PHENOTYPIC AND GENOTYPIC CHARACTERISATION OF JUVENILE MYELOMONOCYTIC LEUKAEMIA



Eleni Louka

A thesis submitted for the degree of Doctor of Philosophy

Hertford College

Trinity Term 2018

ABSTRACT

Paediatric myelodysplastic syndromes and Juvenile Myelomonocytic Leukaemia (JMML) are very rare haematological diseases of childhood, with an estimated incidence of 1 case per million. Due to their rarity there is relatively little known about their natural history and disease biology. As a result, there are often significant challenges in establishing a diagnosis and there are very few treatment options available for these children. The only curative treatment is bone marrow transplantation but even this is associated with a high risk of disease relapse, particularly in JMML where relapse rates are approaching 35%.

In the work described here, I first set up a national observational study to prospectively collect serial patient samples in order to phenotypically and molecularly characterise these patients at diagnosis and to establish the basis for a prospective natural history study. The clinical and molecular characterisation of these diseases allowed me to devise new diagnostic and management algorithms for paediatric MDS and JMML.

In order to understand more about the biology of this group of disorders I decided to focus on JMML as this was the largest distinct group of patients enrolled in the study. My specific aim has been to identify and characterise the disease-propagating population in JMML (the JMML 'leukaemic stem cells, LSC); which is a crucial step in understanding the disease evolution and dynamics of relapse. Prior to this work, there was almost nothing known about the haematopoietic hierarchy in JMML and the impact of JMML-associated mutations on haematopoiesis. With this study I analysed the JMML haematopoietic stem cell and progenitor cell (HSPC) compartments gaining new insights into the architecture of the JMML HSPC and their lineage potential. Using a single cell genotyping approach I was able to map the

JMML mutations onto the haematopoietic hierarchy identifying for the first time the cellular compartment in which the JMML-associated mutations originate. The stem cell capacity of candidate JMML LSC was assessed with in vitro and in vivo stem cell assays establishing the identification of the disease propagating population(s). Molecular techniques, including RNA sequencing, then allowed me to molecularly characterise the JMML LSC and identify putative disease-specific targets and biomarkers that might be utilized in the future for disease burden monitoring and more effective treatments.

ACKNOWLEDGEMENTS

I would like to thank my supervisors Prof Irene Roberts and Prof Adam Mead, based at the WIMM, as well as my supervisor for the clinical part of the project Dr Anupama Rao based at GOSH; firstly for the opportunity to work on such an exciting project as well as for their excellent supervision and support throughout my DPhil. It was a great pleasure and privilege to work with such exceptional scientists and people that have been always very approachable and happy to offer their valuable advice during my time in the laboratory. My DPhil was a fantastic learning experience and my supervisors have become a great inspiration for me. I particularly appreciate my supervisors' dedication and accessibility often outside working hours especially in the year after completing my fellowship, during which I had to return back to clinical practice.

I would like to thank all members of the Roberts and Mead labs for all their help during my project as well as for such an enjoyable three years at the WIMM. In particular, I would like to thank Dr Benjamin Povinelli (Post Doctoral Scientist, based at the Mead lab) for the all the teaching and guidance during my DPhil; his advice as well as hands on help during my long experiments have been incredibly valuable for me in order to achieve submitting my DPhil data for publication. I would also like to especially thank Alba Meira (DPhil student in the Mead lab) and Gemma Buck for all their help on single cell techniques, bioinformatics analysis and friendship during my DPhil.

Next I would like to express my gratitude to all members of the Oxford Molecular Diagnostics lab and especially Prof Anna Schuh for facilitating the set up of the targeted next generation sequencing panel within the Oxford Molecular diagnostic

lab and Dr Angela Hamblin for her immensely valuable help in setting up and running the targeted next generation sequencing (NGS) panel. It is has been a great pleasure to work with such a dedicated clinician and without her help the panel would not have been as successful in delivering high quality data.

The Camelia Botnar lab members at Great Ormond Street Hospital, especially Sarah Inglott and Jane Chalker have contributed greatly in ensuring that samples from GOSH would be send to the WIMM in a timely manner as well as in validating the mutations detected by the NGS panel.

Lastly, I would like to thank all UK tertiary paediatric haematology centres and research teams for contributing samples and clinical information, in order to be able to perform this research project, as well as the UK Childhood Leukaemia bank for contributing archived samples to the project. All WIMM facilities but in particular the WIMM flow cytometry facility for their valuable help with cell sorting . In addition I would like to thank Simon McGowan for his valuable help with developing the NIHR study clinical database.

DECLARATION

All experiments described in this thesis were designed, performed and analysed by myself. Contributions of colleagues for certain experiments are acknowledged in the relevant sections.

CONTENTS

ABSTRACT **2**

ACKNOWLEDGEMENTS **4**

DECLARATION **5**

CONTENTS **6**

LIST OF FIGURES **12**

MAIN ABBREVIATIONS **15**

1. INTRODUCTION **26**

1.1 Paediatric Myelodysplastic Syndromes (MDS) **26**

1.1.1 Definition and incidence of Paediatric MDS 26

1.1.2 Clinical presentation 26

1.1.3 Classification of Paediatric MDS 27

1.1.4 Overview of other secondary Paediatric MDS/ overlap syndromes 31

1.1.5 Molecular and cytogenetic characteristics of Paediatric MDS 39

1.1.6 Differences between adult and paediatric MDS 41

1.1.7 Pathophysiology of Paediatric MDS 43

1.1.8 Current diagnostic pathway for paediatric MDS 44

1.1.9 Treatment of Paediatric MDS 48

1.1.10 Natural history of Paediatric MDS 50

1.2: Juvenile Myelomocytic Leukaemia **51**

1.2.1: Epidemiology, clinical presentation and disease features 51

1.2.2 Hallmarks of JMML 53

1.2.3: JMML diagnostic algorithm 56

1.2.4: The Molecular Landscape of JMML 59
1.2.5 Gene expression studies in JMML 68
1.2.6 Syndromic JMML 70
1.2.7: Management of JMML 75
1.2.8: Natural history and prognosis 78
1.3 AIMS 81
2. METHODS 84
2.1 THE NIHR Paediatric MDS/JMML study 84
2.1.1 Ethics application, procedure and approval 84
2.1.2 Study design/ procedure 85
2.1.3 Deep phenotyping of patients enrolled into the NIHR Paediatric MDS/JMML study 88
2.1.4 Sample processing and cryopreservation 92
2.1.5 DNA extraction 92
2.2 Paediatric MDS/JMML targeted NGS panel 93
2.2.1 Target selection, amplicon design and coverage prediction 93
2.2.2 Library preparation and sequencing 95
2.2.3 Samples alignment and analysis 96
2.2.4 Mutation validation/ Reports generation 98
2.3 Investigation of disease-propagating populations in JMML 99
2.3.1 Phenotypic and functional characterization of JMML HSPCS 99
2.3.2 'Driver' mutation tracing to HSPC compartments in JMML 109
2.3.3 Long-term culture-initiating cell assay protocol 115
2.3.4 In vivo stem cell assays 118
2.3.5 RNA sequencing 121

2.3.6 Quantification and statistical analysis 125

3. THE NIHR PAEDIATRIC MDS/JMML STUDY 127

3.1 Background and purpose of the study 127

3.2 Aims 128

3.3 Demographic and clinical characteristics of children recruited into the NIHR Paediatric MDS/JMML study 130

3.3.1 Demographic data 130

3.3.2 Clinical characteristics 132

3.3.3 Haematological and cytogenetic characteristics 134

3.3.4 Sub-classification of the 17 patients with MDS 137

3.4 Molecular characterisation of the disease-associated mutations using targeted NGS or WGS 138

3.4.1 Identification of pathogenic variants using targeted NGS 138

3.4.2 Identification of pathogenic variants using WGS 141

3.4.3 Phenotype-genotype correlations in JMML 142

3.4.4 Development of algorithms for diagnosis and management 148

3.5 Conclusions and Discussion 152

Chapter 4: PHENOTYPIC AND FUNCTIONAL CHARACTERISATION OF JMML HSPC 154

4: PHENOTYPIC AND FUNCTIONAL CHARACTERISATION OF JMML HSPC . 155

4.1 Background 155

4.1.1 The HSPC compartment in normal haematopoiesis 155

4.1.2 The HSPC compartment in leukaemia 158

4.2 Aim 160

4.3 Experimental approach 161

4.4 Immunophenotypic analysis of JMML HSPC at diagnosis 162

4.4.1 The Lin-CD34+CD38- compartment in JMML 164

4.4.2 The Lin-CD34+CD38+ compartment in JMML 170

4.5 Flow cytometry analysis of JMML HSPC: comparison of BM with PB ... 171

4.6 Flow cytometry analysis of JMML serial samples: effect of disease stage and treatment on HSPC frequency 173

4.7 Lineage output of JMML HSPC 179

4.7.1 JMML HSPC clonogenic capacity in short-term methylcellulose assays 179

4.7.2 Assessment of the megakaryocytic output of JMML HSPC in liquid culture 183

4.7.3 Lymphoid output of JMML HSPC 185

4.8 In vitro stem cell assays of JMML HSPC (long-term culture initiating-cell assay– LTC-IC) 186

4.9 In vivo long term stem cell potential of JMML HSPCs 188

4.9.1 Peripheral blood engraftment of JMML HSPC in NSG mice 189

4.9.2 BM engraftment of JMML HSPC in NSG mice 192

4.9.3 Secondary transplantation of NSG-engrafted JMML HSPC 199

4.10 Discussion and Conclusions 200

5: MOLECULAR CHARACTERISATION OF THE DISEASE PROPAGATING POPULATION IN JMML 204

5.1 Introduction 204

5.1.1: Study of the molecular hierarchy of myeloid neoplasms 204

5.1.2: Study of the molecular hierarchy in JMML 206

5.2 Aims 208

5.3 Experimental approach 208

5.3.1 Mapping of JMML mutations on the haematopoietic hierarchy 208

5.3.2 Molecular characterization of disease-propagating populations using RNA sequencing 209

5.4: Tracing of the JMML mutations on the haematopoietic hierarchy 209

5.4.1 Development and optimisation of the colony genotyping protocol 209

5.4.2 Mutational analysis of bulk patient samples 214

5.4.3 Lineage tracing of patient-specific mutations 215

5.5: Molecular characterisation of disease-propagating populations in JMML using RNA sequencing 233

5.5.1 Bulk RNA sequencing 233

5.5.2 Single cell RNA sequencing of JMML HSPC 246

5.6 Discussion and conclusions 249

6: DISCUSSION 251

6.1 Overview 251

6.2 The NIHR Paediatric MDS/JMML study 252

6.3: Phenotypic and functional characterisation of the JMML HSPC 253

6.4 Molecular characterisation of JMML HSPC 256

6.5 Implications of this work 258

6.6 Limitations of this work 259

6.7 Future work 260

6.8 Conclusions 262

REFERENCES 263

Appendix 1: Example of CoveMi software report 288

Appendix 2: Detailed protocol of next generation sequencing (NGS) targeted panel library preparation 289

Appendix 3: Example of the NIHR study targeted NGS research report 292

Appendix 4: List of monoclonal antibodies used for this project 293

Appendix 5: Primer sequences used for this project 295

LIST OF FIGURES

1.1.1: Paediatric MDS subtypes and overlap syndromes
1.1.2: Distribution of cytogenetic abnormalities detected in cases of refractory cytopenia (RCC)
1.1.3: Routine diagnostic work up for Paediatric MDS cases
1.1.4: Heatmap of clinical characteristics seen in Paediatric MDS overlap cases
1.1.5: Treatment algorithm for Paediatric MDS
1.2.1: RAS pathway targets implicated in the JMML GM-CSG hypersensitivity
1.2.2: Bone marrow aspirate of a child with JMML
1.2.3: Routine diagnostic work up for JMML cases
1.2.4: RAS pathway molecular targets implicated in the pathogenesis of JMML
1.2.5: Kaplan Meier survival curves for JMML patients with single VS secondary mutations
1.2.6: Kaplan Meier survival curves for JMML patients based on their underlying mutation

2.1: The NIHR Paediatric MDS study centres
2.2: Study procedure for sample collection and processing
2.3: The paediatric MDS online database
2.4: Genomic targets for the targeted next generation sequencing panel
2.5: The illumine design studio platform
2.6: Bioinformatics pipeline used for the mutation annotation
2.7: The targeted panel reporting pipeline
2.8: Example of Sanger sequencing validation of mutations detected by the NGS panel
2.9: Example of antibody titration
2.10: Flow cytometry basic gating strategy
2.11: HSPC gating strategy with flow cytometry
2.12: Image animation of the single cell genotyping method
2.13: Colony genotyping analysis: bioinformatics pipeline
2.14: Image animation of the long term initiating cell assays

3.1: The NIHR Paediatric MDS cohort demographic characteristics
3.2: The NIHR Paediatric MDS cohort clinical characteristics
3.3: The impact of cytopenias for the NIHR Paediatric MDS cohort
3.4: The NIHR Paediatric MDS cohort full blood count characteristics
3.5: Cytogenetics and bone marrow characteristics of the NIHR Paediatric MDS cohort
3.6: The Paediatric MDS subtypes identified in the study
3.7: Molecular characterization of the NIHR Paediatric MDS cohort with targeted next generation sequencing
3.8: The rates of mutation detection by the NGS panel and clinical significance
3.9: The clinical characteristics of the JMML cohort
3.10: Phenotypic: genotypic correlations of the JMML cohort
3.11: Serial mutation monitoring of the JMML cohort
3.12: Kaplan Meier survival curve of the JMML cohort

3.13: Revised diagnostic pathway for children with myelodysplasia
3.14: Revised management and surveillance pathway for children with myelodysplasia
4.1 The haematopoietic hierarchy
4.2: Experimental approach for the phenotypic characterization of JMML HSPC
4.3: HSPC analysis of normal controls with flow cytometry
4.4: HSPC analysis of JMMLBM sample compared with controls
4.5: Flow cytometry analysis of the JMML lin-CD34+CD38- compartment compared with controls
4.6: Flow cytometry analysis of the JMML lin-CD34+CD38- compartment compared with controls across different developmental stages
4.7: Flow cytometry analysis of the JMML lin-CD34+CD38- compartment according to the mutation type
4.8: Heatmap of the JMML mutations detected by the targeted NGS panel
4.9: Correlation of HbF expression with the lin-CD34+CD38- CD90+CD45RA+ compartment
4.10: Flow cytometry analysis of the JMML lin-CD34+CD38+ compartment compared with controls
4.11: Flow cytometry analysis of the JMML HSPC compartment in paired BM and PB samples
4.12: Serial samples monitoring of the HSPC compartment in relapse patients
4.13: Frequencies of the HSPC compartment during treatment
4.14: Serial samples monitoring of the HSPC compartment in relapse patients
4.15: The effect of Azacitidine on the JMML HSPC compartment
4.16: Single cell HSPC clonogenicity of JMML compared to normal controls
4.17: Morphological abnormalities of the erythroid colonies detected in JMML
4.18: Morphological abnormalities of the myeloid colonies detected in JMML
4.19: Embryonic globin expression of the JMML erythroid colonies
4.20: Megakaryocytic output of JMML HSPC
4.21: Flow cytometry examples of HSPC colonies from single cell liquid culture assays
4.22: Lymphoid output of the JMML HSPC
4.23: LTC-IC activity of the JMML HSPC
4.24: Profiles of patients used for in xenotraplantation
4.25: Peripheral blood engraftment of the JMML HSPC compared to controls
4.26: Peripheral blood lineage output of the JMML HSPC compared to controls
4.27: Bone Marrow engraftment of the JMML HSPC compared to controls
4.28: Bone Marrow flow cytometry analysis of the JMML HSPC compared to controls
4.29: Survival, bone marrow cellularity and spleen size of the engrafted mice
4.30: Peripheral blood full blood count analysis of the mice recipients
4.31: Analysis of the lin-CD34+CD90+ in the bone marrow of the recipient mice
4.32: Frequencies of the lin-CD34+CD90+ in the bone marrow of the recipient mice
4.33: Secondary transplantations

5.1: Optimisation of the single cell genotyping technique
5.2: Testing of the single cell genotyping technique in different cell numbers and colony types
5.3: Agarose gels of the targets detection verification
5.4: Validation of the mutation detection by single cell genotyping with Sanger sequencing

5.5: Heatmap of the JMML mutations detected by the targeted NGS panel
5.6: Single genotyping of patient ID1
5.7: Quality control of genotyping for patient ID1
5.8: Correlation of the mutational statues with the colony type for patient ID1
5.9: Single genotyping of patient ID15
5.10: Correlation of the mutational statues with the colony type for patient ID15
5.11: Quality control of genotyping for patient ID15
5.12: Single genotyping of patient ID5 at diagnosis
5.13: Quality control of genotyping for patient ID5
5.14: Single genotyping of patient ID5 post transplantation
5.15: Quality control of genotyping for patient ID5 post transplantation
5.16: Single genotyping of patient ID3
5.17: Detection of heterozygous SNPs with single cell genotyping
5.18: Quality control of genotyping for patient ID3
5.19: Differentially expressed genes between JMML LSCs and normal controls
5.20: PCA clustering of the JMML LSCs with normal controls
5.21: Heatmap of differentially expressed genes between JMML and normal HSCs
5.22: Heatmap of differentially expressed genes between JMML HSCs, GMPs, lin-CD34+CD38-CD90+CD45RA+
5.23: Gene expression of cell cycle regulators
5.24: Gene expression of stem cell regulators
5.25: Correlation of gene expression and stem cell capacity of the JMML LSCs
5.26: Impaired erythroid gene expression in JMML LSCs
5.27: Impaired lymphoid gene expression in JMML LSCs
5.28: Gene expression of therapeutic targets in JMML LSCs
5.29: In vitro single cell drug assays
5.30: CD96 expression in JMML by RNA and flow cytometry
5.31: tSNE of single cell lin-CD34+ HSPC in JMML VS normal controls
5.32: Heatmap of differentially expressed genes in single cell lin-CD34+ HSPC that allow cluster formation

6.1: Image animation of the impaired pathways detected in JMML based on the work described in this thesis

MAIN ABBREVIATIONS

ABCB1	ATP binding cassette subfamily B member 1
AEBP2	AE-binding protein 2
AGM	Aorta/gonad/mesonephros region
AKAP12	A-kinase anchor protein 12
A2ML1	Alpha-a macroglobulin like -1
alpha MEM	Minimum Essential Medium (MEM) Alpha Medium
AML	Acute Myeloid Leukaemia
ANKRD26	Ankyrin repeat domain 26
ALL	Acute Lymphoblastic Leukaemia
ARHGAP26	Rho GTPase Activating protein 26
ASXL1	Additional sex combs-like 1 transcription regulator
ATG	Antithymocyte globulin
ATXPC	Ataxia-Pancytopenia syndrome
AZA	Azacitidine
BCL11A	B-cell CLL/Lymphoma 11A
BFU-E	Blast forming unit – erythroid
BIK	BCL2 interacting killer
BM	Bone marrow
BMP4	Bone Morphogenic protein 4
BMT	Bone Marrow tranplantation
bp	Base pairs
BRAF	v-RAF murine sarcoma viral oncoprotein homolog B1
CALCA	Calcitonin/Calcitonin-related polypeptide alpha
CAMT	Congenital Amegakaryocytic thrombocytopenia

CB	Cord blood
CBF	Core binding factor
CBL	Casitas B-Lineage Lymphoma proto-oncogene
CCL3	Chemokine CC motif ligand 3
cDNA	Complementary DNA
CDKN2B	Cyclin dependent kinase inhibitor 2B
CEBPa	CCAAT/Enhancer-binding protein alpha
CFU-G	Colony forming unit – granulocyte
CFU-GEMM	Colony forming unit – granulocyte/erythroid/megakaryocyte/monocyte
CFU-GM	Colony forming unit – granulocyte/monocyte
CFU-M	Colony forming unit – monocyte
CFU-MK	Colony forming unit – megakaryocyte
CFU-MkE	Colony forming unit – megakaryocyte/erythroid
CI	Chief investigator
CML	Chronic myeloid leukaemia
CMML	Chronic myelomonocytic leukaemia
CMP	Common myeloid progenitor
CMV	Cytomegalovirus
CNN3	Calponin 3
CR	Complete remission
CRF	Clinical record form
CRLF2	Cytokine receptor like factor 2
CSC	Cancer stem cell
CSF3R	Colony stimulating factor 3 receptor
dATP	Deoxyadenosine triphosphate

DBA	Diamond Blackfan Anaemia
dCTP	Deoxycytidine triphosphate
DDX41	DEAD box Helicase 41
dGTP	Deoxyguanosine triphosphate
DKC	Dyskeratosis congenita
DMEM	Dulbecco Modified eagle medium
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic Acid
DNMT3A	DNA methyltransferase 3 Alpha
dNTPs	Deoxyribonucleotide triphosphates i.e. dATP, dCTP, dGTP, dTTP
DP	Double positive (Lin-CD34+CD38-CD90+CD45RA+)
DOT1L	DOT1 like histone lysine methyltransferase
dTTP	Deoxythymidine triphosphate (thymidine triphosphate)
EBV	Epstein-Barr virus
EDTA	Ethylene Diamine Triacetic Acid
EZH2	Enhancer of zeste, homololog 2
ELANE	Elastase Neutrophil expressed
EMP	Erythromyeloid progenitor
EPO	Erythropoietin
ETV6	Ets variant 6
EWOG-MDS	European Working Group of Myelodysplastic Syndromes(MDS) in childhood
FACS	Fluorescence activated cell sorting
FA	Fanconi Anaemia
FBC	Full Blood Count
FBS	

	Fetal bovine serum
FGD5	FYVE RhoGEF and PH domain containing 5
FISH	fluorescence in situ hybridization
FL	Fetal liver
FLI1	Friend leukaemia virus integration 1
FLT3	FMS-like tyrosine kinase 3
FTT	Failure to thrive
G6PC3	Glucose-6-Phosphatase Catalytic subunit 3
GATA1	GATA binding protein 1
GATA2	GATA binding protein 2
GAPS	GTP-ase activating proteins
G-CSF	Granulocyte colony-stimulating factor
gDNA	Genomic DNA
GEF	Guanosine exchange factors
GFI1	Growth factor independent 1 transcriptional repressor
GM-CSF	Granulocyte macrophage colony-stimulating factor
GMP	Granulocyte monocyte progenitor
GRB2	Growth factor receptor-bound protein 2
GVHD	Graft versus host disease
GYPE	Glycophorin E
HAX1	HCLS1 associated protein X-1
Hb	Haemoglobin
HDAC9	Histone Deacetylase 9
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HIV	Human immunodeficiency virus
HLF	Hepatic leukaemia factor

HMGA2	High mobility group AT-hook 2
HOPX	HOP homebox
HPLC	High pressure liquid chromatography
HPO	Human phenotype ontology
HRAS	v-ha –RAS Harvey rat sarcoma viral oncogene homolog
HSC	Haematopoietic stem cell
HSPC	Haematopoietic stem and progenitor cells
HSV	Herpes Simplex virus
IBMFS	Inherited bone marrow failure syndromes
ICUS	Idiopathic cytopenias of uncertain significance
IGV	Interactive genomics viewer
IKZF3	IKAROS family zinc finger 3
IL	Interleukin
IRAS	Integrated research application system
IMDM	Iscove's modified Dulbecco's media
JAK2	Janus kinase 2
JAK3	Janus kinase 3
JMML	Juvenile Myelomonocytic leukaemia
KLF1	Kruppel-like factor 1
KRAS	Kirsten rat sarcoma virus
LDB1	LIM domain binding 1
Lin	Lineage
LIN28B	Lin-28 Homolog B
LMO2	LIM domain only 2
LMPP	Lymphoid primed multipotent progenitor
LOH	Loss of heterozygosity

LSC	Leukaemic stem cell
LTC-IC	Long term culture initiating cell
LUC7L2	LUC like 2, pre-mRNA splicing factor
LYL1	Lymphoblastic leukaemia associated haematopoiesis regulator 1
LZTR1	Leukine zipper like transcriptional regulator 1
mAb	Monoclonal antibody
MAP2K1	Mitogen activated protein kinase kinase 1
MCM2	Minichromosome maintenance complex component 2
MCV	Mean Cell Volume
MDS	Myelodysplastic syndromes
MDS-EB	Myelodysplastic syndromes with excess blasts
MDS-EBt	Myelodysplastic syndromes with excess blasts in transformation
MDS-RS	Myelodysplastic syndromes with ring sideroblasts
MDS-Tx	Treatment related myelodysplasia
MDR-AML	MDS related AML
MECOM	MDS1 and EVI1 complex locus
MEIS1	Myeloid ecotropic viral integration site 1 homolog
MEP	Megakaryocyte erythroid progenitor
MKI67	Marker of proliferation KI-67
MLL	Mixed lineage leukaemia
MLP	Myelo-lymphoid progenitors
MNC	Mononuclear cells
MPL	Myeloproliferative Leukaemia protein
MPN	Myeloproliferative neoplasms
MPP	Multipotent progenitor

MRC	Medical research counsil
mRNA	Messenger RNA
MTOR	Mammalian target of rapamycin
NF1	Neurofibromin 1
NGS	Next generation sequencing
NHS	National health service
NIHR	National Institute of Health research
NK	Natural killer cells
NOTCH2	Neurogenic locus notch homolog protein 2
NPM1	Nucleaophosmin/Nucleoplasmin family, member 1
NRAS	Neuroblastoma RAS viral oncogene homolog
NSG	NOD *scid* gamma, NOD-*scid* IL2Rg^{null}, NOD-*scid* IL2Rgamma^{nul}
ORC6	Origin recognition complex, subunit 6
PAX5	Paired box gene 5
PB	Peripheral blood
PBS	Phosphate-buffered saline
PBX1	Pre-B cell leukaemia transcription factor
PCA	Principal component analysis
PCR	Polymerase Chain Reaction
qPCR	Quantitative Polymerase Chain Reaction
PI	Principal investigator
PLK1	Polo-like kinase 1
Plt	Platelets
PIM1	Proto-oncogene Serine/threonine protein kinase PIM1
PR	Partial remission
PRDX2	Peroxiredoxin 2

PTEN	Phosphatase and tensin homolog
PTPN11	Protein Tyrosine Phosphatase, Non Receptor Type 11
RAC2	Ras-Related C3 botulinum toxin substrate 2
RAF1	v-RAF1 murine leukaemia viral oncogene homolog 1
RARB	Retinoic acid receptor beta
RCC	Refractory cytopenias of childhood
R&D	Research and development
REC	Research ethic committees
RNA	Ribonucleic Acid
RPMI	Cell culture media developed at Roswell Park Memorial Institute
RQ-PCR	Real time quantitative Polymerase Chain Reaction
RRAS	Related RAS viral oncogene homolog
RT-PCR	Reverse Transcriptase Polymerase Chain Reaction
RUNX1	Runt related transcription factor 1
SAA	Severe Aplastic anaemia
SAM files	Sequence alignment map files
SAMD9	Sterile Alpha Motif Domain Containing 9
SAMD9L	Sterile Alpha Motif Domain Containing 9 like
SCF	Stem cell factor
SCN	Severe Congenital neutropenia
SDS	Shwachman–Diamond Syndrome
SEM	Standard error of the mean
SETBP1	SET-Binding protein 1
SF3B1	Splicing factor 3B subunit 1
SHOC2	Suppresor of clear , C elegans homolog

SLC2A1	Solute carrier family 2 member 1
SOS1	Son of sevenless homolog 1
SPI1	Spleen focus forming virus proviral integration oncogene
SPR72	Signal recognition particle, 72-KD
SPRED1	Sprouty related EVH1 domain containing protein1
SRSF2	Splicing factory serine/arginine-rich 2
STAT5	Signal transducer and activator of transcription 5A
SUZ12	Suppressor of zeste 12, drosophila homolog
TAL1	T-cell acute lymphocytic leukaemia 1
TET2	TET oncogene family member
TERC	Telomerase RNA component
TERT	Telomerase reverse transcriptase
TNK	Tyrosine kinase inhibitor
TP53	Tumour protein 53
TPO	Thrombopoietin
TRM	Treatment related mortality
tSNE	t-distributed stochastic neighbor embedding
U2AF1	U2 small nuclear RNA auxiliary factor 1
UMI	Unique molecular identifier
UPD	Uniparental disomy
UPN	Unique participant number
VNN2	Vanin 2
WAS	Wiskott-Aldrich-Syndrome
WBC	White blood cell count
WGS	Whole genome sequencing
WHO	World Health Organisation

WIMM	Weatherall Institute of Molecular Medicine
ZRSR2	Zinc finger CCCH domain and RNA binding motif containing serine/arginine rich protein 2

Chapter 1: INTRODUCTION

1. INTRODUCTION

1.1 Paediatric Myelodysplastic Syndromes (MDS)

1.1.1 Definition and incidence of Paediatric MDS

Paediatric myelodysplastic syndromes are a rare group of clonal disorders characterised by inefficient haematopoiesis and frequent progression to acute myeloid leukaemia (AML). The incidence of Paediatric MDS is estimated to be 1-4 cases/ million (1) and together these syndromes constitute 1-3 % of childhood leukaemias (2),(3),(4). Typically, MDS is characterised by cytopenias that persist for more than 3 months, and in children this is predominately thrombocytopenia and neutropenia. At presentation the cytopenias are often mild with a recent review reporting a mean presenting platelet count of $150X10^9$/L, a mean neutrophil count of $1X10^9$/L and anaemia (haemoglobin; Hb <100g/L) in <50% of cases (5).

1.1.2 Clinical presentation

Although MDS in children may present as an incidental finding in asymptomatic children particularly when the cytopenias are mild or moderate, more often children are unwell at presentation. In particular, MDS may present with recurrent life-threatening infections requiring hospitalization; or with bleeding and anaemia, requiring platelet and blood transfusions. The age at presentation and the complications vary hugely among the disease subgroups groups but the majority of cases present around 7 years of age, with an equal distribution among males and females (6),(7),(8),(9). Cases with advanced MDS can present at slightly older age and can be seen more frequently in males (10). Some children present with

quick disease progression to AML, which currently has a 5-year survival of ~ 60-70% (11). To date, there are no evidence-based algorithms to allow clinicians to identify which children are likely to progress early. In addition, disease classification and diagnosis is extremely challenging due to the heterogeneous nature of the syndromes.

1.1.3 Classification of Paediatric MDS

Paediatric MDS is largely divided into two categories: primary MDS and secondary MDS. Primary MDS is defined as myelodysplasia that arises do-novo in an otherwise healthy child (5). Secondary MDS is dysplasia following exposure to chemotherapy or radiotherapy or as part of a pre-existing haematological condition such as one of the inherited bone marrow (BM) failure syndromes. The classification of paediatric MDS (Table 1.1.1) borrows definitions from the adult MDS classification. However certain subtypes are not seen in the adult population, such as Refractory Cytopenia of Childhood (RCC) and Juvenile Myelomonocytic Leukaemia (JMML) (12), (12) (13) and equally MDS with ring sideroblasts and MDS with isolated del(5q) are exceedingly rare in children (1), this makes the adult classification system not suitable to fully characterize the paediatric cases. The updated World Health Organisation (WHO) 2016 classification system for myeloid neoplasms shown in Table 1.1.2 (12) now includes certain paediatric subtypes (RCC and JMML). The definition of MDS-related AML (MDR-AML), (Table 1.1.1) defined as a blast percentage >30%, rather than 20% as defined by the WHO criteria, and MDS with excess blast in transformation (MDS-EB-t) is retained within the Paediatric MDS classification. This aims to highlight that the blast percentage is not sufficient for differentiating AML from MDS, and that this group of children do not benefit from AML-type chemotherapy alone (5). Due to the differences between MDS presenting

in adults and children the classification scheme used in table 1.1.1 established by the European Paediatric MDS Working group (EWOG-MDS), allowing classification of 95% of patients (5).

Table 1.1.1: Classification of Paediatric MDS

Myelodysplastic syndromes (MDS)
• Refractory cytopenia of childhood • MDS with excess blasts (MDS-EB) • MDS with excess blasts in transformation (MDS-EB-t)* • MDR-AML • MDS unclassifiable • Myeloid neoplasms with germline predisposition • Treatment related MDS
Myelodysplastic/Myeloproliferative neoplasms (MDS/MPN)
• Chronic myelomonocytic leukaemia (CMML) • Atypical chronic myeloid leukaemia (aCML), *BCR-ABL1*$^{-}$ • Juvenile myelomonocytic leukaemia (JMML) • MDS/MPN, unclassifiable

Diseases seen only in children shown in red font

*MDS-EB-t: Persistent cytopenias with 20-29% blasts in BM or peripheral blood

** RCC: Persistent cytopenia (> 3 months) with <5% blasts in the BM or <2% blasts in the peripheral blood (this is regarded as a 'low grade' form of MDS)

Table 1.1.2: WHO Classification of myelodysplastic syndromes

Myelodysplastic syndromes (MDS)

• MDS with single lineage dysplasia
• MDS with ring sideroblasts (MDS-RS)
  ○ MDS-RS with single lineage dysplasia
  ○ MDS-RS with multilineage dysplasia
• MDS with multilineage dysplasia
• MDS with excess blasts (MDS-EB)
• MDS with isolated del(5q)
• MDS unclassifiable
• Refractory cytopenia of childhood**
• Myeloid neoplasms with germline predisposition

Myelodysplastic/Myeloproliferative neoplasms (MDS/MPN)
• Chronic myelomonocytic leukaemia (CMML) • Atypical chronic myeloid leukaemia (aCML), *BCR-ABL1*$^{-}$ • Juvenile myelomonocytic leukaemia (JMML) • MDS/MPN with ring sideroblasts and thrombocytosis (MDS/MPN-RS-T) • MDS/MPN, unclassifiable

Diseases seen only in children shown in red font

Despite advances in molecular analysis, the classification of certain children can be very challenging as patients may demonstrate features of more than one subtype of MDS. Figure 1.1.1 outlines the main MDS subtypes and illustrates the overlap of conditions that can be associated with myelodysplasia in childhood.

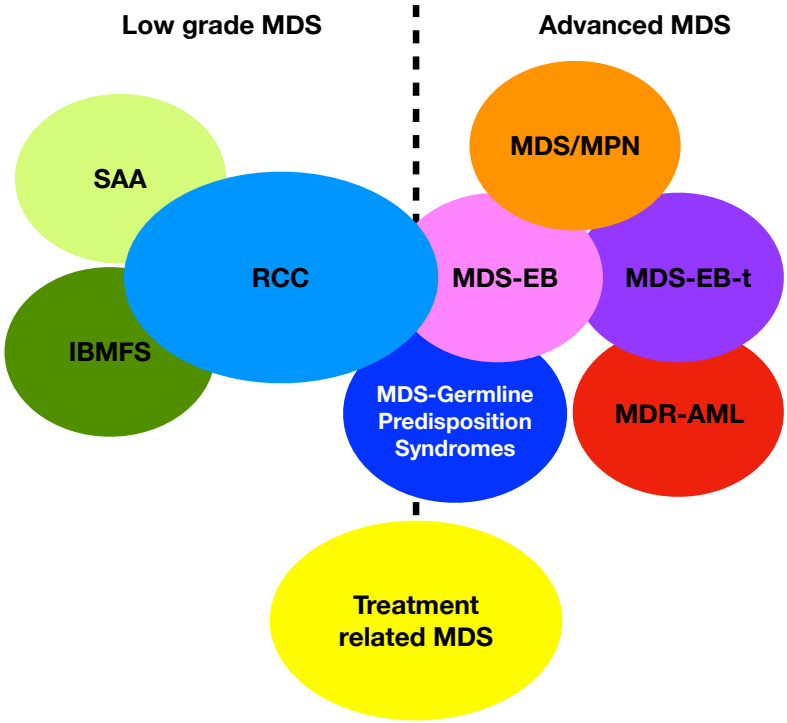


Figure 1.1.1: Landscape of Paediatric MDS (14)and overlap of the different disease subtypes: RCC: Refractory Cytopenia of Childhood, SAA: Severe Aplastic Anaemia, IBMFS: Inherited Bone Marrow Failure Syndromes, MPN: Myeloproliferative disorders, MDS- EB: MDS with excess blasts, MDS- EB-t: MDS with excess blast in transformation, MDR-AML: MDS related Acute Myeloid Leukaemia

1.1.3.1 Refractory cytopenia of childhood (RCC)

This is the most frequent subtype of MDS in childhood and is defined as persistent cytopenia (> 3 months) with <5% BM blasts and <2% blasts in the peripheral blood (5). Males and females are equally affected (15). The median age at presentation is 7-8 years although children can present at any age (7),(16). The main characteristic features of this group is the hypocellular BM, which is present in ~80% of cases (1), (17). RCC can be extremely challenging to differentiate from acquired severe

aplastic anaemia and inherited BM failure syndromes (18) as all of these conditions present with hypocellular bone marrow. The risk of disease progression in children with RCC is largely depended on the presence of cytogenetic abnormalities with 80% of children with abnormal cytogenetics progressing within 6 years from diagnosis (1). The risk of disease progression in patients with normal cytogenetics is not been formally assessed due to the lack of comprehensive long term follow up of the disease.

1.1.3.2 Advanced MDS

Advanced MDS is considered to encompasses two sub-groups: MDS- with excess blasts (MDS-EB) and MDS with excess blasts in transformation (MDS-EB-t). MDS–EB is defined as cytopenia with >2% blasts in the peripheral blood and >5% blasts in the BM. MDS-EB-t is defined as cytopenias with blasts between 20-29% in either the peripheral blood or BM. Children may be stable for weeks or months or progress very quickly to MDR-AML. Most patients present with advanced sub-types of MDS present at a slightly older age than those with RCC or JMML, typically at~ 10 years of age, with males been more frequently affected.

1.1.3.3 Therapy-related myelodysplasia

The incidence of secondary MDS in children following treatment for solid tumours with standard protocols is 5-11% and 1-5% for children treated for ALL (19),(20), (21),(22),(23) .Treatment regimens that include high dose alkylating agents and topoisomerase inhibitors II are associated with a higher risk of developing secondary MDS. The interval between the first treatment and the development of MDS varies from months to years, depending on the type of treatment previously received (24). For example, children that receive topoisomerase II inhibitors tend to present early,

often with MLL rearrangements (25); while those that received alkylating agents or radiotherapy present after 4-7 years, often with cytogenetic aberrations of chromosome 5 and 7. Other cytogenetic abnormalities associated with secondary MDS are 21q abnormalities and loss of 11q, 20q, or 5q often with complex karyotypes and *TP53* or *DNMT3A* mutations (26),(27),(28).

1.1.4 Overview of other secondary Paediatric MDS/ overlap syndromes

A number of inherited and acquired conditions, such as severe aplastic anaemia, are associated with MDS, either as conditions that can progress and develop myelodysplasia or overlap syndromes that display features of two diseases at presentation. The dysplasia has been hypothesised as a reflection of 'stress' haematopoiesis perhaps triggered directly by the genomic instability observed in these conditions and/or secondary to inflammation and an exaggerated immune response mediated by cells of the BM microenvironment. The biology of MDS differs from the inherited IBMFS; with the de novo MDS cases typically associated with evidence of clonal haematopoiesis such as monosomy 7 or *GATA2* haploinsufficiency (29). IBMFS may display evidence of a polyclonal background typically conferred by the genomic instability observed (30). This may be due to loss of apoptotic checkpoints for example in Dyskeratosis Congenica; or ineffective DNA repair in the case of Fanconi Anaemia; these mechanisms can promote the cell survival leading to acquisition of secondary events leading to the development of clonal haematopoiesis and AML (31) . From a clinical point of view, it is important to identify these conditions early, because of the risk of disease progression to AML and/or the need for treatment modification due to the underlying disease as well as implications for family genetic counseling. A summary of the conditions that can predispose children to developing MDS is outlined in Table 1.1.3 (32).

Table 1.1.3: Secondary MDS: conditions that predispose children to MDS.

A. Inherited Bone Marrow Failure Syndromes • Fanconi Anaemia • Diamond- Blackfan Anaemia • Shwachman-Diamond syndrome • Telomere biology disorders • Congenital amegakaryocytic thrombocytopenia • Severe congenital neutropenia
B. Severe Aplastic Anaemia
C. Germline predisposition and pre- existing platelet disorders • Myeloid neoplasms with germline *RUNX1* mutation • Myeloid neoplasms with germline *ANKRD26* mutation • Myeloid neoplasms with germline *ETV6* mutation
D. Myeloid neoplasms with germline predisposition and other organ dysfunction • Myeloid neoplasms with germline *GATA2* mutation • MDS with germline *SAMD9/SAMD9L* mutation • MDS with germline mutations of *SRP72*
E. Other germline predisposition to leukaemia syndromes • Li-Fraumeni syndrome • Bloom's syndrome

1.1.4.1 Inherited BM Failure Syndromes

These conditions often present with cytopenias and dysplasia as many children lack the other clinical pathognomonic features (33). One study of 110 paediatric MDS patients treated by BM transplantation (BMT) found mutations in BM failure-associated genes in 13.6% of patients (34) although a family history of cytopenias or malignancy was only present in 52% of patients tested and physical abnormalities in 24% of patients. Niemeyer et al also reported that 15% of possible MDS cases collected by the EWOG-MDS group turned out to have an unsuspected diagnosis of Fanconi Anaemia (1). Overall it is estimated that around 10% of IBMFS can present due to cytopenias without the presence of a clinical phenotype, but further observational studies are required in order to formally address this point (1). Such patients should be identified prior to BMT to ensure they are treated using appropriate conditioning protocols and that potential family donors are screened for

the disease. Most, if not all, of the inherited BM failure syndromes have a risk of subsequent development of AML which can be as high as 75% by age 18 in patients with Fanconi Anaemia (35).

1.1.4.1.1. Fanconi Anaemia (FA). This is an autosomal recessive, or occasionally X-linked, disorder due to constitutional mutations in one of 21 FA genes necessary for normal DNA interstrand crosslink (36). The estimated incidence of FA is 1/100 000 (37). In addition to BM failure, FA is associated with multiple congenital anomalies, including skeletal anomalies (particularly of the upper limbs), café au lait spots, facial dysmorphology and structural renal defects (38). Haematological manifestations are the predominant finding and MDS is reported in at least 50% of cases (39), often with associated cytogenetic abnormalities (40).

1.1.4.1.2.Dyskeratosis Congenita (DKC). This extremely rare syndrome has an incidence of 1/million (41) and may be autosomal dominant, recessive or X-linked. Typically, patients exhibit a triad of abnormal skin pigmentation, leukoplakia and nail dystrophy as well as BM failure (41). More than 10 DKC genes have so far been identified and all are involved in telomere maintenance (42). MDS is common in DKC and progression to AML has been reported in 25% of cases by 18 years of age (43), (35).

1.1.4.1.3.Shwachman–Diamond Syndrome. This autosomal recessive condition is caused by bi-allelic mutations at the *SBDS* gene (44) ,(45). In addition to BM failure, patients typically have exocrine pancreatic insufficiency, skeletal abnormalities and, sometimes, mental retardation. Patients can present with haematological manifestations and cytopenias without any other gastrointestinal or neurological symptoms (46). Progression to MDS can be seen in 20-30 % of patients, especially

in patients with cytogenetic abnormalities (35),(47) with particularly poor prognosis following bone marrow transplantation (48).

1.1.4.1.4.Severe Congenital Neutropenia (SCN). This heterogeneous group of disorders can be due to mutations in an increasingly large number of genes including *ELANE*, *CSF3R, HAX1, G6PC3, GFI1* and *WAS* with autosomal or X-linked patterns of inheritance (49). Children with these disorders typically present early in childhood with mouth ulcers, poor growth, infections and often gastrointestinal symptoms (50).The risk of evolving to MDS/AML at 10 years is 15-40%, with an increased risk for patients on G-CSF, particularly for those with a poor G-CSF response (51), (52), (53), (51).

1.1.4.1.5.Diamond-Blackfan Anaemia (DBA). Typically DBA presents with red cell anaemia due to red cell aplasia in the first year of life but it may also progress to more global BM failure and/or to MDS and to AML. DBA is caused by mutations in ribosomal protein genes. It may be inherited in an autosomal dominant fashion but there is variable penetrance and many cases are *de novo* (54). Many patients have associated congenital skeletal, cardiac and genitourinary anomalies (55),(56), (57) .

1.1.4.1.6. Congenital Amegakaryocytic Thrombocytopenia (CAMT). This rare autosomal recessive disorder is caused by mutations of the thrombopoietin (TPO) receptor gene *MPL* or in the *TPO* gene itself (58). Thrombocytopenia usually presents soon after birth and is often severe. Although in some cases, thrombocytopenia improves over the first few years, other patients develop progressive BM failure and there is also an increased risk of MDS/AML and acquisition of clonal cytogenetic abnormalities (59) (60). Most patients do not have associated congenital anomalies but cardiac, ocular and brain malformations have been described (58)

1.1.4.2 Severe Aplastic Anaemia (SAA)

SAA is an autoimmune-driven BM failure disorder with an increased risk of MDS (10 -15%) and AML (7%) (61), (62). As the risk appears to be higher in non-severe aplastic anaemia, it has been suggested that some RCC cases may be misdiagnosed as SAA (63),(64) .The risk of MDS is reported to be higher in patients with a poor response to G-CSF(65). Molecular defects in *PIGA, ASXL1, BCOR* or *DNMT3A* genes can often be detected even before disease progression to MDS; (66) however, there is no evidence that the presence of mutations alone indicates disease progression.

1.1.4.3 Germline predisposition syndromes

The use of high throughput sequencing is leading to increasing recognition of causal germline variants in children with MDS. Several of these syndromes present with thrombocytopenia and an autosomal dominant pattern of inheritance as the most prominent clinical feature. These include *ANKRD26*-related thrombocytopenia, *ETV6*-related thrombocytopenia and *RUNX1*-associated familial platelet disorder. ANKRD26-related thrombocytopenia, due to mutations in the *ANKRD26* gene (67), (68), presents with moderate thrombocytopenia and is associated with a 23-fold and 12-fold increased risk of developing MDS or AML, respectively (69). ETV6-associated thrombocytopenia is caused by germline mutations in *ETV6*, a tumour suppressor gene frequently implicated in childhood leukaemias. Patients with germline *ETV6* mutations can present with bleeding, thrombocytopenia and macrocytosis and are often initially treated as autoimmune thrombocytopenia (70),(71). These patients have a high prevalence of haematological malignancies such as ALL, MDS-EB, CMML and mixed phenotype acute leukaemia, as well as

solid tumours. *RUNX1*-associated familial platelet disorder with increased risk of myeloid malignancy (FDP/AML) also presents with mild to moderate thrombocytopenia and may be misdiagnosed as autoimmune thrombocytopenia (72). This condition is due to large deletions, microdeletions or point mutations in the *RUNX1* transcription factor gene (73) and carries a high risk of MDS/AML (~35%) (74) and progression is often associated with a second molecular or cytogenetic abnormality such as trisomy 21 (75).

Other germline mutations tend to present with MDS/AML rather than isolated thrombocytopenia. These include CEBPA-related MDS/AML and familial MDS due to mutations in *GATA2, SPR72, DDX41, SAMD9 and SAM9DL*.

1.1.4.3.1 CEBPA-related MDS/AML This autosomal dominant disease is caused by mutations caused by N-terminal frameshift mutations in the CEBPA transcription factor gene, that result in inability to cause proliferation arrest. At the time of AML progression, patients often acquire a second, in-frame somatic mutation at the C-terminal region which disturbs DNA binding and transcriptional activation (76),(77).Typically patients present in the early twenties and have a high predisposition to AML/MDS. AML developed in this setting carries a more favorable prognosis compared to other AML subtypes (78).

1.1.4.3.2 GATA2 syndromic disorders: Inherited mutations in the GATA2 gene cause a wide spectrum of diseases that include monocytopenia, B-cell lymphopenia, NK-cell deficiency, congenital deafness, lymphoedema (MonoMAC, Emberger syndrome) and cytopenias complicated by an increased risk of developing MDS or AML (79),(80),(81). GATA2 is a transcription factor that has a critical role in haematopoiesis and regulates haematopoietic stem cell survival and self renewal ,

by interacting with a number of other important transcription factors such as SPI1, FLI1, TAL1, LMO2 and RUNX1(82). *GATA2* Mutations may be inherited or somatically acquired (83). Typically patients with *GATA2* disease are older children (~12 years of age) and cases of advanced MDS are more frequent than low risk MDS, with a risk of 15% and 4% respectively (84). In addition, 70% of patients with *GATA2* mutations are also found to have monosomy 7. The majority of germline *GATA2* mutations are de novo events, without any family history. In *GATA2* families the mutations appears to be inherited in an autosomal dominant fashion. Typically mutations result in the loss of the second zinc finger domain or non-coding variants in the + 9.5 Kb regulatory region of *GATA2* resulting in haploinsufficiency (79),(85), (86). The mutations lead to defective stem cell proliferation and defects in lymphatic development (87),(88). One recently published study by the EWOG MDS group, including 500 children with germline *GATA2* mutations, demonstrated that patients with germline *GATA2* mutations do not have a worse prognosis compared with other MDS subtypes; however they do have a higher risk of disease progression and therefore BMT is recommended (84), particularly at the hypocellular MDS phase before severe complications develop. Potential family donors should be screened for *GATA2* mutations. Acquisition of secondary mutations, for example in *ASXL1 NRAS, RUNX1, STAG2, IDH2, TP53* and/or *SETBP* genes *(89)* is associated with disease progression (90),(91).

1.1.4.3.3 SPR72 mutations familial MDS: This autosomal dominant condition is due to mutations in the *SPR72* gene that disturb translation of secretory and extracellular proteins of the endoplasmic reticulum. It presents at any age from childhood to adulthood and it is characterised by cytopenias, BM failure and/or myelodysplasia together with auditory abnormalities (92),(93).

1.1.4.3.4 DDX41-associated familial MDS/AML syndrome This autosomal dominant condition has a high risk in developing MDS (94) and typically presents with advanced disease in late adulthood but familial cases can be diagnosed at an earlier age (95). DDX41 mutations lead to alterations in mRNA splicing and RNA processing. Additional secondary mutations are found in *TP53*, *RUNX1*, and *LUC7L2*; and DDX41 mutations have been associated with del(5q) in ~ 26% of adult MDS cases.

1.1.4.3.5 SAMD9 and SAMD9L mutations and MDS This autosomal dominant condition is caused by mutations in the SAMD9 or SAMD9L genes, that negatively regulate cell proliferation (96),(97),(98). SAMD9L is specifically involved in regulating HSC renewal and differentiation and also plays a role in leucocyte maturation (99). Typically patients have heterozygous missense gain-of-function mutations (100),(99). Patients with inherited *SAMD9* mutations have a syndrome known as MIRAGE syndrome which includes myelodysplasia, infections, growth retardation, adrenal hypoplasia and enteropathy (101),(102). By contrast, patients with inherited *SAMD9L* mutations have Ataxia-Pancytopenia syndrome (ATXPC) which is characterised by myelodysplasia, ataxia due to cerebellar dysfunction and intracellular infections due to B and NK cell deficiency (99). Interestingly, these mutations are lost in the when cells acquire monosomy 7, which has led to the suggestion that this is the first example of human adaption by aneuploidy (101). Since not all patients have non-haematological features, diagnosis should not rest on the presence of clinical manifestations (103). Heterozygous loss of function mutations of both *SAMD9* and *SAMD9L* are reported in the general population and appear to be tolerated (104).

1.1.5 Molecular and cytogenetic characteristics of Paediatric MDS

1.1.5.1 Cytogenetics:

Cytogenetic abnormalities are detected in 30-50% of children with MDS (105). Monosomy 7 is the commonest cytogenetic abnormality, occurring in 25- 30% of cases (105),(106),(107). Trisomy 8 and trisomy 21 are the next most common abnormalities seen. Cases of trisomy 8 mosaicism can be difficult to identify as they might lack any clinical features and therefore cytogenetic analysis should be performed in all cases of suspected myelodysplasia (108),(7),(105),(109). In RCC 80-90% of patients have normal cytogenetics while in advanced MDS abnormal cytogenetics are present in 55-60% of patients (1),(105). Within the RCC group, monosomy 7 is more frequently found in patients with normo- or hyper-cellular BM rather than in the hypocellular MDS (Figure 1.1.2).

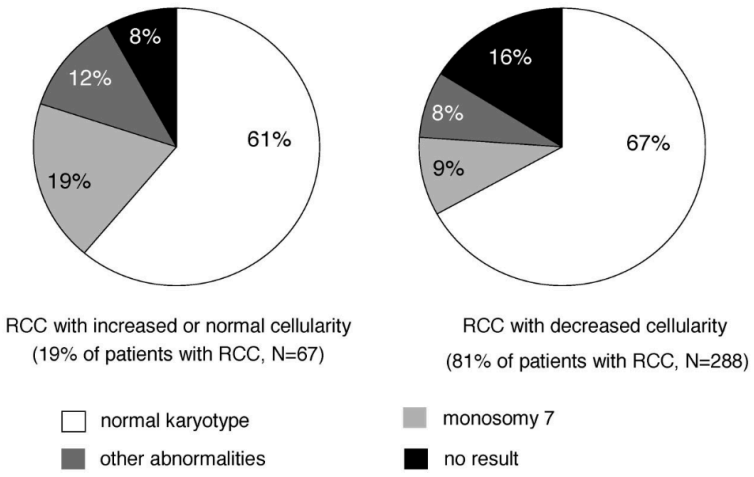


Fig 1.1.2: Niemeyer et al: (1). Association between BM cellularity and cytogenetic abnormalities within the RCC patient group.

Patients with RCC and monosomy 7 are more likely to progress than those without monosomy 7, with previous studies reporting disease progression in 1.9 years from diagnosis, and 80% probability to progress after 6 years from diagnosis. Patients with trisomy 8 or normal karyotype have a much more stable disease course. Sole

gain of chromosome 1q may persist for many years without any evidence of disease progression (110). Spontaneous remission of monosomy 7 is very rare but has been previously reported in infants (111),(112). In children with secondary MDS cytogenetic abnormalities are detected in 76% of patients (105).

1.1.5.2 Molecular characteristics

The molecular characteristics of childhood MDS have only recently started to be investigated (113),(29). A recent study performed by the EWOG-MDS group screened 50 patients (RCC n= 31 and MDS-EB n=19) using an extensive NGS panel of 105 genes identified the molecular basis of RCC in 13% of cases and of advanced MDS in 68% of cases (113). The commonest mutations were in *SETBP1* (18%) while a significant number of patients had mutations in *RUNX1* (6%), *ASXL1* (8%), *PTPN11* (6%), *NRAS* (4%), *KRAS/EZH2/NF1/NPM1* (2% for each). Interestingly, monosomy 7 was associated with *SETBP1* (36%), *ASXL1* (16%) and *RUNX1* mutations (12%). A second study used a whole exome sequencing approach in 46 patients and identified RAS pathway mutations in 45% of patients and *SAMD9* or *SAMD9L* mutations in 17% of cases (29). However, in this study 30% of patients were MDS/MPN patients rather than de novo MDS, which might positively influence the results towards RAS- pathway mutations, due to the very well documented high frequency of RAS mutations in JMML. Germline mutations constitute 30-40% of cases and these can include mutations in *GATA2, RUNX1, PAX5*, *ETV6, DDX41, ANKRD26, SRP72, SAMD9*, *SAMD9-L*, CEBPA and *RUNX1* genes as outlined above (114), (70), (79), (115), (94), as well as mutations associated with BM failure syndromes (see Table 1.1.4) (28),(114).

Table 1.1.4: Genes implicated in MDS associated with inherited BM failure syndromes or other inherited predisposition syndromes

Dyskeratosis congenita	Fanconi Anaemia	Diamond Blackfan Anaemia	Severe Congenital Neutropenia	Other inherited BMF/MDS
CTC1	*FANCA*	*GATA1*	*ELANE*	*ABCB7*
DKC1	*FANCB*	*RPS7*	*G6PC3*	*AK2*
NHP2	*FANCC*	*RPS10*	*GFI1*	*ANKRD26*
NOP10	*BRCA2/FANCD1*	*RPS17*	*HAX1*	*ATR*
RTEL1	*FANCD2*	*RPS19*	*WAS*	*LIG4*
TERC	*FANCE*	*RPS24*	*CSF3R*	*MPL*
TERT	*FANCF*	*RPS26*		*NBN*
TINF2	*FACG*	*RPS27*		*RMRP*
WRAP53	*FANCI*	*RPS27A*		*SBDS*
NOLA3/NOP10	*FANCJ (BRIP1/BACH1)*	*RPS28*		*SRP72*
NOLA2/NHP2	*FANCL*	*RPS29 (TSR2)*		*DDX41*
ACD/TPP1	*FANCM*	*RPL5*		*SAMD9L*
PARN	*FANCN /PALB2*	*RPL11*		*NF1*
NAF1	*FANCO/RAD51C*	*RPL35A*		*BLM*
STN1	*FANCP/SLX4*	*RPL23*		*TP53*
	FANCQ/ERCC4	*RPL26*		*PAX5*
	FANCR/RAD51	*RPL7*		*GATA2*
	FANCS/BRCA1	*RPL31*		*CEBPA*
	FANCT/UBE2T	*RPL36*		*RUNX1*
	FANCU/XRCC2			
	FANCV/REV7			

1.1.6 Differences between adult and paediatric MDS

There are several differences between Paediatric MDS and the adult disease (Table 1.1.5). Firstly the adult disease is far more common than the paediatric MDS with an incidence of 36 patients in every 100 000 cases for adults compared to 1-2 cases/ million for the paediatric cases (116). Adult MDS is a disease of the elderly population and arises after the accumulation of more than 2 or 3 molecular defects frequently arising from epigenetic modifiers and spliceosome genes (117). 30-40% of Paediatric MDS are characterised by inherited predisposition syndromes with a paucity of somatic mutations that are detected very early on in life, as outlined in the

section 1.1.5.2 and there is a different range of cytogenetic abnormalities in the two diseases. Finally, the main form of treatment in children is BMT whereas treatment in adults depends on the clinical symptoms and the patient's performance status.

Table 1.1.5: Differences between MDS in children and adults

Characteristics	Paediatric MDS	Adult MDS
Annual incidence (cases per million)	1-2	40
Age of presentation	7 years	70 years
Inherited predisposition	~ 30-40% of cases	Very rare
Etiology	Majority of cases de novo (primary)	Many cases are secondary to chemotherapy or radiotherapy
Most frequent haematological features at presentation	Neutropenia and thrombocytopenia	Anaemia
Systemic features	Dysmorphism, skeletal, cardiovascular, genitourinary, gastrointestinal	None/ very rare
Distinct disease subgroups	Refractory cytopenia of childhood (RCC) and JMML	MDS with ring sideroblasts
BM cellularity	Hypocellular in the majority of cases	Hypercellular/ normocellular/hypocellular
Number of mutations detected	One or paucity of molecular events	> 2-3 molecular alterations detected per sample
Most frequent molecular abnormalities	RAS pathway mutations, transcription factors (*RUNX1, GATA2*) SETBP1 Inherited BM failure genes (*SBDS, ELANE, TERT, TERC*) Other *ETV6* and SAMD9/SAMD9-L	*TP53, SF3B1, SRSF2, U2AF1, ZRSR2, TET2, DNMT3A, ASXL1, RUNX1*
Commonest cytogenetic abnormalities	Monosomy 7	Chromosome 5q deletion
Aim of treatment	Curative	Usually palliative
Treatment	BMT Blood products Azacitidine Immunosuppressive therapy	Blood products Azacitidine Lenalidominde Immunosuppressive treatments BMT in a small proportion of patients

1.1.7 Pathophysiology of Paediatric MDS

The refractory nature of Paediatric MDS suggests that the disease-propagating population may reside within the haematopoietic stem cell (HSC) or multipotent progenitor compartment, as previously demonstrated in adult MDS (118). However, little is known about the haematopoietic stem cell and progenitor compartments in paediatric cases (119). As the biology and disease evolution in paediatric cases is different from the adult cases, this requires investigation using paediatric samples. Furthermore, it is likely that patients with inherited MDS predisposition syndromes have a different disease evolution than the acquired cases. Monosomy 7 is the most important determinant of disease progression and therefore it is likely that this cytogenetic abnormality contributes to clonal expansion and perhaps leads to acquisition of additional driver effects that lead to the development of MDS. This is supported by a recent study in patients with inherited BM failure syndromes and monosomy 7 which used single cell RNA sequencing to show at a single cell level that the cells with monosomy 7 cluster separately from the healthy controls and were characterised by downregulation of genes associated with DNA repair, apoptosis and immune response (120).

Another mechanism that has been implicated in the pathogenesis of low grade MDS (RCC) is the role of the immune system, given that a proportion of RCC cases respond to pre-BMT immunomodulative therapy. This is supported by results of analysis of TCR Vβ CDR3 heterogeneity which has revealed an oligoclonal pattern in patients with RCC, with an increased frequency of activated CD8+ T cells (121), (122). However, there is no correlation between patients' clinical characteristics

(including response to anti-thymocyte globulin; ATG) and the presence or not of oligoclonality (121),(122),(123).

1.1.8 Current diagnostic pathway for paediatric MDS

As the diagnosis of MDS is complex and often challenging, most specialist centres employ a diagnostic algorithm similar to the one shown in Figure 1.1.3. The majority of patients present with single lineage cytopenia or bicytopenia that persist for more than 3 months and in some cases MCV might be raised similarly to the adult MDS cases. Secondary non-haematological causes of childhood cytopenias should be excluded by performing investigations to identify viral infections; metabolic diseases, such as mevalonate kinase deficiency; mitochondrial disorders (Pearson syndrome); autoimmune conditions and severe nutritional deficiencies.

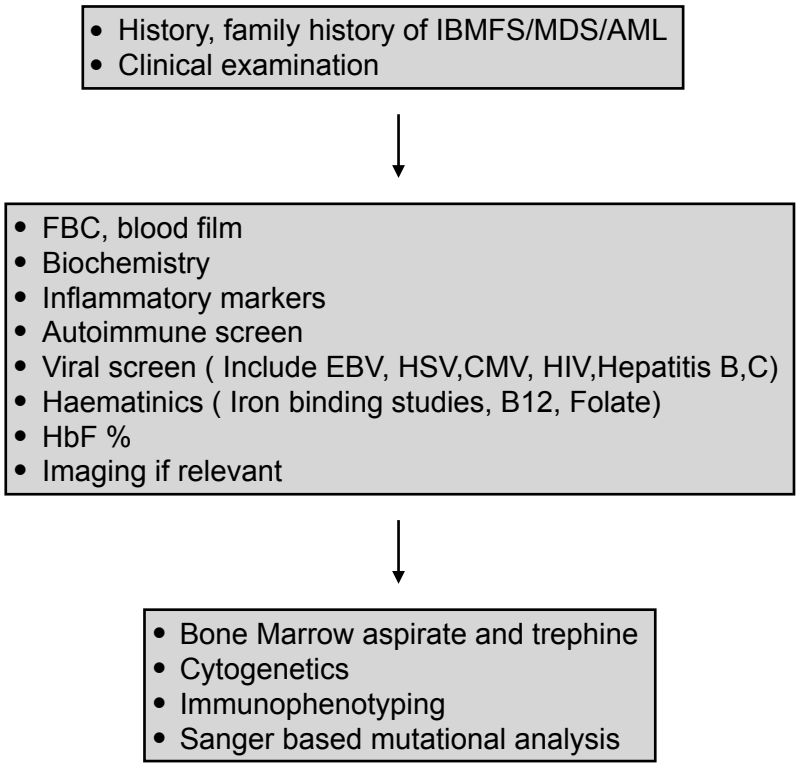


Figure 1.1.3: Routine diagnostic algorithm for Paediatric MDS (adopted from Great Ormond Street Hospital)

It is important to carry out a very detailed clinical examination to identify any dysmorphic features, congenital anomalies, mucocutaneous stigmata of systemic disease (café au lait spots, nail dystrophy, skin hypopigmentation) or other features suggestive of an underlying predisposition syndrome, such as Fanconi anaemia or dyskeratosis congenita. Figure 1.1.4 includes a summary of the key features seen in MDS syndromes with inherited predisposition. Family history of MDS/AML or haematological conditions, relations and phenotypes of the affected family members can also be very important in identifying the underlying disorder.

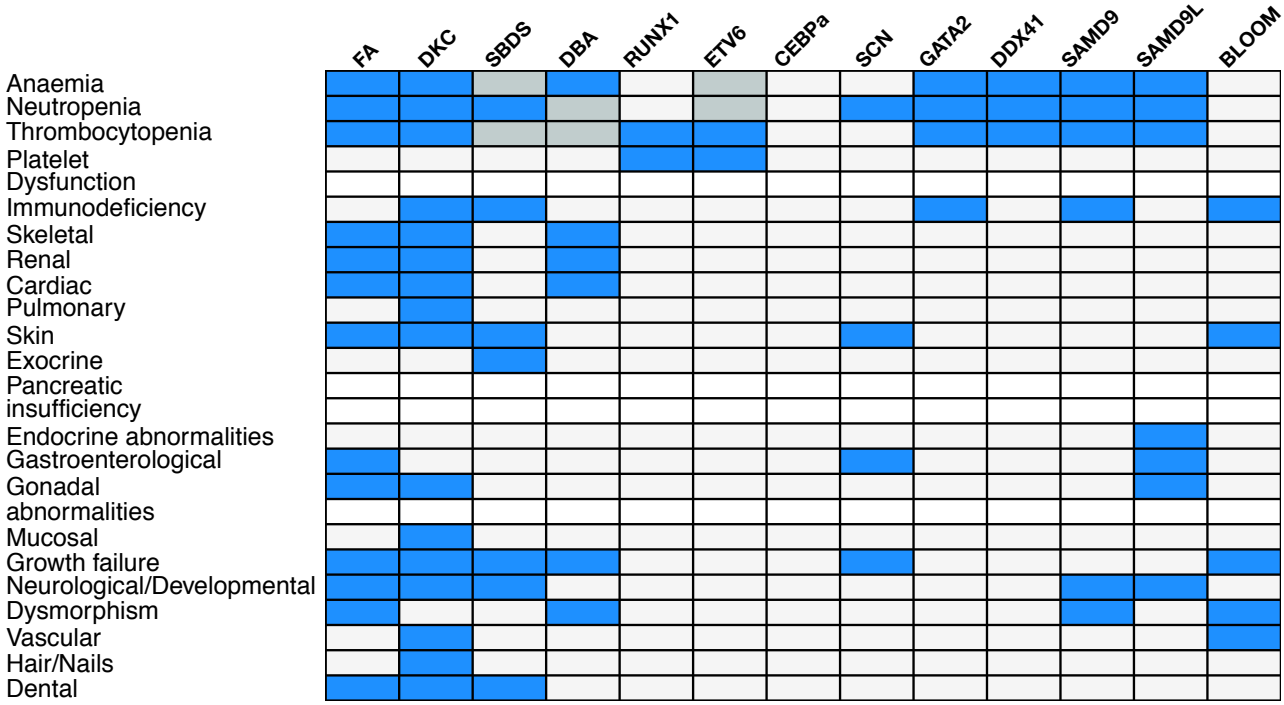


Figure 1.1.4: Comparison of clinical features in Paediatric MDS predisposition syndromes. Column represents a condition as indicated above of the column. Dark blue squares indicate presence symptoms present in disease.Grey squares indicated milder phenotype. Fanconi Anaemia (FA), Dyskeratosis congenital (DKC), Shwachman–Diamond Syndrome (SBDS), Diamond Blackfan Anaemia (DBA), SCN (Severe Congenital Neutropenia)

1.1.8.1 Haematological investigation of Paediatric MDS

As well as cytopenias, many children with MDS have a raised mean corpuscular volume (MCV). Blood films typically show red cell abnormalities such as anisocytosis, tear drop cells and basophilic stippling; leucocyte dysplasia, including hypogranular and/or hyposegmented neutrophils and pseudo-Pelger forms, myeloid precursors, nucleated red cells and circulating blasts; dysthrombopoiesis, such as giant platelets or platelet anisocytosis.

Bone Marrow (BM) aspirates are most often hypocellular but may also be normocellular or hypercellular. BM morphology shows trilineage dysplasia with dyserythropoiesis, dysgranulopoiesis and reduced/absent megakaryocytes or the presence of micromegakaryocytes. Lymphocytes, plasma cells and mast cells may be increased. The presence of dysplasia may not be diagnostic and centralized review and repeat BM aspirate is often recommended for challenging cases (5). The differential diagnosis between MDS, SAA and inherited BM failure can be particularly difficult even for highly specialist pathologists. Differentiation between RCC and SAA may be possible by using morphological criteria established by Baumann et al (17). Differentiation between RCC and IBMFS may not be possible and a combination of clinical, haematological and molecular investigations is often required in order to make the diagnosis. As noted above, the % of blasts determines the type of MDS present. It is important to note that the presence of blasts in one specimen is not sufficient for discriminating between MDS-EB, MDS-EB-t and AML. Differentiation from frank AML cases is essential as the biology and management of these is entirely different. For borderline cases where the % of the blasts is between 20% and 30%, repeat BM aspirate/ biopsy two weeks after the initial examination is

recommend. Often cytogenetics and molecular defects can also be helpful in the identification of AML cases.

Flow cytometry of peripheral blood or BM cells may be helpful in differentiating between SAA, RCC and inherited BM. One study that analysed 81 RCC patients, 17 SAA patients and 7 MDS-EB or MDS-EB-t cases, revealed expansion of the lymphoid compartment in the RCC and SAA cases and a reduction in the myeloid, erythroid and precursor compartments in all disease subgroups except in MDS-EB where an expansion of precursors was seen (124).

1.1.8.2 Cytogenetics and molecular genetics for diagnosis of Paediatric MDS

Common MDS-associated cytogenetic abnormalities are described in section 1.1.5.2. Certain key cytogenetic abnormalities can be helpful in differentiating between MDS and conditions that share similar presenting features. Typical abnormalities in MDS are monosomy 7, trisomy 8 and trisomy 21, whereas inherited BM failure syndromes may have isochrome 7, and t(8;21) or t(15;17) definitively define of AML.

Mutational analysis, using Sanger sequencing or, if available, a targeted NGS panel, is very helpful in identifying clonal disease and establishing the diagnosis. Targeted molecular panels should include targets implicated in myeloid malignancies, *RAS* pathway mutations including *NF1*, *TP53,* transcription factors such as *GATA2* and *RUNX1* (see Chapter 3). For all patients where monosomy 7 is detected, *GATA2* analysis for the mutational hot spots should be performed.

1.1.8.3 Other investigations in Paediatric MDS

In selected patients investigations to identify the BM failure syndromes should be considered**,** e.g. chromosome breakage studies to identify Fanconi anaemia or stool elastase to identify SBDS. Where relevant, PNH should be excluded by flow

cytometry. HPLC to measure the % of HbF is useful, particularly where JMML is suspected. Renal and liver function tests should be checked and imaging of liver, kidneys and spleen should be considered especially if there are deranged renal or liver function tests or organomegaly.

1.1.9 Treatment of Paediatric MDS

BMT is the only curative treatment for children with MDS irrespective of the different sub-types (125),(126),(127). For patients with RCC the timing of BMT is difficult and depends on the presence of cytogenetic abnormalities and the severity of symptoms (32). For RCC patients with relatively stable disease over many years, not requiring any intervention and with no underlying clone, a 'watch and wait' approach is often used. Unfortunately, this means that transformation can occur before any treatment is implemented and once the leukaemia is established, it is more difficult to eradicate the disease and salvage these patients. However, RCC children with monosomy 7,or more complex cytogenetic abnormalities (more than 3 aberrations) or children that are transfusion dependent or have recurrent infections with neutrophil count $<1x10^9$/L are generally regarded as candidates for BMT (Figure 1.1.5) (32). Current treatment of choice for children with advanced MDS is myeloablative BMT, with a 5 year overall survival (OS) of 63%; relapse is the most common cause of treatment failure (128). One cycle of AML–type induction chemotherapy to reduce the pre-transplant blast count can be used in cases of MDR-AML, (32) as this reduces the risk of relapse (10),(129). However, AML chemotherapy in other types of advanced MDS does not appear to have a significant impact on the relapse rates (32). Alternative treatments used in adults, such as differentiating agents or anti-angiogenic drugs are not indicated in children. Ideally a matched sibling donor should be used with the exception of the germline cases, whereas an unrelated

donor should be used. Cord blood stem cells are not recommended for these patients (130). For low grade MDS, in the absence of a suitable donor, immunosuppressive therapy with ATG and cyclosporine A can be used, while awaiting a donor (121), (123) but only in patients that have hypocellular BM and lack cytogenetic abnormalities (131) (131),(132). A recent EWOG MDS has demonstrated an overall survival of 94% and EFS of 88% in RCC patients using low intensity regimens (133),(134). Preliminary data suggest that DNA- demethylating agents, such as azacitidine, which have been used successfully in adult MDS, might also be beneficial in paediatric cases (135).

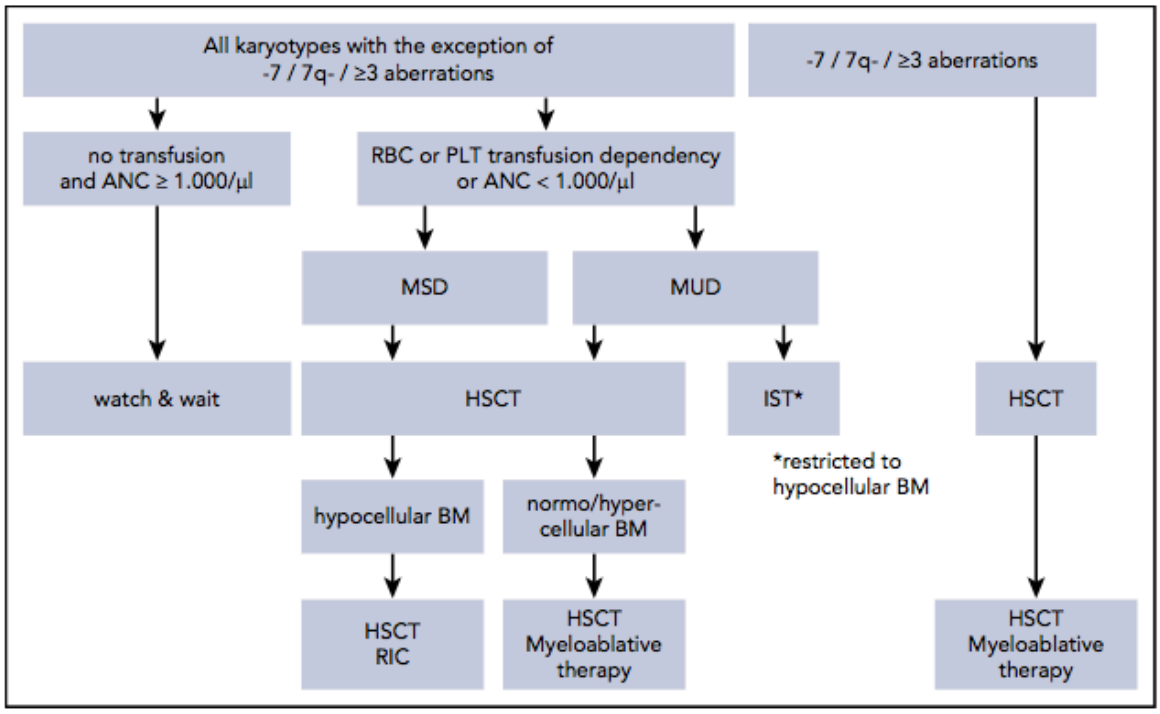


Figure 1.1.5: Indications for Bone Marrow Transplantation in children with Paediatric MDS (32). IST: Immunosuppressive therapy; MDS: matched sibling donor; MUD: matched unrelated donor; RBC: red blood cell; RIC: reduced intensity conditioning

1.1.10 Natural history of Paediatric MDS

Patients with MDS have a very variable and heterogeneous disease course. Many patients have very stable indolent disease for decades and do not require any intervention; with the diagnosis of MDS been uncovered incidentally. At the other end of the spectrum patients present with advanced disease, requiring frequent blood product support and have a very aggressive disease course. To date there are no evidenced based diagnostic and prognostic algorithms that can risk stratify patients and provide guidance on their management, especially for cases that no clonal disease can be identified (136). The adult international prognostic scoring system (IPSS) was applied in paediatric MDS cases, by the EWOG MDS group, but it was found that it is of limited value for paediatric cases (137). To date the most reliable prognostic marker is the presence and complexity of the cytogenetic abnormalities, especially monosomy 7, as it is the strongest independent prognostic marker predicting poor outcome (105). Structural complex abnormalities, defined as 3 or more chromosomal aberrations including at least one structural aberration, are also associated with a very poor prognosis (105).

1.2: Juvenile Myelomonocytic Leukaemia

1.2.1: Epidemiology, clinical presentation and disease features

Juvenile Myelomonocytic Leukaemia (JMML) is very rare aggressive childhood leukaemia characterised by clonal proliferation of monocytic and granulocytic precursors. The disease constitutes 2% of childhood leukaemias (138) and there are no distinct ethnic groups identified as being more susceptible. In the vast majority of cases, the disease is due to mutations in *RAS* pathway genes such as *PTPN11* (Protein Tyrosine Phosphatase, Non Receptor Type 11), *N-RAS* (Neuroblastoma RAS viral oncogene homolog), *K-RAS* (Kirsten rat sarcoma virus), *NF1* (Neurofibromin 1) and *CBL* (Casitas B-Lineage Lymphoma proto-oncogene) (139). In the UK, JMML incidence is estimated to be ~1-2 cases/million(140). JMML carries characteristics of both myeloproliferative and myelodysplastic conditions and is classified as an MDS/MPN overlap neoplasm (12).

JMML typically presents within the first 5 years of life including the neonatal period. The most typical age of presentation is around 2 years of age and it is more frequently seen in males (141),(142). Clinical presentation can be heterogeneous and is characterised by chronic symptoms of poor weight gain, pallor, fever, rashes and recurrent infections. Typical clinical signs of the disease includes: splenomegaly, hepatomegaly, skin rash, bruising or / and bleeding and sometimes non- specific symptoms such as failure to thrive, being generally unwell for weeks and fever (140, 143). Non-haematopoietic tissues such as the lungs, the skin and the intestines, can also be affected due to infiltration by malignant cells which may be fatal in severe cases (129).

Table 1 outlines the WHO diagnostic criteria for JMML. For patients where mutational analysis fails to identify a mutation, high fetal haemoglobin (HbF) , GM-CSF hypersensitivity and/or monosomy 7 can be used in order to make a diagnosis of JMML.

Table 1: JMML WHO diagnostic criteria (139)

I. Major diagnostic criteria: Clinical and haematological features (all 4 features mandatory)
• Peripheral blood monocyte count > 1 x 10^9 /L • Blast percentage in peripheral blood and bone marrow < 20% • Absence of BCR/ABL rearrangement • Splenomegaly
II. Oncogenetic studies (1 finding is sufficient)
• Somatic *PTPN11* or *KRAS* or *NRAS* mutation • Clinical diagnosis of NF1 or germline *NF1* mutation • Germline *CBL* mutation and loss of heterozygosity of CBL
III. Only for patients (10% of the whole number) without any oncogenic parameter, besides the clinical and haematological features listed under I (at least 2 or more of the following criteria have to be fulfilled)
• Hb F increased for age • Presence of myeloid precursors in peripheral blood smear • Monosomy 7 or any other clonal chromosomal abnormality • Spontaneous growth or GM–CSF hypersensitivity in colony assay • Hyperphosphorylation of STAT5

Some of the features of JMML resemble Chronic Myelomonocytic Leukaemia (CMML), a myeloproliferative disease seen in the adult population (144). However the pathogenesis of JMML differs greatly from the adult disease which typically presents in the elderly and is primarily caused by mutations in *SRSF2*, *ASXL1*, *TET2* and *JAK2* genes, and mutations in RAS pathway genes occur less frequently (145) (146), (147), (148).

1.2.2 Hallmarks of JMML

A. GM-CSF signalling pathway hyperactivation: A very well recognised disease feature is hyperactivation of the GM-CSF signalling pathway. This manifests as GM-CSF-independent growth of JMML cells *in vitro* (149), (150). GM-CSF hypersensitivity is thought to be primarily due to the continuous activation of the RAS pathway driven by RAS- pathway activating mutations(151). Recently, mTOR and STAT5 pathway upregulation have also been identified as an additional mechanism of GM-CSF hyperactivation (152). The hyperactivation of the m-TOR pathway is maintained by activation of the p110δ subunit of PI3K, which further promotes the effects of the continuous activation of the RAS pathway (153), (154), (155) (Figure 1.2.1).

While GM–CSF autonomy and hypersensitivity in JMML is currently believed to be predominantly intrinsic to the malignant cells, the role of the bone marrow (BM) microenviroment is of great interest and remains to be elucidated. The only published study on the effect of JMML-associated *PTPN11* mutations on the BM microenviroment indicates that the microenviroment may promote the development and progression of JMML and GM-CSF hypersensitivity also through the excessive production of chemokine CCL3 and other pro-inflammatory cytokines produced by monocytes (156). This study also demonstrated that CCL3 receptor antagonists, such as BX471 and Maraviroc, might be useful clinically in controlling disease progression, especially in Noonan syndrome-associated MPD/JMML. This indicates that study of the microenviroment effect could be of particular interest in JMML, as 30-40% carry germline mutations that may have an impact not only on the haematopoietic cells but almost certainly to the BM microenviroment.

JMML GM-CSF hypersensitivity and autonomy was previously included in the main diagnostic criteria for JMML and was part of the investigations required in order to make a diagnosis. However, since other secondary causes of monocytosis, such as viral infections, can also display this property and the assay cannot be easily standardised among labs, this feature has been removed from the main diagnostic criteria for the disease, although it may still be useful in cases where the diagnosis is unclear (139) .

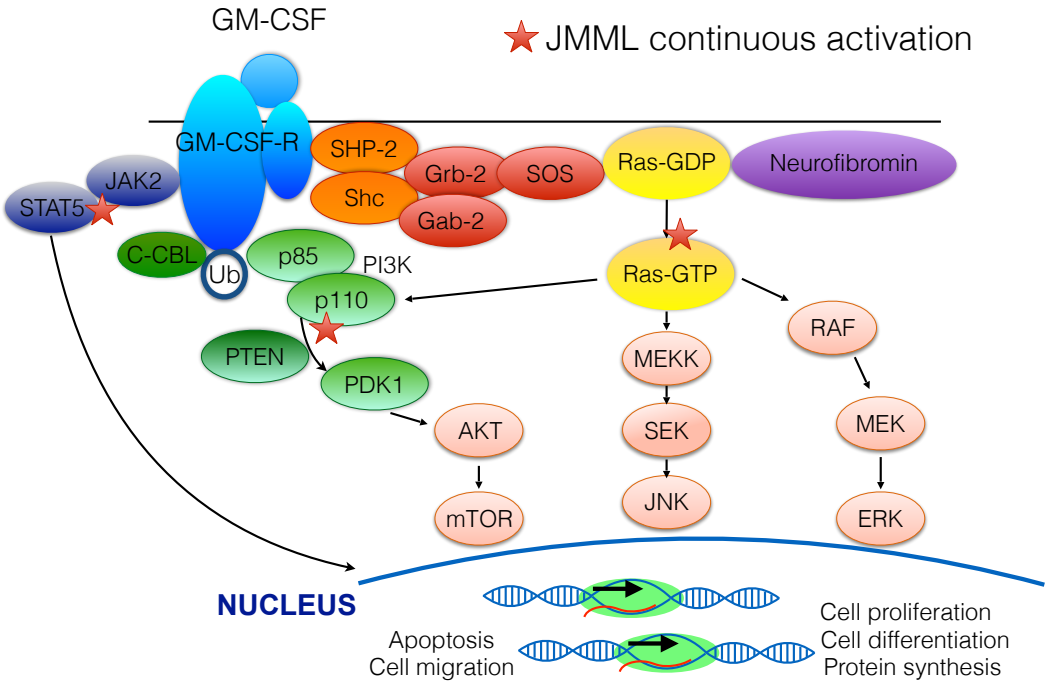


Figure 1.2.1: Molecular pathways implicated in the pathogenesis of JMML. Red star indicates abnormal continuous activation of the implicated pathway.

B. High HbF: High expression of HbF has been noted in JMML for several years (140),(157). This is a disease distinct feature, linked with adverse prognosis. The underlying mechanisms of the HbF dysregulation are currently not fully understood. Very early studies done by Weatherall and others demonstrated that JMML patients had a fetal/neonatal-like pattern of

erythropoiesis that included increased HbF and reduced I antigen expression and HbA2 production in JMML patient samples (158),(159),(160). These studies hypothesised that the fetal signature indicates that the disease must originate in a very primitive haematopoietic stem cell that has arrested in the fetal state, causing adult type Hb to gradually decrease as the disease evolves (161), (162), (163). Another potential mechanism of the raised HbF levels could be the reactivation of the γ-globin gene in response to the anaemia, similar to the one observed in IBMFS (164). In normal fetal development expression of the β-like globin genes switches from γ-globin in the fetus to β-globin transcription in post-natal cells (165). Two of the major transcription factors that regulate globin switching are *KLF1* and *BCL11A*. *KLF1* activates transcription of β-globin and of *BCL11A* which in turn represses γ-globin (166). Investigation of the methylation of these genes was therefore investigated by Fluhr et al (167) as a plausible mechanism to explain the high HbF expression in JMML. This study showed hypermethylation of the ~0.5kb *KLF1* enhancer element in JMML GlyA+ erythroid cells from patients with high HbF while the same region was not methylated in normal controls (cord blood) GlyA+ erythroid cells. This was accompanied by reduced *KLF1* expression in these cells; however *BCL11A* expression was not altered and there was no correlation between the *KLF1* enhancer hypermethylation and the levels of γ-globin expression in patient samples(167). Thus although altered methylation of regulatory regions of *KLF1* may contribute to reactivation of HbF expression, the full mechanism remains to be fully explained. Further studies on the underlying mechanisms of the aberrant HbF expression are required, as this may allow us to unravel the mechanisms that drive very aggressive

JMML cases, given that HbF cases have been shown to be a marker of poor prognosis(157), (168).

1.2.3: JMML diagnostic algorithm

A. Overview: The main diagnostic criteria for JMML have been defined and described by Locatelli and Niemeyer(139). Nevertheless, JMML can be a diagnostic and treatment challenge even in very highly specialist centres due to the heterogeneous clinical features of the disease. Initial suspicion of JMML is usually raised from the peripheral full blood count that typically reveals monocytosis, anaemia with or without thrombocytopenia. Peripheral blood films typically show abnormal monocytes, granulocyte left shift, granulocyte dysplasia, circulating blast cells and sometimes nucleated red cells. The diagnostic criteria for JMML are shown in Table 1 and a suggested diagnostic flow diagram is shown in Figure 1.2.3. Since a wide variety of conditions can cause monocytosis, including chronic infections, chronic inflammatory conditions, inborn errors of metabolism, immunodeficiency, administration of cytokines, leukaemia, pre-leukaemic conditions and myeloproliferative neoplasms, investigation of suspected JMML cases includes exclusion of these secondary causes, eg by screening for CMV, HSV and EBV (169), (170) ,(171). The disease-specific hallmarks, in particular raised HbF, can help in establishing the diagnosis.

The BM aspirate and trephine biopsy typically show hypercellularity with myeloid expansion, left shift, dysplasia of the myeloid series and increased monocytes, monocyte precursors and blast cells (Figure 1.2.2). Erythropoiesis may be decreased and is often dysplasic. Lymphocytes and megakaryocytes are often reduced.

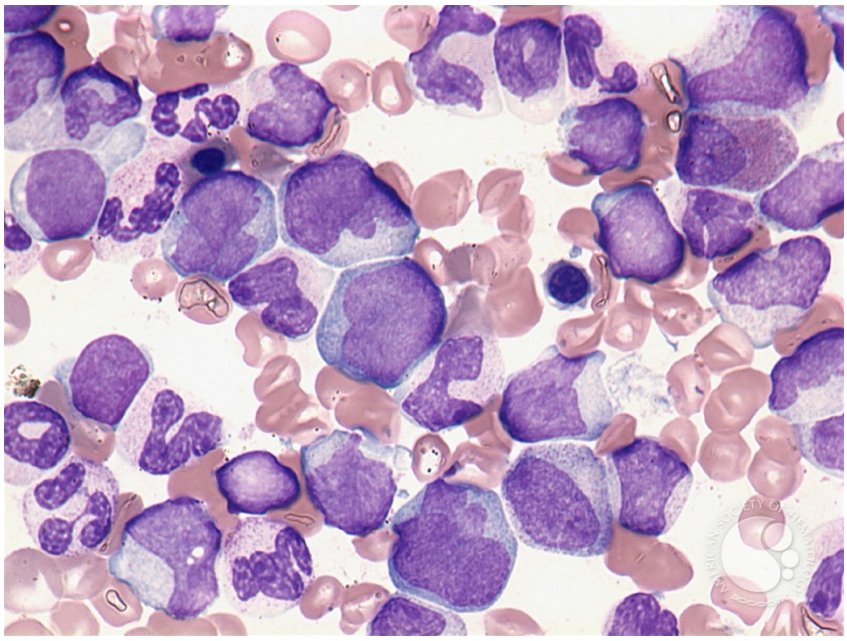


Figure 1.2.2 Maskalk P, Blood Image Bank 2004: Representative example of a bone Marrow aspirate of a patient with JMML. Showing expansion of the myeloid series with left shift and monocytes.

***B. Cytogenetic abnormalities*:** Cytogenetic analysis should be performed in all cases where JMML is suspected, as 36% of cases will have abnormal cytogenetics(140). The commonest cytogenetic abnormality is monosomy 7 (140) which is observed in 26% of JMML cases. Other cytogenetic abnormalities include trisomy 8, trisomy 13 and trisomy 21. Additional copy number variants have been reported in 11q23.3, 17q11.2, 5q31.3 and 12q24.13, at the *CBL*, *NF1*, *PTPN11* and *ARHGAP26* loci respectively (172),(173). In addition, a recent study using RNA sequencing identified some novel fusion genes: ALK-ROS1, DCTN1-ALK, RANBP2-ALK and TBL1XR1–ROS1) in 18% of cases that lacked canonical RAS pathway mutations (174). Similarly to other myeloid malignancies, additional cytogenetic abnormalities can occasionally be detected during blastic transformation although their role in disease progression remains to be confirmed(175).

C. Genomic analysis to identify disease-associated pathogenic mutations: Mutational analysis for *RAS* pathway mutations is increasingly used, not only to establish the diagnosis but also to assist in patients risk stratification and management, as this is depended on the type of *RAS* pathway mutation(139). The availability of clinical grade diagnostic mutational analysis panels is still very limited and designing and applying a reliable panel was one of the aims of my project, as discussed in detail in chapter 3 .

D. Flow cytometry: Although flow cytometry is not a routine diagnostic test, it is sometimes useful in cases which the diagnosis is unclear. Previous reports have demonstrated that JMML blasts cells have an aberrant expression of CD33+, CD34+, CD14+ and $CD38^{low}$ (176), (177). CD141+ a dendritic cell marker has also been reported in JMML cases, but the significance of this needs to be further validated in larger series (178).

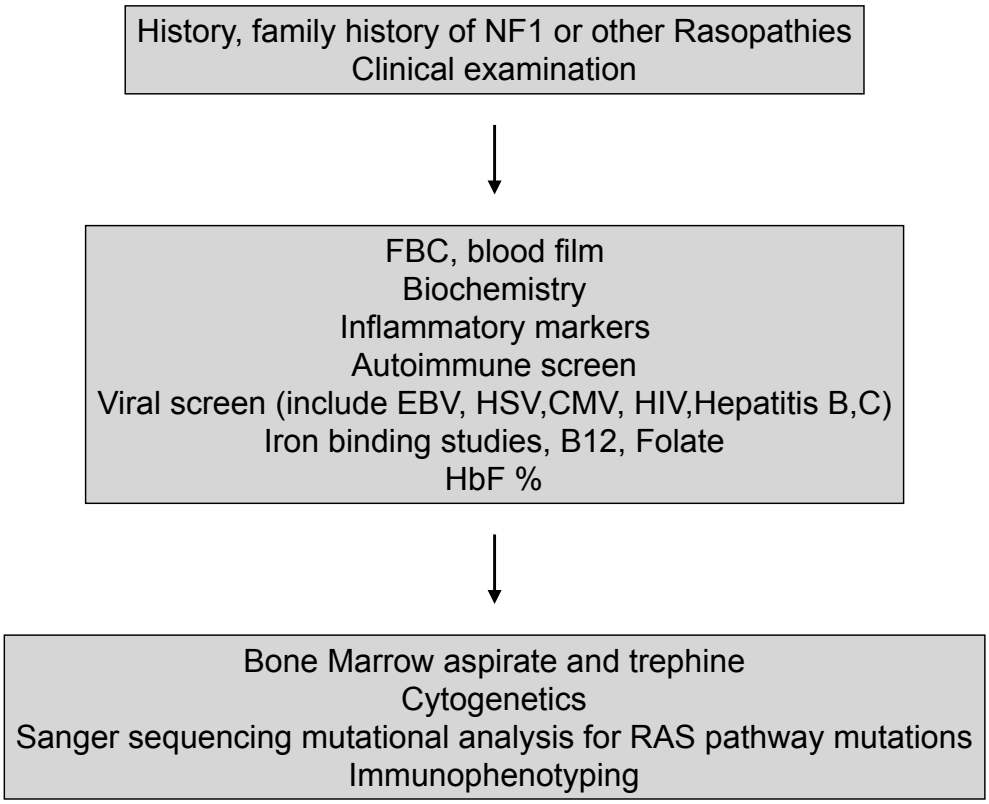


Figure 1.2.3: Diagram showing a typical JMML diagnostic algorithm (modified from the Great Ormond Street Hospital diagnostic algorithm)

1.2.4: The Molecular Landscape of JMML

Due to recent technological advances in targeted next generation sequencing (NGS) molecular abnormalities can now be detected in >90% of JMML patients (179), (141), (140). *RAS* pathway mutations are the commonest mutations(129). The RAS pathway regulates cell growth, cell cycle, differentiation, metabolism and senescence and *RAS*-pathway mutations are detected in ~20% of human tumours (180). Normally the pathway is very tightly regulated by guanosine exchange factors (GEFs) and GTPase activating proteins (GAPs) (181); and activated by binding to receptors on the cell surface, such as the GM-CSF receptor. Ligand binding leads to

autophosphorylation of tyrosine kinase sites of the GM-CSF receptor and the creation of docking sites for adaptor molecules (such as GRB2) which then recruit GEFs leading to activation of RAS and downstream MEK-ERK and mTOR pathways (Figure 1.2.4). In JMML cases the underlying molecular defect leads to defective regulation and continued activation of the RAS pathway. The main genes where mutations are found in JMML are *PTPN11*, *N-* and *K-RAS*, *NF1* and *CBL* and their position in the RAS signalling pathway is shown in Figure 1.2.4.

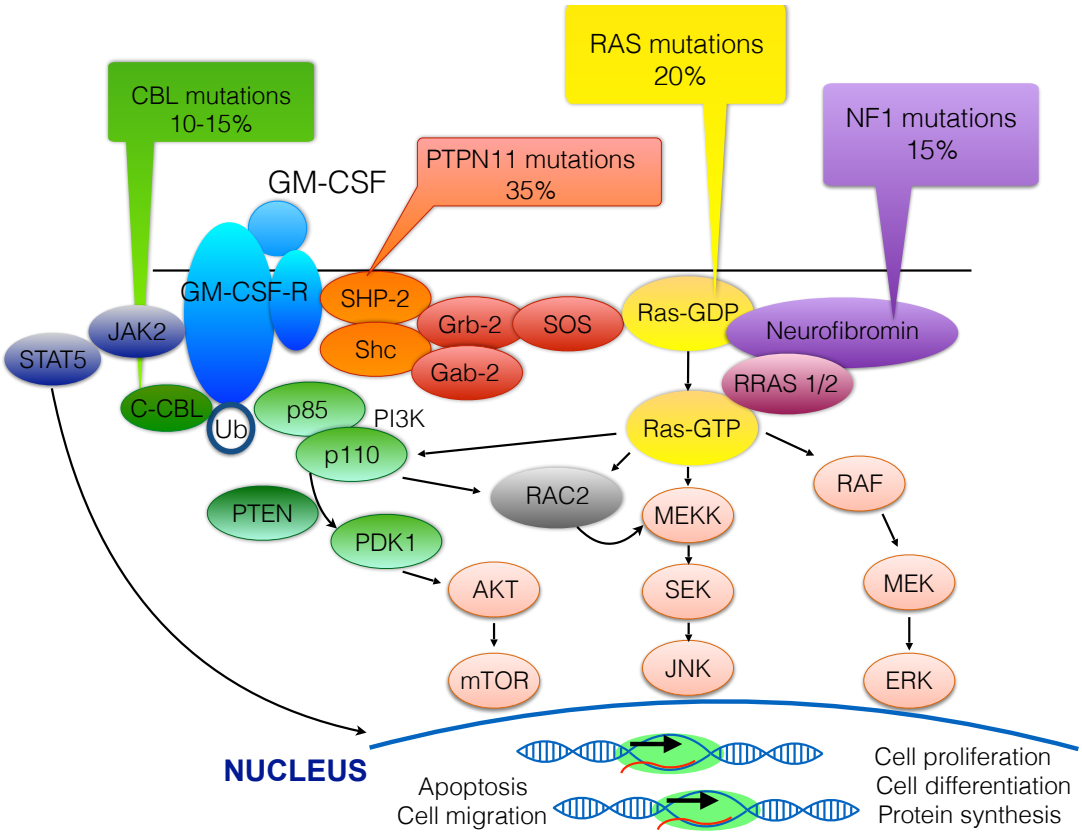


Figure 1.2.4: RAS pathway and annotation of JMML mutations frequencies
The figure outlines the main pathways that are involved in the pathogenesis of JMML and the genes in which mutations have so far been described in patients with JMML.

A. PTPN11: This gene encodes Src-homology tyrosine phosphatase 2 (SHP-2), a protein tyrosine phosphatase which activates downstream pathways that regulate proliferation and differentiation of cells, including haematopoietic cells (182), (183). Mutations in *PTPN11* are often implicated in myeloid malignancies (184) (185) and constitute the most frequent mutation found in JMML (35% of cases)(186), (187), (188) ,(189). *PTPN11* mutations in JMML are most commonly seen in exons 3,8, and 13 and interfere with the interface between the N-terminal Src homology 2 (N-SH2) and catalytic (protein tyrosine phosphatase (PTP) domain). This leads to a gain-of-function effect by impeding the switch between the active and inactive conformation of the protein, favouring continuous activation of SHP-2 and uncontrolled activation of the downstream RAS pathway components (Figure 1.2.4). Patients with somatic *PTPN11* mutations also constitute the group of patients with the highest risk mutations for relapse (139) (184), although germline *PTPN11* mutations have a much more favourable prognosis as outlined in section 1.2.6. Recent studies have shown that depletion of *GAB2*, in transgenic mice, can diminish the effects caused by *PTPN11* JMML mutations in a RAS independent manner, but to date non-RAS JMML driver mutations, in *GAB2* or other mTOR pathway genes like *PTEN* have not been identified (190).

B. K-* and *NRAS* genes: KRAS* mutations are seen in 20%-25% of JMML cases(191). The majority of *K-RAS* mutations are point mutations that disturb GTP-ase activity and maintain RAS in a constant state of activation. *KRAS* mutations are most commonly seen in codons 12, 13 and 61(192). Interestingly, just over 50% of JMML patients with *KRAS* mutations also have

monosomy 7 (193). *N-Ras* mutations are seen in 10% of JMML cases. However unlike most of the other RAS pathway mutations, patients with *NRAS* variants tend to have a good prognosis and a milder disease course (194), (195), (196) (197). *NRAS* patients are only eligible for bone marrow transplantation (BMT) if they acquire additional genetic abnormalities, such as loss of the residual wild type allele (139) (198).

C. NF1: This is a very well known tumour suppressor gene that regulates cell proliferation cell division and modulates microtubule activity. *NF1* mutations can be germline mutations inherited in an autosomal dominant fashion or *de novo* events that occur without any family history of neurofibromatosis type 1 (NF1). Patients with *NF1* mutations have an increased frequency of malignancies; in children the most common malignancies observed are rhabdomyosarcoma, optic pathway glioma, acute lymphoblastic leukaemia, non-Hodgkin lymphoma and JMML(199). It is estimated that the incidence of JMML in patients with NF1 is 200 fold higher than in normal individuals (200). *NF1* mutations are the third commonest cause of JMML (10-15% of cases). Mutations in *NF1* lead to reduced neurofibromin-specific GTPase activity, raised levels of RAS-GTP and anomalous continuous activation of the downstream MEK-ERK pathway (201). The mechanism of JMML leukaemogenesis for the majority of patients with germline NF1 mutations is thought to be through the somatic loss of heterozygosity of the second *NF1* allele through segmental uniparental isodisomy of parts of chromosome 17 (202),(203). One third of patients are found to have compound–heterozygous *NF1* mutations, with one of the 2 variants being somatically

acquired(204).JMML patients with *NF1* mutations have a higher rate of secondary mutations, most likely reflecting the tumour suppressor role of *NF1*(193).

D. CBL: This is an E3 ubiquitin ligase that marks activated receptor and non-receptor tyrosine kinases and other proteins for degradation, and retains adaptor functions(205). *CBL* mutations are found in 10 -15 % of JMML cases (179), (206), (207). The mutations may be loss-of-function mutations through loss of tumour suppressor capacity and protein degradation activity or gain-of-function mutations causing an activated receptor that can no longer be ubiquitinated. *CBL* mutations are homozygous in the majority of cases, but heterozygous variants consisting of splice site mutations and deletions have also been identified. This suggests that *CBL* mutations can also occur as a secondary event in JMML(208), (209). Typical mutations are located through the RING finger domain, affecting codon 371. Similarly to the *NRAS* mutations, *CBL* mutations appear to carry a more favourable prognosis and patients do not require treatment by bone marrow transplantation (BMT) unless they carry a second RAS pathway/other mutation or abnormal cytogenetics. Patients with germline CBL mutations are more likely to spontaneously resolve, and typically present with transient monocytosis, in the majority of times confined to the neonatal period (179).

E. RRAS2 : More recent studies using whole exome sequencing (WES) have identified JMML driver mutations in *RRAS2* in patients with no other previously described RAS pathway mutations; *RRAS* mutations have also been detected

in a JMML patient sample at diagnosis along with a *NRAS* mutation. Interestingly, Ceremsak et al (2016) recently described such a case in which disease progression was propagated by the *RRAS* mutation-bearing clone alone (the *NRAS* mutation was lost)(210), (172). *RRAS2* is a member of the RAS pathway and encodes teratocarcinoma oncogene 21 (TC21) (172),(193) previously described in solid tumours. *RRAS2* appears to act via RAF-independent signalling most likely through *PI3K-AKT*,*RAC* and *JNK* signalling, but further studies are required to investigate this in JMML patients (211), (212).

F. *Secondary mutations*: Traditionally JMML, unlike AML, was thought to be caused by single mutations in RAS pathway genes that are mutually exclusive; however more recent WES studies have demonstrated that secondary mutations can occur in JMML, including second RAS pathway mutations, and these are associated with a poorer prognosis (213), (214). However, unlike AML, patients with more than two molecular defects have not been detected (172), (193). Secondary mutations occur in 100% of NF1 cases and 63% of sporadic JMML cases (172),(193). In JMML, the secondary events can be similar to the primary mutations involved in adult MDS/MPN, such as CMML (215).

Mutations in epigenetic modifiers: Secondary mutations in JMML are most frequently seen in PRC2 polycomb complex core subunits. The commonest PCR2 complex mutations are *ASXL1* mutations, which are found in 8% of cases and *EZH2* mutations which are found in 4% of cases. Copy number variants (CNVs) that result in haploinsufficiency of other PRC2 core subunits,

such as *SUZ12,* and mutations in other PRC2-associated factors, such as *AEBP2, CDYL and JARID2* have also been described(216), (193). *EZH2* mutations are almost always associated with monosomy 7 leading to loss of the normal *EZH2* allele as this is a chromosome 7 gene. It is thought that decreased PRC2 activity, through *ASXL1* mutations or monosomy 7, leads to a decrease in H3K27 trimethylation and increase in acetylation for specific PRC2 oncogenic targets. *DNMT3A* mutations have also been described but are not as frequent as in other haematological malignancies (172).

Spliceosome mutations: Similar to other haematological malignancies, spliceosome machinery mutations, such as in *SRSF2*, *ZRSR2,* have also been described, although based on the variant allele frequency, these are never driver mutations (217). Interestingly it has been shown that in the adult MDS, a very frequent *SRSF2* mutation of codon 95 creates aberrant splicing of *EZH2* , that leads to *EZH2* 'poison exon' resulting to EZH2 mRNA degradation and EZH2 loss even in cases that do not have *EZH2* mutations or monosomy 7 although this mechanism has not yet been reported in JMML (218).

Second RAS- pathway mutations: Recent studies have documented second *RAS* pathway genes in 11-17% of sporadic JMML and 25% of NF1-associated JMML(193),(172). In addition to the main *RAS* pathway genes described above, RAS pathway mutations have been documented in *PDE8A* (a gene that protects RAF1 from inhibitory phosphorylation by protein kinase A) (193); *RAC2* (gain of function mutation); *PLXNB2* which catalyse RRAS inactivation via their GTP-ase activating protein (GAP) domain (219); and *ABI1* mutation a gene that belongs to the multi-molecular complex responsible for *SOS*-mediated *RAS* activation (220) (221). Patients with a second *RAS* pathway

mutation are thought to have the worse prognosis and survival ~37% compared to the other disease subgroups (Figure 1.2.5) (193).

Secondary mutations affecting other pathways: Secondary mutations have been reported in genes encoding signalling molecules such as *SETBP1, JAK3* (213), (222) and *SH2B3*; transcription factors mutations such as *GATA2* and *RUNX1* (172); and in genes normally associated with bone marrow failure, such as somatically acquired *WASP* mutations (223) or a germline *WAS* mutation reported in a patient with an existing *NRAS* mutation (172).

Secondary mutations in JMML are associated with overall worse outcome. The presence of two RAS-pathway mutations are patients identified to have the worse prognosis (Figure 1.2.5) (213), (222), (193). Large studies are required in order to risk stratify the rest of secondary events observed in JMML.

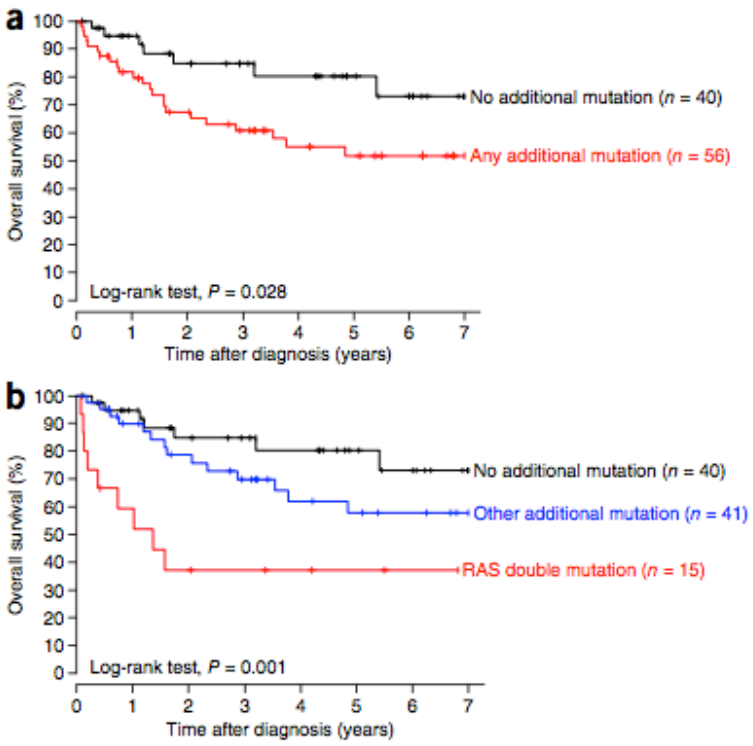


Figure 1.2.5: Caye et al, Nature Genetics, 2015: Kaplan Meier representations of overall survival of 96 JMML patients according to the presence and type of additional mutations (193)

G. Mechanisms of mutation acquisition in JMML: Similarly to other myeloid malignancies, the mutations see in JMML include a mixture of simple point missense, nonsense mutations, deletions and frameshift mutations. However, uniparental disomy is distinct mechanism that leads to loss of heterozygosity (LOH) and, as in other myeloid malignancies (224), (225), it is recognised as the mechanism of carcinogenesis particularly in syndromic JMML cases. Uniparental disomy is the mechanism by which the second normal allele is lost in the majority of patients with constitutional *NF1* and *CBL* mutations (226), (206), (179), (225), (227) and in patients with *KRAS* and *NRAS* mutations at the time of disease progression (228), (198), (229).

H. Timing of mutation acquisition: There are now many reports showing that in a proportion of JMML cases, pathogenic mutations can be detected in neonatal cord blood samples and Guthrie blood spots, indicating that in these cases the mutations are acquired in utero. This will typically lead to a clinical presentation of the disease within the first 2 years of life (Table 1.2.2) (172), (230), (231), (232), (233), (234). Only one of the recent studies (172), has reviewed a large cohort of samples with paired diagnostic BM/blood and cord blood samples. In this study JMML mutations were detected in all of the samples screened, but further studies that exclude patients with germline mutations are required to further address this point.

RAS-pathway mutations are also detected in other myeloid malignancies, for example AML. However, in AML these mutations are often secondary mutations in downstream progenitors and are always seen in association with

other molecular defects (235). These data suggest that the timing of mutation acquisition may determine the disease phenotype and that JMML develops when RAS pathway mutations are acquired before, or soon after, birth.

Table 2: Summary of reports of JMML mutations in neonates, cord blood samples and Guthrie blood spots (172), (230), (231), (232), (233), (234)

Study	Number of cases	Mutation (number of cases)	Development of JMML	Tissue mutation was detected	Syndromic features	Age of presentation
Stieglitz E et al, 2015	5	2 *NRAS* 1 *PTPN11* 1 *NF1* 1 *CBL*	yes	Cord blood	n/a	1-120 months of age
Kothari A et al, 2015	1	*KRAS* (somatic mutation)	yes	Newborn Peripheral blood-somatic	no	newborn
Lee M et al, 2013	1	No report of mutation	yes – died 58 days of age	n/a	yes	Preterm (35 weeks)
Yoshida, R et al , 2004	1	*PTPN11*	yes – but transient disease with spondaneous recovery	Peripheral blood – no germline tested Parents did not carry the mutation	yes (NS)	Preterm (30 weeks)
Cheong JL et al, 2007	1	*PTPN11*	yes – patient died	n/a	yes NS	Preterm (34 weeks)
Matsuda K et al, 2009	7	3 *PTPN11* 3 *NRAS* 1 *KRAS*	yes	Guthrie blood spots	no	Presented with the first 2-19 months of life

1.2.5 Gene expression studies in JMML

Global gene expression studies in JMML patient samples are currently limited. A large European collaborative study used microarray to investigate whether there were any consistent abnormalities of gene expression in JMML cells. They reported the presence of a 'fetal' gene signature characterised by

expression of the *LIN28B* gene which is not normally expressed in haematopoietic cells after birth (109). This 'signature' correlated with high HbF in the same samples (14). MicroRNA studies have also suggested that loss of miR-150-5p may contribute to GM-CSF hypersensitivity through the upregulation of STAT5b pathway (236). Upregulation of mi-R223 and miR-15 has also been reported in human induced pluripotent stem cells (hiPSCs) derived from *PTPN11* mutated JMML cells(237).

Recent genome-wide methylation studies have identified three distinct methylation patterns in JMML patients: a 'hypermethylated' group, that demonstrate an AML-like hypermethylation pattern and is most frequently seen in patients with *PTPN11* mutations; an 'intermediate' group linked to *KRAS* mutations; and a 'low methylation' group associated with *NRAS* mutations(238). Interestingly, the hypermethylated group includes *PTPN11* mutations occurring at the N-terminal SH2 domain which are linked with more aggressive disease and mutations located on the protein tyrosine phosphatase domain (PTP) are located in the intermediated group. The E76K *PTPN11* mutation, which is one of the most activating JMML mutations, has been reported to have a distinct methylation pattern(239). Specific loci with particularly aberrant methylation patterns include *AKAP12, TET1-3, HDAC9* and *DOT1l* (238), (239), (240). Another study has identified upregulation of *DNMT1, DNMT3A, DNMT3B* (239) and silencing of *CDKN2B, RARB, CALCA*, *BMP4* CpG islands through hypermethylation(241). In addition, a further study has reported hypermethylation of the *RASA4* gene, that is, interestingly, located on chromosome 7, in 51% of JMML patients (242). Notably, in patients with monosomy 7 the remaining allele is found to be hypermethylated (242).

1.2.6 Syndromic JMML

A subset of JMML is caused by a family of congenital conditions known as RASopathies. These are common genetic conditions (seen in 1/1000 human births) (243) that share overlapping clinical features such as short stature, cranio-facial dysmorphic features, cardiac malformations, neurodevelopmental delay, musculoskeletal and cutaneous distinct features. The RASopathies include Noonan syndrome, Leopard Syndrome, Noonan syndrome with multiple lentigines, Neurofibromatosis type 1, Costello Syndrome, CBL-syndrome, Cardio-facio-cutaneous syndrome and Capillary malformation-arteriovenous malformation syndrome. As mentioned above, these conditions carry a high risk of malignancy (Table 1.2.3). JMML has been described in the majority of these conditions caused by germline mutations in one of the RAS pathway genes. Germline JMML cases should be distinguished early from the somatically acquired JMML cases as the natural history and management is different and the diagnosis of a RASopathy has important implications for affected families. In patients with significant disease manifestations, the diagnosis of a RASopathy will normally precede the JMML diagnosis; however in patients with milder clinical phenotypes or mosaicism, JMML is often the presenting feature that leads to a diagnosis of the germline disease. Interestingly, JMML has not yet been reported in patients with *BRAF, SOS, SHOC2*, *MAP2K1* or *MAP2K2* mutations(181).

Table 1.2.3: Human RASopathies and their causative genes: (181), (244)

RASopathies	Genes with detected mutations	JMML cases	Other malignancies
Noonan syndrome (NS)	*PTPN11,SOS1, RAF1,KRAS, NRAS,SHOC2, CBL,BRAF, A2ML1,SOS2, LZTR1,RASA2, RRAS, RITI*	Transient MPD, JMML	AML, B-ALL, neuroblastoma, rhabdomyosarcoma,Sertoli cell tumour, astrocytoma, lung,ovarian, colon cancer
Leopard syndrome (Multiple lentigines syndrome) (LS)	*SPRED1*	Not reported	Not reported
Noonan syndrome with multiple lentigines (NS-MLS)	*PTPN11, RAF1, BRAF*	Transient MPD, JMML	AML, B-ALL, MDS,CML, neuroblastoma,rhabdomyosarcoma, Sertoli cell tumour, astrocytoma, lung,ovarian, colon cancer
Neurofibromatosis type 1 (NF1)	*NF1*	JMML	Astrocytoma, peripheral nerve-sheath tumour, rhabdomyosarcoma, neuroblastoma breast cancer, GI stromal tumour , colon cancer, gliomas, lung cancer, ovarian cancer, and many others
CBL-syndrome (CBL-S)	*CBL*	JMML	MDS, MPN, AML, CML, lung cancer
Costello syndrome (CS)	*HRAS*	Not reported	Rhabdomyosarcoma, neuroblastoma, bladder carcinoma, papillomata
Cardio-facio-cutaneous syndrome (CFCS)	*BRAF, MAP2K1, MAP2K2, KRAS*	Not reported	Langerhans cell histiocytosis, malignant melanoma, colon cancer, thyroid cancer , Langerhans cell histiocytosis, hepatoblastoma
Capillary malformation-arteriovenous malformation syndrome (CM-AVMS)	*RASA1*	Not reported	Not reported

MPD: Myeloproliferative disorder, AML: acute myeloid leukaemia, ALL: acute lymphoblastic leukaemia, MPN: myeloproliferative neoplasms, CML: chronic myeloid leukaemia, MDS: Myelodysplasia.

A. Noonan Syndrome

Noonan syndrome (NS) is a relatively common, autosomal dominant syndrome (1/1000-2500 births), that is characterised by distinctive facial anomalies, short stature, heart defects, cryptorchidism, developmental delay and haematological manifestations (245). RAS pathway mutations occur mainly in one of 6 different genes (*PTPN11, SOS1, SOS2, RAF1, BRAF* and *SHOC2*) with heterozygous *PTPN11* mutations seen in more than half of the cases(246),(247). The commonest haematological manifestations include a transient JMML-like myeloproliferative disease (MPD) seen in 5.6% of cases. The majority of cases present very early in the neonatal period. In most cases, spontaneous regression of the monocytosis occurs without any treatment (248) but 10% of cases progress to a more aggressive clinical course that resembles JMML(249),(250),(181). No cytogenetic abnormalities have been identified in NS/MPD cases (251). There are no specific investigations to identify the patients with poor prognosis and GM–CSF hypersensitivity is observed in all cases.

Despite the fact that germline mutations *PTPN11* are potent enough to cause multi-systemic clinical manifestations with often significant clinical consequences, these do not appear to be activating enough to give rise to the aggressive JMML phenotype seen in somatically acquired *PTPN11*-associated JMML(189),(237). Studies have documented that although the NS mutations are often found in same gene as somatically acquired JMML, the position of the mutations is different from the somatic cases. For example, mutations

causing amino acid substitutions in codon 76 (the most commonly mutated PTPN11 codon in JMML are rare in NS (189); whereas mutations in codon 73 (Thr73Ile) are commonly seen in patients with NS and have not been identified in JMML(189). There is only one *PTPN11* mutation (Asp61Gly) that is so far documented in both NS and somatically acquired JMML cases. Patients with this mutation might develop MPD but have a milder clinical course than JMML(251). In addition, the same study suggested that cases with NS/MPD appear to have molecular defects in mutations (Asp61Gly, Ser502Ala, Gly506Pro) associated with decreased fertility and reduced fetal survival. JMML-associated mutations are thought to have a stronger gain-of-function effect that might be embryonically lethal in the case of NS, whereas mutations commonly seen in NS are less activating. There is one report of 'typical' PTPN11 mutations found in fetuses with NS and in all of these cases the patients had a very severe phenotype that resulted in perinatal death (252). Murine studies have shown that homozygous mutations in these positions are embryonic lethal and in the heterozygous state are associated with decreased viability.

NS with multiple lentigines is an autosomal dominant condition caused by germline mutations in *PTPN11*, *RAF1* or *BRAF*. Most individuals have features of NS with an increased risk of hypertrophic cardiomyopathy and multiple lentigines. The *PTPN11* cases have a similar risk to other patients with NS in developing JMML like-MPD or JMML.

B. Neurofibromatosis type 1 (NF1)

Familial NF1 cases have an estimated prevalence of 1/2000- 1/5000. Clinical features of NF1 include café au lait spots, cutaneous neurofibromas or plexiform neurofibromas, axillary or groin freckling, optic pathway glioma, Lisch nodules and bony dysplasias; clinical diagnosis of NF1 can be made on the basis of these clinical findings and/or on the family history. *NF1* mutations have variable expression and expression and clinical features might not present in all patients. Patients with NF1 have a higher risk of developing malignancies, including JMML. Unlike other RASopathies, patients with germline *NF1* mutations have aggressive disease with a higher risk of acquiring additional genetic events and therefore should be treated promptly by BMT.

C. CBL syndrome

Patients with germline *CBL* mutations can present with features similar to other RASopathies. These include developmental delay, impaired growth, dysmorphic facial features, short stature and cryptorchidism (253) (181). In certain cases, it is reported that patients with CBL mutations also display high rates of vascular lesions and autoimmune phenomena, which can become particularly prominent after BMT(254),(139). In the majority of cases, patients with germline *CBL* mutations display a mild JMML–like MPD, similar to the MPD seen in NS(253). Patients with CBL mutations generally have a favourable prognosis and BMT is not generally recommended (139), (227) unless any of the adverse associated features outlined in Table 4 are present.

1.2.7: Management of JMML

BMT is the only curative treatment for JMML. Chemotherapy and hypomethylating agents are useful only for bridging pre- or post-BMT (255). In the interim period, from to diagnosis to transplant, anti-metabolites such as 6-mercaptopurine or cytosine arabinoside, hypomethylating agents, such as azacitidine or splenectomy may be useful (168), (125). AML chemotherapy makes no significant difference to event-free survival, relapse or transplant-related mortality (168),(256) and should only be used for life-threatening symptoms. Splenectomy does not appear to have a significant impact on post-transplant survival or relapse and is now rarely used(168). The most common approach used is 6- mercaptopurine or azacitidine in an effort to control the symptoms of the disease prior to BMT. Currently there is no evidence to favour one approach versus another in symptom control or reduction of relapse following BMT.

Conditioning regimens for JMML BMT have included both myeloablative and sub-myeloabalative therapy. The European Working Group of Myelodysplastic Syndromes in Childhood (EWOG/MDS) reported transplant-related mortality (TRM) of 10-22% following myeloablative chemotherapy including busulfan, cyclophosphamide and melphalan (139) while Japanese investigators typically add fludarabine in place of cyclophosphamide (257),(258). However, the US Children's Oncology Group recently showed that a busulfan/cyclophosphamide/melphalan-based regimen was superior to busulfan/fludaribine (255). The reported overall survival for JMML post-BMT varies from ~ 25-72% depending upon the molecular lesions and the presence of adverse prognostic factors (168),(256),(259), (260),(261), (258). The major

cause of treatment failure remains the high incidence of disease relapse (168). The donor source (family, unrelated, cord blood) does not appear to significantly alter survival (258) however, children transplanted from a donor with 0/1 HLA disparity have a better outcome than those transplanted from donors with greater disparities (256).

The risk of GVHD in JMML is lower than other haematological malignancies and does not represent a major cause for transplant-related mortality (259) perhaps because the patients are nearly all very young and serotherapy with antithymocyte globulin (ATG) is widely employed (139). The EUROCORD-CIBMTR study reported that grade 2-3 GVHD is associated with better survival, possibly because the graft-versus–leukaemia effect may help to prevent relapse (262), (263), (256) . The EWOG-MDS suggested that all JMML cases with the exception of *KRAS* JMML cases, should receive a low-intensity GVHD prophylaxis and immunosuppressive treatment is generally withdrawn 2-3 months after transplant unless there is active GVHD(139). The development of sensitive disease biomarkers that can indicate treatment failure and low level of disease, would be very valuable in order to predict when to institute treatment to enhance the graft-versus-leukaemia effect before clinical evidence of disease relapse emerges.

More recently there have been a number of trials of targeted therapy. The US Children's Oncology group tested the efficacy of a farnesyl transferase inhibitor (Tipifarnib), which blocks prenylation of RAS to the inner surface of the cell membrane, interfering with RAS activation and downstream signalling. Tipifarnib reduces the clonogenicity of JMML cells in vitro (264) and Phase I,II and III trials indicate that the drug is safe. However, despite a 51% initial

response rate, Tipifarnib has failed to significantly reduce relapse rates or to improve overall survival (265),(266). An alternative approach, MEK and mTOR inhibitors, directly inhibit the RAS–pathway hyperactivated pathways seen in JMML can reduce JMML clonogenic potential and GM-CSF hypersensitivity in vitro (267), (190), (153), with promising results in vivo in JMML mouse models and recently investigated in iPCS (induced pluripotent stem cells) cells in vitro (268), (269).Tasia et al has also demonstrated some efficacy of JAK2 inhibitors with JMML iPCS cells (269). A recent study has demonstrated that Dasatinib, a TNK2 inhibitor widely used for the treatment of chronic myeloid leukaemia (CML) can be effective in terminating RAS-pathway hyperactivation in *PTPN11* mutant patients (270). Future studies are required to assess the efficacy of these drugs in the pre- and post-bone marrow transplantation, however currently none of these efforts has lead to the identification of treatments that can substitute bone marrow transplantation. Due to the rare nature of the disease as well as the high relapse rates, the formation of trials for this disease to formally assess the efficacy of these treatments is extremely challenging.

1.2.8: Natural history and prognosis

The disease course can be very heterogeneous. Two thirds of patients have a very indolent disease course (271), whereas the rest demonstrate a very aggressive disease requiring prompt allogeneic HSC transplantation. Previous studies have demonstrated that 9.8% of patients develop blast crisis (272) and 18% of children develop AML (108). Children that present before the age of 2 years, may have a longer disease course, with stable disease or even transient improvement (273). Since the abnormal JMML cells infiltrate the lungs, spleen, kidneys, liver and, rarely, the central nervous system, death usually occurs within the first year of diagnosis (108), (157) if children do not receive any treatment.

Previously, patient risk stratification was primarily based on clinical features at presentation, due to the lack of evidenced-based laboratory methods. With the recent technological advances in mutation tracing, the disease driver mutation can be identified in ~90% of cases, which has allowed better risk stratification based on the molecular diagnosis. *PTPN11* and *NF1* genes representing the highest risk patients and mutations in *NRAS* and *CBL* genes identifying patients with favourable prognosis, that do not require any immediate treatment (195) providing they do not carry other bad prognostic clinical, cytogenetic markers or additional molecular defects (Figure 1.2.6). However, patients within the more favourable groups can also progress and develop a more aggressive disease, therefore all JMML patients need long term follow up. A summary of the typical disease features that indicate a worse prognosis are listed in Table 1.2.4 (274), (140), (157) , (275), (276), (214).

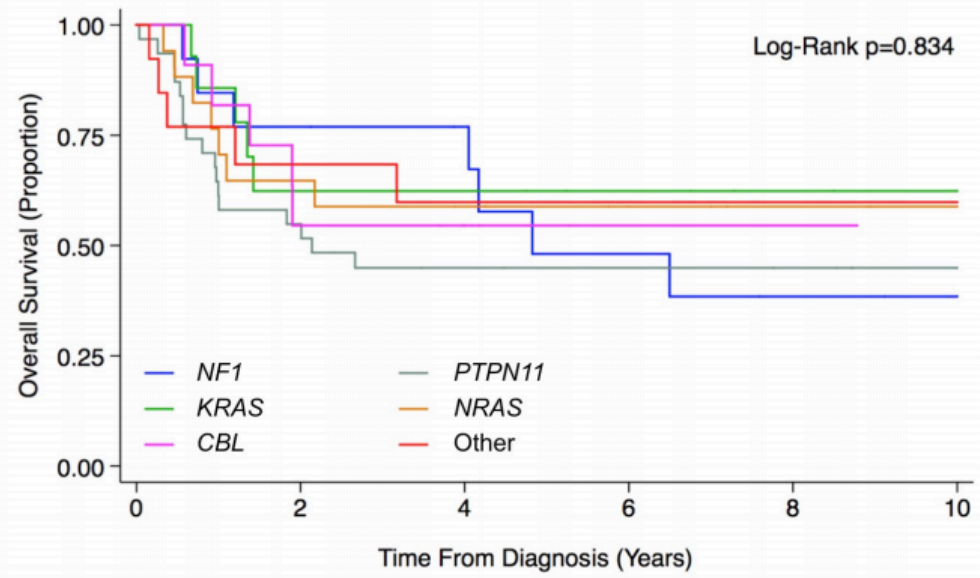


Figure 1.2.6: from Stiegltz E et al, 2015 (172): JMML overall survival based on the different types of mutations

Table 1.2.4: Disease features that predict for more aggressive disease and worse outcome

Factor	High risk
Age at presentation	> 1.4 -2 years
Platelet count	< 33 X 10^9 /L
Blast count	Equal or > 20% BM blasts
HbF	> 10-15% (corrected for age)
Cytogenetics	Monosomy 7 or other complex karyotype
Driver mutation	*PTPN11* or *NF1*
Secondary mutations	Second *RAS*-pathway mutation> Any secondary mutation
Aberrant DNA methylation pattern	Hypermethylation of *BMP4, CALCA, CDKN2B, RARB, RASA* isoform 2 locus

Criteria for treatment response: Initially the criteria set for treatment response included: white cell count, platelet count and liver and spleen size (277), (278). However recent reports have demonstrated that due to the heterogeneous disease presentation this might not be sufficient to evaluate the disease burden in all patients as not all patients will have a total WBC count above the threshold at diagnosis/thrombocytopenia or an enlarged spleen. The most recent, revised definitions of response and relapse are shown in Table 1.2.5.

Table 1.2.5: Current definitions of response and treatment relapse (279)

Complete remission (CR) is defined in the presence of neutrophil engraftment and : 1. Full donor chimerism of unsorted cells from PB or BM and 2. Disappearance of acquired cytogenetic and molecular abnormalities in patients with previously identified abnormality For patients without cytogenetic or acquired molecular abnormalities at diagnosis , who do not achieve full donor chimerism (as defined above), all of the following features of clinical remission must be achieved for definition of CR: 1. Absence of splenomegaly on exam imaging, if splenomegaly is present at diagnosis 2. Absolute white cell count < 15X 10^9 /L 3. < 5% blasts in the BM 4. < 1% myeloid/erythroid precursors including blasts in PB
Relapse is defined if one of the following criteria is fulfilled: 1. clinical JMML features and mixed chimerism > 5% 2. blasts in BM > 5% , total blasts and myeloid/ erythroid precursors in PB > 5% * 3. cytogenetic relapse: if applicable , reappearance of clonal cytogenetic abnormality 4. molecular relapse: if applicable, reappearance of acquired genetic abnormality *exclude other causes of appearance of blasts and myeloid/erythroid precursors in PB such as the regenerating phase after engraftment, severe infections and effect of granulocyte- colony-stimulating factor

1.3 AIMS

Paediatric MDS and JMML are extremely rare and challenging diseases with poor prognosis. This is due to poor disease characterization and lack of disease specific treatments. Prior to this study there was no framework that allowed the study of these conditions in a systematic manner and disease detailed characterisation. The overall aim of my project was to create the foundations of such a framework and develop evidence-based diagnostic and prognostic pathways. Furthermore, I focused in studying the disease biology and haematopoiesis in JMML in order to identify and functionally characterize the disease propagating population.

I hypothesised that correlation of detailed phenotypic information and genetic data will allow better characterization of the disease subgroups allowing evidenced based risk stratification and management. With the specific aim to contribute in identifying new disease modification treatments for JMML, I set out to investigate the haematopoieitc stem cell and progenitor compartments (HSPC) in JMML. I hypothesized due to previous reports discussed in chapter 1.2 and 4, that the JMML leukaemia stem cell (LSC) is located among these compartments and investigated in detail these populations using state of the art single cell molecular techniques and clonogenic assays.

Specific aims of this project:

1. To establish the first UK Paediatric MDS and JMML study that allows the collection of deep clinical phenotypic data and serial patient biological samples.

2. To identify the genetic basis of these conditions, using the best available diagnostic methods including targeted next generation sequencing with a customised gene panel and whole genome sequencing for the negative cases.

3. To use phenotypic:genotypic correlations in order to develop diagnostic and management algorithms for these conditions.

4. To determine the impact of the JMML mutations on haematopoiesis.

5. Identify and characterize the disease propagating population(s) in JMML.

6. To identify new disease specific therapeutic targets in JMML.

Chapter 2:METHODS

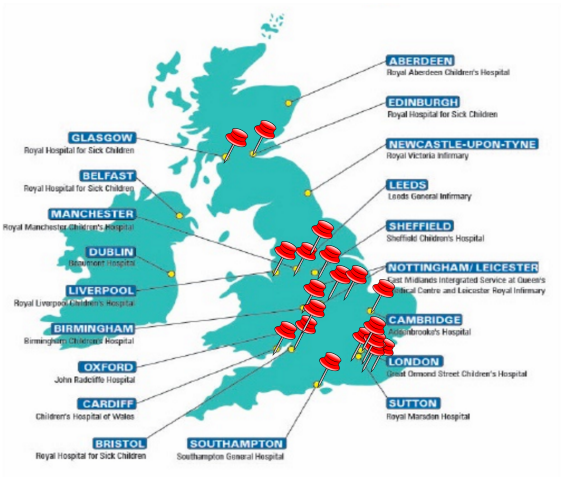
2. METHODS

2.1 THE NIHR Paediatric MDS/JMML study

2.1.1 Ethics application, procedure and approval

The protocol, patient and parent information leaflets, case report forms and consent forms were designed by myself and approved by the sponsor of the study, Great Ormond Street Hospital . The application was submitted using the Integrated Application System (IRAS) and ethical approval for the project was awarded by the NRES London-Brent committee on the 26th of May 2015 (NHS REC reference number: 15/LO/0961).

All 19 UK tertiary paediatric haematology centres were invited to participate in the study. For all centres interested, individual Research and Development (R&D) approval was obtained following local review of the study protocol and procedures. During my 3-year project the majority of UK Paediatric Haematology centres have open to recruitment for this study (Figure 2.1)



1.Great Ormond Street Hospital (GOSH)
2.University College London Hospital (UCLH)
3.John Radcliffe Oxford University Hospital
4.Imperial College Healthcare NHS Trust
5.Royal Marsden Hospital
6.St Georges University Hospital
7.Bristol Royal Hospital for Children/ Bristol University Hospitals
8.Manchester Children's Hospital
9.Leeds Children's Hospital
10.Alder Hey Liverpool Children's Hospital
11.Leicester Children's Hospital/University Hospitals of Leicester
12.Southampton Children's Hospital
13.Noak's Ark Children's Hospital for Wales
14.Glasgow Children's Hospital
15.Sheffield Children's Hospital
16.Addenbrooke's Cambridge University Hospitals
17.Birmingham Children's Hospital
18.Nottingham Children's Hospital
19.Edinburgh Children's Hospital

Figure 2.1: Map showing the Paediatric Haematology Centres and those that were participating in the NIHR Paediatric MDS/JMML study during my project

2.1.2 Study design/ procedure

This was a prospective non-randomized observational study. The core study members included the study chief investigator (CI) Dr Anupama Rao (based at Great Ormond Street Hospital), the co investigators Professor Irene Roberts, Professor Adam Mead and myself, all based at the MRC Weatherall Institute of Molecular Medicine (WIMM), Oxford. A local principal investigator (PI) was allocated for each centre and research nurses and doctors were responsible for recruitment at each centre. Following local R&D approval, meetings with the core research team were organized in order to go through the study objectives and procedures.

The eligibility criteria for the study (Table 2.1) were intentionally broader than the WHO Paediatric MDS/MPN classification in order to allow recruitment of cases where the diagnosis is unclear at presentation. This allowed the inclusion of all possible cases with sustained cytopenias and dysplasia; in order to not only study the characterization of de novo paediatric MDS, but also to allow investigation and development of better diagnostic algorithms for paediatric MDS overlap syndromes and investigate the significance of these conditions in the paediatric population, as such patients represent a large number of cases seen by paediatric haematologists that are often monitored without a conclusive diagnosis or management plan for several years.

All families where patients met the inclusion criteria were invited to participate in the study; and after informed consent samples were collected and sent to the study-coordinating centre (Figure 2.2). The results of all clinical, phenotypic and laboratory investigations were collected by the responsible clinician using a clinical record form; anonymised and sent to the study-coordinating centre. Clinical information was entered on an online secure access database and samples were processed by

myself at the WIMM. Samples collected include peripheral blood (PB) and bone marrow (BM) as well as skin biopsies, hair follicles or nails in order to allow screening of germline material. PB and BM samples were mainly used for cryopreservation of viable cells that allowed downstream functional studies and a small aliquot (0.5- 1ml) of PB and BM was kept for DNA extraction.

Table 2.1: NIHR Paediatric MDS/JMML inclusion and exclusion criteria

Inclusion criteria
1. Age from birth to 21 years of age 2. Probable diagnosis of MDS: a. Sustained (> 3 months) unexplained cytopenia b. Bi/trilineage dysplasia c. Acquired clonal cytogenetic abnormalities in haematopoietic cells d. Bone marrow blasts > 5% 3. Children with syndromic MDS or suspected syndromic MDS 4. Children with secondary MDS (treatment related) 5. Children with Inherited Bone Marrow Failure Syndromes (IBMFS) and features of myelodysplasia 6. Children with suspected familial MDS 7. De novo MDS (hypocellular/hypercellular) 8. JMML 9. MDS/MPS 10. MDS with systemic disease 11. Children with Downs MDS
Exclusion criteria
1. Patients who do not fulfill the inclusion criteria 2. Patients who have not consented for the study 3. Children with known Core Binding Factor (CBF) AML translocations

DNA was screened for known MDS/JMML mutations using a customized targeted NGS panel and mutations identified were validated with Sanger sequencing. Sanger sequencing was performed in the GOSH diagnostic lab in order to allow validation of results by an ISO accredited lab (Appendix 3 for example of research report generated). A research report was generated and sent to the clinical team of each patient. Teleconferences with each centre were organized at regular intervals to discuss the results with each participating centre and obtain further clinical information. For patients with well defined phenotypic features where no mutations were detected with using targeted NGS, DNA aliquots were sent to NIHR Bioresource study in Cambridge (following informed second consent for the NIHR Bioresource study) for whole genome sequencing (WGS). For these cases, DNA was extracted by myself and sent to Illumina in Cambridge, where sequencing and alignment of data to the human genome of reference was performed. Following this the data were data were uploaded to an online software Sapientia (www.sapientia.co.uk), where variants were visualized. Downstream analysis for the identification of the causative variants was performed by myself and discussed at a national MDT organized by myself and led by Prof Irene Roberts.

Serial samples were obtained for each participant at the clinician's discretion and were also used for sequencing in order to detect disease clone evolution. Germline material, whenever available, was also screened for the mutations identified for each patient in order to identify patients with inherited predisposition syndromes. Unfortunately the analysis of germline material was limited to NGS panel as the amount of germline DNA available did not meet in any of the cases WGS requirements.

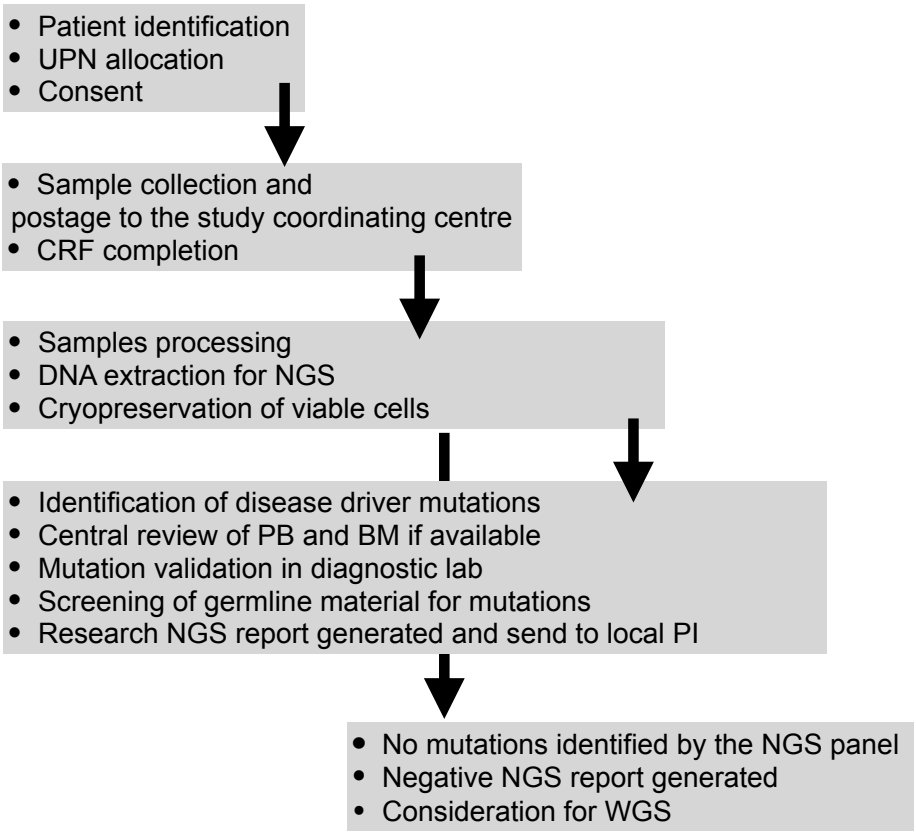


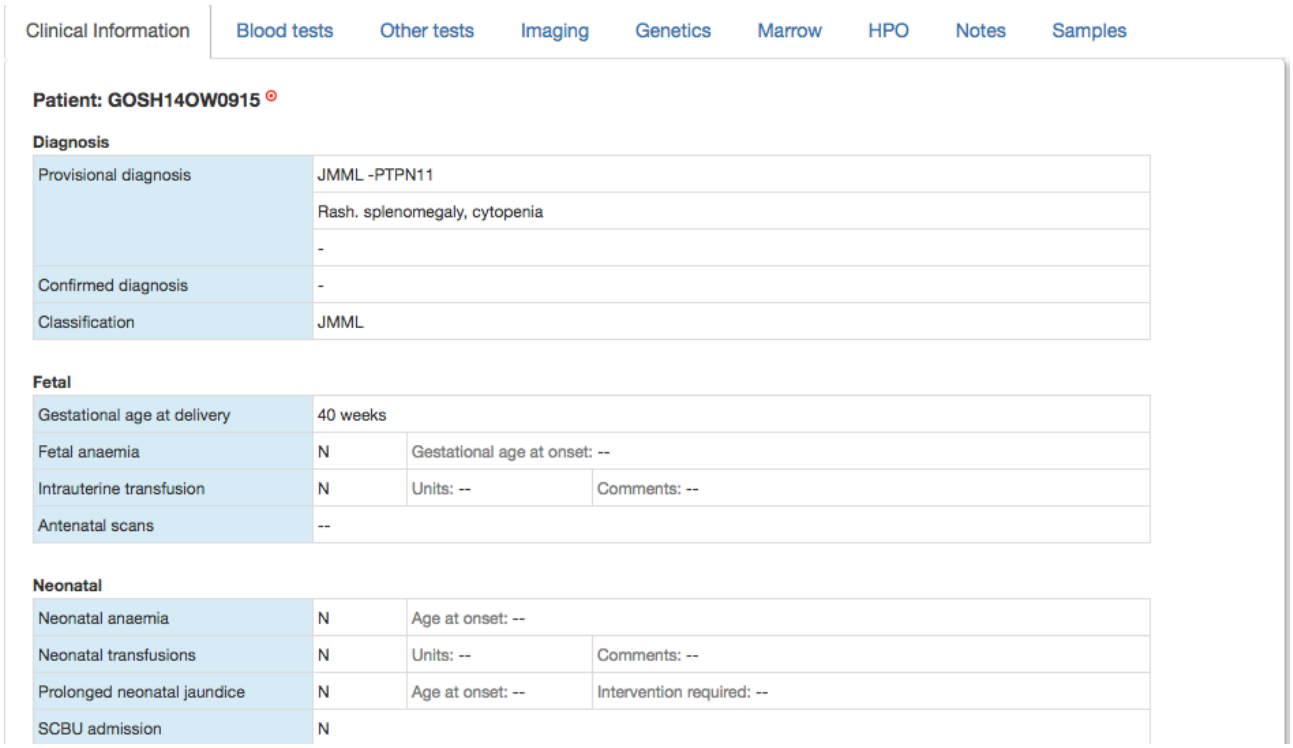
Figure 2.2: Schema showing the design of the NIHR Paediatric MDS/JMML study

2.1.3 Deep phenotyping of patients enrolled into the NIHR Paediatric MDS/JMML study

A customized, restricted-access password protected database was designed in order to capture very detailed phenotypic information for all recruited patients. The clinical information required was designed by myself and recorded on an adapted version of an existing database in the WIMM used for rare unexplained anaemias. This included a detailed clinical history and presenting symptoms, family history, previous medical history and birth history. All investigations and diagnostic reports that were performed locally and a detailed clinical examination by an experienced clinician was performed in order to identify any subtle dysmorphic or syndromic features (Figure 2.3). The phenotypic information was coded using the human phenotype ontology

(HPO) tool and these terms were also captured on the online database. Follow up investigations performed by the local clinical team as well as the results of genetic testing were also securely saved on the online database.

Figure 2.3: Representative screenshots from the NIHR study database (p88-91)



Clinical Information | Blood tests | Other tests | Imaging | Genetics | Marrow | HPO | Notes | Samples

Patient: GOSH14OW0915

Diagnosis

Provisional diagnosis	JMML -PTPN11		
	Rash. splenomegaly, cytopenia		
	-		
Confirmed diagnosis	-		
Classification	JMML		

Fetal

Gestational age at delivery	40 weeks		
Fetal anaemia	N	Gestational age at onset: --	
Intrauterine transfusion	N	Units: --	Comments: --
Antenatal scans	--		

Neonatal

Neonatal anaemia	N	Age at onset: --	
Neonatal transfusions	N	Units: --	Comments: --
Prolonged neonatal jaundice	N	Age at onset: --	Intervention required: --
SCBU admission	N		

Early Childhood

Age at onset of symptoms	--, 10 months		
Failure to thrive	N	Comments: --	
Clotting problems	N	Comments: --	
Platelet-related problems	N	Comments: --	
Developmental delay	N	Comments: --	
Frequent infections	N	Type: --	Comments: --
Transfusions - red cells	--	Units: --	Comments: --
Transfusions - platelets	never	Units: --	Comments: --
Transfusions - FFP	never	Units: --	Comments: --
Transfusions - Cryo	never	Units: --	Comments: --
Pancreatic insufficiency	N		
Gall stones	N	Comments: --	
Venous thromboembolism	N	Date: --	Comments: --
Aplastic crises	N	Date: --	Comments: --
Parvovirus positive	--		
Hepatomegaly	Y	Size: 2cm	
Splenomegaly	Y	Size: 12cm	
Lymphadenopathy	N	Comments: --	
Vaccination history	--		
Other health problems / surgery	4 months old - previous admission for chest infection and wheeze		
Deceased	N	Age at death: --	

Therapy

Splenectomy	N	Date: --	Response: --
Oral antimetabolites	Y	Comments: on 6MP	
Hypomethylating agents	N	Comments: --	
AML chemotherapy	N	Comments: --	
Other therapy	Received BMT Feb 2016		

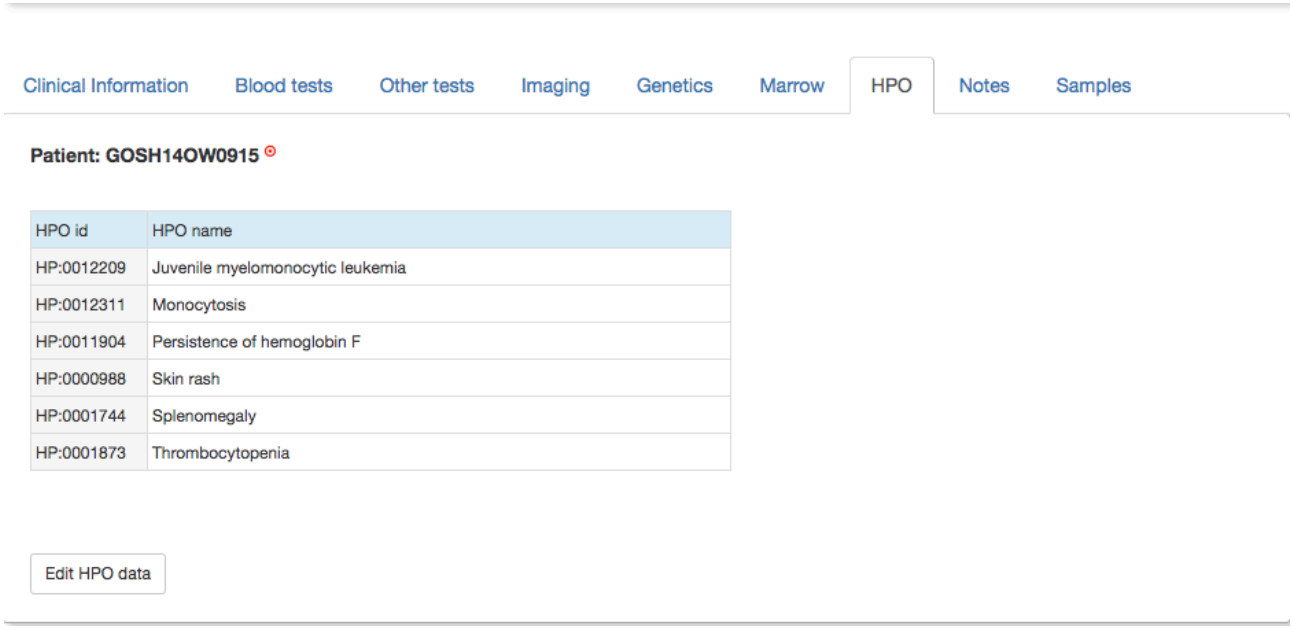
Treatment response

Response	complete remission		
Relapse	--	Age at relapse: --	Comments: --

Patient: GOSH14OW0915

Haematology [+/-]

Date	09-Sep-2015	30-Sep-2015	01-Oct-2015	05-Nov-2015	07-Nov-2015	22-Nov-2015	18-Dec-2015	15-Jan-2016	25-Jan-2016	28-Jan-2016	30-Jan-2016	01-Feb-2016	06-Feb-2016	09-Feb-2016	15-Feb-2016	18-Feb-2016	19-Feb-2016	20-Feb-2016	21-Feb-2016
Stage	diagnosis	follow-up	follow-up	follow-up	follow-up	follow-up	follow-up	follow-up	follow-up	follow-up	follow-up	follow-up	follow-up	follow-up	follow-up	follow-up	follow-up	follow-up	follow-up
On transfusion	N	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--
RBC (x10^{12}/l)	3.14	--	--	--	--	--	--	--	--	--	--	--	--	--	--	3.82	4.08	3.46	3.28
Hb (g/l)	66	96	99	75	68	84	99	96	91	83	81	102	97	77	116	101	110	92	87
HbF (%)	8.6	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--
Hct	0.243	0.296	0.312	0.249	0.224	0.27	0.32	0.29	0.28	0.27	0.29	--	0.29	0.23	0.33	0.292	0.31	0.262	0.248
MCV (fl)	77.4	79.8	81.3	76.9	75.7	75.9	74.9	73	75	75	76.9	72	72	73	77	76.4	76	75.7	75.6
MCH (pg)	21	25.9	25.8	23	23	23.5	23	24	23	23.3	23	23	24	24	26	26.4	27	26.6	26.5
MCHC (g/l)	272	324	317	301	304	310	309	325	317	307	300	321	331	332	344	346	355	351	351
RDW (%)	23.1	--	--	--	--	--	--	--	--	--	--	--	--	--	--	15.2	15	14.8	14.6
Platelets (x10^{9}/l)	68	30	33	77	43	70	242	328	441	415	414	183	17	38	13	78	47	25	18
MPV (fL)	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	9.1	10.1	--	--
WBC (x10^{9}/l)	32.29	46.04	32.38	9.25	7.65	7.27	7.66	7.4	6.17	5.6	2.42	3.5	0.13	0.03	0.76	0.81	1.2	0.6	0.59
Blast (%)	1.29	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--
Chimerism (%)	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--
Neutrophils (x10^{9}/l)	7.75	17.03	12.3	2.96	2.43	2.23	1.51	1.02	2.37	1.2	1.16	1.64	--	--	--	--	--	0.04	0.04
Lymphocytes (x10^{9}/l)	20.67	23.48	12.3	5.37	4.38	4.2	4.58	5.13	3.64	3.89	0.94	1.47	0.7	--	0.58	0.31	0.65	0.23	0.34
Monocytes (x10^{9}/l)	1.61	2.3	3	0.79	0.63	0.5	1.25	0.23	0.13	0.5	0.32	0.27	--	--	0.15	0.48	0.54	0.32	0.21
Eosinophils (x10^{9}/l)	0.32	--	0.65	0.1	0.15	0.26	0.3	0.69	--	--	--	0.01	--	--	0.01	--	--	0.01	--
Basophils (x10^{9}/l)	0.65	0.92	0.65	0.03	0.06	0.01	0.02	0.07	0.03	0.01	--	0.16	--	--	0.02	0.02	0.01	--	--



2.1.4 Sample processing and cryopreservation

Samples were processed as per previously established methods in the lab. In short, the samples were diluted 1:1 with phosphate buffered saline (PBS) (Thermo Fisher, UK) and layered onto 1/3 of the sample volume Ficoll density gradient (Sigma-Aldrich, UK). The sample was centrifuged for 30 min at 400g, at room temperature. The interface mononuclear cell (MNC) layer was collected and washed twice with 1:10 FCS at 300g for 10 minutes. The supernatant was removed and the cells were collected for cryopreservation using 90% fetal bovine serum (FBS) and 10% dimethylsulfoxide (DMSO). For all downstream experiments the cells were carefully thawed and washed twice, at 300g for 10 min.

2.1.5 DNA extraction

DNA from PB, BM and germline material (hair /skin/nails) was extracted using the Qiagen DNeasy blood and tissue kit (Qiagen, Manchester, UK) as per the manufacturer's instructions. In short, the sample was lysed at 56°C for 10 min using

the lysis buffer provided by the kit. The sample was then transferred to a spin column and washed with serial kit buffers in order to remove impurities. DNA was eluted in EB buffer (10 mM Tris-Cl, pH 8.5, Qiagen, Manchester, UK) and was quantified using qubit (Invitrogen Life Technologies, California, USA)

2.2 Paediatric MDS/JMML targeted NGS panel

2.2.1 Target selection, amplicon design and coverage prediction

The genes on the Paediatric MDS/JMML targeted NGS panel were selected on the basis of previous literature reports in order to allow well described mutations in Paediatric MDS to be identified. The panel included myeloid genes commonly implicated in myeloid malignancies in children and adults, transcription factor genes linked with inherited MDS predisposition syndromes and inherited BM failure implicated genes in order to allow the identification of dysplasia developing in the context of an inherited BM failure syndrome (Table 2, Figure 2.4).

RAS pathway	***PTPN11, KRAS, NF1, NRAS, HRAS ,CBL***
Transcription factors	***RUNX1, CEBPA, GATA1, GATA2***
Inherited bone marrow failure genes	***CSF3R, ELANE, G6PC3, GFI1, HAX1, MPL,SBDS, SRP72, WAS***
Epigenetic modifiers	***ASXL1, DNMT3a, EZH2, TET2***
Signalling	***FLT3,ARHGAP26,JAK3,SETBP1,SH3BP1***
Other tumour suppressor genes	***PTEN, TP53***
Spliceosome machinery	***SRSF2, U2AF1***

Figure 2.4: Genes selected for inclusion in the targeted NGS panel

Table 2: Gene targets and exons included in the targeted NGS panel

Gene	Exons included	Literature for previously described mutations
ASXL1	12	(280)
CBL	1-16	(213), (206)
CEBPA	1	(115)
CSF3R	15	(281)
DNMT3a	1-23	(282)
ELANE	2,4,5	(283), (284)
EZH2	1-19	(285)
FLT3	7,14,15,16,20	(286)
GATA1	2,3,4	(287) ,(288)
GATA2	2,3,4,5,6	(289),(79), (290)
GFI1	1,7	(291)
G6PC3	1,3,4,5,6,7	(292), (293)
GRAF/ARHGAP26	1-23	(294)
HAX1	3,4,6,7	(295)
HRAS	1-6	(296)
JAK3	9,10,11,15,17,19,21	(213)
KRAS	1-6	(213), (193)
MPL	2-5,7,10,12	(297)
NF1	1-57	(213)
NRAS	1-7	(213), (193)
PTEN	1-9	(298)
PTPN11	1-16	(213) ,(186)
RUNX1	1-9	(299), (300) ,(301) ,(73)
SBDS	1,2,3,5	(45), (302)
SETBP1	4	(213)
SH3BP1	10	(213)
SRP72	1-19	(93)
SRSF2	1-2	(303), (304)
TET2	1-11	(285)
TP53	1-11	(305)
U2AF1	2,6,7	(303)
WAS	1,2,4,6,7,9,10,11,12	(306)

The panel was designed using the Illumina design studio (http://www.illumina.com/informatics/research/experimental-design/designstudio.html) and contained 32 genes with a total of 916 amplicons (covering 301 exons) ~ 250 base pairs each. Coverage for each target was initially predicted by the Illumina software (Figure 2.5) which allowed adjustments of the panel design in order to

achieve maximum possible coverage per amplicon. In addition, coverage of each base within a target region was assessed using an in-house bioinformatic tool developed at the Oxford Molecular diagnostic lab called Covermi software. (307). Examples of the coverage predicted for each region are shown in Appendix 1.

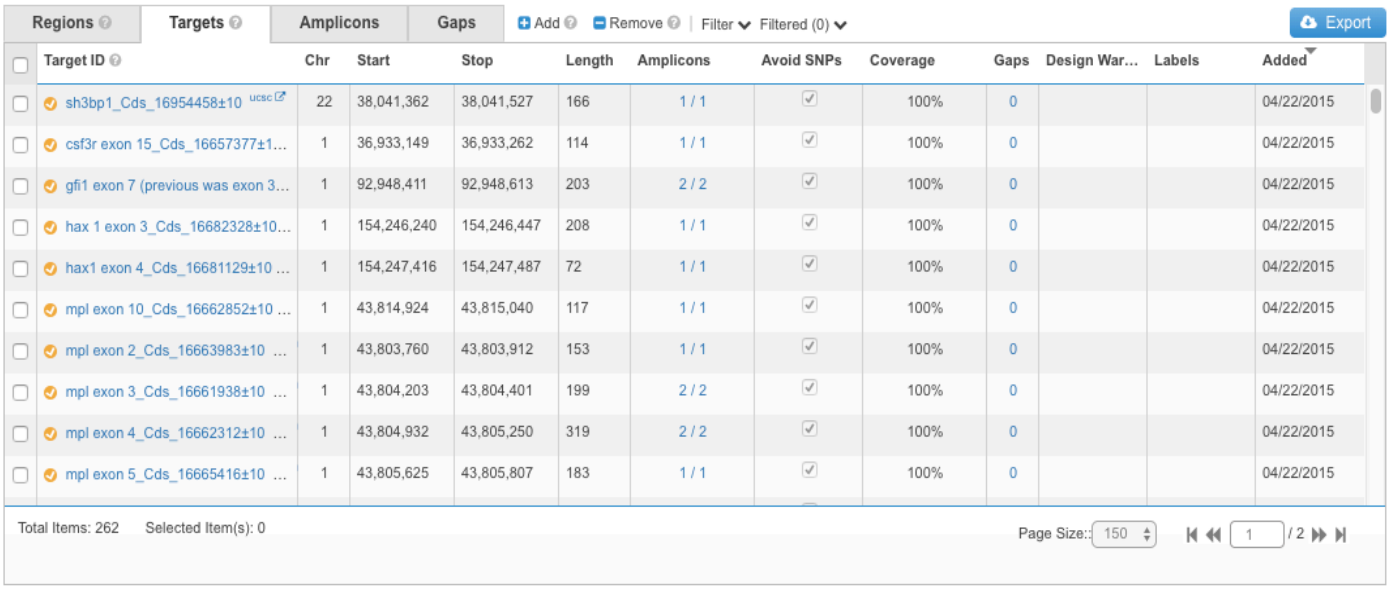


Figure 2.5: Example of Illumina design studio target selection and software prediction of overall coverage and coverage per amplicon.

2.2.2 Library preparation and sequencing

Library preparation and sequencing was performed using the Illumina Truseq Custome Amplicon panel (TSCA) v1.5 kit. With the aim of producing clinical grade results that can be reported back to patients, I adopted the Oxford molecular diagnostic lab's NGS protocol and performed the experiments in the diagnostic lab in order to ensure rigorous checks of the library preparation were followed. The complete protocol for the NGS library preparation is included in Appendix 2. In short, samples were prepared for sequencing and normalized to a concentration of 25ng/ul, as per manufacturer instructions. DNA hybridization with custom probes was performed with the reagents provided by the manufacturer. Unbound oligos were removed and adaptor ligation/extension of the regions of interest was done as per

protocol. Sample indexes were added and PCR amplification followed with the following conditions:

Heated lid: 100°C
Initial activation: 95°C for 3 minutes
Amplification cycling program: (22 cycles)
95°C for 30 seconds
66°C for 30 seconds
72°C for 60 seconds
Final extension: 72°C for 5 minutes
Pause: 10°C ∞

Libraries were purified with Ampure XP beads, normalized and samples were pooled and prepared for sequencing at a concentration of 4nM. Sequencing was performed using the Illumina MiSeq with the v2 300 cycle kit. Initial sequencing experiments were performed by myself. Subsequently the patient's DNA samples were added to the routine diagnostic lab workflow, as due to the rarity of the samples and the aim of reporting results back to clinicians within 8-12 weeks, it was not possible to batch samples and perform individual sequencing runs.

2.2.3 Samples alignment and analysis

Data analysis was performed using in house bioinformatic pipelines at the Oxford Molecular Diagnostic Centre, with the supervision of Dr Angela Hamblin in order to ensure that the results reported back to the families were analysed following rigorous procedures used for diagnostic samples. In short, FASTQ files were processed by the MiSeq Reporter and sequences aligned to the reference genome (hg19), by the Smith-Waterman algorithm and Illumina Somatic Variant caller (1.1.0). Variants were annotated using Illumina VariantStudio (2.2) and were filtered

using the following passing filters: quality >100 and consequence: missense, frameshift, stop gained, stop lost, initiator codon, in frame insertion, in-frame deletion, splice filters (Figure 2.6).

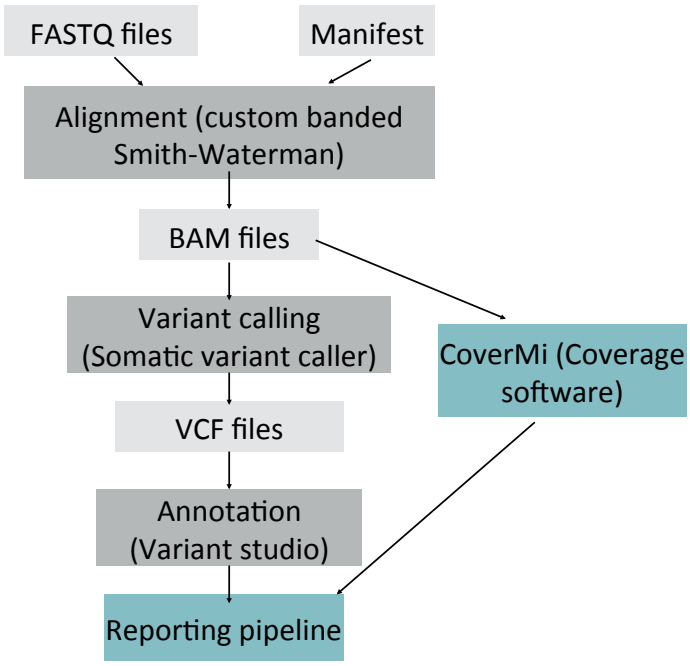


Figure 2.6: Bioinformatics pipeline for sample analysis

Variants were verified by visual check by using the Interactive Genomics Viewer (IGV, 2.3.40) and coverage for each base sequenced and for each sample sequenced was also assessed using the CoverMi software. The overall coverage of each run was estimated at ~96% with a mean depth of 2846 reads/bp. The CoveMi software allowed the rapid identification of samples where disease-associated mutations were poorly sequenced so that re-sequencing could be repeated to ensure that the regions of interest were always screened. Variants identified as possible mutations were compared against public databases, for example COSMIC, for previous reports of somatic mutations; and identification of common polymorphisms that were not filtered by variant studio was done using the Exac

online database (http://exac.broadinstitute.org/) (Figure 2.7), in line with routine practice of the Oxford Molecular diagnostics lab.

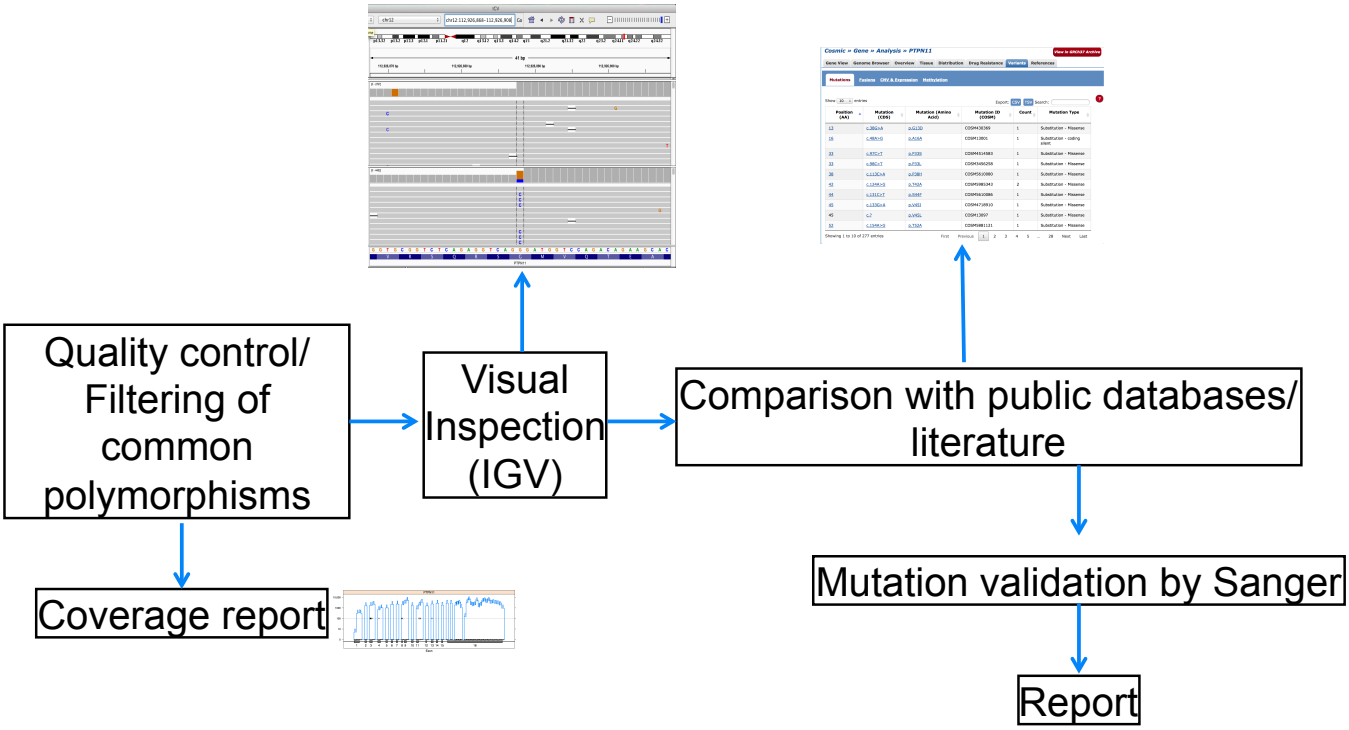


Figure 2.7: Paediatric MDS/JMML NGS panel reporting pipeline

2.2.4 Mutation validation/ Reports generation

Variants were validated by Sanger sequencing at Great Ormond Street Hospital diagnostic lab (Figure 2.8). Research reports for the samples were generated and sent to the referring clinician (see Appendix 3 for examples of NGS reports).



Figure 2.8: Example of Sanger sequencing result used to confirm the presence of a *KRAS* c.38 G>A in a JMML patient sample, that was originally identified by the targeted NGS panel. Arrows indicate the location of the mutation G>A.

2.3 Investigation of disease-propagating populations in JMML

2.3.1 Phenotypic and functional characterization of JMML HSPCS

Samples used in this study were prospectively collected as part of the NIHR Paediatric MDS/JMML Study (as outlined above). In addition, a small number of samples was provided by the UK Childhood Leukaemia Cell Bank (n=5) (NHS REC reference number 16/SW/0219). In the majority of experiments, cord blood (purchased from Zen-Bio, NC,USA) was used as a normal control. Paediatric BM from normal donors was used for a small number of key experiments due to the very limited availability of such material. Samples used were collected after informed consent by Dr D.Iskander (NHS REC reference number 12/LO/0426:Project title Investigation of the molecular and cellular pathogenesis of bone marrow failure syndromes.)

2.3.1.1 Flow cytometry and fluorescence activated cell sorting (FACS)

Frozen viable cells were carefully thawed at room temperature using a water bath. MNC were initially washed with 100% FBS and then FACS buffer: IMDM (Thermo Fisher, UK), 10% FBS and 10ng/ml Dnase (Stem cell Technologies, Cambridge UK). All washes were done at a 1:10 ratio at 300g, 4°C for 10 min. After washing, cells were resuspended in 100 ul of FACS buffer and incubated with a fluorophore-conjugated monoclonal antibody for 20 minutes at 4°C. Following incubation cells were washed and re-suspended in 100-1000 ul of FACS buffer for flow cytometry analysis or cell sorting. Where unconjugated antibodies were used, cells were initially incubated, as described above, with the unconjugated antibody, washed and incubated again with the fluorophore of choice. Details of the antibodies used for this project are included in Appendix 4. The optimal concentration of antibodies was determined after careful titration (Figure 2.9).

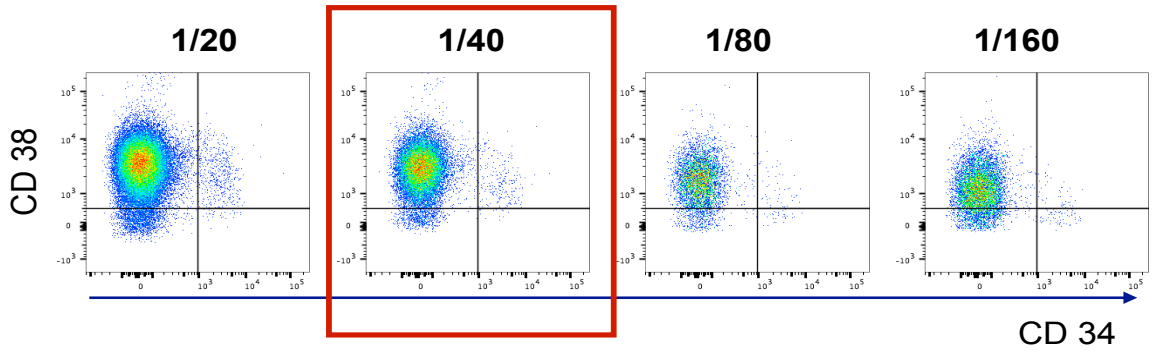


Figure 2.9: Titration of CD38 antibody. 1/40 was identified as the optimal concetration to use for experiments

2.3.1.2 Flow cytometry analysis

Flow cytometry analysis was performed on the LSR Fortessa X50, X20 and LSR II (Becton Dickinson, Oxford, UK). All FACS experiments included single colour-stained controls (CompBeads, BD Biosciences) and Fluorescence Minus One controls

(FMO). Data analysis was performed using FlowJo software (version 10.1, Tree Star Oregon, USA). Gating strategies for each individual experiment is described in the relevant results section. For all experiments this included initial gates of MNC, then single cells and subsequently on viable cells as shown in Figure 2.10.

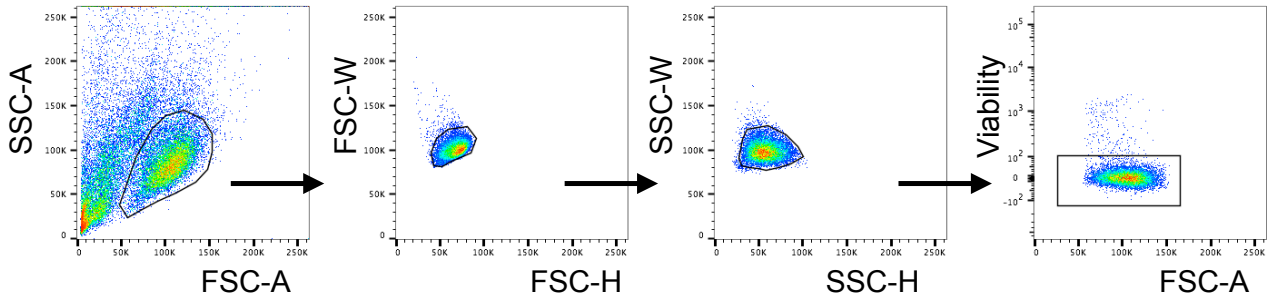


Figure 2.10: Initial gating strategy used for all experiments- label what you are gating at each stage

2.3.1.3 Fluorescence activated cell sorting (FACS)

FACS-sorting of MNC was performed using the FACS Aria II, FACS Aria III or FACS Aria Fusion (Becton Dickinson, Oxford, UK). For bulk sorts, purities of the populations sorted was always assessed at the beginning by reanalysis of the sorted populations. For single cell sorting, all experiments were done using the 'single cell' purity setting. Correct alignment of the sorter was tested at every experiment by sorting BD FACS Accudrop Beads (BD Biosciences) into an eppendorf tube or 96-well plate in order to ensure that all beads were delivered to the bottom of the well/tube. Operation of FACS sorter machines was perfomed by either Benjamin Povinelli or the WIMM flow cytometry facility.

2.3.1.4 Identification of JMML haematopoietic stem and progenitor cells (HSPC)

HSPC populations were isolated by flow cytometry using well established immunophenotypic markers (308), (309). HSC (Haematopoietic stem cells) were

defined as Lin-CD34+CD38-CD90+CD45RA-; MPP (multipotent haematopoietic progenitors) as Lin-CD34+CD38-CD90-CD45RA-; LMPP (lymphoid-primed multipotent progenitors) as Lin-CD34+CD38-CD90-CD45RA+; CMP (common myeloid progenitors as Lin-CD34+CD38+CD123+CD45RA-; MEP (megakaryocyte-erythroid progenitors) as Lin-CD34+CD38+CD123-CD45RA-; and GMP (granulocyte-macrophage progenitors) as Lin-CD34+CD38+CD123+CD45RA+. The antibody panel is shown in Table 2.3 and the gating strategy in Figure 2.11. The antibody panel used in order to eliminate mature downstream erythroid, lymphoid and myeloid progenitors, (310), however monocytic markers such as CD14 were not included in the lineage mix due to expression of these markers on JMML cells (176).

Table 2.3: Antibody panel used for HSPC identification and isolation

Antibody	Fluorochrome
CD38	PE Texas Red
CD34	APC ef 780
CD90	BV421
CD45RA	PE
CD123	PE Cy7
7-AAD (aminoactinomycin D) viability	
Lineage panel CD8 CD20 CD2 CD3 CD16 CD19 CD235a CD66b CD10 CD127	FITC

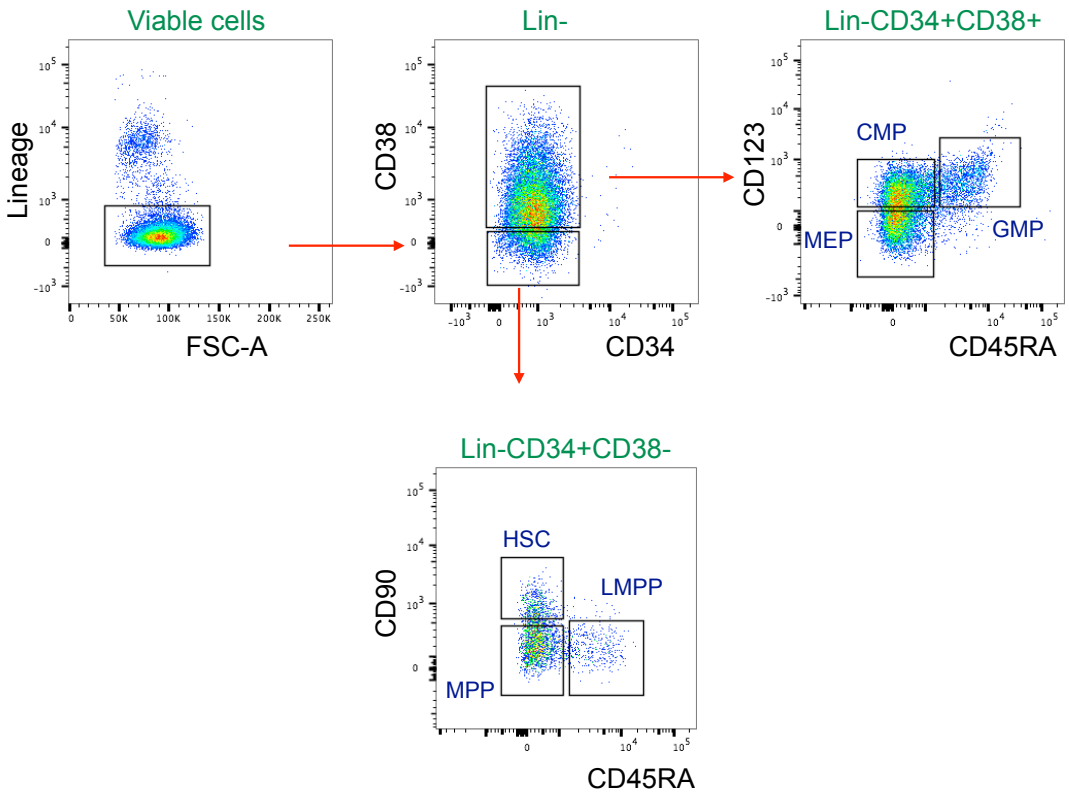


Figure 2.11: Routine gating strategy for HSPC isolation from normal cord blood CD34+ cells. CD34+ mononuclear cells were commercially obtained

2.3.1.5 Clonogenic assays

2.3.1.5.1: Single cell methylcellulose culture assays

Single HSPC were sorted based on their immunophenotypic definitions into 96-well plates with 50 ul of methocult H4435 (Stemcell Technologies, Vancouver, Canada). Methocult H4435 contains essential cytokines for erythroid and myeloid cell growth (SCF, IL-3, IL-6, EPO, G-CSF, GM-CSF). Cells were cultured at 37°C in 5% CO_2 for 2 weeks. Lineage output was assessed on day 7, day 10 and day 14 by evaluating the morphology of the colonies under direct light microscopy and verified by plucking individual colonies and resuspending them in 100ul of PBS for cytospins or analysis by flow cytometry (Table 2.4). Cytospins were prepared on a cytocentrifuge (Thermo Fisher Scientific Cytospin 4 machine, Thermo Fisher Scientific, UK) and stained using the Hematek stain pak (Wright-Giemsa motified stain) with the

Hematek staining system (Bayer HealthCare, Berlin, Germany).

Table 4: Antibodies used to assess the lineage output of clonogenic JMML HSPC

Antibody	Fluorochrome
Viability(AQUA)	BV 510
CD11b	PE Cy7
CD14	PE
CD36	APC
CD90	BV 421
CD235a	FITC
CD34	APC ef 780

For *in vitro drug testing experiments* JQ1 (MedChem Express, final concentration 250 nM), PD-0325901 (Cambridge Bioscience, final concentrations 1uM or 5uM), Fasentin (Sigma-Aldrich, final concentrations 5nM or 20nM) or equivalent DMSO vehicle control were added to the methocult H4435 prior to FACS sorting. Samples were incubated in 37°C in 5% CO_2 and colonies were analysed as described above after 14 days of culture.

GM-CSF hypersensitivity and independence: This was tested by seeding single HSPC into 50ul of methocult H4230 (Stemcell Technologies, Vancouver, Canada) with cytokines (all Peprotech, London, UK) added as shown in Tables 2.5 and 2.6A+B. The mixture was vortexed thoroughly, left for 10-15 minutes at room temperature until the methocult was free of bubbles and 50 ul was plated in 96-well plates and single cells were sorted into each well. Cells were cultured at 37°C in 5% CO_2 for 2 weeks and clonogenic output was assessed as outlined above.

Table 2.5: Custom cytokine cocktail concentration for single cell clonogenic assays

Cytokines	Final concentration
EPO	4 U/ml
SCF	10 ng/ml
IL-3	20 ng/ml
IL-6	10 ng/ml
G-CSF	50 ng/ml
GM-CSF	50 ng/ml
Penicillin /Steptomycin (P/S)	1%

Table 2.6A: Different cytokine combinations

Conditions	Cytokines used
1 (Control)	EPO,SCF,IL-3, IL-6, G-CSF, GM-CSF
2	EPO,SCF,IL-3, IL-6, G-CSF
3	EPO,SCF,IL-3, IL-6
4	SCF,IL-3, IL-6, G-CSF, GM-CSF
5	SCF,IL-3, IL-6, G-CSF
6	SCF,IL-3, IL-6

Table 2.6B: Different cytokine concetrations

Condition	Methocult	IMDM (ul)	1%P/S (ul)	EPO (ul)	SCF (ul)	IL-3 (ul)	IL-6 (ul)	G-CSF (ul)	GM-CSF (ul)	Final volume (ul)
1	H4435	623	77	X	X	X	X	X	X	700
2	H4230	530	77	30	7	14	7	35	X	700
3	H4230	565	77	30	7	14	7	X	X	700
4	H4230	525	77	X	7	14	7	35	35	700
5	H4230	560	77	X	7	14	7	35	X	700
6	H4230	595	77	X	7	14	7	X	X	700

2.3.1.5.2 Megakaryocyte (MK)/Erythroid liquid single cell culture assay

To determine the megakaryocytic output of JMML HSPC, single cells were directly sorted into 96-well plates with 70 ul of erythroid/MK bipotential medium containing Stemspan (Stemcell Technologies) supplemented with TPO (100ng/ml), EPO

(1 IU/ml), SCF (100ng/ml), IL-3 (10 ng/ml), IL-6 (10ng/ml) (all Peprotech, London, UK) and hu LDL (40 mg/ml) (Stemcell Technologies). Single cells were cultured at 37°C in 5% CO_2 for 2 weeks, with 50ul of fresh medium added per well on day 4. Colonies were counted by direct light microscopy, then washed with 100ul of FACS buffer at 300g for 10 min, then stained in the 96-well plate for 20 min at 4°C after the supernatant was removed then resuspended in 150ul of FACS for flow cytometry to determine MK output. The antibody panel used is shown in Table 2.7.

Table 2.7: Antibodies used to assess the Megakaryocytic output of JMML HSPCs

Antibody	Fluorochrome
DAPI (viability)	BV421
CD41	APC
CD42	PE
CD11b	FITC
CD14	FITC
CD34	APCef780
CD235a	PerCyp5.5

2.3.1.5.3 Assessing lymphoid differentiation in MS5 stroma co-cultures

B-lymphoid differentiation of human HSPC is classically measured by co-culturing the HSPC with the murine stromal cell line, MS5.

Culture and passaging of MS5 stromal cells: MS5 cells (DSMZ, Germany) were thawed at 37°C and were transferred to 10 ml of MS5 medium containing a-MEM (Invitrogen UK), 10% heat-inactivated FBS, 1% HEPES (Stem Cell Technologies, Cambridge,UK), 50uM beta-Mercaptoethanol (Sigma-Aldrich,Missouri,USA) , 1% Penicillin/Streptomycin (50U/ml) and 1% L-glutamine). The cells were washed at 400g for 5 min and resuspended in 12ml of MS5 media, transferred to a 25cm^2 flask and grown to 80% confluence before passaging at a 1:3 ratio. In order to passage the cells the excess medium was removed, the flask was washed thoroughly with 5ml

PBS and the cells incubated at 37°C with 0.5% trypsin/EDTA (Invitrogen, Paisley UK). Following trypsinisation, MS5 medium was added and the cells were vigorously washed in order to remove the adherent cells and MS5 medium was added and the cell suspension was divided into further flasks, cryopreserved with 90% FBS and 10% DMSO or plated into 24-well plates for cell-sorting.

Establishing HSPC: MS5 co-cultures: 100 cells from relevant HSPC populations were sorted into 100ul of a 2X cytokine medium containing: FLT3 (10 ng/ml), SCF (20ng/ml), IL2 (10 ng/ml), IL7 (5 ng/ml), all from Peprotech and added to 80% confluent MS5 monolayers for 4 weeks as previously described (311). Co-cultures were disaggregated by vigorous pipetting, passed through a 70 μm filter and replated on fresh MS5 cells every 4-5 days with 100 ul of fresh 1X cytokine medium added per well. The output of mature (CD34-CD19+) B cells was determined by flow cytometry at week 2, 3 and 4 of the assay. Details of the antibodies used are shown in Table 2.8.

Table 2.8: Antibodies used for B-lymphoid output analysis by flow cytometry

Antibody	Fluorochrome
DAPI (viability)	BV421
CD11b	FITC
CD14	FITC
CD34	APC ef 780
CD235a	PerCyp5.5
human CD45	Af 700
CD19	PE
CD56	FITC
CD33	FITC

2.3.1.6 Gene expression analysis for human globin expression in JMML erythroid colonies

Erythroid colonies were picked under direct microscopy and resuspended in 100 ul of PBS. RNA was extracted using the Qiagen RNAeasy Kit (Qiagen, Manchester, UK) according to the manufacturer's instructions. cDNA was produced using AB High Capacity RNA to - cDNA kit (Catalog number: 4368813, Thermo Fisher, UK). RT-PCR conditions for cDNA production are shown in Table 2.9. Quantitative PCR (qPCR) was performed, using Taqman Universal Master mix (catalog number: 40304437, Thermo Fisher UK) and taqman probes for HBB, HBA, HBG and HBE1, as shown in Table 2.10A+B (all Invitrogen).

Table 2.9: RT-PCR condition for cDNA production

Reagents	Concentration	µl
RT Buffer	2x	10
Enzyme Mix	20x	1
RNA<2ug	<222ng/ul	9
total		20
Programme	37°C 60min 95°C 5min 4°C hold	

Table 2.10A: qPCR conditions for taqman probes hybridisation

Reagents	µl	x10
2x TaqMan Master Mix	5.0	50
Probe ID (see table 11B)	0.5	5
H_2O	2.5	25
cDNA	2.0	
total	8.0/well	80

Table 2.10B: Taqman probe IDs used for these reactions

Gene Name	Assay ID
HBA	Hs00361191_g1
HBB	Hs00747223_g1
HBE1	Hs00362216_m1
HBG	Hs00361131_g1
GAPDH	Hs02758991_g1

2.3.2 'Driver' mutation tracing to HSPC compartments in JMML

Mutations were traced back to the cell of origin by mutational analysis of 1) colonies that were grown in methylcellulose from single cells sorted from each population and 2) FACS-sorted single cells from each population.

2.3.2.1 Design of patient-specific primers

Patient-specific primers were designed for downstream amplification of the region of interest using the Primer blast tool website (https://www.ncbi.nlm.nih.gov/tools/primer-blast/). The sequences of primers used can be found in Appendix 5. For every individual cell sorted for genotyping, all index sort data of the MFI for CD34, CD38, CD90, CD45RA and CD123 were recorded. This allowed the immunophenotypic identification of every cell processed and mapping of the JMML mutations onto the haematopoietic hierarchy (Figure 2.12). The protocol used for library preparation was adapted and modified from the Ilumina access array Fluidigm to allow processing and sequencing of high numbers of colonies.

2.3.2.2 Protocol for genotyping of individual plucked HSPC colonies

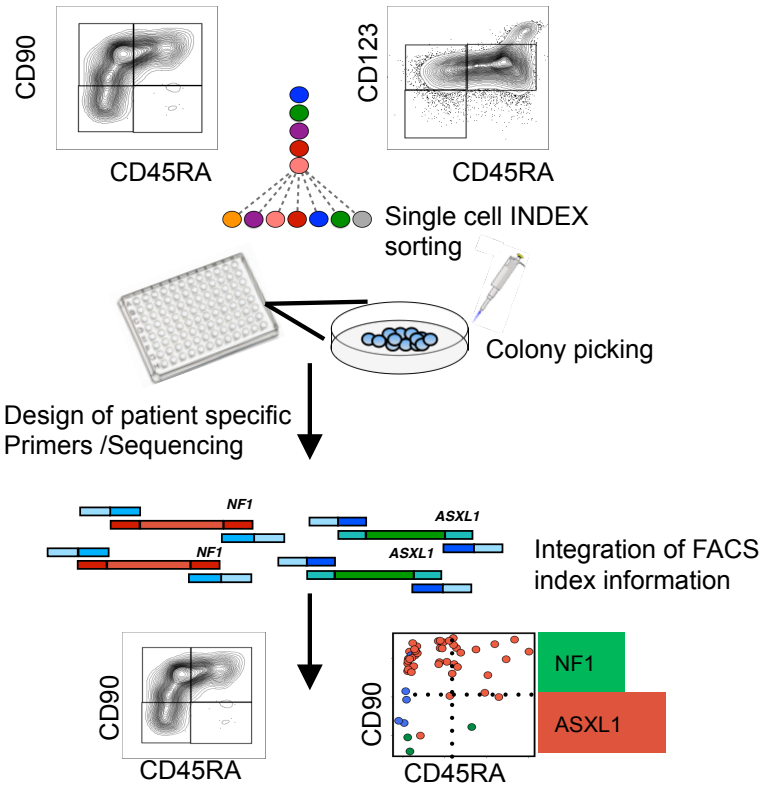


Figure 2.12: Flow diagram of the single cell genotyping protocol: Single cells were FACS sorted into 96 well plates and cultured for 2 weeks. At the end of this period cells were picked and prepared for genotyping.

A. Genotyping protocol from single cell-derived colonies:

Colonies derived from sorted single cells from each population were picked and resuspended in 10 µL PBS and frozen at -80^{0}C for downstream analysis. Samples were thawed and lysed using 15 µL of a custom lysis buffer containing 10% Proteinase K (Qiagen, Manchester, UK) and 0.4% Triton X. Samples were incubated at 56°C for 10 minutes and 75°C for 20 minutes to release genomic DNA and inactivate the Proteinase K, respectively. Regions of interest were amplified using a 3-step nested PCR method with FastStart High Fidelity enzyme (Roche, Basel Switzerland). For the first PCR, 3.6 µL of the lysate was amplified with patient-specific primers covering the regions of interest with a targeted length of ~300bp

amplicons with 23 cycles. For the second PCR step, 1 µL of the PCR1 product was amplified with nested target-specific primers generating ~200bp amplicons for 12 cycles. Details of the PCR reagents and amplification conditions are shown in Table 2.11 A+B+C. Primers for PCR 2 were attached to Access Array adaptors (Fluidigm; Forward adaptor: ACACTGACGACATGGTTCTACA; Reverse adaptor: TACGGTAGCAGAGACTTGGTCT). Lastly, for the third PCR step, Illumina compatible adaptors containing single-direction indexes (Access ArrayTM Barcode Library for Illumina Sequencers-384, Single Direction, Fluidigm) were attached to pre-amplified amplicons from PCR2. This final PCR product was purified with Ampure XP beads (0.8: to 1 ratio beads to product), quantified using Quant-iT Picogreen (Thermo Fisher Scientific), pooled and sequenced on the Illumina MiSeq with 150 bp paired-end reads.

Table 2.11A: Colony genotyping library preparation protocol: PCR 1

Component	Volume per Reaction (µL)
10X FastStart High Fidelity Reaction Buffer without $MgCl_2$ (Roche)	1.0
25 mM $MgCl_2$ (Roche)	1.8
DMSO (Roche)	0.5
10 mM PCR Grade Nucleotide Mix (Roche)	0.2
5 U/ul FastStart High Fidelity Enzyme Blend (Roche)	0.1
PCR Certified Water (TEKnova)	0.4
Total MM	4
DNA	3.6
Primers MM	2.4
Total volume	10

PCR stages	Number of cycles
95°C 10 min	1
95°C 15 seconds 60°C 30 seconds 72°C 1 min	23
72°C 3 min	1
4°C	∞

Table 2.11B: Colony genotyping library preparation protocol PCR 2

Component	**Volume per Reaction (µL)**
10X FastStart High Fidelity Reaction Buffer without $MgCl_2$ (Roche)	1.0
25 mM $MgCl_2$ (Roche)	1.8
DMSO (Roche)	0.5
10 mM PCR Grade Nucleotide Mix (Roche)	0.2
5 U/ul FastStart High Fidelity Enzyme Blend (Roche)	0.1
PCR Certified Water (TEKnova)	3
Total MM	6.6
DNA	1
Primers MM	2.4
Total volume	10
PCR stages	Number of cycles
95°C 10 min	1
95°C 15 seconds 60°C 30 seconds 72°C 1 min	13
72°C 3 min	1
4°C	∞

Table 2.11C: Colony genotyping library preparation protocol:Barcoding PCR

Component	Volume per Reaction (µL)
10X FastStart High Fidelity Reaction Buffer without $MgCl_2$ (Roche)	1.0
25 mM $MgCl_2$ (Roche)	1.8
DMSO (Roche)	0.5
10 mM PCR Grade Nucleotide Mix (Roche)	0.2
5 U/ul FastStart High Fidelity Enzyme Blend (Roche)	0.1
PCR Certified Water (TEKnova)	3
Total MM	6.6
DNA	1
Barcodes	2.4
Total volume	10
PCR stages	Number of cycles
95°C 10 min	1
95°C 15 seconds 60°C 30 seconds 72°C 1 min	11
72°C 3 min	1
4°C	∞

B. Single cell genotyping of FACS-sorted cells

FACS-sorted single cells were processed using an optimized method for the detection of genomic DNA (gDNA) from single cells and a combined approach for genomic and cDNA amplification, as developed by DPhil student Alba Rodriguez Meira et al in the Mead lab. For the gDNA amplification method, single cells were sorted into 4 µL of lysis buffer containing 2.5% Protease (Qiagen, Manchester UK)

and 0.2% Triton X. Cells were lysed as described above and the entire lysate was used for downstream multiplex PCR, with custom patient specific primers. For PCR 1, targeted regions were amplified with SeqAMP enzyme (Clontech, California USA). 1 µL of the first PCR product was further amplified using nested, patient-specific primers with KAPA 2G Robust PCR (KK5702, KAPA Biosystems, UK). This product was further barcoded using the Access Array™ barcoded system, purified with Ampure XP beads and sequenced on a MiSeq (Illumina) as described above.

2.3.2.3 Data analysis

Reads were aligned to the human genome (GRCh37/hg19) using STAR with default settings (version 2.4.2a). Variant calling was performed using mpileup (samtools version 1.1, options -- minBQ 30, --count-orphans, --ignore overlaps) and results were summarized with a custom Python script. Downstream analysis of the variants and annotation of the mutant cells was performed with a custom R script. Each individual amplicon was analysed and samples with coverage <30 reads were excluded from downstream analysis. For each sequencing run, normal controls were also screened for the variants of interest in order to ensure no false positive samples were present. None of the mutations were detected in any of the control samples in any of the sequencing runs. Amplicons covering common heterozygous SNPs for each patient were used to calculate the allelic dropout rate of the assay. Figure 2.13 outlines the main steps used for sample analysis. Scripts used for bioinformatic analysis of the colony genotyping were developed by Alba Meira (DPhil student, Mead lab) Analysis of the quality control samples filtering and variant annotation was led by Alba Meira.

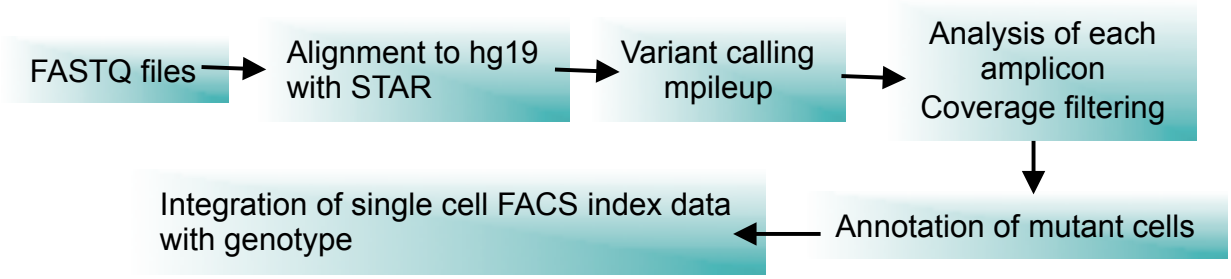


Figure 2.13: Single cell genotyping data analysis pipeline

2.3.3 Long-term culture-initiating cell assay protocol

The LTC-IC assay is regarded as the best in vitro assay of HSC in humans (118). This assay involves co-culture with stromal cell lines for 6 weeks followed by a read out assay in methylcellulose.

2.3.3.1 SL/SL and M210B4 stroma cell line culture and maintenance

Mouse stromal cell lines (M210B4 and SL/SL, both HyClone) were carefully thawed and transferred to 50 ml FALCON tubes containing 10 ml RPMI supplemented with 10% FBS and DMEM supplemented with 15% FBS, respectively. Cells were washed at 400g for 10 min and suspended in 12 ml of medium: RPMI + 10% FCS+100U/ml Penicillin/ Streptomycin+ 2mM L-glutamine $+10^{-4}$ M 2-mercaptoethanol for the M210B4 stroma cell line and DMEM with high Glucose, without L-glutamine + 15% FCS+ 100U/ml Penicillin +100U/ml Streptomycin(final concentration) + 2mM L-glutamine (final concentration) $+10^{-4}$ M 2- mercaptoethanol (final concentration). Cells were cultured at 37°C with 5% CO2 for 1 week. Following one week in culture, cell lines were inspected by direct microscopy, the medium was removed and the cells were washed once with 5ml of PBS before incubating them with 2 ml 0.5% trypsin/EDTA (Invitrogen, Paisley UK) at room temperature. Trypsin-EDTA should

also be at room temperature prior to incubation. Trypsin was spread evenly over the cells, then removed and trypsin and cells were incubated for 5 minutes at 37°C before checking to ensure that trypsin has detached the cells from flask surface. Cells from each line were collected and resuspended in 10mL of RPMI+10%FCS medium (M210B4 cells) and 10mL DMEM+15%FCS (SL/SL cells). Cell lines were split 1:10 and resuspended in 12 ml of the of respective medium. Cell lines should not be kept longer than for 20 passages. After 7 days in culture antibiotic selection was performed with hygromycin (Invitrogen cat number 10687010) and G418 (Geneticin, Invitrogen catalog number 11811-031), as shown in Table 2.12.

Table 2.12: Antibiotic selection of M210B4 and SL/SL cells for LTC-IC cultures

Cells	Medium	G418 Final concentration	G418 Amount added	Hygromycin Final concentration	Hygromycin Amount added
M210B4	20 ml RPMI+10%FCS	0.3 mg/ml	120 ul @50mg/ml	0.06 mg/ml	24 ul @50mg/ml
SL/SL	20 ml DMEM+15%FCS	0.6 mg/ml	240 ul @50mg/ml	0.125mg/ml	50 ul @50mg/ml

2.3.3.2 Establishing co-cultures of M210B4-SL/SL cells with HSPC

24 hours before FACS-sorting M210B4 and SL/SL cells were collected following trypsinisation (as outlined above) and resuspended in 10 ml of the respective medium. Cells were washed at 350g and resuspended in 2ml IMIDM+2% FBS in sterile FACS tubes. Each FACS tube was placed in a 50ml FALCON tube and cells were irradiated at 80 Gy for 15 min. After irradiation, cells were transferred into 20ml IMDM+20% FBS and washed at 350g for 10 min. Supernatant was removed and cells from each cell line were counted using trypan blue. 5000 cells from each cell line in a total volume of 100 ul/well were added to each well of collagen-coated 96-

well plates. Collagen-coated plates were prepared in house using collagen (Stemcell technologies, Vancouver, Canada, catalog number 04902) and were left at 37°C to dry before applying the stromal cell lines. HSPC from JMML patient samples and cord blood controls were isolated by FACS as described above and 100 cells/well were seeded in H5100 Myelocult medium (Stemcell Technologies) supplemented with 10^{-6} hydrocortisone 21-hemisuccinate (Stemcell Technologies) and co-cultured with SL/SL and M210B4 stroma cells. Sorted populations were cultured for 6 weeks, with half medium changes once weekly. Individual wells were harvested after 6 weeks, transferred to Methocult H4435 and colony output evaluated after 2 weeks under direct light microscopy (Figure 2.14). Cytospins of single picked colonies were also used to further classify the clonogenic output.

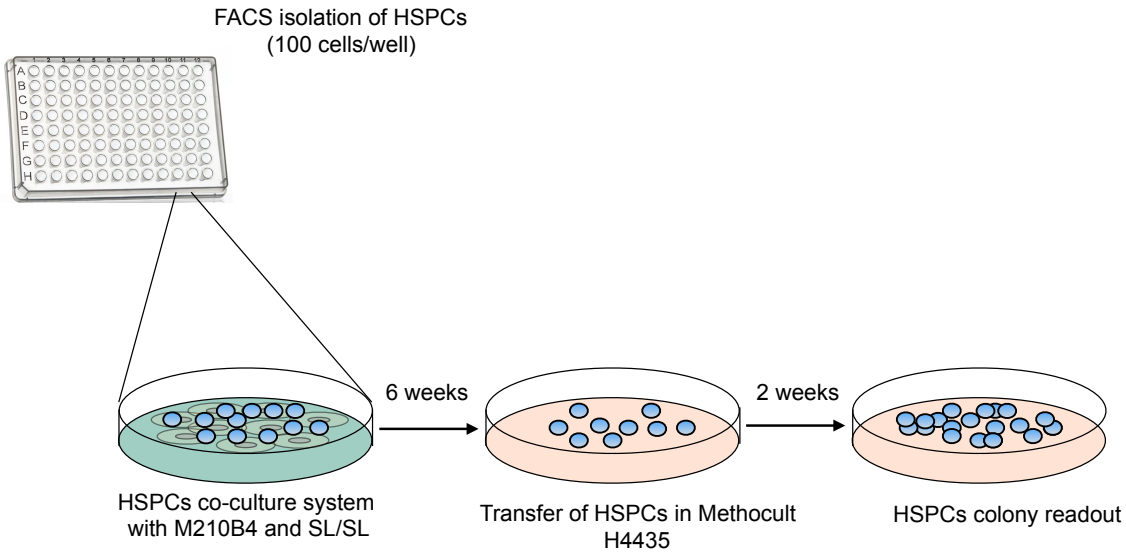


Figure 2.14: Long-term culture-initiating cell assay

2.3.4 In vivo stem cell assays

2.3.4.1 Xenotransplantation protocol

HSC, GMP and +/+ cells were transplanted into NSG (NOD.Cg-Prkdcscid Il2rgtm1Wjl/SzJ), following sublethal irradiation (with two doses of ~1.25Gy delivered 4 hours apart). NSG mice were obtained through the Jackson Laboratory and maintained in individually ventilated cages in a regulated pathogen-free facility. All protocols for in vivo experiments were approved by the UK Home Office Project licence (PPL 30-3103) assigned to Benjamin Povinelli. NSG mice were bred at the Oxford Biomedical Services. 10-14 week old male or female mice were used for these experiments. Prior to the procedure mice were given Baytril in drinking water (1.2mls of Baytril per 250 mls of water), SMA gold baby powder (dry and wet) and transwean. Mice were only used if weight was more than 20grams. HSPC were FACS-sorted in 200ul of PBS+1%FBS and injected into the tail vein 6-8 hours after the last radiation dose as previously described (118). Human cell reconstitution and lineage distribution was monitored by flow cytometry in the peripheral blood at 5, 10,16 or 22 weeks post-transplant. For terminal readouts mice were sacrificed and flow cytometry analysis of BM, spleen and peripheral blood was performed.(See section 2.3.42 and 2.3.4.3 for details and Table 2.13 and 2.14 for the antibodies used for the human cell reconstitution, stem cell and lineage output). MNC from BM and spleen samples were cryopreserved and used for secondary transplantation. Mice were monitored for signs of deteriorating health and weight loss and euthanized if they appeared unwell. Leukaemia was confirmed by the presence of leukaemic cells in the peripheral blood, BM and/or spleen. Femurs were preserved and fixed in neutral buffered formalin and embedded in paraffin for sectioning and staining according to standard procedures.

Irradiation, transplantation, monitoring of the mice at the mouse facility and tissue collection for readouts was performed by Dr Chris Booth and Dr Benjamin Povinelli, as my funding body did not allow to perform direct animal work. Downstream cells processing for cryopreservation, flow cytometry and cell sorting for secondary transplants was performed by myself.

2.3.4.2 Peripheral blood MNC isolation and analysis

Peripheral blood read outs were performed at week 5, 10, 16-22 post transplantation. 9ul of peripheral blood was analysed using a Sysmex counter in order to measure the haemoglobin and numbers of white cells and platelets. The remaining peripheral blood sample was diluted 1:1 with PBS and red cells were lysed for 20 min with ammonium chloride (Stemcell technologies, Vancouver Canada) at 37°C. The supernatant was collected, cells were washed at 350g for 5 min and then incubated for 1-2 min with Dextran (Sigma-Aldrich), mouse and human Fc block was added and the cells were stained for flow cytometry as described above. The antibody panel used for peripheral blood readout analysis is shown in Table 2.13.

Table 2.13: Antibody panels used for peripheral blood readouts in xenograft studies

Antibody	Fluorochrome
human CD45	af700
mice CD45	PE Texas Red
CD19	PE
CD34	ef780
CD235a	PerCyp5.5
CD15	FITC
CD33	FITC
CD66b	FITC
CD11b	FITC
CD14	FITC
DAPI	BV421

2.3.4.3 BM and spleen MNC isolation and analysis

Cells were harvested from bones or spleen using a pestle and mortar. The harvested material was resuspended in 5 ml PBS+1%FBS and filtered through a strainer. Bones and spleen were crushed once more, resuspended in another 5 mls of PBS+1%FBS and passed through the same cell strainer. Cells were washed in PBS+1%FBS, counted using trypan blue and analysed by flow cytometry using the antibody panel shown in Table 10 and/or cryopreserved for secondary xenotransplantation experiments. Fc block was added prior to staining. Splenic cells were routinely analysed with both Table 9 and 10 panels and BM samples were analysed using the panel shown in Table 2.14.

Table 2.14: Antibody panel used for BM and Spleen readouts from transplanted NSG mice

Antibody	Fluorochrome
CD38	PE Texas Red
CD34	APC ef 780
CD90	BV421
CD45RA	PE
CD123	PE Cy7
7-AAD (aminoactinomycin D) viability	
Human CD45	AF700
Mouse CD45	BV510
Lineage panel CD8 CD20 CD2 CD3 CD16 CD19 CD235a CD66b CD10 CD127	FITC

2.3.5 RNA sequencing

2.3.5.1 Bulk RNA sequencing cDNA library preparation/ Taqmentation/Sequencing

FACS-purified HSPC populations were processed for RNA seq using the Smart Seq2 protocol as previously described. (312) Fifty cells were FACS-sorted directly into 4 µl lysis buffer containing 0.2% Triton X-100 (Sigma-Aldrich), RNase inhibitor (Clontech), 2.5 mM dNTPs (Thermo Fisher) and 2.5 µM oligo-dT30VN primer (Biomers.net). Cells were stored at -80°C for a maximum of 2 weeks before library preparation. For library preparation cells were initially lysed at 72°C for 3 min. cDNA was generated using SuperScript II (Invitrogen) as shown in Table 2.15 and was pre-amplified using KAPA HiFi HotStart ReadyMix (KAPA Biosystems) using 19 cycles of amplification (Table 2.16). After PCR amplification, the cDNA libraries were purified with AMPure XP beads (Beckman Coulter) using a ratio of 0.8:1 beads to cDNA, according to the manufacturer's instructions. Post-purification libraries were resuspended in EB buffer (Qiagen, UK). The quality of cDNA traces was assessed by using a High Sensitivity DNA Kit in a Bioanalyzer instrument (Agilent Technologies). Tagmentation and library preparation was performed using the Nextera XT DNA Library Preparation Kit (Illumina) according to the manufacturer's instructions but using 1/4 of the volumes as previously described (313). Samples were sequenced using the Illumina NextSeq 500 platform, generating 75 bp paired-end reads.

Table 2.15: Reverse transcription reagents and conditions

Component	Volume(ul)	Final concentration
SuperScript II reverse transcriptase (200 U l^{-1})	0.50	100 U
RNAse inhibitor (40 U l^{-1})	0.25	10U
Superscript II first-strand buffer (5×)	2.00	10U
DTT (100 mM)	0.5	5mM
Betaine (5 M)	2.00	1 M
$MgCl_2$ (1 M)	0.06	6 mM
TSO (100 M)	0.10	1μM
Nuclease-free water	0.29	
Samples	4.3	
Total volume	10	
RT reaction conditions	1 cycle	42°C 90 min
	11 cycles	50°C 2 min 42°C 2 min
	1 cycle	70°C 15 min
	1 cycle	4°C hold

Table 2.16: PCR preamplification reagents and conditions

Component	Volume(ul)	Final concentration
First-strand reaction	10	
KAPA HiFi HotStart ReadyMix (2×)	12.50	1X
IS PCR primers (10 M)	0.25	0.1 μM
Nuclease-free water	2.25	
PCR conditions	1 cycle	98°C 3 min
	19 cyles	98°C 20 min 67°C 15 s 72°C 6 min
	1 cycle	72°C 5 min
	1 cycle	4°C hold

Bulk RNA library preparation and sequencing was performed by myself under the supervision and help from Dr Benjamin Povinelli.

2.3.5.2 Bulk RNA sequencing data analysis

Data analysis was performed with in house bioinformatics pipelines. Briefly, RNA-sequencing reads from bulk population were aligned to the human genome (GRCh38) using STAR (version 2.4.2a). Aligned reads were counted using featureCounts (version 1.4.5-p1) summarized to gene identifiers from the GRCh38 Ensembl transcriptome. Differential gene expression was done using the DESeq2

package (1.16.1) in R (3.4.1). Plotting expression values of genes was done by calculating transcripts per million (TPM) and plotted in R. Significant differentially expressed genes were selected on the basis of P value < 0.05.Gene-set enrichment analysis was performed using GSEA software (v2.20, http://www.broadinstitute.org/gsea) against Hallmark and curated HSC and cell cycle gene sets obtained through MSigDB using the default settings. Heatmaps were generated using the package pheatmap (1.0.8). Bulk RNA sequencing data analysis was performed by myself with the assistance and training from Dr Benjamin Povinelli. Bioinformatic analysis pipeline of bulk RNA sequencing was developed and led by Dr Benjamin Povinelli.

2.3.5.3 Single cell RNA sequencing library preparation/sequencing

The 10X Chromium platform was used to obtain unbiased RNA sequencing data from JMML HSPC compared to normal control HSPC. To do this 18,333-24,000 Lin-CD34+ single cells from two JMML patient samples (Patient ID1 and ID5) and 2 normal cord blood controls, were sorted and processed for RNA seq as per the manufacturer's instructions. In short, cells were sorted into a total volume of 40 µL and 33.8 µL of the sample was loaded onto the Single Cell 3' Chromium 10X Chip with 66.2 µL of the 10X master mix. cDNA was generated on the Chromium 10X controller. Post RT clean up was performed as per the manufacturer's instructions and the product was amplified with 8 PCR cycles. Post amplification clean up and library construction was performed as per the manufacturer's instructions and samples were sequenced on the Illumina HiSeq4000, generating 75 bp paired-end reads. Library preparation using the Chromium 10 X chromium platform was performed by the WIMM single cell facility.

2.3.5.4 Single cell RNA sequencing data analysis

Reads from 10x genomics single cell RNA-sequencing were aligned and counted using cell ranger (2.0.0) using the hg19 human genome. Unique molecular identifier (UMI) counts were analyzed using Seurat (2.0.1). Genes were filtered by minimum number of cells expressing = 10, greater than 0.15 mean UMI counts, less than 4 mean UMI counts, and greater than 0.5 dispersion. Cells were filtered by greater than 1000 genes detected, less that 4000 genes detected, and less than 5% mitochondrial reads. After filtering 1,127 variable genes and 17,547 cells were used for analysis. tSNE analysis was performed after selecting the most variable principle components by use of the jackStrawPlot feature in Seurat. Clusters were identified by the k-shared nearest neighbor (SSN) approach FindClusters. Figures for single cell analysis were generated in R. Top differentially expressed genes from SSN identified clusters were calculated using FindAllMarkers function in Seurat. Analysis of the single cell RNA seq data was performed by Dr Benjamin Povinelli.

2.3.6 Quantification and statistical analysis

Statistical analyses were performed using GraphPad Prism v7.00 or R v3.2.2. Survival curves were compared using the log-rank test. T-tests, Mann Whitney tests and ANOVA followed by multiple comparisons testing were used to compare experimental groups as indicated in the figure legends.

Chapter 3: The NIHR PAEDIATRIC MDS/JMML STUDY

3. THE NIHR PAEDIATRIC MDS/JMML STUDY

3.1 Background and purpose of the study

Paediatric myelodysplastic syndromes are extremely rare pre-malignant conditions of childhood that are very challenging to manage and have a variable risk of developing AML, which to this date is not very clearly documented (5) .The rarity and lack of comprehensive observational studies on these diseases means that there are no evidence-based diagnostic and management algorithms and this, in turn, may lead to a delayed diagnosis and/or inappropriate management (136). In many cases, the clinical features may be absent or very subtle leading to difficulties in identifying such cases unless molecular testing is used. In the UK, prior to this study, there was no recent information bank or registry for these diseases which systematically documented specific clinical features aligned to a molecular diagnosis, although important work two decades ago had established the prevalence of these disorders in the UK before molecular testing was available (7),(15), (314),(157). Even now, targeted NGS platforms are available in very few paediatric haematology tertiary centres and, prior to my study, there were no targeted panels that specifically screened for the Paediatric MDS/JMML-associated mutations.

As discussed in chapter 1, the molecular landscape of paediatric MDS and JMML is very different from adult MDS and the panels used for adult myeloid malignancies are not appropriate for screening paediatric cases, as a great number of pathogenic variants found in paediatric populations are not included. In the majority of centres, the diagnosis of paediatric MDS/JMML has primarily been based on the peripheral blood counts and peripheral blood (PB) and bone marrow (BM) morphology.

However, BM morphology can be very challenging for these cases, particularly given the overlap with other conditions that present with hypocellular BM such as severe aplastic anaemia and inherited BM failure syndromes (17). In addition 30-40% of cases carry germline mutations which indicates that the impact of misdiagnosing and mismanaging a child can have a direct effect on the management of a whole family.

Bone Marrow transplantation (BMT) is currently the only curative treatment for these conditions (32). JMML cases receive transplantation as soon as possible but for MDS cases this is reserved for patients with significant cytopenias, abnormal cytogenetic studies and severe clinical manifestations while other cases are managed with a 'watch and wait' approach (32). However, lack of knowledge of the natural history of the disease means that there are currently no clinical means of identifying the optimum time for BMT for the cases that do not immediately fulfill the criteria above. Similarly, for JMML, despite increasing knowledge of the molecular landscape of the disease, post-BMT relapse remains the most common cause of treatment failure. Thus, studying of the natural history of JMML and other forms of paediatric MDS should greatly aid the identification, monitoring and treatment of children with these disorders.

3.2 Aims

The overall aim of the NIHR study is to systematically document the clinical presentation, clinical features and molecular abnormalities in paediatric MDS / JMML patients in order to help create evidence-based diagnostic and risk stratification algorithms. Longitudinal follow up and collection of serial patient samples will allow the documentation of the natural history of these diseases,

survival and treatment response, and will provide insights into the correct timing of treatment interventions for this patient group. The observation period of the study is set to be 10 years, which is beyond the timescale of my DPhil. For the duration of my project I aimed to establish the foundations of the study that will allow the achievement of the goals above. My primary objectives in establishing this study, and described in this chapter, were:

1. To create a database that allows collection of deep phenotypic data in order to accurately identify, define and classify the phenotypic characteristics of Paediatric MDS/JMML cases and identify any clinical characteristics that will allow better classification of patients
2. To develop a customised paediatric MDS-specific gene panel to identify the underlying driver mutations of these diseases
3. To perform targeted NGS in serial patient samples to identify disease clonal evolution and monitoring of the response to treatment
4. For cases with well-defined phenotypic characteristics where known mutations have not been identified by targeted NGS, to perform whole genome sequencing (WGS) to identify any novel mutations that are associated with paediatric MDS/JMML.
5. If time allows, to develop phenotype:genotype correlations that may be useful for the development of diagnostic and risk stratification algorithms

3.3 Demographic and clinical characteristics of children recruited into the NIHR Paediatric MDS/JMML study

A total of 68 patients were enrolled into the study between June 2015 and September 2017, following successful ethical approval of the study on 26th of May 2015. Patient recruitment was based on the local clinician's discretion and consultation with the research team at the time of recruitment. PB and BM was collected from all recruited patients and sent to the Weatherall Institute of Molecular Medicine for further analysis. Mononuclear cells (MNC) were isolated and cryopreserved and a small aliquot from each sample was used for DNA isolation as described in chapter 2. Deep phenotypic data were collected from each patient were collected and stored in an online password protected database. All phenotypic information as well as BM aspirate slides and trephines were centrally reviewed by Dr Anupama Rao (study Chief Investigator) and myself. One patient was excluded from the study and downstream analysis following central review of the clinical information and BM morphology which showed that they did not fulfill the study eligibility criteria.

3.3.1 Demographic data

The remaining 67 patients were retrospectively divided into 4 groups following detailed review of the clinical and molecular information collected through the study and largely based on the international Paediatric MDS Classification criteria (5): JMML (n=16), MDS (n=17), Idiopathic Cytopenias of Uncertain Significance (ICUS, n=32) and inherited BM failure syndromes (n=2) (Figure 3.1A). The ICUS group has not been previously described in children and the definition was 'borrowed' from the adult diagnostic definition (315, 316) in order to describe the cohort of children with significant sustained cytopenias with mild BM dysplasia and no other cause identified

for these abnormalities. Our research group (AR, IR, AM, EL) felt that these children represented a distinct group of conditions which often lack cytogenetic abnormalities but are characterised with milder dysplasia (<10% dysplasia, as in adult ICUS) than seen in de novo MDS (>10% dysplastic changes). The diagnosis of an inherited BM failure syndrome was made after molecular analysis and was not known at the time of recruitment- in each case the diagnosis was Schwachman-Diamond syndrome (SBDS). The number of cases of JMML and MDS was higher than expected based on previous population based studies (7),(157). This may be due to missed cases because of the lack of robust molecular diagnostics at the time. The mean age at presentation of the JMML group was 2.1 years, as in previous reports, while the MDS and ICUS groups were older with a mean age of presentation of 5.5 and 5.33 years respectively . The 2 children found to have an inherited BM failure syndrome presented at the age of 5 years and 2 months of age respectively. In line with previous reports, the majority of children with JMML were male while the gender distribution of the MDS group was similar, with a slight female predominance (6) and the majority of children with ICUS were female (Figure 3.1B).

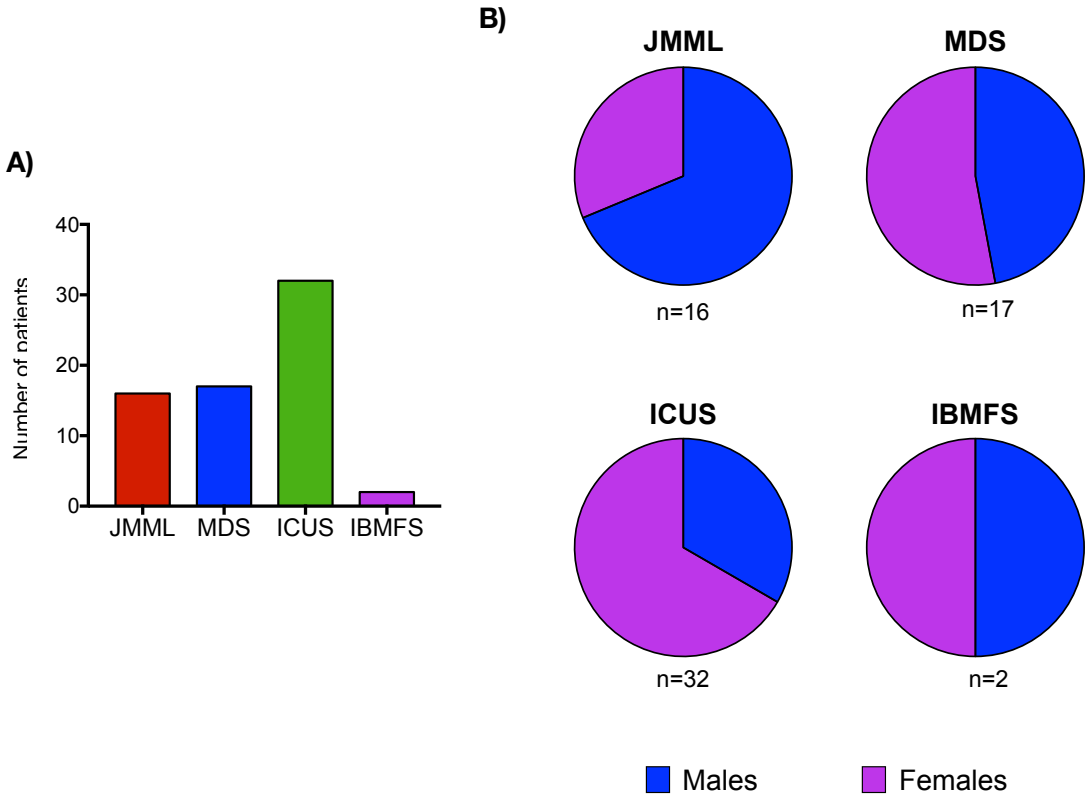


Figure 3.1: A) Bar charts showing the number of patients in each group .B) pie charts showing their gender distribution for each group.

3.3.2 Clinical characteristics

The predominant presenting feature of children with JMML was monocytosis with the exception of one patient who had a relatively low monocyte count ($0.5x10^9$ /l) at presentation and originally presented with thrombocytopenia. That this patient still fulfilled the JMML diagnostic criteria, was also verified by the European Paediatric MDS/JMML working group (EWOG-MDS). The majority of patients with MDS (47.1%) presented with bi-lineage cytopenias (Figure 3.2A) followed by anaemia (23.5%), neutropenia (11.8%) and pancytopenia (17.6%). The majority of patients with ICUS (51.9%) had sustained pancytopenia of unknown cause, neutropenia (25.9%), thrombocytopenia (7.4%) or bi-lineage cytopenia (11.1%). From the two patients

found to have SDS, one was neutropenic (and had a prior diagnosis of autoimmune neutropenia at 14 years of age) while the second patient was pancytopenic with developmental delay and dysmorphic features but without any established diagnosis; neither of these patients had any relevant family history. The overall distribution of positive family history and syndromic features was similar across the different groups (Figure 3.2B,C).

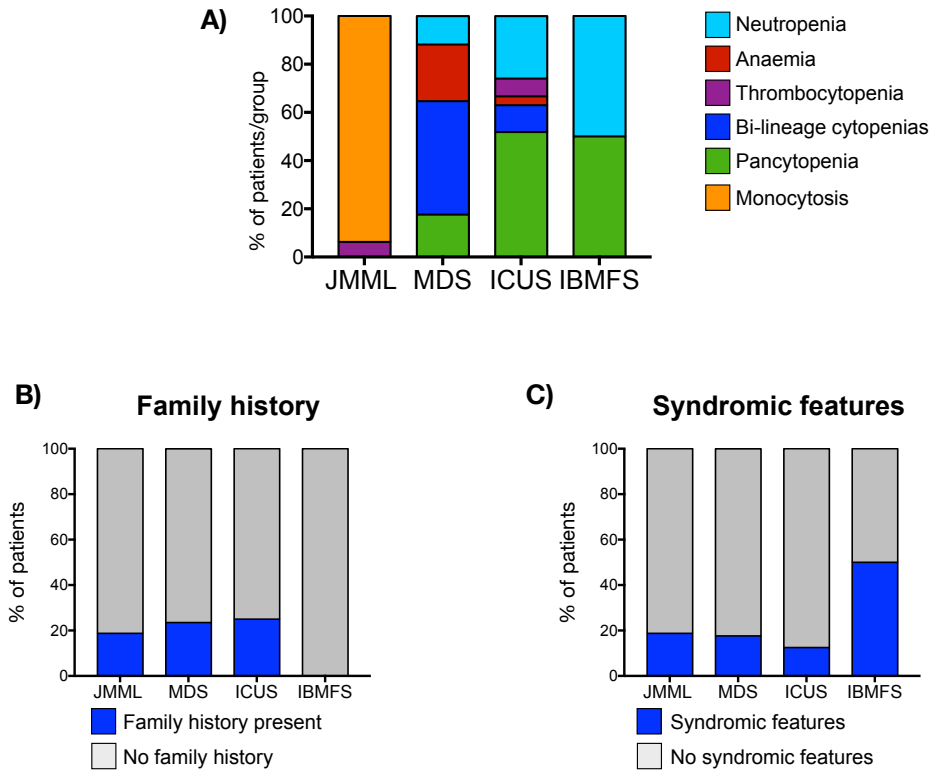


Figure 3.2: A) Main haematological presenting feature for each of the groups. B) Proportion of patients that have any relevant family history, blue bars indicate presence of family history C) Proportion of patients that had syndromic features , blue bars indicate syndromic features.

The frequency of severe or moderately severe infections was highest in the ICUS group (44.4%) as well as being a feature of both patients with SDS, as expected (Figure 3.3A). Interestingly, infection was not a feature of JMML in this series. Failure

to thrive (FTT) was noted in a similar proportion of patients with JMML (25%), MDS (23.5%) and ICUS (18.5%) as well as one of the patients with SDS (Figure 3.3B). It is difficult to assess whether the degree of cytopenias was the primary attribute for the lack of weight gain observed, as multiple other factors such as constitutional growth delay and environmental factors also can also contribute to this. Regarding treatment, around half of the patients with JMML, ICUS and MDS required transfusions (Figure 3.3C). Neither of the 2 patients with SDS were transfused but one was started on treatment with G-CSF. All but 2 patients with JMML were treated by BMT (one with a heterozygous *NRAS* and the other with a heterozygous *CBL* mutation).

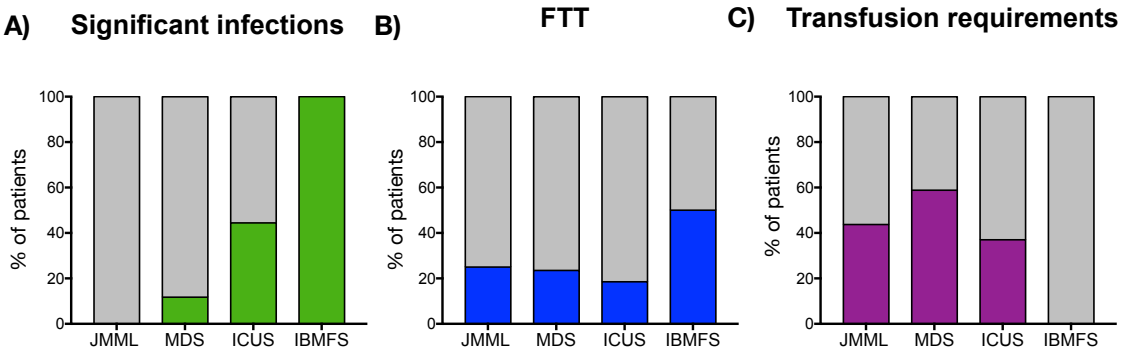


Figure 3.3: A) Percentage of patients with significant infections indicated by the green bars. C) Percentage of patients with growth failure indicated by the blue bars D) Percentage of patients with transfusion requirements indicated by the purple bars

3.3.3 Haematological and cytogenetic characteristics

As expected patients with JMML had a significantly higher total white blood cell (WBC) count compared to the other groups (Figure 3.4A). There was no significant difference in Hb concentrations between the 4 groups although only the JMML patients had a mean Hb concentration below 100 g/l (87.13 SEM g/l; Figure 3.4B), There was no difference in platelet count between the groups apart from the JMML

patients who had significantly lower counts than the ICUS patients and all of the patient groups apart from those with ICUS had a mean platelet count below the normal range (150-400 x 10^9/l; Figure 3.4C). Thrombocytopenia (defined as platelet count < 150 X10^9/l was a presenting feature in 81.25% of JMML patients, 76.47% of those with MDS and 66% of those with ICUS as well as one of the patients with SDS. The neutrophil count was significantly lower in the ICUS group compared with the rest of groups with a mean neutrophil count of 1.1 X10^9/l (Figure 3.4D). Only one JMML patient; 29,4% of MDS; 80% of the ICUS group and both of the IBMFS patients were found to be neutropenic (defined as neutrophil count < 1 X10^9/l). The MCV was increased in the majority of patients; interpretation of the results depends on the age of the patients and I have therefore shown the data for MCV as the proportion of patients in each group with an MCV above the normal range for age in each group (Figure 3.4E).

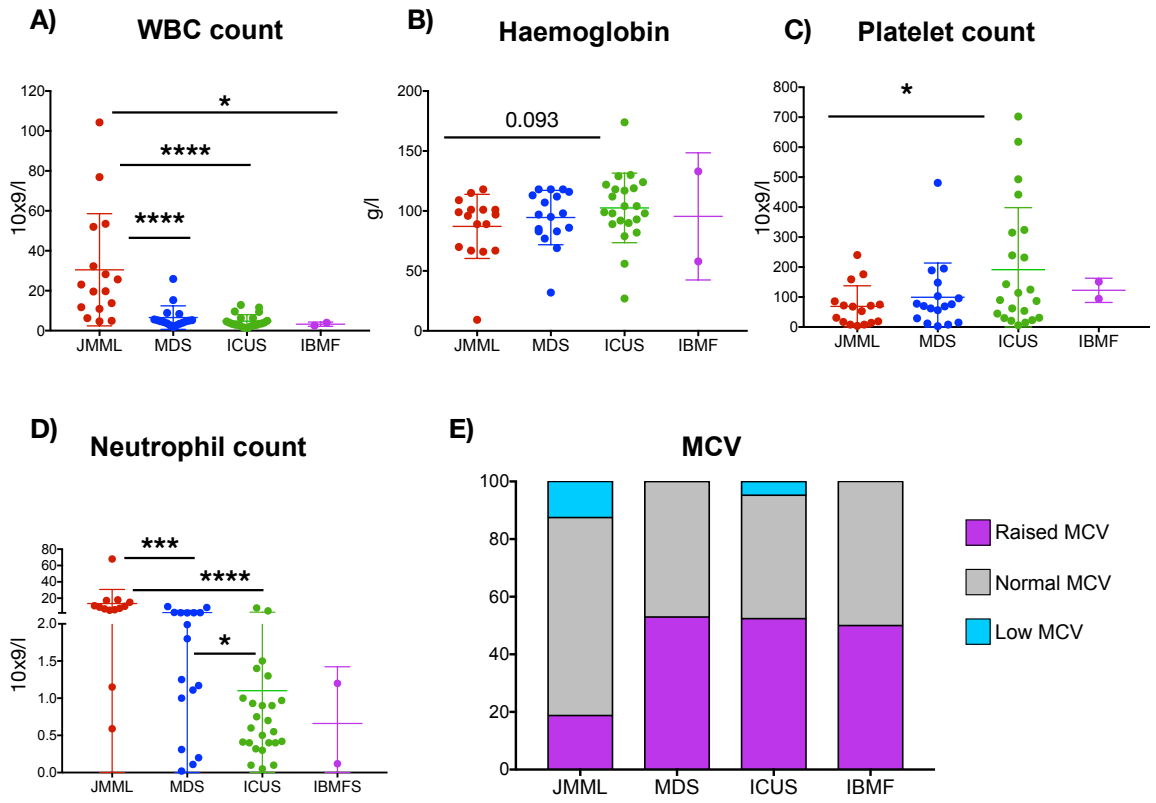


Figure 3.4: The presenting A) WBC count B) Haemoglobin concentration C) platelet count and D) neutrophil count for the different groups. E) Bar graph showing the proportion of each group with an increased MCV for age. * indicates significance p< 0.05

The commonest cytogenetic abnormality detected within this patient group was monosomy 7 which was seen in 1 patient with JMML, 4 patients (23.5%) with MDS, none of the patients with ICUS and one of the 2 patients with SDS (Figure 3.5A). There were no other cytogenetic abnormalities seen in the JMML group or in SDS but 17.6% patients with MDS and 3.12% patients with ICUS had other cytogenetic abnormalities which included *MLL* rearrangement and *RUNX1* triploidy. Correlation of molecular and cytogenetic abnormalities are indicated in Figure 3.7. Overall BM cellularity was assessed as hypercellular in all of the patients with JMML and hypocellular in the vast majority of patients with ICUS (87%) as well as both patients

with SDS (Figure 3.5B). BM cellularity was variable in the MDS patient: hypercellular in 35.2% and hypocellular in 58.82%, as per previous reports(1) .

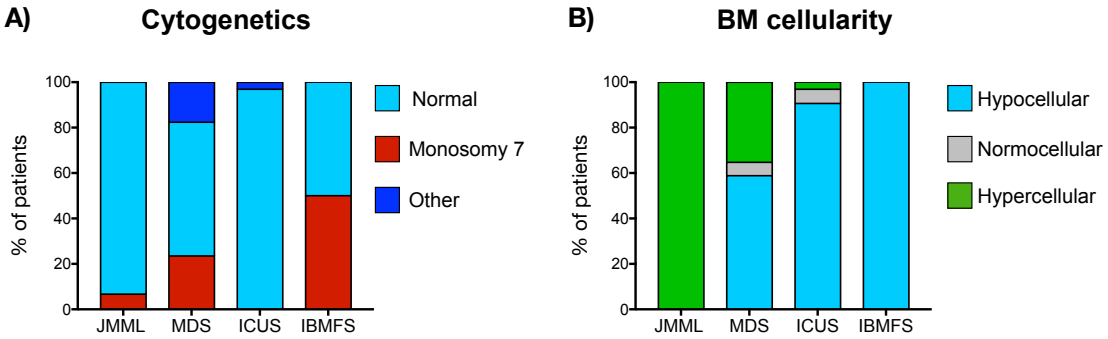


Figure 3.5: A) Cytogenetic abnormities and B) BM cellularity in each group

3.3.4 Sub-classification of the 17 patients with MDS

The MDS group was further divided into four different subgroups based on the current paediatric MDS definitions described in chapter 1. The largest group was classified as Refractory Cytopenia (RC 47%), with 5 cases of MDS with Excess Blasts (MDS-EB 29.4%); 1 case of MDS with Excess Blasts in Transformation (MDS-EBt) and 3 cases of treatment-related MDS (MDS-Tx 18%) (Figure 3.6A). Monosomy 7 was observed in 37% of RC cases and 20% of MDS-EB cases (Figure 3.6B). One patient with RC originally had normal cytogenetics but later developed trisomy 6. More severe cytopenias were observed in the patient with MDS-EBt but the differences were not significant due to the low number of patients per group.

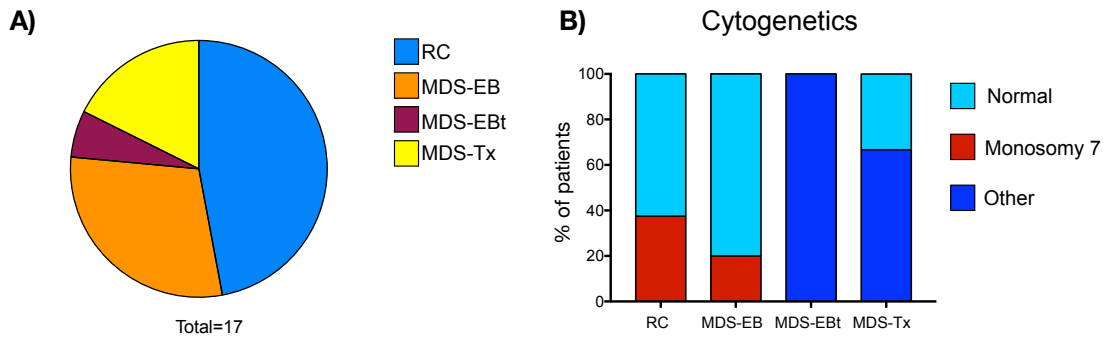


Figure 3.6: A) Pie chart of the different MDS subgroups observed in the NIHR cohort. B) Cytogenetic abnormalities observed in each different subgroup.

3.4 Molecular characterisation of the disease-associated mutations using targeted NGS or WGS.

3.4.1 Identification of pathogenic variants using targeted NGS

DNA from BM (when available) or PB samples from all cases recruited to the project was used for targeted NGS using the Paediatric MDS targeted NGS panel I designed, as described in chapter 2. All positive results from the panel were validated with Sanger sequencing in the Great Ormond Street diagnostic lab and a report was issued to the responsible clinician. The heatmap in Figure 3.7 demonstrates the mutations detected at diagnosis for all of the samples sequenced.

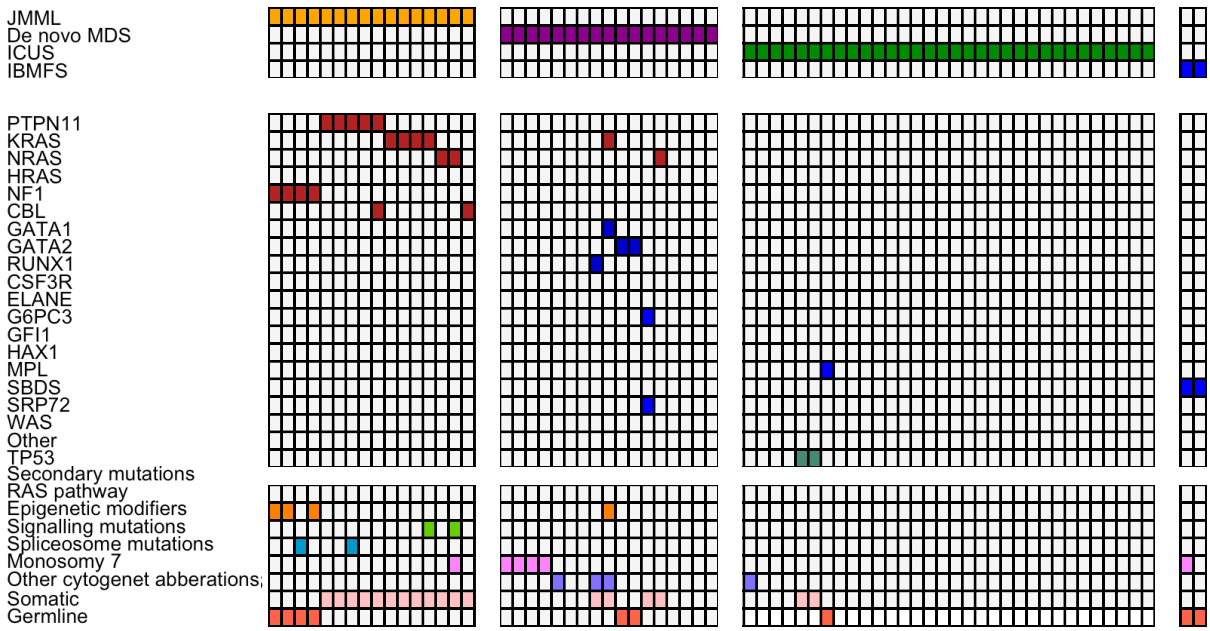


Figure 3.7: Targeted NGS panel results. Each column represents a patient. Coloured boxes represent mutations present. White boxes represent negative results.

All JMML patients screened with the panel were found to have a mutation and this was always a *RAS* pathway mutation, consistent with previous reports (193), (172). Patients with *NF1* were found to have germline mutations whereas the rest of the patients had somatic variants. By contrast, pathogenic mutations in known MDS/JMML-associated genes were found in only 35% of the MDS group and 9% of the ICUS (Figure 3.7 and 3.8A). The majority of mutations detected within the MDS group were mutations in transcription factor genes (TF) such as *GATA1, GATA2* or *RUNX1.* One RC patient was found to have mutations in *SRP72* and *G6PC3.* This patient subsequently progressed and acquired another cytogenetic abnormality with trisomy 6. The patient with MDS-EBt was found to have three different mutations in *GATA1*, *KRAS* and *EZH2*; whereas the MDS-EB patient had *GATA2* and *NRAS* mutations. The findings within the MDS group are consistent with previous reports and as expected the patients with the *GATA2* variants were the patients with the germline mutations; however, certain genes newly associated with paediatric MDS,

such as *SAMD9* and *SAMLD9*, were not linked with paediatric MDS at the time that the panel targets were designed and therefore were not included in the panel (102). Of the small number of mutations identified within the ICUS group, 2 non-related patients had *TP53* mutations and one patient had an *MPL* variant. The 2 patients with SDS had bi-allelic mutations in the *SBDS* gene and, interestingly, neither of these patients any gastroenterological manifestations of the disease and therefore SDS was not previously considered as a possible diagnosis. Secondary mutations were primarily detected in JMML (50%) and MDS (12%) patients (Figure 3.8B).

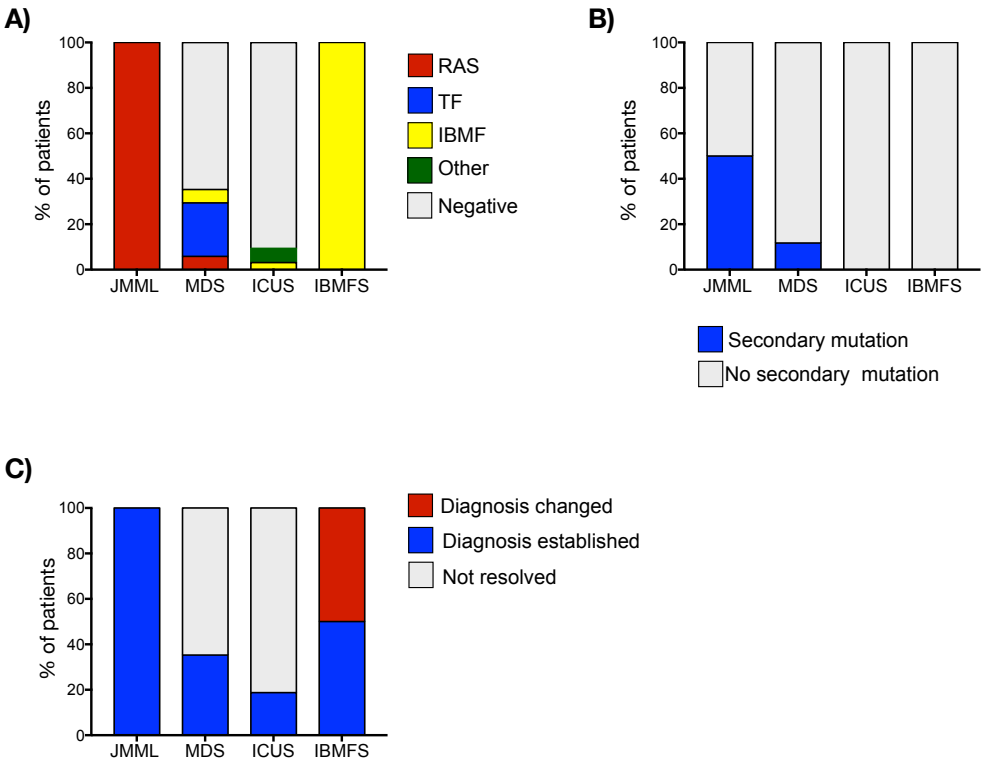


Figure 3.8: A) Different categories of disease initiating mutations detected in each patient group (*RAS,* indicates RAS pathway mutations, TF, mutations found in transcription factor genes, IBMFS, genes involved in inherited BM failure syndromes. B) % of patients where secondary mutations were detected. C) % of patients where a diagnosis was established or changed.* Analysis of whole genome sequencing is still on going beyond this project.

3.4.2 Identification of pathogenic variants using WGS

In order to further investigate cases where the NGS panel had failed to reveal a pathogenic variant, a small number were selected for WGS through the NIHR BioResource Stem and Myeloid Disorders project following additional informed consent from the families. In total 5 ICUS and 4 MDS cases were submitted for further genomic analysis. In addition, 3 JMML cases were also submitted in order to screen for mutations not previously described in JMML. Sequencing of samples was performed by the NIHR Bioresource team and data were uploaded on the Sapientia virtual tool, where variants were visualized and discussed on a national white cell MDT panel organized by myself and led by Prof Irene Roberts. The samples were initially screened for mutations and abnormal alterations focusing on 160 genes that have been previously implicated in stem cell and myeloid disorders. The virtual panel of genes to be used for this analysis was designed by myself as well as the analysis of the possible causative variants for each patient. This led to the identification of pathogenic variants in 3/5 ICUS cases: one patient with long standing neutropenias and gastroenterological granulomas had bi-allelic *CSF3R* mutations, and two further patients with pancytopenias one of which had a pathogenic *MECOM* variant and the third patient had a mitochondrial DNA deletion consistent with Pearson syndrome. The patient with the *CSF3R* mutations had 2 not previously described CSF3R mutations located at the extracellular part of the receptor. A non sense mutation (c.340 C>T) which introduced a premature STOP codon (p.Gln114*) at the IgG –like domain of the receptor which mediates the interaction with G-CSF(317); and a splice site variant (c.1576+1 G>A), proximal to the transmembrane domain of the receptor which is likely to introduce a skip of exon 12 (318). This patient was initially managed on G-CSF due to recurrent infections, mouth ulcers and gastroenterological

complications and poor weight gain and had poor response to G-CSF. Mild bone marrow dysplasia developed at 17 years of age and G-CSF was recently discontinued due to the poor response and the recognized risk of leukaemic transformation (319). Leukaemia risk is more recognized in cases with intracellular CSF3R mutations, however this patient was put under close follow up and annual bone marrow sampling (320).

The second patient had a non-sense c114 C>T *MECOM* variant that is predicted to cause a premature stop codon p.(Gln 372*). This patient had dysmorphic features strong family history of bone marrow failure of childhood requiring bone marrow transplantation. This sample failed to be analysed for extended whole genome sequencing screen, however key regions were specifically targeted and the *MECOM* variant was also found in the other affected family members which deemed the mutation to be pathogenic.

No pathogenic variants have so far been identified in the 4 MDS cases. Similarly, no additional mutations were identified in the JMML patient samples sent for WGS. All negative cases were also screened for any structural variants with no positive results to date, however, the analysis of these cases will still continue for discovery of other coding or non coding variants. It is important to note that germline material was not available from these patients which could have potentially limited the diagnostic yield of the cases investigated. Overall, a molecular diagnosis was established in 35% of the MDS patients and in 12.5% of the ICUS cases (Figure 3.8C).

3.4.3 Phenotype-genotype correlations in JMML

Due to the small number of identified molecular abnormalities in the MDS and ICUS patients, phenotype:genotype correlations were only possible for JMML. Of the16 JMML cases, 5 patients had a *PTPN11* mutation; 4 had an *NF1* mutation; 4 had a

KRAS mutation; 2 patients had an *NRAS* mutation and one patient had a *CBL* primary mutation (Figure 3.9A). Patients across the different mutational groups had comparable monocyte counts and Hb concentration, with the exception of the *CBL* patient that had a lower monocyte count. The majority of patients were thrombocytopenic with the KRAS group demonstrating the most severe thrombocytopenia compared to the rest of the patients (Figure 3.9 B). When the MCV of the JMML cases was assessed I noted that the patients with raised MCV for age carried a *KRAS* mutation, whereas the patients with low MCV carried a *NRAS* mutation; the rest of patients had normal MCV for their age (Figure 3.10C). As expected, syndromic features were concentrated among the *NF1* patient group, due to the familial and systemic nature of the disease (Figure 3.9C). Failure to thrive was a common presentation especially among the poor prognosis mutational groups (*PTPN11*, *NF1* and *KRAS*) especially in the *KRAS* patient group (50% of patients).

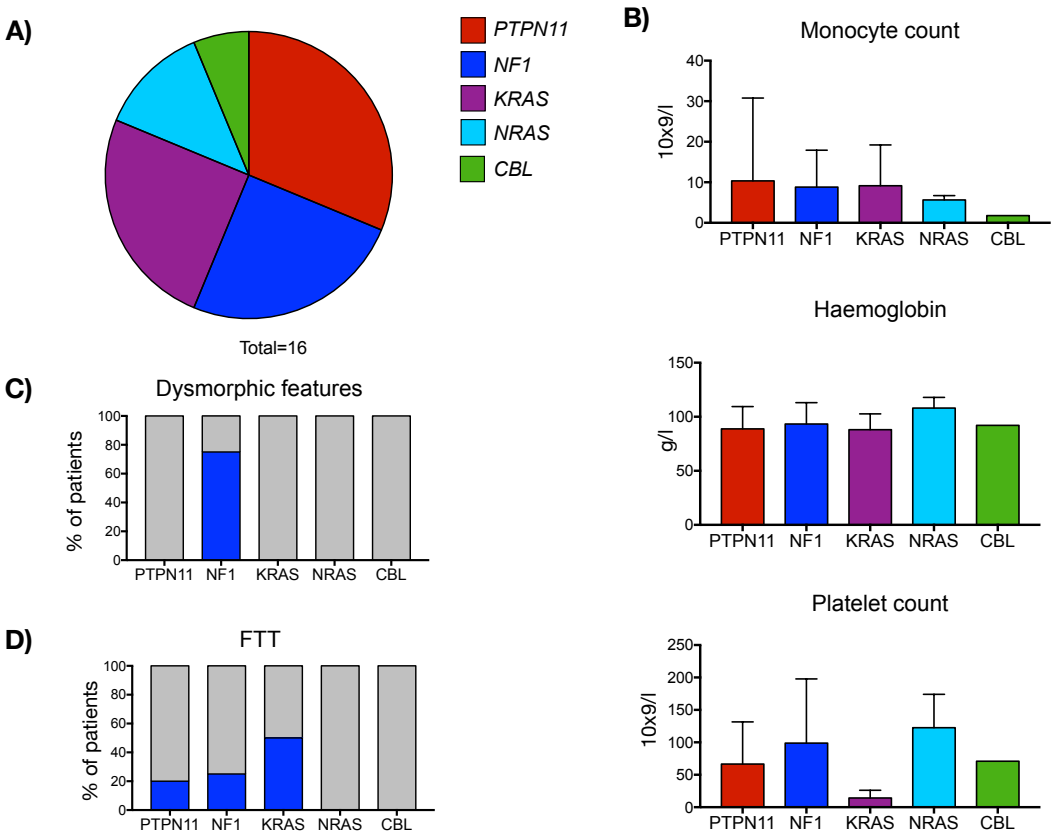


Figure 3.9: A) Pie chart showing the pathogenic mutations in *RAS* pathway genes in the patients with JMML B) Monocyte counts, Hb and platelet counts of the JMML patients C) Percentage of patients with dysmorphic features by mutation type. D) Percentage of patients with failure to thrive (FTT) by mutation type

Overall the mean age at presentation of JMML was around 2 years, however patients in the *KRAS* group were younger, with 3/4 patients presenting at 4 months of age, two of which were carrying the same *KRAS* mutation (*KRAS* c38 G>A) (Figure 3.10A). These patients were also very unwell at presentation with a high bleeding tendency due to thrombocytopenia and coagulopathy with hypofibrinogenaemia, requiring frequent blood product support (Figure 3.10E).

The % HbF was higher in the disease subgroups with the worst prognosis (all patients with *PTPN11* and *NF1* mutation) (Figure 3.10B) consistent with previous reports indicating that high HbF is a marker of poor prognosis (139). A high HbF% was also associated with a higher rate of relapse during the project (Figure 3.10D).

Secondary mutations were observed in 100% of patients with *NF1*, in 40% (2/5) and 25% (1/4) of the *PTPN11* and *KRAS* cases respectively and in 1/2 patients with *NRAS* mutations (Figure 3.10F).

The overall relapse rate was 19% (3/16) with 2 of these 3 patients relapsing within the first few months post-BMT. Serial samples were available on 2 of these patients (ID1 and ID5) as well as in patients that have not yet relapsed (ID2, ID6, ID7 and ID10). Targeted NGS identified relapse at the same time as clinical evidence of relapse in ID2 and ID5 and may not be sensitive enough to detect relapse when the level of disease is very low but further samples would need to be tested to assess this reliably (Figure 3.11A). Pre-BMT patients ID1, ID2, ID 6, ID7 and ID10 received 1-3 cycles of Azacitidine, but this failed to reduce the disease burden except for ID7 who had no detectable mutations following 2 cycles of Azacitidine (Figure 3.11B). The JMML mutations were very stable and unlike AML there were no unstable clones detected that transiently appear during the course of the disease. However larger patient cohorts and longer periods of observation will help to consolidate these results.

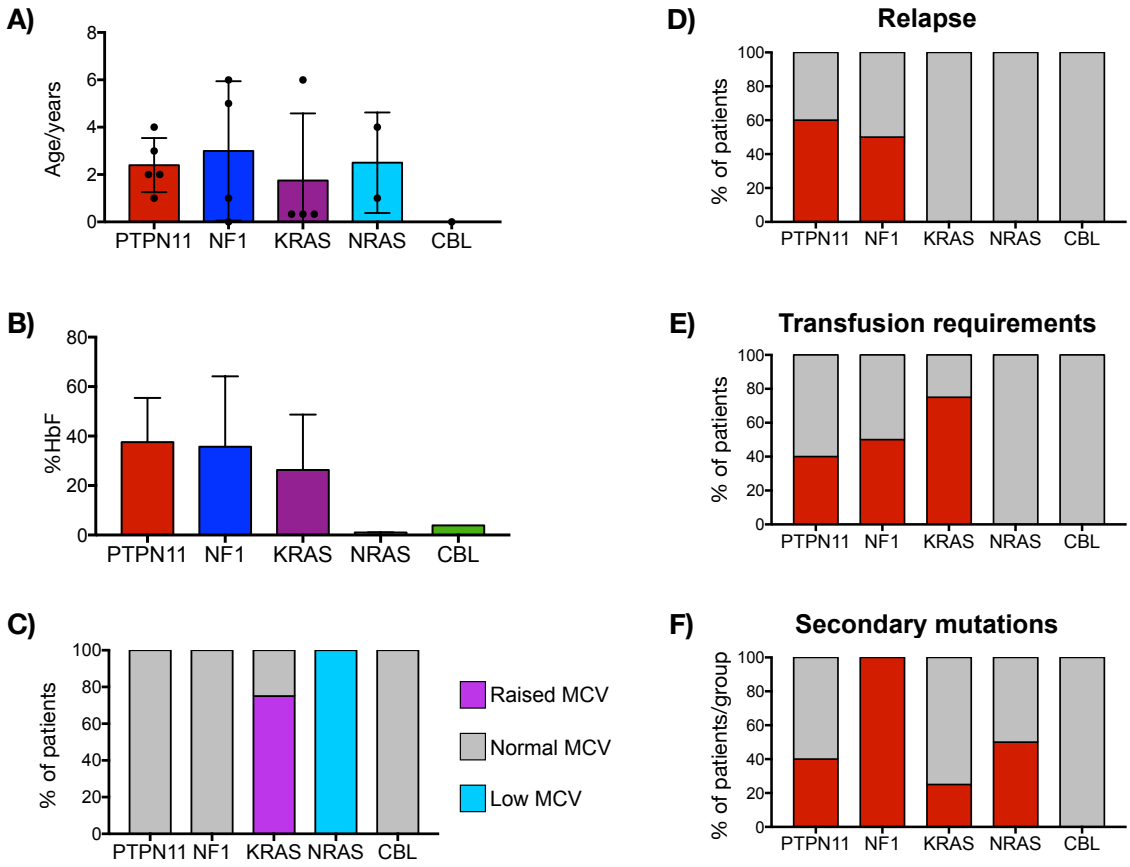


Figure 3.10: The effect of mutation type in JMML on: A) Mean age at presentation B) Mean %HbF at diagnosis C) % of patients with low, normal and high MCV D) % of patients that relapse E) % of patients requiring transfusions and F) the % of patients with secondary mutations

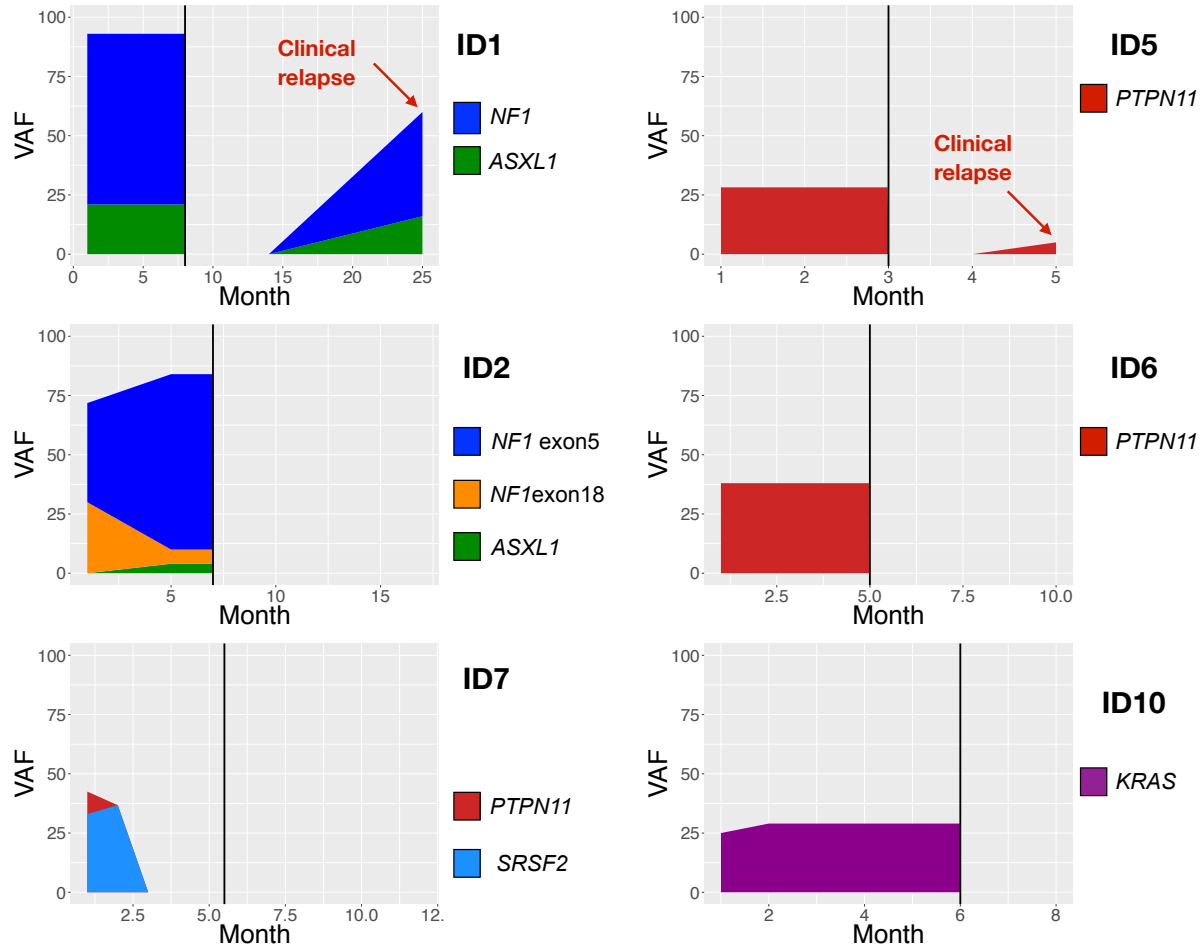


Figure 3.11: Serial sample monitoring with targeted NGS in A) 2 patients who relapsed and B) 4 patients who have not relapsed. The black vertical line indicates when BMT was performed

The overall survival of the JMML patients recruited to the study was 87.5% (2/16). The 2 patients who died both had NF1 variants; one patient died due to disease progression and blastic transformation and the second patient due to severe GvHD (Figure 3.12). Long term follow up of these cases will provide further insights between the disease associated mutations and response to treatment.

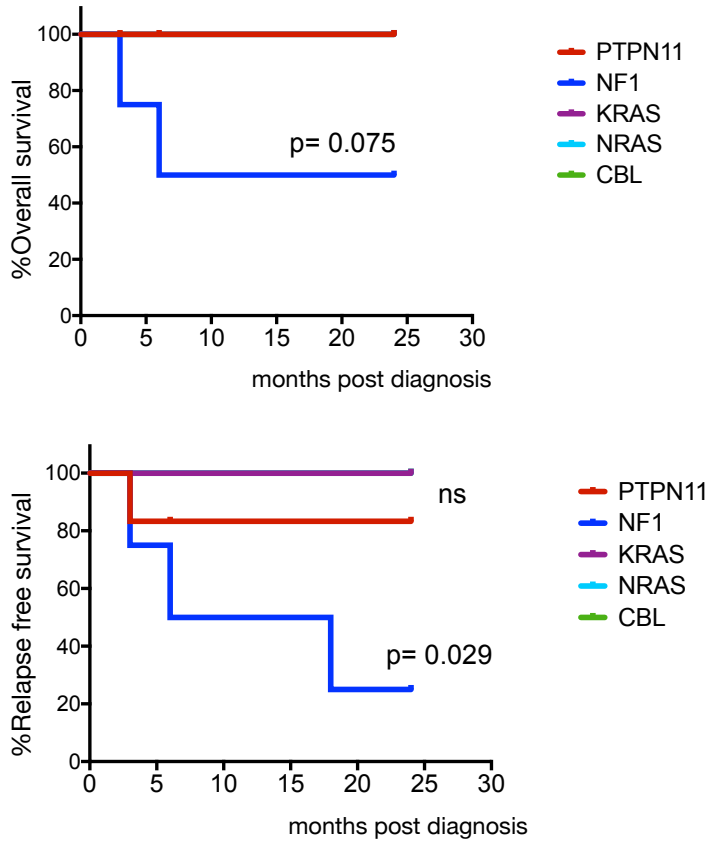


Figure 3.12: Kaplan-Meier survival curve of JMML patients displaying the overall survival and relapse free survival by mutation type.

3.4.4 Development of algorithms for diagnosis and management

3.4.4.1 Diagnosis

It is essential to establish a diagnosis in all children with significant cytopenias in order to plan treatment and provide accurate genetic counseling. A suggested algorithm is shown in Figure 3.13. As part of this, it is also important to determine the precise molecular diagnosis wherever possible; as shown by my data, small targeted NGS panels are currently insufficient for this purpose (apart from in JMML) and high throughput WGS is likely to be needed. Alternatively, targeted panels could be made more extensive to encompass genes responsible for all myeloid malignancies and

will need to be updated as new variants are reported (102),(99),(114),(113),(29),(321),(34)(Figure 3.13). Based on the currently available literature and the paucity of mutations detected by the NGS panel for the ICUS group, an expanded comprehensive targeted NGS panel for the investigation of these children is outlined in Table 3.1.

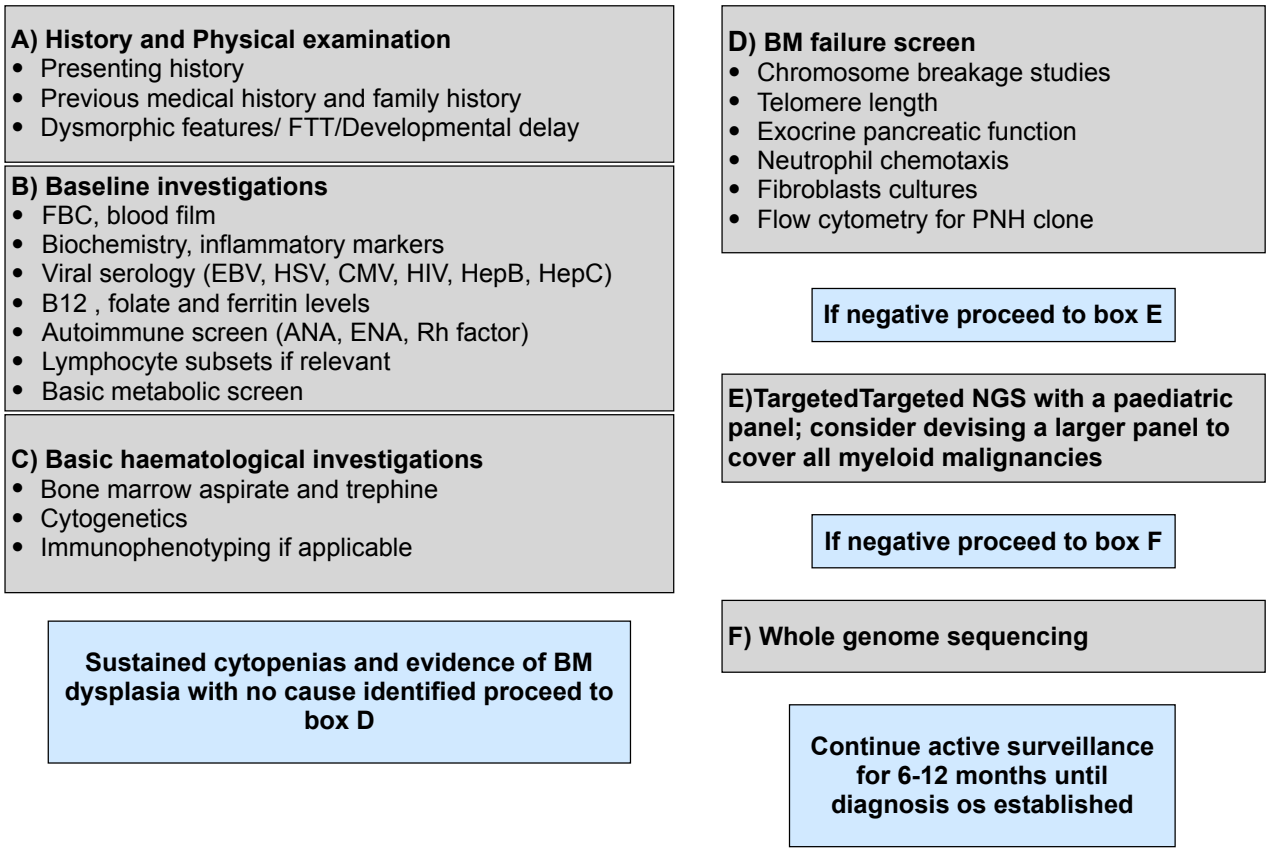


Figure 3.13: Diagnostic algorithm for children with sustained cytopenias and/or dysplasia at the time of writing this thesis

Table 3.1: List of molecular targets that should be included in the investigation of children that present with features of myelodysplasia

Gene targets linked with myeloid neoplasms			
ABL1	*CBL*	*IDH1*	*NRAS*
ASXL1	*CUX1*	*IDH2*	*NOTCH1*
ATRX	*CARL*	*IKZF1*	*PDGFRA*
ARHGAP26	*DNMT3A*	*JAK2*	*PHF6*
BCL2L11	*ETV6*	*JAK3*	*PML*
BCORL1	*EZH2*	*KIT*	*PRPF40B*
BCOR	*FBXW7*	*KRAS*	*PTEN*
BCR	*FTL3*	*KDM6A*	*PTPN11*
BRAF	*GATA 1*	*MYD88*	*RIT1*
CBLB	*GNAS*	*MLL*	*RAD21*
CBLC	*GRAF*	*MET*	*RARA*
CDKN2A	*HRAS*	*NPM1*	*RPS14*
SF3B1	*SMC3*	*SOS1*	*WT1*
SRSF2	*STAG2*	*SF3A1*	*U2AF1*
SETBP1	*SMC1A*	*TET2*	*ZRSR2*
SH3BP1	*SF1*		
BM failure genes			
CTC1	*FANCA*	*SBDS*	*ELANE*
DKC1	*FANCB*	*RPS7*	*G6PC3*
NHP2	*FANCC*	*RPS10*	*GFI1*
NOP10	*BRCA2/FANCD1*	*RPS17*	*HAX1*
RTEL1	*FANCD2*	*RPS19*	*WAS*
TERC	*FANCE*	*RPS24*	*CSF3R*
TERT	*FANCF*	*RPS26*	*RPL23*
TINF2	*FACG*	*RPS27*	*RPL26*
WRAP53	*FANCI*	*RPS27A*	*RPL27*
NOLA3/NOP10	*FANCJ (BRIP1/BACH1)*	*RPS28*	*RPL31*
NOLA2/NHP2	*FANCL*	*RPS29 (TSR2)*	*RPL36*
ACD/TPP1	*FANCM*	*RPL5*	*FANCU/XRCC2*
PARN	*FANCN (PALB2)*	*RPL11*	*FANCV/REV7*
NAF1	*FANCO (RAD51C)*	*RPL35A*	*FANCR/RAD51*
STN1	*FANCP (SLX4)*	*FANCQ (ERCC4)*	*FANCS/BRCA1*
FANCT/UBE2T			
Genes responsible for MDS familial predisposition syndromes			
ABCB7	*NBN*	*SRP72*	*TP53*
AK2	*RMRP*	*DDX41*	*PAX5*
ANKRD26	*MECOM*	*SAMD9L*	*GATA2*
ATR	*NF1*	*SAMD*	*CEBPA*
LIG4	*BLM*	*MPL*	*RUNX1*

3.4.4.2 Management

Guidance for surveillance is a lot more heterogeneous and complex depending on the child's underlying condition. Patients with MDS or JMML that fit the criteria for BMT should proceed to transplant as soon as an appropriate donor is available (32, 139) (Figure 4.14). Follow up post-BMT depends on the underlying condition. Patients with JMML have a very high risk of relapse especially within the first 3 months of treatment. Identification of relapse early is important in order to plan further treatment, including withdrawal of immunosuppression or a second transplant. Asymptomatic patients with RC without a clonal cytogenetic abnormality should be monitored 6- 12 monthly with a BM examination and molecular and cytogenetic screening to detect the development of additional events and disease progression. Asymptomatic children with ICUS should have a FBC every 6 months and consideration of a BM examination on a yearly basis to monitor the degree of dysplasia and disease evolution, until the diagnosis is established (Figure 3.14).

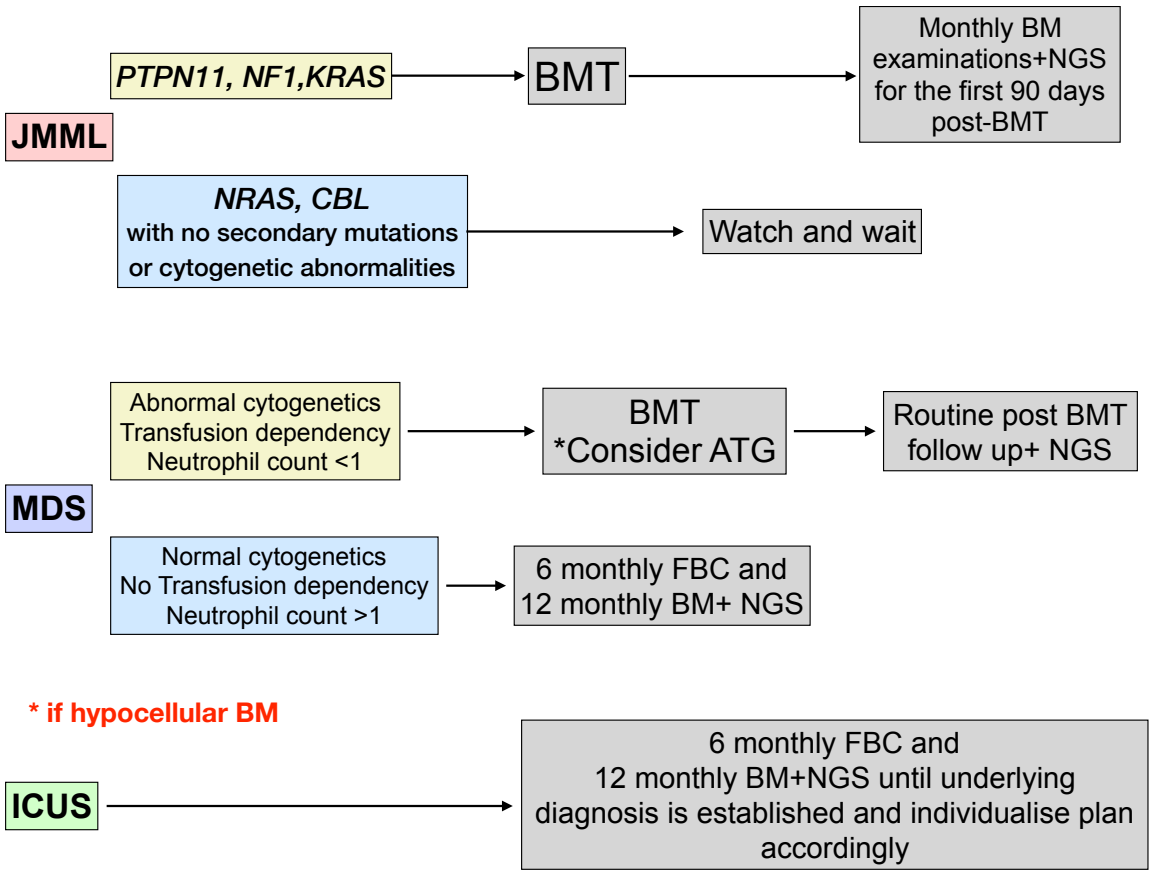


Figure 3.14: Management and follow up algorithm for paediatric myelodysplasia and JMML

3.5 Conclusions and Discussion

Considering the rarity of these conditions and previous reports (7) the estimated recruitment for the study was 10 patients per year given the reported incidence of 1.3 cases/million. However, given that there are 11.4 million children in the UK, the actual recruitment was almost 3 times this estimate. This suggests that the incidence of JMML and MDS in the UK may be higher than originally thought, perhaps due to improved diagnostic tools. Recruitment to the study also showed that in addition to JMML and MDS, there is a poorly recognised group of children with significant cytopenias requiring treatment many of whom do not have an established diagnosis. This ICUS group is characterised by mild dysplastic changes in the BM most likely secondary to molecular defects in genes necessary for normal haematopoiesis, e.g.

CSF3R or *MECOM* mutations. Establishing the diagnosis of these cases can be extremely challenging but is just as important as establishing the diagnosis of *de novo* MDS given the impact on management of these children. I used the term ICUS, normally used in adult patients, as an ‘interim‘ term to characterise this group until a formal diagnosis is established. The cut off of 10% BM dysplasia successfully discriminated patients with *de novo* MDS from the rest of patients with cytopenias, based on the distinct molecular signatures of the two groups. While mutations in *GATA2* were the commonest single group within the MDS patients, no molecular diagnosis was established in the majority of cases of ICUS during the period of my study, although WGS was successful in some cases.

Finally, feedback from the recruiting centres about the value of the study was very positive which in turn translated into the high recruitment numbers. Informal feedback also indicated support for the need for evidence-based diagnostic, surveillance and national management guidelines as well as a specialist forum where such cases can be discussed.

Chapter 4: PHENOTYPIC AND FUNCTIONAL CHARACTERISATION OF JMML HSPC

4: PHENOTYPIC AND FUNCTIONAL CHARACTERISATION OF JMML HSPC

4.1 Background

4.1.1 The HSPC compartment in normal haematopoiesis

Haematopoiesis is organized in a hierarchical manner and in humans is first established at ~ 3 weeks gestation in the yolk sac (322) where it is known as primitive haematopoiesis (323). This gives rise to primitive erythrocytes, macrophages and megakaryocytes. Shortly thereafter, definitive haematopoiesis begins when the first true HSC arise from the haemogenic endothelium in the aorta/gonad/mesonephros (AGM) region of the embryo at ~ 5 weeks and subsequently colonise the fetal liver from day 35 of life (324). At 11-12 weeks HSCs migrate to the bone marrow (BM) but prenatally fetal liver remains the primary site of haematopoiesis until haematopoiesis is fully established in the BM around birth (325) ;the postnatal HSCs pool is maintained through strict regulation of HSC self renewal and differentiation (324).

HSC are characterised by two very distinct features: the ability to produce all of the different types of blood cells (*multipotentiality*) and the ability to *self renew* and maintain the BM stem cell pool (326) .It is over a hundred years since the first report of the clonal architecture of haematopoiesis and observation that haematopoietic precursors could be derived from a single HSC (1909). Subsequently, the self-renewal ability of HSC was recognized following the observation that radiation-induced BM failure could be rescued by injection of BM or spleen cells (327). Identification of the multipotentiality of HSC was made a few years later (328), (329), (330) following the observations of Till et all and Becker et al. HSC are positioned at

the top of the haematopoietic hierarchy (Figure 4.1) and can give rise to all downstream progenitors, leading to the production of different blood cell lineages. Progenitors downstream of HSC have little or no self renewal ability and demonstrate a gradual fate restriction giving rise to specific blood cell lineages (331) (Figure 4.1). Many groups have sought HSC-specific markers which would allow sub-fractionation of HSC. Several markers (eg CD133, a6 integrin) have now been shown to enrich for the frequency of long-term repopulating cells in serial xenograft assays, currently considered the 'gold standard' for enumerating HSC (308), (332), (333) .

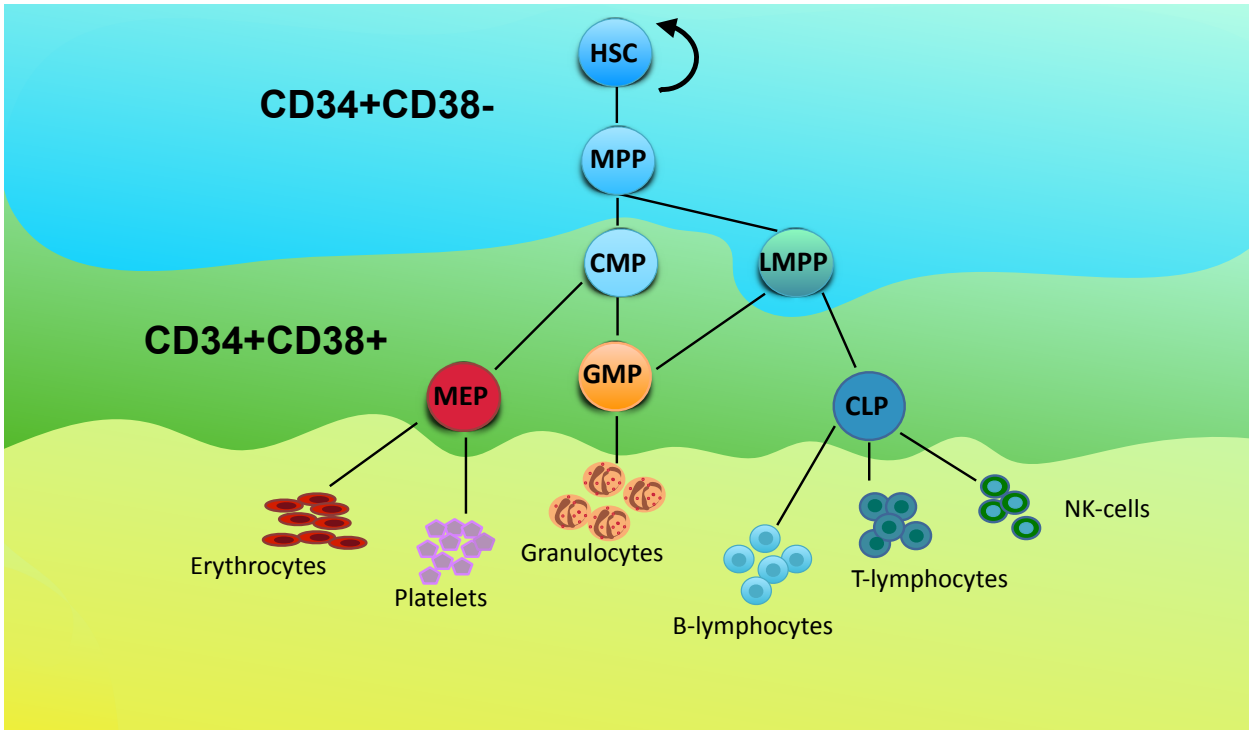


Figure 4.1: A cartoon showing the haematopoietic hierarchy and principal HSPC (recent reviews indicate that the hierarchy is much more complex than shown here (334), (335),(327), (336)

HSC and their progenitors are currently characterised based on their immunophenotypic definition and lineage potential. CD34 is a single pass transmembrance sialomucin protein and represents the fundamental haematopoietic

cell surface marker expressed on all HSPC, constituting ~1-4% of the total human BM cells (337),(338). Further subfraction of HSPC populations has been achieved over time using a variety of cell surface markers such as CD38 is a glycoprotein, involved in cell adhesion and signal transduction (339) which is used to distinguish the HSC and very early progenitors compartment from the more mature myeloid progenitors (340). Weismann and others were able to further purify the HSC compartment using CD90 (Thy-1) and CD45RA (a member of the protein tyrosine phosphatase family) (341),(342). CD90 a conserved cell surface protein normally expressed on T cells was originally discovered as a thymocyte antigen (343), is one of the cell surface markers that aids HSC identification and isolation, along with, more recently discovered antigens such as CD49f (308) an adhesion molecule involved in T cell function and cell signaling (344). HSC are defined as Lin-CD34+CD38-CD90+ and lack CD45RA, the long isoform of CD45 which is implicated in cell differentiations and cell cycle regulation (345), (346). The main multipotent progenitors, immediately downstream of HSC are MPP (Multipotent Progenitors), which have only transient engraftment capability in contrast to HSC, and Lymphomyeloid Multipotent Progenitors (LMPP) or Myelo-lymphoid progenitors (MLP) (347),(309) which have lymphoid and myelomonocytic output, but lack erythroid and megakaryocytic potential. (348),(349), (350),(326). The CD34+CD38+ compartment can be further fractionated by using CD123 (IL3 receptor) and CD45RA (309). A summary of the commonly used markers to identify the main HSPC populations in shown in Table 4.1.

Table 4.1: Immunophenotypic definitions of Haematopoietic stem and progenitor populations (HSPC)

Population	Immunophenotypic definition
Haematopoietic Stem Cell (HSC)	Lin-CD34+CD38-CD90+CD45RA-
Multi-potent Progenitor (MPP)	Lin-CD34+CD38-CD90-CD45RA-
Lymphoid primed Multi-potent Progenitor (LMPP)	Lin-CD34+CD38-CD90-CD45RA+
Common myeloid progenitor (CMP)	Lin-CD34+CD38+CD123+CD45RA-
Megakaryocyte-Erythroid progenitor (MEP)	Lin-CD34+CD38+CD123-CD45RA-
Granulocyte-macrophage progenitor (GMP)	Lin-CD34+CD38+CD123+CD45RA+

4.1.2 The HSPC compartment in leukaemia

In normal haematopoiesis, significant self renewal is observed only within the HSC compartment. However, in leukaemia, such as acute myeloid leukaemia (AML), downstream progenitors acquire abnormal self renewal capacity and display a differentiation block(350). This leads to continuous proliferation; and uncontrolled infiltration of leukaemic cells into the BM, spleen, liver and non haematopoietic organs (351). The structure of haematopoiesis in AML is thought to be best described by one of two models (352). One is the ‘stochastic’ model where leukaemia cells all have equal disease-propagating potential and the heterogeneity observed is dependent on the presence of intrinsic and extrinsic factors, such as the acquisition of secondary genetic defects. The second model is the ‘hierarchical model’ by which

the cancer-propagating stem cell (CSC) in AML is organized in a hierarchical fashion with the CSC at the apex of the hierarchy maintaining the CSC pool growth and also differentiating into leukaemic progenitor cells.

Initially the AML leukaemic stem cells (LSCs) were thought to reside within the Lin-CD34+CD38- compartment (353), (354) however further studies have identified LSCs potential also within the CD34+CD38+ compartment; (355) ,(356) with more mature progenitors such as GMPs appearing to have LSCs activity and ability to propagate the disease in vivo(357), (358),(359),(360).This is now been very well investigated in AML with LMPP-like, GMP-like and CD34negative LSCs being identified in AML(350), (361). By contrast, until the work described in this chapter, little was known about the haematopoietic hierarchy in JMML or the disease-propagating population (JMML-LSCs).

Previous studies have hypothesised that the JMML 'stem cells' must reside within the earliest HSPC compartment defined as Lin-CD34+CD38-; as JMML mutations have been traced in cells of different lineages as well as within the CD34+CD38- compartment (229),(188),(362),(363),(364),(365). JMML xenografts with whole mononuclear cells have shown that the JMML cells can give rise to different types of lineages which suggest that JMML stem cells retain to some degree multi-potent activity and the ability to differentiate(366). However, despite the great advances in understanding the molecular defects that underpin JMML, to date there are no studies that monitor the disease molecular and cytogenetic abnormalities in HSPCs in a serial manner. This precludes further understanding of the cellular origin of JMML; as well as the order of the mutation and cytogenetic events evolution during

disease progression and relapse. The age group affected by JMML, as well as the fetal characteristics of JMML and the fact that the disease-associated mutations can be detected in neonatal and cord blood samples, indicate that the disease may arise in a fetal stem/progenitor cell. One potentially attractive candidate as a target for such transformation is the Erythromyeloid Progenitor Cell (EMP). EMP is a transient population that forms the second wave of primitive murine haematopoiesis responsible for production of platelets, myeloid and erythroid cells before the establishment of definitive HSC (367), (368). EMP are not maintained in adult haematopoiesis and have not yet been reported in humans .In addition, data from mice with activating *PTPN11* JMML mutations in E9.5 yolk sac EMPs demonstrate that mutant EMPs are not sufficient to cause disease when transplanted into recipient mice (369). Thus, the target cell for transformation in JMML remains unknown.

I therefore set out to document the HSPC compartment in human JMML in detail. This chapter describes the immunophenotypic and functional characterisation of JMML HSPC compared with normal controls, in order to document the structure of haematopoiesis in JMML and identify the LSCs responsible for propagating the disease. Multi-potentiality and self-renewal of JMML HSPC was assessed using a combination of short-term and long-term in vitro and in vivo assays.

4.2 Aim

To study the haemopoietic hierarchy in JMML and determine the impact of JMML-associated mutations on:

1. The structure of the haematopoietic hierarchy
2. The lineage output of JMML HSPC

3. The stem cell capacity of JMML HSPC in vitro and in vivo

4.3 Experimental approach

The HSPC compartment in JMML was initially characterised by using flow cytometry to measure the frequency of the classical HSPC populations shown in Figure 4.1 and Table 4.1 and to identify any novel populations which were not present in normal controls. Most of the normal control data was generated from normal cord blood given the difficulty in obtaining normal age-mached BM but for selected experiments it was possible to obtain comparative from normal BM. I then used 'state-of-the-art' single cell and bulk population sorting techniques to perform in vitro short term clonogenic, liquid culture and short- and long-term stromal cell co-culture assays to assess the lineage potential of the different JMML HSPC compartments. Finally, we carried out in vivo assays using a xenotranplantation model to assess the stem cell and leukaemogenic potential of the different JMML HSPC populations (Figure 4.2). Cell sorting for all of these experiments was performed by either Dr Benjamin Povinelli or the WIMM flow cytometry facility. Tail vein injections, mouse monitoring and tissue final terminations were performed by Dr Chris Booth and Dr Benjamin Povinelli. Mouse tissue downstream processing (tissue cryopreservation, flow cytometry and analysis was performed by myself.)

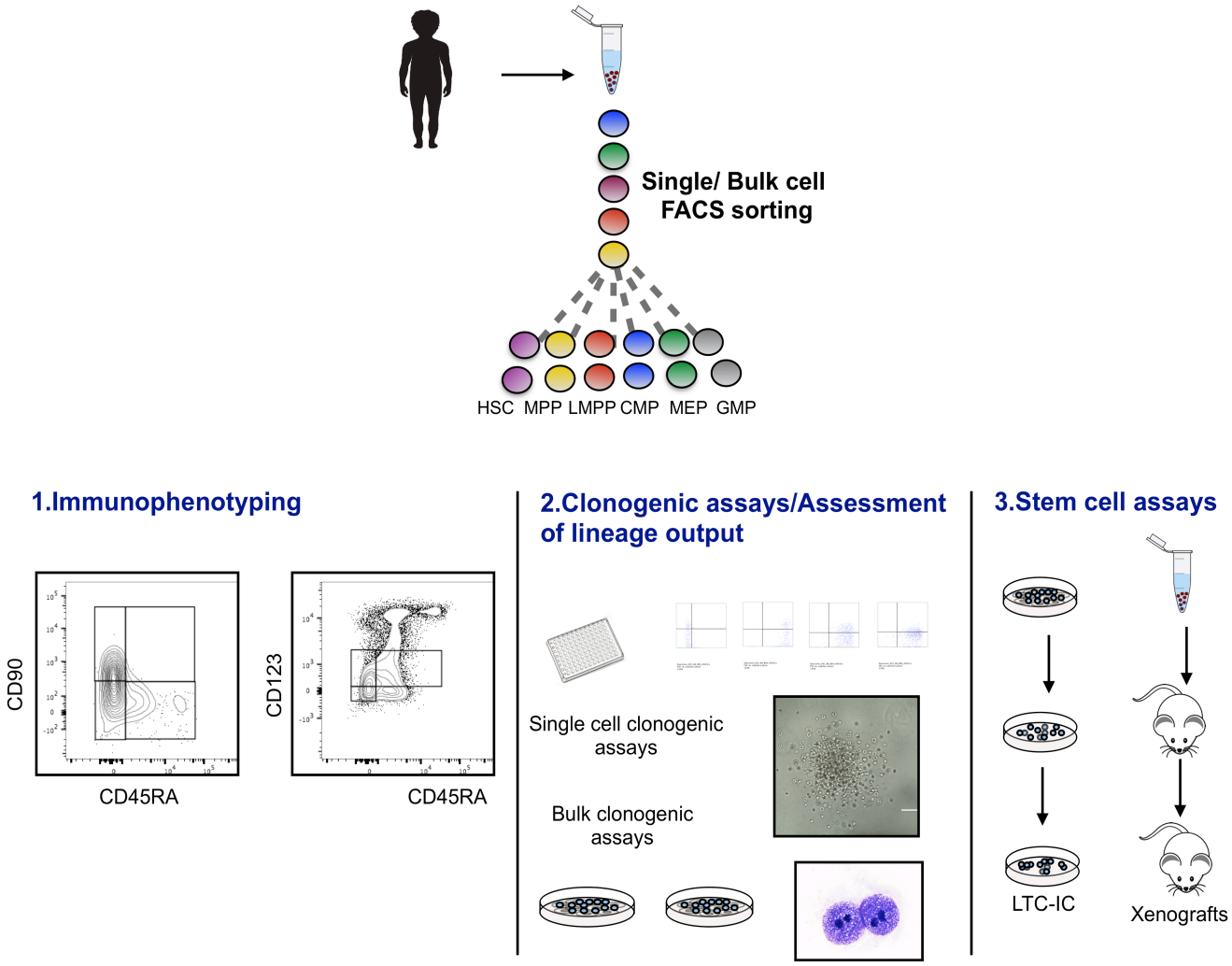


Figure 4.2: Experimental outline of the phenotypic and functional characterization of JMML HSPC

4.4 Immunophenotypic analysis of JMML HSPC at diagnosis

As described above I first analysed primary JMML patient Bone Marrow (BM) and peripheral blood (PB) samples by flow cytometry using the conventional immunophenotypic definitions outlined above (Table 4.1) and calculated the frequency of each progenitor type. The frequency was compared to normal cord blood samples analysed at the same time and to normal paediatric BM obtained from healthy sibling BM donors. Representative flow plots for normal cord blood and BM are shown in Figure 4.3B and C. Compared to both cord blood and paediatric BM, it

can be seen that in JMML the overall architecture is abnormal and that this affects both the Lin-CD34+CD38- compartment which contains HSC, MPP and LMPP, and the Lin-CD34+CD38+ compartment which contains the myeloid progenitors (Figure 4.4).

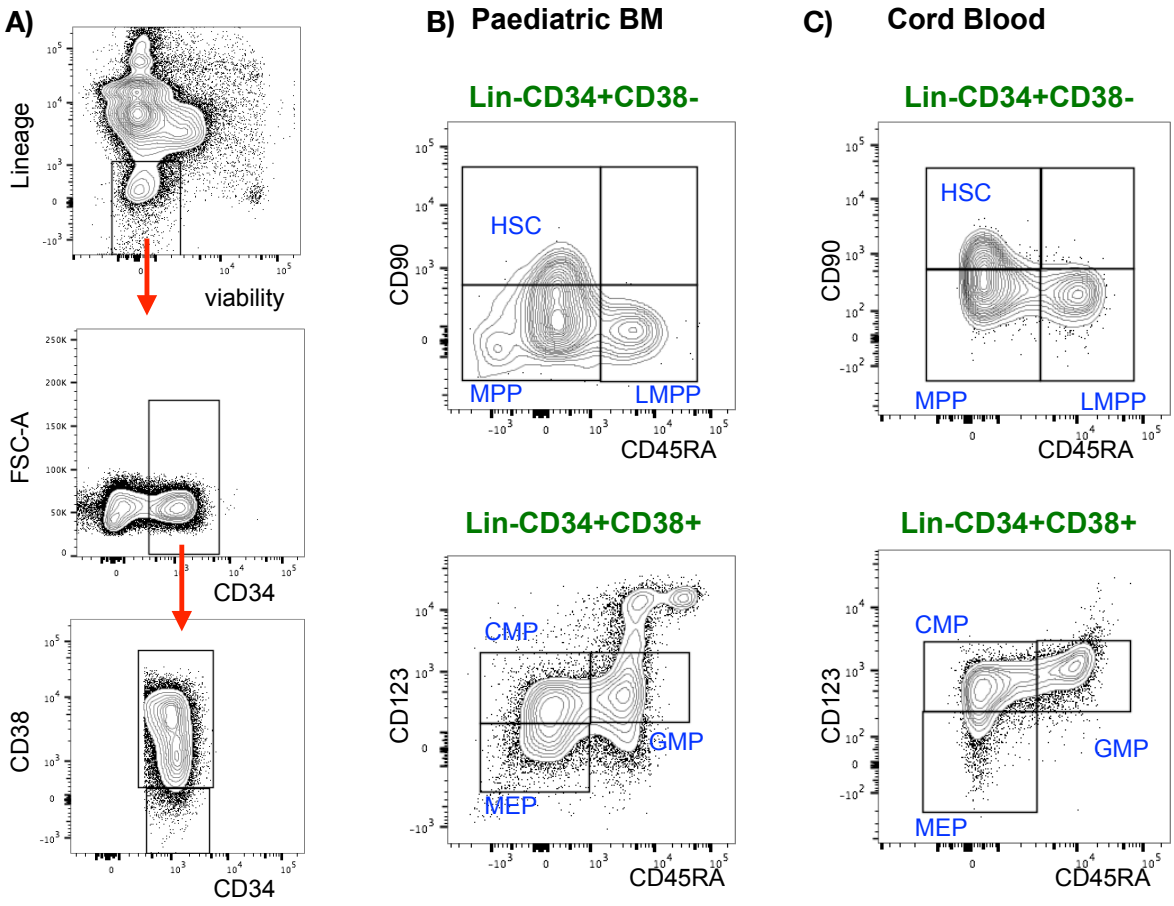


Figure 4.3: Flow cytometry analysis of normal HSPC: A) Gating strategy from top plot gating on Lin negative, viable cells, middle plot gating on CD34+ cells and lower plot representing the CD34+CD38- and CD34+CD38+ compartments B) top plot CD34+CD38- compartment and bottom plot CD34+CD38+ compartment from normal paediatric BM C) top plot CD34+CD38- compartment and bottom plot CD34+CD38+ compartment from cord blood sample from normal donor

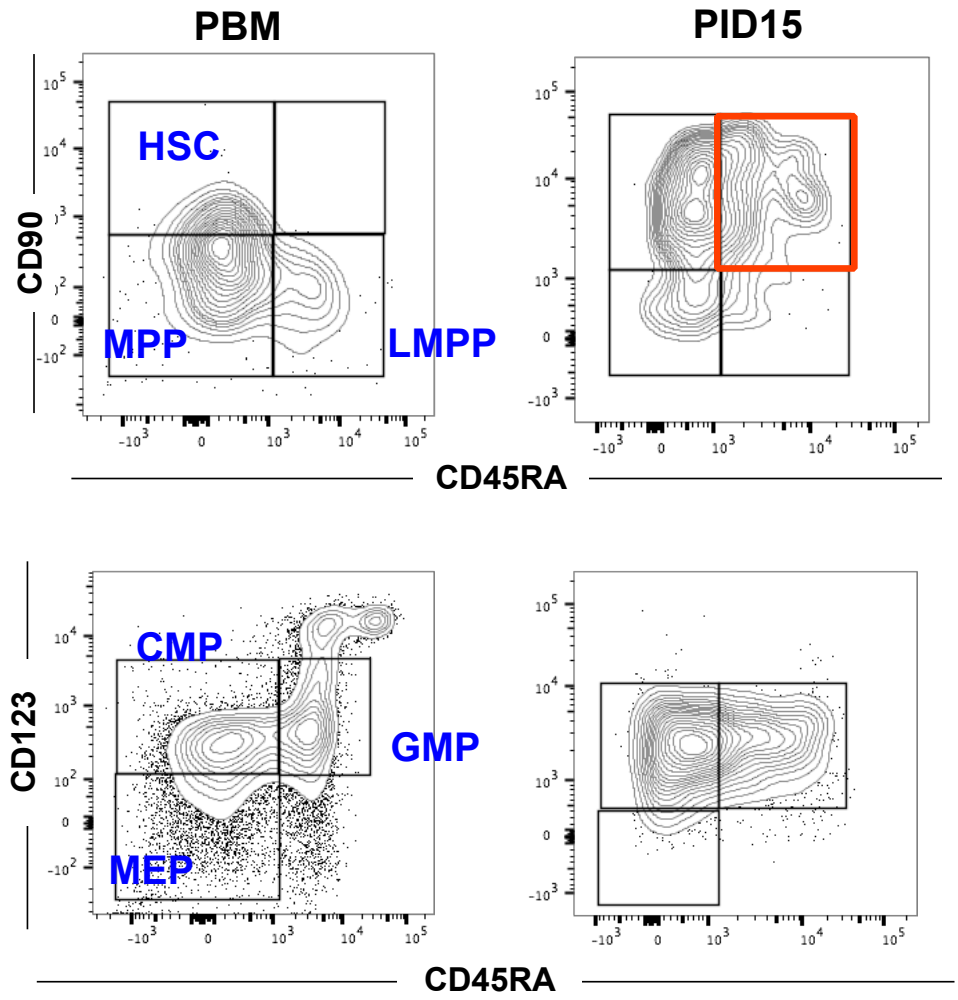


Figure 4.4: Representative example of JMML patient sample Lin-CD34+CD38- compartment (top panel) and Lin-CD34+CD38+ compartment (bottom panel) compared with normal paediatric BM using flow cytometry.

4.4.1 The Lin-CD34+CD38- compartment in JMML

Despite differences in the overall structure between JMML and normal controls, the changes in JMML were heterogeneous and in most cases differed from those reported in AML where the majority of cells are found as a uniform population within the LMPP or GMP gate (350). The most consistent finding in JMML was a significant reduction in MPP ($p<0.0005$; n= 8 vs normal paediatric BM, n=11) and LMPP ($p<0.0112$) (Figure 4.5). The frequency of immunophenotypic HSC was extremely

variable; some patients had almost no detectable cells in the HSC gate, while others had increased numbers of HSC (Figure 4.5). Interestingly, an aberrant fourth population was also present in the majority of the patient samples (5/8) which was **Lin-CD34+CD38-CD90+CD45RA+.** This population has never been described before in normal haematopoiesis and has only been previously reported in adult patients with advanced MDS and monosomy 7 (370).

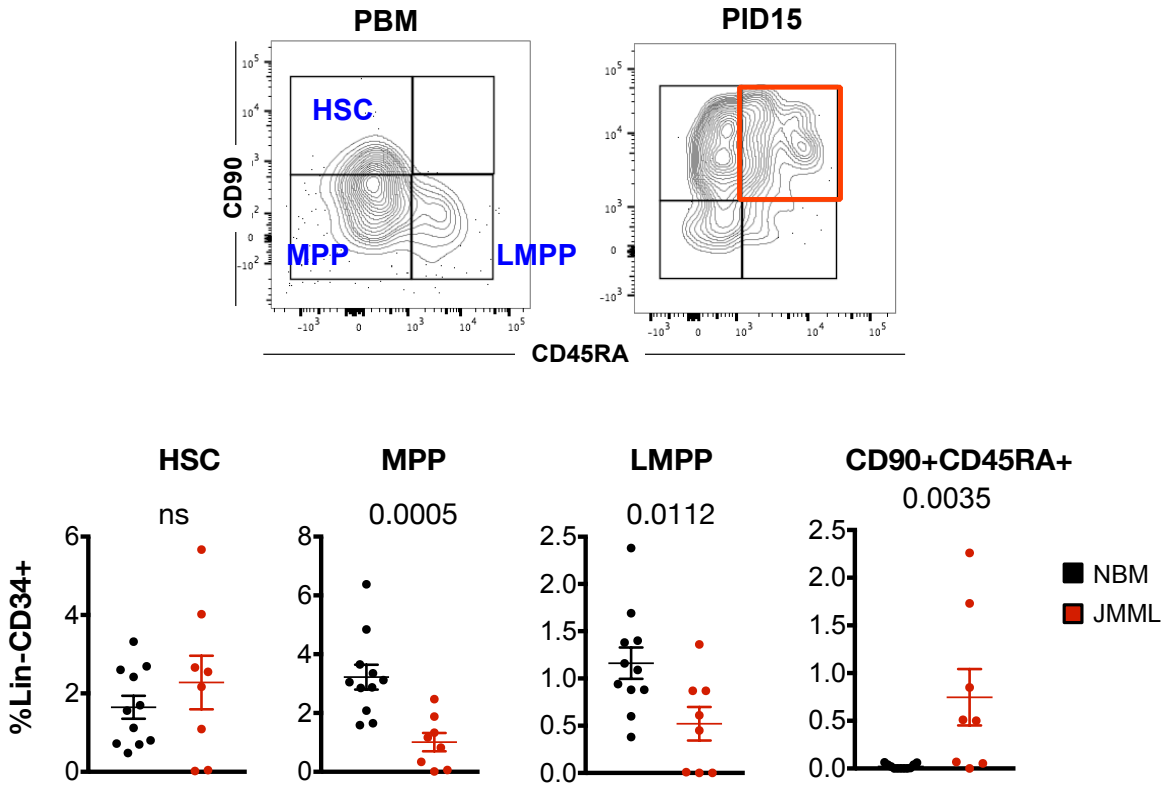


Figure 4.5: Top panel: Flow cytometry examples of the Lin-CD34+CD38- from normal donor and JMML (from left to right)
Bottom panel: Frequencies of the Lin-CD34+CD38- compartment populations in normal controls BM (n= 11), JMML (n= 8) (Analysis of normal Paediatric Bone Marrow and JMML was kindly supervised by Dr Aninida Roy)

I originally thoroughly investigated the presence of this population in normal controls including: cord blood (n=6), paediatric BM (n=11) and analysed normal fetal liver data (n=8; gestation 6-20 weeks) available in our lab through a different project (Dr Anindita Roy). This population was not detected in any of the normal controls from different developmental stages and was found only in JMML patient samples (Figure 4.6).

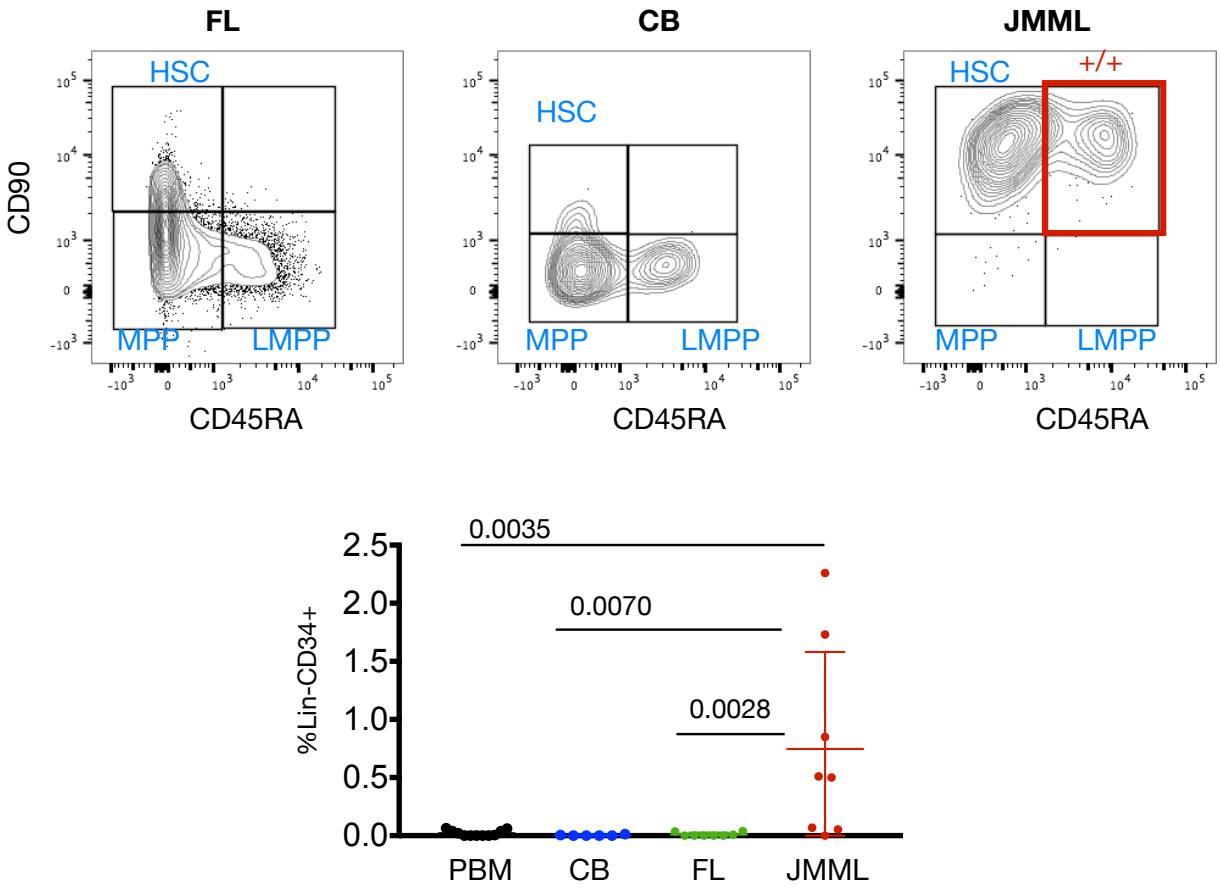


Figure 4.6: Top panel: Flow cytometry examples of the Lin-CD34+CD38- from normal donor samples human fetal liver (FL), cord blood (CB) sample, and JMML (from left to right) Bottom panel: Frequencies of the aberrant Lin-CD34+CD38-CD90+CD45RA+ in normal controls PBM (n=11), CB (n= 6), FL (n= 8), JMML (n=8)

I did note that there was heterogeneity in the CD90+CD45RA+ expression between the different patient samples. Some patients (ID2, ID6, ID9) demonstrated an expansion of the LMPP compartment instead of the aberrant CD90+CD45RA+ population (Figure 4.7); but there were no patients with an expansion of the LMPP as well as the aberrant CD90+CD45RA+ population. A small number of patients (ID8, ID14) had a population with intermediate CD90 expression of ($CD90^{mid}$) population that was also CD45RA+ (Figure 4.7). I investigated whether the mutation type that the patients carried influenced the heterogeneity of the CD90+CD45RA+ expression (Figure 4.7; Figure 4.8). It is interesting to note that 4/5 patient with PTPN11 mutations had an aberrant CD90+CD45RA+ population as did both of the patients with NRAS mutations while none of the 3 patients with KRAS mutations had this population. This suggests that *PTPN11* and *NRAS* mutations may promote the development of the CD90+CD45RA+ population but a larger number of samples would be needed to confirm this, as this did not appear to be statistically significant. However, I did note that the presence of bad prognostic features such as HbF showed some correlation with the detection of the aberrant CD34+CD38-CD90+CD45RA+ population, although this again not statistically significant (Figure 4.8 and Figure 4.9).

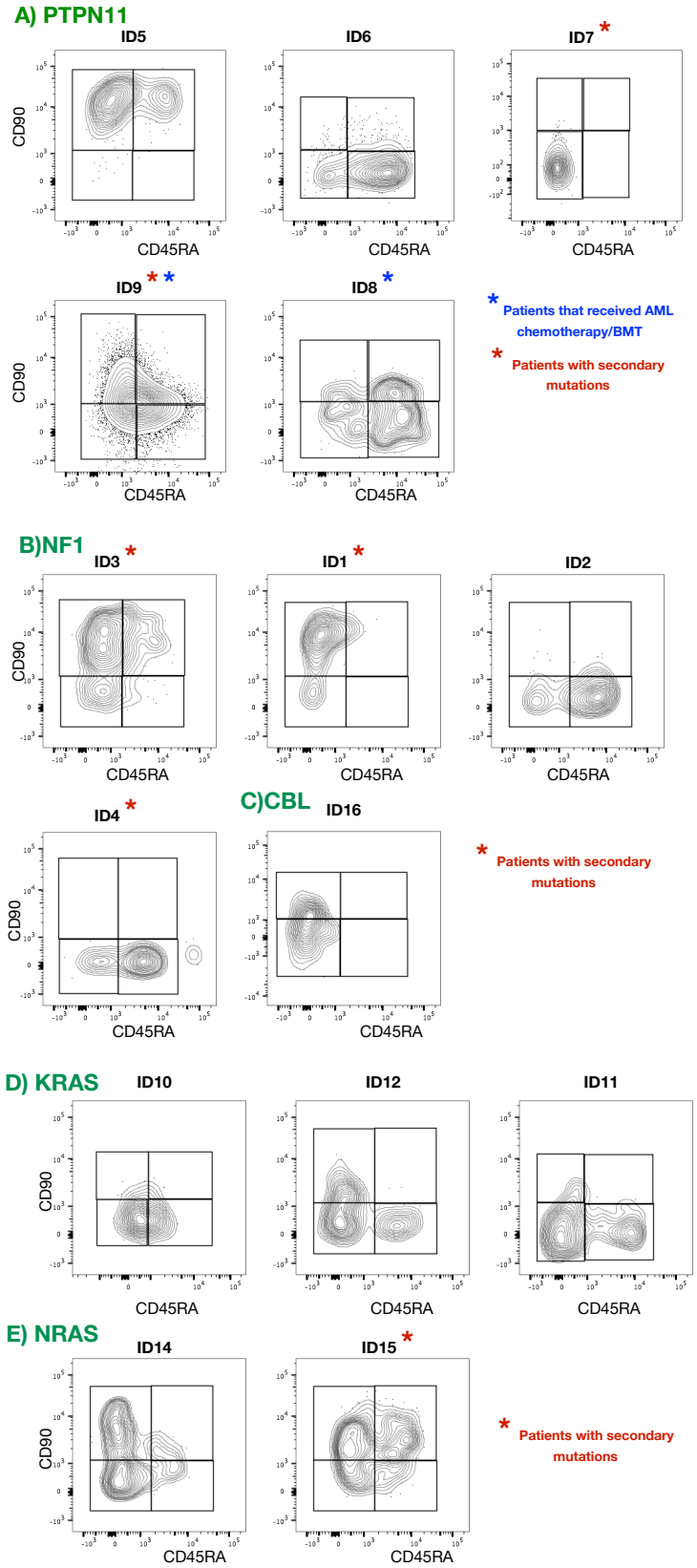


Figure 4.7: Lin-CD34+CD38- profiles of patients. For all patients BM was analysed when this was available, except patient ID 4,6,11 and 16 that PB was used instead. A) Patients with *PTPN11*, B) *NF1*, C) *CBL*, D) *KRAS*, E) *NRAS* mutation

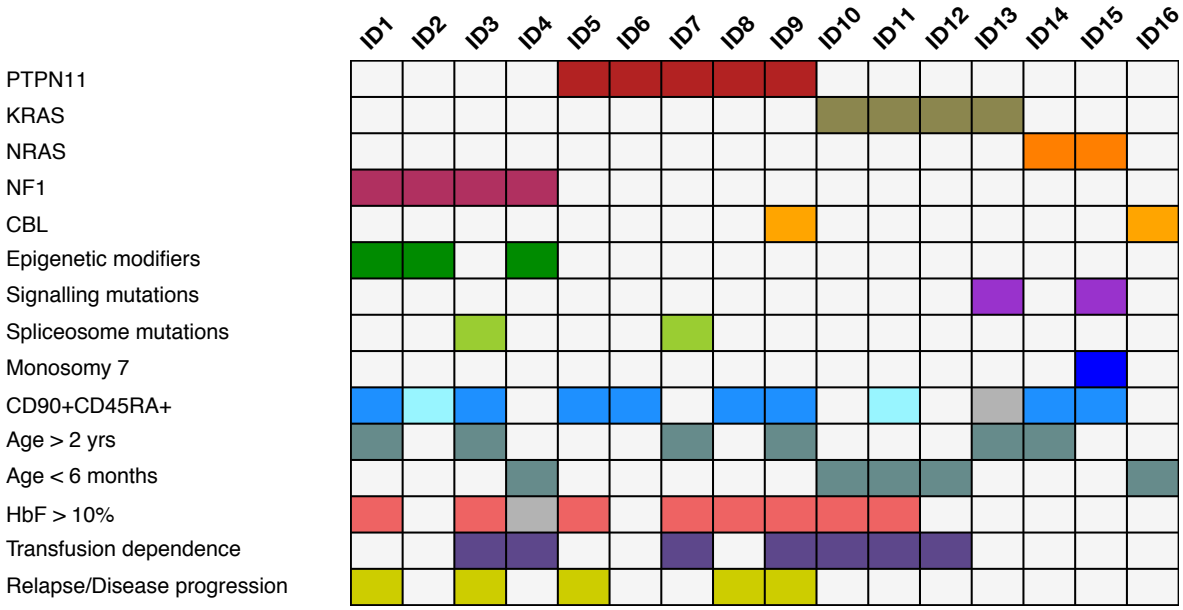


Figure 4.8: Heatmap of all the patients analysed per mutation type, prognostic factors and the presence of CD90+CD45RA+. Light grey squares represent not available data and light blue squares represent patients that have expansion of LMPP instead of CD90+CD45RA+

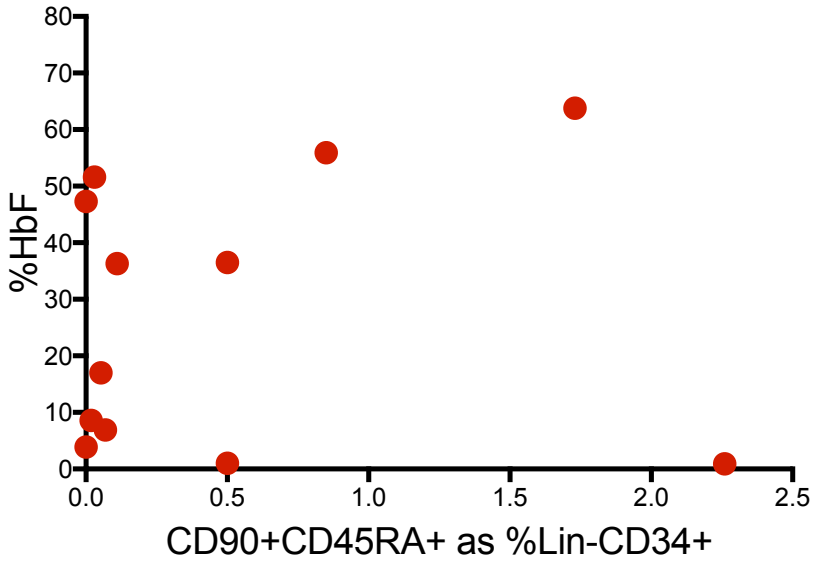


Figure 4.9: Correlation between the %HbF and the presence of the aberrant Lin-CD34+CD38-CD90+CD45RA+

In summary, this analysis indicates that the CD34+CD38- compartment is abnormal in JMML and that the abnormalities are heterogeneous but do not correlate with a specific mutation; but the presence of the aberrant CD90+CD45RA+ appears to have some relationship with other disease bad prognostic features. Analysis of larger patient cohorts will allow to consolidate these results and investigate whether the immunophenophenotype of JMML HSPCs can play a role in the identification of patients with more aggressive disease.

4.4.2 The Lin-CD34+CD38+ compartment in JMML

The changes in the Lin-CD34+CD38+ compartment in JMML samples were also heterogeneous. Several patients had markedly reduced MEP and tended to have a concomitant increase in GMP and/or CMP; however these differences were not statistically significant (Figure 4.10).

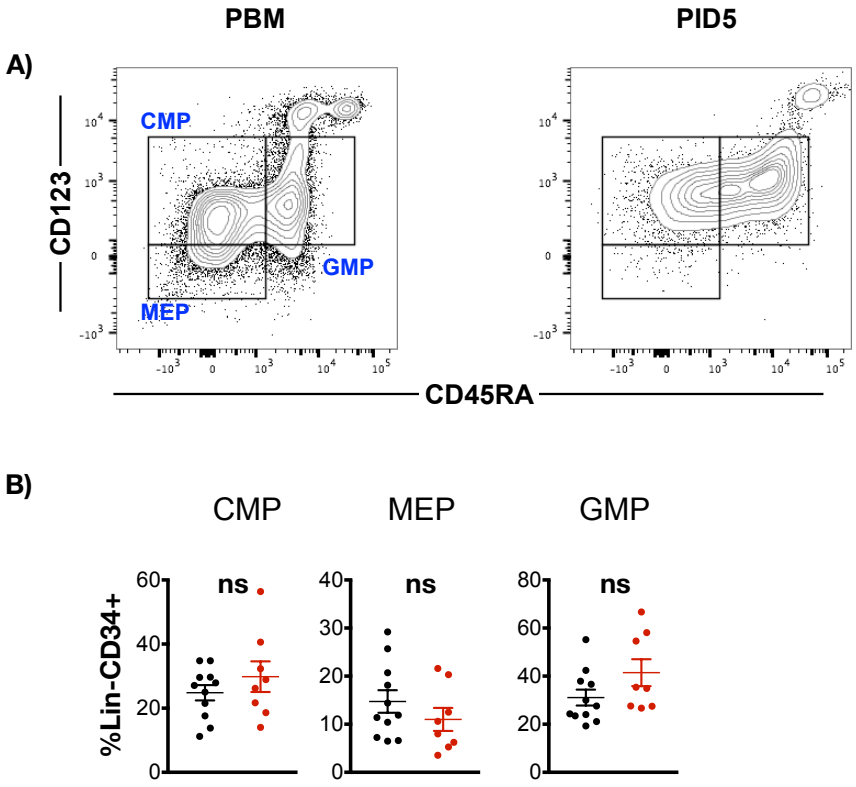


Figure 4.10: A) Flow cytometry analysis of the Lin-CD34+CD38+ compartment compared to controls. B) frequencies of each population expressed as a percentage of Lin-CD34+ cells compared to control (JMML BM n=8, Normal paediatric BM n=11) (Analysis of normal Paediatric Bone Marrows and JMML was supervised by Dr Aninida Roy)

4.5 Flow cytometry analysis of JMML HSPC: comparison of BM with PB

Paired PB and BM collected on the same day were analysed by flow cytometry and compared. There were no significant differences in the frequency of any of the HSPC populations between PB and BM (Figure 4.11). Importantly, the aberrant CD90+CD45RA+population was present in PB as well as BM and the frequency was similar in each paired sample, ie samples with a high frequency of BM CD90+CD45RA+ cells, also had a high frequency in PB (Figure 4.11). There were no statistically significant differences between BM and PB for all populations.

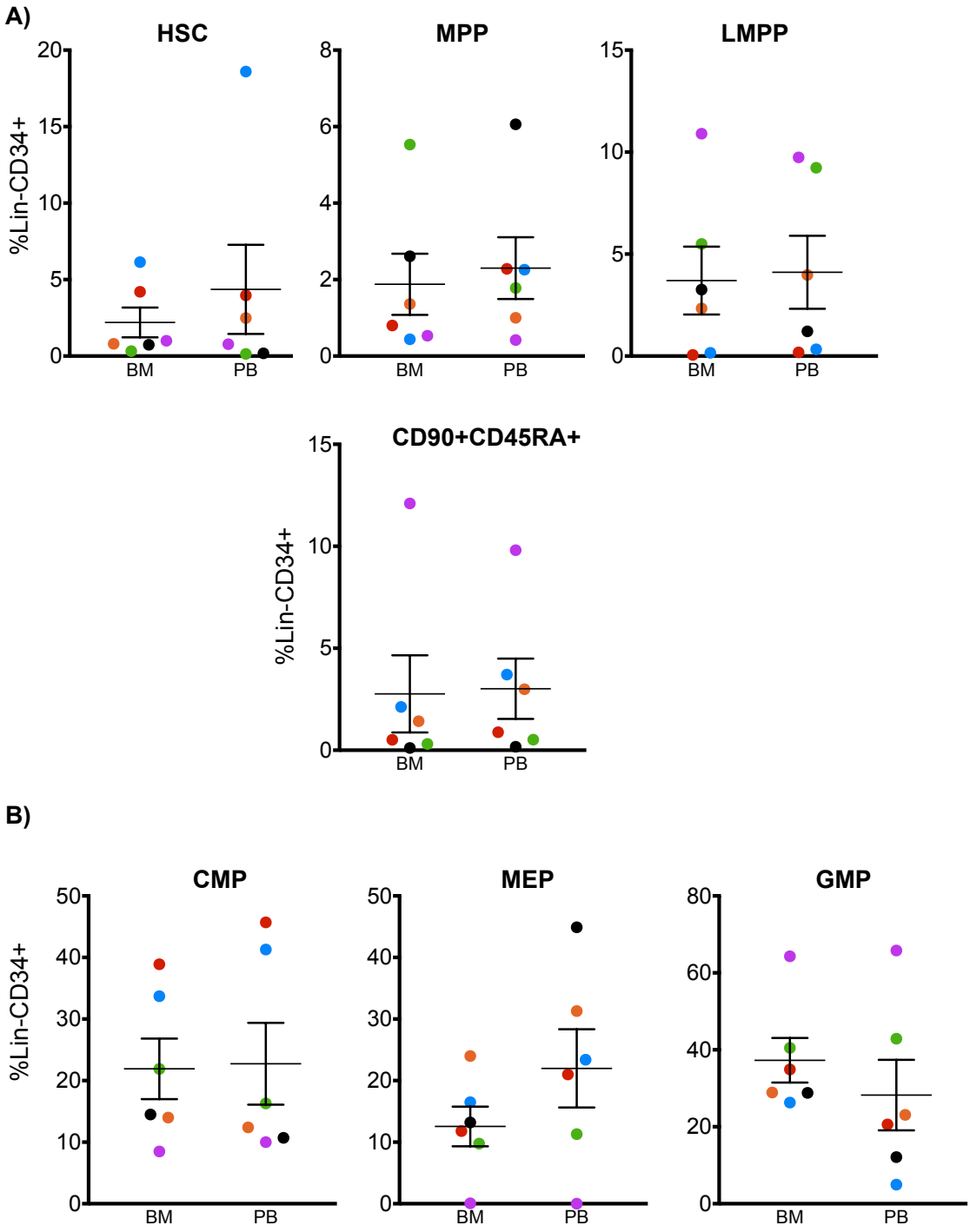


Figure 4.11: Analysis of HSPCs population frequencies expressed as a % of Lin-CD34+ in paired BM and PB patient samples from the same disease timepoint A) analysis of the CD34+CD38- compartment and B) analysis of the CD34+CD38+ compartment (n=6). Each patient is shown in a different colour to allow the frequency of HSPC and PB to be directly compared for each patient.

4.6 Flow cytometry analysis of JMML serial samples: effect of disease stage and treatment on HSPC frequency

Serial samples to allow the effects of disease progression and/or treatment on HSPC frequency were available from 6 patients although these were only available from limited and variable time-points. Of these 2 relapsed (see Figure 4.12) and 4 remain in remission (Figure 4.13). The BMT conditioning regimens were the same in all 6 patients although pre-BMT patients received either azacytidine (AZA;ID 1,2,6,7 and 10) or 6- Mercaptopurine (6-MP; ID5). In 4/5 patients AZA treatment had no effect on disease burden (based on the molecular data from the targeted panel) but in 1 patient (ID7) AZA induced molecular remission although it is not clear whether this would have been sustained since the patient was transplanted shortly afterwards (Figure 4.14).

Considering the impact of BMT, analysis of the 2 patients that relapsed (ID1 and ID5) demonstrates that the CD34+CD38-CD90+CD45RA+ population present at diagnosis was also present at the time of relapse (Figure 4.12A,B; Figure 4.13). In fact, the majority of HSPC at relapse were CD90+CD45RA+, suggesting that this might be the population that propagated the relapse. By contrast, the expanded GMP compartment returned to normal levels post-BMT in patients that have not relapsed (Figure 4.13, Figure 4.14A-D). The only exception to this is patient ID2. In this case, the expanded GMP compartment post-BMT persisted, however, this patient developed severe GVHD and died from this 9 months post-BMT and therefore it is uncertain whether this represents disease burden or the physiological response to GVHD. Interestingly, for the patients that have relapsed no additional mutations were detected at the time of relapse and the VAF at relapse was similar to that at diagnosis (Figure 4.14A).

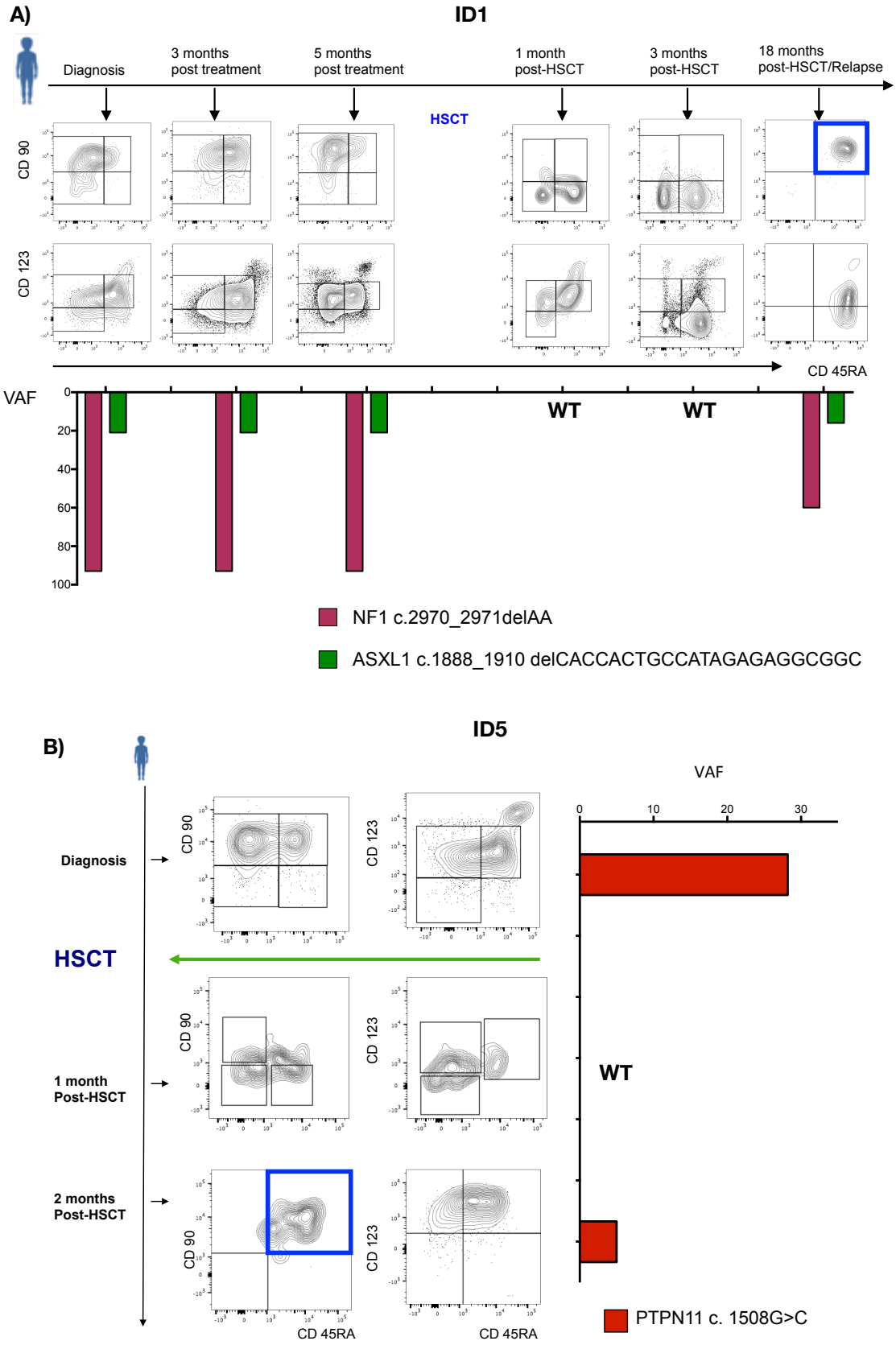


Figure 4.12: Serial flow cytometry and molecular analysis (variant allele frequency, VAF) of BM samples from 2 patients with JMML who relapsed post-BMT. A) Samples for patient ID1, B) Samples for patient ID5 from diagnosis to relapse

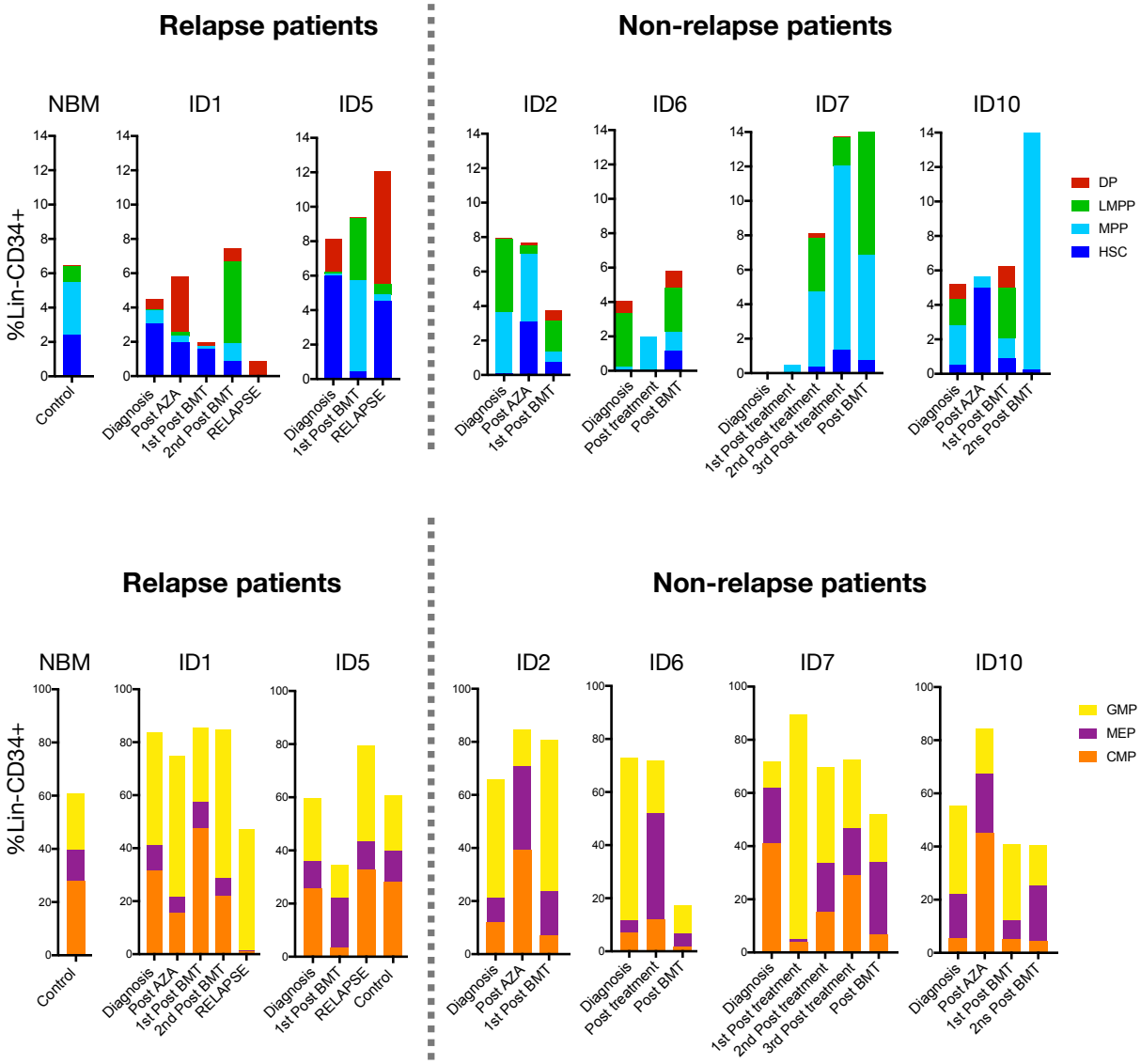
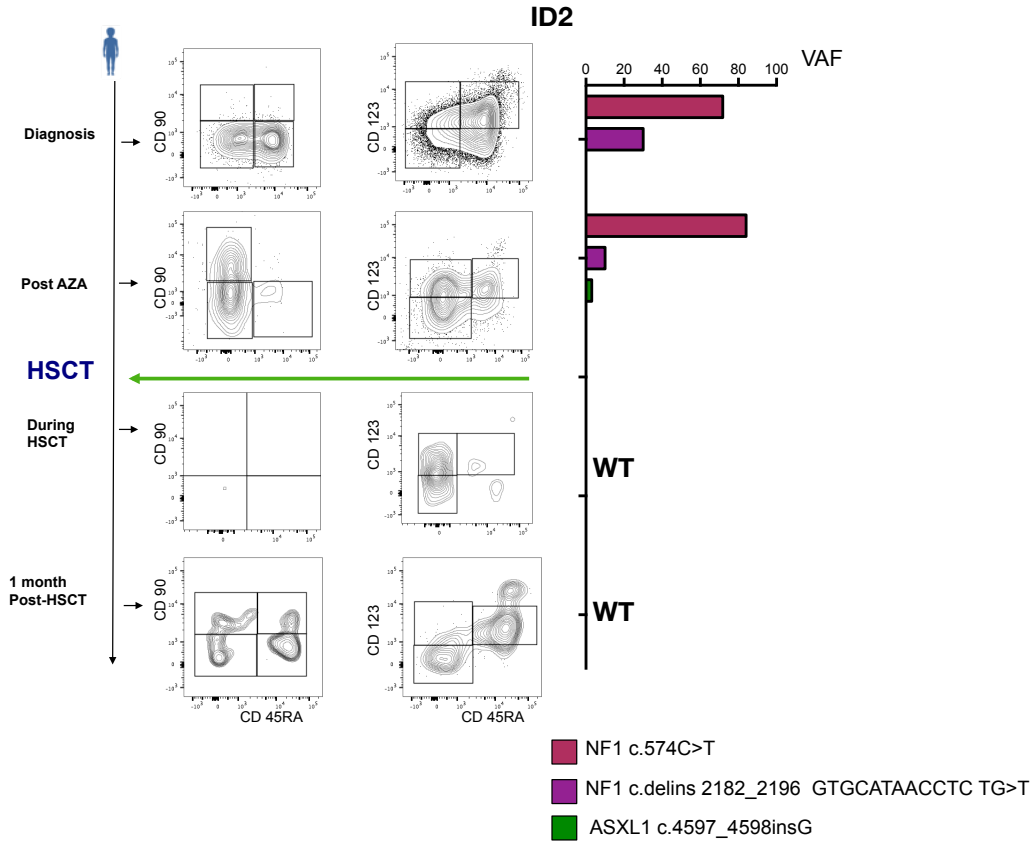
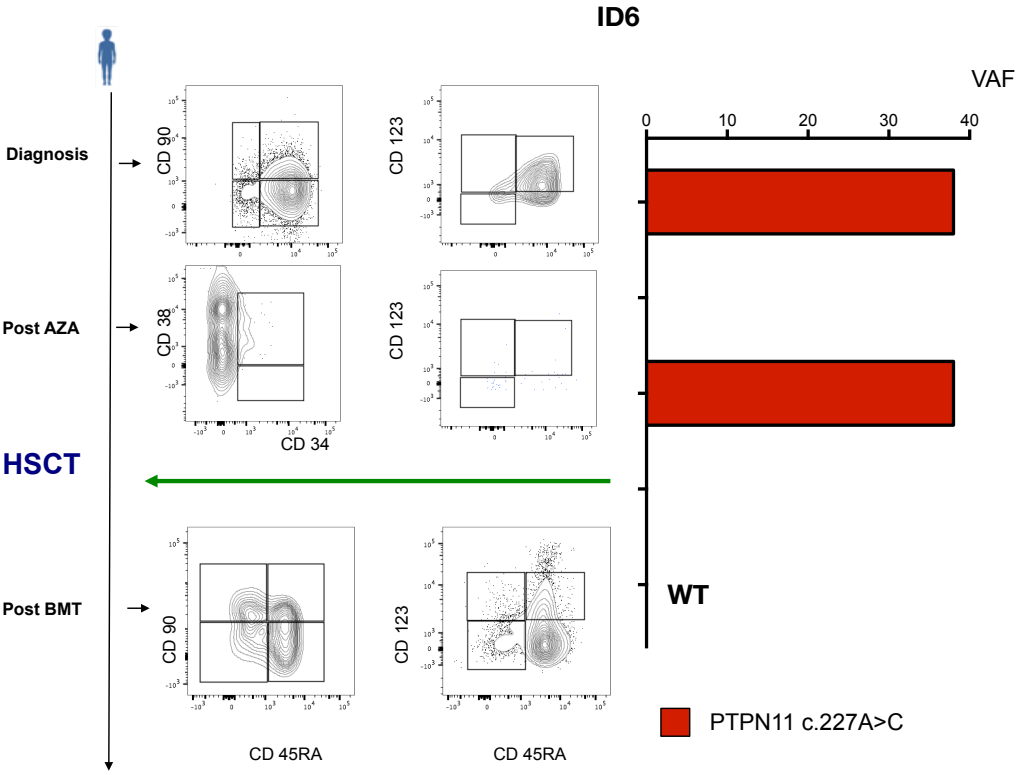
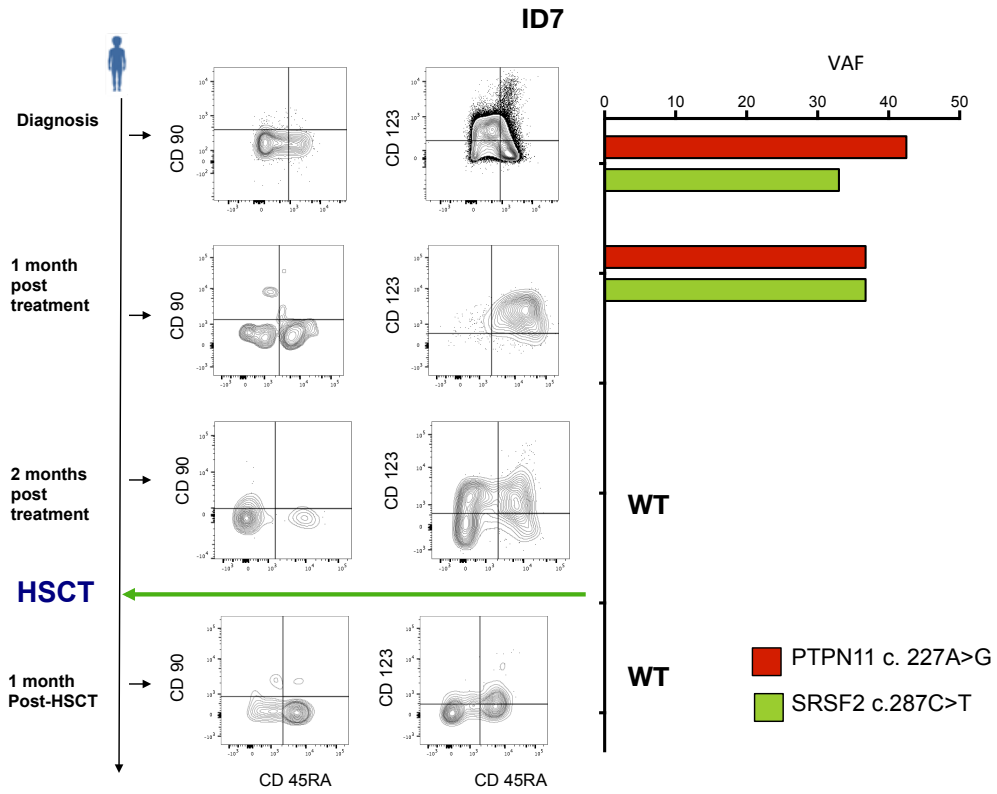


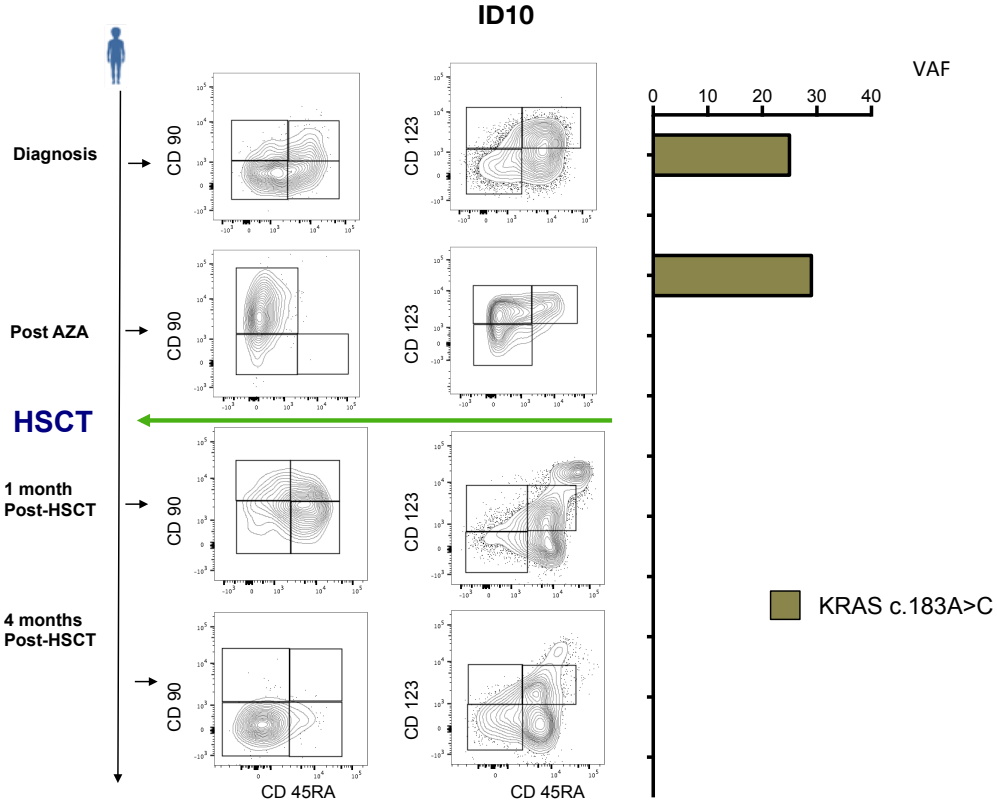
Figure 4.13: Frequencies of JMML HSPCs in serial samples from diagnosis and after bone marrow transplantation. NBM (paediatric bone marrow from healthy controls) Analysis of the CD34+CD38- compartment (top panel) and analysis of the CD34+CD38+ compartment (bottom panel)

Figure 4.14: Serial flow cytometry and molecular analysis (variant allele frequency, VAF) of samples from 4 patients with JMML (ID2, ID6, ID7, ID10) who relapsed post-BMT.






ID7
Diagnosis
1 month post treatment
2 months post treatment
HSCT
1 month Post-HSCT
CD 90
CD 123
CD 45RA
VAF
0
10
20
30
40
50
WT
WT
PTPN11 c. 227A>G
SRSF2 c.287C>T


ID10
Diagnosis
Post AZA
HSCT
1 month Post-HSCT
4 months Post-HSCT
CD 90
CD 123
CD 45RA
VAF
0
10
20
30
40
KRAS c.183A>C

Flow cytometry analysis after azacitidine (AZA) treatment showed a relative expansion of the HSC compartment and reduction of the LMPP compartment (particularly in ID2 and ID10), however this was not significant when all samples were analysed together (HSC: diagnosis vs post-AZA, p= 0.79; LMPP: diagnosis vs post-AZA, p= 0.1255) (Figure 4.15). These results are interesting but since only 5 paired samples were available for analysis during my project, further samples will be needed to assess the impact of AZA on haematopoiesis in JMML.

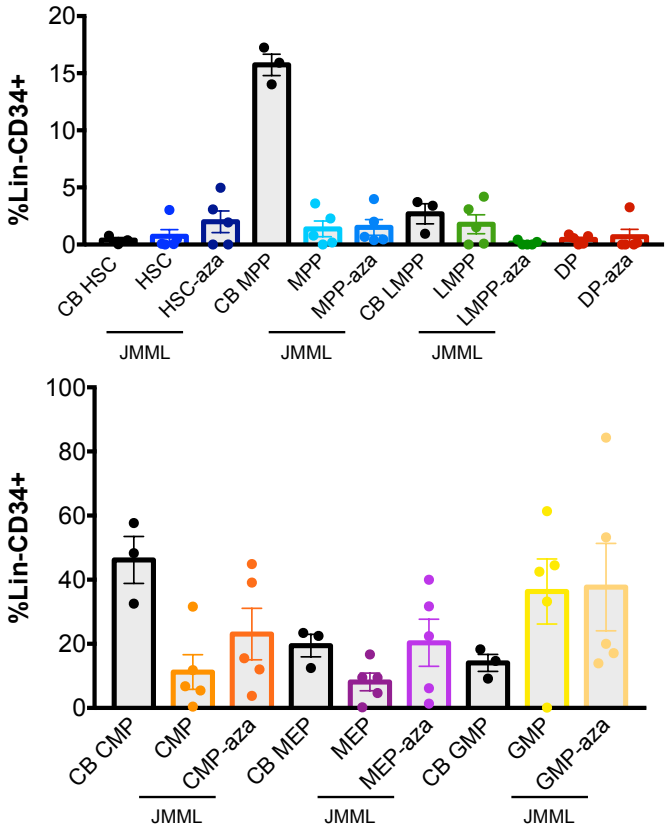


Figure 4.15: HSPC analysis shown as the % of Lin-CD34+ cells in patient samples at diagnosis and after AZA treatment (JMML n=5 paired samples, CB=cord blood control n=3). There were no significant differences between pre- and post-AZA treatment for the frequencies of any of the populations shown.

4.7 Lineage output of JMML HSPC

In order to characterise the JMML HSPC compartments at a functional level, I performed clonogenic assays to assess their myeloid and erythroid potential, liquid cultures for megakaryocytic potential and stromal cell co-cultures for lymphoid potential and compared this to normal controls. Where possible I used normal BM as a control but limited availability meant that for some assays I used normal cord blood.

4.7.1 JMML HSPC clonogenic capacity in short-term methylcellulose assays

Single cells from all HSPC compartments from JMML patient samples at diagnosis as well as normal controls were sorted into 96-well plates and the clonogenic output was assessed after 2 weeks using direct microscopy, cytospins and flow cytometry (details in chapter 2). Overall clonogenicity in all of the HSPC populations was similar in JMML and normal controls (Figure 4.16). The most notable abnormalities in JMML were the reduced erythroid output both compared to normal paediatric BM (Figure 4.16A) and cord blood (Figure 4.16B) controls and the presence of abnormal erythroid colonies in all of the patients (Figure 4.17). Three distinct types of erythroid colonies were seen: 1) 'normal' erythroid colonies similar to typical erythoid colonies from normal controls; 2) 'dysplastic' erythroid colonies that were large and had very poor haemoglobinisation; and 3) 'mixed' erythroid colonies that were typically very small, with very bright haemoglobinisation but also containing a large number of macrophages (Figure 4.17).

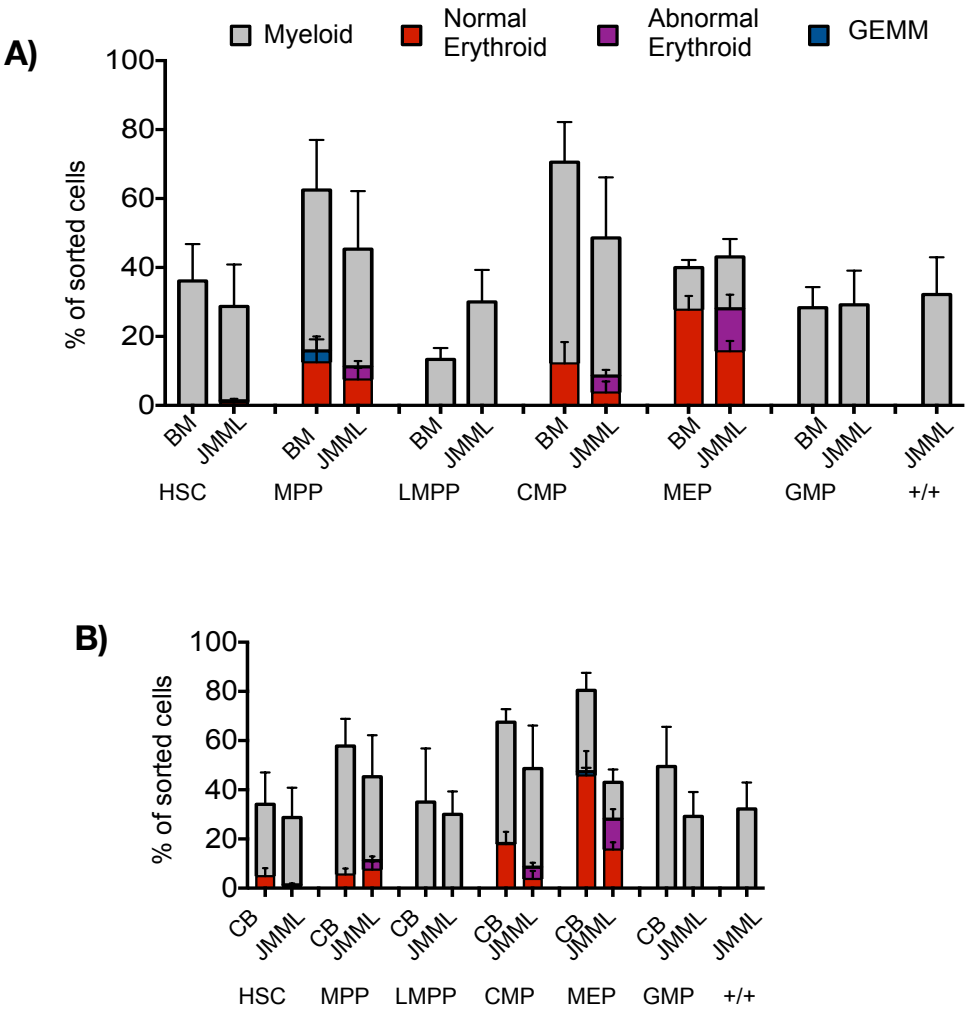


Figure 4.16 A) Clonogenic output of single HSC and progenitor cells sorted into methylcellulose; expressed as the % of sorted cells which generated a colony and the type of colony generated. Mean + SEM for 4 JMML BM samples compared to 3 normal (Paediatric BM; BM) controls. +/+ indicate the aberrant CD34+CD38-CD90+CD45RA+ population

B) Clonogenic output of single HSC and progenitor cells sorted into methylcellulose; expressed as the % of sorted cells which generated a colony and the type of colony generated. Mean + SEM for 4 JMML BM samples compared to 4 normal (cord blood; CB) controls. +/+ indicate the aberrant CD34+CD38-CD90+CD45RA+ population

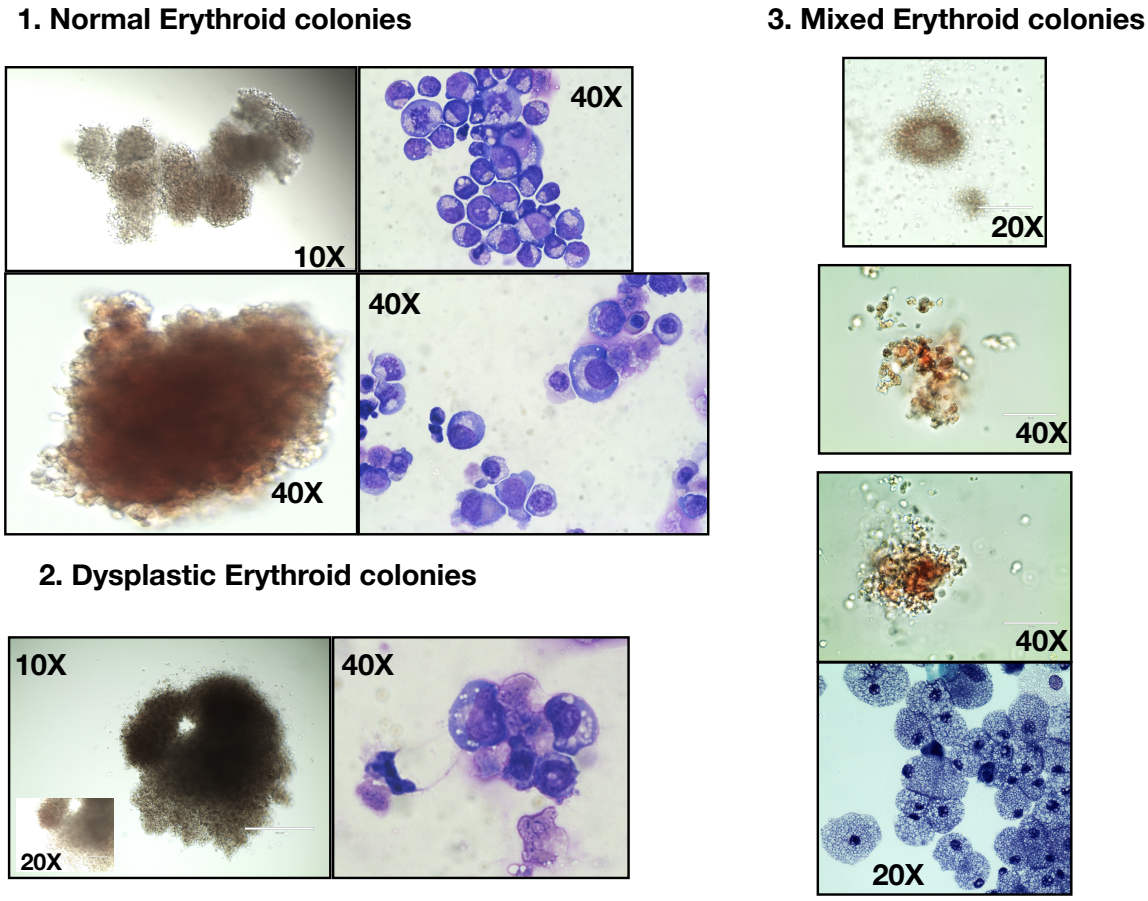


Figure 4.17: Different types of abnormal JMML erythroid colonies photographs under direct microscopy and cytospins preparation: 1) Erythroid colonies with normal appearance; 2) Dysplastic erythroid colonies; 3) Mixed erythroid colonies

The aberrant CD34+CD38-CD90+CD45RA+ population had an exclusively myeloid output and had similar clonogenicity to both normal and JMML HSC (Figure 4.16). In addition, review of the cytospins of the myeloid colonies grown from JMML samples showed that they were composed of morphologically normal and abnormal macrophages (e.g. double nuclei, nuclear bridging, mitotic figures) and, in contrast to normal controls, neutrophils were not present (Figure 4.18). The colony size was heterogeneous but similar to normal controls. (Figure 4.17 and Figure 4.18). Immature MNC were often seen mixed along with the mature macrophages in JMML CFU-M colonies and flow cytometry confirmed that a proportion of cells in plucked colonies retained CD34+ expression (Figure 4.18).

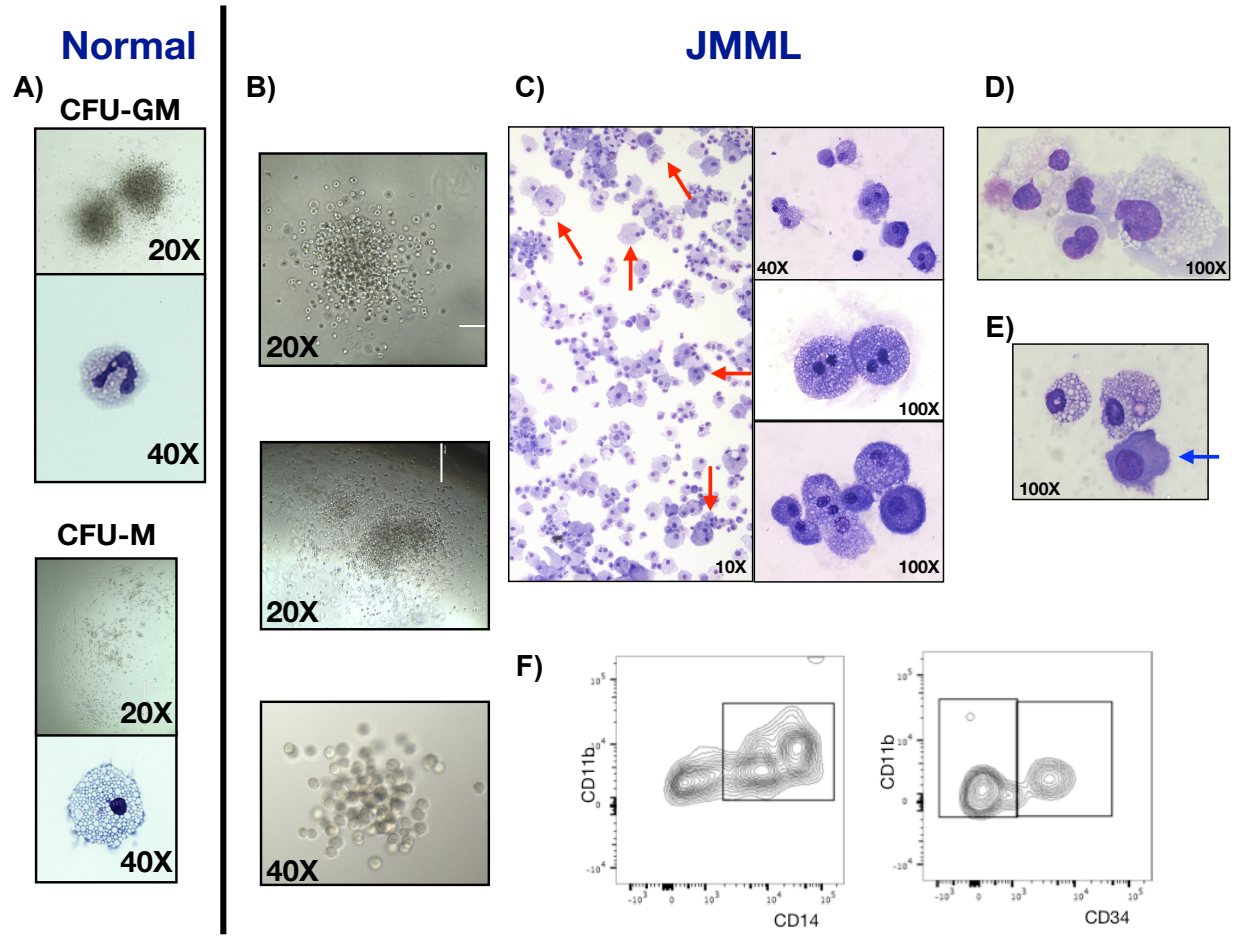


Figure 4.18: A) Representative photographs of myeloid colonies from normal cord blood controls as seen under direct microscopy and photographs cytospins of the same cells. B) Representative photographs of CFU-M JMML colonies as seen under direct microscopy. C) Abnormal features of CFU-M colonies observed in culture from JMML patient sample: Macrophages with double nucleus suggestive of dividing macrophages in culture. Red arrows indicate the abnormal cells in lower magnification. D) nuclear bridging between 2 macrophage cells E) abnormal immature mononuclear cells seen in culture along with mature macrophages (indicated by blue arrows) F) Examples of flow cytometry analysis for JMML colonies demonstrating colonies with myeloid output (CD11b+,CD14+) and myeloid colonies that retained the CD34 antigen.

In the view of the association between raised HbF and JMML, I screened the different types of erythroid colonies for embryonic and fetal globin expression by qPCR (Figure 4.19). While erythroid colonies from normal cord blood expressed both alpha and gamma globin genes as well as a small amount of beta globin, in JMML many of

the colonies also expressed epsilon globin and the level of expression of gamma globin was very high in some colonies.

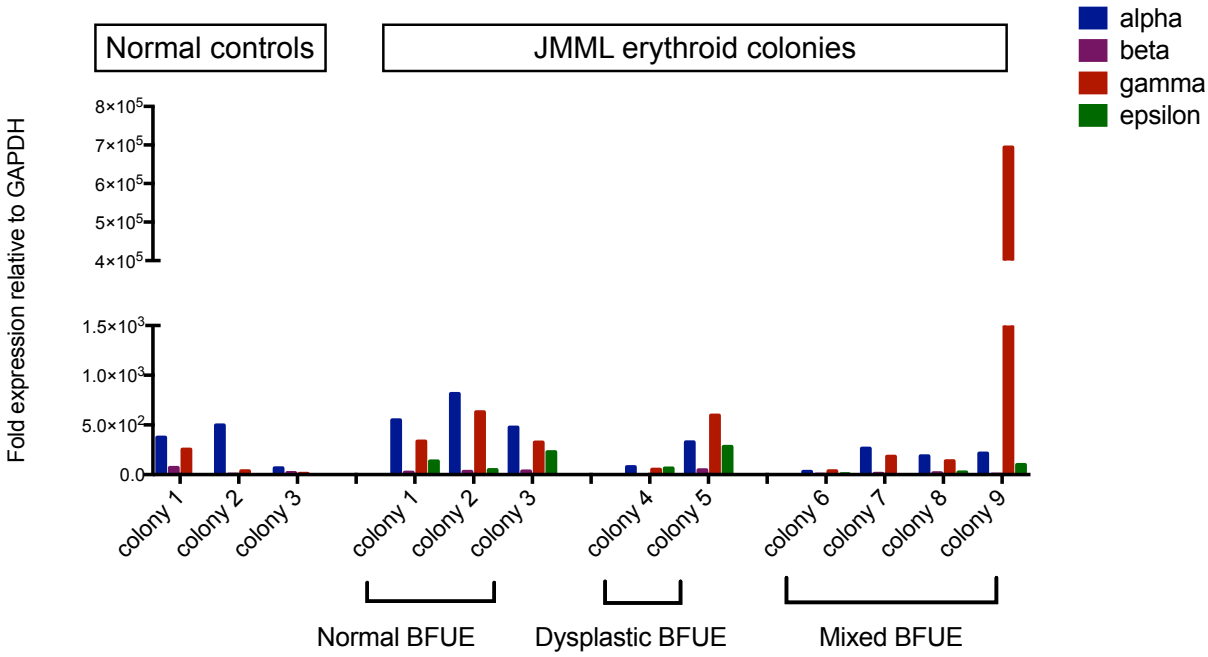


Figure 4.19: Expression of globin genes by qPCR in erythroid JMML colonies compared with controls. (CB = erythroid colonies derived from healthy donors cord blood samples)

4.7.2 Assessment of the megakaryocytic output of JMML HSPC in liquid culture

In order to assess the ability of JMML HSPC to give rise to megakaryocytes, single cells from each HSPC population were sorted into individual wells of a 96-well plate containing medium and cytokines designed to promote megakaryocyte/erythroid differentiation (371). The lineage of the cells generated after 14 days was assessed by flow cytometry. While normal cord blood MPP, CMP and MEP (n=2) gave rise to megakaryocyte and erythroid/megakaryocyte colonies in vitro in this system, the corresponding populations sorted from JMML samples (n=2) failed to give rise to

megakaryocytes (Figure 4.20 and Figure 4.21). Unfortunately, I had insufficient time to repeat these assays in a larger number of patients. In addition, no megakaryocytic cells were generated by the normal cord blood HSC, suggesting that further optimisation of this assay may be needed (e.g. by reducing the amount of erythropoietin which may have promoted erythroid differentiation at the expense of megakaryocyte differentiation).

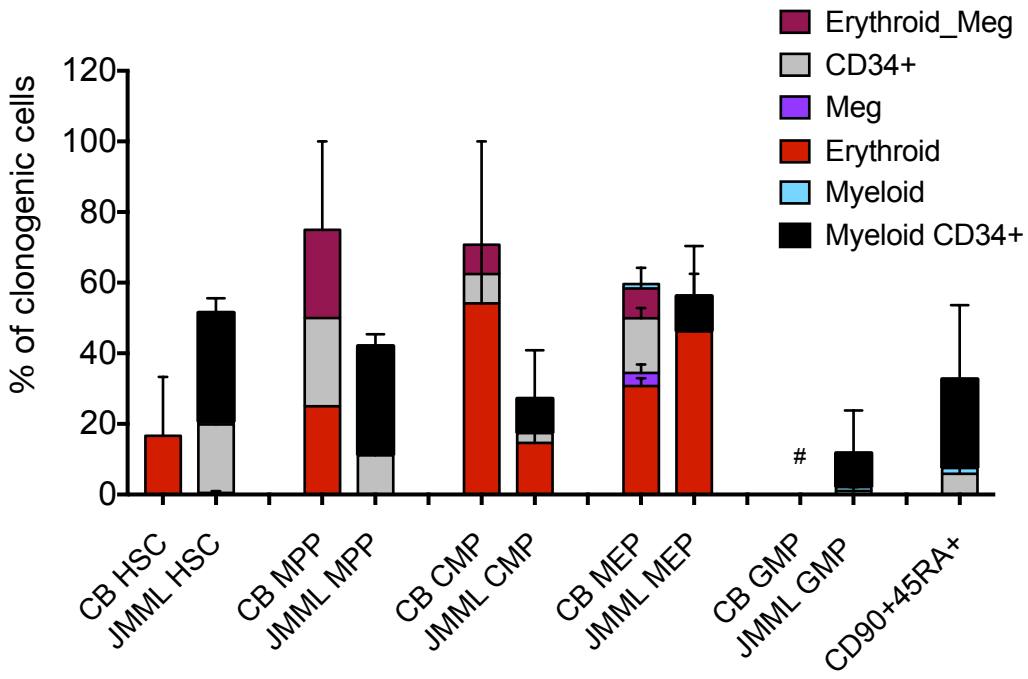


Figure 4.20 Clonogenic output of JMML HSPCs compared with normal controls (JMML n=2, CB n=2) from single cell liquid meg-primed culture . # Indicates no cells sorted for this population in CB

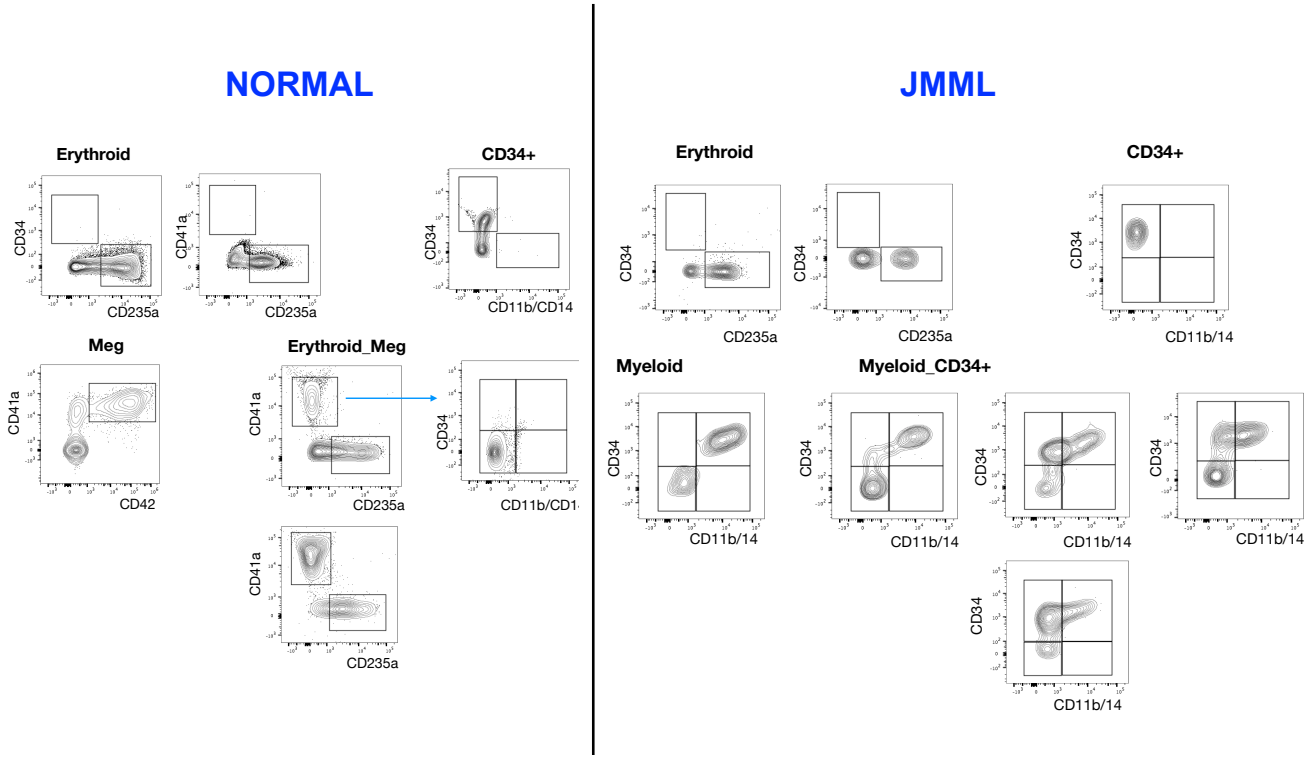


Figure 4.21: Examples of flow cytometry analysis of cord blood HSPCs sorted in single cells and cultured in mega/erythroid cell culture (left). Examples of flow cytometry analysis of JMML HSPCs in single cells and cultured in mega/erythroid cell culture (right) showing lack of megakaryocytic potential

4.7.3 Lymphoid output of JMML HSPC

Lastly, in order to assess the B-lymphoid output of JMML HSPC I performed MS5 co-culture assays, one of the 'gold standard' assays for lymphoid potential (311) . JMML HSC and LMPP (n=4) had reduced B-lymphoid output compared to controls (n=4) (Figure 4.22). The aberrant CD90+CD45RA population, was also able to produce both CD19+ B cells and myeloid (Figure 4.22).

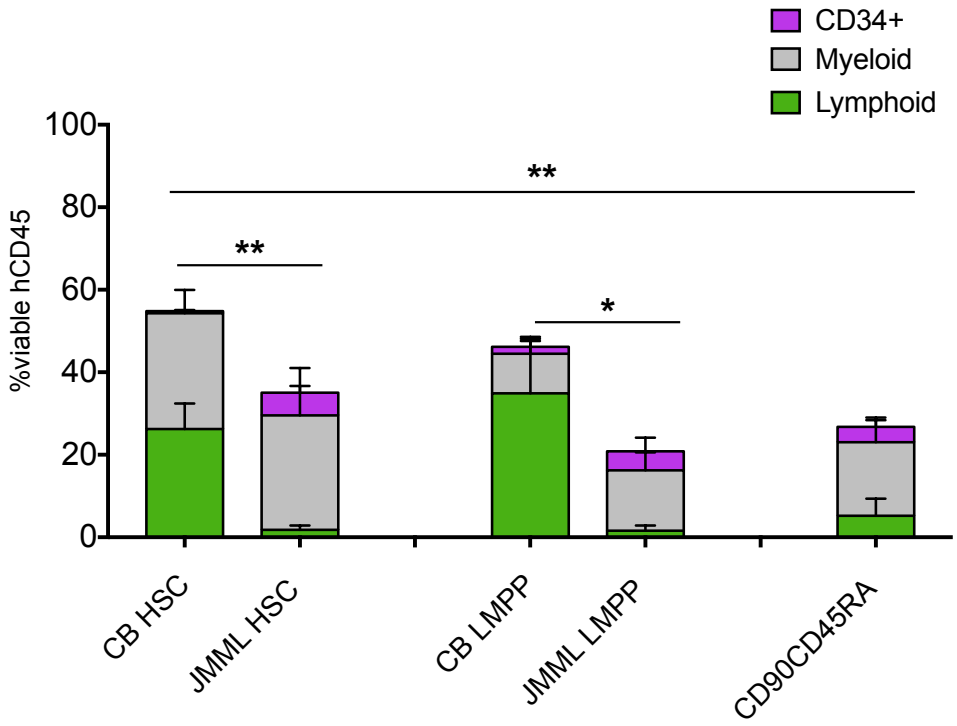


Figure 4.22: Clonogenic output of JMML HSPCs compared with normal controls using the MS5 co-culture cell assay. 100 cells sorted per well and lymphoid output shown after 4 weeks in culture. (JMML n=4, CB n=4) . * indicates significant differences in lymphoid output between populations

4.8 In vitro stem cell assays of JMML HSPC (long-term culture initiating-cell assay– LTC-IC)

In order to assess the stem cell capacity of JMML HSPC in vitro, I used the gold standard LTC-IC in vitro stem cell assay (118) and FACS-sorted 100 cells from each HSPC population into individual wells of pre-prepared stromal cell cultures as described in chapter 2. Each well was inspected weekly and at the end of the 6 weeks co-culture period all positive and negative wells were re-plated in methylcellulose. At the end of the assay all wells were scored for the presences of colonies using direct microscopy as well as cytospins in order to confirm the identify of colonies seen. The JMML HSC compartment appeared to retain self-renewal capacity, and some stem cell capacity was observed in the MPP population but this

was not significantly different from controls (Figure 4.23). Colonies derived from JMML samples were myeloid, whereas colonies derived from CB HSCs had multilineage potential as expected from a normal stem cell (Figure 4. 23). This suggests that although JMML HSC retain stem cell capacity, these cells have a myeloid bias with reduced or absent erythroid differentiation capacity in line with the short term in vitro data. However, one of the limitations of this assay is that it does not allow lymphoid and megakaryocytic output to be assessed.

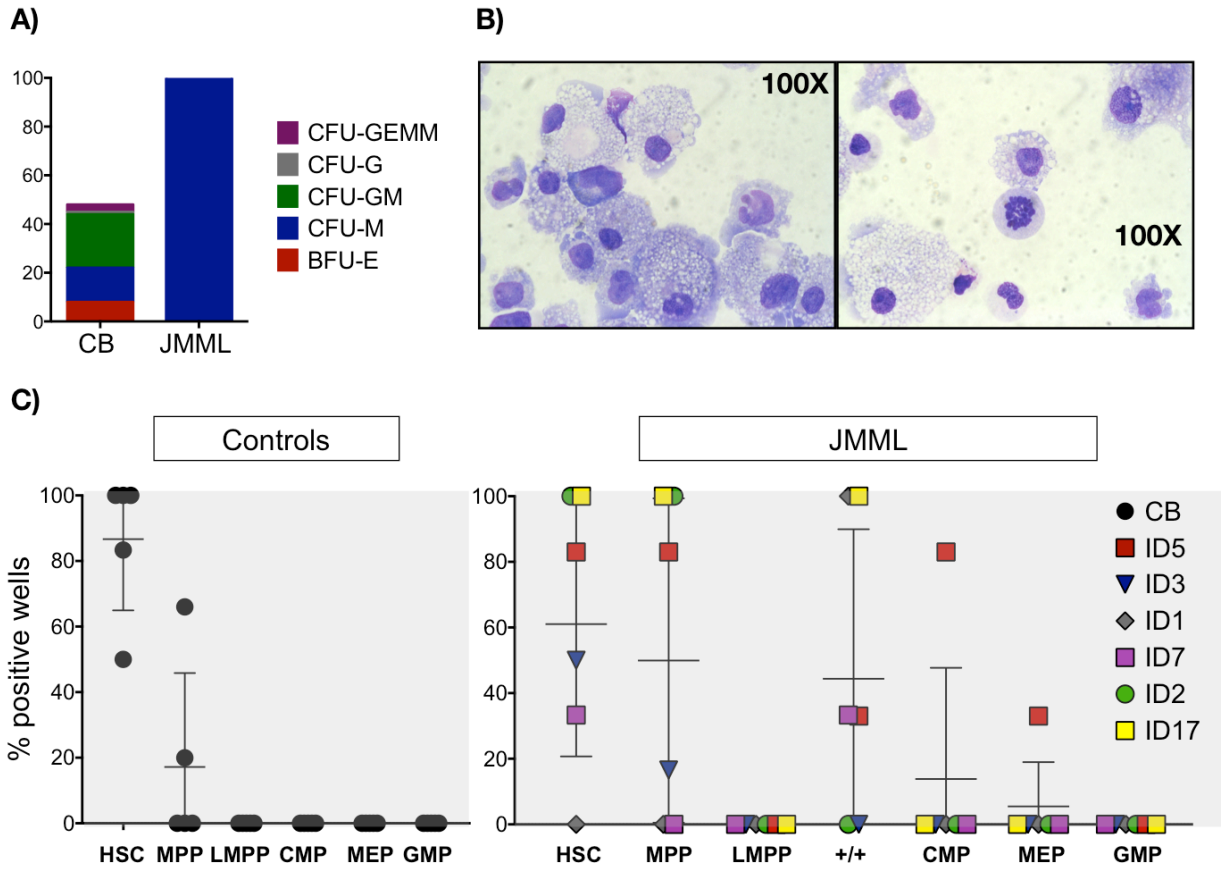


Figure 4.23: A) Different types of colonies observed from CB HSC following 8 weeks in culture with the LTC-IC assay and types of colonies of observed from JMML HSCs following 8 weeks in culture (JMML n=6, CB n= 6) B) Photographs of cytospins of colonies from JMML HSC following 8 weeks in LTC-IC culture. C) HSPC LTC-IC output in CB contros and JMML HSPC.

Although there was some heterogeneity between patients, the JMML HSC and CD90+CD45RA+ populations had the highest stem cell potential compared with controls (Figure 4.23) suggesting that the aberrant CD90+CD45RA+ population has at least some stem cell properties. All but one of the JMML HSC samples (ID1) had LTC-IC activity and, interestingly, in this patient the LTC-IC activity was restricted to the aberrant CD90+CD45RA+ population (Figure 4.23).

4.9 In vivo long term stem cell potential of JMML HSPCs

In order to assess the stem cell capacity of JMML HSC,and the aberrant Lin-CD90+CD45RA+ populations I used an NSG-mouse based xenograft assay previously validated for assessing the stem cell properties of candidate stem cell populations (118) . Flow-sorted HSC and GMP from normal cord blood and HSC, CD90+CD45RA+ and GMP from JMML samples were transplanted into NSG mice via tail vein injection as described in chapter 2. These populations were selected as the most likely compartments to retain stem cell capacity based on the data in 4.8 and chapter 5. GMP were used for transplants as a counterpart of the CD90+CD45RA+ population within the CD34+CD38+ compartment. Equivalent numbers of control and JMML cells were used and transplanted simultaneously.

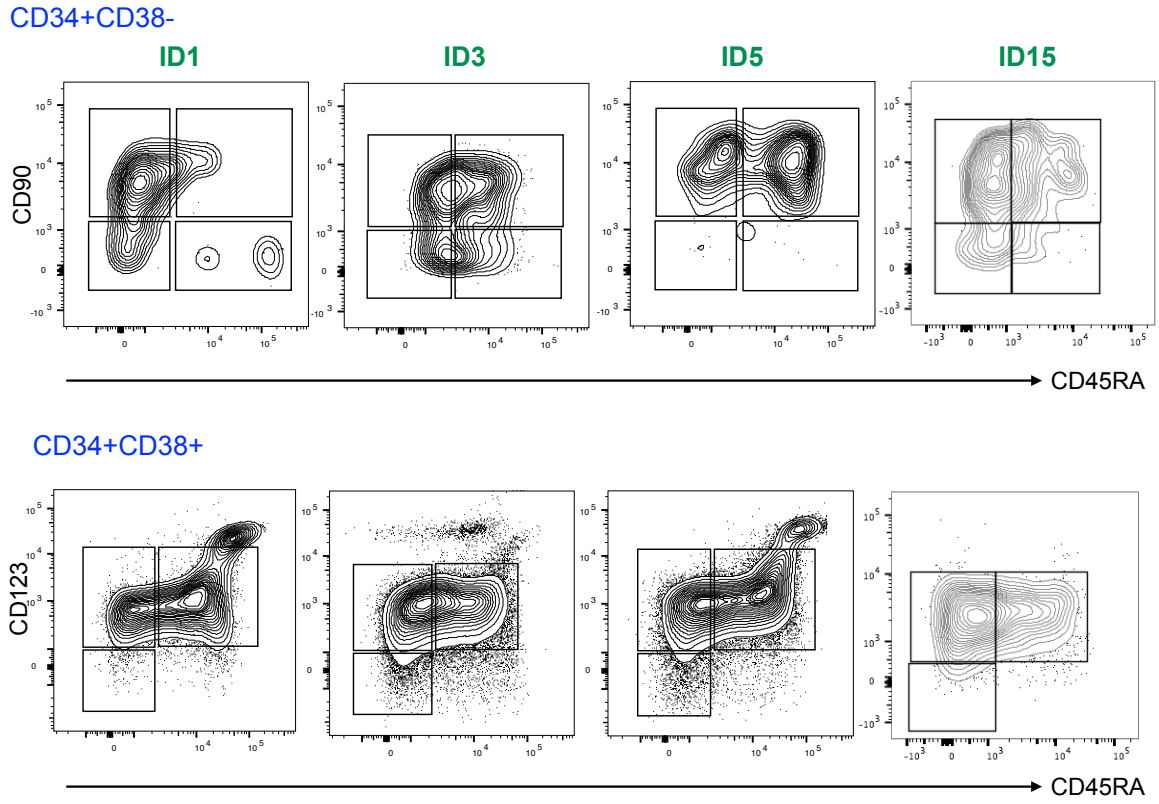


Figure 4.24: Flow cytometry profiles of patients used for xenotransplation. (top panel: CD34+CD38-, bottom panel CD34+CD38+) JMML samples n=4, CB donors n=4 (CB profiles not shown)

4.9.1 Peripheral blood engraftment of JMML HSPC in NSG mice

Engraftment was assessed using flow cytometry at weeks 5, 10 and16-22 post-transplant. JMML HSC from all 4 patients engrafted as shown by the presence of human CD45+ cells in the peripheral blood of the transplanted mice (Figure 4.25). However, engraftment of the aberrant CD90+CD45RA+ and GMP population was heterogeneous among different patient samples. For ID1 the engraftment potential was restricted to the HSC population, whereas in ID3 the CD90+CD45RA and GMP population engrafted (this patient had particularly aggressive disease and died in blastic crisis during BMT due to disease progression); ID5 and ID15 both had HSC and CD90+CD45RA+ engraftment with patient ID5 also demonstrating some late

engraftment from the GMP population (Figure 4.25). For patient ID15 1/3 HSC demonstrated robust human engraftment, and made the recipient unwell resulting to early sacrifice at 13 weeks of age. Transplantation with BM samples from patient ID5 was performed twice: once with 2000 cells per population and once with 250 cells transplanted per population, in order to assess whether fewer cells were enough to propagate the disease in a similar manner, which was indeed the case. A summary of the experiments is shown in Table 4.1.

Table 4.1: Total number of cells transplanted per patient sample

Population	Patient ID	Number of cells
HSC	ID1	250
CD90+CD45RA+		225
GMP		2,000
CD34+		20,000
HSC	ID3	375
CD90+CD45RA+		1,500
GMP		1,500
CD34+		12,500
HSC	ID5	2,000
CD90+CD45RA+		2,000
CD34+CD38+		2,000
CD34+		50,000
HSC	ID5	250
CD90+CD45RA+		250
GMP		250
HSC	ID15	2,000
CD90+CD45RA+		2,000
GMP		2,000
CD34+		12,500

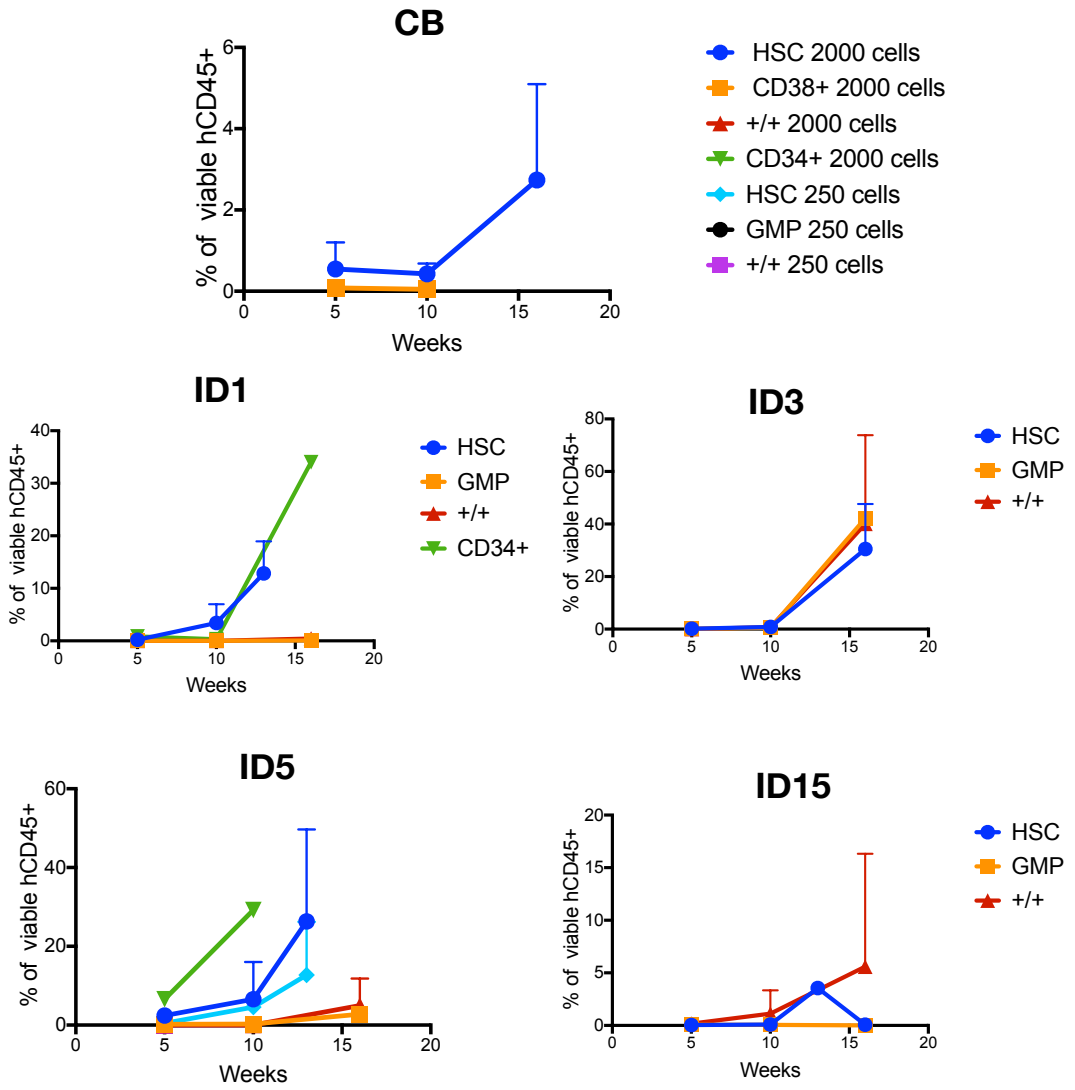


Figure 4.25: Percentage of human cell (CD45) engraftment at different time points during transplantation (weeks 5, week 10 and 16-22, terminal readout) ID indicates the different patient samples that were used for transplantation (JMML n=4, CB n= 4) (Experiment performed with Benjamin Povinelli)

Lineage analysis of the engrafted human cells showed that the majority of cells were myeloid with almost no lymphoid engraftment from the JMML HSC although, interestingly the CD90+CD45RA+ population did produce lymphoid cells in 3/4 of the patient samples (ID3, ID5 and ID15) in line with the in vitro assays (Figure 4.26).

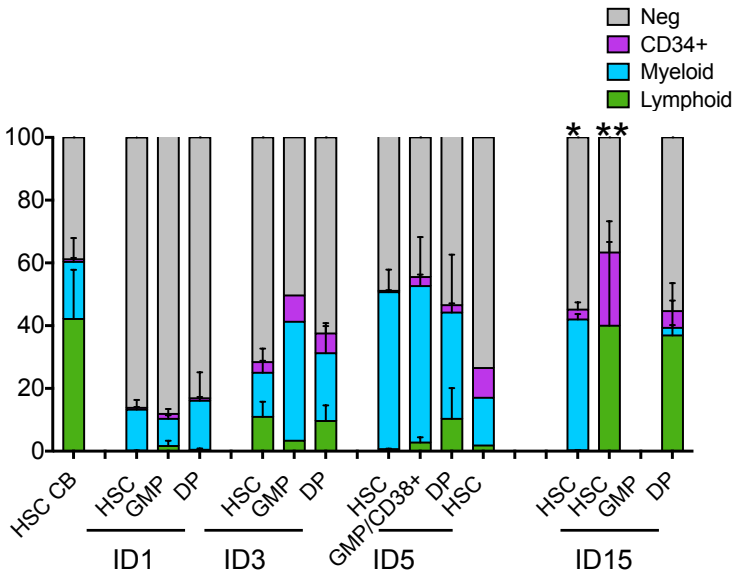


Figure 4.26: Lineage output of the human cells identified in the peripheral blood. * Indicates that became unwell with leukaemia from patient ID15, ** Indicates transplantation with lower number of cells (200). CB samples n=4, JMML samples n=4 (Experiment performed with Benjamin Povinelli)

4.9.2 BM engraftment of JMML HSPC in NSG mice

BM engraftment was consistent with the findings in the peripheral blood. All of the JMML HSC engrafted in the BM, the CD90+CD45RA+ population engrafted in 3/4 patients (ID3, ID5, ID15) and the GMP population in 2/4 patients (ID3, ID5) (Figure 4.27).

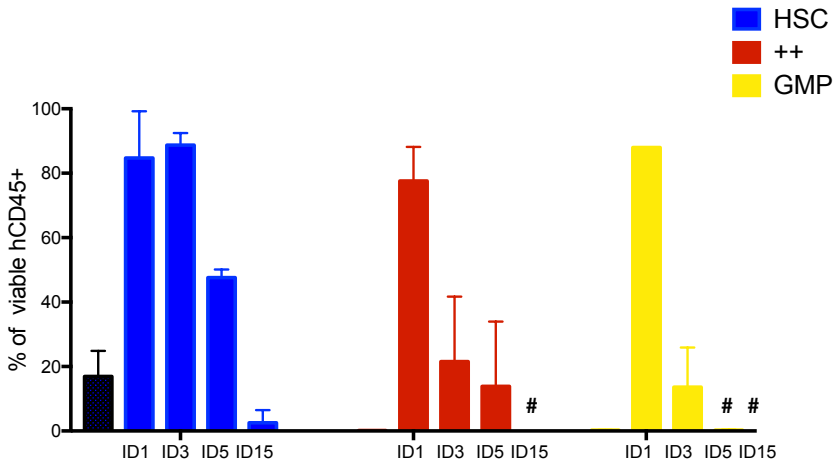


Figure 4.27: Summary of Bone Marrow engraftment of JMML HSPCs (JMML n= 4, cord blood n=4) A) % of viable human CD45+ cells detected per population transplanted. Each bar indicates individual JMML donors. Black bar indicates CB HSC control. # indicates undetectable/very low (< 0.5%) human CD45+ cells detected. (Experiment performed with Benjamin Povinelli)

The HSPC profiles derived from the engrafted NSG mice recapitulated the profiles of the original samples that were used for transplantation (Figure 4.28).

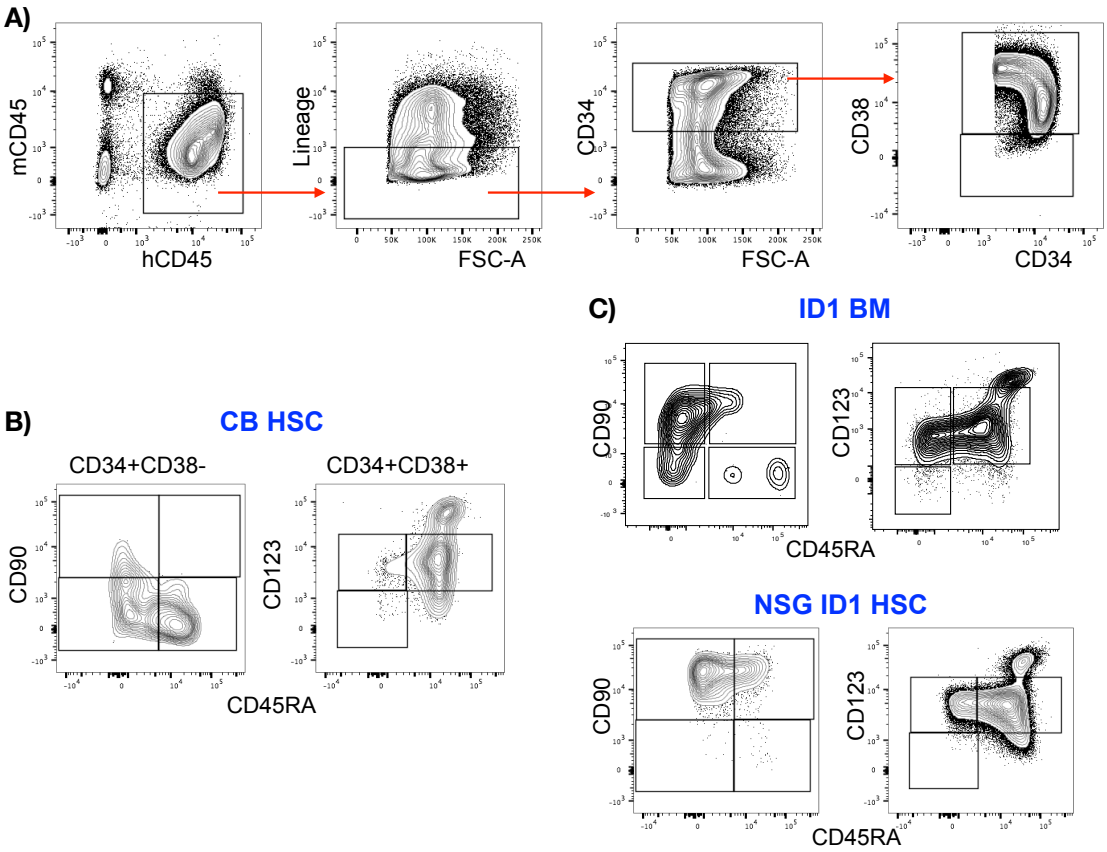


Figure 4.28: A) Gating strategy for analysis of engrafted HSPC in NSG mice. B) Representative example of HSPC profile in BM of mice that received cord blood control (gated on viable hCD45lLin-CD34+ cells. C) Top panel: HSPC profile of patient ID1 primary BM sample at diagnosis used for transplantation.(gated on viable lin-CD34+ cells). Bottom panel: representative example of HSPC profile in BM of mice that received HSC from patient ID1 (gated on viable hCD45+lin-CD34+ cells)

Although there was considerable heterogeneity between the engraftment potential of the different populations between patients, the main findings of these experiments were:

1. When transplanted into NSG mice, JMML HSC give rise to more aggressive disease than CD90+CD45RA+ cells or GMP. All mice receiving JMML HSC developed signs of leukaemia soon after transplant and had to be sacrificed 12-13 weeks post-transplant (Figure 4.29), whereas mice that received CD90+CD45RA+

cells or GMP survived without signs of being unwell until they were terminated at the end of the observation period despite robust engraftment.

2. Engraftment of JMML HSC recapitulated the human disease. Spleen size was increased in the mice engrafted with JMML cells, however the spleens of the mice that received JMML HSC were larger than the spleens of the mice that received the other two populations (Figure 4.29 B). As in the human disease, the mice that received JMML HSC, the BM was hypercellular compared to controls (Figure 4.29 C) and the mice were anaemic and thrombocytopenic at the time of sacrifice (Figure 4.30). Notably, at the end of the observation period, engrafted mice that remained well, including those transplanted with JMML cells, had normal blood counts.

3. Downstream progenitors including the CD90+CD45RA+ population have variable engraftment potential and despite their ability to reconstitute the disease, these do not cause such aggressive proliferation and BM infiltration.

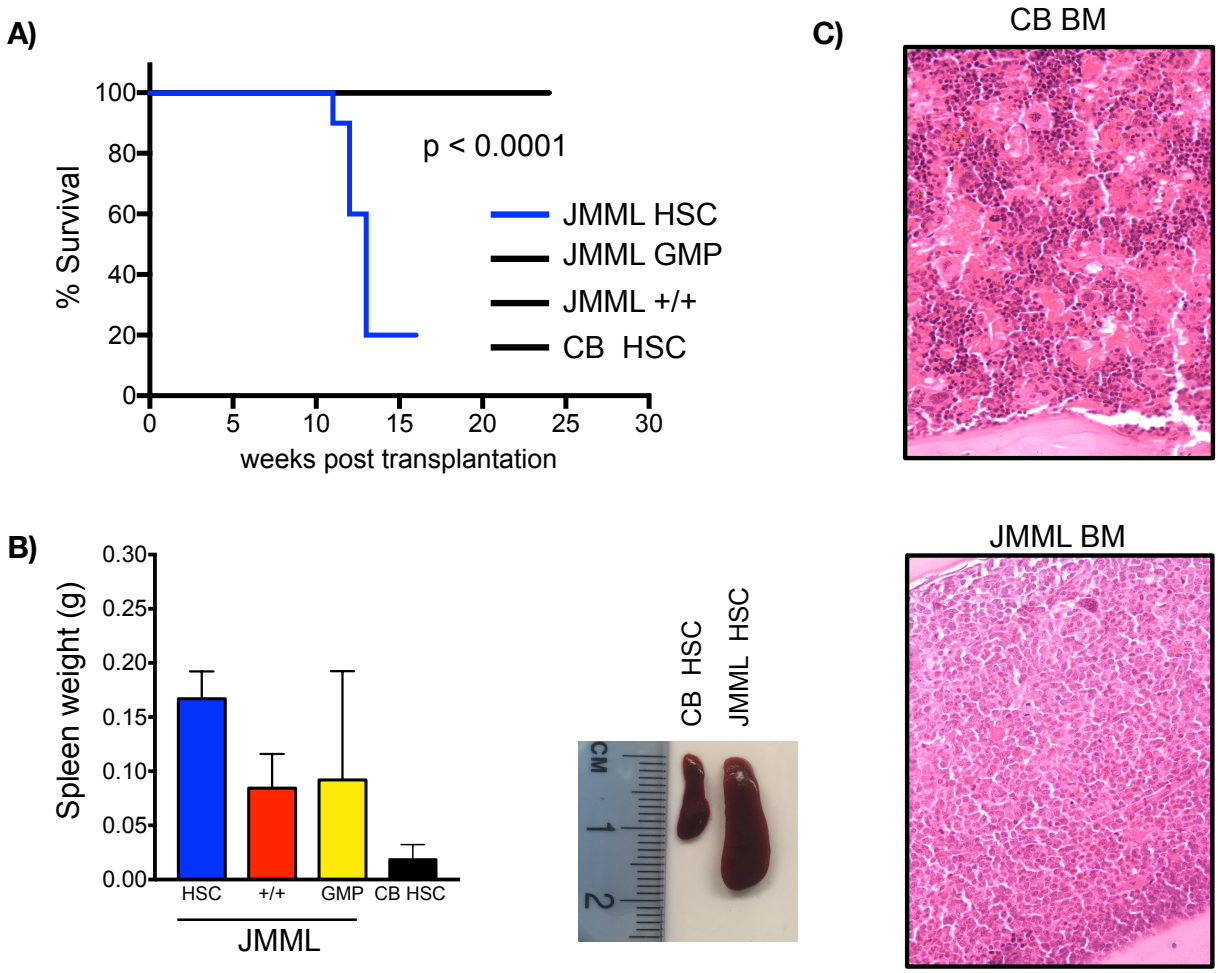


Figure 4.29: A) Survival post-transplantation with JMML HSPC compared to controls. B) Spleen size (in grams (left) and appearance (right) from recipients that received JMML cells compared to controls C) Photomicrograph of Giemsa-stained BM sections from engrafted mice transplanted with normal cord blood (top) or JMML HSC (bottom) (Experiment performed with Benjamin Povinelli)

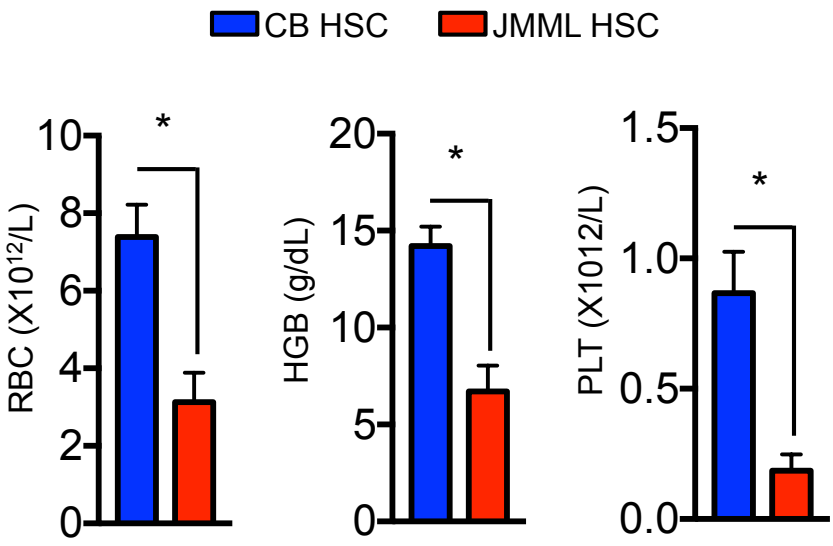


Figure 4.30: Peripheral blood count analysis of sick mice transplanted with JMML HSC at the time of sacrifice compared with CB controls at termination. (Experiment performed with Benjamin Povinelli)

Further analysis of the BM HSPCs from the recipient mice, revealed that it was not possible to extract the hierarchy of the disease propagating population (data not shown). In order to immunophenotypically further characterize the JMML LSC that propagates the disease in mice I performed analysis of the Lin-CD34+ BM compartment without using the CD38 to dissect the early and mature progenitors, since the data above demonstrate that in JMML self renewal capacity in some patients was not retained within the CD34+CD38- compartment. Using this approach I was able to clearly detect that the CD34+CD90+CD45RA+ was robustly detected within the bone marrow cells of the mice that displayed JMML engraftment compared to the controls (Figure 4.31). This analysis demonstrates that in JMML LSCs possibly resides within the CD34+CD90+CD45RA+ compartment. The CD34+CD90+CD45RA+ compartment appeared to be significantly increased in the recipients that received JMML HSCs (Figure 4.32) in line with the observations above the JMML HSC exhibit a much more aggressive disease phenotype.

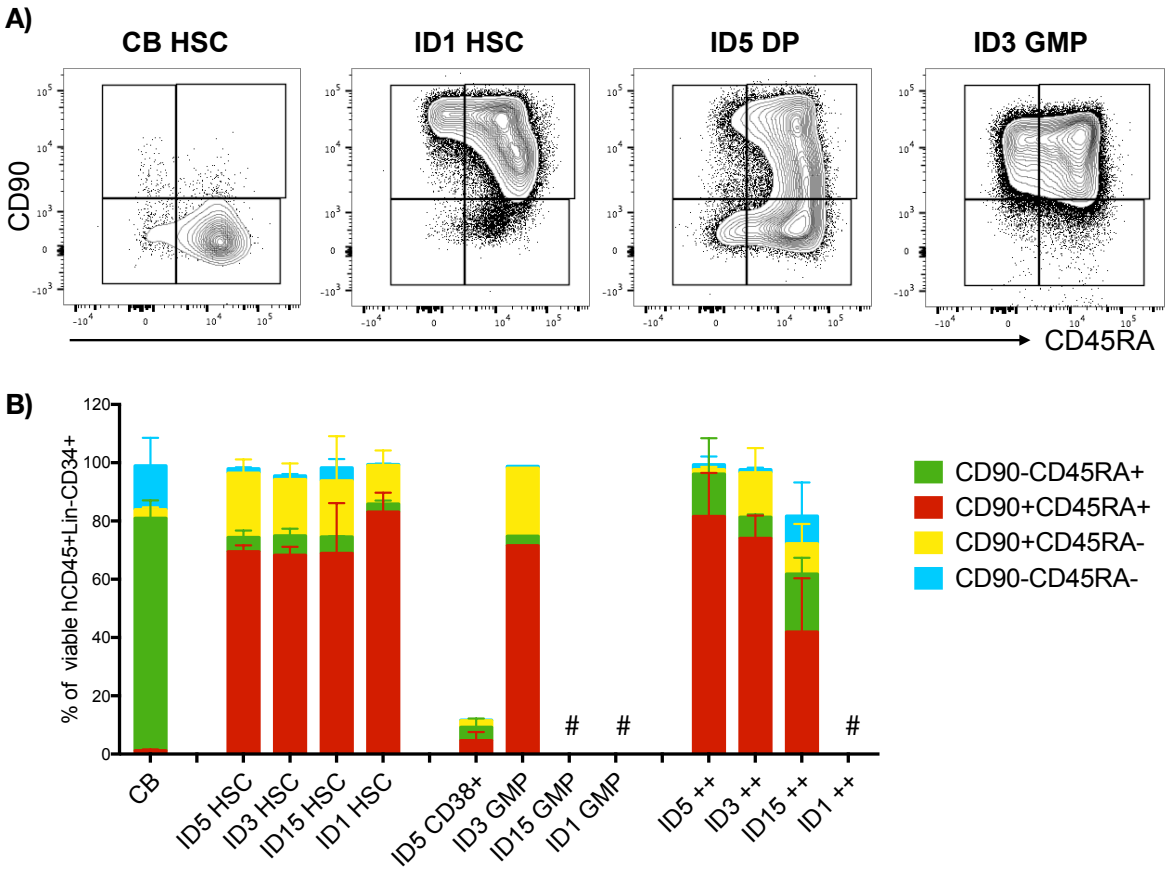


Figure 4.31: A) Profiles of CD34+CD90+CD45RA+ compartment in the bone marrow of mice that received CB HSCs, JMML HSCs, CD90+CD45RA+ and JMML GMPs. B) Summary of frequencies of the different compartments detected when the CD34+ fraction is analysed for CD90 and CD45RA expression, without the CD38 sub fraction

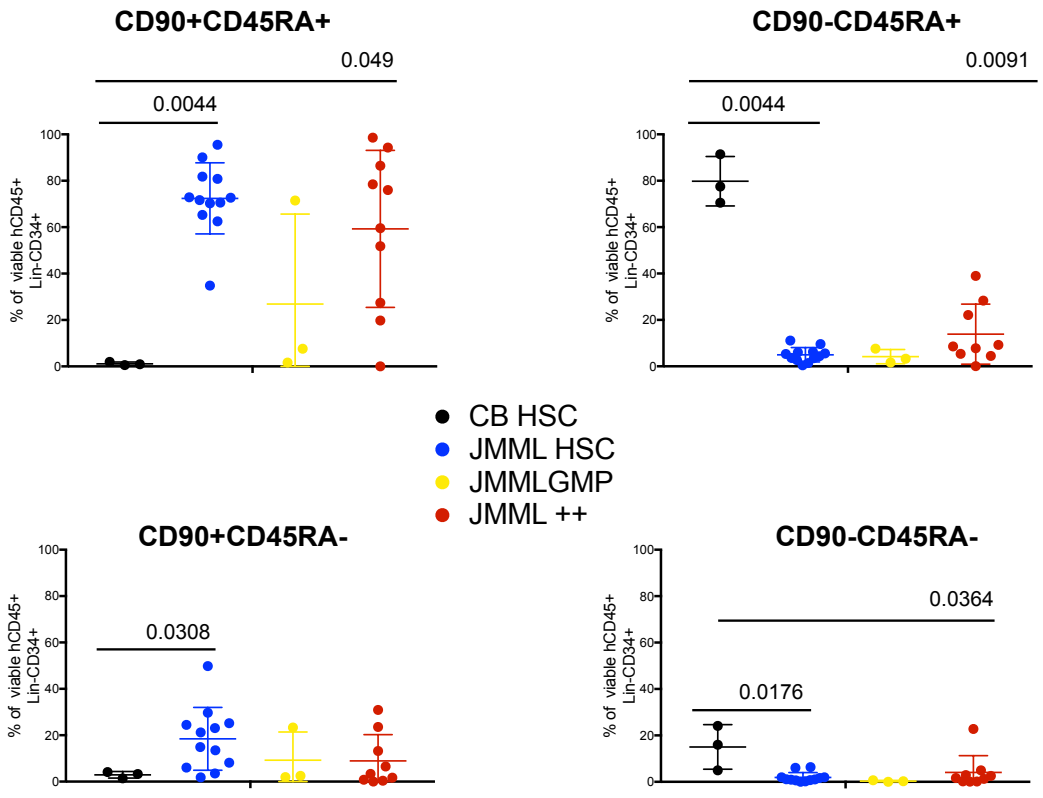


Figure 4.32: Frequencies of engrafted CD34+ human cells in NSG mouse BM based on the expression of CD90 and CD45RA from all populations transplanted .

JMML HSPC engrafted in the NSG mice without any conditioning in contrast to previous reports suggesting poor engraftment of JMML cells, considered to reflect differentiation of the cells post-transplant (372). A recent study by Yoshimi et al (215) have suggested that JMML cells are more likely to engraft in NSG-SGM3 mice that have transgenic expression of human stem cell factor (SCF), granulocyte-macrophage colony-stimulating factors (GM-CSF) and interleukin-3 in the NOD/SCID-IL2rγc^{null} background. However, I did not encounter any engraftment issues using the conventional approach described above. To date this the first study that has transplanted purified JMML sub-populations, rather than unfractionated mononuclear or CD34+ cells.

4.9.3 Secondary transplantation of NSG-engrafted JMML HSPC

CD34+ cells from engrafted NSG mice that received JMML HSC and GMP were used for secondary transplantation in order to more robustly address the question of the disease-propagating population(s) in JMML. The recipients that received cells from primary JMML HSC demonstrated robust engraftment (PB: 73.7%, BM 81%) in secondary mice compared to controls (PB: 0.022%, BM 1.4%) (Figure 4.33A). BM analysis of the engrafted populations revealed a very similar profile to the primary recipients with the CD34+CD90+CD45RA+ population again detected again in the BM of recipient mice (Figure 4.33B). In line with the primary transplantation data, JMML GMP failed to demonstrate sustainable engraftment in secondary mice. This supports the notion that JMML HSC have a superior ability to propagate the disease compared with downstream progenitors, although further replicate experiments would be required to confirm this.

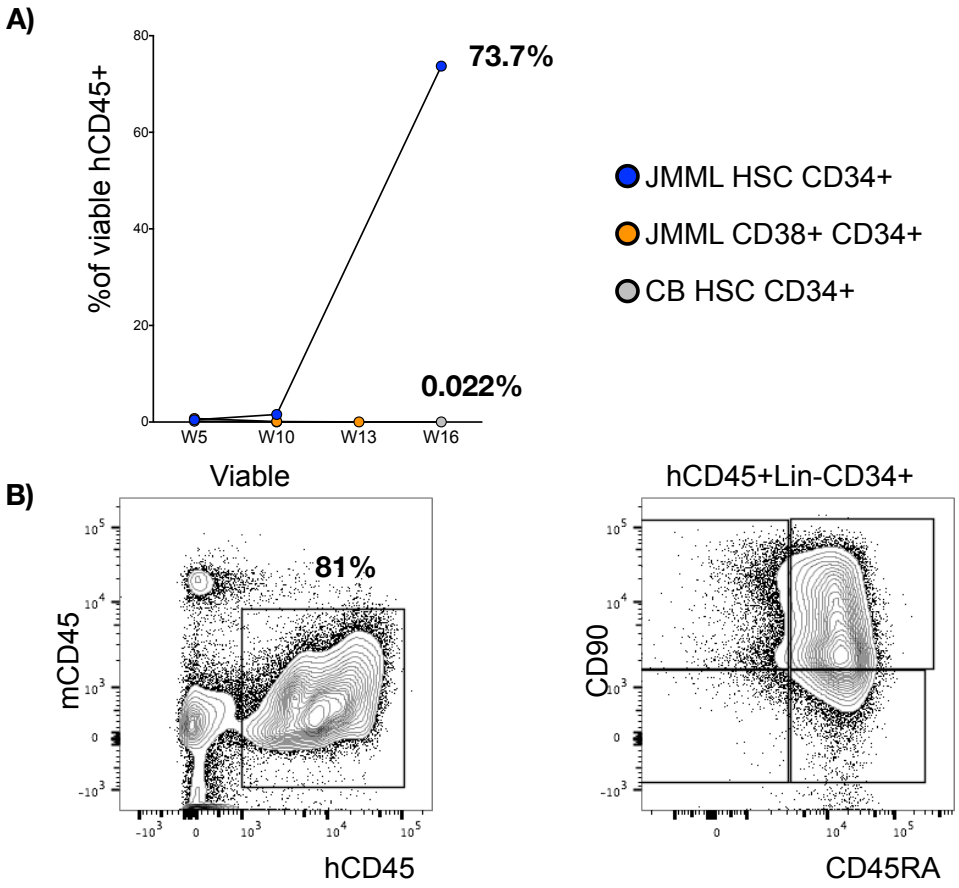


Figure 4.33: A) Engraftment of JMML CD34+ cells in PB following secondary transplantation to NSG mice (shown as the % of human CD45+ cells). B) Profiles of viable lin-CD34+ cells in BM from mice transplanted with JMML CD34+ cells derived following primary transplantation of JMML HSC. (Experiment performed with Benjamin Povinelli)

4.10 Discussion and Conclusions

Prior to this study there was very little known about the architecture of HSPC in JMML. These data demonstrate that JMML HSPC differ from AML; while in AML a series of molecular transformations occur serially in an early progenitor causing differentiation block and uncontrolled proliferation in downstream myeloid progenitors; in JMML HSPC retain the structure of the haematopoietic hierarchy, with all or most HSPC compartments detected in all patient samples. However the architecture of these compartments is abnormal, as shown by expansion of the HSC

and GMP compartments as well as the presence of an aberrant CD34+CD38-CD90+CD45RA+ population that does not exist in normal controls. This population appears to have some correlation with some of the disease bad prognostic features and is present in patients that relapse.

In line with the disease clinical features, the JMML clonogenic output is predominantly myeloid with reduced and abnormal erythroid differentiation, reduced/absent megakaryocytic output and reduced lymphoid potential. The aberrant CD90+CD45RA+ population appears to have bi-lineage potential, which perhaps indicates that this population may derive from the JMML HSC but still sits among the primitive haematopoietic progenitors.

When the stem cell activity of the JMML HSPCs is assessed in vitro and in vivo; it appears that there is inter- patient heterogeneity of the progenitor compartments that display aberrant stem cell activity; however the JMML HSC is the population that consistently gives rise to a more aggressive disease phenotype in all cases. BM analysis of mice transplanted with JMML revealed the presence of the aberrant CD34+CD90+CD45RA+ population in all cases that demonstrated human cell engraftment. Interestingly immunophenotypic analysis of patient serial samples also identify the presence of this population at relapse. One hypothesis might be the CD90+CD45RA+ population is a transformed HSC that has an aberrant CD45RA expression or a GMP with aberrant CD90 expression and low CD38 levels. From my in vitro and in vivo data so far, demonstrating robust lymphoid output from the CD90+CD45A+ population the first hypothesis is more likely but analysis of larger patient cohorts is required.

In line with the refractory nature of the disease and high relapse rates the JMML HSC is transformed and able to propagate all the disease features without additional events in downstream progenitors, unlike AML. Since all downstream progenitors derive from the HSC it is not surprising that some of these still retain leukaemic self renewal capacity and the ability to engraft in xenografts. However, JMML HSC confer a much more aggressive JMML phenotype and are able to propagate the disease in secondary transplants. These data indicate for the first time that the JMML HSC are likely to be the JMML cell of origin, but populations such as CD90+CD45RA+ are also able to propagate the disease in mice and are present at clinical relapse in patient samples. Correlation with molecular data as described in Chapter 5 further supports this hypothesis. Analysis of larger patient cohorts and correlation of LSC heterogeneity with clinical outcomes will allow the opportunity to further validate this point.

Chapter 5: MOLECULAR CHARACTERISATION OF THE DISEASE PROPAGATING POPULATION IN JMML

5: MOLECULAR CHARACTERISATION OF THE DISEASE PROPAGATING POPULATION IN JMML

5.1 Introduction

5.1.1: Study of the molecular hierarchy of myeloid neoplasms

Understanding of leukaemia clonal architecture allows us to better understand the natural history of the disease and may also be useful for disease monitoring during treatment as well as development of better targeted disease specific treatments (373, 374). In AML the clonal architecture of the disease is now better understood, due to the development of sequencing techniques and single cell technology. It is well recognized that AML is often preceded by a pre-malignant phase either in the context of myelodysplasia or development of clonal haematopoiesis that remains clinically silent (375, 376). This model suggests that the first leukaemogenic mutation occurs in a cell that is capable of self-renewal or possibly that the mutation confers self-renewal potential (377). If either of these conditions are not met then the mutation is lost due to downstream differentiation (378). In MDS, Woll et al (118) showed that the HSC harbors the driver mutations that lead to clonal advantage and the development of a premalignant clone. The presence of the disease-initiating mutations in HSC in patients with AML supports this hypothesis (379). These events are thought to be followed by a series of additional mutations in downstream progenitors; conferring self-renewal potential to more mature populations (357, 358) and impairing differentiation. During the pre-leukaemic state, mutant HSC, still produce downstream progenitors, which is why often these conditions can remain clinically silent (380, 381). However, the pre-malignant state (if any) of de-novo AML especially in

childhood is much more difficult to study, due to the lack of premalignant samples to use for investigation.

Studies in AML and ALL suggest that the clonal architecture during treatment and relapse is lot more complex and the relative dominance of different disease sub-clones varies (382-384). Cytogenetic and molecular abnormalities change as the tumour evolves and some mutations may disappear during treatment, until a certain dominant clone acquires specific characteristics that provide a proliferative advantage leading to relapse (385-388). This represents the 'linear' model of disease evolution (385); however a second model, of 'branching' evolution is also described in AML, where multiple different sub-clones are present at diagnosis but a certain sub-clone escapes chemotherapy, acquires additional mutations, expands and propagates relapse (384, 389).

More recent AML studies have started to evaluate, disease burden in the different HSPC compartments at a single cell level. This allows the heterogeneity of leukaemic cell populations to be unraveled, providing new insights into the evolution of the disease-propagating populations in response to treatment (390). This approach also allows the identification of the disease-initiating mutations and an accurate prediction of the order of mutation acquisition. This in turn may help the successful application of targeted therapies, such as IDH2 inhibitors, that can alter the disease prognosis.

In JMML, application of this approach to tracing the clonal architecture of the disease similarly offers the prospect of better understanding the natural history of the disease and may facilitate the development of new generation treatment strategies to specifically target JMML LSC.

5.1.2: Study of the molecular hierarchy in JMML

Next generation sequencing (NGS) studies indicate that each patient with AML has, on average, 13 somatic mutations (391). By contrast, in JMML whole exome sequencing studies have revealed an average of only ~1-2 mutations per patient indicating that the molecular landscape of JMML is far less complex (172). Indeed, unlike AML, a single mutation may be enough to propagate the disease. Single, JMML mutations are frequently detected from neonatal cord blood samples and Guthrie blood spots indicating that the mutations must often be acquired in utero (172, 230, 231, 392). This typically leads to clinical presentation of the disease within the first 2 years of life. Similar/identical RAS-pathway mutations are also detected in other myeloid malignancies, however unlike JMML especially in adult patients, RAS-pathway mutations are not seen in isolation and very rarely represent the disease-initiating mutations in adults (393). Due to this we can hypothesise that the pathogenesis of JMML differs from AML and that the timing of RAS-pathway mutation acquisition during haematopoietic development is likely to be critical in the development of JMML.

In JMML the model of mutation acquisition, is thought to be linear in contrast to other myeloid malignancies where there is loss of mutations and acquisition of different clones. This is based on the fact that in JMML all mutations at diagnosis are stable during disease progression and appear to propagate relapse. However, prior to this study the haematopoietic compartment where the JMML mutations first arise and the impact of the JMML mutations on haematopoiesis was not known and haematopoiesis during treatment and at relapse has not been investigated.

In chapter 4, I described the detailed phenotypic and functional characterization of the candidates for the JMML propagating population based on in vitro and in vivo stem cell assays. However, beyond the phenotypic and functional characterization of disease propagating populations it is important to be able to demonstrate that the disease-associated mutations cluster in the populations demonstrating stem cell activity. Furthermore, the sequence of the disease-associated mutations may give further insights into the hierarchy maintained in JMML and enable us to better understand the natural history.

In this chapter I will describe the molecular characterisaton of JMML HSPC and trace the disease-propagating mutations onto the haemopoieitic hierarchy with the aim of identifying the cellular compartment where JMML-associated mutations first arise. Due to the myelomonocytic nature of the disease, one hypothesis would be that the JMML mutations will be traced to myeloid progenitors such as GMPs as in AML. Or, in line with the findings in chapter 4, the initiating mutation may be found in a very early progenitor and secondary mutations may be found in downstream myeloid progenitors. Tracing of the JMML mutations and mapping these onto the haematopoietic hierarchy will allow to functionally and molecularly identify and define the disease-propagating populations and RNA sequencing of the same population(s) should allow the identification of potential molecular targets for developing new therapies.

5.2 Aims

Aims:

1. Mapping of disease-associated mutations in JMML on the haematopoietic hierarchy
2. Molecular characterization of JMML HSPC using RNA sequencing.

5.3 Experimental approach

5.3.1 Mapping of JMML mutations on the haematopoietic hierarchy

Single cell sequencing techniques were used to screen JMML HSPC for the presence of mutations which had already been identified on bulk cells as described in chapter 3. Primers were designed specifically for each of the patient's mutations (Figure 5.5). To ensure the assay allowed accurate detection of the mutation, cord blood samples from healthy individuals were processed in parallel and sequenced together with the patient samples. In addition, primers for common polymorphisms detected in each patient were designed and multiplexed with the mutation-specific primers in order to ensure that the method was sufficiently sensitive to detect false negative results. For each cell sorted and screened, the immunophenotypic information was stored during FACS-sorting and this allowed me to identify the cellular compartments in which the JMML-associated mutations were detected. (Operation of FACS sorter machine was performed by Dr Benjamin Povinelli (Mead lab) or members the WIMM Flow cytometry facility) Bioinformatic pipelines for the single cell genotyping analysis were developed and led by Alba Meira (DPhil student Mead lab).

5.3.2 Molecular characterization of disease-propagating populations using RNA sequencing

The immunophenotypic, functional and mutation screening I performed allowed me to identify candidate JMML LSC and these populations were specifically selected for RNA sequencing. HSPC were initially sorted in bulk (50 cells of each population) and processed as outlined in chapter 2. Following bioinformatic analysis of the data certain cell surface and molecular pathways were identified as significantly unregulated in JMML LSC. Selected cell surface markers were then validated with flow cytometry and their potential value in for disease monitoring following treatment was investigated by screening patient samples at diagnosis and relapse. As proof of principle, I also carried out preliminary in vitro studies to test the potential value of targeting the top molecular pathways identified by RNA sequencing using pharmaceutical inhibitors that are already commercially available. Operation of FACS sorter machine was performed by Dr Benjamin Povinelli (Mead lab) or members the WIMM Flow cytometry facility. Library preparation for the bulk RNA sequencing was performed by myself under the direct supervision and assistance of Dr Benjamin Povinelli. Bioinformatic pipeline for the bulk and single RNA seq analysis was developed and led by Benjamin Povinelli. Analysis of the bulk RNA seq data presented in this thesis was performed by myself with the assistance of Dr Benjamin Povinelli. Analysis of the single cell RNA seq data was performed by Dr Benjamin Povinelli.

5.4: Tracing of the JMML mutations on the haematopoietic hierarchy

5.4.1 Development and optimisation of the colony genotyping protocol

Single cells from JMML HSPC were index-sorted into methylcellulose and the colonies were picked on day 14 of culture for single cell genotyping based on a modified commercially available Illumina access array protocol. As described in

chapter 2 there are 5 main steps in the protocol: i. cell lysis; ii. PCR amplification of long target sequences (~300 bp) with multiplexing of primers; iii. amplification of each of the targets individually using primer amplifying shorter primer sequences (~180-200bp); iv. barcoding with adaptors containing single direction indexes; and v. sequencing on an Illumina MiSeq.

In order to ensure optimal conditions for sequencing I initially tested 4 different lysis buffers (Table 1) (Figure 5.1A) and selected custom lysis buffer 2 for all of the downstream experiments as this appeared to give to most efficient for the PCR conditions tested. A series of different PCR enzymes was also tested to select the ones that performed better when multiplexing different primers without using the Fluidigm access array specialist-made chips (Figure 5.1B) to allow me to screen large numbers of colonies with custom primers at a low cost. I selected to use the ROCHE PCR kit (ThermoFisher, UK) for reactions as this appeared to give a higher PCR product yield when different combinations of primer multiplexes were used.

Table 1: List of lysis buffers tested

Lysis buffer name
1. CellsDirect 2x reaction mix Invitrogen UK
2. RTL Buffer Qiagen UK
3. Custom lysis buffer 1 (75mM NaCl, 50 mM EDTA, 100 μg/ml SDS) +/- 10% Proteinase K(Qiagen)
4. Custom lysis buffer 2 (0.4% Triton X , +/- Proteinase K (Qiagen)

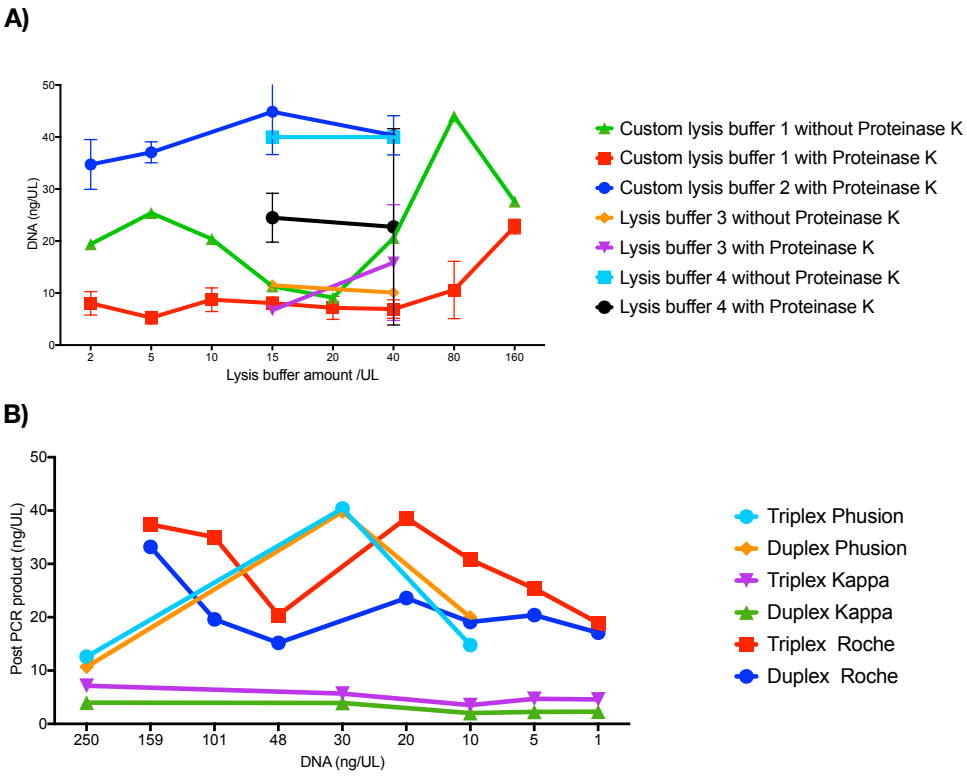


Figure 5.1: A) Post PCR amplification product when different lysis buffers are used at different amounts. B) Post PCR product post-multiplexing 2 and 3 different primer sets with different enzymes.

The optimal reagents and conditions were also tested with different types of fresh and frozen material, as well as different cell numbers in order to ensure that the technique was performing well with small cell numbers and from fresh and frozen material (Figure 5.2), as this was essential when working with small cell numbers.

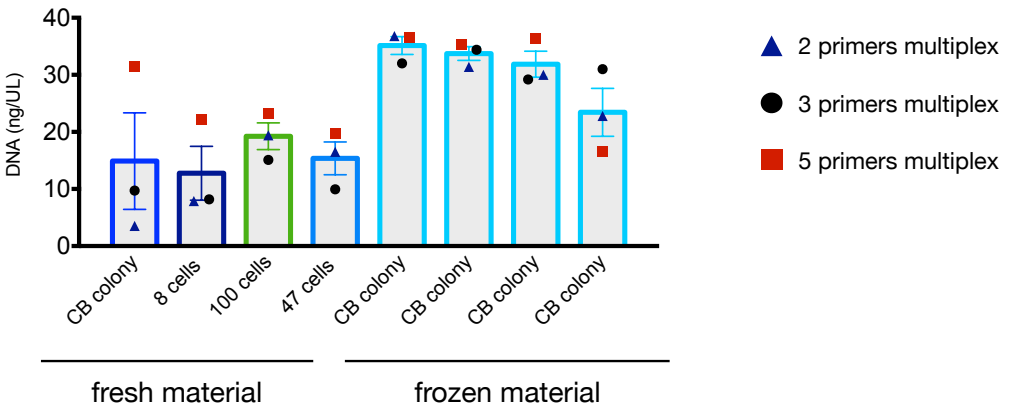
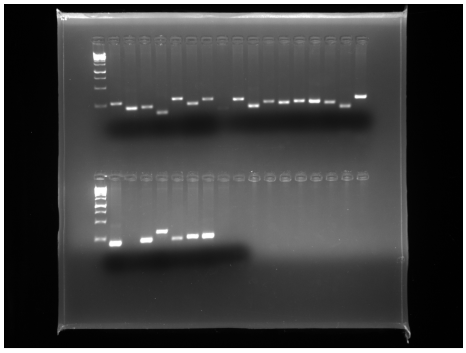


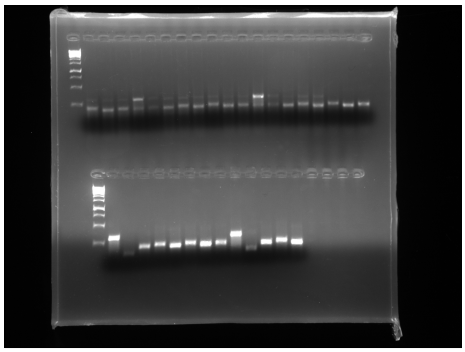
Figure 5.2: Post PCR amplification product from different material and three different multiplex conditions

The specificity of individual primers as well as the off target effects were taken into consideration during primer design to ensure that the sequences used offered the best efficacy for each reaction. Certain regions were more difficult to amplify than others and several primers were designed covering the same region in order to overcome this problem. For each patient sequenced, the primers selected were tested as single primers and the different multiplex combinations were assessed prior to sequencing to ensure that unbalanced amplification was avoided and primer dimers were minimised (Figure 5.3 A and B). Primers melting temperature (generally 59-61°C) was selected as well as the sequence GC context in establishing the optimal multiplexing conditions with the help of online tools: https://www.thermofisher.com/uk/en/home/brands/thermo-scientific/molecular-biology/molecular-biology-learning-center/molecular-biology-resource-library/thermo-scientific-web-tools/tm-calculator.html and https://www.thermofisher.com/uk/en/home/brands/thermo-scientific/molecular-biology/molecular-biology-learning-center/molecular-biology-resource-library/thermo-scientific-web-tools/multiple-primer-analyzer.html.

A)PCR amplification of single primer targets



C)PCR amplification for single targets post PCR 3



B)PCR amplification with different primer multiplex sets

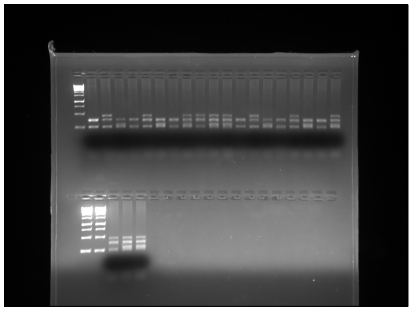
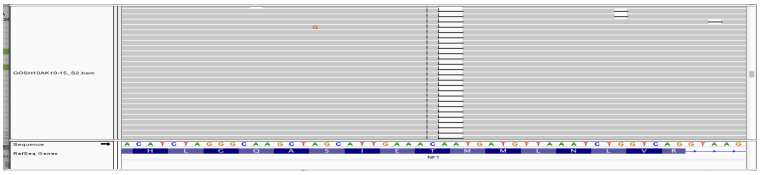


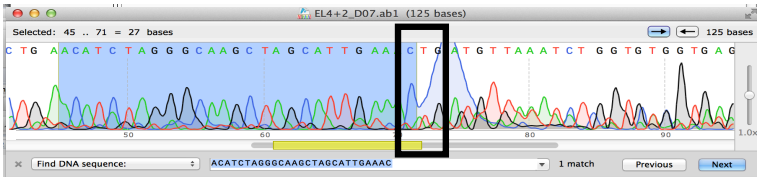
Figure 5.3: A) Image of agarose gel of different individual primers tested using bulk DNA. B) Image of agarose gel of testing different multiplex combinations using bulk DNA. C) Image of agarose gel testing for the presence of individual targets following processing patients colonies with the genotyping method described above

Where satisfactory multiplex combinations could not be established the targets for those regions were amplified individually. Following the multiplex optimization, samples for each individual patient were used to test the efficacy of reactions and once the detection of the patients mutations was confirmed with Sanger sequencing (Figure 5.3C and Figure 5.4) the indexed patient samples were processed for sequencing.

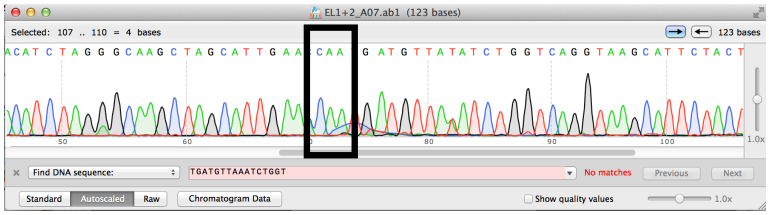
A) NGS of bulk BM of ID1 *NF1* chr: 17 c.2970_2971delAA



Sanger sequencing of *NF1* locus: colony from ID1



Sanger sequencing of *NF1* locus: colony from control



B) Sanger sequencing of rs3217095 chr 7:148504716 AG / A

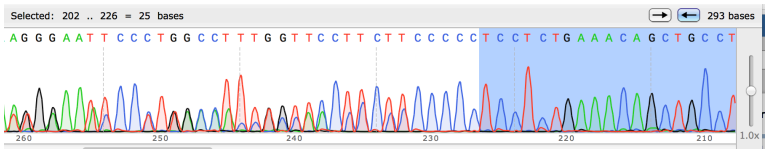


Figure 5.4: A) Image of Interactive genomics viewer (IVG) of patient bulk sample (ID1) following sequencing with targeted NGS showing the detection of an *NF1* mutation. B) Sanger sequencing of a patient-derived colony processed with the genotyping method described above detecting the *NF1* mutation. C) Sanger sequencing of cord blood-derived colony processed as above confirming the absence of mutation in the controls.

5.4.2 Mutational analysis of bulk patient samples

Patients with recurrent JMML mutations, such as *PTPN11* mutations, as well as patients with secondary mutations were selected for single cell genotyping of HSPC. Patient-specific mutations were initially detected by targeted NGS and validated by Sanger sequencing as described in chapter 3. Patients ID1, ID2, ID3, ID5 and ID15 were the patients selected for single cell genotyping (Figure 5.5).

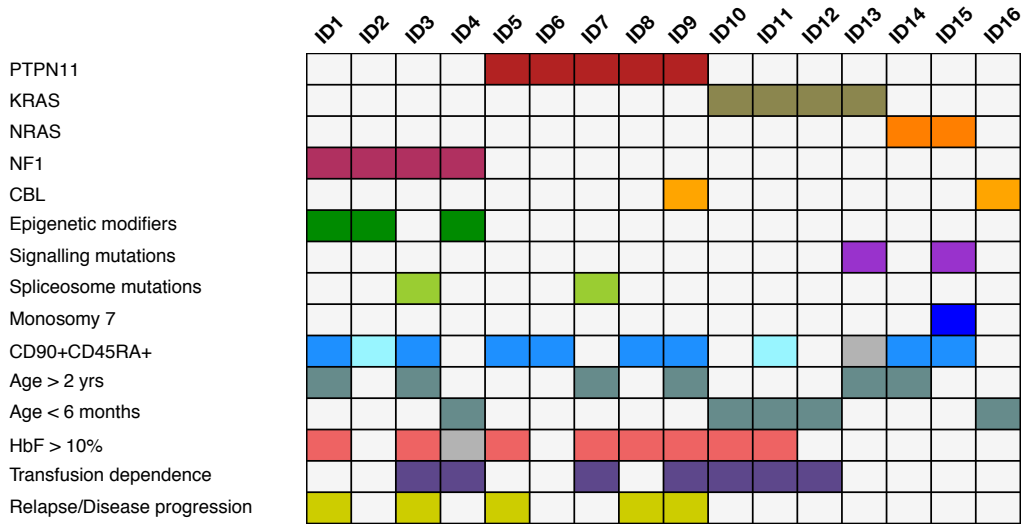


Figure 5.5: Heatmap of patients the cohort of JMML patients. Patients ID1, ID3, ID5 and ID15 were selected for single cell genotyping. Coloured boxes represent the presence of mutations detected through targeted NGS in BM or peripheral blood samples at diagnosis (n=16). Light grey boxes represent data which were not available.

5.4.3 Lineage tracing of patient-specific mutations

5.4.3.1: Colony genotyping for patient ID1

For the first patient screened (ID1), where there were mutations in the *NF1* and *ASXL1* genes, both mutations could be traced to all compartments of the haematopoietic hierarchy (Figure 5.6) including the HSC and the CD90+CD45RA+ compartment. This is in contrast to AML where the initiating mutation gives rise to a pre-leukaemic HSC and additional mutations are then required to produce disease. This in line with the in vivo transplantation data described in chapter 4, is supporting the hypothesis that the JMML HSC is the disease-propagating population. These data are also consistent with previous studies in which secondary mutations in patients with *NF1* mutations were detected in 100% of the cases (193),(172). In this particular patient, screening of the bulk BM samples with the targeted next

generation sequencing panel revealed loss of heterozygosity with a *NF1* deletion detected on both alleles, based on a VAF of 94% and an additional *ASXL1* large deletion with a VAF of 21%. Prior studies have demonstrated that *NF1* mutations alone can give rise to JMML, through the loss of heterozygosity of both *NF1* alleles (202-204). However, these studies had limited ability to detect secondary events, therefore it is difficult to completely exclude the hypothesis that a second event, other than the homozygous loss of *NF1,* is required for the development of JMML. Future comprehensive studies on *NF1* JMML patients will be able to fully address this question whereas recent studies conclusively show that one mutation in other *RAS* pathway genes, such as *PTPN11* or *KRAS,* is enough to give rise to aggressive JMML (193).

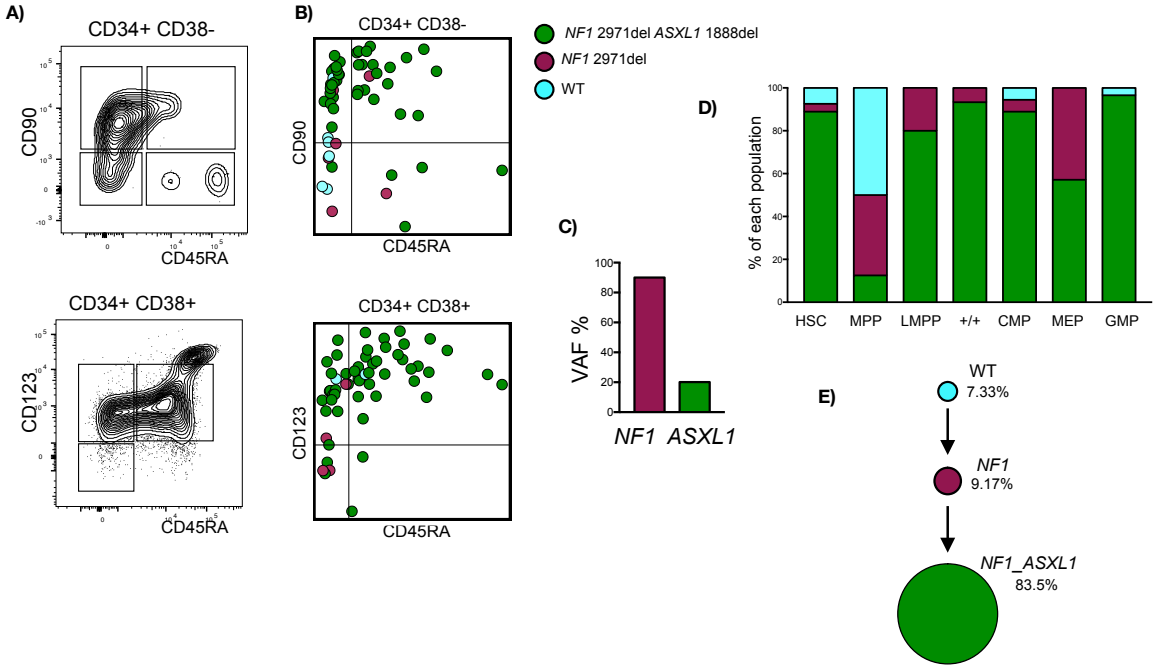


Figure 5.6: A) Flow cytometry analysis of ID1 HSPC at diagnosis. B) Illustration of the distribution of mutations detected in patient ID1. Each point represents single cells plotted based on their index data; the colour indicates the mutational status. Top panel: CD34+CD38- compartment; bottom panel: CD34+CD38+ compartment. C) VAF at diagnosis of each mutation detected in bulk BM samples through targeted NGS. D) Proportion of each subclone detected within each HSPC compartment. E) Inferred model of the order of mutation acquisition in ID1. (Bioinformatic pipeline for analysis of single cell genotyping was developed and led by Alba Meira, DPhil student in Mead lab)

This technique allows us to gain some insights of the order of mutation acquisition in this patient. A clone of cells that carried the *NF1* mutation alone (9.17%) was detected; and single mutant cells were detected in all HSPC compartments, other than the GMPs (Figure 5.6D) consistent with a linear mode of mutation acquisition (Figure 5.6E). Analysis of the patient skin fibroblasts revealed the same *NF1* deletion. Due to this finding I used a rigorous filtering strategy developed by Alba Meira in the Mead lad (described in section 5.4.3.5). Cells that passed the quality control are shown in Figure 5.7 and among these wild type cells were detected with very good coverage rates. These wild type cells were organized in a hierarchical manner leading to the conclusion that this patient is an *NF1* mosaic. NF1 somatic mosaicism has been often reported in cases with very subtle clinical features where JMML was the only presenting feature of NF1 (201),(394) and this was the case for this patient that had no other pathognomonic features of NF1 other than café au lait spots.

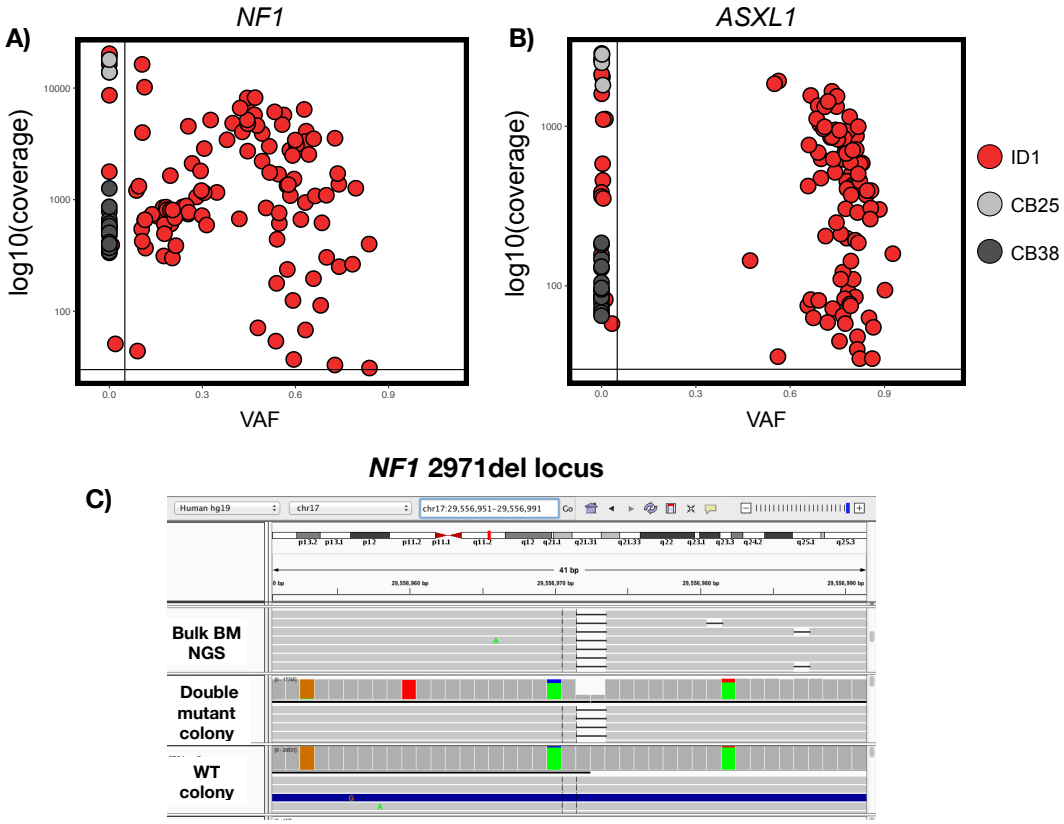


Figure 5.7:Coverage versus variant allelic frequency for every cell used for mutation identification post filtering strategy. A) Coverage for amplicon covering the *NF1* mutation. B) Coverage for amplicon covering the *ASXL1* mutation. C) Examples of samples inspection through IGV. In order from top to bottom: bulk sample following NGS, single cell derived colony displaying the *NF1* mutation; last: single cell derived colony with no mutation detected. (Bioinformatic pipeline for analysis of single cell genotyping was developed by Alba Meira, DPhil student in Mead lab)

To determine whether the abnormal colony types described in chapter 4 correlated with their mutational status, I recorded the colony type for all of the colonies that I genotyped. Figure 5.8 demonstrates the cells sorted and illustrates the mutational status as well as the different colony types observed. Although JMML mutant cells were capable of producing both CFU-M and CFU-GM, the majority of NF1-mutant and NF1/ASXL1 double mutant colonies were CFU-M and the vast majority of CFU-M were 'double mutant'. All three different types of erythroid colonies seen in ID1 were 'double mutant'.

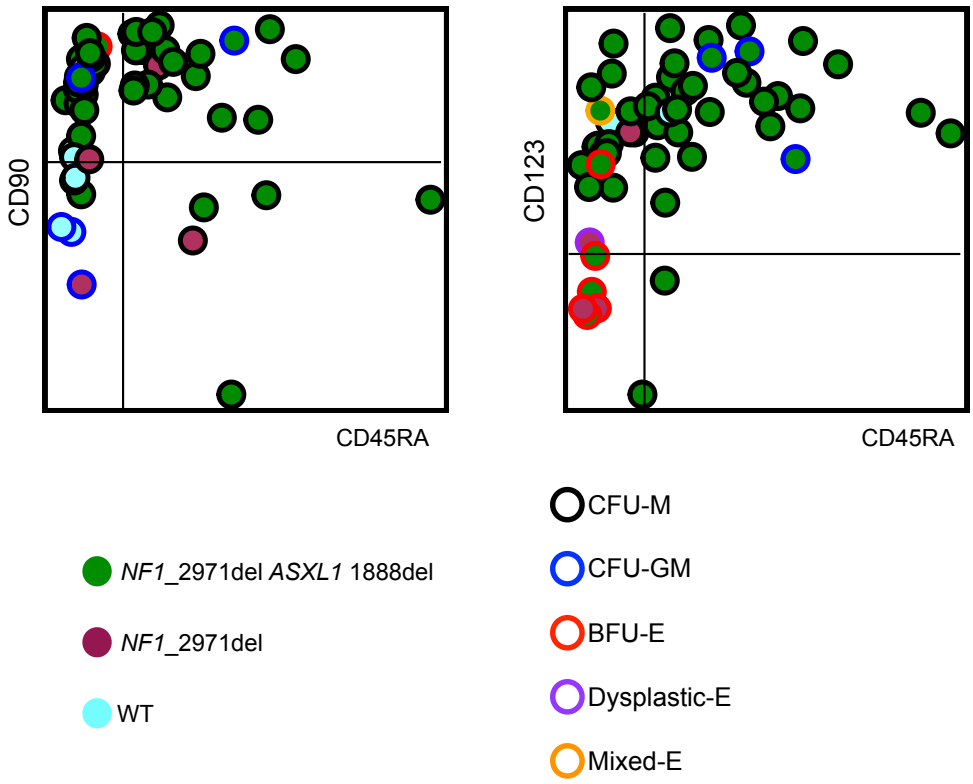


Figure 5.8: Tracing of JMML HSPC mutations on the haematopoietic hierarchy and integration of the genotyping data with the different types of colonies produced by each cell in patient ID1. Single points represent single cells plotted based on their index data, colour indicates the mutational status and the coloured ring indicates the colony type. Left panel: CD34+CD38- compartment and Right panel: CD34+CD38+ compartment.

5.4.3.2 Colony genotyping for patient ID15

This patient had one *NRAS; SETBP1* and also monosomy 7. As for ID1, all of the mutations detected in ID15 could also be traced back to the HSC compartment including the monosomy 7 (Figure 5.9). In contrast to ID1, the sequence of mutation acquisition appeared to follow a 'branching' model (Figure 5.9E): the *NRAS* mutation detected in all cells screened indicating that this was the initiating event with the NRAS dominant clone representing 55% of the cells sequenced while the monosomy 7 appeared to be a secondary event, although traced again back to the HSC. Although it is known that monosomy7 is the commonest cytogenetic abnormality, but it was not previously known whether or not monosomy 7 preceded the RAS-pathway

mutation. The additional *SETBP1* mutation also appears to be a secondary event, that it is again traced back to the HSC compartment. All three clones are detected in all of the HSPC compartments, except MEP where no *NRAS/SETBP1* mutant cells were detected, and GMP, where monosomy 7 was not detected (Figure 5.9).

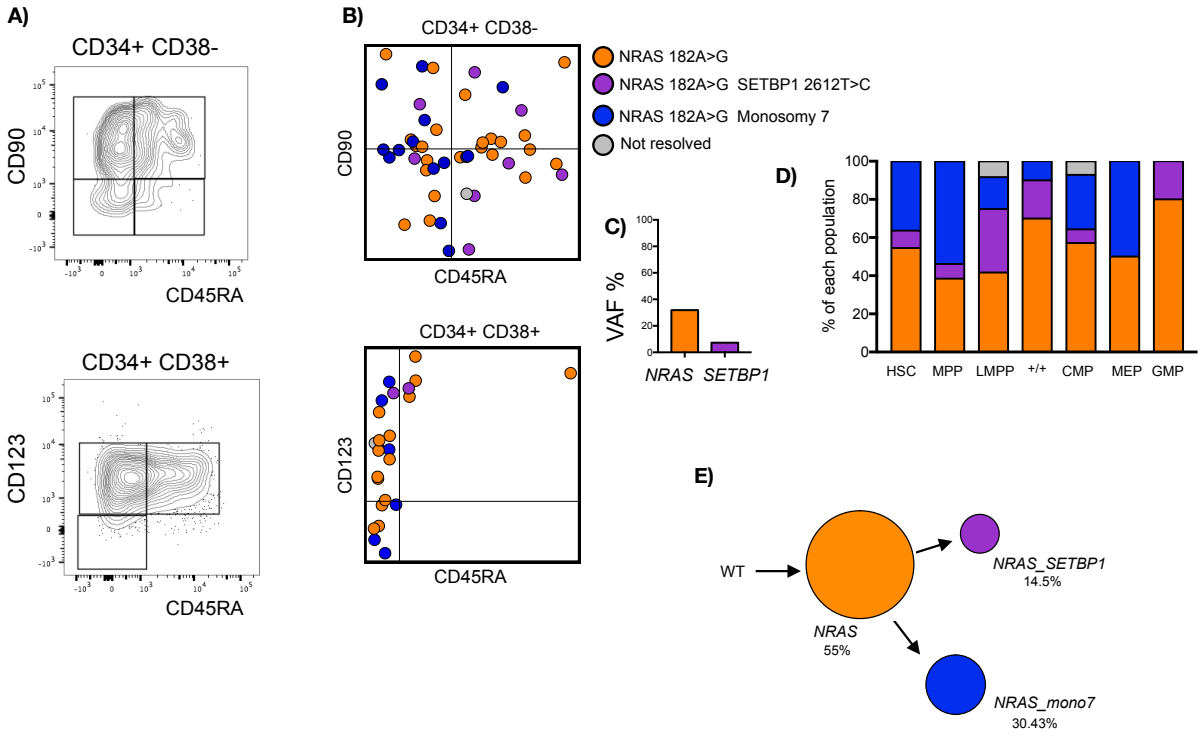


Figure 5.9: A) Flow cytometry analysis of ID15 HSPC at diagnosis. B) Illustration of the distribution of mutations detected in ID15. Single points represent single cells plotted based on their index data, colour indicates the mutational status. Top panel: CD34+CD38- compartment and bottom panel: CD34+CD38+ compartment C) VAF at diagnosis of each mutation detected in bulk BM samples through targeted NGS. D) Inferred model of the order of mutation acquisition in ID15. E) Proportion of each subclone detected within each HSPC compartment. (Bioinformatic pipeline for analysis of single cell genotyping was developed by Alba Meira, DPhil student in Mead lab)

Investigation of the different types of colonies associated with the patients mutations revealed that the 3 different clones were detected in both types of myeloid colonies (Figure 5.10). The erythroid colonies were also always mutant, including the normal in appearance erythroid colonies. Of note, is the fact that the *SETBP1* clone was not

detected among the erythroid output, including the erythroid colonies that were derived from the CMP compartment. This along with the fact that the *SETBP1* is not detected within the MEP compartment may indicate that this clone drives a more myeloid proliferation. However, investigation of further patients with *SETBP1* mutations would be required in order to strengthen this observation.

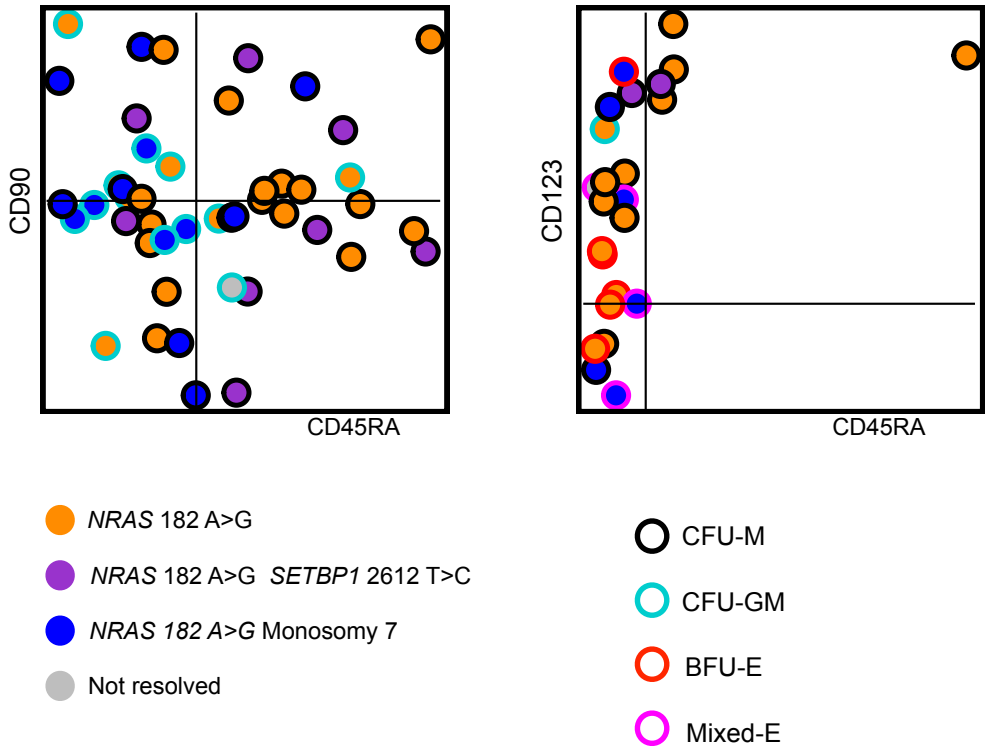


Figure 5.10: ID15: tracing of JMML HSPC mutations on the haematopoietic hierarchy and integration of the genotyping data with the different types of colonies produced by each cell. Single points represent single cells plotted based on their index data, colour indicates the mutational status, coloured rings indicates the colony type. Top panel: CD34+CD38- compartment and bottom panel: CD34+CD38+ compartment.

The sequencing quality and coverage (Figure 5.11) for the samples was also assessed carefully and only the samples that passed the quality control requirements described in section 5.4.3.5 were used for downstream analysis.

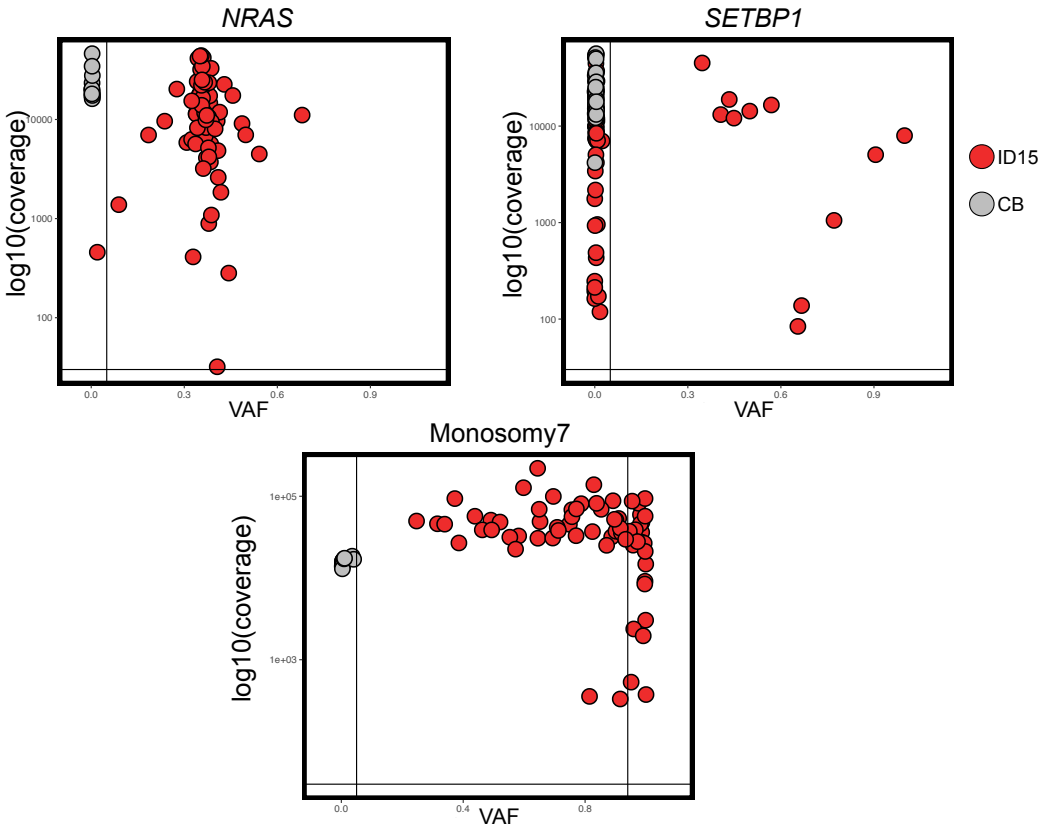


Figure 5.11: Coverage versus variant allelic frequency for every cell used for mutation identification after filtering. Top panels: left coverage for amplicon covering the *NRAS* mutation; right coverage for amplicon covering the *SETBP1* mutation and bottom panel coverage for amplicon screening for monosomy 7 with chromosome 7 common polymorphisms. (Bioinformatic pipeline for analysis of single cell genotyping was developed by Alba Meira, DPhil student in Mead lab)

5.4.3.3 Colony genotyping for patient ID5 at diagnosis and relapse

The third patient I studied had a somatically acquired *PTPN11* mutation and normal cytogenetics (ID5). For this case paired samples from diagnosis and post-BMT were processed. As for patient ID5 the mutation RAS pathway mutation (in this case of *PTPN11*) was traced back to all HSPC compartments screened including HSC, supporting the contention that the disease-propagating mutations are always found in the HSC compartment (Figure 5.12). All the single cells genotyped were mutant and therefore in this case it appears that the mutation is stable and is retained in all HSPC compartments as they differentiate to downstream progenitors (Figure 5.13 A).

The same filtering strategy, as with the previous samples was applied and all cells that did not meet the quality control criteria were not used for downstream analysis. (Figure 5.13 B)

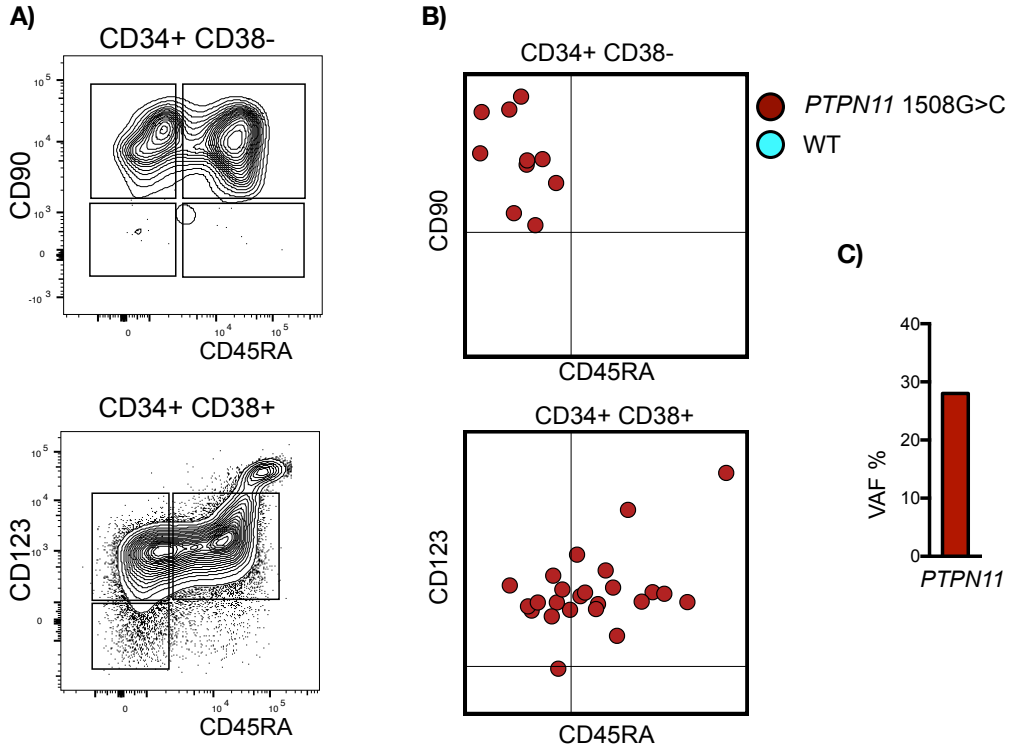


Figure 5.12: A) Flow cytometry analysis of the HSPC compartments in ID5 at diagnosis. B) Illustration of the distribution of mutations. Single points represent single cells plotted based on their index data and colour indicates the mutational status. Top panel: CD34+CD38- compartment and bottom panel: CD34+CD38+ compartment C) VAF at diagnosis of each mutation detected in bulk BM samples through targeted NGS. (Bioinformatic pipeline for analysis of single cell genotyping was developed by Alba Meira, DPhil student in Mead lab)

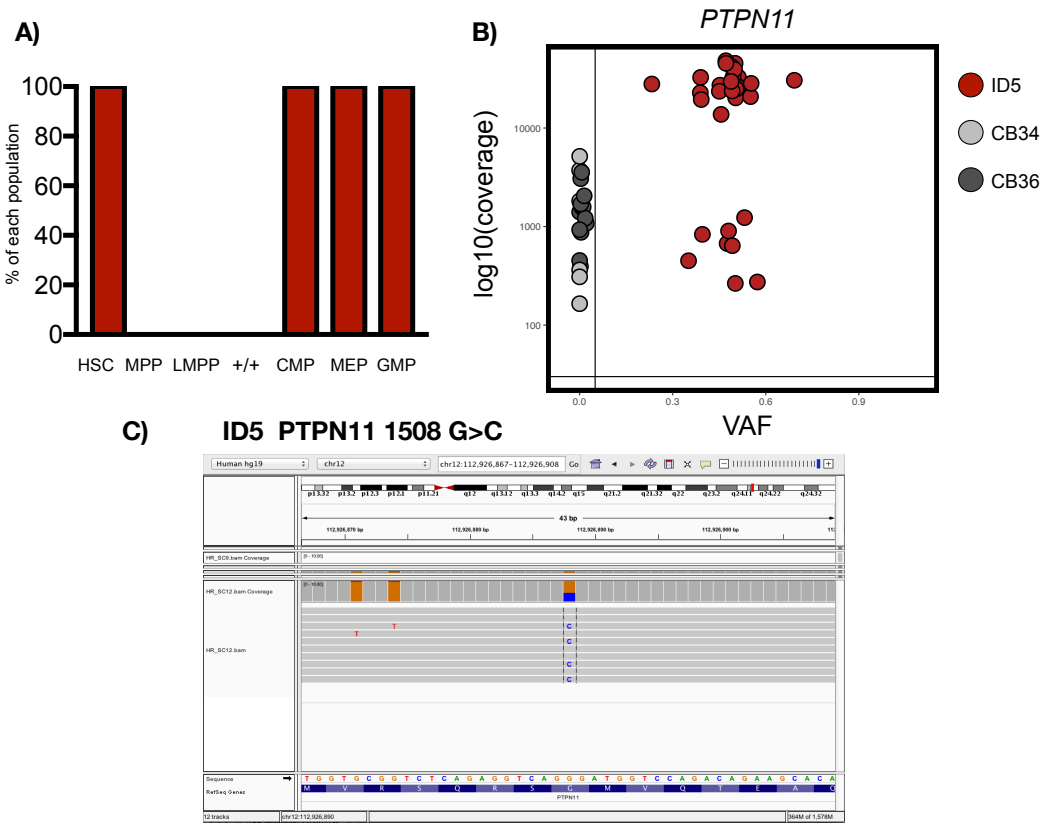


Figure 5.13: A) Proportion of each sub-clone detected within each HSPC compartment. B) Coverage versus VAF for every cell used for mutation identification after filtering: showing coverage for *PTPN11* amplicon. C) Illustration in IVG sequencing of a colony derived from ID5 demonstrating the *PTPN11* mutation

This patient had aggressive disease and relapsed clinically two months after BMT. I used the BM sample one month after BMT (Figure 5.14) to screen for mutations at a single cell level. At this time point there was no evidence of clinical relapse and there was no mutation detected by targeted NGS in the matching bulk sample. However, at a single cell level the mutation was detected in HSPC, including the HSC compartment, propagating relapse that became clinically evident one month later (Figure 5.14). Interestingly no additional mutations were detected at relapse, which demonstrates that in JMML, unlike the typical pattern in AML, the same population that gives rise to the disease, also propagates relapse. This fits more with a linear model of mutation acquisition and it is certainly the case in the patient for the second

patient that relapsed within this cohort. (ID1) (serial sequencing results are previously described in section 4.5).

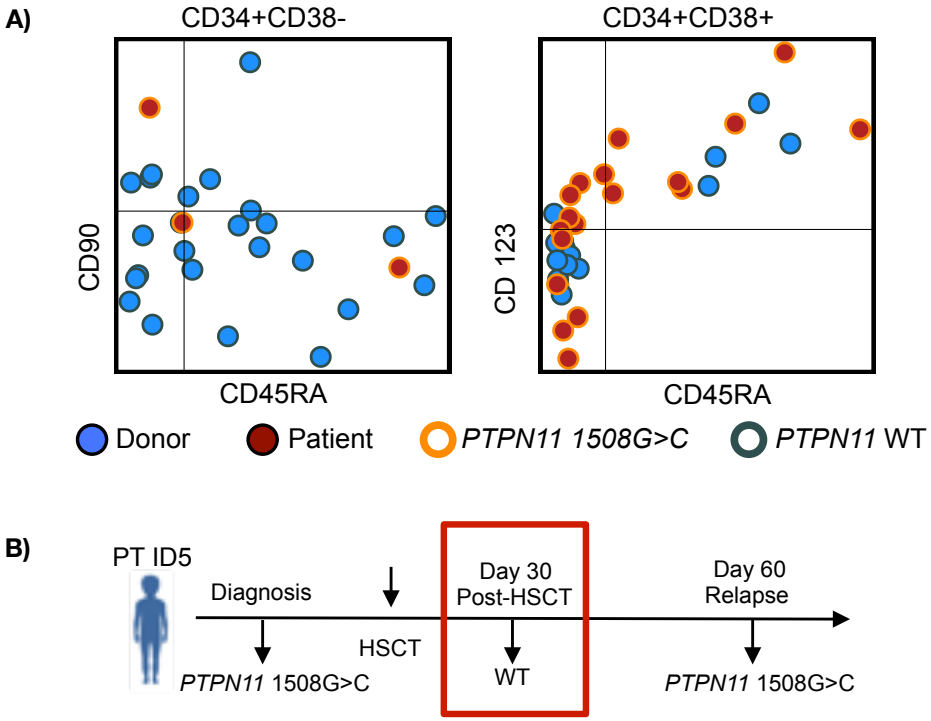


Figure 5.14: A) Illustration of the distribution of mutations detected in ID5. Single points represent single cells plotted based on their index data, colour indicates the mutational status. Left panel: CD34+CD38- compartment and right panel: CD34+CD38+ compartment B) Timeline of mutations detected in bulk BM samples by targeted NGS from diagnosis to relapse (Bioinformatic pipeline for analysis of single cell genotyping was developed by Alba Meira, DPhil student in Mead lab)

In order to be confident that the mutant cells detected were patient derived, along with the amplicons covering the patient's mutation, regions with common polymorphisms detected in the patient as well as the donor were designed and were amplified simultaneously with the patient's mutation targets. This approach allowed me to confidently distinguish between the donor and the patient cells and clearly demonstrated that the mutant cells were the cells were the patient's polymorphisms

were detected and the wild time cells were the cells that carried the donor SNPs (Figure 5.15).

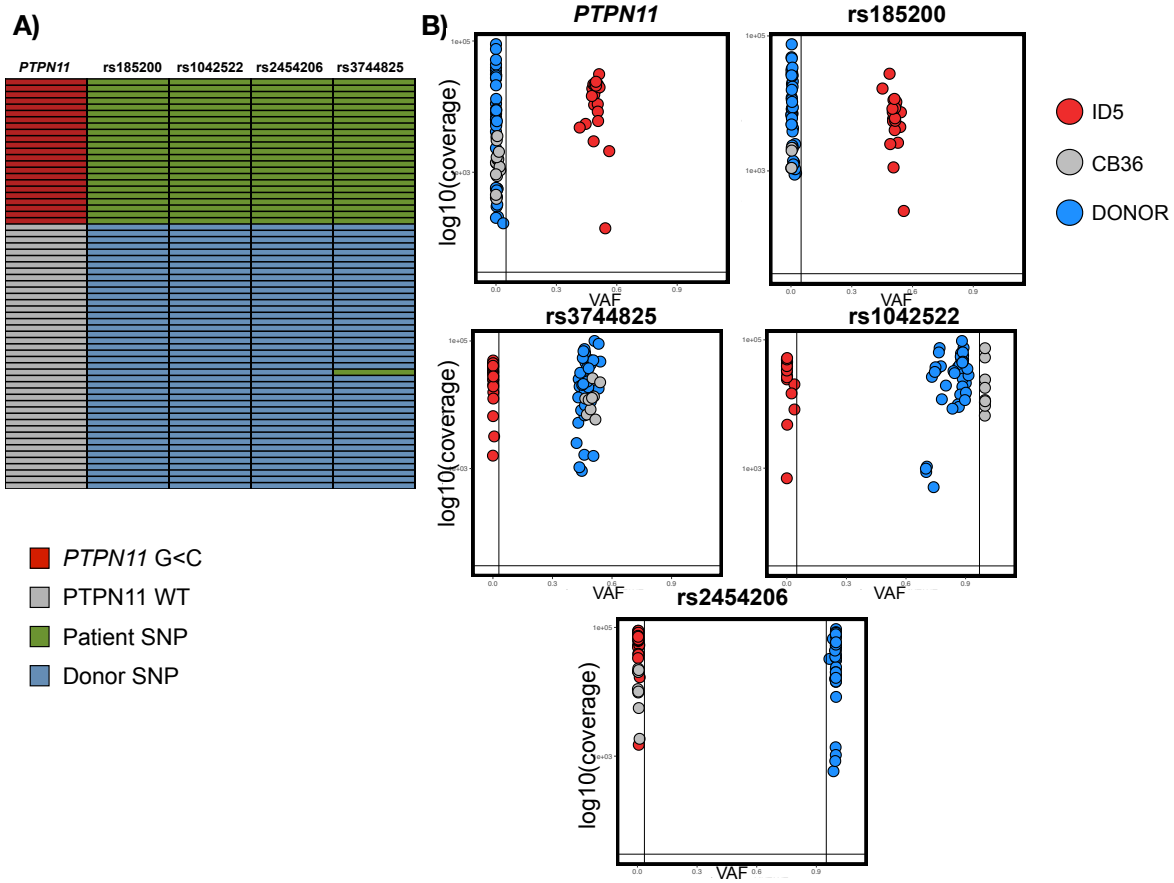


Figure 5.15: A) Heatmap of the 65 single cell-derived colonies analysed showing *PTPN11* mutational status and patient/ and donor-derived SNPs. B) Coverage versus VAF for every cell used for mutation identification after filtering: top panels in B showing coverage for *PTPN11* amplicon and patient SNP. Bottom panel in B showing coverage for donor-derived SNPs (Bioinformatic pipeline for analysis of single cell genotyping was developed and led by Alba Meira, DPhil student in Mead lab)

5.4.3.4 Genotyping of FACS-sorted single cells for ID3

In each case studied above, the JMML mutations appeared to be highly clonal with the majority of cells screened being mutant. However, since the samples analysed were colony-derived, it was important to also screen single cells for mutations rather than colonies in order to ensure that the mutant cells did not have a proliferative

advantage compared to wild type cells resulting in skewing the results towards a highly clonal phenotype. I therefore screened directly sorted single cells rather than colonies, using a method developed by Alba Meira (DPhil student in the Mead lab). Details of the technique are described in chapter 2. I used a sample from a patient (ID3) with a germline *NF1* mutation (*NF1* c.4066 G>T). Sequencing of the diagnostic patient sample revealed two different *NF1* mutations (*NF1* exon 30: *NF1* c.4066 G>T and *NF1* exon 38:*NF1* c.5785 G>T) with a VAF 47% and 46% respectively and an additional SRSF2 mutation (*SRSF2* c. 284 C>A) with a 13% VAF. The *NF1* exon 30 mutation was an inherited mutation detected in other family members. As shown in Figure 5.16 all of the mutations could be traced back to the HSC compartment including the secondary mutations. All cells that passed the quality control requirements were found to be mutant as for the patients described above.

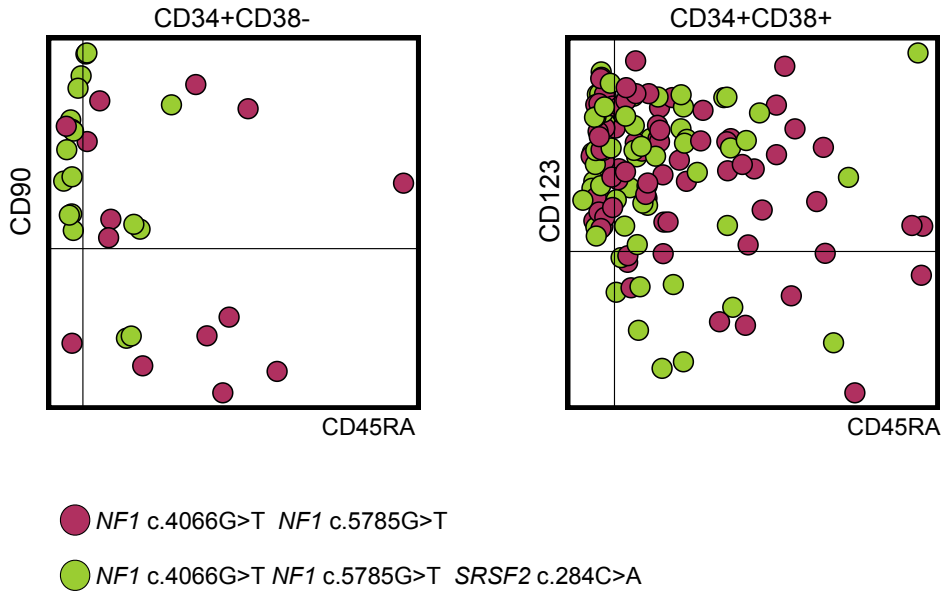


Figure 5.16: Illustration of the distribution of mutations detected in ID3. Single points represent single cells plotted based on their index data and colour indicates the mutational status. Left panel: CD34+CD38- compartment and right panel: CD34+CD38+ compartment (Single cell technique and bioinformatic pipeline for analysis of single cell genotyping was developed and led by Alba Meira, DPhil student in Mead lab)

This particular patient appeared to have a very aggressive disease leading to blastic transformation and death from disease progression; in vivo xenotransplantation experiments using this patient's diagnostic sample showed engraftment from HSC as well as the transformed progenitors. One could hypothesise that in this case the secondary event would be the one that confers the self-renewal capacity to downstream progenitors. However, despite the heterogeneity in LSC activity demonstrated by the *in vivo* stem cell assays, it appears that the disease-associated mutations are very stable and always traced back to the HSC compartment and that the molecular events occurring at the stem cell level allow the JMML cells to differentiate into transformed progenitors although stem cell activity is retained downstream for the majority of the patients.

5.4.3.5: Quality control analysis and filtering strategy

Multiplex PCR techniques, especially those with custom amplicons, can have a high rate of allelic dropouts which can significantly change the results (395). Therefore, specific attention was paid to detection of allelic dropouts and a data filtering strategy that eliminated samples with allelic dropouts.

Detection of false negative results: For each patient sequenced, amplicons covering common polymorphisms were also amplified with the same approach. This allowed the detection of false negative results and enable me to calculate the allelic drop out rates for each sequencing run. Table 2 includes the percentage of allelic dropouts detected per patient following filtering. The filtering strategy was developed with Alba

.Meira (DPhil student in the Mead lab). In short, each individual amplicon sequenced was analysed and those with coverage < 30 reads were excluded from the downstream analysis. Samples in which one of the target amplicons was not detected and samples with allelic dropout (assessed using SNPs) in three or more amplicons were also excluded from further analysis. Since amplification was primer-dependent, following the initial filtering step and coverage cut off was applied for each amplicon, as outlined in Table 5.2, filtering out samples that did not pass these QC requirements. Figure 5.17 demonstrates the coverage and VAFs for all the amplicons used for calculating the allelic dropout rates of the assay. Filtering of the single cell derived data (genotyping for ID3) was more strict in order to eliminate potential bias of the data. All samples where all amplicons did not have coverage of at least 100 reads were removed from the downstream analysis, as well as samples where the mutation hierarchy could not be resolved (Figure 5.18).

Detection of false positive results: For each sequencing run, normal controls were also screened for the variants of interest in order to ensure no false positive samples were present. None of the mutations were detected in any of the control samples in any of the sequencing runs, indicating that there were no false positive results with this assay.

Table 5.2: Allelic dropouts rates per heterozygous SNP sequenced

Amplicon	Donor	Total number of colonies	Coverage	Number of colonies post coverage cutoff	Total colonies post-QC	Allelic dropout
rs4841	PID1	156	30	139	129	5.4%
rs216123	PID1	156	30	149		0%
rs1042522	PID15	157	30	93	71	7%
rs2276599	PID5	93	50	61	34	0%
rs185200	PID5	93	300	73		0%
rs56185013	PID5	93	100	62		0%
rs17319679	PID5	93	50	88		0%

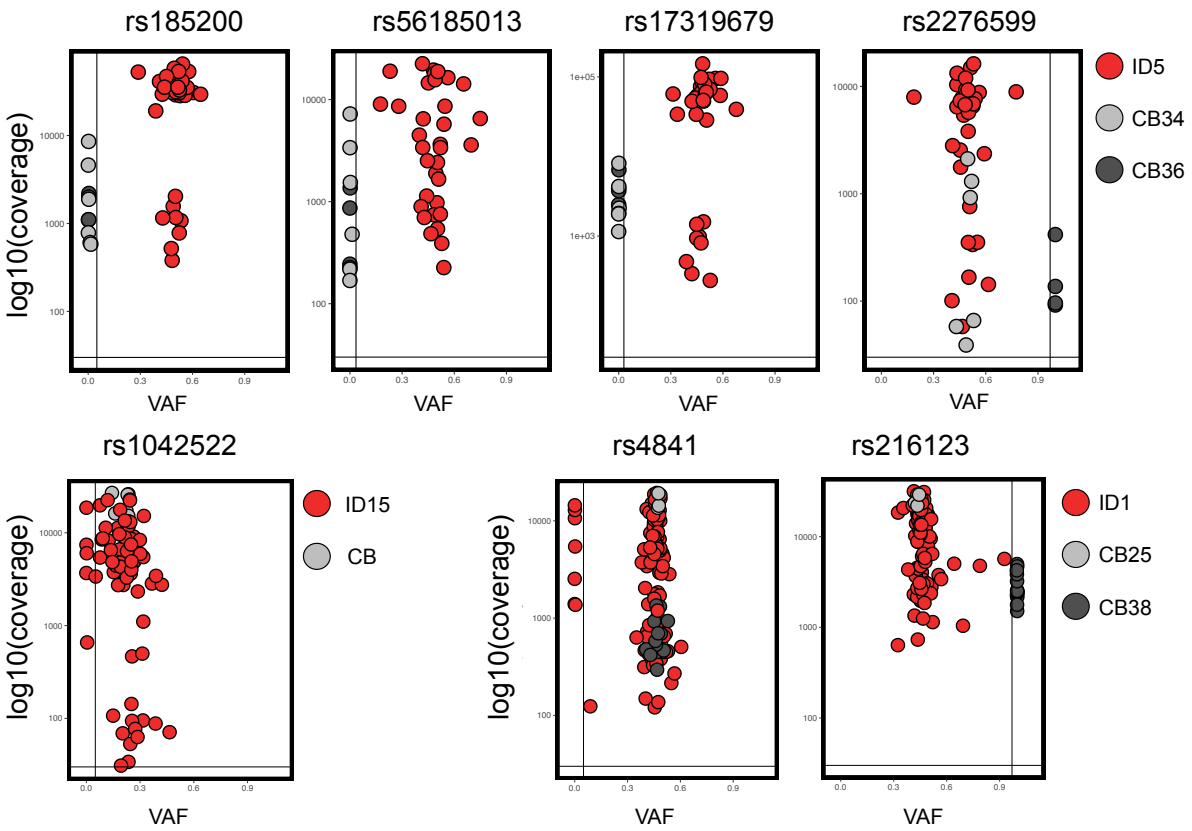


Figure 5.17: Coverage versus VAF plots for every amplicon targeting heterozygous SNPs that was used for the calculation of the allelic drop out rates. (Bioinformatic pipeline for analysis of single cell genotyping was developed and led by Alba Meira, DPhil student in Mead lab)

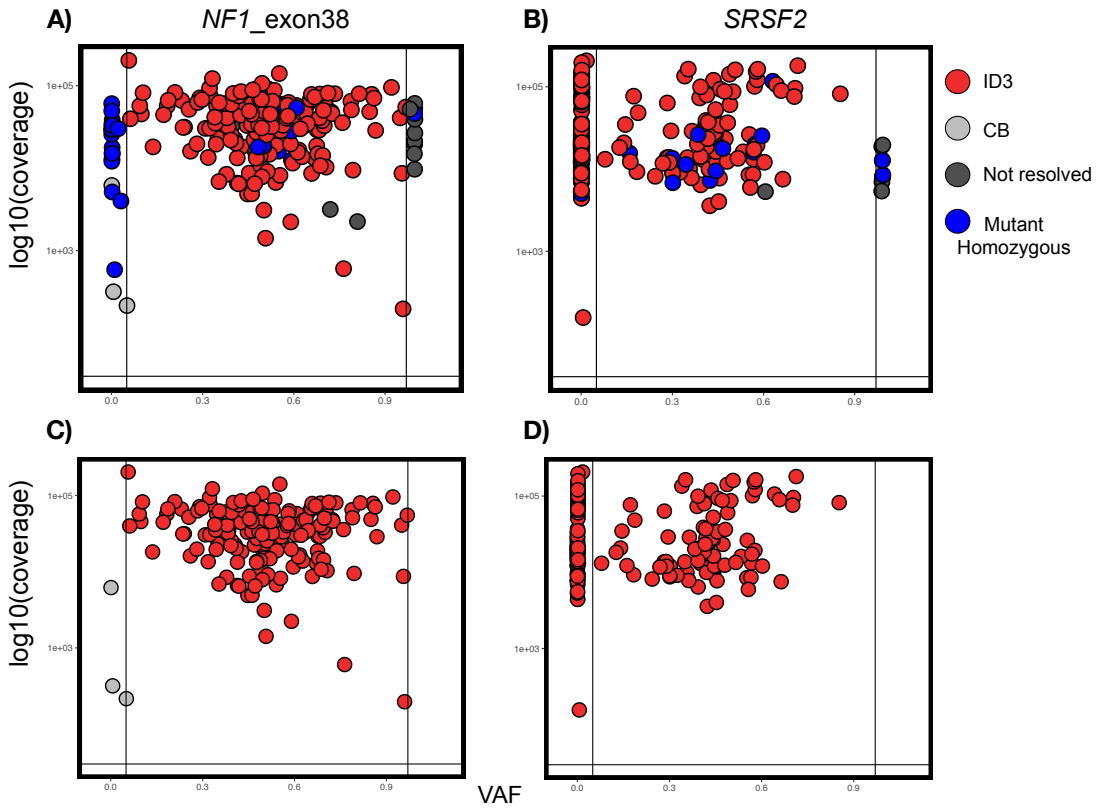


Figure 5.18: Coverage versus VAF plots NF1 and SRSF2 amplicons used for genotyping ID3. A) and C) NF1 amplicon : A) samples post filtering samples with coverage < 100 and C) samples used for data analysis post filtering samples that the haematopoietic hierarchy could not be resolved. B) and D) SRSF2 amplicon : B) samples post filtering samples with coverage < 100 and D) samples used for data analysis post filtering samples that the haematopoietic hierarchy could not be resolved. (Bioinformatic pipeline for analysis of single cell genotyping was developed and led by Alba Meira, DPhil student in Mead lab)

5.5: Molecular characterisation of disease-propagating populations in JMML using RNA sequencing

5.5.1 Bulk RNA sequencing

In order to gain insights into the transcriptome of the JMML disease-propagating cells, I FACS-purified 50 cells of the JMML candidate LSC compartments (HSC, GMP and CD90+CD45RA+ (+/+) cells. The patients ID1, ID3, ID5, ID7, ID15 and one archived JMML sample donated from the Bloodwise Childhood Leukaemia Cell Bank were used and compared with cord blood controls. The patients selected for this analysis had 4 different types of RAS- pathway mutations (*PTPN1, NF1, KRAS, NRAS*) and 4/6 patients had additional secondary mutations in *SRSF2, SETBP1* or *ASXL1* (Figure 5.5). In order to be able to correlate the functional with the molecular data, ID1, ID3, ID5 and ID15 were also used for the in vivo transplantation experiments (chapter 4). All of these patients had an expanded CD90+CD45RA+ compartment, except ID7 who had a CD34+CD38-CD90^{low}CD45RA+ population.

5.5.1.1 Differentially expressed genes between JMML and normal HSPC

RNA sequencing data from the different HSPC sub-populations were analysed for differential gene expression using Deseq2. The numbers of genes that were significantly differentially expressed between the different sub-populations is shown in Figure 5.19. There were 532 differentially expressed genes between the JMML and normal HSC. The transcriptome of JMML HSC was also distinct from the CD90+CD45RA+ population and from JMMl GMP with 229 and 224 differentially expressed genes respectively (Figure 5.19). Interestingly JMML HSC and GMP appeared to be more similar to each other than when normal GMP and HSC were

compared where the number of differentially expressed genes (689) was almost 3-fold higher (Figure 5.19). This is consistent with the presence of JMML-associated mutations in all of these populations although also with the preservation of the haematopoietic hierarchy to some extent in JMML HSC-GMP.

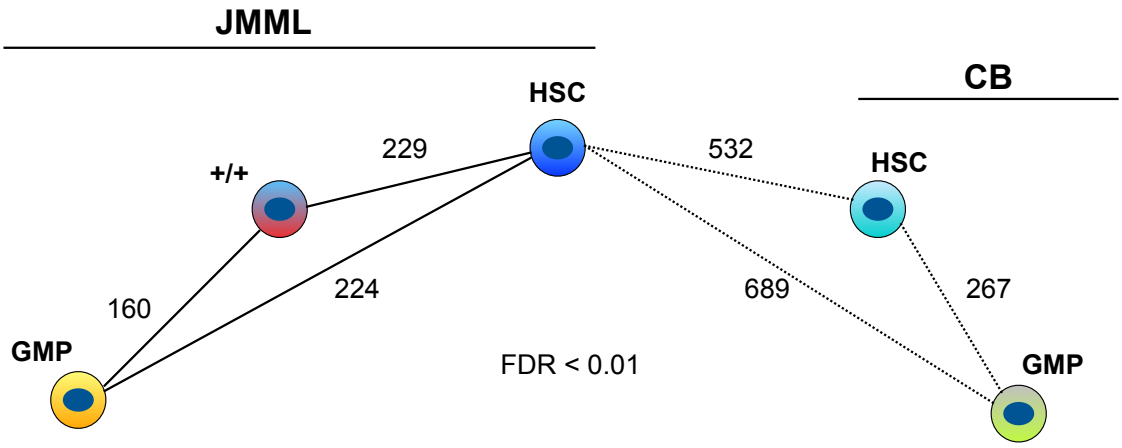


Figure 5.19: Nearest neighbour analysis of purified HSPC from JMML and cord blood. Numbers represent significant differentially expressed genes <0.01 FDR (Bioinformatic analysis was developed and led by Benjamin Povinelli,Mead lab)

Principal component analysis (PCA) demonstrated considerable inter-patient heterogeneity although the JMML HSC formed a relatively distinct cluster close to the normal HSCs while the JMML and normal GMP also tended to cluster together (Figure 5.20). The aberrant CD90+CD45RA+ population clustered with JMML HSC in 3 patients (ID1, ID3, ID17), with JMML GMP in one patient (ID5). For the majority of patients, the JMML HSCs appeared to be distinct from the CB HSCs and the JMML GMPs were positioned in between the GMPs of normal controls and stem cells.

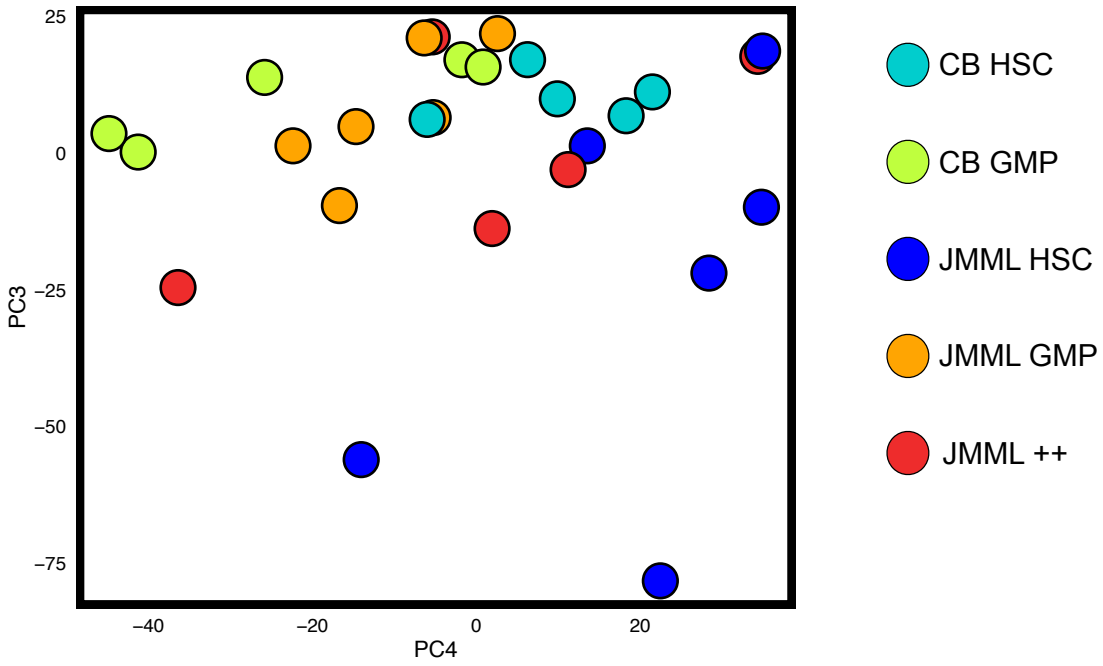


Figure 5.20: Principal component analysis (PCA) of the significantly up regulated genes showing interpatient heterogeneity (PC3 and PC4). (Bioinformatic analysis was developed and led by Benjamin Povinelli, Mead lab)

Nevertheless, differential expression analysis between the JMML and normal HSC did identify distinct clusters (Figure 5.21A) which, on Hallmark GSEA identified genes involved in cycle cell check points, DNA repair, proliferation, inflammation and fatty acid metabolism in line with a population that has undergone malignant transformation (Figure 5.21B). While the transcriptome of the JMML CD90+CD45RA+ population is clearly distinct from JMML HSC and GMP, it appears to share features with that of both the JMML HSC and GMP (Figure 5.22 and 5.23).

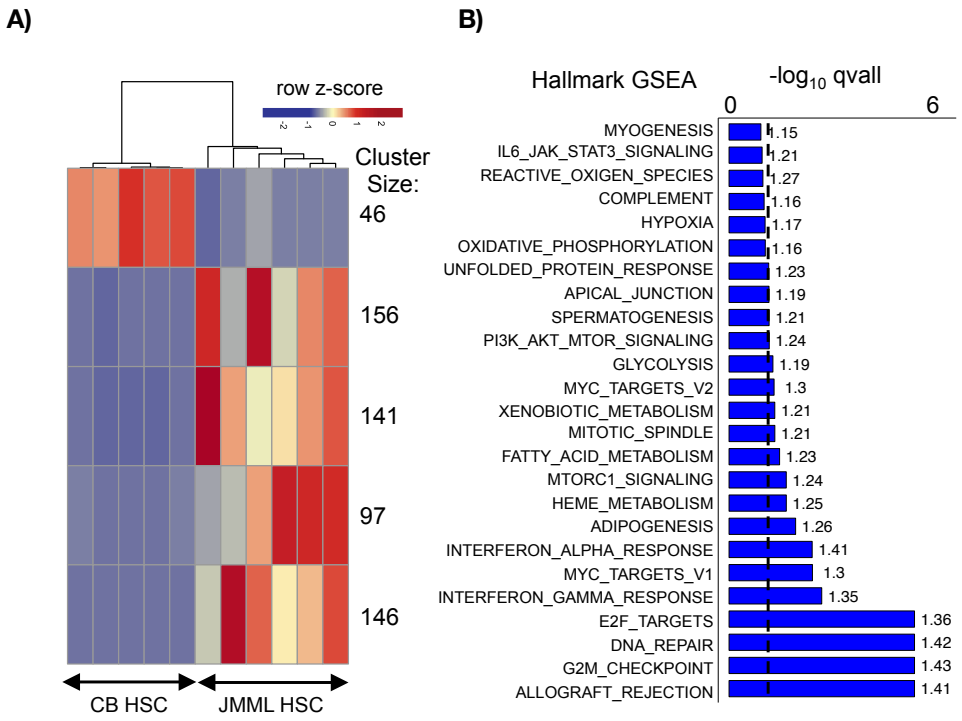


Figure 5.21: A) Heatmap depicting k-means clustering of top differentially expressed genes between JMML HSC and cord blood (CB) HSC with an FDR cutoff < 0.1. Cluster size corresponds to the number of genes within each significant cluster. B) Hallmark GSEA between JMML HSC and CB HSC depicting upregulated gene sets in JMML. Dotted line represents an FDR cutoff of < 0.1; numbers next to the bars represent NES (Normalized Enrichment Score) (Bioinformatic analysis was developed and led by Benjamin Povinelli,Mead lab)

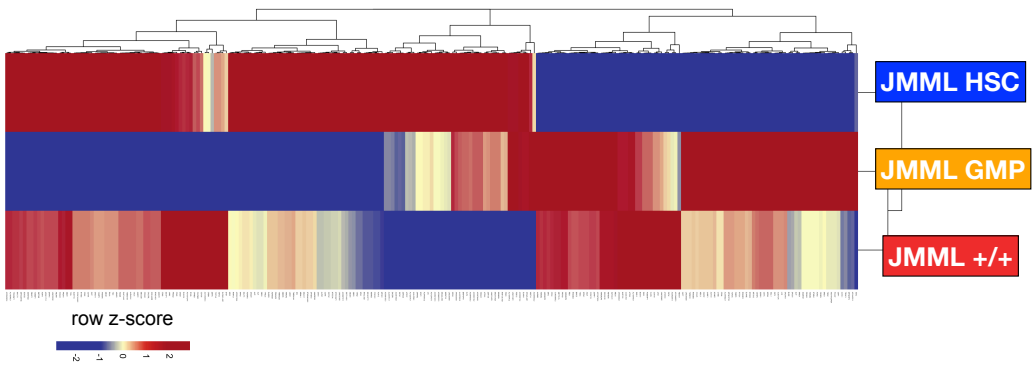


Figure 5.22: The CD90+CD45RA+ shares JMML HSC and JMML GMP signatures.Heatmap of significant differentially expressed genes between the JMML HSC, +/+ and JMML GMPs (Bioinformatic analysis was developed and led by Benjamin Povinelli,Mead lab)

As the GSEA had identified differential expression of cell cycle genes, I investigated the expression of these genes individually. This showed upregulation of some genes between JMML and normal HSC but also upregulation of these genes in the CD90+CD45RA+ and JMML GMP compared to JMML HSC again consistent with some preservation of the haematopoietic hierarchy in JMML (Figure 5.23). For example, expression of proliferation genes such as *ARHGRAF19*, which regulates G1 cell cycle progression displayed higher expression in the JMML CD90+CD45RA+ and JMML GMP compared with the JMML HSCs, suggesting that these populations are cycling more frequently compared with the JMML HSCs. This finding may explain the reduced engraftment potential seen in the vivo experiments from these populations compared with the JMML HSCs.

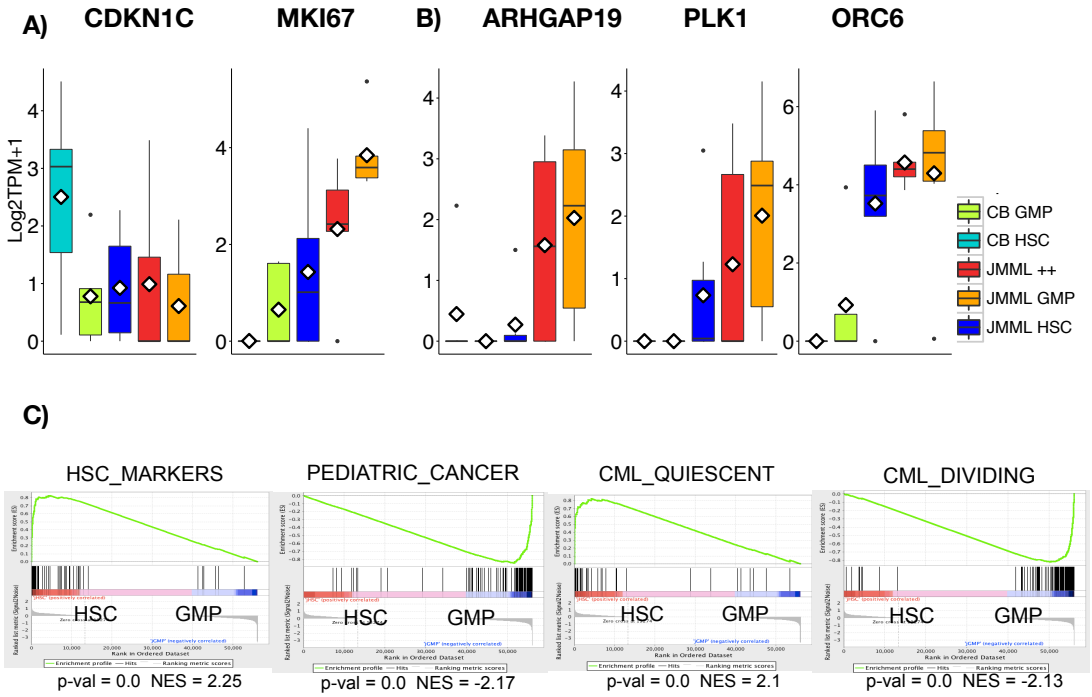


Figure 5.23: A) Cell cycle and B) proliferation, mitosis and DNA replication genes expression in the JMML HSCs and JMML +/+ populations. C) GSEA of JMML HSC vs JMML GMP with upregulation of gene sets for HSC markers and CML quiescence in HSC, upregulation of gene sets for paediatric cancer and CML proliferation GMPs. NES; Normalized enrichment score. (Bioinformatic analysis was developed and led by Benjamin Povinelli,Mead lab)

5.5.1.2 Investigation of lineage trajectories and markers of differentiation in JMML by RNA sequencing

JMML HSC maintained similar levels expression of a number of stem cell genes, including *HLF, FGD5, MEIS1, ABCB1* and *MPL* (Figure 5.24A). Expression of most of these genes was higher in JMML HSC than in the CD90+CD45RA+ population and in these cells was higher than in JMML GMP consistent with the hierarchy of these cells suggested by other molecular and functional data. JMML HSC, CD90+CD45RA+ and GMP also expressed a number of fetal genes consistent with a fetal origin for the LSC although not all fetal genes had increased expression in JMML HSC, e.g. *LIN28B* expression was actually reduced in JMML HSC compared to normal HSC (Figure 5.24B). Previous gene expression study on JMML mononuclear cells displayed high expression of *LIN28B* (14) however these studies were not performed on purified stem cells which might explain the fact that this marker is not expressed on the JMML HSC but it is expressed by downstream progenitors.

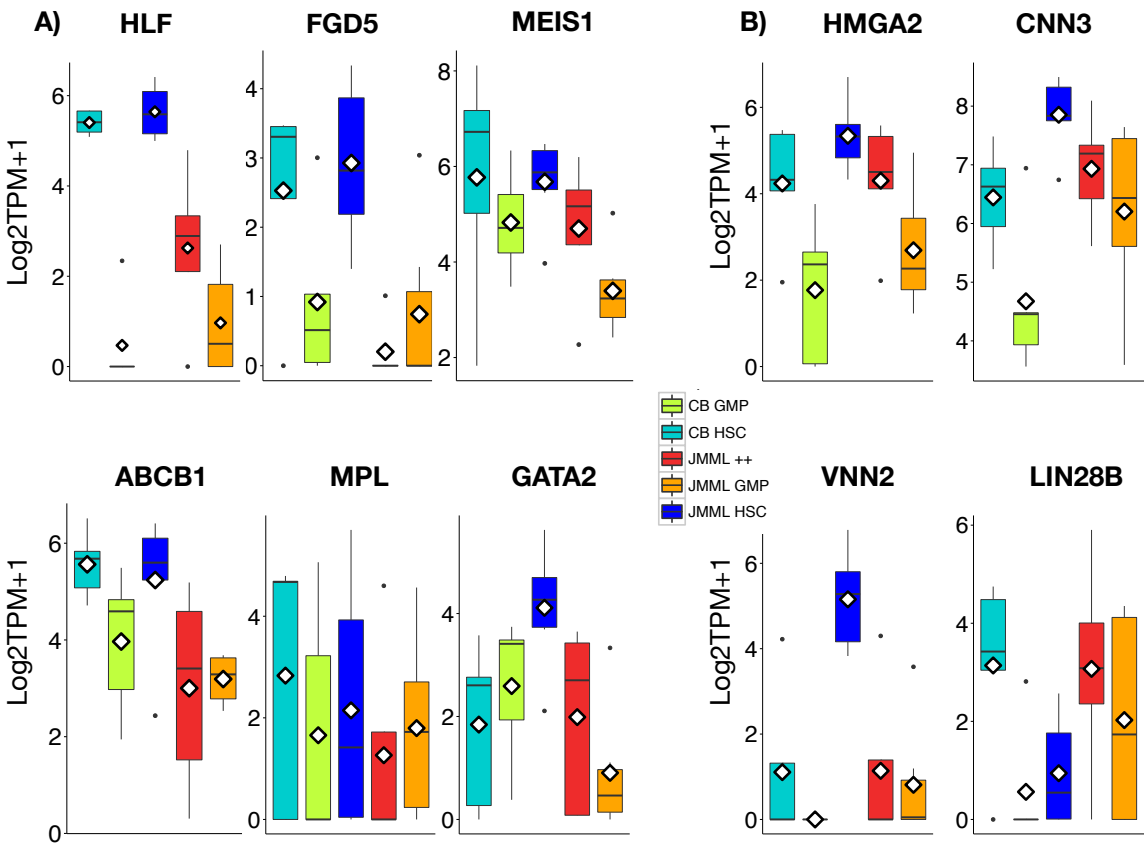
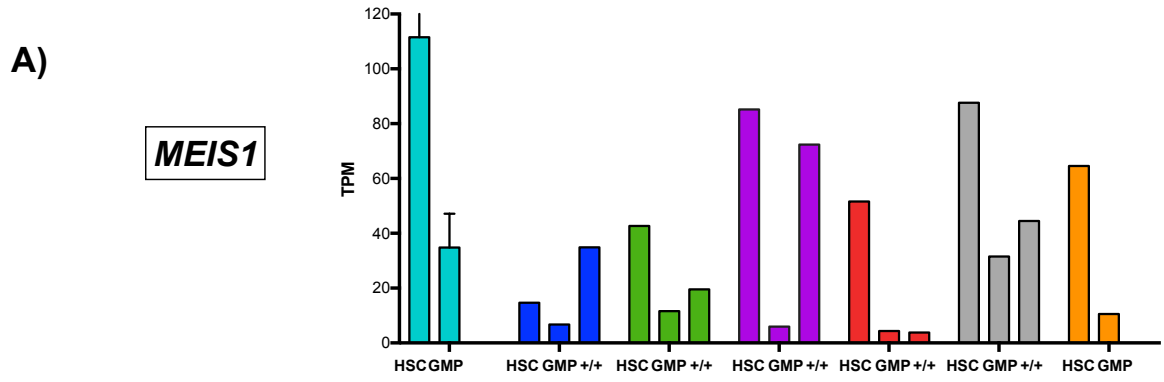


Figure 5.24: Expression of A) known stem cell genes and B) fetal genes by CB HSC (light blue) and GMP (light green) and for JMML HSC (dark blue), CD90+CD45RA+ cells (red) and GMP (orange) for known stem cell genes. (Bioinformatic analysis was developed and led by Benjamin Povinelli,Mead lab)

Although the same overall pattern of expression of stem cell and fetal genes was seen in each patient, there was also significant patient heterogeneity which correlates with the stem cell assays described in chapter 4 (Figure 5.25).



PID	CB	ID1	ID3	ID15	ID5	ID17	ID7
HSC	+	+	+	+	+	NA	NA
GMP	-	-	+	-	+	NA	NA
+/+	NA	-	+	+	+	NA	NA

Engraftment in NSG mice

Figure 5.25: Expression of A) *MEIS1* and B) *HMGA2* genes in JMML HSPC populations for each individual patient with C) matching engraft *in vivo* transplantation data.

There was significant upregulation of a number of erythroid genes, including *GATA1, EPOR* and *GFI1B,* normally driving erythropoiesis from the MPP level downwards (336). JMML HSCs displayed upregulation of these genes compared with control HSC; although expression of other erythroid genes, e.g. *TAL1* and *PBX1,* was not significantly different suggestive of partial impairment of the erythroid pathway (Figure 5.26). Interestingly, *BCL11A* expression was not significantly altered in JMML suggesting that alterations in *BCL11A* expression are unlikely to be the primary mechanism for upregulation of gamma globin expression seen in JMML.

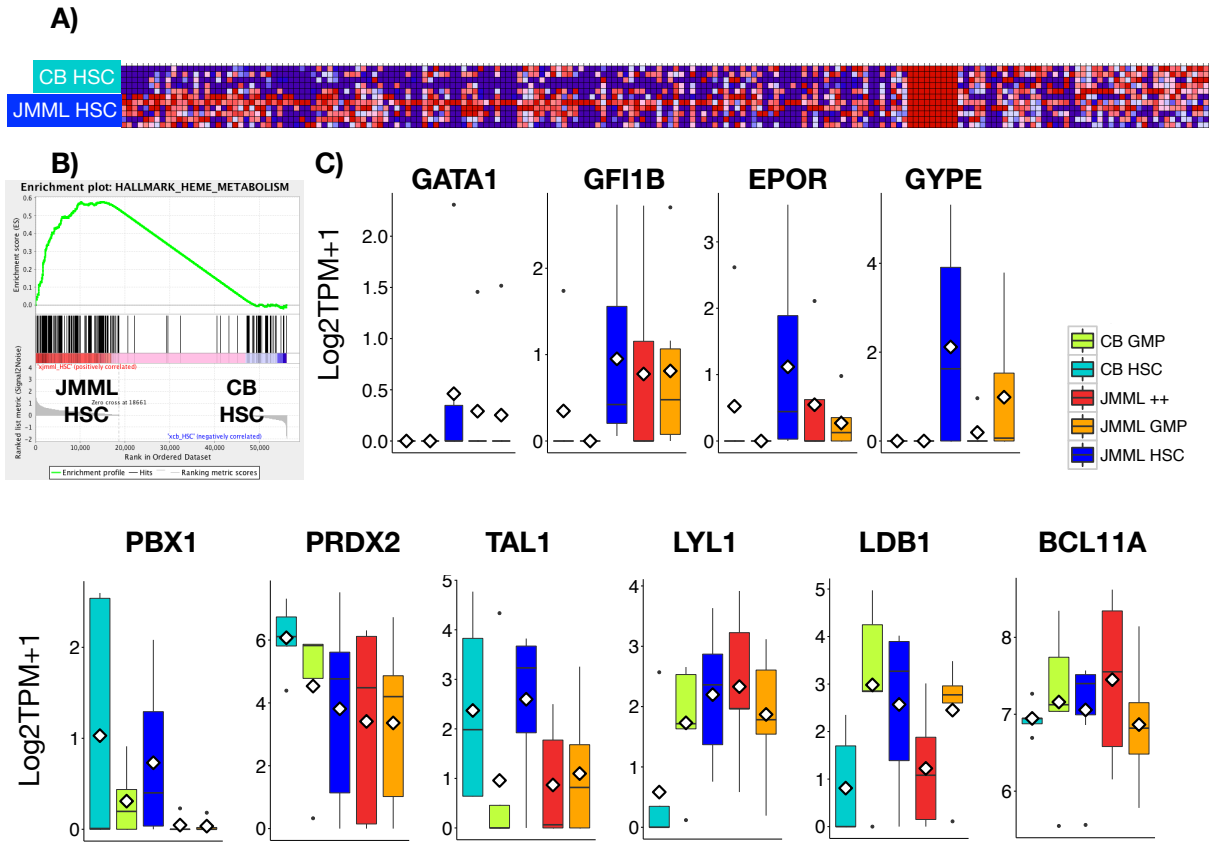


Figure 5.26: A) Heatmap indicating the differentially expressed erythroid genes in JMML vs CB HSC. B) GSEA for erythroid genes showing upregulation in JMML HSC. C) Erythroid gene expression for CB HSC (light blue) and GMPS (light green) and JMML HSC (dark blue), CD90+CD45RA+ (red) and GMP (orange) (Bioinformatic analysis was developed and led by Benjamin Povinelli,Mead lab)

The expression of lymphoid genes was not significantly different in JMML HSC compared to normal HSC or in JMML GMP compared to normal GMP (Figure 5.27). However, similar pattern to the erythroid expression was observed with some lymphoid genes differentially expressed between JMML samples and controls, consistent with the in vitro and in vivo assays reduced lymphoid potential observed.

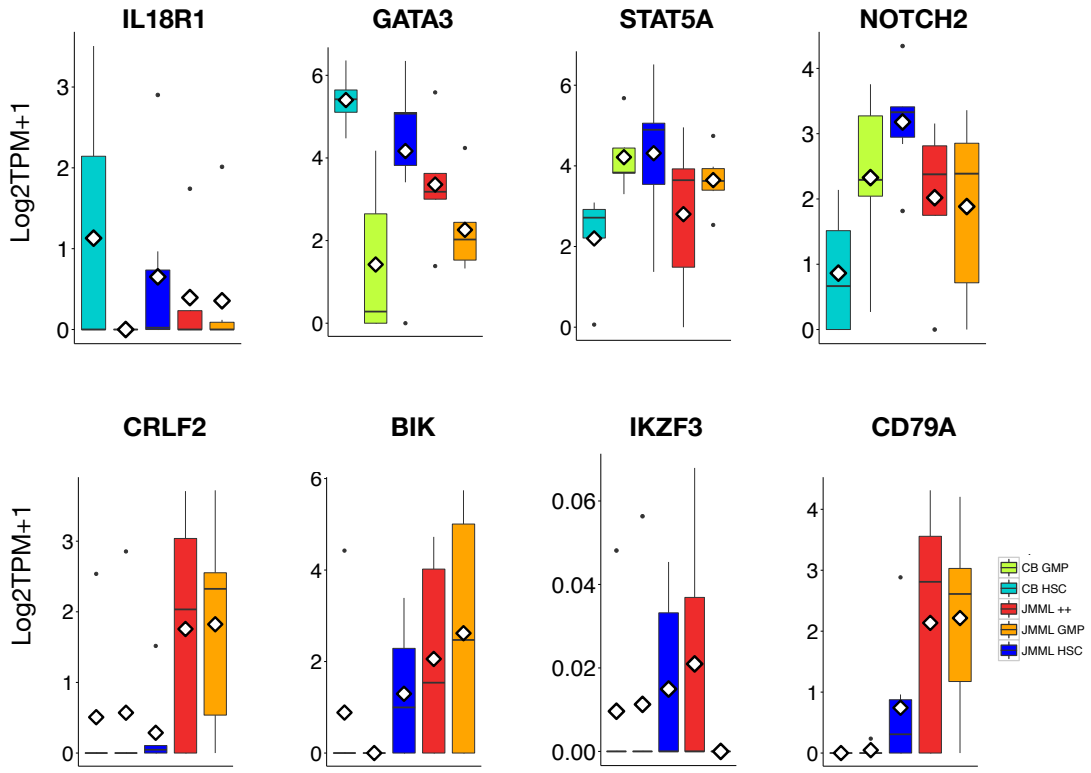


Figure 5.27: Early lymphoid gene expression in CB HSC (light blue) and GMPS (light green) and JMML HSC (dark blue), CD90+CD45RA+ (red) and GMP (orange) (Bioinformatic analysis was developed and led by Benjamin Povinelli,Mead lab)

5.5.1.3 Identification of potential JMML therapeutic targets and/or biomarkers

I also analysed the RNA sequencing dataset to look for changes in expression of known drug target genes. JMML HSC, CD90+CD45RA+ and GMP demonstrated significant upregulation of *mTOR* pathway, *PIM1, MCM2* and *SLC2A1* genes (Figure 5.28). All of these can be targeted with compounds already available for clinical use, including rapamycin (*mTOR inhibitor)*, Fasentin (targets the glucose response transporter GLUT1 encoded by the *SLC2A1* gene) and specific inhibitors being developed as anti-cancer drugs which target *PIM1* and *MCM2* (396, 397) (Figure 5.28) . Fasentin is a new exciting targeted not previously tested in JMML, that can be

applied as an anti-cancer treatments by inhibiting glucose uptake by the cells, which is a mechanism that sensitizes the cells to death receptor stimuli (398).

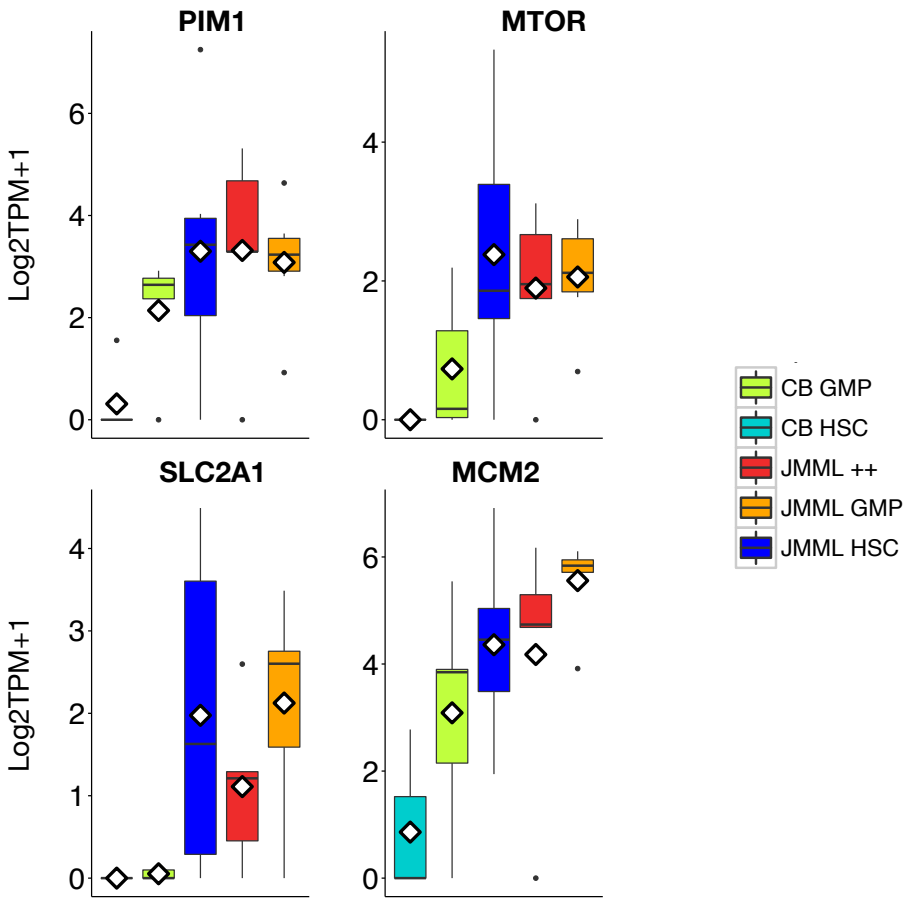


Figure 5.28: Examples of expression of potential therapeutic target genes in CB HSC (light blue) and GMPS(light green) and JMML HSC (dark blue), CD90+CD45RA+ (red) and GMP (orange) (Bioinformatic analysis was developed and led by Benjamin Povinelli,Mead lab)

I selected some of these targets for testing in vitro clonogenic assays as proof of principle. I also tested a RAS pathway inhibitor (PD0325901, a MEK inhibitor) and a bromodomain inhibitor (JQ-1) (Figure 5.29).Bromodomain inhibitor (JQ-1) was also selected to be tested as these family of inhibitors have also shown to reverse RAS-associated transcriptional changes (399), (400). All of these drugs caused dose-dependent inhibition of JMML colony growth although PD0325901 also inhibited the growth of normal cord blood HSPC and may not have a suitable therapeutic 'window'

and further experiments would be required to more rigorously test these compounds *in vitro* and *in vivo*.

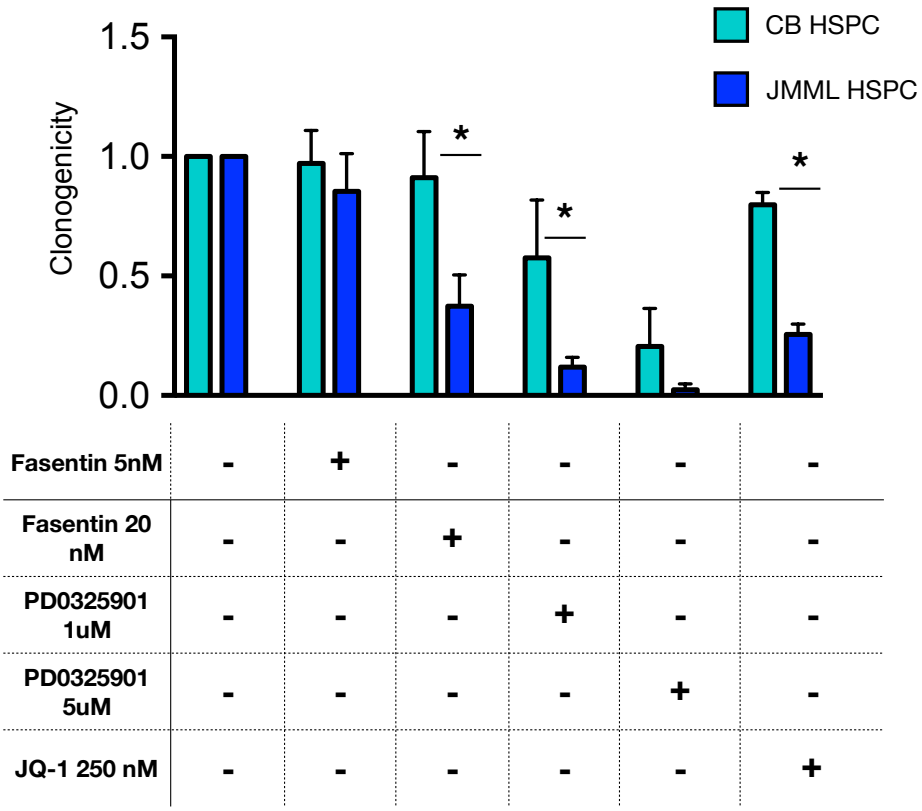


Figure 5.29: The effect of inhibitors of selected target genes on the clonogenic capacity of JMML and CB HSPC in single cell assays (Experiment performed with Benjamin Povinelli)

RNA sequencing also identified up-regulation of the gene encoding the cell surface marker CD96 (Figure 5.30A), previously reported as an AML-specific biomarker (401). *CD96* is a member of the Ig superfamily that is normally expressed on αβ, γδ-T cells and NK cells; and strongly upregulated upon T-cell activation (402) . Further from its role as a targetable epitope recent reports suggest that due to its increasing recognized role as immune check point receptors in T-cells, it can be utilized as anti-cancer immune modulator therapy in a number of RAS-mutated cancers including solid tumours and AML (403), (404). Its role in negatively regulating NK cells function makes CD96 an attractive target for immune modulation that may help promote a

'graft versus leukaemia effect' in JMML (405). The same antibody used in AML was sourced and the JMML patient diagnostic samples were screened for CD96 expression compared with controls. Flow cytometry validated the selective expression of CD96 within the JMML LSC compartment (Lin-CD34+CD90+) compared to CB controls (Figure 5.30B). As proof of principle, I investigated CD96 expression by flow cytometry on paired samples from ID5 at diagnosis and post-BMT. This showed increased cell surface CD96 expression at diagnosis which was also present 30 days post-BMT despite the absence of disease relapse until 60 days post-transplant (Figure 5.30C) suggesting that CD96 may be a useful simple and rapid disease biomarker for JMML although more work is required to confirm this.

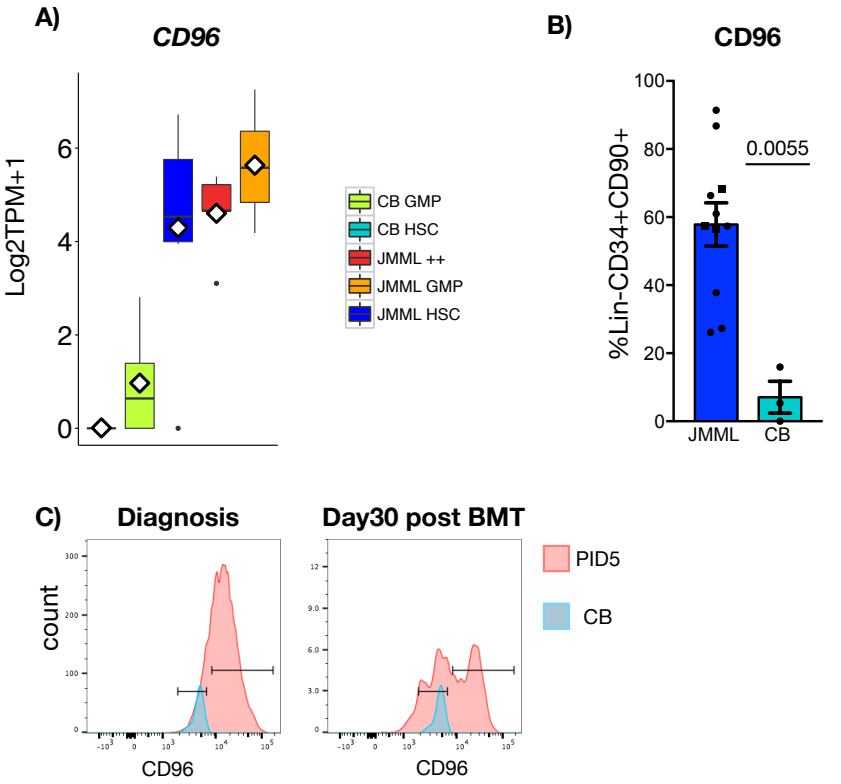


Figure 5.30: A) CD96 gene expression for CB HSC (light blue) and GMPS (light green) and JMML HSC (dark blue), CD90+CD45RA+ (red) and GMP (orange). B) CD96 expression by FACS of CD34+CD90+ cells showing % of Lin-CD34+CD90+CD96+ cells in JMML BM and PB samples at diagnosis (n=11) and CB (n= 3) C) CD96 expression by FACS in patient ID5 at diagnosis and 30 days after BMT.

5.5.2 Single cell RNA sequencing of JMML HSPC

Since the immunophenotypic, functional and molecular studies described above showed inter-patient heterogeneity and it was sometimes difficult to delineate specific sub-populations based on their cell surface CD38 expression, I decided to use an unbiased approach to look at the transcriptome of JMML HSPC. This allowed the investigation of the presence of distinct LSCs clusters among the JMML HSPC cells , despite the immunophenotypic overlap with normal controls. Lin-CD34+ single cells were sorted and processed using the chromium 10X single cell platform. This approach revealed 1127 differentially expressed genes between JMML and normal controls (Figure 5.31A) and further confirmed that the JMML HSPC are transcriptomically distinct from the controls. This approach allowed the identification of distinct clusters not based on the immunophenotypic definitions (5.32A). Each of the clusters is charecterised by overexpression of myeloid genes (*MPO, MS4A3*); fetal genes (*HMGA2*); cell cycle drivers (*MKI67*); eythroid genes (*GATA1*) and leukaemia associated genes (*HOPX*) (5.32B). I used the expression of CD90 and CD38 to identify the stem cluster which was distinct from the cord blood controls, confirming the bulk RNA seq analysis findings with increased expression of genes like *HMGA2* and *CD96* (Figure 5.31B). Similarly to the bulk RNA seq data the JMML clusters were clearly defined by overexpression of genes involved in DNA replication (cluster 3, figure 5.32); mitosis (cluster 5, figure 5.32); embryonic development (cluster 7, figure 5.32) and myeloid differentiation (cluster 6, figure 5.32).

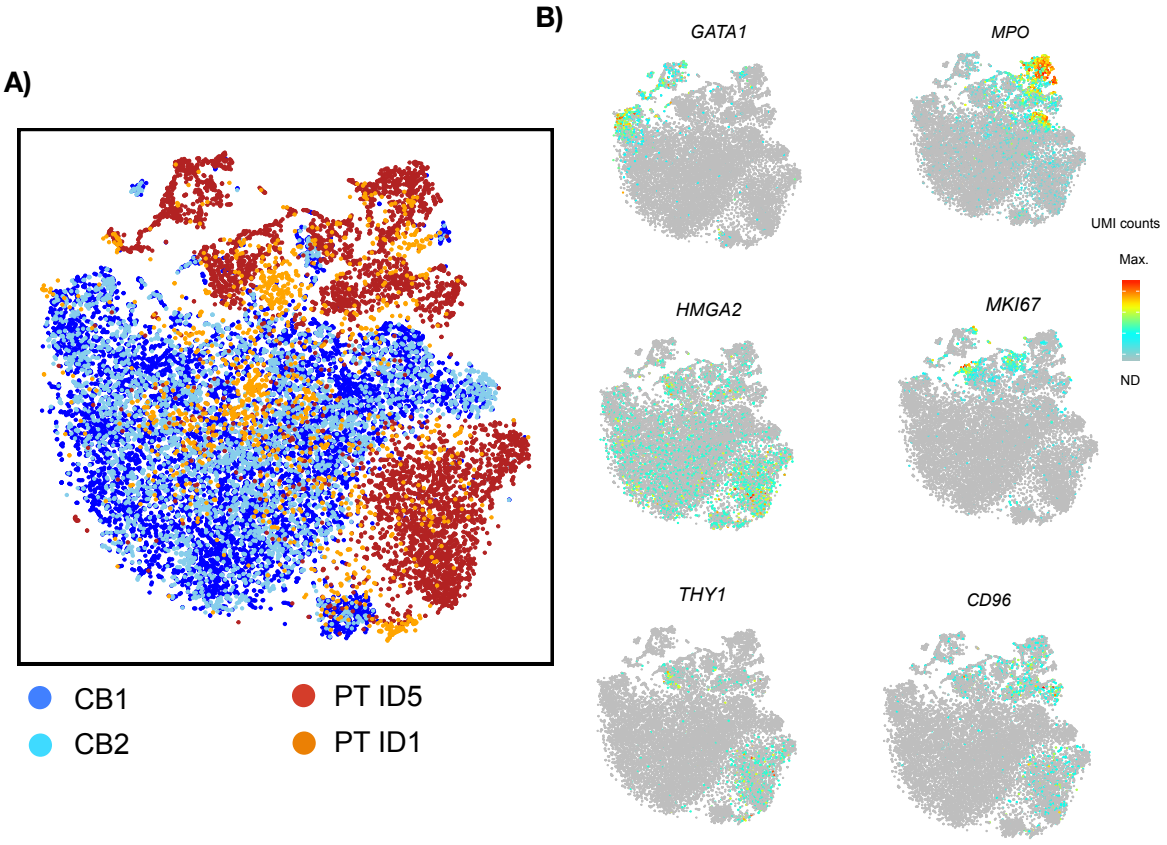


Figure 5.31: A) tSNE plot of Lin-CD34+ single cell RNA-sequencing showing the data from the 17547 single cells that passed filtering from CB controls (shown in blue) and JMML (shown in red) . B) Selected highly variable genes from the top principal components plotted on tSNE plots. Colour scale represents normalized unique molecular identifier counts (UMI, grey not detected, red maximum count) Bioinformatic analysis performed by Benjamin Povinelli

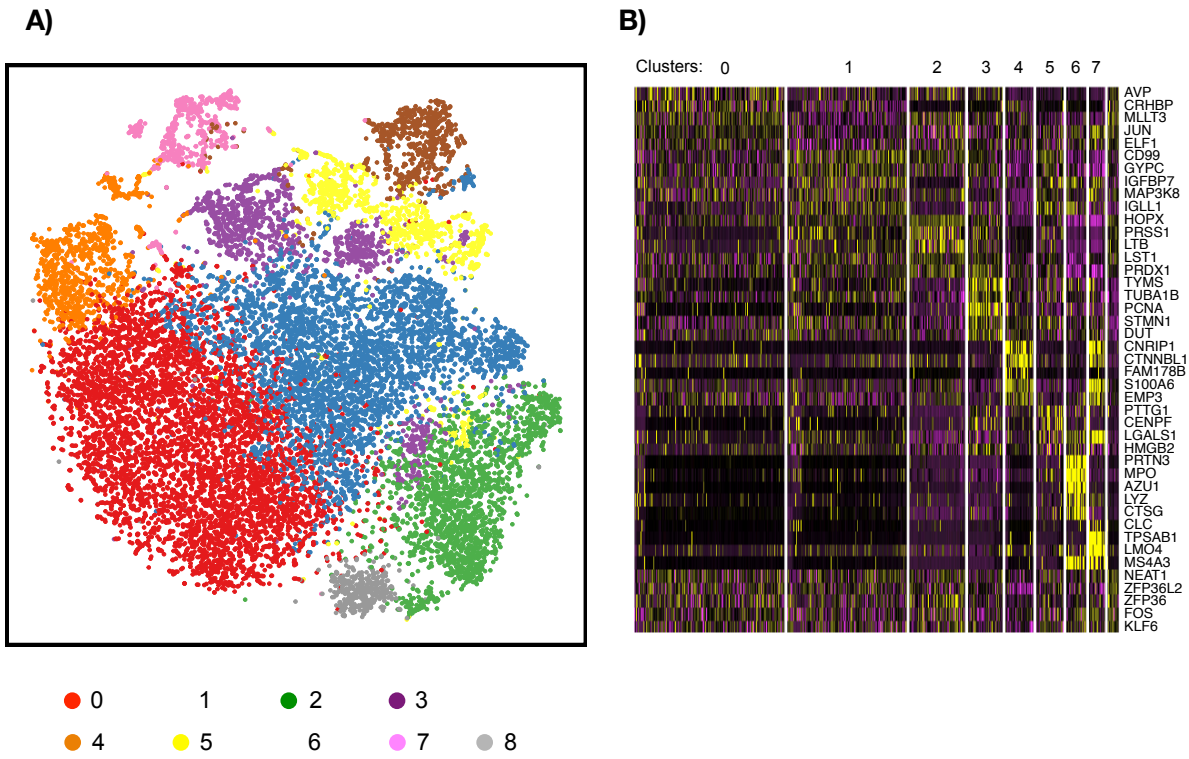


Figure 5.32: A) tSNE plot demonstrating the different clusters identified by k-shared nearest neighbor method. Colours correspond to the different clusters.B) Heatmap illustrating the top 5 differentially upregulated genes that distinguish each cluster. Bioinformatic analysis was performed by Benjamin Povinelli

5.6 Discussion and conclusions

Here, I describe the molecular characterisation of JMML HSPC at the population and single cell level. The data support the conclusions of the immunophenotypic and functional studies and provide some potentially exciting data that may be useful for identifying new drugs for the treatment of JMML or new biomarkers for use in minimal residual disease monitoring.

I used a single cell as well as a population-based approach because it is increasingly clear that single cell sequencing technologies offers several advantages. In particular, it provides the opportunity to better delineate disease-initiating events and, because it removes the bias inherent in sorting defined populations, it can allow the detection of aberrant expression of markers that may be therapeutically manipulated and novel biomarkers. The single cell genotyping technique enabled high throughput screening of a large number of JMML colonies, with high coverage and a relatively low sequencing cost. Purifying single cells from each population allowed me retain the immunophenotypic definition of each cell sorted and back-trace the patient detected mutations to each HSPC compartment. For all patients sequenced the HSC compartment was the population where the JMML driving mutations were identified. This includes primary as well as secondary hits, indicating that JMML does not follow the Knudson' two-hit' hypothesis, unlike other myeloid neoplasms such as AML and CMML. The aberrant CD90+CD45RA+ population was always mutant with the disease-associated mutations in line with a transformed population. Screening of patients with secondary mutations allowed me to identify the model of mutation acquisition in JMML and as demonstrated in section 5.4.3 this a linear in the majority

of cases, with mutations acquired at the HSC level and that are maintained stable over time and propagate relapse, without any additional events propagating relapse. This, along with the data in chapter 4, showed that the HSC were the disease-propagating population in all cases analysed.

Study of the transcriptomic 'signature' of JMML HSPCs demonstrated that the haematopoietic hierarchy in JMML is maintained to some extent with comparable expression of stem cell and fetal genes between the JMML HSCP and controls. However, the JMML HSPCs also have aberrant gene expression signatures which set them apart from the normal counterparts. The differences were predominantly in genes driving proliferation, DNA repair and cell cycle regulators and these patterns appeared to have a hierarchical level of expression with the HSCs at the top, the CD90+CD45RA+ in an intermediate position, followed by the JMML GMP. The identification of disease LSC specific markers, such as CD96 and GLUT-1 overexpression will allow the development of more successful treatment options for JMML in the future.

6: DISCUSSION

6.1 Overview

The cellular architecture and disease-associated mutations of adult myeloid neoplasms are now very well characterised and this has allowed the development of targeted treatment, such as IDH2 inhibitors in AML (406). Paediatric myeloid neoplasms are not so well investigated due to the small number of patients and lack of funding for clinical trials assessing new therapies in paediatric patients. Specifically Paediatric MDS and JMML represent extremely challenging conditions for clinicians both due to the diagnostic difficulties as well as their treatment options and prognosis (136). For several decades there have been no advances in the management of these conditions, although azacitidine has recently started to show some promising results (135) , and BMT remains the only curative treatment, but this fails to eradicate the disease in ~35% of cases (32). In this thesis I investigated in detail the phenotypic and genotypic characteristics of these conditions, with a specific focus on JMML, with the aim of identifying features that could facilitate the application of precision medicine in PMDS and JMML including more improved risk stratification and development of new therapies.

I aimed to establish the first UK Paediatric MDS and JMML study that allowed the collection of deep clinical phenotypic data and serial patient biological samples. This formed the platform that allowed me firstly, to characterise the phenotypic and genotypic features of children with paediatric myelodysplasia in order to improve diagnosis and management of these conditions; and secondly, to determine the impact of JMML-associated, RAS pathway mutations on haematopoiesis. In this

chapter I will summarise the major findings of my thesis and discuss the implications of this work for these diseases.

6.2 The NIHR Paediatric MDS/JMML study

Population-based studies of paediatric MDS performed many years ago provided information about first estimates of the incidence of the disease in the UK (7). However, prior to my project , there was no recent data collection and systematic sample banking to match the phenotypic data had not been performed. The rarity of these diseases (~1 case per centre per year)(7) also meant that in many centres there was little experience to guide the diagnosis and management of these conditions. My project allowed the formation of a UK paediatric MDS/JMML consortium that systematically documented the phenotypic and genotypic characteristics of these diseases creating correlations and allowing better identification of the different disease subtypes.

For the duration of my project, the recruitment of a relatively large number of patients with myelodysplasia or JMML (n=67) to the study allowed the identification of 4 distinct groups: JMML, the MDS group, the ICUS group and the inherited BM failure syndromes. Children with aplastic anaemia can also develop dysplasia but the study's eligibility criteria did not allow the recruitment of such cases.

The presentation of paediatric MDS and ICUS can be very difficult to discriminate due to the comparable presenting features of the two conditions; however the degree of dysplasia (>10% blasts) for the MDS group was sufficient to effectively differentiate between the two groups, as demonstrated by the distinct molecular signatures

detected among each group. The MDS group was characterised primarily by acquired mutations in transcription factors, such as *GATA2,* driving clonal haematopoiesis; whereas the ICUS group was a lot more heterogeneous and included mutations in transcription regulators (such as *MECOM)*; in BM failure genes (such as *CSF3R;)* and children with an underlying metabolic condition (such as Pearson's syndrome). In the future, follow up data, especially looking at the development of clonal haematopoiesis among the MDS and ICUS groups, should be very valuable in validating this point. My study also found that it was not possible to identify the genetic basis of the disease in the majority of patients using the targeted NGS panel that I devised and that more extended sequencing technologies, such as WGS, will be required to fully characterise the ICUS group.

The JMML patient group was clearly characterised by RAS pathway mutations, as previously reported (193), (172). While targeted NGS appears to be of some value in monitoring the disease burden and response to treatment, especially pre-transplant, my data showed this was not always sensitive enough to detect very low level disease, indicating that new disease biomarkers are required. However study of larger patient samples and longer follow up is required to further consolidate this result.

6.3: Phenotypic and functional characterisation of the JMML HSPC

For the rest of my thesis I elected to focus on studying in detail haematopoiesis in JMML as this was the most well defined patient group, with the aim of identifying and characterising the disease-propagating population. In this way, I also hoped to identify new therapeutic targets and disease-specific biomarkers for more accurate

disease burden monitoring. To do this, I analysed primary HSPC from JMML patients and hypothesized that the JMML leukaemic stem cell (LSC) is situated within the HSC compartment, since the disease is refractory to conventional chemotherapy and often myeloabletive BMT is not sufficient to eradicate the disease and since it is the HSC compartment that contains these cells in adult MDS (139, 407).

Analysis of the stem and progenitor compartments in JMML revealed that, unlike AML, the structure of the haematopoietic hierarchy in JMML is usually maintained, although there are severe abnormalities within these compartments. The most prominent abnormality that I found was the presence of an aberrant CD34+CD38-CD90+CD45RA+ population that does not appear to exist in samples from healthy donors across different ages and developmental stages. Assessment of the lineage potential of JMML HSPCs revealed abnormalities across all lineages with extensive myelomonocytic lineage expansion; reduced lymphoid and megakaryocytic potential and severe abnormalities within the erythroid series. These abnormalities were noted in all relevant HSPC compartments, including HSC further implicating this compartment in the pathogenesis of JMML and in line with the disease clinical features. The aberrant CD34+CD38-CD90+CD45RA+ population appeared to have bi-lineage potential indicating that this population may be derived from HSC near the top of the haematopoietic hierarchy but lacking normal erythroid/megakaryocytic and lymphoid output (Figure 6.1).

The stem cell capacity of JMML HSPC was evaluated with in vitro and in vivo assays. This confirmed that cells within the JMML HSC compartment were the population that consistently propagated the disease in all patients assessed with variable engraftment potential detected among downstream progenitors. Mouse recipients of

JMML HSCs demonstrated marked BM infiltration (90-95%) causing the mice to become unwell at ~10 weeks post-transplantation. The recipients of downstream progenitors including the CD34+CD38-CD90+CD45RA+ population demonstrated late engraftment but significant illness to recipients was never observed even in long-term follow up. Unlike AML, where HSC acquire the first genetic (pre-malignant) defect, with further mutations to downstream progenitors required for disease-associated LSC to arise, in JMML, cells within the HSC compartment robustly gave rise to disease in xenografts and propagated a much more aggressive phenotype compared with mice transplanted with downstream progenitors.

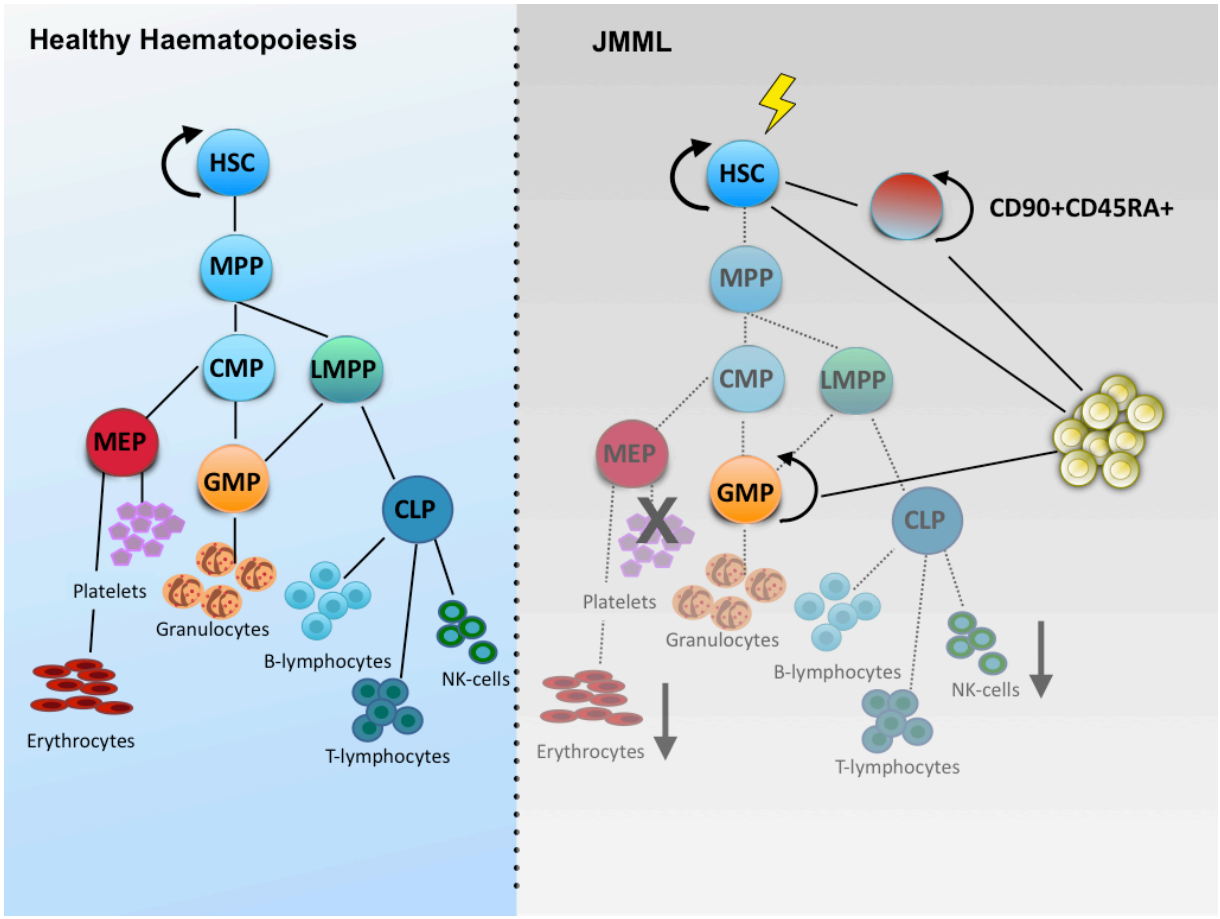


Figure 6.1: Schematic representation of the haematopoietic hierarchy and lineage output in normal haematopoiesis and in JMML

6.4 Molecular characterisation of JMML HSPC

Fundamental to this work was the identification of the origin of JMML mutations in order to validate the stem cell assay results. With the single cell genotyping technique I developed, I was able to map the JMML mutations onto the haematopoietic hierarchy and identify the cellular compartment where the JMML mutations arose. This work demonstrated for the first time that the JMML HSC is the population where all molecular events arose, including secondary mutations. The model of mutation acquisition was linear in the majority of patients tested with only 1 patient showing branching of secondary events. Tracking of mutations post-BMT revealed a similar pattern with the JMML mutations traced back to the HSC and the same molecular events propagating relapse. This together with HSPC analysis of serial patient samples after relapse indicated that the disease-originating population also propagates the relapse carrying the same mutational background with no additional events observed at relapse.

These data, together with the stem cell assays, demonstrate that JMML HSC is the disease-propagating population and I found no evidence, unlike AML, of a pre-leukaemic state harboring the original mutations. In AML pre-leukaemic HSCs can have the initiating mutation, but this does not lead to the disease development. In JMML, single RAS pathway mutations in HSC were often sufficient to cause disease and lead to accelerated myeloid proliferation, impaired erythro-megakaryocytic and poor lymphoid output. This was further validated with RNA sequencing where erythroid and lymphoid signatures at the HSC level were disturbed and upregulation

of myeloid genes, such as *CSF1R,* were observed in JMML. Gene expression data supported the conclusion that the structure of the haematopoietic hierarchy was maintained in JMML, with proportional expression of stem cell, fetal genes in the JMML HSC compared to controls; however additional transcriptomic signatures were detected in the JMML HSPC setting them apart from the normal counterparts, including up-regulation of cell cycle regulators and proliferation and DNA repair genes. The trancriptomic signature of the aberrant CD34+CD38-CD90+CD45RA+ revealed that this population shares molecular HSC as well as GMP signatures, with high expression of genes implicated in cell division, mitosis and proliferation, demonstrating that this is likely to be an HSC-derived transformed progenitor that is mainly characterised by high proliferation potential and myeloid output in line with the disease features.

Focused analysis of uniquely expressed genetic signatures that characterise the JMML LSC lead to the identification of targets that can be utilised therapeutically. During my project the most significant I identified were *mTOR*, *PIM1* and *GLUT-1*, with significant reduction of JMML proliferation and clonogenic capacity in in vitro assays. Furthermore, this work led to identification of aberrant expression of a cell surface marker *CD96* that detected JMML LSC pre and post-BMT, before conventional methods detected relapse, providing the opportunity for CD96 to be used as a biomarker for low level residual disease monitoring.

6.5 Implications of this work

This project has established the foundation of the first UK NIHR Paediatric MDS/JMML study. This study has allowed the collection of deep phenotypic data and the molecular characterisation of children with dysplasia and JMML. The targeted NGS panel that I developed has been used nationally in order to identify disease-associated mutations for these children and following careful validation of the results using a diagnostically accredited lab standards I was able to feed back the findings to clinicians. This allowed the study to actively contribute to the management of these children in real time, thereby providing an immediate translational aspect to my work. The data collected by the study so far has provided the evidence that allowed me to sub-classify this very heterogeneous group of paediatric myelodysplasia and devise diagnostic and management algorithms that may accelerate the diagnosis of these children, as well as tailor surveillance, allowing clinicians to detect disease progression or relapse early. The results from the targeted NGS panel as well as WGS, demonstrated at a national level that the diagnostics of children with possible MDS/JMML needs to evolve further in order to fully understand these conditions. By providing such evidence to Genomics England the study has enable the approval and commissioning of WGS as a first line diagnostic tool in the NHS, which should rapidly improve the diagnostic yield of these cases. Through the experience gained from the study I was also able to advice on the pathway the of the future analysis of these samples nationally and ensure that diagnostics for childhood MDS will continue to evolve. The biological samples collected during my project can also form the foundation of future research in Paediatric MDS and JMML.

My work on JMML haematopoiesis allowed me to address one of the most clinically important questions for this disease, the identification of the disease-propagating population at diagnosis and relapse. This offers the opportunity to identify disease-specific targets and one such target, CD96, seems particularly promising both for minimal residual disease monitoring and/or, given the strategic function of the CD96 receptor on T-cells, as a target for immune checkpoint inhibition (403) especially post-BMT. Targeting of up-regulated pathways, such as *mTOR* and *PIM1*, may also be used pre- or post- BMT to reduce disease burden or eradicate residual clones that have escaped myeloablative transplantation.

6.6 Limitations of this work

One of the major limitations of my project is the relatively small number of samples available to study. Aside from the fact that these diseases are very rare, the prospective nature of the study did not allow the opportunity to expand this work to larger cohorts, limiting the development of phenotypic: genotypic correlations for the duration of my project. Data collected by the study in the future should allow the collection of high quality phenotypic:genotypic data that will address this point. The analysis of JMML patient samples was again limited to the patients recruited to the study, with the addition of one archived JMML patient that was sourced from the Bloodwise Childhood Leukaemia Cell Bank. This has perhaps limited the opportunity for some the results to reach statistical significance, however the patterns observed for all experiments were very consistent allowing me to reach conclusive conclusions. Going forward, incorporation of the samples I collected into a national bank should accelerate research into paediatric MDS, including JMML.

The NIHR study was formed as an observational study aiming not to impact on the patient's routine care, which meant that timepoints for sample collection were non-instructive to clinicians; the treatment patients received were based on the available resources in each centre and timing of treatment and follow up was not consistent among patients. This meant that a great number of MDS patients did not have serial samples collected and monitoring of the JMML patients residual disease was not always consistent among centres especially post-BMT. This led to intervals where disease monitoring was not systematic or consistent, limiting the number of direct comparisons between serial patient samples and assessment of the effects of pre-transplant treatments, such as azacitidine. However, for the purposes of addressing the main hypothesis of my project, diagnostic patient samples were used, thus overcoming the lack of structured sample collection.

6.7 Future work

This thesis has enabled the study of comprehensive characterization of Paediatric MDS and overlap syndromes. The systematic nature of this study has been internationally recognised and following presentation of these data at international conferences the UK was invited to join the European Paediatric MDS/JMML network (EWOG-MDS) and participate in European and International collaborations such as the formation of an international JMML network among Europe and the United States. These collaborations will allow the study of large patient cohorts and should help to advance research in this field.

My work has addressed one of the most fundamental questions in JMML, however many questions remain to be answered. HbF is a well recognised poor prognostic

marker in JMML; but the underlying mechanisms that drive this remain very poorly understood. My project was primarily focused on the mechanisms that lead to aberrant acquisition of stem cell activity and propagate the disease. Understanding the mechanisms that drive high HbF expression may provide further insights in the mechanisms that lead to disease evolution at high risk patients as well as provide an understanding of the underlying mechanisms that set patients to develop a more aggressive disease.

Lineage output appears to be disrupted in JMML HSPC. My RNA sequencing approach was mainly focused on identifying the pathways upregulated in the disease LSC; my project only touched on the transcriptomic signatures associated with impaired erythroid and lymphoid output in JMML. These need to be further understood with the study of the transcriptomic signatures of the downstream progenitors especially MPP, as this appears to be the population in which most of the cells have no mutations detected.

Due to the time limitations of a DPhil project and the duration of serial xenotransplantations, I did not have the opportunity to repeat the serial transplantation experiments in more replicates, as well as proceed in vivo drug testing of some the targets identified by my transcriptomic analysis. In addition, further studies on serial samples will help to validate the value of regular flow cytometry, including the use of CD96, and targeted NGS for monitoring post-BMT.

6.8 Conclusions

Overall this project has set the foundations for ongoing careful and comprehensive characterisation of children with myelodysplasia, creating a national platform for characterisation of these diseases. My work on JMML has led to the identification for the first time of the disease-propagating population and a number of possible druggable targets and/or biomarkers that can be further investigated as tools for developing for future more effective management of this disease.

REFERENCES

1. Niemeyer CM, Baumann I. Classification of childhood aplastic anemia and myelodysplastic syndrome. Hematology American Society of Hematology Education Program. 2011;2011:84-9.
2. Tuncer MA, Pagliuca A, Hicsonmez G, Yetgin S, Ozsoylu S, Mufti GJ. Primary myelodysplastic syndrome in children: the clinical experience in 33 cases. Br J Haematol. 1992;82(2):347-53.
3. Creutzig U, Cantu-Rajnoldi A, Ritter J, Romitti L, Odenwald E, Conter V, et al. Myelodysplastic syndromes in childhood. Report of 21 patients from Italy and West Germany. Am J Pediatr Hematol Oncol. 1987;9(4):324-30.
4. Hasle H. Myelodysplastic syndromes in childhood--classification, epidemiology, and treatment. Leuk Lymphoma. 1994;13(1-2):11-26.
5. Hasle H. Myelodysplastic and myeloproliferative disorders of childhood. Hematology American Society of Hematology Education Program. 2016;2016(1):598-604.
6. Hasle H, Wadsworth LD, Massing BG, McBride M, Schultz KR. A population-based study of childhood myelodysplastic syndrome in British Columbia, Canada. Br J Haematol. 1999;106(4):1027-32.
7. Passmore SJ, Chessells JM, Kempski H, Hann IM, Brownbill PA, Stiller CA. Paediatric myelodysplastic syndromes and juvenile myelomonocytic leukaemia in the UK: a population-based study of incidence and survival. Br J Haematol. 2003;121(5):758-67.
8. Jackson GH, Carey PJ, Cant AJ, Bown NP, Reid MM. Myelodysplastic syndromes in children. Br J Haematol. 1993;84(1):185-6.
9. Sasaki H, Manabe A, Kojima S, Tsuchida M, Hayashi Y, Ikuta K, et al. Myelodysplastic syndrome in childhood: a retrospective study of 189 patients in Japan. Leukemia. 2001;15(11):1713-20.
10. Strahm B, Nollke P, Zecca M, Korthof ET, Bierings M, Furlan I, et al. Hematopoietic stem cell transplantation for advanced myelodysplastic syndrome in children: results of the EWOG-MDS 98 study. Leukemia. 2011;25(3):455-62.
11. Ward E, DeSantis C, Robbins A, Kohler B, Jemal A. Childhood and adolescent cancer statistics, 2014. CA Cancer J Clin. 2014;64(2):83-103.
12. Arber DA, Orazi A, Hasserjian R, Thiele J, Borowitz MJ, Le Beau MM, et al. The 2016 revision to the World Health Organization classification of myeloid neoplasms and acute leukemia. Blood. 2016;127(20):2391-405.
13. Hasle H, Niemeyer CM, Chessells JM, Baumann I, Bennett JM, Kerndrup G, et al. A pediatric approach to the WHO classification of myelodysplastic and myeloproliferative diseases. Leukemia. 2003;17(2):277-82.
14. Helsmoortel HH, Bresolin S, Lammens T, Cave H, Noellke P, Caye A, et al. LIN28B overexpression defines a novel fetal-like subgroup of juvenile myelomonocytic leukemia. Blood. 2016;127(9):1163-72.
15. Kardos G, Baumann I, Passmore SJ, Locatelli F, Hasle H, Schultz KR, et al. Refractory anemia in childhood: a retrospective analysis of 67 patients with particular reference to monosomy 7. Blood. 2003;102(6):1997-2003.
16. Hasegawa D, Chen X, Hirabayashi S, Ishida Y, Watanabe S, Zaike Y, et al. Clinical characteristics and treatment outcome in 65 cases with refractory cytopenia of childhood defined according to the WHO 2008 classification. Br J Haematol. 2014;166(5):758-66.

17. Baumann I, Fuhrer M, Behrendt S, Campr V, Csomor J, Furlan I, et al. Morphological differentiation of severe aplastic anaemia from hypocellular refractory cytopenia of childhood: reproducibility of histopathological diagnostic criteria. Histopathology. 2012;61(1):10-7.
18. Bennett JM, Orazi A. Diagnostic criteria to distinguish hypocellular acute myeloid leukemia from hypocellular myelodysplastic syndromes and aplastic anemia: recommendations for a standardized approach. Haematologica. 2009;94(2):264-8.
19. Bhatia S, Francisco L, Carter A, Sun CL, Baker KS, Gurney JG, et al. Late mortality after allogeneic hematopoietic cell transplantation and functional status of long-term survivors: report from the Bone Marrow Transplant Survivor Study. Blood. 2007;110(10):3784-92.
20. Kushner BH, Cheung NK, Kramer K, Heller G, Jhanwar SC. Neuroblastoma and treatment-related myelodysplasia/leukemia: the Memorial Sloan-Kettering experience and a literature review. Journal of clinical oncology : official journal of the American Society of Clinical Oncology. 1998;16(12):3880-9.
21. Le Deley MC, Leblanc T, Shamsaldin A, Raquin MA, Lacour B, Sommelet D, et al. Risk of secondary leukemia after a solid tumor in childhood according to the dose of epipodophyllotoxins and anthracyclines: a case-control study by the Societe Francaise d'Oncologie Pediatrique. Journal of clinical oncology : official journal of the American Society of Clinical Oncology. 2003;21(6):1074-81.
22. Bhatia S, Sather HN, Pabustan OB, Trigg ME, Gaynon PS, Robison LL. Low incidence of second neoplasms among children diagnosed with acute lymphoblastic leukemia after 1983. Blood. 2002;99(12):4257-64.
23. Pui CH, Ribeiro RC, Hancock ML, Rivera GK, Evans WE, Raimondi SC, et al. Acute myeloid leukemia in children treated with epipodophyllotoxins for acute lymphoblastic leukemia. N Engl J Med. 1991;325(24):1682-7.
24. Smith SM, Le Beau MM, Huo D, Karrison T, Sobecks RM, Anastasi J, et al. Clinical-cytogenetic associations in 306 patients with therapy-related myelodysplasia and myeloid leukemia: the University of Chicago series. Blood. 2003;102(1):43-52.
25. Heuser M. Therapy-related myeloid neoplasms: does knowing the origin help to guide treatment? Hematology American Society of Hematology Education Program. 2016;2016(1):24-32.
26. Ok CY, Patel KP, Garcia-Manero G, Routbort MJ, Fu B, Tang G, et al. Mutational profiling of therapy-related myelodysplastic syndromes and acute myeloid leukemia by next generation sequencing, a comparison with de novo diseases. Leukemia research. 2015;39(3):348-54.
27. Wong TN, Ramsingh G, Young AL, Miller CA, Touma W, Welch JS, et al. Role of TP53 mutations in the origin and evolution of therapy-related acute myeloid leukaemia. Nature. 2015;518(7540):552-5.
28. Gohring G, Thomay K, Schmidt G, Ripperger T, Xu M, Wittner N, et al. A common ancestral DNMT3A-mutated preleukemic clone giving rise to AML and MDS in an adolescent girl. Leuk Lymphoma. 2017;58(3):718-21.
29. Schwartz JR, Ma J, Lamprecht T, Walsh M, Wang S, Bryant V, et al. The genomic landscape of pediatric myelodysplastic syndromes. Nat Commun. 2017;8(1):1557.
30. Bagby GC, Lipton JM, Sloand EM, Schiffer CA. Marrow failure. Hematology American Society of Hematology Education Program. 2004:318-36.
31. Shimamura A. Inherited bone marrow failure syndromes: molecular features. Hematology American Society of Hematology Education Program. 2006:63-71.

32. Locatelli F, Strahm B. How I treat myelodysplastic syndromes of childhood. Blood. 2018;131(13):1406-14.
33. Ortmann CA, Niemeyer CM, Wawer A, Ebell W, Baumann I, Kratz CP. TERC mutations in children with refractory cytopenia. Haematologica. 2006;91(5):707-8.
34. Keel SB, Scott A, Sanchez-Bonilla M, Ho PA, Gulsuner S, Pritchard CC, et al. Genetic features of myelodysplastic syndrome and aplastic anemia in pediatric and young adult patients. Haematologica. 2016;101(11):1343-50.
35. Cada M, Segbefia CI, Klaassen R, Fernandez CV, Yanofsky RA, Wu J, et al. The impact of category, cytopathology and cytogenetics on development and progression of clonal and malignant myeloid transformation in inherited bone marrow failure syndromes. Haematologica. 2015;100(5):633-42.
36. Deakyne JS, Mazin AV. Fanconi anemia: at the crossroads of DNA repair. Biochemistry (Mosc). 2011;76(1):36-48.
37. Rosenberg PS, Tamary H, Alter BP. How high are carrier frequencies of rare recessive syndromes? Contemporary estimates for Fanconi Anemia in the United States and Israel. American journal of medical genetics Part A. 2011;155A(8):1877-83.
38. Kee Y, D'Andrea AD. Molecular pathogenesis and clinical management of Fanconi anemia. The Journal of clinical investigation. 2012;122(11):3799-806.
39. Rosenberg PS, Greene MH, Alter BP. Cancer incidence in persons with Fanconi anemia. Blood. 2003;101(3):822-6.
40. Tonnies H, Huber S, Kuhl JS, Gerlach A, Ebell W, Neitzel H. Clonal chromosomal aberrations in bone marrow cells of Fanconi anemia patients: gains of the chromosomal segment 3q26q29 as an adverse risk factor. Blood. 2003;101(10):3872-4.
41. Dokal I. Dyskeratosis congenita. Hematology American Society of Hematology Education Program. 2011;2011:480-6.
42. Ballew BJ, Savage SA. Updates on the biology and management of dyskeratosis congenita and related telomere biology disorders. Expert Rev Hematol. 2013;6(3):327-37.
43. Shimamura A, Alter BP. Pathophysiology and management of inherited bone marrow failure syndromes. Blood Rev. 2010;24(3):101-22.
44. Myers K. Shwachman-Diamond Syndrome. In: Adam MP, Ardinger HH, Pagon RA, Wallace SE, Bean LJH, Stephens K, et al., editors. GeneReviews((R)). Seattle (WA)1993.
45. Boocock GR, Morrison JA, Popovic M, Richards N, Ellis L, Durie PR, et al. Mutations in SBDS are associated with Shwachman-Diamond syndrome. Nat Genet. 2003;33(1):97-101.
46. Dror Y, Durie P, Ginzberg H, Herman R, Banerjee A, Champagne M, et al. Clonal evolution in marrows of patients with Shwachman-Diamond syndrome: a prospective 5-year follow-up study. Exp Hematol. 2002;30(7):659-69.
47. Donadieu J, Fenneteau O, Beaupain B, Beaufils S, Bellanger F, Mahlaoui N, et al. Classification of and risk factors for hematologic complications in a French national cohort of 102 patients with Shwachman-Diamond syndrome. Haematologica. 2012;97(9):1312-9.
48. Lindsley RC, Saber W, Mar BG, Redd R, Wang T, Haagenson MD, et al. Prognostic Mutations in Myelodysplastic Syndrome after Stem-Cell Transplantation. N Engl J Med. 2017;376(6):536-47.
49. Horwitz MS, Corey SJ, Grimes HL, Tidwell T. ELANE mutations in cyclic and severe congenital neutropenia: genetics and pathophysiology. Hematol Oncol Clin North Am. 2013;27(1):19-41, vii.

50. Boztug K, Klein C. Genetics and pathophysiology of severe congenital neutropenia syndromes unrelated to neutrophil elastase. Hematol Oncol Clin North Am. 2013;27(1):43-60, vii.
51. Rosenberg PS, Zeidler C, Bolyard AA, Alter BP, Bonilla MA, Boxer LA, et al. Stable long-term risk of leukaemia in patients with severe congenital neutropenia maintained on G-CSF therapy. Br J Haematol. 2010;150(2):196-9.
52. Rosenberg PS, Alter BP, Bolyard AA, Bonilla MA, Boxer LA, Cham B, et al. The incidence of leukemia and mortality from sepsis in patients with severe congenital neutropenia receiving long-term G-CSF therapy. Blood. 2006;107(12):4628-35.
53. Rosenberg PS, Alter BP, Link DC, Stein S, Rodger E, Bolyard AA, et al. Neutrophil elastase mutations and risk of leukaemia in severe congenital neutropenia. Br J Haematol. 2008;140(2):210-3.
54. Vlachos A, Rosenberg PS, Atsidaftos E, Alter BP, Lipton JM. Incidence of neoplasia in Diamond Blackfan anemia: a report from the Diamond Blackfan Anemia Registry. Blood. 2012;119(16):3815-9.
55. Balaban EP, Buchanan GR, Graham M, Frenkel EP. Diamond-Blackfan syndrome in adult patients. Am J Med. 1985;78(3):533-8.
56. Vlachos A, Muir E. How I treat Diamond-Blackfan anemia. Blood. 2010;116(19):3715-23.
57. Fargo JH, Kratz CP, Giri N, Savage SA, Wong C, Backer K, et al. Erythrocyte adenosine deaminase: diagnostic value for Diamond-Blackfan anaemia. Br J Haematol. 2013;160(4):547-54.
58. Ballmaier M, Germeshausen M. Congenital amegakaryocytic thrombocytopenia: clinical presentation, diagnosis, and treatment. Semin Thromb Hemost. 2011;37(6):673-81.
59. Maserati E, Panarello C, Morerio C, Valli R, Pressato B, Patitucci F, et al. Clonal chromosome anomalies and propensity to myeloid malignancies in congenital amegakaryocytic thrombocytopenia (OMIM 604498). Haematologica. 2008;93(8):1271-3.
60. Leguit RJ, van den Tweel JG. The pathology of bone marrow failure. Histopathology. 2010;57(5):655-70.
61. Socie G, Henry-Amar M, Bacigalupo A, Hows J, Tichelli A, Ljungman P, et al. Malignant tumors occurring after treatment of aplastic anemia. European Bone Marrow Transplantation-Severe Aplastic Anaemia Working Party. N Engl J Med. 1993;329(16):1152-7.
62. Ohara A, Kojima S, Hamajima N, Tsuchida M, Imashuku S, Ohta S, et al. Myelodysplastic syndrome and acute myelogenous leukemia as a late clonal complication in children with acquired aplastic anemia. Blood. 1997;90(3):1009-13.
63. Fuhrer M, Rampf U, Baumann I, Faldum A, Niemeyer C, Janka-Schaub G, et al. Immunosuppressive therapy for aplastic anemia in children: a more severe disease predicts better survival. Blood. 2005;106(6):2102-4.
64. Hasle H, Niemeyer CM. Advances in the prognostication and management of advanced MDS in children. Br J Haematol. 2011;154(2):185-95.
65. Kojima S, Ohara A, Tsuchida M, Kudoh T, Hanada R, Okimoto Y, et al. Risk factors for evolution of acquired aplastic anemia into myelodysplastic syndrome and acute myeloid leukemia after immunosuppressive therapy in children. Blood. 2002;100(3):786-90.

66. Babushok DV, Perdigones N, Perin JC, Olson TS, Ye W, Roth JJ, et al. Emergence of clonal hematopoiesis in the majority of patients with acquired aplastic anemia. Cancer Genet. 2015;208(4):115-28.
67. Pippucci T, Savoia A, Perrotta S, Pujol-Moix N, Noris P, Castegnaro G, et al. Mutations in the 5' UTR of ANKRD26, the ankirin repeat domain 26 gene, cause an autosomal-dominant form of inherited thrombocytopenia, THC2. Am J Hum Genet. 2011;88(1):115-20.
68. Necchi V, Balduini A, Noris P, Barozzi S, Sommi P, di Buduo C, et al. Ubiquitin/proteasome-rich particulate cytoplasmic structures (PaCSs) in the platelets and megakaryocytes of ANKRD26-related thrombo-cytopenia. Thromb Haemost. 2013;109(2):263-71.
69. Noris P, Favier R, Alessi MC, Geddis AE, Kunishima S, Heller PG, et al. ANKRD26-related thrombocytopenia and myeloid malignancies. Blood. 2013;122(11):1987-9.
70. Zhang MY, Churpek JE, Keel SB, Walsh T, Lee MK, Loeb KR, et al. Germline ETV6 mutations in familial thrombocytopenia and hematologic malignancy. Nat Genet. 2015;47(2):180-5.
71. Noetzli L, Lo RW, Lee-Sherick AB, Callaghan M, Noris P, Savoia A, et al. Germline mutations in ETV6 are associated with thrombocytopenia, red cell macrocytosis and predisposition to lymphoblastic leukemia. Nat Genet. 2015;47(5):535-8.
72. Owen CJ, Toze CL, Koochin A, Forrest DL, Smith CA, Stevens JM, et al. Five new pedigrees with inherited RUNX1 mutations causing familial platelet disorder with propensity to myeloid malignancy. Blood. 2008;112(12):4639-45.
73. Beri-Dexheimer M, Latger-Cannard V, Philippe C, Bonnet C, Chambon P, Roth V, et al. Clinical phenotype of germline RUNX1 haploinsufficiency: from point mutations to large genomic deletions. Eur J Hum Genet. 2008;16(8):1014-8.
74. Jongmans MC, Kuiper RP, Carmichael CL, Wilkins EJ, Dors N, Carmagnac A, et al. Novel RUNX1 mutations in familial platelet disorder with enhanced risk for acute myeloid leukemia: clues for improved identification of the FPD/AML syndrome. Leukemia. 2010;24(1):242-6.
75. Preudhomme C, Renneville A, Bourdon V, Philippe N, Roche-Lestienne C, Boissel N, et al. High frequency of RUNX1 biallelic alteration in acute myeloid leukemia secondary to familial platelet disorder. Blood. 2009;113(22):5583-7.
76. Pabst T, Eyholzer M, Haefliger S, Schardt J, Mueller BU. Somatic CEBPA mutations are a frequent second event in families with germline CEBPA mutations and familial acute myeloid leukemia. Journal of clinical oncology : official journal of the American Society of Clinical Oncology. 2008;26(31):5088-93.
77. Taskesen E, Bullinger L, Corbacioglu A, Sanders MA, Erpelinck CA, Wouters BJ, et al. Prognostic impact, concurrent genetic mutations, and gene expression features of AML with CEBPA mutations in a cohort of 1182 cytogenetically normal AML patients: further evidence for CEBPA double mutant AML as a distinctive disease entity. Blood. 2011;117(8):2469-75.
78. Tawana K, Wang J, Renneville A, Bodor C, Hills R, Loveday C, et al. Disease evolution and outcomes in familial AML with germline CEBPA mutations. Blood. 2015;126(10):1214-23.
79. Hahn CN, Chong CE, Carmichael CL, Wilkins EJ, Brautigan PJ, Li XC, et al. Heritable GATA2 mutations associated with familial myelodysplastic syndrome and acute myeloid leukemia. Nat Genet. 2011;43(10):1012-7.
80. Pasquet M, Bellanne-Chantelot C, Tavitian S, Prade N, Beaupain B, Larochelle O, et al. High frequency of GATA2 mutations in patients with mild chronic neutropenia

evolving to MonoMac syndrome, myelodysplasia, and acute myeloid leukemia. Blood. 2013;121(5):822-9.
81. Micol JB, Abdel-Wahab O. Collaborating constitutive and somatic genetic events in myeloid malignancies: ASXL1 mutations in patients with germline GATA2 mutations. Haematologica. 2014;99(2):201-3.
82. Collin M, Dickinson R, Bigley V. Haematopoietic and immune defects associated with GATA2 mutation. Br J Haematol. 2015;169(2):173-87.
83. Fisher KE, Hsu AP, Williams CL, Sayeed H, Merritt BY, Elghetany MT, et al. Somatic mutations in children with GATA2-associated myelodysplastic syndrome who lack other features of GATA2 deficiency. Blood Adv. 2017;1(7):443-8.
84. Wlodarski MW, Hirabayashi S, Pastor V, Stary J, Hasle H, Masetti R, et al. Prevalence, clinical characteristics, and prognosis of GATA2-related myelodysplastic syndromes in children and adolescents. Blood. 2016;127(11):1387-97; quiz 518.
85. Dickinson RE, Griffin H, Bigley V, Reynard LN, Hussain R, Haniffa M, et al. Exome sequencing identifies GATA-2 mutation as the cause of dendritic cell, monocyte, B and NK lymphoid deficiency. Blood. 2011;118(10):2656-8.
86. Ostergaard P, Simpson MA, Connell FC, Steward CG, Brice G, Woollard WJ, et al. Mutations in GATA2 cause primary lymphedema associated with a predisposition to acute myeloid leukemia (Emberger syndrome). Nat Genet. 2011;43(10):929-31.
87. Rodrigues NP, Janzen V, Forkert R, Dombkowski DM, Boyd AS, Orkin SH, et al. Haploinsufficiency of GATA-2 perturbs adult hematopoietic stem-cell homeostasis. Blood. 2005;106(2):477-84.
88. Lim KC, Hosoya T, Brandt W, Ku CJ, Hosoya-Ohmura S, Camper SA, et al. Conditional Gata2 inactivation results in HSC loss and lymphatic mispatterning. The Journal of clinical investigation. 2012;122(10):3705-17.
89. Wang X, Muramatsu H, Okuno Y, Sakaguchi H, Yoshida K, Kawashima N, et al. GATA2 and secondary mutations in familial myelodysplastic syndromes and pediatric myeloid malignancies. Haematologica. 2015;100(10):e398-401.
90. Bodor C, Renneville A, Smith M, Charazac A, Iqbal S, Etancelin P, et al. Germ-line GATA2 p.THR354MET mutation in familial myelodysplastic syndrome with acquired monosomy 7 and ASXL1 mutation demonstrating rapid onset and poor survival. Haematologica. 2012;97(6):890-4.
91. Stieglitz E, Loh ML. Pediatric MDS: GATA screen the germline. Blood. 2016;127(11):1377-8.
92. Yoshimi A, Toya T, Kawazu M, Ueno T, Tsukamoto A, Iizuka H, et al. Recurrent CDC25C mutations drive malignant transformation in FPD/AML. Nat Commun. 2014;5:4770.
93. Kirwan M, Walne AJ, Plagnol V, Velangi M, Ho A, Hossain U, et al. Exome sequencing identifies autosomal-dominant SRP72 mutations associated with familial aplasia and myelodysplasia. Am J Hum Genet. 2012;90(5):888-92.
94. Li R, Sobreira N, Witmer PD, Pratz KW, Braunstein EM. Two novel germline DDX41 mutations in a family with inherited myelodysplasia/acute myeloid leukemia. Haematologica. 2016;101(6):e228-31.
95. Polprasert C, Schulze I, Sekeres MA, Makishima H, Przychodzen B, Hosono N, et al. Inherited and Somatic Defects in DDX41 in Myeloid Neoplasms. Cancer Cell. 2015;27(5):658-70.
96. Nagamachi A, Matsui H, Asou H, Ozaki Y, Aki D, Kanai A, et al. Haploinsufficiency of SAMD9L, an endosome fusion facilitator, causes myeloid malignancies in mice mimicking human diseases with monosomy 7. Cancer Cell. 2013;24(3):305-17.

97. Li CF, MacDonald JR, Wei RY, Ray J, Lau K, Kandel C, et al. Human sterile alpha motif domain 9, a novel gene identified as down-regulated in aggressive fibromatosis, is absent in the mouse. BMC Genomics. 2007;8:92.
98. Wang Q, Zhai YY, Dai JH, Li KY, Deng Q, Han ZG. SAMD9L inactivation promotes cell proliferation via facilitating G1-S transition in hepatitis B virus-associated hepatocellular carcinoma. Int J Biol Sci. 2014;10(8):807-16.
99. Tesi B, Davidsson J, Voss M, Rahikkala E, Holmes TD, Chiang SCC, et al. Gain-of-function SAMD9L mutations cause a syndrome of cytopenia, immunodeficiency, MDS, and neurological symptoms. Blood. 2017;129(16):2266-79.
100. Chen DH, Below JE, Shimamura A, Keel SB, Matsushita M, Wolff J, et al. Ataxia-Pancytopenia Syndrome Is Caused by Missense Mutations in SAMD9L. Am J Hum Genet. 2016;98(6):1146-58.
101. Narumi S, Amano N, Ishii T, Katsumata N, Muroya K, Adachi M, et al. SAMD9 mutations cause a novel multisystem disorder, MIRAGE syndrome, and are associated with loss of chromosome 7. Nat Genet. 2016;48(7):792-7.
102. Davidsson J, Puschmann A, Tedgard U, Bryder D, Nilsson L, Cammenga J. SAMD9 and SAMD9L in inherited predisposition to ataxia, pancytopenia, and myeloid malignancies. Leukemia. 2018;32(5):1106-15.
103. Pastor VB, Sahoo SS, Boklan J, Schwabe GC, Saribeyoglu E, Strahm B, et al. Constitutional SAMD9L mutations cause familial myelodysplastic syndrome and transient monosomy 7. Haematologica. 2018;103(3):427-37.
104. Lek M, Karczewski KJ, Minikel EV, Samocha KE, Banks E, Fennell T, et al. Analysis of protein-coding genetic variation in 60,706 humans. Nature. 2016;536(7616):285-91.
105. Gohring G, Michalova K, Beverloo HB, Betts D, Harbott J, Haas OA, et al. Complex karyotype newly defined: the strongest prognostic factor in advanced childhood myelodysplastic syndrome. Blood. 2010;116(19):3766-9.
106. Martinez-Climent JA. Molecular cytogenetics of childhood hematological malignancies. Leukemia. 1997;11(12):1999-2021.
107. Nowell PC. Cytogenetics of preleukemia. Cancer genetics and cytogenetics. 1982;5(3):265-78.
108. Luna-Fineman S, Shannon KM, Atwater SK, Davis J, Masterson M, Ortega J, et al. Myelodysplastic and myeloproliferative disorders of childhood: a study of 167 patients. Blood. 1999;93(2):459-66.
109. Maserati E, Aprili F, Vinante F, Locatelli F, Amendola G, Zatterale A, et al. Trisomy 8 in myelodysplasia and acute leukemia is constitutional in 15-20% of cases. Genes Chromosomes Cancer. 2002;33(1):93-7.
110. Quentin S, Cuccuini W, Ceccaldi R, Nibourel O, Pondarre C, Pages MP, et al. Myelodysplasia and leukemia of Fanconi anemia are associated with a specific pattern of genomic abnormalities that includes cryptic RUNX1/AML1 lesions. Blood. 2011;117(15):e161-70.
111. Parker TM, Klaassen RJ, Johnston DL. Spontaneous remission of myelodysplastic syndrome with monosomy 7 in a young boy. Cancer genetics and cytogenetics. 2008;182(2):122-5.
112. Mantadakis E, Shannon KM, Singer DA, Finklestein J, Chan KW, Hilden JM, et al. Transient monosomy 7: a case series in children and review of the literature. Cancer. 1999;85(12):2655-61.
113. Pastor V, Hirabayashi S, Karow A, Wehrle J, Kozyra EJ, Nienhold R, et al. Mutational landscape in children with myelodysplastic syndromes is distinct from

adults: specific somatic drivers and novel germline variants. Leukemia. 2017;31(3):759-62.
114. Zhang MY, Keel SB, Walsh T, Lee MK, Gulsuner S, Watts AC, et al. Genomic analysis of bone marrow failure and myelodysplastic syndromes reveals phenotypic and diagnostic complexity. Haematologica. 2015;100(1):42-8.
115. Smith ML, Cavenagh JD, Lister TA, Fitzgibbon J. Mutation of CEBPA in familial acute myeloid leukemia. N Engl J Med. 2004;351(23):2403-7.
116. Sekeres MA. The epidemiology of myelodysplastic syndromes. Hematol Oncol Clin North Am. 2010;24(2):287-94.
117. Papaemmanuil E, Gerstung M, Malcovati L, Tauro S, Gundem G, Van Loo P, et al. Clinical and biological implications of driver mutations in myelodysplastic syndromes. Blood. 2013;122(22):3616-27; quiz 99.
118. Woll PS, Kjallquist U, Chowdhury O, Doolittle H, Wedge DC, Thongjuea S, et al. Myelodysplastic syndromes are propagated by rare and distinct human cancer stem cells in vivo. Cancer Cell. 2014;25(6):794-808.
119. Ostendorf BN, Flenner E, Florcken A, Westermann J. Phenotypic characterization of aberrant stem and progenitor cell populations in myelodysplastic syndromes. PLoS One. 2018;13(5):e0197823.
120. Zhao X, Gao S, Wu Z, Kajigaya S, Feng X, Liu Q, et al. Single-cell RNA-seq reveals a distinct transcriptome signature of aneuploid hematopoietic cells. Blood. 2017;130(25):2762-73.
121. de Vries AC, Langerak AW, Verhaaf B, Niemeyer CM, Stary J, Schmiegelow K, et al. T-cell receptor Vbeta CDR3 oligoclonality frequently occurs in childhood refractory cytopenia (MDS-RC) and severe aplastic anemia. Leukemia. 2008;22(6):1170-4.
122. Aalbers AM, van den Heuvel-Eibrink MM, Baumann I, Beverloo HB, Driessen GJ, Dworzak M, et al. T-cell receptor Vbeta skewing frequently occurs in refractory cytopenia of childhood and is associated with an expansion of effector cytotoxic T cells: a prospective study by EWOG-MDS. Blood cancer journal. 2014;4:e209.
123. Sloand EM, Wu CO, Greenberg P, Young N, Barrett J. Factors affecting response and survival in patients with myelodysplasia treated with immunosuppressive therapy. Journal of clinical oncology : official journal of the American Society of Clinical Oncology. 2008;26(15):2505-11.
124. Aalbers AM, van den Heuvel-Eibrink MM, Baumann I, Dworzak M, Hasle H, Locatelli F, et al. Bone marrow immunophenotyping by flow cytometry in refractory cytopenia of childhood. Haematologica. 2015;100(3):315-23.
125. Yusuf U, Frangoul HA, Gooley TA, Woolfrey AE, Carpenter PA, Andrews RG, et al. Allogeneic bone marrow transplantation in children with myelodysplastic syndrome or juvenile myelomonocytic leukemia: the Seattle experience. Bone Marrow Transplant. 2004;33(8):805-14.
126. Woods WG, Barnard DR, Alonzo TA, Buckley JD, Kobrinsky N, Arthur DC, et al. Prospective study of 90 children requiring treatment for juvenile myelomonocytic leukemia or myelodysplastic syndrome: a report from the Children's Cancer Group. Journal of clinical oncology : official journal of the American Society of Clinical Oncology. 2002;20(2):434-40.
127. Anderson JE, Appelbaum FR, Fisher LD, Schoch G, Shulman H, Anasetti C, et al. Allogeneic bone marrow transplantation for 93 patients with myelodysplastic syndrome. Blood. 1993;82(2):677-81.
128. Basquiera AL, Pizzi S, Correas AG, Longo PG, Goldman WC, Prates MV, et al. Allogeneic hematopoietic stem cell transplantation in pediatric myelodysplastic

syndromes: a multicenter experience from Argentina. Pediatric blood & cancer. 2015;62(1):153-7.
129. Chang TY, Dvorak CC, Loh ML. Bedside to bench in juvenile myelomonocytic leukemia: insights into leukemogenesis from a rare pediatric leukemia. Blood. 2014;124(16):2487-97.
130. Madureira AB, Eapen M, Locatelli F, Teira P, Zhang MJ, Davies SM, et al. Analysis of risk factors influencing outcome in children with myelodysplastic syndrome after unrelated cord blood transplantation. Leukemia. 2011;25(3):449-54.
131. Yoshimi A, van den Heuvel-Eibrink MM, Baumann I, Schwarz S, Simonitsch-Klupp I, de Paepe P, et al. Comparison of horse and rabbit antithymocyte globulin in immunosuppressive therapy for refractory cytopenia of childhood. Haematologica. 2014;99(4):656-63.
132. Yoshimi A, Baumann I, Fuhrer M, Bergstrasser E, Gobel U, Sykora KW, et al. Immunosuppressive therapy with anti-thymocyte globulin and cyclosporine A in selected children with hypoplastic refractory cytopenia. Haematologica. 2007;92(3):397-400.
133. Strahm B, Locatelli F, Bader P, Ehlert K, Kremens B, Zintl F, et al. Reduced intensity conditioning in unrelated donor transplantation for refractory cytopenia in childhood. Bone Marrow Transplant. 2007;40(4):329-33.
134. Strahm B AM, Bierings M, et al. Reduced intensity conditioning for children and adolescents with refractory cytopenia of childhood. Bone Marrow Transplant. 2017;52(S103).
135. Cseh AM, Niemeyer CM, Yoshimi A, Catala A, Fruhwald MC, Hasle H, et al. Therapy with low-dose azacitidine for MDS in children and young adults: a retrospective analysis of the EWOG-MDS study group. Br J Haematol. 2016;172(6):930-6.
136. Glaubach T, Robinson LJ, Corey SJ. Pediatric myelodysplastic syndromes: they do exist! Journal of pediatric hematology/oncology. 2014;36(1):1-7.
137. Hasle H, Baumann I, Bergstrasser E, Fenu S, Fischer A, Kardos G, et al. The International Prognostic Scoring System (IPSS) for childhood myelodysplastic syndrome (MDS) and juvenile myelomonocytic leukemia (JMML). Leukemia. 2004;18(12):2008-14.
138. Emanuel PD, Shannon KM, Castleberry RP. Juvenile myelomonocytic leukemia: molecular understanding and prospects for therapy. Mol Med Today. 1996;2(11):468-75.
139. Locatelli F, Niemeyer CM. How I treat juvenile myelomonocytic leukemia. Blood. 2015;125(7):1083-90.
140. Niemeyer CM, Arico M, Basso G, Biondi A, Cantu Rajnoldi A, Creutzig U, et al. Chronic myelomonocytic leukemia in childhood: a retrospective analysis of 110 cases. European Working Group on Myelodysplastic Syndromes in Childhood (EWOG-MDS). Blood. 1997;89(10):3534-43.
141. Niemeyer CM, Kratz CP. Paediatric myelodysplastic syndromes and juvenile myelomonocytic leukaemia: molecular classification and treatment options. Br J Haematol. 2008;140(6):610-24.
142. Loh ML. Recent advances in the pathogenesis and treatment of juvenile myelomonocytic leukaemia. Br J Haematol. 2011;152(6):677-87.
143. Loh ML. Recent advances in the pathogenesis and treatment of juvenile myelomonocytic leukaemia. British Journal of Haematology. 2011;6:677–87
.
144. Bennett JM, Catovsky D, Daniel MT, Flandrin G, Galton DA, Gralnick H, et al. The chronic myeloid leukaemias: guidelines for distinguishing chronic granulocytic, atypical

chronic myeloid, and chronic myelomonocytic leukaemia. Proposals by the French-American-British Cooperative Leukaemia Group. Br J Haematol. 1994;87(4):746-54.
145. Yoshida K, Sanada M, Shiraishi Y, Nowak D, Nagata Y, Yamamoto R, et al. Frequent pathway mutations of splicing machinery in myelodysplasia. Nature. 2011;478(7367):64-9.
146. Patnaik MM, Lasho TL, Finke CM, Hanson CA, Hodnefield JM, Knudson RA, et al. Spliceosome mutations involving SRSF2, SF3B1, and U2AF35 in chronic myelomonocytic leukemia: prevalence, clinical correlates, and prognostic relevance. Am J Hematol. 2013;88(3):201-6.
147. Delhommeau F, Dupont S, Della Valle V, James C, Trannoy S, Masse A, et al. Mutation in TET2 in myeloid cancers. N Engl J Med. 2009;360(22):2289-301.
148. Gelsi-Boyer V, Trouplin V, Adelaide J, Bonansea J, Cervera N, Carbuccia N, et al. Mutations of polycomb-associated gene ASXL1 in myelodysplastic syndromes and chronic myelomonocytic leukaemia. Br J Haematol. 2009;145(6):788-800.
149. Emanuel PD, Bates LJ, Castleberry RP, Gualtieri RJ, Zuckerman KS. Selective hypersensitivity to granulocyte-macrophage colony-stimulating factor by juvenile chronic myeloid leukemia hematopoietic progenitors. Blood. 1991;77(5):925-9.
150. Emanuel PD, Bates LJ, Zhu SW, Castleberry RP, Gualtieri RJ, Zuckerman KS. The role of monocyte-derived hemopoietic growth factors in the regulation of myeloproliferation in juvenile chronic myelogenous leukemia. Exp Hematol. 1991;19(10):1017-24.
151. Chan RJ, Leedy MB, Munugalavadla V, Voorhorst CS, Li Y, Yu M, et al. Human somatic PTPN11 mutations induce hematopoietic-cell hypersensitivity to granulocyte-macrophage colony-stimulating factor. Blood. 2005;105(9):3737-42.
152. N K. Single-cell pro ling identi es aberrant STAT5 activation in myeloid malignancies with speci c clinical and biologic correlates. CANCER CELL. 2008.
153. Goodwin CB, Li XJ, Mali RS, Chan G, Kang M, Liu Z, et al. PI3K p110delta uniquely promotes gain-of-function Shp2-induced GM-CSF hypersensitivity in a model of JMML. Blood. 2014;123(18):2838-42.
154. Padron E, Painter JS, Kunigal S, Mailloux AW, McGraw K, McDaniel JM, et al. GM-CSF-dependent pSTAT5 sensitivity is a feature with therapeutic potential in chronic myelomonocytic leukemia. Blood. 2013;121(25):5068-77.
155. Emanuel. Hallway gossip between Ras and PI3K pathways. Blood. 2014:2751-3.
156. Dong L, Yu WM, Zheng H, Loh ML, Bunting ST, Pauly M, et al. Leukaemogenic effects of Ptpn11 activating mutations in the stem cell microenvironment. Nature. 2016;539(7628):304-8.
157. Passmore SJ, Hann IM, Stiller CA, Ramani P, Swansbury GJ, Gibbons B, et al. Pediatric myelodysplasia: a study of 68 children and a new prognostic scoring system. Blood. 1995;85(7):1742-50.
158. Weatherall DJ, Edwards JA, Donohoe WT. Haemoglobin and red cell enzyme changes in juvenile myeloid leukaemia. Br Med J. 1968;1(5593):679-81.
159. Maurer HS, Vida LN, Honig GR. Similarities of the erythrocytes in juvenile chronic myelogenous leukemia to fetal erythrocytes. Blood. 1972;39(6):778-84.
160. Weatherall DJ, Clegg JB, Wood WG, Callender ST, Sheridan BL, Pritchard J. Foetal erythropoiesis in human leukaemia. Nature. 1975;257(5528):710-2.
161. Dover GJ, Boyer SH, Zinkham WH, Kazazian HH, Jr., Pinney DJ, Sigler A. Changing erythrocyte populations in juvenile chronic myelocytic leukemia: evidence for disordered regulation. Blood. 1977;49(3):355-65.

162. Gahr M, Schroter W. The pattern of reactivated fetal erythropoiesis in bone marrow disorders of childhood. Acta Paediatr Scand. 1982;71(6):1013-8.
163. Papayannopoulou T, Nakamoto B, Anagnou NP, Chui D, Dow L, Sanders J. Expression of embryonic globins by erythroid cells in juvenile chronic myelocytic leukemia. Blood. 1991;77(12):2569-76.
164. Alter BP, Rosenberg PS, Day T, Menzel S, Giri N, Savage SA, et al. Genetic regulation of fetal haemoglobin in inherited bone marrow failure syndromes. Br J Haematol. 2013;162(4):542-6.
165. Sankaran VG, Orkin SH. The switch from fetal to adult hemoglobin. Cold Spring Harb Perspect Med. 2013;3(1):a011643.
166. Siatecka M, Bieker JJ. The multifunctional role of EKLF/KLF1 during erythropoiesis. Blood. 2011;118(8):2044-54.
167. Fluhr S, Krombholz CF, Meier A, Epting T, Mucke O, Plass C, et al. Epigenetic dysregulation of the erythropoietic transcription factor KLF1 and the beta-like globin locus in juvenile myelomonocytic leukemia. Epigenetics. 2017;12(8):715-23.
168. Locatelli F, Nollke P, Zecca M, Korthof E, Lanino E, Peters C, et al. Hematopoietic stem cell transplantation (HSCT) in children with juvenile myelomonocytic leukemia (JMML): results of the EWOG-MDS/EBMT trial. Blood. 2005;105(1):410-9.
169. Moritake H, Ikeda T, Manabe A, Kamimura S, Nunoi H. Cytomegalovirus infection mimicking juvenile myelomonocytic leukemia showing hypersensitivity to granulocyte-macrophage colony stimulating factor. Pediatric blood & cancer. 2009;53(7):1324-6.
170. Herrod HG, Dow LW, Sullivan JL. Persistent epstein-barr virus infection mimicking juvenile chronic myelogenous leukemia: immunologic and hematologic studies. Blood. 1983;61(6):1098-104.
171. Lorenzana A, Lyons H, Sawaf H, Higgins M, Carrigan D, Emanuel PD. Human herpesvirus 6 infection mimicking juvenile myelomonocytic leukemia in an infant. Journal of pediatric hematology/oncology. 2002;24(2):136-41.
172. Stieglitz E, Taylor-Weiner AN, Chang TY, Gelston LC, Wang YD, Mazor T, et al. The genomic landscape of juvenile myelomonocytic leukemia. Nat Genet. 2015;47(11):1326-33.
173. Magnusson MK, Meade KE, Brown KE, Arthur DC, Krueger LA, Barrett AJ, et al. Rabaptin-5 is a novel fusion partner to platelet-derived growth factor beta receptor in chronic myelomonocytic leukemia. Blood. 2001;98(8):2518-25.
174. Murakami N, Okuno Y, Yoshida K, Shiraishi Y, Nagae G, Suzuki K, et al. Integrated molecular profiling of juvenile myelomonocytic leukemia. Blood. 2018.
175. Sakashita K, Matsuda K, Koike K. Diagnosis and treatment of juvenile myelomonocytic leukemia. Pediatr Int. 2016;58(8):681-90.
176. Hasegawa D, Bugarin C, Giordan M, Bresolin S, Longoni D, Micalizzi C, et al. Validation of flow cytometric phospho-STAT5 as a diagnostic tool for juvenile myelomonocytic leukemia. Blood cancer journal. 2013;3:e160.
177. Kotecha N, Flores NJ, Irish JM, Simonds EF, Sakai DS, Archambeault S, et al. Single-cell profiling identifies aberrant STAT5 activation in myeloid malignancies with specific clinical and biologic correlates. Cancer Cell. 2008;14(4):335-43.
178. Meyerson H, Awadallah A, Blidaru G, Osei E, Schlegelmilch J, Egler R, et al. Juvenile myelomonocytic leukemia with prominent CD141+ myeloid dendritic cell differentiation. Hum Pathol. 2017;68:147-53.
179. Niemeyer CM, Kang MW, Shin DH, Furlan I, Erlacher M, Bunin NJ, et al. Germline CBL mutations cause developmental abnormalities and predispose to juvenile myelomonocytic leukemia. Nat Genet. 2010;42(9):794-800.

180. Bos JL. ras oncogenes in human cancer: a review. Cancer Res. 1989;49(17):4682-9.
181. Niemeyer CM. RAS diseases in children. Haematologica. 2014;99(11):1653-62.
182. Neel BG, Gu H, Pao L. The 'Shp'ing news: SH2 domain-containing tyrosine phosphatases in cell signaling. Trends Biochem Sci. 2003;28(6):284-93.
183. Tartaglia M, Niemeyer CM, Shannon KM, Loh ML. SHP-2 and myeloid malignancies. Current opinion in hematology. 2004;11(1):44-50.
184. Loh ML, Vattikuti S, Schubbert S, Reynolds MG, Carlson E, Lieuw KH, et al. Mutations in PTPN11 implicate the SHP-2 phosphatase in leukemogenesis. Blood. 2004;103(6):2325-31.
185. Tartaglia M, Martinelli S, Cazzaniga G, Cordeddu V, Iavarone I, Spinelli M, et al. Genetic evidence for lineage-related and differentiation stage-related contribution of somatic PTPN11 mutations to leukemogenesis in childhood acute leukemia. Blood. 2004;104(2):307-13.
186. Tartaglia M, Niemeyer CM, Fragale A, Song X, Buechner J, Jung A, et al. Somatic mutations in PTPN11 in juvenile myelomonocytic leukemia, myelodysplastic syndromes and acute myeloid leukemia. Nat Genet. 2003;34(2):148-50.
187. Tartaglia M. Somatic mutations in PTPN11 in juvenile myelomonocytic leukemia, myelodysplastic syndromes and acute myeloid leukemia. Nat Genet. 2003;34:148-50.
188. Flotho C, Valcamonica S, Mach-Pascual S, Schmahl G, Corral L, Ritterbach J, et al. RAS mutations and clonality analysis in children with juvenile myelomonocytic leukemia (JMML). Leukemia. 1999;13(1):32-7.
189. Kratz CP, Niemeyer CM, Castleberry RP, Cetin M, Bergstrasser E, Emanuel PD, et al. The mutational spectrum of PTPN11 in juvenile myelomonocytic leukemia and Noonan syndrome/myeloproliferative disease. Blood. 2005;106(6):2183-5.
190. Liu W, Yu WM, Zhang J, Chan RJ, Loh ML, Zhang Z, et al. Inhibition of the Gab2/PI3K/mTOR signaling ameliorates myeloid malignancy caused by Ptpn11 (Shp2) gain-of-function mutations. Leukemia. 2017;31(6):1415-22.
191. Kalra R, Paderanga DC, Olson K, Shannon KM. Genetic analysis is consistent with the hypothesis that NF1 limits myeloid cell growth through p21ras. Blood. 1994;84(10):3435-9.
192. Braun BS, Tuveson DA, Kong N, Le DT, Kogan SC, Rozmus J, et al. Somatic activation of oncogenic Kras in hematopoietic cells initiates a rapidly fatal myeloproliferative disorder. Proc Natl Acad Sci U S A. 2004;101(2):597-602.
193. Caye A, Strullu M, Guidez F, Cassinat B, Gazal S, Fenneteau O, et al. Juvenile myelomonocytic leukemia displays mutations in components of the RAS pathway and the PRC2 network. Nat Genet. 2015;47(11):1334-40.
194. Flotho C, Kratz CP, Bergstrasser E, Hasle H, Stary J, Trebo M, et al. Genotype-phenotype correlation in cases of juvenile myelomonocytic leukemia with clonal RAS mutations. Blood. 2008;111(2):966-7; author reply 7-8.
195. Matsuda K, Shimada A, Yoshida N, Ogawa A, Watanabe A, Yajima S, et al. Spontaneous improvement of hematologic abnormalities in patients having juvenile myelomonocytic leukemia with specific RAS mutations. Blood. 2007;109(12):5477-80.
196. Miyauchi J, Asada M, Sasaki M, Tsunematsu Y, Kojima S, Mizutani S. Mutations of the N-ras gene in juvenile chronic myelogenous leukemia. Blood. 1994;83(8):2248-54.
197. Doisaki S, Muramatsu H, Shimada A, Takahashi Y, Mori-Ezaki M, Sato M, et al. Somatic mosaicism for oncogenic NRAS mutations in juvenile myelomonocytic leukemia. Blood. 2012;120(7):1485-8.

198. Matsuda K, Nakazawa Y, Sakashita K, Shiohara M, Yamauchi K, Koike K. Acquisition of loss of the wild-type NRAS locus with aggressive disease progression in a patient with juvenile myelomonocytic leukemia and a heterozygous NRAS mutation. Haematologica. 2007;92(11):1576-8.
199. Patil S, Chamberlain RS. Neoplasms associated with germline and somatic NF1 gene mutations. Oncologist. 2012;17(1):101-16.
200. Stiller CA, Chessells JM, Fitchett M. Neurofibromatosis and childhood leukaemia/lymphoma: a population-based UKCCSG study. Br J Cancer. 1994;70(5):969-72.
201. L S. Mutations of the NF1 gene in children with juvenile myelomonocytic leukemia without clinical evidence of neurofibromatosis, type 1. Blood. 1998;92:267-72.
202. Shannon KM, O'Connell P, Martin GA, Paderanga D, Olson K, Dinndorf P, et al. Loss of the normal NF1 allele from the bone marrow of children with type 1 neurofibromatosis and malignant myeloid disorders. N Engl J Med. 1994;330(9):597-601.
203. Miles DK, Freedman MH, Stephens K, Pallavicini M, Sievers EL, Weaver M, et al. Patterns of hematopoietic lineage involvement in children with neurofibromatosis type 1 and malignant myeloid disorders. Blood. 1996;88(11):4314-20.
204. Steinemann D, Arning L, Praulich I, Stuhrmann M, Hasle H, Stary J, et al. Mitotic recombination and compound-heterozygous mutations are predominant NF1-inactivating mechanisms in children with juvenile myelomonocytic leukemia and neurofibromatosis type 1. Haematologica. 2010;95(2):320-3.
205. Schmidt MH, Dikic I. The Cbl interactome and its functions. Nat Rev Mol Cell Biol. 2005;6(12):907-18.
206. Loh ML, Sakai DS, Flotho C, Kang M, Fliegauf M, Archambeault S, et al. Mutations in CBL occur frequently in juvenile myelomonocytic leukemia. Blood. 2009;114(9):1859-63.
207. Muramatsu H, Makishima H, Jankowska AM, Cazzolli H, O'Keefe C, Yoshida N, et al. Mutations of an E3 ubiquitin ligase c-Cbl but not TET2 mutations are pathogenic in juvenile myelomonocytic leukemia. Blood. 2010;115(10):1969-75.
208. Strullu M, Caye A, Cassinat B, Fenneteau O, Touzot F, Blauwblomme T, et al. In hematopoietic cells with a germline mutation of CBL, loss of heterozygosity is not a signature of juvenile myelo-monocytic leukemia. Leukemia. 2013;27(12):2404-7.
209. Bunda S, Qin K, Kommaraju K, Heir P, Ohh M. Juvenile myelomonocytic leukaemia-associated mutation in Cbl promotes resistance to apoptosis via the Lyn-PI3K/AKT pathway. Oncogene. 2015;34(6):789-97.
210. Ceremsak JJ, Yu A, Esquivel E, Lissewski C, Zenker M, Loh ML, et al. Germline RRAS2 mutations are not associated with Noonan syndrome. J Med Genet. 2016;53(11):728.
211. Rong R, He Q, Liu Y, Sheikh MS, Huang Y. TC21 mediates transformation and cell survival via activation of phosphatidylinositol 3-kinase/Akt and NF-kappaB signaling pathway. Oncogene. 2002;21(7):1062-70.
212. Graham SM, Oldham SM, Martin CB, Drugan JK, Zohn IE, Campbell S, et al. TC21 and Ras share indistinguishable transforming and differentiating activities. Oncogene. 1999;18(12):2107-16.
213. Sakaguchi H, Okuno Y, Muramatsu H, Yoshida K, Shiraishi Y, Takahashi M, et al. Exome sequencing identifies secondary mutations of SETBP1 and JAK3 in juvenile myelomonocytic leukemia. Nat Genet. 2013;45(8):937-41.

214. Stieglitz E, Troup CB, Gelston LC, Haliburton J, Chow ED, Yu KB, et al. Subclonal mutations in SETBP1 confer a poor prognosis in juvenile myelomonocytic leukemia. Blood. 2015;125(3):516-24.
215. Yoshimi A, Balasis ME, Vedder A, Feldman K, Ma Y, Zhang H, et al. Robust patient-derived xenografts of MDS/MPN overlap syndromes capture the unique characteristics of CMML and JMML. Blood. 2017;130(4):397-407.
216. Puda A, Milosevic JD, Berg T, Klampfl T, Harutyunyan AS, Gisslinger B, et al. Frequent deletions of JARID2 in leukemic transformation of chronic myeloid malignancies. Am J Hematol. 2012;87(3):245-50.
217. Jankowska. Mutational spectrum analysis of chronic myelomonocytic leukemia includes genes associated with epigenetic regulation: UTX, EZH2, and DNMT3A. Blood. 2011:3923 -41.
218. Kim E, Ilagan JO, Liang Y, Daubner GM, Lee SC, Ramakrishnan A, et al. SRSF2 Mutations Contribute to Myelodysplasia by Mutant-Specific Effects on Exon Recognition. Cancer Cell. 2015;27(5):617-30.
219. Worzfeld T, Swiercz JM, Senturk A, Genz B, Korostylev A, Deng S, et al. Genetic dissection of plexin signaling in vivo. Proc Natl Acad Sci U S A. 2014;111(6):2194-9.
220. Innocenti M, Frittoli E, Ponzanelli I, Falck JR, Brachmann SM, Di Fiore PP, et al. Phosphoinositide 3-kinase activates Rac by entering in a complex with Eps8, Abi1, and Sos-1. J Cell Biol. 2003;160(1):17-23.
221. Offenhauser N, Borgonovo A, Disanza A, Romano P, Ponzanelli I, Iannolo G, et al. The eps8 family of proteins links growth factor stimulation to actin reorganization generating functional redundancy in the Ras/Rac pathway. Mol Biol Cell. 2004;15(1):91-8.
222. Bresolin S. Mutations of SETBP1 and JAK3 in juvenile myelomonocytic leukemia: a report from the Italian AIEOP study group. Oncotarget. 2016:28914-9.
223. Coppe A, Nogara L, Pizzuto MS, Cani A, Cesaro S, Masetti R, et al. Somatic mutations activating Wiskott-Aldrich syndrome protein concomitant with RAS pathway mutations in juvenile myelomonocytic leukemia patients. Hum Mutat. 2018;39(4):579-87.
224. Dunbar AJ, Gondek LP, O'Keefe CL, Makishima H, Rataul MS, Szpurka H, et al. 250K single nucleotide polymorphism array karyotyping identifies acquired uniparental disomy and homozygous mutations, including novel missense substitutions of c-Cbl, in myeloid malignancies. Cancer Res. 2008;68(24):10349-57.
225. Sanada M, Suzuki T, Shih LY, Otsu M, Kato M, Yamazaki S, et al. Gain-of-function of mutated C-CBL tumour suppressor in myeloid neoplasms. Nature. 2009;460(7257):904-8.
226. Muraoka M, Okuma C, Kanamitsu K, Ishida H, Kanazawa Y, Washio K, et al. Adults with germline CBL mutation complicated with juvenile myelomonocytic leukemia at infancy. J Hum Genet. 2016;61(6):523-6.
227. Matsuda K, Taira C, Sakashita K, Saito S, Tanaka-Yanagisawa M, Yanagisawa R, et al. Long-term survival after nonintensive chemotherapy in some juvenile myelomonocytic leukemia patients with CBL mutations, and the possible presence of healthy persons with the mutations. Blood. 2010;115(26):5429-31.
228. Kato M, Yasui N, Seki M, Kishimoto H, Sato-Otsubo A, Hasegawa D, et al. Aggressive transformation of juvenile myelomonocytic leukemia associated with duplication of oncogenic KRAS due to acquired uniparental disomy. The Journal of pediatrics. 2013;162(6):1285-8, 8 e1.

229. Tokuda K, Eguchi-Ishimae M, Iwabuki H, Kawakami S, Tauchi H, Ishii E, et al. Lineage-dependent skewing of loss of heterozygosity (LOH) of KRAS gene in a case of juvenile myelomonocytic leukemia. Eur J Haematol. 2015;94(2):177-81.
230. Kothari. Aggressive congenital juvenile myelomonocytic leukemia associated with somatic KRAS p.G13D mutation and concurrent germline IGF1R duplication. Leuk Lymphoma 2015;56:1175-8.
231. Lee M. Juvenile myelomonocytic leukemia in a premature neonate mimicking neonatal sepsis. Pediatric Neonatol. 2013.
232. Matsuda K, Sakashita K, Taira C, Tanaka-Yanagisawa M, Yanagisawa R, Shiohara M, et al. Quantitative assessment of PTPN11 or RAS mutations at the neonatal period and during the clinical course in patients with juvenile myelomonocytic leukaemia. Br J Haematol. 2010;148(4):593-9.
233. Yoshida R, Miyata M, Nagai T, Yamazaki T, Ogata T. A 3-bp deletion mutation of PTPN11 in an infant with severe Noonan syndrome including hydrops fetalis and juvenile myelomonocytic leukemia. American journal of medical genetics Part A. 2004;128A(1):63-6.
234. Cheong JL, Moorkamp MH. Respiratory failure, juvenile myelomonocytic leukemia, and neonatal Noonan syndrome. Journal of pediatric hematology/oncology. 2007;29(4):262-4.
235. Bolouri H, Farrar JE, Triche T, Jr., Ries RE, Lim EL, Alonzo TA, et al. The molecular landscape of pediatric acute myeloid leukemia reveals recurrent structural alterations and age-specific mutational interactions. Nat Med. 2018;24(1):103-12.
236. Leoncini PP, Bertaina A, Papaioannou D, Flotho C, Masetti R, Bresolin S, et al. MicroRNA fingerprints in juvenile myelomonocytic leukemia (JMML) identified miR-150-5p as a tumor suppressor and potential target for treatment. Oncotarget. 2016;7(34):55395-408.
237. Mulero-Navarro S, Sevilla A, Roman AC, Lee DF, D'Souza SL, Pardo S, et al. Myeloid Dysregulation in a Human Induced Pluripotent Stem Cell Model of PTPN11-Associated Juvenile Myelomonocytic Leukemia. Cell Rep. 2015;13(3):504-15.
238. Stieglitz E, Mazor T, Olshen AB, Geng H, Gelston LC, Akutagawa J, et al. Genome-wide DNA methylation is predictive of outcome in juvenile myelomonocytic leukemia. Nat Commun. 2017;8(1):2127.
239. Lipka DB, Witte T, Toth R, Yang J, Wiesenfarth M, Nollke P, et al. RAS-pathway mutation patterns define epigenetic subclasses in juvenile myelomonocytic leukemia. Nat Commun. 2017;8(1):2126.
240. Wilhelm T, Lipka DB, Witte T, Wierzbinska JA, Fluhr S, Helf M, et al. Epigenetic silencing of AKAP12 in juvenile myelomonocytic leukemia. Epigenetics. 2016;11(2):110-9.
241. Olk-Batz C, Poetsch AR, Nollke P, Claus R, Zucknick M, Sandrock I, et al. Aberrant DNA methylation characterizes juvenile myelomonocytic leukemia with poor outcome. Blood. 2011;117(18):4871-80.
242. Poetsch AR, Lipka DB, Witte T, Claus R, Nollke P, Zucknick M, et al. RASA4 undergoes DNA hypermethylation in resistant juvenile myelomonocytic leukemia. Epigenetics. 2014;9(9):1252-60.
243. Rauen KA. The RASopathies. Annu Rev Genomics Hum Genet. 2013;14:355-69.
244. Aoki Y, Matsubara Y. Ras/MAPK syndromes and childhood hemato-oncological diseases. Int J Hematol. 2013;97(1):30-6.
245. Tartaglia M, Gelb BD. Noonan syndrome and related disorders: genetics and pathogenesis. Annu Rev Genomics Hum Genet. 2005;6:45-68.

246. Tartaglia M, Kalidas K, Shaw A, Song X, Musat DL, van der Burgt I, et al. PTPN11 mutations in Noonan syndrome: molecular spectrum, genotype-phenotype correlation, and phenotypic heterogeneity. Am J Hum Genet. 2002;70(6):1555-63.
247. Tartaglia M, Mehler EL, Goldberg R, Zampino G, Brunner HG, Kremer H, et al. Mutations in PTPN11, encoding the protein tyrosine phosphatase SHP-2, cause Noonan syndrome. Nat Genet. 2001;29(4):465-8.
248. Silvio F, Carlo L, Elena B, Nicoletta B, Daniela F, Roberto M. Transient abnormal myelopoiesis in Noonan syndrome. Journal of pediatric hematology/oncology. 2002;24(9):763-4.
249. Fukuda M, Horibe K, Miyajima Y, Matsumoto K, Nagashima M. Spontaneous remission of juvenile chronic myelomonocytic leukemia in an infant with Noonan syndrome. Journal of pediatric hematology/oncology. 1997;19(2):177-9.
250. Bader-Meunier B, Tchernia G, Mielot F, Fontaine JL, Thomas C, Lyonnet S, et al. Occurrence of myeloproliferative disorder in patients with Noonan syndrome. The Journal of pediatrics. 1997;130(6):885-9.
251. Strullu M, Caye A, Lachenaud J, Cassinat B, Gazal S, Fenneteau O, et al. Juvenile myelomonocytic leukaemia and Noonan syndrome. J Med Genet. 2014;51(10):689-97.
252. Mason-Suares H, Toledo D, Gekas J, Lafferty KA, Meeks N, Pacheco MC, et al. Juvenile myelomonocytic leukemia-associated variants are associated with neo-natal lethal Noonan syndrome. Eur J Hum Genet. 2017;25(4):509-11.
253. Perez B, Mechinaud F, Galambrun C, Ben Romdhane N, Isidor B, Philip N, et al. Germline mutations of the CBL gene define a new genetic syndrome with predisposition to juvenile myelomonocytic leukaemia. J Med Genet. 2010;47(10):686-91.
254. Naramura M, Jang IK, Kole H, Huang F, Haines D, Gu H. c-Cbl and Cbl-b regulate T cell responsiveness by promoting ligand-induced TCR down-modulation. Nat Immunol. 2002;3(12):1192-9.
255. Dvorak CC, Satwani P, Stieglitz E, Cairo MS, Dang H, Pei Q, et al. Disease burden and conditioning regimens in ASCT1221, a randomized phase II trial in children with juvenile myelomonocytic leukemia: A Children's Oncology Group study. Pediatric blood & cancer. 2018.
256. Locatelli F, Crotta A, Ruggeri A, Eapen M, Wagner JE, Macmillan ML, et al. Analysis of risk factors influencing outcomes after cord blood transplantation in children with juvenile myelomonocytic leukemia: a EUROCORD, EBMT, EWOG-MDS, CIBMTR study. Blood. 2013;122(12):2135-41.
257. Yabe M, Sako M, Yabe H, Osugi Y, Kurosawa H, Nara T, et al. A conditioning regimen of busulfan, fludarabine, and melphalan for allogeneic stem cell transplantation in children with juvenile myelomonocytic leukemia. Pediatr Transplant. 2008;12(8):862-7.
258. Yabe M, Ohtsuka Y, Watanabe K, Inagaki J, Yoshida N, Sakashita K, et al. Transplantation for juvenile myelomonocytic leukemia: a retrospective study of 30 children treated with a regimen of busulfan, fludarabine, and melphalan. Int J Hematol. 2015;101(2):184-90.
259. Manabe A, Okamura J, Yumura-Yagi K, Akiyama Y, Sako M, Uchiyama H, et al. Allogeneic hematopoietic stem cell transplantation for 27 children with juvenile myelomonocytic leukemia diagnosed based on the criteria of the International JMML Working Group. Leukemia. 2002;16(4):645-9.
260. Smith FO, King R, Nelson G, Wagner JE, Robertson KA, Sanders JE, et al. Unrelated donor bone marrow transplantation for children with juvenile myelomonocytic leukaemia. Br J Haematol. 2002;116(3):716-24.

261. Korthof ET, Snijder PP, de Graaff AA, Lankester AC, Bredius RG, Ball LM, et al. Allogeneic bone marrow transplantation for juvenile myelomonocytic leukemia: a single center experience of 23 patients. Bone Marrow Transplant. 2005;35(5):455-61.
262. Yoshimi A, Niemeyer CM, Bohmer V, Duffner U, Strahm B, Kreyenberg H, et al. Chimaerism analyses and subsequent immunological intervention after stem cell transplantation in patients with juvenile myelomonocytic leukaemia. Br J Haematol. 2005;129(4):542-9.
263. Inagaki J, Fukano R, Nishikawa T, Nakashima K, Sawa D, Ito N, et al. Outcomes of immunological interventions for mixed chimerism following allogeneic stem cell transplantation in children with juvenile myelomonocytic leukemia. Pediatric blood & cancer. 2013;60(1):116-20.
264. Emanuel PD, Snyder RC, Wiley T, Gopurala B, Castleberry RP. Inhibition of juvenile myelomonocytic leukemia cell growth in vitro by farnesyltransferase inhibitors. Blood. 2000;95(2):639-45.
265. Widemann BC, Arceci RJ, Jayaprakash N, Fox E, Zannikos P, Goodspeed W, et al. Phase 1 trial and pharmacokinetic study of the farnesyl transferase inhibitor tipifarnib in children and adolescents with refractory leukemias: a report from the Children's Oncology Group. Pediatric blood & cancer. 2011;56(2):226-33.
266. Stieglitz E, Ward AF, Gerbing RB, Alonzo TA, Arceci RJ, Liu YL, et al. Phase II/III trial of a pre-transplant farnesyl transferase inhibitor in juvenile myelomonocytic leukemia: a report from the Children's Oncology Group. Pediatric blood & cancer. 2015;62(4):629-36.
267. Gandre-Babbe S, Paluru P, Aribeana C, Chou ST, Bresolin S, Lu L, et al. Patient-derived induced pluripotent stem cells recapitulate hematopoietic abnormalities of juvenile myelomonocytic leukemia. Blood. 2013;121(24):4925-9.
268. Akutagawa J, Huang TQ, Epstein I, Chang T, Quirindongo-Crespo M, Cottonham CL, et al. Targeting the PI3K/Akt pathway in murine MDS/MPN driven by hyperactive Ras. Leukemia. 2016;30(6):1335-43.
269. Tasian SK, Casas JA, Posocco D, Gandre-Babbe S, Gagne AL, Liang G, et al. Mutation-specific signaling profiles and kinase inhibitor sensitivities of juvenile myelomonocytic leukemia revealed by induced pluripotent stem cells. Leukemia. 2018.
270. Jenkins C, Luty SB, Maxson JE, Eide CA, Abel ML, Togiai C, et al. Synthetic lethality of TNK2 inhibition in PTPN11-mutant leukemia. Sci Signal. 2018;11(539).
271. Arico M, Biondi A, Pui CH. Juvenile myelomonocytic leukemia. Blood. 1997;90(2):479-88.
272. Honda Y, Tsuchida M, Zaike Y, Masunaga A, Yoshimi A, Kojima S, et al. Clinical characteristics of 15 children with juvenile myelomonocytic leukaemia who developed blast crisis: MDS Committee of Japanese Society of Paediatric Haematology/Oncology. Br J Haematol. 2014;165(5):682-7.
273. Castro-Malaspina H, Schaison G, Passe S, Pasquier A, Berger R, Bayle-Weisgerber C, et al. Subacute and chronic myelomonocytic leukemia in children (juvenile CML). Clinical and hematologic observations, and identification of prognostic factors. Cancer. 1984;54(4):675-86.
274. Flotho C, Kratz CP, Bergstrasser E, Hasle H, Stary J, Trebo M, et al. Genotype-phenotype correlation in cases of juvenile myelomonocytic leukemia with clonal RAS mutations. Blood. 2008;111(2):966-7; author reply 7-8.
275. Yoshida N, Yagasaki H, Xu Y, Matsuda K, Yoshimi A, Takahashi Y, et al. Correlation of clinical features with the mutational status of GM-CSF signaling pathway-related genes in juvenile myelomonocytic leukemia. Pediatr Res. 2009;65(3):334-40.

276. Ahn JS, Yang DH, Jung SH, Park HC, Moon JH, Sohn SK, et al. A comparison of bortezomib, cyclophosphamide, and dexamethasone (Vel-CD) chemotherapy without and with thalidomide (Vel-CTD) for the treatment of relapsed or refractory multiple myeloma. Annals of hematology. 2012;91(7):1023-30.
277. Bergstraesser E, Hasle H, Rogge T, Fischer A, Zimmermann M, Noellke P, et al. Non-hematopoietic stem cell transplantation treatment of juvenile myelomonocytic leukemia: a retrospective analysis and definition of response criteria. Pediatric blood & cancer. 2007;49(5):629-33.
278. Chan RJ, Cooper T, Kratz CP, Weiss B, Loh ML. Juvenile myelomonocytic leukemia: a report from the 2nd International JMML Symposium. Leukemia research. 2009;33(3):355-62.
279. Niemeyer CM, Loh ML, Cseh A, Cooper T, Dvorak CC, Chan R, et al. Criteria for evaluating response and outcome in clinical trials for children with juvenile myelomonocytic leukemia. Haematologica. 2015;100(1):17-22.
280. Sugimoto Y, Muramatsu H, Makishima H, Prince C, Jankowska AM, Yoshida N, et al. Spectrum of molecular defects in juvenile myelomonocytic leukaemia includes ASXL1 mutations. Br J Haematol. 2010;150(1):83-7.
281. Forbes LV, Gale RE, Pizzey A, Pouwels K, Nathwani A, Linch DC. An activating mutation in the transmembrane domain of the granulocyte colony-stimulating factor receptor in patients with acute myeloid leukemia. Oncogene. 2002;21(39):5981-9.
282. Haferlach T, Nagata Y, Grossmann V, Okuno Y, Bacher U, Nagae G, et al. Landscape of genetic lesions in 944 patients with myelodysplastic syndromes. Leukemia. 2014;28(2):241-7.
283. Horwitz M, Benson KF, Person RE, Aprikyan AG, Dale DC. Mutations in ELA2, encoding neutrophil elastase, define a 21-day biological clock in cyclic haematopoiesis. Nat Genet. 1999;23(4):433-6.
284. Dale DC, Person RE, Bolyard AA, Aprikyan AG, Bos C, Bonilla MA, et al. Mutations in the gene encoding neutrophil elastase in congenital and cyclic neutropenia. Blood. 2000;96(7):2317-22.
285. Wang J, Ai X, Gale RP, Xu Z, Qin T, Fang L, et al. TET2, ASXL1 and EZH2 mutations in Chinese with myelodysplastic syndromes. Leukemia research. 2013;37(3):305-11.
286. Gratias EJ, Liu YL, Meleth S, Castleberry RP, Emanuel PD. Activating FLT3 mutations are rare in children with juvenile myelomonocytic leukemia. Pediatric blood & cancer. 2005;44(2):142-6.
287. Nichols KE, Crispino JD, Poncz M, White JG, Orkin SH, Maris JM, et al. Familial dyserythropoietic anaemia and thrombocytopenia due to an inherited mutation in GATA1. Nat Genet. 2000;24(3):266-70.
288. Freson K, Matthijs G, Thys C, Marien P, Hoylaerts MF, Vermylen J, et al. Different substitutions at residue D218 of the X-linked transcription factor GATA1 lead to altered clinical severity of macrothrombocytopenia and anemia and are associated with variable skewed X inactivation. Human molecular genetics. 2002;11(2):147-52.
289. Hsu AP, Sampaio EP, Khan J, Calvo KR, Lemieux JE, Patel SY, et al. Mutations in GATA2 are associated with the autosomal dominant and sporadic monocytopenia and mycobacterial infection (MonoMAC) syndrome. Blood. 2011;118(10):2653-5.
290. Mansour S, Connell F, Steward C, Ostergaard P, Brice G, Smithson S, et al. Emberger syndrome-primary lymphedema with myelodysplasia: report of seven new cases. American journal of medical genetics Part A. 2010;152A(9):2287-96.

291. Person RE, Li FQ, Duan Z, Benson KF, Wechsler J, Papadaki HA, et al. Mutations in proto-oncogene GFI1 cause human neutropenia and target ELA2. Nat Genet. 2003;34(3):308-12.
292. Boztug K, Appaswamy G, Ashikov A, Schaffer AA, Salzer U, Diestelhorst J, et al. A syndrome with congenital neutropenia and mutations in G6PC3. N Engl J Med. 2009;360(1):32-43.
293. Banka S, Chervinsky E, Newman WG, Crow YJ, Yeganeh S, Yacobovich J, et al. Further delineation of the phenotype of severe congenital neutropenia type 4 due to mutations in G6PC3. Eur J Hum Genet. 2011;19(1):18-22.
294. Borkhardt A, Bojesen S, Haas OA, Fuchs U, Bartelheimer D, Loncarevic IF, et al. The human GRAF gene is fused to MLL in a unique t(5;11)(q31;q23) and both alleles are disrupted in three cases of myelodysplastic syndrome/acute myeloid leukemia with a deletion 5q. Proc Natl Acad Sci U S A. 2000;97(16):9168-73.
295. Smith BN, Ancliff PJ, Pizzey A, Khwaja A, Linch DC, Gale RE. Homozygous HAX1 mutations in severe congenital neutropenia patients with sporadic disease: a novel mutation in two unrelated British kindreds. Br J Haematol. 2009;144(5):762-70.
296. Aoki Y, Niihori T, Kawame H, Kurosawa K, Ohashi H, Tanaka Y, et al. Germline mutations in HRAS proto-oncogene cause Costello syndrome. Nat Genet. 2005;37(10):1038-40.
297. Tijssen MR, di Summa F, van den Oudenrijn S, Zwaginga JJ, van der Schoot CE, Voermans C, et al. Functional analysis of single amino-acid mutations in the thrombopoietin-receptor Mpl underlying congenital amegakaryocytic thrombocytopenia. Br J Haematol. 2008;141(6):808-13.
298. Aggerholm A, Gronbaek K, Guldberg P, Hokland P. Mutational analysis of the tumour suppressor gene MMAC1/PTEN in malignant myeloid disorders. Eur J Haematol. 2000;65(2):109-13.
299. Song WJ, Sullivan MG, Legare RD, Hutchings S, Tan X, Kufrin D, et al. Haploinsufficiency of CBFA2 causes familial thrombocytopenia with propensity to develop acute myelogenous leukaemia. Nat Genet. 1999;23(2):166-75.
300. Michaud J, Wu F, Osato M, Cottles GM, Yanagida M, Asou N, et al. In vitro analyses of known and novel RUNX1/AML1 mutations in dominant familial platelet disorder with predisposition to acute myelogenous leukemia: implications for mechanisms of pathogenesis. Blood. 2002;99(4):1364-72.
301. Walker LC, Stevens J, Campbell H, Corbett R, Spearing R, Heaton D, et al. A novel inherited mutation of the transcription factor RUNX1 causes thrombocytopenia and may predispose to acute myeloid leukaemia. Br J Haematol. 2002;117(4):878-81.
302. Calado RT, Graf SA, Wilkerson KL, Kajigaya S, Ancliff PJ, Dror Y, et al. Mutations in the SBDS gene in acquired aplastic anemia. Blood. 2007;110(4):1141-6.
303. Hirabayashi S, Flotho C, Moetter J, Heuser M, Hasle H, Gruhn B, et al. Spliceosomal gene aberrations are rare, coexist with oncogenic mutations, and are unlikely to exert a driver effect in childhood MDS and JMML. Blood. 2012;119(11):e96-9.
304. Takita J, Yoshida K, Sanada M, Nishimura R, Okubo J, Motomura A, et al. Novel splicing-factor mutations in juvenile myelomonocytic leukemia. Leukemia. 2012;26(8):1879-81.
305. Saito S, Matsuda K, Taira C, Sano K, Tanaka-Yanagisawa M, Yanagisawa R, et al. Genetic analysis of TP53 in childhood myelodysplastic syndrome and juvenile myelomonocytic leukemia. Leukemia research. 2011;35(12):1578-84.
306. Derry JM, Ochs HD, Francke U. Isolation of a novel gene mutated in Wiskott-Aldrich syndrome. Cell. 1994;79(5):following 922.

307. Roy NB, Wilson EA, Henderson S, Wray K, Babbs C, Okoli S, et al. A novel 33-Gene targeted resequencing panel provides accurate, clinical-grade diagnosis and improves patient management for rare inherited anaemias. Br J Haematol. 2016;175(2):318-30.
308. Notta F, Doulatov S, Laurenti E, Poeppl A, Jurisica I, Dick JE. Isolation of single human hematopoietic stem cells capable of long-term multilineage engraftment. Science. 2011;333(6039):218-21.
309. Majeti R, Park CY, Weissman IL. Identification of a hierarchy of multipotent hematopoietic progenitors in human cord blood. Cell Stem Cell. 2007;1(6):635-45.
310. Attar A. Changes in the Cell Surface Markers During Normal Hematopoiesis: A Guide to Cell Isolation. Global Journal of Hematology and Blood Transfusion. 2014;1:20-8.
311. Roy A, Cowan G, Mead AJ, Filippi S, Bohn G, Chaidos A, et al. Perturbation of fetal liver hematopoietic stem and progenitor cell development by trisomy 21. Proc Natl Acad Sci U S A. 2012;109(43):17579-84.
312. Picelli S, Faridani OR, Bjorklund AK, Winberg G, Sagasser S, Sandberg R. Full-length RNA-seq from single cells using Smart-seq2. Nat Protoc. 2014;9(1):171-81.
313. Giustacchini A, Thongjuea S, Barkas N, Woll PS, Povinelli BJ, Booth CAG, et al. Single-cell transcriptomics uncovers distinct molecular signatures of stem cells in chronic myeloid leukemia. Nat Med. 2017;23(6):692-702.
314. Webb DK, Passmore SJ, Hann IM, Harrison G, Wheatley K, Chessells JM. Results of treatment of children with refractory anaemia with excess blasts (RAEB) and RAEB in transformation (RAEBt) in Great Britain 1990-99. Br J Haematol. 2002;117(1):33-9.
315. Malcovati L, Cazzola M. The shadowlands of MDS: idiopathic cytopenias of undetermined significance (ICUS) and clonal hematopoiesis of indeterminate potential (CHIP). Hematology American Society of Hematology Education Program. 2015;2015:299-307.
316. Valent P, Bain BJ, Bennett JM, Wimazal F, Sperr WR, Mufti G, et al. Idiopathic cytopenia of undetermined significance (ICUS) and idiopathic dysplasia of uncertain significance (IDUS), and their distinction from low risk MDS. Leukemia research. 2012;36(1):1-5.
317. Ward AC, Gits J, Majeed F, Aprikyan AA, Lewis RS, O'Sullivan LA, et al. Functional interaction between mutations in the granulocyte colony-stimulating factor receptor in severe congenital neutropenia. Br J Haematol. 2008;142(4):653-6.
318. Zhang H, Means S, Schultz AR, Watanabe-Smith K, Medeiros BC, Bottomly D, et al. Unpaired Extracellular Cysteine Mutations of CSF3R Mediate Gain or Loss of Function. Cancer Res. 2017;77(16):4258-67.
319. Beel K, Vandenberghe P. G-CSF receptor (CSF3R) mutations in X-linked neutropenia evolving to acute myeloid leukemia or myelodysplasia. Haematologica. 2009;94(10):1449-52.
320. Ai J, Druhan LJ, Loveland MJ, Avalos BR. G-CSFR ubiquitination critically regulates myeloid cell survival and proliferation. PLoS One. 2008;3(10):e3422.
321. Bluteau O, Sebert M, Leblanc T, Peffault de Latour R, Quentin S, Lainey E, et al. A landscape of germ line mutations in a cohort of inherited bone marrow failure patients. Blood. 2018;131(7):717-32.
322. Kondo M, Wagers AJ, Manz MG, Prohaska SS, Scherer DC, Beilhack GF, et al. Biology of hematopoietic stem cells and progenitors: implications for clinical application. Annu Rev Immunol. 2003;21:759-806.
323. Tavian M, Hallais MF, Peault B. Emergence of intraembryonic hematopoietic precursors in the pre-liver human embryo. Development. 1999;126(4):793-803.

324. Mikkola HK, Orkin SH. The journey of developing hematopoietic stem cells. Development. 2006;133(19):3733-44.
325. Charbord P, Newton I, Voillat L, Schaal JP, Herve P. The purification of CD34 cells from human cord blood: comparison of separation techniques and cytokine requirements for optimal growth of clonogenic progenitors. Br J Haematol. 1996;94(3):449-54.
326. Doulatov S, Notta F, Eppert K, Nguyen LT, Ohashi PS, Dick JE. Revised map of the human progenitor hierarchy shows the origin of macrophages and dendritic cells in early lymphoid development. Nat Immunol. 2010;11(7):585-93.
327. Doulatov S, Notta F, Laurenti E, Dick JE. Hematopoiesis: a human perspective. Cell Stem Cell. 2012;10(2):120-36.
328. Becker AJ, Mc CE, Till JE. Cytological demonstration of the clonal nature of spleen colonies derived from transplanted mouse marrow cells. Nature. 1963;197:452-4.
329. Till JE, McCulloch EA, Siminovitch L. A Stochastic Model of Stem Cell Proliferation, Based on the Growth of Spleen Colony-Forming Cells. Proc Natl Acad Sci U S A. 1964;51:29-36.
330. Till JE, McCulloch EA, Siminovitch L. Isolation of Variant Cell Lines during Serial Transplantation of Hematopoietic Cells Derived from Fetal Liver. J Natl Cancer Inst. 1964;33:707-20.
331. Bryder D, Rossi DJ, Weissman IL. Hematopoietic stem cells: the paradigmatic tissue-specific stem cell. Am J Pathol. 2006;169(2):338-46.
332. Y C. The AC133+CD38 , but not the rhodamine-low, phenotype tracks LTC-IC and SRC function in human cord blood ex vivo expansion cultures
. Blood. 2010.
333. Takahashi M, Matsuoka Y, Sumide K, Nakatsuka R, Fujioka T, Kohno H, et al. CD133 is a positive marker for a distinct class of primitive human cord blood-derived CD34-negative hematopoietic stem cells. Leukemia. 2014;28(6):1308-15.
334. Chen L, Kostadima M, Martens JHA, Canu G, Garcia SP, Turro E, et al. Transcriptional diversity during lineage commitment of human blood progenitors. Science. 2014;345(6204):1251033.
335. Karamitros D, Stoilova B, Aboukhalil Z, Hamey F, Reinisch A, Samitsch M, et al. Single-cell analysis reveals the continuum of human lympho-myeloid progenitor cells. Nat Immunol. 2018;19(1):85-97.
336. Notta F, Zandi S, Takayama N, Dobson S, Gan OI, Wilson G, et al. Distinct routes of lineage development reshape the human blood hierarchy across ontogeny. Science. 2016;351(6269):aab2116.
337. Civin CI, Strauss LC, Brovall C, Fackler MJ, Schwartz JF, Shaper JH. Antigenic analysis of hematopoiesis. III. A hematopoietic progenitor cell surface antigen defined by a monoclonal antibody raised against KG-1a cells. J Immunol. 1984;133(1):157-65.
338. Krause DS, Fackler MJ, Civin CI, May WS. CD34: structure, biology, and clinical utility. Blood. 1996;87(1):1-13.
339. Orciani M, Trubiani O, Guarnieri S, Ferrero E, Di Primio R. CD38 is constitutively expressed in the nucleus of human hematopoietic cells. J Cell Biochem. 2008;105(3):905-12.
340. Hao QL, Shah AJ, Thiemann FT, Smogorzewska EM, Crooks GM. A functional comparison of CD34 + CD38- cells in cord blood and bone marrow. Blood. 1995;86(10):3745-53.

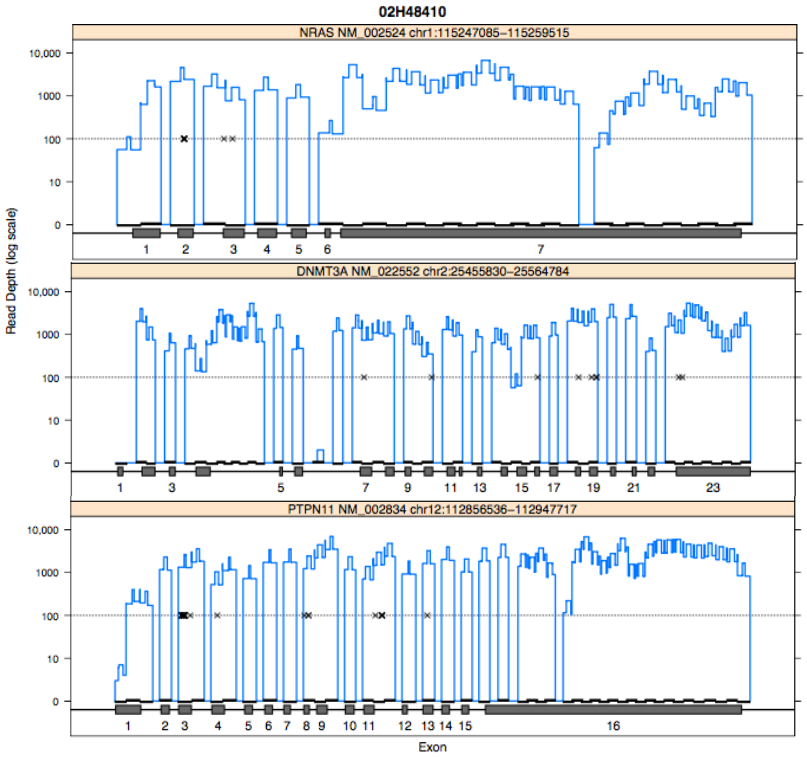
341. Baum CM, Weissman IL, Tsukamoto AS, Buckle AM, Peault B. Isolation of a candidate human hematopoietic stem-cell population. Proc Natl Acad Sci U S A. 1992;89(7):2804-8.
342. Craig W, Kay R, Cutler RL, Lansdorp PM. Expression of Thy-1 on human hematopoietic progenitor cells. J Exp Med. 1993;177(5):1331-42.
343. Reif AE, Allen JM. The Akr Thymic Antigen and Its Distribution in Leukemias and Nervous Tissues. J Exp Med. 1964;120:413-33.
344. Tachibana I, Bodorova J, Berditchevski F, Zutter MM, Hemler ME. NAG-2, a novel transmembrane-4 superfamily (TM4SF) protein that complexes with integrins and other TM4SF proteins. J Biol Chem. 1997;272(46):29181-9.
345. Beverley PC. Functional analysis of human T cell subsets defined by CD45 isoform expression. Semin Immunol. 1992;4(1):35-41.
346. Beverley PC, Daser A, Michie CA, Wallace DL. Functional subsets of T cells defined by isoforms of CD45. Biochem Soc Trans. 1992;20(1):184-7.
347. Christensen JL, Weissman IL. Flk-2 is a marker in hematopoietic stem cell differentiation: a simple method to isolate long-term stem cells. Proc Natl Acad Sci U S A. 2001;98(25):14541-6.
348. Adolfsson J, Mansson R, Buza-Vidas N, Hultquist A, Liuba K, Jensen CT, et al. Identification of Flt3+ lympho-myeloid stem cells lacking erythro-megakaryocytic potential a revised road map for adult blood lineage commitment. Cell. 2005;121(2):295-306.
349. Yoshida T, Ng SY, Zuniga-Pflucker JC, Georgopoulos K. Early hematopoietic lineage restrictions directed by Ikaros. Nat Immunol. 2006;7(4):382-91.
350. Goardon N, Marchi E, Atzberger A, Quek L, Schuh A, Soneji S, et al. Coexistence of LMPP-like and GMP-like leukemia stem cells in acute myeloid leukemia. Cancer Cell. 2011;19(1):138-52.
351. da Silva-Coelho P, Kroeze LI, Yoshida K, Koorenhof-Scheele TN, Knops R, van de Locht LT, et al. Clonal evolution in myelodysplastic syndromes. Nat Commun. 2017;8:15099.
352. Dick JE. Stem cell concepts renew cancer research. Blood. 2008;112(13):4793-807.
353. Bonnet D, Dick JE. Human acute myeloid leukemia is organized as a hierarchy that originates from a primitive hematopoietic cell. Nat Med. 1997;3(7):730-7.
354. Ishikawa F, Yoshida S, Saito Y, Hijikata A, Kitamura H, Tanaka S, et al. Chemotherapy-resistant human AML stem cells home to and engraft within the bone-marrow endosteal region. Nat Biotechnol. 2007;25(11):1315-21.
355. Taussig DC, Miraki-Moud F, Anjos-Afonso F, Pearce DJ, Allen K, Ridler C, et al. Anti-CD38 antibody-mediated clearance of human repopulating cells masks the heterogeneity of leukemia-initiating cells. Blood. 2008;112(3):568-75.
356. Yoshimoto G, Miyamoto T, Jabbarzadeh-Tabrizi S, Iino T, Rocnik JL, Kikushige Y, et al. FLT3-ITD up-regulates MCL-1 to promote survival of stem cells in acute myeloid leukemia via FLT3-ITD-specific STAT5 activation. Blood. 2009;114(24):5034-43.
357. Cozzio A, Passegue E, Ayton PM, Karsunky H, Cleary ML, Weissman IL. Similar MLL-associated leukemias arising from self-renewing stem cells and short-lived myeloid progenitors. Genes Dev. 2003;17(24):3029-35.
358. Huntly BJ, Shigematsu H, Deguchi K, Lee BH, Mizuno S, Duclos N, et al. MOZ-TIF2, but not BCR-ABL, confers properties of leukemic stem cells to committed murine hematopoietic progenitors. Cancer Cell. 2004;6(6):587-96.

359. Krivtsov AV, Twomey D, Feng Z, Stubbs MC, Wang Y, Faber J, et al. Transformation from committed progenitor to leukaemia stem cell initiated by MLL-AF9. Nature. 2006;442(7104):818-22.
360. So CW, Karsunky H, Passegue E, Cozzio A, Weissman IL, Cleary ML. MLL-GAS7 transforms multipotent hematopoietic progenitors and induces mixed lineage leukemias in mice. Cancer Cell. 2003;3(2):161-71.
361. Quek L, Otto GW, Garnett C, Lhermitte L, Karamitros D, Stoilova B, et al. Genetically distinct leukemic stem cells in human CD34- acute myeloid leukemia are arrested at a hemopoietic precursor-like stage. J Exp Med. 2016;213(8):1513-35.
362. Cooper LJ, Shannon KM, Loken MR, Weaver M, Stephens K, Sievers EL. Evidence that juvenile myelomonocytic leukemia can arise from a pluripotential stem cell. Blood. 2000;96(6):2310-3.
363. Altman AJ, Palmer CG, Baehner RL. Juvenile "chronic granulocytic" leukemia: a panmyelopathy with prominent monocytic involvement and circulating monocyte colony-forming cells. Blood. 1974;43(3):341-50.
364. Amenomori T, Tomonaga M, Yoshida Y, Kuriyama K, Matsuo T, Jinnai I, et al. Cytogenetic evidence for partially committed myeloid progenitor cell origin of chronic myelomonocytic leukaemia and juvenile chronic myeloid leukaemia: both granulocyte-macrophage precursors and erythroid precursors carry identical marker chromosome. Br J Haematol. 1986;64(3):539-46.
365. Inoue S, Shibata T, Ravindranath Y, Gohle N. Clonal origin of erythroid cells in juvenile chronic myelogenous leukemia. Blood. 1987;69(3):975-6.
366. Nakamura Y, Ito M, Yamamoto T, Yan XY, Yagasaki H, Kamachi Y, et al. Engraftment of NOD/SCID/gammac(null) mice with multilineage neoplastic cells from patients with juvenile myelomonocytic leukaemia. Br J Haematol. 2005;130(1):51-7.
367. KE M. Distinct Sources of Hematopoietic Progenitors Emerge before HSCs and Provide Functional Blood Cells in the Mammalian Embryo. Cell Rep 2015:1892- 904.
368. E GP. Tissue-resident macrophages originate from yolk-sac-derived erythro-myeloid progenitors. Nature. 2015:547- 51.
369. SP T. Yolk sac erythromyeloid progenitors expressing gain of function PTPN11 have functional features of JMML but are not sufficient to cause disease in mice. Dev Dyn. 2017;246(12):1001 -14.
370. Dimitriou M, Woll PS, Mortera-Blanco T, Karimi M, Wedge DC, Doolittle H, et al. Perturbed hematopoietic stem and progenitor cell hierarchy in myelodysplastic syndromes patients with monosomy 7 as the sole cytogenetic abnormality. Oncotarget. 2016;7(45):72685-98.
371. Psaila B, Barkas N, Iskander D, Roy A, Anderson S, Ashley N, et al. Single-cell profiling of human megakaryocyte-erythroid progenitors identifies distinct megakaryocyte and erythroid differentiation pathways. Genome Biol. 2016;17:83.
372. Sakashita K, Kato I, Daifu T, Saida S, Hiramatsu H, Nishinaka Y, et al. In vitro expansion of CD34(+)CD38(-) cells under stimulation with hematopoietic growth factors on AGM-S3 cells in juvenile myelomonocytic leukemia. Leukemia. 2015;29(3):606-14.
373. McGranahan N, Swanton C. Clonal Heterogeneity and Tumor Evolution: Past, Present, and the Future. Cell. 2017;168(4):613-28.
374. Alizadeh AA, Aranda V, Bardelli A, Blanpain C, Bock C, Borowski C, et al. Toward understanding and exploiting tumor heterogeneity. Nat Med. 2015;21(8):846-53.
375. Wiemels JL, Cazzaniga G, Daniotti M, Eden OB, Addison GM, Masera G, et al. Prenatal origin of acute lymphoblastic leukaemia in children. Lancet. 1999;354(9189):1499-503.

376. Hong D, Gupta R, Ancliff P, Atzberger A, Brown J, Soneji S, et al. Initiating and cancer-propagating cells in TEL-AML1-associated childhood leukemia. Science. 2008;319(5861):336-9.
377. Weissman I. Stem cell research: paths to cancer therapies and regenerative medicine. JAMA. 2005;294(11):1359-66.
378. Grove CS, Vassiliou GS. Acute myeloid leukaemia: a paradigm for the clonal evolution of cancer? Dis Model Mech. 2014;7(8):941-51.
379. Jan M, Snyder TM, Corces-Zimmerman MR, Vyas P, Weissman IL, Quake SR, et al. Clonal evolution of preleukemic hematopoietic stem cells precedes human acute myeloid leukemia. Sci Transl Med. 2012;4(149):149ra18.
380. Shlush LI, Zandi S, Mitchell A, Chen WC, Brandwein JM, Gupta V, et al. Identification of pre-leukaemic haematopoietic stem cells in acute leukaemia. Nature. 2014;506(7488):328-33.
381. Shlush LI, Mitchell A. AML evolution from preleukemia to leukemia and relapse. Best Pract Res Clin Haematol. 2015;28(2-3):81-9.
382. Anderson K, Lutz C, van Delft FW, Bateman CM, Guo Y, Colman SM, et al. Genetic variegation of clonal architecture and propagating cells in leukaemia. Nature. 2011;469(7330):356-61.
383. Notta F, Mullighan CG, Wang JC, Poeppl A, Doulatov S, Phillips LA, et al. Evolution of human BCR-ABL1 lymphoblastic leukaemia-initiating cells. Nature. 2011;469(7330):362-7.
384. Ding L, Ley TJ, Larson DE, Miller CA, Koboldt DC, Welch JS, et al. Clonal evolution in relapsed acute myeloid leukaemia revealed by whole-genome sequencing. Nature. 2012;481(7382):506-10.
385. Kronke J, Bullinger L, Teleanu V, Tschurtz F, Gaidzik VI, Kuhn MW, et al. Clonal evolution in relapsed NPM1-mutated acute myeloid leukemia. Blood. 2013;122(1):100-8.
386. Bachas C, Schuurhuis GJ, Assaraf YG, Kwidama ZJ, Kelder A, Wouters F, et al. The role of minor subpopulations within the leukemic blast compartment of AML patients at initial diagnosis in the development of relapse. Leukemia. 2012;26(6):1313-20.
387. Kern W, Haferlach T, Schnittger S, Ludwig WD, Hiddemann W, Schoch C. Karyotype instability between diagnosis and relapse in 117 patients with acute myeloid leukemia: implications for resistance against therapy. Leukemia. 2002;16(10):2084-91.
388. Ottone T, Zaza S, Divona M, Hasan SK, Lavorgna S, Laterza S, et al. Identification of emerging FLT3 ITD-positive clones during clinical remission and kinetics of disease relapse in acute myeloid leukaemia with mutated nucleophosmin. Br J Haematol. 2013;161(4):533-40.
389. Mullighan CG, Phillips LA, Su X, Ma J, Miller CB, Shurtleff SA, et al. Genomic analysis of the clonal origins of relapsed acute lymphoblastic leukemia. Science. 2008;322(5906):1377-80.
390. Quek L, David MD, Kennedy A, Metzner M, Amatangelo M, Shih A, et al. Clonal heterogeneity of acute myeloid leukemia treated with the IDH2 inhibitor enasidenib. Nat Med. 2018.
391. Cancer Genome Atlas Research N, Ley TJ, Miller C, Ding L, Raphael BJ, Mungall AJ, et al. Genomic and epigenomic landscapes of adult de novo acute myeloid leukemia. N Engl J Med. 2013;368(22):2059-74.
392. K M. Quantitative assessment of PTPN11 or RAS mutations at the neonatal period and during the clinical course in patients with juvenile myelomonocytic leukaemia. Br J Haematol. 2010;148:593 -9.

393. Papaemmanuil E, Gerstung M, Bullinger L, Gaidzik VI, Paschka P, Roberts ND, et al. Genomic Classification and Prognosis in Acute Myeloid Leukemia. N Engl J Med. 2016;374(23):2209-21.
394. Tinschert S. Tinschert, S. et al. Segmental neuro bromatosis is caused by somatic mutation of the neuro bromatosis type 1 (NF1) gene. Eur. J. Hum. Genet. 8, 455–459 (2000). Eur J Hum Gent 2000:455-9.
395. Blais J, Lavoie SB, Giroux S, Bussieres J, Lindsay C, Dionne J, et al. Risk of Misdiagnosis Due to Allele Dropout and False-Positive PCR Artifacts in Molecular Diagnostics: Analysis of 30,769 Genotypes. J Mol Diagn. 2015;17(5):505-14.
396. Horiuchi D, Camarda R, Zhou AY, Yau C, Momcilovic O, Balakrishnan S, et al. PIM1 kinase inhibition as a targeted therapy against triple-negative breast tumors with elevated MYC expression. Nat Med. 2016;22(11):1321-9.
397. Cheung CHY, Hsu CL, Chen KP, Chong ST, Wu CH, Huang HC, et al. MCM2-regulated functional networks in lung cancer by multi-dimensional proteomic approach. Sci Rep. 2017;7(1):13302.
398. Wood TE, Dalili S, Simpson CD, Hurren R, Mao X, Saiz FS, et al. A novel inhibitor of glucose uptake sensitizes cells to FAS-induced cell death. Mol Cancer Ther. 2008;7(11):3546-55.
399. Booth CAG, Barkas N, Neo WH, Boukarabila H, Soilleux EJ, Giotopoulos G, et al. Ezh2 and Runx1 Mutations Collaborate to Initiate Lympho-Myeloid Leukemia in Early Thymic Progenitors. Cancer Cell. 2018;33(2):274-91 e8.
400. De Raedt T, Beert E, Pasmant E, Luscan A, Brems H, Ortonne N, et al. PRC2 loss amplifies Ras-driven transcription and confers sensitivity to BRD4-based therapies. Nature. 2014;514(7521):247-51.
401. Hosen N, Park CY, Tatsumi N, Oji Y, Sugiyama H, Gramatzki M, et al. CD96 is a leukemic stem cell-specific marker in human acute myeloid leukemia. Proc Natl Acad Sci U S A. 2007;104(26):11008-13.
402. Wang PL, O'Farrell S, Clayberger C, Krensky AM. Identification and molecular cloning of tactile. A novel human T cell activation antigen that is a member of the Ig gene superfamily. J Immunol. 1992;148(8):2600-8.
403. Dougall WC, Kurtulus S, Smyth MJ, Anderson AC. TIGIT and CD96: new checkpoint receptor targets for cancer immunotherapy. Immunol Rev. 2017;276(1):112-20.
404. Georgiev H, Ravens I, Papadogianni G, Bernhardt G. Coming of Age: CD96 Emerges as Modulator of Immune Responses. Front Immunol. 2018;9:1072.
405. Staudinger M, Humpe A, Gramatzki M. Strategies for purging CD96(+) stem cells in vitro and in vivo: New avenues for autologous stem cell transplantation in acute myeloid leukemia. Oncoimmunology. 2013;2(6):e24500.
406. Stein EM, DiNardo CD, Pollyea DA, Fathi AT, Roboz GJ, Altman JK, et al. Enasidenib in mutant IDH2 relapsed or refractory acute myeloid leukemia. Blood. 2017;130(6):722-31.
407. Woll S, Schlitter AM, Dhaene K, Roller M, Esposito I, Sahin U, et al. Claudin 18.2 is a target for IMAB362 antibody in pancreatic neoplasms. Int J Cancer. 2014;134(3):731-9.

Appendix 1: Example of CoveMi software report



X represent locations of previously described mutations

Appendix 2: Detailed protocol of next generation sequencing (NGS) targeted panel library preparation

TSCA Library preparation

Prior to starting

- Prepare paperwork – plate plan incl. sample ID, panel to be tested and indices to be used for sample barcoding/multiplexing.

Sample QC

- A260/A280 ratio is tested on the nanodrop ≥1.7.
- Concentration is calculated using the Qubit dsDNA BR kit.
- Normalise for a 250ng input - 10µl @25ng/µl (lower parameter = 150ng - 10µl @15ng/µl). Normalise into a labelled 96-well plate. Use internal control provided with kit and H2O for –ive. Store in the fridge/freezer until required.

SAFE STOPPING POINT – Keep plates in fridge overnight or freezer for longer periods.

Hybridisation of Oligos – ~15 minutes hands-on, Incubation 80 minutes

- Prior to starting switch on Hybex (95°C), heat sealer. Thaw OHS2, ACD1, ACP1 (RT 30 minutes before required) and the panel (~15 minutes before required).

1. Thaw, mix and spin samples (1min at 1000g).
2. Label a new 96-well HYP plate. Add 5µl of the appropriate panels to the corresponding wells. Add 5µl of ACP1 (control panel) to the internal control and the No Template Control wells.
3. Add 5µl ACD1 to the internal control well. Add 5µl H2O to the internal control well.
4. Add 10µl of H2O to No Template Control well.
5. Transfer 10µl of DNA to corresponding wells on the HYP plate. Pipette mix (x10) before adding to the plate.
6. Check OHS2 reagent – make sure precipitates/crystals have dissolved. Vortex if necessary.
7. Add 35µl of OHS2 to each sample, pipette mix.
8. Heat seal plate with aluminium foil lid. Then spin 1min at 1000g.
9. Incubate in Hybex at 95°C for one minute then re-set the temperature to 40°C and leave to cool. This takes 1 hour 20 minutes but return after 1 hour to prepare column plate. Before leaving SW1 from the fridge and warm at RT (for at least 30 minutes).

Removal of unbound oligos – 25 minutes

1. Remove UB1 and ELM4 and warm to RT (30 minutes before use). Switch incubation oven on and set at 37°C.
2. Assemble filter plate and label plate.
3. Check SW1 to make sure precipitates/crystals have dissolved.
4. Add 45µl of SW1 to sample, cover the filter plate with the lid and centrifuge at 2100g for 10 minutes.
5. Empty the waste from the MIDI plate onto paper towels. Pre-wash is now complete.
6. Once the Hybex has reached 40°C for at least 1 minute remove plate and spin 1000g for 1 minute.
7. Transfer hybridisation product to the filter plate (pierce foil lid with tips).
8. Replace lid, spin 2100-2400g for 5 minutes (up to 10 minutes), until liquid has passed through plate.
9. Empty the waste onto paper towels.
10. Add 45µl SW1 to each sample. Replace lid and spin 2100-2400g for 5 minutes. Empty waste onto paper towels. Repeat this for a total of 2 SW1 washes (second wash spin lasting 10 minutes).

11. Drain the filter by spinning at 2100-2400g for 5-10 minutes until liquid has completely passed through. Empty waste from MIDI plate onto paper towels.

12. Dispense 45µl UB1 buffer into sample wells. Replace lid and spin 2100-2400g for 5-10 minutes. Check the filter is not wet and discard waste from MIDI plate onto paper towels (make sure all waste has been removed). Proceed immediately to Extension/Ligation.

Extension/Ligation – Hands-on 5 minutes, Incubation 45 minutes

- Oven should have reached 37°C and ELM4 should have completely thawed.

1. Add 45µl ELM4 to each sample.

2. Seal the plate with a non-heat sealing aluminium foil and cover with the lid. Incubate at 37°C for 45 minutes. Thaw PMM2 in the fridge for next step.

Removal of unbound TSCA oligos – Hands-on 30 minutes, incubation ~30 minutes, PCR 85-105 minutes.

1. Remove i5 and i7 indices from the freezer.

2. Prepare fresh 50mM NaOH solution (allow at least 30µl per sample including controls). For 1ml: 995µl of H2O and 5µl of 10M NaOH. Vortex.

3. In a newly labelled 96-well plate transfer 4µl of the appropriate i5 index to the corresponding wells. Once finished replace i5 white caps with new clean ones.

4. Transfer 4µl of the appropriate i7 index into the corresponding wells. Pipette mix x5 to mix. Replace orange caps with new ones. Cover the plate with a foil seal and store in the fridge until required.

5. Mastermix: PMM2 - 29.2µl per reaction (includes overage), TDP1 – 0.58µl per reaction (includes overage). Do not add TDP1 until libraries have been re-suspended as TDP1/PMM2 is only stable for 10 minutes.

6. Remove plate from incubator, remove foil seal and replace with filter plate lid. Spin 2100-2400g for 5-10 minutes.

7. Empty MIDI plate waste onto paper towels.

8. Add 25µl of 50mM NaOH to each sample – ensure pipette tips come in contact with the membrane. Incubate at RT for 5 minutes.

9. Remove index plate and PMM2 aliquot from fridge and remove TDP1 from freezer. Mix TDP1 and transfer calculated amount to PMM2 aliquot (if more than 32 samples in a run make one batch TDP1/PMM2 at a time then aliquot before making another etc).

10. Add 22µl of TDP1/PMM2 mastermix to each sample well of index plate, pipette mix x5.

11. After 5 minutes incubation transfer 20µl of sample eluted from filter plate to corresponding wells of index plate (tilt filter plate as removing). Pipette mix x10. Repeat for all samples/controls.

12. Heat seal index plate with aluminium foil lid. Spin plate 1000g for 1 minute. Place on thermocycler – calculate no. of cycles based on amplicon size.

    Heated lid: 100°C
    Initial activation: 95°C for 3 minutes
    Amplification cycling program: 95°C for 30 seconds
    For x cycles 66°C for 30 seconds
    72°C for 60 seconds
    Final extension: 72°C for 5 minutes
    Pause: 10°C ∞

    SAFE STOPPING POINT: Either store plate on thermocycler overnight or freezer 5 for up to one month.

    Library QC

    Before starting gel remove AMPure XP beads from fridge to equilibrate to RT for 30 minutes.
    Also remove MiSeq cartridge to thaw in water if sequencing on the same day. Set Hybex to 96°C.

- Prepare 3% agarose gel, 6g agarose in 200ml of 0.75xTBE (contains EthBR). Can prepare in advance and keep in buffer at RT or fridge 11. Remove plate from storage, vortex and spin 1000g for 1 minute. In a new 96-well plate add 5µlloading gel in corresponding wells, transfer 5µl PCR product. Pipette mix. Load 12µl into gel, on each row add 5µl of 50bp DNA ladder. Run at 130-150V for 30-45 minutes. Image.

Library clean-up (perform whilst gel is running) – Hands-on: 20 minutes, Processing: 30 minutes

1. Mix XP beads to ensure homogenous solution. Prepare fresh 80% EtOH (allow at least 500µl per sample).

2. In a new MIDI plate add the appropriate amount of XP beads to the sample wells to be used (dependant on size). 250bp amplicon = 45µl of XP.

3. Transfer entire remaining volume of PCR product to the MIDI plate with beads. Seal with a capmat and shake 1800rpm for at least 2 minutes.

4. Incubate at RT without shaking for 10 minutes.

5. Place plate on magnet for at least 2 minutes or until supernatant has cleared.

6. Whilst on the magnet remove the capmat and remove and discard the supernatant.

7. Whilst on the stand add 200µl 80% EtOH to samples, incubate for 30 seconds. Remove supernatant and discard. Repeat this step for a total of 3 EtOH washes.

8. Remove any remaining supernatant with 20µl pipette tips. Air dry beads for 3-5 minutes (up to a maximum of 10 minutes). Take care not to over-dry the beads as this can affect the yield – make sure there are no cracks!

9. Remove plate from magnetic stand; add 30µl of EBT to each sample. Cover with capmat, shake at 1800rpm for 2 minutes. Incubate at RT for 2 minutes without shaking.

10. Place the plate on the magnetic stand for 2 minutes or until the supernatant is clear. Transfer eluate (20µl – do not recover all – take care not to take any beads) to a new clean 96-well plate.

Normalise from qubit readings (dsDNA BR kit)

- 1:200 dilution working solution. Standards (x2): 190µl working solution to 10µl standard. Sample: 197µl working solution to 3µl sample.

- Convert concentration data to nM (for 250bp amplicons mol. Weight = 231,000mol/g.)

$$\text{DNA concentration (ng/µl)/molecular weight (mol/g)} \times 10^6$$

- Dilute and normalise libraries to 4nM, to a final volume of 10µl.

$$\text{Volume library} = (\text{final concentration} \times \text{final volume})/\text{concentration library}$$

- Normalise samples into a new 96-well plate.

SAFE STOPPING POINT: -20 FREEZER UNTIL REQUIRED

- Pool

SAFE STOPPING POINT: -2O FREEZER UNTIL REQUIRED.

Sequencing on the MiSeq

- Pre-wash MiSeq – takes 20 minutes.

- Once cartridge is thawed, invert 10 times to mix – tap to remove air bubbles.

- Mix 5µl library and 5µl NaOH and denature in hybex at 96°C for 5 minutes to give ssDNA for sequencing. Place straight into ice after incubation.

- Calculate final library concentration and bring up to 20pM with HT1 buffer.

- Make plate plan then create sample sheet. Load library into cartridge, insert on MiSeq and run.

Appendix 3: Example of the NIHR study targeted NGS research report





OXFORD/GOSH PAEDIATRIC MYELODYSPLASTIC SYNDROMES RESEARCH STUDY

Paediatric MDS Targeted NGS Panel Research Results

PLEASE NOTE THIS TEST IS STILL UNDERGOING VALIDATION AND IS NOT YET ACCREDITED FOR DIAGNOSTIC USE. IT IS ESSENTIAL THAT RESULTS ARE CONFIRMED USING AN ALTERNATIVE UKAS ISO ACCREDITED METHOD IF THEY ARE TO BE USED FOR CLINICAL MANAGEMENT.

Patient Study ID: OX16020716 Referring clinician: Georgina Hall

DOB: 01/03/2016 Sample date: 14/07/2016

Sex: Female Report date: 12/08/2016

Clinical summary: Juvenile myelomonocytic leukaemia; new diagnosis

Sample type: Peripheral blood

RESULTS

SUMMARY
Variants detected of predicted clinical significance:

KRAS c.38G>A p.Gly13Asp

INTERPRETATION
Variants detected of predicted clinical significance:

KRAS c.38G>A resulting in p.Gly13Asp amino acid change. This variant has previously been reported as somatically acquired in juvenile myelomonocytic leukaemia (Matsuda et al. Spontaneous improvement of hematologic abnormalities in patients having juvenile myelomonocytic leukemia with specific RAS mutations. Blood. 2007; 109:5477).

Technical information

1. DNA was extracted from this peripheral blood sample in the Oxford Molecular Diagnostic Centre (sample reference number G181107L).
2. Sequencing was performed within the Oxford Molecular Diagnostic Centre.

Appendix 4: List of monoclonal antibodies used for this project

Antibodies	Company	Ordering information
Mouse CD45 PE-Texas Red	Invitrogen	Clone: 30-F11 Cat#: MCD4517 RRID: AB_10392557
Human CD45 Alexa Flour 700	BioLegend	Clone: HI130 Cat#: 304024 RRID: AB_493761
Human CD45 Alexa Flour 700	BD	Clone: 2D1 Cat#: 56-9459-42
CD19 PE	BD	Clone: SJ25C1 Cat#: 345789
CD19 APC	BioLegend	Clone: HIB19 Cat#: 302212
CD15 FITC	BD	Clone: MMA Cat#: 347423
CD33 FITC	BD	Clone: HIM3-4 Cat#: 555626
CD66b FITC	BioLegend	Clone: G10F5 Cat#: 305104 RRID: AB_314496
CD11b FITC	eBioscience	Clone: ICRF44 Cat#: 11-0118-42 RRID: AB_1582242
CD14 FITC	eBioscience	Clone: 61D3 Cat#: 11-0149-41 RRID: AB_10597445
CD34 APC-eFlour 780	eBioscience	Clone: 4H11 Cat#: 47-0349-42 RRID: AB_2573956
CD34 PerCP Cy5.5	BioLegend	Clone: 581 Cat#: 343522
CD235a PerCP Cy5.5	BioLegend	Clone: HIR2 Cat#: 306614 RRID: AB_10683170
CD38 PE-Texas Red	invitrogen	Clone: HIT2 Cat#: MHCD3817 RRID: AB_10392545
CD90 BV421	BioLegend	Clone: 5 E 10 Cat#: 328122 RRID: AB_2561420
CD45RA PE	eBioscience	Clone: HI100 Cat#: 12-0458-41 RRID: AB_10717397
CD123 PE Cy7	BioLegend	Clone: 6H6 Cat#: 306010 RRID: AB_493576
Mouse CD45 BV510	BioLegend	Clone: 30-F11 Cat#: 103138 RRID: AB_2563061
CD8 FITC	BioLegend	Clone: RPA-T8 Cat#: 301006 RRID: AB_314124
CD20 FITC	BioLegend	Clone: 2H7 Cat#: 302304 RRID: AB_314252

CD2 FITC	BD	Clone: RPA-2.10 Cat#: 555326
CD3 FITC	BD	Clone: SK7 Cat#: 345763
CD16 FITC	eBioscience	Clone: eBioCB16 Cat#: 11-0168-41 RRID: AB_10804882
CD19 FITC	eBioscience	Clone: HIB19 Cat#: 11-0199-42 RRID: AB_10669461
CD235a FITC	eBioscience	Clone: HIR2 Cat#: 11-9987-82 RRID: AB_465477
CD10 FITC	BioLegend	Clone: HI10a Cat#: 312208 RRID: AB_314919
CD10 PE Cy7	eBioscience	Clone: ebioCB-CALLA Cat#: 25-0106-42
CD127 FITC	eBioscience	Clone: RDR5 Cat#: 11-1278-42 RRID: AB_1907343
CD73 BV421	BioLegend	Clone: AD2 Cat#: 344008
CD96	Laboratory of Martin Gramatzki	Clone: TH-111
Donkey Anti-mouse PE	eBioscience	Cat# 12-4010-82 RRID: AB_11063706
CD41a APC	eBioscience	Clone: HIP8 Cat#: 17-0419-42 RRID: AB_2573144
CD42b PE	eBioscience	Clone: HIP1 Cat# 12-0429042 RRID: AB_10852864
CD56 FITC	eBioscience	Clone: TULY56 Cat#: 11-0566-42 RRID: AB_2572458
CD56 PE	eBioscience	Clone: CMSSB Cat#: 12-0567-41 RRID: AB_10598372
CD45RA BV650	BioLegend	Clone: HI100 Cat# 304136 RRID: AB_2563653

Appendix 5: Primer sequences used for this project

Primer name	Chromosomal location	Gene	Primer forward sequence	Primer reverse sequence
NF1_exon22_PID1_1	chr 17:29556971	NF1	ACACTGACGACATGGTTCTACAggcattgatggcaaatcatta	TACGGTAGCAGAGACTTGGTCTtttcaatgctagcttgcccta
NF1_exon22_PID1_2	chr 17:29556971	NF1	ACACTGACGACATGGTTCTACACCAATACTCAATTTGTAGAACAAACC	TACGGTAGCAGAGACTTGGTCTTTCTGCTACATTTCAGTAGAATGCTT
NF1_exon22_PID1_3	chr 17:29556971	NF1	ACACTGACGACATGGTTCTACAAACCATAGCTATAATGAAGAACTTGCT	TACGGTAGCAGAGACTTGGTCTTTTCTGCTACATTTCAGTAGAATGC
NF1_exon22_PID1_4	chr 17:29556971	NF1	ACACTGACGACATGGTTCTACATGAAGGCAGCTCTGAACATC	TACGGTAGCAGAGACTTGGTCTCTACCAGTATCAGTGTGTAAGAGGTTT
ASXL1_exon12_PID1	chr20: 31022402	ASXL1	ACACTGACGACATGGTTCTACAggaccctcgcagacattaaa	TACGGTAGCAGAGACTTGGTCTctctgccacctccctcatc
rs216123	chr5:149460553	CSF1R	ACACTGACGACATGGTTCTACAtcccagtgatagagcccagt	TACGGTAGCAGAGACTTGGTCTaaggtagcgttgttggtgct
rs4841	chr5:149826526	RPS14	ACACTGACGACATGGTTCTACAtcatcctccaccatgactga	TACGGTAGCAGAGACTTGGTCTgcagccaacatagcagcata
rs2228458	chr5:150518988	ANXA6	ACACTGACGACATGGTTCTACAcagggtgggtaggatgatga	TACGGTAGCAGAGACTTGGTCTtggttaccttgccgtagagg
rs772774529	chr5:150639456	GM2A	ACACTGACGACATGGTTCTACAccatcccagctcagtagctt	TACGGTAGCAGAGACTTGGTCTggctcaccttcagaggagaa
rs113313832	chr5:151043750	SPARC	ACACTGACGACATGGTTCTACAgtacatcgccctggatgagt	TACGGTAGCAGAGACTTGGTCTgaccaagagagcggagagtg
rs2302427	chr7:148525904	EZH2	ACACTGACGACATGGTTCTACACGGTGATCCTCCAGATCTTT	TACGGTAGCAGAGACTTGGTCTTTTGTGGAGTTGGTGAATGC
rs 41277437	chr7:148511171	EZH2	ACACTGACGACATGGTTCTACAAGCGGCTCCACAAGTAAGAC	TACGGTAGCAGAGACTTGGTCTGAAGGTCAAAACCGCTTTCC
rs143671377	chr17:29553512	NF1	ACACTGACGACATGGTTCTACACAGGATAGTGCAGCAGGATG	TACGGTAGCAGAGACTTGGTCTATGGCAACCAGAACAGCTTC
NRASMUT_LONG	chr1:115256529	NRAS	CCTTCGCCTGTCCTCATGTA	TGGTGGGTTTTCTTTTATGTAGGGT
NRASMUT_MISEQ	chr1:115256529	NRAS	ACACTGACGACATGGTTCTACACGCCTGTCCTCATGTATTGGT	TACGGTAGCAGAGACTTGGTCTAGCATTGCATTCCCTGTGGT
SETBP1_MUT_LONG	chr18: 42531917	SETBP1	ATACACAGTGGAACCTGGAAGC	GTTCTTTGTGCTGGTGTCGGA
SETBP1_MUT_MISEQ	chr18: 42531917	SETBP1	ACACTGACGACATGGTTCTACAAAAGGAAATCACGCTGTCCCC	TACGGTAGCAGAGACTTGGTCTGTTGTCCAGGGAGCAGAAATCA
RS2072407_LONG	chr7:148508833	EZH2	TTATCAGTGCCTTACCTCTCCAC	GAAGCTGTGTGCCCAATTACTG
RS2072407_MiSeq	chr7:148508833	EZH2	ACACTGACGACATGGTTCTACATCAGATGGTGCCAGCAATAGA	TACGGTAGCAGAGACTTGGTCTGAGGATTTACAGTGATAGCTTTTGT
tp53_rs1042522_LONG	chr17:7579472	TP53	AAGTCACAGACTTGGCTGTCC	CCCGGACGATATTGAACAATGG
RS1042522	chr17:7579472	TP53	ACACTGACGACATGGTTCTACACTGGGAAGGGACAGAAGATGAC	TACGGTAGCAGAGACTTGGTCTCAATGGTTCACTGAAGACCCAG
rs1642785	chr17:7579801	TP53	AAGACAAGAGCAGAAAGTCAGTCC	CCAGGGTTGGAAGTGTCTCAT
rs1642785	chr17:7579801	TP53	ACACTGACGACATGGTTCTACACCAGAACGTTGTTTTCAGGAAGT	TACGGTAGCAGAGACTTGGTCTGCAGTCAGATCCTAGCGTCG
ID5_PTPN11_LONG	chr12:112926888	PTPN11	ttctgggatcattggtattggtc	tatccaagaggcctagcaagag
ID5_PTPN11_MiSeq	chr12:112926888	PTPN11	ACACTGACGACATGGTTCTACAGACTGCGATATTGACGTTCCCA	TACGGTAGCAGAGACTTGGTCTgagaatccgcatgccagcc
rs185200_LONG	chr5:142254679	ARHGAP26	ACCTGCATTTCTGGGTCATCC	AGAGCTCAAAACATTCAAATGGCA
rs185200_Miseq	chr5:142254679	ARHGAP26	ACACTGACGACATGGTTCTACACAAGGCCCATGACACTCCTT	TACGGTAGCAGAGACTTGGTCTCCTTGCTCCCTCAACACGAA
rs2276599_LONG	chr2:25469913	DNMT3A	TCTGAGTAGTGAGGGTGGCA	TACGGTAGCAGAGACTTGGTCTGTAAGCCTCGGCAAACAAGG
rs2276599_Miseq	chr2:25469913	DNMT3A	ACACTGACGACATGGTTCTACAGGTGGAGAGAAAGGAGTCGC	TACGGTAGCAGAGACTTGGTCTGTAAGCCTCGGCAAACAAGG
rs17319679_LONG	chr4: 106164980	TET2	AAATACGGCACGCTCACCA	AACTGAGGCCATGTGGTTACA

rs17319679_LONG	chr4: 106164980	TET2	ACACTGACGACATGGTTCTACA CCTCTTTGCGGCCACTGAT	TACGGTAGCAGAGACTTGGTCTACTG CCAAACAGGCATTTTGAT
rs56185013_LONG	chr4: 106160133	TET2	GCCCCATCTTCAGGCTCTAC	TAGTCACCCAAGAGCTCCGA
rs56185013_MiSeq	chr4: 106160133	TET2	ACACTGACGACATGGTTCTACA TTTGCCCGGATCCATCAGAAG	TACGGTAGCAGAGACTTGGTCTTCC GATGGAGATGCACAAGAT
TET2 intronic_rs2454206_LO NG	chr4: 106196951	TET2	GGAGCCACCTCTAGATTACCAC	TTGTGTAAGCCTCCTTGGACAC
TET2 intronic_rs2454206_MIS EQ	chr4: 106196951	TET2	ACACTGACGACATGGTTCTACA TACCACCCAATCTGAGCAATCC	TACGGTAGCAGAGACTTGGTCTTATC ATGGTTAAGAGCTGGAAGCA
tp53_rs1042522_LONG	chr17: 7579472	TP53	AAGTCACAGACTTGGCTGTCC	CCCGGACGATATTGAACAATGG
tp53_rs1042522_MISEQ	chr17: 7579472	TP53	ACACTGACGACATGGTTCTACA CTGGGAAGGGACAGAAGATGA C	TACGGTAGCAGAGACTTGGTCTCAAT GGTTCACTGAAGACCCAG
setbp1_rs3744825_PBM T_LONG	chr18:42532606	SETBP1	TGGAATGCCTTACACATCAATG C	TCCTCTTATGGAGCCGACCA
setbp1_rs3744825_PBM T_MISEQ	chr18:42532606	SETBP1	ACACTGACGACATGGTTCTACA AGCTCCTTTGTACCTATCGCAC	TACGGTAGCAGAGACTTGGTCTGCC CATGCTAACAGGTCCC
rs258819_LONG	chr5:142593652	ARHGAP 26	TTTTCACACAGCACACCGTTC	TTTTCCCTCGGGAGGAGCTTTG
rs258819_MiSeq	chr5:142593652	ARHGAP 26	ACACTGACGACATGGTTCTACA TTTTCACACAGCACACCGTTC	TACGGTAGCAGAGACTTGGTCTGAG GAGCTTTGAAAATGCCAGTTA