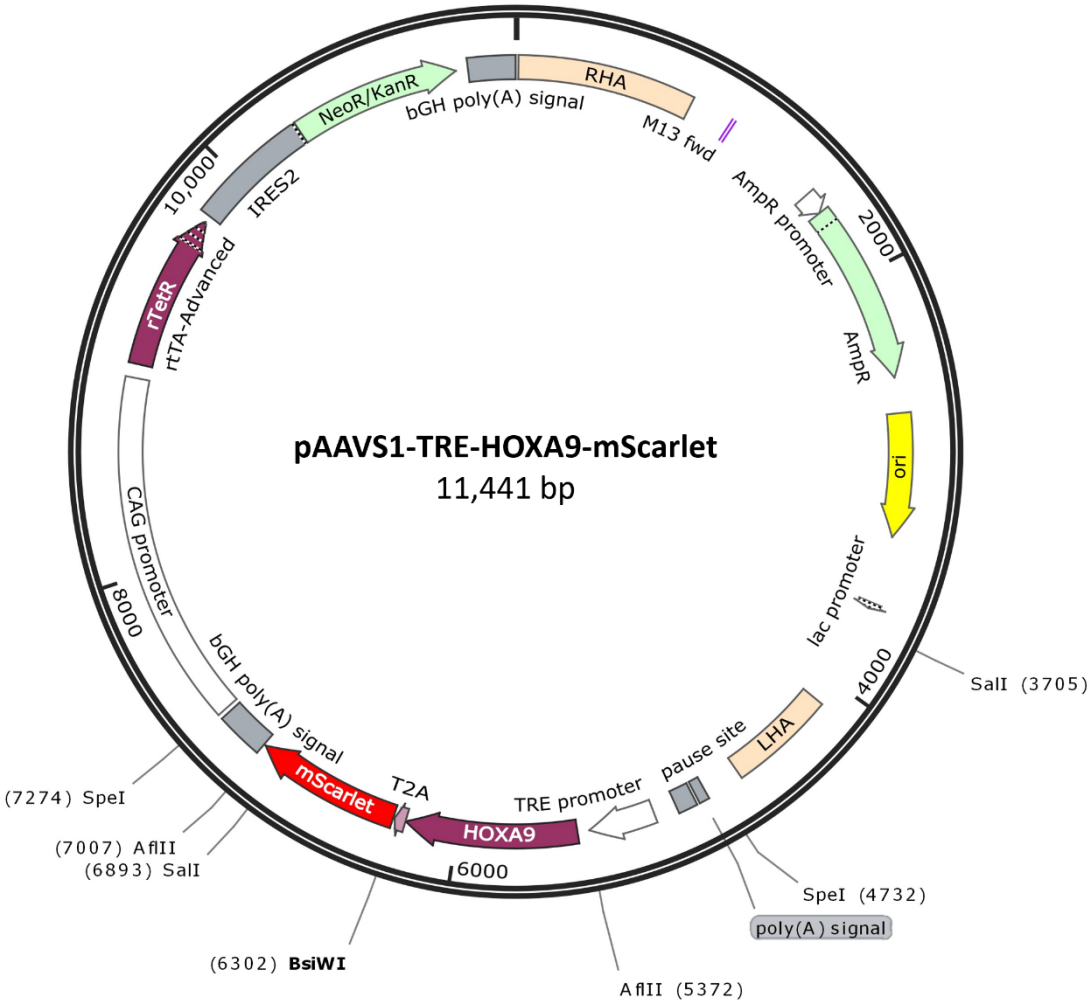


Figure 3.8: Map of CRISPR donor template plasmid, pAAVS1-TRE-HOXA9-mScarlet

SnapGene map of pAAVS1-TRE-HOXA9-mScarlet. Sites of the restriction enzymes used for diagnostic digest (SalI, SpeI and BsiWI) are shown. AflII sites introduced for excision of HOXA9-T2A-mScarlet fragment are also shown.

3.3.3 Transient transfection of pAAVS1-TRE-HOXA9-mScarlet

To test if pAAVS1-TRE-HOXA9-mScarlet can produce *HOXA9* and mScarlet upon DOX stimulation, 1 µg or 3 µg of pAAVS1-TRE-HOXA9-mScarlet was transfected via nucleofection into RPAT hiPSCs. After 24h, the transfected cells were treated with 1 µg/mL DOX, a concentration commonly used in systems with the same TRE promoter (tightTRE) (Tang et al., 2013). Transfected cells were then imaged using the Operetta® fluorescence microscope and samples were taken for RNA extraction and qPCR 24h later. mScarlet expression was only observed in cells treated with DOX (Figure 3.9A), suggesting high sensitivity of the inducible system to DOX. QPCR similarly showed *HOXA9* was only elevated upon DOX treatment, with higher expression seen in cells treated with 3 µg donor plasmid (Figure 3.9B). Quantification of mScarlet images (see Figure 4.2 for method) revealed a similar expression trend to *HOXA9* qPCR (Figure 3.9C), demonstrating mScarlet expression is indicative of HOXA9 and thus a suitable approach to assess HOXA9 expression in live cells.

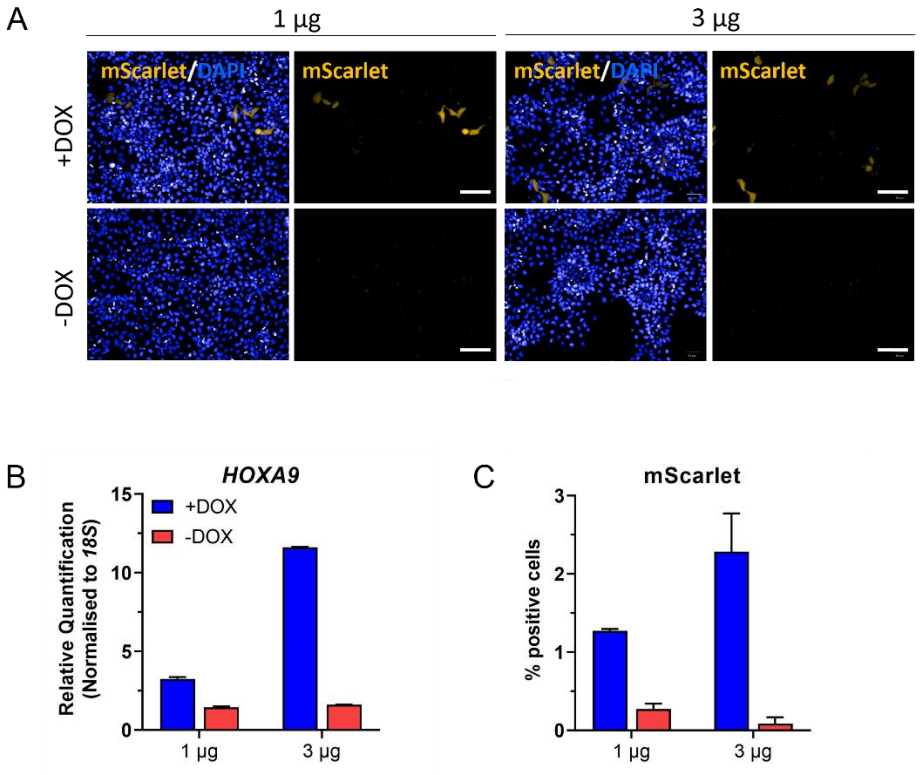


Figure 3.9: Transient transfection of inducible HOXA9 donor plasmid

(A). Fluorescence microscope images taken using the Operetta® imaging system of RPAT hiPSCs 48h after transfection with 1 µg or 3 µg of pAAVS1-TRE-HOXA9-mScarlet and 24h after treatment with 1 µg/mL DOX. Nuclei are co-stained with DAPI. Scale bar = 100 µm. (B) QPCR for HOXA9. (C) Quantification of mScarlet expression using Columbus™ analysis software. Results are mean ± SD, n = 3 technical replicates. For mScarlet quantification, each technical replicate represents 10 fields of view.

3.3.4 Survival analysis

Prior to CRISPR transfection, the concentration of G418 antibiotic to be used for clonal selection was determined. A G418 kill-curve analysis was performed to identify the minimum concentration of G418 which will kill all cells after 72h. RPAT hiPSCs were treated with a range of G418 concentrations, from 25 – 200 µg/mL. These concentrations were selected based on previous publications which used concentrations ranging from 25 µg/mL to 175 µg/mL in hiPSC lines (Cerbini et al., 2015; Oceguera-Yanez et al., 2016; Reinhardt et al., 2013). A concentration of 50 µg/mL was determined to kill all cells after 72h (Figure 3.10) and was chosen for use in clonal selection.

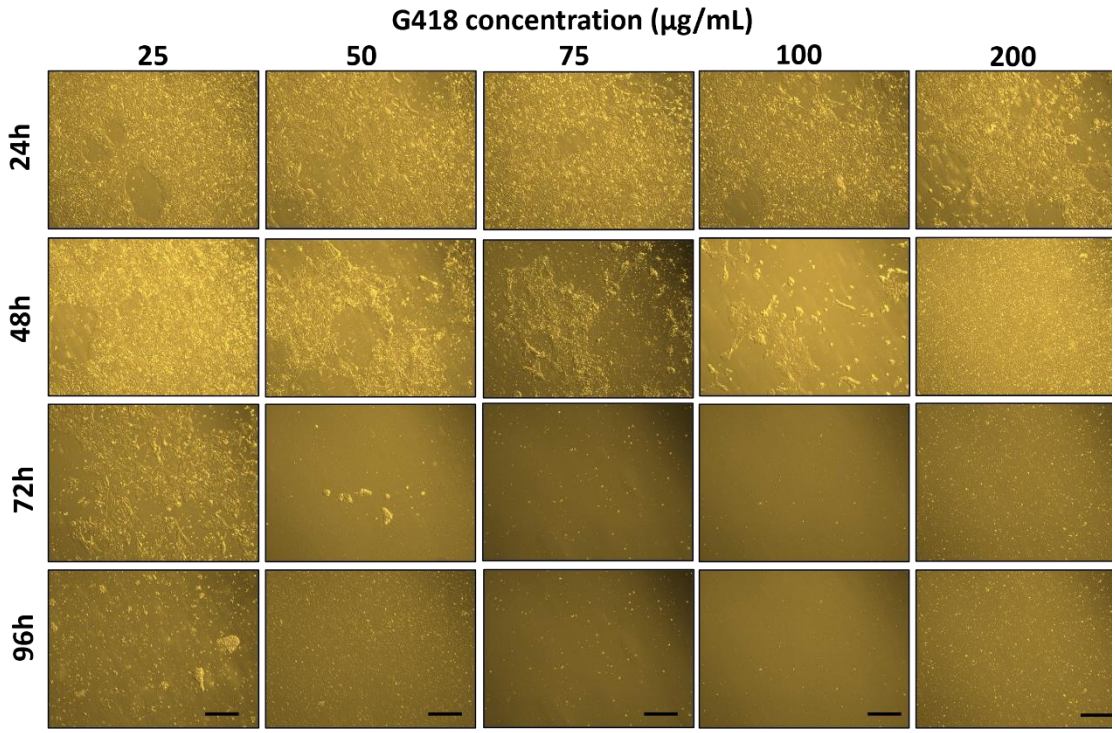


Figure 3.10: Identification of optimal G418 concentration for selection of targeted clones

Representative light microscope images of RPAT hiPSCs treated with different concentrations (25 – 200 g/mL) of G418 antibiotic every 24h for 4 days. Scale bar = 200 µm.

3.3.5 Targeting of the AAVS1 locus

The Cas9n D10A plasmid, two gRNA plasmids and pAAVS1-TRE-HOXA9-mScarlet were introduced into RPAT hiPSCs (p18) using the Amaxa™ 4D-Nucleofector™ and seeded in 6-well plates (see Figure 3.11 for workflow). After 72h, transfected cells were passaged again and seeded into 6-well plates at a range of different densities (1:10 – 1:30). Untransfected cells were also seeded at the same densities to serve as a negative control. The following day G418 selection was started. Significant cell death in untransfected wells was observed after 72h but treatment was maintained for up to 6 days, by which time compact, large, round colonies had formed in several of the transfected cell wells (Figure 3.11). Individual clones were picked and seeded into a single well of a 96-well plate. Clones were subsequently expanded to 12-well plates, where genomic DNA (gDNA) samples were taken for genotyping and clones were cryopreserved. G418 at 50 µg/mL was maintained in the media throughout expansion.

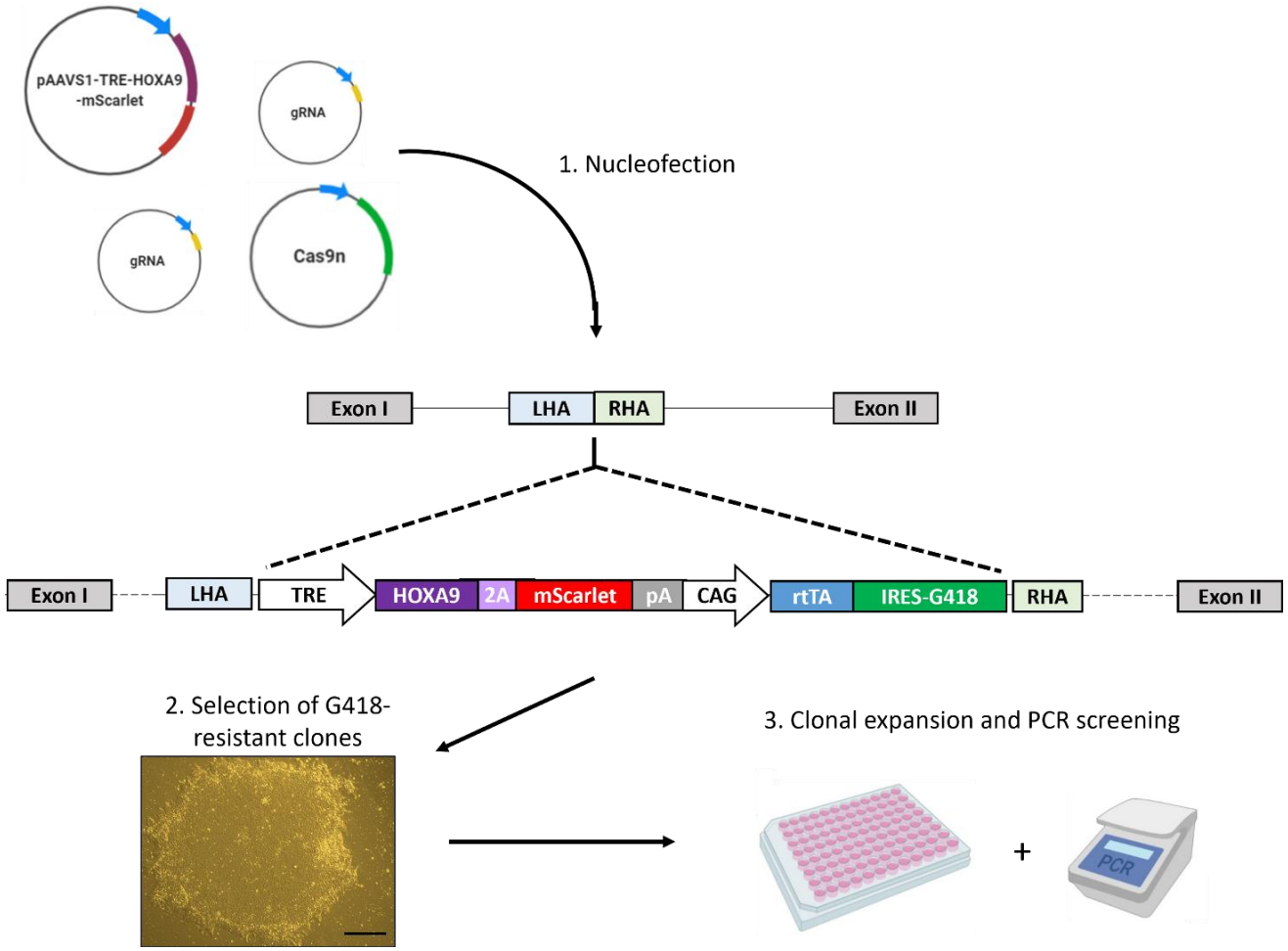


Figure 3.11: CRISPR/Cas9n workflow

The inducible HOXA9 donor, gRNAs and Cas9n D10A plasmids were introduced into the AAVS1 locus of RPAT hiPSCs via nucleofection. Targeted clones were identified through G418 selection and expanded prior to genotyping PCR to verify correct integration. Scale bar = 200 µm.

Optimisation of genotyping PCR primers for 5' and 3' integration of the cassette into the AAVS1 locus was performed using genomic DNA from the inducible Shox2 reporter cell line and unedited RPAT hiPSCs (Supplementary Figure 1A, B). Following expansion and selection, the surviving clones were screened using the optimised genotyping primers (Figure 3.12A). All clones were determined to have the 933 bp 5' integration band (Figure 3.12B), while 12/14 clones had the 2204 bp 3' integration band (Figure 3.12C). To confirm the bands correspond to correct integration, the gel bands from clone number 54 were excised, purified and sequenced. Sequencing across 5' (Figure 3.12D) and 3' integration sites (Figure 3.3E) confirmed insertion of the inducible HOXA9 cassette.

To determine if both or one allele had been targeted, the AAVS1 locus was PCR amplified using primers optimised in unedited RPATs and the heterozygous inducible Shox2 reporter cell line (Figure 3.13A, B). The AAVS1 locus was amplified in all clones, demonstrating all clones are heterozygous for the cassette (Figure 3.13C). Attempts to PCR amplify the edited allele were unsuccessful, likely due to difficulties in amplifying the large product (8156 bp).

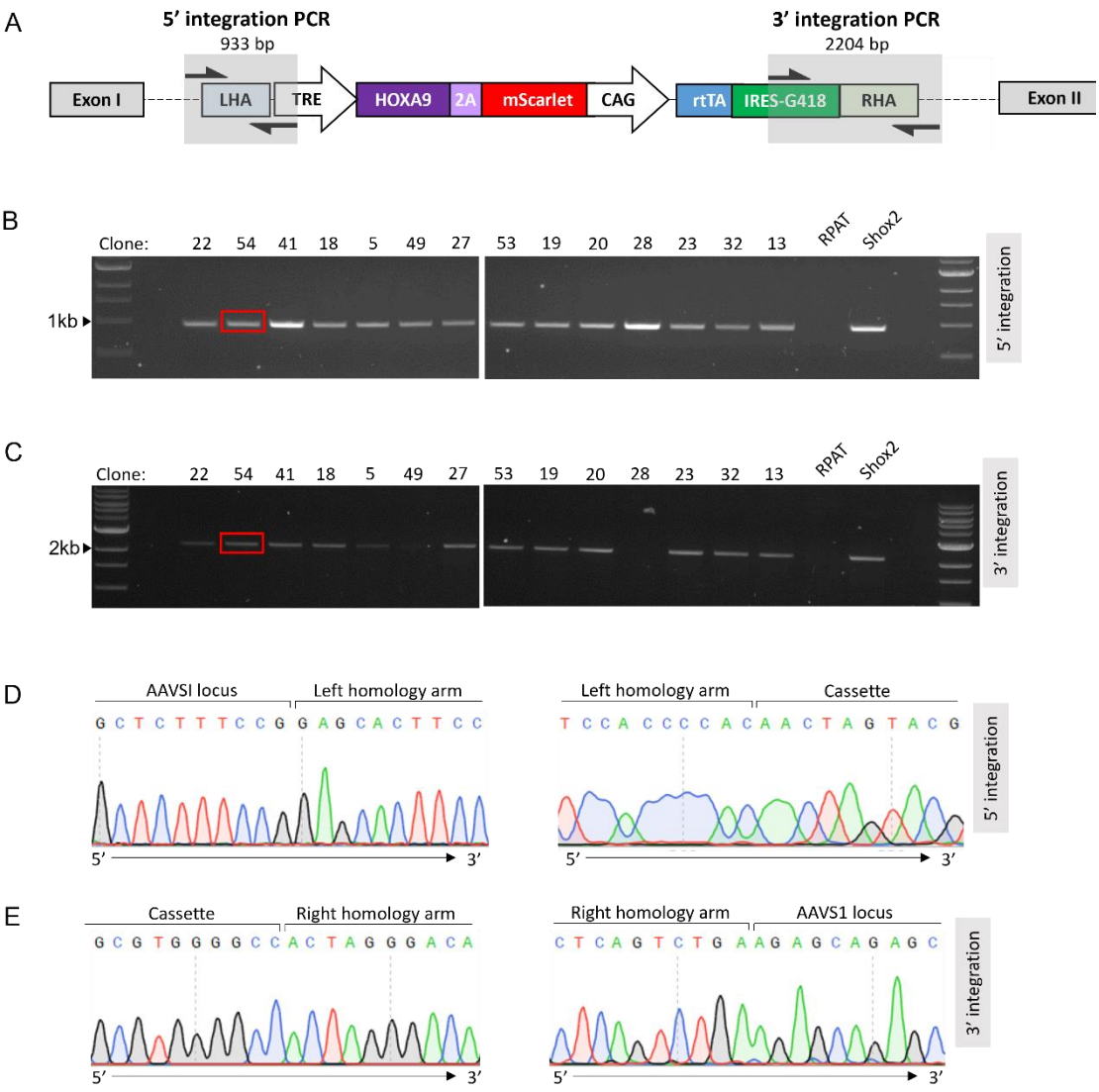


Figure 3.12: Genotyping PCR and sequencing identifies correctly targeted clones

(A) Schematic demonstrating location of the 5' and 3' integration primers and expected PCR product sizes. (B) 5' integration PCR showing all clones have the correct integration product. (C) 3' integration PCR showing 12/14 clones have the correct integration product. (D - E) Representative chromatogram from sequencing of clone 54 (red box) confirms integration of the cassette into the AAVS1 locus at the 5' and 3' integration sites.

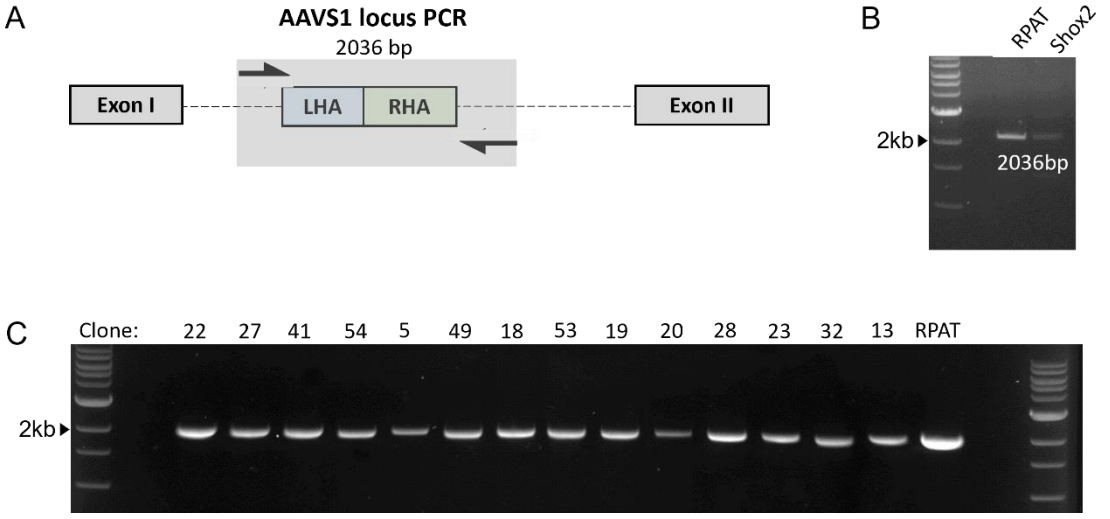
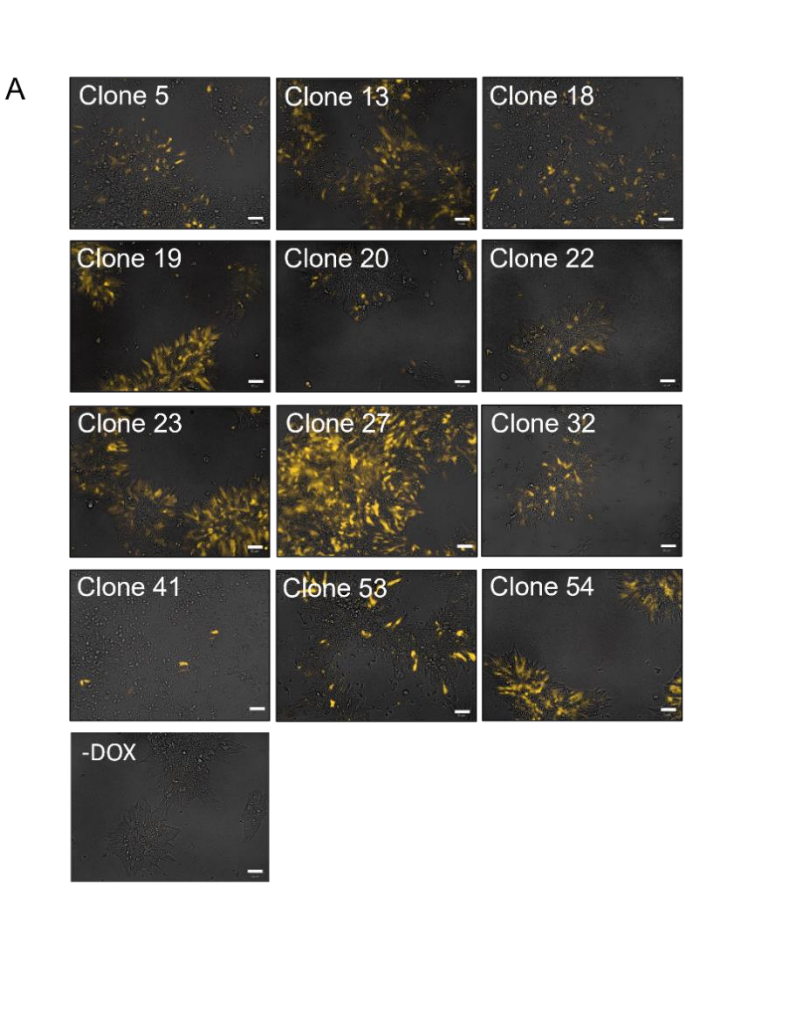


Figure 3.13: All clones are heterozygous

(A) Schematic demonstrating location of AAVS1 locus primers used to determine if an allele has been untargeted. (B) PCR testing of AAVS1 locus primers using genomic DNA from unedited RPAT and heterozygous inducible Shox2 reporter cell lines. (C) PCR of all clones with AAVS1 locus primers.

3.3.6 Inducible HOXA9 clones are functional and specific for DOX

The ability of DOX to induce upregulate of HOXA9 and mScarlet was tested in all twelve correctly targeted inducible HOXA9 (iA9) clones. iA9 clones were treated with 1 µg/mL DOX for 24h, after which clones were imaged using the Operetta® imaging system to identify mScarlet expression and RNA was extracted for qPCR analysis of *HOXA9* expression. All clones were sensitive and specific for DOX, only expressing mScarlet following DOX treatment (Figure 3.14A). Quantification of mScarlet expression using Columbus™ analysis software (Figure 3.15) showed high variability between different clones, with the proportion of cells determined to express mScarlet ranging from 0.9% to 44.7% (Figure 3.14C). The intensity of mScarlet expression was also iA9 clone dependent (Figure 3.14D). QPCR for *HOXA9* revealed a similar expression trend to mScarlet for most of the clones, apart from iA9 clones 19, 20 and 27 (Figure 3.14B), possibly as a result of discrepancies in RNA and protein expression levels.



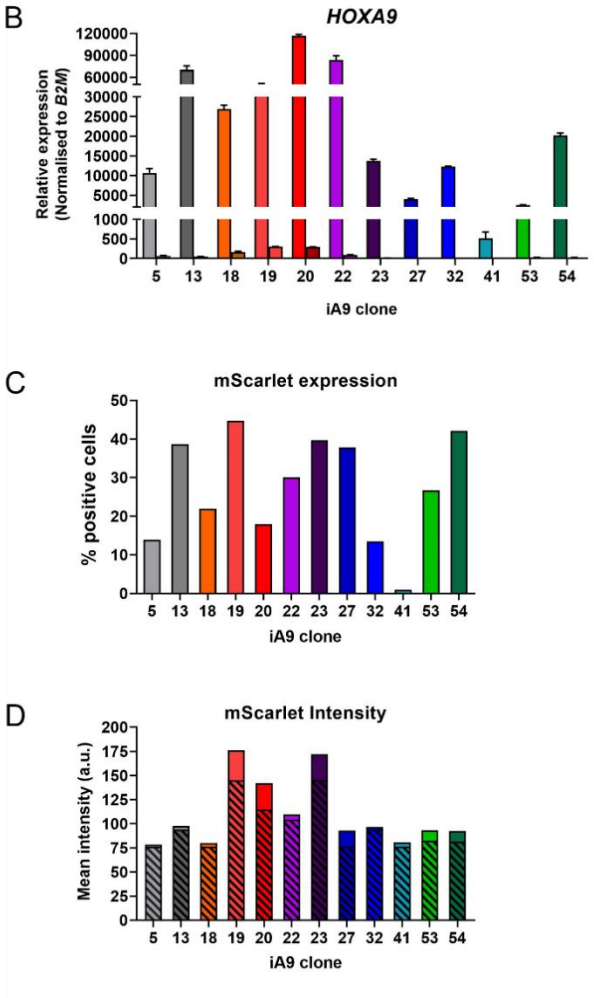


Figure 3.14: All clones upregulate *HOXA9* and mScarlet in the presence of DOX

(A) Representative images of each clone 24h after treatment with 1 µg/mL DOX. An uninduced (-DOX) clone is also shown. Scale bar = 50 µm. (B) QPCR for HOXA9 following 24h treatment with or without 1 µg/mL DOX. Results are mean ± SD, n = 3 technical replicates. (C) Quantification of mScarlet images after DOX treatment performed using Columbis™ analysis software. (D) Mean intensity of mScarlet expression. Bars with diagonal black lines indicate intensity threshold. Results are based on analysis of 10 fields of view.

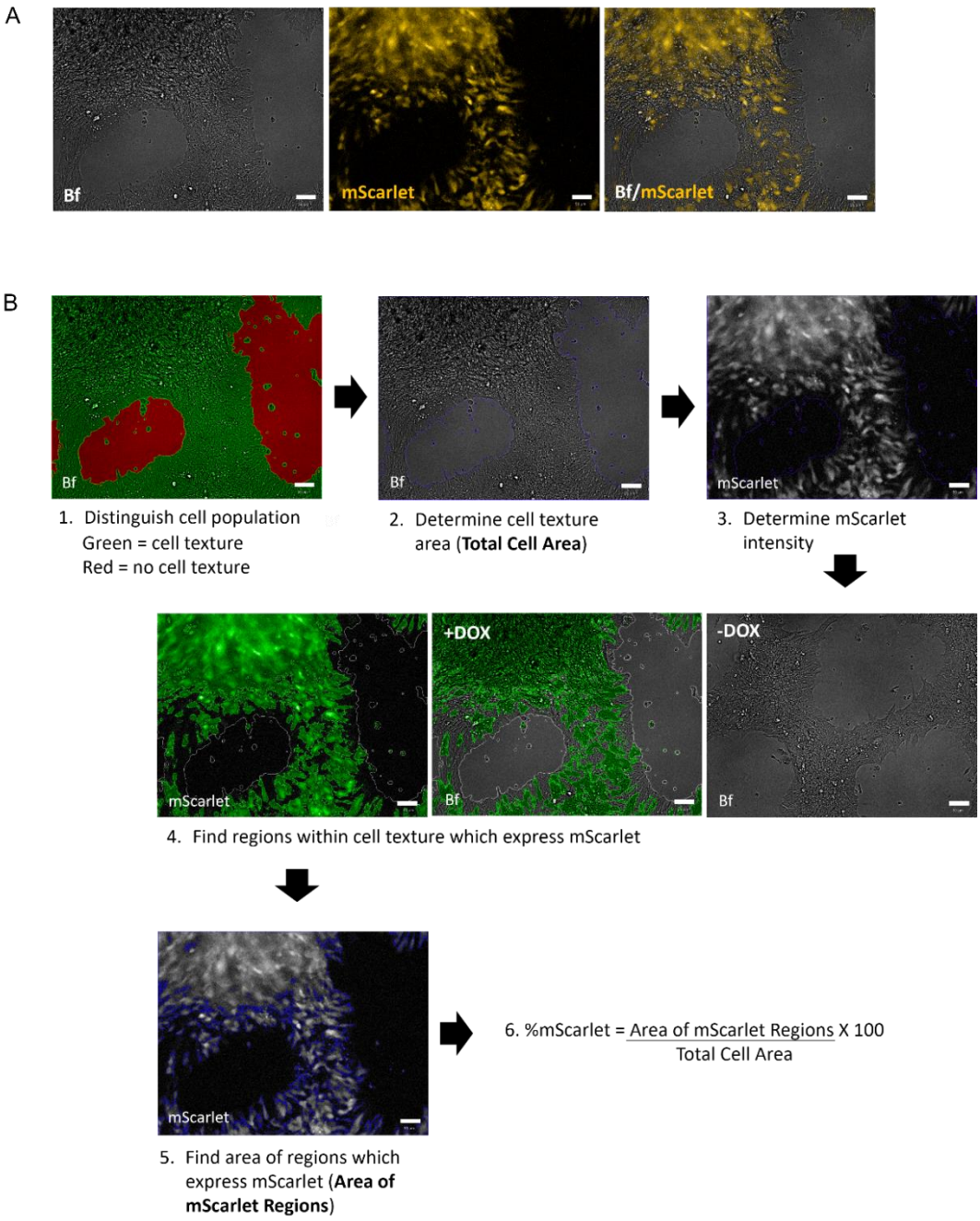


Figure 3.15: Quantification of mScarlet images using Columbus™

(A) Example image of iA9 clone 27 following treatment with 1 µg/mL DOX for 24h. The image represent a single filed of view. (B) Overview of quantification steps. First, the cell area is calculated (steps 1 – 2). Next, based on mScarlet intensity (step 3), regions within the cell population which express mScarlet are identified (step 4). mScarlet is only detected if cells have been treated with DOX (+DOX vs –DOX images). The area of mScarlet-expressing regions is then calculated (step 5). The percentage of mScarlet expressing cells is calculated (step 6). Scale bar = 50 µm.

As inducible HOXA9 (iA9) clone 54 exhibited both high *HOXA9* and mScarlet expression (42%), further assessments and experiments were initially performed on this clone. To determine if HOXA9 protein was produced, immunostaining was performed on iA9 clone 54 iPSCs after 24h DOX treatment and compared to control (-DOX). Due to the poor selection and specificity of HOXA9-specific antibodies, advantage was taken of the fact that the 2A peptide (from T2A) remains fused to HOXA9 following translation. A specific antibody against the self-cleaving peptide demonstrated that, similar to mScarlet, 2A is only expressed in DOX-treated cells (Figure 3.17A). Consistent with the expression profile of HOXA9, 2A remained in the nucleus while mScarlet was expressed in both the nucleus and cytoplasm. Quantification of staining revealed mScarlet (Figure 3.17B, 24.7%) and 2A (Figure 3.17C, 23.6%) were expressed at similar levels, demonstrating mScarlet is an appropriate indicator of HOXA9 expression.

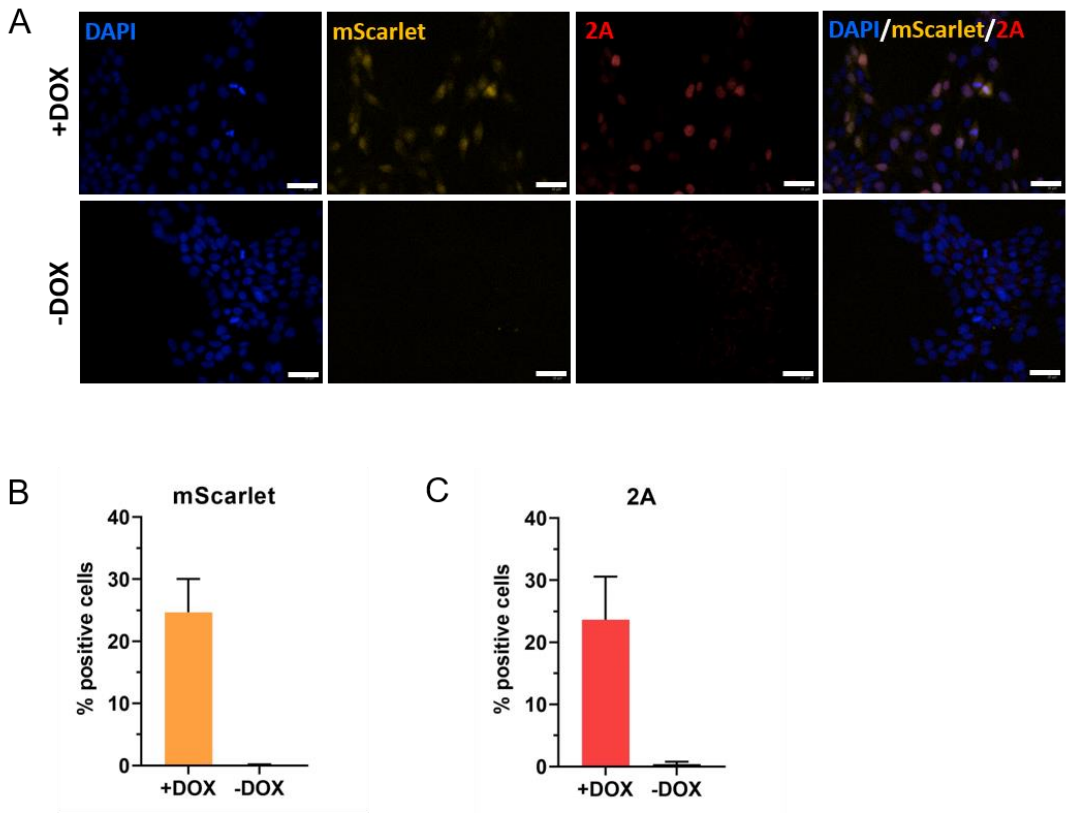


Figure 3.16: HOXA9 and mScarlet proteins induced expressed at similar levels following DOX treatment

(A) Representative images of iA9 clone 54 iPSCs 24h after +/- DOX treatment and staining for 2A. Nuceli were co-stained with DAPI. (B – C) Quantification of mScarlet expression (B) and 2A staining (C) using Columbus™ software. Results are from 2 independent biological replicates with 3 technical replicates performed for each. Mean ± SD is calculated based on technical replicates. Scale bar = 40 µm.

3.3.7 Inducible HOXA9 clones retain their pluripotency following DOX treatment

The pluripotency of iA9 clone 54 iPSCs was assessed via immunostaining for OCT4 and NANOG and qPCR for *OCT4*, *NANOG* and *SOX2* following 24h +/- 1 µg/mL DOX treatment. Immunostaining (Figure 3.17A) and qPCR (Figure 3.17B) demonstrated the iA9 clone is pluripotent and that addition of DOX or upregulation of HOXA9 does not affect pluripotency, in agreement with a previous study (Domingo-Reines et al., 2017). QPCR was performed relative to unedited RPAT iPSCs, with RPAT treated +/- DOX also included as a control to verify DOX treatment does not adversely affect pluripotency. iA9 clone 54 iPSCs also retained a normal karyotype following gene editing (Figure 3.18C).

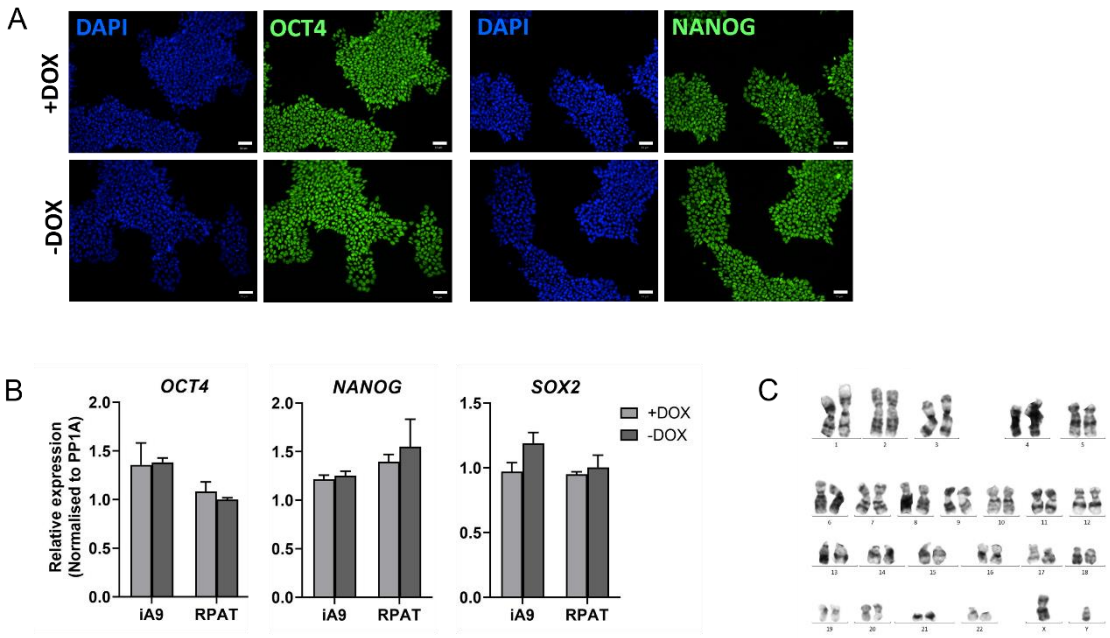


Figure 3.17: iA9 clone 54 iPSCs retain pluripotency in response to DOX

(A) Representative images of iA9 clone 54 iPSCs immunostained for OCT4 and NANOG 24h after +/- DOX treatment. Nuceli were co-stained with DAPI. Scale bar = 50 µm. (B) QPCR for *OCT4*, *NANOG* and *SOX2* in iA9 clone 54 and RPAT iPSCs treated +/- DOX for 24h. Results were normalised to untreated RPAT iPSCs. Results are mean ± SD, n = 3 technical replicates. (C) Representative karyogram of iA9 clone 54 iPSCs.

3.3.8 Loss of DOX sensitivity following mesoderm specification in iA9 clone differentiation to haematopoietic cells

The ability of iA9 clones to maintain their sensitivity to DOX during lineage differentiation was assessed by differentiation of iA9 clone 54 iPSCs to haematopoietic progenitors using the STEMdiff™ Hematopoietic kit (Figure 3.18). iA9 clone 54 differentiation cultures were treated once with 1 µg/mL DOX at different timepoints throughout haematopoietic differentiation. 24h after DOX induction, mScarlet and *HOXA9* expression was measured via fluorescence microscopy and qPCR, respectively. During differentiation, robust expression of both mScarlet and *HOXA9* was detected on day 3, however, this steadily declined with DOX no longer able to induce *HOXA9* or mScarlet expression following treatment on day 6 (Figure 3.19A – C). Mesendoderm markers *MIXL1* and *T* were elevated on day 3 of differentiation indicating DOX sensitivity declines following mesendoderm induction (Figure 3.19D). Cultures not treated with DOX (DOX-

) did not produce HOXA9 or mScarlet, suggesting the overexpression system remains tightly regulated by DOX.

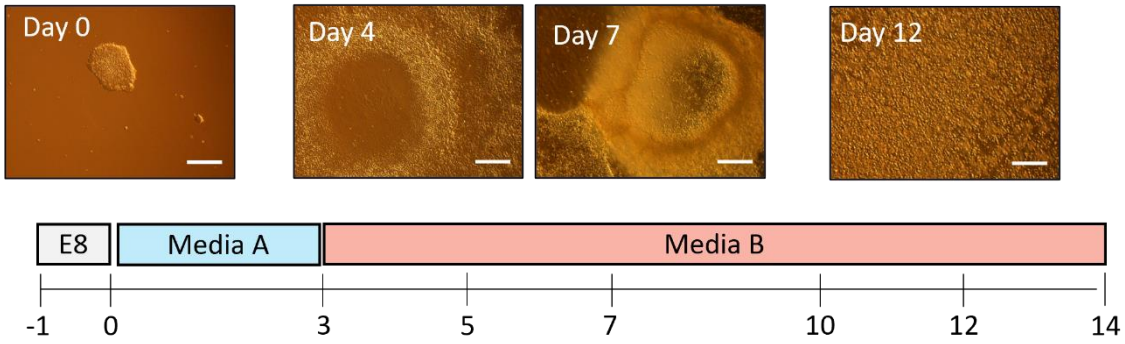


Figure 3.18: STEMdiff™ Hematopoietic kit differentiation overview

Overview of hPSC differentiation to haematopoietic cells using the STEMdiff™ Hematopoietic kit from STEMCELL Technologies. Human PSCs are seeded at low densities in clumps on Matrigel-coated 12-well plates in E8 medium. After 24h (day 0), only wells which contain 16 – 40 colonies ~200 µm in diameter are selected to continue with differentiation and media is changed to Media A for the first 3 days. On day 3, media is changed to Media B, with subsequent media changes performed on the days shown. Suspension cells are harvested on day 12 and assessed for haematopoietic potential. Representative light microscope images of RPAT hiPSCs during differentiation are shown. Scale bar = 200 µm.

Notably, the differentiation also failed to yield any suspension cells and therefore haematopoietic potential could not be fully assessed. HOXA9 expression in hPSC-based differentiations is reported to be gradually upregulated as haematopoietic commitment proceeds (Ramos-Mejía et al., 2014). However, *HOXA9* expression did not change from day 7 onwards in either +/- DOX-treated cultures. Upregulation of *KDR* and haematopoietic-specific marker *RUNX1c* suggests some haematopoietic pathways were stimulated. However, it was concluded that a more effective and reproducible differentiation procedure would be required to further assess the functionality of the iA9 system in hiPSC-derived haematopoiesis.

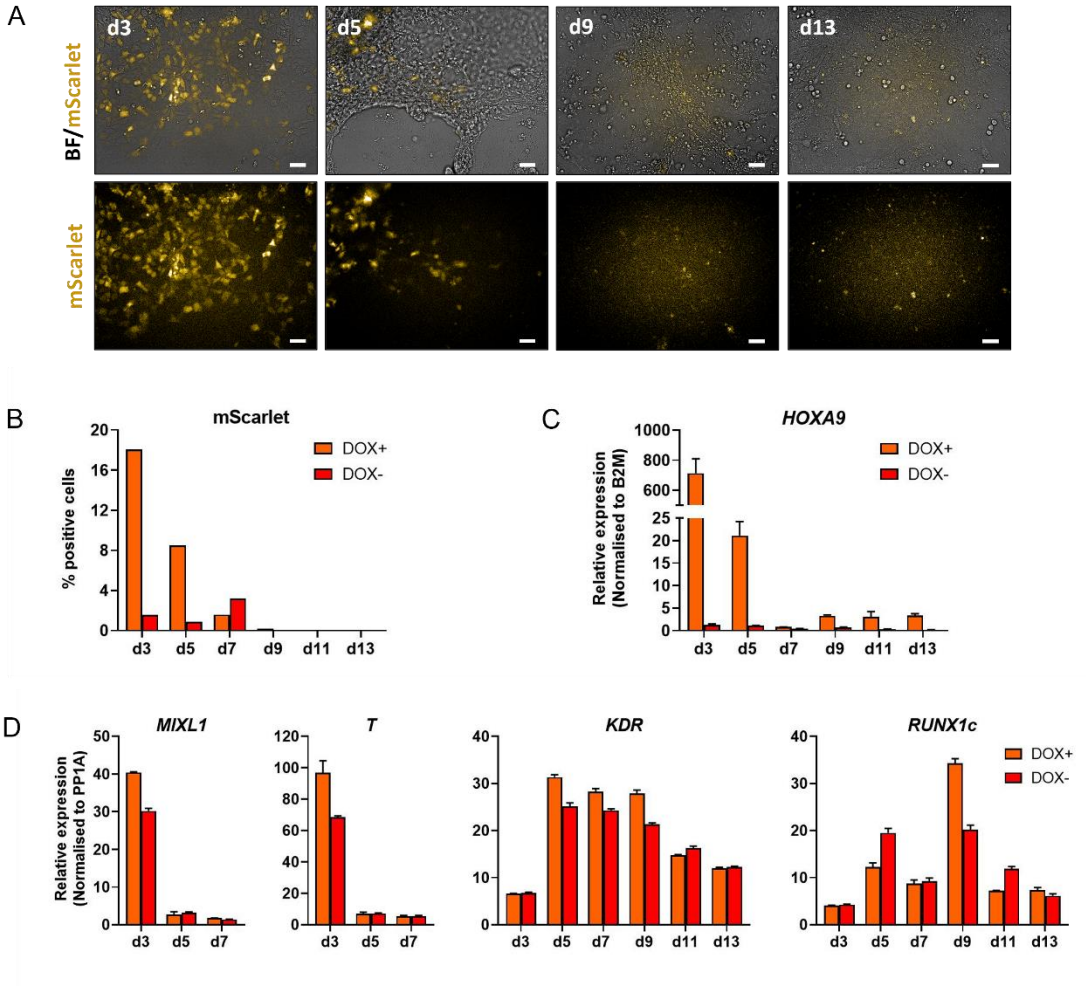


Figure 3.19: DOX mediated *HOXA9*/mScarlet induction is not maintained throughout haematopoietic differentiation

(A) Representative images of iA9 clone 54 over the course of haematopoietic differentiation using the STEMdiff™ kit. iA9 cells were treated for 24h with 1 µg/mL DOX and imaged at the timepoints indicated. Scale bar = 50 µm. (B) Quantification of mScarlet imaging using Columbus™ software. (C) QPCR for *HOXA9* 24h after DOX treatment. (D) QPCR for mesendoderm (*MIXL1* and *T*), endothelial (*KDR*) and hematopoietic (*RUNX1c*) markers 24h after DOX treatment. Results are mean ± SD, n = 3 technical replicates.

3.4 Discussion

In this chapter, an inducible HOXA9 expression system targeted to the AAVS1 locus was generated in hiPSCs. Following CRISPR/Cas9-mediated gene editing, 12 clones (out of an initial 54) were identified to be correctly targeted to one allele of the AAVS1 locus by PCR and Sanger sequencing. All iA9 clones produced *HOXA9* and mScarlet upon stimulation with 1 µg/mL DOX for 24h, however expression was heterogenous and highly variable between clones. A preliminary differentiation of one iA9 clone (54) to haematopoietic cells to assess maintenance of DOX mediated *HOXA9*/mScarlet induction throughout differentiation revealed a progressive decline in DOX responsiveness as the differentiation proceeded.

To generate our inducible HOXA9 model we used a donor template, pAAVS1-TRE-Shox2a-Nluc, readily available in our lab as a background vector upon which we could build our model. The pAAVS1-TRE-Shox2a-Nluc vector had several desirable features; it was targeted to the AAVS1 locus, facilitating stable targeted integration, and it contained the optimised rtTA-M2 variant and second generation tightTRE promoter (Bäckman et al., 2004). A similar Tet-On system (consisting of a tightTRE-driven transgene and rtTA-M2) was previously demonstrated to facilitate robust transgene expression in hESCs following DOX treatment (Xia et al., 2008). However, upon differentiation to neural, cardiac and endocrine cells, DOX induction did not elicit the same response as in hESCs, with DOX-induced transgene expression ranging from 20 – 55% (Xia et al., 2008). This was later attributed to the lentivirus-mediated random integration approach of the study (Qian et al., 2014). The AAVS1 locus contains an insulator element which is postulated to protect integrated transgenes from trans-activation or repression (Ogata et al., 2003). Therefore, although our system may be susceptible to transgene silencing during differentiation, we postulated targeting of our inducible system to the AAVS1 locus would mitigate against such possibilities.

The pAAVS1-TRE-Shox2a-Nluc vector contained all the necessary components to create a targeted, stable DOX-inducible expression system in hiPSCs. However, due to the lack of restriction enzyme sites flanking the Shox2a-T2A-Nluc element, it was not readily adaptable for use with other transgenes. Two SpeI sites up- and downstream of Shox2a-T2A-Nluc were instead used to excise the Shox2a-T2A-Nluc element. A Fusion PCR approach was initially planned to generate our HOXA9-T2A-mScarlet insert fragment from three different plasmids. However, addition of 36 bp of the T2A to mScarlet to create a complementary overhang for Fusion PCR was not possible, likely due to difficulties in PCR amplifying the T2A which possesses high GC content (72%). An

alternative approach was then taken, whereby a BsiWI site was incorporated at the 5' end of mScarlet in place of T2A, while the full 54 bp T2A and a BsiWI site were added to the 3' end of HOXA9. The HOXA9-T2A-mScarlet fragment with flanking SpeI sites was then generated by Fusion PCR and restriction enzyme digestion and ligation. Gel purification of the final TRE-HOXA9-mScarlet fragment yielded insufficient product for subsequent cloning into pAAVS1-TRE-Shox2a-Nluc background vector. Attempts to PCR amplify the product were unsuccessful, likely due to its ~2.5 kb length or structure. However, the product was successfully sub-cloned into pMiniT 2.0 vector, yielding sufficient high quantities of TRE-HOXA9-mScarlet following amplification for excision and subsequent integration into the backbone vector to produce pAAVS1-TRE-HOXA9-mScarlet.

The Cas9n and gRNA plasmids required to target the AAVS1 locus were kindly donated by Dr Jamie Bhagwan. The gRNAs were designed and optimised using the Guide Design Resources online tool (https://zlab.bio/guide-design-resources; Zhang lab) achieving a targeting efficiency of 19% with no observable off-target events (Bhagwan, 2018). The use of gRNA pairs rather than single gRNAs was found to greatly reduced the potential for off-targets, with the predicted occurrence reduced to zero using the two gRNAs selected (Bhagwan, 2018). To generate our inducible HOXA9 iPSC line, the gRNAs provided by Dr Jamie Bhagwan (Figure 3.1, gRNA #1 and #2) were used to target Cas9n to the AAVS1 locus, facilitating integration of the TRE-HOXA9-mScarlet cassette after induction of the DNA break. Given the thorough validation of the gRNAs previously performed (Bhagwan, 2018), off-target events were not further investigated.

As our inducible system was designed based on the pAAVS1-TRE-Shox2a-Nluc vector present in the lab, we were also able to access gDNA isolated from the previous inducible Shox2a reporter hESC line to facilitate prompt optimisation of 5' and 3' integration primers. This helped to streamline the genotyping process and led to quick identification of correctly targeted clones. Furthermore, the inducible Shox2a reporter is monoallelically targeted which we hoped would enable us to identify primers which amplify the edited and unedited alleles. However, despite numerous attempts, PCR of the edited allele was never achieved. The cassette incorporated into the AAVS1 locus in the inducible Shox2a reporter is 8865 bp and 8156 bp in the iA9 clones. This large size or associated secondary structure likely contributed to the difficulties in amplification. The 2036 bp untargeted allele was readily amplified. Using primers targeting the unedited allele, all clones were determined to be monoallelically targeted. Of note, gDNA from

RPAT hiPSCs produced more prominent bands than the targeted clones; further implying one allele of the AAVS1 locus is edited in the iA9 clones.

mScarlet was tagged to HOXA9 to facilitate assessment of HOXA9 expression in live cells by fluorescence microscopy. Each iA9 clone was treated with 1 µg/mL DOX for 24h to test its ability to induce HOXA9 and mScarlet. Amongst the twelve correctly targeted iA9 clones, no clone was able to completely induce mScarlet (i.e. 100% expression) following DOX treatment. When mScarlet expression in each iA9 clone was quantified, expression was found to range considerably amongst the different iA9 clones, from 0.9% to 44.7%. This range was mostly reflected in the corresponding *HOXA9* qPCR, with iA9 clones which induced high mScarlet expression tending to produce high *HOXA9* expression in qPCR assays and vice versa. The variation in *HOXA9*/mScarlet expression amongst the twelve iA9 clones has been previously observed in our lab in a similar system, whereby transgene expression in different subclones was found to range from 32 – 96%, despite correct targeting of the transgene to the AAVS1 locus in each subclone (Bhagwan et al., 2020b). As such, we hypothesised the variation in *HOXA9*/mScarlet expression observed amongst the iA9 clones may be clone specific.

mScarlet expression was quantified using Columbus™ software which relies on fluorescence intensity to distinguish between positive and negative cells. The background intensity of mScarlet (97.3±35.5 a.u.) was very similar to the intensity of mScarlet-expressing cells (109.3±33.1 a.u.). As a result, accurate determination of mScarlet expression was greatly hindered, with the possibility of inclusion of false positives or negatives during quantitative analysis greatly increased. Inaccurate quantification of mScarlet expression may be another explanation for the expression variability observed amongst iA9 clones, in particular for iA9 clones 19, 20 and 22 where mScarlet expression was not found to mirror that of *HOXA9* qPCR. Thus, further optimisation of DOX treatment conditions is required to minimise mScarlet background and increase the accuracy of mScarlet image quantification.

Inefficient clone selection is an alternative explanation for the expression variability amongst iA9 clones. Following CRISPR transfection, cells were seeded as single cells and allowed to expand in the presence of G418 antibiotic. The subsequent colonies should be derived from targeted cells but it is possible untargeted cells grew in unison, generating a mixed population or mosaic clone. However, indicators of possible mosaicism were not detected (Oceguera-Yanez et al., 2016); genotyping PCRs did not yield any secondary bands and chromatogram traces were clear. Additionally, issues

with protein translation may have impacted mScarlet production. To assess the suitability of mScarlet as a surrogate for HOXA9, iA9 clone 54 iPSCs were stained for the T2A tag and mScarlet after DOX treatment. Quantification of staining demonstrated both proteins were equally expressed, indicating the T2A peptide was efficiently cleaved during translation. As such, poor translation or inefficient T2A cleavage are unlikely to be responsible for the varied transgene expression observed following DOX induction in iA9 clones.

In a preliminary experiment to assess if DOX activity was maintained during differentiation, iA9 clone 54 was differentiated to haematopoietic cells using the STEMdiff™ Hematopoietic kit. *HOXA9*/mScarlet expression was stimulated by treatment with 1 µg/mL DOX. Induction was performed at different timepoints throughout differentiation, with each treatment only lasting 24h. In differentiating cells, maximum expression of both *HOXA9* and mScarlet was observed on d3, with progressive decline in inducibility thereafter. Expression of key mesendoderm markers *MIXL1* and *T* on d3 suggests DOX induction ability is hampered following mesoderm patterning. A hallmark of *in vitro* haematopoietic differentiation is the detachment of haematopoietic progenitors from the underlying endothelial layer into suspension. Differentiation cultures treated with or without DOX, failed to produce suspension cells and so haematopoietic potential could not be assessed. Expression of *KDR* and *RUNX1c* suggest some patterning towards the haematopoietic lineage occurred, however ultimately the differentiation was unsuccessful.

Although transgene silencing during differentiation has previously been shown for the tightTRE/rtTA-M2 system, the authors demonstrated measurable retention of inducible transgene expression in some differentiated cell types (Xia et al., 2008). The retention observed in this study was likely as a result of lentivirus-mediated random integration of the Tet-On system, with some integration sites being silenced upon differentiation and others remaining accessible. Stable integration of our inducible HOXA9 system into a single site in the genome may, unexpectedly, be disadvantageous to our model as complete loss of transgene induction was observed from d7 onwards. As such, possible silencing within our inducible system warrants further investigation.

The apparent silencing of the inducible HOXA9 system observed may be specific to iA9 clone 54. Variation in *HOXA9* and mScarlet expression amongst the twelve iA9 clones is evident in hiPSCs; therefore, it is possible this variation is maintained during differentiation and certain clones may be more efficient at retaining DOX responsiveness.

Our lab has previously demonstrated hPSC clones which possessed high transgene expression in pluripotent cells were more likely to maintain transgene expression in differentiated cardiomyocytes. However, although some clones were more likely to retain expression, reduced or complete loss of transgene expression was also evident in differentiated cells for several clones (Bhagwan et al., 2020b). This indicates transgene expression in hiPSCs is not a predictor of continued transgene expression in differentiated cell types.

CRISPR/Cas9 editing resulted in twelve correctly targeted iA9 clones which each varied in their response to DOX. The inducible HOXA9 cassette was only targeted to one allele in all clones. However, for the purposes of transgene overexpression, biallelic targeting is not necessary. Higher expression of HOXA9 and mScarlet can likely be achieved by altering the concentration of DOX, as has been previously reported (Das et al., 2016). The preliminary haematopoietic differentiation indicated DOX inducibility may not be retained in differentiated cell types. Therefore, to identify an iA9 clone which could retain DOX-inducibility we aimed to screen iA9 clones for response to DOX throughout haematopoietic differentiation. However, a major caveat of this approach was the STEMdiff™ kit used to model haematopoiesis was inefficient and not amenable to optimisation (discussed in detail in Chapter 5). Screening iA9 clones with an unreliable protocol could result in falsely identifying a clone as uninducible. As an alternative, we opted to screen the iA9 clones via cardiac differentiation; for which differentiation protocols are well-established.

Chapter 4: Cardiac differentiation proof of concept study

4.2 Introduction

Doxycycline-inducible HOXA9-mScarlet (iA9) reporter iPSC lines were generated in the previous chapter. While all iA9 iPSC clones produce HOXA9 and mScarlet expression following DOX treatment, this response was not maintained during preliminary differentiation to haematopoietic cells. Further verification of additional iA9 clones was hindered by the failure of the haematopoietic differentiation procedure. To overcome this, generation of iPSC-derived cardiomyocytes was explored as a surrogate procedure to screen for iA9 clones which retain DOX responsiveness during differentiation. This chapter also investigates the role DNA methylation may play in reducing DOX responsiveness.

4.2.1 Cardiomyocyte generation *in vivo* and *in vitro*

Cardiomyogenesis begins with the formation of mesoderm during gastrulation. Nodal signalling in the proximal epiblast co-ordinates with BMP4, Wnt3A, and inhibitory Nodal and Wnt signalling from adjacent tissues to direct posterior epiblast cells to form the primitive streak and undergo an epithelial-to-mesenchymal transition (EMT). The resultant mesoendoerm is marked by upregulation of *T* and *EOMES*, which drive formation of lateral plate mesoderm (Burridge et al., 2011). Subsequent upregulation of cardiac master regulator MESP1 in mesoderm and inhibitory Wnt signalling promote the formation of the first and second heart fields. The heart fields represent cardiac progenitor populations; with further development and differentiation into cardiomyocytes governed by inhibitory Wnt signalling (Spater et al., 2014) and transcription factor families *NKX*, *MEF2*, *GATA, TBX* and *HAND* (Olson et al., 2006).

Generation of hPSC-derived cardiomyocytes relies on recapitulating *in vivo* embryonic cardiac development. Early successful differentiation procedures employed sequential treatment of Activin A and BMP4 at the start of differentiation to mimic Nodal/BMP4 signalling gradients (Kattman et al., 2010; Laflamme et al., 2007). Other differentiation procedures utilise the bi-phasic role the Wnt/β-catenin pathway plays during cardiomyogenesis. Wnt3A or small molecule Wnt agonists are used to direct differentiating hPSCs to mesoderm, while cardiac progenitor formation is induced by inhibiting Wnt with small molecules (Burridge et al., 2014; Paige et al., 2010). The role the extracellular matrix plays during primitive streak patterning has also been harnessed to improve cardiac differentiation protocols. Addition of a Matrigel-overlay at the start of

hPSC cardiac differentiation promotes EMT and subsequent cardiomyocyte derivation (Celiz et al., 2015; Zhang et al., 2013).

4.2.1.1 HOX genes in cardiac development

The role of anterior HOX genes in cardiac development is well described. Hoxa1, Hoxa3 and Hoxb1 are expressed in cardiac progenitors of the second heart field and downregulated upon further differentiation (Bertrand et al., 2011). In particular, Hoxb1 is an essential regulator of early cardiac development. Loss of Hoxb1 leads to abnormal development of the second heart field in mouse embryos and a differentiation block of cardiac progenitors in a mouse ESC model of myogenesis (Stefanovic et al., 2020). Roles for posterior HOX genes in early cardiac development have also been postulated. Hoxa9 and Hoxa10 are upregulated in a mesoderm sub-population of the $Nkx2.5^{+}$ cardiac crescent in mouse embryos (Behrens et al., 2013). Overexpression of Hoxa10 impairs differentiation of Nkx2.5 progenitors into $Flk1^{+}Pdgfr\alpha^{+}$ cardiac mesoderm and subsequent differentiation into cardiomyocytes. In contrast, Hoxa13 overexpression did not affect cardiomyocyte formation, suggesting the defect is attributed to Hoxa10 and not all posterior Hox genes (Behrens et al., 2013). In adult cardiomyocytes, upregulation of Hoxa9 has been observed in murine *in vitro* and *in vivo* models of cardiac hypertrophy. In addition, treatment of mouse primary cardiomyocytes with isoprenaline results in cardiac hypertrophy, with upregulation of Hoxa9 and downregulation of cardio-protective protein PRMT5 observed. Overexpression of PRMT5 represses Hoxa9, indicating a direct interaction between the two proteins (Cai et al., 2020). Interaction between Hoxa9 and lncRNAs associated with cardiac hypertrophy, including PEG10 (Wen et al., 2019) and UCA1 (Zhou et al., 2018) have also been established.

4.2.2 DNA methylation

DNA methylation is involved in repression of gene expression. It occurs when a methyl group from S-adenyl methionine is transferred to the fifth carbon of a cytosine residue, forming 5-methyl cytosine (5-mC). The reaction is catalysed by a family of DNA methyltransferases (DNMTs), namely DNMT3a and DNMT3b, which are involved in *de novo* methylation, and DNMT1 which functions during replication to copy methylation patterns to newly synthesised DNA, thus maintaining methylation (Bird, 2002; Li and Zhang, 2014; Moore et al., 2013). DNA methylation predominately occurs on cytosine residues which are proceeded by a guanine residue (known as CpG dinucleotide) (Laurent et al., 2010). In somatic cells, 70 – 80% of all CpG sites are methylated, with exceptions including CpG islands; regions of high CpG content approx. 1 kb in length (Li and Zhang, 2014). Approximately 70% of promoters in the human genome contain CpG

islands which, depending on the cell context, are largely unmethylated (Saxonov et al., 2006).

During mammalian embryogenesis, cells undergo two waves of global demethylation and methylation. Prior to implantation of the zygote, a wave of demethylation leads to removal of methyl groups from the parental genome. This is followed by extensive *de novo* methylation mediated by DNMT3a and DNMT3b, whereby almost all CpG sites are methylated, with the exception of CpG islands which are protected by *cis*-acting regulatory elements (Cedar and Bergman, 2012; Reizel et al., 2018). Further methylation or demethylation events during embryogenesis and in adult tissues are either gene- or tissue-specific. Several key genes involved in maintaining pluripotency initially possess an unmethylated profile and become methylated upon differentiation resulting in gene repression (Straussman et al., 2009). OCT3/4 is an essential protein for pluripotency, such that self-renewal of mESCs is abrogated by the loss of Oct3/4, while overexpression stimulates mESC differentiation (Niwa et al., 2000). Bisulphite sequencing of mESCs demonstrated that Oct3/4 is unmethylated in undifferentiated cells but undergoes *de novo* methylation upon differentiation. Although Oct3/4 is first downregulated by transcriptional repressors during differentiation, it is *de novo* methylation mediated by Dnmt3a and Dnmt3b and driven by methyltransferase G9a which induce irreversible silencing of Oct3/4 in differentiating cells and somatic tissues (Epsztejn-Litman et al., 2008). Furthermore, methylation of Oct3/4 following differentiation represents a tissue-specific event. In an analysis of differentially methylated regions in hESCs and somatic tissues, hypomethylated regions in hESCs were most highly enriched in pluripotency transcription factors OCT4, SOX2 and NANOG. Similarly, haematopoietic transcription factors SPI and TAL1 were enriched in hypomethylated regions in CD34 cells isolated from human blood (Ziller et al., 2013). Thus, DNA methylation is a dynamic process with roles in regulating gene expression in developmental- and cell-specific contexts.

4.2.2.1 Mapping DNA methylation

Studies of DNA methylation status utilises techniques which enable identification of 5-mC patterns at either specific sites or across the whole genome. Bisulphite sequencing and methylated DNA immunoprecipitation (MeDIP) are unbiased, reliable quantitative methods to detect 5-mC (Li and Zhang, 2014). In bisulphite sequencing, sodium bisulphite is used to convert unmethylated cytosines to uracil, while methylated cytosines remain protected. As such, methylated cytosines can be distinguished from unmethylated. Subsequent PCR and sequencing enables direct determination of all methylated cytosines within the target region (Li and Tollefsbol, 2011). Bisulphite

sequencing can also be combined with whole genome sequencing to facilitate genome wide mapping of methylated CpG sites (Hansen et al., 2012).

The methylation status of the whole genome or specific regions can also be assessed using MeDIP. In MeDIP, the starting material (termed Input), consisting of both methylated and unmethylated DNA, is prepared by randomly shearing total genomic DNA into short fragments approx. 400 bp in length. The MeDIP fraction is then isolated by immunoprecipitation with antibodies specific for 5-mC (Weber et al., 2005). Genome-wide methylation can subsequently be characterised with microarrays; whereby the Input and MeDIP fraction are labelled with dyes and co-hybridised with microarrays of CpG rich regions. Alternatively, methylation within specific regions of the genome can be estimated by qPCR (Lisanti et al., 2012). Short sequences within the region of interest (typically 150 – 200 bp in length) are amplified in both the Input and MeDIP fraction and the relative enrichment of the target sequence between the MeDIP fraction and Input is determined. Reactions for genes of known methylation status can be run as controls (Mohn et al., 2009).

4.3 Aims

The primary aim of this chapter was to identify an iA9 clone which maintains sensitivity to DOX throughout differentiation. To this end, iA9 clones were screened for DOX sensitivity by differentiation to cardiomyocytes with HOXA9 induced every 2 days of differentiation. HOXA9 overexpression in cardiomyocytes is implicated in cardiac hypertrophy; however, no adverse effect has been shown during cardiac differentiation (see Section 4.2.1.1). Therefore, the screening approach should not impede cardiac differentiation. As discussed in Chapter 3, there is minimal difference in background mScarlet fluorescence intensity and the intensity of mScarlet-expressing cells, making identification of true mScarlet-expressing cells challenging. Therefore, an additional aim of this chapter was to investigate if amplification of the mScarlet signal would enhance mScarlet intensity and increase the percentage of positively expressing cells. It has been postulated that silencing of the Tet-On system during differentiation may be responsible for the loss of DOX induction observed. To this end, we aimed to investigate such a possibility by assessing the DNA methylation status of specific sites within the targeted AAVS1 locus using MeDIP.

4.4 Results

4.4.1 Optimisation of DOX treatment conditions

The high background intensity of mScarlet hindered accurate quantification of mScarlet expression in iA9 iPSCs. Additionally, all 12 iA9 clones expressed mScarlet at low to moderate levels (<50% cells mScarlet+) following 24h treatment with 1 µg/mL DOX (Figure 3.12). To facilitate more efficient assessment of mScarlet expression, DOX treatment conditions were optimised, with a focus on identifying conditions which increased the number of mScarlet expressing cells and mScarlet intensity. To verify mScarlet expression, HOXA9/2A was also stained. The impact of a single DOX treatment at three different concentrations (0.5, 1 or 2 µg/mL DOX) on mScarlet and 2A expression over a 72h period was assessed. iA9 clone 54 iPSCs were treated with DOX and stained for 2A and DAPI 24h, 48h and 72h post-treatment, followed by imaging using the Operetta® imaging system (Figure 4.1). The percentage of mScarlet- and 2A-positive cells was calculated relative to DAPI nuclear staining using Columbus™ analysis system (Figure 4.2). After 24h, all three DOX concentrations yielded similar levels of mScarlet (22.6 – 30.5%, Figure 4.3A) and 2A expression (20.6 – 23.3%, Figure 4.3B). After 48h, both mScarlet and 2A expression significantly decreased in cells treated with 0.5 µg/mL (0.9% mScarlet+, $p < 0.0001$; 0.3% 2A+, $p < 0.001$) and 1 µg/mL DOX (7.5% mScarlet+, $p < 0.0001$; 2% 2A+, $p < 0.001$). Cells treated with 2 µg/mL DOX maintained mScarlet and 2A expression over 48h, with a reduction in expression not observed until 72h after treatment (Figure 4.3A, B). No changes in mScarlet intensity were observed amongst the three DOX concentrations tested, ranging from 57.3 a.u. to 58.8 a.u. and minimal difference was observed over background intensity, 51 a.u. (Figure 4.4B).

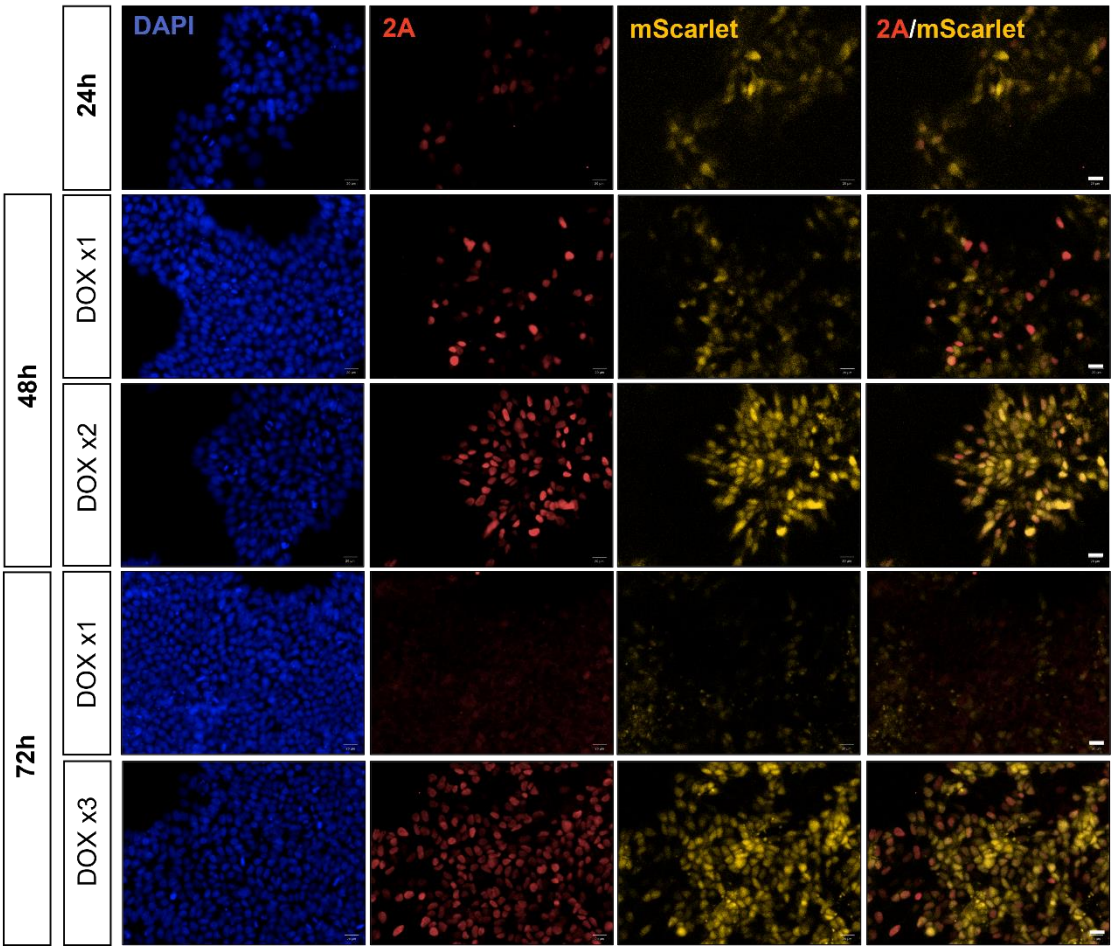


Figure 4.1: mScarlet and 2A staining in DOX-treated iA9 clone 54 iPSCs

Representative images taken using the Operetta® imaging system, of clone 54 iPSCs treated with 2 µg/mL DOX and stained for DAPI, 2A (surrogate for HOXA9) and mScarlet. IPSCs were treated once (DOX x1) with 2 µg/mL DOX and imaged 24h, 48h and 72h after treatment. IPSCs were also treated with DOX multiple times; either two (48h DOX x2) or three times (72h DOX x3) with 2 µg/mL DOX at 24h intervals. Scale bar = 20 µm.

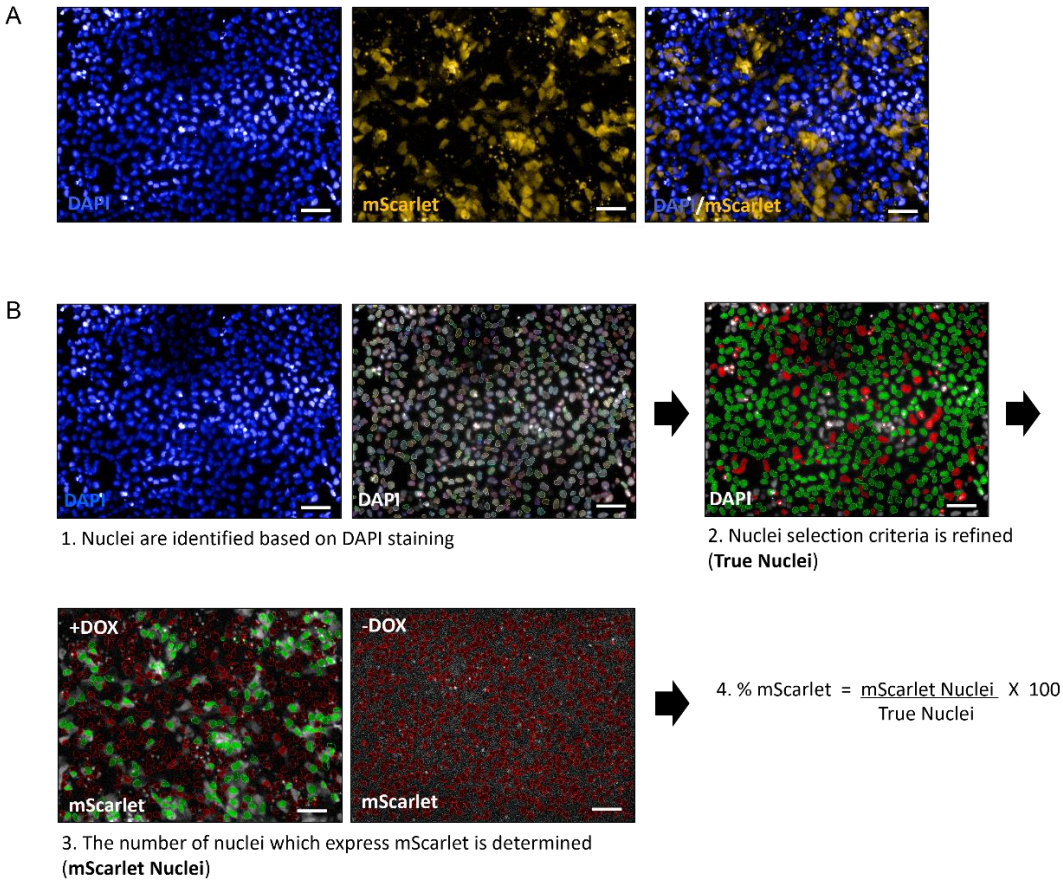


Figure 4.2: Quantification of fluorescence staining relative to nuclear staining using Columbus™

Example of mScarlet quantification in cells co-stained with DAPI. (A) Representative image of iA9 clone 18 on d2 of cardiac differentiation following treatment with 1 µg/mL DOX for 48h. The image represents a single field of view. (B) Overview of Columbus™ quantification steps. First, nuclei are identified based on DAPI staining (step 1). DAPI area and intensity are calculated from which nuclei can be further refined (step 2). Two nuclei close together may be counted as one or artefacts may be mistakenly identified as nuclei by the Columbus™ software. These are eliminated from nuclei selection criteria (red nuclei), while the number of nuclei which meet the criteria are calculated (True Nuclei shown in green). Next, mScarlet intensity is determined (step 3) and background intensity is calculated using a negative sample (-DOX). The number of nuclei which express mScarlet at an intensity greater than background is calculated (mScarlet Nuclei shown in green). Finally, the percentage of mScarlet expression is calculated (step 4). Scale bar = 50 µm.

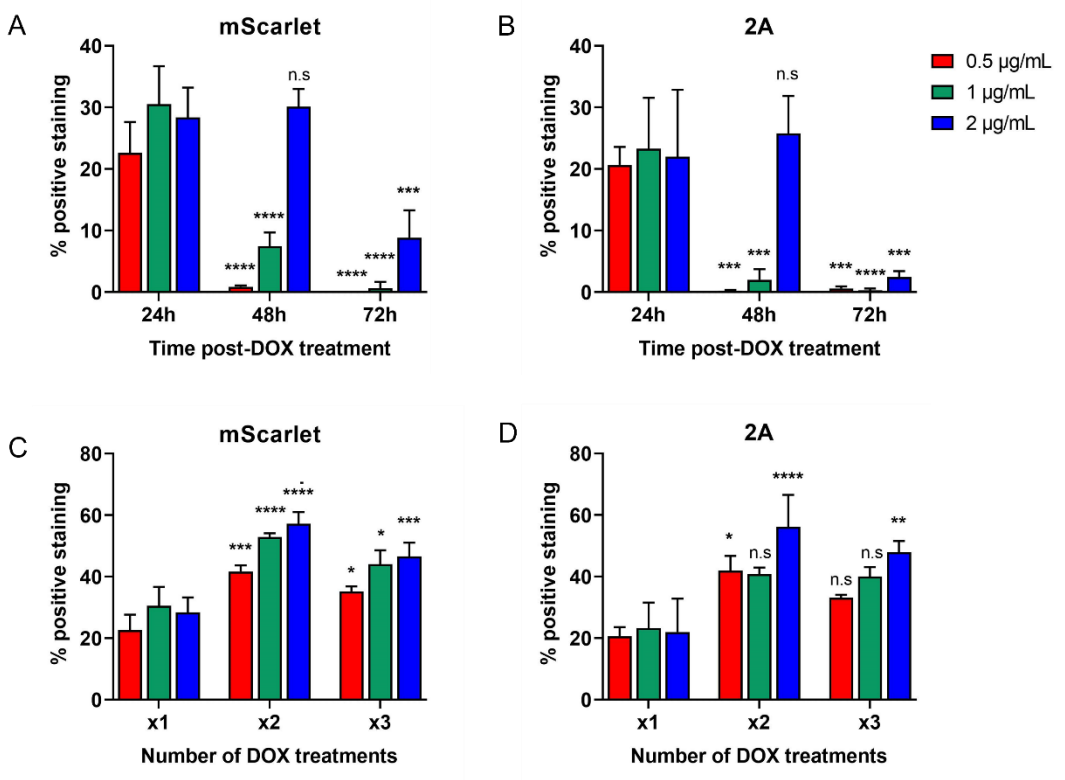


Figure 4.3: DOX concentration and number of treatments impact the level of mScarlet and 2A expression

Quantification of mScarlet and 2A staining using Columbus™ software. (A - B) iA9 clone 54 iPSCs treated once with 0.5, 1 or 2 µg/mL DOX followed by imaging, staining and quantification of mScarlet (A) and 2A (B) 24h, 48h and 72h post-treatment. (C – D) iA9 iPSCs treated one (x1), two (x2) or three (x3) times with 0.5, 1 or 2 µg/mL DOX at 24h intervals followed by imaging, staining and quantification of mScarlet (C) and 2A (D) 24h after last treatment. Results are mean ± SD, n = 3 technical replicates, each technical replicate represents a sum of 20 fields of view; *p < 0.05, **p < 0.01, ***p < 0.001, ****p < 0.0001 and n.s non-significant compared to 24h DOX treatment, 2-way ANOVA, Tukey's multiple comparisons test.

As a single DOX treatment at the concentrations tested did not impact mScarlet/2A expression level or mScarlet intensity, we next investigated the impact of successive treatments on these parameters. iA9 clone 54 iPSCs were treated with one, two or three pulses of DOX at 24h intervals for up to 72h and stained for 2A and DAPI 24h after the last pulse (Figure 4.1). All three DOX concentrations were tested. Treated cells were imaged using the Operetta® and the number of mScarlet- and 2A-expressing cells was calculated as before (Figure 4.2). Multiple DOX treatments induced significantly higher mScarlet expression than single treatment at all concentrations (0.5 µg/mL: 22.6% x1, 41.7% x2 $p < 0.001$, 35.2% x3 $p < 0.05$; 1 µg/mL: 30.5% x1, 52.9% x2 $p < 0.0001$, 44.1% x3 $p < 0.05$; 2 µg/mL: 28.4% x1, 57.2% x2 $p < 0.0001$, 46.6% x3 $p < 0.001$). There was

a trend for the highest mScarlet induction in cells treated twice with DOX, with 2 µg/mL DOX eliciting the highest response at 57.2% mScarlet-positive (Figure 4.3C). However, this trend was not significant. Similar to mScarlet, quantification of 2A-stained revealed two treatments with 2 µg/mL DOX achieved the highest number of 2A-expressing cells (22% x1 vs 56.1% x2 $p < 0.0001$). Two treatments with 0.5 µg/mL DOX also produced significantly more 2A-positive cells than a single treatment (20.6% x1 vs 41.9% x2 $p < 0.05$). The same was observed for 1 µg/mL DOX (23.2% x1 vs 40.9% x2 n.s.), however the difference was not found to be statistically significant (Figure 4.3D).

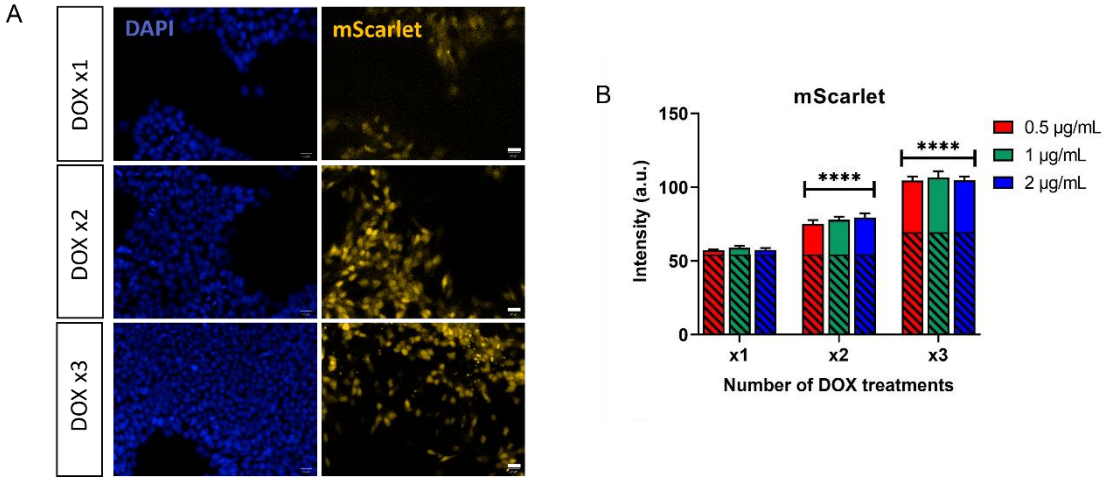


Figure 4.4: mScarlet intensity increases with successive DOX treatments

(A) Representative images of iA9 clone 54 iPSCs treated one, two or three times with 1 µg/mL DOX at 24h intervals. Scale bar = 20 µm. (B) Intensity of mScarlet staining following single (x1) or multiple treatments (x2 or x3) with 0.5, 1 or 2 µg/mL DOX. Bars with black diagonal lines demonstrate the intensity of mScarlet background fluorescence. Results are mean ± SD, n = 3 technical replicates, each technical replicate represents a sum of 20 fields of view; ****$p < 0.0001$ compared to x1 DOX treatment, 2-way ANOVA, Tukey's multiple comparisons test.

The intensity of mScarlet expression significantly increased with successive DOX treatments and was not found to be concentration dependent. Little variation in mScarlet fluorescence intensity was observed between the three DOX concentrations tested, regardless of the number of times cells were treated. Compared with single treatment (ranging from 57.2 – 58.8 a.u.), mScarlet intensity was significantly greater following multiple DOX treatments (ranging from 74.9 – 79.4 a.u. for two and 104.7 – 106.5 a.u. for three treatments) at all concentrations (Figure 4.4B). In comparison with baseline intensity (60 ± 8.7 a.u.), multiple DOX treatments yielded substantially higher intensity values than single treatment. The influence of multiple DOX treatments on mScarlet

expression could also be detected qualitatively, with the presence of mScarlet-expressing cells more apparent following two or three treatments (Figure 4.4A). Thus, in comparison with single DOX treatment, multiple DOX treatments results in higher induction of mScarlet expression, with both a greater number of mScarlet-expressing cells and mScarlet intensity observed. No statistically significant difference in the level of mScarlet or 2A induction was identified when cells were treated twice or three times with DOX; however, there was a trend for higher mScarlet expression following two DOX treatments. As such, two successive DOX treatments were incorporated into future experiments.

4.4.2 Loss of inducible HOXA9-mScarlet expression during cardiac differentiation of iA9 clones

iA9 clones were differentiated to cardiomyocytes to identify clones which have the capacity to remain responsive to DOX throughout differentiation. iA9 clones were treated twice with DOX over a 48h period, followed by nuclei staining with Hoechst and live imaging to detect mScarlet with the Operetta®. Samples were subsequently collected for *HOXA9* qPCR. Treatment was initiated on d0, d2, d4, d6 or d8 of differentiation. Cardiomyocytes were characterised by spontaneous contraction and/or staining for α-actinin on d15 (Figure 4.5).

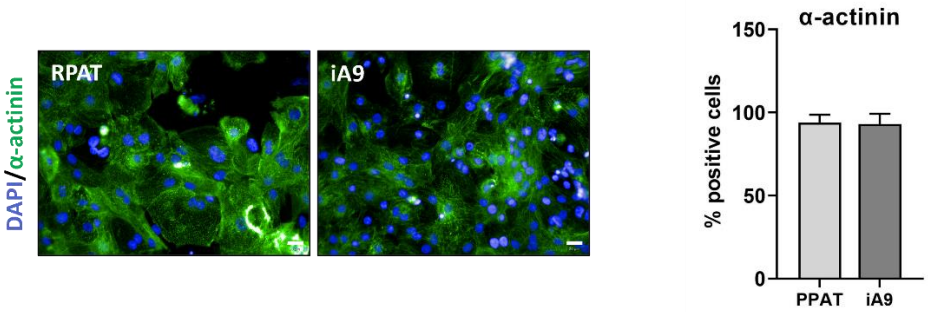


Figure 4.5: iA9 clones can be differentiated to cardiomyocytes

Representative images of RPAT- and iA9 clone 54-derived cardiomyocytes stained for sarcomertic protein α-actinin (left panel) and quantification of staining using Columbus™ software (right panel). RPAT-derived cardiomyocytes are shown for reference. RPAT and iA9 clone iPSCs can be efficiently differentiated to cardiomyocytes. Scale bar = 20 µm. Results are mean ± SD, n = 2 technical replicates, each technical replicate represents a sum of 20 fields of view.

To increase the possibility of identifying an inducible clone during differentiation, we aimed to stimulate maximal mScarlet expression. As two DOX treatments at 2 µg/mL resulted in the highest number of mScarlet- and 2A-expressing cells in iPSCs, we initially differentiated the iA9 clones to cardiomyocytes as described above using this concentration. However, after the first 48h induction on d2, substantially lower cell density was evident in cultures treated with DOX, but not in uninduced cultures (Figure 4.6A). Reduced cell number was assessed quantitatively by Hoechst nuclear staining (Figure 4.6B) and was not observed at any other time-point following treatment with 2 µg/mL DOX. Three clones (iA9 clone 20, 22 and 23) were differentiated using 2 µg/mL DOX to induce gene expression, with the same effect observed on d2 in the three clones. DOX concentration was lowered to 1 µg/mL, leading to a slight reduction in cell number in comparison with uninduced cells on d2, although an improved outcome compared with the higher DOX concentration (Figure 4.6B). The findings suggest possible toxicity of DOX or HOXA9 at the higher DOX concentration. Subsequent experiments used DOX at 1 µg/mL.

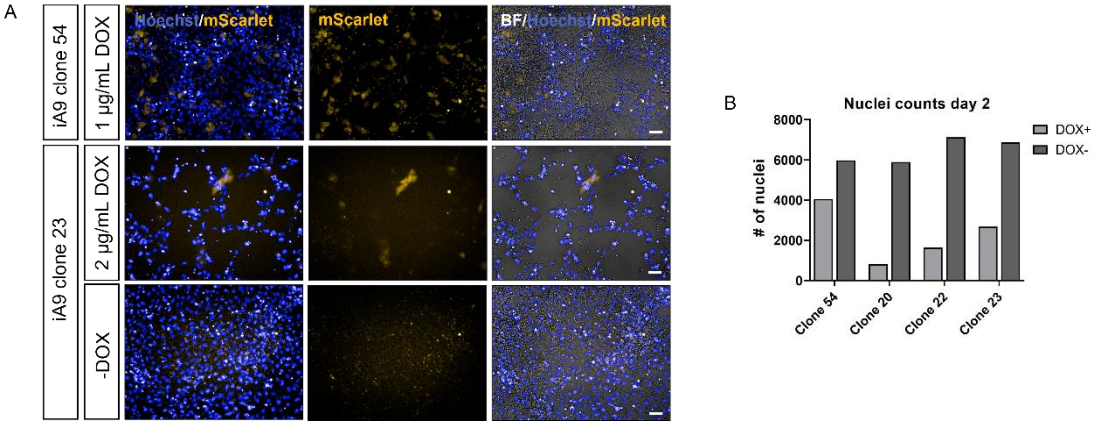


Figure 4.6: DOX at 2 µg/mL DOX reduces total cell number on day 2 of cardiac differentiation

(A) Representative images of iA9 clone 54 and clone 23 on d2 of cardiac differentiation. iA9 clone 54 was treated twice with 1 µg/mL DOX over 24h intervals. iA9 clone 23 was treated under the same conditions as clone 54 with 2 µg/mL DOX or received no DOX (-DOX). Nuclei are stained with Hoechst. Scale bar = 50 µm. (B) Total number of nuclei/well on d2 of cardiac differentiation following treatment +/- 1 µg/mL DOX (clone 54) or +/- 2 µg/mL DOX (clone 20, 22 and 23). Nuclei were counted using Columbus™ software. Quantification performed on 20 fields of view.

In total, seven of the twelve iA9 clones identified as being correctly targeted (Figure 3.12) were differentiated. All differentiated clones revealed the same pattern; gradual loss of inducible *HOXA9* (Figure 4.7) and mScarlet expression (Figure 4.8) as cardiac differentiation progressed. The highest induction of both *HOXA9* and mScarlet during differentiation was detected on d2, with exception of clone 20 and 22 where d4 showed the highest mScarlet and *HOXA9* induction, respectively. Subsequent qPCR showed mesendoderm marker *MIXL1* was upregulated on d2 of differentiation (Figure 4.9). For most of the clones, upregulation of HOXA9 did not influence *MIXL1* expression; however, for two iA9 clones (clone 5 and 54), *MIXL1* was more highly expressed in uninduced cells, while for one clone (clone 22), higher expression was observed in DOX treated cells. The remaining five iA9 clones of the twelve correctly targeted clones were not differentiated as the data suggested the loss of DOX responsiveness is not clone-specific as originally hypothesised but rather an issue with the gene editing strategy.

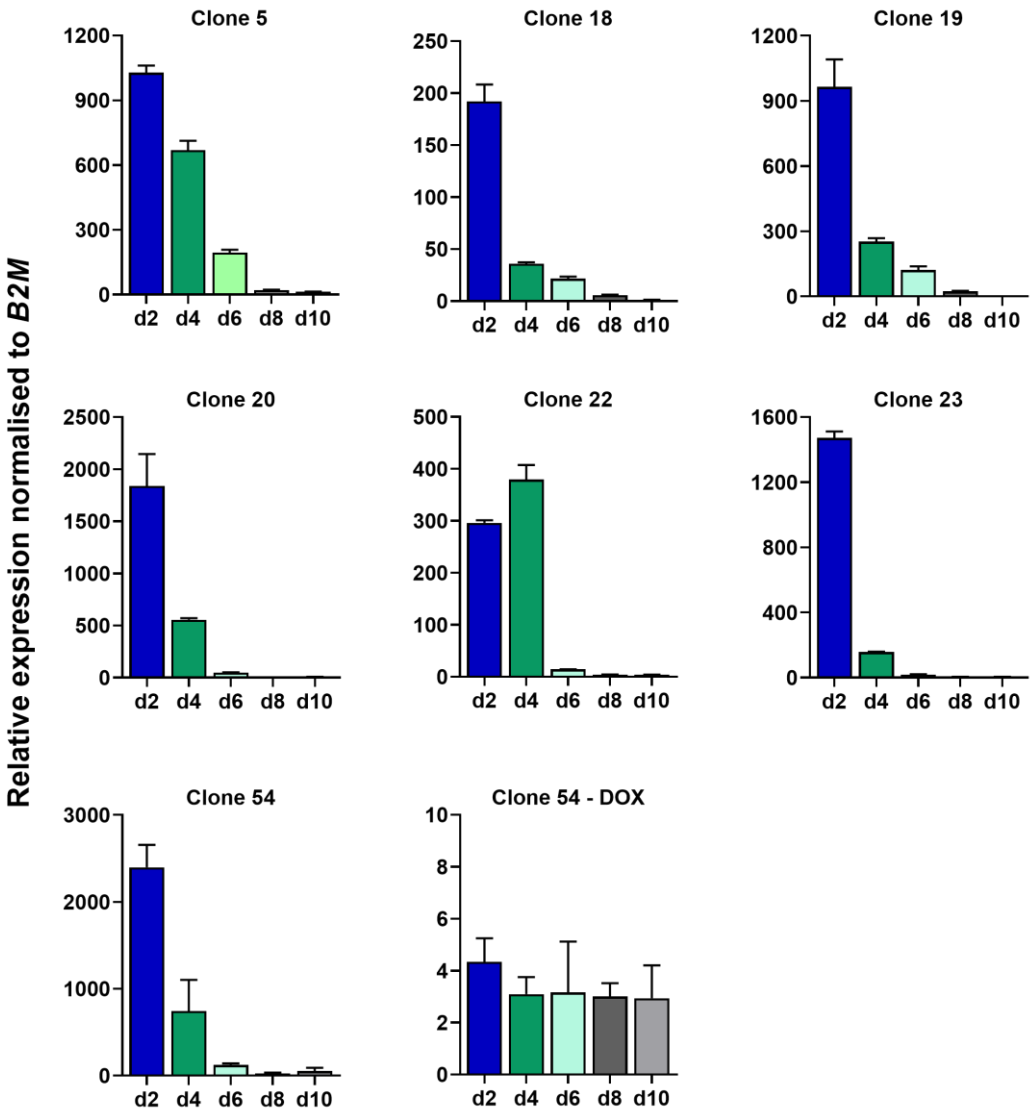


Figure 4.7: The ability of DOX to induce HOXA9 decreases as cardiac differentiation proceeds

QPCR for *HOXA9* in seven different iA9 clones during cardiac differentiation. Clones were treated twice with 1 or 2 µg/mL DOX at 24h intervals on d0, d2, d4, d6 or d8, with samples taken for qPCR after 48h of treatment. *HOXA9* expression is normalised to housekeeping gene *PP1A* and shown relative to uninduced iPSCs. A representative graph of endogenous *HOXA9* expression throughout cardiac differentiation in uninduced cells (clone 54 –DOX) is shown. Results are mean ± SD, n = 3 technical replicates.

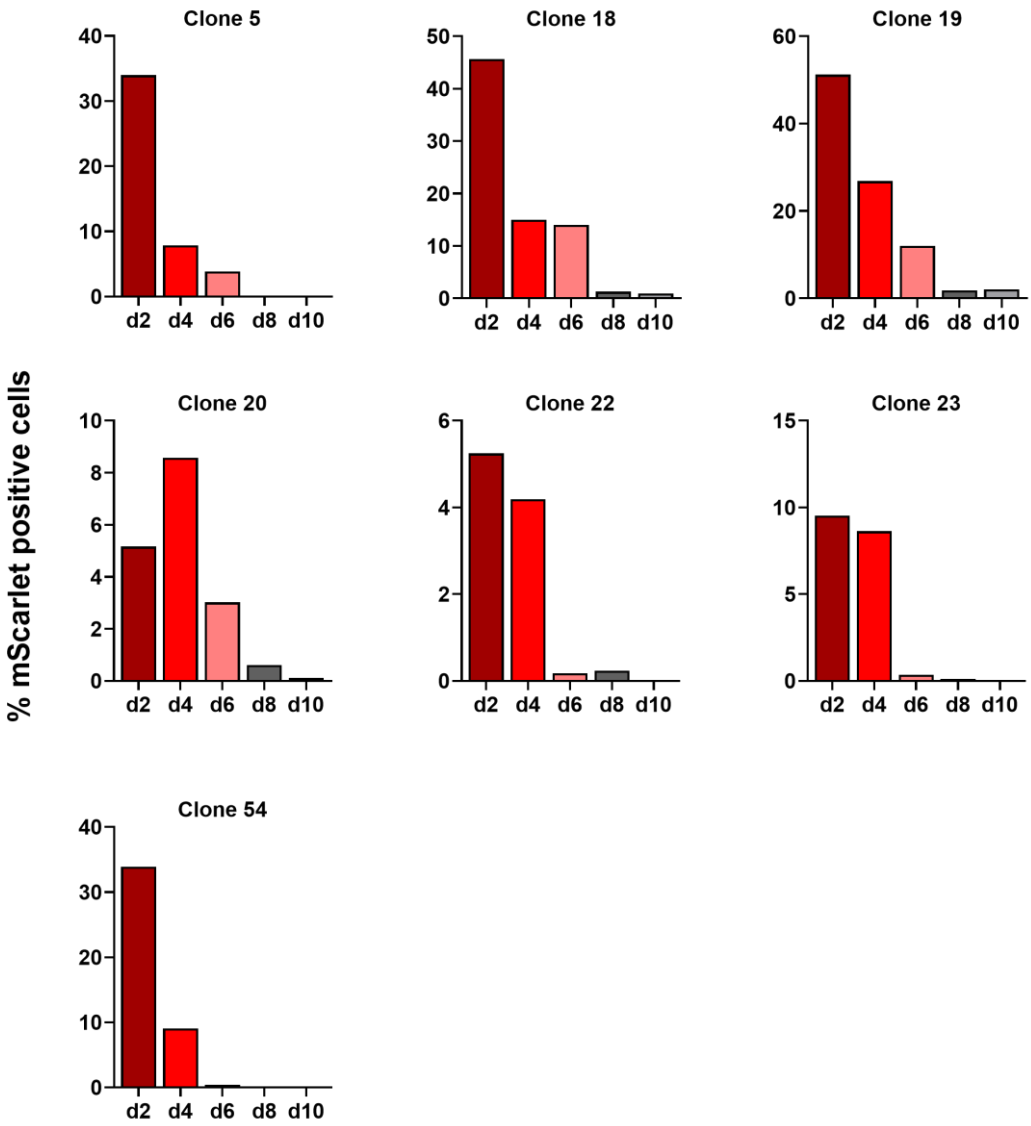


Figure 4.8: mScarlet expression reduces as cardiac differentiation progresses

Quantification of mScarlet induction during cardiac differentiation using Columbus™ software. iA9 clones were treated twice with 1 or 2 µg/mL DOX at 24h intervals on d0, d2, d4, d6 or d8 of cardiac differentiation, and imaged after 48h of treatment. Results represent quantification of 20 fields of view.

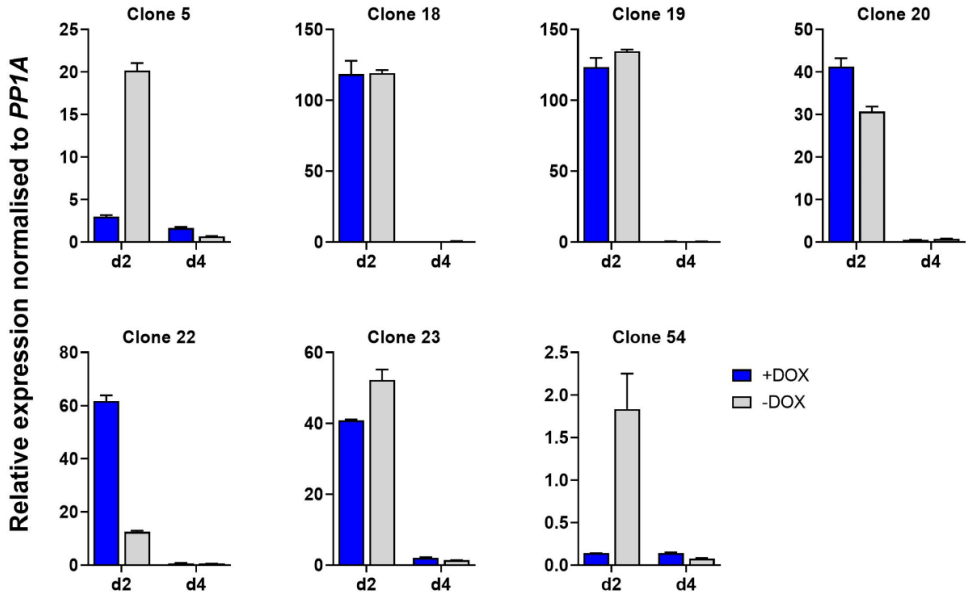


Figure 4.9: Day 2 of cardiac differentiation corresponds to mesendoderm formation

QPCR for mesoderm marker *MIXL1* on d2 and d4 of cardiac differentiation in cells treated +/- 1 or 2 µg/mL DOX twice at 24h intervals starting on d0 or d2 of differentiation. *MIXL1* expression is normalised to housekeeping gene *PP1A* and shown relative to uninduced iPSCs. Results are mean ± SD, n = 3 technical replicates.

4.4.3 DNA methylation identified within the AAVS1 locus and TRE promoter during cardiac differentiation

Loss of transgene expression during differentiation of Tet-On systems in hPSCs has been attributed to the integration approach used (Qian et al., 2014) and specific silencing of the tet responsive element (Mandegar et al., 2016; Ordovas et al., 2015). In addition, despite its reported protective properties, methylation of promoters integrated into the AAVS1 locus has been directly linked to loss of transgene expression (Klatt et al., 2020; Ordovas et al., 2015). Therefore, we decided to assess the methylation status of the TRE promoter and specific sites within the AAVS1 locus. Methylation status was determined using MeDIP, which uses an immunoprecipitation approach to isolate methylated DNA and qPCR as a means of identifying sites which exhibit methylation enrichment. iA9 clone 5 iPSCs were differentiated to cardiomyocytes and MeDIP was performed on samples taken at three different time-points. Samples were taken on d0, d2 and d6; corresponding to states proposed to possess no methylation (d0), partial methylation (d2) and hypermethylation (d6). In preparation of gDNA for immunoprecipitation, gDNA was randomly sheared to fragments of approx. 400 bp in length. Since the qPCR

template sequence was unknown, several different primer pairs were designed to amplify the TRE promoter and sites of interest in the AAVS1 locus (primer pairs which successfully amplified target DNA are shown in Figure 4.10). TRE primer pairs were designed using SnapGene, while previous publications were used to inform the design of AAVS1 primers. Ordovas *et al.* (2015) found regions upstream of their targeted integration site in the AAVS1 locus were hypermethylated in hESCs, resulting in spread of methylation marks to the integrated cassette and inhibition of transgene expression. Thus, we designed several primer pairs to amplify upstream of the target site, specifically sites surrounding an EcoRI site at 2756 bp (see Figure 4.10 for illustration of approx. location of primers). Hypomethylation was also observed by Ordovas *et al.* (2015) at the start of the AAVS1 locus, following *PPP1R12C* exon 1; therefore, we designed primers to amplify this region as a control. In addition, we amplified UBE2B, a gene of known non-methylated status, to act as negative control in qPCR reactions (Mohn et al., 2009).

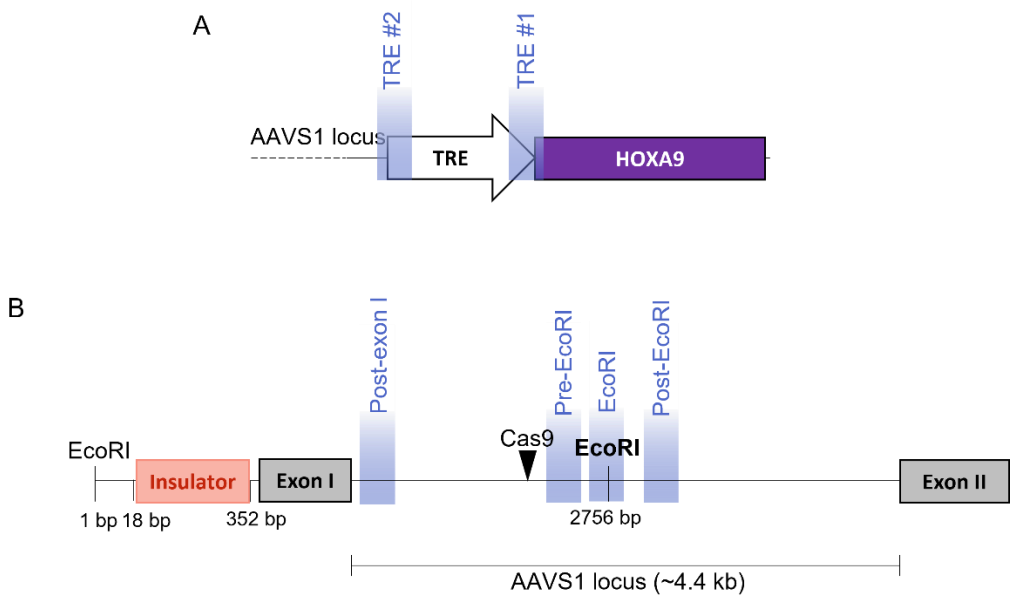


Figure 4.10: Representation of regions examined for DNA methylation within the TRE promoter and AAVS1 locus

Schematic depicting regions assessed for DNA methylation via MeDIP and qPCR in the TRE promoter (A) and AAVS1 locus (B). Approx. location of MeDIP primers are indicated by shaded blue boxes. (A) Methylation at the start (TRE #2 primers) and end of the TRE promoter (TRE #1 primers) was examined. (B) The AAVS1 locus and surrounding genomic regions are shown. Nucleotide numbering starts from the EcoRI site as was demonstrated in the first characterisation of the locus by Kotin *et al.* (1992). The insulator element identified by Ogata *et al.* (2003), location of the Cas9 target site (black arrow) and the EcoRI site at 2756 bp found to be hypermethylated in hESCs by Ordovas *et al*. (2015) are also depicted. QPCR primers were designed to amplify upstream of *PPP1R12C* exon I (Post-exon I), across the EcoRI site (EcoRI) and upstream (Post-EcoRI) and downstream (Pre-EcoRI) of the EcoRI site.

QPCR reactions were performed with gDNA isolated at different stages of the MeDIP procedure. Input control gDNA was collected after sonication of gDNA and prior to immunoprecipitation. As such, it represented a whole genome sample, comprising of both methylated and unmethylated gDNA fragments. Amplification of Input control gDNA in a qPCR reaction demonstrates the primer pair used binds to the gDNA fragment, regardless of methylation status (Figure 4.11A – C). Failure to amplify the Input control or generation of multiple PCR products indicates a primer pair is unreliable and data produced was not included in analysis. Immunoprecipitation results in isolation of the MeDIP fraction, containing only methylated gDNA. By normalisation of MeDIP to Input samples, the presence of methylation within the specific region targeted can be deduced. By this approach, the start of the AAVS1 locus (post-exon I primers) was found not to be methylated (Figure 4.11E), consistent with a previous assessment that the region is

hypomethylated (Ordovas et al., 2015). One site within the AAVS1 locus (EcoRI site primers) was enriched for methylation on d2 of cardiac differentiation (Figure 4.11E). Furthermore, the TRE promoter was shown to be enriched on d6 of cardiac differentiation (Figure 4.11D). This preliminary study suggests methylation in both the TRE promoter and AAVS1 locus are contributing to loss of *HOXA9*/mScarlet induction as differentiation progresses.

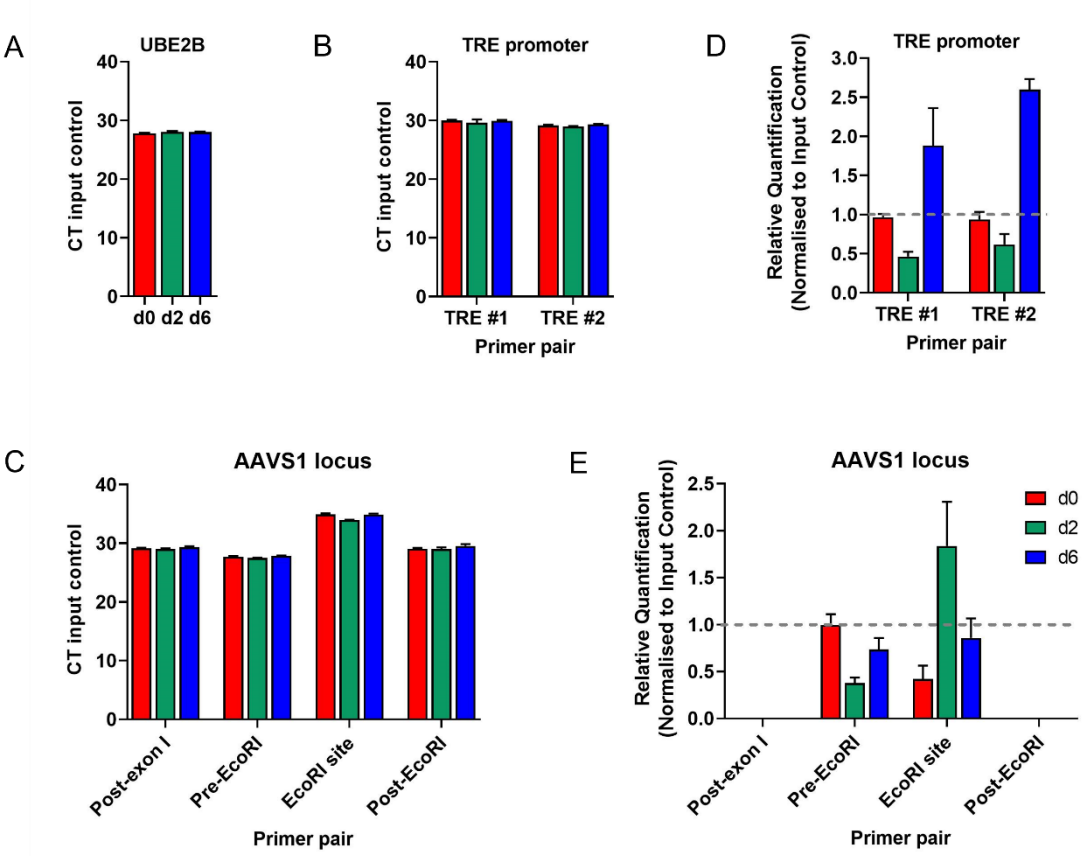


Figure 4.11: iA9 clone 5 is enriched for methylation on d2 and 6 of cardiac differentiation

(A – C) Bar graphs demonstrating Input control CT values in qPCR of target regions within the UBE2B gene (A), TRE promoter (B) and AAVS1 locus (C) in iA9 clone 5 cardiac differentiation on d0, d2 and d6. (D - E) qPCR for methylated regions with the TRE promoter (D) and AAVS1 locus (E) on d0, d2 and d6 of cardiac differentiation. Regions amplified are represented by their primer pair. Dotted grey line indicates the threshold for methylation enrichment. Results are mean ± SD, n = 3 technical replicates.

4.4.4 *HOXA9*-mScarlet expression can be maintained by continuous DOX treatment

It has previously been reported that transgene loss in systems affected by silencing can be minimised by prolonged stimulation of the TRE promoter (Mandegar et al., 2016). Therefore, we assessed if continuous DOX treatment throughout cardiac differentiation could ameliorate the loss in *HOXA9*/mScarlet induction observed in iA9 clones. iA9 clone 5 iPSCs were differentiated to cardiomyocytes and were either treated continuously with 1 µg/mL DOX from d-2 (DOX continuous) or received two DOX treatments at 24h intervals (DOX 48h) every two days beginning on d-2. As before, mScarlet imaging and qPCR was performed on samples throughout differentiation. The highest induction of *HOXA9* and mScarlet was observed on d0 in both conditions (mScarlet+: 71.6% DOX continuous; 75.8% DOX 48h). The ability of DOX to induce *HOXA9*/mScarlet was weaker on d2 in both conditions (mScarlet+: 36.7% DOX continuous; 33.9% DOX 48h). Subsequent decline in inducible *HOXA9*/mScarlet expression was minimal in DOX continuous conditions, with an mScarlet induction level of 23.4% obtained on d10 (Figure 4.12C). However, in 48h DOX samples, *HOXA9*/mScarlet induction progressively declined, with induced expression no longer detectable from d6 onwards (Figure 4.12A – C). This finding further supports the MeDIP data and suggests silencing of the TRE promoter plays a large role in the loss of inducibility observed in our system.

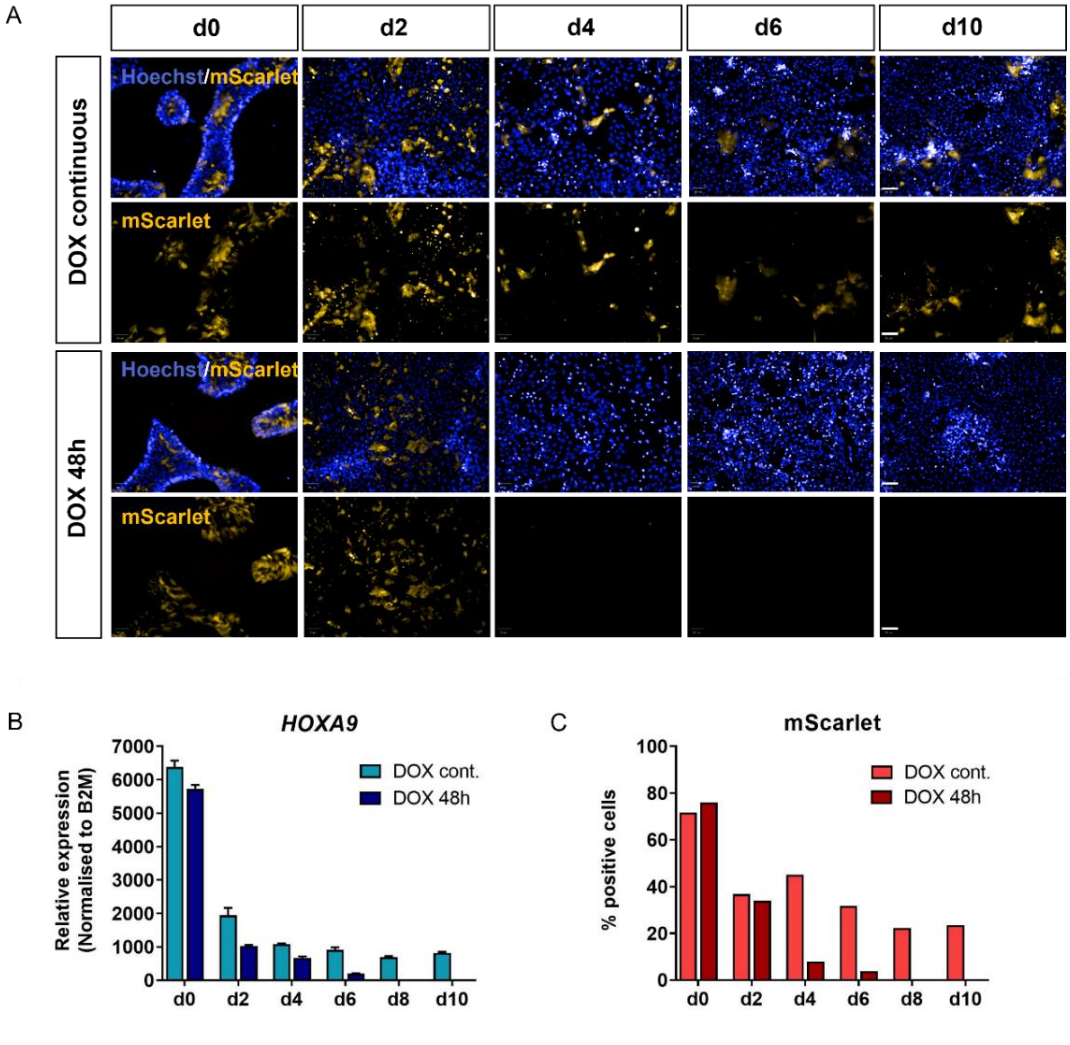


Figure 4.12: Inducible *HOXA9*/mScarlet expression can be maintained by continuous DOX treatment

iA9 clone 5 iPSCs were treated continuously with 1 µg/mL DOX, starting on d-2 (DOX cont.), or with two 1 µg/mL DOX treatments at 24h intervals on d-2, d0, d2, d4, d6 or d8 of cardiac differentiation (DOX 48h). (A) Representative images of DOX cont. and DOX 48h treatments throughout cardiac differentiation. Nuclei were co-stained with Hoechst 33341. Scale bar = 50 µm. (B - C) QPCR for HOXA9 (B) and quantification of mScarlet expression (C) in DOX cont. and DOX 48h cultures. Results are mean ± SD, n = 3 technical replicates.

4.5 Discussion

In this chapter, cardiac differentiation was used to screen for iA9 clones which could retain inducible *HOXA9*/mScarlet expression throughout differentiation. In all clones differentiated, inducible *HOXA9*/mScarlet expression declined as differentiation progressed, with induction not feasible after d6. Preliminary investigation of the methylation status of the AAVS1 locus and the integrated inducible HOXA9 cassette suggested both sites become methylated during cardiac differentiation.

We first optimised DOX treatment conditions to enhance the level and intensity of mScarlet expression in induced cells. As discussed in Chapter 3, the high background activity of mScarlet fluorescence hindered accurate identification of mScarlet-expressing cells, and consequently quantification of mScarlet expression. Compared with the single DOX treatment initially tested, multiple DOX treatments demonstrated a concentration-independent increase in mScarlet intensity with minimal increase in background. Single DOX treatment resulted in complete loss of the mScarlet and 2A signal within 72h for cells treated with 0.5 or 1 µg/mL DOX, and 96h in cell treated with 2 µg/mL DOX, similar to other TRE-driven transgenes (Xia et al., 2008). Multiple DOX treatments increased the number of mScarlet- and 2A-expressing cells at all concentrations, however, the effect of multiple treatments on mScarlet/2A degradation rate was not assessed. Two treatments of 2 µg/mL DOX resulted in the highest mScarlet and 2A induction at 57.2% and 56.1%, respectively. Surprisingly, three DOX treatments did not yield higher mScarlet or 2A expression at all concentrations tested, suggesting prolonged or excessive stimulation with DOX does not enhance transgene induction. The results also showed a close association between HOXA9/2A and mScarlet expression, demonstrating mScarlet is a suitable indicator of HOXA9 protein expression. The aim of this DOX optimisation experiment was to increase the mScarlet signal to facilitate more accurate measurement of mScarlet expression. As such, stimulation of the inducible system with higher DOX concentrations or experiments to identify concentrations which induced maximal HOXA9/mScarlet expression were not attempted.

Seven of the twelve correctly targeted iA9 clones were differentiated to cardiomyocytes and DOX induction, and the effect it had on *HOXA9*/mScarlet expression, was tested at different timepoints. During differentiation, the highest *HOXA9*/mScarlet expression in most iA9 clones was measured on d2. iA9 clones 20 and 22 did not follow this trend; during differentiation the highest *HOXA9* expression in clone 22 was observed on d4, while for clone 20 the highest mScarlet expression was found on d4. Both of these clones were treated with 2 µg/mL DOX which coincided with a reduction in cell number on d2 of differentiation, likely accounting for the discrepancy in expression observed. It is unclear

why 2 µg/mL DOX elicited this response on d2 but the DOX concentration, rather than HOXA9, may be responsible. In iPSC, no statistically significant difference was observed in 2A expression following treatment with 1 or 2 µg/mL DOX suggesting the level of HOXA9 induced did not differ between the two DOX concentrations. Additionally, 1 µg/mL DOX treatments did not stimulate large cell loss on d2 further suggesting HOXA9 is unlikely to play a role in the cell loss observed.

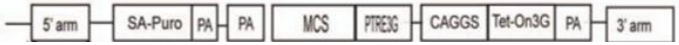
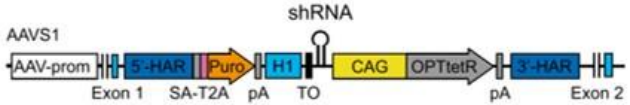
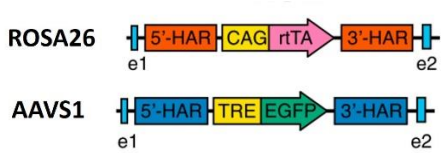
DOX is an antibiotic which acts by inhibiting mitochondrial protein synthesis. At a concentration of 1 µg/mL, DOX has been shown to decrease proliferation and modify metabolism in several human cell lines (Ahler et al., 2013). In contrast, at the same concentration DOX was found to promote the survival and proliferation of hPSCs, hPSC-derived neurons (Chang et al., 2014) and iPSC-derived endoderm (Peaslee et al., 2021). Activin A signalling is required to differentiate iPSC to definitive endoderm; however, the high dosage of Activin A required also leads to high cell death. Addition of DOX at approx. 1 µg/mL during definitive endoderm patterning was found to promote cell survival by inducing AKT signalling (Peaslee et al., 2021). The insulin/AKT pathway has dual roles during cardiomyogenesis. In iPSC differentiation to cardiomyocytes, insulin is required at late stages of differentiation, however, during early differentiation insulin signalling reduces cell viability and proliferation. Interestingly, insulin does not effect mesendoderm formation, assessed via *MIXL1*, *T* and *EOMES* expression, but prevents the formation of MESP1+ cardiac mesoderm and differentiation to cardiomyocytes (Lian et al., 2013). Therefore, the reduction in cell number observed following treatment with 2 µg/mL DOX may be due to DOX-mediated stimulation of AKT signalling. In support of this, *MIXL1* expression did not differ in iA9 clone 20 and 23 d2 cultures treated +/-DOX. Furthermore, 2 µg/mL did not have a negative effect on cell number at later time-points, when insulin signalling has an active regulatory role in cardiac differentiation. As DOX-treated cells were lysed for qPCR analysis following fluorescence imaging, the impact treatment with 2 µg/mL DOX may have on cardiac differentiation could not be fully assessed.

As discussed, during differentiation most iA9 clones demonstrated highest induction of *HOXA9/*mScarlet expression on d2 of differentiation. *MIXL1* qPCR indicated this time-point corresponds to the mesendoderm stage of cardiac development. In most iA9 clones, *MIXL1* expression was similar between DOX-induced and uninduced cultures. For clone 5, higher *MIXL1* was detected in uninduced cells. Similarly, iA9 clone 54 showed very low *MIXL1* expression, with highest expression in uninduced cells. During embryonic development, mesendoderm represents a transient state, where Nodal and Wnt signalling gradients direct mesendoderm cells to mesoderm or definitive endoderm lineages (Van Vliet et al., 2012). *In vitro* hPSC-based cardiac differentiation models

reflects this; differentiating cells are specified from a transient population with upregulated MIXL1 to MESP1+ mesoderm and NKx2.5+ cardiac mesoderm (Den Hartogh et al., 2016). As such, the low level *MIXL1* expression in DOX-induced clone 5 and 54 cells may be because cells have already passed through this intermediate state and are committed to mesoderm fate. Furthermore, we have found upregulation of HOXA9 promotes mesoderm specification under haematopoietic differentiation conditions (discussed in Chapter 6) and it is possible a similar process occurs during cardiac differentiation.

A limitation of these experiments is DOX responsiveness was only tested during differentiation, not in iPSCs. In Chapter 3, *HOXA9* and mScarlet expression in all twelve iA9 clones following 24h treatment with 1 µg/mL DOX was measured by qPCR and quantification of fluorescence imaging, respectively (Figure 3.14). In iPSCs, the seven iA9 clones which were differentiated in this chapter exhibited >10,000 upregulation in *HOXA9* relative to uninduced iPSCs (Figure 3.14B) In contrast, during cardiac differentiation the maximum *HOXA9* upregulation in each clone ranged from approx. 200 to 2500 relative to uninduced and undifferentiated iPSCs. As DOX treatment conditions differed in iPSC and cardiac differentiation induction experiments, a direct comparison cannot be made. However, in iA9 clone 5, DOX induction experiments were performed in iPSCs and during differentiation. Here, the ability of DOX to induce *HOXA9* was shown to be weaker in d2 cardiac cells than iPSCs, with the level of *HOXA9* induced differing by approx. 5.5-fold. Interestingly, mScarlet quantification did not reveal such substantial differences. mScarlet expression induced in iPSCs ranged from 13.9% to 44.7% (Figure 3.14C), while in cardiac differentiation maximal induced mScarlet expression of each clone ranged from 5.2% to 51.2%. Inaccurate quantification of mScarlet in iPSCs may account for this discrepancy, which is further supported by the approx. 40% reduction in mScarlet induced on d0 versus mScarlet induced on d2 in iA9 clone 5 differentiation. Thus, the data suggests the ability of DOX to induced *HOXA9*/mScarlet is weaker in differentiating cells than iPSCs. Importantly, the data also demonstrates that DOX is still able to induce relatively robust *HOXA9* and mScarlet expression on d2 of differentiation.

Table 4.1: Summary of Tet-On inducible expression systems which demonstrate robust transgene expression in different cell type

Tet-On system & integration site	Construct schematic	Targeting method	Cell types transgene expressed in	Reference
TRE3G drives multiple cloning site (MCS); CAG drives rtTA (Tet-On3G); Targeted to AAVS1 locus		TALENs	hESCs and differentiated neurons	Qian *et al.*, 2014
H1 promoter drives TetO (H1-TO) and shRNA; CAG drives optimised rtTA (OPTtetR); Targeted to AAVS1 locus		CRISPR/ Cas9	hESCs, 3 germ layers and 13 differentiated cell types, including cardiomyocytes, neurons and hepatocytes	Bertero *et al.*, 2016
CAG drives rtTA expression in ROSA26 locus; TRE3G drives EGFP expression in AAVS1 locus		ZFNs	hPSCs and differentiated neurons, skeletal myocytes and oligodendrocytes	Pawlowski *et al.*, 2017
TRE3G drives dCas9-KRAB (CRISPRi); CAG drives rtTA (TetON); Targeted to AAVS1 locus		TALENs	hPSCs	Hazelbaker *et al.*, 2020
TRE3G drives RUNX1-ETO; CAG drives rtTA; Targeted to AAVS1 locus		CRISPR/ Cas9	hESCs and differentiated definitive hematopoietic progenitors	Nafria *et al.*, 2020

Cardiac differentiation of iA9 clones demonstrated the loss of *HOXA9*/mScarlet induction observed during differentiation was not clone-specific as initially hypothesised (Bhagwan et al., 2020a). Although the MeDIP and continuous DOX treatment experiments are preliminary studies, they provide support for the hypothesis that the TRE promoter and AAVS1 locus become methylated as differentiation progresses, hindering the ability of DOX to adequately induce transgene expression. However, in contrast with our results several studies have demonstrated Tet-On systems targeted to the AAVS1 locus facilitate tightly regulated robust transgene expression in hPSCs and differentiated cell types (Bertero et al., 2016; Hazelbaker et al., 2020; Nafria et al., 2020b; Pawlowski et al., 2017; Qian et al., 2014). In comparison with our Tet-On system, these studies differ in their targeting approach; orientation of integrated elements; and in the Tet responsive element used (studies are summarised in Table 4.1). In these studies, transgene expression was driven by either the TRE3G promoter or a single tet operon (tetO) under control of the H1 promoter, while the tightTRE promoter was utilised in our system. On first inspection, the tightTRE and TRE3G promoters appear very similar; the promoters are 76.4% identical, both comprise of seven TetOs fused to a minimal CMV promoter, with each TetO separated by a 36 bp sequence (see Figure 4.13 for comparison of the different Tet responsive elements). However, the promoters differ in the location of the seven TetOs, which are placed upstream of the minimal CMV promoter in the tightTRE promoter, while they are placed downstream in TRE3G. Single nucleotide differences are also evident in the TATA box and TFIIb of the minimal CMV promoter. The TATA box and TFIIb nucleotide sequences in TRE3G were an early modification made to tet responsive elements in order to improve their efficiency (Loew et al., 2010). These differences may account for the loss of transgene induction observed in our system in comparison with other published Tet-On systems.

tightTRE (315 bp)

5' – GAGTTTACTCCCTATCAGTGATAGAGAACGTATGTCGAGTTTACTCCCTATCAGTGATAGA
GAACGATGTCGAGTTTACTCCCTATCAGTGATAGAGAACGTATGTCGAGTTTACTCCCTATCAG
TGATAGAGAACGTATGTCGAGTTTACTCCCTATCAGTGATAGAGAACGTATGTCGAGTTTATCC
CTATCAGTGATAGAGAACGTATGTCGAGTTTACTCCCTATCAGTGATAGAGAACGTATGTCGAG
GTAGGCGTGTACGGTGGGAGGCCTATATAAGCAGAGCTCGTTTAGTGAACCGTCAGATCGCC
- 3'

TRE3G (379 bp)

3' – GAGTTTACTCCCTATCAGTGATAGAGAACGTATGAAGAGTTTACTCCCTATCAGTGATAGA
AAACGTATGCAGACTTTACTCCCTATCAGTGATAGAGAACGTATAAGGAGTTTACTCCCTATCA
GTGATAGAGAACGTATGACCAGTTTACTCCCTATCAGTGATAGAGAACGTATCTACAGTTTACT
CCCTATCAGTGATAGAGAACGTATATCCAGTTTACTCCCTATCAGTGATAGAGAACGTATAAGC
TTTAGGCGTGTACGGTGGGCGCCTATAAAAGCAGAGCTCGTTTAGTGAACCGTCAGATCGCC
TGGAGCAATTCCACAACACTTTTGTCTTATACCAACTTTCCGTACCACTTCCTACCCTCGTAAA -
5'

H1-TO (84 bp)

5' – TCGCTATGTGTTCTGGGAAATCACCATAAACGTGAAATGTCTTTGGATTTGGGAATCTTAT
AAGTTCCCTATCAGTGATAGAGA – 3'

	tightTRE	TRE3G	H1-TO
Tet operon	TCCCTATCAGTGATAGAGA		
CMV initiator	CGTCAGA		
TFIIB	GGGAGGCC	GGG-CGCC	
TATA	TATATAA	TATAAAA	
H1 promoter			TCGCTATGTGTTCTGGGAAA

Figure 4.13: Tet responsive elements

The sequence of 3 different Tet responsive elements is shown, with key elements highlighted and further described in the table. TRE3G and tightTRE (Qian et al., 2014) are 76.4% similar, each possessing 7 tet operons (highlighted in grey) and a minimal CMV promoter. The CMV promoter comprises of a CMV initiator (red), TFIIB (blue) and TATA box (green). TFIIB and TATA sequences in TRE3G and tightTRE differ by 1 base as shown in the table, with differences highlighted in red. Major differences between the two promoters include the length and orientation of Tet operons, which are located upstream of minimal CMV promoter in tightTRE and downstream in TRE3G. H1-TO (Bertero et al., 2016) differs substantially from TRE3G and tightTRE, possessing a single Tet operon which is driven by the H1 promoter (highlighted in purple).

Silencing of transgene expression has been observed in other studies using the tightTRE promoter. Robust transgene expression was achieved in a lentivirus-mediated tightTRE driven system in hESC, however, transgene expression was reduced by 50 – 80% following differentiation to neurons, cardiomyocytes and endocrine cells (Xia et al., 2008). In contrast, stable integration of a Tet-On system into the AAVS1 locus did not

lead to loss of transgene expression in hESC-derived neurons (Qian et al., 2014). The loss in transgene expression observed in the former study was attributed to silencing of some but not all randomly integrated cassettes after differentiation. However, in the AAVS1-targeting study, TRE3G was used to drive transgene expression, indicating transgene loss after differentiation observed in Xia et al. (2008) is unlikely to be solely due to the random integration targeting approach and the tightTRE element may also have a role. In support of this theory, Piggybac-mediated random integration of a TRE3G Tet-On system was able to achieve sustained transgene expression following hPSC differentiation to cardiomyocytes (Randolph et al., 2017). Interestingly, in an inducible dCas9-KRAB expression system targeted to the AAVS1 locus, silencing was observed during cardiac differentiation despite the transgene being driven by TRE3G (Mandegar et al., 2016). The system was still able to achieve knockdown of targeted genes; however, generation of dCas9-KRAB proteins was reduced by ~20% and 50 – 80% in cardiac progenitors and cardiomyocytes, respectively. Therefore, although the Tet response element used in inducible systems plays a prominent role in transgene silencing, other factors are also involved.

Our preliminary MeDIP data suggests the AAVS1 locus is methylated on d2 of cardiac differentiation. As already discussed, methylation in the AAVS1 locus has been demonstrated in hESCs by bisulphite sequencing and postulated to also be involved in transgene silencing in endoderm lineage differentiation (Ordovas et al., 2015). Furthermore, methylation of promoters targeted to the AAVS1 locus has been observed. In an AAVS1-targeted expression system, all CpG dinucleotides present in the EF1α promoter were found to be methylated by bisulphite sequencing in hiPSC-derived hepatocytes resulting in loss of EF1α-driven transgene expression (Nitzahn et al., 2021). Interestingly, immediately after targeting, the EF1α promoter was not found to be methylated in iPSCs, however, methylation was introduced in iPSCs after several passages. Similarly, Tet-On systems in which the rtTA was driven by EF1α failed to express the transgene in hPSCs; however, robust expression was achieved when the CAG promoter was used (Pawlowski et al., 2017). Promoter methylation has also been demonstrated with tissue-specific promoters. Three different myeloid promoters targeted to the AAVS1 locus were shown by bisulphite sequencing to be initially methylated in hPSC as expected, yet following differentiation to myeloid cells the promoters remained methylated, ranging from 8% for one promoter to 80% methylated for another (Klatt et al., 2020). DNA methylation is known to fluctuate during lineage specification (Reizel et al., 2018), however, these studies demonstrate methylation can arise without differentiation and the AAVS1 locus may mediate inhibition of integrated transgenes or promoters.

It has previously been demonstrated that reduced gene expression can trigger silencing. The Tet-Off system was used to illustrate this concept in mouse cells edited to express a human transgene (Oyer et al., 2009). DOX was used to knockdown an integrated transgene, however, once DOX was removed, cells showed variable ability to recover transgene expression. Bisulphite sequencing revealed the TRE promoter of cells which did not express the transgene post-DOX treatment were methylated. Furthermore, ChIP demonstrated the cells also exhibited a histone profile consistent with repressed transcription. Our continuous DOX treatment experiment supports this concept. Continuous stimulation of the TRE promoter prevented complete loss of *HOXA9*/mScarlet induction, suggesting the transient application of our inducible system may contribute to the observed loss.

Of note, we did not investigate methylation of the CAG-rtTA element as a potential contributor to the loss of inducible *HOXA9*/mScarlet during differentiation. The CAG promoter has shown variable expression in human cell lines (Qin et al., 2010); however, the same study also found the EF1α promoter supports strong expression of rtTA, which conflicts with other studies performed in hPSCs (Nitzahn et al., 2021; Pawlowski et al., 2017). A limitation of our study is the stability of rtTA expression throughout differentiation was not assessed (for example, by qPCR). However, given the prominence of CAG-driven rtTA elements in successful Tet-On systems targeted to the AAVS1 locus (see Table 4.1), it is unlikely it has a role in silencing in our system.

Evidently several factors are contributing to the loss of DOX-mediated *HOXA9*/mScarlet expression during cardiac differentiation. Preliminary experiments suggest methylation of the TRE promoter and AAVS1 locus are both responsible for the expression loss observed. A review of the literature further supports this hypothesis. The tightTRE promoter has been demonstrated to be a poor promoter for Tet-On systems and promoter-specific methylation was found to occur in the AAVS1 locus, despite its reported protective features. Although DOX-mediated induction of *HOXA9* and mScarlet is weaker on d2 of cardiac differentiation than in iPSCs, robust expression is still observed. Therefore, the inducible HOXA9 expression system still permits study of HOXA9 overexpression in early mesoderm, and further assessment of how it impacts differentiation to mesoderm-derived lineages.

Chapter 5: Optimisation of haematopoietic differentiation of hiPSCs

5.1 Introduction

The DOX-inducible HOXA9-mScarlet (iA9) cell line was found to not retain its inducible activity in differentiated cell types in the previous chapter. After mesendoderm induction, a rapid reduction in DOX responsiveness was observed in iA9 cardiac differentiation cultures. The iA9 cell line was generated to establish the function of HOXA9 during haematopoietic development; specifically, through temporal overexpression of HOXA9 at specific time-points in a hiPSC-based model of haematopoiesis. While the iA9 cell line does not permit HOXA9 induction throughout differentiation, the function of HOXA9 in mesendoderm, and thus haematopoietic lineage commitment, can be studied. This chapter focuses on optimisation of hiPSC differentiation to haematopoietic progenitors to provide an *in vitro* haematopoiesis model suitable for HOXA9 overexpression experiments.

5.1.1 Human PSC-based haematopoietic differentiation

Human PSC-derived models of haematopoiesis aim to recapitulate the events of *in vivo* embryonic haematopoiesis. Embryonic haematopoiesis (discussed in detail in chapter 1) takes place in three successive waves of haematopoietic activity and across different sites in the developing embryo. Briefly, extra-embryonic mesoderm is first specified from the primitive streak (PS) and subsequently generates the blood islands in the yolk sac (Palis et al., 1999). The first and second waves of haematopoiesis initiate in the blood islands, with the first wave producing predominately primitive erythrocytes, while more specified cell types, namely EMPs and lymphoid progenitors, are generated in the second wave (Palis et al., 1999; Yoshimoto et al., 2011). Haematopoietic cells originate from bipotent haemogenic endothelial cells, which are derived from mesodermal cells (Cortes et al., 1999). The third wave of haematopoiesis occurs in the embryo proper in the AGM region of the dorsal aorta. Here, haemogenic endothelial cells lining the ventral wall of the dorsal aorta undergo an EHT to produce HSCs (Kissa and Herbomel, 2010). The AGM is a temporary site for haematopoiesis and produces low number of HSCs (Gekas et al., 2005; Rybtsov et al., 2016). Differentiation and expansion of HSCs takes place in the yolk sac (Kumaravelu et al., 2002), placenta (Gekas et al., 2005) and foetal liver, the major site of haematopoiesis during foetal life (Rybtsov et al., 2016).

In hPSCs, PS and mesoderm induction is achieved by co-ordinated activities of the TGFβ, Wnt and BMP signalling pathways. As is the case *in vivo*, PS formation is directed by Activin/Nodal and Wnt signalling and represents a transient state with mesoderm and endoderm differentiation potential, characterised by upregulation of *T*, *MIXL1* and *FOXF1* (Nostro et al., 2008; Tan et al., 2013). Short-term BMP4 is a strong mesoderm inducer and acts synergistically with Activin/Nodal and Wnt to promote specification of KDR+ mesoderm (Cerdan et al., 2012; Pearson et al., 2008; Zhang et al., 2008). VEGF is necessary to specify HE cells and its activity can be enhanced by FGF2 (Uenishi et al., 2014; Wang et al., 2012). Haematopoietic commitment of the HE is marked by loss of endothelial characteristics and acquisition of a haematopoietic signature, notably the surface marker CD43 (Vodyanik et al., 2006). Multipotent $CD34^+CD43/CD45^+$ haematopoietic progenitors emerge upon further culture with supportive cytokines such as SCF, FLT3, IL3, IL6 and TPO (Kardel and Eaves, 2012). Prior to haematopoietic commitment, the HE can be further specified to arterial-type HE, reminiscent of the arterial endothelium lining the ventral wall of the aorta *in vivo* (Lange et al., 2021). Arterial-type HE is characterised by Notch dependency and upregulation of arterial-specific genes, such as *SOX17*, *HEY2* and *CXCR4*. (Gama-Norton et al., 2015; Uenishi et al., 2018). Arterialisation of hPSC-derived HE is considered essential to generate haematopoietic progenitors with lympho-myeloid potential and replication of intra-embryonic definitive haematopoiesis (Park et al., 2018; Uenishi et al., 2018).

In embryonic development, blood cell formation takes place in three distinct programs: primitive, definitive EMP and definitive HSC haematopoiesis. Contrary to embryonic development, these phases can occur simultaneously in hPSC-derived haematopoiesis models (Ackermann et al., 2015). The three programs share several similarities; all haematopoietic cells develop from KDR+ mesoderm, transition through $CD34^+CD43^-$ endothelial-derived intermediates and acquire CD43 expression as they are generated (Canu and Ruhrberg, 2021; Ditadi et al., 2015; Garcia-Alegria et al., 2018). Several surface marker and gene expression patterns have been identified which can aid in discrimination between the programs *in vitro*. CD235a expression in KDR+ mesoderm indicates mesoderm fated to the primitive program (Sturgeon et al., 2014). Endothelium with haematopoietic potential can be distinguished from vascular endothelium by lack of CD73 (Choi et al., 2012) and CD184 (Ditadi et al., 2015), while HE with lympho-myeloid potential is further distinguished by expression of arterial-specific markers DLL4 and CXCR4 (Uenishi et al., 2018). Functional assays have also been developed to discriminate between specified cell types. Primitive erythroid progenitors are the first haematopoietic lineage derived during hESC differentiations and this stage can be

recapitulated in the blast colony-forming cell (BL-CFC) assay (Choi et al., 1998; Kennedy et al., 2007). In this assay, KDR+ cells seeded in methylcellulose with haematopoietic cytokines develop into either small erythroid colonies, comprising of primitive erythrocytes, or large colonies consisting of erythroid and macrophage cells (Kennedy et al., 2007). Thus, the BL-CFC assay can be used to assess early acquisition of haematopoietic potential, indicative of the primitive program. T lymphoid differentiation potential is a key feature of definitive programs and can be assayed by co-culture of haematopoietic progenitors on OP9 stroma (Kennedy et al., 2012). The globin expression profile of haematopoietic progenitor-derived erythroid colonies in the CFU assay can also distinguish primitive and definitive progenitors. Erythroid colonies which express embryonic (ε or ζ) or foetal (γ) globin are indicative of primitive erythrocytes generated from primitive progenitors (Kennedy et al., 2007; Yang et al., 2017a). Expression of adult (α or β) globins and evidence of globin switching from embryonic/foetal to adult indicates differentiation of mature erythroid cells, derived from definitive progenitors (Ng et al., 2016; Yang et al., 2017a).

Endogenous signalling pathways can also be manipulated during hPSC-based differentiations to promote definitive haematopoiesis. During mesoderm specification, the primitive program is temporarily dependent on Activin/Nodal signalling, and consequently, short-term inhibition of Activin/Nodal signalling inhibits primitive haematopoiesis and promotes the definitive program (Kennedy et al., 2012). The mesoderm stage of hPSC-based differentiation is also sensitive to Wnt signalling. Brief inhibition of Wnt signalling from d2 – d3 of differentiation promotes upregulation of CD235a in KDR+ mesoderm, while activation of Wnt signalling prevents the emergence of CD235a+ cells, subsequently promoting derivation of CD34+ progenitors with T lymphoid potential (Sturgeon et al., 2014). Similarly, in hESC-derived BL-CFCs, addition of Wnt3A supresses primitive haematopoiesis (Gertow et al., 2013). A combined approach, with both Activin/Nodal inhibition and Wnt induction during early differentiation, results in multipotent haematopoietic progenitors which develop from a $SOX17^{+}RUNX1c^{-}CD34^{+}CD90^{+}$ endothelium, phenotypically similar to human AGM (Ng et al., 2016). Thus, modulation of Activin/Nodal and Wnt signalling pathways during mesoderm patterning can be used to suppress primitive haematopoiesis, leading to promotion of the definitive program.

EMP and HSC haematopoietic programs can be distinguished *in vitro* based on Notch dependency and *HOXA* expression patterns. Haematopoietic progenitors can be generated in both a Notch-dependent and independent manner. In Notch-dependent

haematopoiesis, progenitors possess greater *in vitro* haematopoietic potential than Notch-independent progenitors, and arise from HE which has undergone arterialisation (Uenishi et al., 2018), akin to endothelial cells of the dorsal aorta from which HSCs first develop during normal development (De Bruijn et al., 2000). Intra-embryonic haematopoietic potential can be identified by upregulation of HOXA genes in hPSC-derived HE, while low or absent *HOXA* expression signifies extra-embryonic potential (Dege et al., 2020). The best approach to distinguish between the two definitive programs remains transplantation to immunodeficient mice. Only *bona fide* HSCs are able to reconstitute the entire blood system of immunodeficient mice (Zovein et al., 2008), providing an unambiguous means of determining HSC potential.

A diverse range of protocols exist for differentiation of haematopoietic progenitors from hPSCs, primarily through co-culture on stromal cell lines, monolayer-based differentiation on extracellular matrices (ECM) or three-dimension embryoid body (EB) differentiation (Georgomanoli and Papapetrou, 2019). Additionally, the precise method of hPSC differentiation varies. Morphogen-directed differentiation, ectopic overexpression of transcription factors and forward reprogramming of hPSCs to haematopoietic cell-types have all been applied to generated haematopoietic progenitors, although with varying functional activity (see Table 5.1).

Table 5.1: Summary of key hPSC-derived haematopoietic differentiation protocols

Approach	Cell Source	TFs/conditions	Lineages	Engraftment *in vivo*	References
Directed differentiation (EB)	hESC	6 µM SB d2 – d4 5% O_2 d0 – d8	Erythroid, Myeloid, T Lymphoid	Not attempted	Kennedy et al. 2012
Directed & TF-driven differentiation	hPSC-derived CD34+CD45+	HOXA9, ERG, RORA, SOX4, MYB	Erythroid, Myeloid, Lymphoid	Short-term	Doulatov et al. 2013
Directed differentiation (EB)	hPSC	3 µM CHIR d2 – d3 5% O_2 d0 – d8	Erythroid, Myeloid, NK, T Lymphoid	Not attempted	Sturgeon et al. 2014
Directed differentiation (monolayer)	hPSC	Tenascin C 5% O_2 d0 – d5	Erythroid, Myeloid, T Lymphoid	Unsuccessful	Uenishi et al. 2014
Directed differentiation (EB)	hESC	3 - 4 µM SB/3 µM CHIR d2 – d4 Spin EB d0 – d7/8 Matrigel d7/8+	Erythroid, Myeloid, T Lymphoid	Unsuccessful	Ng et al. 2016
Directed & TF-driven differentiation	hPSC-derived HE	ERG, HOXA5, HOXA9, HOXA10, LCOR, RUNX1, SPI1	Myeloid, Erythroid, Megakaryocytic, T-cell, B-cell	Long-term	Sugimura et al. 2017
Directed & TF-driven differentiation	hiPSC-derived blood cells	MLL-AF4	Myeloid, Erythroid, T-cell, B-cell	Long-term	Tan et al. 2018
Directed differentiation (monolayer)	hiPSC	STEMdiff™ Hematopoietic kit Matrigel 3 µM SB/3 µM CHIR d2 – d3	Erythroid, Myeloid	Unsuccessful	Ruiz et al. 2019
Directed & TF-driven differentiation	hiPSC	SCL, LMO2, GATA2, ETV2	Myeloid, Erythroid, Megakaryocytic, NK-cell	Unsuccessful	Lange et al. 2020

5.1.2 Directed differentiation

Directed differentiation approaches have been employed in stromal co-culture, EB-, and monolayer-based platforms. Co-cultures with stromal cell lines, such as OP9 and E4ORF1, rely on intrinsic cues from the stromal microenvironment to direct hPSC differentiation (Choi et al., 2012; Rafii et al., 2013; Vodyanik et al., 2006). However, dependence on the microenvironment limits external control over haematopoietic cell fate, which can in turn reduce differentiation reproducibility and efficiency. Stromal co-culture systems often use animal-derived serum, making them susceptible to batch-to-batch variation which can also influence differentiation characteristics and efficiency. (Kardel and Eaves, 2012; Uenishi et al., 2014).

Monolayer-based approaches use serum-free, chemically defined media and ECM proteins to support haematopoietic differentiation. Commonly used ECM substrates include Matrigel (Cao et al., 2019; Mills et al., 2014; Niwa et al., 2011), vitronectin (Jeong et al., 2020; Shen et al., 2021) and collagen IV (Jung et al., 2021; Uenishi et al., 2018). Tenascin C, a protein identified in OP9 systems, has also been shown to support hPSC haematopoietic differentiation (Uenishi et al., 2014). In monolayer directed differentiation approaches, hPSC are seeded as colonies at low density and extrinsic growth factors, cytokines and small molecules are used to induce differentiation to $CD34^+CD43/CD45^+$ haematopoietic progenitors. Progenitors produced predominately exhibit erythro-myeloid potential (Mills et al., 2014; Niwa et al., 2011), although some lymphoid potential has been demonstrated (Hansen et al., 2018; Uenishi et al., 2014). Further differentiation of monolayer-derived progenitors to monocytes, macrophages and DCs has also been achieved (Cao et al., 2019; Yanagimachi et al., 2013).

Numerous EB differentiation protocols exist, differing in their serum and O_2 requirements, EB size, and EB formation approach (Hong et al., 2010). EBs can be produced via dissociation of hPSC as clumps and culture in suspension on bacterial-grade plates or as hanging drop on the lid of a petri dish (Kurosawa, 2007). EBs produced by these approaches tend to be cultured in groups, which is reported to enhance haematopoietic development due to paracrine signalling between EBs (Hong et al., 2010). However, these approaches can lead to heterogeneity in EB size and morphology, reducing reproducibility (Hong et al., 2010). Alternatively, EBs can be formed by forced aggregation of hPSC dissociated as single cells, such as the spin EB method in which hPSCs are centrifuged in 96-well U-bottom plates to create EBs of uniform size (Ng et al., 2005). Such EB directed differentiation protocols aim to reproduce the spatial organisation of the early embryo and induce differentiation to the haematopoietic lineage

by administration of extrinsic morphogens, mimicking endogenous signalling cues (Dege and Sturgeon, 2017; Kennedy et al., 2012; Sturgeon et al., 2014). EB and monolayer approaches have also been combined, whereby PSCs are initially differentiated as EBs to haemato-endothelial cells, at which point EBs are transferred to Matrigel-coated plates. EBs are either transferred directly (Ng et al., 2016) or dissociated and seeded as single cells on the Matrigel plate (Ditadi et al., 2015), thus allowing generation of a supportive adherent endothelial niche for haematopoietic progenitor cell derivation.

EB- and monolayer-based directed differentiation protocols share a similar stepwise process. In both, hPSC are first directed to T+ mesendoderm and $KDR^{+}CD34^{+}$ haemogenic mesoderm (Pearson et al., 2015). A haemogenic endothelium intermediate is then specified, identified by co-expression of CD31, CD34, CD144 and KDR and lack of CD43 and CD45 (Garcia-Alegria et al., 2018). HE cells subsequently undergo EHT, characterised as a morphological change whereby flat, adherent HE cells become round suspension cells, characteristic of haematopoietic cells (Lancrin et al., 2012). EHT and haematopoietic commitment are illustrated molecularly by downregulation of endothelial genes and upregulation of haematopoietic-specific transcription factors, such as RUNX1, *GATA2, GFI1* and *GFI1b* (Lange et al., 2021).

Haematopoietic cells capable of long-term multilineage engraftment upon transplantation has not been achievable with current defined directed differentiation protocols, suggesting *bona fide* HSCs are not generated (Canu and Ruhrberg, 2021). As such, monolayer and EB-based approached most likely recapitulate yolk sac-derived EMP haematopoiesis.

5.1.3 Transcription factor-driven differentiation

A combination of directed and transcription factor-driven differentiation approaches have shown more success at generating engraftable haematopoietic progenitors than directed differentiation alone. Human PSC-derived $CD34^{+}CD45^{+}$, which initially lacked reconstitution ability, acquired short-term erythro-myeloid engraftment in NSG mice when factors *ERG, HOXA9, RORA, SOX4* and *MYB*, were overexpressed (Doulatov et al., 2013). In another study, ectopic expression of *SCL, GATA2, LMO2* and *ETV2* in hPSC-derived $CD144^{+}CD73^{-}$ HE produced $CD34^{+}CD45^{+}$ haematopoietic progenitors which co-expressed CD90 and CD49f and lacked CD45RA and CD38 expression. Despite their HSC-like immunophenotype, progenitors showed a myeloid bias when transplanted to NSG mice and were not able to achieve long-term engraftment (Lange et al., 2020). Successful long-term engraftment has been achieved by overexpression of

MLL-AF4 in hiPSC-derived haematopoietic cells. Initially, transplanted mice exhibited erythroid, myeloid, B and T cell reconstitution, however, after 16 weeks only lymphoid cells persisted, suggesting possible leukaemic transformation of the transplanted cells (Tan et al., 2018). Long-term multilineage reconstitution without adverse effects was achieved by overexpression of *ERG, HOXA5, HOXA9, HOXA10, LCOR, RUNX1* and *SPI1* in hiPSC-derived HE. However, haematopoietic progenitors were generated following transplantation, suggesting the *in vivo* environment was crucial for formation of engraftable HSPCs (Sugimura et al., 2017). Thus, transcription factor-driven differentiation approaches can reproduce intra-embryonic definitive HSC haematopoiesis, although a completely defined *in vitro* approach, which does not require genetic modification, has not yet been achieved.

5.1.4 Forward reprogramming

Forward reprogramming of hPSCs to haematopoietic progenitors has also been achieved by overexpression of key endothelial and/or haematopoietic transcription factors. Transduction of hPSCs with *ETV2* and *GATA2* produced $CD43^+$ haematopoietic cells within 7 days that possessed multilineage differentiation potential in the CFU assay. Similarly, CD43+ cells were produced in the same timeframe by reprogramming hPSCs with *GATA2, TAL1* and *LMO2*, however, differentiation potential was limited to erythroid and megakaryocyte lineages (Elcheva et al., 2014). Interestingly, forward reprogramming of non-haematopoietic endothelial cell line HUVEC with *FOSB1, GFI1, RUNX1* and *SPI1* generated multipotent haematopoietic progenitors which could reconstitute primary and secondary recipients (Sandler et al., 2014). This study demonstrates forward reprogramming can be used to generate multipotent, engraftable haematopoietic progenitors, suggesting, with an appropriate combination of transcription factors, it will be feasible in hPSCs.

5.2 Aims

The primary aim of this chapter was to optimise an hPSC-based haematopoietic differentiation protocol which could be used to study the effects of inducible HOXA9 overexpression. In doing so we also aimed to determine distinct gene and surface marker expression patterns at different time-points during differentiation. This would enable us to mark and characterise differentiation progression. Furthermore, we aimed to streamline the procedure by identifying early measures to predict differentiation efficiency.

5.3 Results

5.3.1 Method selection and optimisation

A monolayer haematopoietic differentiation procedure using the STEMdiff™ Hematopoietic kit was first adopted. The STEMdiff™ Hematopoietic kit facilitates differentiation of hPSCs in chemically defined media in serum- and feeder-free conditions to $CD34^+CD43^+CD45^+$ haematopoietic progenitors. The STEMdiff™ differentiation procedure is relatively straight-forward; hPSCs are disaggregated as clumps and seeded at low density in Matrigel-coated 12-well plates in E8 medium. The following day, media is changed to Media A for 3 days, and then Media B until day 12 of differentiation, at which point floating cells are collected and assessed for haematopoietic potential (see Figure 3.18 for overview). Confluent RPAT hiPSCs were passaged using the recommended dissociation agent ReLeSR and seeded in E8 medium containing 10 µM Rock inhibitor to aid cell attachment. Cells were seeded based on splitting ratios as counting clumps to seed at specific numbers was found to not be accurate. ReLeSR is an enzyme-free dissociation reagent designed for clump passaging of hPSCs. However, even with extension of ReLeSR incubation time, dissociation of RPAT hiPSCs was very challenging, with complete dissociation not achievable. To be suitable for differentiation, aggregates need to have formed compact colonies ~200 µm in diameter 24h after seeding (manufacturer's protocol). ReLeSR resulted in mostly dispersed colonies that tended to be larger than 200 µm in diameter and which formed monolayers rather than colonies after 2 – 3 days in culture (Figure 5.1A). In clump passaging, cell-cell adhesions are maintained and Rock inhibitor should not be required (Beers et al., 2013). Therefore, Rock inhibitor was removed from newly seeded cultures, resulting in compact colonies of the desired size after 24h (Figure 5.1A). Although, seeding without Rock inhibitor improved the morphology of aggregates, issues with dissociating hiPSCs using ReLeSR persisted.

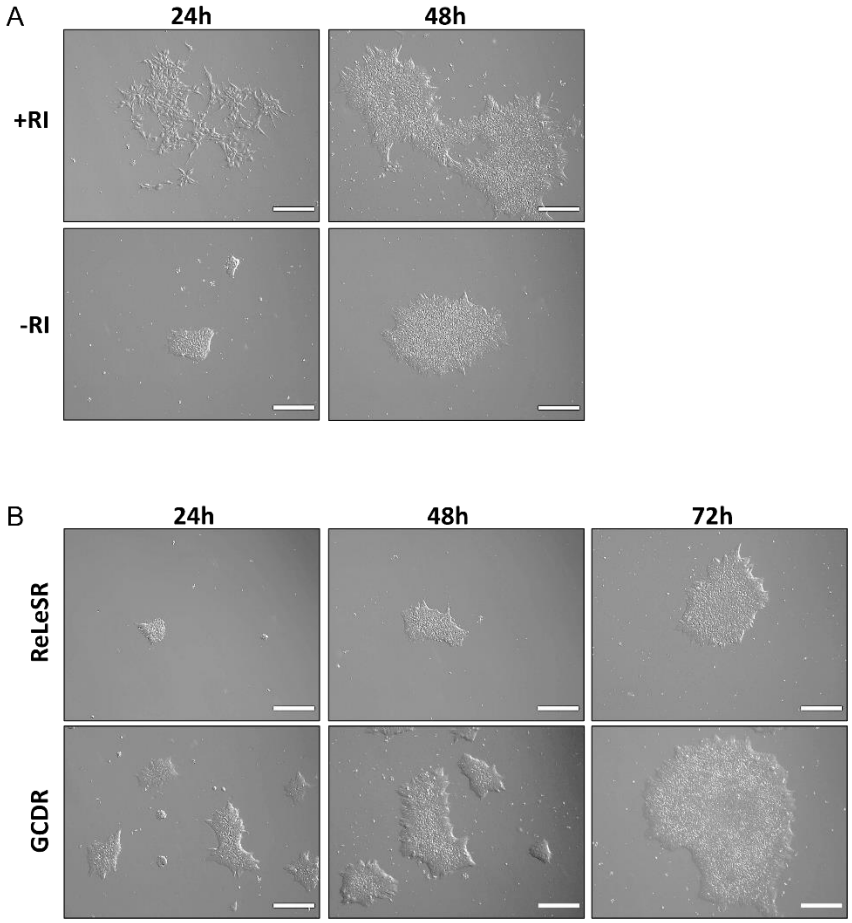


Figure 5.1: Optimisation of aggregate seeding conditions for STEMdiff™ Hematopoietic differentiation

Optimisation of RPAT hiPSC colony culture. (A) Representative light microscope images of RPAT hiPSCs treated +/- 10 µM Rock Inhibitor (RI) for 24h following dissociation with ReLeSR. (B) Representative light microscope images of RPAT hiPSC 24h, 48h and 72h post-dissociation with either ReLeSR or GCDR. Cells were seeded at a split ratio of 1:12 and seeded in 6-well plates. Scale bar = 200 µm.

RPAT hiPSCs are routinely passaged as single cells using TrypLE. Therefore, it was hypothesised RPATs may first require conditioning to colony passage in order to achieve complete dissociation. GCDR is an enzyme-free dissociating agent which requires scraping of cells from the surface to achieve dissociation, making it a harsher agent than ReLeSR. GCDR was used to condition RPAT hiPSCs, with the intention of then using ReLeSR following conditioning to generate aggregates of optimal size for differentiation. All RPAT hiPSCs were detached following GCDR incubation and scraping. Aggregates formed 24h post-seeding were 100 – 200 µm in diameter and after 3 – 4 days in culture formed large colonies with tightly compact centres (Figure 5.1B). Despite the harsher

process, cells could be passaged several times in GCDR without adverse effects in colony quality. Furthermore, tweaking of colony density was more readily achieved using GCDR. The STEMdiff™ procedure outlines 16 – 40 aggregates/well are necessary for haematopoietic differentiation. As GCDR was more efficient at dissociating cells, splitting ratios which could yield the desired number of aggregates were easier to predict. As such, GCDR was used in all experiments requiring clump passaging.

5.3.2 STEMdiff™ Hematopoietic kit produces $CD34^+CD38^-$ progenitors at varying efficiencies

RPAT hiPSC were dissociated using GCDR and seeded for haematopoietic differentiation in Matrigel-coated 12-well plates. Only wells which contained the optimal number of colonies (16 – 40 colonies) between 50 and 200 µm in diameter 24h after seeding were selected for differentiation, which was initiated by change to Media A (d0). Floating cells were produced by d12 and were assessed for haematopoietic potential by flow cytometry for markers CD34 and CD38 (Figure 5.2A) and colony formation ability using the CFU assay. In a comparison of 12 independent RPAT hiPSC differentiation experiments, suspension cells produced on d12 differed greatly in their CD34/CD38 expression (Figure 5.2B). Differentiations were found to vary in the total number of suspension cells produced (Figure 5.2C) and some differentiations failed to generate suspension cells (differentiation #1, #5 and #6). Furthermore, differentiation experiments which produced suspension cells, produced predominately non-hematopoietic $CD34^-CD38^-$ cells. The $CD34^+CD38^-$ hematopoietic progenitor fraction varied amongst the 12 differentiation experiments performed, ranging from 2 – 23.9% of cells assayed (Figure 5.2B). The CFU assay suggested most progenitors lacked haematopoietic potential. Haematopoietic colonies were only produced with suspension cells from four differentiation experiments (#3, #4, #7 and #9); the remaining differentiations failed to produce suspension cells with colony-forming ability (Figure 5.2E). Colonies formed were mainly of the granulocyte-macrophage (GM) sub-type, suggesting the RPAT-derived suspension cells possess limited or myeloid-restricted haematopoietic activity.

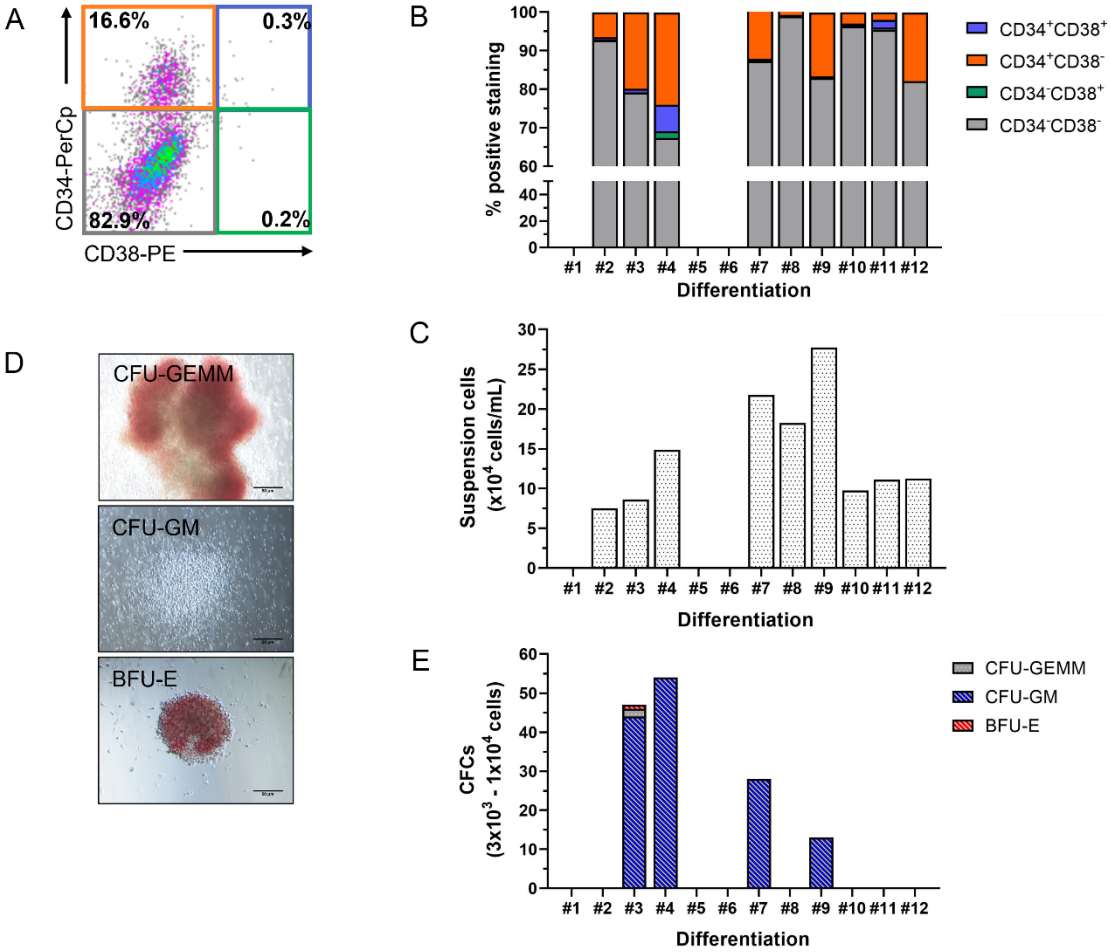


Figure 5.2: Haematopoietic differentiation using STEMdiff™ Hematopoietic kit produces CD34+CD38- haematopoietic progenitors at varying efficiencies

RPAT hiPSC differentiation to haematopoietic progenitors using the STEMdiff Hematopoietic kit. Blank columns in bar graphs indicate differentiations which failed to produce suspension cells/haematopoietic colonies. (A) Gating strategy used to identify different cell populations defined by CD34 and CD38 expression from suspension cells collected on d12. (B) CD34/CD38 expression profile of suspension cells from 12 independent differentiations (#1 - #12). (C) Number of suspension cells/mL collected from each differentiation. (D) Representative images of haematopoietic colonies formed in MethoCult™ Classic H4434 or Optimum H4034 (STEMCELL Technologies) after 10 days. Scale bar = 50 µm. (E) Total colony forming cells (CFCs) produced by suspension cells from 12 independent differentiations.

5.3.3 Spin EB differentiation shows poor mesoderm differentiation

The STEMdiff™ Hematopoietic kit was found to be variable and inconsistent, therefore, we opted to try a different procedure entirely. As other monolayer-based differentiation protocols similarly rely on clump passaging of hPSCs, an EB-based protocol was adopted. The spin EB method of EB formation permits generation of uniform EBs of reproducible size and robust haematopoietic differentiation potential (Ng et al., 2005, 2016). Andrew Elefanty's lab have generated a spin EB-based differentiation procedure which uses growth factors and small molecules to direct differentiation of hESCs to SOX17+ haemogenic endothelium, from which multipotent RUNX1c^{+}CD34^{+}CD43^{+} haematopoietic progenitors subsequently arise (Ng et al., 2016). The procedure follows a stepwise process; EBs are first formed in U-bottom 96-well plates by centrifugation and ratio splitting; whereby one T25 of hPSCs seeds three plates, constituting 60 EBs/plate (outer wells are excluded). EBs are then differentiated to haemato-vascular mesoderm, at which point, EBs are transferred to Matrigel-coated 6-well plates. EBs attach to the underlying Matrigel and haemogenic endothelial cells and haematopoietic progenitors are subsequently generated.

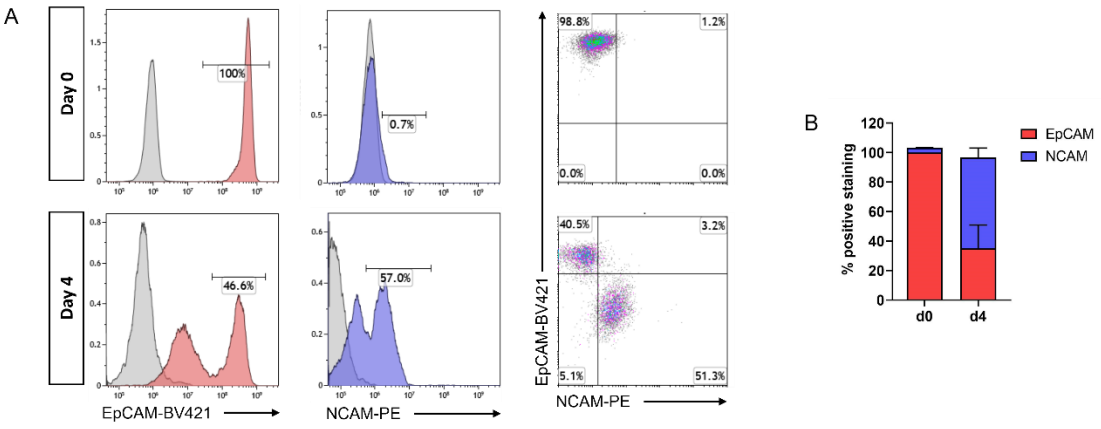


Figure 5.3: Spin EB haematopoietic differentiation shows poor mesoderm specification

(A) Representative histograms and dot plots of flow cytometric analysis of pluripotency marker EpCAM and pan-mesoderm marker NCAM on d0 and d4 of RPAT hiPSC differentiation to haematopoietic progenitors using the spin EB protocol. Unstained controls are overlaid in grey. (B) Quantification of EpCAM- and NCAM-expressing cells in d0 and d4 EBs. Results are mean ± SD, n = 2 biological replicates.

The spin EB haematopoietic differentiation procedure developed in Elefanty's lab was adopted. The differentiation procedure reported uses a homemade differentiation media

APEL (Ng et al., 2008) which is now commercially available (APEL2, STEMCELL Technologies #05275). APEL2 is an animal component-free, defined media that differs from homemade APEL due to the absence of bovine serum albumin and protein-free hybridoma. Due to convenience and to eliminate batch-to-batch variability associated with animal-derived media components, all spin EB differentiations were performed using APEL2. Initial attempts to generate haematopoietic progenitors from RPAT hiPSCs failed to produce suspension cells. Flow cytometry was performed on d0 and d4 for pluripotency marker EpCAM/CD236 and pan-mesoderm marker NCAM. Loss of EpCAM and upregulation of NCAM denotes commitment to a mesoderm-derived fate and marks a population which can generate multiple mesoderm lineages, including haematopoietic, endothelial, cardiac and mesenchymal (Evseenko et al., 2010). EpCAM expression was reduced by 65% from d0 to d4, while NCAM expression increased by 62% (Figure 5.3A, B). EpCAM- and NCAM-positive cells formed distinct populations on d4 (Figure 5.3A), indicating cells either remained pluripotent or underwent mesoderm commitment under the conditions used. The large fraction of pluripotent cells on d4, alongside lack of suspension cells at the end of differentiation, suggests mesoderm commitment was inefficient at the start of differentiation, resulting in the poor differentiation observed.

5.3.4 Spin EB differentiation of feeder- and Matrigel-cultured hPSCs do not produce suspension cells

The spin EB differentiation procedure was first published in 2016 (Ng et al., 2016), however, a research article was later published by the Elefanty group using an amended spin EB differentiation procedure (Nafria et al., 2020b), with a follow-up Protocols paper published later (Nafria et al., 2020a). The two procedures differed in growth factor and cytokine concentrations used and in preparation of hPSCs prior to differentiation (see Table 5.2 for comparison of the two procedures). However, both procedures describe differentiation of hESCs grown on mouse embryonic feeder (MEF) layers. In contrast, RPAT hiPSCs are maintained on Matrigel in E8 media. The differences in routine hPSC culture in the published work and our own may have contributed to the initial poor differentiation observed; reinforced by our observation that problems with the differentiation process begin at an early stage. We received a batch of H9 hESC dual reporter cells which are cultured on MEFs and routinely differentiated using the reported protocol (Nafria et al., 2020b); kindly donated by Dr Monica Nafria. This permitted a direct comparison of spin EB differentiation of cells grown on feeders and RPAT hiPSCs cultured on Matrigel. Notably, the H9 dual reporter cell line carries a mCherry gene in the *SOX17* locus and a GFP gene in the *RUNX1* locus, which is driven by the RUNX1 distal promoter (denoted hereafter as S/R-H9). SOX17 signalling is required for

development of the arterial endothelium (Clarke et al., 2013; Jung et al., 2021), while RUNX1 has an essential role in EHT (Bertrand et al., 2010; Boisset et al., 2010; Lancrin et al., 2009). The RUNX1 distal promoter produces the RUNX1c isoform which is active in definitive haematopoietic progenitors (Sroczynska et al., 2009). As such, GFP expression marks RUNX1c-expressing cells and cells fated to the definitive haematopoietic program. Thus, differentiation progression and specification of the definitive haematopoietic program can be monitored *in vitro* using fluorescent imaging and flow cytometry.

Table 5.2: Comparison of spin EB differentiation procedures

	Ng *et al.* (2016)		**Nafria *et al.* (2020)**
Day -1	Passage hPSCs with EDTA 1:1 into new T25		
Day 0 (am)	E8 media change		Passage hPSCs with EDTA 1:1 into new T25
Day 0 (pm)	Passage hPSCs with Accutase, centrifuge to form spin EBs and initiate differentiation Seed in ratios; 1 x T25 into 3 x 96-well plates (600 EBs)		
	20 ng/mL **10 ng/mL** 25 ng/mL 25 ng/mL 10 ng/mL	BMP4 **AA** VEGF SCF FGF2 **CHIR**	20 ng/mL **7.5 ng/mL** 25 ng/mL 25 ng/mL 10 ng/mL **0.5 µM**
Day 2	3 – 4 uM 3 uM	SB CHIR	3.5 uM 3 uM
Day 4	MESODERM		
	20 ng/mL 50 ng/mL 50 ng/mL **30 ng/mL** 10 ng/mL	**BMP4** VEGF SCF **IGF2** FGF2	**5 ng/mL** 50 ng/mL 50 ng/mL **10 ng/mL** 10 ng/mL
Day 7/8	HAEMATOVASCULAR MESODERM		
Day 9	50 ng/mL **50 ng/mL** **50 ng/mL** 25 ng/mL **25 ng/mL** **25ng/mL** 20 ng/mL 10 ng/mL	**BMP4** VEGF **SCF** **IL3** IL6 **TPO** **FLT3** IGF2 FGF2	**20 ng/mL** 50 ng/mL **100 ng/mL** **25 ng/mL** 25 ng/mL **50 ng/mL** **100 ng/mL** 20 ng/mL 10 ng/mL
Day 14+		**SCF** **FLT3** **TPO** **IL3** IL6	**100 ng/mL** **100 ng/mL** **50 ng/mL** **25 ng/mL** 25 ng/mL
Day 18	HAEMATOPOIETIC PROGENITORS		

RPAT and S/R-H9 cells were differentiated according to the procedure in Nafria et al. (2020) and progression was monitored by flow cytometry for stage-specific markers. On d4 of differentiation, both cell lines yielded >85% NCAM+ cells (Figure 5.4A, C). As haematopoietic cells are specified from KDR+ mesoderm (Pearson et al., 2015), KDR expression was assessed. RPAT EBs possessed a greater proportion of KDR+ cells than S/R-H9 EBs on d4 (34% vs 10.9%, Figure 5.4A, C). Notably, all KDR+ cells co-expressed NCAM, demonstrating KDR+ cells arise in cells committed to the mesoderm lineage. The haemogenic mesoderm can be identified by co-expression of KDR and CD34 (Kennedy et al., 2012). CD34 was not detected in EBs from either cell line suggesting cells have not yet transitioned to this state (Figure 5.4C). In the early human embryo $KDR^{+}CD34^{-}$ expression marks mesodermal precursors with endothelial differentiation potential (Cortes et al., 1999). Therefore, d4 EBs likely represent mesodermal precursors to the HE.

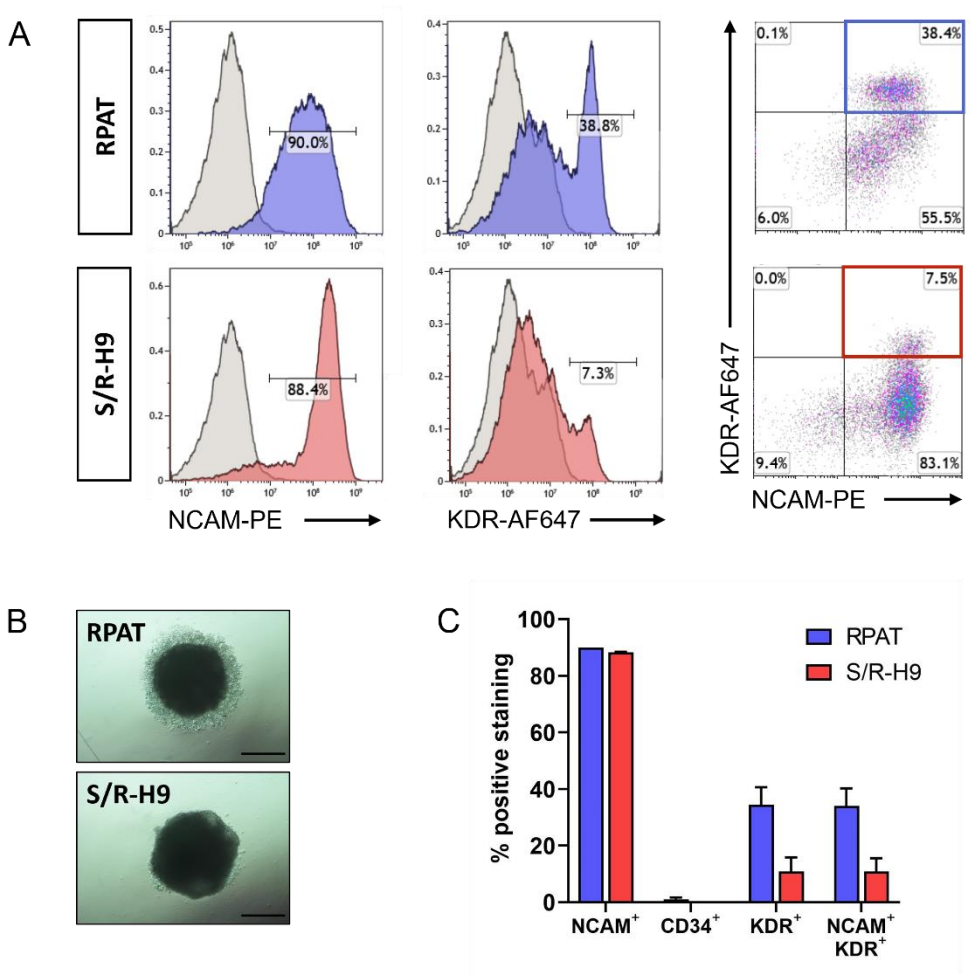


Figure 5.4: RPAT and S/R-H9 cell lines can produce NCAM+KDR+ mesoderm

(A) Flow cytometry analysis of NCAM and KDR expression in d4 EBs from RPAT and S/R-H9 cell lines. Representative histogram plots show NCAM/KDR expression in comparison with an isotype control (grey). Representative dot plots demonstrate NCAM/KDR double-positive population, highlighted by blue/red box. (B) Representative light microscope images of d4 EBs. Scale bar = 200 µm. (C) Quantitative summary of flow cytometry analysis for NCAM, KDR, CD34 and $NCAM^+KDR^+$ populations in d4 EBs produced from the two cell lines. Results are mean ± SD, n = 2 technical replicates.

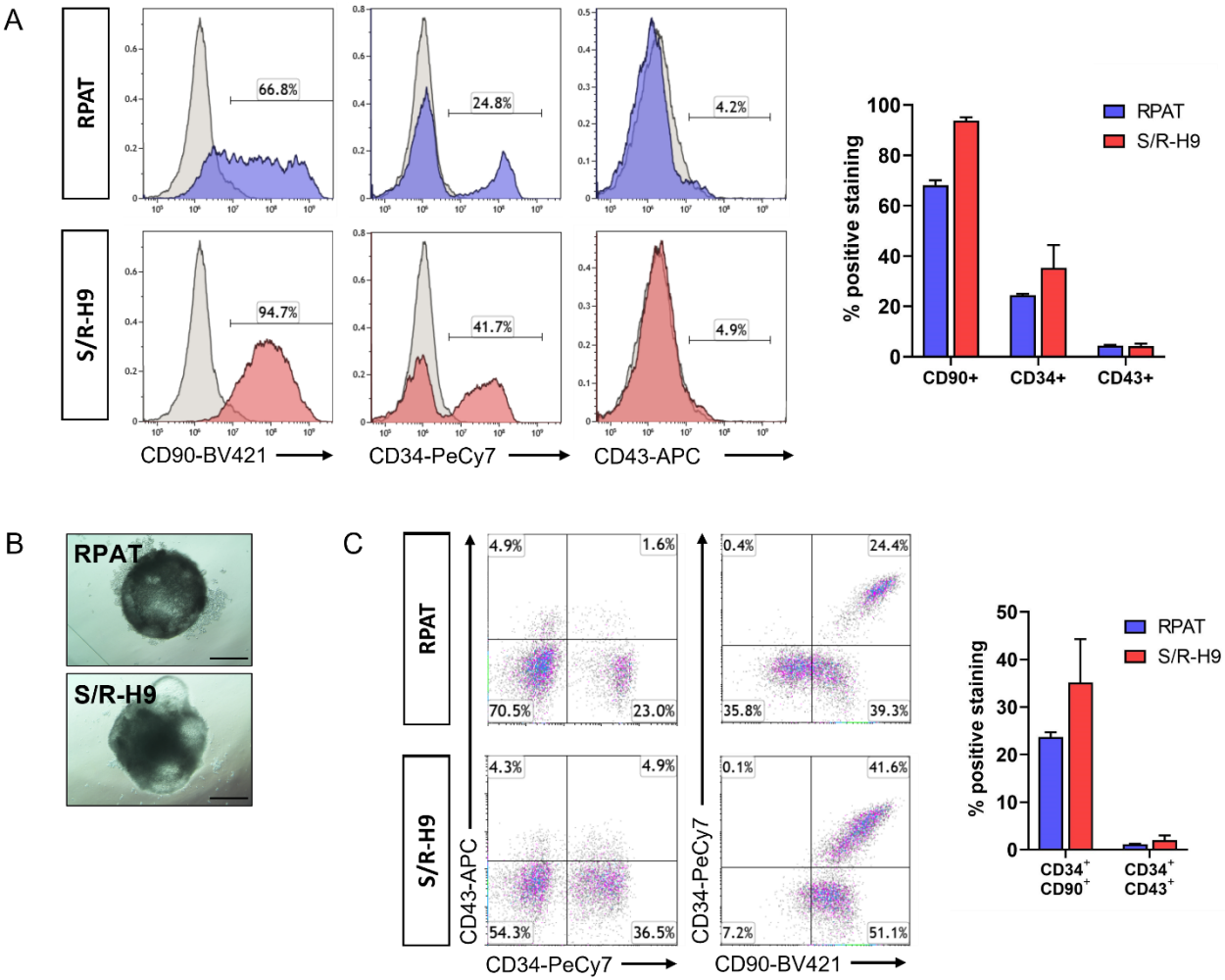


Figure 5.5: RPAT and S/R-H9 cells differ in CD90 and CD34 expression on day 7 of differentiation

(A) Flow cytometric analysis of CD90, CD43 and CD34 expression in d7 EBs from RPAT and S/R-H9 cell lines. Histogram plots show expression of each marker alongside their respective isotype control (grey). Quantitative summary of flow cytometry analysis shown in bar graph. (B) Representative light microscope images of d7 EBs. Scale bar = 200 µm. (C) Representative dot plots for CD34/CD90 and CD34/CD43 co-expression profiles. Quantitative summary of flow cytometry analysis shown in bar graph. Results are mean ± SD, n = 2 technical replicates.

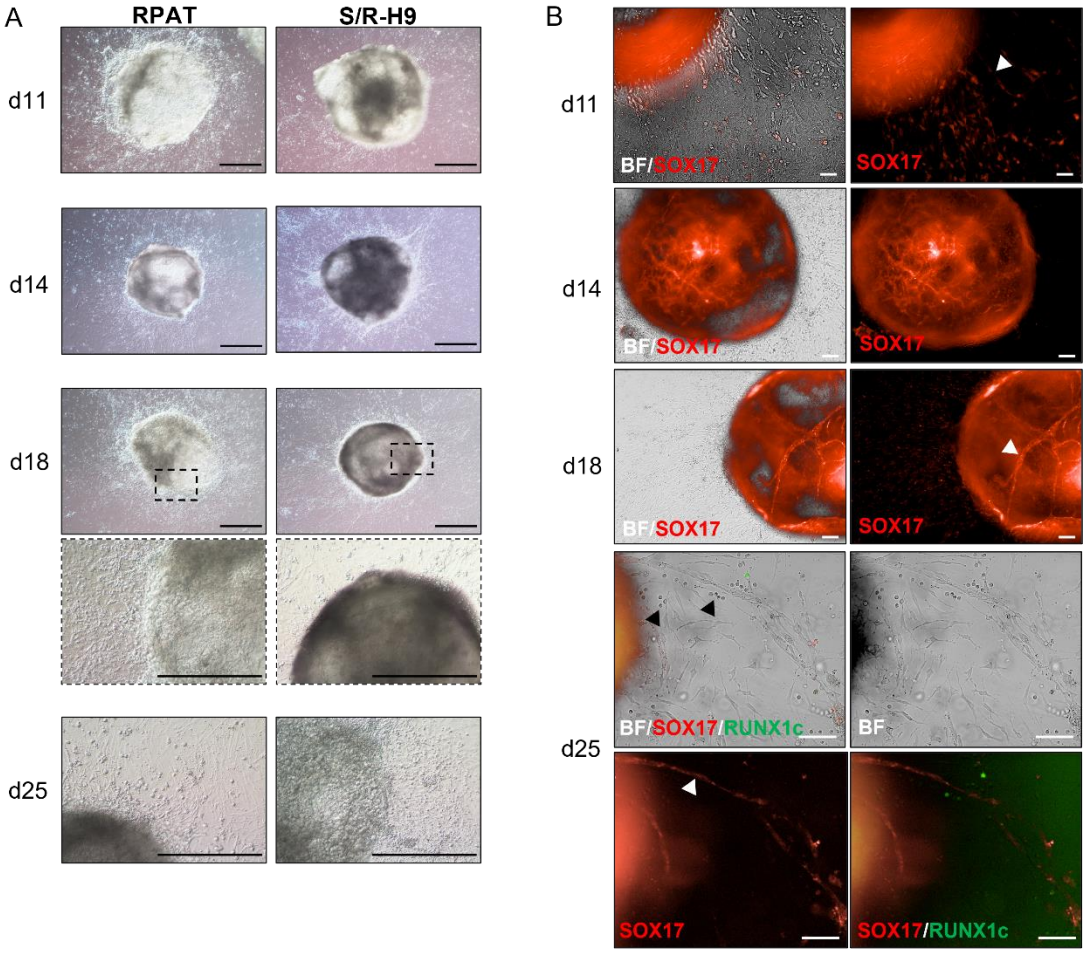


Figure 5.6: Differentiation of both RPAT and S/R-H9 cell lines failed to produce suspension cells

(A) Representative brightfield imaging of RPAT and S/R-H9 EBs from day 11 – 25 following transfer to Matrigel-coated plates. Scale bar = 500 µm. (B) Representative fluorescence imaging of S/R-H9 EBs in adherent culture. SOX17 marks endothelial cells (indicated by white arrow). RUNX1c marks haematopoietic progenitor cells (indicated by black arrow). Scale bar = 50 µm.

Differentiation was next assessed on d7 prior to EB transfer to adherent culture. EBs from both cell lines were assayed via flow cytometry for CD90, CD34 and CD43 expression. CD34 and CD90 are expressed in both endothelial and HSPC-like cells (Guibentif et al., 2017; Prashad et al., 2015). Upregulation of CD43 marks commitment to the haematopoietic lineage (Vodyanik et al., 2006) and can discriminate between the two cell types (Nafria et al., 2020a). In d7 EBs from both cell lines CD43 expression was minimal (<5%; Figure 5.5A). CD34 and CD90 were more highly expressed in EBs from S/R-H9s (35.3% CD34, 93.8% CD90) than RPATs (24.4% CD34, 68.2% CD90, Figure

5.5A). All CD34+ cells co-expressed CD90, while few CD34+ cells co-expressed CD43 (Figure 5.5C). As such, d7 EBs possess endothelial identity but have not yet been restricted to the haematopoietic lineage and likely represent haemato-vascular mesoderm.

After transfer to adherent culture, EBs were monitored for the appearance of floating round haematopoietic cells. Brightfield imaging of RPAT and S/R-H9 cultures from d11 – d25 revealed EBs attached to the underlying Matrigel and produced a stromal layer, albeit without generating suspension cells (Figure 5.6A). Fluorescence microscopy of S/R-H9 cultures initially identified a SOX17+ stromal layer on d11 which developed into SOX17-expressing vessels detected within the EB from d14 onwards (Figure 5.6B). Close to the EBs, several small clusters (3 – 4 cells) of round cobblestone-like cells characteristic of haematopoietic cells were detected. Some of these clusters contained GFP+ cells indicative of RUNX1c expression. However, these cells were very rare and of insufficient quantity to characterise further. Overall, the differentiation potential of RPAT and S/R-H9 cells was similar; both produced haemato-vascular-like cells yet failed to generate suspension cells. As such, the maintenance conditions of hPSCs prior to differentiation does not appear to influence the outcome.

5.3.5 Cross titration of BMP4 and Activin A identified conditions which produce CD34^{+}CD43/45^{+} haematopoietic progenitors

BMP4 and Activin/Nodal signalling pathways play key roles in orchestrating early differentiation of hPSCs. BMP4 and Activin/Nodal signalling gradients are dosage sensitive, with high Activin A stimulating endoderm differentiation and low or moderate BMP4 directing cells to either a cardiac or haematopoietic fate (Cerdan et al., 2012; Tan et al., 2013). Endogenous BMP4 and Activin/Nodal signalling can vary amongst different hiPSC and hESC lines (Kattman et al., 2010), which in turn influences haematopoietic differentiation potential (Ramos-Mejia et al., 2010). As such, optimisation of supplemental BMP4/Activin A concentration is important to ensure efficient differentiation to the lineage of interest. RPAT hiPSC were differentiated according to Ng et al. (2016) differentiation procedure and BMP4 and Activin A concentrations were cross titrated to identify the optimal concentration for RPAT hiPSC differentiation. Three different concentrations of BMP4 (5, 10 or 20 ng/mL) and Activin A (5, 7.5 or 10 ng/mL) were tested, and differentiation progression was monitored via flow cytometry on d4, d7 and d15 of differentiation.

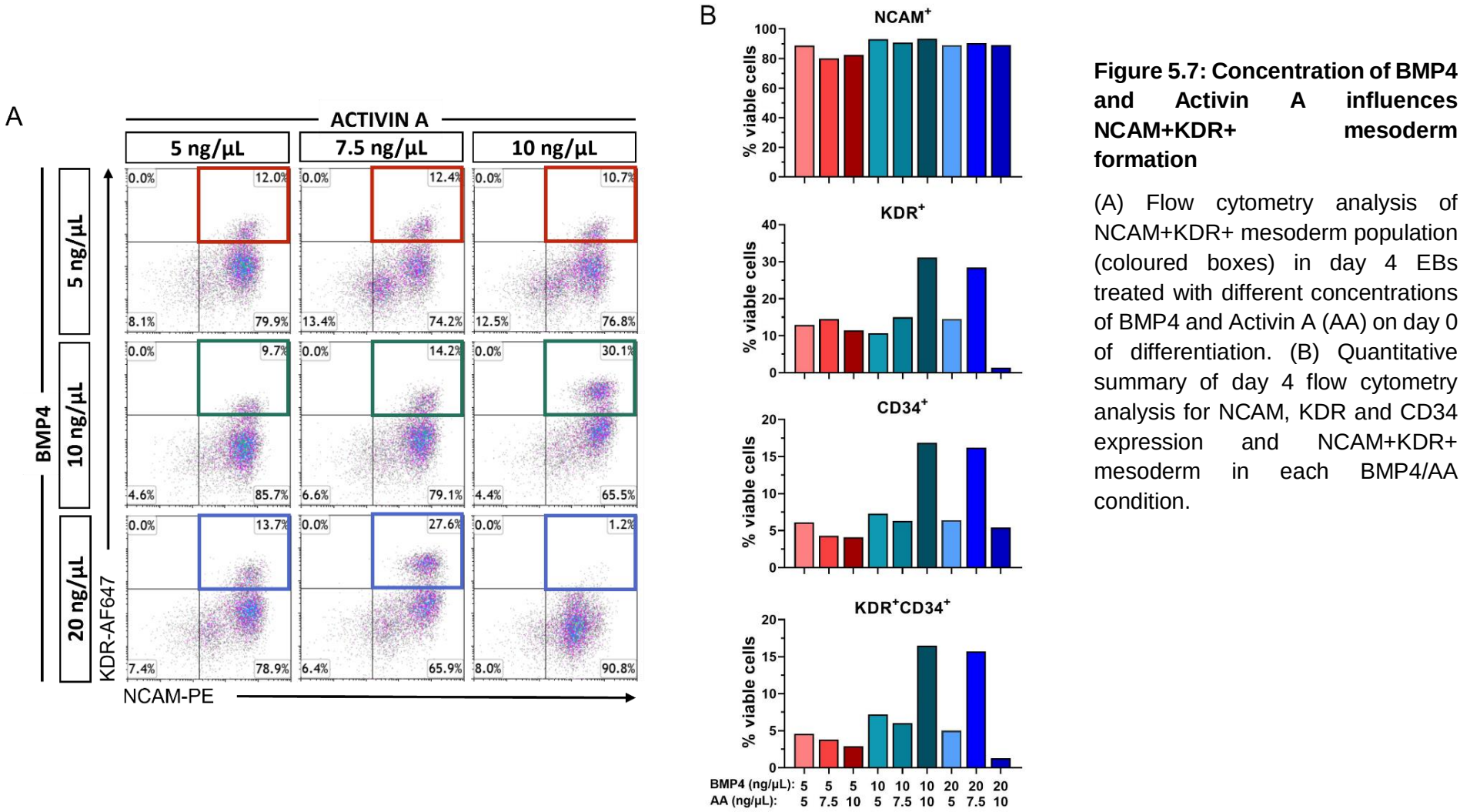
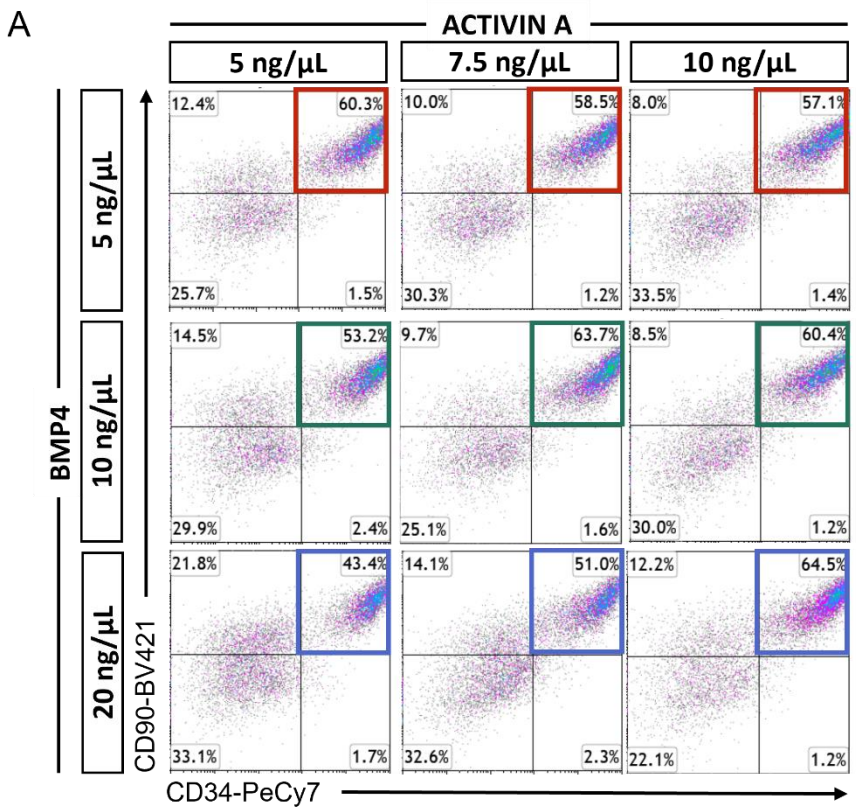


Figure 5.7: Concentration of BMP4 and Activin A influences NCAM+KDR+ mesoderm formation

(A) Flow cytometry analysis of NCAM+KDR+ mesoderm population (coloured boxes) in day 4 EBs treated with different concentrations of BMP4 and Activin A (AA) on day 0 of differentiation. (B) Quantitative summary of day 4 flow cytometry analysis for NCAM, KDR and CD34 expression and NCAM+KDR+ mesoderm in each BMP4/AA condition.



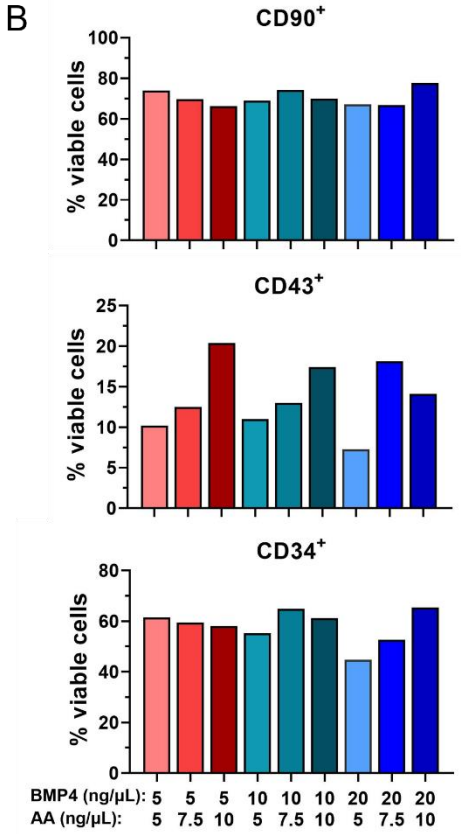


Figure 5.8: CD43 expression in d7 EBs varied with BMP4 and Activin A concentration

(A) Flow cytometry analysis of haemato-vascular marker ($CD34^{+}CD90^{+}$) expression in day 7 EBs treated with different concentrations of BMP4 and AA on d0 of differentiation. (B) Flow cytometry analysis of CD90, CD43 and CD34 expression in day 7 EBs from each BMP4/AA condition.

Regardless of the BMP4/Activin A (B/AA) concentration used, all d4 EBs were >80% NCAM+ and KDR-expressing cells were only identified in the NCAM+ population (Figure 5.7A, B). KDR+ expression varied, from 1.2% (20/10 ng/mL B/AA) to 30.1% (10/10 ng/mL B/AA). Similarly, CD34 expression varied amongst different conditions tested, ranging from 4.1% (20/10 ng/mL B/AA) to 16.9% (10/10 ng/mL B/AA). All CD34+ cells co-expressed KDR, a phenotype which is indicative of haemogenic mesoderm (Kennedy et al., 2012).

In d7 EBs, CD90 expression was relatively low in all BMP4/Activin A conditions, ranging from 66.8% to 77.7% (Figure 5.8B). Conversely, CD34 expression was high in all conditions, ranging from 44.8 – 65.4%. As observed previously (Figure 5.5C), all CD34+ cells co-expressed CD90 (Figure 5.8A), consistent with acquisition of a haemato-vascular phenotype. CD43 expression was the most variably expressed marker, ranging from 7.3% in 20/5 ng/mL B/AA to 20.4% in of 5/10 ng/mL B/AA (Figure 5.8B).

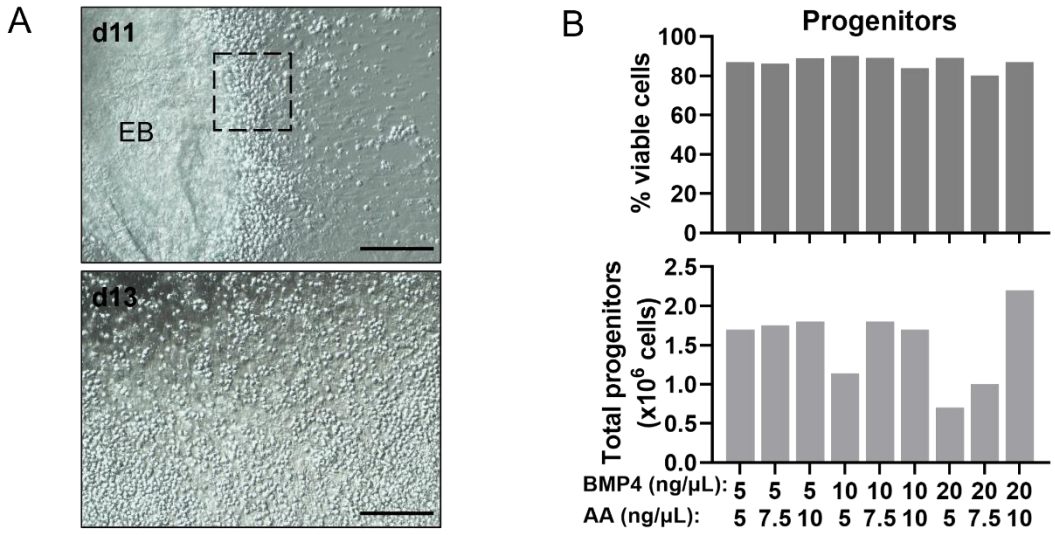


Figure 5.9: All BMP4/Activin A conditions produce suspension cells

(A) Representative light microscopy images of RPAT hiPSCs on d11 and d13 of spin EB differentiation. On d11, the EB is labeled and a region with bright round progenitor cells differentiating from the EB is shown for reference by a black box. Scale bar = 200 µm. (B) Viability and total number of suspension cells harvested on day 15 from one well of a 6-well plate. Viability was determined by Zombie NIR staining.

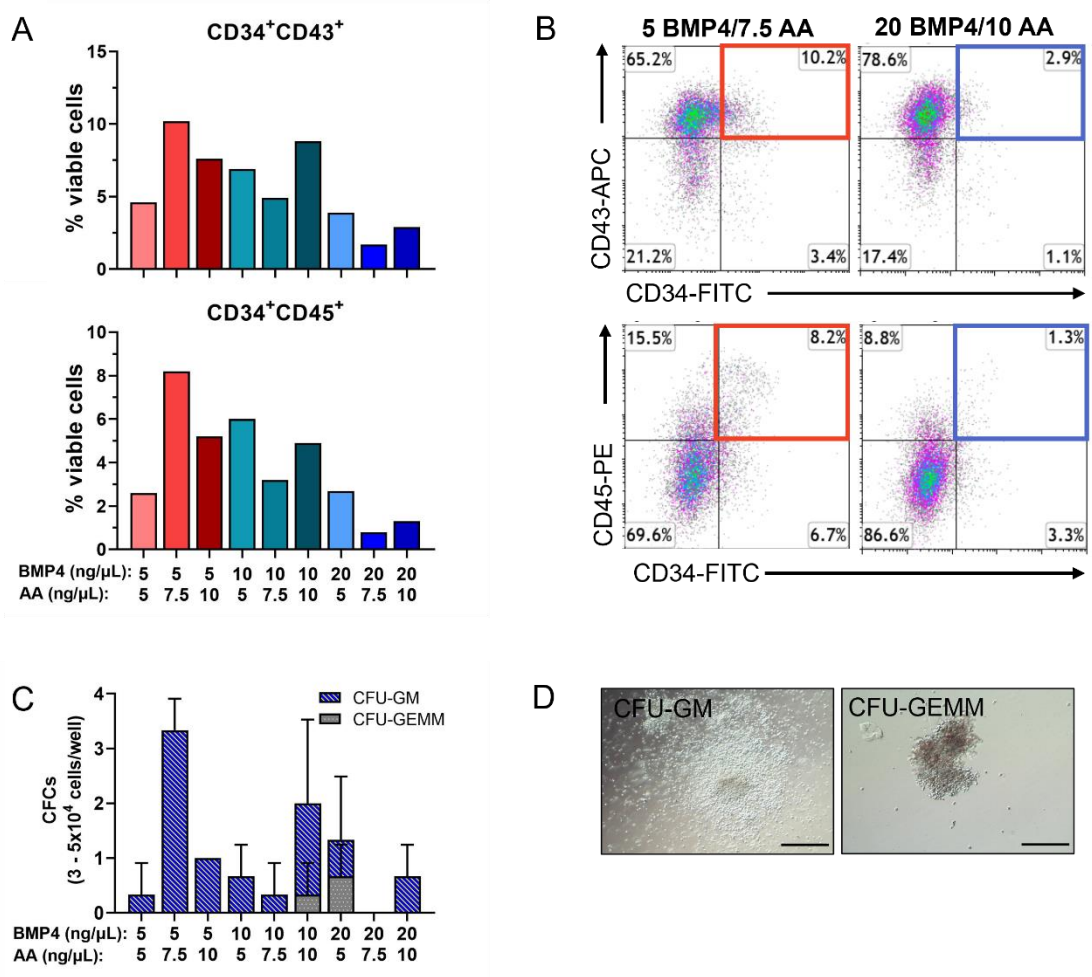


Figure 5.10: Flow cytometry analysis and CFU assay demonstrate 5 ng/mL BMP4/7.5 ng/mL Activin A yields cells with the greatest haematopoietic potential

(A) Flow cytometry analysis of haematopoietic progenitor cells ($CD34^+CD43^+/CD45^+$) harvested on day 15 from each BMP4/AA condition. (B) Representative dot plots demonstrating gating for $CD34^+CD43^+/CD45^+$ populations. (C) Haematopoietic colony-forming potential of d15 suspension cells collected from each BMP4/AA condition. Results mean ± SD, n = 3 technical replicates. (D) Representative images of CFU-GM and CFU-GEMM colonies formed following 14 days culture in MethoCult™ Optimum H4304 (STEMCELL Technologies). Scale bar = 50 µm.

Following transfer to adherent culture, round cells were observed emanating from the adherent EB on d11 (Figure 5.9A). Cells continued to emerge and began to detach from the underlying stromal layer and enter into suspension on d13. By d15, large quantities of cells in suspension were detected, which were harvested and assayed for haematopoietic marker expression using flow cytometry and haematopoietic potential via the CFU assay. In total, 0.7 – 2.2 x 10^6 suspension cells were collected (from 1 well

of a 6-well plate), with >80% of cells viable as assessed by staining for Zombie NIR (Figure 5.9B). Expression of $CD34^+CD43/CD45^+$ haematopoietic progenitors varied amongst the different BMP4/Activin A conditions (Figure 5.10A, B), with the highest expression observed when EBs were treated with 5/7.5 ng/mL B/AA (10.2% $CD34^+CD43^+$; 8.2% $CD34^+CD45^+$) and the lowest with 20/7.5 ng/mL B/AA (1.7% $CD34^+CD43^+$; 0.8% $CD34^+CD45^+$). Few colonies were formed in the CFU assay, with GM colonies predominant (Figure 5.10C, D). However, colony formation followed a similar trend to the $CD34^+CD43/CD45^+$ expression profile; whereby, suspension cells which gave rise to the highest number of colonies in the CFU assay also exhibited the highest marker expression. The combination of low BMP4 (5 ng/mL) and median Activin A (7.5 ng/mL) yielded the highest number of haematopoietic progenitors and showed the greatest colony-forming capabilities in the CFU assay. Thus, this condition was chosen for future experiments.

5.3.6 SB/CHIR treatment is essential to specify haematopoietic progenitors

The addition of SB and CHIR on d2 of differentiation is purported to suppress primitive haematopoiesis, resulting in stimulation of AGM-like definitive haematopoiesis (Ng et al., 2016). To determine the role of SB/CHIR in directing haematopoietic fate, RPAT hiPSCs were differentiated with or without addition of SB/CHIR on d2 of differentiation (+/- SB/CHIR). Light microscopy imaging on d4 showed +SB/CHIR-treated EBs were larger than -SB/CHIR EBs (Figure 5.11B). EB size was quantified using ImageJ (Figure 5.11A), revealing EBs treated with SB/CHIR increased in size by approx. 1.2-fold from d2 to d4, while untreated EBs remained unchanged (Figure 5.11C). Flow cytometry of d4 EBs showed NCAM expression was similar in both treatments (82.5% vs 71.6%; Figure 5.12A, B). However, KDR+ cells were not detected in untreated EBs (Figure 5.12A, B), suggesting SB/CHIR treatment may be necessary to confer haematopoietic potential.

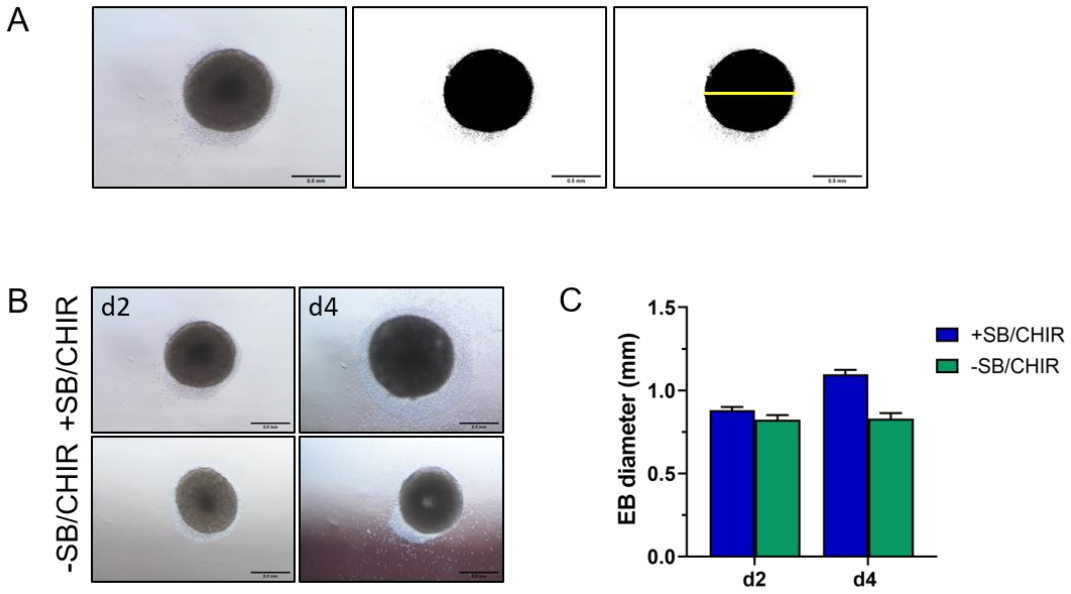


Figure 5.11: SB/CHIR treatment on day 2 of differentiation enhances EB size

(A) Diameter of EBs measured using ImageJ. Images taken at 4X on a light microscope were converted to binary images and the diameter of the EB was measured using the measurement tool (yellow line). Six measurements were taken per EB and averaged. (B) Representative images of EBs on day 2 and day 4 of differentiation. Scale bar = 500 µm. (C) EB diameter measured on day 2 and day 4 using ImageJ. EBs were treated +/-SB/CHIR on day 2 of differentiation. Results are from 2 independent biological replicates with 3 technical replicates performed for each. Mean ± SD is calculated based on technical replicates.

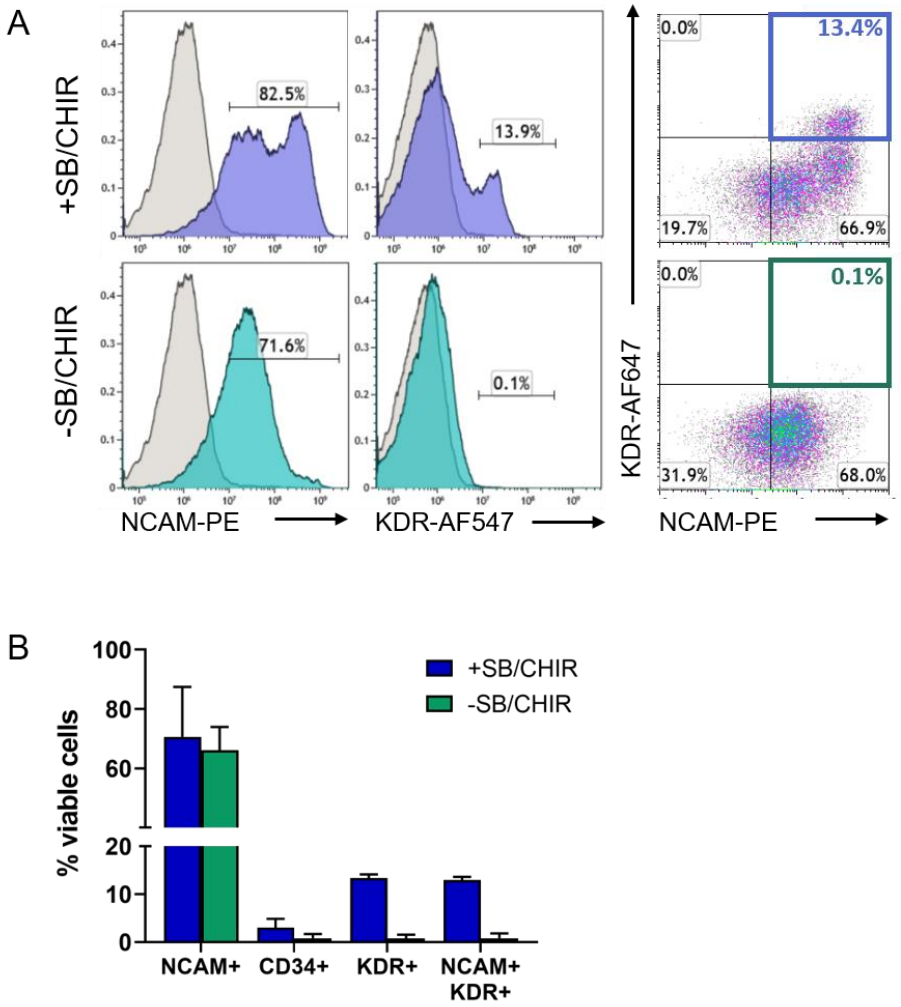


Figure 5.12: Addition of SB/CHIR on day 2 of differentiation is essential for specification of KDR+ mesoderm

(A) Flow cytometry analysis of day 4 EBs treated +/-SB/CHIR on day 2 of differentiation. Histograms show NCAM and KDR expression with isotype control overlaid in grey. Dot plots show $NCAM^+KDR^+$ mesoderm (coloured box). (B) Quantitative summary of day 4 flow cytometry analysis for NCAM, CD34 and KDR expression and $NCAM^+KDR^+$. Results are mean ± SD, n = 2 biological replicates.

CD90, CD34 and CD43 expression was greater in d7 SB/CHIR-treated EBs compared to untreated (Figure 5.13B). Haemato-vascular mesoderm can be identified by co-expression of CD90 and CD34 and lack of CD43 (Ng et al., 2016) and can also broadly be recognised as $CD34^+CD43^-$ (Lange et al., 2021). These two populations were assessed in d7 EBs. The percentage of $CD34^+CD90^+CD43^-$ and $CD34^+CD43^-$ populations tended to be higher in SB/CHIR-treated cells than untreated (Figure 5.13C, D). SB/CHIR-treated cultures also possessed a higher proportion of $CD34^+CD43^+$ cells

compared to untreated (Figure 5.13D). To assess if the haematopoietic program had been initiated, qPCR was performed on d4 and d7 EBs for key haematopoietic regulators *RUNX1c* and *GATA2*. Expression of the RUNX1c isoform in hPSC-based haematopoietic differentiation tracks haematopoietic commitment and is first upregulated in haemato-endothelial progenitors (Navarro-Montero et al., 2017). GATA2 possesses similar functions but is upregulated earlier in hPSC differentiations; being first expressed in mesoderm, with its expression progressively increasing as haematopoietic differentiation proceeds (Castano et al., 2019). Consistent with this expression profile, in SB/CHIR-treated cultures *RUNX1c* was upregulated in d7 EBs, while *GATA2* was initially expressed in d4 EBs and upregulated further by d7 (Figure 5.13E). In untreated cultures, *RUNX1c* was not expressed, while *GATA2* was expressed in d7 EBs, although at a lower level than in SB/CHIR treated cultures (Figure 5.13E). Identification of the $CD34^{+}CD90^{+}CD43^{-}$ population by flow cytometry and upregulation of *RUNX1c* and *GATA2* by qPCR suggests SB/CHIR-treated d7 EBs represent a haemato-vascular mesoderm. However, the high expression of CD43 also suggests some cells have undergone haematopoietic commitment.

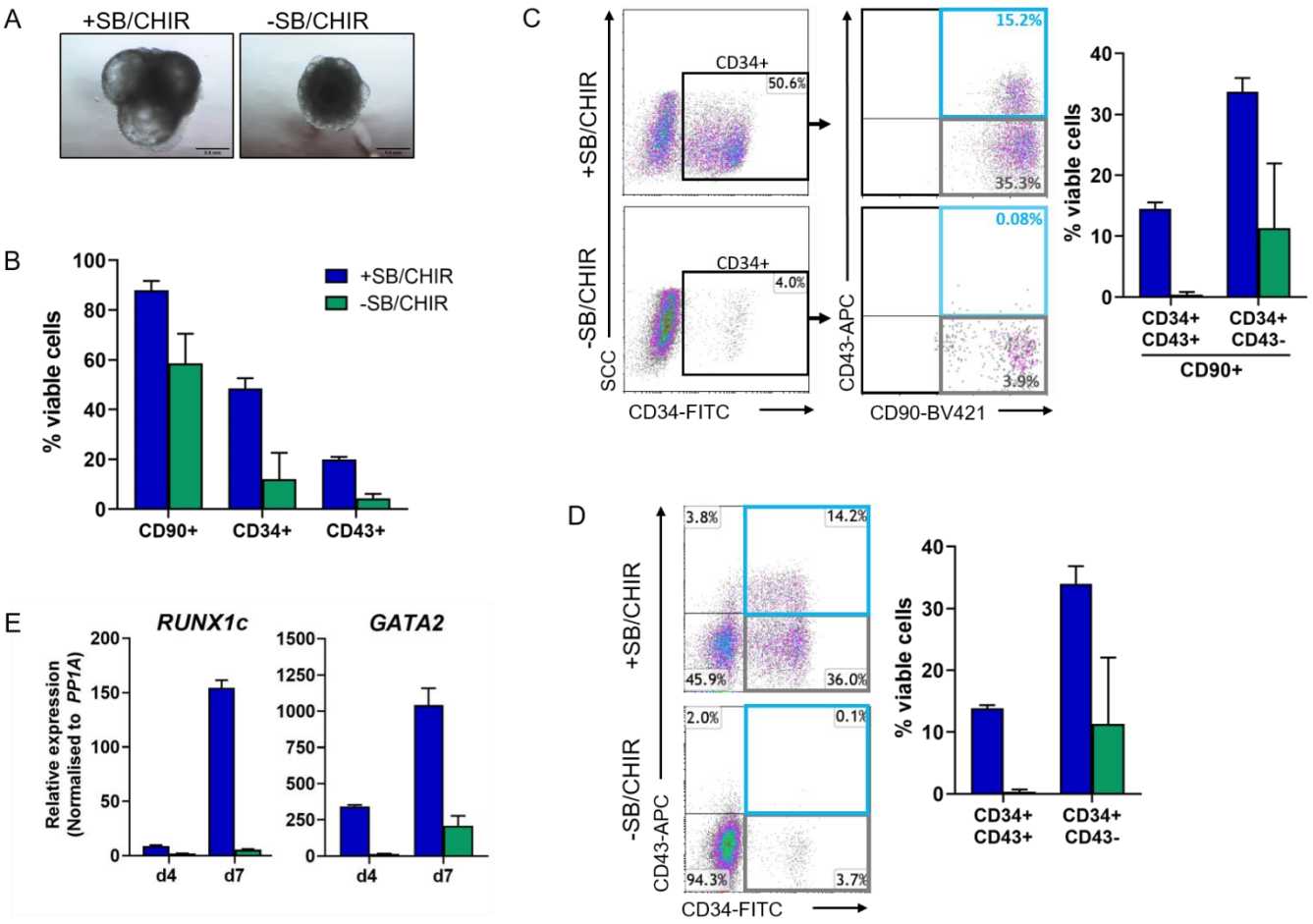


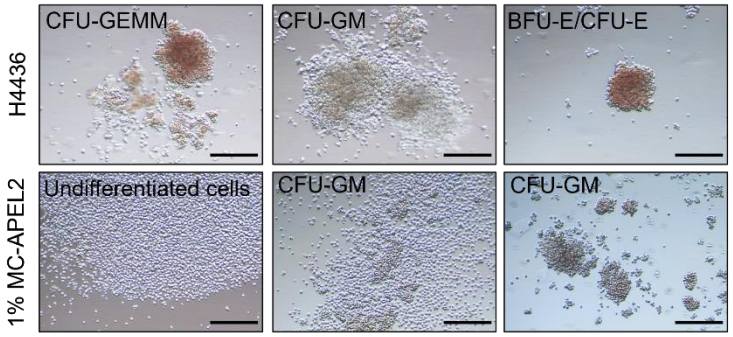
Figure 5.13: Day 2 SB/CHIR treatment enhances the formation of haemato-vascular mesoderm

(A Representative images of d7 EBs. Scale bar = 500 µm. (B) Flow cytometry analysis of CD90, CD34 and CD43 expression in d7 EBs. (C) Gating strategy for identification of $CD34^{+}CD90^{+}CD43^{+/-}$ population in d7 EBs, with quantitative summary shown in bar graph. (D) Representative dot plots showing $CD34^{+}CD43^{+/-}$ populations in d7 EBs, with quantitative summary shown in bar graph. Mean ± SD, n = 2 biological replicates. (E) QPCR for haematopoietic genes *RUNX1c* and *GATA2* in d4 and d7 EBs treated +/- SB/CHIR. Expression shown relative to housekeeping gene *PP1A* and normalised to undifferentiated hiPSCs. Results are mean ± SD, n = 3 technical replicates.

A

d14 suspension cells				
Methylcellulose (MC):	MethoCult™ SF H4636 (STEMCELL Technologies)		1% MC-APEL2 + EPO, GM-CSF, VEGF, IL3, IL6, TPO, FLT3	
Plate type:	24-well	12-well	24-well	12-well
Seeding density:	$3x10^3$ cells/well $5x10^3$ cells/well	$1x10^4$ cells/well $3x10^4$ cells/well	$3x10^3$ cells/well $5x10^3$ cells/well	$1x10^4$ cells/well $3x10^4$ cells/well
Outcome:	CFU-GEMM, CFU-GM and CFU-E/BFU-E colonies produced at both densities	Overconfluent with colonies merging (not counted)	Few colonies evident (CFU-GMs), mostly undifferentiated single cell aggregates (not counted)	No colonies evident, undifferentiated single cell aggregates at centre of well (not counted)

B



C

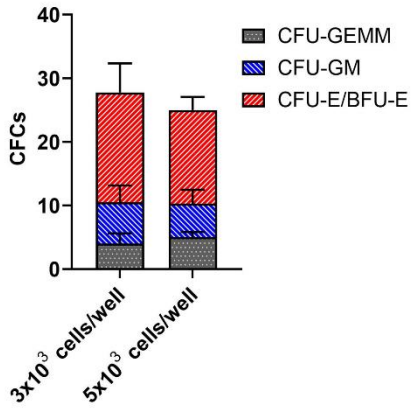


Figure 5.14: Optimisation of CFU assay conditions

(A) Table describing the outcome and conditions during test of two different methylcellulose (MC) media – H3636 (STEMCELL Technologies) and 1% MC prepared in APEL2 media supplemented with growth factors and cytokines (1% MC-APEL2). (B) Representative images of CFU-GEMM, GFU-GM and CFU-E/BFU-E colonies generated by d14 suspension cells cultured in H4636 in 24-well plates (upper panel). Representative images of undifferentiated single cell aggregates and CFU-GMs formed from d14 cells cultured in 1% MC-APEL2 in 24-well plates (lower panel). Scale bar = 200 µm. (C) Quantification of CFCs formed following culture of d14 suspension cells in H4436 in 24-well plates at a density of $3x10^3$ or $5x10^3$ cells/well. Results are mean ± SD, n = 4 technical replicates.

By d14, round suspension cells had been generated by EBs treated with SB/CHIR and were harvested for flow cytometry and optimisation of the CFU assay. Previous methylcellulose (MC) formulations used yielded poor results; thus, two different MC media were compared to determine which could most reliably produce CFCs. MethoCult™ SF H4636 produced by STEMCELL Technologies has been utilised by several groups to assess CFU potential (Cao et al., 2019; Lundin et al., 2020). H4636 is a defined, serum-free MC media designed to support the differentiation of hPSC-derived haematopoietic progenitors, however, growth factors and cytokines used in H4636 are at proprietary concentrations. Alternatively, Ng et al. (2016) employed their own methylcellulose formulation, whereby 1% MC was prepared in APEL media supplemented with EPO, GM-CSF, VEGF, IL3, IL6, TPO and FLT3 (1% MC-APEL). Haematopoietic differentiation potential of d14 suspension cells was assessed in both MC compilations, with different plate types (24-well or 12-well low attachment plates), and seeding densities compared (Figure 5.14A). After 14 days in culture, colonies were scored. Suspension cells seeded in 1% MC-APEL2 mainly formed undifferentiated single cell aggregates at the centre of the well (Figure 5.14B). When seeded in 24-well plates, suspension cells were able to form myeloid colonies, although these were quite rare. The large density of floating single cells hampered colony identification in 24-well plates and colonies were not counted (Figure 5.14B). In contrast, d14 suspension cells seeded in H4636 in 24-well plates had robust colony-forming potential, differentiating to both myeloid and erythroid sub-types (Figure 5.14B). The two densities tested yielded a similar frequency of colonies (Figure 5.14C). Cells seeded in 12-well plates in H4636 also produced robust differentiation, however, wells were overconfluent at both seeding densities tested and colonies often merged making accurate determination of colony number not possible. Thus, future CFU assays were performed using H4636 in 24-well plates with cells seeded at a density of $3x10^3$ cells/well.

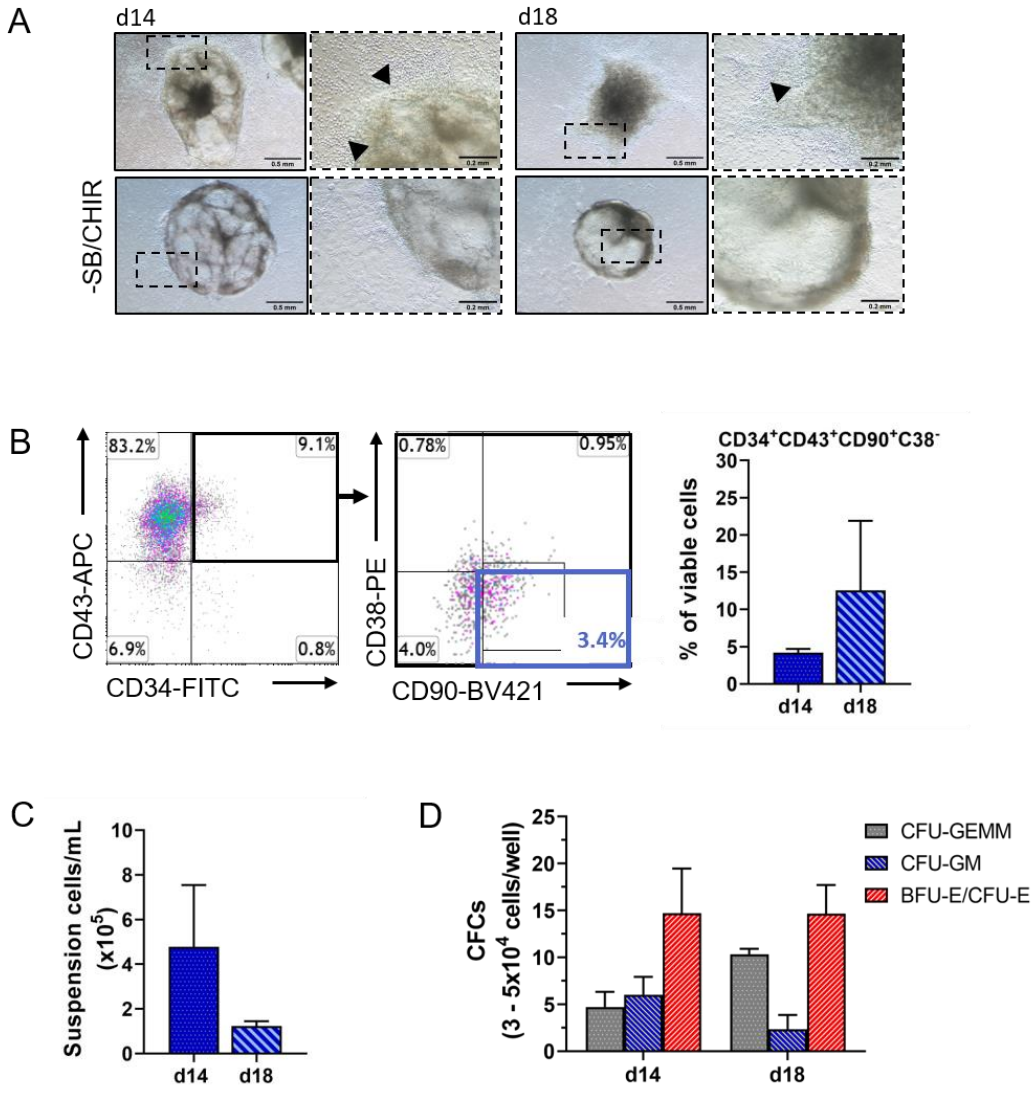


Figure 5.15: SB/CHIR-treated cultures generate two populations of CD34+CD43+CD90+CD38- haematopoietic progenitors

(A) Representative images of EBs taken at 4X (scale bar = 500 µm) and 10X (scale bar = 200 µm) on d14 and d18. The black arrow indicates haematopoietic progenitors emanating from EB. (B) Flow cytometric dot plots demonstrating gating strategy for identification of $CD34^{+}CD43^{+}CD90^{+}CD38^{-}$ HSPC-like cells. Bar charts show percentage of HSPC-like cells in SB/CHIR-treated cultures on d14 and d18. (C) The number of suspension cells/mL collected from +SB/CHIR cultures on d14 and d18. Results are mean ± SD, n = 2 biological replicates. (C) Haematopoietic colony-forming ability of suspension cells harvested from SB/CHIR treated cultures on d14 and d18 following culture in MethoCult™ SF H4636 (STEMCELL Technologies) for 14 days. Results are from 2 independent biological replicates with 3 technical replicates performed for each. Mean ± SD is calculated based on technical replicates.

Suspension cells were produced by d14 in SB/CHIR-treated cultures and not from untreated cultures (Figure 5.15A) Continued addition of differentiation media led to further generation of round suspension cells in SB/CHIR-treated cultures, which were harvested on d18 (Figure 5.15A). Continued media supplementation did not yield suspension cells beyond d18. Suspension cells harvested on d14 and d18 from +SB/CHIR cultures were assessed using the CFU assay and flow cytometry. Under the optimised CFU assay conditions, suspensions cells from d14 and d18 cultures differentiated to erythroid, myeloid and mixed colonies. Both d14 and d18 suspension cells showed similar erythroid and myeloid differentiation propensity, however, a greater proportion of immature CFU-GEMM colonies were generated by d18 cells than d14 (Figure 5.15D).

In previous experiments suspension cells had been assessed for $CD34^+CD38^-$ or $CD34^+CD43^+/CD45^+$ expression. These phenotypes mark haematopoietic progenitors but do not distinguish if progenitors resemble yolk sac-derived primitive progenitors, EMPs or AGM-derived HSPCs. To better characterise the suspension cells produced, cells were assessed for a CD34+ population which co-expressed CD43 and CD90 and lacked CD38 expression; a profile which demarcates haematopoietic cells with a HSPC-like phenotype (Dou et al., 2016; Prashad et al., 2015). A $CD34^+CD43^+CD90^+CD38^-$ population was identified in both d14 and d18 cells (Figure 5.15B). A greater number of suspension cells were collected from d14 cultures than d18 (0.6 vs 0.2 x 10^5 cells/mL, Figure 5.15C). However, a greater proportion of d18 suspension cells possessed the HSPC-like phenotype $CD34^+CD43^+CD90^+CD38^-$ than d14 (12.6% vs 4.2%, Figure 5.15B), suggesting the extended culture enhances HSPC-like cell production. Overall, only cultures treated with SB/CHIR were able to produce haematopoietic progenitors demonstrating this step is essential in haematopoietic differentiation.

5.3.7 Spin EB differentiation yields two functionally different haematopoietic progenitor populations

SB/CHIR treatment is essential to generate haematopoietic progenitors. However, the precise timing of Activin/Nodal inhibition and Wnt stimulation can impact the phenotype and functionality of haematopoietic progenitors produced (Kennedy et al., 2012; Sturgeon et al., 2014; Tan et al., 2013). In the previous spin EB differentiation experiments, SB/CHIR was administered between 44h and 48h after differentiation initiation. Previous experiments have also generated a relatively high proportion of CD43-expressing cells (> 5%) in d7 EBs, suggesting the primitive program has not been adequately supressed (Nafria et al., 2020a). SB/CHIR was administered 38h or 40h post-

differentiation initiation to assess if earlier modulation of Activin and Wnt signalling would impact haemato-vascular populations and haematopoietic progenitor formation.

EBs treated with SB/CHIR at 38h or 40h increased in size from d2 to d4 of differentiation but no difference was observed between the two timepoints (Figure 5.16A). Flow cytometry for NCAM and KDR expression in d4 EBs showed >85% of cells were NCAM+ when SB/CHIR was added at either timepoint, of which approx. 11% co-expressed KDR (Figure 5.16B). In d7 EBs, the CD90+, CD34+ and CD43+ populations were comparable between the two timepoints (Figure 5.16C). There was a trend for a higher proportion of $CD34^{+}CD90^{+}CD43^{-}$ haemato-vascular cells (43.5% 38h vs 32.8% 40h, n.s.) and lower haematopoietic committed $CD34^{+}CD90^{+}CD43^{+}$ cells (5.3% 38h vs 6.7% 40h, n,s.) when EBs were treated at 38h compared to 40h, although this was not statistically significant (Figure 5.16D).

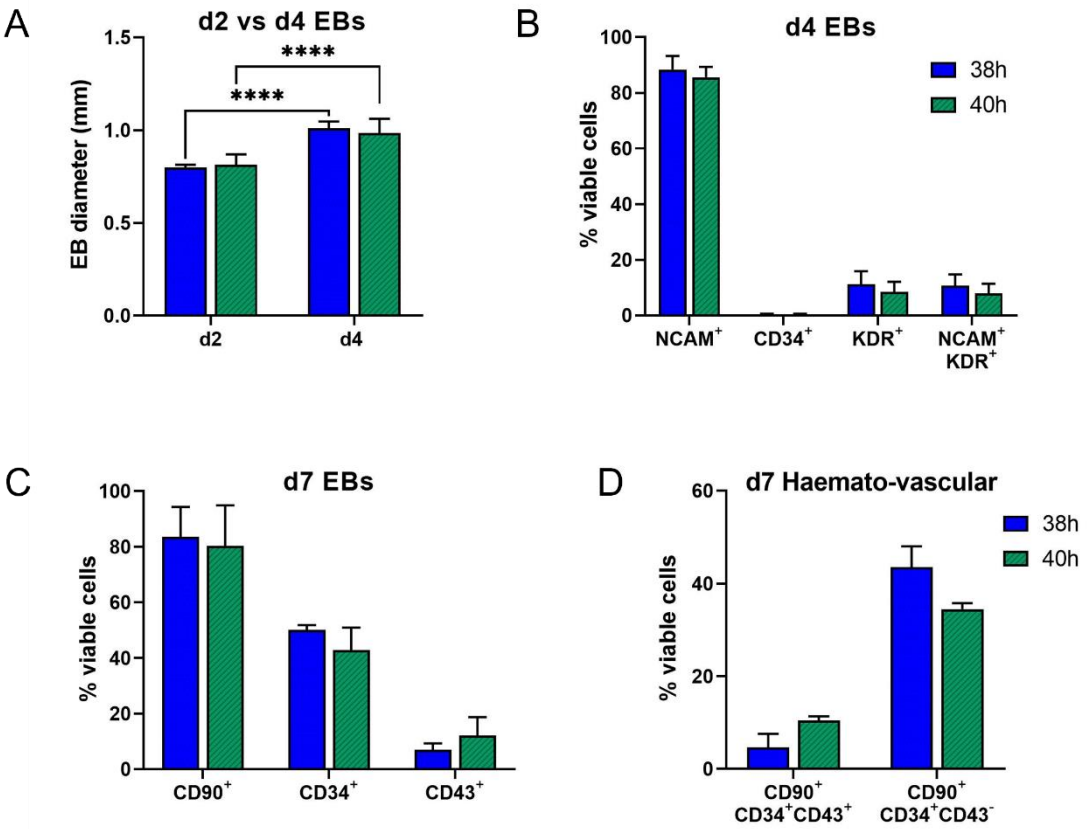


Figure 5.16: Addition of SB/CHIR 38h or 40h post-seeding does not affect the efficiency of mesoderm or haemato-vascular cell formation

(A) EB diameter on d2 and d4 of EBs treated with SB/CHIR 38h or 40h post-seeding. Results are from 3 biological replicates with 3 technical replicates performed for each. Mean ± SD is calculated based on technical replicates; $p < 0.0001$, 2-way ANOVA, Tukey's multiple comparisons test. (B) Flow cytometry analysis of NCAM, CD34 and KDR expression and NCAM+KDR+ mesoderm in d4 EBs. Results are mean ± SD, n = 3 biological replicates. (C) Flow cytometry analysis of CD90, CD34 and CD43 expression in d7 EBs. Mean ± SD, n = 3 biological replicates. (D) Flow cytometry analysis of haemato-vascular ($CD34^+CD90^+CD43^-$) and haematopoietic committed ($CD34^+CD90^+CD43^+$) populations in d7 EBs treated with SB/CHIR at 38h and 40h post-seeding. Mean ± SD, n = 3 biological replicates.

Gene expression patterns in d4 and d7 EBs were also examined by qPCR. Changes in gene expression were interpreted relative to undifferentiated RPAT hiPSCs. As expected, mesendoderm marker *MIXL1* was upregulated in d4 EBs from both conditions and downregulated by d7 (Figure 5.17A). The expression of five HOXA genes was also examined, namely *HOXA3, HOXA5, HOXA7, HOXA9* and *HOXA10.* HOXA expression in d4 EBs was minimal, although all HOXA genes examined were upregulated by d7

(Figure 5.17B). *HOXA5* and *HOXA10* were significantly higher in cultures treated with SB/CHIR at 38h compared to 40h. The remaining HOXA genes showed no difference in expression between the two conditions. CDX genes regulate HOX expression (Deschamps and van Nes, 2005). CDX1 and CDX2 regulate mesendoderm development, while CDX4 is required to specify $KDR^{+}CD235a^{-}$ definitive mesoderm (Creamer et al., 2017). Consistent with their role in mesendoderm, *CDX1* and *CDX2* were upregulated in d4 EBs, with *CDX1* significantly downregulated by d7, while *CDX2* expression remained unchanged (Figure 5.17C). *CDX4* expression level was equivalent to undifferentiated RPAT hiPSCs in d4 EBs treated at 38h but was significantly lower in d7 EBs and d4 EBs treated at 40h (Figure 5.17C).

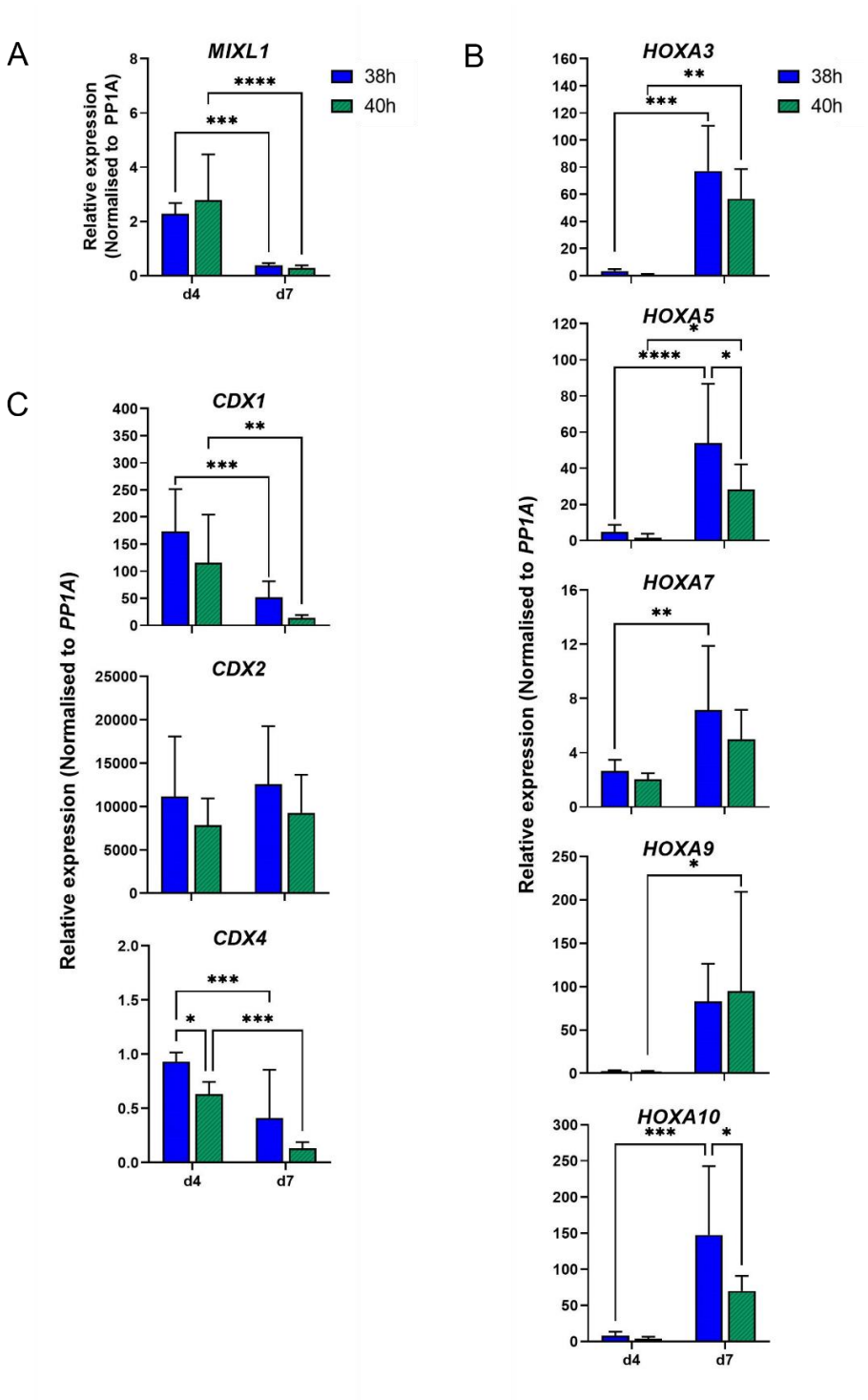


Figure 5.17: HOXA signature upregulated by d7 of differentiation

QPCR analysis of mesendoderm marker *MIXL1* (A), HOXA genes (B) and CDX genes (C) in d4 and d7 EBs treated with SB/CHIR at 38h or 40h post-seeding. Expression calculated relative to undifferentiated RPAT hiPSCs. Results are mean ± SD, n = 3 biological replicates; *p < 0.05, **p < 0.01, ***p < 0.001 and ****p < 0.0001, 2-way ANOVA, Tukey's multiple comparisons test.

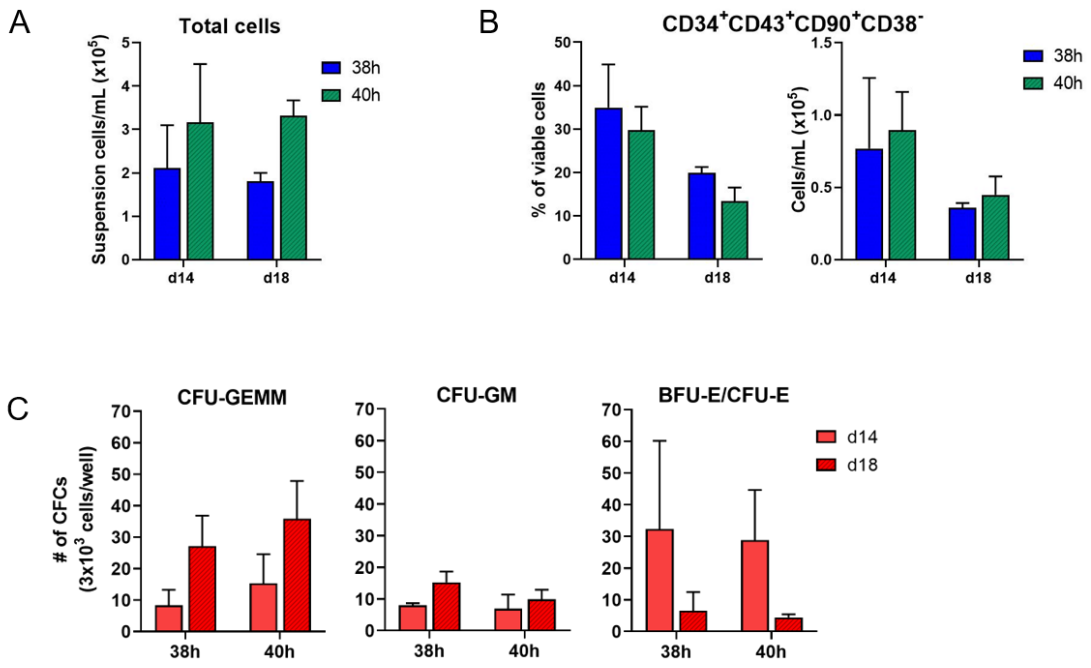


Figure 5.18: Haematopoietic progenitors with different colony forming abilities are produced with addition of SB/CHIR at 38h or 40h of differentiation

(A) The number of suspension cells/mL harvested on d14 and d18. Results are mean ± SD, n = 3 biological replicates. (B) Flow cytometry analysis of percentage and absolute number (cells/mL) of HSPCs harvested on d14 and d18 in cultures treated with SB/CHIR 38h or 40h post-seeding. Mean ± SD, n = 3 biological replicates. (C) Haematopoietic colony-forming ability of suspension cells harvested on d14 and d18. Colonies were counted after 14 days in MethoCult™ SF H4636 (STEMCELL Technologies). Results are mean ± SD from 2 - 3 biological replicates with 3 technical replicates performed for each.

Suspension cells were harvested from both treatment groups on d14 and d18. No significant difference in the number of suspension cells was observed at each timepoint, however, a trend for higher numbers of suspension cells in the 40h treatment group compared to 38h was observed (Figure 5.18A). In contrast to the previous experiment, a greater fraction of d14 suspension cells possessed the HSPC-like phenotype $CD34^+CD43^+CD90^+CD38^-$ than d18 cells both when assessed by flow cytometry and when the absolute number of $CD34^+CD43^+CD90^+CD38^-$ cells/mL was calculated (Figure 5.18B). In the CFU assay, BFU-E/CFU-E colonies were the main colony subtype generated by d14 cells from both treatments. CFU-GM and CFU-GEMM were produced at a lower frequency and no difference in differentiation potential was observed between 38h or 40h treatment groups (Figure 5.18C). In contrast, CFU-GEMM was the main

colony subtype produced by all d18 cells. Day 18 cells from both groups produced fewer BFU-E/CFU-E colonies and a greater proportion of CFU-GEMMs than d14 cells. However, the differences were not found to be statistically significant. The CFU assay suggests a functional difference exists between d14 and d18 suspension cells, which is not impacted by the time of SB/CHIR treatment. Day 14 cells appeared biased toward erythroid differentiation, while d18 cells yield predominately CFU-GEMMs, suggesting d18 suspension cells are of a more immature phenotype.

Haematopoietic specification was also monitored throughout differentiation by qPCR for haematopoietic transcription factors: *HOXA9*, *GATA2*, *RUNX1c*, *SPI1* and *TAL1* (Figure 5.19). Changes in gene expression were interpreted relative to undifferentiated RPAT hiPSCs. No statistically significant difference in gene expression was identified between cultures treated with SB/CHIR at 38h or 40h in all genes and timepoints assessed. TAL1 is a master regulator of mesoderm and haemato-endothelial specification (Org et al., 2015; Real et al., 2012). Consistent with this role, *TAL1* was first expressed in d4 EBs and its expression increased as differentiation proceeded. *GATA2* demonstrated a similar expression profile. *RUNX1c* was first expressed in d7 EBs and was further upregulated at d14, where it remained similarly elevated at d18. *SPI1* is directly regulated by RUNX1 (Hoogenkamp et al., 2009) and its expression closely followed *RUNX1c*. Upregulation of *HOXA9* in hESC differentiations parallels haematopoietic commitment (Ramos-Mejía et al., 2014). *HOXA9* expression reached maximum levels in d7 EBs, but expression dropped at d14. *HOXA9* expression appeared elevated again in d18 cells, however, they were not statistically significant from d4 expression levels.

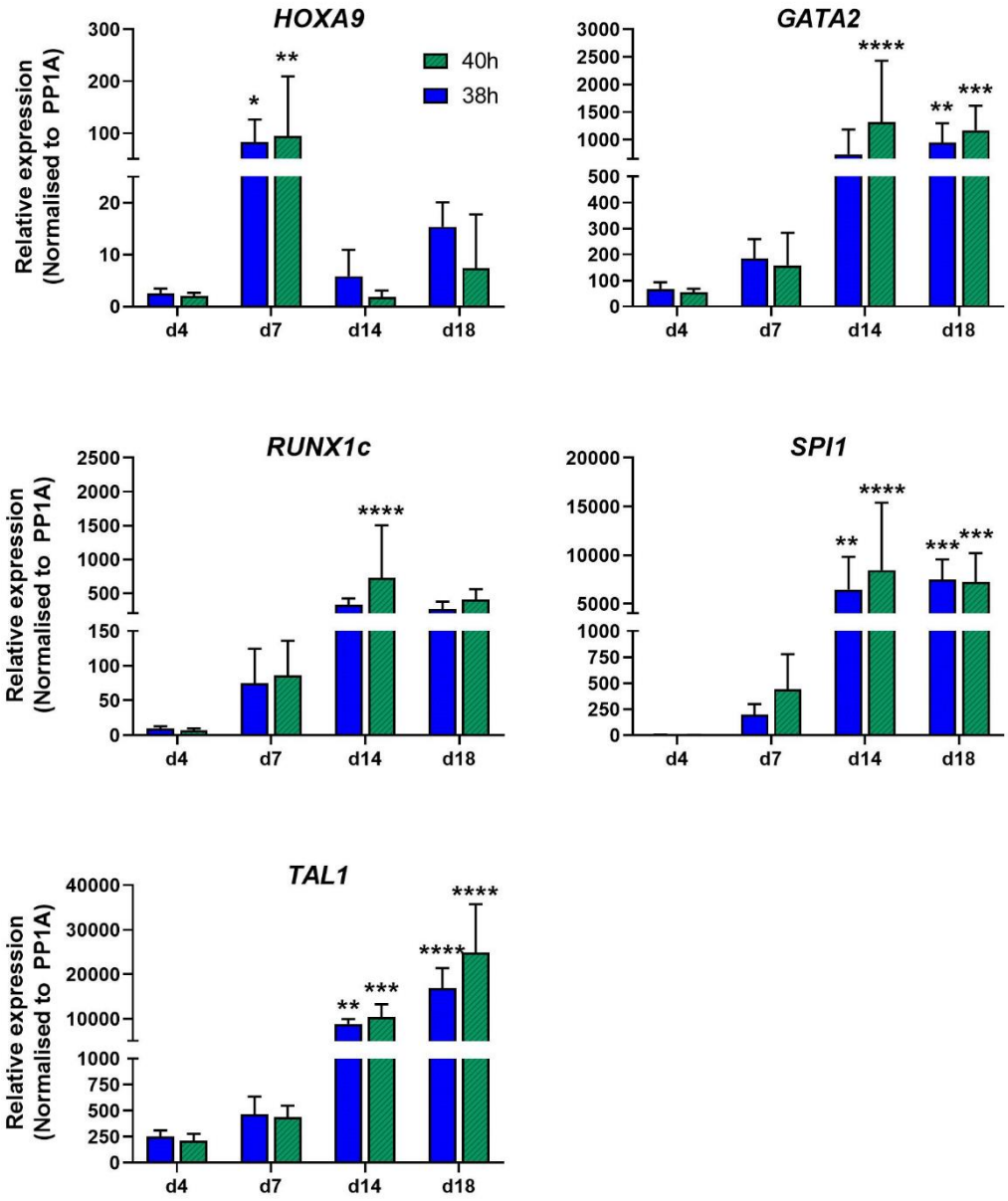


Figure 5.19: qPCR reveals gradual acquisition of key haematopoietic transcription factors as differentiation progresses

QPCR analysis of *HOXA9* and haematopoietic master regulator genes (*GATA2, RUNX1c, TAL1* and *SPI1*) expression on days 4, 7, 14 and 18 of differentiation in cultures treated with SB/CHIR at 38h or 40h post-seeding. Expression calculated relative to undifferentiated RPAT hiPSCs. Results are mean ± SD, n = 3 biological replicates; *p < 0.05, **p < 0.01, ***p < 0.001 and ****p < 0.0001 compared to d4 38h or 40h, 2-way ANOVA, Tukey's multiple comparisons test.

5.4 Discussion

In this chapter, a hiPSC haematopoietic differentiation model was optimised using RPATs. A monolayer-based approach was first adopted using the commercial STEMdiff™ Hematopoietic kit. However, generation of haematopoietic progenitors was found to be inconsistent and inefficient via this approach and an alternative EB-based differentiation procedure was tested. Following optimisation, the spin EB differentiation readily produced multipotent haematopoietic progenitors and represents a suitable *in vitro* model to further study the role of HOXA9 during haematopoiesis.

A monolayer-based differentiation approach was first chosen as they allow relatively straightforward tracking and imaging of differentiating cells, which can be complicated in EBs due to their 3D structure. This was particularly important for this project as mScarlet fluorescence was intended to be used as a surrogate for HOXA9 expression in iA9 cells, facilitating assessment of HOXA9 expression in real-time via fluorescence imaging. The STEMdiff™ Hematopoietic kit was seen as a suitable method to generate haematopoietic progenitors due to its relatively simple design. While performing the differentiation itself was quite straight-forward, generating iPSC colonies of the appropriate size and number for efficient differentiation proved very challenging. Generation and seeding of hPSC colonies represents a crucial step in monolayer-based differentiations, with it previously been noted that higher seeding densities can supress haematopoietic progenitor specification (Cao et al., 2020). ReLeSR was initially used to dissociate RPAT iPSCs, however GCDR proved more efficient. Although ReLeSR is the recommended dissociation agent, it does not appear to be suitable for all cell lines as other studies using the STEMdiff™ Hematopoietic kit have used GCDR (Rai et al., 2020) or EDTA (Ruiz et al., 2019) to dissociate iPSCs as clumps. Since GCDR facilitated complete dissociation of RPAT iPSCs it helped provide some standardisation to colony seeding, as seeding ratios were easier to replicate. However, the number of colonies seeded per well and the size of colonies remained highly variable. Seeding of iPSC as single cells with rock inhibitors which can support their survival and growth, such as Revitacell (Hansen et al., 2018), may be a better approach to generate reproducible uniform colony sizes and densities.

The STEMdiff™ Hematopoietic kit has shown success in generation of haematopoietic progenitors which are subsequently differentiated to macrophages (Matsuo et al., 2021) and microglia (Rai et al., 2020). However, in published studies using the STEMdiff™ kit, generation of haematopoietic progenitors with demonstrable functional haematopoietic activity has required additional optimisation steps. In one study, $CD34^{+}CD38^{-}$

haematopoietic progenitors with robust erythroid-myeloid CFU potential were produced following incorporation of an Activin/Nodal inhibition and Wnt induction step into the differentiation process (Ruiz et al., 2019). Similarly in another study, brief 24h Wnt stimulation on differentiation d0 and delayed differentiation initiation until 4 days post-seeding was used to generate $CD34^{+}CD43^{+}$ progenitors (Fortschegger et al., 2021). Although direct comparisons with unmodified STEMdiff™-derived haematopoietic progenitors were not made in either study, they suggest that further optimisation of the STEMdiff™ Hematopoietic kit is required for optimal results.

The 12 differentiation experiments performed with RPAT hiPSCs showed high variability in haematopoietic progenitor generation and differentiation potential in the CFU assay. As such, the STEMdiff™ Hematopoietic kit must be optimised in order to achieve reproducible and efficient differentiation of the RPAT hiPSC line. The proprietary nature of the differentiation media made optimisation challenging. Media A is reported to contain bFGF, BMP4 and VEGF while Media B also contains SCF, FLT3 and TPO (Ruiz et al., 2019). As was later identified in Spin EB differentiations, RPAT hiPSC are sensitive to the dosage of BMP4 and Activin A, with haematopoietic differentiation enhanced when BMP4 was added at low concentrations. Therefore, it is possible the concentration of growth factors and cytokines used in the STEMdiff™ media may be limiting for optimal differentiation of RPAT and potentially other hiPSCs. Without knowledge of their concentration, modification of growth factor and cytokine concentrations would have been very challenging and was not attempted.

The poor reproducibility of the STEMdiff™ differentiation particularly affected interpretation of CFU assay results. To produce reliable results, the CFU assay requires optimisation of plating, seeding and methylcellulose conditions. CFU assays were performed in 12 well-plates with MethoCult™ Classic (H4434, STEMCELL Technologies) and/or MethoCult™ Optimum (H4034, STEMCELL Technologies) at a range of seeding densities. Although 9/12 differentiations produced $CD34^{+}CD38^{-}$ progenitor cells, only 4/12 differentiations generated haematopoietic colonies. It was unclear if failure to generate haematopoietic colonies was as a result of inappropriate assay conditions or if suspension cells lacked differentiation potential. Colonies produced were restricted to the myeloid lineage, with suspension cells from only one differentiation were capable of producing erythroid cells. This may reflect a myeloid bias of the haematopoietic progenitors or inability of the CFU assay to support differentiation of other colony sub-types. Without a reliable source of haematopoietic progenitors, optimisation of the assay conditions was very challenging.

The STEMdiff™ Hematopoietic kit differentiation efficiency can be improved by modulation of Activin/Nodal and/or Wnt signalling pathways as previously discussed. In other monolayer-based differentiations variability can be influenced by the substrate used, in particular, the frequency of haematopoietic progenitors was found to be more consistent when differentiations were performed on vitronectin compared to Matrigel (Jeong et al., 2020). Future experiments using the STEMdiff™ Hematopoietic kit should consider these modifications for more efficient and reproducible results.

The poor reproducibility of the STEMdiff™ kit prompted the change to the spin EB differentiation procedure. With this method, EBs were seeded at a defined, reproducible density, with EB size consistent across different experiments as demonstrated by quantification of EB diameter using ImageJ. Furthermore, as experiments comparing EBs treated with or without SB/CHIR showed, only EBs which increased in size from d2 to d4 were able to generate haematopoietic progenitors, thus providing a means of predicting haematopoietic potential at an early stage of differentiation. Using this measure, differentiation of EBs which do not increase in size can be discontinued, saving on costly media and time.

All differentiation components of the spin EB procedure were clearly defined, therefore, when optimisation was required, it was relatively straightforward. Several features of the spin EB differentiation procedure outlined in the published work were adapted for RPAT differentiation including, timing of SB/CHIR addition; concentration of BMP4/Activin A; use of commercial APEL2 media; and differentiation of Matrigel-cultured hPSCs. EBs formed from hPSCs maintained on MEFs or in MEF-conditioned media reportedly have a higher haematopoietic output (Kardel and Eaves, 2012) and it was initially hypothesised the lack of progenitors observed in the first experiments were due to the use of Matrigel to maintain RPAT hiPSCs. This hypothesis was tested in a comparative differentiation of RPATs and MEF-maintained S/R-H9 cells, kindly donated by Dr Monica Nafria. Both cell lines generated robust NCAM+ mesendoderm, although RPAT EBs produced approx. 5-fold more $NCAM^{+}KDR^{+}$ cells than S/R-H9 EBs on d4, suggesting the differentiations were proceeding at different rates. However, cell cycle analysis was not performed to assess the validity of this observation. S/R-H9 EBs also generated a greater proportion of $CD34^{+}CD90^{+}$ haemato-endothelial cells than RPATs. Expression of SOX17-mCherry further demonstrated the strong endothelial induction in S/R-H9 cells. Despite this, S/R-H9 cells did not generate haematopoietic progenitors, likely due to inefficient haematopoietic commitment following transfer to adherent culture. RPATs had reduced endothelial potential compared to S/R-H9s and similarly did not generate

haematopoietic progenitors. However, in this case poor mesoderm patterning or haemato-endothelium induction were likely responsible for the lack of success. Unfortunately, replication of this experiment was not possible. Live S/R-H9 cells were gifted from Dr Monica Nafria, intended to be differentiated straightaway. Expansion of the S/R-H9 cells or transition of RPATs to MEF culture was not possible due to lack of resources to maintain cells on MEFs.

As BMP4 and Activin A signalling plays a central role in haematopoietic differentiation (Cerdan et al., 2012; Tan et al., 2013), BMP4 and Activin A (B/AA) were cross titrated to determine the optimal concentration of B/AA for haematopoietic differentiation of RPATs. All 9 B/AA conditions tested produced suspension cells, although the proportion of $CD34^+CD43^+/CD45^+$ haematopoietic progenitors and colony-forming potential varied amongst the different conditions. In general, low concentrations of BMP4 (5 or 10 ng/mL) generated the most $CD34^+CD43^+/CD45^+$ cells, while Activin A did not appear to have a major effect, unless combined with high BMP4, which negatively impacted progenitor formation. Interestingly, the B/AA concentrations used in previous experiments (20/10 ng/mL B/AA or 20/7.5 ng/mL B/AA, see Table 5.2) resulted in the lowest yield of $CD34^+CD43^+$ (<3%) and $CD34^+CD45^+$ (<1.5%) progenitors. As such, high BMP4 signalling is inappropriate for optimal haematopoietic differentiation of RPAT hiPSC, and use of such a high concentration in previous experiments was likely a contributing factor in their poor outcome.

The spin EB differentiation procedure incorporates a SB/CHIR treatment step from d2 to d4 of differentiation which promotes the formation of $SOX17^+RUNX1c^-CD34^+CD90^+$ endothelium, phenotypically similar to human AGM. In the absence of SB/CHIR, hESCs differentiated via the spin EB procedure still generate haematopoietic progenitors; however, differentiation is more characteristic of primitive haematopoiesis (Ng et al., 2016). To better understand the role SB/CHIR plays in directing definitive and primitive haematopoiesis, RPAT hiPSCs were differentiated with or without the addition of SB/CHIR from d2 – d4. We initially hypothesised that in the absence of SB/CHIR, the exogenous growth factors and cytokines added (BMP4, Activin A, VEGF, SCF and FGF2) would be sufficient to direct differentiating cells towards the haematopoietic lineage, possibly generating primitive progenitors. Surprisingly, cultures which did not receive SB/CHIR presented with a haematopoietic defect from an early stage. Unlike SB/CHIR-treated EBs, untreated EBs did not increase in size from d2 to d4 of differentiation; KDR was not upregulated on d4; the haemato-vascular population on d7 was substantially reduced; and haematopoietic transcription factors were not

upregulated on d4 or d7. Additionally, haematopoietic progenitors were not produced from d14 or d18 cultures. The data suggests a complete defect in haematopoiesis, in both the primitive and definitive programs, in the absence of SB/CHIR. Both haematopoietic programs transition through KDR+ mesoderm and $CD34^{+}CD43^{-}$ HE (Ditadi et al., 2015; Sturgeon et al., 2014), yet neither state was evident in -SB/CHIR EBs at the timepoints assayed. Early upregulation of CD43 in d7 EBs is indictive of primitive haematopoiesis (Ng et al., 2016), however, CD43+ cells were not detected in d7 -SB/CHIR EBs. Together this data suggests SB/CHIR is essential for haematopoietic differentiation of RPAT hiPSC via the spin EB differentiation procedure.

A limitation of this experiment is haematopoietic potential was not assessed prior to d14. The first haematopoietic cells generated both *in vitro* and *in vivo* have limited differentiation ability, mainly restricted to primitive erythrocytes. This can be recapitulated in the BL-CFC assay (Kennedy et al., 2007). HPSC-derived KDR+ cells with primitive potential are endowed with BL-CFC activity which is lacked by definitive KDR+ cells (Sturgeon et al., 2014). In -SB/CHIR cultures, assessment of haematopoietic potential of d4 EBs or earlier using the BL-CFC assay would be a better approach to determine if primitive haematopoiesis occurs in the absence of SB/CHIR. Additionally, CD235a specifically marks cells with primitive potential and evaluation of its expression would also help establish if +/-SB/CHIR directs differentiation of primitive or definitive haematopoiesis (Ditadi et al., 2015; Sturgeon et al., 2014).

Haematopoietic cells produced during hPSC-based differentiations are heterogenous populations. SB/CHIR-treated cultures generated haematopoietic progenitors on d14 and d18, although with poor efficiency. Less than 5% of d14 cells expressed HSPC-like surface markers $CD34^{+}CD43^{+}CD90^{+}CD38^{-}$ and <10% were $CD34^{+}CD43^{+}$. Haematopoietic differentiation does not yield singular cell types; at a given timepoint, a population of cells may simultaneously possess primitive and definitive potential or comprise of both vascular and haemogenic endothelium. Early upregulation of CD43 during HE specification suggests commitment to the primitive program (Nafria et al., 2020a; Ng et al., 2016). In d7 EBs, high CD43 expression (20%) was evident in +SB/CHIR cultures, with most CD43+ cells co-expressing CD90 (14.5%) and/or CD34 (13.9%), surface marker expression profiles associated with haematopoietic progenitors. In the BMP4/Activin A cross-titration experiment, some conditions also possessed high CD43 expression on d7 (>15%), which tended to correlate with low yields of $CD34^{+}CD43^{+}/CD45^{+}$ progenitors on d15. This data suggested SB/CHIR treatment was

not adequately suppressing primitive haematopoiesis during mesoderm patterning and some cells continued to differentiate towards a primitive fate.

SB/CHIR was administered at earlier timepoints (38h and 40h post-seeding) than previously reported (44 – 48h post-seeding) to assess if earlier Wnt activation/Activin inhibition would improve primitive haematopoiesis suppression. Little difference was observed between the two timepoints in the various assays performed throughout differentiation. Both timepoints produced $CD34^{+}CD43^{+}CD90^{+}CD38^{-}$ HSPC-like cells, with a higher proportion of d14 cells (29.8 – 34.9%) than d18 cells (13.4 – 19.9%) expressing the HSPC-like immunophenotype. In d7 EBs, <15% of cells expressed CD43 and <10% co-expressed CD43, CD34 and CD90, suggesting primitive haematopoiesis was better supressed under the conditions applied. However, without a direct comparison with EBs treated with SB/CHIR at later timepoints (44 – 48h), more efficient primitive fate suppression or enhanced definitive haematopoietic differentiation could not be concluded.

Inclusion of SB/CHIR in the spin EB differentiation procedure promotes acquisition of a *HOXA* signature, hypothesised to endow haematopoietic progenitors with repopulating activity in immunodeficient mice (Ng et al., 2016). The *HOXA* and *CDX* expression profile was examined in d4 and d7 EBs. HOXA genes are reported to first be upregulated on d4 and remain elevated by d7 (Ng et al., 2016). In RPAT differentiation, HOXA genes were not upregulated until d7. *CDX1* and *CDX2* were upregulated on d4, consistent with their role in mesendoderm development (Creamer et al., 2017), however, *CDX4* was not expressed. CDX4 has a central role in specification of mesoderm with definitive haematopoietic potential (Creamer et al., 2017), suggesting promotion of definitive fate specifically may not be a function of SB/CHIR and SB/CHIR may primarily act to supress primitive haematopoiesis. In another study, upregulation of HOXA genes in HE was found solely in cells fated to the intra-embryonic-like definitive program (Dege et al., 2020). Only $CD34^{+}CD43^{-}$ HE sorted from $KDR^{+}CD235a^{-}$ mesoderm exhibited high *HOXA* expression and generated $CD4^{+}CD8^{+}$ T lymphocytes following culture on OP9-DLL4. $CD34^{+}CD43^{-}$ HE sorted from $KDR^{+}CD235a^{+}$ mesoderm showed low or no *HOXA* expression and although $CD56^{+}CD45^{+}$ natural killer cells could be generated in co-culture with OP9-DLL4, $CD4^{+}CD8^{+}$ T lymphocytes were not produced. *HOXA* expression was not measured in $KDR^{+}CD235a^{+/-}$ mesoderm, but the results suggest *HOXA* expression in HE is indicative of the definitive program, with high and low expression recapitulating intra-embryonic and extra-embryonic haematopoiesis, respectively (Dege and Sturgeon, 2017; Dege et al., 2020).

A key feature of haematopoietic progenitors is their ability to differentiate into myeloid and erythroid colonies in the CFU-assay. Lack of an optimised and efficient CFU assay was a major drawback of early experiments. CFU assays performed on suspension cells from the STEMdiff™ Hematopoietic kit and BMP4/Activin A cross-titration experiments used MethoCult™ Optimum or Classic. The results indicated suspension cells from these experiments had limited and mostly myeloid colony forming potential. Due to differences in haematopoietic progenitor derivation and CFU assay set-up conditions, comparison with published work is difficult. However, as an example, one study was able to generate up to 200 colonies with diverse differentiation potential when hiPSC-derived haematopoietic progenitors were cultured at $4x10^4$ cells/mL in MethoCult™ Optimum, although the size of culture dish used was not noted (Tsui et al., 2021). Our optimisation of CFU assay conditions demonstrated the methylcellulose media can have a major impact on colony formation, and thus, interpretation of results. Seeding d14 suspension cells at $3x10^3$ or $5x10^3$ cells/well in 24-well plates in either MethoCult™ SF H4636 or 1% MC-APEL2 supplemented with haematopoietic cytokines had drastically different results despite the similarities in assay set-up. H4636 supported robust colony formation, while most cells seeded in 1% MC-APEL2 failed to differentiate. Although the results of CFU assays performed prior to use of H4636 tended to follow the same pattern as the flow cytometry data, they likely do not accurately reflect the true colony forming potential of suspension cells.

As H4636 supported more robust differentiation, we were able to decipher different functional activities of d14 and d18 suspension cells, however, differences were not determined to be statistically significant. Only two biological repeats were collected for CFU-GEMM and BFU-E/CFU-E counts, lowering the power of statistical tests. Day 14 suspension cells harvested from cultures treated with SB/CHIR at 38h or 40h tended to produce more erythroid colonies (BFU-E/CFU-E) than d18 cells, while d18 cells generated more immature mixed erythoid-myeloid colonies (CFU-GEMMs) than d14 cells. The increase in CFU-GEMMs generated from d18 cells was particularly notable as these colonies are considered quite rare amongst hPSC-derived $CD34^+CD45^+$ progenitors (Doulatov et al., 2013). Their presence suggests d18 cells may represent true multipotent cells, being able to give rise to both myeloid and erythroid lineages. Conversely, d14 progenitors are biased toward erythroid differentiation. Despite their similar surface marker phenotype, the differences observed suggest d14 and d18 cells are not the same and represent two different haematopoietic progenitor populations, worthy of further study.

Identification of predictors of haematopoietic differentiation efficiency can help improve productivity. As discussed, increase in EB size from d2 to d4 of differentiation was found to be associated with haematopoietic potential, providing an early measure to assess differentiation efficiency. On review of all flow cytometry data for d4, low – medium (2 – 15%) $NCAM^+KDR^+$ expression was also associated with haematopoietic potential. In the BMP4/Activin A cross-titration experiment, conditions which produced high (>25%) or very low (<2%) $NCAM^+KDR^+$ populations, also generated few (<5%) $CD34^+CD45^+$ haematopoietic progenitors and suspension cells which lacked CFU activity on d15. Similarly, -SB/CHIR cultures lacked haematopoietic potential and did not upregulate KDR on d4. Haematopoietic progenitors were not produced in the S/R-H9 and RPAT comparative differentiation; however, RPAT d4 EBs expressed 5-fold more $NCAM^+KDR^+$ cells than S/R-H9 (34% RPAT vs 10.9% S/R-H9) and later showed poor haemato-vascular mesoderm potential. All other differentiations performed successfully generated haematopoietic progenitors, albeit at varying frequencies, and contained 2 – 15% $NCAM^+KDR^+$ cells on d4. Together this data suggests the d4 timepoint represents mesoderm in the early stages of specification to haemogenic mesoderm. When assayed at this timepoint, a high proportion of $NCAM^+KDR^+$ cells or absence of KDR indicates the differentiation is not proceeding at the optimal rate for haematopoietic specification under the conditions outlined.

Chapter 6: Transient HOXA9 overexpression during mesoderm patterning

6.1 Introduction

The primary aim of this thesis was to establish the function of HOXA9 at specific stages of haematopoietic development in an iPSC model of haematopoiesis. To this end, the DOX-inducible HOXA9-mScarlet reporter system was generated. As described in Chapter 3 and 4, DOX-mediated induction of HOXA9/mScarlet was found to reduce during differentiation, with substantial reduction in induction observed following mesendoderm specification in cardiac differentiation and in a preliminary haematopoietic differentiation experiment. Thus, use of this model was limited to the study of HOXA9 function during early differentiation. Primitive and definitive haematopoietic fates are purported to diverge during mesoderm patterning. Upregulation of CD235a in KDR+ mesoderm marks a population with BL-CFC activity, and is considered to be an indicator of primitive haematopoiesis (Ditadi et al., 2015; Kennedy et al., 2012; Sturgeon et al., 2014). Conversely, haematopoietic progenitors with definitive erythro-myeloid and T lymphoid potential have been shown to arise from $KDR^{+}CD235a^{-}$ mesoderm (Dege et al., 2020; Luff et al., 2022). As such, mesoderm represents an appropriate and potentially critical state to study the function of HOXA9 in haematopoiesis and its role in haematopoietic fate decisions.

To model haematopoiesis, a monolayer-based approach using the STEMdiff Hematopoietic kit was initially chosen. As discussed in Chapter 5, this differentiation approach proved inefficient and unreliable, and the spin EB differentiation procedure was subsequently adopted (Ng et al., 2016). The spin EB method was extensively characterised and was shown to recapitulate key stages of *in vivo* embryonic haematopoiesis. Co-expression of NCAM and KDR in d4 EBs was found to mark mesoderm with haematopoietic potential. A bipotent population with haematopoietic and endothelial differentiation potential, described as $CD90^{+}CD34^{+}CD43^{-}$, was formed on d7. Finally, $CD34^{+}CD43^{+}CD90^{+}CD38^{-}$ HSPC-like cells were generated in two waves, on d14 and d18.

6.2 Aims

The concentration of DOX used in previous experiments was based on the response DOX induced in iPSCs. However, induction of *HOXA9*/mScarlet is not retained in mesoderm-derived lineages. As such, the first aim of this chapter was to determine the optimal DOX concentration to overexpress HOXA9 in iPSC-derived mesendoderm. iA9 clone 54 (here after referred to as iA9) was used in all experiments described in this chapter. This clone had been shown to be pluripotent and possessed a normal karyotype. Additionally, it retained relatively robust HOXA9 and mScarlet induction in mesoderm populations. As has been observed with cardiac differentiations, the ability of DOX to induce *HOXA9*/mScarlet was also shown in Chapter 3 to reduce during differentiation with the STEMdiff™ Hematopoietic kit. DOX responsiveness during spin EB haematopoietic differentiation had previously not been investigated. Therefore, the second aim of this chapter was to assess DOX inducibility during spin EB differentiation. The third and final aim of this chapter was to investigate the impact of transient HOXA9 overexpression during mesoderm patterning on mesoderm and haematopoietic progenitor formation. To this end, iA9 iPSCs were differentiated following the spin EB differentiation procedure outlined in Chapter 5. Flow cytometry and qPCR were used to monitor differentiation progression while RNA-sequencing was performed to identify transcriptome changes induced by HOXA9. The erythro-myeloid differentiation potential of progenitors generated was also determined using the CFU assay.

6.3 Results

6.3.1 Optimisation of DOX treatment conditions in iPSC-derived mesendoderm

In chapter 4, DOX treatment conditions were optimised in iPSCs, with the aim of determining the optimal concentration and duration of DOX to amplify the mScarlet signal and facilitate mScarlet detection via fluorescence microscopy and quantification via mScarlet intensity measurements. Two pulses of 2 µg/mL DOX was found to produce the highest expression of mScarlet in iPSCs, however, this treatment condition was later found to reduce cell number during cardiac differentiation (Figure 4.6). Given the discrepancies in response to DOX in iPSCs and during cardiac differentiation, we decided to re-access the optimal DOX concentration and treatment duration to upregulate HOXA9, specifically in iPSC-derived mesendoderm. iA9 iPSCs were differentiated as monolayers to mesendoderm (identified by *MIXL1* expression) and were

treated with five different concentrations of DOX for 2 days (ranging from 0.25 – 2 µg/mL). The impact of successive DOX treatments on induction was also assessed and cells either received one pulse of DOX (DOX x1) or two pulses at 24h intervals (DOX x2). iA9-derived mesendoderm was stained for HOXA9/2A and DAPI and imaged using the Operetta® imaging system (Figure 6.1A). The percentage of mScarlet- and 2A-positive cells was calculated relative to DAPI nuclear staining using the Columbus™ analysis system. DOX was able to induce robust mScarlet and 2A expression at all concentrations tested. (Figure 6.1A, B). There was a clear increase in protein expression as DOX concentration was increased from 0.25 µg/mL to 1 µg/mL, however, a further increase in DOX concentration did not appear to induce substantially more protein expression. (Figure 6.1B). The fraction of mScarlet+ cells was higher than 2A+ cells at all concentrations, with maximum expression of both proteins identified following a single pulse of 1.5 µg/mL DOX (62.7% mScarlet+; 32.9% 2A+). No significant difference in mScarlet/2A induction was observed when cells were treated with either one or two pulses of DOX at all concentrations tested.

HOXA9 gene expression was assessed via qPCR and quantified relative to undifferentiated iPSCs. As was observed with mScarlet and 2A staining, *HOXA9* expression increased as DOX concentration increased up until 1 µg/mL, where higher DOX concentration did not lead to further increases in *HOXA9* expression (Figure 6.1C). Formation of mesendoderm represents the first step in iPSC specification to the haematopoietic lineage and as such expression of mesendoderm marker MIXL1 was also examined by qPCR. *MIXL1* expression was determined relative to uninduced mesendoderm. There was a trend for higher *MIXL1* expression following a single DOX treatment at 0.5 µg/mL or 1 µg/mL, with higher levels of *MIXL1* not observed at other concentrations (Figure 6.1C). Taking into consideration gene expression and fluorescence imaging data, a single treatment of 1 µg/mL DOX was selected for future experiments.

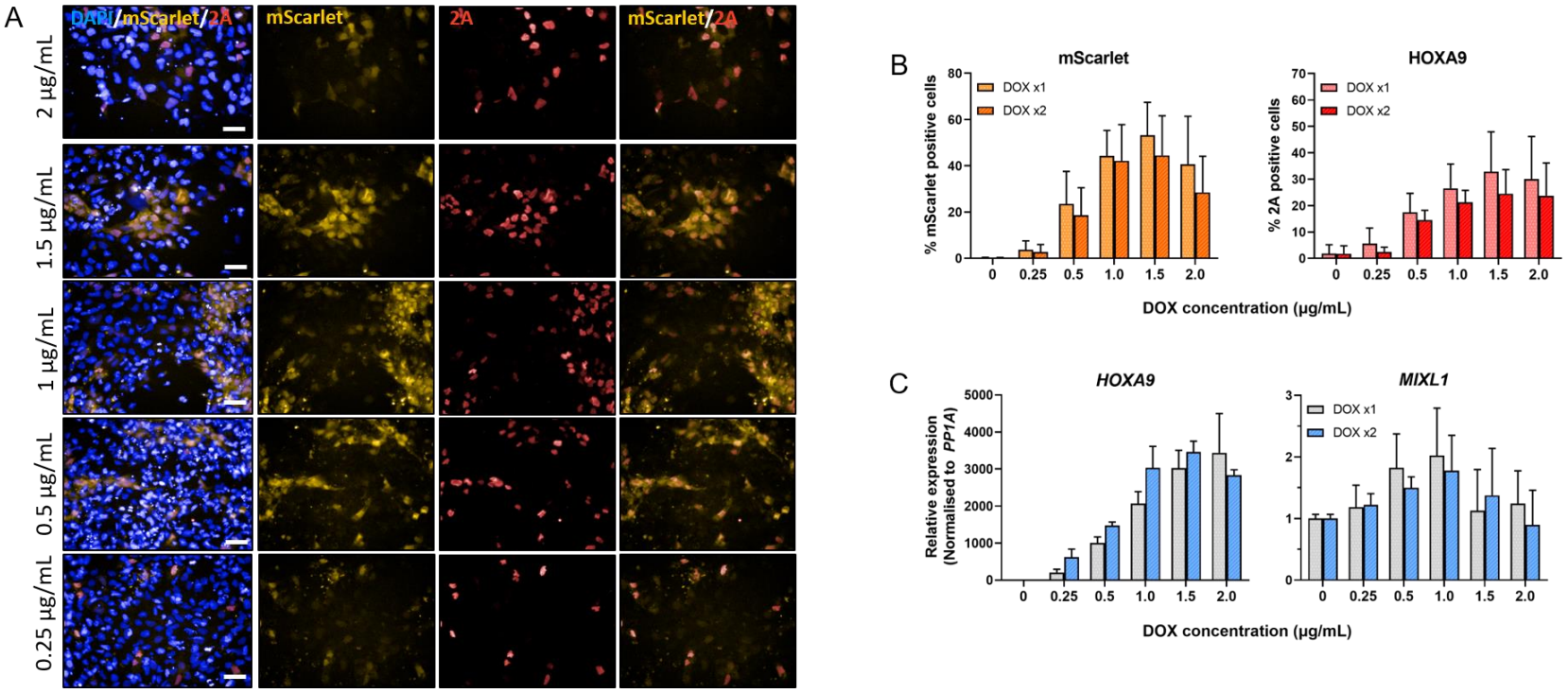


Figure 6.1: Optimisation of DOX treatment conditions in iPSC-derived mesendoderm

(A) Representative images of iA9-derived mesendoderm treated with different concentrations of DOX for 2 days. Scale bar = 40 µm. (B) Quantification of mScarlet and 2A/HOXA9 fluorescence images. (C) QPCR for *HOXA9* and *MIXL1* normalised to housekeeping gene *PP1A*. *HOXA9* expression is quantified relative to undifferentiated iPSCs. *MIXL1* expression is shown relative to uninduced mesendoderm (0 µg/mL). Results are from 2 independent biological replicates with 3 technical replicates performed for each. Mean ± SD is calculated based on technical replicates.

6.3.2 DOX inducibility is not maintained throughout spin EB differentiation

The ability of DOX to induce *HOXA9*/mScarlet expression was previously shown to be limited to iPSCs and early stages of haematopoietic differentiation using the STEMdiff™ Hematopoietic kit. While monolayer and EB-based differentiation procedures direct hPSCs through similar stages, the environment in which cells are differentiated is not the same. Access to nutrients and oxygen differs between the two culture systems, as does cell-cell interactions, whereby cells within the EB microenvironment form 3D interactions, while cells in monolayer systems form additional interactions with the extracellular matrix. The two systems also differ in differentiation duration; haematopoietic progenitors are generated in 12 days using the STEMdiff™ kit and in 18 days using the spin EB procedure. Due to these differences, the ability of DOX to induce HOXA9 expression throughout haematopoietic differentiation was again tested using the spin EB differentiation model. In a preliminary experiment, iA9 iPSCs were differentiated using the spin EB differentiation procedure and treated with a single pulse of 1 µg/mL DOX on d2, d3, d6 or d11 of differentiation. HOXA9 induction was measured by qPCR for *HOXA9* and *mScarlet* on d4, d7 and d12 of differentiation. Haematopoietic potential of d14 and d18 suspension cells was assessed via flow cytometry and the CFU assay (see Figure 6.2A for experimental plan).

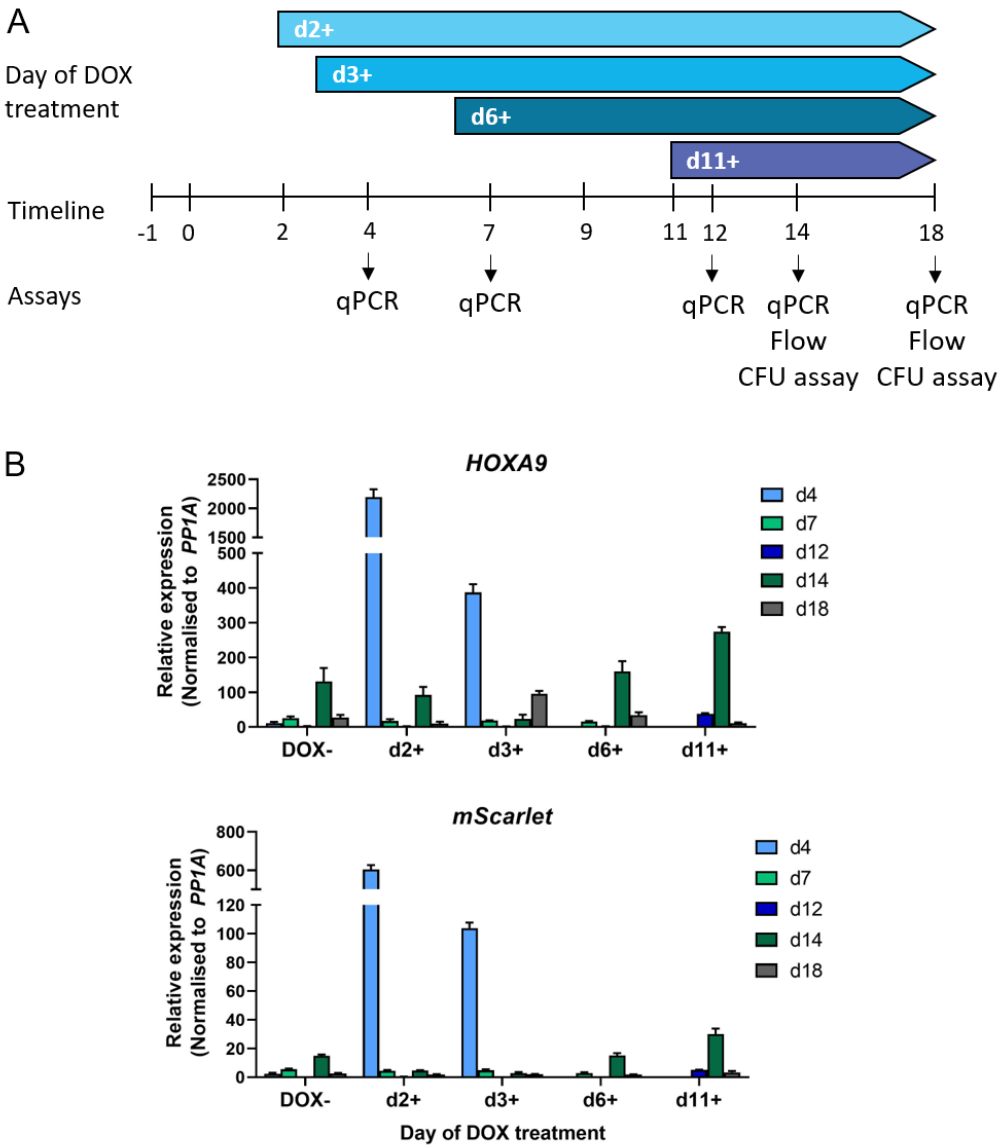


Figure 6.2: Inducible *HOXA9/mScarlet* expression is not retained after d3 of spin EB haematopoietic differentiation

(A) Overview of experimental plan to test the ability of DOX to upregulate *HOXA9* throughout spin EB haematopoietic differentiation. Differentiating cultures were treated with a single pulse of 1 µg/mL DOX on d2, d3, d6 or d11 of differentiation. *HOXA9* and *mScarlet* expression was quantified by qPCR on d4, d7, d12, d14 and d18. Haematopoietic progenitors generated on d14 and d18 were further assessed via flow cytometry and CFU assay. (B) QPCR for *HOXA9* and *mScarlet* at indicated timepoints following treatment of iA9 cultures with a single pulse of 1 µg/mL DOX on d2, d3, d6 or d11. Results are normalised to housekeeping gene *PP1A* and shown relative to undifferentiated iA9 iPSCs. Results are mean ± SD, n = 3 technical replicates.

As HOXA9 is reported to be upregulated during hESC-based haematopoietic differentiation (Ramos-Mejía et al., 2014), an uninduced (DOX-) control was differentiated as an indicator of endogenous *HOXA9* expression and *mScarlet* expression was measured by qPCR as an indicator of induced *HOXA9* expression. A single pulse of DOX on d2 or d3 of spin EB differentiation resulted in *HOXA9* and *mScarlet* upregulation, as measured by qPCR on d4. DOX treatment on d6 or d11 did not upregulate *HOXA9* or *mScarlet* when measured 24h after induction (Figure 6.2B). In d2 and d3 DOX induced cultures *HOXA9* expression was not maintained following the initial upregulation, with *HOXA9* or *mScarlet* no longer detected by d7. Although, d2 DOX-induced cultures also exhibited upregulation of *HOXA9* on d14, this was similarly observed in uninduced, d6 DOX-induced and d11 DOX-induced cultures suggesting it reflects endogenous *HOXA9* expression.

Suspension cells were generated on d14 and d18 in uninduced and DOX-induced cultures. The $CD34^+CD43^+$ haematopoietic fraction was similar in all treatment conditions, as was the total number of $CD34^+CD43^+$ cells produced (Figure 6.3A). The haematopoietic potential of d14 suspension cells from the different DOX induction conditions varied greatly, however, d18 cells from all conditions showed a similar colony-forming propensity (Figure 6.3B). Thus, although the ability of DOX to induce *HOXA9/mScarlet* expression is impaired after d3 of spin EB differentiation, short-term *HOXA9* upregulation or DOX treatment does not negatively impact haematopoietic progenitor formation.

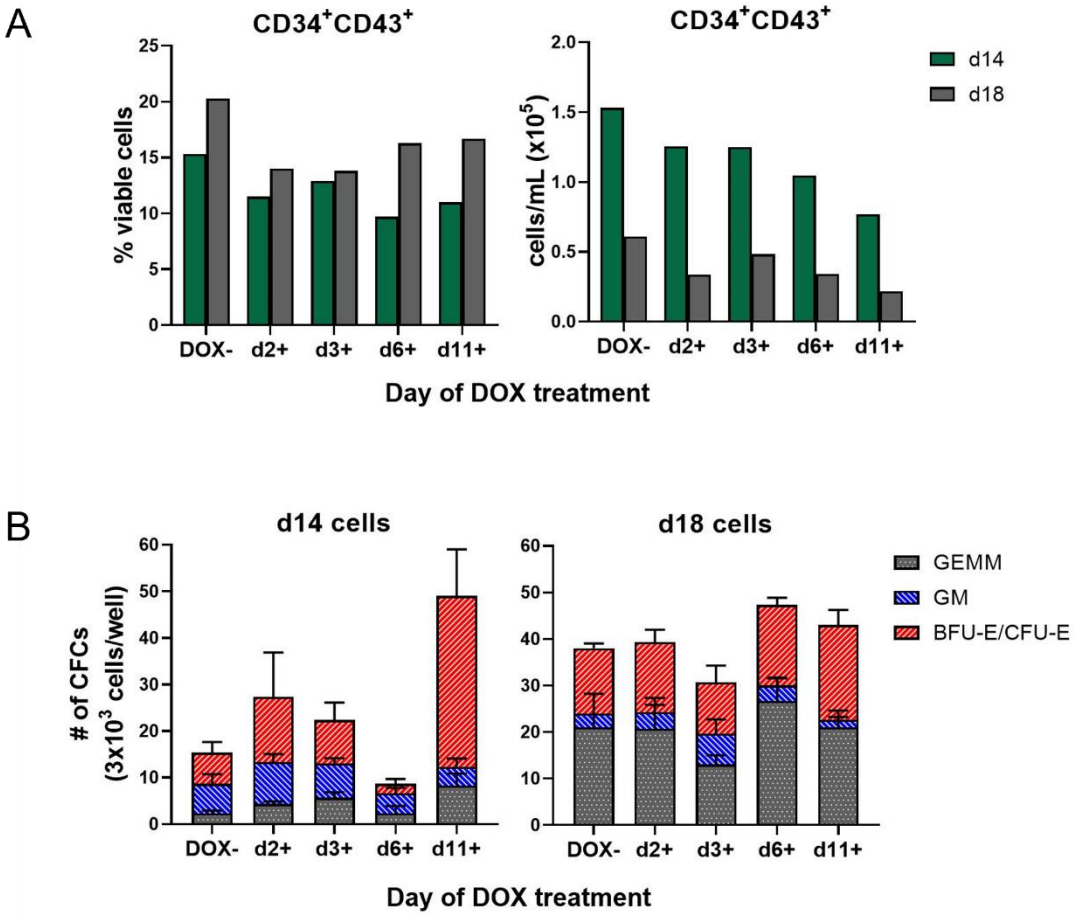


Figure 6.3: Spin EB differentiation of iA9 iPSCs yields haematopoietic progenitors

Flow cytometric analysis and assessment of haematopoietic colony-forming ability of suspension cells harvested on d14 and d18 from iA9 cultures treated with a single pulse of DOX on d2, d3, d6 or d11 of differentiation and uninduced cultures (DOX-). (A) Percentage and absolute number of $CD34^+CD43^+$ harvested from all conditions. (B) Assessment of haematopoietic colony formation following culture of d14 and d18 suspension cells in MethoCult™ SF H4636 for 14 days. Results are mean ± SD, n = 3 technical replicates.

6.3.3 Transient overexpression of HOXA9 during mesoderm patterning

To further establish the function of HOXA9 in mesoderm patterning, iA9 cells were differentiated to haematopoietic progenitors under four different conditions. The impact of transient HOXA9 overexpression on mesoderm and haematopoietic progenitor formation was then assessed. In Chapter 5, SB/CHIR treatment on d2 of differentiation was determined to be essential for efficient haematopoietic differentiation. Two timepoints of SB/CHIR addition were further investigated, 38h and 40h post-seeding, yet both were found to similarly promote haematopoietic differentiation. Given the importance of SB/CHIR for differentiation, the same timepoint was chosen for induction of HOXA9, in this case 40h post-seeding. iA9 cells were seeded for differentiation and, after 40h, differentiating cells were treated with 1 µg/mL DOX alone (SC-DOX+) or together with SB/CHIR (SC+DOX+). In parallel, iA9 cells were differentiated either receiving no DOX or SB/CHIR treatments (SC-) or with addition of SB/CHIR at 40h post-seeding (SC+), forming negative and positive controls of haematopoietic differentiation, respectively (see Figure 6.4 for summary of experiment plan). The inducible HOXA9 expression system was designed with HOXA9 tagged to mScarlet, thereby facilitating assessment of HOXA9 induction in live cells. However, due to the difficulties associated with imaging 3D structures, HOXA9 expression was not assessed by live fluorescence microscopy during iA9 differentiations as initially planned. HOXA9 induction was instead determined by *HOXA9* and *mScarlet* qPCR and mScarlet flow cytometry.

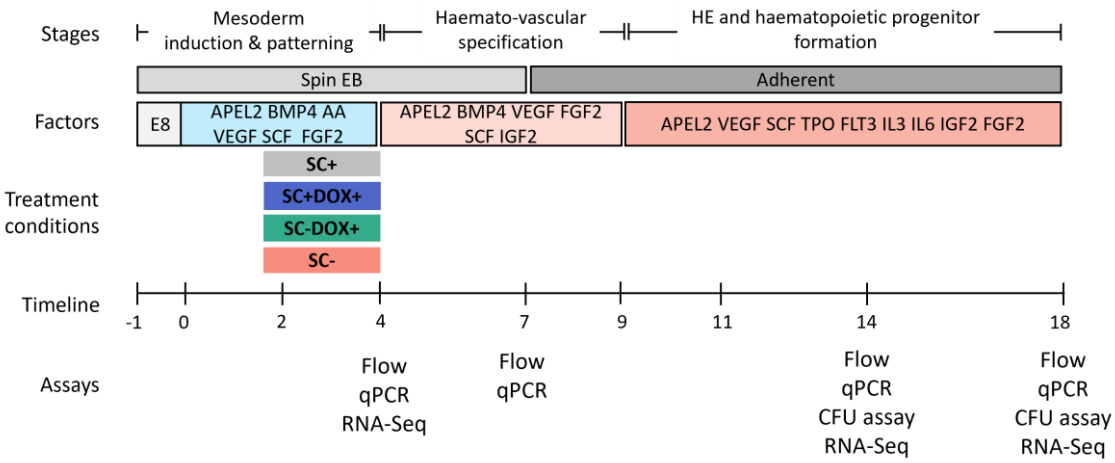


Figure 6.4: Experimental plan to establish the function of HOXA9 in mesoderm patterning and haematopoietic differentiation.

Summary of the developmental stages, culture format, growth factors and timeline during the differentiation of iPSCs to haematopoietic progenitors using the spin EB differentiation procedure. The four different treatment conditions employed to study the function and effect of transient HOXA9 overexpression at 40h post-seeding are indicated. The assays used to assess the impact of HOXA9 overexpression during mesoderm patterning are indicated, as are the timepoints the assays were performed at.

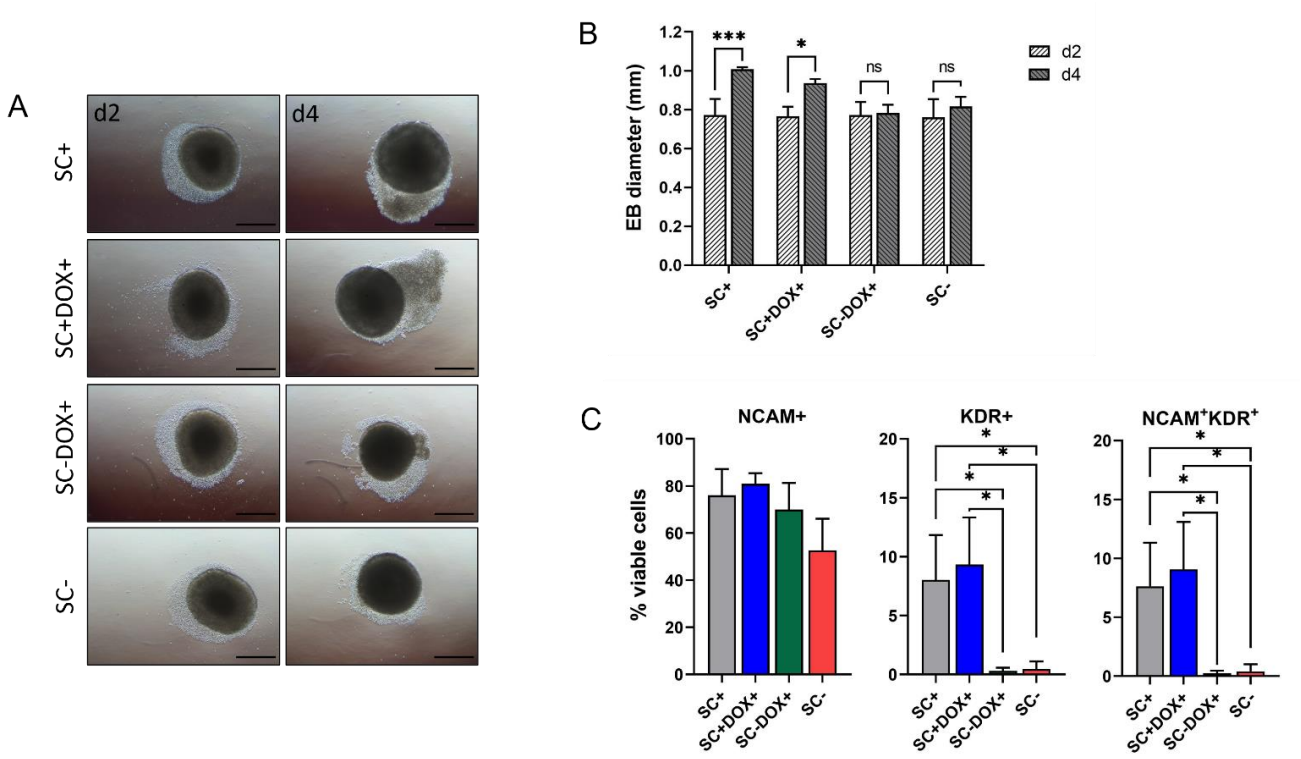


Figure 6.5: SB/CHIR treatment is necessary to form NCAM+KDR+ mesendoderm

(A) Representative images of EBs from each treatment conditions taken on d2 and d4 of differentiation. Scale bar = 0.5 mm. (B) Quantification of EB diameter using ImageJ on d2 and d4 of differentiation. Results are mean ± SD from 3 biological replicates with 3 technical replicates performed for each; *p < 0.05, ***p < 0.001, 2-way ANOVA, Sidak's multiple comparisons test. (C) Flow cytometry analysis of NCAM and KDR expression and $NCAM^{+}KDR^{+}$ mesendoderm in d4 EBs in all conditions. Results are mean ± SD, n = 3 biological replicates; *p < 0.05, One-way ANOVA, Tukey's multiple comparisons test.

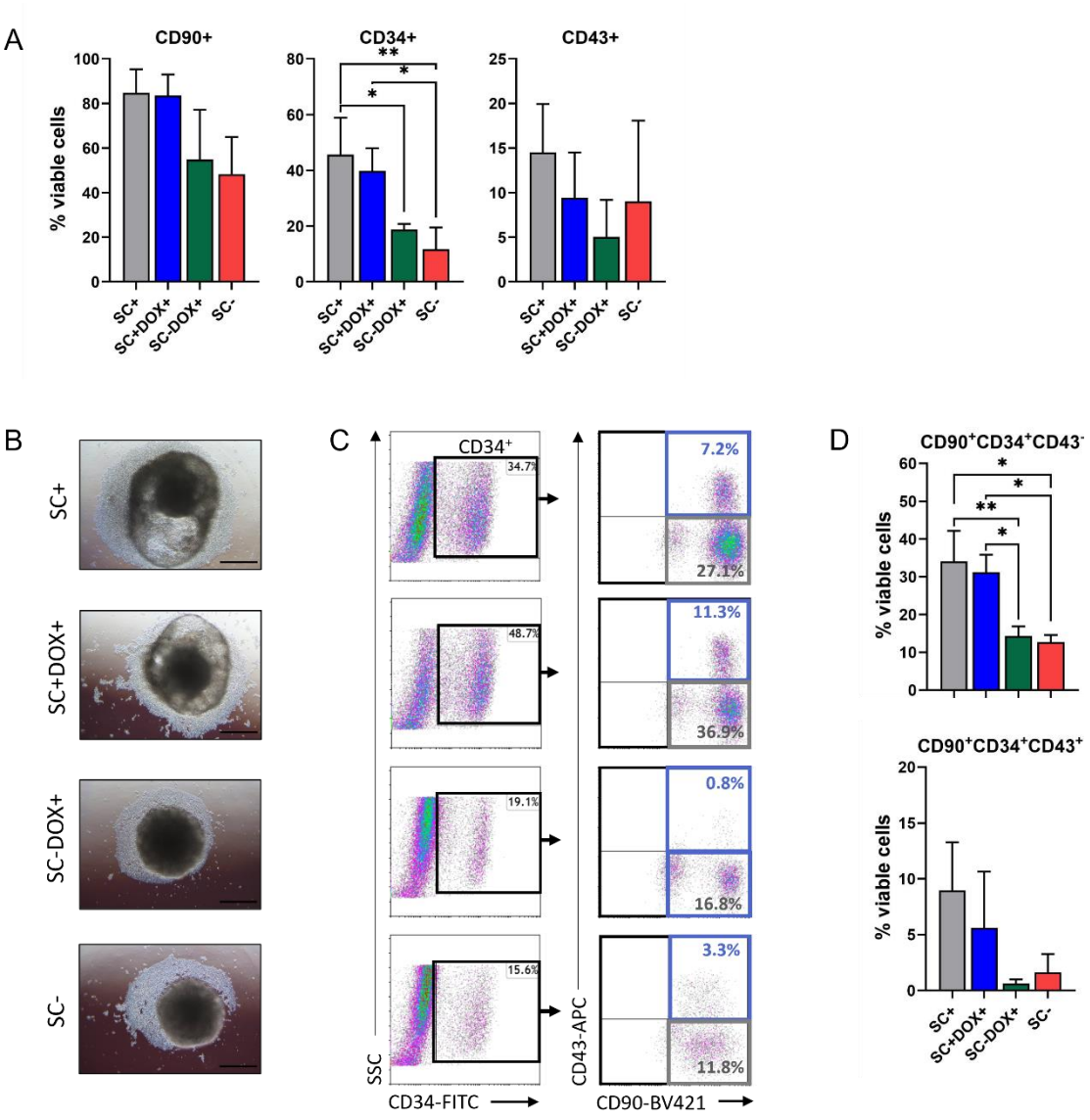


Figure 6.6: Formation of haemato-vascular cells is enhanced in cultures treated with SB/CHIR

(A) Flow cytometric analysis of CD90+, CD34+ and CD43+ populations from all treatments on d7. Results are mean ± SD, n = 3 biological replicates; *p < 0.05, and **p < 0.01, One-way ANOVA, Tukey's multiple comparisons test. (B) Representative images of d7 EBs from each treatment condition. Scale bar = 0.5 mm. (C - D) Representative flow plots demonstrating gating used to identify $CD90^{+}CD34^{+}C43^{+/-}$ populations in d7 EBs (C) and quantitative summary of flow cytometry analysis (D). Results are mean ± SD, n = 3 biological replicates; *p < 0.05, and **p < 0.01, One-way ANOVA, Tukey's multiple comparisons test.

Two methods of identifying d4 EB cultures with haematopoietic potential were described in Chapter 5. An increase in EB size from d2 to d4 of differentiation and formation of an $NCAM^+KDR^+$ population on d4 were associated with EBs which later produced haematopoietic progenitors. EBs from SC+ and SC+DOX+ cultures significantly increased in size from d2 to d4, while SC- and SC-DOX+ conditions did not (Figure 6.5A, B). EBs from all four conditions expressed the pan-mesoderm marker NCAM, although only SB/CHIR-treated cultures upregulated KDR on d4 (Figure 6.5C).

Flow cytometric analysis of d7 EBs showed all conditions expressed CD90, CD34 and CD43 but varied in the extent of expression (Figure 6.6A). No significant difference was observed in the proportion of CD90+ and CD43+ cells amongst the four conditions, however, there was a trend for higher expression of both markers in the SB/CHIR-treated groups. CD34 expression was found to be significantly higher in SC+ and SC+DOX+ groups compared to SC- and SC-DOX+ (Figure 6.6A). The proportion of haemato-vascular cells, described as $CD90^+CD34^+CD43^-$, was significantly greater in SB/CHIR-treated cultures compared to SC-DOX+ and SC- (Figure 6.6C, D). In SC+ and SC+DOX+ groups the fraction of early committed $CD90^+CD34^+CD43^+$ progenitors was relatively low (8.9±4.3% SC+; 5.6±5% SC+DOX+, Figure 6.6C, D). The analysis of CD90, CD34 and CD43 expression profiles suggests the haemato-vascular program was efficiently stimulated in SB/CHIR-treated cultures, yet absent in cultures which were not treated with SB/CHIR (SC- and SC-DOX+ groups).

QPCR was performed on d4 and d7 EBs from all conditions to assess *HOXA9/mScarlet* expression. As expected *HOXA9* and *mScarlet* were most highly expressed in SC+DOX+ and SC-DOX+ d4 EBs (Figure 6.7B). In particular, *HOXA9* was significantly higher in SC-DOX+ d4 EBs in comparison with SC+DOX+. Flow cytometry for mScarlet was also performed to validate the qPCR expression profiles. Surprisingly, mScarlet remained elevated in d7 EBs of SC+DOX+ and SC-DOX+ groups, with no difference observed in expression on d4 and d7 (Figure 6.7A). *HOXA9/mScarlet* mRNA was not detected in d7 qPCRs suggesting the high mScarlet protein levels reflect the high stability of the fluorescent protein and does not directly correlate with HOXA9 induction.

QPCR was also performed for mesendoderm marker *MIXL1* and the CDX genes. *MIXL1* was upregulated in all d4 conditions and downregulated by d7 (Figure 6.7B). SB/CHIR treatment is reported to upregulate *CDX* and H*OXA* expression (Ivanovs et al., 2017; Ng et al., 2016). Consistent with this role *CDX1* and *CDX2* were significantly upregulated in SC+ and SC+DOX+ d4 EBs only (Figure 6.7B). In d7 EBs, *CDX1* expression was

downregulated, while *CDX2* remained elevated. *CDX4* was elevated in d4 SC+DOX+ EBs compared to the other groups, however the difference was not found to be statistically significant (Figure 6.7B).

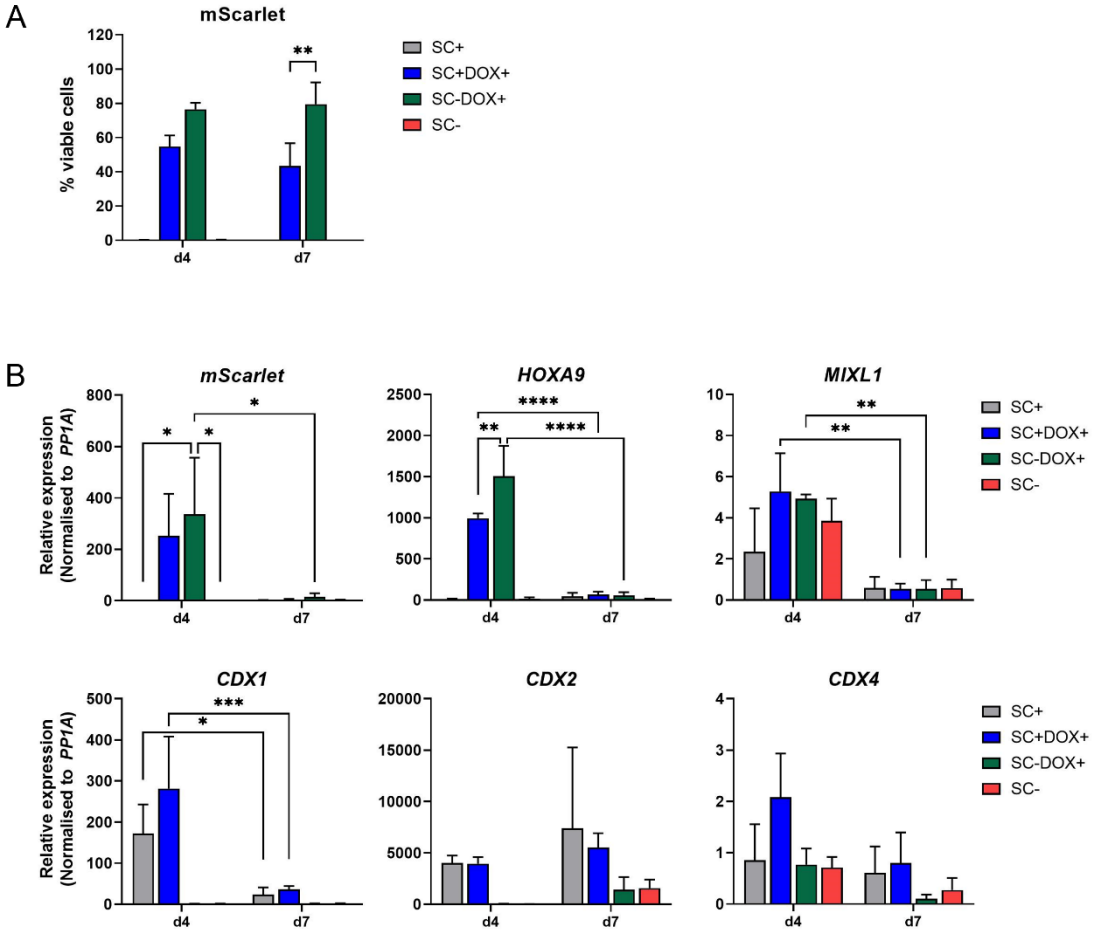


Figure 6.7: DOX induction at 40h post-seeding produces robust HOXA9 expression on d4 of differentiation

Flow cytometric analysis of mScarlet expression in d4 and d7 EBs from all treatment conditions. Results are mean ± SD, n = 3 biological replicates; *p < 0.05, 2-way ANOVA, Tukey's multiple comparisons test. (B) QPCR for *mScarlet*, *HOXA9,* mesendoderm marker *MIXL1* and CDX genes in d4 and d7 EBs from all conditions. Results are normalised to housekeeping gene *PP1A* and shown relative to undifferentiated iA9 iPSCs. Results are mean ± SD from 3 biological replicates with 3 technical replicates performed for each; *p < 0.05, **p < 0.01, ***p < 0.001, ****p < 0.0001, 2-way ANOVA, Tukey's multiple comparisons test.

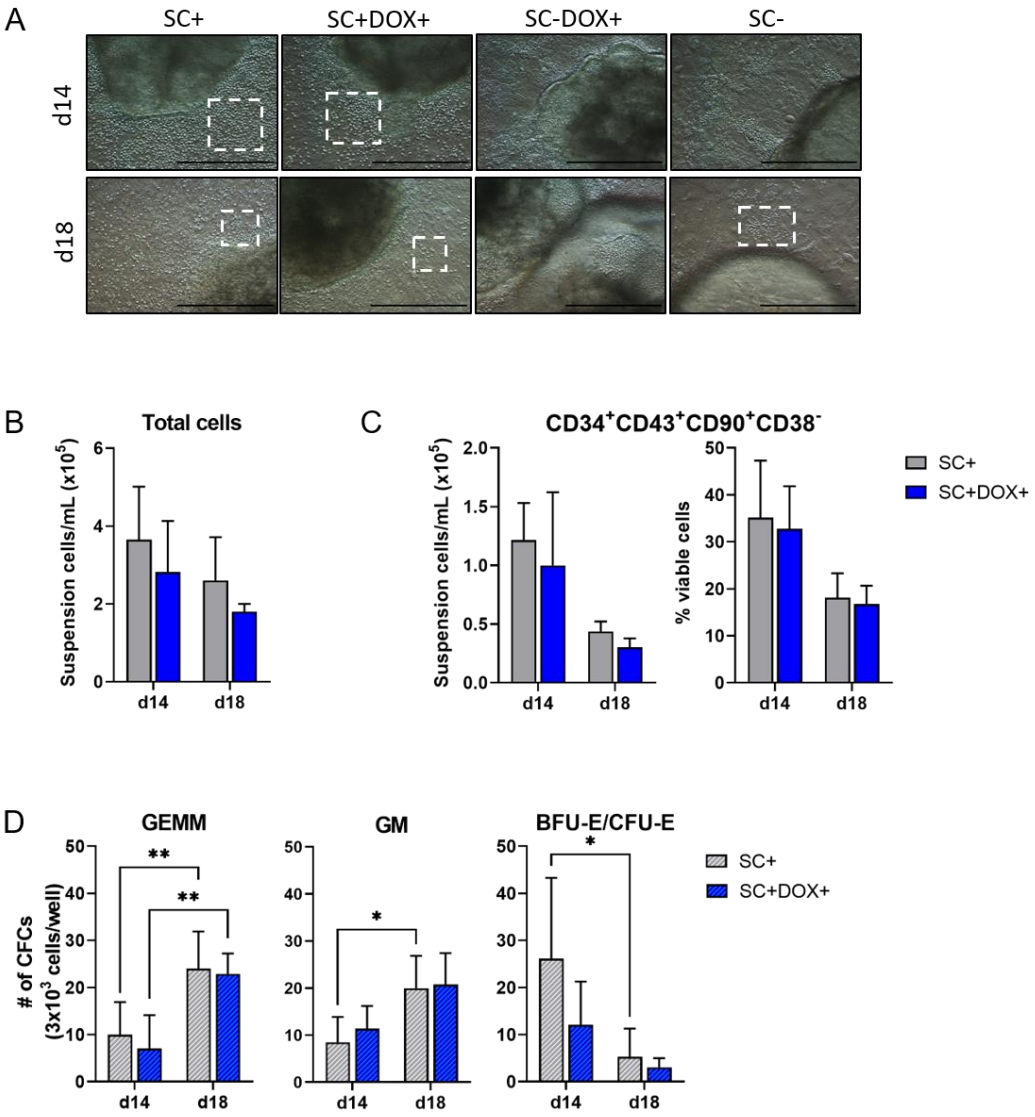


Figure 6.8: Transient HOXA9 overexpression at 40h post-seeding does not affect the formation or functionality of HSPCs in SB/CHIR treated cultures.

(A) Representative light microscopy images of all treatment conditions on d14 and d18. Dotted white boxes indicate clusters of progenitor cells, both still adhered to the underlying stromal layer and in suspension. Scale bar = 200 µm. (B) Total number of viable suspension cells (x10^5 cells/mL) harvested on d14 and d18 from SC+ and SC+DOX+ cultures. Viability was determined by Zombie NIR staining. (C) Flow cytometry analysis showing the percentage and absolute numbers of $CD34^+CD43^+CD90^+CD38^-$ haematopoietic progenitors generated from SC+ and SC+DOX+ d14 and d18 cultures. Results are mean ± SD, n = 3 biological replicates. (D) Haematopoietic colony-forming potential of suspension cells collected on day 14 and day 18 from SC+ and SC+DOX+ cultures. Colony formation was assessed after 14 days culture on MethoCult™ SF H4636. Results are mean ± SD from 3 biological replicates with 3 technical replicates performed for each; *$p < 0.05$, **$p < 0.01$, 2-way ANOVA, Tukey's multiple comparisons test.

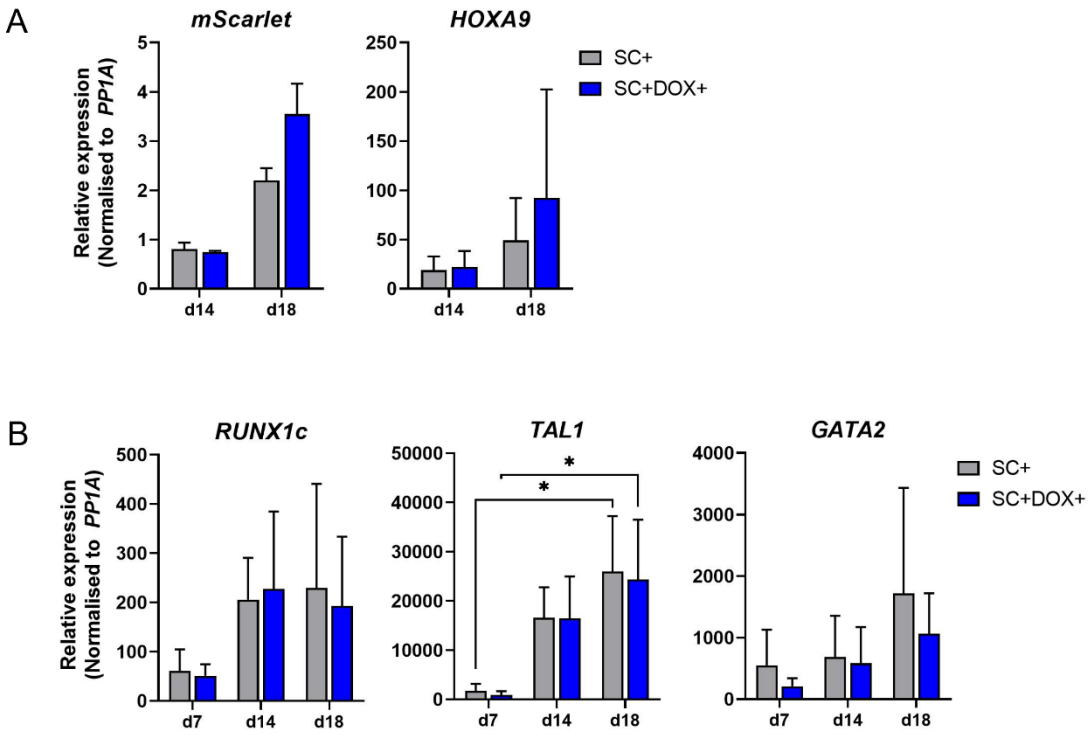


Figure 6.9: Similar gene expression profile observed in d14 and d18 SC+ and SC+DOX+ suspension cells

QPCR for *mScarlet* and *HOXA9* in d14 and d18 suspension cells harvested from SC+ and SC+DOX+ cultures. (B) QPCR for haematopoietic transcription factors in d7 EBs and d14 and d18 suspension cells harvested from SC+ and SC+DOX+ cultures. Results are normalised to housekeeping gene *PP1A* and shown relative to undifferentiated iA9 iPSCs. Results are mean ± SD from 3 biological replicates with 3 technical replicates performed for each; *p < 0.05, 2-way ANOVA, Tukey's multiple comparisons test.

On d14 and d18 of differentiation, suspension cells were harvested from SC+ and SC+DOX+ cultures for further analysis. SC- and SC-DOX+ cultures did not give rise to suspension cells. In some cultures, round, adherent, cobblestone-like cells morphologically similar to haematopoietic progenitors were evident, however, these cells remained attached to the underlying stroma and were not further assessed (Figure 6.8A).

The flow cytometry and qPCR data from d4 and d7 EBs suggest SC-DOX+ cultures possess an early haematopoietic defect similar to that observed in SC-, untreated cultures. SC- and SC-DOX+ culture conditions direct cells towards a mesoderm fate, as indicated by NCAM and *MIXL1* expression, but conditions do not support induction of a

vascular program as evidenced by lack of KDR expression in d4 EBs and the low proportion of $CD90^+CD34^+CD43^-$ cells on d7. The absence of suspension cells in SC- and SC-DOX+ cultures further supports this observation.

Suspension cells harvested from SC+ and SC+DOX+ cultures possessed a similar HSPC-like immunophenotype when assessed by flow cytometry. No difference was identified in the number and proportion of $CD34^+CD43^+CD90^+CD38^-$ HSPC-like cells collected from SC+ and SC+DOX+ cultures on d14 or d18 (Figure 6.8C). However, the d14 timepoint tended to yield a greater number and higher proportion of HSPC-like cells than the d18 timepoint in both conditions (Figure 6.8C). This was further reflected when the total number of suspension cells harvested at each timepoint was determined, although a significant difference in total cell number was not found between the two timepoints (Figure 6.8B). QPCR for key haematopoietic regulators *RUNX1c*, *TAL1* and *GATA2* did not identify differences in gene expression between d14 and d18 timepoints or SC+ and SC+DOX+ groups (Figure 6.9B). No significant difference was found in *HOXA9* expression between SC+ and SC+DOX+ groups at either timepoint, although there was a trend for higher *HOXA9* expression in d18 cells (Figure 6.9A). mScarlet expression was also measured. Although, a minute fold change was observed in d18 cultures, it was observed in both SC+ and SC+DOX+ groups and thus is unlikely to represent continued transcription of *HOXA9/mScarlet* induced at 40h post-seeding (SC+ fold change 2.1±0.25; SC+DOX+ fold change 3.5±0.6, Figure 6.9A).

Despite their similar HSPC-like immunophenotype and gene expression profile, d14 and d18 cells demonstrated statistically significant differences in haematopoietic potential in the CFU assay, as was similarly observed in Chapter 5. Suspension cells harvested on d14 had a more limited haematopoietic differentiation potential than d18 cells, predominately generating BFU-E/CFU-E colonies (Figure 6.8D). In contrast, d18 suspension cells had a broader differentiation capacity, producing greater numbers of CFU-GMs than d14 cells and significantly more immature erythro-myeloid CFU-GEMM colonies (Figure 6.8D). Together this data suggests a distinct functional difference between progenitors isolated at the two timepoints.

Comparison of SC+ and SC+DOX+ treatment groups revealed no apparent difference in colony formation ability or HSPC immunophenotype at either timepoint, suggesting transient overexpression of HOXA9 during early haematopoietic differentiation does not influence the functionality or immunophenotype of the resulting progenitor cells.

6.3.4 Transcriptome analysis of mesoderm populations

To better understand the mesoderm populations generated on d4 and to investigate possible transcriptome differences between d14 and d18 suspension cells, RNA sequencing was performed on these samples. Transcriptome analysis was performed on RNA isolated from all d4 treatment conditions and on RNA isolated from suspension cells collected from SC+ and SC+DOX+ cultures on d14 and d18. As SC- and SC-DOX+ cultures did not yield suspension cells, they were not included in d14 and d18 sequencing experiments. Principal component analysis (PCA) was performed to investigate the relationship between the different sample groups (A). Day 4 samples clustered away from d14 and d18 samples in two distinct groups, SB/CHIR+ (SC+ and SC+DOX+) and SB/CHIR- (SC- and SC-DOX+). Day 14 and d18 samples clustered based on timepoint not DOX treatment, consistent with the functional differences observed between d14 and d18 cells in the CFU assay. Unsupervised hierarchical clustering was performed and heatmaps revealed major expression differences between d4 samples and d14 and d18 samples, reflecting the transcriptional differences between mesodermal and progenitor cells (Supplementary Figure 2).

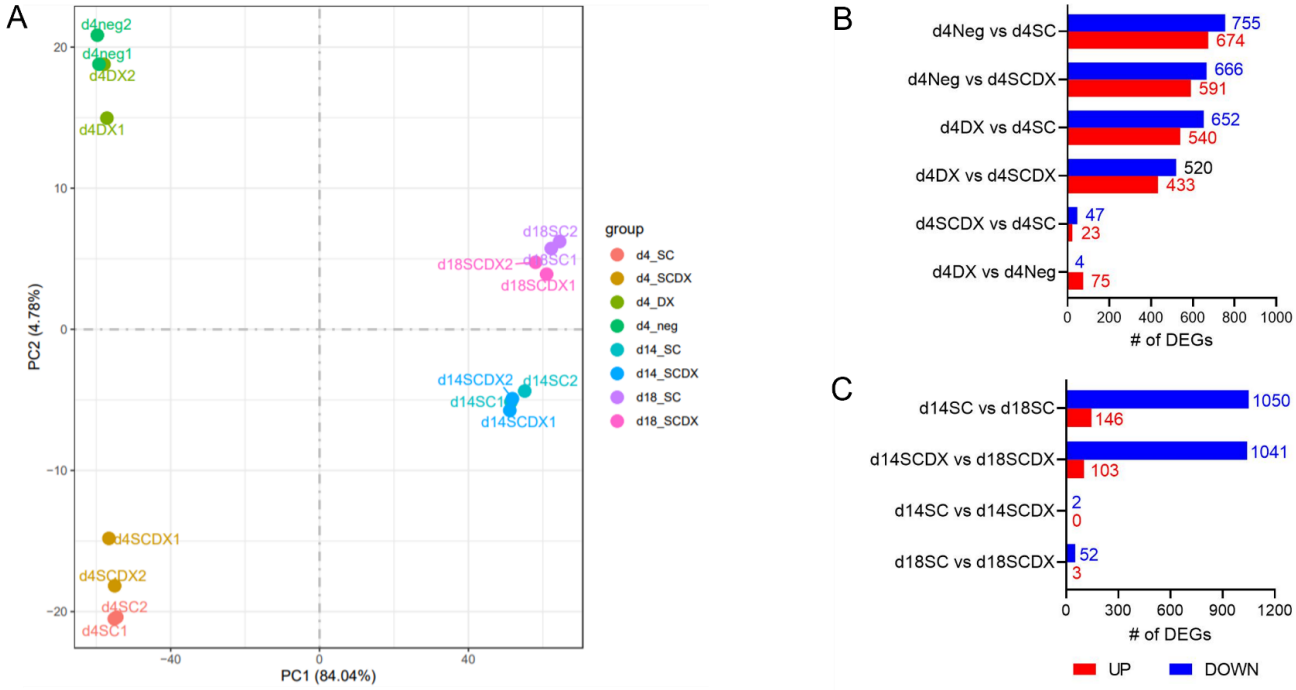


Figure 6.10: Relationship between d4, d14 and d18 cells under the different treatment conditions used

(A) PCA plot demonstrates the relationship between d4, d14 and d18 populations following treatment with SC+ (SC), SC+DOX+ (SCDX), SC-DOX+ (DX) or SC- (Neg) at 40h post-seeding. (B – C) The number of downregulated and upregulated differential expressed genes (DEGs) identified between comparative treatments on d4 (B) and d14 and d18 (C). DEGs were identified using DESeq2 and based on a threshold of $\log_2$FoldChange > 1.5 and padj < 0.05. Results are shown for 2 biological replicates.

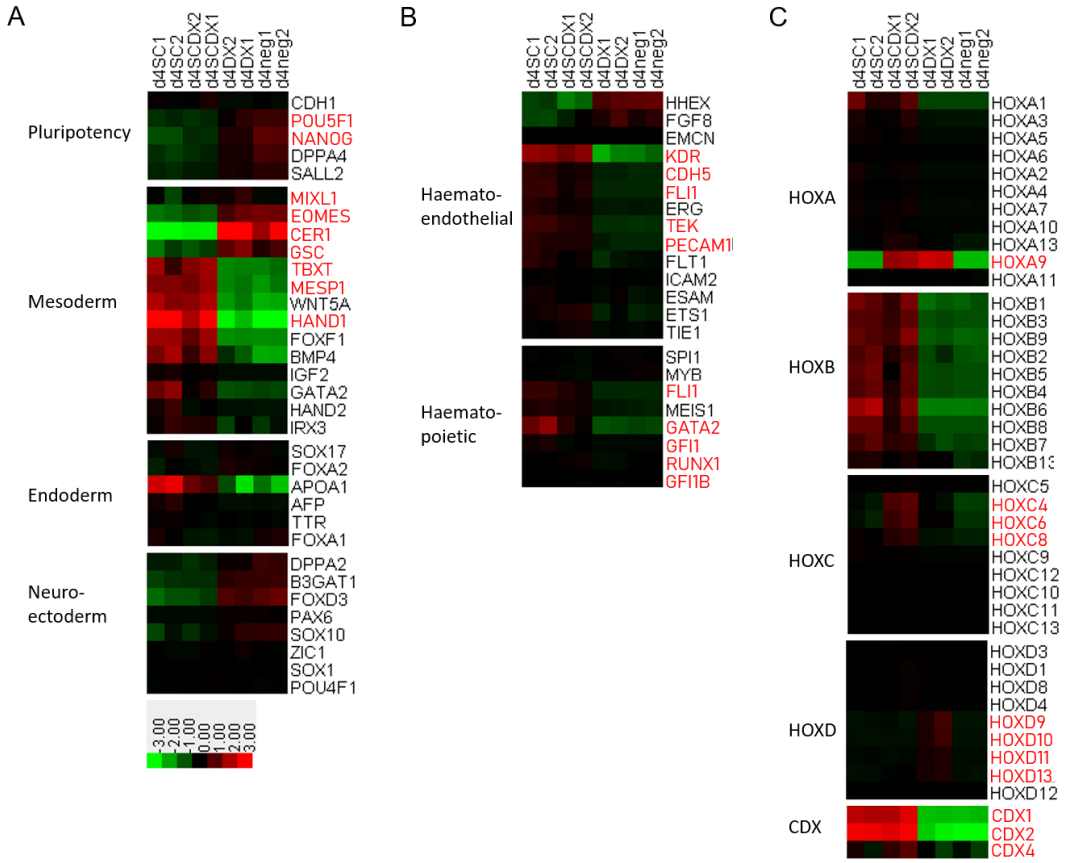


Figure 6.11: Gene expression changes induced in d4 mesoderm

Heat maps visualising the expression of genes associated with pluripotency and the three germ layers (A), haemato-endothelial and haematopoietic genes (B) and HOX cluster genes (C) in d4 populations treated with SC+, SC+DOX+, SC-DOX+ or SC- at 40h post-seeding. Genes highlighted in red are discussed further in the main text. Heat maps were made using Cluster 3.0 and TreeView. Gene expression level shown as normalised $\log_2$ FPKM and shown for two independent experiments.

To further characterise d4 samples, the expression of several selected genes important in early differentiation processes were examined. Cluster 3.0 was used to group d4 samples based on expression of genes related to pluripotency and germ layer differentiation and results were visualised as heat maps made using TreeView. Surprisingly, pluripotency genes, such as *NANOG* and *OCT4/POU5F1*, remained elevated in SC- and SC-DOX+ groups (Figure 6.11A). Several neuroectoderm markers were also highly expressed in these samples. In contrast, SC+ and SC+DOX+ cultures

predominately expressed genes related to mesoderm development, such as *HAND1*, *TBXT* and *MESP1*. Mesoderm associated genes upregulated in SC- and SC-DOX+ groups included *EOMES*, *CER1* and *GSC*, which are specifically associated with mesendoderm development. The expression of genes associated with the haemato-vascular and haematopoietic programs was also examined (Figure 6.11B). Endothelial markers *KDR*, *PECAM1*, *TEK* and *CDH5* were expressed in SB/CHIR-treated cultures yet absent in SC- and SC-DOX- samples. Similarly, expression of haematopoietic transcription factors was only observed in SB/CHIR-treated cultures. Interestingly, *FLI1* and *GATA2* appeared more highly expressed in SC+ samples than SC+DOX+, while *RUNX1* and *GFI1/GFI1B* were not expressed in either treatment condition. Lack of *RUNX1* and *GFI1/GFI1B* expression is consistent with their role in haematopoietic development; RUNX1 and GFI1/GFI1B function following mesoderm formation, in specification of haematopoietic fate and suppression of endothelial fate, while *FLI1* and *GATA2* are expressed earlier in development (Goode et al., 2016; Lancrin et al., 2012). The gene expression signatures observed correlated with the flow cytometric analysis performed on d4 samples; suggesting SB/CHIR-treated cultures on d4 represent mesoderm with haemato-vascular differentiation potential, while cultures not treated with SB/CHIR are more heterogenous and not restricted to mesoderm-specific lineages.

The expression of the 39 HOX genes and their regulators, the CDX genes (Deschamps and van Nes, 2005), were also examined in all conditions on d4 (Figure 6.11C). As expected *HOXA9* was highly expressed in SC+DOX+ and SC-DOX+ cultures. HOXB genes were also highly expressed in cultures treated with SB/CHIR as has been previously reported (Ng et al., 2016). In agreement with the d4 qPCR data, *CDX1* and *CDX2* were upregulated in cultures treated with SB/CHIR but not in SC- untreated or SC-DOX+ cultures. Additionally, *CDX4* expression appeared more elevated in SC+DOX+ cultures compared to the other treatment groups. A slight upregulation in *CDX4* was similarly observed in this treatment group in qPCR analysis. Transient overexpression of HOXA9 had a minimal impact on expression of the selected genes evaluated. *HOXA9* overexpression in SB/CHIR-treated cultures resulted in upregulation of *CDX4, HOXC4, HOXC6* and *HOXC8.* Surprisingly, HOXA9 overexpression in the absence of SB/CHIR did not stimulate the same genes, with *HOXD9, HOXD10, HOXD11* and *HOXD13* upregulated in SC-DOX+ cultures. This suggests HOXA9 has different activity when it is expressed in SB/CHIR-induced mesoderm with haematopoietic potential and mesoderm which lacks haematopoietic differentiation potential (untreated SC- mesoderm). Additionally, co-stimulation of SB/CHIR pathways impacts the HOXA9-induced transcriptome.

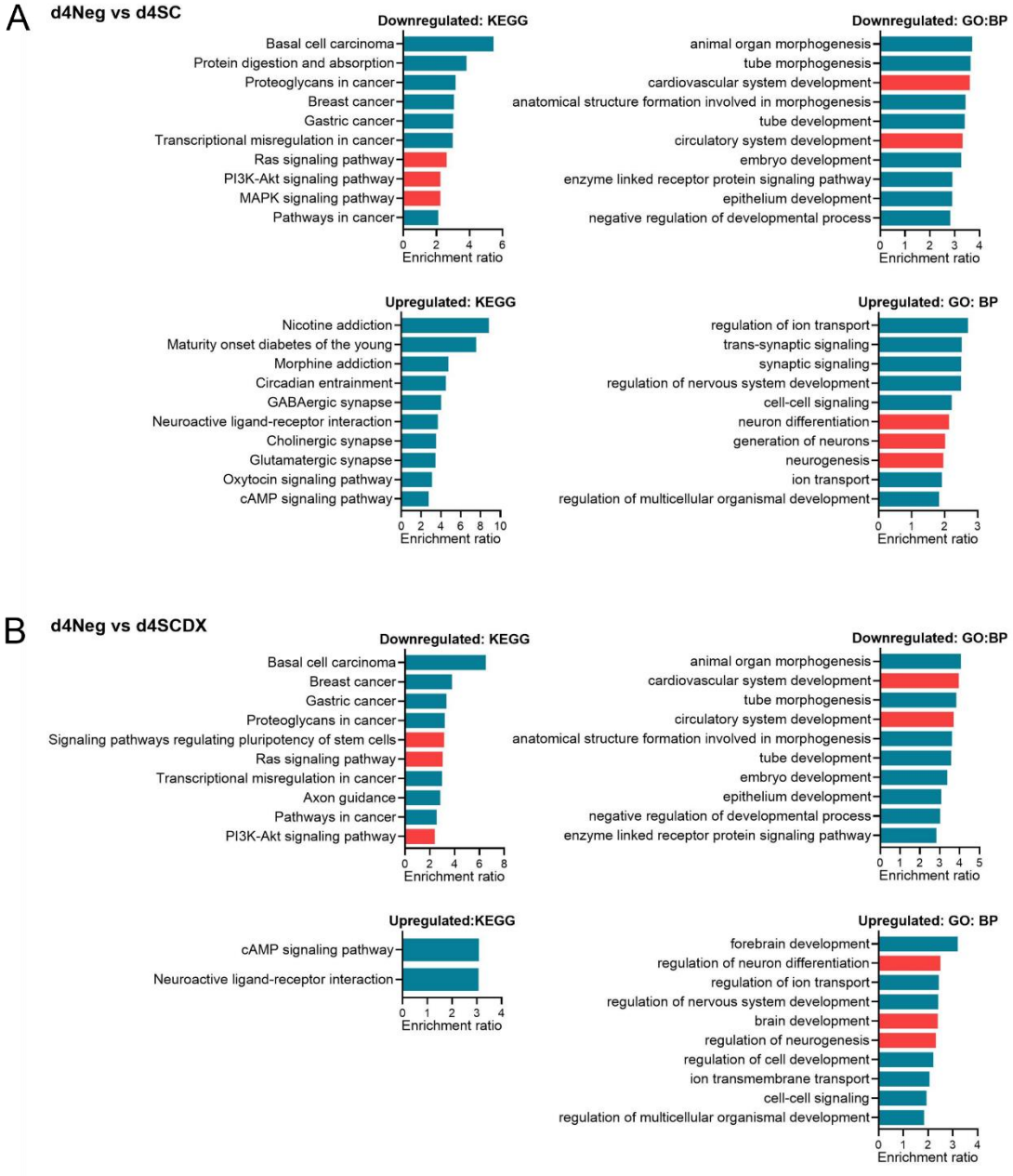


Figure 6.12: Pathway analysis of differentially expressed genes in d4 SC- cultures versus d4 SC+ and SC+DOX+ cultures

The top 10 enriched downregulated and upregulated KEGG and GO pathways in d4 SC- cultures (d4Neg) compared with (A) SC+ cultures (d4SC) and (B) SC+DOX+ cultures (d4SCDX). Pathway analysis was performed using WebGestalt on differential expressed genes with a $\log_2$FoldChange > 1.5 and padj < 0.05. Results represent two independent experiments. FDR < 0.05 in enriched pathways. Pathways of interest are highlighted in red.

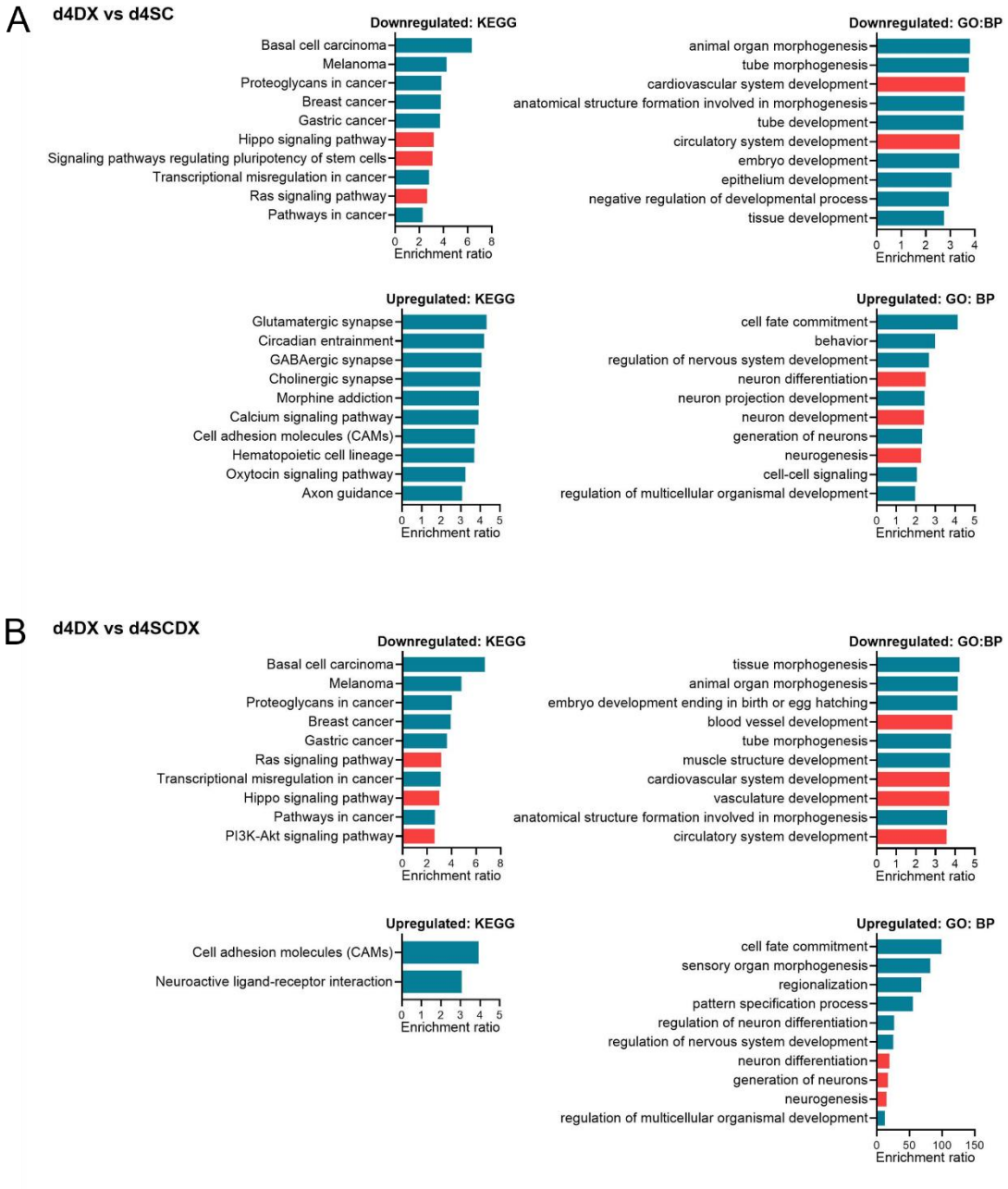


Figure 6.13: Pathway analysis of differentially expressed genes in d4 SC-DOX+ cultures versus d4 SC+ and SC+DOX+ cultures

The top 10 enriched downregulated and upregulated KEGG and GO pathways in d4 SC-DOX+ cultures (d4DX) versus (A) SC+ cultures (d4SC) and (B) SC+DOX+ cultures (d4SCDX). Pathway analysis was performed using WebGestalt on differential expressed genes with a $\log_2$FoldChange > 1.5 and padj < 0.05. Results represent two independent experiments. FDR < 0.05 in enriched pathways. Pathways of interest are highlighted in red.

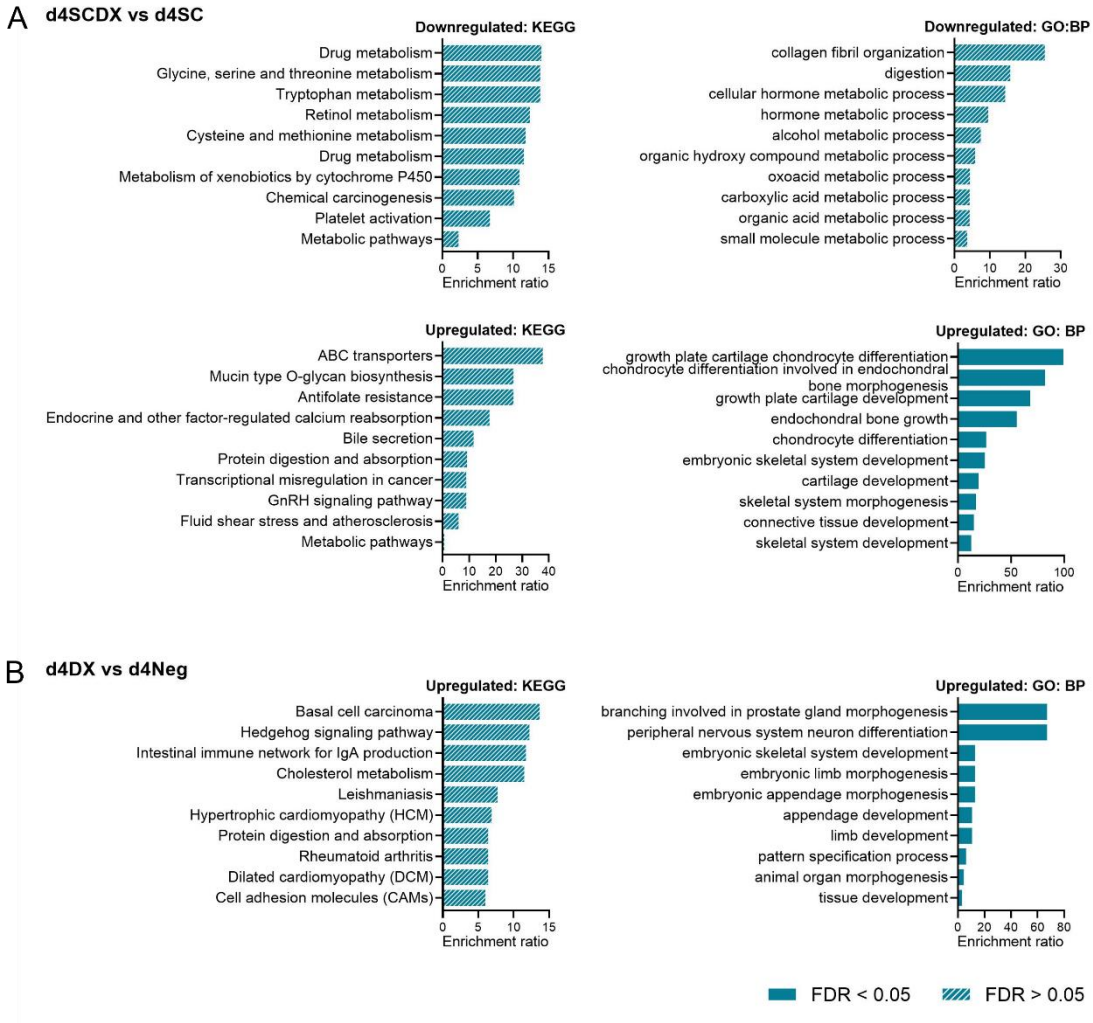


Figure 6.14: Pathway analysis of differentially expressed genes in +SB/CHIR and -SB/CHIR groups treated +/- DOX

(A) The top 10 enriched downregulated and upregulated KEGG and GO pathways in d4 SC+DOX+ cultures (d4SCDX) versus SC+ cultures (d4SC). (B) The top 10 enriched downregulated and upregulated KEGG pathways in d4 SC-DOX+ cultures (d4DX) versus SC- cultures (d4Neg). Pathway analysis was performed using WebGestalt on differential expressed genes with a $\log_2$FoldChange > 1.5 and padj < 0.05. Results represent two independent experiments. FDR < 0.05 indicates enriched pathways.

Differential gene expression analysis was performed using DESeq2 software on d4 samples from all treatment conditions. Differentially expressed genes (DEGs) were filtered based on a $\log_2$FoldChange > 1.5 and padj < 0.05 (see Supplementary Table 1 for total number of DEGs identified before and after application of threshold). Based on these criteria, few differentially expressed genes were identified between SC+ and SC+DOX+ groups and SC- and SC-DOX+ groups (B), in line with the PCA analysis (see Supplementary Table 2 for list of all DEGs found). However, when differential gene expression analysis was performed on -SB/CHIR groups (SC- or SC-DOX+) in comparison with +SB/CHIR groups (SC+ or SC+DOX+) 433 – 674 upregulated and 520 – 755 downregulated genes were identified (B). To further establish any interactions between the differentially expressed genes and possible biological roles, pathway enrichment analysis was performed using WebGestalt (http://www.webgestalt.org/). KEGG and GO pathway analysis was performed on DEGs, with the minimum number of genes required to define a category set to 5. BH multiple test adjustment was performed, and enriched categories were identified based on FDR < 0.05, unless otherwise indicated.

As was observed with PCA, SB/CHIR treatment was the main driver of pathway enrichment segregation. KEGG and GO analysis was performed on DEGs determined to be downregulated or upregulated in -SB/CHIR groups (SC- and SC-DOX+) compared to +SB/CHIR groups (SC+ and SC+DOX+). Pathway analysis was performed between each comparison; d4Neg vs d4SC (Figure 6.12A), d4Neg vs d4SCDX (Figure 6.12B), d4DX vs d4SC (Figure 6.13A) and d4DX vs d4SCDX (Figure 6.13B). Enriched pathways showed a high degree of similarity between -SB/CHIR groups and +SB/CHIR groups, regardless of whether cells had been treated with DOX. In -SB/CHIR groups, GO pathways related to neural differentiation and nervous system development were associated with upregulated DEGs, while pathways associated with cell development were enriched in downregulated DEGs, for example ‘tube morphogenesis’, ‘circulatory system development’ and ‘embryo development’. Similarly, KEGG pathways specific to brain function, such as ‘GABAergic synapse’ and ‘Neuroactive ligand receptor interaction’ were enriched in upregulated DEGs in -SB/CHIR groups, while downregulated DEGs in -SB/CHIR groups were associated with signalling pathways with roles in PSC maintenance and differentiation, such as ‘PI3K-Akt signaling pathway’ and ‘Ras signaling pathway’.

To decipher the impact of short-term HOXA9 expression, pathway analysis was also performed on DEGs enriched in SC+DOX+ samples compared with SC+ (d4SCDX vs

d4SC) and SC-DOX+ samples compared with SC- untreated samples (d4DX vs d4Neg). In both comparisons only GO pathways were determined to be significantly enriched, based on a cut-off of FDR < 0.05. DEGs identified amongst SC+DOX+ and SC+ samples were upregulated in pathways associated with mesoderm differentiation to non-haematopoietic and non-endothelial lineages, such as chondrocyte and skeletal development (Figure 6.14A). Upregulated DEGs in SC-DOX+ samples were enriched in pathways associated with limb development and neuron differentiation (Figure 6.14B).

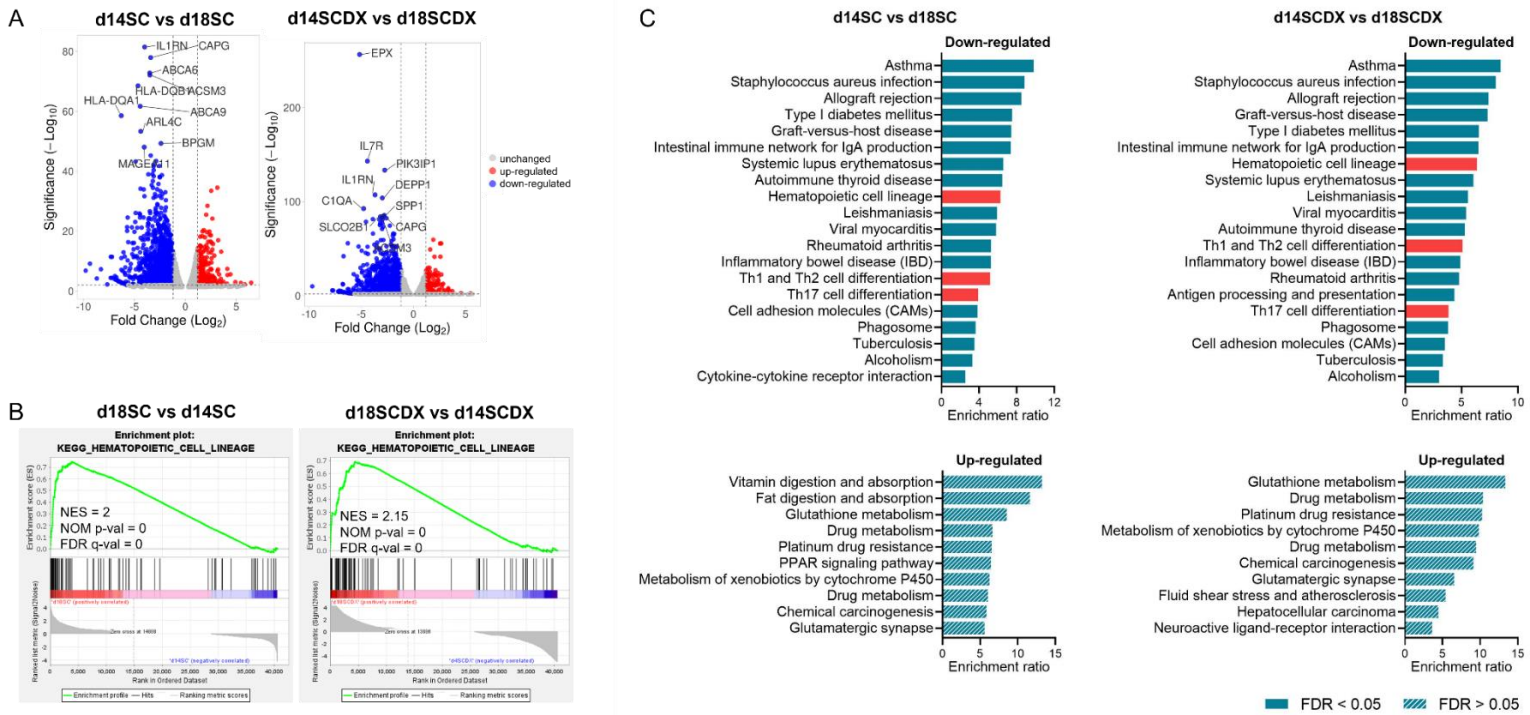


Figure 6.15: GSEA and KEGG pathway analysis of d14 and d18 suspension cells

(A) Volcano plots generated using VolcaNoseR showing all DEGs in d14 versus d18 SC+ and SC+DOX+ populations. Upregulated genes (red) and downregulated genes (blue) are indicated based on a threshold of log_2FoldChange > 1.5 and p-value < 0.05. The top 10 most differentially expressed genes are labelled. (B) GSEA enrichment plot of KEGG pathway 'Hematopoietic cell lineage' in d18 and d14 SC+ and SC+DOX+ populations. (C) Enriched KEGG pathways among differentially expressed genes in comparison of d14 and d18 SC+ cells (d14SC vs d18SC) and d14 and d18 SC+DOX+ cells (d14SCDX vs d18SCDX). Results represent two independent experiments. FDR < 0.05 in enriched pathways. Pathways of interest are highlighted in red.

6.3.5 Transcriptome analysis of haematopoietic populations

Suspension cells generated on d14 and d18 were functionally distinct. To identify if transcriptome differences were also apparent between the two suspension cell populations, RNA-seq analysis was performed. PCA and differential gene expression analysis indicated d14 and d18 cells have unique transcriptome profiles. Under the criteria previously outlined, a total of 1196 and 1144 differential expressed transcripts were identified in a comparison of d14 and d18 cells from SC+ and SC+DOX+ groups, respectively (C, Figure 6.15A; see Supplementary Table 3 for the number of DEGs identified before and after application of threshold). The majority of these transcripts were downregulated in d14 cells compared with d18, with 1050 and 1041 downregulated DEGs identified in SC+ and SC+DOX+ groups, respectively. In comparison with d18 cells, KEGG pathway enrichment analysis revealed d14 cells were significantly downregulated in pathways related to the hematopoietic cell lineage and Th1, Th2 and Th17 cell differentiation (Figure 6.15C). Several enriched pathways related to metabolism were identified in DEGs upregulated in d14 cells in comparison with d18, however, these pathways were not determined to be significantly enriched based on FDR < 0.05. Gene set enrichment analysis (GSEA) further indicated d18 cells were enriched in genes associated with the haematopoietic cell lineage (Figure 6.15B).

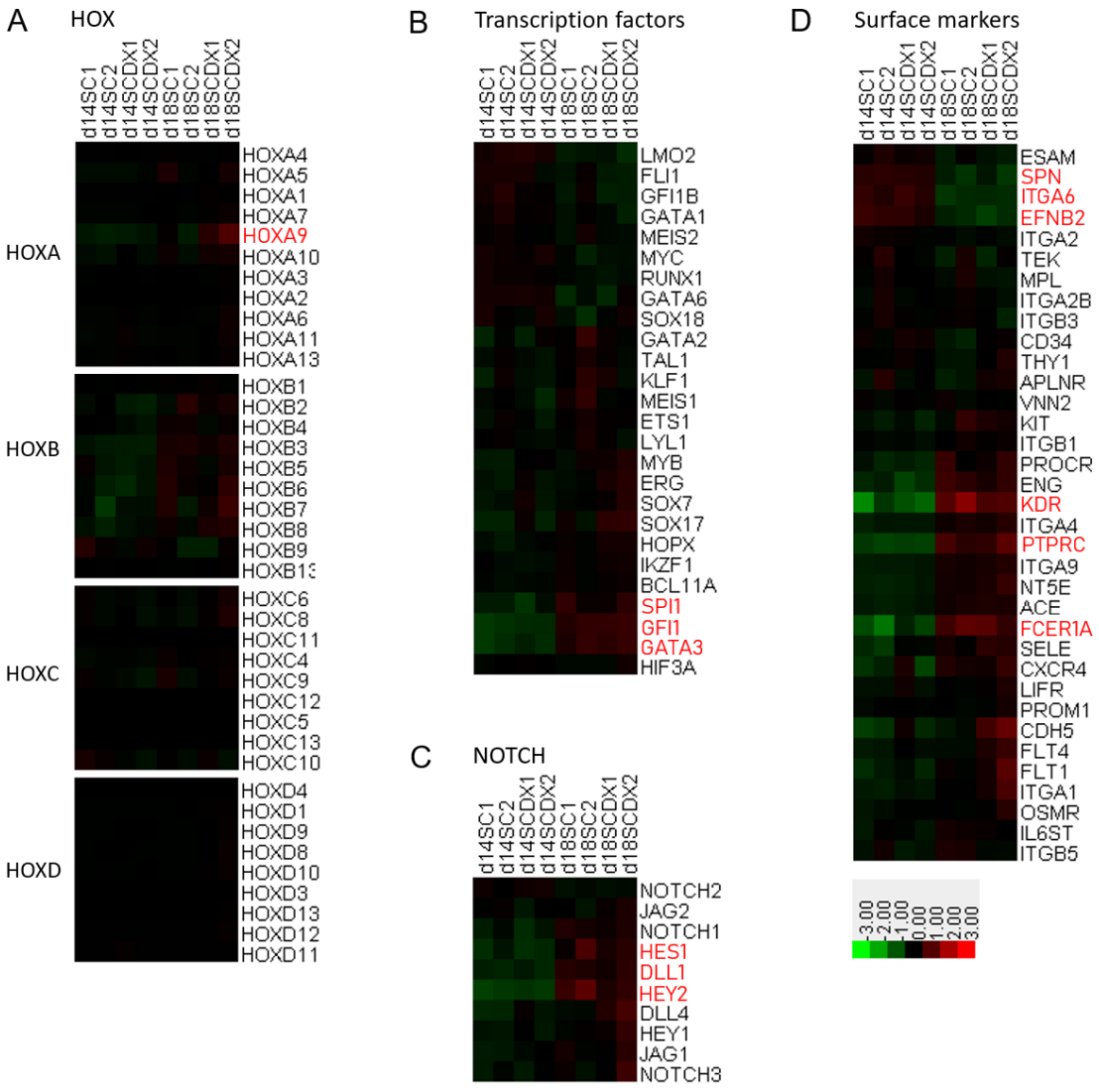


Figure 6.16: Gene expression of key haematopoietic transcription factors and surface markers in d14 and d18 suspension cells

Heat maps visualising the expression of HOX cluster genes (A), haematopoietic transcription factors (B), haematopoietic cell surface markers (C) and Notch signalling genes (D) in d14 and d18 suspension cells from SC+ and SC+DOX+ groups. Genes highlighted in red are discussed further in the text. Heat maps were made using Cluster 3.0 and TreeView. Gene expression level shown as normalised $\log_2$ FPKM and shown for two independent experiments.

The expression of key haematopoietic transcription factors and cell surface markers in the RNA-seq analysis of d14 and d18 populations was examined. Heat maps were made using Cluster 3.0 and visualised with TreeView. Distinct gene expression profiles were observed, with timepoint rather than DOX treatment driving the expression differences, in line with PCA analysis. Cell surface markers associated with endothelial cells were upregulated in both populations, although the expression pattern varied. High expression of *SPN/CD43*, *EFNB2* and *ITGA6* were found in d14 populations, whereas some of the most highly expressed transcripts in d18 populations included *KDR*, *PTPRC/CD45* and *FCER1A* (Figure 6.16D). HE and haematopoietic progenitors share expression of several endothelial (including *KDR, EFN2B* and *LMO2*) and HE genes (such as *ESAM*, *CDH5, SOX17* and *KIT*) (Angelos et al., 2018) which is reflected in the d14 and d18 populations. Similar to the expression profile of haematopoietic cell surface markers, transcription factor expression differed between d14 and d18 populations. Key haematopoietic regulators *FLI1, GFIIB* and *GATA1* appeared more expressed in d14 cells, whereas *SPI1, GFI1* and *GATA3* were more highly expressed in d18 cells (Figure 6.16B). GFI1 and GFI1B both function in repression of endothelial fate during EHT (Thambyrajah et al., 2016) and are expressed in HSCs (Osawa et al., 2002; Zeng et al., 2004). Gfi1b has additional functions in erythropoiesis (Osawa et al., 2002), while Gfi1 is implicated in the generation of B lymphocytes and granulocytes in the mouse bone marrow (Spooner et al., 2009; Zhang et al., 2019). Similarly, GATA1 functions in erythroid and megakaryocyte differentiation and GATA3 is required for T cell differentiation (Gao et al., 2015). Together, this expression pattern suggests d14 and d18 suspension cells are two distinct cell types, with d14 and d18 suspension cells primed for erythroid and lymphoid differentiation, respectively. In support of this distinction, Notch signalling was found to be unique to d18 cells, with Notch ligand *DLL1* and downstream targets *HES1* and *HEY2* expressed in d18 populations (Figure 6.16C).

The HOXA cluster was not expressed in either d14 or d18 populations (Figure 6.16A), consistent with the well-established *HOXA*- profile of hPSC-derived progenitors (Dou et al., 2016; Fidanza et al., 2020; Ivanovs et al., 2017; Ng et al., 2016). Several HOXB genes appeared upregulated in d18 populations, an observation which has previously been described in an endothelial-enriched $CD34^{+}CD90^{+}$ fraction of hPSC-derived haematopoietic progenitors (Motazedian et al., 2020).

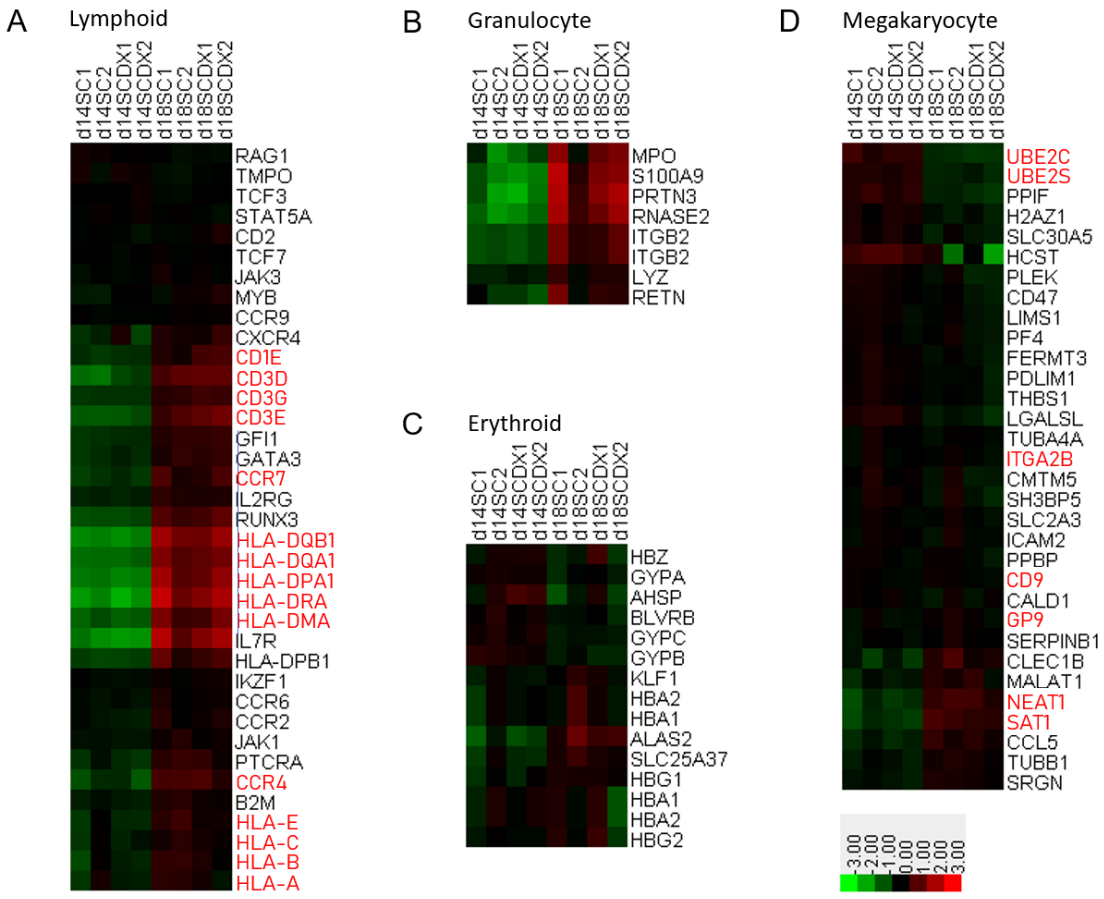


Figure 6.17: Expression of lineage-specific genes in d14 and d18 suspension cells

Heat maps visualising the expression of lymphoid- (A), granulocyte- (B), erythroid- (C) and megakaryocyte- (D) associated genes in d14 and d18 suspension cells from SC+ and SC+DOX+ groups. Genes highlighted in red are discussed further in the text. Heat maps were made using Cluster 3.0 and TreeView. Gene expression level shown as normalised $\log_2$ FPKM and shown for two independent experiments.

Haematopoietic lineage commitment is proposed to occur prior to EHT (Guibentif et al., 2017). As such we investigated the presence of haematopoietic lineage-committed progenitors in the d14 and d18 populations by assessing the expression pattern of genes connected to lymphoid- (Motazedian et al., 2020), erythroid-, megakaryocyte- and granulocyte-committed cells (Fidanza et al., 2020) in our RNA-seq analysis. Strikingly, lymphoid- and granulocyte-related genes were highly expressed in d18 populations (Figure 6.17A, B). In particular, MHC class I (*HLA-A, HLA-B, HLA-C* and *HLA-E*) and MHC class II molecules (*HLA-DRA, HLA-DMA, HLA-DQA1* and *HLA-DQB1*) were upregulated in d18 populations. MHC class II molecules are expressed in the human

paediatric thymus (Collado et al., 2013) and mouse MHC class I molecules are purported to have roles in maturation of HSCs in the AGM (Kieusseian et al., 2012). Several T cell receptor subunits (*CD1E, CD3D, CD3G* and *CD3E*) were also upregulated in d18 populations, along with genes related to chemokine signalling (*CCR7* and *CCR4*) (Figure 6.17A). Chemokine signalling is active in T cell development (Lancaster et al., 2018) and CCR7 and CCR4 are both implicated in T cell migration and maturation in the thymus (Cowan et al., 2014).

The expression patterns of erythroid and megakaryocyte associated genes was also examined. The megakaryocyte gene profile investigated had previously been identified in single cell RNA-seq analysis of hPSC-derived $CD34^{+}CD235a^{-}$ haematopoietic progenitors, whereby three distinct clusters of megakaryocyte-committed progenitors were identified (Fidanza et al., 2020). While low expression of key megakaryocyte genes (*GP9, PF4, ITGA2B* and *CD9*) was found in both d14 and d18 populations, genes related to two distinct megakaryocyte clusters were upregulated uniquely in d14 cells or d18 cells. *UBE2C* and *UBE2S* identified in one cluster were upregulated in d14 cells, while *NEAT1* and *SAT1*, found in a different cluster, were expressed in d18 cells (Figure 6.17D), suggesting d14 and d18 population possess different megakaryocyte potential. The erythroid expressing profile was highly variable, with different expression observed for specific genes amongst biological replicates (Figure 6.17C).

Table 6.1: DEGs in d14 and d18 cultures treated +/-DOX at 40h

Samples	Expression	DE genes (log_2FoldChange > 1.5; padj < 0.05)
d14SC vs d14SCDX	DOWN	NSFP1, RGPD6
	UP	-
d18SC vs d18SCDX	DOWN	AMBP, APOA4, FGG, FG1, AHSG, BCHE, APOB, PRND, EGFLAM, AFP, PLG, RBP4, CHST1. APOA1, APOC3, FGB, ALB, TACR1, HPX, VTN, APOA2, S100A14, SERPINA1, BGN, IL33, CXorf36, SLIT3, PCDH1, ADAMTS9, ADGRL4, ROBO4, RPL29P11, RAMP2, FBLIM1, HOXA9, FBLN1, CLEC14A, DLL4, IGDCC3, CLDN5, ITGA1, EDN1, GPX8, ADGRF5, COL15A1, HTRA1, RHOJ, MCAM, FLT4, RGS5, CDH5
	UP	XAF1, OAS2, FAM157B

Although, suspension cells harvested on d14 and d18 from DOX-treated and untreated cultures were shown to possess a similar immunophenotype and functional activity in flow cytometry and CFU assays (Figure 6.8), DEG analysis was performed between the two groups to determine if transcriptional differences were evident. Based on the previously outlined DEG criteria, only 2 DEGs were identified in comparison of d14 SC+ and SC+DOX+ treated cells, while 55 DEGs were identified in comparison of the treatment groups in d18 cells (C), with 52 genes found to be significantly downregulated in d18 SC+ cells compared to d18 SC+DOX+ cells (see Table 6.1 for list of DEGs). Significantly downregulated DEGs included *HOXA9*, *DLL4* and haematopoietic surface markers (*CDH5, FLT4*, and *ITGA1*), indicative of their upregulation in d18 SC+DOX+ groups (Figure 6.16C). KEGG and GO pathway analysis was performed on the 52 downregulated genes (Figure 6.18). Enriched GO pathways included pathways related to 'blood vessel development', 'angiogenesis' and the extracellular matrix, reflecting the identification of endothelial genes, extracellular matrix components (*VTN, COL15A1*) and genes which regulate blood vessel formation (*ROBO4, ITGA1*) in downregulated DEGs (Dai et al., 2019; Yamamoto et al., 2015). GO pathway 'chylomicron remodeling' and KEGG pathway 'Cholesterol metabolism' were both enriched, likely in response to identification of several differentially expressed apolipoprotein genes, which function in cholesterol metabolism (Oguro, 2019).

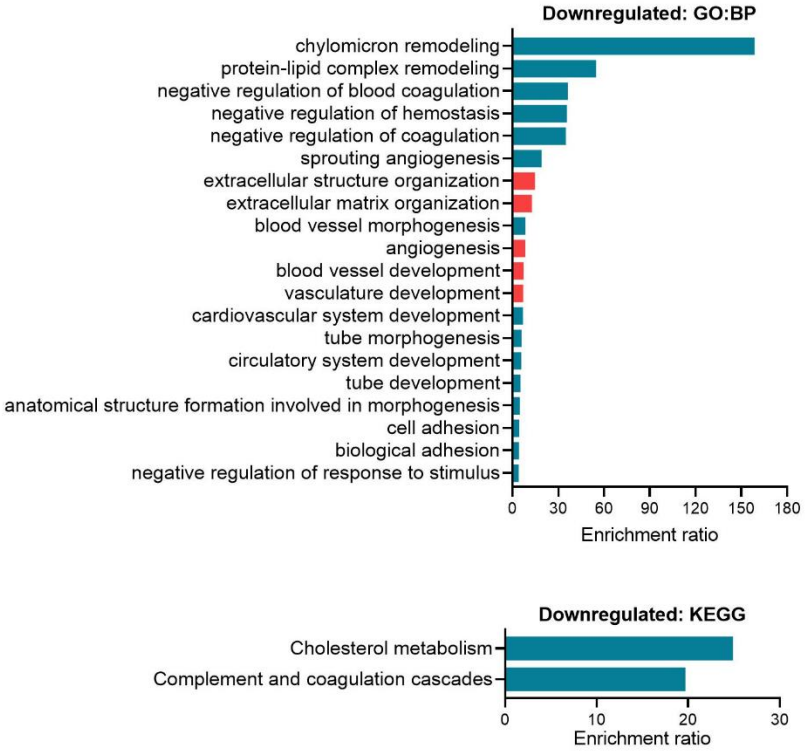


Figure 6.18: Pathways associated with downregulated genes in d18SC+ cells versus d18 SC+DOX+ cells

KEGG and GO pathways enriched in downregulated differentially expressed genes identified between d18 SC+ and d18 SC+DOX+ cells. Pathway analysis was performed using WebGestalt on differential expressed genes with a $\log_2$FoldChange > 1.5 and padj < 0.05. Results represent two independent experiments. FDR < 0.05 in enriched pathways. Pathways of interest are highlighted in red.

6.4 Discussion

In this chapter, the impact of transient HOXA9 overexpression during mesoderm patterning on mesoderm and haematopoietic progenitor formation was investigated. The spin EB haematopoietic differentiation model requires modulation of Activin and Wnt signalling, performed using small molecules SB and CHIR, to generate haematopoietic progenitors. Overexpression of HOXA9 in the absence of SB/CHIR was insufficient to direct cells towards a haemato-vascular fate and generate haematopoietic progenitors. In the presence of SB/CHIR, transient HOXA9 overexpression promoted haematopoietic development, similar to uninduced SB/CHIR-treated cultures. RNA-seq analysis revealed minor transcriptome differences between suspension cells harvested on d18 from SC+ and SC+DOX+ cultures. However, the transcriptional network and

haematopoietic potential of suspension cells generated on d14 and d18 showed great diversification.

6.4.1 Untreated cultures have impaired mesoderm differentiation compared with SB/CHIR-treated cultures

The impact of transient HOXA9 overexpression on mesoderm induction in an efficient (SC+) and defective (SC-) model of haematopoiesis was first assessed. Overexpression of HOXA9 in SB/CHIR-treated or untreated cultures at 40h post-seeding resulted in a >1000-fold increase in *HOXA9* expression on d4, relative to uninduced iPSCs. Despite this high HOXA9 expression level, DOX-induced and uninduced cultures showed minimal differences in d4 flow cytometry experiments and transcriptome analysis. SB/CHIR treatment was the main driver of diversification between the different d4 conditions analysed. No significant difference in the ability of SC+ and SC+DOX+ cultures to generate $NCAM^{+}KDR^{+}$ mesoderm on d4 or $CD90^{+}CD34^{+}CD43^{-}$ haemato-vascular mesoderm on d7 was observed, while SC- and SC-DOX+ cultures both lacked the ability to generate either population. RNA-seq analysis revealed similar patterns. SB/CHIR treatment was shown to influence sample grouping in PCA plots and pathway enrichment in KEGG and GO pathway analysis. In contrast to –SB/CHIR groups, SB/CHIR-treated groups were enriched in pathways related to pluripotency and differentiation. KEGG pathways related to Ras/MAPK signalling, PI3K-Akt signalling, Hippo signalling and pathways regulating PSCs were upregulated in both SC+ and SC+DOX+ groups in comparison with SC- and SC-DOX+. These signaling pathway play essential roles in regulating PSCs and their differentiation (Singh et al., 2012). In PSCs, PI3K signalling promotes sustained self-renewal, while MAPK signalling promotes differentiation and stimulates canonical Wnt signalling (Singh et al., 2012). ERK/MAPK signalling is also essential for mesendoderm induction (Yiangou et al., 2019) and at later stages of hPSC-based differentiation promotes the specification of arterial HE (Park et al., 2018). Wnt signaling also appears to be upregulated in +SB/CHIR groups. The KEGG pathway ‘Signalling pathways regulating PSC’ was enriched in +SB/CHIR groups and predominately comprises of Wnt family genes, including Wnt3A, Wnt5A and Wnt11, which are known to regulate and promote mesoderm and haematopoietic development (Gertow et al., 2013; Vijayaragavan et al., 2009). As such, the upregulation of pathways associated with PSC and differentiation provides evidence that the transcriptome of SB/CHIR-treated cells are primed for differentiation.

Interestingly, upregulated GO pathways in -SB/CHIR groups compared with +SB/CHIR groups were predominately associated with neural differentiation and neurogenesis,

reflecting the upregulation in neuroectoderm genes in cluster analysis. Gene expression patterns examined in RNA-seq analysis of d4 populations revealed that while several pluripotency markers remained elevated in -SB/CHIR groups, mesendoderm-specific markers (*MIXL1, EOMES* and *CSR1*) were also upregulated. Enriched downregulated KEGG pathways in –SB/CHIR groups versus +SB/CHIR were associated with morphogenesis and structural changes, including development of cardiovascular, circulatory, and vascular systems. Thus, SC+ and SC+DOX+ cultures appear to have greater propensity to differentiate to mesoderm-specific lineages than -SB/CHIR groups. This observation is consistent with flow cytometry data, whereby -SB/CHIR groups can generate NCAM+ mesoderm but cannot produce NCAM+KDR+ haemogenic mesoderm.

6.4.2 Transient HOXA9 overexpression does not specifically promote haemogenic mesoderm formation

Overexpression of HOXA9 either in the presence or absence of SB/CHIR did not elicit a strong response in comparison with uninduced cultures as assessed by flow cytometry, qPCR and RNA-seq. However, some transcriptome differences were observed, most notably in the HOX expression profile. Overexpression of HOXA9 in +SB/CHIR cultures led to upregulation of several HOXC genes, while HOXA9 overexpression in SC- cultures resulted in upregulation of HOXD genes. Pathway analysis revealed GO terms related to bone and cartilage development and GO terms relating to limb development were upregulated in SC+DOX+ and SC-DOX+ groups, respectively, in comparison with their uninduced counterpart. This pathway enrichment reflects the role of HOXC genes in cartilage development (Kruger and Kappen, 2010) and HOXD genes in regulating limb development during embryogenesis (Zakany and Duboule, 2007; Zákány et al., 2004). Furthermore, significantly higher *HOXA9* expression was detected by qPCR in SC-DOX+ cultures than SC+DOX+, despite cultures been treated for the same duration and with the same DOX concentration. This suggests the cell context in which HOXA9 is expressed influences both the downstream targets of HOXA9 and the level of HOXA9 induced.

All d4 conditions generated a NCAM+ population, a population which has previously been shown to possess wide mesoderm lineage differentiation capabilities; being able to generate cardiac, haematopoietic, endothelial, osteogenic, chondrogenic and smooth muscle cell types (Evseenko et al., 2010). Continuous overexpression of HOXA9 has been shown to promote haematopoietic progenitor formation from hPSCs (Ramos-Mejía et al., 2014; Zeng et al., 2021). Here, we have shown short-term HOXA9 overexpression during mesoderm induction is not sufficient to direct haematopoietic differentiation.

SB/CHIR treatment is required to generate $NCAM^+KDR^+$ mesoderm with haematopoietic potential and generation of this population is not influenced by HOXA9 overexpression. As pathway analysis comparisons between d4 DOX-induced and uninduced cultures has shown, HOXA9 overexpression promotes enrichment of diverse mesoderm pathways, including pathways associated with cartilage, limb and skeletal development. This potentially broad activity of HOXA9 may hinder haemogenic mesoderm development as early haematopoietic transcription factors *GATA2* and *FLI1* expression were observed to be more highly expressed in SC+ than SC+DOX+ cultures. Thus, HOXA9 overexpression during early haematopoietic differentiation appears to promote mesoderm formation but it does not specifically act to promote haemogenic mesoderm formation.

6.4.3 Haematopoietic progenitors generated on d18 possess a gene signature and functional activity similar to extra-embryonic EMPs

Transcriptome analysis and the CFU assay demonstrated suspension cells generated on d14 and d18 of spin EB haematopoietic differentiation are molecularly and functionally distinct. KEGG pathway analysis indicated d18 populations, in contrast with d14, were upregulated in genes related to the haematopoietic cell lineage and helper T cell development. Investigation of gene expression differences revealed d18 populations highly expressed MHC class I and II molecules, T cell receptor subunits and chemokine receptor signalling components; all associated with T cell development in the thymus (Motazedian et al., 2020). Notch signalling is required for T cell differentiation (Bertrand et al., 2010; Hadland et al., 2004) and was also upregulated specifically in d18 populations. While this data suggests lymphoid potential is unique to d18 suspension cells, a caveat of this research is functional assessment of lymphoid potential, for example via co-culture of suspension cells with OP9-DLL4 (Mohtashami et al., 2010), was not performed. As such, lymphoid potential of d18 suspension cells cannot be accurately deduced.

Assessment of lymphoid potential is an important measure of haematopoietic activity as it can be used to determine if differentiating cultures recapitulate primitive or definitive haematopoiesis. Lymphoid potential was originally proposed to be unique to intra-embryonic definitive haematopoiesis (Kennedy et al., 2012), however, it is now accepted both extra-embryonic and intra-embryonic definitive programs possess this potential (Fidanza et al., 2020; Motazedian et al., 2020). However, the two definitive programs can be distinguished based on the functionality of the resulting progenitors (Dzierzak and Bigas, 2018). For example, in the embryo, natural killer (NK) cells are generated in both

waves of definitive haematopoiesis. This can be recapitulated using *in vitro* models of haematopoiesis, whereby hPSC-derived $HOXA^{-/low}CD34^{+}$ progenitors were observed to generate $CD56^{+}CD45^{+}$ NK cells but not $CD4^{+}CD8^{+}$ T cells in OP9-DLL4 co-culture. In contrast, $HOXA^{+}CD34^{+}$ progenitors first generated $CD4^{+}CD8^{+}$ T cells followed by NK cells, reflecting the different developmental waves observed *in vivo* (Dege et al., 2020). As such, differentiation to specific T lymphoid lineages, such as NK cells, can determine if hPSC-derived progenitors have definitive haematopoietic potential and, more specifically, if their phenotype resembles extra- or intra-embryonic progenitors.

Notch signaling and *HOXA* expression can also be used to distinguish between extra- and intra-embryonic haematopoietic programs. HSCs are specified from arterial endothelium in the dorsal aorta where Notch signalling is active (North et al., 2002). Notch inhibition and OP9-DLL4 co-culture assays have shown hPSC-derived arterial-type $CXCR4^{+}DLL4^{+}$ HE requires Notch signalling to specify haematopoietic progenitors with erythro-myeloid and lymphoid potential. In comparison, Notch-independent $CXCR4^{+/-}DLL4^{-}$ HE has reduced T cell differentiation propensity and CFU activity, reflecting the restricted differentiation potential of extra-embryonic progenitors (Jung et al., 2021; Uenishi et al., 2018). Thus, formation of Notch-dependent HE can be used as indicator of induction of the intra-embryonic program.

Upregulation of HOXA genes is unique to the intra-embryonic program (Dege et al., 2020; Ivanovs et al., 2017; Ng et al., 2016). In the mouse, HOXA genes were found to be specifically upregulated in arterial HE versus yolk sac HE (Gao et al., 2018). Additionally, EMPs derived from E9.5 mouse embryos lacked *HOXA* expression while common lymphoid progenitors of the adult mouse bone marrow expressed HOXA genes (Dege et al., 2020). Here, the expression of *HOXA9* was monitored by qPCR at different timepoints during differentiation. The whole HOXA cluster was not monitored by qPCR, however, RNA-seq analysis on d4, d14 and d18 suggests the HOXA cluster was not upregulated at the given timepoints. In qPCR analysis, no significant difference in *HOXA9* expression was observed between d4 and d7 SC+ cultures. While, in DOX treated groups, SC+DOX+ and SC-DOX+, *HOXA9* was significantly more highly expressed in d4 samples compared to d7, indicating *HOXA9* expression induced at 40h was not sustained. Although a high proportion of d7 EBs were identified as $CD34^{+}CD90^{+}CD43^{-}$, a marker profile attributed to human AGM (Luff et al., 2022; Ng et al., 2016), without measurement of *HOXA* expression in d7 EBs, it is not possible to conclude if differentiating cultures are indictive of definitive HE.

Of note, while *HOXA9* (and *mScarlet*) mRNA levels were shown to reduce from d4 to d7 in DOX-treated EBs, mScarlet protein remained elevated. mScarlet possesses a high fluorescence lifetime (Bindels et al., 2016), and while inducible mScarlet expression has not previously been assessed in Tet-On systems, TRE-driven GFP expression remained elevated up to 96h after DOX treatment (Xia et al., 2007). In Chapter 4, mScarlet degradation was assessed in iPSCs, where it was found to disappear within 72h of 1 µg/mL DOX treatment. From d0 to d7 of haematopoietic differentiation, a complete media change was only made once on d4, which likely contributed to the extended mScarlet expression observed. These findings highlight the importance of determining protein degradation rates under the conditions and in the cell types in which the protein will be studied.

As discussed, it is not possible to determine if the spin EB haematopoietic differentiation reproduces the intra-embryonic haematopoietic program based on the assays performed in this thesis. Due to the absence of *HOXA* expression in d14 and d18 suspension cells, it is unlikely to do so. However, d18 suspension cells do possess several features of definitive extra-embryonic haematopoietic progenitors, EMPs. In the CFU assay, d18 suspension cells exhibited robust erythro-myeloid differentiation potential and transcriptome analysis revealed a lymphoid gene signature was evident in d18 populations. Granulocyte differentiation potential is another indicator of extra-embryonic definitive haematopoiesis, with granulocyte differentiation potential previously observed in mouse yolk sac-derived EMPs yet not in early erythroid progenitors produced in the first haematopoietic wave (McGrath et al., 2015). Granulocyte-associated genes were upregulated only in d18 populations and d18 suspension cells possessed greater CFU-GM activity in the CFU assay than d14 cells, consistent with the proposed EMP phenotype of d18 cells. Conversely, the lack of a lymphoid gene signature, prevalence of BFU-E/CFU-E colonies in the CFU assay and upregulation of erythroid/megakaryocyte regulatory genes *GFI1B* and *GATA1* suggest d14 suspension cells are an *in vitro* equivalent of primitive haematopoietic progenitors.

6.4.4 Early transient overexpression of HOXA9 produced minor changes in the transcriptome profile of d18 haematopoietic progenitors

Transient HOXA9 overexpression during mesoderm patterning in SB/CHIR-treated cultures was shown to not significantly impact the number or immunophenotype of haematopoietic progenitor formation. However, RNA-seq analysis indicated transient HOXA9 overexpression induced some changes in the transcriptome of the resultant progenitors. DEG and pathway analysis showed d18 SC+DOX+ suspension cells were

enriched in endothelial-related genes with possible roles in blood vessel development and angiogenesis. It remains unclear as to why DOX treatment induced this response, however, HOXA9 expression has previously been described in endothelial cell types. In adult endothelial cells, HOXA9 has roles in regulating endothelial cell migration (Bruhl et al., 2004) and endothelial cell activation in response to inflammation (Trivedi et al., 2008). Additionally, HOXA9 is expressed in haemogenic endothelium lining the mouse dorsal aorta (Gao et al., 2018). Given this prior connection, it is possible HOXA9 overexpression during mesoderm patterning may promote endothelial-specific lineages and/or blood vessel development.

Chapter 7: General Discussion

The aim of this thesis was to establish the function of HOXA9 at distinct stages of haematopoietic differentiation in an hPSC model of haematopoiesis. To this end, CRISPR/Cas9 was used to generate a DOX-inducible HOXA9 overexpression system in hiPSCs. Despite the robust response induced by DOX in iPSCs, DOX-mediated upregulation of HOXA9 was not maintained throughout cardiac or haematopoietic differentiation and was found to reduce as differentiation progressed, with rapid reduction in inducible expression observed following mesendoderm specification in both lineages. As a result, the impact of transient HOXA9 overexpression specifically during mesoderm patterning was the focus of this thesis. Transient HOXA9 overexpression during mesoderm patterning in spin EB haematopoietic differentiation was found to promote the generation of all mesoderm subtypes, although specific enhancement of haemogenic mesoderm was not observed. The spin EB differentiation generated functionally and transcriptionally distinct haematopoietic progenitors in two waves, however, early transient overexpression of HOXA9 was not found to alter the immunophenotype or functionality of either subtype of progenitors.

The progressive loss of DOX responsiveness during differentiation to mesoderm-derived lineages was an unexpected finding of this thesis. This loss was attributed to methylation of elements within the integrated inducible HOXA9 cassette and the targeted AAVS1 locus. We have provided preliminary MeDIP data that demonstrates methylation was enriched at a specific site within the AAVS1 locus and within the TRE promoter on d2 and d6 of cardiac differentiation, respectively. However, further repeats of this experiment are necessary to confirm this observation, not only in the cardiac lineage but also in the haematopoietic lineage. Additionally, as MeDIP only indicates the presence of methylation, bisulphite sequencing should be performed to confirm the presence of methylation marks (Li and Tollefsbol, 2011; Li and Zhang, 2014). Attempts to demethylate iA9 differentiation cultures, thereby possibly facilitating DOX-mediated HOXA9/mScarlet induction, were not performed. In iPSCs, demethylation agents, such as 5-azacytidine, have been shown to both promote and inhibit iPSC self-renewal and differentiation, depending on the concentration used (Yoon et al., 2006). Methylation is a dynamic process and fluctuates during lineage specification (Reizel et al., 2018). As such, non-specific demethylation could be detrimental to haematopoietic development and was not considered a suitable approach to rectify possible methylation-induced silencing.

The impact of possible methylation within the AAVS1 locus is unclear. Interestingly, only d2 cardiac mesendoderm was enriched for methylation in the AAVS1 locus, with methylation in the AAVS1 locus not identified on d6. Moreover, DOX-mediated induction of *HOXA9* and mScarlet expression was still identified on d2 of cardiac and on d4 of spin EB haematopoietic differentiation, with rapid reduction in DOX induction activity evident thereafter. Therefore, although silencing within the AAVS1 locus may be responsible for the reduced induction of *HOXA9* and mScarlet observed in early iA9 differentiation cultures, it does not completely hinder expression of integrated transgenes. This is supported by the literature, where numerous studies have demonstrated robust expression of transgenes integrated in the AAVS1 locus (Bertero et al., 2016; Hazelbaker et al., 2020; Nafria et al., 2020b; Pawlowski et al., 2017; Qian et al., 2014). Nonetheless, methylation within the AAVS1 locus warrants further investigation.

The tightTRE promoter employed in this thesis appears to be particularly susceptible to DNA methylation, whereas alternative iterations of the Tet responsive element, such as the TRE3G promoter, are not associated with transgene silencing (Hazelbaker et al., 2020; Nafria et al., 2020b; Qian et al., 2014). The inducible HOXA9-mScarlet donor vector (pAAVS1-TRE-HOXA9-mScarlet) was designed with the intention of being reused in other inducible transgene systems. AflII sites were incorporated flanking the HOXA9-T2A-mScarlet element, facilitating straightforward excision of the element and replacement with an alternative transgene. However, in their current form, both the HOXA9-mScarlet donor vector and iA9 cell line have very limited use. Both are restricted to use in transgene overexpression in iPSCs or during early differentiation. RNA-seq analysis of d4 spin EB differentiations suggested induction of HOXA9 is feasible in endoderm and neuroectoderm cell types, as well as mesoderm which has been demonstrated in this thesis. In future inducible transgene expression studies, the TRE3G promoter should be used to drive transgene expression. Alternatively, other inducible systems, such as the tamoxifen-ERT2 system, could be employed (Jackson et al., 2016; Yang et al., 2017b). In addition, incorporation of insulator elements or ubiquitin chromatin opening elements (UCOEs) to circumvent against possible transgene silencing should be considered (Klatt et al., 2020; Ordovas et al., 2015; Pfaff et al., 2013).

Our efforts to establish a hiPSC-based model of haematopoiesis highlight the necessity to optimise differentiation conditions, regardless as to whether procedures adopted utilise commercially derived products or are based on the published literature. Characterisation of differentiation progression is central to optimisation, and in PSC haematopoietic model systems this is mainly performed using surface markers or gene

expression profiles associated with specific developmental stages of *in vivo* embryonic haematopoiesis. However, no standard surface markers or approach exists to characterise hPSC-derived haematopoiesis and several different expression profiles have been described. For example, specification of mesoderm with haematopoietic potential in early PSC differentiations can be identified by formation of $KDR^{+}CD34^{+}$ (Kennedy et al., 2012), $KDR^{+}PDGFR\alpha^{+}$ (Kataoka et al., 2011) or $KDR^{+}CD235a^{+/-}$ populations (Ditadi et al., 2015; Sturgeon et al., 2014), while specification of haemato-vascular mesoderm can be further distinguished by $KDR^{+}APLNR^{+}PDGFR\alpha^{-}$ (Choi et al., 2012) or $CD90^{+}CD34^{+}CD43^{-}$ (Ng et al., 2016) expression profiles. Haematopoietic cells are derived from HE in both yolk-sac and AGM haematopoiesis, which can be broadly characterised by co-expression of CD144, CD34, KDR and CD31 and lack of CD43 (Garcia-Alegria et al., 2018; Lange et al., 2021). Several expression profiles have been used to distinguish HE indictive of intra-embryonic haematopoiesis modelling based on acquisition of an arterial signature, which include expression of SOX17 (Jung et al., 2021; Ng et al., 2016), DLL4 (Park et al., 2018; Uenishi et al., 2018) and/or CXCR4 (Jung et al., 2021) and lack of CD73 and CD184 (Ditadi et al., 2015). The immunophenotype of PSC-derived haematopoietic progenitors is similarly quite broad; progenitors can be identified by co-expression of CD34 and CD45, and/or CD43 or more specific marker profiles such as co-expression of CD90, GPI-80 and CD44 and lack of CD38 and CD45RA (Lange et al., 2021; Prashad et al., 2015).

While this wide plethora of characterisation profiles reflects both the heterogenous nature and the ever-changing understanding of embryonic haematopoiesis, it also makes cross-comparison of different PSC haematopoietic differentiation models challenging. This thesis describes the first application of hPSC-based haematopoietic differentiation in our lab. A challenge we experienced during optimisation was determining how our spin EB differentiation procedure fitted in with the wider literature. A consensus on the expected marker profiles during haematopoietic differentiation would help characterise and assess new haematopoietic differentiation procedure and allow comparison with the published literature.

The spin EB differentiation procedure described in this thesis produces phenotypes and cell types reminiscent of extra-embryonic definitive haematopoiesis. During differentiation, we noted several key developmental events in our spin EB differentiation which are indicative of extra-embryonic haematopoiesis: acquisition of KDR in a subset of mesoderm with haematopoietic potential; formation of $CD90^{+}CD34^{+}CD43^{-}$ haemato-vascular cells; and formation of haematopoietic progenitors with a

$CD34^+CD43^+CD90^+CD38^-$ immunophenotype. Spin EB differentiation also produced two distinct progenitor sub-types. On d14, suspension cells with an apparent erythroid-biased, were generated, while continued culture yielded suspension cells on d18 with erythro-myeloid potential and a lymphoid gene signature, consistent with derivation of EMPs. Together this suggests the spin EB differentiation model described in this thesis is an appropriate model of extra-embryonic definitive haematopoiesis. Furthermore, we demonstrated the spin EB differentiation is applicable to hiPSCs routinely maintained on Matrigel, having previously only been performed in hPSCs cultured on MEFs (Nafria et al., 2020b, 2020a; Ng et al., 2016). Thus, this broadens the utility of the spin EB procedure, to both Matrigel and MEF maintained hPSCs.

Despite the challenges encountered in generation of the inducible HOXA9 system, a conditional HOXA9 overexpression approach remains a suitable model to study the function of HOXA9 in haematopoiesis. Using RNA-seq we were able to establish some unique changes induced by short-term overexpression of HOXA9 in early differentiation. Transient HOXA9 overexpression was shown to promote pan-mesoderm specification and induce transcriptome changes in mesoderm populations when overexpressed both in the background and absence of SB/CHIR signalling pathways. Furthermore, although brief HOXA9 overexpression during mesoderm patterning was largely shown to not impact the formation and functionality of haematopoietic progenitors, some small transcriptome differences were identified. Use of more in-depth analysis tools, such as single cell RNA-seq, will likely identify further transcriptome differences induced by HOXA9.

The findings of this thesis suggest overexpression of HOXA9 during mesoderm patterning promotes specification of all mesoderm sub-types and does not substantially effect haemogenic mesoderm or haematopoietic progenitor formation and fate specification. Of note, this conclusion specifically relates to short-term transient HOXA9 overexpression during mesoderm development; investigation of HOXA9 function at later stages of haematopoietic development may better establish HOXA9's role in haematopoiesis. Recently, single cell analysis of human embryos identified HOXA9 as part of a HSC signature $RUNX1^+HOXA9^+MLLT3^+MECOM^+HLF^+SPINK2^+$ which could distinguish HSCs from other progenitor populations during embryonic development (Calvanese et al., 2022). Individual studies of other HSC signature genes have established prominent roles for these genes in HSC function, such as MLLT3, which has been shown to regulate both HSC maintenance and maturation (Calvanese et al., 2019), and Hlf, which specifically marks the development of HSCs and not other progenitors

(Yokomizo et al., 2019). HOXA9 may have a similar role to these HSC signature genes. Recent work has also proposed the low reconstitution of hPSC-derived HSPCs in mouse transplantation assays is attributed to impairments in hPSC-derived HSPC maintenance and maturation (Calvanese et al., 2022; Fidanza et al., 2020). Given that *HOXA9* has been identified as one of the most differentially expressed genes in a comparison of hiPSC-derived and foetal liver HSPCs (Fidanza et al., 2020), it is possible HOXA9 plays a crucial role in functional maturity of HSCs. In addition, a role for HOXA9 in HSC maintenance and self-renewal is well-established in the mouse haematopoietic system (Lawrence et al., 2005; Thorsteinsdottir et al., 2002) and in hPSC-derived progenitors HOXA9 has been reported to confer self-renewal and multilineage differentiation potential in *in vitro* assays (Doulatov et al., 2013). In conclusion, the literature supports a prominent role for HOXA9 in the specification and maintenance of HSCs and also suggests HOXA9 has crucial roles during late haematopoietic differentiation. Investigation of HOXA9 function at late stages of haematopoietic differentiation will help determine how HOXA9 may influences these processes.

References

Ackermann, M., Liebhaber, S., Klusmann, J., and Lachmann, N. (2015). Lost in translation: pluripotent stem cell-derived hematopoiesis. EMBO Mol. Med. *7*, 1388–1402. https://doi.org/10.15252/emmm.201505301.

Ahler, E., Sullivan, W.J., Cass, A., Braas, D., York, A.G., Bensinger, S.J., Graeber, T.G., and Christofk, H.R. (2013). Doxycycline Alters Metabolism and Proliferation of Human Cell Lines. PLoS One *8*, 8–14. https://doi.org/10.1371/journal.pone.0064561.

Alcalay, M., Tiacci, E., Bergomas, R., Bigerna, B., Venturini, E., Minardi, S.P., Meani, N., Diverio, D., Bernard, L., Tizzoni, L., et al. (2005). Acute myeloid leukemia bearing cytoplasmic nucleophosmin shows a distinct gene expression profile characterized by of genes involved in maintenance. Blood *106*, 899. https://doi.org/10.1182/blood-2005-02-0560.Supported.

Alharbi, R.A., Pettengell, R., Pandha, H.S., and Morgan, R. (2013). The role of HOX genes in normal hematopoiesis and acute leukemia. Leukemia *27*, 1000–1008. https://doi.org/10.1038/leu.2012.356.

Angelos, M.G., Abrahante, J.E., Blum, R.H., and Kaufman, D. s. (2018). Single cell resolution of human hemato-endothelial cells degines transcriptional signatures of hemogenic endothelium. Stem Cells *36*, 206–217. https://doi.org/10.1002/stem.2739.Single.

Antonchuk, J., Sauvageau, G., and Humphries, R.K. (2001). HOXB4 overexpression mediates very rapid stem cell regeneration and competitive hematopoietic repopulation. Exp. Hematol. *29*, 1125–1134. https://doi.org/10.1016/S0301-472X(01)00681-6.

Armstrong, S.A., Staunton, J.E., Silverman, L.B., Pieters, R., Den Boer, M.L., Minden, M.D., Sallan, S.E., Lander, E.S., Golub, T.R., and Korsmeyer, S.J. (2002). MLL translocations specify a distinct gene expression profile that distinguishes a unique leukemia. Nat. Genet. *30*, 41–47. https://doi.org/10.1038/ng765.

Aubrey, B.J., Kelly, G.L., Kueh, A.J., Brennan, M.S., O'Connor, L., Milla, L., Wilcox, S., Tai, L., Strasser, A., and Herold, M.J. (2015). An Inducible Lentiviral Guide RNA Platform Enables the Identification of Tumor-Essential Genes and Tumor-Promoting Mutations InVivo. Cell Rep. *10*, 1422–1432. https://doi.org/10.1016/j.celrep.2015.02.002.

Bach, C., Buhl, S., Mueller, D., García-Cuéllar, M.P., Maethner, E., and Slany, R.K. (2010). Leukemogenic transformation by HOXA cluster genes. Blood *115*, 2910–2918. https://doi.org/10.1182/blood-2009-04-216606.

Bäckman, C.M., Zhang, Y., Hoffer, B.J., and Tomac, A.C. (2004). Tetracycline-inducible expression systems for the generation of transgenic animals : a comparison of various inducible systems carried in a single vector. *139*, 257–262. https://doi.org/10.1016/j.jneumeth.2004.05.012.

Beers, J., Gulbranson, D.R., George, N., Siniscalchi, L.I., Jones, J., Thomson, J.A., and Chen, G. (2013). Passaging and colony expansion of human pluripotent stem cells by enzyme-free dissociation in chemically defines culture condictions. NIH-PA Author Manuscr. *7*, 2029–2040. https://doi.org/10.1038/nprot.2012.130.Passaging.

Behrens, A.N., Iacovino, M., Lohr, J.L., Ren, Y., Zierold, C., Harvey, R.P., Kyba, M., Garry, D.J., and Martin, C.M. (2013). Nkx2-5 Mediates Differential Cardiac Differentiation Through Interaction with Hoxa10. Stem Cells Dev. *22*, 2211–2220. https://doi.org/10.1089/scd.2012.0611.

Bertero, A., Pawlowski, M., Ortmann, D., Snijders, K., Yiangou, L., Cardoso de Brito, M., Brown, S., Bernard, W.G., Cooper, J.D., Giacomelli, E., et al. (2016). Optimized inducible shRNA and CRISPR/Cas9 platforms for in vitro studies of human development using hPSCs. Development *143*, 4405–4418. https://doi.org/10.1242/dev.138081.

Bertrand, J.Y., Chi, N.C., Santoso, B., Teng, S., Stainier, D.Y.R., and Traver, D. (2010). Haematopoietic stem cells derive directly from aortic endothelium during development. Nature *464*, 2–6. https://doi.org/10.1038/nature08738.

Bertrand, N., Roux, M., Ryckebüsch, L., Niederreither, K., Dollé, P., Moon, A., Capecchi, M., and Zaffran, S. (2011). Hox genes define distinct progenitor sub-domains within the second heart field. Dev. Biol. *353*, 266–274. https://doi.org/10.1016/j.ydbio.2011.02.029.

Bhagwan, J. (2018). CRISPR-Cas9 gene targeting of human induced pluripotent stem cells to generate a palette of optical reporter cell lines towards investigating hypertrophic cardiomyopathy.

Bhagwan, J.R., Mosqueira, D., Chairez-Cantua, K., Mannhardt, I., Bodbin, S.E., Bakar, M., Smith, J.G.W., and Denning, C. (2020a). Isogenic models of hypertrophic

cardiomyopathy unveil differential phenotypes and mechanism-driven therapeutics. J. Mol. Cell. Cardiol. *2019*, 104265. https://doi.org/10.1016/j.meegid.2020.104265.

Bhagwan, J.R., Collins, E., Mosqueira, D., Bakar, M., Smith, J.G.W., Denning, C., Johnson, B.B., and Thompson, A. (2020b). Variable expression and silencing of CRISPR-Cas9 targeted transgenes identifies the AAVS1 locus as not an entirely safe harbour [version 2 ; peer review : 2 approved , 1 not approved]. 1–27. .

Bhatlekar, S., Fields, J.Z., and Boman, B.M. (2018). Role of HOX Genes in Stem Cell Differentiation and Cancer. Stem Cells Int. *2018*. .

Bibikova, M., Carroll, D., Segal, D.J., Trautman, J.K., Smith, J., Kim, Y.-G., and Chandrasegaran, S. (2001). Stimulation of Homologous Recombination through Targeted Cleavage by Chimeric Nucleases. Mol. Cell. Biol. *21*, 289–297. https://doi.org/10.1128/mcb.21.1.289-297.2001.

Bijl, J., Thompson, A., Ramirez-solis, R., Krosl, J., Grier, D.G., Lawrence, H.J., and Sauvageau, G. (2006). Analysis of HSC activity and compensatory Hox gene expression profile in Hoxb cluster mutant fetal liver cells. Blood *108*, 116–122. https://doi.org/10.1182/blood-2005-06-2245.Supported.

Bindels, D.S., Haarbosch, L., Van Weeren, L., Postma, M., Wiese, K.E., Mastop, M., Aumonier, S., Gotthard, G., Royant, A., Hink, M.A., et al. (2016). MScarlet: A bright monomeric red fluorescent protein for cellular imaging. Nat. Methods *14*, 53–56. https://doi.org/10.1038/nmeth.4074.

Bird, A. (2002). DNA methylation patterns and epigenetic memory. 6–21. https://doi.org/10.1101/gad.947102.6.

Boch, J., Scholze, H., Schornack, S., Landgraf, A., Hahn, S., Kay, S., Lahaye, T., Nickstadt, A., and Bonas, U. (2009). Breaking the code of DNA binding specificity of TAL-type III effectors. Science (80-.). *326*, 1509–1512. https://doi.org/10.1126/science.1178811.

Boisset, J., Cappellen, W. Van, Andrieu-soler, C., Galjart, N., Dzierzak, E., and Robin, C. (2010). In vivo imaging of haematopoietic cells emerging from the mouse aortic endothelium. Nature *464*, 116–120. https://doi.org/10.1038/nature08764.

Bruhl, T., Urbich, C., Aicher, D., Acker-Palmer, A., Zeiher, A.M., and Dimmeler, S. (2004). Homeobox A9 Transcriptionally Regulates the EphB4 Receptor to Modulate Endothelial Cell Migration and Tube Formation. Circ. Res. *94*, 743–751. https://doi.org/10.1161/01.RES.0000120861.27064.09.

De Bruijn, M.F.T.R., Speck, N.A., Peeters, M.C.E., and Dzierzak, E. (2000). Definitive hematopoietic stem cells first develop within the major arterial regions of the mouse embryo. EMBO J. *19*, 2465–2474. https://doi.org/10.1093/emboj/19.11.2465.

Bruveris, F.F., Ng, E.S., Leitoguinho, A.R., Motazedian, A., Vlahos, K., Sourris, K., Mayberry, R., Mcdonald, P., Azzola, L., Davidson, N.M., et al. (2020). Human yolk sac-like haematopoiesis generates RUNX1-, GFI1- and/or GFI1B-dependent blood and SOX17-positive endothelium. Development *147*. https://doi.org/10.1242/dev.193037.

Burridge, P.W., Keller, G., Gold, J.D., and Wu, J.C. (2011). Production of De Novo Cardiomyocytes : Human Pluripotent Stem Cell Differentiation and Direct Reprogramming. Stem Cell *10*, 16–28. https://doi.org/10.1016/j.stem.2011.12.013.

Burridge, P.W., Matsa, E., Shukla, P., Lin, Z.C., Churko, J.M., Ebert, A.D., Lan, F., Diecke, S., Huber, B., Mordwinkin, N.M., et al. (2014). Chemically defined generation of human cardiomyocytes. *11*. https://doi.org/10.1038/nmeth.2999.

Buske, C., Feuring-buske, M., Antonchuk, J., Rosten, P., Hogge, D.E., Eaves, C.J., and Humphries, R.K. (2001). Overexpression of HOXA10 perturbs human lymphomyelopoiesis in vitro and in vivo. *97*, 2286–2292. .

Buske, C., Feuring-Buske, M., Abramovich, C., Spiekermann, K., Eaves, C.J., Coulombel, L., Sauvageau, G., Hogge, D.E., and Humphries, R.K. (2002). Deregulated expression of HOXB4 enhances the primitive growth activity of human hematopoietic cells. Blood *100*, 862–868. https://doi.org/10.1182/blood-2002-01-0220.

Cai, S., Liu, R., Wang, P., Li, J., Xie, T., and Wang, M. (2020). PRMT5 Prevents Cardiomyocyte Hypertrophy via Symmetric Dimethylating HoxA9 and Repressing HoxA9 Expression. *11*, 1–11. https://doi.org/10.3389/fphar.2020.600627.

Cai, Z., De Bruijn, M., Ma, X., Dortland, B., Luteijn, T., Downing, J.R., and Dzierzak, E. (2000). Haploinsufficiency of AML1 affects the temporal and spatial generation of hematopoietic stem cells in the mouse embryo. Immunity *13*, 423–431.

https://doi.org/10.1016/S1074-7613(00)00042-X.

Calvanese, V., Nguyen, A.T., Bolan, T.J., Vavilina, A., Su, T., Lee, L.K., Wang, Y., Lay, F.D., Magnusson, M., Crooks, G.M., et al. (2019). MLLT3 governs human haematopoietic stem-cell self-renewal and engraftment. Nature *576*, 281–286. https://doi.org/10.1038/s41586-019-1790-2.

Calvanese, V., Capellera-Garcia, S., Ma, F., Fares, I., Liebscher, S., Ng, E.S., Ekstrand, S., Aguadé-Gorgorió, J., Vavilina, A., Lefaudeux, D., et al. (2022). Mapping human haematopoietic stem cells from haemogenic endothelium to birth. Nature *604*, 534–540. https://doi.org/10.1038/s41586-022-04571-x.

Canu, G., and Ruhrberg, C. (2021). First blood: the endothelial origins of hematopoietic progenitors. Angiogenesis *24*, 199–211. https://doi.org/10.1007/s10456-021-09783-9.

Cao, X., Yakala, G.K., van den Hil, F.E., Cochrane, A., Mummery, C.L., and Orlova, V. V. (2019). Differentiation and Functional Comparison of Monocytes and Macrophages from hiPSCs with Peripheral Blood Derivatives. Stem Cell Reports *12*, 1282–1297. https://doi.org/10.1016/j.stemcr.2019.05.003.

Cao, X., Hil, F.E. Van Den, Mummery, C.L., and Orlova, V. V (2020). Generation and Functional Characterization of Monocytes and Macrophages Derived from Human Induced Pluripotent Stem Cells. Curr. Protoc. Stem Cell Biol. *52*, 1–31. https://doi.org/10.1002/cpsc.108.

Castano, J., Aranda, S., Bueno, C., Calero-nieto, F.J., Mejia-ramirez, E., Mosquera, J.L., Blanco, E., Wang, X., Prieto, C., Zabaleta, L., et al. (2019). GATA2 Promotes Hematopoietic Development and Represses Cardiac Differentiation of Human Mesoderm. Stem Cell Reports *13*, 515–529. https://doi.org/10.1016/j.stemcr.2019.07.009.

Cedar, H., and Bergman, Y. (2012). Programming of DNA Methylation Patterns. Annu. Rev. Biochem. *81*, 97–117. https://doi.org/10.1146/annurev-biochem-052610.

Celiz, A.D., Smith, J.G.W., Patel, A.K., Hook, A.L., Rajamohan, D., George, V.T., Flatt, L., Patel, M.J., Epa, V.C., Singh, T., et al. (2015). Discovery of a Novel Polymer for Human Pluripotent Stem Cell Expansion and Multilineage Differentiation. 4006–4012. https://doi.org/10.1002/adma.201501351.

Cerbini, T., Funahashi, R., Luo, Y., Liu, C., Park, K., Rao, M., Malik, N., and Zou, J. (2015). Transcription activator-like effector nuclease (TALEN)-mediated CLYBL targeting enables enhanced transgene expression and one-step generation of dual reporter human induced pluripotent stem cell (iPSC) and neural stem cell (NSC) lines. PLoS One *10*, 1–18. https://doi.org/10.1371/journal.pone.0116032.

Cerdan, C., Mcintyre, B.A.S., Mechael, R., Levadoux-martin, M., Yang, J., Lee, J.B., and Bhatia, M. (2012). Activin A Promotes Hematopoietic Fated Mesoderm Development Through Upregulation of Brachyury in Human Embryonic Stem Cells. Stem Cells Dev. *21*, 2866–2877. https://doi.org/10.1089/scd.2012.0053.

Chanda, B., Ditadi, A., Iscove, N.N., and Keller, G. (2013). Retinoic acid signaling is essential for embryonic hematopoietic stem cell development. Cell *155*, 215–227. https://doi.org/10.1016/j.cell.2013.08.055.

Chang, M.Y., Rhee, Y.H., Yi, S.H., Lee, S.J., Kim, R.K., Kim, H., Park, C.H., and Lee, S.H. (2014). Doxycycline enhances survival and self-renewal of human pluripotent stem cells. Stem Cell Reports *3*, 353–364. https://doi.org/10.1016/j.stemcr.2014.06.013.

Chen, M.J., Yokomizo, T., Zeigler, B., Dzierzak, E., and Speck, A. (2009). Runx1 is required for the endothelial to hematopoietic cell transition but not thereafter. Nature *457*, 887–891. https://doi.org/10.1038/nature07619.Runx1.

Chen, Y., Cao, J., Xiong, M., Petersen, A.J., Dong, Y., Tao, Y., Huang, C.T.L., Du, Z., and Zhang, S.C. (2015). Engineering Human Stem Cell Lines with Inducible Gene Knockout using CRISPR/Cas9. Cell Stem Cell *17*, 233–244. https://doi.org/10.1016/j.stem.2015.06.001.

Chen, Z., Ren, X., Xu, X., Zhang, X., Hui, Y., Liu, Z., Shi, L., Fang, Y., Ma, L., Liu, Y., et al. (2018). Genetic Engineering of Human Embryonic Stem Cells for Precise Cell Fate Tracing during Human Lineage Development. Stem Cell Reports *11*, 1257–1271. https://doi.org/10.1016/j.stemcr.2018.09.014.

Cho, S.W., Kim, S., Kim, J.M., and Kim, J.S. (2013). Targeted genome engineering in human cells with the Cas9 RNA-guided endonuclease. Nat. Biotechnol. *31*, 230–232. https://doi.org/10.1038/nbt.2507.

Cho, S.W., Kim, S., Kim, Y., Kweon, J., Kim, H.S., Bae, S., and Kim, J. (2014). Analysis

of off-target effects of CRISPR / Cas-derived RNA-guided endonucleases and nickases. Genome Res. *24*, 132–141. https://doi.org/10.1101/gr.162339.113.Freely.

Choi, K., Kennedy, M., Kazarov, A., Papadimitriou, J.C., and Keller, G. (1998). A common precursor for hematopoietic and endothelial cells. Development *125*, 725–732. https://doi.org/10.1242/dev.125.4.725.

Choi, K.D., Vodyanik, M.A., Togarrati, P.P., Suknuntha, K., Kumar, A., Samarjeet, F., Probasco, M.D., Tian, S., Stewart, R., Thomson, J.A., et al. (2012). Identification of the Hemogenic Endothelial Progenitor and Its Direct Precursor in Human Pluripotent Stem Cell Differentiation Cultures. Cell Rep. *2*, 553–567. https://doi.org/10.1016/j.celrep.2012.08.002.

Clarke, R.L., Yzaguirre, A.D., Yashiro-ohtani, Y., Bondue, A., Blanpain, C., Pear, W.S., Speck, N.A., and Keller, G. (2013). The expression of Sox17 identifies and regulates haemogenic endothelium. Nat. Cell Biol. *15*, 502–510. https://doi.org/10.1038/ncb2724.

Collado, J.A., Alvarez, I., Ciudad, M.T., Espinosa, G., Canals, F., Pujol-Borrell, R., Carrascal, M., Abian, J., and Jaraquemada, D. (2013). Composition of the HLA-DR-associated human thymus peptidome. Eur. J. Immunol. *43*, 2273–2282. https://doi.org/10.1002/eji.201243280.

Cong, L., Ran, F.A., Cox, D., Lin, S., Barretto, R., Hsu, P.D., Wu, X., Jiang, W., and Marraffini, L. a (2013). Multiplex Genome Engineering Using CRISPR/Cas Systems. Science *339*, 819–823. https://doi.org/10.1126/science.1231143.Multiplex.

Corada, M., Orsenigo, F., Morini, M.F., Pitulescu, M.E., Bhat, G., Nyqvist, D., Breviario, F., Conti, V., Briot, A., Iruela-Arispe, M.L., et al. (2013). Sox17 is indispensable for acquisition and maintenance of arterial identity. Nat. Commun. *4*. https://doi.org/10.1038/ncomms3609.

Cortes, F., Debacker, C., Peault, B., and Labastie, M. (1999). Differential expression of KDR/VEGFR-2 and CD34 during mesoderm development of the early human embryo. Mech. Dev. *83*, 161–164. .

Cowan, J.E., McCarthy, N.I., Parnell, S.M., White, A.J., Bacon, A., Serge, A., Irla, M., Lane, P.J.L., Jenkinson, E.J., Jenkinson, W.E., et al. (2014). Differential Requirement for CCR4 and CCR7 during the Development of Innate and Adaptive αβT Cells in the Adult

Thymus. J. Immunol. *193*, 1204–1212. https://doi.org/10.4049/jimmunol.1400993.

Creamer, J.P., Dege, C., Ren, Q., Ho, J.T.K., Valentine, M.C., Druley, T.E., Sturgeon, C.M., Louis, S., and Louis, S. (2017). Human definitive hematopoietic specification from pluripotent stem cells is regulated by mesodermal expression of CDX4. Blood *129*, 2988–2993. https://doi.org/10.1182/blood-2016-11-749382.The.

Crooks, B.G.M., Fuller, J., Petersen, D., Izadi, P., Malik, P., Pattengale, P.K., Kohn, D.B., and Gasson, J.C. (1999). Constitutive HOXA5 Expression Inhibits Erythropoiesis and Increases Myelopoiesis From Human Hematopoietic Progenitors. 519–528. .

D'Souza, S.L., Elefanty, A.G., and Keller, G. (2005). SCL/Tal-1 is essential for hematopoietic commitment of the hemangioblast but not for its development. Blood *105*, 3862–3870. https://doi.org/10.1182/blood-2004-09-3611.

Daga, A., Podesta, M., Capra, M.C., Piaggio, G., Frassoni, F., and Corte, G. (2000). The retroviral transduction of HOXC4 into human CD34+ cells induces an in vitro expansion of clonogenic and early progenitors. Exp. Hematol. *28*, 569–574. https://doi.org/10.1016/S0301-472X(00)00135-1.

Dai, C., Gong, Q., Cheng, Y., and Su, G. (2019). Regulatory mechanisms of Robo4 and their effects on angiogenesis. Biosci. Rep. *39*, 1–18. https://doi.org/10.1042/BSR20190513.

Dard, A., Reboulet, J., Jia, Y., Bleicher, F., Duffraisse, M., Vanaker, J.M., Forcet, C., and Merabet, S. (2018). Human HOX Proteins Use Diverse and Context-Dependent Motifs to Interact with TALE Class Cofactors. Cell Rep. *22*, 3058–3071. https://doi.org/10.1016/j.celrep.2018.02.070.

Das, A.T., Tenenbaum, L., and Berkhout, B. (2016). Tet-On Systems For Doxycycline-inducible Gene Expression. Curr. Gene Ther. *16*, 156–167. https://doi.org/10.2174/1566523216666160524144041.

Dege, C., and Sturgeon, C.M. (2017). Directed differentiation of primitive and definitive hematopoietic progenitors from human pluripotent stem cells. J. Vis. Exp. *2017*, 1–9. https://doi.org/10.3791/55196.

Dege, C., Fegan, K.H., Creamer, J.P., Fehniger, T.A., Palis, J., Sturgeon, C.M., Dege,

C., Fegan, K.H., Creamer, J.P., Berrien-elliott, M.M., et al. (2020). Potently Cytotoxic Natural Killer Cells Initially Emerge from Erythro-Myeloid Progenitors during Mammalian Development. Dev. Cell *53*, 229-239.e7. https://doi.org/10.1016/j.devcel.2020.02.016.

Deltcheva, E., Chylinski, K., Sharma, C.M., Gonzales, K., Chao, Y., Pirzada, Z.A., Eckert, M.R., Vogel, J., and Charpentier, E. (2011). CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III. Nature *471*, 602–607. https://doi.org/10.1038/nature09886.

Deschamps, J., and van Nes, J. (2005). Developmental regulation of the Hox genes during axial morphogenesis in the mouse. Development *132*, 2931–2942. https://doi.org/10.1242/dev.01897.

Di-Poï, N., Koch, U., Radtke, F., and Duboule, D. (2010). Additive and global functions of HoxA cluster genes in mesoderm derivatives. Dev. Biol. *341*, 488–498. https://doi.org/10.1016/j.ydbio.2010.03.006.

Dickson, G.J., Lappin, T.R., and Thompson, A. (2009). Complete Array of HOX Gene Expression by RQ-PCR. *538*. https://doi.org/10.1007/978-1-59745-418-6.

Ditadi, A., Sturgeon, C.M., Tober, J., Awong, G., Kennedy, M., Yzaguirre, A.D., Azzola, L., Ng, E.S., Stanley, E.G., French, D.L., et al. (2015). Human definitive haemogenic endothelium and arterial vascular endothelium represent distinct lineages. Nat. Cell Biol. *17*, 580–591. https://doi.org/10.1038/ncb3161.

Domingo-Reines, J., López-Ornelas, A., Montes, R., Romero, T., Rodriguez-Llamas, J.L., Lara-Rodarte, R., González-Pozas, F., Ayllon, V., Menendez, P., Velasco, I., et al. (2017). Hoxa9 and EGFP reporter expression in human Embryonic Stem Cells (hESC) as useful tools for studying human development. Stem Cell Res. *25*, 286–290. https://doi.org/10.1016/j.scr.2017.08.004.

Dou, D.R., Calvanese, V., Sierra, M.I., Nguyen, A.T., Minasian, A., Saarikoski, P., Sasidharan, R., Ramirez, C.M., Zack, J.A., Crooks, G.M., et al. (2016). Medial HOXA genes demarcate haematopoietic stem cell fate during human development. Nat. Cell Biol. *18*, 595–606. https://doi.org/10.1038/ncb3354.

Doulatov, S., Vo, L.T., Chou, S.S., Kim, P.G., Arora, N., Li, H., Hadland, B.K., Bernstein, I.D., Collins, J.J., Zon, L.I., et al. (2013). Induction of multipotential hematopoietic

progenitors from human pluripotent stem cells via respecification of lineage-restricted precursors. Cell Stem Cell *13*, 459–470. https://doi.org/10.1016/j.stem.2013.09.002.

Durston, A.J. (2012). Hox Genes: Master Regulators of the Animal Bodyplan. Embryogenesis https://doi.org/10.5772/37007.

Dzierzak, E., and Bigas, A. (2018). Blood Development : Hematopoietic Stem Cell Dependence and Independence. Stem Cell *22*, 639–651. https://doi.org/10.1016/j.stem.2018.04.015.

Dzierzak, E., and Speck, N.A. (2008). Of lineage and legacy: The development of mammalian hematopoietic stem cells. Nat. Immunol. *9*, 129–136. https://doi.org/10.1038/ni1560.

Elcheva, I., Brok-Volchanskaya, V., Kumar, A., Liu, P., Lee, J.H., Tong, L., Vodyanik, M., Swanson, S., Stewart, R., Kyba, M., et al. (2014). Direct induction of haematoendothelial programs in human pluripotent stem cells by transcriptional regulators. Nat. Commun. *5*. https://doi.org/10.1038/ncomms5372.

Ema, H., and Nakauchi, H. (2000). Expansion of hematopoietic stem cells in the developing liver of a mouse embryo. Blood *95*, 2284–2288. https://doi.org/10.1182/blood.v95.7.2284.

Epsztejn-Litman, S., Feldman, N., Abu-remaileh, M., Shufaro, Y., Gerson, A., Shinkai, Y., Cedar, H., Bergman, Y., Ueda, J., and Deplus, R. (2008). De novo DNA methylation promoted by G9a prevents reprogramming of embryonically silenced genes. Nat. Struct. Mol. Biol. *15*, 1176–1183. https://doi.org/10.1038/nsmb.1476.

Evseenko, D., Zhu, Y., Schenke-Layland, K., Kuo, J., Latour, B., Ge, S., Scholes, J., Dravid, G., Li, X., MacLellan, W.R., et al. (2010). Mapping the first stages of mesoderm commitment during differentiation of human embryonic stem cells. Proc. Natl. Acad. Sci. U. S. A. *107*, 13742–13747. https://doi.org/10.1073/pnas.1002077107.

Fidanza, A., Stumpf, P.S., Ramachandran, P., Tamagno, S., Babtie, A., Lopez-Yrigoyen, M., Taylor, A.H., Easterbrook, J., Henderson, B.E.P., Axton, R., et al. (2020). Single-cell analyses and machine learning define hematopoietic progenitor and HSC-like cells derived from human PSCs. Blood *136*, 2893–2904. https://doi.org/10.1182/blood.2020006229.

Fischbach, N.A., Rozenfeld, S., Shen, W., Fong, S., Chrobak, D., Ginzinger, D., Kogan, S.C., Radhakrishnan, A., Le Beau, M.M., Largman, C., et al. (2005). HOXB6 overexpression in murine bone marrow immortalizes a myelomonocytic precursor in vitro and causes hematopoietic stem cell expansion and acute myeloid leukemia in vivo. Blood *105*, 1456–1466. https://doi.org/10.1182/blood-2004-04-1583.

Fortschegger, K., Husa, A., Schinnerl, D., Nebral, K., and Strehl, S. (2021). Expression of RUNX1-JAK2 in Human Induced Pluripotent Stem Cell-Derived Hematopoietic Cells Activates the JAK-STAT and MYC Pathways. Int. J. Mol. Sci. *22*, 1–18. .

Frame, J.M., McGrath, K.E., and Palis, J. (2013). Erythro-Myeloid Progenitors: “definitive” hematopoiesis in the conceptus prior to the emergence of hematopoietic stem cells. Blood Cells, Mol. Dis. *51*. https://doi.org/10.1016/j.bcmd.2013.09.006.Erythro-Myeloid.

Fu, Y., Foden, J.A., Khayter, C., Maeder, M.L., Reyon, D., Joung, J.K., and Sander, J.D. (2013). High-frequency off-target mutagenesis induced by CRISPR-Cas nucleases in human cells. Nat. Biotechnol. *31*, 822–827. https://doi.org/10.1038/nbt.2623.

Fuller, J.F., Mcadara, J., Yaron, Y., Sakaguchi, M., Fraser, J.K., and Gasson, J.C. (1999). Characterization of HOX Gene Expression During Myelopoiesis: Role of HOX A5 in Lineage Commitment and Maturation. Blood *93*, 3391–3400. .

Gama-Norton, L., Ferrando, E., Ruiz-Herguido, C., Liu, Z., Guiu, J., Islam, A.B.M.M.K., Lee, S.-U., Yan, M., Guidos, C.J., Lopez-Bigas, N., et al. (2015). Notch signal strength controls cell fate in the haemogenic endothelium. Nat. Comm https://doi.org/10.1038/ncomms9510.

Gao, J., Chen, Y.H., and Peterson, L.A.C. (2015). GATA family transcriptional factors: Emerging suspects in hematologic disorders. Exp. Hematol. Oncol. *4*, 1–7. https://doi.org/10.1186/s40164-015-0024-z.

Gao, X., Xu, C., Asada, N., and Frenette, P.S. (2018). The hematopoietic stem cell niche : from embryo to adult. Development *145*. https://doi.org/10.1242/dev.139691.

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

https://doi.org/10.1016/j.stemcr.2018.09.013.

Garcia-Fernàndez, J. (2005). Hox, ParaHox, ProtoHox: Facts and guesses. Heredity (Edinb). *94*, 145–152. https://doi.org/10.1038/sj.hdy.6800621.

Gekas, C., Dieterlen-lièvre, F., Orkin, S.H., Mikkola, H.K.A., and Marne, N. (2005). The Placenta Is a Niche for Hematopoietic Stem Cells. Dev. Cell *8*, 365–375. https://doi.org/10.1016/j.devcel.2004.12.016.

Georgomanoli, M., and Papapetrou, E.P. (2019). Modeling blood diseases with human induced pluripotent stem cells. Dis. Model. Mech. *12*. https://doi.org/10.1242/dmm.039321.

Gertow, K., Hirst, C.E., Yu, Q.C., Ng, E.S., Pereira, L.A., Davis, R.P., Stanley, E.G., and Elefanty, A.G. (2013). WNT3A promotes hematopoietic or mesenchymal differentiation from hESCs depending on the time of exposure. Stem Cell Reports *1*, 53–65. https://doi.org/10.1016/j.stemcr.2013.04.002.

Gilmour, J., Assi, S.A., Jaegle, U., Kulu, D., van de Werken, H., Clarke, D., Westhead, D.R., Philipsen, S., and Bonifer, C. (2014). A crucial role for the ubiquitously expressed transcription factor Sp1 at early stages of hematopoietic specification. Dev. *141*, 2391–2401. https://doi.org/10.1242/dev.106054.

Glass, W.G., Mcdermott, D.H., Lim, J.K., Lekhong, S., Yu, S.F., Frank, W.A., Pape, J., Cheshier, R.C., and Murphy, P.M. (2006). CCR5 defi ciency increases risk of symptomatic West Nile virus infection. *203*, 35–40. https://doi.org/10.1084/jem.20051970.

Golub, T.R., Slonim, D.K., Tamayo, P., Huard, C., Gaasenbeek, M., Mesirov, J.P., Coller, H., Loh, M.L., Downing, J.R., Caligiuri, M.A., et al. (1999). Molecular classification of cancer: class discovery and class prediction by gene expression with self-organizing maps: methods and applications to hematopoietic differentiation. Proc Nat Acad Sci USA *96*, 2907–2912. https://doi.org/10.1073/pnas.96.6.2907.

Gomez-Ospina, N., Scharenberg, S.G., Mostrel, N., Bak, R.O., Mantri, S., Quadros, R.M., Gurumurthy, C.B., Lee, C., Bao, G., Suarez, C.J., et al. (2019). Human genome-edited hematopoietic stem cells phenotypically correct Mucopolysaccharidosis type I. Nat. Commun. *10*, 1–14. https://doi.org/10.1038/s41467-019-11962-8.

Gomez Perdiguero, E., Klapproth, K., Schulz, C., Busch, K., Azzoni, E., Crozet, L., Garner, H., Trouillet, C., De Bruijn, M.F., Geissmann, F., et al. (2015). Tissue-resident macrophages originate from yolk-sac-derived erythro-myeloid progenitors. Nature *518*, 547–551. https://doi.org/10.1038/nature13989.

González, F., Zhu, Z., Shi, Z.D., Lelli, K., Verma, N., Li, Q. V., and Huangfu, D. (2014). An iCRISPR platform for rapid, multiplexable, and inducible genome editing in human pluripotent stem cells. Cell Stem Cell *15*, 215–226. https://doi.org/10.1016/j.stem.2014.05.018.

Goode, D.K., Obier, N., Vijayabaskar, M.S., Lie-A-Ling, M., Lilly, A.J., Hannah, R., Lichtinger, M., Batta, K., Florkowska, M., Patel, R., et al. (2016). Dynamic Gene Regulatory Networks Drive Hematopoietic Specification and Differentiation. Dev. Cell *36*, 572–587. https://doi.org/10.1016/j.devcel.2016.01.024.

Gossen, M., and Bujard, H. (1992). Tight control of gene expression in mammalian cells by tetracycline-responsive promoters. Proc. Natl. Acad. Sci. U. S. A. *89*, 5547–5551. https://doi.org/10.1073/pnas.89.12.5547.

Gossen, M., Freundlieb, S., Bender, G., Müller, G., Hillen, W., and Bujard, H. (1995). Transcriptional activation by tetracyclines in mammalian cells. Science (80-.). *268*, 1766–1769. https://doi.org/10.1126/science.7792603.

Grier, D.G., Thompson, A., Kwasniewska, A., McGonigle, G.J., Halliday, H.L., and Lappin, T.R. (2005a). The pathophysiology of HOX genes and their role in cancer. J. Pathol. *205*, 154–171. https://doi.org/10.1002/path.1710.

Grier, D.G., Thompson, A., Kwasniewska, A., McGonigle, G.J., Halliday, H.L., and Lappin, T.R. (2005b). The pathophysiology of HOX genes and their role in cancer. J. Pathol. *205*, 154–171. https://doi.org/10.1002/path.1710.

Guibentif, C., Rönn, R.E., Böiers, C., Lang, S., Saxena, S., Soneji, S., Enver, T., Karlsson, G., and Woods, N.B. (2017). Single-Cell Analysis Identifies Distinct Stages of Human Endothelial-to-Hematopoietic Transition. Cell Rep. *19*, 10–19. https://doi.org/10.1016/j.celrep.2017.03.023.

Gupta, P., Gurudutta, G.U., Saluja, D., and Tripathi, R.P. (2009). PU.1 and partners: Regulation of haematopoietic stem cell fate in normal and malignant haematopoiesis. J.

Cell. Mol. Med. *13*, 4349–4363. https://doi.org/10.1111/j.1582-4934.2009.00757.x.

Hadland, B.K., Huppert, S.S., Kanungo, J., Xue, Y., Jiang, R., Gridley, T., Conlon, R.A., Cheng, A.M., Kopan, R., and Longmore, G.D. (2004). A requirement for Notch1 distinguishes 2 phases of definitive hematopoiesis during development. Blood *104*, 3097–3105. https://doi.org/10.1182/blood-2004-03-1224.

Hammad, M., Rao, W., Smith, J.G.W., Anderson, D.G., Langer, R., Young, L.E., Barrett, D.A., Davies, M.C., Denning, C., and Alexander, M.R. (2016). Identification of polymer surface adsorbed proteins implicated in pluripotent human embryonic stem cell expansion. Biomater. Sci. *4*, 1381–1391. https://doi.org/10.1039/c6bm00214e.

Van Handel, B., Montel-Hagen, A., Sasidharan, R., Nakano, H., Ferrari, R., Boogard, C.J., Schredelseker, J., Wang, Y., Hunter, S., Org, T., et al. (2012). Scl Represses Cardiomyogenesis in Prospective Hemogenic Endothelium and Endocardium. Cell *150*, 590–605. https://doi.org/10.1016/j.cell.2012.06.026.

Hansen, K.D., Langmead, B., and Irizarry, R.A. (2012). BSmooth : from whole genome bisulfite sequencing reads to differentially methylated regions. Genome Biol. *13*, 1–10. .

Hansen, M., Varga, E., Aarts, C., Wust, T., and Kuijpers, T. (2018). Efficient production of erythroid , megakaryocytic and myeloid cells , using single cell-derived iPSC colony di ff erentiation. Stem Cell Res. *29*, 232–244. https://doi.org/10.1016/j.scr.2018.04.016.

Den Hartogh, S.C., Wolstencroft, K., Mummery, C.L., and Passier, R. (2016). A comprehensive gene expression analysis at sequential stages of in vitro cardiac differentiation from isolated MESP1-expressing- mesoderm progenitors. Sci. Rep. *6*, 1–16. https://doi.org/10.1038/srep19386.

Hazelbaker, D.Z., Beccard, A., Angelini, G., Mazzucato, P., Messana, A., Lam, D., Eggan, K., and Barrett, L.E. (2020). A multiplexed gRNA piggyBac transposon system facilitates efficient induction of CRISPRi and CRISPRa in human pluripotent stem cells. Sci. Rep. *10*, 1–10. https://doi.org/10.1038/s41598-020-57500-1.

Hockemeyer, D., Soldner, F., Beard, C., Gao, Q., Mitalipova, M., Dekelver, R.C., Katibah, G.E., Amora, R., Boydston, E.A., Zeitler, B., et al. (2009). Efficient targeting of expressed and silent genes in human ESCs and iPSCs using zinc-finger nucleases. Nat. Biotechnol. *27*, 851–857. https://doi.org/10.1038/nbt.1562.

Holkers, M., Maggio, I., Liu, J., Janssen, J.M., Miselli, F., Mussolino, C., Recchia, A., Cathomen, T., and Gonçalves, M.A.F.V. (2013). Differential integrity of TALE nuclease genes following adenoviral and lentiviral vector gene transfer into human cells. Nucleic Acids Res. *41*. https://doi.org/10.1093/nar/gks1446.

Holland, P.W.H., Booth, H.A.F., and Bruford, E.A. (2007). Classification and nomenclature of all human homeobox genes. BMC Biol. *5*, 1–28. https://doi.org/10.1186/1741-7007-5-47.

Hong, S.H., Werbowetski-Ogilvie, T., Ramos-Mejia, V., Lee, J.B., and Bhatia, M. (2010). Multiparameter comparisons of embryoid body differentiation toward human stem cell applications. Stem Cell Res. *5*, 120–130. https://doi.org/10.1016/j.scr.2010.04.007.

Hoogenkamp, M., Lichtinger, M., Krysinska, H., Lancrin, C., Clarke, D., Williamson, A., Mazzarella, L., Ingram, R., Jorgensen, H., Fisher, A., et al. (2009). Early chromatin unfolding by RUNX1: A molecular explanation for differential requirements during specification versus maintenance of the hematopoietic gene expression program. Blood *114*, 299–309. https://doi.org/10.1182/blood-2008-11-191890.

Hsu, P.D., Scott, D.A., Weinstein, J.A., Ran, F.A., Konermann, S., Agarwala, V., Li, Y., Fine, E.J., Wu, X., Shalem, O., et al. (2013). DNA targeting specificity of RNA-guided Cas9 nucleases. Nat. Biotechnol. *31*, 827–832. https://doi.org/10.1038/nbt.2647.

Huang, Y., Sitwala, K., Bronstein, J., Sanders, D., Dandekar, M., Collins, C., Robertson, G., MacDonald, J., Cezard, T., Bilenky, M., et al. (2012). Identification and characterization of Hoxa9 binding sites in hematopoietic cells. Blood *119*, 388–398. https://doi.org/10.1182/blood-2011-03-341081.

Iacovino, M., Chong, D., Szatmari, I., Hartweck, L., Rux, D., Caprioli, A., Cleaver, O., and Kyba, M. (2011). HoxA3 is an apical regulator of haemogenic endothelium. Nat. Publ. Gr. *13*, 72–79. https://doi.org/10.1038/ncb2137.

Irion, S., Luche, H., Gadue, P., Fehling, H.J., Kennedy, M., and Keller, G. (2007). Identification and targeting of the ROSA26 locus in human embryonic stem cells. Nat. Biotechnol. *25*, 1477–1482. https://doi.org/10.1038/nbt1362.

Ivanovs, A., Rybtsov, S., Welch, L., Anderson, R.A., Turner, M.L., and Medvinsky, A. (2011). Highly potent human hematopoietic stem cells first emerge in the intraembryonic

aorta-gonad-mesonephros region. J. Exp. Med. *208*, 2417–2427. https://doi.org/10.1084/jem.20111688.

Ivanovs, A., Rybtsov, S., Anderson, R.A., Turner, M.L., and Medvinsky, A. (2014). Identification of the niche and phenotype of the first human hematopoietic stem cells. Stem Cell Reports *2*, 449–456. https://doi.org/10.1016/j.stemcr.2014.02.004.

Ivanovs, A., Rybtsov, S., Ng, E.S., Stanley, E.G., Elefanty, A.G., and Medvinsky, A. (2017). Human haematopoietic stem cell development: from the embryo to the dish. Development *144*, 2323–2337. https://doi.org/10.1242/dev.134866.

Jackson, M., Axton, R.A., Taylor, A.H., Wilson, J.A., Gordon-Keylock, S.A.M., Kokkaliaris, K.D., Brickman, J.M., Schulz, H., Hummel, O., Hubner, N., et al. (2012). HOXB4 can enhance the differentiation of embryonic stem cells by modulating the hematopoietic niche. Stem Cells *30*, 150–160. https://doi.org/10.1002/stem.782.

Jackson, M., Ma, R., Taylor, A.H., Axton, R.A., Easterbrook, J., Kydonaki, M., Olivier, E., Marenah, L., Stanley, E.G., Elefanty, A.G., et al. (2016). Enforced Expression of HOXB4 in Human Embryonic Stem Cells Enhances the Production of Hematopoietic Progenitors but Has No Effect on the Maturation of Red Blood Cells. Stem Cells Transl. Med. *5*, 981–990. https://doi.org/10.5966/sctm.2015-0324.

Jeong, S., An, B., Kim, J.H., Han, H.W., Kim, J.H., Heo, H.R., Ha, K.S., Han, E.T., Park, W.S., and Hong, S.H. (2020). BMP4 and perivascular cells promote hematopoietic differentiation of human pluripotent stem cells in a differentiation stage-specific manner. Exp. Mol. Med. *52*, 56–65. https://doi.org/10.1038/s12276-019-0357-5.

Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J.A., and Charpentier, E. (2012). A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. Science (80-.). *337*, 816–821. https://doi.org/10.1126/science.1225829.

Jore, M.M., Lundgren, M., Duijn, E. Van, Bultema, J.B., Westra, E.R., Waghmare, S.P., Wiedenheft, B., Pul, Ü., Wurm, R., Wagner, R., et al. (2019). Articles Structural basis for CRISPR RNA-guided DNA recognition by Cascade. Nat. Publ. Gr. *18*. https://doi.org/10.1038/nsmb.2019.

Jung, H.S., Uenishi, G., Park, M.A., Thomson, J.A., Ong, I.M., Slukvin, I.I., Jung, H.S., Uenishi, G., Park, M.A., Liu, P., et al. (2021). SOX17 integrates HOXA and arterial

programs in hemogenic endothelium to drive definitive lympho- myeloid hematopoiesis. Cell Rep. *34*, 108758. https://doi.org/10.1016/j.celrep.2021.108758.

Kaimakis, P., De Pater, E., Eich, C., Kartalaei, P.S., Kauts, M.L., Vink, C.S., Van Der Linden, R., Jaegle, M., Yokomizo, T., Meijer, D., et al. (2016). Functional and molecular characterization of mouse Gata2-independent hematopoietic progenitors. Blood *127*, 1426–1437. https://doi.org/10.1182/blood-2015-10-673749.

Kallunki, T., Barisic, M., Jäättelä, M., and Liu, B. (2019). How to Choose the Right Inducible Gene Expression System for Mammalian Studies? Cells *8*, 1–16. https://doi.org/10.3390/cells8080796.

Kardel, M.D., and Eaves, C.J. (2012). Modeling human hematopoietic cell development from pluripotent stem cells. Exp. Hematol. *40*, 601–611. https://doi.org/10.1016/j.exphem.2012.04.001.

Kataoka, H., Hayashi, M., Nakagawa, R., Tanaka, Y., Izumi, N., Nishikawa, S., Jakt, M.L., Tarui, H., and Nishikawa, S.I. (2011). Etv2/ER71 induces vascular mesoderm from Flk1 +PDGFRα + primitive mesoderm. Blood *118*, 6975–6986. https://doi.org/10.1182/blood-2011-05-352658.

Kattman, S.J., Witty, A.D., Gagliardi, M., Dubois, N.C., Niapour, M., Hotta, A., and Ellis, J. (2010). Stage-Specific Optimization of Activin / Nodal and BMP Signaling Promotes Cardiac Differentiation of Mouse and Human Pluripotent Stem Cell Lines. Stem Cell *8*, 228–240. https://doi.org/10.1016/j.stem.2010.12.008.

Kauts, M.L., Vink, C.S., and Dzierzak, E. (2016). Hematopoietic (stem) cell development — how divergent are the roads taken? FEBS Lett. *590*, 3975–3986. https://doi.org/10.1002/1873-3468.12372.

Kennedy, M., D'Souza, S.L., Lynch-Kattman, M., Schwantz, S., and Keller, G. (2007). Development of the hemangioblast defines the onset of hematopoiesis in human ES cell differentiation cultures. Blood *109*, 2679–2687. https://doi.org/10.1182/blood-2006-09-047704.The.

Kennedy, M., Awong, G., Sturgeon, C.M., Ditadi, A., LaMotte-Mohs, R., Zúñiga-Pflücker, J.C., and Keller, G. (2012). T Lymphocyte Potential Marks the Emergence of Definitive Hematopoietic Progenitors in Human Pluripotent Stem Cell Differentiation Cultures. Cell

Rep. *2*, 1722–1735. https://doi.org/10.1016/j.celrep.2012.11.003.

Kessel, M., and Gruss, P. (1990). Murine developmental control genes. Science (80-.). *249*, 374–379. https://doi.org/10.1126/science.1974085.

Kieusseian, A., de la Grange, P.B., Burlen-Defranoux, O., Godin, I., and Cumano, A. (2012). Immature hematopoietic stem cells undergo maturation in the fetal liver. Dev. *139*, 3521–3530. https://doi.org/10.1242/dev.079210.

Kim, H., and Kim, J.S. (2014). A guide to genome engineering with programmable nucleases. Nat. Rev. Genet. *15*, 321–334. https://doi.org/10.1038/nrg3686.

Kim, Y.G., Cha, J., and Chandrasegaran, S. (1996). Hybrid restriction enzymes: Zinc finger fusions to Fok I cleavage domain. Proc. Natl. Acad. Sci. U. S. A. *93*, 1156–1160. https://doi.org/10.1073/pnas.93.3.1156.

Kissa, K., and Herbomel, P. (2010). Blood stem cells emerge from aortic endothelium by a novel type of cell transition. Nature *464*, 112–116. https://doi.org/10.1038/nature08761.

Klatt, D., Cheng, E., Hoffmann, Di., Santilli, G., Thrasher, A.J., Brendel, C., and Schambach, A. (2020). Differential Transgene Silencing of Myeloid-Specific Promoters in the AAVS1 Safe Harbor Locus of Induced Pluripotent Stem Cell-Derived Myeloid Cells. Hum. Gene Ther. *31*, 199–210. https://doi.org/10.1089/hum.2019.194.

Kmita, M., and Duboule, D. (2003). Organizing axes in time and space; 25 years of colinear tinkering. Science (80-.). *301*, 331–333. https://doi.org/10.1126/science.1085753.

Ko, K.H., Kwan Lam, Q.L., Zhang, M., Yen Wong, C.K., Chun Lo, C.K., Kahmeyer-Gabbe, M., Tsang, W.H., Tsang, S.L., Chan, L.C., Sham, M.H., et al. (2007). Hoxb3 deficiency impairs B lymphopoiesis in mouse bone marrow. Exp. Hematol. *35*, 465–475. https://doi.org/10.1016/j.exphem.2006.10.014.

Kotin, R.M., Linden, R.M., and Berns, K.I. (1992). Characterization of a preferred site on human chromosome 19q for integration of adeno-associated virus DNA by non-homologous recombination. EMBO J. *11*, 5071–5078. .

Koyunlar, C., Gioacchino, E., Vadgama, D., de Looper, H., Zink, J., Hoogenboezem, R.,

Havermans, M., Bindels, E., Dzierzak, E., Touw, I.P., et al. (2022). Gata2-regulated Gfi1b expression controls endothelial programming during endothelial-to-hematopoietic transition. Cold Spring Harb. Lab. https://doi.org/10.14341/pmpe-2022-10.

Kroon, E., Krosl, J., Thorsteinsdottir, U., Baban, S., Buchberg, A.M., and Sauvageau, G. (1998). Hoxa9 transforms primary bone marrow cells through specific collaboration with Meis1a but not Pbx1b. EMBO J. *17*, 3714–3725. https://doi.org/10.1093/emboj/17.13.3714.

Kruger, C., and Kappen, C. (2010). Expression of cartilage developmental genes in Hoxc8-and Hoxd4-transgenic mice. PLoS One *5*. https://doi.org/10.1371/journal.pone.0008978.

Kumaravelu, P., Hook, L., Morrison, A.M., Ure, J., Zhao, S., and Zuyev, S. (2002). Quantitative developmental anatomy of definitive haematopoietic stem cells/long-term repopulating units (HSC/RUs): role of the aorta-gonad-mesonephros (AGM) region and the yolk sac in colonisation of the mouse embryonic liver. Development *129*, 4891–4899.
.

Kurosawa, H. (2007). Methods for inducing embryoid body formation: in vitro differentiation system of embryonic stem cells. J. Biosci. Bioeng. *103*, 389–398. https://doi.org/10.1263/jbb.103.389.

Kyba, M., Perlingeiro, R.C.R., and Daley, G.Q. (2002). HoxB4 confers definitive lymphoid-myeloid engraftment potential on embryonic stem cell and yolk sac hematopoietic progenitors. Cell *109*, 29–37. https://doi.org/10.1016/S0092-8674(02)00680-3.

Lacaud, G., and Kouskoff, V. (2017). Hemangioblast, hemogenic endothelium, and primitive versus definitive hematopoiesis. Exp. Hematol. *49*, 19–24. https://doi.org/10.1016/j.exphem.2016.12.009.

Laflamme, M.A., Chen, K.Y., Naumova, A. V, Muskheli, V., Fugate, J.A., Dupras, S.K., Reinecke, H., Xu, C., Hassanipour, M., Police, S., et al. (2007). Cardiomyocytes derived from human embryonic stem cells in pro-survival factors enhance function of infarcted rat hearts. *25*, 1015–1024. https://doi.org/10.1038/nbt1327.

Lancaster, J.N., Li, Y., and Ehrlich, L.I.R. (2018). Chemokine-Mediated Choreography of

Thymocyte Development and Selection. Trends Immunol. *39*, 86–98. https://doi.org/10.1016/j.it.2017.10.007.

Lancrin, C., Sroczynska, P., Stephenson, C., Allen, T., Kouskoff, V., and Lacaud, G. (2009). The haemangioblast generates haematopoietic cells through a haemogenic endothelium stage. Nat. Lett. *457*, 892–896. https://doi.org/10.1038/nature07679.

Lancrin, C., Mazan, M., Stefanska, M., Patel, R., Lichtinger, M., Costa, G., Vargel, Ö., Wilson, N.K., Möröy, T., Bonifer, C., et al. (2012). GFI1 and GFI1B control the loss of endothelial identity of hemogenic endothelium during hematopoietic commitment. Blood *120*, 314–322. https://doi.org/10.1182/blood-2011-10-386094.

Lange, L., Hoffmann, D., Schwarzer, A., Ha, T.C., Philipp, F., Lenz, D., Morgan, M., and Schambach, A. (2020). Inducible Forward Programming of Human Pluripotent Stem Cells to Hemato-endothelial Progenitor Cells with Hematopoietic Progenitor Potential. Stem Cell Reports *14*, 122–137. https://doi.org/10.1016/j.stemcr.2019.11.005.

Lange, L., Morgan, M., and Schambach, A. (2021). The hemogenic endothelium: a critical source for the generation of PSC-derived hematopoietic stem and progenitor cells. Cell. Mol. Life Sci. *78*, 4143–4160. https://doi.org/10.1007/s00018-021-03777-y.

Lappin, T.R.J., Grier, D.G., Thompson, A., and Halliday, H.L. (2006). HOX genes: Seductive science, mysterious mechanisms. Ulster Med. J. *75*, 23–31. .

Laurent, L., Wong, E., Li, G., Huynh, T., Tsirigos, A., Ong, C.T., Low, H.M., Sung, K.W.K., Rigoutsos, I., Loring, J., et al. (2010). Dynamic changes in the human methylome during differentiation. Genome Res. *20*, 320–331. https://doi.org/10.1101/gr.101907.109.

Lawrence, H.J., Helgason, C.D., Sauvageau, G., Fong, S., Izon, D.J., Humphries, R.K., and Largman, C. (1997). Mice bearing a targeted interruption of the homeobox gene HOXA9 have defects in myeloid, erythroid, and lymphoid hematopoiesis. Blood *89*, 1922–1930. .

Lawrence, H.J., Christensen, J., Fong, S., Hu, Y., Weissman, I., Humphries, R.K., Largman, C., Lawrence, H.J., Christensen, J., Fong, S., et al. (2005). Loss of expression of the Hoxa-9 homeobox gene impairs the proliferation and repopulating ability of hematopoietic stem cells. Blood *106*, 3988–3994. https://doi.org/10.1182/blood-2005-05-2003.

Lebert-Ghali, C.-E., Fournier, M., Kettyle, L., Thompson, A., Sauvageau, G., and Bijl, J.J. (2016). Hoxa cluster genes determine the proliferative activity of adult mouse hematopoietic stem and progenitor cells. Blood *127*, 87–91. https://doi.org/10.1182/blood-2015-02-626390.The.

Lebert-Ghali, C.-E., Thompson, A., Melichar, H.J., and Bijl, J.J. (2017). Targeted deletion of the Hoxa cluster affects B lymphopoiesis through depletion of early lymphoid progenitors. Exp. Hematol. *50*, 84–89. https://doi.org/10.1016/j.exphem.2017.03.003.

Lebert-Ghali, C.E., Fournier, M., Dickson, G.J., Thompson, A., Sauvageau, G., and Bijl, J.J. (2010). HoxA cluster is haploinsufficient for activity of hematopoietic stem and progenitor cells. Exp. Hematol. *38*, 1074-1086.e5. https://doi.org/10.1016/j.exphem.2010.07.006.

Li, E., and Zhang, Y. (2014). DNA Methylation in Mammals. Cold Spring Harb. Prespectives Biol.

Li, Y., and Tollefsbol, T.O. (2011). DNA methylation detection: Bisulfite genomic sequencing analysis. Methods Enzymol. *791*, 11–21. https://doi.org/10.1007/978-1-61779-316-5.

Lian, X., Zhang, J., Zhu, K., Kamp, T.J., and Palecek, S.P. (2013). Insulin inhibits cardiac mesoderm, not mesendoderm, formation during cardiac differentiation of human pluripotent stem cells and modulation of canonical wnt signaling can rescue this inhibition. Stem Cells *31*, 447–457. https://doi.org/10.1002/stem.1289.

Lichtinger, M., Ingram, R., Hannah, R., Müller, D., Clarke, D., Assi, S.A., Lie-A-Ling, M., Noailles, L., Vijayabaskar, M.S., Wu, M., et al. (2012). RUNX1 reshapes the epigenetic landscape at the onset of haematopoiesis. EMBO J. *31*, 4318–4333. https://doi.org/10.1038/emboj.2012.275.

Lindblad, O., Chougule, R.A., Moharram, S.A., Kabir, N.N., Sun, J., Kazi, J.U., and Rönnstrand, L. (2015). The role of HOXB2 and HOXB3 in acute myeloid leukemia. Biochem. Biophys. Res. Commun. *467*, 742–747. https://doi.org/10.1016/j.bbrc.2015.10.071.

Lisanti, S., von Zglinicki, T., and Mathers, J.C. (2012). Standardization and quality controls for the methylated DNA immunoprecipitation technique. Epigenetics *7*, 615–

625. https://doi.org/10.4161/epi.20028.

Liu, F., Walmsley, M., Rodaway, A., and Patient, R. (2008). Fli1 Acts at the Top of the Transcriptional Network Driving Blood and Endothelial Development. Curr. Biol. *18*, 1234–1240. https://doi.org/10.1016/j.cub.2008.07.048.

Liu, R., Paxton, W.A., Choe, S., Ceradini, D., Martin, S.R., Horuk, R., MacDonald, M.E., Stuhlmann, H., Koup, R.A., and Landau, N.R. (1996). Homozygous defect in HIV-1 coreceptor accounts for resistance of some multiply-exposed individuals to HIV-1 infection. Cell *86*, 367–377. https://doi.org/10.1016/S0092-8674(00)80110-5.

Lizama, C.O., Hawkins, J.S., Schmitt, C.E., Bos, F.L., Zape, J.P., Cautivo, K.M., Pinto, H.B., Rhyner, A.M., Yu, H., Donohoe, M.E., et al. (2015). Repression of arterial genes in hemogenic endothelium is sufficient for haematopoietic fate acquisition. Nat. Commun. https://doi.org/10.1038/ncomms8739.

Loew, R., Heinz, N., Hampf, M., Bujard, H., and Gossen, M. (2010). Improved Tet-responsive promoters with minimized background expression. BMC Biotechnol. *10*, 1–13. https://doi.org/10.1186/1472-6750-10-81.

Lombardo, A., Genovese, P., Beausejour, C.M., Colleoni, S., Lee, Y.L., Kim, K.A., Ando, D., Urnov, F.D., Galli, C., Gregory, P.D., et al. (2007). Gene editing in human stem cells using zinc finger nucleases and integrase-defective lentiviral vector delivery. Nat. Biotechnol. *25*, 1298–1306. https://doi.org/10.1038/nbt1353.

Lombardo, A., Cesana, D., Genovese, P., Di Stefano, B., Provasi, E., Colombo, D.F., Neri, M., Magnani, Z., Cantore, A., Lo Riso, P., et al. (2011). Site-specific integration and tailoring of cassette design for sustainable gene transfer. Nat. Methods *8*, 861–869. https://doi.org/10.1038/nmeth.1674.

Luff, S.A., Creamer, J.P., Valsoni, S., Dege, C., Scarfò, R., Dacunto, A., Cascione, S., Randolph, L.N., Cavalca, E., Merelli, I., et al. (2022). Identification of a retinoic acid-dependent haemogenic endothelial progenitor from human pluripotent stem cells. Nat. Cell Biol. *24*, 616–624. https://doi.org/10.1038/s41556-022-00898-9.

Lundin, V., Sugden, W.W., Theodore, L.N., Sousa, P.M., Han, A., Chou, S., Wrighton, P.J., Cox, A.G., Ingber, D.E., Goessling, W., et al. (2020). YAP Regulates Hematopoietic Stem Cell Formation in Response to the Biomechanical Forces of Blood Flow. Dev. Cell

52, 446-460.e5. https://doi.org/10.1016/j.devcel.2020.01.006.

Luo, Z., Rhie, S.K., and Farnham, P.J. (2019). The enigmatic hox genes: Can we crack their code? Cancers (Basel). *11*. https://doi.org/10.3390/cancers11030323.

Magnusson, M., Brun, A.C.M., Miyake, N., Larsson, J., Ehinger, M., Bjornsson, J.M., Wutz, A., Sigvardsson, M., and Karlsson, S. (2007a). HOXA10 is a critical regulator for hematopoietic stem cells and erythroid/megakaryocyte development. Hematopoiesis *109*, 3687–3696. https://doi.org/10.1182/blood-2006-10-054676.The.

Magnusson, M., Brun, A.C.M., Lawrence, H.J., and Karlsson, S. (2007b). Hoxa9/hoxb3/hoxb4 compound null mice display severe hematopoietic defects. Exp. Hematol. *35*, 1421–1428. https://doi.org/10.1016/j.exphem.2007.05.011.

Makarova, K.S., Haft, D.H., Barrangou, R., Brouns, S.J.J., Charpentier, E., Horvath, P., Moineau, S., Mojica, F.J.M., Wolf, Y.I., Yakunin, A.F., et al. (2011). Evolution and classification of the CRISPR-Cas systems. Nat. Rev. Microbiol. *9*, 467–477. https://doi.org/10.1038/nrmicro2577.

Makarova, K.S., Wolf, Y.I., Alkhnbashi, O.S., Costa, F., Shah, S.A., Saunders, S.J., Barrangou, R., Brouns, S.J.J., Charpentier, E., Haft, D.H., et al. (2015). An updated evolutionary classification of CRISPR-Cas systems. Nat. Publ. Gr. *13*, 722–736. https://doi.org/10.1038/nrmicro3569.

Mali, P., Yang, L., Esvelt, K.M., Aach, J., Guell, M., DiCarlo, J.E., Norville, J.E., and Church, G.M. (2013). RNA-guided human genome engineering via Cas9. Science (80-.). *339*, 823–826. https://doi.org/10.1126/science.1232033.

Mallo, M., Wellik, D.M., and Deschamps, J. (2010). Hox genes and regional patterning of the vertebrate body plan. Dev. Biol. *344*, 7–15. https://doi.org/10.1016/j.ydbio.2010.04.024.

Mandegar, M.A., Huebsch, N., Frolov, E.B., Shin, E., Truong, A., Olvera, M.P., Chan, A.H., Miyaoka, Y., Holmes, K., Spencer, C.I., et al. (2016). CRISPR Interference Efficiently Induces Specific and Reversible Gene Silencing in Human iPSCs. Cell Stem Cell *18*, 541–553. https://doi.org/10.1016/j.stem.2016.01.022.

Mann, R.S., Lelli, K.M., and Joshi, R. (2009). Hox Specificty: Unique Roles for Cofactors

and Collaborators. Curr. Top. Dev. Biol. *88*, 63–101. https://doi.org/10.1016/S0070-2153(09)88003-4.Hox.

Mark, M., Rijli, F.M., and Chambon, P. (1997). Homeobox Genes in Embryogenesis and Pathogenesis. Pediatr. Res. *42*, 421–429. .

Matsuo, K., Lepinski, A., Chavez, R.D., Barruet, E., Pereira, A., Moody, T.A., Ton, A.N., Sharma, A., Hellman, J., Tomoda, K., et al. (2021). ACVR1 R206H extends inflammatory responses in human induced pluripotent stem cell-derived macrophages ☆. Bone *153*, 116129. https://doi.org/10.1016/j.bone.2021.116129.

McGrath, K.E., Frame, J.M., Fegan, K.H., Bowen, J.R., Conway, S.J., Catherman, S.C., Kingsley, P.D., Koniski, A.D., and Palis, J. (2015). Distinct Sources of Hematopoietic Progenitors Emerge before HSCs and Provide Functional Blood Cells in the Mammalian Embryo. Cell Rep. *11*, 1892–1904. https://doi.org/10.1016/j.celrep.2015.05.036.

Merabet, S., and Mann, R.S. (2016). To Be Specific or Not: The Critical Relationship Between Hox And TALE Proteins. Trends Genet. *32*, 334–347. https://doi.org/10.1016/j.tig.2016.03.004.

Mikkola, H.K.A., Klintman, J., Yang, H., Hock, H., Schlaeger, T.M., Fujiwara, Y., and Orkin, S.H. (2003). Haematopoietic stem cells retain long-term repopulating activity and multipotency in the absence of stem-cell leukaemia SCL/tal-1 gene. Nature *421*, 547–551. https://doi.org/10.1038/nature01345.

Miller, J.C., Holmes, M.C., Wang, J., Guschin, D.Y., Lee, Y.L., Rupniewski, I., Beausejour, C.M., Waite, A.J., Wang, N.S., Kim, K.A., et al. (2007). An improved zinc-finger nuclease architecture for highly specific genome editing. Nat. Biotechnol. *25*, 778–785. https://doi.org/10.1038/nbt1319.

Miller, J.C., Tan, S., Qiao, G., Barlow, K.A., Wang, J., Xia, D.F., Meng, X., Paschon, D.E., Leung, E., Hinkley, S.J., et al. (2011). A TALE nuclease architecture for efficient genome editing. Nat. Biotechnol. *29*, 143–150. https://doi.org/10.1038/nbt.1755.

Mills, J.A., Paluru, P., Weiss, M.J., Gadue, P., and French, D.L. (2014). Hematopoietic Differentiation of Pluripotent Stem Cells in Culture. In Hematopoietic Stem Cells Protocols, pp. 181–194.

Minegishi, N., Ohta, J., Yamagiwa, H., Suzuki, N., Kawauchi, S., Zhou, Y., Takahashi, S., Hayashi, N., Engel, J.D., and Yamamoto, M. (1999). The mouse GATA-2 gene is expressed in the para-aortic splanchnopleura and aorta-gonads and mesonephros region. Blood *93*, 4196–4207. https://doi.org/10.1182/blood.v93.12.4196.

Mohn, F., Weber, M., Schubeler, D., and Roloff, T.-C. (2009). Methylated DNA Immunoprecipitation (MeDIP). In DNA Methylation: Methods and Protocols, pp. 55–64.

Mohtashami, M., Shah, D.K., Nakase, H., Kianizad, K., Petrie, H.T., and Zúñiga-Pflücker, J.C. (2010). Direct Comparison of Dll1- and Dll4-Mediated Notch Activation Levels Shows Differential Lymphomyeloid Lineage Commitment Outcomes. J. Immunol. *185*, 867–876. https://doi.org/10.4049/jimmunol.1000782.

Moignard, V., MacAulay, I.C., Swiers, G., Buettner, F., Schütte, J., Calero-Nieto, F.J., Kinston, S., Joshi, A., Hannah, R., Theis, F.J., et al. (2013). Characterization of transcriptional networks in blood stem and progenitor cells using high-throughput single-cell gene expression analysis. Nat. Cell Biol. *15*, 363–372. https://doi.org/10.1038/ncb2709.

Moore, L.D., Le, T., and Fan, G. (2013). DNA methylation and its basic function. Neuropsychopharmacology *38*, 23–38. https://doi.org/10.1038/npp.2012.112.

Mosqueira, D., Mannhardt, I., Bhagwan, J.R., Lis-slimak, K., Katili, P., Scott, E., Hassan, M., Prondzynski, M., Harmer, S.C., Tinker, A., et al. (2018). CRISPR / Cas9 editing in human pluripotent stem cell-cardiomyocytes highlights arrhythmias , hypocontractility , and energy depletion as potential therapeutic targets for hypertrophic cardiomyopathy. *44*, 3879–3892. https://doi.org/10.1093/eurheartj/ehy249.

Motazedian, A., Bruveris, F.F., Kumar, S. V, Schiesser, J. V, Chen, T., Ng, E.S., Chidgey, A.P., Wells, C.A., Elefanty, A.G., and Stanley, E.G. (2020). Multipotent RAG1+ progenitors emerge directly from haemogenic endothelium in human pluripotent stem cell-derived haematopoietic organoids. Nat. Cell Biol. *22*, 60–73. https://doi.org/10.1038/s41556-019-0445-8.

Müller, A.M., Medvinsky, A., Strouboulis, J., Grosveld, F., and Dzierzakt, E. (1994). Development of hematopoietic stem cell activity in the mouse embryo. Immunity *1*, 291–301. https://doi.org/10.1016/1074-7613(94)90081-7.

Muragaki, Y., Mundlos, S., Upton, J., and Olsen, B.R. (1996). Altered growth and branching patterns in synpolydactyly caused by mutations in HOXD13. Science (80-.). *272*, 548–551. https://doi.org/10.1126/science.272.5261.548.

Nafria, M., Bonifer, C., Guy, E., Ng, S., Nafria, M., Bonifer, C., Stanley, E.G., and Ng, E.S. (2020a). Protocol for the Generation of Definitive Hematopoietic Progenitors from Human Pluripotent Stem Cells Protocol for the Generation of Definitive Hematopoietic Progenitors from Human Pluripotent Stem Cells. STAR Protoc. *1*. https://doi.org/10.1016/j.xpro.2020.100130.

Nafria, M., Keane, P., Ng, E.S., Stanley, E.G., Elefanty, A.G., and Bonifer, C. (2020b). Expression of RUNX1-ETO Rapidly Alters the Chromatin Landscape and Growth of Early Human Myeloid Precursor Cells. Cell Rep. *31*, 107691. https://doi.org/10.1016/j.celrep.2020.107691.

Nakajima-takagi, Y., Osawa, M., Oshima, M., Takagi, H., Miyagi, S., Endoh, M., Endo, T.A., Takayama, N., Eto, K., Toyoda, T., et al. (2013). Role of SOX17 in hematopoietic development from human embryonic stem cells. Blood *121*, 447–458. https://doi.org/10.1182/blood-2012-05-431403.There.

Nakamura, T., Largaespada, D.A., Lee, M.P., Johnson, L.A., Ohyashiki, K., Toyama, K., Chen, S.-J., Willman, C.L., Chen, I.-M., Feinberg, A.P., et al. (1996). Fusion of the nucleoporin gene NUP98 to HOXA9 by the chromosome translocation t(7;11)(p15;p15) in human myeloid leukaemia. Nat. Genet. *12*, 154–158. .

Navarro-Montero, O., Ayllon, V., Lamolda, M., López-Onieva, L., Montes, R., Bueno, C., Ng, E., Guerrero-Carreno, X., Romero, T., Romero-Moya, D., et al. (2017). RUNX1c Regulates Hematopoietic Differentiation of Human Pluripotent Stem Cells Possibly in Cooperation with Proinflammatory Signaling. Stem Cells *35*, 2253–2266. https://doi.org/10.1002/stem.2700.

Ng, E.S., Davis, R.P., Azzola, L., Stanley, E.G., and Elefanty, A.G. (2005). Forced aggregation of defined numbers of human embryonic stem cells into embryoid bodies fosters robust, reproducible hematopoietic differentiation. Blood *106*, 1601–1603. https://doi.org/10.1182/blood-2005-03-0987.

Ng, E.S., Davis, R., Stanley, E.G., and Elefanty, A.G. (2008). A protocol describing the use of a recombinant protein-based, animal product-free medium (APEL) for human

embryonic stem cell differentiation as spin embryoid bodies. Nat. Protoc. *3*, 768–776. https://doi.org/10.1038/nprot.2008.42.

Ng, E.S., Azzola, L., Bruveris, F.F., Calvanese, V., Phipson, B., Vlahos, K., Hirst, C., Jokubaitis, V.J., Yu, Q.C., Maksimovic, J., et al. (2016). Differentiation of human embryonic stem cells to HOXA + hemogenic vasculature that resembles the aorta-gonad-mesonephros. Nat. Biotechnol. *34*, 1168–1179. https://doi.org/10.1038/nbt.3702.

Nitzahn, M., Truong, B., Khoja, S., Vega-crespo, A., and Le, C. (2021). CRISPR-Mediated Genomic Addition to CPS1 Deficient iPSCs is Insufficient to Restore Nitrogen Homeostasis. *94*, 545–557. .

Niwa, A., Heike, T., Umeda, K., Oshima, K., Kato, I., Sakai, H., Suemori, H., Nakahata, T., and Saito, M.K. (2011). A Novel Serum-Free Monolayer Culture for Orderly Hematopoietic Differentiation of Human Pluripotent Cells via Mesodermal Progenitors. PLoS One *6*, e22261. https://doi.org/10.1371/journal.pone.0022261.

Niwa, H., Miyazaki, J., and Smith, A.G. (2000). Quantitative expression of Oct-3 / 4 defines differentiation , dedifferentiation or self-renewal of ES cells. Nat. Genet. *24*, 2–6. .

North, T., Gu, T.L., Stacy, T., Wang, Q., Howard, L., Binder, M., Marín-Padilla, M., and Speck, N.A. (1999). Cbfa2 is required for the formation of intra-aortic hematopoietic clusters. Development *126*, 2563–2575. https://doi.org/10.1242/dev.126.11.2563.

North, T.E., de Bruijn, M.F.T.R., Stacy, T., Talebian, L., Lind, E., Robin, C., Binder, M., Dzierzak, E., Speck, N.A., Bjornson, C.R.., et al. (2002). Runx1 expression marks long-term repopulating hematopoietic stem cells in the midgestation mouse embryo. Immunity *16*, 661–672. https://doi.org/10.1016/s1074-7613(02)00296-0.

Nostro, M.C., Cheng, X., Keller, G.M., and Gadue, P. (2008). Wnt, Activin, and BMP Signaling Regulate Distinct Stages in the Developmental Pathway from Embryonic Stem Cells to Blood. Cell Stem Cell *2*, 60–71. https://doi.org/10.1016/j.stem.2007.10.011.

Nottingham, W.T., Jarratt, A., Burgess, M., Speck, C.L., Cheng, J.F., Prabhakar, S., Rubin, E.M., Li, P.S., Sloane-Stanley, J., Kong-A-San, J., et al. (2007). Runx1-mediated hematopoietic stem-cell emergence is controlled by a Gata/Ets/SCL-regulated enhancer. Blood *110*, 4188–4197. https://doi.org/10.1182/blood-2007-07-100883.

Oberlin, E., Hafny, B. El, Petit-Cocault, L., and Souyri, M. (2010). Definitive human and mouse hematopoiesis originates from the embryonic endothelium: A new class of HSCs based on VE-cadherin expression. Int. J. Dev. Biol. *54*, 1165–1173. https://doi.org/10.1387/ijdb.103121eo.

Oceguera-Yanez, F., Kim, S. Il, Matsumoto, T., Tan, G.W., Xiang, L., Hatani, T., Kondo, T., Ikeya, M., Yoshida, Y., Inoue, H., et al. (2016). Engineering the AAVS1 locus for consistent and scalable transgene expression in human iPSCs and their differentiated derivatives. Methods *101*, 43–55. https://doi.org/10.1016/j.ymeth.2015.12.012.

Ogata, T., Kozuka, T., and Kanda, T. (2003). Identification of an insulator in AAVS1, a preferred region for integration of adeno-associated virus DNA. J. Virol. *77*, 9000–9007. https://doi.org/10.1128/JVI.77.16.9000.

Oguro, H. (2019). The roles of cholesterol and its metabolites in normal and malignant hematopoiesis. Front. Endocrinol. (Lausanne). *10*, 1–15. https://doi.org/10.3389/fendo.2019.00204.

Okuda, T., Van Deursen, J., Hiebert, S.W., Grosveld, G., and Downing, J.R. (1996). AML1, the target of multiple chromosomal translocations in human leukemia, is essential for normal fetal liver hematopoiesis. Cell *84*, 321–330. https://doi.org/10.1016/S0092-8674(00)80986-1.

Olson, E.N., Nes, V., and Olson, E.N. (2006). Gene Regulatory Networks in the Evolution and Development of the Heart Published by : American Association for the Advancement of Science Linked references are available on JSTOR for this article : Gene Regulatory Networks in the Evolution and Development. *313*, 1922–1927. .

Ordovas, L., Boon, R., Pistoni, M., Chen, Y., Wolfs, E., Guo, W., Sambathkumar, R., Bobis-wozowicz, S., Helsen, N., Vanhove, J., et al. (2015). Efficient Recombinase-Mediated cassette Exchange in hPSCs to Study the Hepatocyte Lineage Reveals AAVS1 Locus-Mediated Transgene Inhibition. Stem Cell Reports *5*, 918–931. https://doi.org/10.1016/j.stemcr.2015.09.004.

Org, T., Duan, D., Ferrari, R., Montel-Hagen, A., Van Handel, B., Kerényi, M.A., Sasidharan, R., Rubbi, L., Fujiwara, Y., Pellegrini, M., et al. (2015). Scl binds to primed enhancers in mesoderm to regulate hematopoietic and cardiac fate divergence. EMBO J. *34*, 759–777. https://doi.org/10.15252/embj.201490542.

Osawa, M., Yamaguchi, T., Nakamura, Y., Kaneko, S., Onodera, M., Sawada, K.I., Jegalian, A., Wu, H., Nakauchi, H., and Iwama, A. (2002). Erythroid expansion mediated by the Gfi-1B zinc finger protein: Role in normal hematopoiesis. Blood *100*, 2769–2777. https://doi.org/10.1182/blood-2002-01-0182.

Oyer, J.A., Chu, A., Brar, S., and Turker, M.S. (2009). Aberrant epigenetic silencing is triggered by a transient reduction in gene expression. PLoS One *4*. https://doi.org/10.1371/journal.pone.0004832.

Paige, S.L., Osugi, T., Afanasiev, O.K., Pabon, L., Reinecke, H., and Murry, C.E. (2010). Endogenous Wnt / b -Catenin Signaling Is Required for Cardiac Differentiation in Human Embryonic Stem Cells. *5*. https://doi.org/10.1371/journal.pone.0011134.

Palis, J., and Yoder, M.C. (2001). Yolk-sac hematopoiesis: The first blood cells of mouse and man. Exp. Hematol. *29*, 927–936. https://doi.org/10.1016/S0301-472X(01)00669-5.

Palis, J., Robertson, S., Kennedy, M., Wall, C., and Keller, G. (1999). Development of erythroid and myeloid progenitors in the yolk sac and embryo proper of the mouse. Development *126*, 5073–5084. https://doi.org/10.1242/dev.126.22.5073.

Panagopoulos, I., Isaksson, M., Billström, R., Strömbeck, B., Mitelman, F., and Johansson, B. (2003). Fusion of the NUP98 gene and the homeobox gene HOXC13 in acute myeloid leukemia with t(11;12)(p15;q13). Genes Chromosom. Cancer *36*, 107–112. https://doi.org/10.1002/gcc.10139.

Papapetrou, E.P., and Schambach, A. (2016). Gene insertion into genomic safe harbors for human gene therapy. Mol. Ther. *24*, 678–684. https://doi.org/10.1038/mt.2016.38.

Park, M.A., Kumar, A., Jung, H.S., Uenishi, G., Moskvin, O. V., Thomson, J.A., and Slukvin, I.I. (2018). Activation of the Arterial Program Drives Development of Definitive Hemogenic Endothelium with Lymphoid Potential. Cell Rep. *23*, 2467–2481. https://doi.org/10.1016/j.celrep.2018.04.092.

de Pater, E., Kaimakis, P., Vink, C.S., Yokomizo, T., Yamada-Inagawa, T., van der Linden, R., Kartalaei, P.S., Camper, S.A., Speck, N., and Dzierzak, E. (2013). Gata2 is required for HSC generation and survival. J. Exp. Med. *210*, 2843–2850. https://doi.org/10.1084/jem.20130751.

Pattanayak, V., Lin, S., Guilinger, J.P., Doudna, J.A., Liu, D.R., Biology, C., Division, B., and Berkeley, L. (2014). High-throughput profiling of off-target DNA cleavage reveals RNA-programmed Cas9 nuclease specificity. Nat. Biotechnol. *31*, 839–843. https://doi.org/10.1038/nbt.2673.High-throughput.

Pawlowski, M., Ortmann, D., Bertero, A., Tavares, J.M., Pedersen, R.A., Vallier, L., and Kotter, M.R.N. (2017). Inducible and Deterministic Forward Programming of Human Pluripotent Stem Cells into Neurons, Skeletal Myocytes and Oligodendrocytes. Stem Cell Reports *8*, 803–812. https://doi.org/10.1016/j.stemcr.2017.02.016.

Pearson, S., Sroczynska, P., Lacaud, G., and Kouskoff, V. (2008). The stepwise specification of embryonic stem cells to hematopoietic fate is driven by sequential exposure to. Development *135*, 1525–1535. https://doi.org/10.1242/dev.011767.

Pearson, S., Cuvertino, S., Fleury, M., Lacaud, G., and Kouskoff, V. (2015). In vivo repopulating activity emerges at the onset of hematopoietic specification during embryonic stem cell differentiation. Stem Cell Reports *4*, 431–444. https://doi.org/10.1016/j.stemcr.2015.01.003.

Peaslee, C., Esteva-Font, C., Su, T., Munoz-Howell, A., Duwaerts, C.C., Liu, Z., Rao, S., Liu, K., Medina, M., Sneddon, J.B., et al. (2021). Doxycycline Significantly Enhances Induction of Induced Pluripotent Stem Cells to Endoderm by Enhancing Survival Through Protein Kinase B Phosphorylation. Hepatology *74*, 2102–2117. https://doi.org/10.1002/hep.31898.

Pellenz, S., Phelps, M., Tang, W., Hovde, B.T., Sinit, R.B., Fu, W., Li, H., Chen, E., and Monnat, R.J. (2019). New Human Chromosomal Sites with “Safe Harbor” Potential for Targeted Transgene Insertion. Hum. Gene Ther. *30*, 814–828. https://doi.org/10.1089/hum.2018.169.

Pfaff, N., Lachmann, N., Ackermann, M., Kohlscheen, S., Brendel, C., Maetzig, T., Niemann, H., Antoniou, M.N., Grez, M., Schambach, A., et al. (2013). A ubiquitous chromatin opening element prevents transgene silencing in pluripotent stem cells and their differentiated progeny. Stem Cells *31*, 488–499. https://doi.org/10.1002/stem.1316.

Pineault, K.M., and Wellik, D.M. (2014). Hox genes and limb musculoskeletal development. Curr. Osteoporos. Rep. *12*, 420–427. https://doi.org/10.1007/s11914-014-0241-0.

Pineault, N., Helgason, C.D., Lawrence, H.J., and Humphries, R.K. (2002). Differential expression of Hox, Meis1, and Pbx1 genes in primitive cells throughout murine hematopoietic ontogeny. Exp. Hematol. *30*, 49–57. https://doi.org/10.1016/S0301-472X(01)00757-3.

Porcher, C., Swat, W., Rockwell, K., Fujiwara, Y., Alt, F.W., and Orkin, S.H. (1996). The T cell leukemia oncoprotein SCL/tal-1 is essential for development of all hematopoietic lineages. Cell *86*, 47–57. https://doi.org/10.1016/S0092-8674(00)80076-8.

Prashad, S.L., Calvanese, V., Magnusson, M., Katri, H., Prashad, S.L., Calvanese, V., Yao, C.Y., Kaiser, J., and Wang, Y. (2015). GPI-80 Defines Self-Renewal Ability in Hematopoietic Stem Cells during Human Development. Cell Stem Cell *16*, 80–87. https://doi.org/10.1016/j.stem.2014.10.020.

Qian, K., Huang, C.T.-L., Chen, H., Blackbourn IV, L.W., Chen, Y., Cao, J., Yao, L., Sauvey, C., Du, Z., and Zhang, S.-C. (2014). A Simple and Efficient System for Regulating Gene Expression in Human Pluripotent Stem Cells and Derivatives. Stem Cells *32*, 1230–1238. https://doi.org/10.1038/jid.2014.371.

Qin, J.Y., Zhang, L., Clift, K.L., Hulur, I., Xiang, A.P., Ren, B.Z., and Lahn, B.T. (2010). Systematic comparison of constitutive promoters and the doxycycline-inducible promoter. PLoS One *5*, 3–6. https://doi.org/10.1371/journal.pone.0010611.

Rafii, S., Kloss, C.C., Butler, J.M., Ginsberg, M., Gars, E., Lis, R., Zhan, Q., Josipovic, P., Ding, B. Sen, Xiang, J., et al. (2013). Human ESC-derived hemogenic endothelial cells undergo distinct waves of endothelial to hematopoietic transition. Blood *121*, 770–780. https://doi.org/10.1182/blood-2012-07-444208.

Rai, M.A., Hammonds, J., Pujato, M., Mayhew, C., and Roskin, K. (2020). Comparative analysis of human microglial models for studies of HIV replication and pathogenesis. Retrovirology 1–21. https://doi.org/10.1186/s12977-020-00544-y.

Ramos-Mejia, V., Melen, G.J., Sanchez, L., Gutierrez-Aranda, I., Ligero, G., Cortes, J.L., Real, P.J., Bueno, C., and Menendez, P. (2010). Nodal/Activin Signaling Predicts Human Pluripotent Stem Cell Lines Prone to Differentiate Toward the Hematopoietic Lineage. Am. Soc. Gene Cell Ther. *18*, 2173–2181. https://doi.org/10.1038/mt.2010.179.

Ramos-Mejía, V., Navarro-Montero, O., Ayllón, V., Bueno, C., Romero, T., Real, P.J.,

and Menendez, P. (2014). HOXA9 promotes hematopoietic commitment of human embryonic stem cells. Blood *124*, 3065–3075. https://doi.org/10.1182/blood-2014-03-558825.

Ran, F.A., Hsu, P.D., Wright, J., Agarwala, V., Scott, D.A., and Zhang, F. (2013a). Genome engineering using the CRISPR-Cas9 system. Nat. Protoc. *8*, 2281–2308. https://doi.org/10.1038/nprot.2013.143.

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., et al. (2013b). Double nicking by RNA-guided CRISPR cas9 for enhanced genome editing specificity. Cell *154*, 1380–1389. https://doi.org/10.1016/j.cell.2013.08.021.

Randolph, L.N., Bao, X., Zhou, C., and Lian, X. (2017). PiggyBac system for human pluripotent stem cells and derivatives. Sci. Rep. 1–8. https://doi.org/10.1038/s41598-017-01684-6.

Raza-Egilmez, S.Z., Jani-Sait, S.N., Grossi, M., Higgins, M.J., Shows, T.B., and Aplan, P.D. (1998). NUP98-HOXD13 gene fusion in therapy-related acute myelogenous leukemia. Cancer Res. *58*, 4269–4273. .

Real, P.J., Ligero, G., Ayllon, V., Ramos-Mejia, V., Bueno, C., Gutierrez-Aranda, I., Navarro-Montero, O., Lako, M., and Menendez, P. (2012). SCL/TAL1 regulates hematopoietic specification from human embryonic stem cells. Mol. Ther. *20*, 1443–1453. https://doi.org/10.1038/mt.2012.49.

Reinhardt, P., Schmid, B., Burbulla, L.F., Schöndorf, D.C., Wagner, L., Glatza, M., Höing, S., Hargus, G., Heck, S.A., Dhingra, A., et al. (2013). Genetic correction of a lrrk2 mutation in human iPSCs links parkinsonian neurodegeneration to ERK-dependent changes in gene expression. Cell Stem Cell *12*, 354–367. https://doi.org/10.1016/j.stem.2013.01.008.

Reizel, Y., Sabag, O., Skversky, Y., Spiro, A., Steinberg, B., Bernstein, D., Wang, A., Kieckhaefer, J., Li, C., Pikarsky, E., et al. (2018). Postnatal DNA demethylation and its role in tissue maturation. Nat. Commun. *9*, 1–11. https://doi.org/10.1038/s41467-018-04456-6.

Van Rensburg, R., Beyer, I., Yao, X.Y., Wang, H., Denisenko, O., Li, Z.Y., Russell, D.W.,

Miller, D.G., Gregory, P., Holmes, M., et al. (2013). Chromatin structure of two genomic sites for targeted transgene integration in induced pluripotent stem cells and hematopoietic stem cells. Gene Ther. *20*, 201–214. https://doi.org/10.1038/gt.2012.25.

Rice, K.L., and Licht, J.D. (2007). HOX deregulation in acute myeloid leukemia. J. Clin. Invest. *117*, 865–868. https://doi.org/10.1172/JCI31861.

Robin, C., Bollerot, K., Mendes, S., Haak, E., Crisan, M., Cerisoli, F., Lauw, I., Kaimakis, P., Jorna, R., Vermeulen, M., et al. (2009). Human Placenta Is a Potent Hematopoietic Niche Containing Hematopoietic Stem and Progenitor Cells throughout Development. Cell Stem Cell *5*, 385–395. https://doi.org/10.1016/j.stem.2009.08.020.

Roney, I.J., Rudner, A.D., Couture, J., and Kærn, M. (2016). Improvement of the reverse tetracycline transactivator by single amino acid substitutions that reduce leaky target gene expression to undetectable levels. Nat. Publ. Gr. 1–8. https://doi.org/10.1038/srep27697.

Ruiz, J.P., Chen, G., Haro Mora, J.J., Keyvanfar, K., Liu, C., Zou, J., Beers, J., Bloomer, H., Qanash, H., Uchida, N., et al. (2019). Robust generation of erythroid and multilineage hematopoietic progenitors from human iPSCs using a scalable monolayer culture system. Stem Cell Res. *41*, 101600. https://doi.org/10.1016/j.scr.2019.101600.

Rybtsov, S., Ivanovs, A., Zhao, S., and Medvinsky, A. (2016). Concealed expansion of immature precursors underpins acute burst of adult HSC activity in foetal liver. Development *143*, 1284–1289. https://doi.org/10.1242/dev.131193.

Sadelain, M., Papapetrou, E.P., and Bushman, F.D. (2012). Safe harbours for the integration of new DNA in the human genome. Nat. Rev. Cancer *12*, 51–58. https://doi.org/10.1038/nrc3179.

Samokhvalov, I.M., Samokhvalova, N.I., and Nishikawa, S.I. (2007). Cell tracing shows the contribution of the yolk sac to adult haematopoiesis. Nature *446*, 1056–1061. https://doi.org/10.1038/nature05725.

Sandler, V.M., Lis, R., Liu, Y., Kedem, A., James, D., Elemento, O., Butler, J.M., Scandura, J.M., and Rafii, S. (2014). Reprogramming human endothelial cells to haematopoietic cells requires vascular induction. Nature *511*, 312–318. https://doi.org/10.1038/nature13547.

Sanghez, V., Luzzi, A., Clarke, D., Kee, D., Beuder, S., Rux, D., Osawa, M., Madrenas, J., Chou, T.F., Kyba, M., et al. (2017). Notch activation is required for downregulation of HoxA3-dependent endothelial cell phenotype during blood formation. PLoS One *12*, 1–18. https://doi.org/10.1371/journal.pone.0186818.

Sapranauskas, R., Gasiunas, G., Fremaux, C., Barrangou, R., Horvath, P., and Siksnys, V. (2011). The Streptococcus thermophilus CRISPR / Cas system provides immunity in Escherichia coli. *39*, 9275–9282. https://doi.org/10.1093/nar/gkr606.

Sauvageau, G., Lansdorp, P.M., Eaves, C.J., Hogge, D.E., Dragowska, W.H., Reid, D.S., Largman, C., Lawrence, H.J., and Humphries, R.K. (1994). Differential expression of homeobox genes in functionally distinct CD34+ subpopulations of human bone marrow cells. Proc. Natl. Acad. Sci. U. S. A. *91*, 12223–12227. https://doi.org/10.1073/pnas.91.25.12223.

Sauvageau, G., Thorsteinsdottir, U., Hough, M.R., Hugo, P., Lawrence, H.J., Largman, C., and Humphries, R.K. (1997). Overexpression of HOXB3 in hematopoietic cells causes defective lymphoid development and progressive myeloproliferation. Immunity *6*, 13–22. https://doi.org/10.1016/S1074-7613(00)80238-1.

Saxonov, S., Berg, P., and Brutlag, D.L. (2006). A genome-wide analysis of CpG dinucleotides in the human genome distinguishes two distinct classes of promoters.

Schmittgen, T.D., and Livak, K.J. (2008). Analyzing real-time PCR data by the comparative CT method. Nat. Protoc. *3*, 1101–1108. https://doi.org/10.1038/nprot.2008.73.

Schulz, C., Perdiguero, E.G., Chorro, L., Szabo-Rogers, H., Cagnard, N., Kierdorf, K., Prinz, M., Wu, B., Jacobsen, S.E.W., Pollard, J.W., et al. (2012). A lineage of myeloid cells independent of myb and hematopoietic stem cells. Science (80-.). *335*, 86–90. https://doi.org/10.1126/science.1219179.

Shen, J., Zhu, Y., Zhang, S., Lyu, S., Lyu, C., Feng, Z., Hoyle, D.L., Wang, Z.Z., and Cheng, T. (2021). Vitronectin-activated αvβ3 and αvβ5 integrin signalling specifies haematopoietic fate in human pluripotent stem cells. Cell Prolif. *54*, 1–13. https://doi.org/10.1111/cpr.13012.

Shimamoto, T., Tang, Y., Naot, Y., Nardi, M., Brulet, P., Bieberich, C.J., and Takeshita,

K. (1999). Hematopoietic progenitor cell abnormalities in Hoxc-8 null mutant mice. J. Exp. Zool. *283*, 186–193. https://doi.org/10.1002/(SICI)1097-010X(19990201)283:2<186::AID-JEZ9>3.0.CO;2-0.

Shivdasanl, R.A., Mayer, E.L., and Orkin, S.H. (1995). Absence of blood formation in mice lacking the T-cell leukaemia oncoprotein tal-1/SCL. Nature *373*, 432–434. https://doi.org/10.1038/373432a0.

Singh, N.P., and Krumlauf, R. (2022). Diversification and Functional Evolution of HOX Proteins. Front. Cell Dev. Biol. *10*, 1–16. https://doi.org/10.3389/fcell.2022.798812.

Singh, A.M., Reynolds, D., Cliff, T., Ohtsuka, S., Mattheyses, A.L., Sun, Y., Menendez, L., Kulik, M., and Dalton, S. (2012). Signaling network crosstalk in human pluripotent cells: A Smad2/3-regulated switch that controls the balance between self-renewal and differentiation. Cell Stem Cell *10*, 312–326. https://doi.org/10.1016/j.stem.2012.01.014.

Slattery, M., Riley, T., Liu, P., Abe, N., Gomez-Alcala, P., Dror, I., Zhou, T., Rohs, R., Honig, B., Bussemaker, H.J., et al. (2011). Cofactor binding evokes latent differences in DNA binding specificity between hox proteins. Cell *147*, 1270–1282. https://doi.org/10.1016/j.cell.2011.10.053.

Spater, D., Hansson, E.M., Zangi, L., and Chien, K.R. (2014). How to make a cardiomyocyte. Development *2030*, 4418–4431. https://doi.org/10.1242/dev.091538.

Spencer, D.H., Young, M.A., Lamprecht, T.L., Helton, N.M., Fulton, R., Laughlin, M.O., Fronick, C., Magrini, V., Demeter, R.T., and Miller, C.A. (2015). Epigenomic analysis of the HOX gene loci reveals mechanisms that may control canonical expression patterns in AML and normal hematopoietic cells. Leukemia 1279–1289. https://doi.org/10.1038/leu.2015.6.

Spooner, C.J., Cheng, J.X., Pujadas, E., Laslo, P., and Singh, H. (2009). A Recurrent Network Involving the Transcription Factors PU.1 and Gfi1 Orchestrates Innate and Adaptive Immune Cell Fates. Immunity *31*, 576–586. https://doi.org/10.1016/j.immuni.2009.07.011.

Spyropoulos, D.D., Pharr, P.N., Lavenburg, K.I.M.R., Jackers, P., Papas, T.S., Ogawa, M., and Watson, D.K. (2000). Hemorrhage, Impaired Hematopoiesis, and Lethality in Mouse Embryos Carrying a Targeted Disruption of the Fli1 Trancription Factor. Mol. Cell.

Biol. *20*, 5643–5652. .

Sroczynska, P., Lancrin, C., Kouskoff, V., and Lacaud, G. (2009). The differential activities of Runx1 promoters define milestones during embryonic hematopoiesis. Blood *114*, 5279–5289. https://doi.org/10.1182/blood-2009-05-222307.

Stanulović, V.S., Cauchy, P., Assi, S.A., and Hoogenkamp, M. (2017). LMO2 is required for TAL1 DNA binding activity and initiation of definitive haematopoiesis at the haemangioblast stage. Nucleic Acids Res. *45*, 9874–9888. https://doi.org/10.1093/nar/gkx573.

Stefanovic, S., Laforest, B., Desvignes, J., Lescroart, F., Argiro, L., Maurel-zaffran, C., Salgado, D., Plaindoux, E., Bono, C. De, Pazur, K., et al. (2020). Hox -dependent coordination of mouse cardiac progenitor cell patterning and differentiation. 1–32. .

Straussman, R., Nejman, D., Roberts, D., Steinfeld, I., Blum, B., Benvenisty, N., Simon, I., Yakhini, Z., and Cedar, H. (2009). Developmental programming of CpG island methylation profiles in the human genome. Nat. Struct. Mol. Biol. *16*, 564–571. https://doi.org/10.1038/nsmb.1594.

Sturgeon, C.M., Ditadi, A., Awong, G., Kennedy, M., and Keller, G. (2014). Wnt signaling controls the specification of definitive and primitive hematopoiesis from human pluripotent stem cells. Nat. Biotechnol. *32*, 554–561. https://doi.org/10.1038/nbt.2915.

Sugimura, R., Jha, D.K., Han, A., Soria-Valles, C., Da Rocha, E.L., Lu, Y.F., Goettel, J.A., Serrao, E., Rowe, R.G., Malleshaiah, M., et al. (2017). Haematopoietic stem and progenitor cells from human pluripotent stem cells. Nature *545*, 432–438. https://doi.org/10.1038/nature22370.

Suzuki, A., Ito, Y., Sashida, G., Honda, S., Katagiri, T., Funjino, T., Nakamura, T., and Ohyashiki, K. (2002). t(7;11)(p15;p15) chronic myeloid leukaemia developed into blastic transformation showing a novel NUP98/HOXA11 fusion. Br. J. Haematol. *116*, 170–172. .

Swiers, G., Baumann, C., O'Rourke, J., Giannoulatou, E., Taylor, S., Joshi, A., Moignard, V., Pina, C., Bee, T., Kokkaliaris, K.D., et al. (2013). Early dynamic fate changes in haemogenic endothelium characterized at the single-cell level. Nat. Commun. *4*, 2924. https://doi.org/10.1038/ncomms3924.

Szczepek, M., Brondani, V., Büchel, J., Serrano, L., Segal, D.J., and Cathomen, T. (2007). Structure-based redesign of the dimerization interface reduces the toxicity of zinc-finger nucleases. Nat. Biotechnol. *25*, 786–793. https://doi.org/10.1038/nbt1317.

Takeshita, K., Bollekens, J.A., Hijiya, N., Ratajczak, M., Ruddle, F.H., and Gewirtz, A.M. (1993). A homeobox gene of the Antennapedia class is required for human adult erythropoiesis. Proc. Natl. Acad. Sci. U. S. A. *90*, 3535–3538. https://doi.org/10.1073/pnas.90.8.3535.

Tan, J.Y., Sriram, G., Rufaihah, A.J., Neoh, K.G., and Cao, T. (2013). Efficient Derivation of Lateral Plate and Paraxial Mesoderm Subtypes from Human Embryonic Stem Cells Through GSKi-Mediated Differentiation. Stem Cells Dev. *22*, 1893–1906. https://doi.org/10.1089/scd.2012.0590.

Tan, Y.-T., Ye, L., Xie, F., Beyer, A.I., Muench, M.O., Wang, J., Chen, Z., Liu, H., Chen, S.-J., and Kan, Y.W. (2018). Respecifying human iPSC-derived blood cells into highly engraftable hematopoietic stem and progenitor cells with a single factor. Proc. Natl. Acad. Sci. *115*, 2180–2185. https://doi.org/10.1073/pnas.1718446115.

Tang, J.C.Y., Szikra, T., Kozorovitskiy, Y., Teixiera, M., Sabatini, B.L., Roska, B., and Cepko, C.L. (2013). A nanobody-based system using fluorescent proteins as scaffolds for cell-specific gene manipulation. Cell *154*, 928–939. https://doi.org/10.1016/j.cell.2013.07.021.

Tavian, M., Hallais, M.F., and Péault, B. (1999). Emergence of intraembryonic hematopoietic precursors in the pre-liver human embryo. Development *126*, 793–803. https://doi.org/10.1242/dev.126.4.793.

Tavian, M., Robin, C., Coulombel, L., and Péault, B. (2001). The human embryo, but not its yolk sac, generates lympho-myeloid stem cells: Mapping multipotent hematopoietic cell fate in intraembryonic mesoderm. Immunity *15*, 487–495. https://doi.org/10.1016/S1074-7613(01)00193-5.

Teichweyde, N., Kasperidus, L., Carotta, S., Kouskoff, V., Lacaud, G., Horn, P.A., Heinrichs, S., and Klump, H. (2018). HOXB4 Promotes Hemogenic Endothelium Formation without Perturbing Endothelial Cell Development. Stem Cell Reports *10*, 875–889. https://doi.org/10.1016/j.stemcr.2018.01.009.

Thambyrajah, R., Mazan, M., Patel, R., Moignard, V., Stefanska, M., Marinopoulou, E., Li, Y., Lancrin, C., Clapes, T., Möroÿ, T., et al. (2016). GFI1 proteins orchestrate the emergence of haematopoietic stem cells through recruitment of LSD1. Nat. Cell Biol. *18*, 21–32. https://doi.org/10.1038/ncb3276.

Thorsteinsdottir, U., Mamo, A., Kroon, E., Jerome, L., Bijl, J., Lawrence, H.J., Humphries, K., and Sauvageau, G. (2002). Overexpression of the myeloid leukemia – associated Hoxa9 gene in bone marrow cells induces stem cell expansion. Blood *99*, 121–129. https://doi.org/10.1182/blood.V99.1.121.

Tober, J., Koniski, A., McGrath, K.E., Vemishetti, R., Emerson, R., De Mesy-Bentley, K.K.L., Waugh, R., and Palis, J. (2007). The megakaryocyte lineage originates from hemangioblast precursors and is an integral component both of primitive and of definitive hematopoiesis. Blood *109*, 1433–1441. https://doi.org/10.1182/blood-2006-06-031898.

Tober, J., Yzaguirre, A.D., Piwarzyk, E., and Speck, N.A. (2013). Distinct temporal requirements for Runx1 in hematopoietic progenitors and stem cells. Dev. *140*, 3765–3776. https://doi.org/10.1242/dev.094961.

Trivedi, C.M., Patel, R.C., and Patel, C. V. (2008). Differential regulation of HOXA9 expression by nuclear factor kappa B (NF-κB) and HOXA9. Gene *408*, 187–195. https://doi.org/10.1016/j.gene.2007.11.001.

Tsai, F.Y., Keller, G., Kuo, F.C., Weiss, M., Chen, J., Rosenblatt, M., Alt, F.W., and Orkin, S.H. (1994). An early haematopoietic defect in mice lacking the transcription factor GATA-2. Nature *371*, 221–226. https://doi.org/10.1038/371221a0.

Tsui, M., Min, W., Ng, S., Dobbs, K., Notarangelo, L.D., Dror, Y., and Grunebaum, E. (2021). The Use of Induced Pluripotent Stem Cells to Study the Effects of Adenosine Deaminase Deficiency on Human Neutrophil Development. Front. Immunol. *12*, 1–12. https://doi.org/10.3389/fimmu.2021.748519.

Uenishi, G., Theisen, D., Lee, J., Kumar, A., Raymond, M., Vodyanik, M., Swanson, S., Stewart, R., Thomson, J., and Slukvin, I. (2014). Tenascin C Promotes Hematoendothelial Development and T Lymphoid Commitment from Human Pluripotent Stem Cells in Chemically Defined Conditions. Stem Cell Reports *3*, 1073–1084. https://doi.org/10.1016/j.stemcr.2014.09.014.

Uenishi, G.I., Jung, H.S., Kumar, A., Park, M.A., Hadland, B.K., McLeod, E., Raymond, M., Moskvin, O., Zimmerman, C.E., Theisen, D.J., et al. (2018). NOTCH signaling specifies arterial-type definitive hemogenic endothelium from human pluripotent stem cells. Nat. Commun. *9*. https://doi.org/10.1038/s41467-018-04134-7.

Urlinger, S., Baron, U., Thellmann, M., Hasan, M.T., Bujard, H., and Hillen, W. (2000). Exploring the sequence space for tetracycline- dependent transcriptional activators : Novel mutations yield expanded range and sensitivity. *97*, 7963–7968. .

Urnov, F.D., Miller, J.C., Lee, Y.L., Beausejour, C.M., Rock, J.M., Augustus, S., Jamieson, A.C., Porteus, M.H., Gregory, P.D., and Holmes, M.C. (2005). Highly efficient endogenous human gene correction using designed zinc-finger nucleases. Nature *435*, 646–651. https://doi.org/10.1038/nature03556.

Vijayaragavan, K., Szabo, E., Bossé, M., Ramos-Mejia, V., Moon, R.T., and Bhatia, M. (2009). Noncanonical Wnt Signaling Orchestrates Early Developmental Events toward Hematopoietic Cell Fate from Human Embryonic Stem Cells. Cell Stem Cell *4*, 248–262. https://doi.org/10.1016/j.stem.2008.12.011.

Van Vliet, P., Wu, S.M., Zaffran, S., and Puceat, M. (2012). Early cardiac development : a view from stem cells to embryos. Cardiovasc. Res. *96*, 352–362. https://doi.org/10.1093/cvr/cvs270.

Vodyanik, M.A., Thomson, J.A., and Slukvin, I.I. (2006). Leukosialin (CD43) defines hematopoietic progenitors in human embryonic stem cell differentiation cultures. Blood *108*, 2095–2106. https://doi.org/10.1182/blood-2006-02-003327.Supported.

Wadman, I.A., Osada, H., Grütz, G.G., Agulnick, A.D., Westphal, H., Forster, A., and Rabbitts, T.H. (1997). The LIM-only protein Lmo2 is a bridging molecule assembling an erythroid, DNA-binding complex which includes the TAL1, E47, GATA-1 and Ldb1/NLI proteins. EMBO J. *16*, 3145–3157. https://doi.org/10.1093/emboj/16.11.3145.

Wang, C., Tang, X., Sun, X., Miao, Z., Lv, Y., and Yang, Y. (2012). TGF β inhibition enhances the generation of hematopoietic progenitors from human ES cell-derived hemogenic endothelial cells using a stepwise strategy. Cell Res. *22*, 194–207. https://doi.org/10.1038/cr.2011.138.

Wang, L., Menendez, P., Shojaei, F., Li, L., Mazurier, F., Dick, J.E., Cerdan, C., Levac,

K., and Bhatia, M. (2005). Generation of hematopoietic repopulating cells from human embryonic stem cells independent of ectopic *HOXB4* expression. J. Exp. Med. *201*, 1603–1614. https://doi.org/10.1084/jem.20041888.

Wareing, S., Mazan, A., Pearson, S., Göttgens, B., Lacaud, G., and Kouskoff, V. (2012). The Flk1-Cre-Mediated Deletion of ETV2 Defines Its Narrow Temporal Requirement during Embryonic Hematopoietic Development. Stem Cells *30*, 1521–1531. https://doi.org/10.1002/stem.1115.

Warren, A.J., Colledge, W.H., Carlton, M.B.L., Evans, M.J., Smith, A.J.H., and Rabbitts, T.H. (1994). The Oncogenic Cysteine-rich LIM domain protein Rbtn2 is essential for erythroid development. Cell *78*, 45–57. https://doi.org/10.1016/0092-8674(94)90571-1.

Weber, M., Davies, J.J., Wittig, D., Oakeley, E.J., Haase, M., Lam, W.L., and Schübeler, D. (2005). Chromosome-wide and promoter-specific analyses identify sites of differential DNA methylation in normal and transformed human cells. Nat. Genet. *37*, 853–862. https://doi.org/10.1038/ng1598.

Wellik, D.M. (2007). Hox patterning of the vertebrate axial skeleton. Dev. Dyn. *236*, 2454–2463. https://doi.org/10.1002/dvdy.21286.

Wen, Z., Li, S., Shui, X., Tang, L., Zheng, J., and Chen, L. (2019). LncRNA PEG10 aggravates cardiac hypertrophy through regulating HOXA9. *23*, 281–286. .

Wiedenheft, B., Sternberg, S.H., and Doudna, J.A. (2012). RNA-guided genetic silencing systems in bacteria and archaea. Nature *482*, 331–338. https://doi.org/10.1038/nature10886.

Wilson, N.K., Foster, S.D., Wang, X., Knezevic, K., Schütte, J., Kaimakis, P., Chilarska, P.M., Kinston, S., Ouwehand, W.H., Dzierzak, E., et al. (2010). Combinatorial transcriptional control in blood stem/progenitor cells: Genome-wide analysis of ten major transcriptional regulators. Cell Stem Cell *7*, 532–544. https://doi.org/10.1016/j.stem.2010.07.016.

Xia, X., Zhang, Y., Zieth, C.R., and Zhang, S.-C. (2007). Transgenes Delivered by Lentiviral Vector are Suppressed in Human Embryonic Stem Cells in a Promoter-Dependent Manner. Stem Cells Dev. *16*, 167–176. https://doi.org/10.1038/jid.2014.371.

Xia, X., Ayala, M., Thiede, B.R., Zhang, S., and Ph, D. (2008). In Vitro- and In Vivo-Induced Transgene Expression in Human Embryonic Stem Cells and Derivatives. Stem Cells *26*, 525–533. https://doi.org/10.1634/stemcells.2007-0710.

Yamamoto, H., Ehling, M., Kato, K., Kanai, K., Van Lessen, M., Frye, M., Zeuschner, D., Nakayama, M., Vestweber, D., and Adams, R.H. (2015). Integrin β1 controls VE-cadherin localization and blood vessel stability. Nat. Commun. *6*, 1–14. https://doi.org/10.1038/ncomms7429.

Yanagimachi, M.D., Niwa, A., Tanaka, T., Honda-ozaki, F., Nishimoto, S., Murata, Y., Yasumi, T., Ito, J., Tomida, S., Oshima, K., et al. (2013). Robust and Highly-Efficient Differentiation of Functional Monocytic Cells from Human Pluripotent Stem Cells under Serum- and Feeder Cell-Free Conditions. PLoS One *8*, 1–9. https://doi.org/10.1371/journal.pone.0059243.

Yang, C.-T., Ma, R., Axton, R.A., Jackson, M., Taylor, A.H., Fidanza, A., Marenah, L., Frayne, J., Mountford, J.C., and Forrester, L.M. (2017a). Activation of KLF1 Enhances the Differentiation and Maturation of Red Blood Cells from Human Pluripotent Stem Cells. Stem Cells *35*, 886–897. https://doi.org/10.1002/stem.2562.

Yang, D., Zhang, X., Dong, Y., Liu, X., Wang, T., Wang, X., Geng, Y., Fang, S., Zheng, Y., Chen, X., et al. (2015). Enforced expression of Hoxa5 in haematopoietic stem cells leads to aberrant erythropoiesis in vivo Enforced expression of Hoxa5 in haematopoietic stem cells leads to aberrant erythropoiesis in vivo. Cell Cycle *14*, 612–620. https://doi.org/10.4161/15384101.2014.992191.

Yang, F., Liu, C., Chen, D., Tu, M., Xie, H., Sun, H., Ge, X., Tang, L., Li, J., Zheng, J., et al. (2017b). CRISPR/Cas9-loxP-Mediated Gene Editing as a Novel Site-Specific Genetic Manipulation Tool. Mol. Ther. - Nucleic Acids *7*, 378–386. https://doi.org/10.1016/j.omtn.2017.04.018.

Yiangou, L., Grandy, R.A., Osnato, A., Ortmann, D., Sinha, S., and Vallier, L. (2019). Cell cycle regulators control mesoderm specification in human pluripotent stem cells. J. Biol. Chem. *294*, 17903–17914. https://doi.org/10.1074/jbc.RA119.008251.

Yokomizo, T., Watanabe, N., Umemoto, T., Matsuo, J., Harai, R., Kihara, Y., Nakamura, E., Tada, N., Sato, T., Takaku, T., et al. (2019). Hlf marks the developmental pathway for hematopoietic stem cells but not for erythro-myeloid progenitors. J. Exp. Med. *216*,

1599–1614. https://doi.org/10.1084/jem.20181399.

Yoon, B.S., Yoo, S.J., Lee, J.E., You, S., Lee, H.T., and Yoon, H.S. (2006). Enhanced differentiation of human embryonic stem cells into cardiomyocytes by combining hanging drop culture and 5-azacytidine treatment. Differentiation *74*, 149–159. https://doi.org/10.1111/j.1432-0436.2006.00063.x.

Yoshimoto, M., Montecino-Rodriguez, E., Ferkowicz, M.J., Porayette, P., Shelley, W.C., Conway, S.J., Dorshkind, K., and Yoder, M.C. (2011). Embryonic day 9 yolk sac and intra-embryonic hemogenic endothelium independently generate a B-1 and marginal zone progenitor lacking B-2 potential. Proc. Natl. Acad. Sci. U. S. A. *108*, 1468–1473. https://doi.org/10.1073/pnas.1015841108.

Yoshimoto, M., Porayette, P., Glosson, N.L., Conway, S.J., Carlesso, N., Cardoso, A.A., Kaplan, M.H., and Yoder, M.C. (2012). Autonomous murine T-cell progenitor production in the extra-embryonic yolk sac before HSC emergence. Blood *119*, 5706–5714. https://doi.org/10.1182/blood-2011-12-397489.

Zakany, J., and Duboule, D. (2007). The role of Hox genes during vertebrate limb development. Curr. Opin. Genet. Dev. *17*, 359–366. https://doi.org/10.1016/j.gde.2007.05.011.

Zákány, J., Kmita, M., and Duboule, D. (2004). A dual role for Hox genes in limb anterior-posterior asymmetry. Science (80-.). *304*, 1669–1672. https://doi.org/10.1126/science.1096049.

Zeng, H., Yücel, R., Kosan, C., Klein-Hitpass, L., and Möröy, T. (2004). Transcription factor Gfi1 regulates self-renewal and engraftment of hematopoietic stem cells. EMBO J. *23*, 4116–4125. https://doi.org/10.1038/sj.emboj.7600419.

Zeng, J., Yi, D., Sun, W., Liu, Y., Chang, J., Zhu, L., Zhang, Y., and Pan, X. (2021). Overexpression of HOXA9 upregulates NF- κ B signaling to promote human hematopoiesis and alter the hematopoietic differentiation potentials. Cell Regen. *10*, 1–12. .

Zhang, J., Klos, M., Wilson, G.F., Herman, A.M., Lian, X., Raval, K.K., Barron, M.R., Hou, L., Soerens, A.G., Yu, J., et al. (2013). Extracellular Matrix Promotes Highly Efficient Cardiac Differentiation of Human Pluripotent Stem Cells: The Matrix Sandwich

Method. Circ. Res. *111*, 1125–1136. https://doi.org/10.1161/CIRCRESAHA.112.273144.Extracellular.

Zhang, P., Li, J., Tan, Z., Wang, C., Liu, T., Chen, L., Yong, J., Jiang, W., and Sun, X. (2008). Short-term BMP-4 treatment initiates mesoderm induction in human embryonic stem cells. Blood *111*, 1933–1941. https://doi.org/10.1182/blood-2007-02-074120.

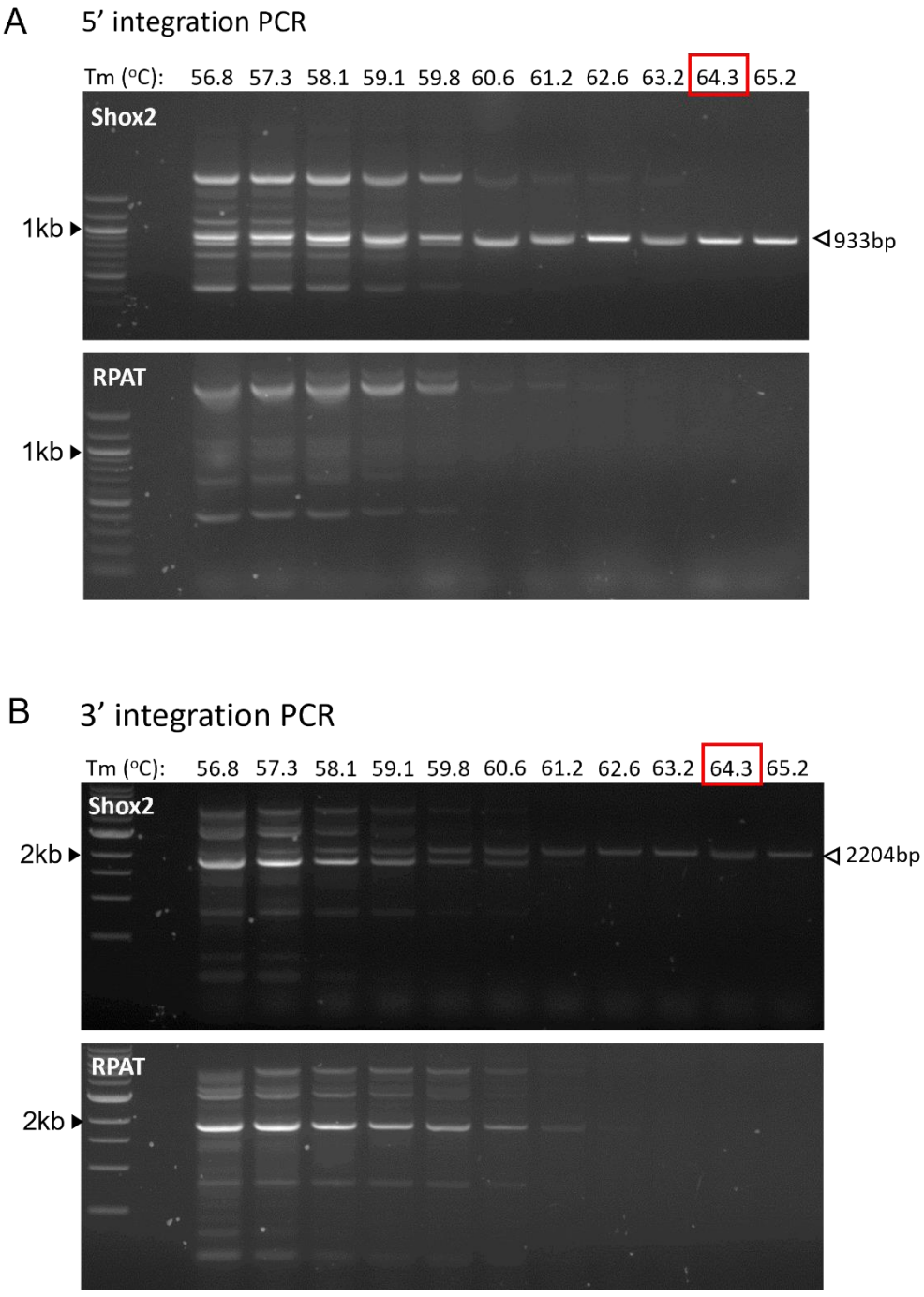
Zhang, Y., Hu, N., and Dong, F. (2019). Gfi1-Mediated Repression of c-Fos, Egr-1 and Egr-2, and Inhibition of ERK1/2 Signaling Contribute to the Role of Gfi1 in Granulopoiesis. Sci. Rep. *9*, 1–9. https://doi.org/10.1038/s41598-018-37402-z.

Zhou, G., Li, C., Feng, J., Zhang, J., and Fang, Y. (2018). lncRNA UCA1 Is a Novel Regulator in Cardiomyocyte Hypertrophy through Targeting the miR-184/HOXA9 Axis. CardioRenal Med. *230011*, 130–139. https://doi.org/10.1159/000487204.

Ziller, M.J., Gu, H., Muller, F., Donaghey, J., Tsai, L.T., Kohlbacher, O., De Jager, P.L., Rosen, E.D., Bennett, D.A., Bernstein, B.E., et al. (2013). Charting a dynamic DNA methylation landscape of the human genome. Nature *500*, 477–481. https://doi.org/10.1038/nature12433.

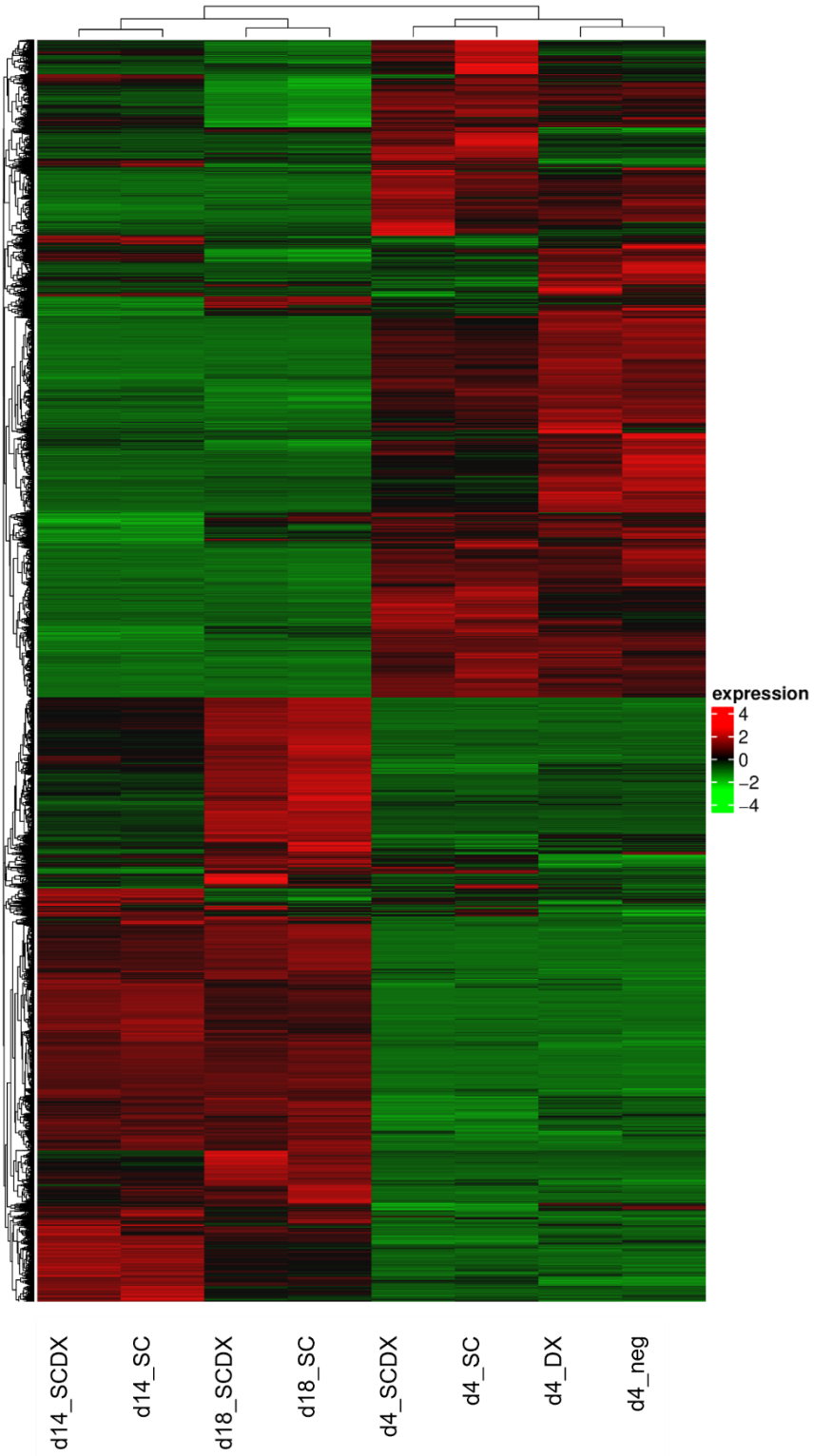
Zovein, A.C., Hofmann, J.J., Lynch, M., French, W.J., Turlo, K.A., Yang, Y., Becker, M.S., Zanetta, L., Dejana, E., Gasson, J.C., et al. (2008). Fate Tracing Reveals the Endothelial Origin of Hematopoietic Stem Cells. Cell Stem Cell *3*, 625–636. https://doi.org/10.1016/j.stem.2008.09.018.

Appendix I: Supplementary Figures



Supplementary Figure 1: Optimisation of genotyping PCR primers.

Gradient PCR of 5' (A) and 3' (B) integration primers to determine optimal primer annealing temperature using genomic DNA from inducible Shox2 and RPAT cell lines. Selected optimal annealing temperatures are highlighted by the red boxes and bands of the expected product size are indicated by the white arrow.



Supplementary Figure 2: Unsupervised hierarchal clustering of d4, d14 and d18 samples

Unsupervised clustering performed on all samples using Cluster Profiler. Gene expression level shown as normalised log2 FPKM. Results are from 2 biological replicates.

Appendix II: Supplementary Tables

Supplementary Table 1: Number of DEGs in d4 samples comparisons under indicated thresholds

Comparison	log_2FoldChange > 0		log_2FoldChange > 1.5	
	p-value < 0.05		padj < 0.05	
	DOWN	UP	DOWN	UP
d4Neg vs d4SC	3222	1959	755	674
d4Neg vs d4SCDX	3072	3099	666	591
d4DX vs d4SC	2567	2474	652	540
d4DX vs d4SCDX	2512	1585	520	433
d4SCDX vs d4SC	751	510	47	23
d4DX vs d4Neg	833	808	4	75

Supplementary Table 2: List of DEGs in d4 SC+DOX+ versus d4 SC+ cultures and d4 SC-DOX+ versus d4 SC- cultures

Samples	DE genes (log_2FoldChange > 1.5; padj < 0.05)	
d4SCDX vs d4SC	DOWN	PRODH2, COL3A1, TECRL, LUM, GUCY1A1, TDO2, OR7E35P, PGA4, DPYD, ALDH1A1, HABP2, PRR9, UGT2B7, SERPINF2, GSTA1, FRZB, CBS, AKR1D1, GSTA2, S100A14, MGAM, AP005202.1, LHX8, BCO1, LINC02413, TRPC5, PITX1, GJB1, SAMSN1, MYL4, FGG, TBX2-AS1, RSPO3, P2RY6, RASGRF2, DDC, LINC02381, KLB, PAMR1, LINC01695, SLC39A5, CST4, SOAT2, ITIH5, APOA1, SLC28A2, GREM2
	UP	**HOXA9**, GALNT16, GSX1, GNRH2, **HOXC6**, AC011330.1, CALB1, PAX9, SMOC1, NPPC, **HOXC4**, TRAC, ABCG1, ADAMTS15, TSHZ2, CLSTN2, ABCG2, NOV, VWA2, SOX21, PCDH19, COL14A1, SIX1
d4DX vs d4Neg	DOWN	AC131392.1, SPANXA2-OT1, GPR151, HECW1
	UP	**HOXA9, HOXC6, HOXD10**, ARX, NTRK2, **HOXD9**, C2CD4A, **HOXC4**, C2CD4B, SMOC1, CLSTN2, GNRH2, RGS4, WNT16, ZNF503-AS2, DANT1, TBX2, GALNT16, LYPD1, NSFP1, ANGPT2, CAPN6, **HOXD13**, LINC01467, BAALC, PTCH2, CNTNAP3P2, SGPP2, DAPL1, PAG1, UNC5C, RGS5, GLIS3, TMEM100, ADAMTS5, LIPG, ZNF503, FAM84A, CNTNAP3, ITGA4, PLA2G7, STAMBPL1, LRP2, ATP6V0D2, COL11A2, ALPK2, SLC1A1, MTX1P1, SLIT2, APOLD1, DYRK3, TGFBI, STAG3L1, FILIP1L, CNTNAP3B, OAF, KCNV1, HLA-DPA1, CALB1, NCAM1, ARID5B, HOXB13, SCML1, ST8SIA1, HNMT, BMP4, MAGED4, ZNF703, RNASE1, ATP8B1, SULT1C2, DAAM2, PDZRN3, MYL3, SPOCK1

Supplementary Table 3: Number of DEGs in d14 and d18 sample comparisons under indicated thresholds

Comparison	log_2FoldChange > 0		log_2FoldChange > 1.5	
	p-value < 0.05		padj < 0.05	
	DOWN	UP	DOWN	UP
d14SC vs d18SC	3297	2702	1050	146
d14SCDX vs d18SCDX	3509	3096	1041	103
d14SC vs d14SCDX	176	171	2	0
d18SC vs d18SCDX	414	230	52	3

Appendix III: Publications

Abuhantash, M., **Collins, E**. and Thompson, A., 2021. Role of the HOXA cluster in HSC emergence and blood cancer. Biochemical Society Transactions, 49(4), pp.1817-1827.

Bhagwan, J., **Collins, E**., Mosqueira, D., Bakar, M., Johnson, B., Thompson, A., Smith, J. and Denning, C. (2019). Variable expression and silencing of CRISPR-Cas9 targeted transgenes identifies the AAVS1 locus as not an entirely safe harbour. F1000Research, 8, p.1911.

Collins, E. and Thompson, A. (2018). HOX genes in normal, engineered and malignant hematopoiesis. The International Journal of Developmental Biology, 62(11-12), pp.847-856.