MOLECULAR AND CELLULAR PATHOGENESIS OF THE ERYTHROID DEFECT AND THE THERAPEUTIC EFFECT OF GLUCOCORTICOIDS IN DIAMOND-BLACKFAN ANAEMIA

by

Deena Mohsen Iskander

Centre for Haematology. Department of Medicine, Imperial College London

A thesis submitted for the degree of Doctor of Philosophy

2017

To my family

'We shall not cease from exploration, and the end of all our exploring will be to arrive where we started and know the place for the first time.'

T. S. Eliot

ABSTRACT

Diamond-Blackfan Anaemia (DBA) is a rare inherited bone marrow (BM) failure syndrome characterised by selective anaemia, congenital anomalies and predisposition to malignancy. DBA is caused by monoallelic, loss-of-function mutations in ribosomal protein (RP) genes. Delineation of the precise erythroid defect underpinning anaemia in DBA has been hampered by a lack of markers that define cells giving rise to burst- and colony-forming unit-erythroid (BFU-E and CFU-E) colonies, i.e., the clonogenic assays that quantify early and late erythroid progenitor (EP) potential respectively. By combining flow-cytometry, cell-sorting and single cell clonogenic assays, I identify Lin-CD34+CD38+CD45RA-CD123-CD71+CD41a-CD105-CD36- BM cells as early EP (EEP) and Lin-CD34+/-CD38+CD45RA-CD123-CD71+CD41a-CD105+CD36+ cells as late EP (LEP) giving rise to BFU-E and CFU-E respectively. By applying these definitions to DBA, I reveal, for the first time, that both quantitative and qualitative defects in EEP and LEP contribute to defective erythropoiesis in DBA and its restoration by corticosteroids (CS). I also demonstrate that, while an EP defect associated with profound loss of erythroblasts (EB) is present in transfusion-dependent (TD) patients with *RPS* gene-associated DBA, EP/EB are relatively preserved in TD patients with *RPL* gene variants. Transcriptome profiling by RNA-sequencing of FACS-purified EB from *RPL* genotype patients compared with age-matched controls, provides a unique dataset showing a distinct DBA molecular signature, characterised by p53 and inflammatory signalling pathways. Remarkably, the cellular differences between *RPL* and *RPS* gene-associated DBA correlate with divergent clinical phenotypes: patients with *RPL* rather than *RPS* gene pathogenic variants are more likely to present with anaemia at an older age, and to respond to CS. Finally, although DBA is rare, I demonstrate its wider utility as a paradigm of disordered erythroid commitment, providing insight into normal haematopoiesis. The strategies described for prospective isolation of

EP/EB constitute an original contribution to knowledge that will facilitate more incisive study of normal and aberrant erythropoiesis.

ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to my supervisor and mentor, Professor Tassos Karadimitris, who has helped me to flourish and consistently supported me to strive for the highest standards, leading by example. I am also extremely grateful to Professor Irene Roberts for initially giving me the opportunity to do this project and for her encouragement and invaluable contribution to this work. Thank you also to Dr Josu de la Fuente: this project would not have been possible without his personal support, and his unfailing commitment to patients with DBA.

I am grateful to many members of the Department of Haematology, both past and present. Special thanks to Dr Beth Psaila as working alongside her was the most fulfilling and enjoyable part of my PhD. Also thanks to Dr Aristeidis Chaidos and Dr Mark Robinson, who designed the pipelines and advised on the bioinformatics analysis of the RNA-seq data. I wish to acknowledge Professor Naresh and Dr Pritesh Truvadi, who performed the immunohistochemistry on BM trephine sections and Dr Richard Szydlo, who assisted with statistical analysis of the clinical data. I also thank the Molecular Pathology Department at Hammersmith Hospital, in particular Dr Gareth Gerrard and Dr Mary Alikian who analysed the data from next generation sequencing of ribosomal protein genes and reported the variants identified in patients. Also thanks to Dr Katerina Goudevenou, Dr Catherine Garnett (*Paresh Vyas lab, Oxford*), Dr Valentina Caputo, Dr Andi Roy, Mr Nikos Trasanidis, Dr Neha Bhatnagar and Dr David O'Connor for their advice regarding specific techniques, and to all the lab members who helped to keep my primary cell samples safe in the LN tanks! In addition I am thankful to the John Goldman centre for cellular therapy, the MRC Genomics and Flow Cyometry Facilities and the NIHR BRC Imaging and FACS Facility, for their helpful guidance.

This project relied almost exclusively on human samples thus would not have been possible without the trust and generosity of the patients and their families. Observing their fortitude inspired to continue with this project through difficult times. I am also thankful to members of the Paediatric Haematology team at St Mary's Hospital, especially clinical nurse specialist Yvonne Harrington and Dr Clare Miller, both of whom proved to be completely dependable. Also thanks to Dr Leena Karnik for consenting the normal donors and to all the SpRs who collected research samples.

I am immensely grateful for the financial support I have received to do this project, above all from Bloodwise, who funded my Clinical Research Fellowship and from the NIHR, DBA UK and the Imperial College Healthcare Charity.

Last but not least, I am indebted to my friends and family (especially Mohsen & Laila Iskander) for their practical help and, even more, for their unconditional support and faith in me over so many years. Above all, thanks to Chris and Chloe Skeoch; I don't think I could have had a more loyal, encouraging and patient family. Chris: you even convinced me that you are genuinely interested in erythroid commitment, so this thesis is dedicated to you.

DECLARATION

I confirm that the work contained in this thesis is my own and that, where collaborations were provided, they are explicitly acknowledged.

Deena Iskander

November 2017

LIST OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

AD	Autosomal dominant
AF405	Alexa Fluor 405
ALL	Acute lymphoblastic leukaemia
AML	Acute myeloid leukaemia
APC	Allophycocyanin
APCeF780	Allophycocyanin eFluor780
AR	Autosomal recessive
ATAC	Assay for Transposase Accessible Chromatin
B-ALL	B acute lymphoblastic leukaemia
bp	Base pair
BCC	Basal cell carcinoma
BFU-E	Burst forming unit-erythroid
BM	Bone marrow
BV	Brilliant violet
CD	Cluster of differentiation
cDNA	Complementary deoxyribonucleic acid
CFU-E	Colony forming unit-erythroid
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
ChIP	Chromatin immunoprecipitation
CIN	Carcinoma in situ
CLL	Chronic lymphocytic leukaemia
CMP	Common myeloid progenitor

CON	Control
CRIPSR	Clustered Regularly Interspaced Short Palindromic Repeats
CS	Corticosteroids
CSD	Corticosteroid dependent
CSR	Corticosteroid responsive
CSI	Corticosteroid intolerant
DAPI	4',6-diamidino-2-phenylindole
DBA	Diamond-Blackfan anaemia
DBAR	DBA registry of North America
DEX	Dexamethasone
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
eADA	Erythrocyte adenosine deaminase
ELISA	Enzyme linked immunosorbent assay
EEP	Early erythroid progenitor
EP	Erythroid progenitor
EPO	Erythropoietin
FACS	Fluorescence- activated cell sorting
FBC	Full blood count
FBS	Fetal bovine serum
F	Female
FcR	Fc receptor
FITC	Fluorescein isothiocyanate
FL	Fetal liver
FMO	Fluorescence minus one
GA	Gestational age
GAS-6	Growth arrest-specific protein 6
G-CSF	Granulocyte-Macrophage Colony Stimulating Factor
gDNA	Genomic deoxyribonucleic acid
G-MCSF	Granulocyte-macrophage colony-stimulating factor

GI	Gastrointestinal
GMP	Granulocyte-macrophage progenitor
GR	Glucocorticoid receptor
Hb	Haemoglobin
HIF	Hypoxia-inducible factor
HSC	Haematopoietic stem cell
HSCT	Haematopoietic stem cell transplantation
HSP	Heat shock protein
HSPC	Haematopoietic stem/progenitor cell
IBMFS	Inherited bone marrow failure syndrome
ICHT	Imperial College Healthcare Trust
ID	Identifier
IL-	Interleukin
IMDM	Iscove's Modified Dulbecco's Medium
IUGR	Intrauterine growth restriction
IVF	*in vitro* fertilisation
m(s)	Month(s)
LEP	Late erythroid progenitor
LMPP	Lymphoid-primed multipotent progenitor
LTIC	Long-term initiating cells
M	Male
MCV	Mean cell volume
MDM2	Mouse double minute 2
MDS	Myelodysplastic syndrome
MEP	Megakaryocyte erythroid progenitor
MGG	May-Grünwald Giemsa
MkP	Megakaryocyte progenitor
MPP	Multipotent progenitor
mRNA	Messenger ribonucleic acid
NGS	Next generation sequencing

NHS	National Health Service
PECy7	Phycoerythrin-Cyanine7
PB	Peripheral blood
PBS	Phosphate Buffer Saline
PBST	Phosphate Buffer Saline with Tween 20
PCR	Polymerase chain reaction
PE	Phycoerythrin
PerCPCy5.5	Peridinin-chlorophyll- protein complex-cyanine 5.5
PID	Pre-implantation diagnosis
Plts	Platelets
PPAR	Peroxisome proliferator-activated receptor
Pt	Patient
PR	Primary refractory
P/S	Penicillin + streptomycin
RBC	Red blood cell
RD	Reagent diluent
RPL	Ribosomal protein large subunit
RPS	Ribosomal protein small subunit
RNA	Ribonucleic acid
rRNA	Ribosomal RNA
RNA-seq	RNA-sequencing
qRT-PCR	Quantitative real time polymerase chain reaction
SCF	Stem cell factor
SGA	Small for gestational age
Sh	Short hairpin
Si	Small interfering
SNT	Corticosteroid not tried
SR	Secondary refractory
TALEN	Transcription activator-like effector nucleases
TD	Transfusion dependent

TEC	Transient erythroblastopenia of childhood
TF	Transcription factor
TPO	Thrombopoietin
UTR	Untranslated region
WCC	White blood cell count
wk(s)	Week(s)
yr(s)	Year(s)

1 INTRODUCTION

1.1 Diamond-Blackfan anaemia

1.1.1 Epidemiology

Diamond–Blackfan anaemia [DBA Mendelian Inheritance in Man (MIM): 205900] is a rare inherited bone marrow failure syndrome (IBMFS) first described by Diamond and Blackfan in 1937 and originally referred to as chronic erythrocytosis imperfecta (Diamond & Blackfan 1938) . DBA is estimated to occur in 1-2 per 100 000 live births, hence it is one of the more common IBMFS, although definitive incidences of these conditions are unknown (Dokal & Vulliamy 2010). Like all IBMFS, DBA is characterised by a triad of bone marrow failure, cancer predisposition and systemic congenital abnormalities. Males and females are equally affected and there is no apparent predilection for a particular ethnic group. Although a viral aetiology was previously suggested by the clustering of birth months observed by Ball et al (Ball et al. 1996) this link was not substantiated by subsequent population studies (Ramenghi et al. 1999).

1.1.2 Molecular genetics and pathogenesis

Most cases of DBA are caused by mono allelic mutations or deletions of genes that encode ribosomal proteins (RP). DBA is therefore one of the 'ribosomopathies' (Luzzatto & Karadimitris 1998), a term coined to describe a group of IBMFS also comprising Shwachman-Diamond syndrome, X-linked dyskeratosis congenita (DC) and cartilage hair hypoplasia (Table 1.1). The majority of pathogenic mutations- 60-80% depending upon the population studied (Campagnoli et al. 2004)- occur *de novo*. The remainder are inherited as an autosomal dominant (AD) trait; importantly once a mutation is detected in a proband, even if it arose sporadically, there is a 50% risk of vertical transmission. Very rarely, X-linked inheritance of mutated *GATA-1* (an erythroid-specific transcription factor) or *TSR2* (a RP chaperone), are implicated in DBA (Gripp et al. 2014; Sankaran et al. 2012). Although putative autosomal recessive

cases (AR) have been described (Campagnoli et al. 2004), these were not formally proven, rather presumed on the basis of consanguineous parents and/or affected siblings with seemingly unaffected parents. Possible alternative explanations to AR inheritance include i) parental germ line/gonadal mosaicism or ii) a related bone marrow failure syndrome that shares clinical features with DBA (Table 1.1). In the majority of international cohorts, pathogenic genetic variants are identified in 50-75% of patients (Gerrard et al. 2013; Boria et al. 2010). DBA shows markedly variable expressivity: the clinical phenotype ranges from asymptomatic carriers to patients with severe anaemia and congenital malformations. This incomplete penetrance exists even within multiplex families comprising individuals with the same mutation/deletion (Carlston et al. 2017). It has therefore been assumed that additional genetic or environmental factors modify the clinical expression of the underlying genetic defect, although no such modifiers have been identified as yet.

Table 1.1 Classification of inherited and acquired bone marrow failure syndromes*

Disease	Mode of inheritance	Chromosomal locus	Gene	Clinical manifestations
Diamond–Blackfan anaemia	AD	See Table 1.2		See text
	X-linked recessive	Xp11	*TSR2*	
	X-linked recessive	Xp11	*GATA-1*	
Fanconi anaemia	AR/X-linked	21 different genes on 13 autosomes + X chromosome implicated to date		Gradual onset BMF, skeletal abnormalities most commonly of the radius and thumb, skin lesions, renal and urinary tract malformations and gonadal dysfunction

Dyskeratosis congenita	X-linked recessive	Xq28	*DKC1*	Lacy reticulated skin, nail dystrophy, and mucosal leukoplakia in the first years of life; Development of BMF
	AD	3q26	*TERC*	
	AD or AR	5p15	*TERT*	
	AR	17p13	*CTC1*	
	AD	14q12	*TINF2*	
	AR	15q14	*NOP10*	
	AR	17p13	*WRAP53/TCAB1*	
	AD or AR	16q2	*ACD/TPP1*	
	AD or AR	20q13	*RTEL1*	
	AR	16p13	*PARN*	
	AR	5q35	*NHP2*	
Shwachman–Diamond syndrome	AR	7q11	*SBDS*	Neutropenia, exocrine pancreatic insufficiency, metaphyseal dysostosis
Amegakaryocytic thrombocytopenia	AR	1p34	c-*mpl* (thrombopoietin receptor)	Absent megakaryocytes in bone marrow, late BMF
Thrombocytopenia with absent radii syndrome	AR	1q21	*RBM8A*	Bilateral radial aplasia, lower limb anomalies, cow's milk intolerance, renal anomalies, and cardiac anomalies

Congenital thrombocytopenia and radius–ulna synostosis	AD	7p15–p14	*HOXA11*	Aplastic anaemia, proximal radius–ulna synostosis, clinodactyly, syndactyly, hip dysplasia and sensorineural hearing loss
	De novo	3q26	*MECOM*	
Cartilage-hair hypoplasia	AR	9p13	*RMRP*	Short tubular bones, sparse hair, severe immunodeficiency, macrocytic, anaemia
Pearson marrow–pancreas syndrome	Mitochondrial	Usually from nt 8469 to nt 13447	Contiguous genes deleted	Pancreatic exocrine dysfunction, sideroblastic anaemia
ACQUIRED				
Idiopathic				
Transient erythroblastopenia of childhood (TEC)				
Radiation				
Drugs and chemicals				
Regular: cytotoxic, benzene				
Idiosyncratic: chloramphenicol, NSAIDs, anti-epileptics, gold				
Viruses				
Epstein–Barr virus				
Hepatitis				
Parvovirus				
Human immunodeficiency virus				
Cytomegalovirus				
Human T-lymphotropic virus-1				

Human herpes virus-6
Immune diseases
Thymoma (5-10% develop pure red cell aplasia)
Myasthenia gravis
Systemic lupus erythematous
Pregnancy
Paroxysmal nocturnal haemoglobinuria

*AD, autosomal dominant; AR, autosomal recessive; NSAID, non-steroidal anti-inflammatory drug; nt, nucleotide. *(Luzzatto & Karadimitris 2010)*

The first pathogenic genetic variant identified in DBA was a heterozygous mutation in *RPS19* resulting from a *de novo* balanced translocation*:* 46,XX,t(X;19)(p21;q13) (Draptchinskaia et al. 1999). To date, *RPS19* remains the most common gene involved in DBA, accounting for 25% of cases (Lipton et al. 2006). However, since its identification, causal variants in 21 other proteins of both the small (RPS) and large (RPL) ribosomal subunits have been described, many in the last few years, including *RPS24*, *RPS26*, *RPL5*, *RPL35A* and *RPL11* (Table 1.2). A wide variety of genetic variants are reported including missense, nonsense, splice site mutations, frameshift insertion/deletions, and whole allele deletions (Gazda et al. 2004) (Farrar et al. 2011). The latter sometimes occur due to structural variants, for example in the case of 3q27.2 microdeletion syndrome (Alkhunaizi et al. 2017), in which the deleted region encompasses *RPL35A;* or 19q13.2 microdeletion involving *RPS19* (Yuan et al. 2016; Cario et al. 1999). In these contexts, DBA is part of a more complex clinical phenotype that may include developmental delay, congenital malformations and immunodeficiency, depending on the contiguous genes deleted.

The finding that the whole RP gene is deleted in some cases of DBA, strongly argued that haploinsufficiency, as opposed to dominant negative mutants, is the underlying pathogenetic mechanism in most cases of DBA (Hamaguchi et al. 2002; Gustavsson et al. 1998). In the case of whole allele deletions, it is intuitive that this results in a decrease in messenger ribonucleic acid (mRNA) of the corresponding gene to 50—60% of normal and haploinsufficiency. Similarly for splicing defects that result in exon skipping or nonsense/ frameshift deletions resulting in a premature stop codon, reduction of RPS19 protein has been confirmed, secondary to nonsense-mediated decay of the abnormal transcript (Chatr-Aryamontri et al. 2004; Gazda et al. 2004), a physiological RNA surveillance process that prevents the expression of aberrant peptides (Frischmeyer et al. 2002). To explain how RP function is altered in the presence of a missense mutation, where the mRNA level of the mutated gene is predicted to be normal, three mechanisms have been postulated:

1) In most cases the missense variant is likely to lead to decreased stability of the

protein and impaired incorporation into the ribosome or loss of its function (Angelini et al. 2007). Missense mutations in *RPS19* have also been shown to affect its nucleolar localisation (Da Costa 2003). In the case of this mechanism, one would predict that other cells (all of which are affected by the same RP gene mutation) might be much less affected than erythrocytes, because while the former can compensate for decreased stability through increased synthesis of the RP protein, the latter lack the organelles required for protein synthesis.

2) In some cases, the amino acid replacement may affect an important functional domain of the protein. This hypothesis is supported by evidence that missense mutations cluster in a 'mutational hotspot' between codons 52 and 62 o*f RPS19* (Willig, Draptchinskaia, et al. 1999a);

3) Some missense mutations may exert a dominant negative effect, as suggested by a transgenic murine model expressing an *RPS19* missense mutation present in multiple DBA pedigrees (Devlin et al. 2010).

In addition to the more common genetic lesions described above, it has been shown that intronic variants (e.g. upstream of the ATG start site/ in the promoter region of *RPS19*) (CrÃ tien et al. 2010) can have a deleterious impact upon mRNA production and lead to DBA. As knowledge of the function of the non-coding genome expands, it is likely that more variants in RP gene regulatory areas will be implicated in DBA.

In summary, there is now firm evidence that DBA is underpinned by haploinsufficiency of ribosomal proteins. How aberration in such a global cellular process triggers tissue-specific disease manifestations will be discussed in section 1.3.3. The utility of understanding these mechanisms extends beyond DBA, to an expanding spectrum of diseases associated with RP gene mutations. For example, germline inheritance of mutated RPS23 has been linked to dysmorphic features and neurological impairment in the absence of a haematological abnormality (Paolini et al. 2017). Furthermore, the role of somatic RP gene mutations in the biology of sporadic cancers is increasingly

being recognised: in 5q- syndrome, a subtype of myelodysplastic syndrome (MDS), characterised by somatically acquired heterozygous loss of *RPS14* on chromosome 5q (Ebert et al. 2008); in multiple myeloma (MM; RPL5, RPL10) (Hofman et al. 2017); T-cell acute lymphoblastic leukaemia (T-ALL; RPL5, RPL10; (De Keersmaecker et al. 2012) and chronic lymphocytic leukaemia (CLL; RPS15; (Landau et al. 2015). In many cases these somatically acquired mutations interfere with the function of the tumour suppressor p53.

1.1.3 Presentation and clinical features

DBA presents at birth in 25% (Campagnoli et al. 2004) and by the age of 1 year in 90% of patients (Clinton & Gazda 2009), usually with lethargy, pallor, prolonged jaundice or cardiorespiratory complications. Although there are a few case reports of non-immune hydrops fetalis secondary to DBA (E. G. Zhang et al. 2011; Da Costa 2013; McLennan et al. 1996; Dunbar et al. 2003), in general fetal erythropoiesis appears to be relatively spared, for reasons that are as yet unknown.

Typical peripheral blood (PB) features at diagnosis are normochromic, macrocytic anaemia, reticulocytopenia, increased haemoglobin F (HbF) and elevated erythrocyte adenosine deaminase (eADA), an enzyme involved in purine metabolism. The anaemia is usually severe (<50g/L) by the time of presentation. Although DBA has classically been described as isolated anaemia resulting from pure red cell aplasia, in fact leucopenia, thrombocytosis or thrombocytopenia are not uncommon at presentation (Giri et al. 2000).

With disease progression, patients may develop bi- or tri-lineage hypoplasia; indeed thrombocytopenia and neutropenia occurred in 15% of patients in one study (Campagnoli et al. 2004). Typically, the bone marrow (BM) at diagnosis is normocellular with a paucity or absence of erythroblasts (specifically erythroblasts <5% of nucleated cells on BM smears; Figure 1:1). Some case series describe normal numbers of erythroblasts at presentation (Wiernik et al. 2012; Pospisilova et al. 2012).

Progressive BM hypocellularity is also increasingly recognised in many patients (Giri et al. 2000).

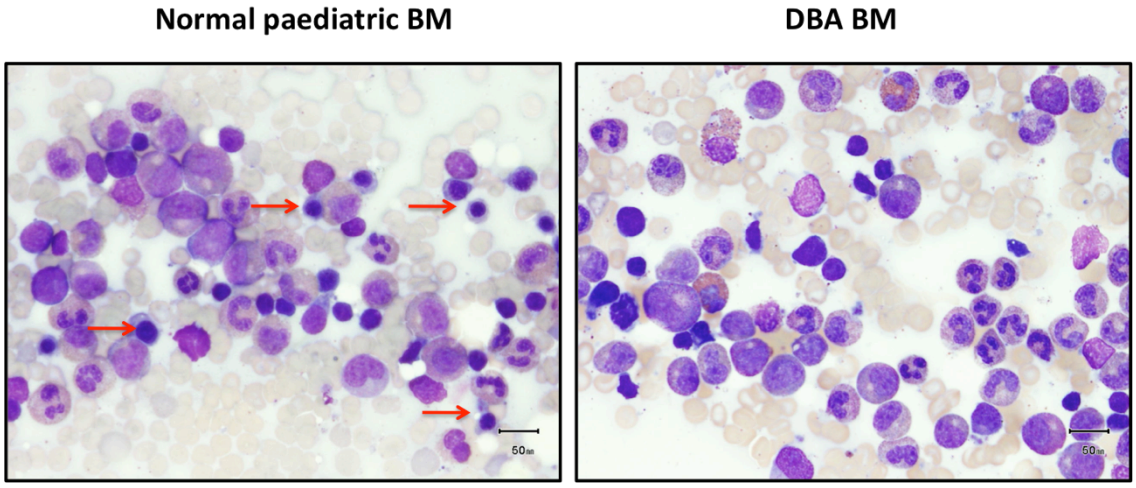


Figure 1:1 Selective paucity of erythroid precursors in DBA bone marrow. Left-hand side image shows a bone marrow aspirate (BMA) from a healthy child aged 8 years (sibling donor for HSCT). Red arrows highlight erythroblasts. Right-hand side image shows a BMA from transfusion-dependent patient with DBA aged 5 years. There is a remarkable paucity of erythroid precursors and myeloid cells predominate. Photographs taken with Nikon eclipse E400 inverted microscope and camera.

Table 1.2 List of genes known to harbour pathogenic variants causing Diamond-Blackfan anaemia

Gene name	Genetic Locus	% of cases^	No. of pedigrees (References)
RPS19	19q13.2	25	Several families (Campagnoli et al. 2004) (Gazda et al. 2008) (Ichimura et al. 2016) (Farrar et al. 2008) (Boria et al. 2010) (Gerrard et al. 2013) (Doherty et al. 2010)
RPL5	1p22.1	6.6	
RPS26	12q13.2	6.4	
RPL11	1p36.11	4.8	
RPL35A	3q29	3	
RPS10	6p21.31	2.6	
RPS24	10q22.3	2	
RPS17	15q25.2	1	1 patient (Cmejla et al. 2007); 2 patients, 1 family (Gerrard et al. 2013); 1 patient (Gazda et al. 2008); 1 patient (Song et al. 2010)
RPS7	2p25.3	<1	2 patients, 1 family (Ichimura et al. 2016)
RPS29	14q21.3	<1	2 multiplex families (Mirabello et al. 2014)
RPL26	17p13.1	<1	1 patient (Gazda et al. 2012)
RPS27	1q21.3	<1	1 patient(R. Wang et al. 2015)
RPL27A	11p15.4	<1	2 patients, 1 family (Gazda et al. 2008)
RPL27	17q21.1-21.2	<1	1 patient (R. Wang et al. 2015)
RPL15	3p24.2	<1	1 patient (Landowski et al. 2013)
RPL31	2q11.2	<1	1 patient (Farrar & Dahl 2011)
RPS28	19p13.2	<1	2 families (Gripp et al. 2014)
RPL35	9q33.3	<1	1 family (Mirabello et al. 2017)

RPS15	19p13.3	<1	1 patient (Cmejla et al. 2000)
RPL36	19p13.3	<1	1 patient (Gazda et al. 2008)
RPL18	19q13.33	<1	2 patients, 1 family (Mirabello et al. 2017)
RPS28	19p13.2	<1	2 patients, 2 families (Gripp et al. 2014)
TSR2	Xp11.22	<1	1 patient (Gripp et al. 2014)
*GATA-1**	Xp11.23	1	7 patients, 2 families (Hollanda et al. 2006), 2 patients, 1 family (Sankaran et al. 2012), 1 patient (Parrella et al. 2014)

^(Clinton & Gazda 2009) * Clinical and haematological features not typical of DBA

System	Specific abnormalities
Cardiac	Patent ductus arteriosus, atrial/ ventricular septal defect, coarctation of the aorta, complex cardiac defects
Musculoskeletal	Triphalangeal, duplicated or bifid thumbs, radial hypoplasia, clinodactly, flattened thenar eminence, syndactyly
Craniofacial	Cleft lip/palate, high arched palate, Pierre-Robin anomaly, Cathie facies, choanal atresia, craniosynostosis, microtia, epicanthus, ptosis, microcephaly, strabismus, cataract, glaucoma
Neck and shoulders	Webbed neck, short neck, Sprengel deformity, Klippel-Feil malformation
Genitourinary	Hypospadias, undescended testis, horseshoe kidney, absent kidney

Figure 1:2 Range of congenital anomalies described in DBA (Ball et al. 1996). Photograph shows abnormal thumbs in a U.K. patient with DBA secondary to a monoallelic deletion of *RPL5 (patient consent obtained).*

In addition to the haematological manifestations, developmental abnormalities (Figure 1:2) are observed in at least 30-50% of patients with DBA. Multisystem co-morbidities are also widely reported, though their exact prevalence remains unknown. These include faltering growth, marked by short stature (reported in 30% of patients), endocrine disorders (Lahoti et al. 2016) and dental disease (Ozden et al. 2011). There are occasional reports of other clinical manifestations such as learning difficulties (Vlachos & E. Muir 2010) though these have not yet been quantified or described systematically in a large cohort of patients. It is often difficult to discern whether clinical co-morbidities are constitutional and/or secondary to disease complications and treatments.

Finally, patients with DBA have an increased risk, relative to the general population, of cancer. In 2012 the incidence of malignancy in the DBA Registry (DBAR) was estimated at ~20% by age 46y, comprising solid tumours in 15% and MDS/AML in 5% of patients (Vlachos et al. 2012). Latterly it has been estimated to be even higher (Table 1.3), approximating the risk of cancer in DC (40% by age 50 years) but still lower than that in FA (where the risk of MDS/AML alone is 52% by the age of 40 years (Alter et al. 2009). The median age at diagnosis of a solid tumour for patients with DBA who have not received HSCT is 35 yrs (range, 11-63yrs) (Vlachos et al. 2016). The predilection for certain malignancies and the mechanisms of neoplastic transformation are not yet understood.

Table 1.3 Incidence of neoplasia in the DBAR. Data from 702 patients with 12,376 person-years of follow-up.

Neoplasia		No of patients	O/E Ratio
All cancers		37	4.75
Solid cancers	Osteosarcoma	4	42
	Colorectal adenocarcinoma	9	45

	Other*	13	
Haematological	MDS	8	352
	AML	3	29

** 3 patients with >1 cancer. The Observed/Expected (O/E) Ratio was calculated using surveillance and epidemiology data for cancer from the National Institute of Health, USA (Vlachos et al. 2016).*

1.1.4 Diagnostic criteria

Diagnostic criteria for DBA include the presence of classical haematological and clinical features and identification of a pathogenic gene mutation (Table 1.4). The presence of congenital anomalies, age <1y, raised eADA and macrocytosis all suggest a diagnosis of DBA, rather than transient erythroblastopenia of childhood (TEC), which is the main differential diagnosis (Vlachos et al. 2008; Glader et al. 1983). This is an important distinction as the latter condition usually resolves spontaneously within a few months. Macrocytosis and raised HbF are indicators of stress erythropoiesis and therefore they are not specific to DBA and also occur in FA. eADA, on the other hand, is 84% sensitive and 95% specific for DBA at diagnosis (Fargo et al. 2012). Although eADA has been reported to be raised in AML, ALL (Glader et al. 1983) and myeloproliferative neoplasms (Storch et al. 1981), it is not elevated in congenital anaemias (other than DBA) nor in the 5q- subtype of MDS (Narla et al. 2016). Interestingly a phenotype characterised by an isolated elevated eADA has been identified in healthy family members of index cases (49 out of 149 in one study; (Willig et al. 1998). While a proportion of these cases are likely to be asymptomatic carriers of a DBA-causing mutation, in some pedigrees, raised eADA has been observed to segregate separately from the pathogenic mutation (Fargo et al. 2012).

A recent paper suggested that elevated glutathione synthase could serve as an additional marker for DBA. When tested in isolation, it was less sensitive than eADA, however when measured simultaneously with eADA, this improved the sensitivity of DBA diagnosis from 77 to 100% with 100% specificity for DBA versus healthy controls

(Utsugisawa et al. 2016). This study was limited by a small sample size of 22 patients and the exclusion of patients with other congenital anaemias; it is therefore unlikely that this test will be integrated into routine diagnostic practice.

In specialised DBA centres, genetic diagnosis using targeted next generation sequencing (NGS) of RP genes has replaced traditional Sanger sequencing. This approach, complemented by additional techniques such as multiplex ligation-dependent probe amplification (MLPA) for validation of computationally identified deletions, has led to an improvement in the pick-up rate of mutations from 40% to around 70% of patients (Gerrard et al. 2013) and has led to the identification of new candidate RP genes (Table 1.2).

Whilst great progress has been made in characterizing the RP variants underlying the development of DBA, the molecular basis of the disease in 30% of cases is yet to be elucidated. A recent study discovered that 8 of 173 patients with presumed DBA but no pathogenic RP gene mutation, in fact harboured mitochondrial DNA deletions similar to those seen in the Pearson marrow-pancreas syndrome (Gagne et al. 2014)). Given the clinical overlap between the IBMFs and the accessibility of genetic testing in routine clinical practice, a child presenting with congenital anaemia should be tested for multiple conditions, using a panel encompassing multiple genes (N. B. A. Roy et al. 2016). Furthermore, cases in which a mutation is not identified by targeted multigene assays are candidates for whole genome sequencing with the aim of novel gene discovery, as exemplified by the Genomics England 100,000 Genomes Project (Griffin et al. 2017).

Table 1.4 International consensus on the diagnostic criteria for DBA (Vlachos et al. 2013).

CLASSICAL DBA	NON-CLASSICAL DBA	POSSIBLE DBA
All of diagnostic criteria:	**1 diagnostic criterion & reported mutation**	**3 diagnostic criteria & family history**
Age <1yr		**OR 2 diagnostic criteria +3 minor criteria:**
Macrocytic anaemia with no other significant cytopenias		Raised eADA
Reticulocytopenia		Classical congenital anomaly
Normocellular bone marrow with paucity of erythroid precursors		Raised HbF
		No evidence of another IBMFS

1.1.5 Treatment modalities

The management of DBA focuses on treatment of the anaemia and includes blood product support, corticosteroids, and haematopoietic stem cell transplantation (HSCT) in selected cases. Occasionally DBA is detected incidentally, when a full blood count is performed for another reason. In these cases, where the anaemia is too mild to warrant treatment, patients can be educated on their higher cancer risk and managed by surveillance alone, to pre-empt any deterioration in their haematological status.

1.1.5.1 Blood transfusions

Chronic transfusions with packed red blood cells are the mainstay of treatment for moderate to severe anaemia in DBA. In most cases, transfusion is required initially while a diagnosis is reached and until it is deemed safe to commence corticosteroids (CS). eADA may normalise in patients on blood transfusion (Willig et al. 1998) so ideally this test should be performed prior to transfusion in any patient with congenital anaemia. Usually transfusions are administered every 4-6 weeks, with the aim of maintaining the Hb above a particular threshold, bespoke for each individual to avoid symptomatic anaemia and to facilitate growth.

Iron overload is common in transfusion-dependent (TD) patients, usually occurring in the liver but occasionally also in the heart and endocrine organs (Vlachos et al. 2008). Notably, patients with DBA are more prone to iron overload than individuals with other TD anaemias and the risk heightens after 20 red blood cell units (corresponding to approximately 6 months of transfusions). Indeed, hepatic iron overload has been detected in children as young as 2 years of age (Berdoukas et al. 2013) and can persist even after discontinuation of transfusions. Data from 5 patients suggest that this propensity for iron loading is explained by release of non-transferrin bound iron from overloaded macrophages (a process generic to TD anaemias) compounded by low utilization of transferrin iron due to the paucity of erythroblasts (unique to DBA) (Porter et al. 2014). It is therefore important to monitor patients for evidence of iron overload, using a combination of serum ferritin measurements, hepatic ferriscan, cardiac magnetic resonance imaging (MRI) and liver histology. Patients receiving >100mls/kg/year red cells should be commenced on iron chelation- usually deferasirox (Exjade®) or in some cases subcutaneous desferrioxamine infusion (Vlachos & E. Muir 2010).

1.1.5.2 Corticosteroids

Corticosteroids (CS), usually in the form of prednisolone, are the first- line therapy for DBA and have been used since the 1950s (Calvert & Robson 1956). Unless there is

a compelling reason not to trial CS, most patients are started on high dose prednisolone (2mg/kg/day) at the age of 12 months. This delay allows for the administration of immunisations and mitigates the detrimental effects of CS on musculoskeletal development in infancy. Initially 80% of patients respond to CS, with a surge in reticulocytes followed by a rise in Hb, leading to transfusion-independence. However, in 40-50% of these responders, CS treatment is discontinued, either because the dose required for therapeutic benefit risks unacceptable side effects or because the haematological response is lost (Lipton et al. 2006). This may be triggered by transition into puberty, pregnancy or by use of female hormones for gender reassignment (*Vlachos A, unpublished*) but in other cases there is no clear precipitant. Occasionally patients respond to a second steroid trial (Willig, Niemeyer, et al. 1999b), however, this is only recommended in cases where the initial trial was deemed inadequate or prior to HSCT (Vlachos & E. Muir 2010).

High dose methylprednisolone (30-100 mg/kg/day) has been considered as a second-line treatment option in prednisolone-resistant cases on the basis of one study where 4 out of 10 patients demonstrated an initial response (Bernini et al. 1995). As a caveat, only 1 of these individuals subsequently achieved haematological remission, thus dissuading clinicians from adopting methylprednisolone into the routine management of DBA.

Ultimately, 30-40% of patients with DBA can be safely maintained on CS in the long-term, characteristically at a very low dose of <1mg/kg on alternate days. Whilst CS-responsiveness is deemed a good outcome in patients with DBA, CS therapy is by no means a panacea. Side effects in children include growth retardation, myopathy, osteoporosis, avascular necrosis, severe infections, psychosis, cataracts, hypertension, and diabetes mellitus (Lipton et al. 2006; C. Muir et al. 2017). In order to develop alternative and better-tolerated drugs for DBA, it is necessary to solve

one of the unsolved riddles in DBA: the mechanisms of CS-responsiveness and -refractoriness in some patients. This will be discussed further in section 1.3.4.

1.1.5.3 Allogeneic haematopoietic stem cell transplantation

HSCT is the only curative therapy for DBA and was used as early as 1976 (August & King 2005). Indications for HSCT include transfusion-dependence and steroid-resistance, or those with progressive BM failure (BMF) (Vlachos et al. 2008). Historically HSCT was *not* recommended after the age of 10 years (Lipton et al. 2006), however, poor outcomes in older patients were confounded by the severity of iron overload (Lucarelli 2008) and HSCT can be performed >10 years as long as iron chelation is optimal. Although there are some data to support the use of unrelated donors, these are reserved for patients with BMF and no sibling donor. Stem cells should be derived from BM, in combination with umbilical cord blood (UCB), if the latter is necessary to achieve an adequate cell dose (Vlachos et al. 2001). As yet there is no international consensus on the optimal conditioning regimen; however, the aim is to minimise both the risks of morbidity associated with intensive conditioning and the residual RP haploinsufficient cells in the BM, as these have an increased risk of transformation to MDS/AML. Mortality following HSCT is estimated at 5-10%, depending on the experience of the centre. The main complications of HSCT are acute and chronic GvHD, which occur in 57% and 21% patients respectively (Fagioli et al. 2014). Importantly, HSCT cures only the haematological manifestations of DBA and does not modify either the congenital abnormalities or the future risk of solid tumours.

1.1.6 Natural history

As outlined above, in the long-term, approximately 30-40% of patients remain on a very low dose of CS and 40-60% are TD or undergo HSCT. The remaining 10-20% of patients enter into a clinical remission by the age of 25 years. Remission is defined as transfusion independence for 6 months or more and has been observed during transfusion therapy or following discontinuation of CS. 15% of patients in remission

will later relapse (Lipton et al. 2006). Some will then prove resistant to CS, even if they were previously CS-responsive, therefore in general, CS are maintained life-long at the lowest dose required for a therapeutic effect.

A recent report from the DBAR showed a median overall survival of 56 years (95% confidence interval: 49 – 62 years; (Vlachos et al. 2012). Overall actuarial survival, reported in 2006, was 75.1% +/-4.8% at 40 years of age overall, but was significantly increased for steroid-dependent (86.7% +/- 7.0%) versus transfusion-dependent (57.2% +/-8.9%) patients (Lipton et al. 2006). The increased mortality in the TD group is primarily due to the complications of HSCT and iron overload. As supportive therapy improves and patients with DBA can anticipate longer lives, it becomes increasingly important to understand the drivers of malignant transformation so that the risk of cancer can be minimised.

1.2 Normal erythropoiesis

1.2.1 Ontogeny

Blood production in humans commences at a gestational age (GA) of approximately 3 weeks in the yolk sac (Huyhn et al. 1995). A transient wave of primitive haematopoiesis, dominated by erythropoiesis, is followed by the onset, of definitive haematopoiesis in the aorta-gonad mesonephros, a structure forming part of the dorsal aorta (Charbord et al. 1996). In accordance with this, haematopoietic stem/progenitor cells (HSPCs) have been isolated from human embryos at 5 weeks gestation (Tavian et al. 1999) with haematopoietic stem cell (HSC) production becoming robust at approximately 11 weeks (Ottersbach et al. 2010). Fetal haematopoiesis is established in fetal liver (FL) at GA 5-8 weeks and in BM at 8-12 weeks, although BM becomes the major site of haematopoiesis only after birth, when FL haematopoiesis ceases (Medvinsky & Dzierzak 1996).

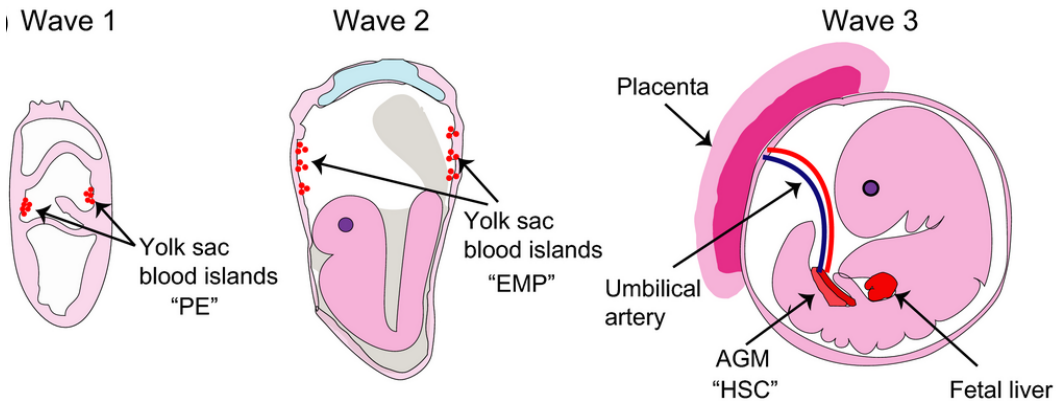


Figure 1:3 The ontogeny of mammalian erythropoiesis. The first wave occurs in the yolk sac blood islands and is characterised by the development of primitive erythroblasts (PE) that express embryonic globin chains. The second wave involves migration of erythro-myeloid progenitors (EMP) from the yolk sac to the fetal liver (FL). These generate definitive erythroblasts that express fetal haemoglobins. In the third wave, HSC egress from the haemogenic endothelium in the aorto-gonad mesonephros (AGM) to the FL, and ultimately to the BM. *[Image reproduced, with permission, from* (Nandakumar et al. 2016)*].*

1.2.2 The path from stem cells to red blood cells

Anaemia in DBA is the result of insufficient red blood cell (RBC) production. RBCs are the major cellular constituent of PB and are specially adapted for gas transport, principally oxygen carriage, by binding to the Fe2+ present in the heme component of haemoglobin. Haemoglobin is a tetrameric protein consisting of heme bound to equimolar amounts of four polypeptide chains (Table 1.5). β and γ chains are highly homologous, however the structure of the latter renders deoxyHbF less amenable than deoxyHbA to stabilization by 2,3-diphosphoglycerate (2,3-DPG). As a consequence, the oxygen dissociation curve is left-shifted and HbF has a higher affinity for oxygen, protecting the fetus from hypoxia. In conditions such as acute or chronic anaemia or recovery from HSCT, steady-state erythropoiesis is substituted by stress erythropoiesis which allows for a more rapid generation of RBCs. Activation of stress erythropoiesis is often accompanied by increased expression of the Hb

Gamma (*HBG)* gene with a resultant increase in HbF. Possible mechanisms include induction of *HBG* by stem cell factor (SCF), erythropoietin (EPO), glucocorticoids (GC) and transforming growth factor (TGF)β (Aerbajinai et al. 2009; Gabbianelli 2002) or expansion of erythroid progenitor cells that prematurely undergo terminal differentiation and are committed to producing γ-globin chain (Blau et al. 1993).

Table 1.5 Normal haemoglobin variants

Type of Hb	Globin subunits	Globin gene symbol; chromosome	Contexts in which Hb is present
Adult haemoglobin (HbA)			
HbA_1- 97%	2 α chains	HBA1, HBA2; 11	Adult life
	β chains	HBB; 16	
HbA_2- 3%	2 α chains	HBA1, HBA2; 11	
	2 δ chains	HBD; 16	
Fetal haemoglobin (HbF) <1%			
	2 α chains	HBA1, HBA2; 11	Fetal life ≈ 6 months post birth Stress erythropoiesis
	2 γ chains	HBG1, HBG2; 16	

Erythropoiesis refers to the developmental process by which RBC are generated from haematopoietic stem cells (HSC). While it is well established that HSC lie at the apex of the erythropoietic hierarchy and that RBC are at the terminally differentiated nadir, our understanding of the intermediary developmental pathway is continuously evolving (Orkin 2000) and varies in fetal and adult life (A. Roy et al. 2012). A classical model of adult haematopoiesis (black arrows, Figure 1.4) posits that HSC, also known as long-term HSC, develop into RBC via a series of cell types: multipotent progenitors (MPP; also known as short-term HSC), common myeloid progenitors (CMP) and megakaryocyte erythroid progenitors (MEP). Granulocyte-

macrophage progenitors (GMP) and lymphoid-primed multipotent progenitors (LMPP) arise at distinct branches of the hierarchy and are not involved in erythropoiesis (Doulatov et al. 2012). These HSPC are defined according to 2 major criteria 1) the cell surface expression of specific cluster of differentiation (CD) molecules, otherwise known as their 'immunophenotype' and 2) their ability, demonstrable in *in vitro* or *in vivo* model systems, to produce cells of particular haematopoietic lineages (Table 1.6) Traditionally, the *in vitro* assessment of the functional output of HSPC is performed by plating cells in a semi-solid medium, such as methylcellulose, supplemented with haematopoietic growth factors, and then enumerating the formation of haematopoietic colonies at defined time points. Colonies exhibit distinct morphology depending upon their lineage. Cell fate can also be studied in liquid culture systems, utilising different permutations of growth factors and cytokines to support specific lineages, or by xenotransplantation assays.

Such experiments have shown that transition along the pathway from HSC to terminally differentiated cell types is accompanied by decreasing proliferative/self-renewal capacity and increasing commitment to one particular lineage (Doulatov et al. 2012). For instance, according to the classical model, HSC and MPP can give rise to cells of any haematopoietic lineage (Table 1.6). However, it is not clear to what extent each HSPC population comprises true multi- or bipotent cells versus a mixture of unipotent cells with different lineage specifications. Accumulating evidence supports alternate models of haematopoiesis in which proximal cell types can differentiate into distal ones directly, without passage through each step in the classical developmental hierarchy (Adolfsson et al. 2005; Notta et al. 2016).

Table 1.6 Characteristics of human haematopoietic stem and progenitor cells

HSPC population	Immunophenotype*	Potency	Long-term haematopoietic reconstitution
HSC	Lin-CD34+CD38-CD45RA-CD90+	Multi- L,M, MK, E	+
MPP	Lin-CD34+CD38-CD45RA-CD90-	Multi- L,M, MK, E	_
LMPP	Lin-CD34+CD38- CD45RA+	Multi- L, M	_
CMP	Lin-CD34+CD38+CD45RA-CD123lo	Multi- M, MK, E	_
GMP	Lin-CD34+CD38+CD45RA+CD123lo	Bi-M	_
MEP	Lin-CD34+CD38+CD45RA-CD123-	Bi- MK, E	_

** Putative additional markers have been reported but different groups variably use these. E: erythroid; L: lymphoid; Lin: lineage; M: myeloid encompassing granulocyte/monocyte/macrophage lineages; MK: megakaryocyte.*

In the case of erythropoiesis, while there is evidence that HSC, MPP, CMP and MEP (but *not* GMP and LMPP) can differentiate into erythroblasts, the precise point at which exclusive erythroid commitment arises is unknown. Classically MEP have been thought to comprise bipotent cells that have the potential to generate either red blood cells or platelets (Klimchenko et al. 2009), thus erythroid commitment has been ascribed to an 'erythroid progenitor (EP)' stage, distal to MEP.

When human BM is cultured in methylcellulose, 2 distinct erythroid colony types are formed- burst forming unit-erythroid (BFU-E) and colony forming unit-erythroid (CFU-E). BFU-E are large colonies comprising eponymous 'bursts' of hundreds-thousands of cells while CFU-E are smaller uni- or bi-focal colonies made up of tens-a few hundred cells (Gregory & Eaves 1978). These 2 colony types suggest the

existence of at least 2 distinct erythroid progenitor types. Indeed, studies in mouse FL have led to the identification of 2 immunophenotypcially distinct cell populations that enrich for BFU-E and CFU-E colonies, respectively (Flygare et al. 2011). The identification of analogous populations in humans has lagged behind, precluding in depth investigation of both normal and aberrant erythropoiesis.

Erythroid progenitors (EP) differentiate into erythroid precursors (synonymous with the term erythroblasts; EB), which, under physiological conditions, are restricted to the BM. While E**P** cannot be identified morphologically on BM specimens, E**B** encompass several morphologically distinct cell types termed pro-EB, basophilic EB, polychromatic EB and orthochromatic EB (Figure 1:4). The passage of pro-EB through these sequential stages is accompanied by decreasing cell size, increasing nuclear condensation and haemoglobinisation. Subsequent enucleation gives rise to reticulocytes, which are able to exit the BM into the peripheral circulation. Although unable to transcribe mRNA, reticulocytes contain mitochondria and ribosomes and are highly translationally active. In fact, they synthesise up to a third of their final haemoglobin content in the 48 to 72 hour period before they lose ribosomes, organelles, and mRNA, becoming mature RBC (Warren 2017). Under normal conditions, RBC survive in the blood for approximately 120 days.

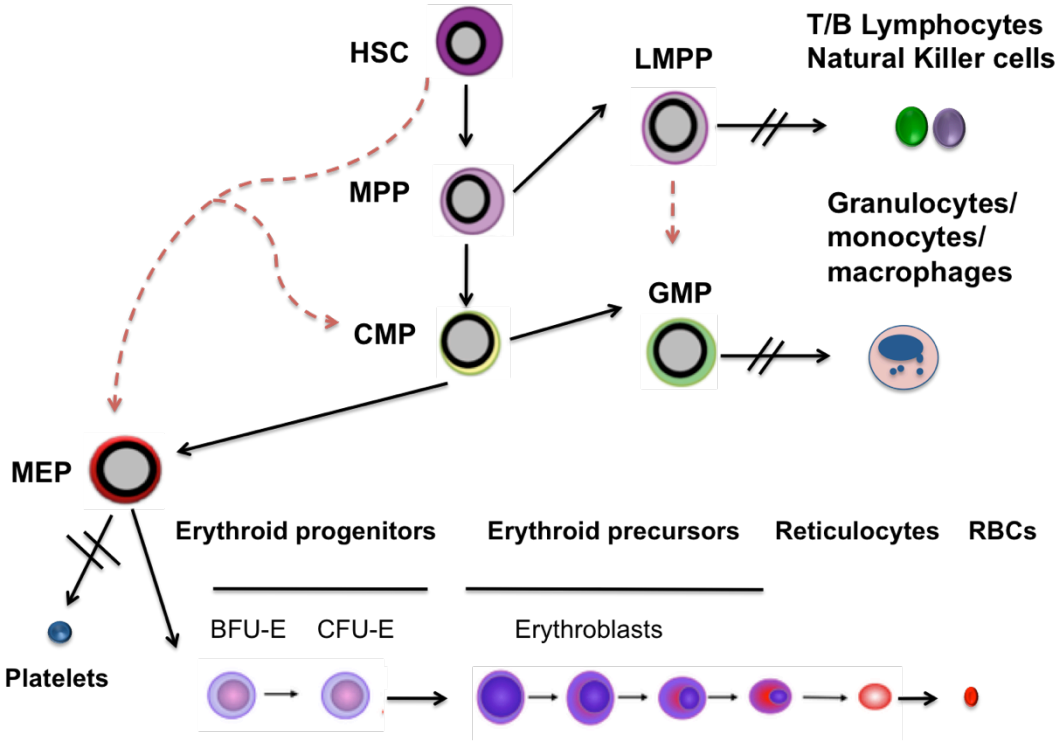


Figure 1:4 Schematic diagram of the haematopoietic and erythropoietic hierarchies (Akashi et al. 2000; Majeti et al. 2007). HSC differentiate through several progenitor stages to produce terminally differentiated cells. Arrows demonstrate the lineage relationships between distinct populations (black arrows: classical model; red arrows: alternative model). For simplicity, some progenitor and precursors in the non-erythroid lineages have been omitted (indicated by interrupted arrows). In erythropoiesis, progenitors pass step-wise through a series of precursor stages in the BM. Enucleation produces reticulocytes, which egress from the BM and mature into RBC that populate the PB (Hattangadi et al. 2011).

1.2.3 Regulation of erythropoiesis at multiple levels

To meet the daily demand for $2X10^{11}$ RBC while keeping production in check, erythropoiesis is tightly regulated by a myriad of cellular and molecular processes that involve an intricate balance between self-renewal/proliferation, differentiation and apoptosis.

1.2.3.1 Growth factors

Growth factors are vital for the regulation of erythroid cell function (Table 1.7). The production of EPO, the major erythroid cytokine, is dynamic and forms part of an autoregulatory feedback loop. Anaemia reduces renal oxygenation, stimulating the transcription factor (TF) Hypoxia-inducible factor (HIF)2α, which increases transcription of genes with hypoxia-response elements, such as EPO, increasing erythroid output (Haase 2013). Once EPO interacts with its receptor (EpoR) on erythroid progenitors, the complex is endocytosed and degraded (Walrafen et al. 2005). In addition, the restoration of adequate renal oxygenation leads to degradation of HIF by the proteasome in renal EPO-producing cells, thus negatively regulating EPO production (Haase 2013). Both these mechanisms serve to prevent RBC overproduction.

EPO has multifaceted roles in promoting erythropoiesis. It has been shown to induce nuclear localisation of p21 (a marker of p53 activation) which allows cell cycle arrest and differentiation of EP (Varricchio & A. R. Migliaccio 2014). EPO signalling also upregulates Bcl-X_L (Dolznig et al. 2002) and downregulates membrane-bound Fas, preventing the induction of apoptosis in developing erythroid cells. Finally, EB differentiation is facilitated by EPO-mediated activation of caspase-3 (Zermati et al. 2001; Boehm et al. 2013) .

Table 1.7 Growth factors and cytokines involved in erythropoiesis*

GF/ cytokine	Site of production	Receptor	Role in erythropoiesis	Other roles in haematopoiesis
SCF	BM stromal cells	c-kit	Critical, non redundant role Proliferation and survival from BFU-E till ≈orthochromatic EB Delays E differentiation Prevents apoptosis Synergises with EPO	Self-renewal and survival of HSC Thrombopoiesis Myelopoiesis
EPO	Renal cortical interstitial cells Hepatocytes (in moderate/severe hypoxia)	EpoR	Critical, non redundant role Proliferation and survival Differentiation Prevents apoptosis	Erythroid-specific
IL-3	T cells, mast cells and eosinophils	IL-3Rα	EP proliferation	Thrombopoiesis Immune function
IL-6	T cells and macrophages	IL-6R	EP proliferation	Thrombopoiesis B-lymphopoiesis
TPO	Liver,	TPOR/c-	EP proliferation	Thrombopoiesis

	kidney, skeletal muscle	MPL		
IGF1/ insulin	Liver/ pancreas	IGF1R/ insulin receptor	EP proliferation Terminal differentiation	Lymphopoiesis
Thyroid hormone	Anterior pituitary gland	TRα and TRβ	EB proliferation Terminal erythroid differentiation and enucleation	B-lymphopoiesis
Oestrogen	Ovaries, testes, liver and adrenals	ER	Suppresses EPO Synergises with GC	Negatively regulates lymphopoiesis
GC	Adrenal cortex	GR	See text	

(Vandekerckhove et al. 2009; Muta et al. 1995; Gao et al. 2017; Huang et al. 2010; Metcalf 2008; Sato et al. 2000; Leberbauer 2005)*.**

EP: erythroid progenitor; EPO: erythropoietin; ER: oestrogen receptor; GC: glucocorticoids; GF: growth factor; IGF-1: insulin growth factor 1; IL-: interleukin; R: receptor; SCF: stem cell factor; TPO: thrombopoietin; TR: thyroid hormone receptor.

1.2.3.2 The erythroblastic island

The role of BM stromal cells in the production of crucial erythroid growth factors, in part explains why, in both FL and adult BM, erythroid development has evolved to occur within a specialised niche, termed the 'erythroblastic island' (Palis 2017). This morphologically identifiable structure comprises a central macrophage surrounded by EB at different stages of maturation (Figure 1:5). The island facilitates cross-talk between adjacent erythroid cells, necessary for cell survival and control (Hanspal

1997). For instance, more mature EB, which express the death ligand FasL, can induce apoptosis of immature Fas-expressing EB, to regulate RBC number (De Maria, Testa, et al. 1999a). SCF or high doses of EPO reduce sensitivity of EB to FasL and limit apoptosis, irrespective of the size of the mature EB pool (Liu et al. 2006). This function is important in the context of stress erythropoiesis, when there is a higher than normal demand for RBC. The erythroblastic island also allows the interaction of macrophages with EB through various cell-cell adhesion molecules, including erythroid macrophage protein 1 (EMP1) on the EB surface (Soni et al. 2006). These interactions are critical for enucleation, after which the extruded nuclei are engulfed and degraded by the central macrophages (Yoshida et al. 2005). Therefore, although erythroid differentiation can occur *in vitro* in the absence of macrophages, the production of enucleated RBC is enhanced by the symbiotic interaction with macrophages that occurs *in vivo* (Rhodes et al. 2008).

Erythropoiesis is also *negatively* regulated in a paracrine fashion by inflammatory cytokines. TNF-α and IL-1 can suppress *EPO* transcription, via NFkB signalling (Frede et al. 1997) and they promote caspase-mediated cleavage of GATA-1 (De Maria, Zeuner, et al. 1999b). IFN-γ interferes at the CMP stage to drive cell fate towards the myeloid lineage, decreasing the numbers of MEP and BFU-E (Libregts et al. 2011). IFN-γ also increases the number of Fas receptors leading to apoptosis of EB (Hom et al. 2015). At the stage of haemoglobin synthesis, inflammatory cytokines can cause functional iron deficiency in EB via secretion of hepcidin (Macdougall & Cooper 2005).

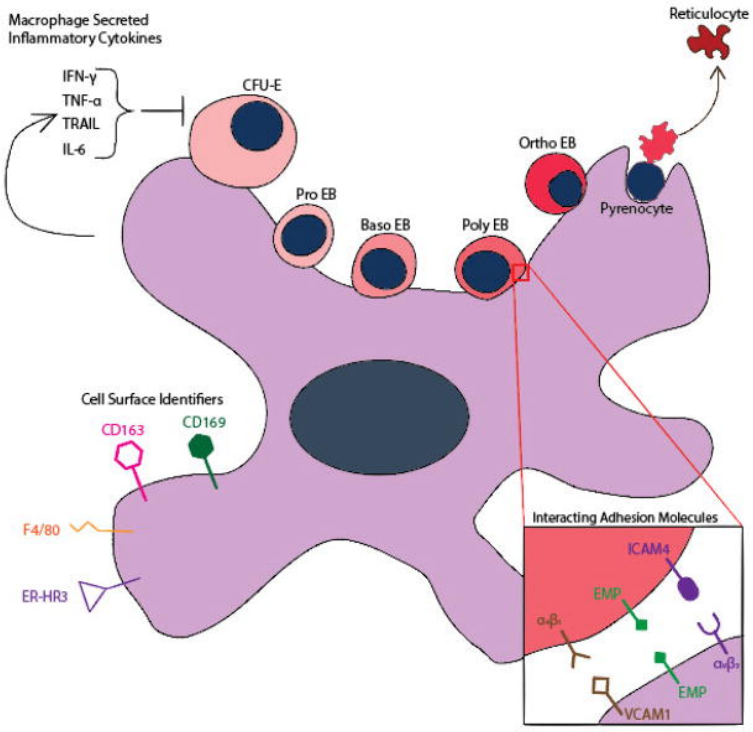


Figure 1:5 The 'erythroblastic island'. Erythroid cell differentiation occurs within a specialised BM niche termed the erythroblastic island. EB adhere to a central macrophage via α4β1, erythroid macrophage protein 1 (EMP1), and Intercellular adhesion molecule 4 (ICAM4). The central macrophage mediates these interactions through VCAM1, EMP and αvβ3 respectively. Other cell surface molecules expressed by the macrophage are CD163 and CD169, F4/80, and ER-HR3. In inflammation, the macrophage secretes cytokines, particularly IFN-γ, TNF-α, TRAIL, and IL-6, which inhibit erythropoiesis. [*Image reproduced, with permission from* (Hom et al. 2015)].

1.2.3.3 Transcriptional and translational control of erythropoiesis

Erythropoiesis is orchestrated by a complex network of transcription factors, including GATA-1, SCL/TAL1, c-myc and KLF1 (which are DNA binding) and LMO2 and LDB1 (which are non DNA binding) (Cantor & Orkin 2002). GATA-1 serves as the master regulator of erythroid differentiation, cell cycle and survival. Importantly, GATA-1 is protected from caspase-3-mediated cleavage and inactivation during erythroid differentiation by its physical interaction with HSP70, a constitutively expressed heat shock protein (Ribeil et al. 2007). Another crucial erythroid-specific

transcription factor, EKLF/KLF1, is expressed from the stage of CMP (Siatecka & Bieker 2011) and promotes erythroid lineage commitment. Overexpression of KLF1 induces greater levels of erythropoiesis than megakaryopoiesis from MEP (Olopade et al. 1992). Conversely, knockdown of *KLF1* favours megakaryocyte formation and prevents the transition of CFU-E to proEB (Frontelo et al. 2007), highlighting the central role of KLF-1 in determining erythroid cell fate.

There is increasing recognition of the role for microRNAs (miRNA) and long non-coding RNAs (lncRNAs) in regulating erythropoiesis. For example, in the human K562 cell line it has been shown that miR-24 suppresses translation and perturbs erythroid differentiation by binding to the 3′ UTR of human activin type I receptor ALK4 (Q. Wang et al. 2008) and in a murine FL cellular model, knockdown of several lncRNAs that are regulated by G,ATA-1, KLF1 and TAL1, impairs erythroid differentiation (Alvarez-Dominguez et al. 2014).

As well as regulation at a transcriptional level, the concerted process of erythroid differentiation is also co-ordinated at a translational level. Most importantly, control of the translation of α- and β-globin mRNAs maintains equimolar ratios of the encoded α- and β-globin subunits of haemoglobin. An example of such translational control is the regulation of eiF2a, an erythroid specific kinase, by heme. When heme is deficient, eIF2aP phosphorylates eukaryotic initiation factor 2 alpha (EIF2a), suppressing protein synthesis hence maintaining a balance between globin and heme (Chen 2014). A recent study employed RNA-seq integrated with Ribo-seq in murine FL to define the translational landscape across different stages of erythropoiesis (Alvarez-Dominguez et al. 2017). The authors demonstrate that translation of *cis*-regulatory elements, known as upstream open reading frames, within the 5′ untranslated region (UTR) of specific mRNAs, can affect downstream translation e.g. inhibiting translation of BCL11A, which is necessary for adult Hb switching and which is downregulated during erythroid differentiation. Furthermore, the action of *trans*-acting lineage-specific RNA-binding proteins, such

as RBM38, modulates the translation efficiency of various mRNAs in a dynamic fashion throughout erythroid differentiation. For example, although β-globin mRNA is induced at 24 hours of murine FL *in vitro* erythroid differentiation, translation is fully established only at 33 hours (Alvarez-Dominguez et al. 2017).

Finally, autophagy is increasingly being appreciated as key catabolic pathway involved in erythroid development. Specifically, during reticulocyte differentiation autophagy is needed to degrade mitochondria and ribosomes, allowing RBC to form. Accordingly, mice knocked out for genes encoding essential components of the autophagy pathway (e.g. Atg7) develop anaemia (J. Zhang et al. 2009).

1.3 Disease mechanisms in DBA and the therapeutic effect of glucocorticoids

1.3.1 Model systems in DBA

Given the obvious challenges in studying EP derived from patients with DBA, model systems have been used to gain insight into disease pathogenesis. In mice, complete loss of *RPS19* by doxycycline-induced small interfering (si)-RNA or by genetic knockout results in early embryonic lethality. RPS19 $^{+/-}$ mice although viable, lack haematopoietic defects (Matsson et al. 2004). Surprisingly *Rps19* mRNA and protein levels from spleens of control and mutant animals are comparable therefore the authors propose that monoallelic loss of *RPS19* is fully compensated for at the mRNA level with preservation of erythropoiesis. It is not clear why such compensatory mechanisms fail to occur in patients with DBA. Transgenic mice with inducible expression of an *RPS19* mutant have also been generated (Devlin et al. 2010), although again these exhibit a rather mild haematological phenotype. Conditional knockout mouse models have been used to study the tissue-specific effects of RP haploinsufficiency at distinct stages of development (McGowan et al. 2008).

Humanised mouse models have been of limited use in DBA as human EP do not reliably engraft and consequently, few human RBC are identifiable in the PB of these mice. Finally, murine FL cells *ex vivo* have been utilised for cellular models of DBA (Flygare et al. 2011).

Zebrafish models of DBA also exist, initially generated with antisense oligonucleotide morpholinos; the advantage of this approach is that the morpholino dose can be titrated to achieve 50% mRNA reduction and RP haploinsufficiency, thus mimicking human disease. Additionally, it is possible to image embryonic development *in vivo*. One such zebrafish model of *RPS19*-deficiency displays defective erythropoiesis, congenital anomalies and cancer predisposition, akin to the abnormalities seen in patients with DBA. While the latter is rescued by genetic knockout/pharmacological inhibition of *p53*, impaired erythropoiesis persists to some degree (Danilova et al. 2008). More recently, gene editing tools such as transcription activator-like effector nucleases (TALEN) and Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR)- Cas9, offer the prospect of generating elegant animal models in which a DBA-causing mutation can be induced and corrected in specific cell types.

An advantage of the animal model systems above is that they overcome the inherent challenge of working with limited sources of primary cells from patients who exhibit clinical and genetic heterogeneity. The disadvantage is that they do not faithfully recapitulate human disease and there are interspecies differences, as specifically demonstrated in the case of transcriptional and epigenetic events underlying murine and human haematopoiesis (Pishesha et al. 2014).

Finally, some groups have studied DBA using human cellular models, such as *in vitro* culture of human CD34+ progenitors, in which short hairpin (sh) or siRNA are employed to knockdown a particular RP (Dutt et al. 2011; Ebert 2005), then lentiviral vectors are employed for ectopic expression and correction of the disease phenotype. Limitations of this approach include imprecise levels of the

haploinsufficient protein and technical difficulty in expanding, and especially enucleating, these cells in culture. DBA patient cells have occasionally been studied *ex vivo* (Gazda et al. 2006) or *in vitro* (Ohene-Abuakwa et al. 2005). Alternatively they have been used to generate lymphoblastoid cell lines (Aspesi et al. 2017) or induced pluripotent stem cells (iPSCs) (Garçon et al. 2013; Ge et al. 2015), providing an unlimited source of cells for genetic manipulation and pharmacological studies. For example, work in iPSCs has recently led to the identification of SMER28, a small-molecule inducer of autophagy, as a potential therapeutic agent in DBA (Doulatov et al. 2017).

1.3.2 Cellular pathogenesis of erythroid failure in DBA

Historical data suggest that impaired haematopoiesis in DBA may not be confined to the erythroid lineage. Although long-term culture initiating cells (LTC-IC), an *in vitro* measure of HSC number, are similar in frequency in normal and DBA BM, the average clonogenic efficiency per LTC-IC is significantly lower in DBA patients (Giri et al. 2000). It has also been shown that the myeloid output from the CD34+ fraction of DBA BM is diminished in some patients (Santucci et al. 1999). Finally Hamaguchi et al. showed that in RPS19+/- patients, proliferation of the multipotent CD34+CD38- BM compartment improved upon enforced expression of *RPS19* with lentiviral vectors (Hamaguchi 2003). Despite the above evidence, the erythroid defect is the most severe and the most consistent haematopoietic abnormality in DBA.

Plating total BM in methylcellulose has demonstrated absent or reduced growth of BFU-E in patients with newly diagnosed or transfusion/corticosteroid- treated DBA (Nathan, Clarke, et al. 1978a; Bagnara et al. 1991; Ramenghi et al. 1999; Alter et al. 1992). In addition to reduced erythroid progenitor frequencies, impaired differentiation of BFU-E and CFU-E has been demonstrated *in vitro* (Lipton et al. 1986) (Ohene-Abuakwa et al. 2005). This defect has been attributed to insensitivity to EPO (Halperin et al. 1989) although to date, no abnormalities in the EPOR or its

signalling pathway have been identified. Colony number and/or size were improved by addition of SCF (Bagnara et al. 1991; Alter et al. 1992) or IL-3 (Halperin et al. 1989), but this *in vitro* effect did not correlate with a therapeutic effect *in vivo*. Similarly, *in vitro* cultures showed variable responses to prednisolone or dexamethasone (DEX) that did not reliably correlate with patient responsiveness to CS (Chan, Saunders & Freedman 1982a; Chan, Saunders & Freedman 1982b; Ohene-Abuakwa et al. 2005). Of note, some patients have been shown to have normal progenitor frequencies (using their functional colony output as a surrogate marker) (Freedman et al. 1976; Ohene-Abuakwa et al. 2005). The determinants of these divergent haematological phenotypes and the reason for erythroid failure arising at a specific point in the hierarchy (Figure 1:4), have not yet been explained.

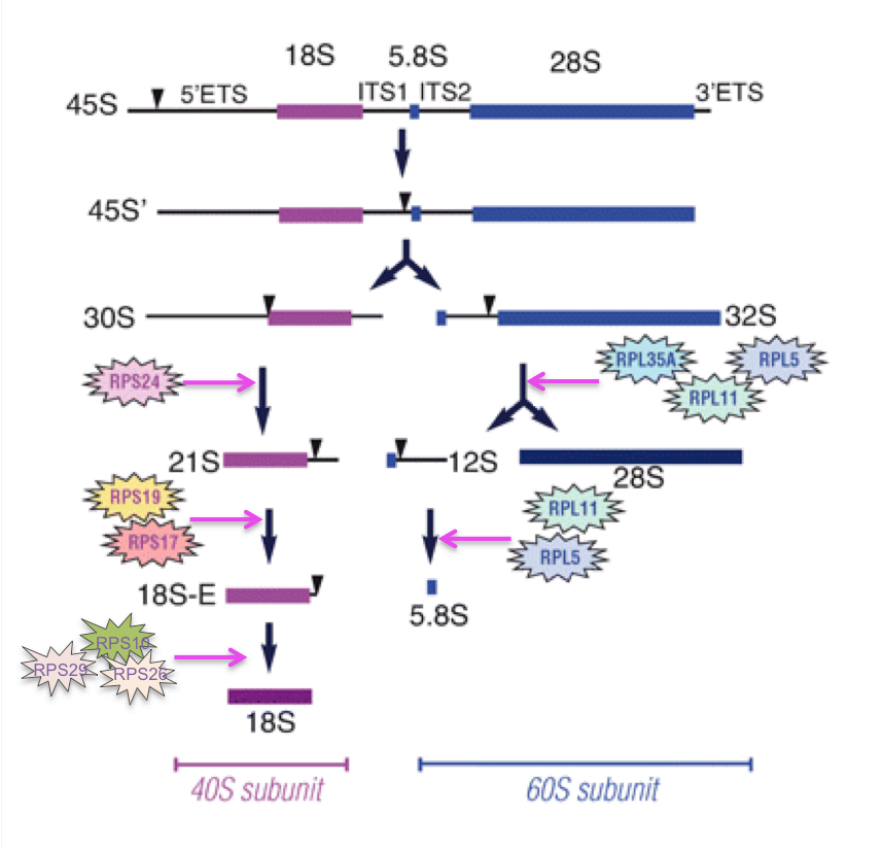
Historically it was postulated that the cellular mechanism for the erythroid defect in DBA was autoimmune mediated destruction of EP (Hoffman et al. 1976) (Steinberg-Shemer et al. 2016). However, subsequent evidence challenged this by demonstrating that DBA T cells do not inhibit normal erythropoiesis (Nathan, Hillman, et al. 1978b) and conversely, normal T cells cannot improve DBA clonogenic activity (Freedman & Saunders 1978). Latterly, apoptosis of mutated EP has been asserted as the main mechanism underpinning anaemia. This was initially based upon the observation that, compared with normal controls, DBA BM deprived of EPO *in vitro* generated fewer BFU-E and CFU-E colonies, and that these showed an accelerated increase in DNA fragmentation compatible with apoptosis (Perdahl et al. 1994). Notably, there is no evidence for EPO deprivation *in vivo*- in fact EPO levels are high- so in the experiments above, increased apoptosis should have arisen even in EPO replete conditions. Still, it remains possible that EPO insensitivity is implicated in the DBA erythroid defect and that this culminates in apoptosis or cell cycle arrest. In line with this notion, there is ample evidence that DBA pathogenesis involves p53, a TF that regulates genes involved in apoptosis, cell cycle and senescence. Immunohistochemistry for p53 protein in sections of DBA BM showed a tendency towards increased nuclear staining in erythroid cells, though findings were

inconclusive given the low number of biological replicates (Dutt et al. 2011). Several models of RP deficiency (1.3.1) have confirmed p53-mediated haematopoietic and developmental defects with reversal of the DBA phenotype by abrogation of p53 (Jaako et al. 2011; Danilova et al. 2008; Dutt et al. 2011). It has also been shown that cells of the erythroid lineage are particularly sensitive to *p53* up-regulation and that activation of p53 induces premature erythroid maturation (Dutt et al. 2011). Studies have also revealed p53-independent mechanisms (Singh et al. 2014; Danilova et al. 2008) for cell cycle arrest and induction of apoptosis.

1.3.3 The link between defective ribosomes and erythroid failure

The precise mechanisms by which RP gene mutations cause DBA remain unclear and will be explored in this section. The major role of RP is ribosome assembly and protein synthesis. After their folding in the endoplasmic reticulum, RP migrate to the nucleolus, where they form the ribosomal subunits in association with ribosomal RNA (rRNA). The small subunit is termed 40S and the large one 60S; together they form a ribosome that is 20nm in size. As well as being structural components of the mature ribosomal subunits, RP are also required in the nucleolus for processing of rRNA. The steps involved in ribosome biogenesis include 1) rRNA transcription using RNA polymerase I and III, 2) processing by cleavage of rRNA transcripts 3) structural modifications of both rRNA and RP molecules, 4) assembly of rRNA and select RPs in the nucleolus 5) re-export into the cytoplasm for addition of more RPs to form ribosomal subunits for mRNA translation. Depletion of particular RP therefore causes a defect in the maturation of specific precursor rRNA molecules (Figure 1:6). These abnormalities are detected by quantifying rRNA transcripts, either by northern blotting or capillary electrophoresis (Quarello et al. 2016). Polysome profiling using a sucrose gradient can also be informative; small (RPS) or large (RPL) ribosomal subunit defects are reflected by a lower amplitude of the 40S and 60S peaks, respectively (Horos & M Lindern 2012; Moore et al. 2010). Both approaches assist in the diagnosis of DBA by confirming the presence of an rRNA processing defect and also by helping to assign pathogenicity to novel RP gene variants.

To initiate mRNA translation, simultaneous translocation of mRNA and tRNA through the ribosome is facilitated by RP. The RPS is involved in decoding mRNA using the adapter tRNA while the RPL is required for peptide synthesis and assembly of the polypeptide.



Mutated/deleted RP gene	Expected rRNA profile	
RPS19/S17	↑21S↓18S	High 28s: 18s ratio relative to controls (≈2.1)
RPS24	↑30S↓21S	
RPS26/S10/S29	↑18S-E↓18S	
RPL5/L11	↑28S↓5.8S	Low 28s: 18s ratio relative to controls (≈1.3)
RPL5/L11/L35a/L31/L27	↑32s↓28s	

Figure 1:6 Ribosomal structure and rRNA processing defects in the presence of distinct RP gene mutations/deletions (Juli et al. 2016; Quarello et al. 2010; Danilova & Gazda 2015). 18S, 5.8S and 28S rRNAs are transcribed by PolI in the

nucleolus as segments of a long precursor pre-rRNA, which are flanked by two externally transcribed spacers 5'ETS and 3'ETS and separated by two internally transcribed spacers, ITS1 and ITS2. 5S rRNA (not shown) is transcribed independently by PolIII in the nucleus. The pre-rRNA assembles with accessory factors and a subset of RPs, forming a secondary structure necessary for the correct folding, modification and cleavage of pre-rRNA. Arrows above pre-rRNA indicate cleavage sites. The main cleavage pathway involves removal of the 5'ETS and cleavage in the ITS1 site to give rise to the 21S and 32S precursors to the 40S and 60S subunits. 5S rRNA incorporates into pre-60S subunit. Subunits are then exported to the cytoplasm. Once in the cytoplasm, small and large subunits undergo removal of remaining accessory factors and incorporation of missing RPs. A functional ribosome forms after transcribed mRNA binds to the 40S subunit, triggering conjoining with the 60S subunit. Over 200 accessory factors, including helicases, nucleases, small nucleolar RNAs, chaperones and transporters, associate with the maturing ribosomal subunits at various steps. Each stage of rRNA processing depends on particular RPs, depicted by star shapes and pink arrows. Mutations affecting an RP gene will therefore interrupt a specific step in ribosomal biogenesis resulting in a specific rRNA profile as shown in the table below the diagram. *[Image adapted from* (Dianzani & Loreni 2008)*,* with permission from http://www.haematologica.org*]*

How impairment in ribosome function, a global process, leads to the relatively selective phenotypes in DBA is not clear. Mutations in at least 15 RP genes underpin the Minute phenotype of *Drosophila melanogaster*, characterised by delayed larval development, thin bristles and small body size (Kongsuwan et al. 1985; Marygold et al. 2007). This phenotype is postulated to result from insufficient protein synthesis during particular developmental stages that require high cell growth. Commensurate with this theory, the synthesis of Hb in erythropoiesis imposes a huge demand on the cellular translational apparatus. A putative mechanism for the demise of EB in DBA is inadequate globin synthesis leading to accumulation of toxic free heme

(Figure 1.7). Indeed, Feline leukaemia virus causes a DBA-like phenotype in cats by impairing function of its receptor, the heme exporter, Flvcr. Conditional knockout of *Flvcr* in adult mice phenocopies DBA with macrocytic anaemia and differentiation arrest at the proEB stage (McGowan & Mason 2011). Finally, there is some evidence that alternatively spliced forms of *FLVCR1*, which disrupt FLVCR1 protein expression and function, are more common in the CD71+ BM fraction of patients with DBA relative to normal controls (Rey et al. 2008).

The suggestion that DBA results from a general reduction in translation is also supported by the finding that lymphocytes from DBA patients have reduced translation activity (Cmejlova et al. 2006), coupled with the recent recognition that low T/B cells and hypogammaglobulinaemia are a feature of DBA in a proportion of patients (Iskander 2012). Another possibility is that RP defects impair the translation of specific mRNAs exerting cell- and genotype-specific disease phenotypes. Recent evidence for ribosome specialisation, whereby the constituents of a ribosome impact upon its translational activity, provides support for this hypothesis. Indeed, yeast ribosomes deficient in RPS26 have recently been shown to preferentially translate mRNAs from stress response pathways (Ferretti et al. 2017).

It has also been suggested that the pathogenesis of DBA might be linked to the fact that RP have a multitude of extra-ribosomal functions, as described in *E. coli*, *Drosophila melanogaster* and mammals. In the case of RPS19, protein dimers are chemotactic for human monocytes (Shibuya et al. 2001) and free cytoplasmic RPS19 interacts with fibroblast grow factor 2 (FGF2) (Soulet et al. 2001), and with the serine/ threonine kinase Pim-1, which arises downstream of EPO signalling (Chiocchetti et al. 2005). Perturbation of RP levels may therefore affect cellular functions directly, independent of secondary effects on ribosome function and translation. RPL11 and RPL5 regulate p53 through an inhibitory physical association with mouse double minute 2 (MDM2). MDM2 binds to and ubiquitinates p53, promoting its proteasomal degradation. Failure to synthesise complete ribosomes as

a result of RP haploinsufficiency is postulated to lead to accumulation of free RPL5 and RPL11, their leakage from the nucleolus, sequestration of MDM2, p53 activation and apoptosis (Figure 1:7) (Y. Zhang et al. 2003; Dai & Lu 2004). A similar mechanism is predicted to exist in human cells, involving RPL5/11 and HDM2 (Fumagalli et al. 2009). Although not yet proven, haploinsufficiency has been shown to cause rRNA processing defects and upregulation of p53 target genes (Danilova et al. 2008; Dutt et al. 2011).

Aside from their involvement in the MDM2-p53 axis, RPL11 and RPL5 have recently been shown to have a non-redundant role in the maintenance of nucleosomal structure, while other RPs are not required for the maintenance of nucleolar integrity (Nicolas et al. 2016). The biological and clinical significance of this variable need for RP in controlling nucleolar structure, is yet to elucidated; given that alterations in nucleolar size and shape are a feature of cancers (Quin et al. 2014) one could speculate that the nucleolar destabilisation induced by RPL5/11 haploinsufficiency might somehow interfere with cellular proliferation, representing another putative pre-translational pathogenic mechanism in DBA.

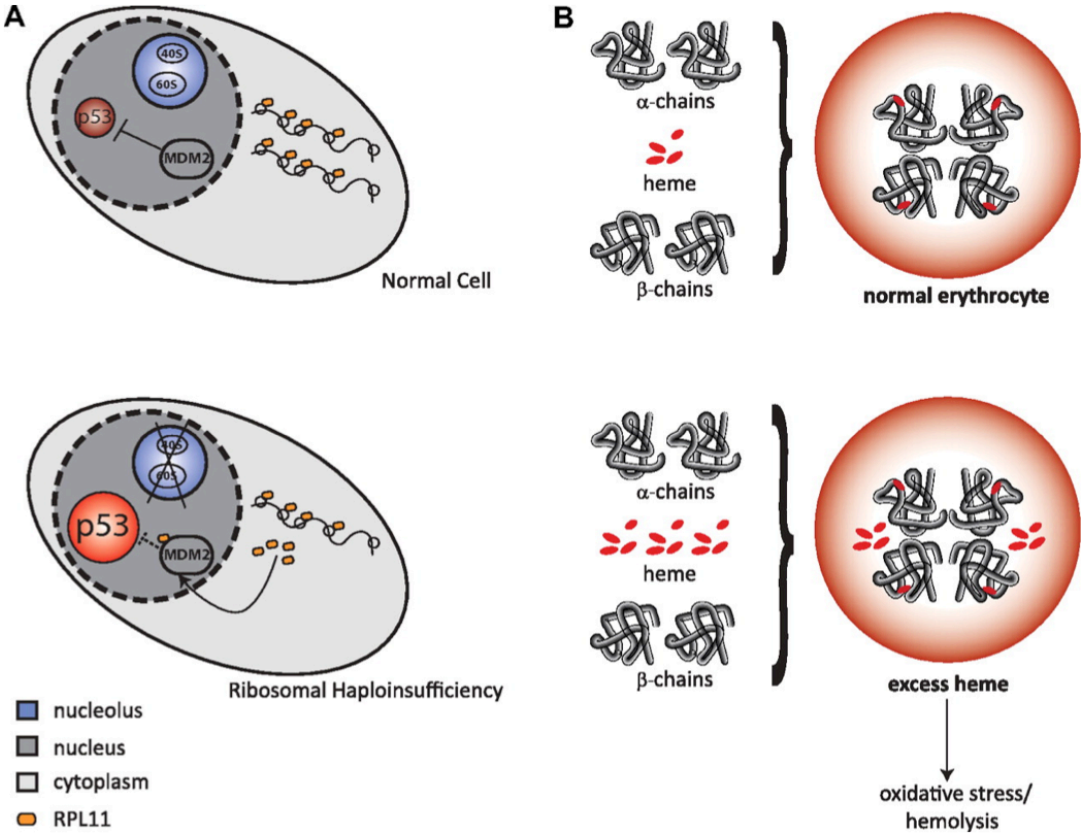


Figure 1:7 Possible mechanisms for erythroid tropism in DBA. **A)** In normal cells, where ribosome assembly is intact, MDM2 interacts with p53 preventing its activation. In RP haploinsufficient cells, impaired ribosome synthesis allows free RPL11 or RPL5 to exit the nucleolus and bind to MDM2, preventing its inhibitory effect on p53. **B)** In a normal RBC/erythrocyte, heme is sequestered within the haemoglobin molecule. RP deficiency reduces translation of globin chains so heme is in excess, which in turn can promote oxidative stress and apoptosis of EP. [*Image reproduced, with permission from* (Narla & Ebert 2010)*]*

1.3.4 Therapeutic effects of glucocorticoids in DBA

Glucocorticoids (GCs) are lipophilic hormones that regulate various physiological responses and developmental processes by binding to and modulating the transcriptional activity of their cognate intracellular GC receptor (GR). GR is encoded by the *NR3C1* gene: in humans alternative splicing of *NR3C1* leads to the production of 5 isoforms, GRα, GRβ, GRγ, GRP, and GRA. The roles of GRγ, GRP, and GRA are not

clearly established. GRα and GRβ are produced by alternative splicing of exon 9 and are highly homologous. While GRα, which exists as a monomer in the cytoplasm of the cell, binds GC and has a crucial role in transcriptional regulation, GRβ is a dominant negative isoform. It exhibits poor GC binding and, by forming heterodimers with GRα, impairs the ability of the latter to bind GC (Varricchio & A. R. Migliaccio 2014). It is not known which isoforms prevail in cells of the erythroid lineage nor whether their expression and distribution vary in FL, UCB and BM. Upon ligand binding, GRα (henceforth called GR) is activated and, on dissociating from its complex with HSP90, translocates from the cytoplasm to the nucleus. Here it forms homodimers that bind to specific *cis* DNA motifs (glucocorticoid responsive elements- GRE) in promoters and enhancers of target genes (Biddie & Hager 2009) (George et al. 2009) (Figure 1:8). In general this activates target gene transcription although repression can also occur as a result of direct GR binding (Lim et al. 2015). There are thousands of GRE in the genome but in any cell type only a small minority are bound by the GR (So et al. 2007). Binding of the GR to DNA is transient and dynamic, and is often pre-determined by pioneer transcription factors (TF) that ensure accessibility of the GR to regulatory areas (Biddie et al. 2011). Therefore GR can bind to DNA either directly (as a dimer) or via other TF (as a monomer), such as AP-1 (Biddie et al. 2011) or GATA-1 (Chang et al. 1993), activating or repressing transcription. Transcriptional activation or repression is also determined by further interaction of the GR with histone modifying complexes, e.g. histone acetyltransferases and deacetylases and chromatin remodellers, such as SWI/SNF (John et al. 2011) (So et al. 2007). Therefore, higher order chromatin structure and chromatin accessibility have a profound effect on the cell- and gene-specific transcriptional effects of GC.

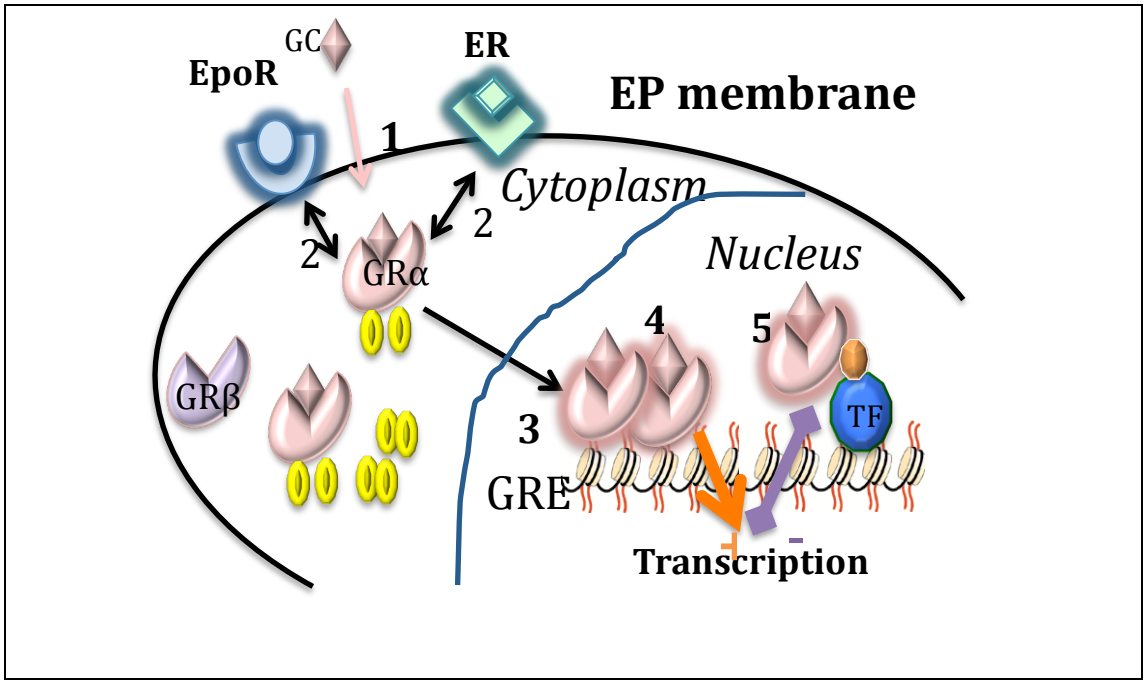


Figure 1:8 The cellular response to GR activation. 1. GC traverse the erythroid progenitor membrane by passive diffusion 2. Upon ligand binding to cytoplasmic GRα, the receptor is activated and interacts with neighbouring receptors via shared molecules, such as STA5, within their respective signalling pathways (Stellacci et al. 2009). 3. Dissociation from its chaperone proteins HSP90, allows GR nuclear translocation. 4. Here it dimerises and binds to genes via their GC response elements (GRE), bringing about transcription or transrepression. 5. GR can also bind indirectly to DNA via TF and epigenetic regulators. *[Adapted from*(Zhou & Cidlowski 2005)*]*

Although GC are the single most effective group of drugs for the treatment of DBA, there is limited knowledge of their mechanism of action in erythropoiesis. The first important consideration is that GC have a role in physiological erythropoiesis as highlighted by the clinical observation that patients with corticosteroid deficiency due to Addison's disease develop chronic anaemia while those with GC overproduction (Cushing's disease) exhibit erythrocytosis. GC are known to be critical for the physiological response to stress erythropoiesis (Bauer et al. 1999).

Much of the existing knowledge about the action of GC has been derived from murine studies. GC have been shown to influence erythropoiesis in a cell autonomous manner, i.e. they target EP directly, by enhancing their proliferation (Bauer et al. 1999; Kolbus 2003) and blocking their maturation in response to EPO and SCF (Dolznig et al. 2006). Mice expressing a GR modified to trans-repress but not trans-activate, fail to increase erythroid clonogenic activity and Hb levels in response to stress stimuli (Bauer et al. 1999), suggesting that transcriptional activation is a prerequisite to enhanced BFU-E self-renewal. Indeed, recent work involving chromatin immunoprecipitation (ChIP)-sequencing, shows that GC activate a transcriptional programme that enhances proliferation of murine FL EP by synergising with both HIF1α (Flygare et al. 2011) and peroxisome proliferator-activated receptor (PPARα; (Lee et al. 2015). The latter discovery proved particularly tantalising as it suggested a role for PPARα agonists in DBA. Using microarray expression profiling coupled with ChIP-seq, it has also been shown that the GC effect in fetal erythropoiesis is mediated by an RNA-binding protein, zinc-finger 36 protein-like 2 (Zfp36l2; (L. Zhang et al. 2013). On transcriptional upregulation by GC, Zfp36l2, by binding to the ATTTA 3' UTR motifs of target genes, downregulates expression of genes essential for erythroid differentiation. The relevance of this finding in DBA has not yet been investigated, however one can speculate that if GC delay terminal differentiation of EP *in vivo,* this could lead to an expansion of the erythroid progenitor pool and in turn, an increase in the overall erythroid output and an increase in Hb.

The ability of GC to maintain EP in proliferation has been exploited to develop liquid cultures that permit massive amplification of primary human EB derived from PB, BM or UCB MNCs (human erythroid mass amplification: HEMA) (G. Migliaccio et al. 2011; Leberbauer 2005; M von Lindern et al. 1999). Both normal (M von Lindern et al. 1999) and RP-deficient CD34+ HSPC (Ebert 2005) display enhanced clonogenicity and delayed differentiation *in vitro* when treated with the synthetic GC dexamethasone (DEX).

The proliferative effect has been presumed to be due to the direct action of GC on CD34+ erythroid progenitors, as already demonstrated in mice. Gene expression profiling has been used to study the transcriptional impact of GC on *RPS19* -deficient CD34+ HSPC. Importantly, GC do not restore expression in EP of the RP gene mutated in DBA (Ebert 2005) suggesting that correction of anaemia in response to GC in DBA is independent of RP gene expression. Importantly, gene expression was assessed at 18 hours post-exposure to GC, therefore it is likely that several of the observed changes reflected secondary transcriptional events.

Finally, it remains unclear why patients with DBA (even family members with the same mutation) exhibit varying responses to GC. In other disease contexts, such as inflammatory bowel disease (Iudicibus 2011) and acute lymphoblastic leukaemia (ALL) (Kaspers et al. 2009), resistance to CS therapy has been attributed to multiple genetic polymorphisms in the GR or in other allied molecules involved in the GC response, or epigenetic modifications such as diminished somatic methylation of the *caspase-1* promoter, leading to caspase-mediated degradation of the GR (Paugh et al. 2015). Thus far there is no clear-cut evidence for GR polymorphisms (Varricchio et al. 2011; Varricchio & A. R. Migliaccio 2014) or any other mechanisms leading to steroid resistance in DBA.

In aggregate, the full cellular and transcriptional programme(s) modulated by GC in human EB/EB have not yet been defined. Delineation of the mechanisms that promote erythropoiesis in normal progenitors is likely to be important for understanding the therapeutic effects of GC in DBA and the factors underpinning CS refractoriness. In turn, this could facilitate the development of novel therapies that directly target erythropoiesis, show efficacy in steroid-resistant patients and/or to synergise with GC, mitigating GC-related side effects.

1.3.5 Mechanism-based experimental therapies in DBA

Historically various treatments have been tried in DBA. Despite the absence of evidence that DBA is associated with a deficiency of the haematopoietic growth factors that promote erythropoiesis (Bagnara et al. 1991; Draptchinskaia et al. 1999), recombinant EPO, IL-3, SCF and TPO have been tried therapeutically in patients. EPO does not appear to be beneficial; in one series, for example, no reticulocyte or haemoglobin responses were observed in 9 patients treated with doses as high as 2000 U/kg per day (Niemeyer et al. 1991). By contrast, 10-20% of patients appear to respond to rhIL-3, as evidenced by three studies (Ball et al. 1995; Olivieri et al. 1994; Gillio et al. 1993), however two patients developed deep vein thromboses necessitating discontinuation of treatment, (Olivieri et al. 1994). To date there is one case report describing an adult with DBA with anaemia responsive to eltrombopag, a TPO receptor agonist (Winkler et al. 2016), however larger trials are pending.

Other experimental therapies have included valproate (Jabr & Taher 2006), the prolactin agonist metoclopramide (Abkowitz et al. 2002; Leblanc et al. 2007) and ciclosporin (Anur et al. 2009; Alessandri et al. 2000; Jabr & Taher 2006), however all were inferior to CS and their use was limited by side effects. Great excitement was sparked by recent experimental evidence that, through activation of the rapamycin-sensitive mTOR/RPS6 pathway, the amino acid L-leucine could enhance ribosomal translational activity in RP haploinsufficient cells (Boultwood et al. 2013), including DBA cells with *RPS19* mutations (Jaako et al. 2012). Moreover, L-leucine was shown to rescue both the haematological and developmental DBA phenotypes in a zebrafish RPS19+/- model (Payne et al. 2012). In a subsequent clinical trial, clinical remissions were induced by L-Leucine in some CS-resistant patients though the response rate was disappointing at <10% (4 in 46 patients; *Vlachos A, unpublished*). Still, L-Leucine is a well-tolerated agent with minimal side effects and subjective beneficial effects on quality of life (Pospisilova et al. 2012) therefore arguably it should be tried in TD patients. Finally, a clinical trial is underway in TD adults with DBA for sotatercept, a recombinant, chimeric protein consisting of the extracellular

domain of human activin receptor type IIa and the Fc domain of human IgG1. By binding activin A and other members of the transforming growth factor-β superfamily of cytokines, it inhibits their function. This agent has been shown to be effective in several anaemia models (Suragani et al. 2014), including an RPL11-deficient zebrafish (Ear et al. 2015). Interim results are awaited. Overall, experimental therapies have yielded variable results and their use has been limited by responses in only a small proportion of patients and an inability to predict responders *a priori*.

Identification of RP haploinsufficiency as the underlying defect in DBA suggested that increased expression of the corresponding RP gene had the potential to cure DBA, paving the way for gene therapy. Hamaguchi et al. have shown that forced expression of *RPS19* in CD34+ cells derived from *RPS19*-mutated patients, using oncoretroviral vectors, significantly ameliorated erythroid colony formation *in vitro* (Hamaguchi et al. 2002). Furthermore, $CD34^{+}$ cells from an RPS19-deficient DBA patient transduced with *RPS19* had a proliferative advantage over nontransduced RPS19-deficient cells after transplantation into immunodeficient mice (Flygare 2005) and could rescue BMF in mice *in vivo* (Debnath et al. 2017). With the development of novel gene editing technologies, namely CRISPR-Cas9, gene therapy has become an even more tangible future prospect for the treatment of DBA, although the large number of distinct mutations in different genes poses a logistical hurdle.

1.4 Concluding remarks

Despite recognition of DBA 70 years ago, many aspects of this disease are only beginning to be understood. While the advent of new technologies, such as NGS, has improved the sensitivity and accuracy of genetic diagnosis, approximately 30% of cases remain genetically undiagnosed. Also unknown is the molecular basis for the clinical heterogeneity observed in this disease, including the remarkably variable penetrance, even within families. Although patient registries have provided novel

insights into the epidemiology and biology of the disease, to date no robust associations have been made between the underlying RP gene mutation and the haematological phenotype or clinical outcome. The unifying mechanism of erythroid failure that links distinct mutations is yet to be elucidated, as are the cellular and molecular mechanisms by which erythropoiesis fails in DBA. Finally, although CS remain the cornerstone of DBA treatment, very little is known about how they stimulate human erythropoiesis. Understanding the clinical, cellular and molecular determinants of steroid-responsiveness versus refractoriness are an important pre-requisite for the development of novel therapies.

1.5 Aims of this work

The overall aim of this project is to characterise the cellular and molecular aspects of the erythroid defect in DBA, building upon the existing knowledge summarised above. The specific aims are to:

- Evaluate the clinical and genetic spectrum of DBA in a large cohort of patients in the UK
- Physically isolate and characterise erythroid progenitors (EP) enriched in BFU-E and CFU-E activity in human BM
- Dissect the phenotypic and functional cellular architecture of early haematopoiesis and erythropoiesis in DBA using the findings from aim 2
- Elucidate potential molecular mechanisms for any abnormalities identified in aim 3, using transcriptome profiling of purified erythroid populations
- Determine the cellular and molecular effects of glucocorticoids in normal and DBA erythropoiesis, focusing on patients who are corticosteroid-responsive

2 MATERIALS AND METHODS

2.1 Human samples

2.1.1 Study design

The National Research Ethics Service (ref 12/LO/0426) approved this study prior to its commencement. Samples were collected following written informed consent from participants or their parents/guardians where appropriate. An example of one of the consent forms used is shown in Appendix A.

Bone marrow (BM) and peripheral blood (PB) samples were obtained from individuals attending the Paediatric or adult haematology services at Imperial College Healthcare Trust (ICHT), London. Ages of participants ranged from 9 months to 43yrs therefore, where possible, control and disease samples were age- (and sex-) matched in each experiment.

Healthy control BM was derived from sibling or parental donors of children undergoing HSCT for red cell disorders, such as sickle cell disease or thalassemia. Many of the donors themselves, while healthy, were carriers of alpha/beta thalassaemia or sickle cell trait. These samples were still utilised given the difficulty in acquiring paediatric BM. However, where possible, at least one true haematologically normal individual was included as a biological replicate in the control group in each experiment.

The indications for BM aspiration/biopsy in patients with DBA included i) confirmation of diagnosis, ii) investigation of new cytopenias iii) autologous BM harvesting prior to HSCT or iv) 5-yearly disease surveillance to monitor cellularity. For all patients enrolled in the study, the diagnosis of DBA was verified by consulting historical records and by performing family interviews. Targeted NGS was used to screen for RP gene and *GATA-1* mutations, as previously described (Gerrard et al. 2013).

2.1.2 Sample Collection

1-10mls of BM aspirate (BMA) were collected into a syringe using aseptic technique. To collect BM plasma, 1-2mls of BMA was placed in an EDTA tube (BD) and processed 30 minutes (and no later than 4 hours) after collection, as described below. To collect BM cells, 1-10mls of BMA was placed in 2ml RPMI solution, supplemented with 1% P/S, 1% L-glutamine and 20iU/ml preservative-free heparin. Samples were processed within 48 hours from collection, as described below. All samples were anonymised on receipt and assigned a unique identification (ID) number.

2.1.3 Umbilical cord blood

10-20ml umbilical cord blood (UCB) was aspirated from the umbilical vein as soon as possible after cutting the umbilical cord, usually after elective caesarean section, using 21G needles and immediately transferred to a 50ml Falcon (Becton-Dickinson, UK) containing preservative-free heparin at a concentration of 10-50 IU/ml (CP Pharmaceuticals, UK). All pregnant women gave written informed consent for the collection and use of UCB. Samples were anonymised before research use.

2.2 Cellular assays

2.2.1 Isolation of BM plasma

1-2mls BMA was collected in a BD tube containing EDTA anticoagulant. The tube was then spun at 1500g at 4^0C for 12 minutes. The upper plasma layer was aspirated, subjected to a repeat spin at 4^0C, and then stored immediately in 75-250µl aliquots at -80^0C.

2.2.2 Isolation of mononuclear cells (MNCs)

Mononuclear cells (MNCs) were isolated by Ficoll-Hypaque (Sigma-Aldrich, Dorset, UK) density centrifugation as per the manufacturer's instructions. Briefly, BMA was

passed through a 40µm cell strainer (Becton-Dickinson, Oxford, UK) into a 50 ml Falcon tube (Becton Dickinson), using the blunt end of a plastic syringe plunger as a pestle to get a single-cell suspension. This was diluted with Phosphate Buffer Saline (PBS) at a sample: PBS ratio of 1:3 (vol/vol). Diluted BMA was then layered over 1.077 Lymphoprep (Sigma-Aldrich, Dorset, UK), which had been left to equilibrate to RT prior to use at a sample: Lymphoprep ratio of 1:1 or 2:1 (vol/vol). This was immediately spun at 400g for 30 minutes at RT, with deceleration set at 0. After removal of the upper BM plasma layer, the interface MNC layer was aspirated, washed in PBS at 500g for 15 minutes, resuspended in PBS and enumerated. Subsequently, MNCs were used for flow cytometric analysis, subjected to CD34 selection or cryopreserved, as determined by downstream experiments.

To isolate PBMCs, the protocol above was used with the exception that the sample was collected into an EDTA tube (BD Biosciences) and PB was not passed through a cell strainer prior to dilution in PBS.

2.2.3 Cell enumeration

Cell counting was performed, in triplicate, by trypan blue exclusion or using an electronic cell counter (Countess II FL, Life Technologies). When combined with flow cytometric analysis, cell counting was performed using AccuCheck Counting Beads (Invitrogen, Paisley, UK), as per manufacturer's instructions. Briefly, a fixed number of beads (10µl) was added to each sample immediately prior to flow cytometry. The number of beads was then used to calculate the absolute cell number.

2.2.4 CD34 positive cell selection

CD34+ cells were isolated from MNCs using the MiniMACS Separator kit as per the manufacturer's instructions (Miltenyi Biotech Ltd, Surrey, UK). Briefly, cells were incubated for 30 minutes with FcR blocker and CD34 microbeads (Miltenyi Biotech

Ltd, Surrey UK), washed and passed once through a positive selection column. MS and LS columns were used for total MNC numbers <1x10^8 or >1x10^{8,} respectively. The typical purity and yield of the selected population was usually >85% following a single column passage. In the majority of experiments, the CD34+ population was further purified by fluorescence activated cell sorting (FACS; Figure 2:1).

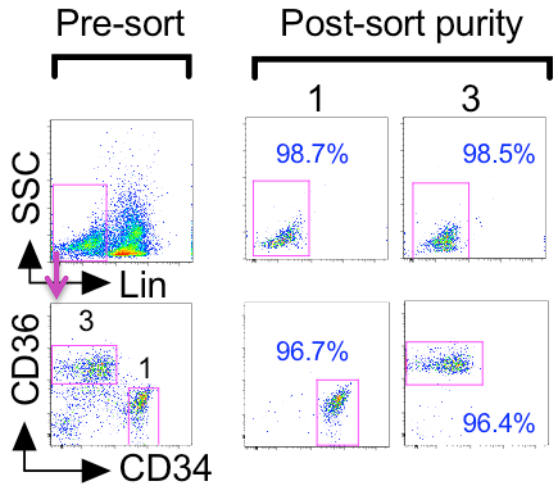


Figure 2:1 Post-sort purity. Two erythroblast populations (labelled 1 and 3) were FACS-sorted from human BM MNCs into RoboSep then reanalysed by flow cytometry to check the purity of FACS-based isolation.

2.2.5 Cell cryopreservation

Cellular fractions were resuspended in freezing medium, constituting 90% Fetal Bovine Serum (FBS; Sigma-Aldrich, Dorset, UK) supplemented with 10% (vol/vol) dimethyl sulfoxide (DMSO; Sigma-Aldrich), at a maximum concentration of 1x10^7 cells/ml freezing medium. Immediately after resuspension in freezing medium, cells were cryopreserved at -80^{0}C in a Mr.Frosty^{TM} freezing vessel (Nalgene) containing 2-Propanol (Sigma-Aldrich). Within 12-48 hours samples were transferred to liquid nitrogen for long-term storage.

2.2.6 Cell thawing

Thaw medium was made up as follows: Iscove's Modified Dulbecco's Medium (IMDM, Sigma-Aldrich) + 20% FBS+ 1% Penicillin-Streptomycin (P/S, Sigma-Aldrich). 1000u/ml DNase (Calbiochem, San Diego, USA) was added when thawing high cell densities of >5x10^6/ml. Thaw medium was stored at 4^0C for up to 1 week then warmed to 37^0C immediately prior to use. Cryopreserved vials from liquid nitrogen were immersed in a 37^0C water bath for 1-3 minutes. The cell suspension was then transferred to a 50ml falcon tube, to which pre-warmed thaw medium was added drop wise over 5 minutes. Samples were spun at 300g at RT, washed then used in further experiments.

2.2.7 In vitro erythroid culture systems

2.2.7.1 Liquid culture systems

Total CD34+ or FACS-isolated HSPC subpopulations were cultured using protocols A or B), depending upon the experiment. In both protocols, cells were plated in 96 well round-bottom plates and concentration was maintained at <1x10^6/ml by partial medium changes every 2-3 days. For transition to new phases, cells were washed and resuspended in fresh media. The cultures were incubated at 37°C with 5% CO2 for up to 21 days, depending upon the experiment. At different time points, cells were harvested for analysis.

To test the effect of CS *in vitro,* culture medium was supplemented with DEX (D4902; Sigma-Aldrich). DEX was dissolved in ethanol to give a 2mM stock solution. Serial dilutions in IMDM/ StemSpan achieved a final concentration of 100nM-1uM. Control cultures were treated with an 0.0001% ETOH vehicle control.

2.2.7.2 Semi-solid culture systems: methylcellulose colony-forming assays

Clonogenic assays were performed with FACS-isolated cells plated in MethoCult™ GF H4034 Optimum (Stem Cell Technologies, Vancouver, Canada), which supports both erythroid and myeloid colony development. Contents are listed in Appendix B: Constituents of reagents and buffers. Thawed medium was supplemented with 2% P/S and aliquots were stored at -20^{0}C.

Table 2.1 Protocols for liquid erythroid culture.

	Protocol A; serum free (Dutt et al. 2011; Ebert 2005)**.**	**Protocol B** (Griffiths et al. 2012; Douay & Giarratana 2009)**,**
Base medium	Serum-free expansion medium (StemSpan; Stem cell technologies) + penicillin/streptomycin (P/S; Sigma-Aldrich Aldrich) 100units/ml, stem cell factor (SCF) 100ng/ml, Interleukin (IL)-3 10ng/ml	Iscove's Modified Dulbecco's Medium (IMDM) + FBS 2%, AB serum 3%, P/S 100units/ml, human insulin 10μg/ml + EPO 3units/ml
D0-7	+ Erythropoietin (EPO) 0.5units/ml. +/- 10ng/ml IL-6 * +/-40ug/ml Lipids* +2mM L-glutamine*	+ SCF 10ng/ml, IL-3 10ng/ml + human holo transferrin 200mg/ml;
D8-11	+ EPO 4units/ml.	+IL-3 10ng/ml + transferrin 1mg/ml
D12 onwards	As per D8-D11	+ Transferrin 1mg/ml

**In selected experiments only, as detailed in relevant chapters. EPO was purchased from R&D systems, FBS, Lipids, L-glutamine, AB serum and insulin from Sigma-Aldrich and the remaining cytokines from PeproTech.*

For bulk clonogenic assays, 100 cells were FACS-sorted directly into 1ml medium in 1.5ml Eppendorf microcentrifuge tubes, vortexed, left to settle for 5 minutes, then transferred to 24-well plates using a 16 gauge blunt end needle (Stem Cell Technologies, Vancouver, Canada). Where cell numbers were not limiting, populations were plated in technical duplicate. For single-cell clonogenic assays, a single cell was sorted into 100μl medium in each well of a flat-bottomed 96-well plate. Visible bubbles were lanced with a 25-gauge needle (VWR International Ltd, Leicestershire, UK) and methylcellulose-containing wells were surrounded by a halo of wells containing PBS. Cultures were incubated at 37°C in a fully humidified 5% CO_2 atmosphere.

Colonies were scored and photographed using an inverted microscope (Evos x1 Core) at days 7 and 12-14, according to established criteria (Gregory & Eaves 1977; A. Roy et al. 2012). Specifically, large or small dense colonies of erythroblasts in 'bursts' were counted on d12-14 as BFU-E and considered to be the progeny of early erythroid progenitors while small uni- or bi-centric clusters of erythroblasts were counted on day 7 as CFU-E and regarded as the progeny of late erythroid progenitors. On day 14, single colonies were plucked from the semisolid medium under direct light microscopy and collected in Eppendorf microcentrifuge tubes containing 1ml IMDM/1% FBS. Individual colony suspensions were then analysed by flow-cytometry (as described above and in Appendix C: Flow cytometry panels and monoclonal antibodies) or used for cytospin preparations to allow morphological determination of the cellular composition of each colony.

2.2.8 Preparation of cells for microscopy using cytospins

FACS-sorted or cultured cells were suspended in RoboSep or IMDM supplemented with 2% FBS, at a concentration of approximately 2-5x10^5 cells/200μl. Cells were cytocentrifuged at 4000 rpm for 5 min onto Superfrost slides, using a Shandon Cytospin 2 (Fisher Scientific, Loughborough, UK). Slides were then air dried, fixed in 100% methanol and stained with May-Grünwald Giemsa (MGG; Sigma-Aldrich,

Dorset, UK). Working solutions of May-Grünwald (MG) and Giemsa (G) stains were made by diluting stocks with pH 6.8 buffered distilled water at 1:1 and 1:10 (vol/vol), respectively. Slides were immersed in MG stain for 7 minutes then directly transferred to G stain for 20 minutes. Slides were then washed in distilled water by gentle agitation for 3 x 2 minutes. After air-drying, slides were mounted with PTX mountant (Sigma-Aldrich, Dorset, UK) and covered with a 50x25mm coverslip (VWR International Ltd, Leicestershire, UK). Cytospin preparations were photographed using a Nikon eclipse E400 inverted microscope and camera. Original magnification is shown on each image.

2.3 Flow cytometry

2.3.1 Staining for cell surface markers

To assess HSPC, EP and EB populations, I designed 4 distinct multiparameter flow cytometry panels (Appendix C) according to the following principles:

1. suitability for the laser and filter configurations of both the LSR Fortessa Analyser and the Aria III/BF Fusions cell sorters
2. use of the same antibody clones as those published for the detection of analogous cell populations, where possible
3. use of strong fluorophores for weakly expressed antigens, and vice versa
4. use of fluorophores that are spectrally similar for different cell subpopulations that would be gated and analysed separately e.g. avoiding use of APC and FITC for 2 surface markers that are co-expressed on cell types

Flow cytometry was performed on freshly isolated or on thawed samples, as indicated in relevant chapters. Cells were suspended in RoboSep buffer (Stem Cell Technologies) and incubated with FcR blocker (Miltneyi Biotech) for 3 minutes at room temperature (RT). Cells were then stained with a panel of up to 13 fluorophore-conjugated monoclonal antibodies (Appendix C: Flow cytometry panels and monoclonal antibodies) for 20 minutes at 4°C. After washing, cells were stained

with 4',6-diamidino-2-phenylindole (DAPI; Sigma-Aldrich) for dead cell exclusion, immediately prior to data acquisition on a 4-laser BD LSR Fortessa. As DAPI could not be used in conjunction with the panels for FACSAria II or Aria Fusion cell sorters, stained samples were washed then stained with the Live/Dead Fixable Aqua Dead Stain Kit (Life Technologies) for 30 minutes at RT.

For isolation of cell subpopulations, cells were FACS-sorted into 1.5ml Eppendorf microcentrifuge tubes (for bulk cell culture assays), PCR tubes (for RNA-seq) or 96 well flat-bottom plates (for single cell culture assays). Cells were tightly gated to ensure discrete populations of cells similar in maturity and lineage. The sort purity was assessed by recovery of sorted cells and was consistently >96% (Figure 2.1).

2.3.2 Analysis

Data were acquired on a BD LSR Fortessa, a FACS Aria II cell sorter or an Aria Fusion cell sorter using FACSDIVA^{TM} software v8.0.1, and analysed using FlowJo software (Tree Star). Compensation settings for multiparameter analyses were set using single-stained beads (OneComp eBeads, Thermo Fisher Scientific) for fluoropohore-conjugated monoclonal antibodies and cells for DAPI. Cell doublets and non-viable (DAPI or BV510 positive) cells were excluded. During panel design and optimisation, gates were set with fluorescence-minus-one plus isotype controls. Otherwise, gates were routinely set with unstained controls as limited primary cell numbers, combined with the complexity of staining panels, precluded the use of fluorescence-minus-one controls for each sample and each fluorophore.

2.3.3 Cell death and cell cycle assays

To measure cell death, cells were stained in Annexin V Binding Buffer (eBioscience) with an antibody against Annexin V (Appendix C: Flow cytometry panels and monoclonal antibodies), as a positive cell surface marker for early and late apoptosis.

After washing, DAPI (1µl/100µl of 5ng/ml stock) was added as an intracellular marker for necrosis.

For cell cycle analysis, cells were harvested, washed twice and fixed using Fixation/Permeabilization Buffer (eBioscience). Cells were resuspended in DAPI x1 solution (Appendix C: Flow cytometry panels and monoclonal antibodies) at a concentration of $1x10^6$cells/ml and incubated for 30 minutes at 4^0C. Analysis was performed using FlowJo software.

2.3.4 Optimisation of HSPC staining panel

As there are limited published data describing the cellular architecture of haematopoiesis in BM from healthy children, in the first instance I compared paediatric BM (PBM) with adult BM (ABM; Figure 2:2). Observed relative frequencies of HSPC in ABM (Manz et al. 2002; Pang et al. 2013) were in line with published data. Overall PBM is similar to ABM, in terms of the frequencies of myeloid progenitors and the presence of immature T lymphoid progenitors within the CD38+ fraction (Figure 2:2: highlighted with red gates).

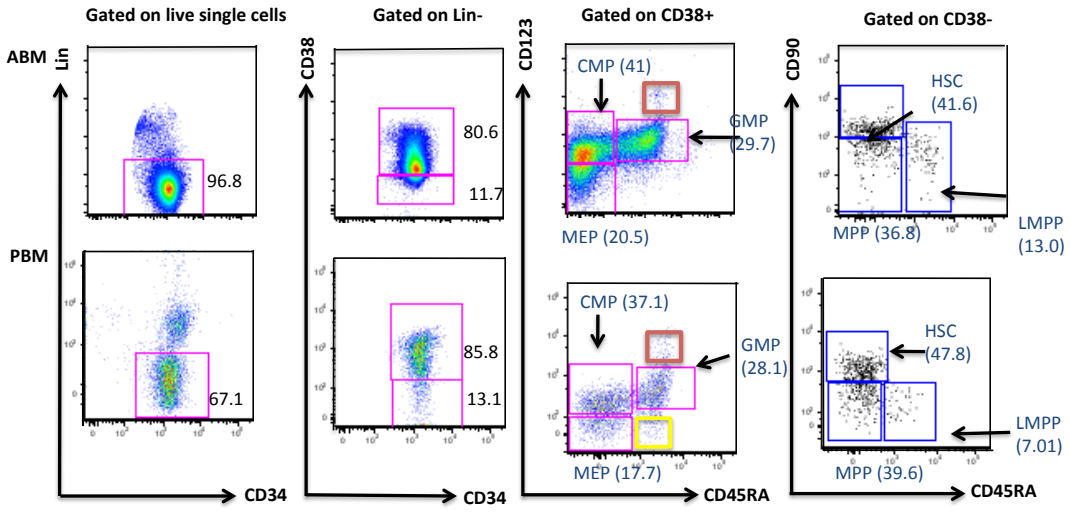


Figure 2:2 Patterns of staining for HSPC populations in normal paediatric and adult BM. CD34+ cells derived from ABM or PBM (n=1) were enriched by immunomagnetic

bead selection and the haematopoietic progenitor populations analysed by multiparameter flow cytometry. Lin+ cells represent terminally differentiated cells of the lymphocyte/monocyte/neutrophil/erythroid or megakaryocyte lineages therefore these were excluded. Lin-CD34+ live single cells were classified as shown: within the CD34+CD38+ population CD123loCD45RA- Common Myeloid Progenitors (CMP), CD123-CD45RA- Megakaryocyte-Erythroid Progenitors (MEP) and CD123loCD45RA+ Granulocyte-Macrophage Progenitors (GMP) fractions were identified; the CD34+CD38-/lo population was subdivided into CD90+CD45RA- Haematopoietic Stem Cells (HSC), CD90-CD45RA- MultiPotent Progenitors (MPP) and CD90-CD45RA+ Lymphoid-primed MultiPotential Progenitors (LMPP). The gating strategy used is highlighted in each plot and by the labels above the plots. Gates were determined using unstained cells as a negative control and therefore vary in each sample. Antigens are labelled on the x and y axes. Numbers refer to the % of the parent population. A red gate highlights CD123hiCD45RA+ T lymphoid progenitors and a yellow gate highlights an initially unknown population, later shown to be early B lymphoid progenitors.

In PBM I observed an unexpected cell population with the following immunophenotype: CD34+CD38+CD123-CD45RA+ (Figure 2:2: highlighted with yellow gate). As CD19, but not CD10, is included in the manufactured cocktail of antibodies for haematopoietic lineages, and given that FSC versus SSC demonstrated that these cells are small with low organelle content, I hypothesised that they could be early B progenitors that express CD10 but not yet CD19 (Haddad et al. 2004). Staining for CD10 confirmed that the majority of cells in this gate were indeed CD10+. I therefore optimised the established HSPC panel for PBM, to minimise effects on the frequencies of myeloid progenitors within the gates of interest, by adding CD10 to the lineage cocktail.

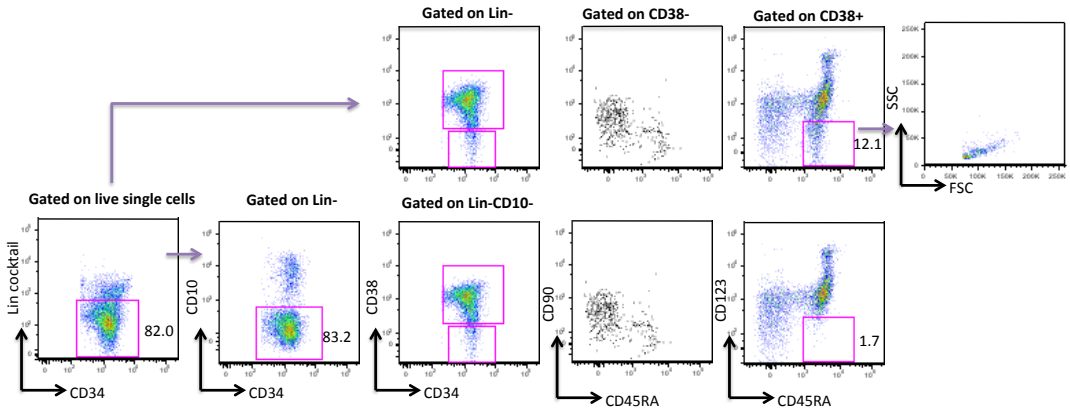


Figure 2:3 Modified gating strategy to exclude CD10+ B progenitors from normal PBM. Representative BM from 7y female shown. Traditional HSPC staining approach is shown in the top panel. In the modified approach (bottom panel) CD10 is added to lineage markers, which are negatively gated out. In this example, this strategy reduced the CD34+CD38+CD123-CD45RA+ population from 12.1 to 1.7% of the CD38+ fraction.

2.4 Nucleic acid techniques

2.4.1 RNA extraction

Total RNA from $1x10^2$-$1x10^5$ cells was extracted using the NucleoSpin RNA XS kit (Macherey-Nagel, Germany), as per the manufacturer's instructions. Briefly, cells were suspended or FACS-sorted into lysis buffer and homogenised by vortexing before freezing at -80^{0}C. After thawing on ice, RNA was extracted using a microcolumn, including removal of DNA contaminations by on-column digestion with rDNase. RNA was then eluted in 12.5µl RNase-free H_20 and stored at -80^{0}C or used immediately in cDNA synthesis.

2.4.2 cDNA synthesis

cDNA was synthesised with the Superscript III first-strand RT-PCR System (Invitrogen, Paisley, UK) as per the manufacturer's instructions. Total RNA was denatured and annealed by incubating with dNTPs and oligo$(dT)_{20}$ at 65^{o}C for 5 minutes then cooled on ice for 1 minute. First-strand cDNA was synthesized by incubating the sample

with RT buffer, $MgCl_2$, RNase inhibitor, dithiothreitol (DTT) and Superscript III reverse transcriptase enzyme, at 50^{o}C for 50 minutes. Termination of the reaction was achieved by incubating at 85^{o}C for 5 minutes. Finally, the RNA template was removed by incubating with RNase H at 37^{0}C for 20 minutes. Synthesized cDNA was stored at-20^{o}C or was utilized immediately in quantitative reverse transcription polymerase chain reaction (qRT-PCR)

2.4.3 qRT-PCR

To measure gene expression, qRT-PCR of template cDNA was performed using Taqman Fast Advanced Master mix and Taqman Gene Expression Assays (Appendix D. RT-PCR Taqman Gene Expression ID), as per the manufacturer's instructions. Reactions were run in a MicroAmp Optical 96-Well Reaction Plate (Life Technologies) in the Applied Biosystems StepOnePlus RT-PCR System as follows:

UNG denaturation:	Hold at 50^{0}C for 2 minutes
Polymerase activation:	Hold at 95^{0}C for 20 seconds
PCR cycling (40 cycles):	Denature at 95^{0}C for 3 seconds
	Anneal at 60^{0}C for 30 seconds

Expression was calculated relative to the reference genes, *ACTB* or *GAPDH*, using the ΔΔCT method (Schmittgen & Livak 2008).

2.4.4 Nucleic acid quantification

For the methods outlined above, nucleic acids were quantified and assessed qualitatively using a NanoDrop 2000 (Thermo Fisher Scientific). Optimal DNA and RNA quality was reflected by A260/280 ratios of ≥1.8 and 2.0, respectively.

2.5 Protein assays

2.5.1 Enzyme-linked immunosorbent assay (ELISA)

To measure the concentration of the secreted protein, Growth Arrest-Specific 6 (GAS-6), in human BM plasma, a sandwich ELISA was performed using the Human

Gas6 DuoSet ELISA Kit and the DuoSet Ancillary Reagent Kit 2 (Bio-Techne Ltd, Oxfordshire, UK), as per the manufacturer's instructions. A 96 well flat-bottom plate was coated overnight at RT with 100µl of GAS-6 specific Capture Antibody (diluted in PBS). Following 3 washes with Wash Buffer (PBS containing 0.01 % Tween-20; PBST) plates were blocked with 300µl of Reagent Diluent (RD) at RT for 1 hour. After washing, plates were ready for sample addition. Human serum was thawed on ice and serially diluted in RD to 1 in 10, 1 in 20 and 1 in 40. 100µl of each dilution was plated in duplicate. Plates also included control wells containing RD alone and 7 standards of known concentration, all in duplicate. Plates were left to incubate for 2 hours at RT. Following 3 washes with PBST, 100µl GAS-6 specific Detection Antibody (diluted in RD) was added to each well. Plates were incubated for 2 hours at RT, washed 3 times with PBST then 100µl working solution Streptavadin-HRP was added to each well and incubated in the dark for 20 minutes at RT. After washing, 100µl Substrate Solution was added and colour was allowed to develop for 20 minutes at RT. The reaction was then stopped with 50µl Stop Solution and gentle agitation of the plate. To determine the optical density of each well, absorbances were read immediately, using a microplate reader set to 450nm. For each sample, GAS-6 concentration was determined by interpolation from the standard curve, correction for the dilution factor and calculation of the mean concentration of the different dilutions.

2.5.2 Immunohistochemistry

Immunohistochemisty (IHC) was performed on archived BM paraffin sections by Dr. Pritesh Truvadi, under the supervision of Professor Kikkeri Naresh, Department of Histopathology, Imperial College Healthcare NHS Trust. 1mm sections were cut and stained in a Leica BOND-III Automated IHC Stainer. Automated steps included: deparaffinisation and treatment with 3% hydrogen peroxide to quench endogenous peroxidase activity, antigen retrieval using EDTA (0.001 mol/L), pH 8.0 (Invitrogen, San Francisco, CA) at 96 °C for 30 minutes, washing in Tris buffer (Covance, Dedham,

MA), incubation with anti-GATA1 (D52H6) rabbit monoclonal antibody (Cell Signaling Technology, Danvers, MA) at a 1:200 dilution for 20 minutes at RT, incubation for 30 minutes with a horseradish peroxidase (HRP)–labelled polymer conjugated to goat anti–rabbit immunoglobulin antibody (PowerVision; Leica Microsystems, Buffalo Grove, IL), staining with diaminobenzidine (DAB; Dako, Carpinteria, CA) as a chromogen, enhancement with 1% copper sulphate solution, washing, counterstaining with Mayer's haematoxylin, dehydration, and mounting.

The pattern and intensity of staining was evaluated as follows: Grade 1 staining was defined as the homogenous deep staining observed in megakaryocytes in normal BM; Grade 2 staining was defined as the speckled, light staining observed in eosinophils (and the occasional mast cells) in normal BM.

To allow more reliable assessment of GATA-1 staining intensity in the erythroid lineage, the identity of EB was confirmed by double staining for Glycophorin C (GLYC)- a cell membrane marker of early and late EB. Co-staining was performed in the Leica BOND-III using a monoclonal mouse anti-human antibody (clone RET40f, Dako, Carpinteria, CA) at a 1 in 150 dilution. A red stain was used as a chromogen. GLYC positive cells were identified by their deep red perinuclear staining.

2.6 Transcriptome analysis by RNA-sequencing

2.6.1 Methodological approach

To prepare double stranded DNA libraries for RNA-sequencing, $1x10^3$-$8x10^4$ cells (according to availability in individual samples) were FACS-sorted into lysis buffer (NucleoSpin RNA XS kit, Macherey-Nagel), vortexed, snap frozen on dry ice and stored at -80^{0}C. On thawing, RNA was extracted (as described above) and RNA quality assessed using the Agilent RNA 6000 Pico Kit (Agilent Bioanalyser), aiming for a RNA Integrity Number (RIN) of >7. Bioanalyser traces were used to determine approximate RNA concentrations. Though the Qubit RNA HS Kit (Thermo Fisher Scientific) provides more accurate quantification, some samples were below its

lower limit of detection (250pg/µl). RNA was stored at -80^{0}C until library preparation using the NEBNext Poly(A) mRNA Magnetic Isolation Module, NEBNext Directional RNA First and Second Strand Synthesis Modules and the NEBNext® Ultra™ II DNA Library Prep Kit for Illumina® (New England Biolabs, Hertfordshire, UK), as per the manufacturer's version 1.5 protocol, optimised for <100ng total RNA input. First, mRNA was isolated from total RNA and fragmented, followed by first and second strand cDNA synthesis. Directionality was achieved by addition of Actinomycin D (Sigma-Aldrich, Dorset, UK). Double-stranded cDNA was purified and stored at -20'C until End Prep of the cDNA library, adaptor ligation and PCR enrichment, using a unique Truseq barcode for each sample, from the NEBNext Multiple Oligos for Illumina Primer (Index Primers Set 1). Following optimisation, 13-16 PCR cycles were performed for 1-10ng total input RNA. The PCR reaction was purified, eluted in 15ul 10mM TrisHCl then stored at -20^{0}C. All clean up steps were performed using Agencourt AMPure XP Beads (Beckman Coulter), which were allowed to equilibrate to RT before use. 80% ETOH was made up fresh each time with molecular grade ETOH and RNase- free H_20 (Sigma-Aldrich).

Where possible, normal and disease samples were paired and processed in parallel to prevent confounding batch effects.

cDNA library quality was assessed using the Agilent DNA HS Kit (Agilent Bioanalyser). In cases where an adaptor-dimer peak was visible at 128bp, a repeat clean up step was performed. Libraries were quantified by the Qubit dsDNA HS Assay Kit (Thermo Fisher Scientific) and diluted to concentrations of 2-2.5nM to allow equimolar pooling of samples for sequencing. Pooled samples were submitted to the Medical Research Council Genomics Facility for sequencing using the Illumina HiSeq 2500 platform followed by demultiplexing to obtain paired-end 100bp reads.

Table 2.2 Summary of protocol used for preparation of libraries for RNA-sequencing

	Protocol for EB
Cell source	Fresh/ thawed BM CD34- or MNC
Population	FACS-isolated E3 (Erythroid panel; Appendix C)
No. of cells per replicate	$1x10^3$-$8x10^4$
Technology	RNA-extraction→ PolyA mRNA selection (NEB PolyA Isolation Kit)→Library preparation (NEB Ultra v1.5/v2 RNA Kit)
Quality control steps prior to sequencing	RNA: RNA Pico cDNA library: HS DNA (Agilent Bioanalyser)
Strand specificity	Yes
ERCC RNA spike-in control	No
No. of samples multiplexed per lane	4-6

2.6.2 Bioinformatics analysis

Myself and Dr. Aristeidis Chaidos performed bioinformatic analysis of RNA-seq data using R scripts provided by Dr. Aristeidis Chaidos. The approach (Figure 2:4) was as follows: raw reads obtained from RNA-seq experiments were subjected to quality control using FAST-QC software then aligned using "STAR 2.5.3a" (Dobin et al. 2012) against the GRCh38 human reference genome, employing the default settings. Raw counts at gene level were obtained using the "Rsubread" Bioconductor package (Liao et al. 2013). Data normalisation and differential expression analysis across different groups were performed using the "DESeq2" Bioconductor package (Love et al. 2014). A combined fold change (FC) of 0.6 (log2-scale) and Padj 0.05 cutoff was applied to obtain significantly up- or down-regulated genes. Fold change (log2-scaled) values were z-score standardised, clustered and visualised as heatmaps within R

environment. Gene networks and upstream regulators were identified through the use of Ingenuity Pathways Analysis (Ingenuity® Systems, www.ingenuity.com).

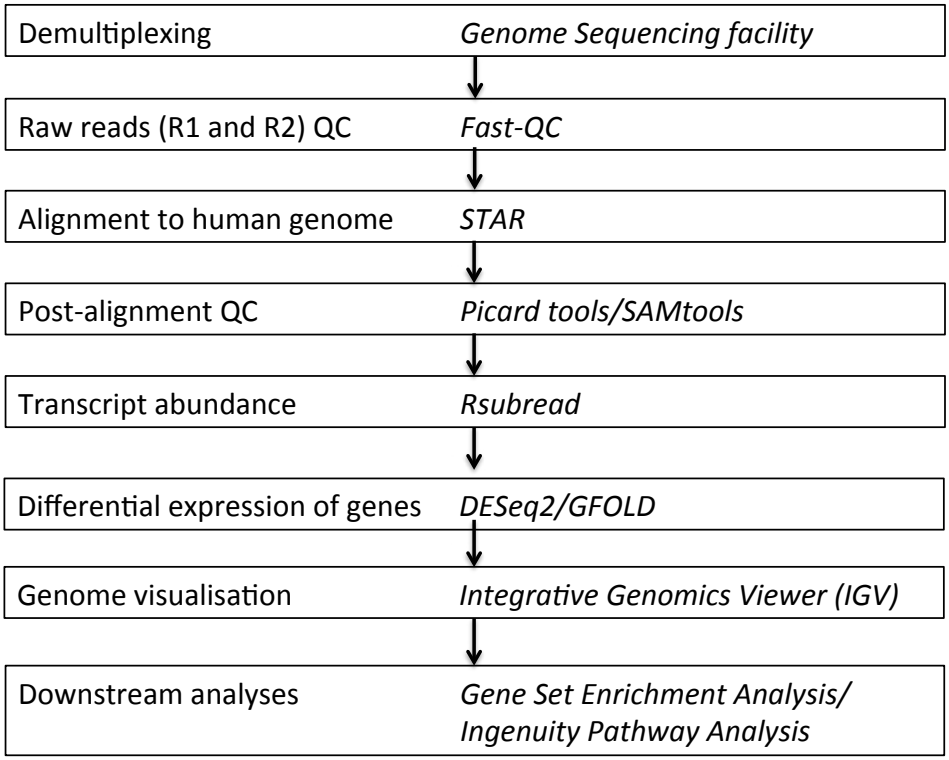


Figure 2:4 Bioinformatics analysis pipeline for RNA-seq data

2.7 Statistical analysis

Data were analysed and graphs generated using Excel (Microsoft) and Prism GraphPad for Mac (v6.0). Descriptive statistics show mean ± SEM, unless otherwise stated in figure legends. Differences in characteristics between groups were assessed using the unpaired Student *t*-test, Mann-Whitney U test or Kruskal-Wallis tests as appropriate. The chi-square test was used to determine differences between categorical variables. To evaluate the relationship between potential prognostic factors e.g. mutation, and clinical outcome, logistical regression analysis was performed by Dr Richard Szyldo, using SPSS (Version 22, IBM Corp., Armonk, NY) computer software. For all analyses, a two-tailed p value of less than 0.05 was considered significant.

3 DIAMOND-BLACKFAN ANAEMIA IN THE UNITED KINGDOM: GENETIC AND CLINICAL CHARACTERISTICS

3.1 Introduction and aims

As for all rare diseases, accurate description of the clinical and laboratory phenotype of DBA requires the generation and maintenance of a patient registry. The largest international registry is currently the DBA North American registry (DBAR), comprising over 600 patients (Lahoti et al. 2016). Many insights into DBA have stemmed from registry data, including the identification of new disease associations such as endocrine disorders (Lahoti et al. 2016), and quantification of the risk of neoplasia (Vlachos et al. 2012).

St Mary's Hospital, Imperial College Healthcare Trust (ICHT), is the main centre in the United Kingdom (U.K) for non-malignant HSCT (15-20 sibling donor BM transplants for inherited red cell disorders/year) and is the national tertiary referral centre for patients with DBA. One advantage of this centralised system of care is consistent approaches to clinical coding, diagnostic investigations and management, allowing robust comparisons both between different patients and within individual patients over time. The last update from the U.K registry was in 2004 (Orfali et al. 2004). Since then, follow-up of existing patients has continued, more patients have been notified to ICHT, and genetic testing has advanced considerably. I aimed to evaluate the clinical and genetic spectrum of DBA in the U.K, specifically to:

- provide comprehensive genetic and clinical characterisation of patients from whom I would be collecting tissue samples
- systematically characterise the genetic and phenotypic spectrum of the disease, thus identifying any novel features
- investigate the presence of any novel genotype/phenotype correlations

3.2 Methodological approach

I collected and analysed data from all 128 cases of DBA notified or referred to St Mary's Hospital, ICHT over the last 10 years. 5 asymptomatic carriers of an RP gene mutation i.e., not categorised as DBA patients (all mothers of probands) were excluded from the analysis. In the first instance I verified the diagnosis of DBA, using published diagnostic criteria outlined in Table 1.4 (Vlachos et al. 2013). NGS of 80 RP genes plus *GATA-1* was performed as previously described (Gerrard et al. 2013). Pathogenicity was assigned to variants as per ACGS criteria; clearly pathogenic [CPV], likely pathogenic [LPV], variant of unknown significance [VUS], unlikely pathogenic variant (UPV) and clearly not pathogenic (Wallis 2013). In summary, variants were regarded as CPV if previously documented as pathogenic in the Leiden Open (source) Variant Database (LOVD) (Boria et al. 2008). New null variants (nonsense, frameshift, canonical ±1 or 2 splice sites, initiation codon, frameshift or whole allele deletions) in known DBA genes, i.e., genes known to harbour variants that cause DBA, were also classified as CPV. RP missense mutations and non-canonical splice site defects were classified as likely deleterious as per the *in silico* platform Mutation Taster for functional prediction (Schwarz et al. 2014). Mutations in RP genes not previously implicated in DBA were classified as VUS, pending further studies.

In 5/128 cases, referred as DBA, a definite diagnosis of DBA could not be confirmed due to the unavailability of diagnostic investigations such as mean cell volume (MCV), reticulocytes and BM in addition to a lack of a CPV or LPV. In 4 of these cases, NGS did not detect a variant and the other case could not be tested as he was deceased. These 5 patients were therefore excluded from the study leaving a total of 123.

Data were collected retrospectively from historic electronic medical records (this was the only source of information for 2/123 patients who had died prior to the commencement of the study) and prospectively from family interviews in the paediatric haematology clinic and from referring clinicians and laboratories at other

NHS sites throughout the U.K. The majority (117/121) of living patients had undergone at least one assessment at St Mary's Hospital over the last 5 years. The 4 remaining cases had not been assessed on site but were referred for consultation or for genetic testing. In these cases, data were collected from the referring clinical team.

Available data were not uniform for all patients therefore descriptive statistics were calculated according to the total number for whom data were complete for a given variable, rather than the full cohort. Statistic analysis was performed as described in materials and methods. Specifically, comparisons between genotype and phenotype were assessed using Mann-Whitney U or Kruskal-Wallis tests as appropriate. All variables found to be significant in univariate analyses (using Pearson Chi-square test) were included in a multivariable stepwise logistic regression analysis. For all analyses a two-tailed p value of less than 0.05 was considered significant.

3.3 Results

3.3.1 Demographics

The U.K Cohort comprises 123 patients from 117 families (i.e., 6 multiplex families with more than 1 case of DBA). Patients were born over a 48-year period (1967-2016), which corresponds to an incidence of approximately 4 per million live births. The median age at the time of analysis was 12.5yrs with a range of 1.1-50.0yrs. As previously reported (Vlachos et al. 2012), the disease shows no predilection for a particular ethnic group (Figure 3:1A). However, in the U.K cohort, as opposed to other registries (Pospisilova et al. 2012; Lipton et al. 2006) there are more females than males (71:52, ratio of 1.3:1). Notably, there is a striking sex imbalance between RPS and RPL genotypes with the proportion of females 17/46 (36.9%) and 27/33 (81.8%) in the 2 respective groups. Contrary to a previous observation from the U.K. registry (Orfali et al. 2004), there is no observed increase in incidence of DBA in individuals born in winter (Figure 3:1B).

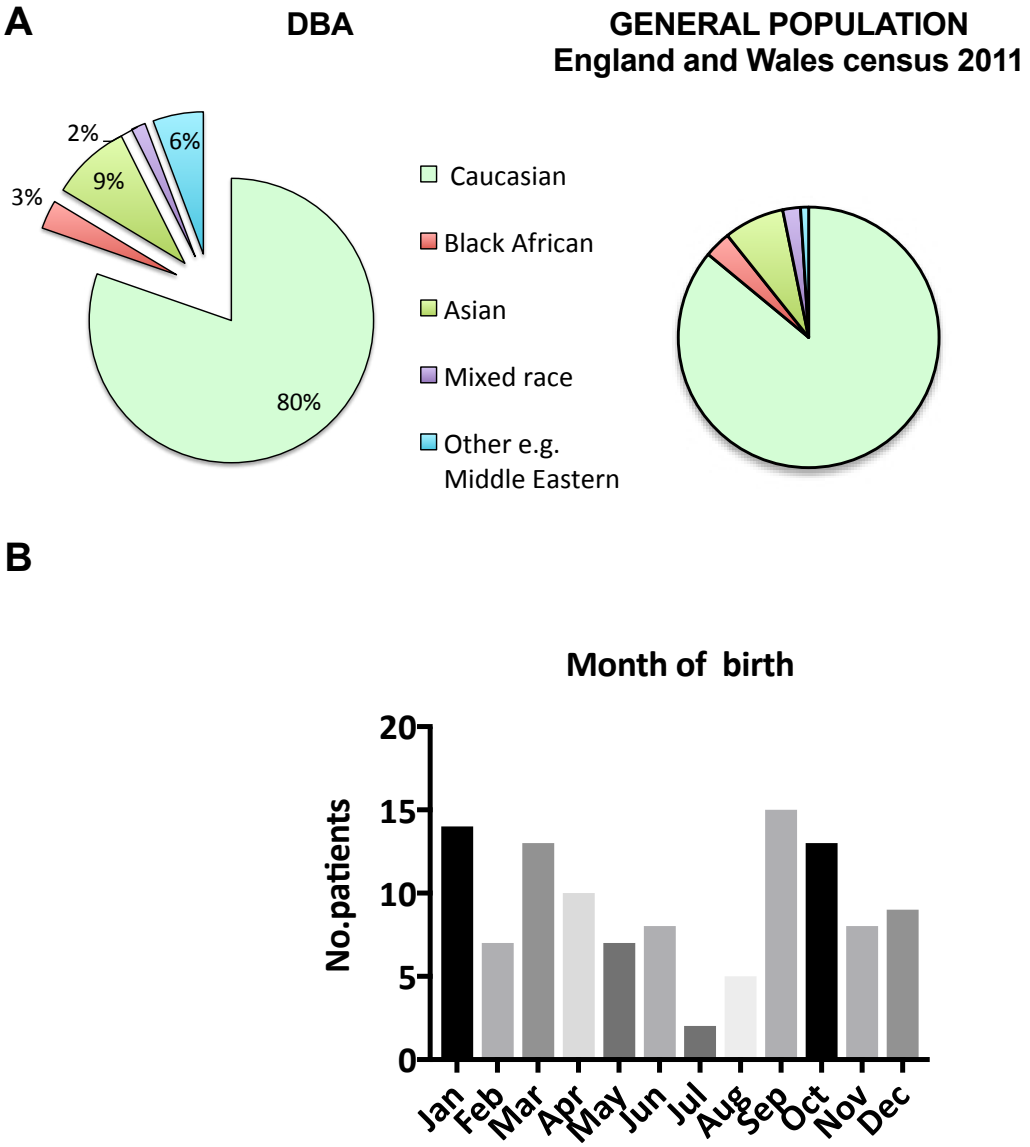


Figure 3:1 Demographics of 123 patients with DBA in the U.K. **A)** Ethnic diversity of patients with DBA corresponds to that of the U.K population. Relative proportions of different ethnic groups are shown for patients with DBA (left) and for the general population in England and Wales (right; Office of National Statistics 2016) **B)** Birth month distribution does not suggest any associations between DBA births and seasons.

3.3.2 Genetic analysis

NGS was performed in 112/123 cases (Figure 3.2A). In the remaining cases, genetic screening is awaited (n=8), patients have died prior to the opportunity for screening (n=2) or have been transplanted with no available pre-transplant DNA (n=1).

A heterozygous CPV/LPV, including deletion of an RP gene, was identified in 79/112 (70.5%) of tested cases. A further 15/112 (13.4%) of cases harboured variants of unknown significance (VUS: Table 3.1). Functional validation studies are required to confirm whether these VUS are implicated DBA phenotype. No DNA variant was identified by NGS in 18/112 (16.1%) cases. Of these, 13 have consented to enrolment in the NHS England 100 000 Genomes Project (Griffin et al. 2017) and results are awaited. No mutations in *GATA-1* were identified.

As previously reported (Boria et al. 2010), of *all* cases of DBA, *RPS19* is the most common gene mutated (21%), followed by *RPL5* (16%), *RPS26* (12%), and *RPL11* (11%). Together these 4 genes account for 67/79 (85%) of *genetically diagnosed* DBA cases in our cohort. Hence genotype/phenotype studies were performed by comparing these 4 genes and by comparing RPL, RPS and no known mutation (NKM) cases overall. The incidence of mutations in *RPL5/11* genes is higher in the U.K than in other international cohorts [27% vs 12% (Clinton & Gazda 2009)]. As previously reported (Farrar & Dahl 2011), pathogenic variants (Figure 3:2B) are most commonly nonsense**,** splice site mutations and frameshift insertions/deletions. All are heterozygous, in line with the notion that RP gene haploinsufficiency underpins DBA. Interestingly, one index case has a heterozygous splice variant in *RPL5* (c.3+3G>C) that is present in homozygosity in her asymptomatic father, suggesting that it is unlikely to be pathogenic, although the fact that pathogenic variants can arise in haematologically normal first degree relatives needs to be borne in mind.

Table 3.1 Characteristics of VUS/UPV identified in 15 probands

ID	Gene(s)	Variant coding sequence	Effect	Variant type	Parental screening	Novel aspects
007	*RPL5*	c.3+3G>C	p.?	Splice site	Homozygous in asymptomatic father, not present in mother	Known gene, UPV
070	*RPS7*	c.562T>C	p.Leu133Ser	Missense	Not performed	Known gene, VUS
100	*RPS29* *RPL13*	*RPS29* c.139G>A *RPL13* c.497C>T	p.Ala47Thr p.Ala166Val	Missense Missense	RPS29 variant present in asymptomatic father	Known gene; VUS Novel gene, VUS
064	*RPL15*	c.466T>G	p.Ser141Ala	Missense	Not performed	Known gene, VUS
031	*RPS17* & *RPS15*	*RPS17* *c.*312A>G *RPS15* c.244G>C	TBC p.Asp82His	Missense Missense	RPS17 variant present in asymptomatic father, RPS15 variant present in asymptomatic mother	Known genes, VUS
18	*RPL13*	c.25G>A	p.?	Regulatory region ^	Not performed	Novel gene, VUS
035	*RPS3A*	c.737C>A	p.Thr246Lys	Missense	Negative in both parents	Novel gene, VUS
057	*RPL18*	c.*18G>A	p.?	Regulatory region	Not performed	Novel gene, VUS
068	*RPL17*	c.87G>A/	p.?	Regulatory region	Not performed	Novel gene, VUS
109	*RPL19*	c.125G>A	p.Arg42Gln	Missense	Not performed	Novel gene,

						VUS
118+119	*RPS6*	c.152G>A	p.Phe72Tyr	Missense	Present in affected mother	Novel gene, VUS
020	*RPL11*	c.406-407insGGGACAGGT		In-frame insertion	Negative in both parents	Known gene, VUS
039	*RPS19*	c.71+5G>T	p.?	Splice site	Negative in both parents	Known gene, VUS
058	*RPL6* & *RPS12*	*RPL6* c.131+69C>T; *RPS12* c.15-9G>C	p.?; p.?	Splice site Splice site	Negative in both parents	Novel genes, VUS

DNA variants were detected by NGS and validated by Sanger sequencing. Some variants are of unknown significance (VUS) or unlikely to be pathogenic (UPV) although they occur in genes known to be involved in DBA. Other variants are in 'novel' DBA genes i.e., RP genes not previously implicated in DBA. ^Pt also has ring chromosome 14 syndrome. Human Genome Variation Society nomenclature used.

To date, genetic screening of both parents has been performed in 39 of the 117 families, in which there is a known mutation in the index case(s). Of these, 28 (71.8%) mutations occur *de novo* while 11 (39.3%) are inherited; in 6 (15.4%) of the latter, the parent is a silent 'carrier' without anaemia while in the remaining 5 (12.8%), the parent also has a history of anaemia. The latter figure is in line with the published figure of 10-15% familial cases (Alter & Young 2003). There were no differences in clinical features between familial and non-familial cases.

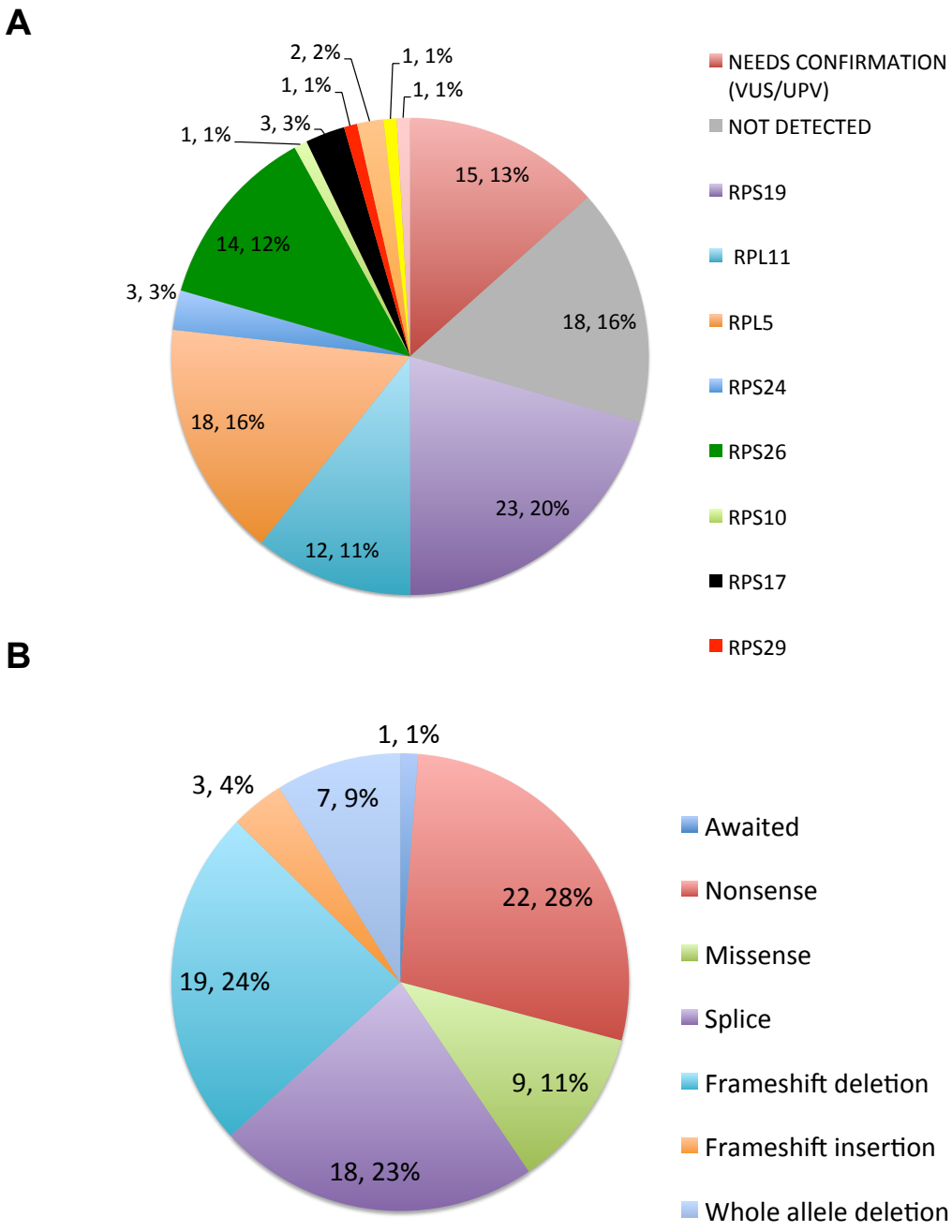


Figure 3:2 Spectrum of RP gene mutations in the U.K cohort. A) NGS of 80 RP genes + *GATA-1* was performed in 112 cases. Shown are frequencies (%) of RP genes in which pathogenic genetic variants (CPV/LPV) were identified. Mutations and deletions were confirmed by Sanger sequencing + Multiplex-Ligation Probe Dependent Amplification, respectively. **B)** Frequencies of types of genetic variants identified across all RP genes (n=79).

In a further 2 families, there is a possibility of gonadal/germline mosaicism, a phenomenon whereby somatic cells carry 2 normal alleles but a proportion of gametes carry a mutated allele**.** In family 1 (Figure 3.3) a known pathogenic nonsense mutation in *RPS24* was identified in the affected child. Peripheral blood screening for this mutation by Sanger sequencing in both parents was negative therefore it was assumed to be a *de novo* case. However a subsequent pregnancy by *in vitro* fertilisation (IVF) with pre-implantation diagnosis (PID) showed the presence of the same mutation in 7/22 embryos, suggesting germline mosaicism. In family 2 (Figure 3.3), a novel missense variant in *RPS19* was detected by NGS in 2 siblings with DBA, however the parents were found to be negative for the variant by Sanger sequencing, again suggesting the possibility of germline mosaicism. In both families, NGS screening of the parents should be performed to exclude low-level (<10%) somatic mosaicism that may have been missed by the relatively insensitive Sanger sequencing method.

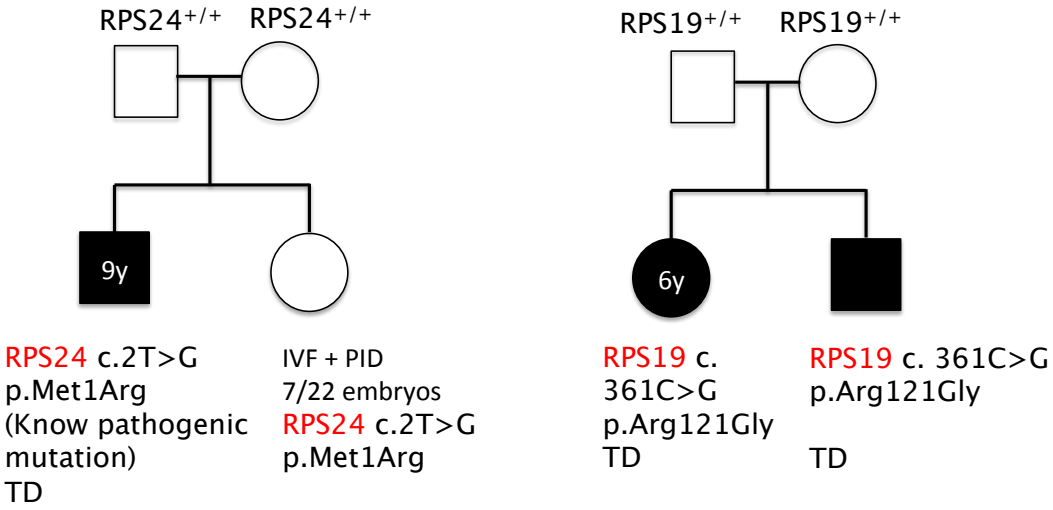


Figure 3:3 Pedigrees for 2 DBA families in which inheritance pattern is suggestive of germline mosaicism. Gene, exon and nucleotide variant and amino acid change all shown**.** *IVF- in vitro fertilisation; PID- preimplantation diagnosis; TD= transfusion dependent.*

3.3.3 Clinical features

3.3.3.1 Birth and disease presentation

Although historically 90% of cases of DBA have been reported to present in infancy (Vlachos & E. Muir 2010), only 73.9% of cases in our cohort presented within the first year of life. Most early cases presented with symptoms directly related to anaemia such as respiratory distress, lethargy, poor feeding or pallor. Apart from Hb, other presenting investigations were poorly documented though where available, MCV and eADA were elevated in 18/22 (81.8%) and 37/39 (94.9%) of patients, respectively. In line with previous observations (Fargo et al. 2012), there are 2 probands with first-degree relatives who have high eADA levels but who do not carry the pathogenic mutation detected in the corresponding proband.

For the first time, my data show an association between RPL genotype and later age of presentation (Figure 3:5: median of 6 weeks in *RPS* patients versus 1 year in *RPL* patients, P=0.002). Both *RPL11* and *RPL5* cases appear to present later than RPS gene-associated cases, though the delay is more marked for *RPL11* mutations (Figure 3:4A & Figure 3.6B). Age of presentation is bimodal in these groups, although there are no clear factors that delineate early and late presenters. There was no difference in the level of Hb by the time patients presented with symptoms of anaemia (Figure 3:4B & Figure 3:6C).

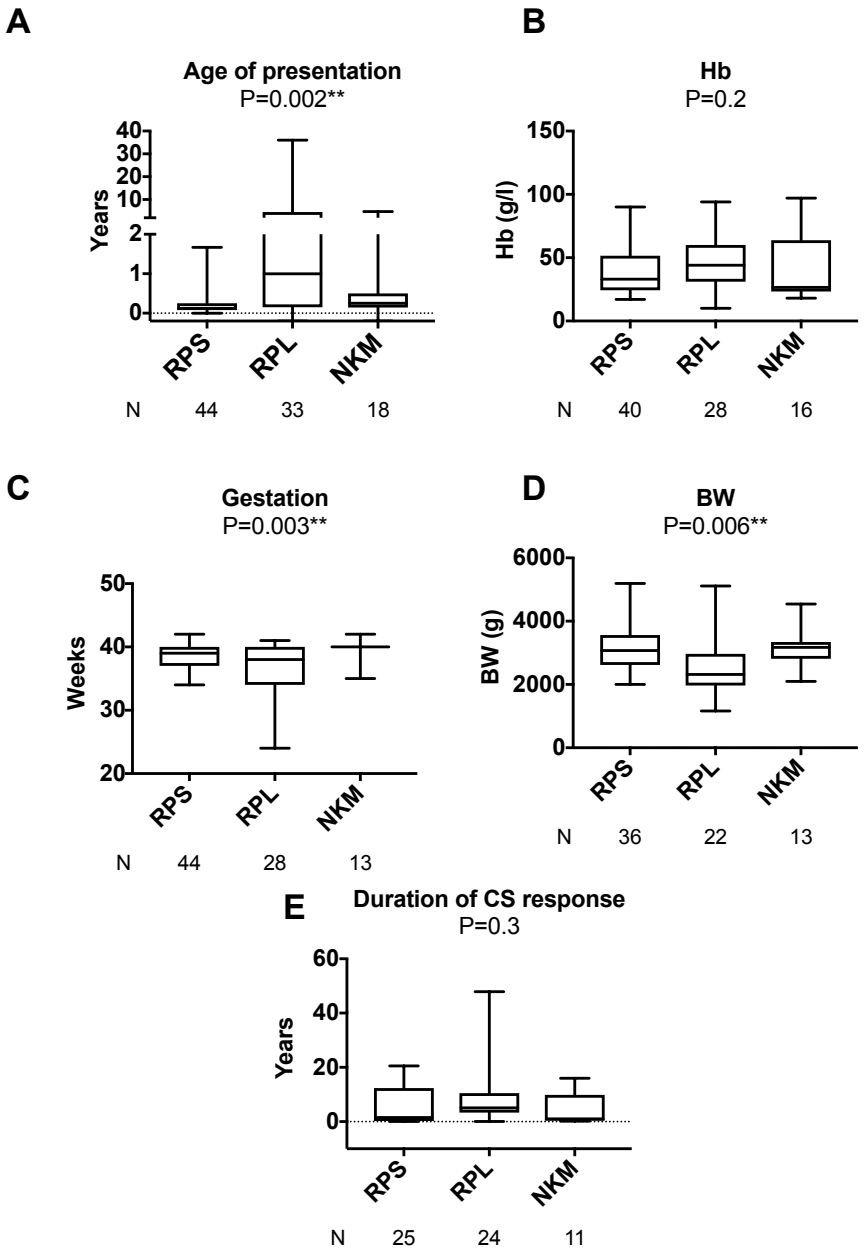
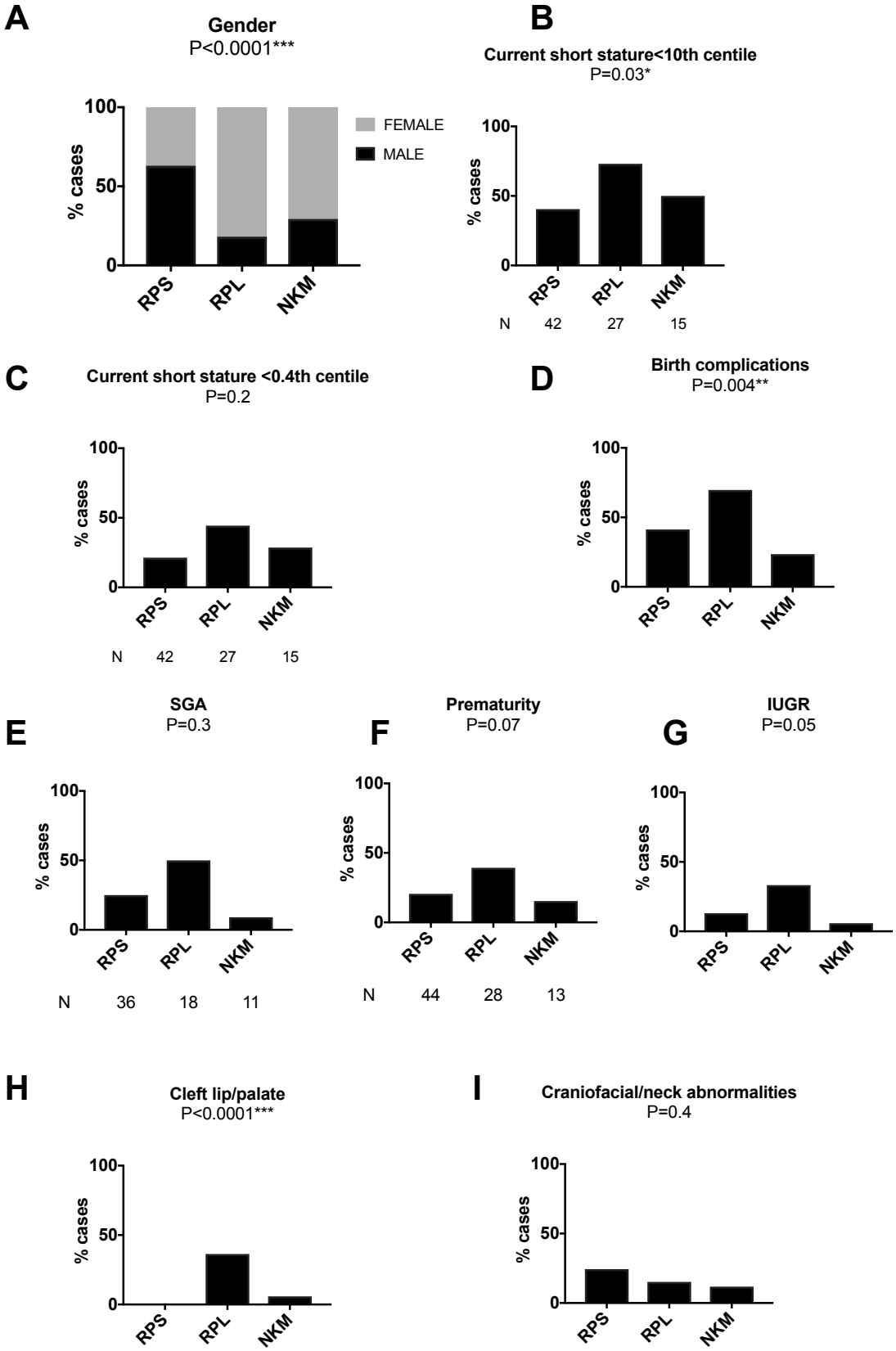
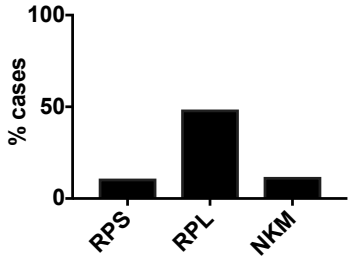


Figure 3:4 Clinical parameters according to genotype I. Box and whisker plots show median and minimum to maximum values. No. of patients with available data (N) shown for each genotypic subgroup in each graph. A three-way comparison was performed between patients with *RPS*, *RPL* pathogenic variants and those in whom genetic basis is unknown (NKM). P values were generated using the Kruskal-Wallis test and reflect the significance of differences between the 3 groups.


A
Gender
P<0.0001***
FEMALE
MALE
% cases
100
50
0
RPS
RPL
NKM
B
Current short stature<10th centile
P=0.03*
% cases
100
50
0
RPS
RPL
NKM
N 42 27 15
C
Current short stature <0.4th centile
P=0.2
% cases
100
50
0
RPS
RPL
NKM
N 42 27 15
D
Birth complications
P=0.004**
% cases
100
50
0
RPS
RPL
NKM
E
SGA
P=0.3
% cases
100
50
0
RPS
RPL
NKM
N 36 18 11
F
Prematurity
P=0.07
% cases
100
50
0
RPS
RPL
NKM
N 44 28 13
G
IUGR
P=0.05
% cases
100
50
0
RPS
RPL
NKM
H
Cleft lip/palate
P<0.0001***
% cases
100
50
0
RPS
RPL
NKM
I
Craniofacial/neck abnormalities
P=0.4
% cases
100
50
0
RPS
RPL
NKM

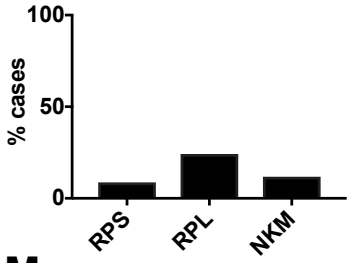
J

Thumb/radial abnormalities
P<0.0001***



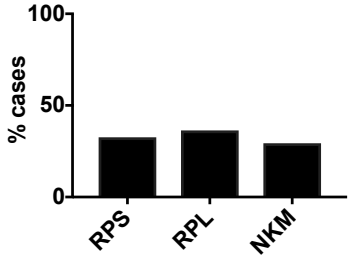
K

Other limb abnormalities
P=0.1



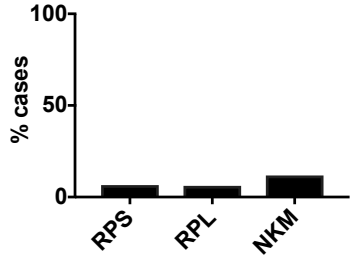
L

Congenital cardiac abnormalities
P=0.9



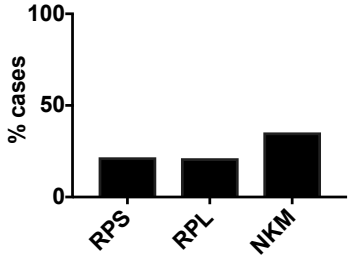
M

Congenital genitourinary abnormalities
P=0.4



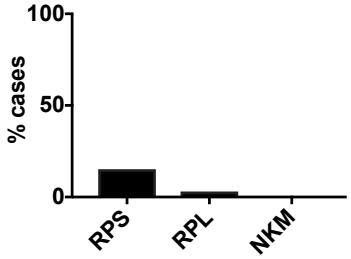
N

Skeletal abnormalities
P=0.5



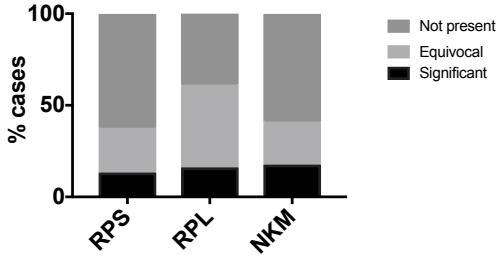
O

Hernias
P=0.06



P

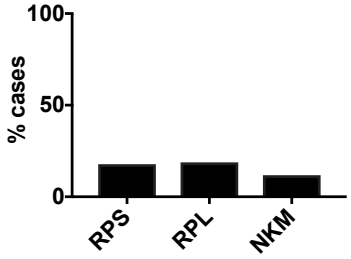
Infections
P=0.3



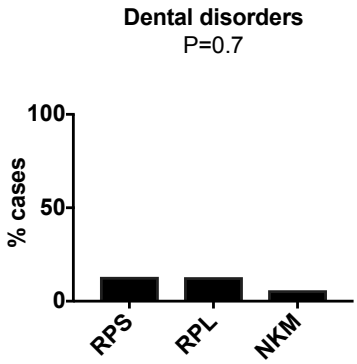
N 45 33 18

Q

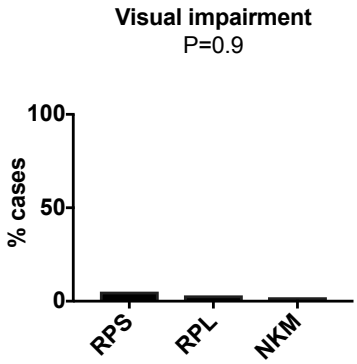
Endocrinopathies
P=0.8



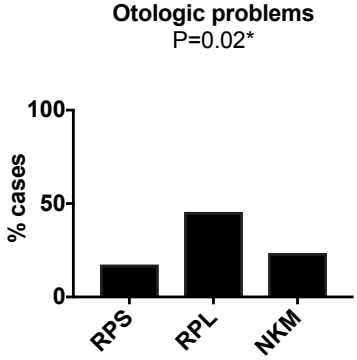
R


Dental disorders
P=0.7
% cases
100
50
0
RPS
RPL
NKM

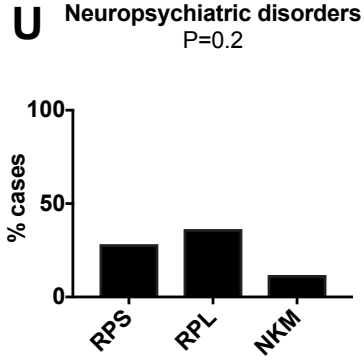
S


Visual impairment
P=0.9
% cases
100
50
0
RPS
RPL
NKM

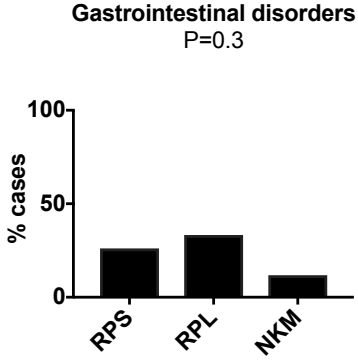
T


Otologic problems
P=0.02*
% cases
100
50
0
RPS
RPL
NKM

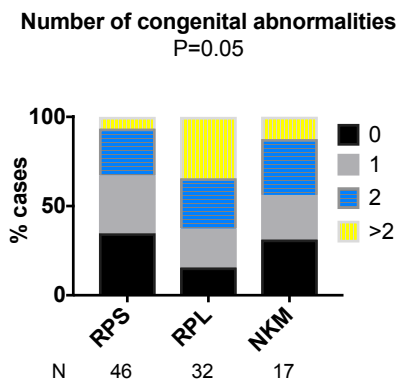
U


Neuropsychiatric disorders
P=0.2
% cases
100
50
0
RPS
RPL
NKM

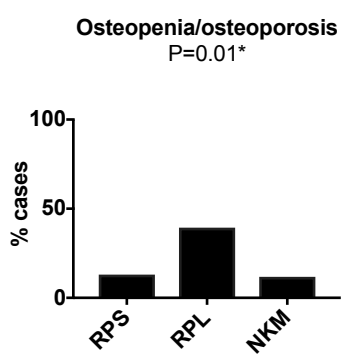
V


Gastrointestinal disorders
P=0.3
% cases
100
50
0
RPS
RPL
NKM

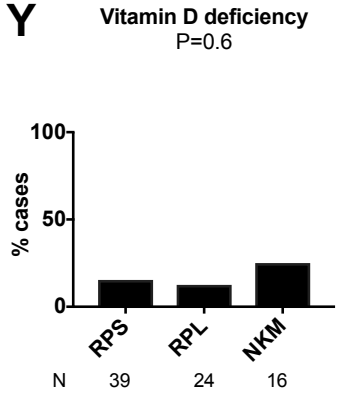
W


Number of congenital abnormalities
P=0.05
% cases
100
50
0
0
1
2
>2
RPS
RPL
NKM
N 46 32 17

X


Osteopenia/osteoporosis
P=0.01*
% cases
100
50
0
RPS
RPL
NKM

Y


Vitamin D deficiency
P=0.6
% cases
100
50
0
RPS
RPL
NKM
N 39 24 16

Z

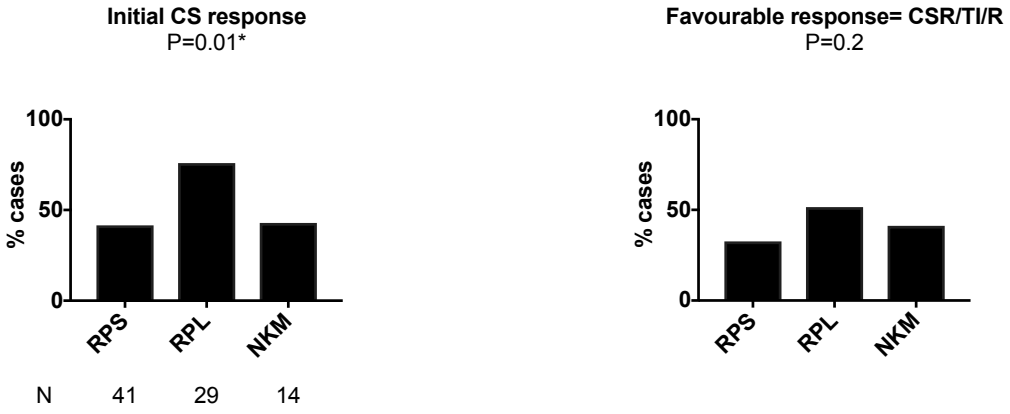
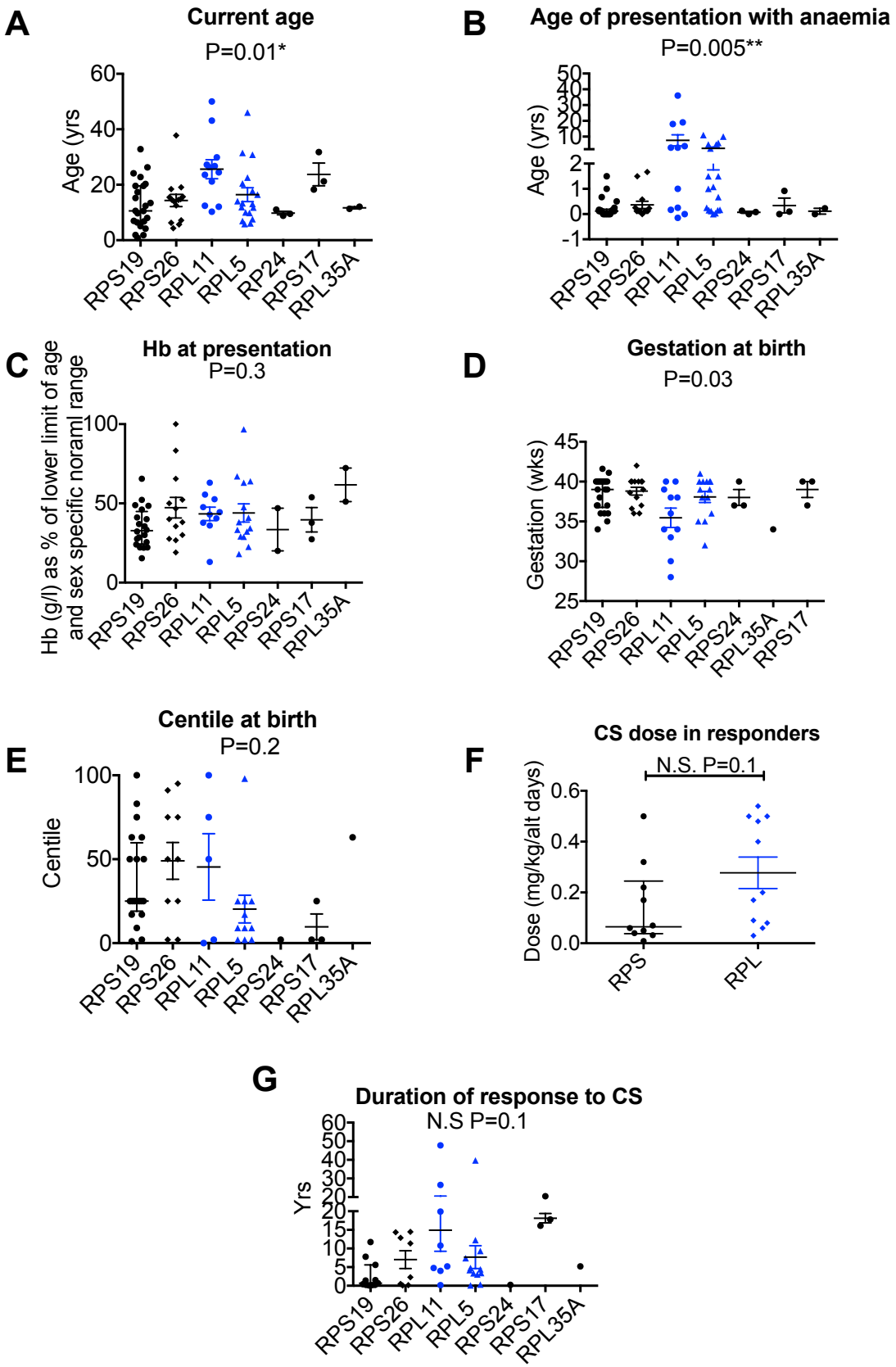
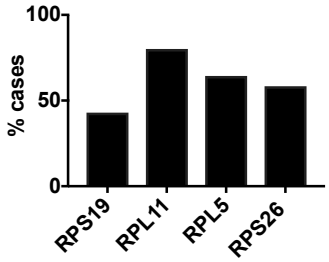


Figure 3:5 Clinical parameters according to genotype II. A three-way comparison was performed between patients with *RPS* pathogenic variants, *RPL* pathogenic variants and those in whom genetic basis is unknown (NKM). Column plots show proportions of cases in each genotypic group. P values for a three-way comparison were determined using the Pearson chi-squared test. *RPS* n=46 *RPL* n=33 and NKM n=18, unless otherwise indicated. *BW: birth weight; CSR: corticosteroid responsive; Hb: haemoglobin; IUGR: intrauterine growth restriction; SGA: small for gestational age; TI: transfusion-independent; R: remission.*


A
Current age
P=0.01*
Age (yrs
60
40
20
0
RPS19
RPS26
RPL11
RPL5
RP24
RPS17
RPL35A
B
Age of presentation with anaemia
P=0.005**
Age (yrs)
50
40
30
20
10
2
1
0
-1
RPS19
RPS26
RPL11
RPL5
RPS24
RPS17
RPL35A
C
Hb at presentation
P=0.3
Hb (g/l) as % of lower limit of age and sex specific noraml range
100
50
0
RPS19
RPS26
RPL11
RPL5
RPS24
RPS17
RPL35A
D
Gestation at birth
P=0.03
Gestation (wks)
45
40
35
30
25
RPS19
RPS26
RPL11
RPL5
RPS24
RPL35A
RPS17
E
Centile at birth
P=0.2
Centile
100
50
0
RPS19
RPS26
RPL11
RPL5
RPS24
RPS17
RPL35A
F
CS dose in responders
N.S. P=0.1
Dose (mg/kg/alt days)
0.6
0.4
0.2
0.0
RPS
RPL
G
Duration of response to CS
N.S P=0.1
Yrs
60
50
40
30
20
20
15
10
5
0
RPS19
RPS26
RPL11
RPL5
RPS24
RPS17
RPL35A

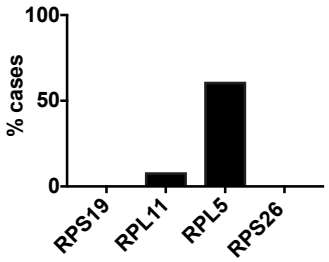
H

Short stature <10th centile
P=0.2



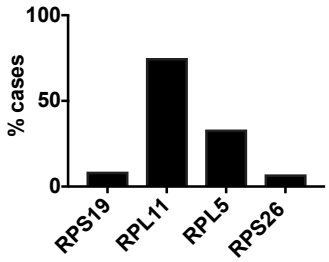
I

Cleft lip/palate
P<0.0001***



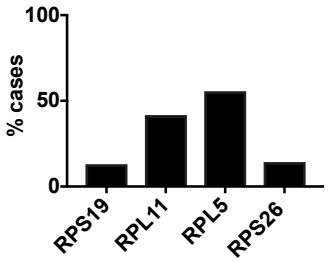
J

Thumb/radial abnormalities
P<0.0001***



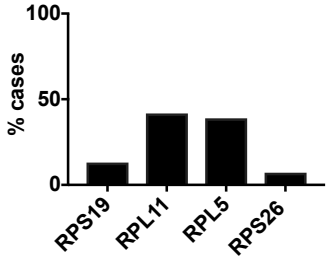
K

Otologic problems
P=0.01*



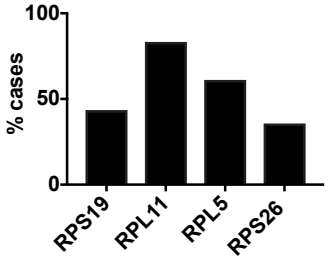
L

Osteopenia/osteoporosis
P=0.04*



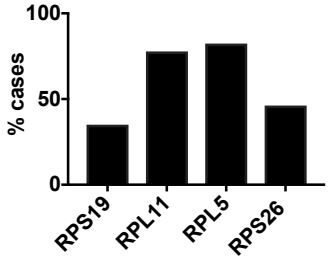
M

Birth complications
P=0.06



N

Initial CS response
P=0.01*



O

Favourable response= CSR/TI/R
P=0.03*

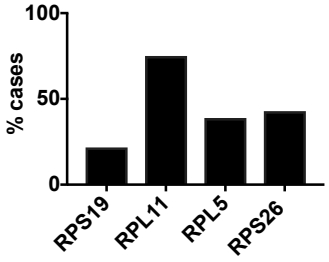


Figure 3:6 Genotype-phenotype correlations. Graphs show clinical parameters according to RP gene, for genes that are mutated/deleted in >1 patient. P values were calculated using the Kruskal-Wallis test (3.6A-E&G), the unpaired students t test (3.6F) and the Pearson chi-square test (3.6H-O). A) Patients with *RPL11* mutations were older at the time of analysis. B) Both *RPL11* and *RPL5* patients present with DBA at a later age though there is no difference in presenting Hb, which is generally below 50g/l (C) *RPL11* is associated with prematurity (D) while there is a possible association between *RPL5* and low birth weight (E). The latter is not significant when all 4 RP genes are included in the analysis as shown. F) In patients maintained on CS, the maintenance dose is similar in *RPS* and *RPL* groups. G) *RPS19* patients respond to steroids for a shorter duration than other genotypes, in particular *RPL11,* though differences are not statistically significant across all groups H-O) Significant differences between RPL and RPS genotypes (Figures 3.4&5) were further analysed in the 4 most common RP gene groups.

Data from the U.K cohort support the previous observation (Boria et al. 2010) that there is a higher incidence of birth complications in DBA relative to the general population (Table 3.2) The most common indication for complicated modes of delivery, either caesarean section (CS) or induced/forceps-assisted vaginal delivery, is fetal distress, although the indication for delivery was only available in 36 of 55 cases. Pregnancy is a recognised cause of worsening anaemia in women with DBA. Accordingly, elective CS was used to deliver two fetuses with DBA due to maternal DBA and anaemia exacerbation. Fetal complications in the U.K cohort include 4 cases of hydrops fetalis, which occurred in 2 patients with *RPS19* mutations (1 missense and 1 nonsense), 1 patient with a nonsense mutation in R*PL15* and 1 genetically undiagnosed patient; prematurity (birth at <37 weeks gestation); intrauterine growth restriction (IUGR); and small for gestational age (SGA) neonates (birth weight <10th centile for gestational age). Furthermore, there is a positive correlation between *RPL* genotype and one or more of the above complications

(P=0.004; Figure 3:5D-G) Examination of individual genes suggests that patients with *RPL11* are most likely to be born prematurely (P=0.03) and *RPL5* with low birth weight (P=0.2,Figure 3.6D&E). However, as the latter difference is not statistically significant, further work is required to delineate this further.

Table 3.2 Clinical, haematological and biochemical features of DBA in the U.K.

Parameter	N with available data	U.K cohort N (%) or median [range]
Delivery		
Spontaneous vaginal delivery	105	50 (47.6)
Caesarean section (CS; elective or emergency)		14 (13.3)
Assisted (forceps delivery or induced)		41 (39.0)
Indication for CS or assisted delivery		
• Fetal indication	36	22 (61.1)
• Maternal indication		14 (38.9)
Birth weight (kg)	36	2.9 [0.85-5.2]
Birth complications		
• Hydrops fetalis	123	4 (3.3)
• Prematurity	108	27 (25)
• SGA	87	29 (33.3)
• IUGR	107	24 (22.4)
• None	123	66 (53.7)
Neonatal issues other than anemia e.g. NICU/seizures	123	20 (16.2)
Presenting features		
Age at presentation with anaemia	116	2.6 months [22/40 in utero – 36years]
Presentation at birth		22 (19.0)
Presentation within 3 months		73 (62.9)
Presentation within 1 yr		85 (73.3)

Haemoglobin at presentation (g/L)	103	42 [10-97]
MCV at presentation (fL)	22	99.9 [68-119]
Absolute reticulocyte count at presentation (x109/L)	35	5.65 [0-39]
eADA at presentation	39	188 [63-1466]
HbF (current or at presentation)	68	0.7 [0.2-27.4]
Congenital anomalies		
● None	119	33 (27.7)
● Single anomaly		36 (30.2)
● 2 anomalies		29 (24.4)
● Multiple (≥3) anomalies		21 (17.6)
Cardiac		37 (31.1)
Radial abnormalities		27 (22.9)
Other limb e.g. clinodactyly, toe joint displacement, toe clawing		15 (12.7)
Cleft palate & Pierre-Robin anomaly		14 (11.8)
Other craniofacial e.g. Cathie facies (Cathie 1950), high arched palate, webbed neck, choanal atresia, microcephaly , craniosynostosis		23 (19.3)
Genitourinary		8 (6.7)
Axial skeleton (confirmed on imaging)		**31 (25.2)^**
● Spina bifida		3 (2.5)
● Cervical spine fusion		5 (4.2)
● Abnormal spinal curvature		13 (11.0)
● Sacral dimple requiring surgical correction		1 (0.8)
● Cervical erosion or incomplete elements		8 (6.8)

• Cervical ribs		3 (2.5)
Short stature at last measurement		
< 10th centile	102	55 (53.9)
< 0.4th centile		29 (28.4)
Co-morbidities		
Osteopenia/Osteoporosis	123	24 (19.5)
Vitamin D deficiency	94	79 (84.0)
Endocrine dysfunction	123	22 (17.9)"
• Adrenal insufficiency		5
• Thyroid dysfunction		7
• Growth Hormone deficiency		4
• Hypogonadism		6
• Accelerated puberty		1
• Hypoparathyroidism		1
• Prolactinoma		1
Poor dentition (hypomineralisation, poor enamel and decay)		16 (13.0)
Grommets/hearing aids /hearing loss/external auditory canal atresia/perforated tympanic membrane		28 (22.8)
Strabismus		8 (6.5)
Cataract		1 (0.8)
Unilateral reduced visual acuity		1 (0.8)
Hernia		10 (8.1)
• Umbilical hernia		1
• Inguinal hernia		8
• Both		1
Neuropsychiatric disorders		30 (24.4)
• Seizures		1

• Delayed motor development		3
• Delayed linguistic development		3
• Global developmental delay		3
• Mood disorder – anxiety, depression, behavioural disorder		8
• Learning difficulties /autism		12
Gastrointestinal disorders		31 (25.2)*
• Chronic diarrhoea/constipation (pseudo-obstruction in 1 case)		10
• Lactose intolerance requiring alternative diet		7
• Poor feeding and weight loss requiring NGT/PEG feeding		7
• Oesophagitis/GORD		10
• Colitis		1
• Eosinophilic oesophagitis and stricture		1
Serious/recurrent infections		
• Unequivocal	117	17 (14.5)
• Equivocal		33 (28.2)
• Not present		67 (57.3)

^2 abnormalities co-exist in 2 patients; ">1 disorder in 5 patients; "*>1 disorder in 3 patients*

3.3.3.2 Physical anomalies

In line with an SGA rate of 33.3% and likely compounded by the iatrogenic effects of therapies, short stature <10th centile for age at most recent measurement is present in 55/102 (53.9%) of patients. Of these, 29/102 (28.4%) of patients have significant short stature <0.4th centile. There is no relationship between height and specific

treatment regimens (P=0.695). Commensurate with the propensity of RPL patients to have an earlier gestation and lower birth weight, this group have a tendency to present with shorter stature later in life, both <10th centile (P=0.01) and <0.4th centile (P=0.04 when only RPS and RPL groups compared: Figure 3:5B&C). This difference is observed independent of treatment regimen.

Congenital abnormalities have been reported to occur in 47% of cases of DBA (Vlachos et al. 2001). Our study demonstrates at least one anomaly (not including short stature) in 72% of patients (Table 3.2). This may reflect more precise delineation of DBA from other congenital anaemias, due to the advances in genetic testing, and also more comprehensive investigations e.g. the majority of patients in the U.K cohort have undergone pelvic and cardiac ultrasound scanning. Importantly, anomalies were not counted if they were equivocal in terms of diagnosis e.g. Sprengel's deformity of the scapula (n=33), 'possible' spina bifida (n=2); or defects for which it was not clear whether they were congenital or acquired e.g. hydrocephalus (n=2).

Congenital cardiac and craniofacial (including cleft palate) defects are the most common, both occurring in 31.1% of patients, followed by thumb/radial abnormalities -including tri phalangeal, duplicated or bifid thumbs and radial hypoplasia (22.9%), and genitourinary (6.7%) abnormalities. In addition to these defects classically associated with DBA, abnormalities of the spine and axial skeleton are present in 25.2% of patients. There is no apparent correlation between RPL versus RPS genotype and skeletal defects (P=0.5; Figure 3:5N). However, in line with previously published data (Gazda et al. 2008), *RPL* versus *RPS* genetic variants are associated with multiple as opposed to single or no congenital abnormalities (P=0.05; Figure 3.5W). *RPL5* also correlates specifically with cleft palate (P<0.001;Figure 3.6I) and *RPL11* with thumb/radial abnormalities (P<0.001; Figure 3.6J), confirming prior studies (Gazda et al. 2008). Cleft palate is not observed in any patients with non-*RPL5/11* genes while thumb abnormalities were present in 3

patients with *RPS19* and 3 patients with *RPS26* pathogenic variants. By contrast, we did not confirm previous reports (Quarello et al. 2010) suggesting a link between RPL genotype and cardiac defects (P=0.9; Figure 3:5L), nor a link between *RPS26* pathogenic variants and genitourinary abnormalities, although the latter occur infrequently in the U.K cohort (Figure 3:5M).

3.3.3.3 Medical co-morbidities

As expected, iron overload is common in transfusion-dependent patients with DBA, but did not occur in steroid-dependent patients with no significant transfusion history. The number of recognised medical problems associated with DBA has increased over the last decade to include endocrine (Lahoti et al. 2016), dental (Ozden et al. 2011) and otologic (Kalejaiye et al. 2016; Gripp et al. 2014) manifestations. All three were prevalent in the U.K cohort (Table 3.2; Figure 3:5Q,R,T) as were some new features, not previously quantified in DBA: hernias (8.1%), neuropsychiatric (24.6%) and gastrointestinal (GI) disorders (25.2%). In 3 cases the onset of GI disorders was temporally related to use of the chelating agent deferasirox therefore these were excluded, as were 2 clear cases of corticosteroid-induced mood disturbance. All endocrine abnormalities were enumerated although in some cases they are likely to be attributable to iron overload (which can lead to hypopituitarism) or to steroid therapy (due to adrenal insufficiency). An investigation into associations between disease complications and genotypes revealed that 7/10 patients with hernias have mutations in RPS-associated genes (P=0.06 when RPS, RPL and NKM compared and P=0.01 when only *RPS* and *RPL* groups compared; Figure 3:5O), including 2 siblings with an inactivating mutation in *RPS17* and bilateral inguinal hernias and another 1 patient with an *RPS17* mutation. A further patient with a hernia has an *RPS17* VUS awaiting validation and only 1 of the patients with hernia has a mutation in *RPL5*. In addition, a novel association was identified between RPL versus RPS genotype and otologic problems (P=0.02; Figure 3:5T). This may in part be confounded by the co-existence of cleft palate, which predisposes to

chronic ear infection. Indeed 11 of the 13 patients with cleft palate have grommets or hearing loss, although otologic problems also arose in 7 patients with *RPL5/11* mutations but without clefts, suggesting an independent effect of genotype. *RPL* vs *RPS* genotype is also associated with a higher incidence of osteoporosis and osteopenia (P=0.01; Figure 3:5X) though this difference may be confounded by the older age of these patients (Figure 3.6A). The underlying genetic defect was not associated with either neuropsychiatric (P=0.2; Figure 3:5U) or gastrointestinal disorders (P=0.3; Figure 3:5V).

During the 1791 person-years encompassed by this study (calculated as the date of presentation with anaemia until the date of death or last follow-up), there were 2 incidences of malignancy in 2 different patients: B-ALL, basal cell carcinoma (BCC) and cervical intraepithelial neoplasia (CIN), as well as 1 case of myelodysplastic syndrome (MDS). The lower incidence in our cohort (2/123=1.1%) compared with the North American DBA registry (18/608=2.9%; (Vlachos et al. 2016)) may be explained by differences in the median ages of the 2 cohorts (12y versus 18y respectively) and the shorter follow-up of our patients (1791 vs 9458y). Two deaths have occurred in the U.K cohort: one caused by infection secondary to chelation-induced agranulocytosis and one due to idiopathic pulmonary syndrome post-HSCT.

3.3.4 Clinical outcomes

The natural history of DBA in the U.K cohort is outlined in Table 3.3. In summary, 42% of patients are TD, 25% are CS-dependent, 9% are in remission, 4% have never required treatment for anaemia, 2% have died and 18% have received HSCT. Indications for the latter include transfusion dependence (n=19), progressive BMF and/or MDS (n=2) and a family request for HSCT in a case of steroid- responsive anaemia with occasional transfusion requirements (n=1). HSCT approaches and outcomes in the U.K cohort have been reported elsewhere (O 039 Boyle et al. 2016).

Additionally, 3 patients failed a metoclopramide trial, 1 a trial of ciclosporin and 1 a trial of EPO. To our knowledge, L-leucine has not been used in any U.K patients.

A lower proportion of patients in the U.K cohort (53.4%; Table 3.3) than the previously reported (approximately 70%) (Ball et al. 1996 ; Vlachos & E. Muir 2010) responded initially to high dose CS, defined as a sustained increase in Hb and reticulocytes during the first 4 weeks of therapy and consequently transfusion independence (Ball et al. 1996). However, similar proportions of U.K and North American patients – 31% versus 37% (Vlachos et al. 2014)- maintained a normal Hb on long-term CS. Three patients who were initially primary refractory to CS were subjected to a second trial but none responded, while a further 2 patients who lost their CS response on weaning were successfully reinitiated onto CS (Table 3.4). In 8/33 patients, although successfully maintained on CS, transfusion support has been required intermittently e.g. during puberty, pregnancy, following infection or with no clear precipitant.

Table 3.3 Clinical outcomes in the U.K compared with published data.

Parameter	N with available data	U.K cohort N (%) or median [range]	Italy cohort {Campagnoli:2004vm}
CS trial			
• Performed		106 (86.2)	
• Not performed		11 (8.9)	
• Not known/ documented		6 (4)	
Response to CS trial	106		
Primary refractory		49 (46.2)	
Initially responsive		57 (53.4)	
○ Responsive but stopped due to side-effects		6 (5.7)	
○ Responsive and dependent		33 (31.1)	
○ Remission on withdrawal		4 (3.8)	
○ Secondary refractory		14 (13.3)	
Transfusion-dependent		51 (41.5)	42.5%
● CS failure		41	
● CS intolerance		4	
● CS not yet trialled		6	
CS- responsive		31 (25.2)	19%
Anaemia not requiring treatment		6 (4.9)	NA
Remission post-CS		5 (4.1)	21%*
Remission post-transfusions (spontaneous remission)		6 (4.9)	
HSCT		22 (17.9)	9.4%
Deceased		2 (1.6)	8.2%

** not clear if remission occurred post CS*

Table 3.4 Clinical outcomes in patients who underwent 2 CS trials.

ID	Current age (y)	RP gene mutated	Age at first CS trial	Response	Reason for CS discontinuation	Indication for second CS trial	Age at second CS trial	Response
001	11.2	*RPS10*	4y	CSR	Remission for 4y	Worsening anaemia at puberty	10y	CSR
067	20.4	*RPL5*	Childhood; exact age NA	CSR	Remission for >10y	Neutropenia and worsening anaemia	16y	CSR
090	37.7	*RPS26*	10y	CSR	Remission on discontinuation after 1y	Worsening anaemia	17y	CSR and repeat remission on discontinuation after 6m
074	9.5	*RPS24*	9m	CSR	Remission on discontinuation after 1y	Pre-HSCT	3y	PR
092	5.6	*RPS26*	1y	Stopped at 3/52	CSI-neuromuscular toxocity	Inadequate first trial	2y	PR
065	17.7	*RPL5*	12y	CSR for 6/12 then CSI	CSI-Bowel perforation after 6/12	Previously CSR	15y	CSR but discontinued 6/12 later as intolerant of side effects
066	10.2	*RPL5*	3y	CSR then lost respons	SR	Pt request	5y	CSR

				e on weaning				
087	4.3	*RPS26*	1y	CSR but secondary refractory after 1m on weaning - recurrent infections	SR	Lost response in context of infection so second trial attempted	2y	CSR
022	32.8	*RPS19*	5y	PR	PR	Pt request	10y	PR
081	11.9	*RPL5*	6m	PR	PR	Pt request	Childhood	PR
062	10	*RPL5*	5m	CSR but stopped due to effects on growth	CSI	Worsening anaemia	5y	PR
40	9.4	ND	1y	PR	PR	Not sure of dose used in first trial	NA	PR

CSI: corticosteroid intolerant; CSR: corticosteroid responsive; NA: not available; ND: not detected; PR: primary refractory; SR: secondary refractory.

Next, associations between patient factors and clinical outcomes were investigated (Figure 3:5C). In keeping with the different clinical phenotypes observed between RPL and RPS genotypes, differential responses to treatment were also identified in these 2 groups. Specifically, both *RPL5*- and *RPL11*-mutated patients were more likely than RPS-mutated patients to exhibit an initial response to a steroid trial (P=0.01; Figure 3:5 Clinical parameters according to genotype II). Importantly there was a high rate of secondary loss of response to CS in *RPL5* patients (Figure 3.6N&O) therefore only patients with *RPL11* mutations show a more favourable disease outcome (P=0.03; Figure 3.60), defined as CS-responsiveness (CSR), remission (R) or never having required any treatment/treatment independence (TI). Conversely, patients with non-*RPL11* lesions, in particular *RPS19,* are more likely to be

dependent upon transfusions or HSCT**,** regimens that are known to be associated with higher morbidity and mortality (Lipton et al. 2006)**.**

3.4 Discussion

This is a retrospective single centre cohort study using prospectively acquired clinical and laboratory data from 123 unselected patients with DBA in the U.K. A particular strength of this study is that patients have been assessed at a single centre and have been subjected to the same clinical coding strategy and investigations. In particular, the comprehensive NGS-based genetic screening approach detected at least 70% of CPV/LPV mutations, compared with average pick-up rates of 50-60% for most existing multi-gene panels (Farrar et al. 2011; Doherty et al. 2010). This allowed robust genetic annotation to complement the comprehensive clinical annotation of patients.

Importantly, this analysis of a large cohort has confirmed several aspects of published experience: the spectrum of congenital anomalies and medical co-morbidities (Table 1.2); the fact that *RPS19* is the most commonly mutated gene in DBA (Figure 3.2); and known associations between *RPL5/11* and multiple congenital defects, and specifically with cleft lip/thumb abnormalities, respectively (Figure 3.6I-J). However, importantly this analysis has highlighted several novel disease features, pertaining to epidemiology and the role of gender, clinical features, and genotype-phenotype correlations. These new insights are discussed below.

In the U.K cohort, 78% of mutations have arisen *de novo,* with 22% confirmed cases of AD inheritance. One limitation is that both parents have been screened in only 39 of the 117 families. Interestingly, maternal transmission is implicated in 10 of the 11 confirmed pedigrees with AD inheritance, in keeping with an observation tentatively made by another group (Carlston et al. 2017). The reason for this is unclear. In 4

cases of sporadic DBA reported in a case series (Orfali et al. 2004), parental DNA was available and contained informative markers for determining the paternal origin of the sporadic mutation. In all 4 cases, the pathogenic mutation was inherited from the paternal germ line. The incidental finding in this study of 2 pedigrees with putative germline mosaicism suggests that the incidence of heritable DBA may be under-estimated. It was not possible to determine whether mosaicism occurred in the mother or father's gametes in either family nor to fully exclude the possibility of low level somatic mosaicism below the lower limit of Sanger sequencing. Although in family 2 (Figure 3.3), the *RPS19* missense variant has not previously been associated with DBA, it is the only variant detected by NGS in both affected cases in the family and is therefore likely to be pathogenic, though confirmatory studies are required. Notably, there are two previous reports of inheritance by gonadal mosaicism (Cmejla et al. 2000; Gazda et al. 2004), one of which describes 2 affected sisters with a missense variant in *RPS19*, absent from both their parents (Gazda et al. 2004).

The finding that eADA segregated separately from the pathogenic mutation in 2 pedigrees raises the possibility that polygenic factors are implicated in the elevation of eADA and potentially in the other phenotypic features of DBA.

Although *RPS19* remains the most commonly implicated gene in DBA, as a result of improved detection of other RP gene variants, *RPS19* variants account for 20.5% of U.K DBA cases as opposed to the historical figure of 25% of cases (Willig, Draptchinskaia, et al. 1999a) The reason for a higher incidence of RPL mutations in the U.K compared with other cohorts (Gazda et al. 2008 (Clinton & Gazda 2009)) is not clear.

Interestingly, DNA variants were identified in only 20 of the 80 RP genes screened. Of these, 11 are known DBA genes while 8 are novel. Further work will involve validation of the 8 putative mutations in novel RP genes (2 in 1 patient; Table 3.1). It is also important to understand why DBA-causing variants cluster in a handful of RP

genes, while others are spared. As a case in point, a mutation in *RPS14* has never been reported in DBA, though somatic mono-allelic LOF variants in this gene are associated with a variant of MDS, the 5q- syndrome (Ebert et al. 2008). In addition, studies using the *Drosophila megaloblaster* model show that not all RP gene mutations result in the Minute phenotype (Marygold et al. 2007); this has been presumed to be due to disparate expression of distinct RP.

No bigenic/biallelic cases were detected in the UK series and there were no detectable mutations in *GATA-1*, consistent with the absence of dyserythropoiesis, a characteristic BM morphological finding in the context of *GATA-1* mutations. Whole exome/genome approaches are likely to be required to identify the pathogenic lesions in the 18 patients in whom no mutations were identified by targeted NGS. Candidate genes include *TSR2* (a chaperone of RPS26), which has been implicated in one family with X-linked DBA (Gripp et al. 2014) as well as other ribosome assembly proteins. Two of the genetically undiagnosed probands are from consanguineous pedigrees, raising the possibility of autosomal recessive inheritance, although this inheritance pattern has not yet been genetically confirmed in DBA. The presence of as yet undetected RP gene variants is also feasible given that our RP gene targeted sequencing approach does not cover gene regulatory areas in their entirety. Another potential outcome is that some patients will be found to have an IBMFS other than DBA, given the shared features of these conditions (Table 1.1). Indeed, it has recently been recognised that the mitochondrial disease Pearson marrow-pancreas syndrome may masquerade as DBA (Gagne et al. 2014) and also that 6% of the genetically undiagnosed cases of DBA in the German registry are caused by biallelic compound heterozygoous mutations in *CECR1,* leading to a form of ADA2 deficiency that is characterised by pure red cell aplasia, normal MCV/eADA and hypogammaglobulinaemia, in the absence of a vasculopathy (Szvetnik et al. 2017).

This study elucidated some new clinical features of DBA, including abnormalities of the axial skeleton, GI and neuropsychiatric disorders. Interestingly, in one patient the

neurological deficit predated the onset of anaemia by a few months. Although learning difficulties have been reported in DBA, so far this has exclusively been in patients with large deletions encompassing additional genes flanking the deleted RP gene (Campagnoli et al. 2004 (Alkhunaizi et al. 2017). In contrast, in this study there is no apparent association between large deletions, i.e., whole allele deletions versus nucleotide deletions, and specific clinical features, such as developmental delay.

One limitation of my study is that not all patients were subjected to the same range of investigations, for example spinal imaging was performed only in patients with back pain or a clinical suspicion of scoliosis, and endocrine tests were tailored to those with relevant risk factors, symptoms and signs. Thus the incidences reported in Table 3.2 may be underestimates. Prospective longitudinal studies are now warranted to better characterise these co-morbidities and to confirm whether they are intrinsic to DBA or arise as complications of treatment.

Unexpectedly, my study shows a propensity for *RPL*-affected patients to be female while gender is more evenly distributed in the *RPS* group. There are various potential explanations for this intriguing observation: one is increased loss of male fetuses with *RPL* gene mutation. The risk of miscarriage is known to be higher in mothers of patients with DBA (Giri et al. 2017; Ball et al. 1996) though it is not yet known whether particular genotypes are more susceptible to miscarriage. Another possibility is that males with *RPL* gene mutations exhibit a milder phenotype than females and are therefore under diagnosed. While it has previously been suggested that physical anomalies occur more frequently in males than females (Ball et al. 1996), this observation was not confirmed in the current study (P=0.7). It is known that a proportion of patients undergoing repair of isolated congenital cardiac defects and cleft palate, have unexplained macrocytosis (*Vlachos A, unpublished*), which could be a feature of subclinical DBA, but the genotypes and gender mix of these putative non-classical DBA patients are yet to be determined.

In terms of clinical presentation, the U.K DBA cohort has a high prevalence of peri- and neo-natal complications, as previously noted by Willig *et al* (Willig, Niemeyer, et al. 1999b) and more recently in the DBAR (Giri et al. 2017). Neonatal complications such as seizures and ventilatory problems were associated with anaemia at birth in 10 cases whereas Hb was normal in the remaining 10 cases, which instead presented with overt DBA later. One area for further exploration is the reason why anaemia is uncommon in fetal life in DBA, as reflected by the low incidence of hydrops fetalis, both in our cohort (Table 3.2) and in the published literature (E. G. Zhang et al. 2011; Da Costa et al. 2013; McLennan et al. 1996; Dunbar et al. 2003).

Traditionally DBA has been termed 'classical', characterised by presentation in infancy with severe anaemia, or 'non-classical', characterised by an older age of onset with a milder haematological phenotype. Despite the high frequency of birth complications and the presence of anaemia at birth in 19.0% of cases (Table 3.2), one quarter of patients in the U.K cohort presented with sustained anaemia *after* 1y. This higher proportion is explained by the higher frequency of RPL-gene associated patients in the U.K cohort and the novel finding that RPL genotype is an independent predictor of later age of presentation to haematologists. Age of presentation is in fact bimodal with a propensity for RPS-gene associated DBA to present <18m and RPL-gene associated DBA either < or >18m. Consistent with this novel observation, to my knowledge all published case reports describing atypical presentations of DBA in adulthood, refer to patients with pathogenic variants in *RPL5* or *RPL11* (Carlston et al. 2017; Narla et al. 2016; Flores Ballester et al. 2015; Steinberg-Shemer et al. 2016), supporting this notion that so-called 'non-classical' versus 'classical' DBA are in fact the clinical manifestations of *RPL* and *RPS* genotypes respectively.

Despite the later manifestation of the haematological defect in patients with pathogenic variants in RPL genes, they exhibit a propensity for congenital anomalies, in line with published experience. Specifically, RPL are more likely than RPS-gene associated casesor patients with unknown mutations to have multiple rather than

single defects. Cleft palate and thumb/radial abnormalities are associated with *RPL5* and *RPL11*, respectively. This study also revealed novel associations between *RPL*, variants and hearing difficulties as well as short stature.

With regards to treatment regimens, a higher proportion of patients in the U.K cohort than in the DBAR were primarily refractory to CS (46.2% vs. 20%), despite the use of appropriately high doses (2mg/kg/d) for least 4 weeks at induction of CS therapy. There are 3 potential explanations for this discrepancy: 1) use of different therapeutic regimens and different definitions of CS responsiveness across international centres, 2) true differences in CS responsiveness in distinct geographical areas, consequent upon genetic or environmental factors or 3) ascertainment bias resulting from more complex (and more likely to be refractory) cases of DBA being notified to ICHT. The latter possibility cannot be excluded, although to our knowledge the cohort encompasses most of the cases in the U.K and the remission rate corroborates published experience (Alter & Young 2003). Specifically, 11/123 patients are currently in haematological remission and there is a history of transient treatment independence in an additional 6/123 patients (median length of remission 3y, range 0.5- 13y). Classically remission has been thought to arise more commonly in CS-responsive patients (Vlachos & E. Muir 2010) however in this study equal proportions of patients entered remission while on blood transfusions and on CS therapy. A further 6/123 patients with a median age of 13.6y (range 3.1-43.1y) have never required regular transfusions or CS therapy for anaemia, reflecting the known variable penetrance of the disease, and also suggesting that the ICHT cohort is not biased towards severe disease.

To date, there has been little evidence for an association between genotype and haematological phenotype nor steroid responsiveness. In the Canadian registry a higher proportion of patients with *RPL11* versus *RPS19* were recently reported to have mild anaemia not requiring treatment though in total there were only 11 and 7 patients in each category (Arbiv et al. 2017). In contrast, the French group have

postulated that patients with *RPL5/11* causal variants have a more severe phenotype on the basis that they have a higher incidence of congenital malformations- however it is important to distinguish between haematological and non-haematological features. Indeed in the French cohort there is no difference in the rate of treatment independence in the R*PS19* group (10/66=15.2 %) versus the *RPL5/11* group (6/41=14.6%). In our cohort a higher proportion of *RPL11* (3/13=23.1%) versus *RPS19* (1/24=4.2) patients have not required treatment, though the difference is not significant (P=0.1).

Importantly, this work demonstrates that a pathogenic mutation in *RPL11* or *RPL5* rather than in *RPS19* or other RPS genes, predicts for a higher rate of steroid responsiveness and, in the case of *RPL11*, a more favourable long-term outcome (Figure 3.6N-O). This was defined as CS-responsiveness, remission or not requiring treatment, given that these clinical outcomes have been associated with improved survival (Lipton et al. 2006). Specifically the overall actuarial survival at greater than 40 years of age has been reported to be 86·7% for corticosteroid- versus 57·2% for transfusion-dependent patients (Vlachos & E. Muir 2010). Overall my data suggest a novel genotype-phenotype model of DBA in which *RPS19* and *RPL11* deficiency are prototypes of classical and non-classical DBA respectively.

These findings support the need for an international collaborative approach to further characterise these differences, to confirm the clinical phenotypes associated with the rarer RP genes involved in DBA, and to elucidate their underlying mechanisms, with the long-term aim of providing patients with individualised therapies and prognostication.

4 CHARACTERISATION OF HAEMATOPOIESIS AND ERYTHROPOIESIS IN DBA BY PROSPECTIVE ISOLATION OF HUMAN ERYTHROID PROGENITORS

4.1 Introduction and aims

As outlined in Chapter 1, normochromic, macrocytic anaemia and reticulocytopenia are the haematological hallmarks of DBA. These stem from a selective paucity of erythroid precursors/EB in the BM (Vlachos et al. 2014). Indeed, absence or reduction of EB to <5% of BM nucleated cells is one of the major diagnostic criteria. However, BM examination does not reveal where along the haematopoietic hierarchy (Figure 1:4) the erythroid defect arises.

Animal models of DBA and *in vitro* studies of human unfractionated BM or total CD34+ cells have suggested that red cell aplasia in DBA arises from a defect in EP (Nathan, Clarke, et al. 1978a; Devlin et al. 2010), This defect has been defined functionally by a reduced ability of DBA BM to generate burst-forming unit-erythroid (BFU-E) and colony-forming unit-erythroid (CFU-E) colonies (Gregory & Eaves 1977; Hattangadi et al. 2011). For example, an early study showed absent BFU-E and reduced CFU-E frequencies in 7 patients with DBA who had relapsed after a period of remission (Nathan, Clarke, et al. 1978a). On this basis, the prevailing concept is that red cell aplasia arises from an early erythropoietic defect, between the BFU-E and CFU-E stages (Da Costa et al. 2010). Supporting this notion, a reduction in BFU-E activity has been reported in human CD34+ cells infected with *RPS19* shRNAs (Ebert 2005).

By contrast, other studies have demonstrated normal BFU-E/CFU-E frequencies with impaired CFU-E differentiation (Lipton et al. 1986). A defect in the late precursor stage of erythropoiesis has also been demonstrated in mouse models (Devlin et al. 2010; Jaako et al. 2011) and in an *in vitro* erythroid culture generated from DBA patient PBMCs. Specifically, erythroid development, assessed by cell number and cell surface markers of differentiation, was normal until the proEB stage, followed by impaired proliferation and differentiation (Ohene-Abuakwa et al. 2005). This was

attributed to relative EPO insensitivity of mature EB populations, a mechanism that has also been put forward by other investigators (Lipton et al. 1986). The identification of BFU-E/CFU-E immunophenotypes in murine fetal liver (Flygare 2005) has facilitated important insights into erythropoiesis (Flygare et al. 2011; L. Zhang et al. 2013) whereas in humans, definitions of corresponding EP are limited, precluding their direct study in normal and aberrant erythropoiesis. I hypothesised that knowledge of the cell surface markers that define the BM cells giving rise to BFU-E and CFU-E would allow elucidation of the relative roles of qualitative and quantitative defects in DBA erythropoiesis.

Although DBA classically causes an isolated anaemia, thrombocytosis, thrombocytopenia or neutropenia are not uncommon in patients (Willig, Niemeyer, et al. 1999b) and occur in animal models of RP haploinsufficiency (Danilova et al. 2011; Jaako et al. 2011). From these observations we can infer a more generalised disturbance of haematopoiesis, supported by granulo-macrophage defects observed in long-term BM cultures (Santucci et al. 1999) and impairment of myeloid (M) and megakaryocyte (MK) formation in DBA patient-derived iPSCs (Garçon et al. 2013). In one study the frequency of CFU-GM progenitors was normal in 10 patients within one year of diagnosis but reduced in 7/14 patients 3 years after diagnosis, suggesting a worsening of multipotent progenitor function with age (Casadevall et al. 1994). Otherwise, the development and functional characteristics of HSPC in DBA are poorly characterised.

In this chapter I aimed to define the cellular landscape of haematopoiesis in DBA, and to dissect the erythropoietic defect in DBA by identification and prospective isolation of human EP that are analogous to BFU-E and CFU-E progenitors.

4.2 Methodological approaches

BM samples were obtained from patients with DBA and healthy BM donors under approval by the National Research Ethics Service, and following written informed consent, as described in Section 2.1.1. Clinical details and laboratory investigations for normal donors/controls (CON) and patients with DBA are summarized in Table 4.1. More detailed clinical information regarding patients with DBA is listed in Table 4.2.

Table 4.1 Clinical and haematological characteristics of normal donors and patients with DBA.

	CON N=23	DBA N=47
Median age (range)	7.6y (1.6-43y)	8.2y(0.8-34y)
Gender (M:F)	1:1	1:1
Indication for BM	Donor for bone marrow transplant	32 surveillance, 8 pre HSCT, 6 cytopenias, 1 diagnostic
Haemoglobin electrophoresis	8 normal, 6 HbAS, 5 βTT, 4 αTT	Raised HbF consistent with DBA, no haemoglobinopathies
Median Hb g/l (range)	113 (92-143)	105(67-146)*
Median WBC count x10^9/l (range)	7.2 (4.5-14.5)	7.7 (2.8-12.6)
Median platelet count x10^9/l (range)	272 (210-448)	293 (113-502)
Treatment	NA	See Table 4.2

**Blood results and treatment regimens pertain to the time of BM sampling, not diagnosis.*

Flow-cytometry, cell sorting, clonogenic assays, liquid cultures and qPCR were performed as described in Chapter 2. Antibodies for flow cytometry and TaqMan qPCR probes are listed in Appendices C&D.

I initially focused my analysis on TD patients (who underwent BM sampling at varying time points from diagnosis) as I anticipated that any HSPC defects inherent in DBA would be more evident in this subgroup, compared with patients who did not require any treatment for anaemia or who were successfully treated with CS. Also, BM sampling was more commonly performed in TD patients e.g. prior to HSCT. In Chapter 5 I will explore the effects of CS on haematopoiesis, however here I compare TD DBA patients with CON individuals and with patients with DBA who are treatment- independent or are in remission (TI/R).

Given that, as discussed in Chapter 3, there are differences in haematological phenotypes between RPL and RPS genotypes, and because of the limited numbers of BM samples from TD *RPL* patients, initially I focused my study on patients with mutations/deletions in RPS genes and those with no known mutation (NKM); either because NGS had not yet been performed (N=1) or because no mutation was detected on screening (N=3).

Descriptive statistics in this chapter show mean ± S.E.M. The unpaired Student *t*-test or Mann-Whitney U test were used as appropriate (GraphPad Prism v6). Differences between groups are not significant (N.S) unless otherwise indicated. Number of asterix indicate the level of significance as follows: **P*<05; ***P*<.01; *** *P*<.001; **** *P*<.0001

Table 4.2 Clinical characteristics of patients with DBA from whom BM samples were collected.

Sample ID	Sex Male=1 Female =0	Inherited mutation	Ethnicity	Age at presentation with persistent anaemia (yrs)	Age at time of BM (yrs)	DNA variant	Presenting Hb (g/l)	eADA (40-100nmol/h/mgHb)	Congenital anomalies	Age at time of BM (yrs)	Treatment at time of BM (incl. CS dose in mg/kg/alt days)	CS response	Hb (g/l)	WBC x109/L	Plts x109/L	Neut x109/L
152	1	Yes= mother asymptomatic carrier	1	Birth	1.16	RPS19 C[57DELA]:[=] P.[ALA20Profs*9]:(=)	42		Hypospadias	1.16	TD	SNT	125	9.1	251	1.7
159	1		4	0.15	1.06	2.82MB DELETION INCL RPS19	20		ASD, craniosynostosis	1.06	TD	PR	112	12.1	355	5.3
157	1	Possible germline mosaic	1	Birth	8.19	RPS24 1 c.2T>G p.Met1Arg	61		R choanal atresia, PDA, clinodactly, Sprengel's, phimosis	8.19	TD	PR	120	7.7	179	6.1
156	1		1	0.17	1.19	RPS19 c.412-2A>C	30		Cathie facies, short neck, ASD	1.19	TD	SNT	111	8.3	170	1.4
144	1		2	0.12	13.04	RPS26 c.55C>T P.GLN19*	24		None	13.04	TD	PR	146	6.3	286	3.5
123	As for 144			0.12	8.7					12.3	TD					
111	1	No	1	0.67	17.19	RPS19 c.328delC p.Leu110*	24		Thumb defect, Sprengel's	17.19	TD	PR	126	8.9	202	7.9
143	0		1	0.04	1.50	NT- Awaited	58		PDA, PFO	1.50	TD	PR	103	8.6	253	1.4
124	1		3	0.25	16.60	ND		190	Hydronephrosis L kidney , microcephaly, Cathie facies	16.60	CSR	CSR	67	3.2	188	2.8
141	1		1	1.17	2.28	Not confirmed: RPS29 c.139G>C VUS	*54*	HIGH	None	2.28	TD	PR	75	3.5	295	8.2
155	0		4	0.17	17.00	RPS19 initiation codon atg1att	59		None	17.00	Remission	CSR--> Remission	124	4.1	181	1.7
110	0	No	1	1.08	2.73	RPL5 whole allele deletion	36	186	Wide fonatenlle, reduction tehnar eminence	2.73	TI	SR	87	5.3	275	2
162	As for 110			1.08	5.6						TD					
103	0	No	1	1.42	8.07	RPS10 c.337C>T p.Arg113*	17	184	Sprengel's	8.07	Remission	CSR	78	4.9	323	1.2
92	0		1	1.50	34.36	RPS26 splice donor c.3+1 G>T	40		Sprengel's, vestigial thumb L hand	34.36	Remission	CSR-- >Remission	120	4.4	335	2.7
87	1		1	0.23	9.48	Not confirmed: RPS3A c.737C>A VUS in new gene	32		Cathie facies, Sprengel's, mild S shaped scoliosis of lower thoracic/lumbar spine	9.48	CSR	CSR	97	5.5	279	3.8

Sample ID	Sex Male=1 Female =0	Inherited mutation	Ethnicity	Age at presentation with persistent anaemia (yrs)	Age at time of BM (yrs)	DNA variant	Presenting Hb (g/l)	eADA (40-100nmol/h/mgHb)	Congenital anomalies	Age at time of BM (yrs)	Treatment at time of BM (incl. CS dose in mg/kg/alt days)	CS response	Hb (g/l)	WBC x109/L	Plts x109/L	Neut x109/L
116	0	No	1	11.00	12.64	RPL5 c.175_176del GA	70	191	ASD, cleft palate, spinal abnormalities	12.64	TD	CSI	104	5.4	255	4.9
137	1	No	1	0.25	10.92	RPL5 c.166-169del			Cleft palate	10.92	TD	SR	ua	ua	ua	ua
153	0	No	3	0.25	4.89	RPL11 c.204delT	40		None	4.89	CSR	CSR	108	11.4	418	11.2
129	0		1	0.17	19.69	RPL11c.267-268insG	50		ASD, triphalangeal thumb	19.69	TD	SR	123	6.7	338	4
102	As for 102			0.17	18					18	TD					
104	0		1	Birth	3.37	RPS19 c.3G>A p.Met1Ile	48		VSD, Cathie facies, high arched palate, wide neck	3.37	TD	PR	75	6.4	135	4.6
126	0		1	1.50	5.34	RPL5 c.175-176delGA	50		Cleft palae, Pierre-Robin jaw, reduced thenar eminence	5.34	TD	SNT	102	9.9	502	5.2
72	0		1	0.08	15.93	RPS19 exon 3 splice donor c172+2T>G	22		Thenar hypoplasia	15.93	TD	PR	92	5	217	1.9
73	1	No	1	Birth	6.95	RPS19 c.3G>A p.Met1Ile	63		Clindodactyly, Sprengel's, failure of fusion of posterior elements of C6	6.95	TD	PR	91	12.6	333	9.9H
70	0		1	1.58	5.60	ND	70	101	Hypoplastic L thumbs, absent R thumb, dysplastic middle phalanxes of little fingers, cleft palate, micrognathia, shortened forearms	5.60	TI	SNT	101	5	113	1.3
79	1	No	5	Birth	1.58	RPS19 c.167G>T p.Arg56Leu	52	171	PDA and PFO	1.58	TD	PR	105	9.7	297	7.9
131	0	No	1	0.15	2.49	RPL5 c.176_177delAT	29		Small PFO with persistent L to R shunt	2.49	TD	SR	72	6.2	349	1.1
115	0		1	0.19	23.05	RPL11 c.44-45delTTinsCCCATC p.Lys15Profs*41			None	23.05	CSR	CSR	89	7.9	456	7
76	0	No	1	0.12	8.40	RPL5 c.664C>T {q222*}	26		Triphalangeal thumb, Pierre-Robin anomaly, fusion of cervical vertebrae	8.40	TD	PR	87	5.6	353	2.3
145	0	Possible germline mosaic	1	0.12	4.14	RPS19 c.361C>G p.Arg121Gly	30	172	Short neck	4.14	TD	SR	130	7.7	218	3.5

Sample ID	Sex Male=1 Female =0	Inherited mutation	Ethnicity	Age at presentation with persistent anaemia (yrs)	Age at time of BM (yrs)	DNA variant	Presenting Hb (g/l)	eADA (40-100nmol/h/mgHb)	Congenital anomalies	Age at time of BM (yrs)	Treatment at time of BM (incl. CS dose in mg/kg/alt days)	CS response	Hb (g/l)	WBC x109/L	Plts x109/L	Neut x109/L
60	1	No	1	0.23	8.2	RPL35A whole allele deletion	46		None	8.2	CSR	CSR	92	2.8	176	0.7
88	0	No	1	0.33	17	RPL5 c.625_626insG p.Tyr210Leufs*7	87	261	None	17	CSR	CSR	113	2.8	268	1
20	AS FOR 088			0.33	15.6					15.6	TI					
112	1		1	0.33	9.3	RPL11 c.475_476delAA p.Lys159Argfs*	34	305	PDA, thumb abnormalities	9.3	CSR	CSR	101	6.6	329	2.2
90	1	Y- MOTHER AND 2 SIBS	1	0.12	12.1	RPS26 c.3+1G>T, p.0?	73	138	Duplicate terminal phalanx toe, Sprengel's deformity	12.1	CSR	CSR	115	6.3	426	4.1
97	0	Elevated eADA in mother but NKM	1	0.23	3	ND	69	251	Webbed neck	3	CSR	CSR	124	11.3	220	6.8
69	1	No	1	0.75	1.9	RPS29 c.63-3C>A p.0?	25	293	Scoliosis	1.9	TD	SR	97	6.5	324	1.9
71	1		1	0.12	9.75	RPS19 c.3G>A p.Met1?	47		Scoliosis	9.75	TD	SR	116	10.9	291	7.3
136	1	No	1	Birth	19.6	RPS19 c.167G>T p.Arg56Leu	52	171	PDA, PFO	19.6	TD	SNT	105	9.7	302	7.9
89	1	Yes- MOTHER AND 2 SIBS DBA	1	0.12	0.8	RPS26 c.3+1G>T, p.0?	59		None	0.8	TD	SNT	121	12	278	4.8
94	0		1	0.23	6	ND	19		Sprengel's deformity	6	TD	SR	112	10.5	447	3.5
96	1	Yes- MOTHER DBA	1	0.15	15.2	RPS26 c.3+1G>A p.0?	38		None	15.2	TD	PR	96	7.7	247	5
99	0		1	0.25	2.3	RPS26 whole allele deletion	42		Sprengel's deformity	2.3	TD	SNT	109	12.5	401	4.5
105	0	No	1	0.12	4	RPS26 c.6_9delAAAG p.Lys4Glufs*40	32		VSD	4	TD	SR	124	10.9	319	4.9
107	1	No	3	Birth	3.1	RPS26 c.176_177TC>AA p.Phe59*	68		None	3.1	TD	PR	114	12.5	394	3.5
73	1		1	Birth	6.9	RPS19 c.3G>A p.Met1?	63		Clinodactyly	6.9	TD	PR	91	12.6	333	9.5

Hb- haemoglobin; H-high; Neut- neutrophils; ND- not determined; Plts- platelets; PDA- Patent Ductus Arteriosus; PFO- Patent Foramen Ovale; RP- ribosomal protein; SNT- steroids not tried; CSR: corticosteroid responsive= Hb > 8-9 g/dL, without the need for transfusions; PR- steroid unresponsive, primary refractory= steroid response (Hb > 8-9 g/dL) not achieved within 4 weeks therapy of prednisolone 2 mg/kg/daily; SR- steroid unresponsive, secondary refractory= transfusion requirement on weaning steroid dose below 2mg/kg; TD- transfusion dependent; VSD-Ventricular Septal Defect; WCC- white blood cell count. Ethnicity codes as follows: 1= Caucasian, 2=Black African, 3=Asian, 4=Mixed, 5=Other.

4.3 Results

4.3.1 Characterisation of HSPC populations in normal and DBA BM

4.3.1.1 Immunophenotypic characterisation of HSPC populations in paediatric BM

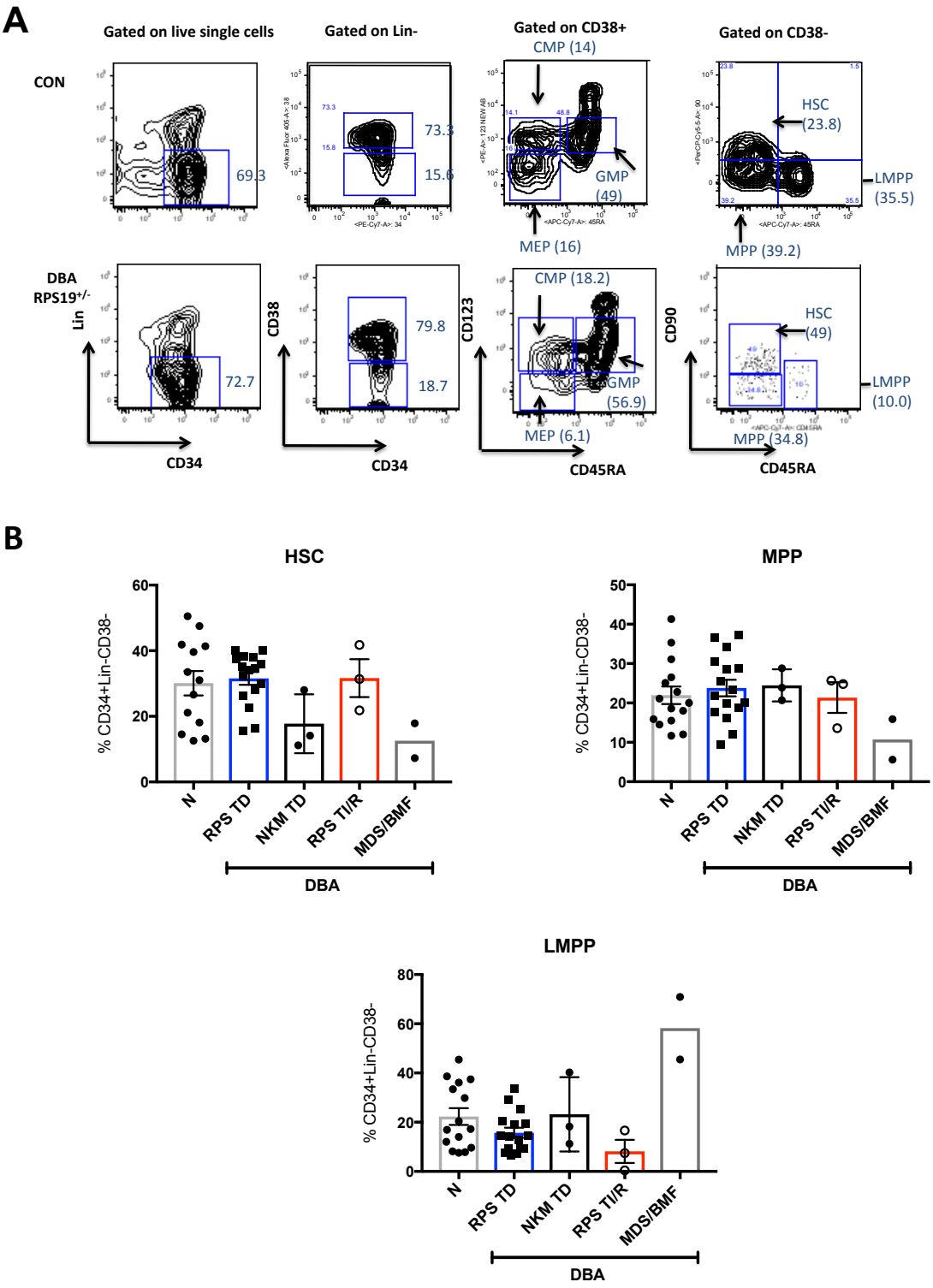
Established immunophenotypes and an accepted gating hierarchy (Table 1.6) were used to define and quantify HSPC populations in normal paediatric BM (PBM). The staining panel was optimised as described in Section 2.3.4. In brief, live (DAPI negative) single cells were positively gated for CD34 and negatively gated for a cocktail of lineage markers that stain for mature neutrophils, monocytes, T, B, NK and erythroid cells (Figure 2:2). The remaining population was then subdivided into CD38+ and CD38- fractions. CD45RA and CD123 were used to define MEP, CMP and GMP (so called 'myeloid progenitors') within the CD38+ fraction while CD45RA and CD90 were used to define HSC, MPP and LMPP within the CD38- fraction.

4.3.1.2 Preservation of the CD34+CD38- HSC compartment in TD DBA

Using the approach above, the frequencies of HSC, MPP and LMPP were measured in freshly isolated CD34+ cells from DBA BM and were compared with age-matched controls. Frequencies are expressed as a % of the parent CD34+Lin-CD38- fraction.

I found that the CD34+CD38- HSC compartment, comprising HSC, MPP and LMPP is preserved in TD DBA in terms of progenitor frequencies (Figure 4.1A&B). Notably, while relative HSPC frequencies are normal, absolute numbers are likely to be lower than in normal BM given the high prevalence of hypocellularity in DBA BM. By contrast, in the 2 patients with BMF/MDS, HSC, MPP, CMP, and MEP, were all severely depleted while LMPP and GMP were relatively spared (Figure 4.1B&C & 4.2). These data suggest that evolution to BMF and MDS in DBA is accompanied by loss of multi-/oligo-potent progenitors at the apex of the haematopoietic hierarchy. However, it is important to note that NGS screening did not identify RP gene

mutations in either patient leaving open the possibility that IBMFS other than DBA might underpin the widespread HSPC defects observed.



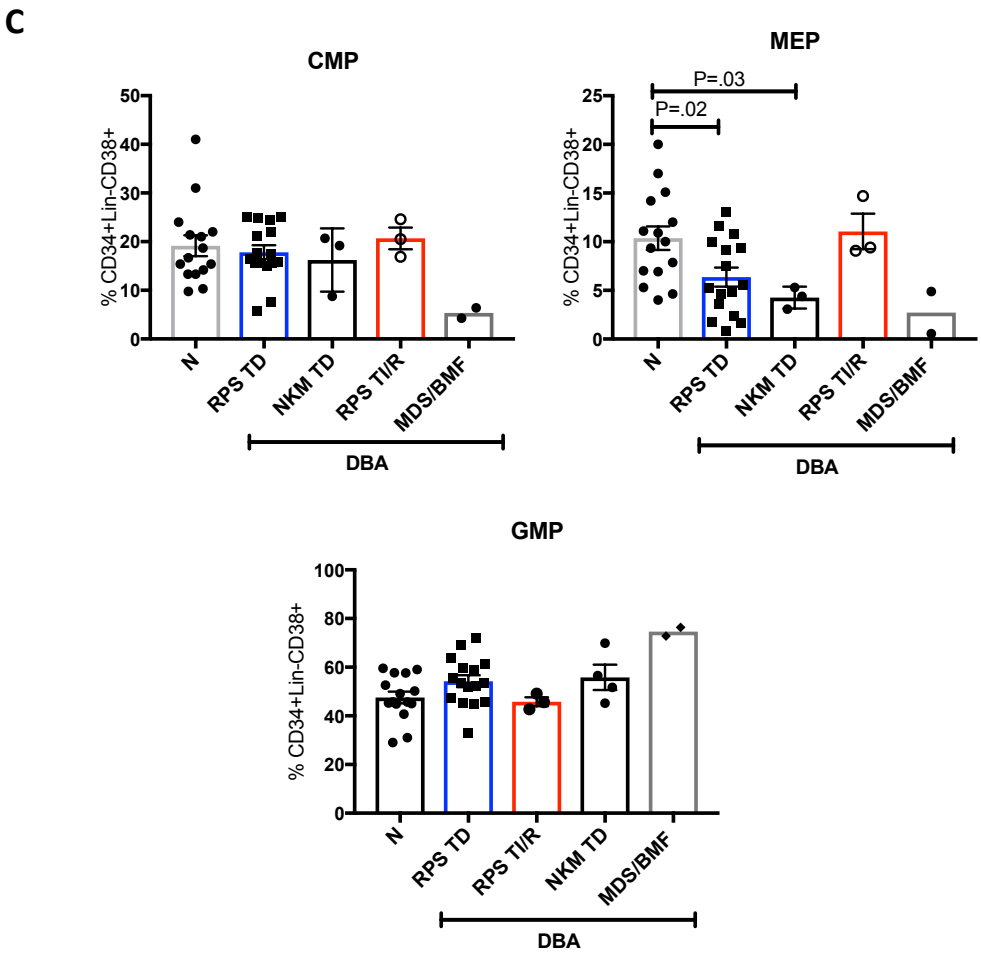


Figure 4:1 Normal relative frequencies of HSC, MPP, LMPP, CMP and GMP with selective reduction of MEP in DBA. A) Representative flow cytometry plots from CON and TD DBA (4y old male, with whole allele deletion of *RPS19)* BM **B)** Frequencies of HSC/ MPP/ LMPP progenitors, expressed as % of CD34+Lin-CD38- cells **C)** Frequencies of CMP/ GMP/ MEP progenitors, expressed as % of CD34+Lin-CD38+ cells, in BM of CON (n=15), *RPS* TD DBA (n=16), NKM TD DBA (n=3), TI/R DBA (n=4) and DBA complicated by MDS/BMF (n=2). There was no significant difference in the numbers of HSC, MPP, LMPP, CMP or GMP between normal control PBM and any of the DBA patient sub-groups. MEP were significantly reduced in TD DBA patients with *RPS* mutations or with no known mutations (P=0.02 and P=0.03 respectively). Patients with MDS/BMF (n=2) have reduced HSC, MPP, CMP and MEP compared to normal PBM.

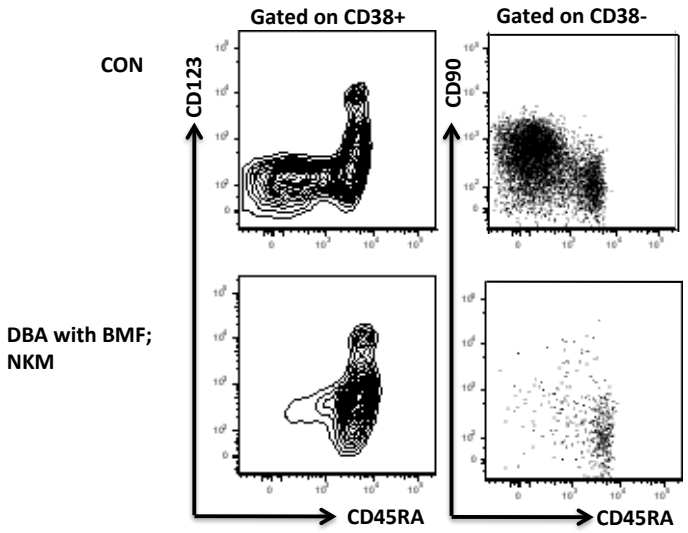


Figure 4:2 Perturbation of early haematopoiesis in patient with DBA and MDS/BMF. Representative flow cytometry plots from control and patient with DBA, pancytopenia and BM myelodysplastic features, showing striking paucity of HSC, MPP, CMP and MEP, with *relative* increases in LMPP and GMP.

4.3.1.3 MEP are quantitatively and functionally abnormal in DBA

As previously discussed, established models of human haematopoiesis suggest that CMP generate GMP and MEP (Manz et al. 2002). The latter overlap with a CD45RA-CD123- population enriched in committed EP (Doulatov et al. 2012; Hattangadi et al. 2011; A. Roy et al. 2012)**.** To study the erythropoietic defect in DBA, I first examined the frequency of CMP, MEP and GMP in BM from *RPS* TD patients. I found that CMP and GMP progenitor frequencies are similar to age-matched controls however MEP are significantly (p=0.02) reduced (Figure 4:1A&C).

Next I assessed the function of myeloid progenitor populations by FACS-sorting them in duplicate from CON (N=5) and DBA RPS TD BM (N=4). As expected, CON and DBA GMP generated myeloid [granulocyte/macrophage (CFU-G/M/GM)] colonies

exclusively (Figure 4:3). Both CMP and MEP gave rise to erythroid (BFU-E and CFU-E) and myeloid colonies. Compared with CON, DBA MEP gave rise to significantly fewer BFU-E and CFU-E ($P<0.05$) and generated small, poorly haemoglobinised colonies containing <40 cells, which I termed 'erythroid cell clusters'. These failed to develop into fully-formed BFU-E, even after 16 days. DBA CMP also produced fewer normal BFU-E colonies and more erythroid clusters than CON CMP, though the difference was not significant, perhaps because of the relatively low erythroid output of CON CMP (four-fold reduced efficiency compared with MEP; Figure 4:3). The myeloid output of DBA CMP, MEP and GMP was comparable to that of normal PBM.

4.3.2 Characterisation and prospective isolation of normal human EP

4.3.2.1 Immunophenotypic characterisation of MEP allowing prospective isolation of human EP

Having confirmed an erythroid defect within MEP (Figure 4:3), I next sought to precisely characterise this defect by devising a flow cytometry gating strategy for identification and prospective isolation of human EP. Specifically, I first identified surface markers relevant to murine and/or human erythropoiesis/megakaryopoiesis (Table 0.3; Terszowski 2005; Flygare et al. 2011; Merryweather-Clarke et al. 2011; Wangen et al. 2014) and then assessed their expression on the MEP subfraction of normal BM (Figure 4:4). This delineated 3 populations: CD71-CD41a-, CD71+CD41a- and CD71+CD41a+ MEP. I hypothesized that these would enrich for M, E and MK populations, respectively, because CD71 is a known erythroid lineage marker and CD41a a known megakaryocyte lineage marker. The CD71+CD41a- subfraction was further subdivided using 2 other erythroid markers, CD36 and CD105, (Figure 4:4) with the aim of discerning distinct EP populations.

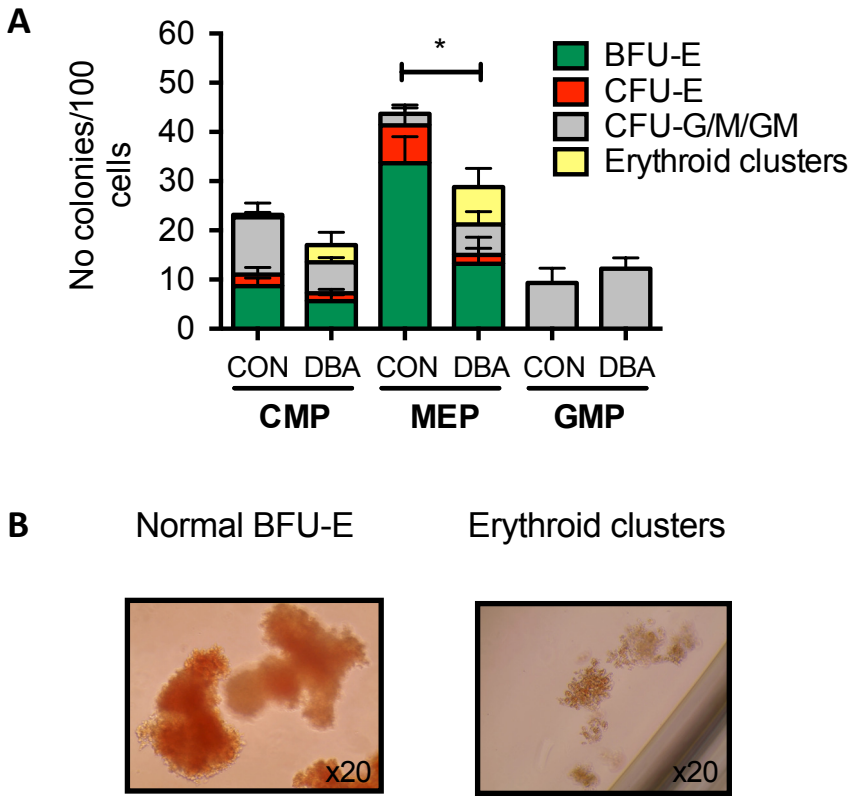


Figure 4:3 Defective production of erythroid colonies from DBA MEP. A) Clonogenic capacity of flow-sorted CMP, MEP and GMP from purified, BM-derived CD34+ cells from transfusion-dependent DBA patients (n=5) and age–matched controls (CON, n=4) **B)** Abnormal erythroid colonies, termed erythroid clusters, are observed exclusively in DBA cultures.

Table 4.3 Cell surface markers relevant to human and/or murine erythropoiesis

CD molecule	Expression in haematopoietic lineages	Function
CD34	HSPC, endothelial cells, MSC	Glycoprotein; cell adhesion
CD38	HSPC, T/B/NK cells	Glycoprotein; cell adhesion
CD45RA	Myeloid, T/B cells	Isoform of CD45; protein tyrosine phosphatase
CD123	HSPC (low expression), Myeloid, DC	IL-3 receptor α chain (low affinity subunit of IL-3R)

CD71	E, MK, Myeloid, proliferating cells/ T/B cells	Glycoprotein; transferrin receptor
CD41a	MK/Plt	Integrin αIIb; plt adhesion
CD36	E, Plt, Myeloid	Plt glycoprotein 4/ thrombospondin receptor; cell adhesion
CD105	E, EC, MSC	Endoglin; part of TGFβR complex
CD235a	E	Glycophorin A (GYPA); sialoglycoproteins bearing MNS blood group antigens

DC: dendritic cells; E: erythroid; EC: endothelial cells; HSPC: haematopoietic stem and progenitor cells; MK: megakaryocyte; MSC: mesenchymal stem cells; NK: natural kill cells; Plt: platelets.

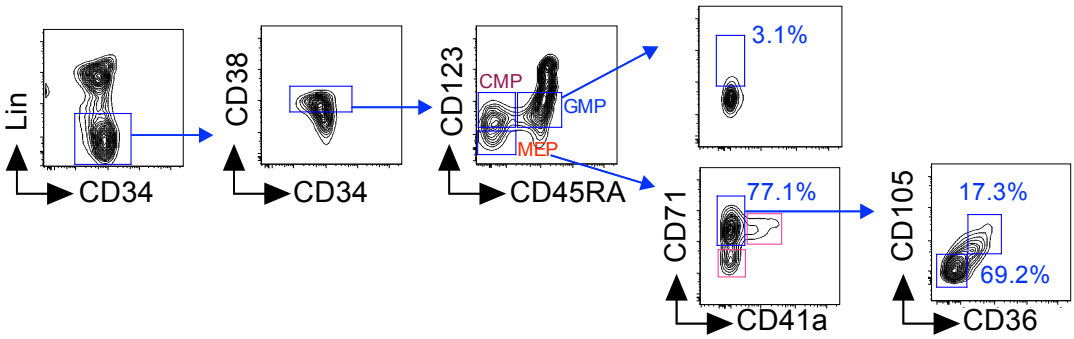


Figure 4:4 Characterisation of MEP using CD71, CD41a, CD105 and CD36. EP populations within MEP (defined by flow-cytometry as Lin-CD34+CD38+CD45RA-CD123-) were characterised by their expression of CD71, CD36, CD105 and CD41a. This demarcated 3 populations: CD71-CD41a-, CD71+CD41a+ and CD71+CD41a- . The latter population could be further subdivided by CD36 and CD105 expression. The above antigens were negligibly expressed on GMP.

Bulk and single cell flow sorting into methylcellulose medium revealed that both CD105-CD36- and CD105+CD36+ subpopulations of CD71+CD41a- MEP, generated erythroid but not myeloid colonies (Figure 4:5A). Moreover, CD105-CD36- cells generated almost exclusively large and small BFU-E with a purity of 96% and clonogenicity of 66.8% in single cell assays (Figure 4:5B). I termed these CD71+CD41a-CD105-CD36- cells early EP (EEP). In contrast, I termed CD71+CD41a-CD105+CD36+ cells late EP (LEP), as I postulated that they would lie downstream of the EEP since they generated CFU-E but not BFU-E colonies with a purity of 97% and clonogenicity of 72.3% (Figure 4:5B). Notably, the frequencies of BFU-E and CFU-E derived from EEP and LEP (corrected for their clonogenicity at a single cell level) were similar to BFU-E and CFU-E frequencies derived from age-matched, total BM-derived CD34+ cells (Figure 4:6), suggesting that my sorting strategy captures the majority of EP activity in CD34+ cells (Iskander et al. 2015). Indeed, very few E colonies were generated from either CD71-CD41a- or CD71+CD41a+ MEP (Figure 4:5A). While the former comprised a mixture of M and E lineage progenitors, the latter gave rise to only occasional BFU-E, and was later shown by our group to enrich for unipotent MK progenitors (Psaila et al. 2016).

4.3.2.2 Differences in EP immunophenotypes determined by ex vivo and in vitro approaches

While working on data, a report was published by Li and colleagues outlining a sorting strategy for prospective isolation of BFU-E and CFU-E progenitors (Li et al. 2014). UCB-derived CD34+ were cultured *in vitro* and sorted into methylcellulose at different time points throughout the culture to identify specific stages of erythroid differentiation that enriched for BFU-E and CFU-E EP. EP were shown to be CD45+ therefore I examined this marker and confirmed that both EEP and LEP are indeed CD45+ (Figure 4:7A). However, Li et al also reported that BFU-E are CD34+ and CFU-E CD34-. By contrast, my data show that CFU-E are highly enriched in a CD34+ BM population (Figure 4:5). To address this difference directly, I compared the

clonogenic output of CD34+ and CD34- Lin-CD38+CD123-CD45RA-CD71+CD36+CD105+ cells side-by-side (Figure 4:7B). Although the BM CD34- population contained CFU-E activity, this was only 25% of the CFU-E activity of the CD34+ population (Figure 4:7C), suggesting that my approach captures 75% of the total CFU-E activity.

To explain the above differences I explored whether cells with shared immunophenotypes behave differently *in vitro* and *ex vivo*. I found that EEP generated following a 4 day liquid culture of CD34+Lin- BM, when plated in methylcellulose gave a lower yield and purity of BFU-E colonies, as compared with the corresponding EEP population sorted *ex vivo* on day 0 (Figure 4:8). Conversely, it is likely that culturing HSPC populations *in vitro* as opposed to analysing them directly *ex vivo* can influence the expression of cell surface markers. Taken together, the data above suggest that *in vivo*, CFU-E lose expression of CD34 as they mature.

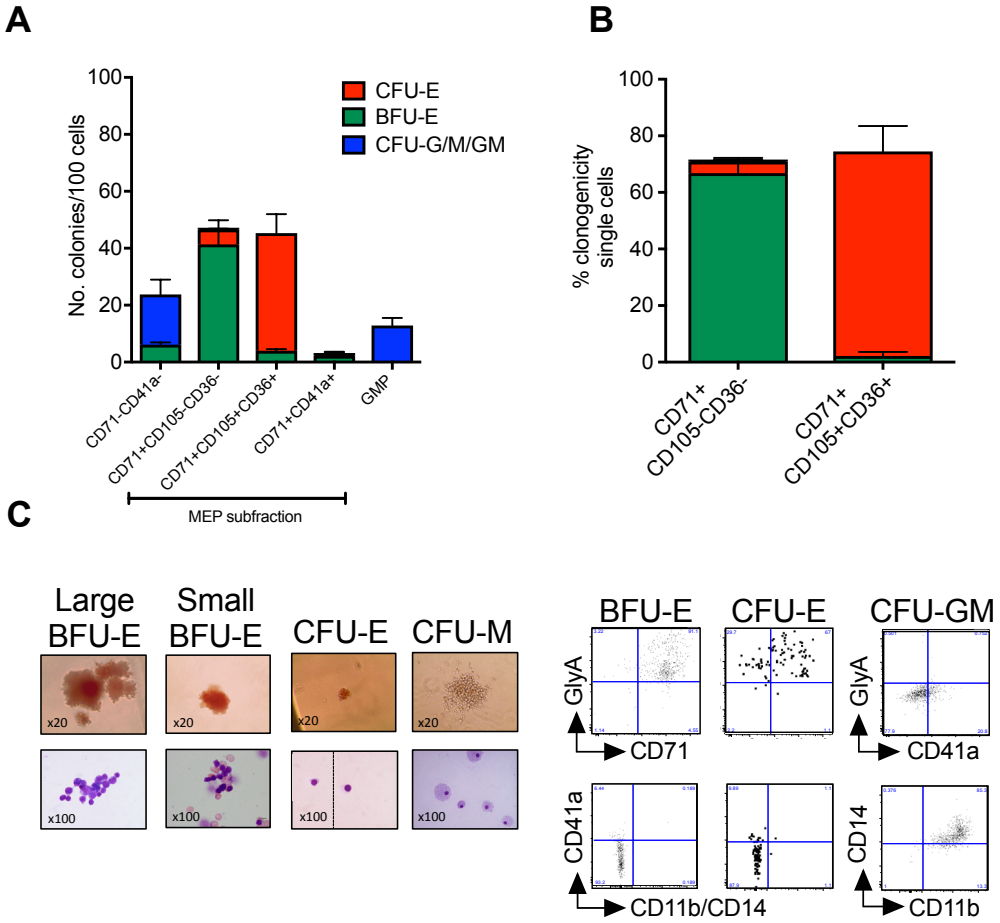


Figure 4:5 Prospective isolation of human BM populations highly enriched in BFU-E and CFU-E activity. A) CD71a-CD41a- (n=6), CD71+CD41a-CD105-CD36- (n=6) and CD71+CD41a-CD105+CD36+ (n=5) and CD71a+CD41a+ (n=6) subpopulations of MEP were sorted into complete methylcellulose containing cytokines supportive of erythroid and myeloid development. GMP were sorted concurrently as a control. Haematopoietic colonies were identified and scored as described in Section 2.2.7.2 **B)** Single cell clonogenic assays of CD71+CD41a-CD105-CD36- MEP (EEP; n=5) and CD71+CD41a-CD105+CD36+ MEP (LEP; n=3). % single cell clonogenicity is defined as % of wells in which a colony grew following flow-sorting of 1 cell per well. **C)** Morphologic distinction between large BFU-E, small BFU-E and CFU-E derived from flow-sorted erythroid progenitors. CFU-M derived from flow-sorted GMP also shown. Large (early) BFU-E colonies are composed of large and/or multiple clusters of EB while small (late) BFU-E colonies comprise small uni- or bi-focal clusters of

>200 EB. BFU- E are usually visible by Day 7 (D7) of culture but increase in size and haemoglobinisation by D14. CFU-E colonies are composed of a unifocal cluster of <200 EB; they appear by D7 of culture and persist till D14. To validate colony scoring, single (BFU-E) or pooled (2-3 CFU-E) colonies were plucked on D14 and analysed by flow cytometry. As expected, BFU-E and CFU-E express CD71 and GlyA but not CD41a (MK lineage markers) nor CD11b/14 (monocyte/granulocyte markers). Loss of CD71 (a marker of EB maturity) is observed in CFU-E but not BFU-E colonies demonstrating that CFU-E are more differentiated. CFU-GM are CD11b/14 positive, GlyA/CD41a negative with negative/low CD71 expression. Colony identity was further confirmed by cytospin. Morphologically, BFU-E are composed of late EB and reticulocytes. CFU-E colonies consist of nucleated EB and occasional reticulocytes. CFU-M are composed of macrophages, as shown.

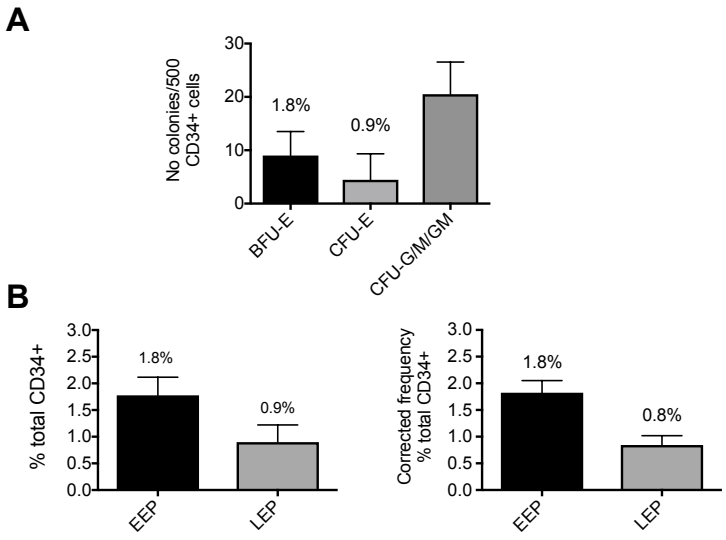
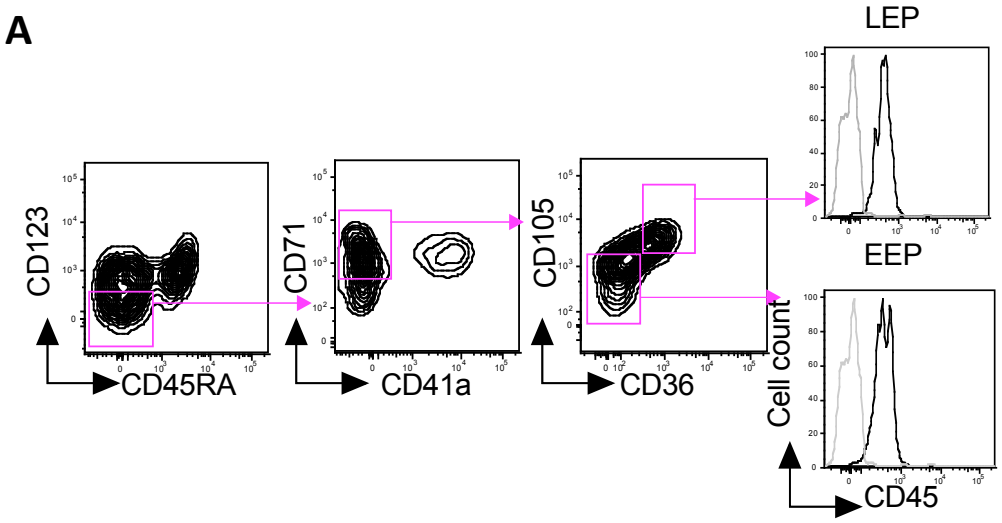
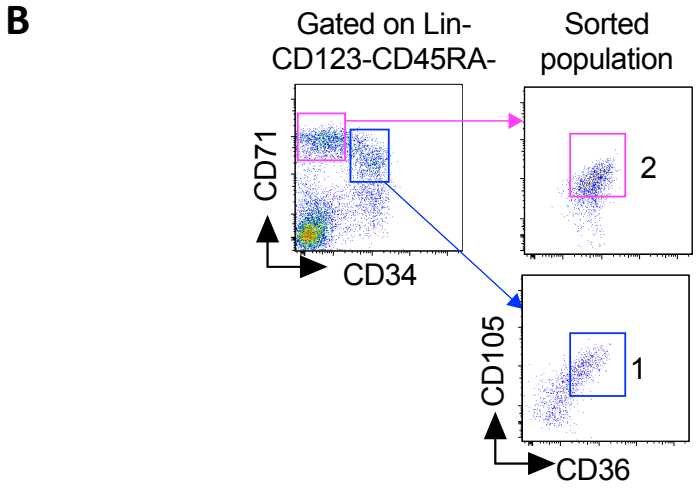
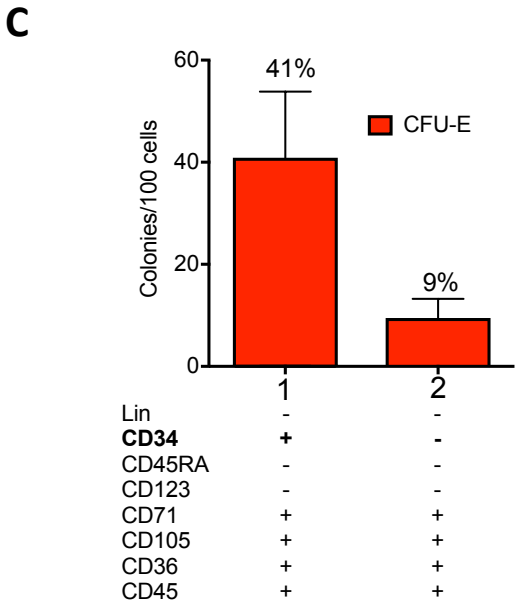


Figure 4:6 Novel EP sorting strategy captures most of BFU-E and CFU-E activity in the CD34+ fraction of normal BM A) Frequencies of BFU-E, CFU-E and CFU-G/M/GM colonies in total CD34+ cells derived from control paediatric BM (n=7). Values above bars show colony frequencies expressed as percentages of total CD34+ cells. **B)** Frequencies of EEP-derived BFU-E and LEP-derived CFU-E according to their immunophenotype (*left*: n=8 including 4 of the donors represented in A) and corrected for their clonogenicity at a single cell level (*right*: n=6 for EEP and n=5 for

LEP), by multiplying the EEP and LEP frequencies by their respective cloning efficiencies, determined using single-cell clonogenic assays.


A
CD123
CD45RA
CD71
CD41a
CD105
CD36
LEP
EEP
Cell count
CD45


B
Gated on Lin-
CD123-CD45RA-
Sorted
population
CD71
CD34
2
CD105
1
CD36


C
Colonies/100 cells
60
40
20
0
41%
9%
CFU-E
1
2
Lin - -
CD34 + -
CD45RA - -
CD123 - -
CD71 + +
CD105 + +
CD36 + +
CD45 + +

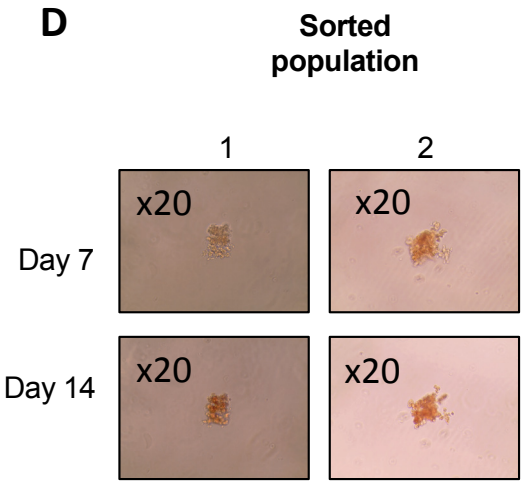
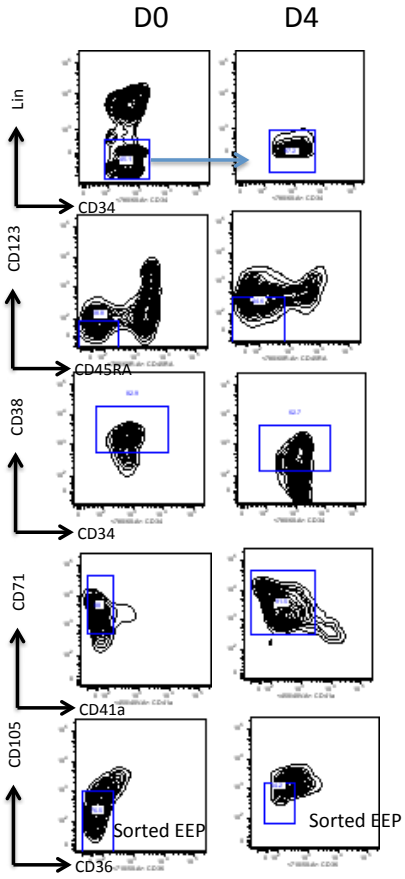

D
Sorted
population
1
2
Day 7
Day 14
x20
x20
x20
x20

Figure 4:7 CFU-E lose expression of CD34 as they mature A) Expression of CD45 on EEP and LEP subpopulations of normal BM. Overlay histograms show labelled cells (black line) and unstained negative control cells (grey line). Both EEP and LEP are CD45 positive. **B)** CD34+ (population 1) and CD34- (population 2) subfractions of the Lin-CD123-CD45RA-CD71+CD36+CD105+ compartment of BMMNCs were FACS-sorted and plated in methylcellulose in duplicate **C)** Colonies were scored in 3 independent experiments using control BM **D)** Representative CFU-E colonies generated by each of the 2 populations, photographed on days 7 and 14, demonstrating similar morphology.



	D0 *ex vivo*	D3 *in vitro*
EEP as % CD34+Lin-	10.7	9.6
Colony-forming capacity (%) = number of colonies/ 100 single cells	86	66

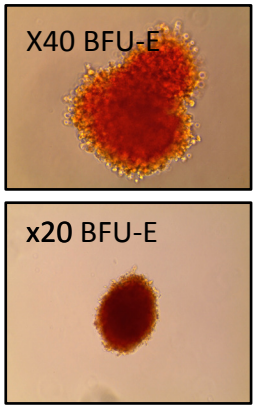
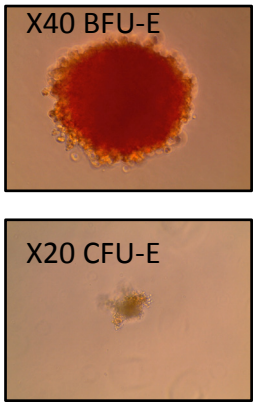
plated		
• **BFU-E**	86	42
• **CFU-E**	0	20
• **CFU-G/GM**	0	4
EEP purity (%)	100	63.6
Examples of colony morphology		

Figure 4:8 Reduced colony-forming capacity and purity of *in vitro* versus *ex vivo* EEP. EEP were sorted *ex vivo* from fresh BM into methylcellulose on D0. Concurrently CD34+Lin- cells were sorted into liquid culture (Protocol A, with IL-6; **Table 2.1**). 4 days later, *in vitro* generated EEP were sorted into methylcellulose. The table shows a comparison of the progeny of *ex vivo* and *in vitro* EEP.

4.3.2.3 Dissection of the hierarchical relationship between EEP and LEP

Consistent with EEP and LEP being distinct progenitor populations, I found that *ex vivo* sorted EEP are small, blast-like cells with a higher nuclear-cytoplasmic ratio than the larger LEP (Figure 4:9), similar to murine fetal EP (Flygare et al. 2011).

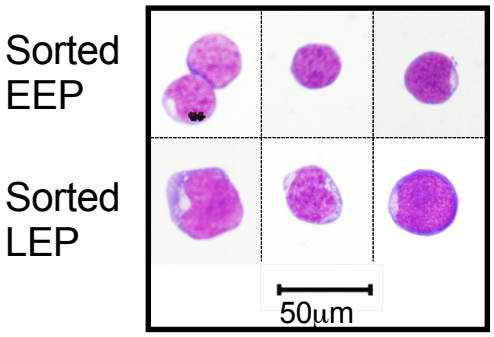


Figure 4:9 Distinct morphology of EEP and LEP. May-Grünwald Giemsa staining of flow-sorted EEP and LEP from control adult BM.

Furthermore, under liquid culture conditions, EEP had a higher proliferative capacity than LEP (P<.05; Figure 4:10A) and a higher fraction of erythroid precursors i.e., EB derived from EEP were in S phase (P<.05; Figure 4:10B) consistent with LEP arising downstream of EEP in the erythropoietic hierarchy. While flow-purified EEP generated LEP in short-term liquid cultures, LEP were unable to generate EEP (Figure 4:10C). Additionally, although both populations generated purely EB (i.e., CD34-CD71+GlyA+CD14-CD11b-CD41a-), LEP differentiated more rapidly as shown by a higher fraction of GlyA+ (glycophorin A) cells lacking CD105 (loss of CD105 correlates with EB maturity; Wangen et al. 2014) on day 11 (Figure 4:10D). Finally, these populations exhibited distinct gene expression patterns (Figure 4.11): a) in line with their immunophenotypes expression of *CD36* progressively increased from EEP to EB and b) consistent with findings in murine (Suzuki 2003) and human (Li et al. 2014) EP, while expression of the MK-E transcription factor *GATA-1* (Doré & Crispino 2011) increased from EEP to EB, *GATA-2* was downregulated at the LEP stage (Figure 4:10A).

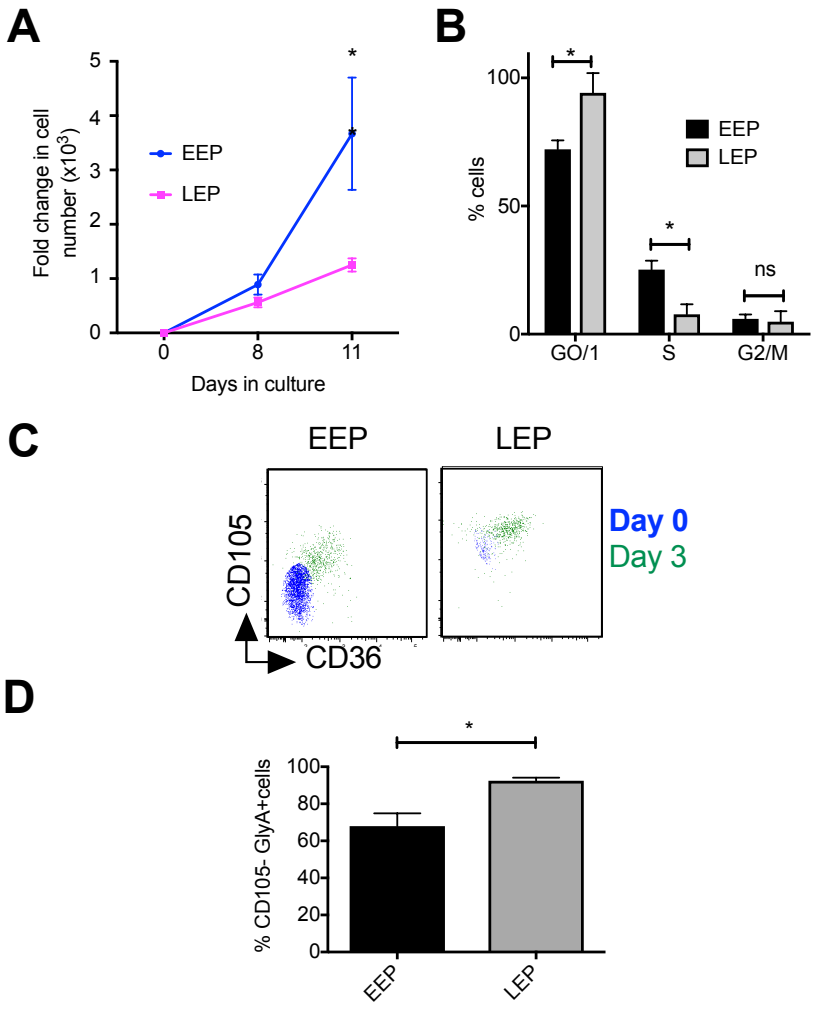


Figure 4:10 EEP lie upstream of LEP in the erythropoietic hierarchy. A) Proliferative capacity of flow-sorted EEP and LEP in a longitudinal liquid culture (Protocol B; **Table 2.2**) assessed by cell counting (n=4). Cells on day 11 were EB i.e., CD34-CD71+GlyA+ with negative staining for myeloid (CD14, CD11b) and MK (CD41a) markers. **B)** Concurrent cell cycle analysis on day 11 shows a higher number of EB derived from EEP in S phase with fewer in $G_{0/1}$ compared with EB derived from LEP (P<.05). **C)** Flow-sorted EEP and LEP derived from CON ABM were cultured in the presence of EPO, IL-3, IL-6 and SCF (Protocol A; **Table 2.1**). Three days later, flow-cytometric analysis showed that EEP acquired expression of CD105 and CD36 while LEP gained higher expression of CD105 and CD36 (blue dots denote day 0, green dots day 3). Plots shown are representative of 2 independent experiments. **D)** Frequency of late

EB (CD105-GlyA+) generated by day 11 of erythroid liquid culture from sorted EEP and LEP (n=4).

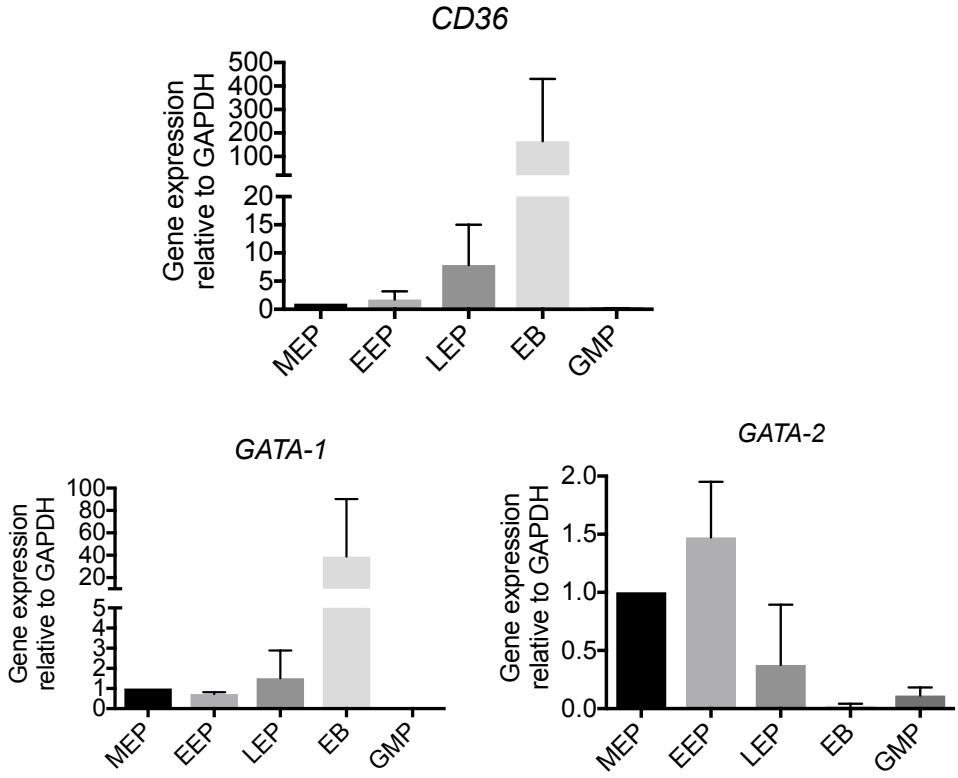


Figure 4:11 EEP and LEP exhibit distinct transcriptional profiles. mRNA expression levels of *CD36*, *GATA-1* and *GATA-2* were determined by quantitative real-time PCR in purified GMP, MEP, EEP, LEP and EB (Lin-CD34-CD36hiCD71+GlyA+ EB), derived from 3 control BM samples. Transcript levels in each cell population were normalised to *GAPDH*. The levels shown in the graphs are expressed relative to the corresponding gene expression in MEP. In line with published data, GMP do not express *CD36* and *GATA-1* but express a low level of *GATA-2*.

4.3.3 Cumulative defects in the frequency and function of DBA EP

Next, I sought to elucidate the EP defect in DBA using my novel EP definitions. To account for the potential heterogeneity within the DBA group, as mentioned in Section 4.2, I focussed on TD patients with *RPS* gene pathogenic genetic variants. I

found that the frequency of EEP was reduced in DBA BM (*P*<.001;Figure 4:12A) which was mainly attributable to decreased frequency of CD71+CD41a- cells (Figure 4:12B&C). In clonogenic assays, DBA EEP generated fewer colonies than controls (*P*<.05**;** Figure 4:13A), consistent with quantitative and qualitative defects in DBA EEP, which were even more pronounced when EEP frequency and their clonogenic output were considered together (Figure 4:13B). Single cell assays confirmed the reduced clonogenicity of DBA EEP (*P*<.01;Figure 4:13C) and revealed heterogeneity within EEP in individual patients; while some EEP formed BFU-E of normal size and morphology, others produced abnormal, poorly haemoglobinised erythroid clusters (Figure 4:13D). Consistent with these findings, DBA EEP generated fewer EB than controls in an erythroid liquid culture system (*P*<.05;Figure 4:14).

Similar abnormalities were found in the frequency and clonogenicity of DBA LEP, although morphologically the colonies were similar to those produced by normal PBM (Figure 4:15). Additional studies of LEP function in DBA were precluded by their low frequency. Both the frequency (Figure 4:12) and the single cell clonogenic myeloid output of the CD71-CD41a- subfraction of MEP (encompassing mainly myeloid and some erythroid progenitors) were normal in TD DBA BM though again fewer E colonies were produced from this population (Figure 4:16).

While CFU-E do not circulate except under pathological conditions, BFU-E circulation in the peripheral blood is physiological (Dzierzak & Philipsen 2013). I therefore measured the immunophenotypic frequency of CD34+CD71+ EP in PBMCs from 3 TD *RPS* DBA patients, compared with controls (Figure 4:17). This was reduced three-fold in DBA, corresponding to the four-fold reduction in EEP present in DBA BM (Figure 4:12).

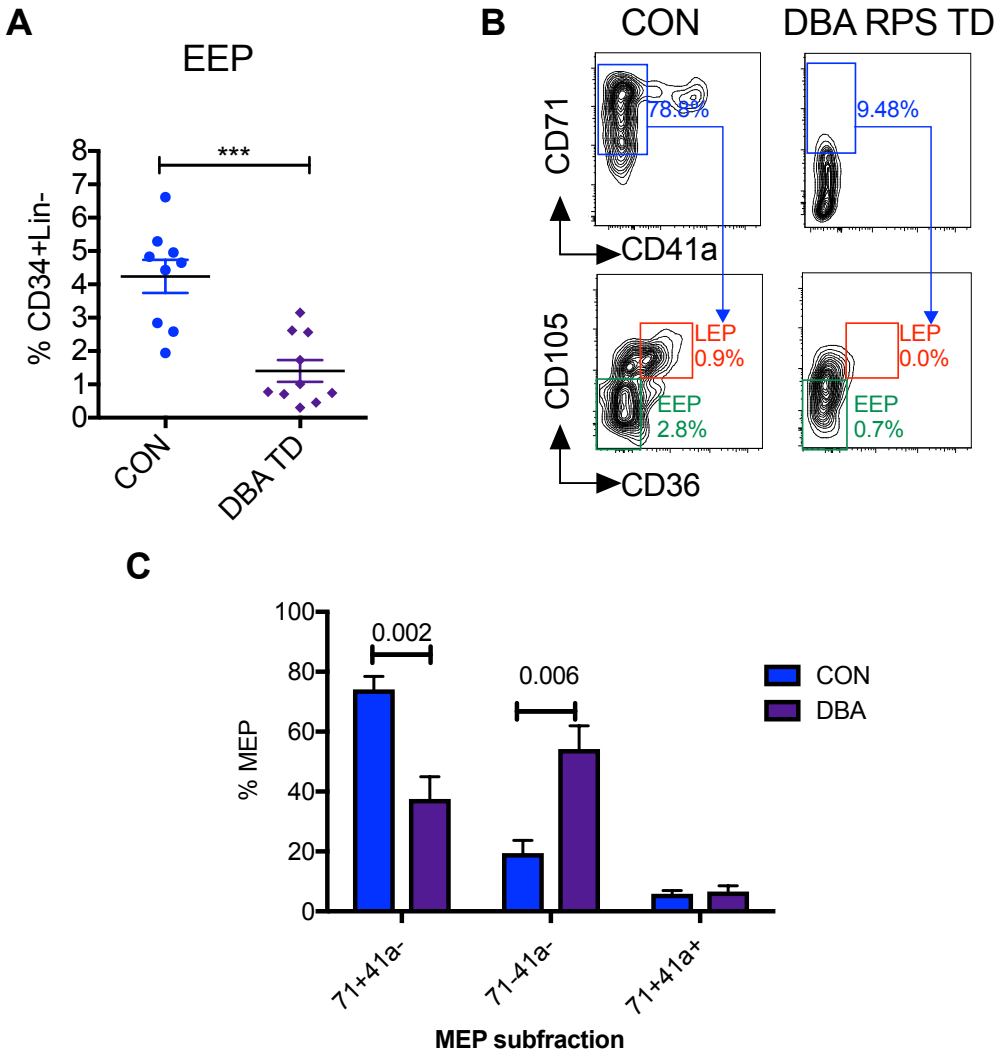


Figure 4:12 Depletion of EEP in TD DBA. A) Frequency of BM EEP in TD DBA (n=10) and control BM (n=9), as assessed by flow-cytometry. **B)** Representative plots from a patient with transfusion-dependent (TD) DBA and an age-matched control are shown. Values refer to frequency of CD71+CD41a- cells as % of MEP and frequency of EEP and LEP as % of CD34+Lin- population. Note that Figure 4.6B shows EEP and LEP as % of total CD34+ cells **C)** Relative frequencies of CD71+CD41a-, CD71-CD41a- and CD71+CD41a+ subfractions of MEP. In DBA, CD71+CD41a- EP are reduced with a relative increase in the proportion of CD71-CD41a- cells.

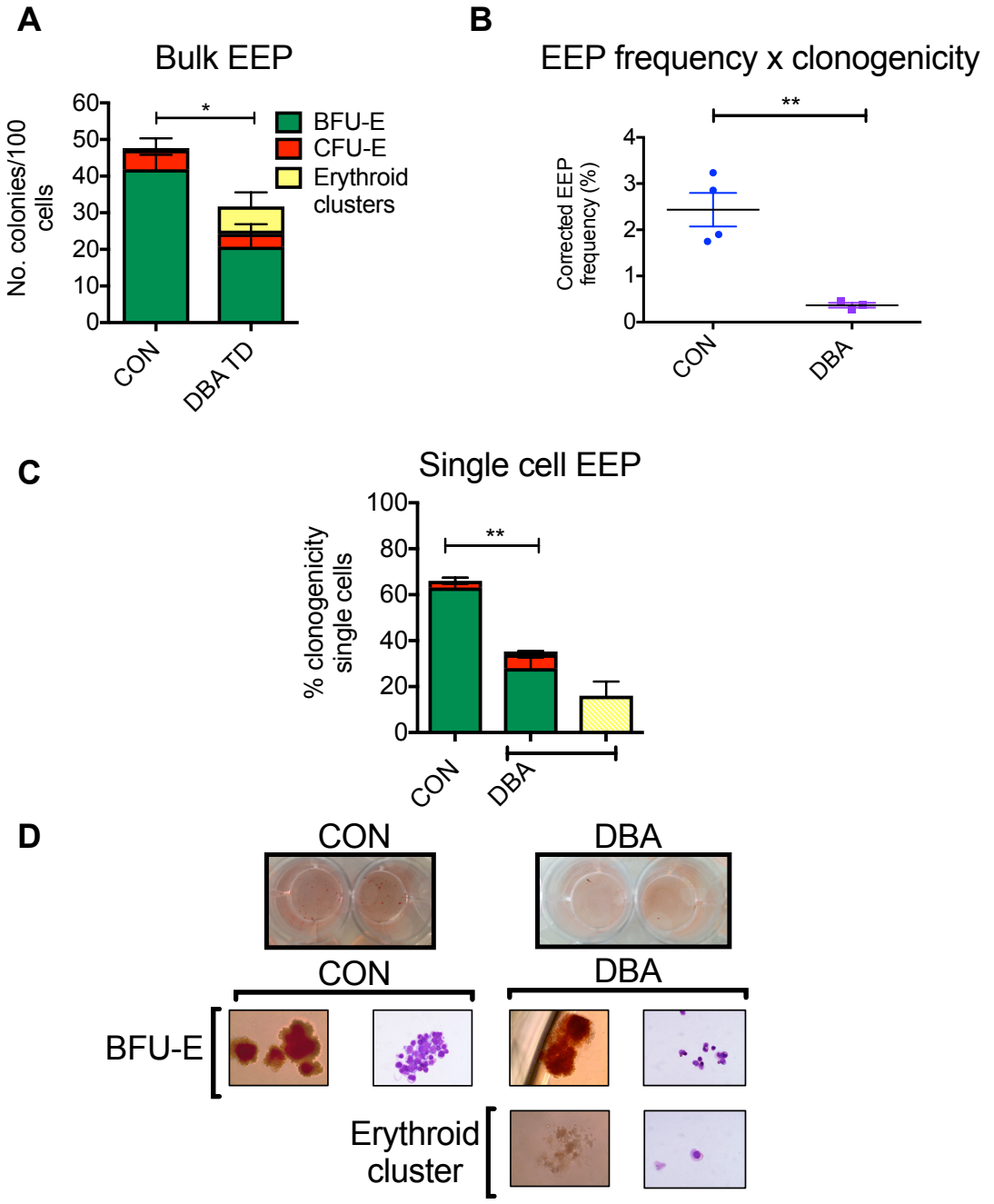


Figure 4:13 Impaired differentiation of EEP in TD DBA. A) Frequency of BFU-E colonies generated from 100 flow-sorted DBA and control EEP (n=4). **B)** Corrected EEP frequency in CON and TD DBA i.e., a product of EEP immunophenotypic frequency in CD34+Lin- cells and the EEP clonogenicity at a single cell level (n=4). **C)** Clonogenicity of TD DBA (n=4), and control EEP (n=6) flow-sorted as single cells. **D)** Macroscopic (*upper panel*) images of control and TD DBA EEP sorted in duplicate

wells and microscopic (*lower panel*) morphology of colonies derived from single cell sorting of DBA and control EEP and LEP in methylcellulose.

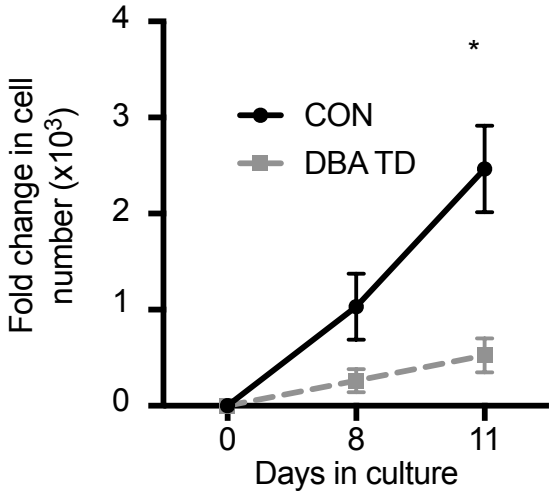


Figure 4:14 Impaired EEP proliferation in TD DBA. The proliferative capacity of flow-sorted DBA and control EEP (n=4) was assessed by cell counting in a longitudinal liquid culture (Protocol B; Table 2:1). Cells on day 11 were a pure population of EB i.e., CD34-CD71+GlyA+CD14-CD11b-CD41a-.

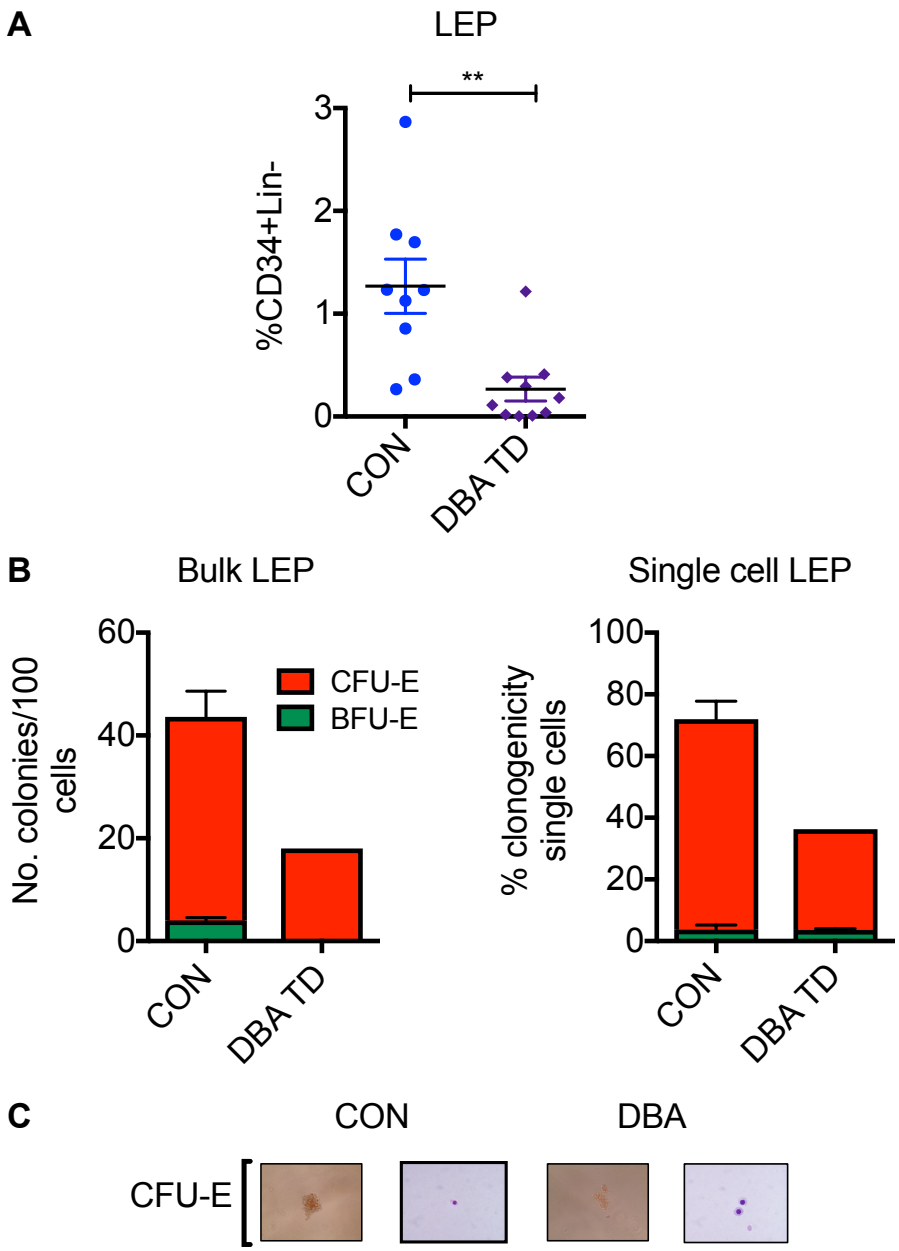


Figure 4:15 Quantitative and qualitative LEP defects in TD DBA. A) Frequency of LEP in TD DBA (n=10), and control BM (n=9) as assessed by flow-cytometry **B)** Clonogenicity of TD DBA (n=2; due to insufficient LEP in TD DBA BMs for sorting), and control (n=4 for bulk, 5 for single cells) LEP after sorting and plating of 100 cells in bulk or 50 single cells. **C)** Microscopic and morphological appearances of CFU-E colonies derived from N and DBA LEP.

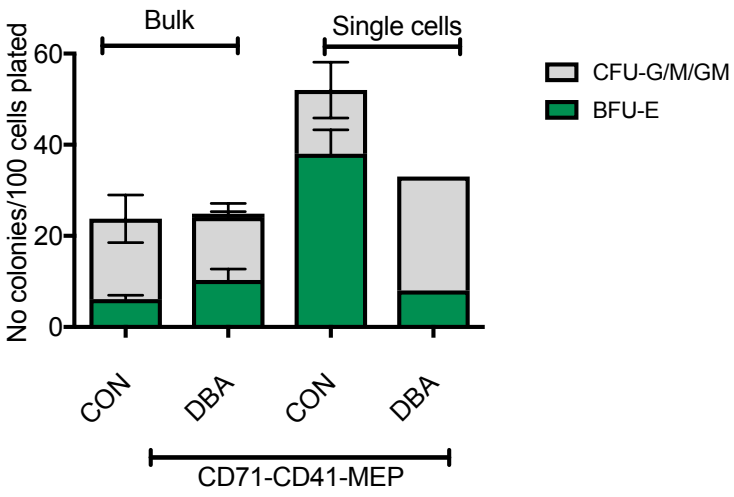


Figure 4:16 Clonogenicity of CD71-CD41a- subset of MEP. 100 cells were FACS-isolated in bulk or as single cells from TD DBA (n=5 bulk, n=2 single cells) and CON (n=4 bulk, n=3 single cells) BM. Colonies were scored on D14.

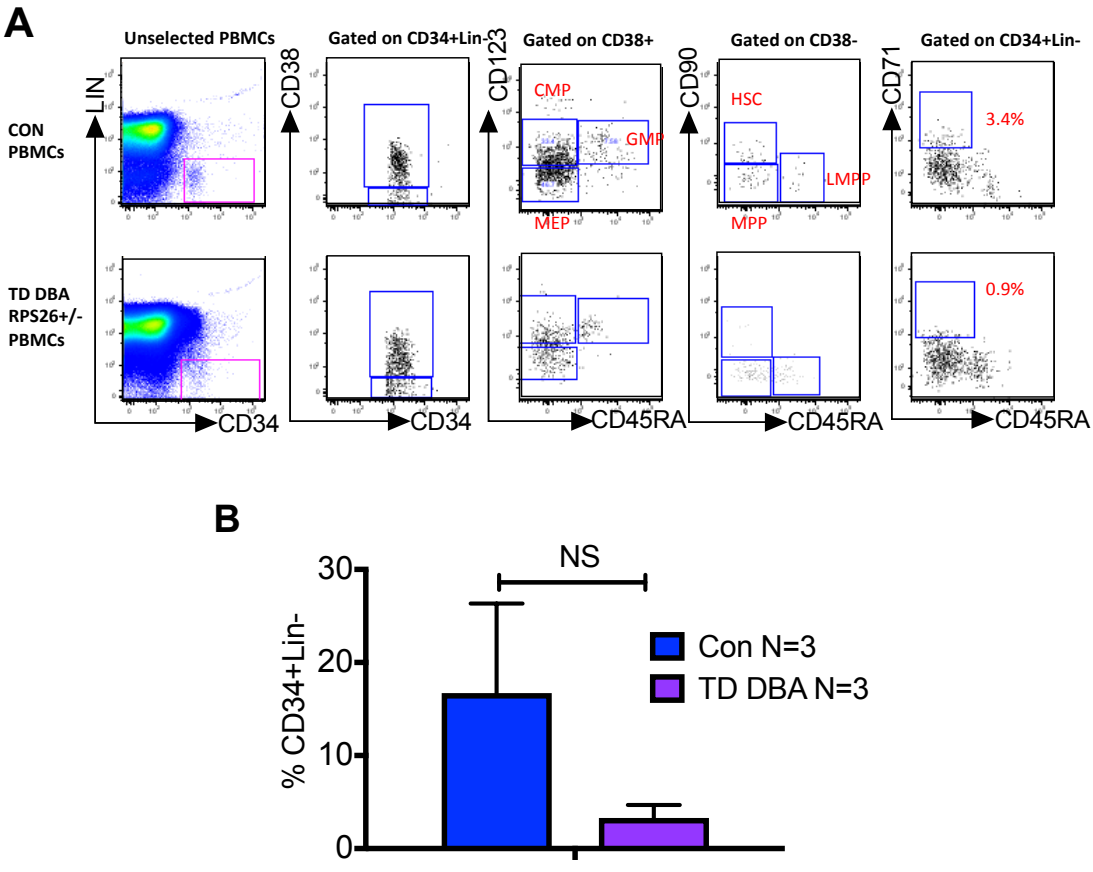


Figure 4:17 Circulating CD71+ EP are reduced in TD DBA. A) Immunophenotypes of CD34+Lin- HSPC in unselected PBMCs from a normal CON (26yM) and a TD patient with DBA (6yF). **B)** Cumulative data from 3 biological replicates.

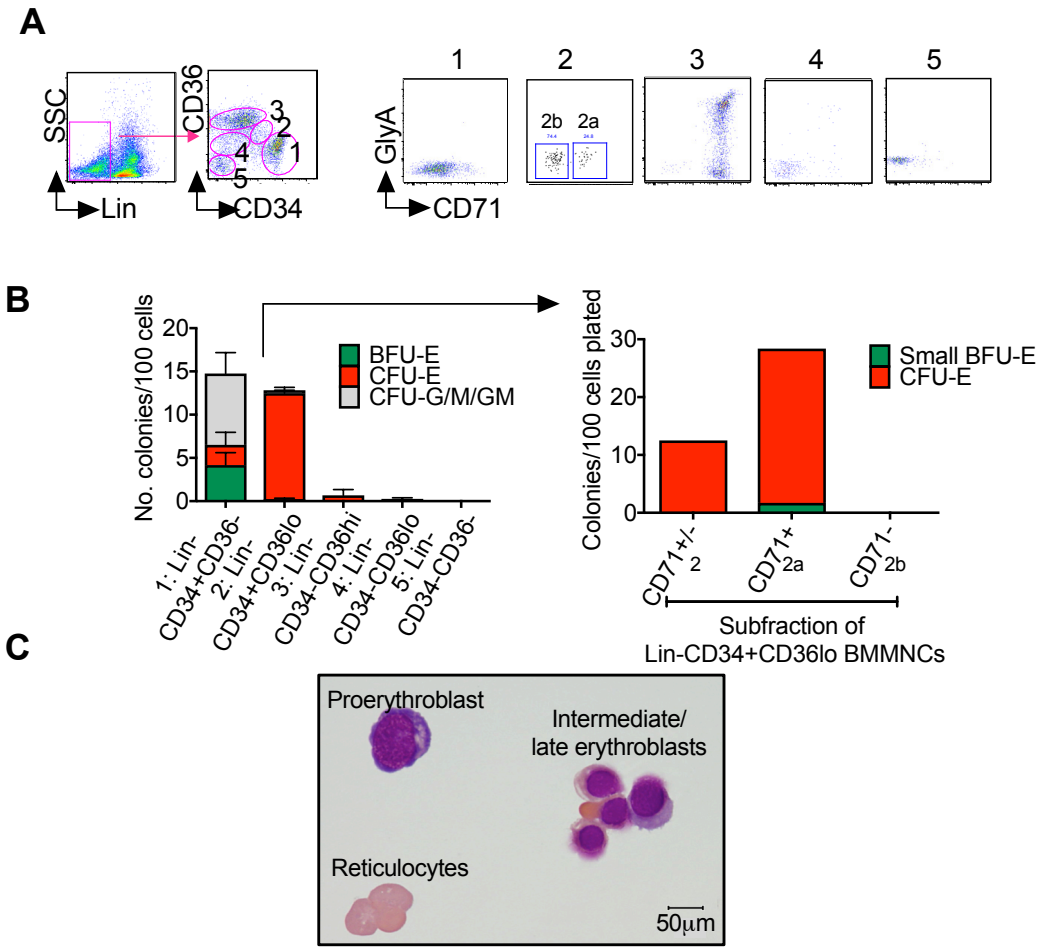


Figure 4:18 EB reside in the Lin-CD34^{-}CD36^{hi}CD71^{hi} compartment of BM. A) After exclusion of non-erythroid lineage cells, the remaining compartment of BMMNCs isolated from CON PBM (n=3) was divided into 5 populations according to expression of CD34 (a marker of haematopoietic progenitors) and CD36 (an early erythroid marker). These populations were also analysed for their expression of CD71 and GLYA (markers of erythroid differentiation) **B)** Each of the above populations was FACS-sorted and plated in methylcellulose at different cell densities. CD34+ populations 1-2 were plated at a density of 100 cells/plate while CD34- populations 3-5 were plated at densities of 100 and 1000 cells/plate. In 1 experiment, further

sorting of the 71^{lo} (2a) and 71- (2b) compartments of population 2 was performed. Haematopoietic colonies were scored on D14 **C)** Population 1 (Lin-CD34+CD36-) contains cells with both myeloid and erythroid potential while population 2 (Lin-$CD34^{lo}CD36^{lo}$) enriches purely for CFU-E and small BFU-E. Within population 2, CFU-E enrich in the $CD71^{lo}$ BM compartment, with a cloning efficiency of 35%. **D)** Sorting of population 3 cells (Lin-CD34-CD36hiCD71hiGlyA+/-) and staining with May-Grünwald Giemsa confirmed that this BM compartment is composed of a pure population of EB at a range of differentiation stages, from proEB through to anuclear reticulocytes.

Finally, to investigate whether the reduced frequency of EEP and LEP in TD DBA is associated with a reduced frequency of EB *in vivo*, I examined this population in fresh BM samples from the same patients studied above. To define EB, I sorted BM cells into five erythroid lineage negative populations (Figure 4:18A)**.** Consistent with my previous data above, Lin-CD34+CD36-71-/+ cells (population 1) generated a heterogeneous mixture of erythroid and myeloid colonies, while Lin-CD34+CD36lo71loGlyA- cells (population 2) enriched for CFU-E colonies. As per published immunophenotypic descriptions of EB (Merryweather-Clarke et al. 2011), Lin-CD34-CD36hiCD71hiGlyA-/+ cells (population 3) had negligible colony-forming capacity (Figure 4:18B) and comprised proEB to orthochromatic EB with occasional reticulocytes (Figure 4:18C)**.** Compared with CON, population 3, comprising EB, was virtually absent in TD *RPS* DBA.

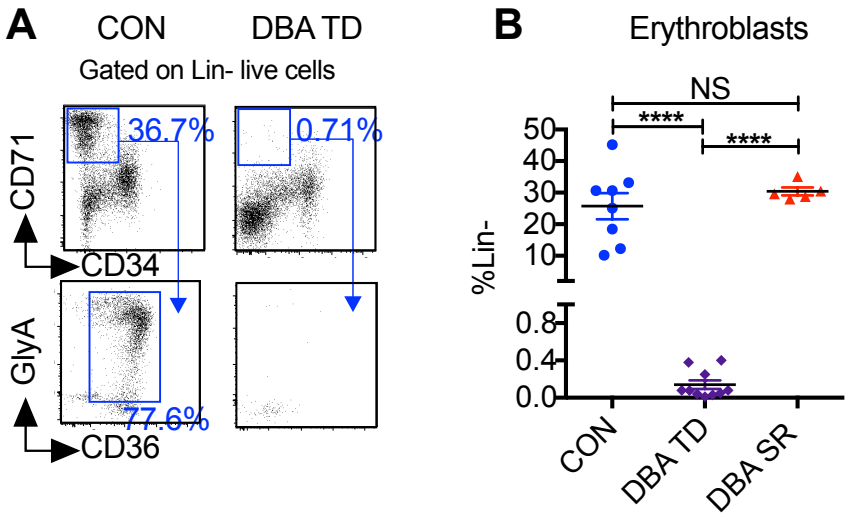


Figure 4:19 Depletion of EB in TD *RPS* DBA BM. A) Representative flow-cytometry plots from a CON and a TD patient with DBA. BM EB arising downstream of EP were identified as Lin- (not shown) CD34-CD71+ BMMNCs (upper plots) co-expressing CD36 and variable expression of GlyA (lower plots) **B)** Cumulative data from CON (n=8) and TD (n=10) DBA patients.

4.3.4 Purification of putative EP upstream of MEP; implications for the erythroid defect in DBA

Although the data above demonstrate that CD71+ MEP enrich for EP, recent studies in human haematopoietic cells have suggested that erythroid and megakaryocyte lineage specification occur upstream of MEP, at either the stem or multipotential progenitor stage (Sanjuan-Pla et al. 2013; Notta et al. 2016). These data posit an alternative model of haematopoiesis in which committed EP can directly arise from the HSC or MPP (Figure 1.4).

To explore this possibility and its pertinence to DBA, I first examined expression of the E/MK markers already used to characterise MEP within the CMP population, which lies immediately upstream of MEP in the classical model of haematopoiesis. Analogous to MEP, CMP can be subdivided into 3 populations using CD71 and CD41a

and the CD71a+CD41a- population can be further separated into 2 sub compartments using CD105 and CD36 (Figure 4:20). The lineage specificity of these subsets was the same as their MEP counterparts, with the exception that CD71 expression allowed precise discrimination of E from myeloid progenitors within CMP (Figure 4:21A versus Figure 4:16). In addition, BFU-E and CFU-E could be purified from the CD105-CD36- and CD105+CD36+ subfractions of CMP, with no significant differences from MEP in terms of clonogenicity or purity (Table 4.4).

Table 4.4 Comparison of E colony output from MEP and CMP subfractions of BM.

HSPC population	Subfraction	Clonogenicity on bulk sorting (mean %, SD)	Purity (%)
MEP	CD71+CD41a-CD105-CD36-	41.4+/-10.0	87.8
	CD71+CD41a-CD105+CD36+	41.3+/-11.6	90.4
CMP	CD71+CD41a-CD105-CD36-	36.0+/-12.8	94.2
	CD71+CD41a-CD105+CD36+	37.9+/-12.0	98.0

Next, I compared the colony-forming capacity of these populations in CON and TD *RPS* DBA BM. Both BFU-E and CFU-E colonies were reduced in number in DBA with abnormal erythroid clusters arising in place of normal BFU-E, though these differences lacked statistical significance (Figure 4:21A). Functional impairment of DBA EP populations within CMP therefore remains equivocal, however there was a clear decrease in the frequencies of BFU-E and CFU-E progenitors within the total CD38+ compartment, encompassing both CMP and MEP (Figure 4:21B&C).

Given the utility of CD71 as an E marker, I subsequently designed a new multiparameter flow cytometry panel to resolve the expression of E/MK markers on primitive CD34+CD38- HSPC (Figure 4:22A). This demonstrated that CD71 is expressed on a low frequency of cells within MPP (but not HSC or LMPP). In 1 preliminary experiment I confirmed that the progeny of CD71+ MPP are mainly BFU-

E (20% BFU-E, 3% CFU-G/M/GM) although the colony-forming efficiency was lower than from CD71+CD38+ EP in the same sample (33% BFU-E, 3% CFU-E). As with CMP, the CD71- subfraction of MPP gave rise to myeloid colonies exclusively (22% CFU-G/M/GM). Finally, I showed that these early CD38-CD71+MPP, enriched in erythroid activity, are reduced by almost three-fold in DBA BM compared with age-matched controls (Figure 4:22B). While it was not feasible to study this rare population further using DBA primary cells, these data raise the possibility that the origins of the erythroid defect in DBA lie in the primitive CD38- fraction, at the inception of erythroid lineage commitment.

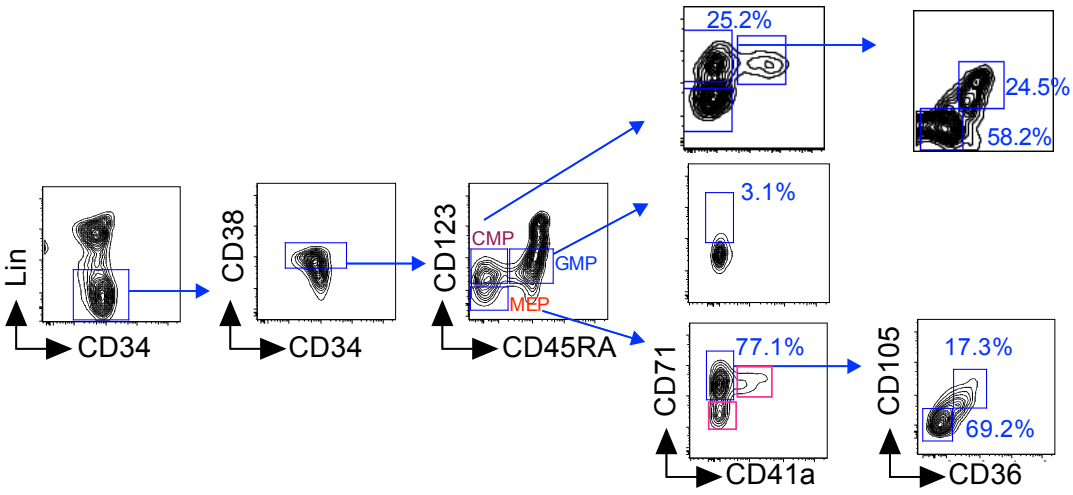


Figure 4:20 Characterisation of CMP using erythroid/megakaryocyte markers. Plots shown are from 1y7m normal male BM and are representative of 14 controls.

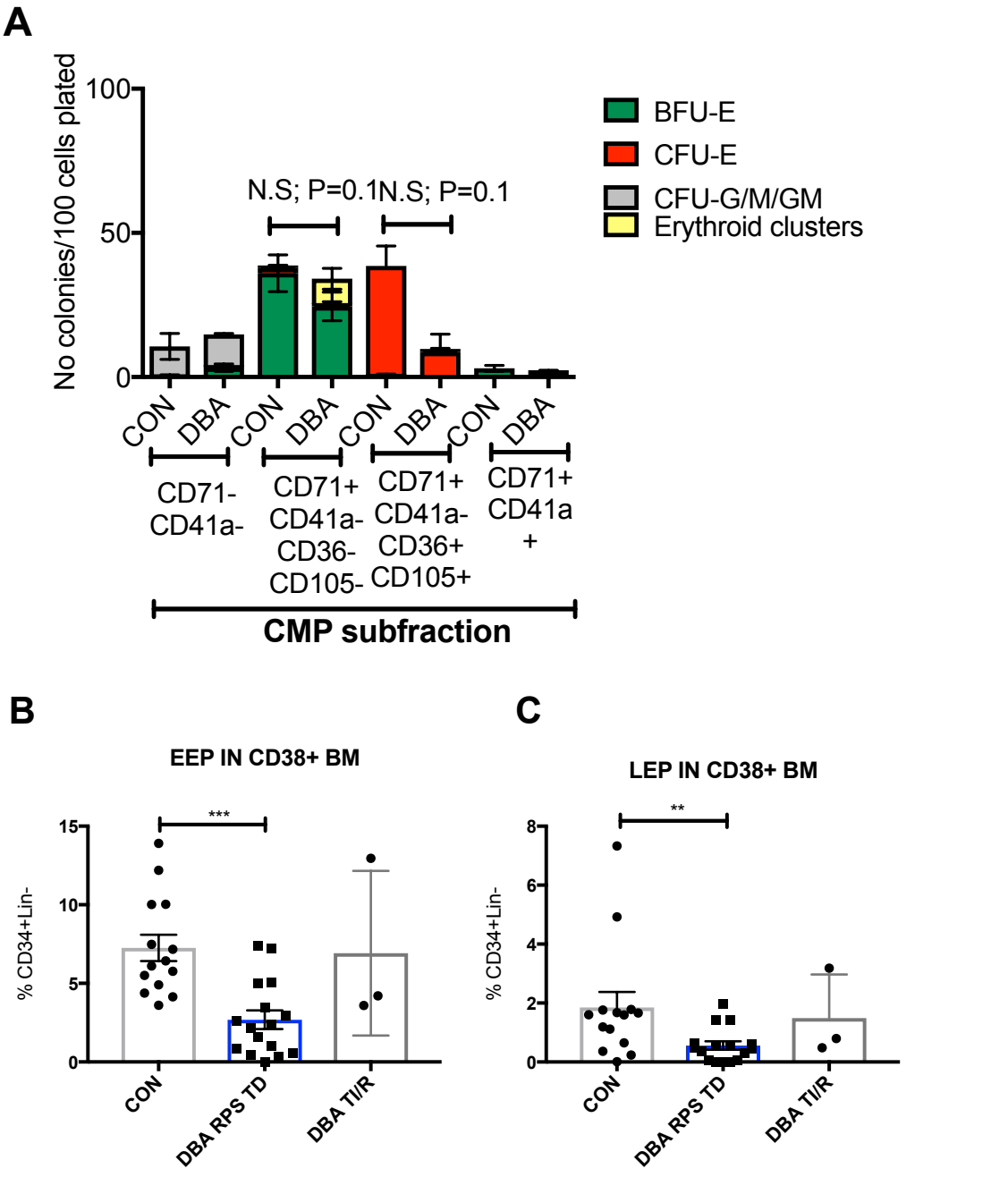
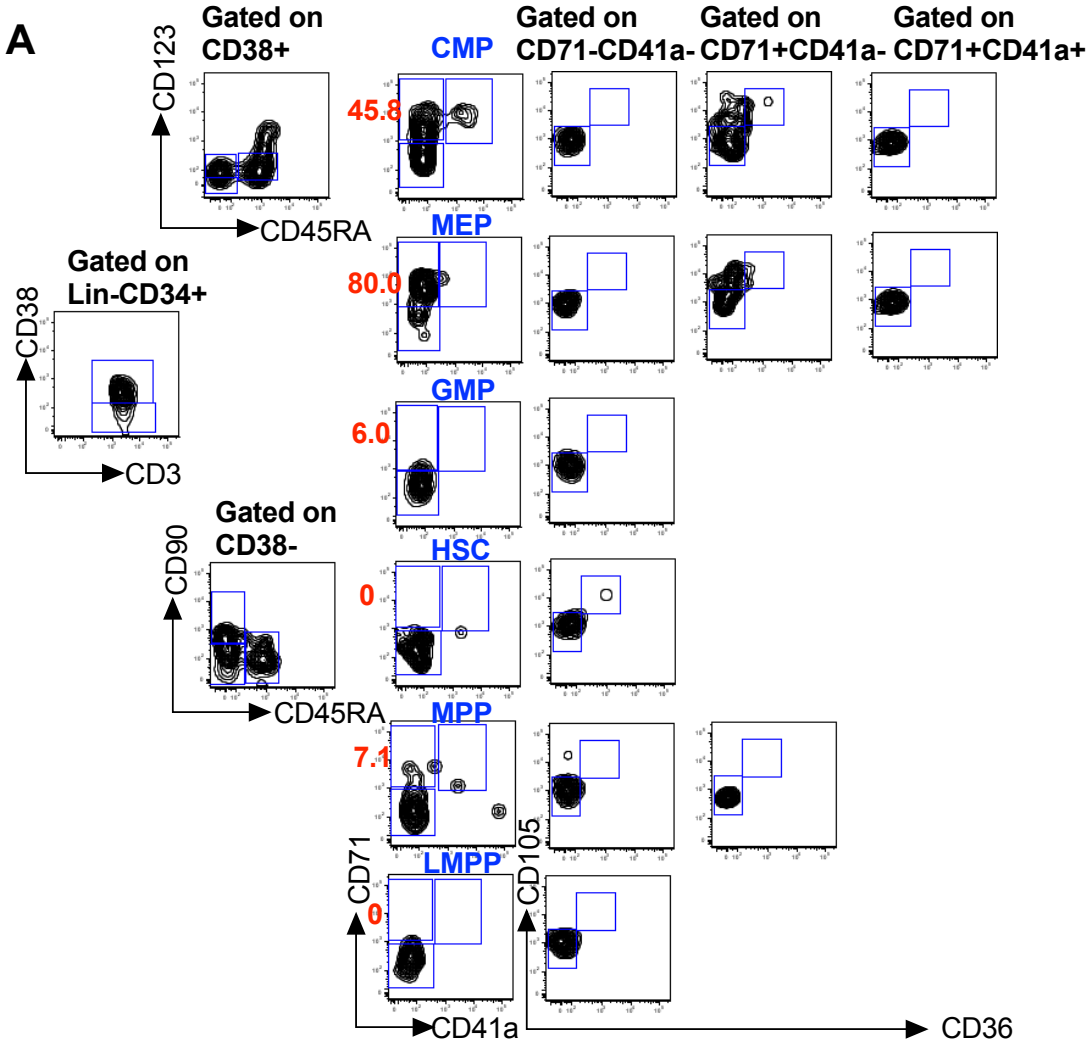


Figure 4:21 Purification of erythroid and myeloid progenitors from CMP and defective production of erythroid colonies from DBA CMP A) Clonogenic capacity of CMP subfractions flow-sorted in bulk from TD DBA patients (n=5) and age–matched controls (CON, n=4) **B**) Frequency of BM EEP and **C)** LEP in the CD38+ fraction (i.e. in CMP+MEP as GMP do not contain EP) of TD RPS DBA (n=16) and CON BM (n=14), as assessed by flow-cytometry.


A
Gated on CD38+
CD123
CD45RA
CMP
MEP
GMP
HSC
MPP
LMPP
45.8
80.0
6.0
0
7.1
0
Gated on Lin-CD34+
CD38
CD3
Gated on CD38-
CD90
CD45RA
Gated on CD71-CD41a-
Gated on CD71+CD41a-
Gated on CD71+CD41a+
CD71
CD105
CD41a
CD36

B

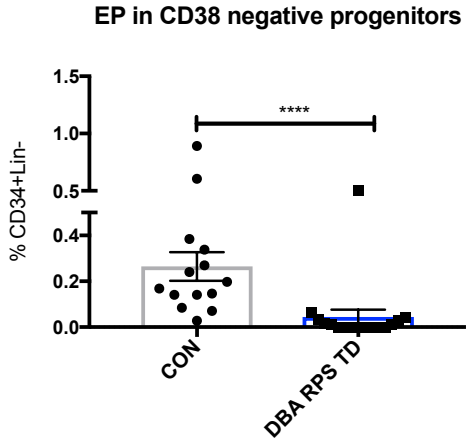

EP in CD38 negative progenitors
% CD34+Lin-
1.5
1.0
0.5
0.4
0.2
0.0
****
CON
DBA RPS TD

Figure 4:22 The CD71+ MPP compartment is reduced In TD DBA A) Expression of CD71, CD41a, CD105 and CD36 in HSPC populations. Representative plots from normal BM shown **B)** Cumulative data from CON and TD EP in the CD38- fraction (i.e. MPP) of TD *RPS* DBA (n=16) and CON BM (n=14), as assessed by flow-cytometry.

4.4 Discussion

To date, limited access to primary samples from patients with DBA and the marked genetic and clinical heterogeneity that typify the disease have precluded detailed characterisation of the haematopoietic hierarchy and elucidation of the precise nature of the erythroid defect. As a first step, I used well-established HSPC immunophenotypes to present the first *ex vivo* analysis of HSPC populations in DBA PBM from TD patients with mutations in *RPS* genes, and age-matched controls. Although I found preservation of HSC, MPP and LMPP frequencies in DBA, further work could include functional assessment of these populations, by clonogenic assays, stromal cell cultures to assess lymphoid output and xenotransplantation assays.

In terms of the myeloid progenitor compartment, for the first time I show that GMP and CMP frequencies are normal in TD DBA BM, whereas MEP are quantitatively reduced and have impaired ability to generate erythroid colonies *in vitro.* To explore the erythroid defect in MEP further, I characterised MEP using a unique combination of erythroid and megakaryocyte markers, with the aim of prospectively isolating BM EP populations, which had not previously been defined in humans.

4.4.1 Prospective isolation of human EP from normal BM

Using FACS-isolation coupled with morphological, clonogenic, liquid culture and RT-PCR assays, I demonstrate that EP reside in the CD71+CD41a- subfraction of MEP.

Consistent with this, our group (in collaboration with others) published transcriptional profiles derived from single cells showing that CD71+CD41a- and CD71+CD41+ MEP are E and MK biased, respectively (Psaila et al. 2016). I also demonstrate that BFU-E can be distinguished from CFU-E by their expression of CD36 and CD105. Specifically, CD105-CD36^{-} CD71+CD41a-MEP (termed 'EEP') generate BFU-E with 67% clonogenicity at a single cell level while CD105+CD36+ CD71+CD41a MEP (termed 'LEP') give rise to CFU-E with 58% clonogenicity. Both populations yield their respective E colonies with >95% purity (Iskander et al. 2015). As the morphological distinction between BFU-E and CFU-E colonies is arguably subjective, I provide further evidence that my candidate phenotypes enrich for 2 distinct populations. As well as exhibiting distinct morphological features and transcriptional profiles, their behaviour in culture is consistent with EEP arising upstream of LEP in the haematopoietic hierarchy.

Interestingly, another sorting strategy for BFU-E and CFU-E progenitors was also published in the last 3 years (Li et al. 2014). While my work corroborated the investigators' definition of BFU-E progenitors, there was an important distinction in the immunophenotype of CFU-E progenitors, which they described as CD34-. The reasons for this difference are unclear, however it may relate to the fact that *Li et al* reached their candidate phenotypes based on *in vitro* cultured UCB CD34+ cells, whereas my approach involved direct *ex vivo* analysis of BM haematopoietic cells. As CFU-E activity has previously been localised to the CD34+ subtraction of BM (Strauss et al. 1986; Pereira et al. 2007), my investigation focused on CD34+ selected cells *a priori*. However, in subsequent experiments I demonstrate 4-fold enrichment in CFU-E activity in flow-sorted Lin-**CD34+**CD71+CD41a-CD105+CD36+ versus Lin-**CD34-** CD71+CD41a-CD105+CD36+ cells. This is consistent with the work of Strauss *et a*l (Strauss et al. 1986), which showed that the vast majority of CFU-activity is contained in CD34$^{hi/med}$ and not CD34- cells. More recently, Weissman's group also reported that BM-derived CD34+CD71+CD105+MEP enrich for CFU-E activity (Mori et al. 2015), consistent with the phenotypes I report. Ultimately, the existence of a

range of sorting strategies for HSPC, and now EP (Figure 4:24), point to the fact that *in vivo*, the situation is likely to be more complex than can be recapitulated *in vitro* or even *ex vivo*. As opposed to 2 distinct EP populations, erythroid maturation *in vivo* plausibly involves a gradual trajectory from BFU-E to CFU-E to proEB via multiple intermediary progenitor cell types.

Next I show that characterisation of CMP in normal BM, with the same repertoire of markers used on MEP, also purifies BFU-E and CFU progenitors. Furthermore, CD71 expression within CMP permits robust separation of erythroid from myeloid progenitors. This finding has 2 important ramifications; one conceptual and one practical. First, it casts doubt over the existence *per se* of a common myeloid progenitor with dual potency for the myeloid and erythroid lineages. Other groups have recently provided evidence supporting the hypothesis that 'CMP' in fact comprises a mixture of unipotent MK, E and myeloid progenitors (Miyawaki et al. 2017; Notta et al. 2016). Second, it suggests that functionally homogeneous cell populations can be isolated from the combined CMP and MEP subfractions, increasing the absolute numbers of cells available for my subsequent experiments (Chapter 5).

Importantly, my data reveal that BFU-E progenitors also reside in the CD71- MEP population (Figure 4:16). Here they cannot be purified from GM progenitors using the collection of cell surface markers included in my study. Whether these E colonies represent the progeny of multi-/bi-potent progenitors or are the earliest unipotent EP remains to be resolved. It is likely that recent advances in single cell technologies, for example single cell immunophenotyping, transcriptome profiling and xenotransplantation assays, will facilitate these discoveries. In parallel, it will be possible to apply even greater precision to mapping out the haematopoietic defects in DBA.

In summary, CD71 is a useful marker for erythroid commitment that arises early in the haematopoietic hierarchy, at the level of MPP. Within CD71+CD41a- EP, differentiation from CD38- to CD38+ is associated with increasing purity and clonogenic efficiency of EEP. The transition from EEP to LEP is marked by acquisition of CD105 and CD36 with gradual los of CD34 (Figure 4:24). BFU-E colonies can also be generated from HSC (which are CD71-), and the CD71- subfraction of MEP (Figure 4:21), suggesting the existence of as yet unknown surface markers that will help in the isolation of erythroid-biased progenitors from more primitive HSPC subsets.

4.4.2 Progenitor cell mechanisms underpinning the erythroid defect in DBA

As described above, my first observation was a selective reduction in E colonies generated from MEP in DBA BM. By contrast MEP-derived myeloid colonies are preserved in DBA. Although classically MEP refers to progenitors lacking in GM potential, GM progeny from MEP has also been observed by other groups (Edvardsson et al. 2006) (Mori et al. 2015). This may be attributable to the fact that MEP is classically a negatively defined population, which is particularly difficult to discriminate from CMP and MPP using CD123 and CD38 expression alone. The novel sorting strategy above helps to circumvent this problem.

As the methylcellulose medium used in these experiments does not support MK development, the MK output of MEP in DBA was not specifically assessed. Notably, I did not observe any differences in HSPC frequency/function according to age in DBA (range 10m-15y2m), however the numbers of primary BM samples available for different age groups of patients are not sufficient to comprehensively address this question.

To validate my strategy for characterisation of normal EP within MEP, and to provide insight into the pathogenesis of DBA, I characterised DBA BM using the same gating approach described above. A strength of the work is that DBA PBM samples were

compared with age-matched controls. However, many of the CON BM samples were collected from patients with a haemoglobinopathy trait, which may lead to subtle erythroid defects. To mitigate this, I included at least one true normal control (i.e. no haemoglobinopathy trait) as a biological replicate in each experiment, for example when measuring frequencies of EP and EB populations (Figure 4:23).

For the first time, I present evidence that EEP and LEP are reduced quantitatively in DBA BM and PB from TD *RPS* patients. Moreover, I demonstrate, using FACS-isolated single cells, that EP are functionally impaired. This manifests as failure to produce a colony or production of an abnormal colony in semi-solid culture, or reduced proliferation in liquid culture. The heterogeneity of colony morphology within a single patient and the existence of some apparently normal BFU-E/CFU-E is an intriguing observation as it suggests that a single cell –omics approach in DBA HSPC may yield fascinating data with therapeutic implications.

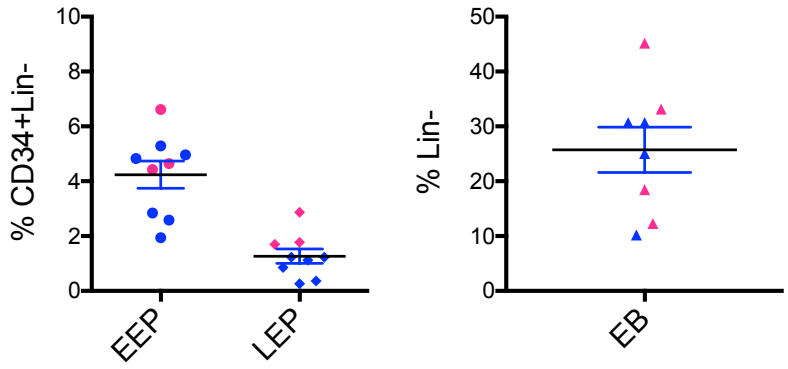
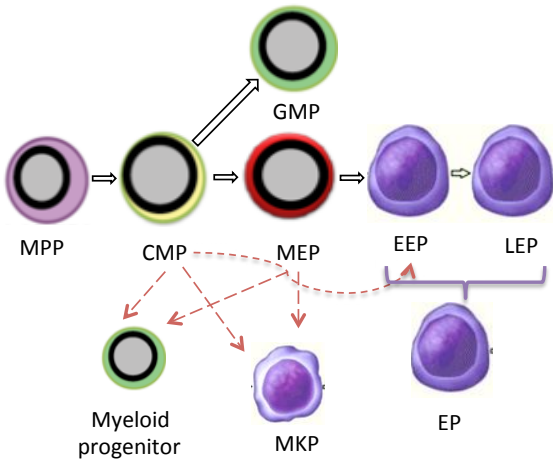


Figure 4:23 Normal donor EP/EB frequencies, showing comparable data from haematologically normal donors (pink symbols) and donors with sickle cell/ thalassaemia trait (blue symbols).

In conclusion, I have developed a novel approach allowing the prospective identification, enumeration and isolation of human BM EP that give rise directly and selectively to BFU-E and CFU-E. Moreover, using this approach I have demonstrated

that in RPS gene-associated DBA, cumulative defects in the frequency and function of early and late EP are likely to underpin the striking paucity of EB.

Finally, the finding that EEP are reduced in frequency suggests that the erythroid defect in DBA has its origins upstream of these EP populations. Pertaining to this I demonstrate a three-fold reduction in the CD71+ MPP subpopulation in DBA BM. If supported by more detailed studies, this would constitute the first evidence for impaired erythroid lineage commitment in primitive HSPC as a cellular mechanism in DBA.



	Lin	CD 34	CD 38	CD45 RA	CD 90	CD12 3	CD 71	CD41 a	CD 36	CD10 5	Tissue: REFS
MPP	-	+	-	-	-	-/LO	-/+	-/+	-	-	UCB: (Majeti et al. 2007), BM: current work
GMP	-	+	-	+	NA	LO	-	-	-	-	BM: (Manz et al. 2002), current work
CMP	-	+	+	-		LO	-/+	-/+	-/+	-/+	BM: (Manz et al. 2002), current work
MEP	-	+	+	-		-	-/+	-/+	-/+	-/+	BM: (Manz et al. 2002), current work
Bipotent MEP	-	+	MID	-		NA	NA	-	-	NA	Also FLT3-MPL+. GCSF-mobilised PB/BM: (Sanada et al. 2016)
Myeloid progenitor	-	+	+	-		-/LO	-	-	-	-	BM: current work
MKP	-	+	+	-		-/LO	LO	+	-/+	-	BM: current work, (Mori et al. 2015) GCSF-mobilised

											PB/ BM: (Psaila et al. 2016)
	-	+	+	-		LO	NA	+	NA	NA	BM: (Miyawaki et al. 2017)
EP	-	+	+	-		-/LO	HI	-	-/+	-/+	BM: current work, (Mori et al. 2015)
BFU-E/ EEP	-	+	HI	-		-/LO	+	-	-	-	BM: (Iskander et al. 2015)
	-	+	NA	NA-		-	LO	NA	-	NA	PB: (Li et al. 2014)
CFU-E/ LEP	-	+	HI	-		-/LO	+	-	+	+	BM: (Iskander et al. 2015)
	-	-	NA	-NA		-	HI	NA	+	NA	PB: (Li et al. 2014)

Figure 4:24 Different strategies for the prospective isolation of HSPC from MPP to LEP. A) The origins of EP according to classical (block arrows) and alternative (red dashed arrows) models of haematopoiesis. Developmental relationships between populations are shown by arrows. Alternative models suggest that unipotent myeloid, MK and E progenitors can develop directly from MEP or from CMP (without passing through bi- or multi-potent stages). It is not yet known whether MPP also encompass unipotent myeloid, MK and/or E progenitors. **B)** Immunophenotypic characterisation of HSPC populations from MPP to LEP, according to my data and the work of other investigators. Similarities and differences between strategies used to FACS-isolate these populations in recent studies are shown. Red font highlights the specific contributions from my work. [Cell images available from: http://epomedicine.com/wp-content/uploads/2017/09/erythropoiesis.jpg]

5 POTENTIAL MECHANISMS OF ERYTHROID FAILURE AND THE ROLE OF GLUCOCORTICOID THERAPY IN DBA

5.1 Introduction and aims

In Chapter 3 I presented clinical data showing that the haematological phenotype in DBA is influenced by the underlying RP gene mutated. Specifically, patients with pathogenic variants in *RPL5/11* present with anaemia later and display a higher initial response rate and longer duration of response to CS therapy compared with patients with pathogenic variants in *RPS* genes. Subsequently, In Chapter 4, I demonstrated that in RPS gene-associated TD DBA, EEP are reduced in frequency and function. In turn there is a severe loss of LEP and near absence of EB. However, these cellular defects do not explain the heterogeneity in haematological phenotype i.e., RPS vs RPL gene-associated DBA, which raises the question of whether the cellular defect in RPL gene-associated DBA is distinct. To address this question, in this chapter I explore the hypothesis that RPL versus RPS genotype influences the HSPC phenotype thereby contributing to the divergent clinical features observed in these 2 groups. I also explore the molecular mechanisms underpinning erythroid failure in RPL gene-associated DBA, by analysing the transcriptome of RPL gene-mutated EB.

Finally, I aim to define the effect of GC on progenitor cells *ex vivo* and *in vitro.* While GC-mediated stimulation of erythropoiesis has been studied comprehensively in normal mouse FL (Bauer et al. 1999; Leberbauer 2005; Flygare et al. 2011; L. Zhang et al. 2013), there is a scarcity of data from human systems and, in particular, data specific for DBA.

5.2 Methodological approach

All methods were performed as described in Chapter 2 and reagents used are listed in the Appendices. Descriptive statistics in this chapter show mean±S.E.M. The unpaired Student *t*-tes or Mann-Whitney U were used as appropriate (GraphPad Prism v6). Where there are no significant differences between groups, P values are

not shown. Clinical details associated with normal and DBA samples are show in Tables 4.1 and 4.2.

5.3 Results

5.3.1 Divergent progenitor cell phenotypes in *RPS* and *RPL* genotypes

Using the immunophenotypic approaches outlined in Chapter 4, normal (n=14), TD *RPS* DBA (n=13) and TD *RPL* DBA (n=6) BM samples were compared in terms of HSPC populations. While there are no significant differences in HSPC frequencies between *RPS* and *RPL* TD patients, MEP are *not* significantly reduced in patients with the RPL5/ RPL11 genotype, in contrast to patients with the RPS genotype or those with no known mutation (NKM; Figure 5:1).

Next I examined the frequency of CD71+ EP within the CD38- and CD38+ subfractions of BM (Figure 5:2). Although EEP are reduced in both DBA subgroups compared to CON, patients with *RPL5/11* variants have higher frequencies of LEP (p=0.006). Given that 4 of the 5 patients with RPL variants included in the study are female, I also compared the relative frequencies of EEP and LEP in females and males with RPS variants (Figure 5:3). Again, there is no difference suggesting that female gender *per se* is not associated with increased EP frequencies in DBA.

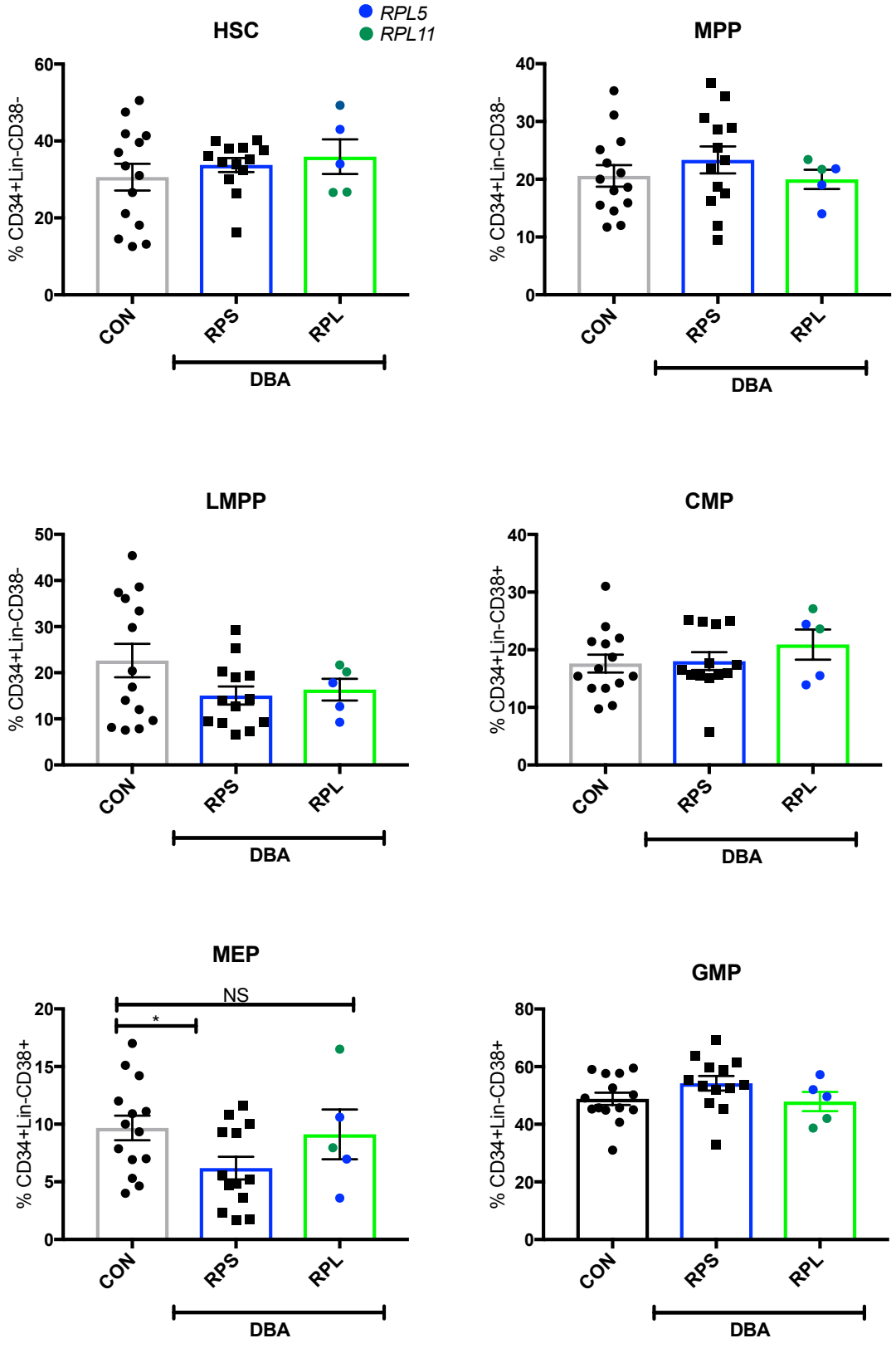

RPL5
RPL11
HSC
% CD34+Lin-CD38-
60
40
20
0
CON
RPS
RPL
DBA
MPP
% CD34+Lin-CD38-
40
30
20
10
0
CON
RPS
RPL
DBA
LMPP
% CD34+Lin-CD38-
50
40
30
20
10
0
CON
RPS
RPL
DBA
CMP
% CD34+Lin-CD38+
40
30
20
10
0
CON
RPS
RPL
DBA
MEP
NS
*
% CD34+Lin-CD38+
20
15
10
5
0
CON
RPS
RPL
DBA
GMP
% CD34+Lin-CD38+
80
60
40
20
0
CON
RPS
RPL
DBA

Figure 5:1 MEP are quantitatively reduced in *RPS* but not *RPL* TD DBA BM. Shown are frequencies of CMP/ GMP/ MEP progenitors, expressed as % of CD34+Lin-CD38+ cells, and HSC/MPP/LMPP progenitors, expressed as % of CD34+Lin-CD38- cells in BM of CON (n=14), *RPS* TD DBA (n=13) and *RPL* TD DBA (n=5). In the RPL group, green symbols indicate a patient with a frameshift insertion in *RPL11* (2 BMs performed 2 years apart in the same patient) while blue symbols indicate 3 patients with *RPL5* pathogenic variants.

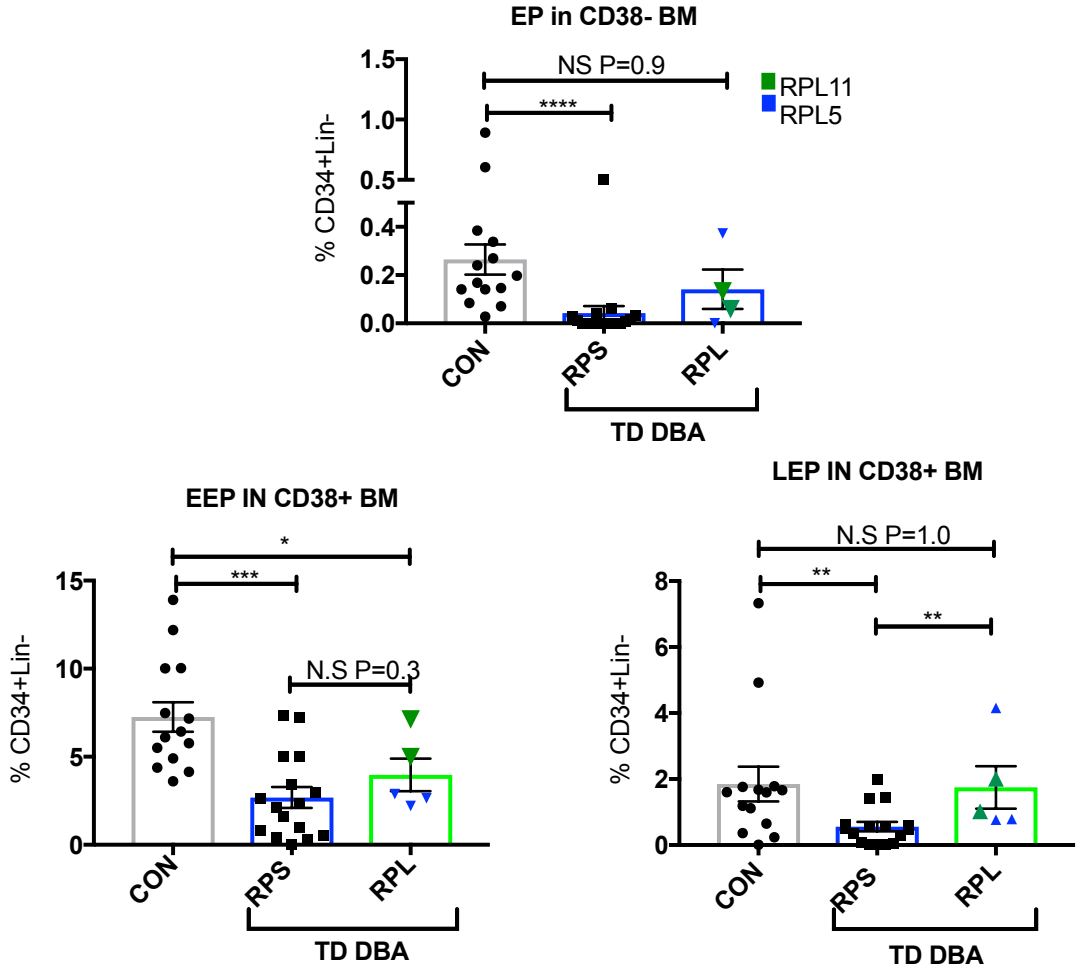


Figure 5:2 LEP are increased in TD DBA secondary to mutations in *RPL5/11,* compared with *RPS* gene variants. Frequencies of CD71+MPP, EEP and LEP in the CD38+ fraction (i.e. CMP+MEP) of TD *RPS* DBA (n=16), TD *RPL* DBA (n=5) and CON BM (n=14), as assessed by flow-cytometry.

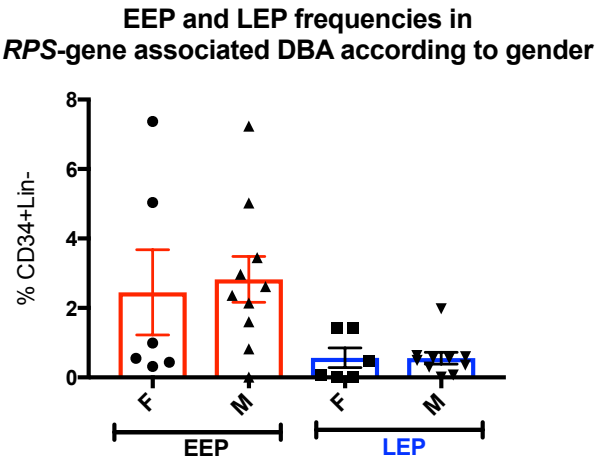
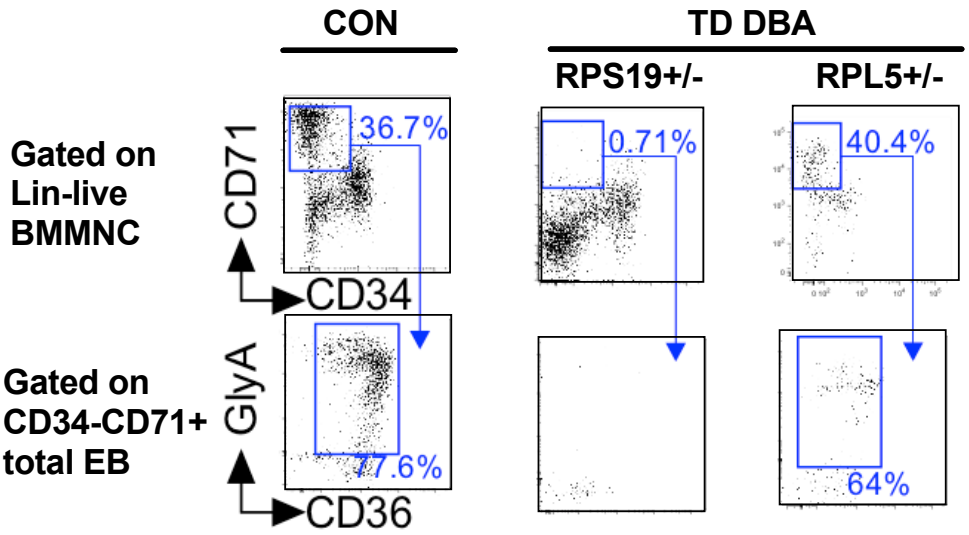


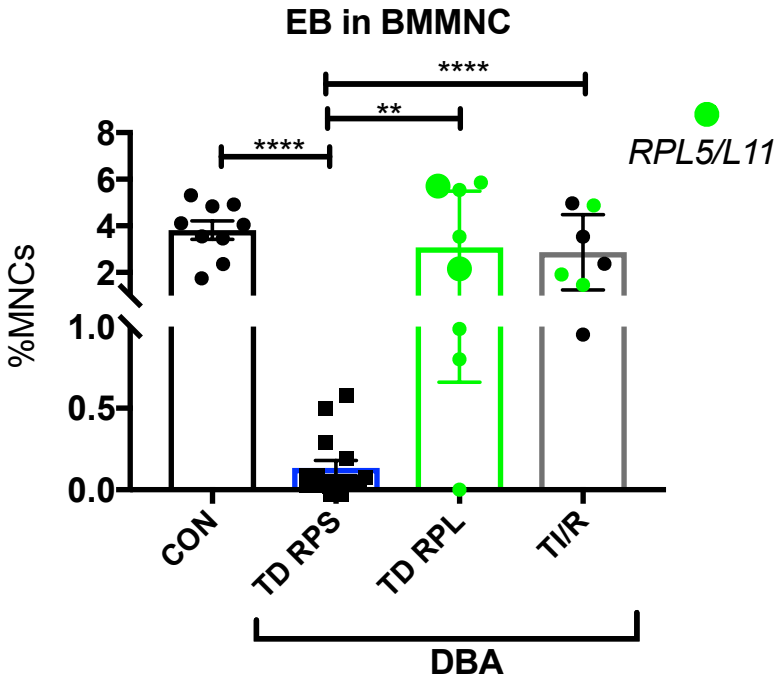
Figure 5:3 No influence of gender on EEP and LEP frequencies in BM of TD patients with DBA secondary to RPS-gene variants. Frequencies of EEP and LEP in the CD38+ fraction (i.e. CMP+MEP) of female (n=7) and male (n=11) patients with TD RPS- gene associated DBA, as assessed by flow-cytometry.

Having shown a preservation of MEP and increase in EP frequencies in TD *RPL* DBA relative to TD *RPS* DBA, next I enumerated mature erythroid precursors/ EB (immunophenotypcially defined as Lin-CD34-CD71+CD36+GlyA-/+), in fresh BM samples *ex vivo.* This revealed 2 markedly distinct phenotypes in TD patients with DBA (Figure 5:4A): while patients with *RPS*-gene associated DBA have virtually no EB, in patients with *RPL*-gene associated DBA, EB are present, suggesting a degree of preservation of mature erythropoiesis (Figure 5:4B). To validate these findings, I collected differential cell count data from bone marrow aspirates (BMA) reports in patients with DBA over the last 10 years (Figure 5:4C&D). In line with the flow-cytometric analysis, I found that EB frequency was significantly higher in RPL compared to RPS gene-associated DBA BM (Figure 5:4C). Similarly, the M:E ratio was significantly lower in TD RPL rather than RPS-associated DBA (Figure 5:4D). These data are consistent with the less severe haematological phenotype in RPL- versus RPS-gene associated DBA (described in Chapter 3).

A


CON
TD DBA
RPS19+/-
RPL5+/-
Gated on Lin-live BMMNC
CD71
CD34
36.7%
0.71%
40.4%
Gated on CD34-CD71+ total EB
GlyA
CD36
77.6%
64%

B


EB in BMMNC
****
**
****
RPL5/L11
%MNCs
8
6
4
2
1.0
0.5
0.0
CON
TD RPS
TD RPL
TI/R
DBA

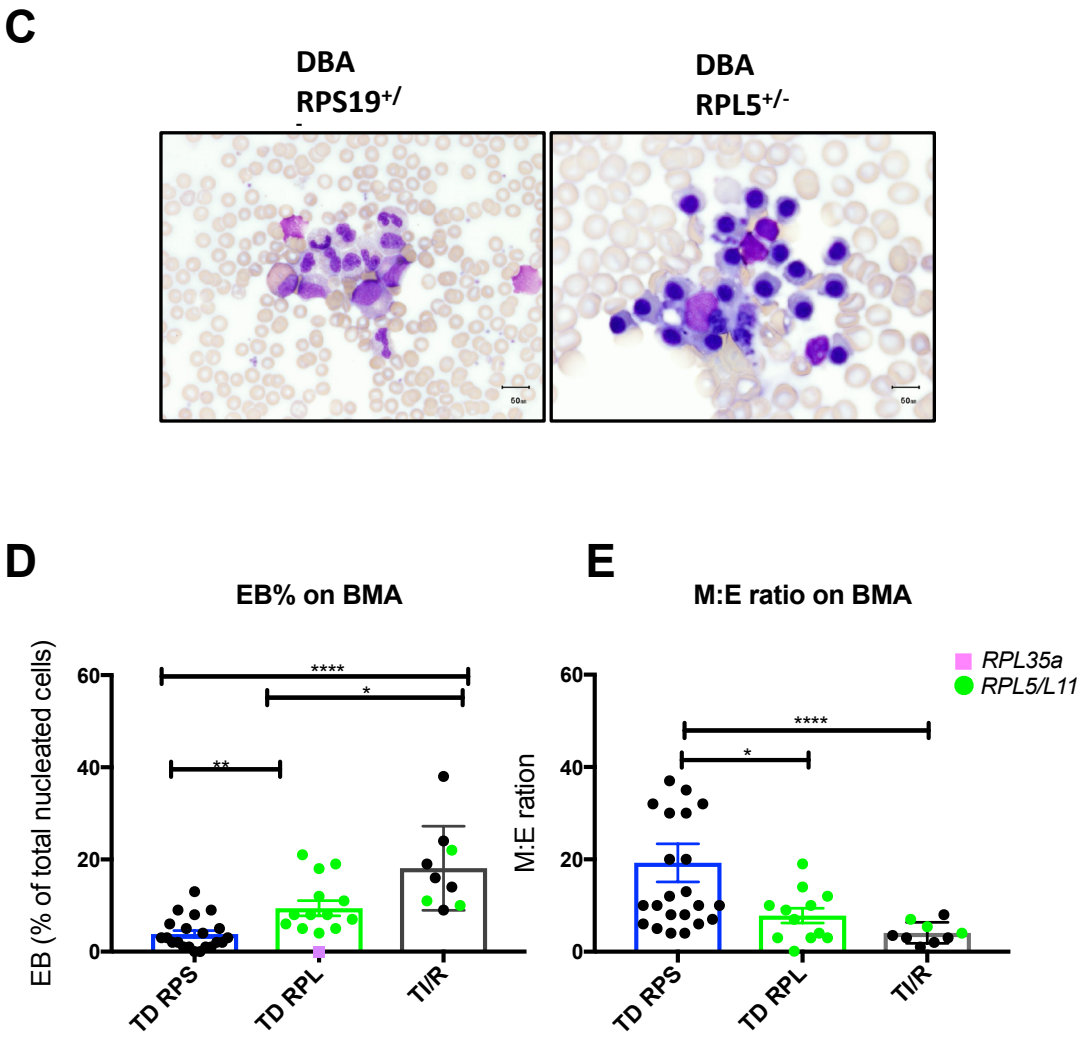


Figure 5:4 Relative preservation of EB in RPL TD DBA. A) Representative flow cytometry plots showing Lin-CD34-CD71+CD36+GlyA-/+ EB in fresh BM MNCs from CON and TD *RPL* patient with DBA but absent in patient with *RPS* TD DBA **B)** EB frequencies, expressed as % of total MNCs, in BM from CON (n=9), TD *RPS* (n=15), TD *RPL* (n=8) and TI/R (n=7) DBA BM. EB are reduced in TD *RPS* DBA relative to CON (as shown in Chapter 4), and increased in TD *RPL* and TI/R relative to TD *RPS DBA* **C)** Representative morphological appearances of BMA (stained with haematoxylin and eosin) in a patient with a mutation in *RPS19* (left), showing plentiful myeloid cells with scarce EB and another patient with a mutation in *RPL5* (right), where EB at all stages of differentiation can be seen clustered around a macrophage, forming an EB island. **D)** EB, expressed as a % of total nucleated cells, enumerated by reporting

clinicians on slide preparations of BMA from TD *RPS* (n=21), TD *RPL* (n=12), and TI/R DBA (n=9). Single TD patient with *RPL35a* deletion and TI/R patients with pathogenic variants in *RPL5/11* are highlighted by coloured symbols as shown in the figure legend. **E)** Ratio of absolute numbers of myeloid to erythroid cells in BMA, as determined by reporting clinicians. Patients with *RPL*-gene associated DBA have a higher % of residual EB and a lower M:E ratio than those with *RPS*-gene associated DBA.

As chronic blood transfusion has been reported to suppress endogenous erythropoiesis, I verified that TD patients with pathogenic variants in *RPS* genes had not been exposed to a longer duration of transfusion than *RPL* cases (6.8y±5.7 versus 8.0y±6.7 respectively). I also confirmed that EB are not absent in transfusion-dependent anaemias *per se,* such as thalassaemia major and other congenital anaemias; rather this is a specific feature of RPS-gene associated DBA (Figure 5:5).

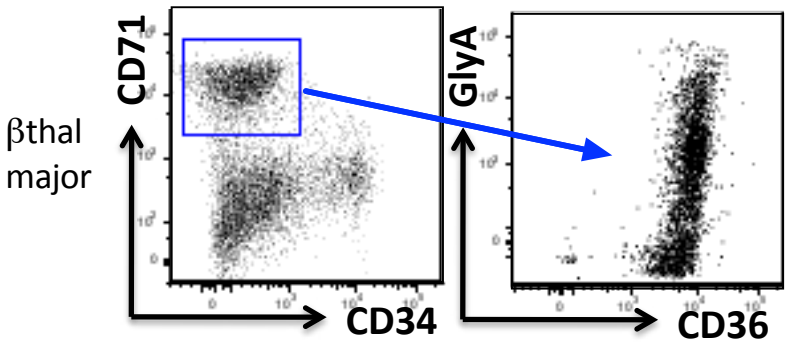


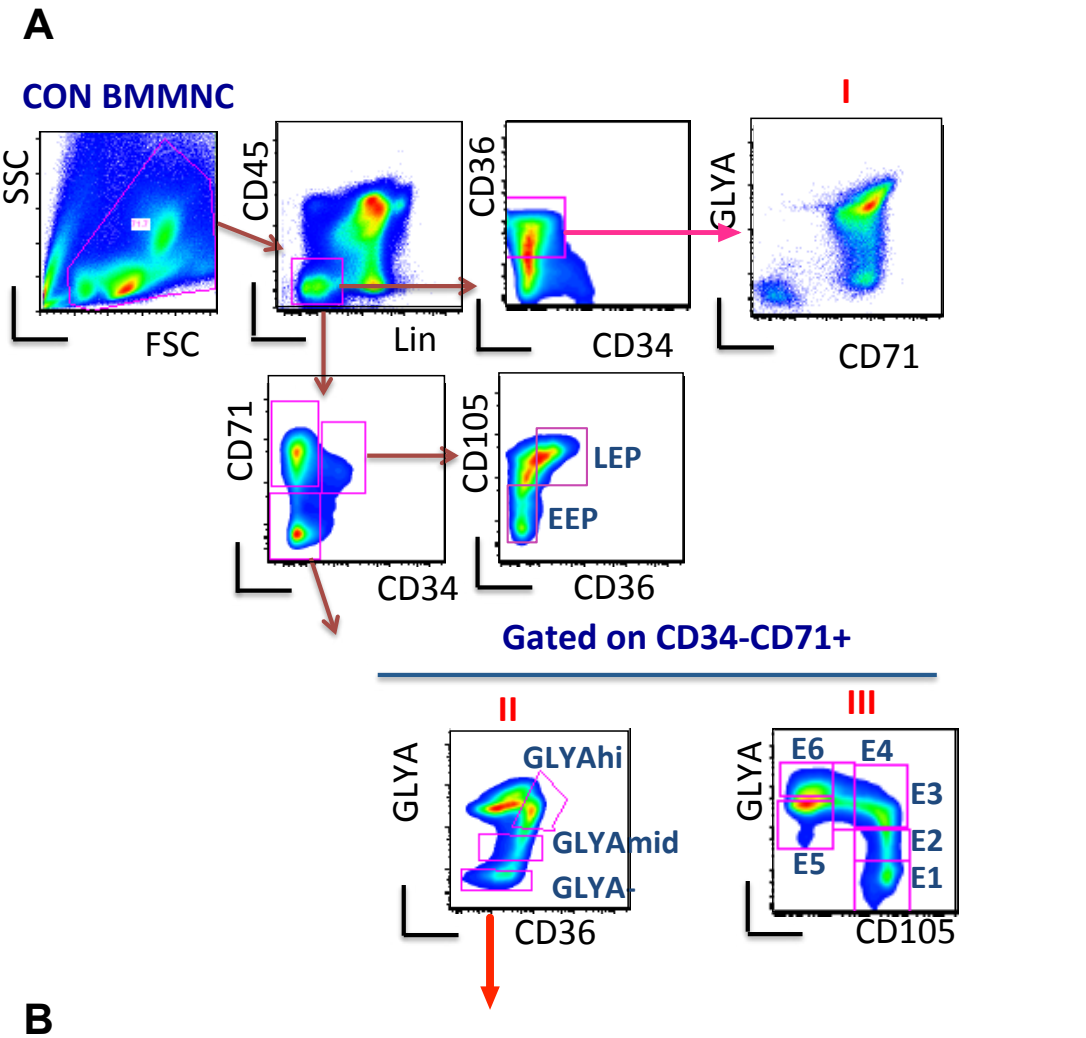
Figure 5:5 EB differentiation arrest is a specific feature of DBA. Representative plots from BMMNC of an 8y old child with transfusion-dependent β-thalassaemia major, showing plentiful CD34-CD71+CD36+GlyA-/+ EB.

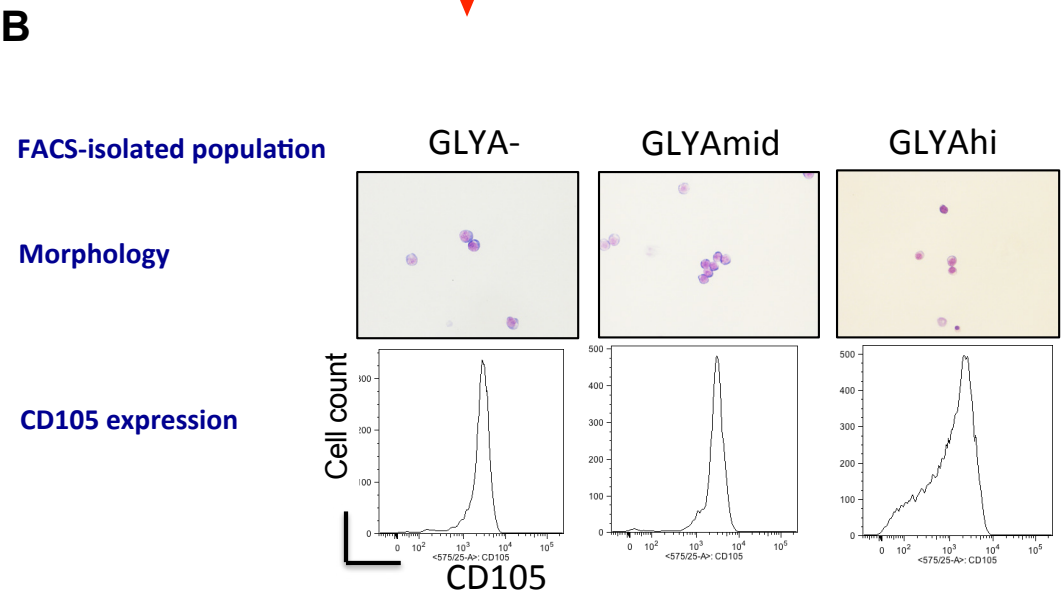
5.3.2 Prospective isolation of distinct EB stages in normal and DBA BM

Given my original observation that EB are present In RPL patients with DBA, yet these patients are still reliant upon 4 weekly transfusions, I hypothesised that, in this

subgroup of patients, erythropoietic defects at later stages of the hierarchy contribute to erythroid failure. A prerequisite to more precise characterisation of this defect is an ability to identify EB at equivalent stages of differentiation in normal and DBA BM.

Analogous to the lack of an immunophenotyping strategy for prospective isolation of human EP, until recently there were limited data on identification of distinct EB stages (pro, basophilic, orthochromatic and polychromatic EB). In one paper, transcriptionally distinct populations, from CFU-E to Late-E, were identified on the basis of CD71 and GLYA expression within CD36+ E lineage cells (Merryweather-Clarke et al. 2011). Here, homogenous cell populations were achieved by prospective isolation at different time points during an erythroid cell culture of human PBMCs. When applied to BM samples, which encompass a heterogeneous mixture of EB at different stages of differentiation, the pattern of CD71/GLYA expression does not clearly demarcate individual stages (Figure 5:6A: I). I then investigated whether CD36/GlyA expression would allow more precise delineation by examining the morphology of FACS-isolated populations (Figure 5:6A: II). I found that the CD36+GlyAhi fraction comprises a heterogeneous mixture of EB stages from basophilic EB to occasional reticulocytes. Within this population there is bimodal expression of CD105, another erythroid marker (Figure 5:6B). I therefore designed a new sorting strategy that delineates 6 distinct EB populations (E1-E6) according to CD105 and GlyA expression within CD71+ E lineage cells (Figure 5:6A: III and Figure 5:6C).


A
CON BMMNC
SSC
FSC
CD45
Lin
CD36
CD34
I
GLYA
CD71
CD71
CD34
CD105
CD36
LEP
EEP
Gated on CD34-CD71+
II
GLYA
GLYAhi
GLYAmid
GLYA-
CD36
III
GLYA
E6
E4
E3
E2
E5
E1
CD105


B
FACS-isolated population
GLYA-
GLYAmid
GLYAhi
Morphology
CD105 expression
Cell count
CD105

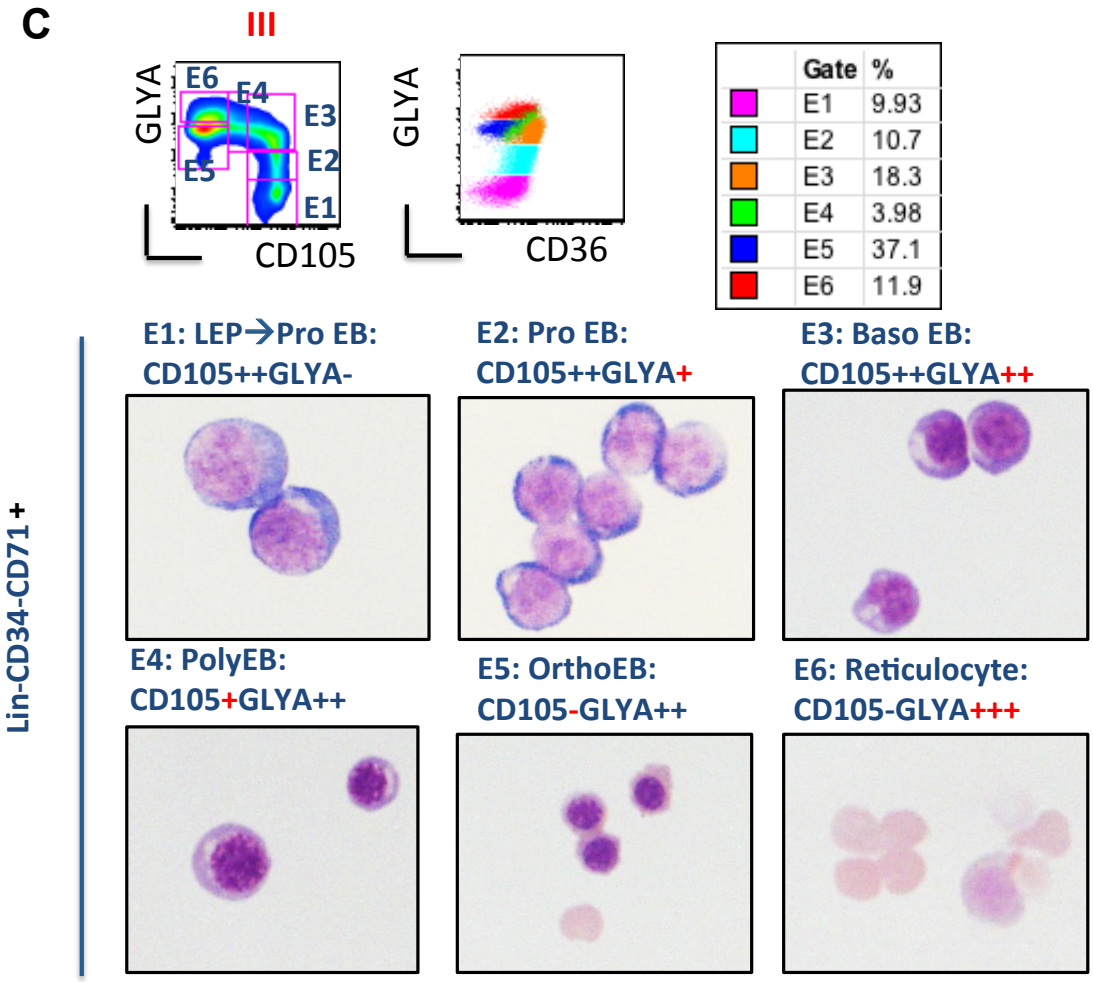


Figure 5:6 FACS-based strategy for prospective isolation of EB subpopulations A) Representative flow cytometry plots from a normal BM, showing 3 potential gating strategies (I-III) for prospective isolation of EB populations. First non-erythroid cells were excluded by gating on a population negative for lineage and CD45 cell surface markers. CD71 expression within the CD36+ population did not delineate distinct populations (strategy I). Within the CD34+CD71+ population, EEP and LEP could be delineated as expected. To characterise more mature EB, the CD34-CD71+ population was stratified using either CD36/GLYA (strategy II) or CD105/GLYA (strategy III). **B)** Morphology of cells FACS-isolated from GlyAlo, mid and hi subpopulations of CD36+ cells (strategy II). GlyAhi cells are heterogeneous in terms of morphology and show varying expression of CD105 **C)** Distinct morphology and immunophenotypes of cells FACS-isolated from E1-E6 (strategy III) using CD105 and

GlyA are shown. CD36 expression increases from pro EB to baso EB then decreases at the poly EB stage. The table shown relative proportions of these EB stages in the normal BM shown, expressed as a % of the total EB.

Using the approach outlined above, I assessed whether the distribution of distinct EB stages varies in normal BM and that from TD patients with DBA secondary to *RPL5/11* pathogenic variants. E3 was specifically reduced in DBA. A recent study showed accelerated E differentiation of RPS19+/- murine FL cells (Sjögren et al. 2015) and interestingly my data show that more mature populations appear to be increased, however this was not significant due to marked variability between biological replicates so firm conclusions cannot be reached.

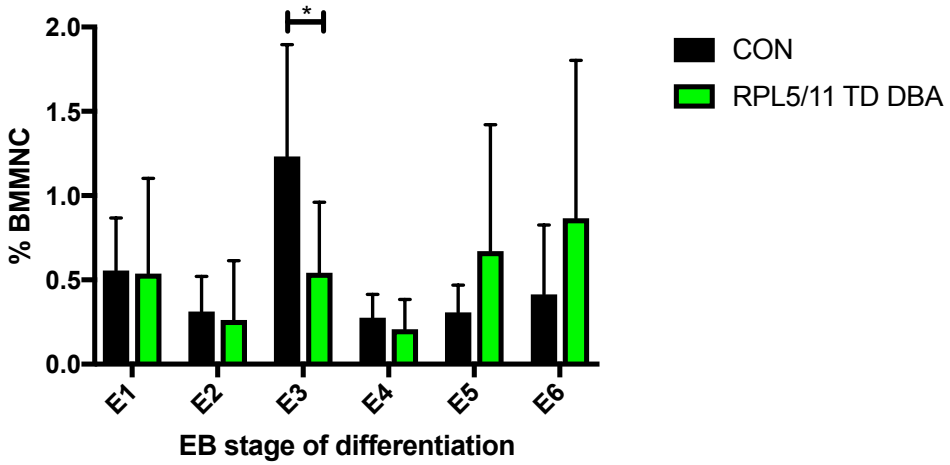


Figure 5:7 Relative proportions of EB subpopulations in CON (n=9) and TD RPL-gene associated DBA (n=7), as assessed by flow cytometry.

5.3.3 Molecular mechanisms of late erythroid failure in RPL gene-associated DBA

As discussed in Chapter 1, the later phases of erythropoiesis are difficult to model *in vitro*. In particular, enucleation relies upon physical interactions between EB and macrophages/stromal factors. Therefore, rather than performing cellular studies, I

used the limited available RPL samples to examine the transcriptional profile of purified *ex vivo* FACS-isolated EB, by RNA-seq. My specific aim was to compare gene expression in normal and DBA EB to gain insight into the genes and gene networks underpinning the late erythroid failure in these patients.

Given that EB in CON and DBA BM comprise different relative proportions of the distinct differentiation stages, it was important to examine gene expression in equivalent stages from CON and DBA BM. I chose to study basophilic EB (population E3) because it is a frequent EB subpopulation in normal BM and is reduced in DBA BM (Figure 5:6C). Furthermore basoEB are nucleated with active gene transcription.

I compared FACS-isolated E3 ($1x10^3$-$8x10^4$ cells per sample) from 3 healthy controls (3.5yM, 13yF, 18yF) with 3 patients with TD DBA caused by pathogenic variants in *RPL5/RPL11* (Table 5.1). Samples were therefore age- and sex- matched. I also performed RNA-seq on 1 patient pre and post CS, who was TI prior to starting CS, although the data were not included in the differential gene expression analysis as there is only a single replicate in these conditions. Patients with RPS-gene associated DBA were excluded as EB numbers are negligible in this subgroup.

Table 5.1 Summary of clinical characteristics of patients included in EB RNA-seq.

Patient ID	RP gene pathogenic variant	Age (y)	Gender	Treatment at time of BM	CS-response
044	*RPL11* c.267-268insG	16	F	TD	SR- lost response after 4y
116	*RPL5* c.175_176delGA	16	F	TD	CSI-stopped within 2m due to side effects
137	*RPL5 c.*166-	11	M	TD	PR

	169 del				
020	*RPL5*	15	F	TI; Hb 90g/L	CSR
088	c.625_626ins G	17		CSR: Hb 113g/L	CSR

BasoEB were FACS-isolated from thawed BM samples and total RNA extracted. An RNA Integrity score >8 was confirmed (Figure 5:8A) prior to mRNA isolation and preparation of cDNA libraries (Figure 5:8B). The technical quality of the RNA-sequencing data was further validated by FastQC software and bioinformatics analysis was performed as described in Section 2.6.2.

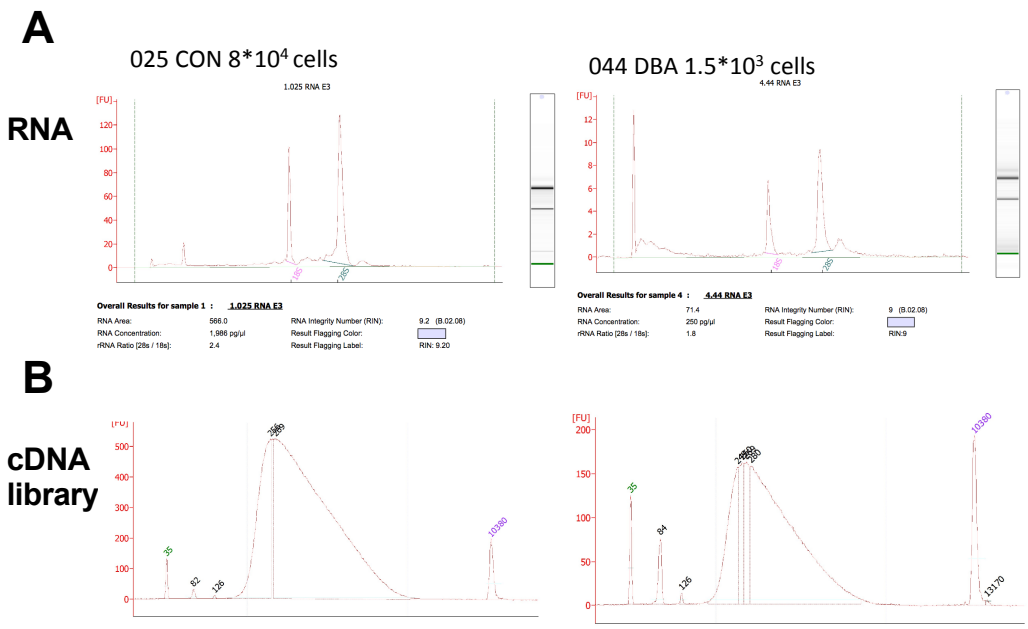
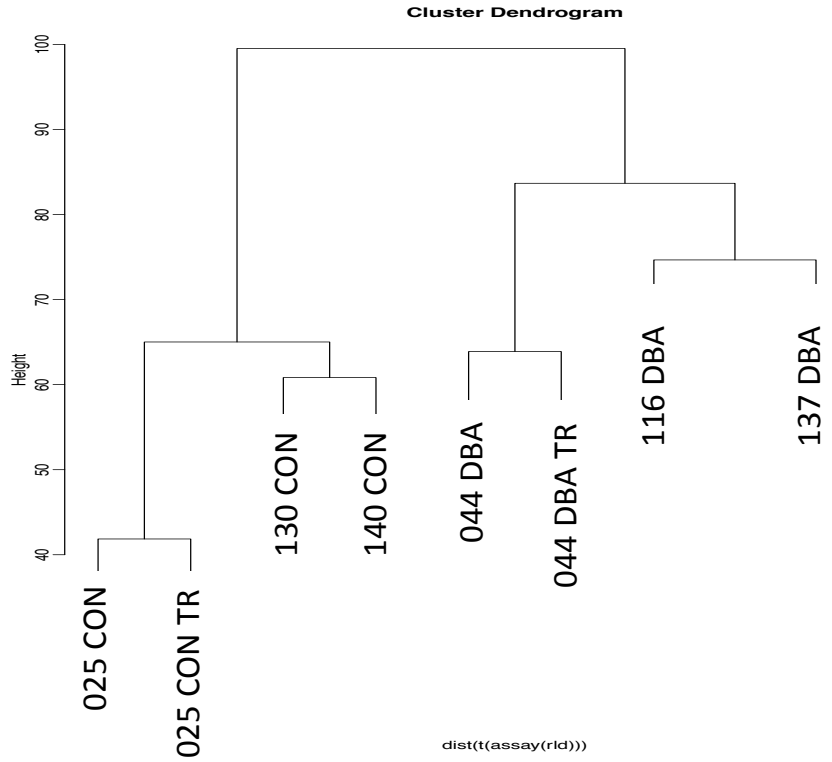


Figure 5:8 Quality control steps during preparation of cDNA libraries for RNA-sequencing. Example Agilent bioanalyser traces generated from a CON and DBA sample are shown for **A)** RNA, showing the 28S and 18S rRNA peaks and **B)** cDNA showing a prominent peak at approximately 300bp representing the library for sequencing. Small peaks at 80 and 130 bp show excess primers and primer dimers respectively. Note 32S rRNA peak in *RPL5*-mutated sample.

5.3.3.1 DBA EB have a unique transcriptional signature

Hierarchical clustering (Figure 5:9) showed clear separation between CON and DBA samples, with high correlation between technical replicates. Comparison of CON and TD DBA samples (n=3) revealed 1709 differentially expressed genes [adjusted P value (Padj) <0.05, log2 fold change: 0.6) with 1011 genes being regulated (Figure 5:10). Gene expression data were validated by: i) confirming 50% reduction in the mRNA of the specific RP gene mutated in each patient ii) confirming upregulation of eADA and dysferlin, two genes encoding proteins that are known to be highly expressed in DBA RBC (Figure 5:11). Dysferlin is a protein recently found to be aberrantly expressed on the RBC membrane of non-transfused patients with DBA (Pesciotta et al. 2014) and in my dataset was one of the 50 most differentially expressed genes. Interestingly *HBG2*, a gene encoding gamma globin (Table 5.2) was highly upregulated in DBA EB, compatible with their 'erythropoietic stress' phenotype.



B

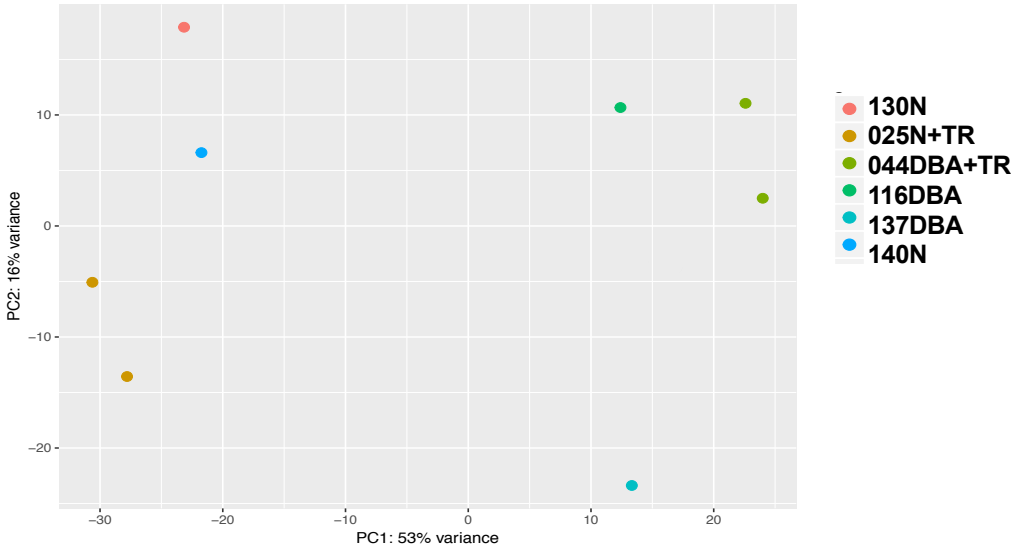


Figure 5:9 Unsupervised hierarchical clustering of normal and DBA EB transcriptomes, generated by RNA-seq. A) Unsupervised hierarchical clustering of samples by expression levels of the most variant features shows partitioning (top dendrogram split) predominantly by disease status (i.e, CON vs DBA). Technical replicates (TR) i.e., cDNA libraries prepared from the same RNA stock but in separate batches and sequenced on different runs, clustered together as expected. **B)** Principal component analysis (PCA) depiction of gene expression data, showing separation between CON and DBA samples when distributed along PC1 and PC2.

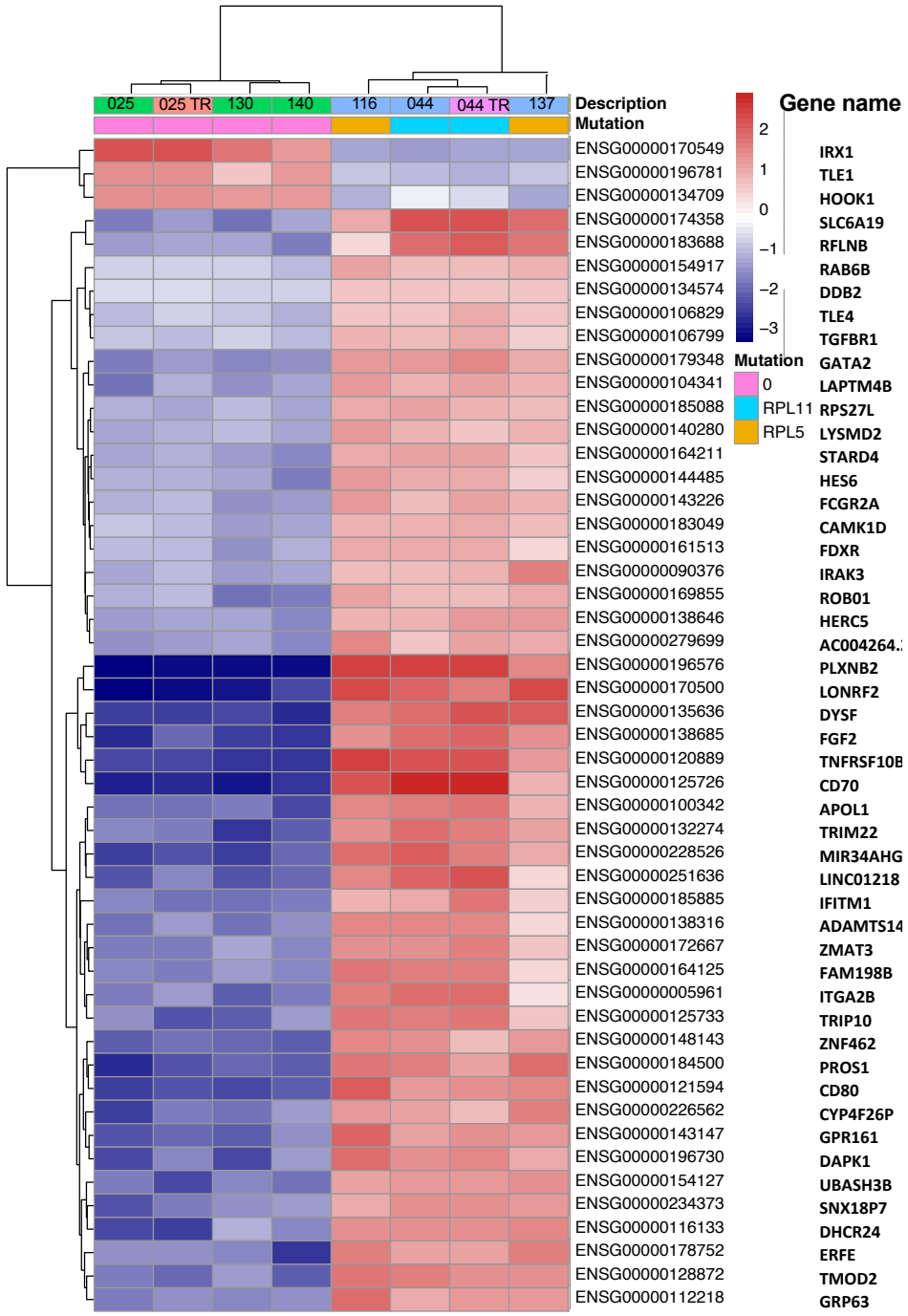


Figure 5:10 Heat map showing 50 most differentially expressed genes between CON and DBA EB. Each column is an individual sample and each row is a gene. Corresponding Ensmebl gene IDs and gene names are shown on the right. Red and blue colours denote increasing and decreasing gene expression, respectively. Differential gene expression analysis between CON and TD DBA (n=3) was performed using DESEQ2. 47 of the 50 most DGE are upregulated in DBA relative to CON.

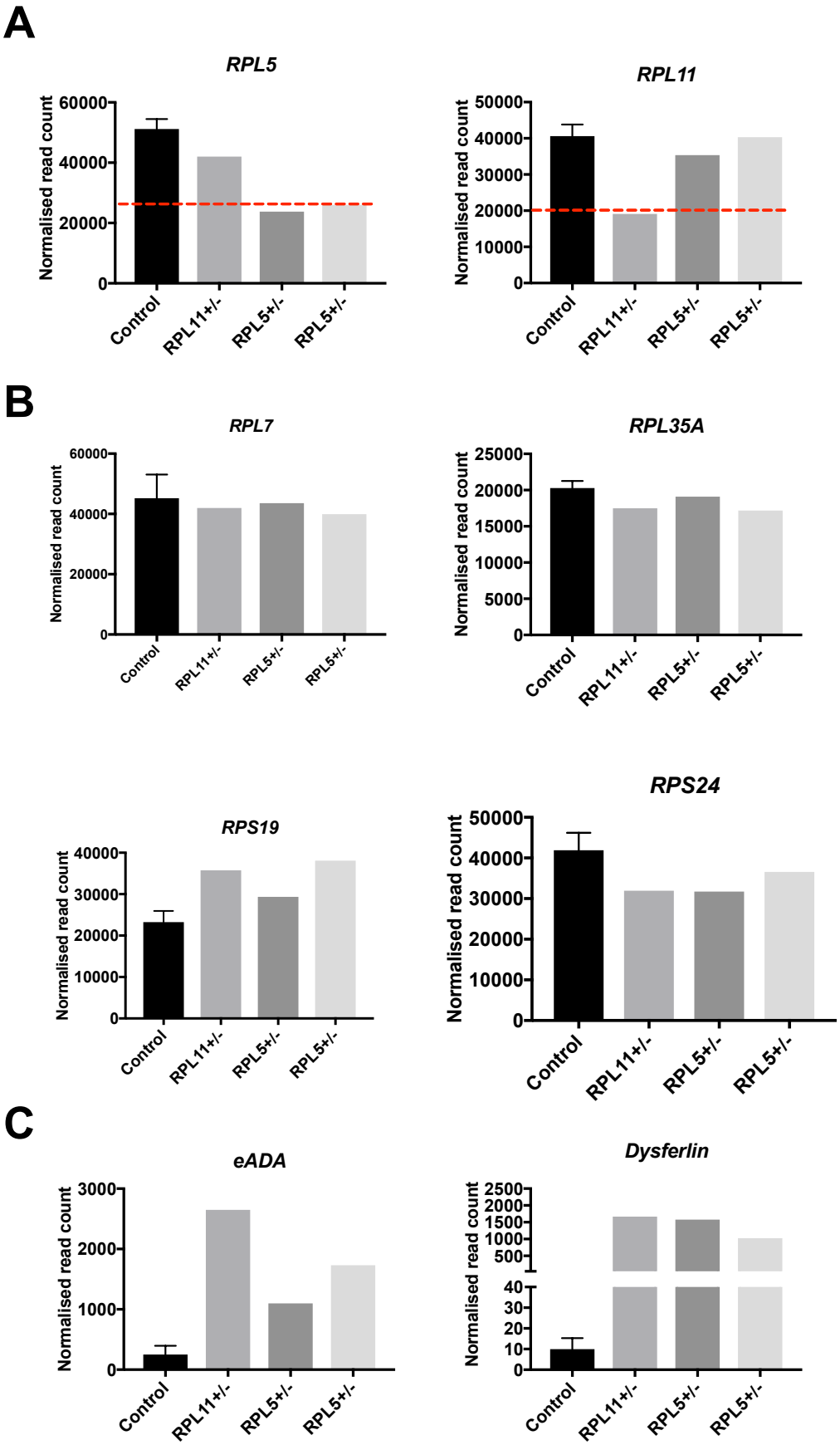

A
RPL5
Normalised read count
60000
40000
20000
0
Control
RPL11+/-
RPL5+/-
RPL5+/-
RPL11
Normalised read count
50000
40000
30000
20000
10000
0
Control
RPL11+/-
RPL5+/-
RPL5+/-
B
RPL7
Normalised read count
60000
40000
20000
0
Control
RPL11+/-
RPL5+/-
RPL5+/-
RPL35A
Normalised read count
25000
20000
15000
10000
5000
0
Control
RPL11+/-
RPL5+/-
RPL5+/-
RPS19
Normalised read count
40000
30000
20000
10000
0
Control
RPL11+/-
RPL5+/-
RPL5+/-
RPS24
Normalised read count
50000
40000
30000
20000
10000
0
Control
RPL11+/-
RPL5+/-
RPL5+/-
C
eADA
Normalised read count
3000
2000
1000
0
Control
RPL11+/-
RPL5+/-
RPL5+/-
Dysferlin
Normalised read count
2500
2000
1500
1000
500
40
30
20
10
0
Control
RPL11+/-
RPL5+/-
RPL5+/-

Figure 5:11 Validation of RNA-seq by comparison of normalised read counts for genes of interest A) Confirmation of reduced gene expression of mutated RP gene in each patient with DBA. Red dashed line highlights 50% gene expression relative to CON. **B)** Selectivity of 50% reduction in mutated gene confirmed by examining read counts of other RPL and RPS genes **C)** eADA and dysferlin upregulated in all DBA samples, relative to CON. Elevated peripheral blood eADA is a specific marker for DBA and dysferlin has recently been shown to be significantly upregulated on the red cell membrane in patients with DBA (Pesciotta et al. 2014).

Table 5.2 Expression of globin genes in DBA EB, determined by RNA seq

Gene ID	Log2 FC	Padj
HBA1	-0.206626196	0.748062079
HBA2	0.12637226	0.938796224
HBB	0.26838008	0.621714768
HBG1	1.006332759	0.209065135
HBG2*	1.44100182	0.036074823

** HBG2 (Haemoglobin Subunit Gamma 2), which encodes for Yglobin chains, is significantly upregulated in DBA (n=3) relative to CON EB (n=3).*

5.3.3.2 p53-mediated apoptosis/cell cycle arrest

To identify gene networks involved in DBA erythroid failure, the gene expression data were analysed using Ingenuity Pathway Analysis software (Ingenuity® Systems, www.ingenuity.com). p53-mediated cell death is a well-established mechanism of E failure in DBA, although there are very limited data from primary BM cells (Dutt et al. 2011). Ingenuity Pathway Analysis (IPA) analysis highlighted p53 protein as one of the top upstream regulators that could account for the differential gene expression between CON and DBA EB (e.g. increased *p21/CDKN1* and *BAX* expression). Conversely DEX was identified as the most significant upstream suppressor of the transcriptomic features in DBA cells, consistent with its therapeutic effect in DBA.

Top Upstream Regulators		
Upstream Regulator	p-value of overlap	Predicted Activation
TNF	2.97E-32	Activated
IFNA2	5.11E-29	Activated
IFNG	1.21E-25	Activated
TP53	2.03E-25	Activated
dexamethasone	3.27E-23	

Figure 5:12 Top five upstream regulators accounting for differential gene expression between CON and DBA TD samples (n=3)**.** P53 and inflammatory cytokines were identified as the top 'activators' of the DBA transcriptome while dexamethasone was identified as the top 'repressor'.

5.3.3.3 Pro-inflammatory signature in DBA EB

As highlighted by the IPA analysis (Figure 5:12), inflammatory pathways mediated by IFN and TNF were highly upregulated in DBA EB, as were MHC molecules involved in the immunoproteasome. Given the proximity between EB and immune cells [specifically, CD163+ macrophages (Falchi et al. 2015; Hom et al. 2015)] during erythroid development, I hypothesised that EB/macrophage interactions may be involved in the observed inflammatory phenotype observed in DBA. While the major adhesion molecules required for these interactions were not differentially expressed in DBA, growth-arrest specific protein- 6 (GAS6) was significantly reduced (Log2FC -1.4, adj P=0.02). This molecule, secreted by EB in response to EPO, has been shown in mice to enhance EPO signalling, and to bind to macrophages, preventing the release of molecules such as IL-10 and TNFα that inhibit erythropoiesis (Angelillo-Scherrer et al. 2008). Moreover, in a transgenic mouse model of chronic anaemia secondary to EPO deficiency, GAS6 synergised with EPO to correct anaemia (Angelillo-Scherrer et al. 2008). Reduced GAS6 secretion from DBA EB may plausibly be involved in erythroid failure in DBA. To account for other cellular sources of GAS6 other than EB , I investigated the concentration of this secreted molecule in whole BM plasma from patients with DBA and normal controls (Figure 5:13). This was suggestive of reduced GAS6 in the DBA BM microenvironment but, given the variation in normal controls, more samples are needed to confirm or refute this.

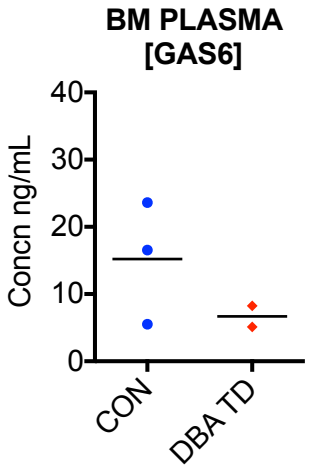


Figure 5:13 GAS6 concentration in BM plasma. BM plasma was extracted from normal CON (n=3) and TD patients with DBA (n=2) and the concentration of GAS6, a secreted molecule, was determined by ELISA. There is a suggestion of reduced [GAS6] in DBA but further replicates are required.

5.3.3.4 Normal expression of GATA-1 protein and its target genes in DBA EB

There are currently conflicting data on whether reduced expression of GATA-1 as a result of its defective translation is at the heart of DBA pathogenesis. Data so far have been generated from RP knockdown cellular models (Ludwig et al. 2014), EB generated *in vitro* from DBA PBMCs (O'Brien et al. 2017; Gastou et al. 2017) and CD34+CD71+CD45RA- EP purified from DBA BM samples from 3 *RPS19*-mutated patients in remission (Gazda et al. 2006; Ludwig et al. 2014). I therefore interrogated my transcriptome dataset for expression of *GATA-1* itself and known GATA-1 target genes (Ludwig et al. 2014; Ferreira et al. 2005). As shown (Table 5.3) the majority of GATA-1 regulated genes were not significantly altered in DBA EB, with the exception of *GATA-2* (which was activated) and *EPOR* (which was repressed). While these data challenge previous observations that GATA-1 target genes are repressed in DBA erythroid cells due to defective GATA-1 translation (Ludwig et al. 2014), it is important to note that these analyses are fraught with limitations. The expression

levels of genes regulated by GATA-1 do not provide any direct information about GATA-1 protein expression nor its capacity for chromatin binding. Many of these genes are co-regulated by other TF.

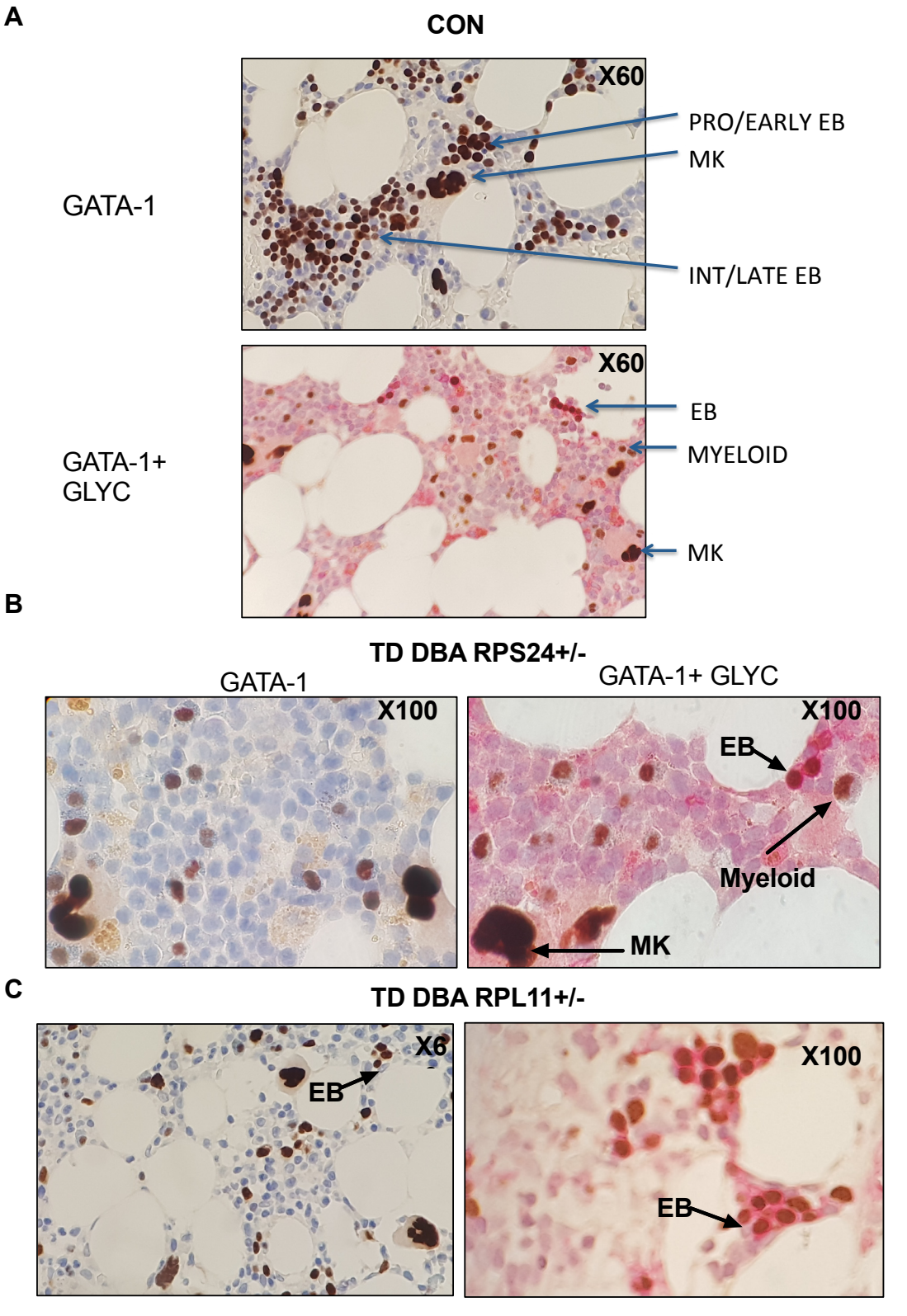
Table 5.3 Expression of GATA-1 target genes

Gene	Log2FC	Padj
GATA-1	-0.064840727	0.907374592
GATA-2	3.80067232	6.04E-61
ALA-S1	-0.350485377	0.245006041
ALA-D	-0.36347138	0.285305025
PBG-D/HMBS	-0.047286208	0.940454534
ABCme/Abcb10	-0.077225284	0.898484195
EPO	-0.248307488	NA
EPOR	-0.53865227	0.028311337
BCLxL/BCL2l-1	-0.657743548	0.009239249
MaFK	0.819428863	0.050764797
Myc	0.011510308	0.993597188
ALAS2	0.773088052	0.004475646

RNA-seq was used to measure gene expression in purified DBA EB (n=3) relative to CON (n=3). Log2 FC and adjusted P value are shown for each gene. Significantly differentially expressed genes are highlighted in red.

To validate my finding that GATA-1 is stably expressed in stage- matched DBA EB, using a more direct measurement, I assessed its expression directly by immunohistochemistry in paraffin-embedded BM trephine sections from patients with TD DBA, compared with individuals with unrelated haematological disorders in remission (n=2) or DBA in remission post-HSCT (n=1). Under normal conditions GATA-1 is expressed in MK, EB and occasional myeloid cells (eosinophils & mast cells; Figure 5:14A). Expression is weaker in myeloid compared with E/MK cells, and

decreases in late EB relative to early/intermediate EB. To verify GATA-1 staining in EB specifically, I used a co-stain for Glycophorin C (GYPC) which is expressed on E lineage cells from ProEB to RBC (Figure 5:14A). In TD DBA caused by RPS gene pathogenic variants, there were very few (50-100) EB available for assessment, however those present exhibited dark brown staining. This was the same level of intensity as neighbouring MK (Figure 5:14B), and the same as in control samples (Figure 5:14D). Although RPS BM showed a predominance of cells staining weakly for GATA-1, examination of their nuclear morphology coupled with the GATA-1/GYPC co-stain confirmed that these were indeed myeloid, likely eosinophils. In TD DBA caused by RPL gene pathogenic variants, EB and MK GATA-1 staining were also equivalent to normal controls (Figure 5:14D). In summary, neither transcriptome profiling nor immunohistochemical staining of primary erythroid precursors from patients with DBA provided evidence for erythroid failure resulting from reduced GATA-1 protein.


A
CON
X60
PRO/EARLY EB
MK
INT/LATE EB
GATA-1
X60
EB
MYELOID
MK
GATA-1+
GLYC
B
TD DBA RPS24+/-
GATA-1
GATA-1+ GLYC
X100
X100
EB
Myeloid
MK
C
TD DBA RPL11+/-
X6
EB
X100
EB

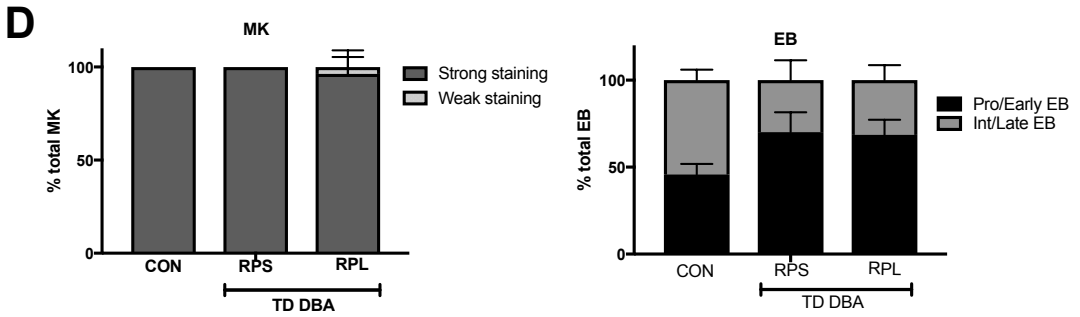


Figure 5:14 Normal GATA-1 expression in DBA BM. Formalin fixed paraffin embedded (FFPE) BM trephine sections were stained by immunohistochemistry for GATA-1 protein alone, or GATA-1 plus GLYC. **A)** Dark brown homogeneous nuclear GATA-1 staining is seen in pro/early EB. Light brown speckled staining is seen in int/late E and myeloid precursors. GLYC staining forms a red perinuclear rim in EB, whereas MK and myeloid cells are negative for GLYC. **B)** Representative images of staining in RPS-genotype BM. Very few EB observed but these exhibit normal GATA-1 staining. Most of GATA-1 cells are GLYC-negative myeloid cells. **C)** Representative images of staining in RPL-genotype patients, showing normal GATA-1 expression in EB. **D)** Equal GATA-1 staining intensity in MK and EB of CON (n=3), TD RPS (n=5) and TD RPL (n=6) DBA BM. In the bottom right panel, a distinction is made between early and late EB to show that DBA patients did not have an increase in late EB, which, under normal circumstances, stain only weakly for GATA-1.

5.3.4 Mechanisms of CS therapy in DBA

5.3.4.1 CS restore the frequency and function of DBA EP ex vivo

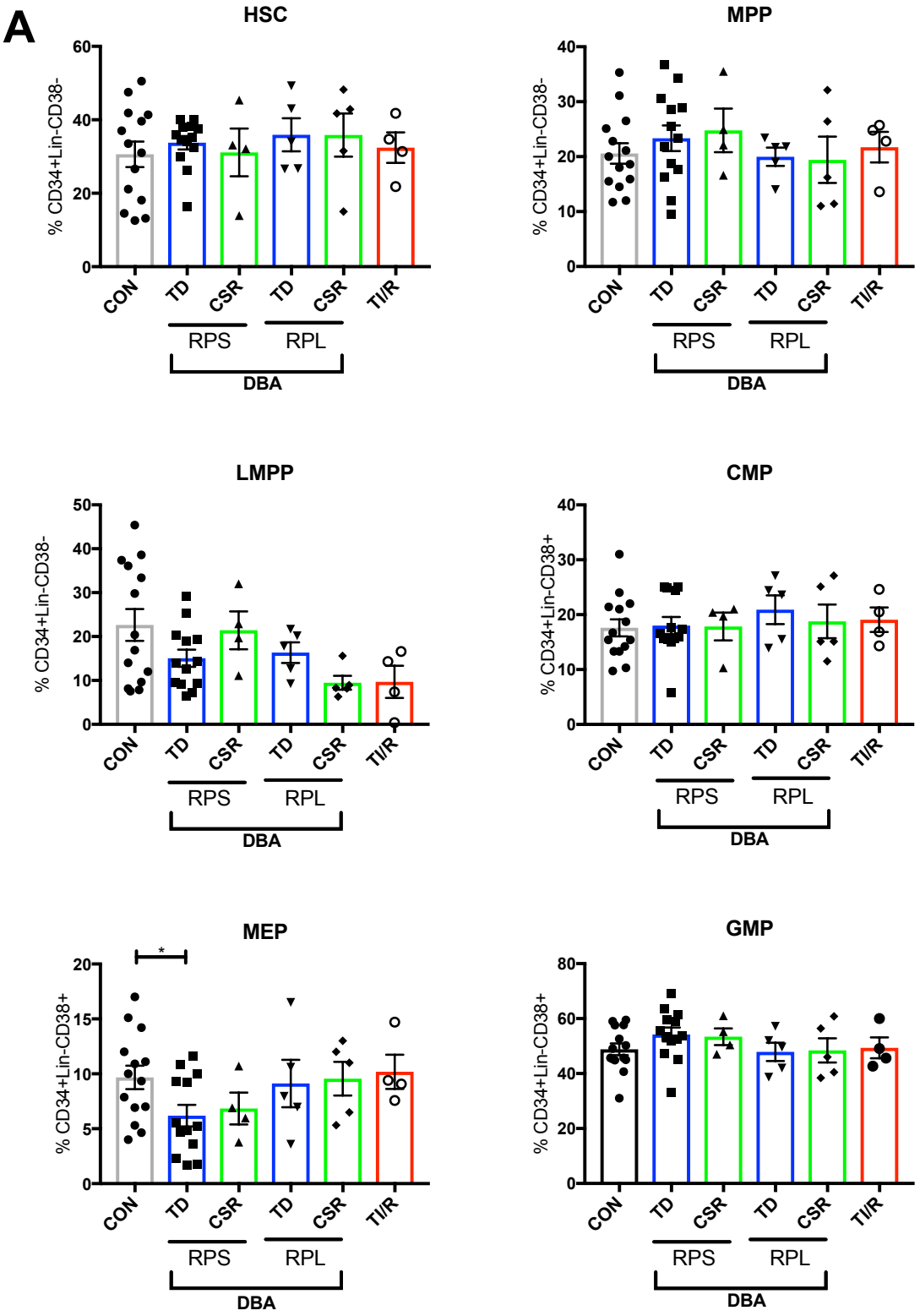
Having presented a detailed study of erythropoiesis in TD DBA patients, the aim of my work in this section was to investigate the cellular and molecular effects of CS in CSR patients with DBA. First I investigated the frequencies of HSPC and EP populations in DBA BM from patients treated with CS at the time of BM sampling. These data show that CD71+ MPP are significantly increased in CSR compared with

TD patents with the RPS genotype. Although the frequency of total MEP is not increased (Figure 5:15A), the EEP subpopulation of MEP is increased in CSR relative to TD RPS DBA (Figure 5:15B). CS do not appear to significantly alter the EP frequencies in DBA patients with the RPL genotype; this will be discussed further in Section 5.4.

After showing that EP are quantitatively improved in some CSR patients, I next assessed their function, using single cell clonogenic culture assays (Figure 5:16). These data suggest that the therapeutic effect of GC in DBA, is at least in part, attributable to enhancement of EEP function, as reflected by increased clonogenicity and a higher number of normal BFU-E as opposed to erythroid clusters. The effect on LEP clonogenicity was variable but sample numbers are small.

In line with this improvement in EEP frequency and function, CSR patients with DBA display a significant increase in the number of EB, as determined by both flow cytometry and BM morphology (Figure 5:17). Importantly there was no evidence for increased GATA-1 protein expression in the BM of CSR patients with DBA, as shown by equal intensity of GATA-1 immunohistochemical staining (Figure 5:18).

In one patient, I also confirmed by RNA-seq that CS do not ameliorate the erythroid defect by increasing mRNA expression of the mutated RP gene in cells of the E lineage (Figure 5:19). While *HbG2* expression is only mildly raised in this patient (consistent with the fact that her HbF level is normal), eADA remained elevated (Table 5.1) despite successful treatment with CS therapy, in line with clinical observations that these parameters do not revert on CS (Vlachos & E. Muir 2010).


A
HSC
% CD34+Lin-CD38-
60
40
20
0
CON
TD
CSR
TD
CSR
TI/R
RPS
RPL
DBA
MPP
% CD34+Lin-CD38-
40
30
20
10
0
CON
TD
CSR
TD
CSR
TI/R
RPS
RPL
DBA
LMPP
% CD34+Lin-CD38-
50
40
30
20
10
0
CON
TD
CSR
TD
CSR
TI/R
RPS
RPL
DBA
CMP
% CD34+Lin-CD38+
40
30
20
10
0
CON
TD
CSR
TD
CSR
TI/R
RPS
RPL
DBA
MEP
% CD34+Lin-CD38+
*
20
15
10
5
0
CON
TD
CSR
TD
CSR
TI/R
RPS
RPL
DBA
GMP
% CD34+Lin-CD38+
80
60
40
20
0
CON
TD
CSR
TD
CSR
TI/R
RPS
RPL
DBA

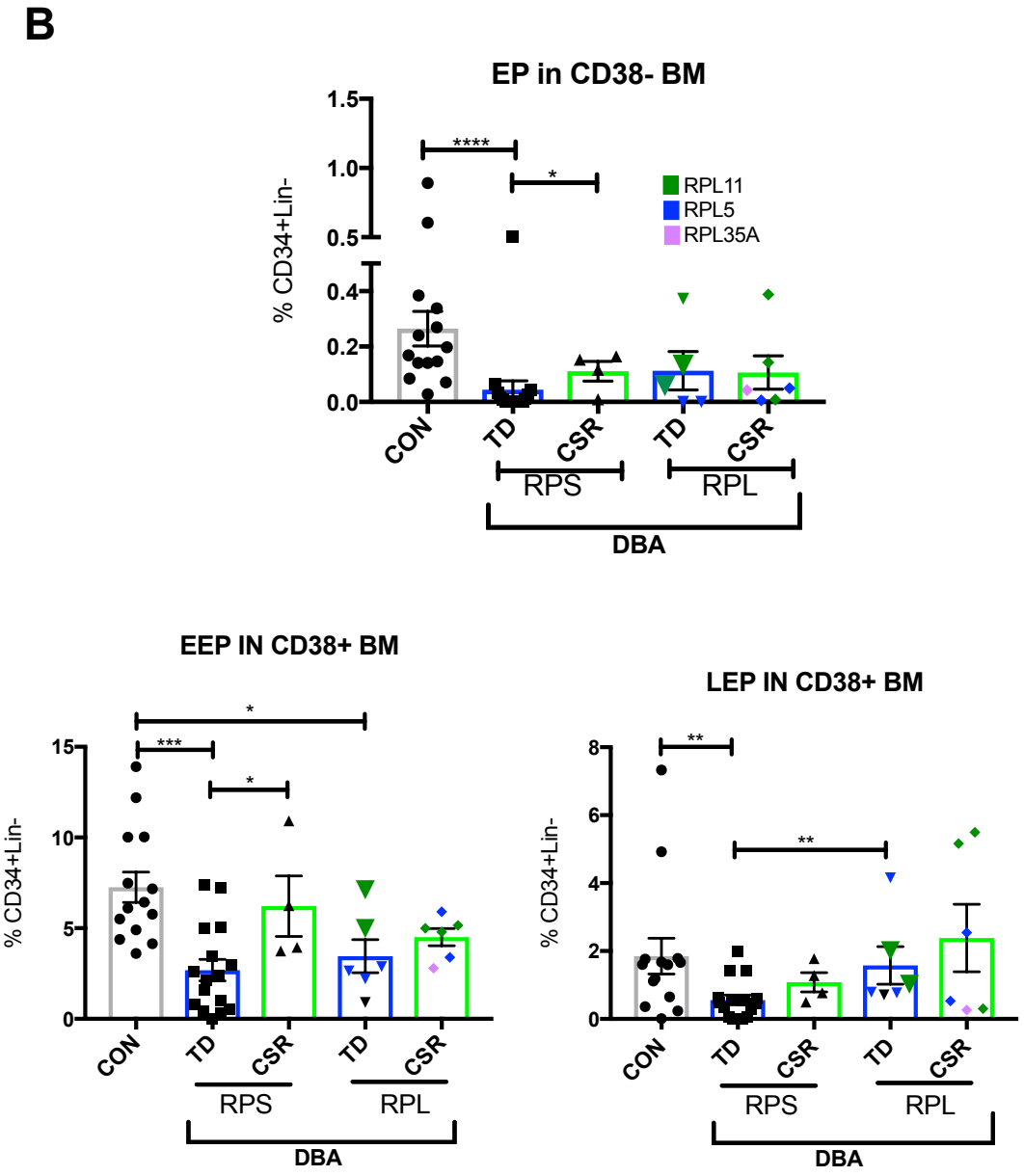


Figure 5:15 Impact of CS on HSPC and EP frequencies in patients with DBA. A) Frequencies of CMP/ GMP/ MEP progenitors, expressed as % of CD34+Lin-CD38+ cells and of HSC/MPP/LMPP progenitors, expressed as % of CD34+Lin-CD38- cells were measured by flow cytometry in BM of CON (n=14), *RPS* TD (n=13), *RPS* CSR (n=4), *RPL* TD (n=5), *RPL* CSR (n=5) and TI/R DBA (n=4). HSPC populations are preserved in CSR patients **B)** EP frequencies, measured by flow cytometry, in the CD38- and CD38+ subfractions of BM of CON (n=14), *RPS* TD (n=16), *RPS* CSR (n=4),

RPL TD (n=6) and *RPL* CSR (n=6). EP in CD38- BM and EEP are increased in CSR patients with RPS-gene associated DBA.

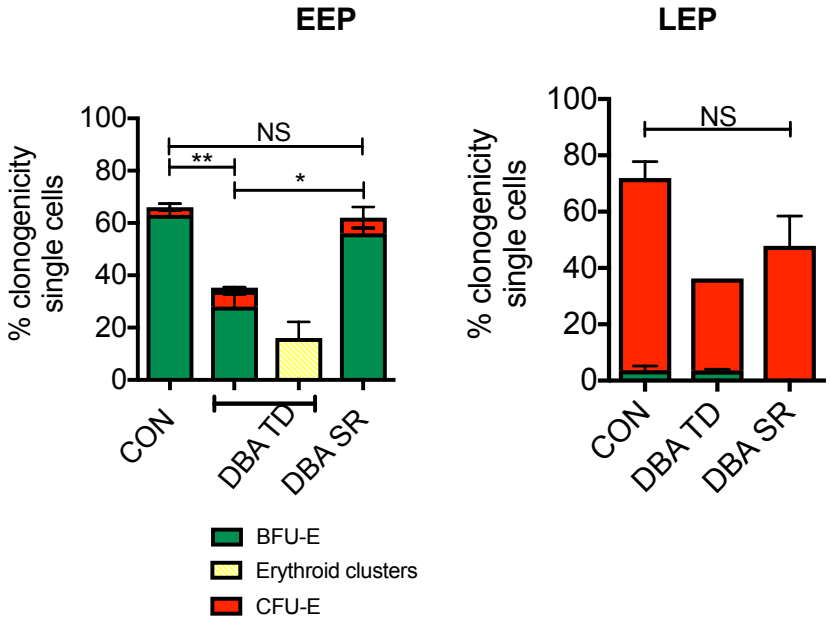


Figure 5:16 CS restore EEP function in DBA. Clonogenicity of TD DBA (n=4 EEP, n=2 LEP), SR DBA (n=3 EEP and LEP) and CON EP (n=6 EEP, n=5 LEP), flow-sorted as single cells.

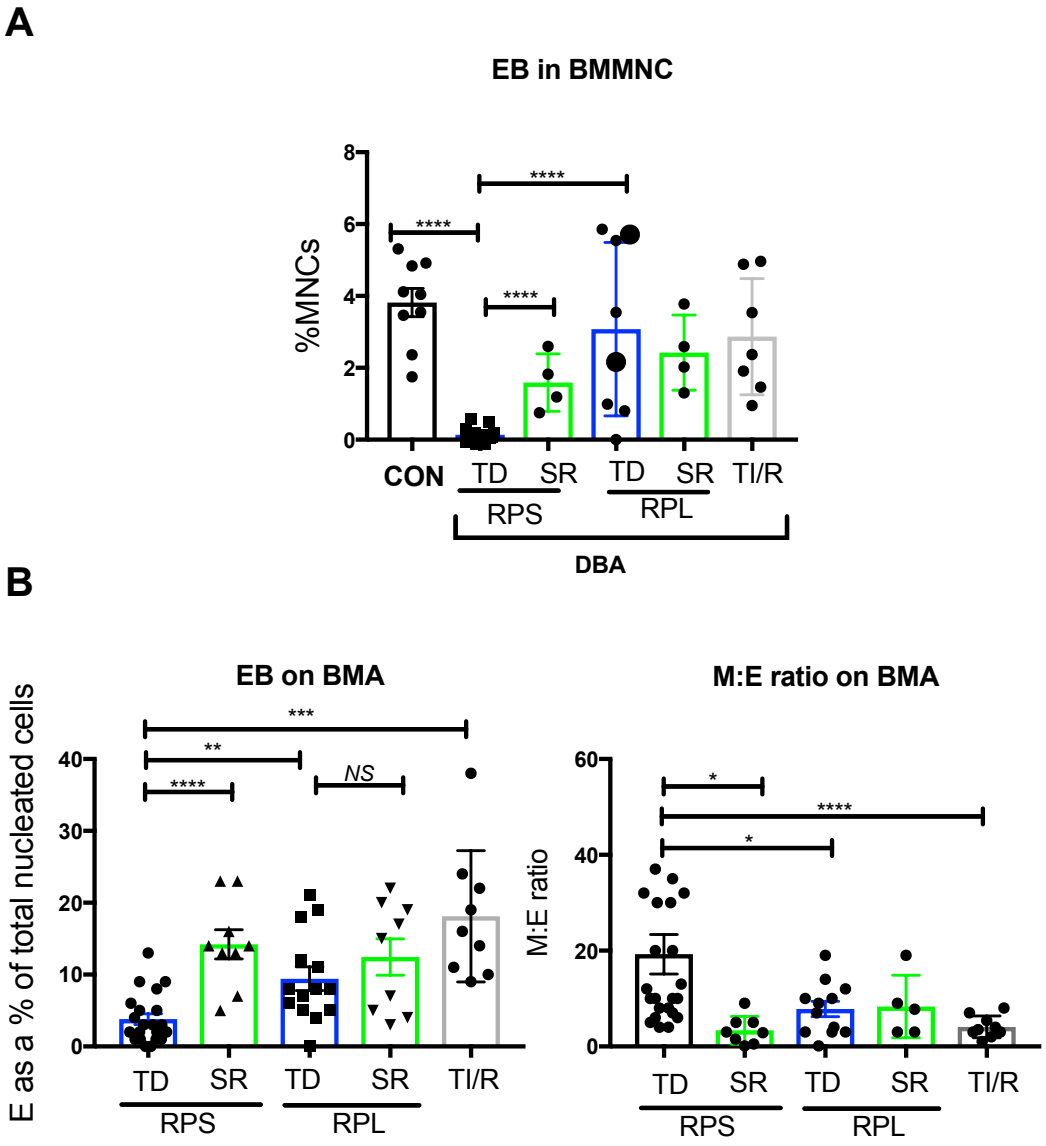


Figure 5:17 Increased EB frequencies in CSR DBA, as measured by **A)** flow cytometry in BM of CON (n=9), *RPS* TD (n=15), *RPS* CSR (n=4), *RPL* TD (n=8), *RPL* CSR (n=4) and TI/R DBA (n=7) and **B)** enumeration on BMA of *RPS* TD (n=21), *RPS* CSR (n=9), *RPL* TD (n=14), *RPL* CSR (n=9) and TI/R DBA (n=9).

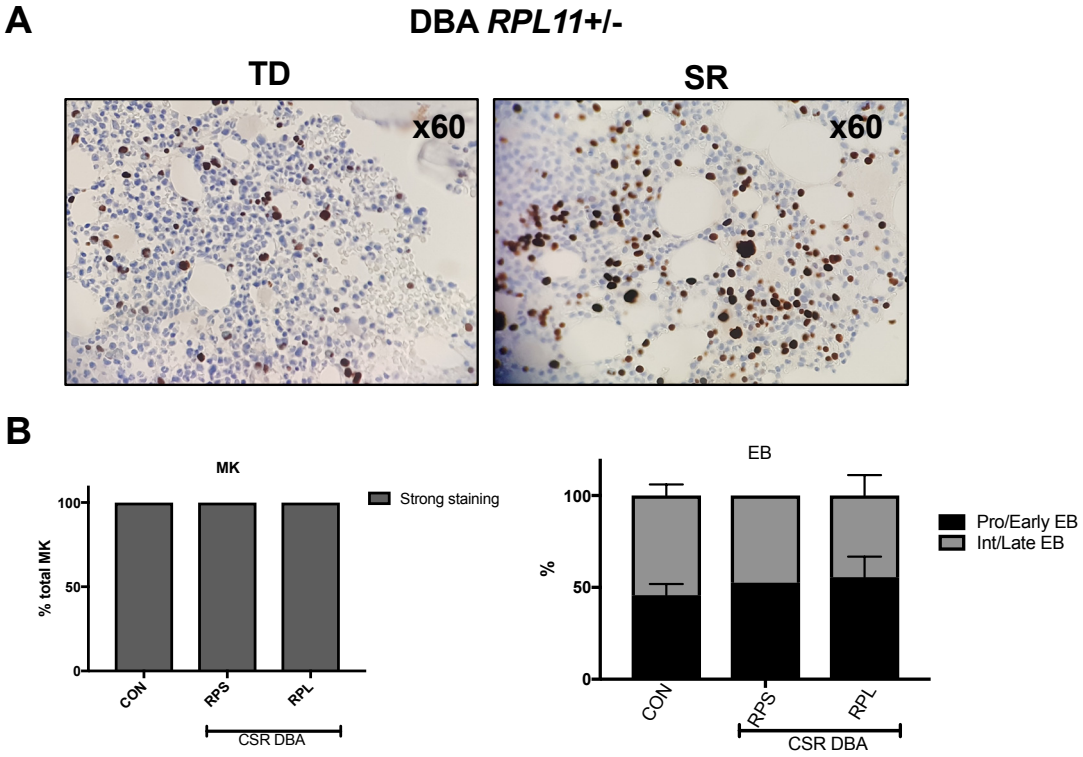


Figure 5:18 Normal expression of GATA-1 in CSR patients with DBA. A) BM sections pre – and post- CS in the same patient show an increase in EB frequency but no change in the intensity of GATA-1 immunohistochemical staining **B)** Normal GATA-1 staining intensity in MK and EB of CON (n=3), CSR RPS (n=2) and CSR RPL (n=4) DBA BM.

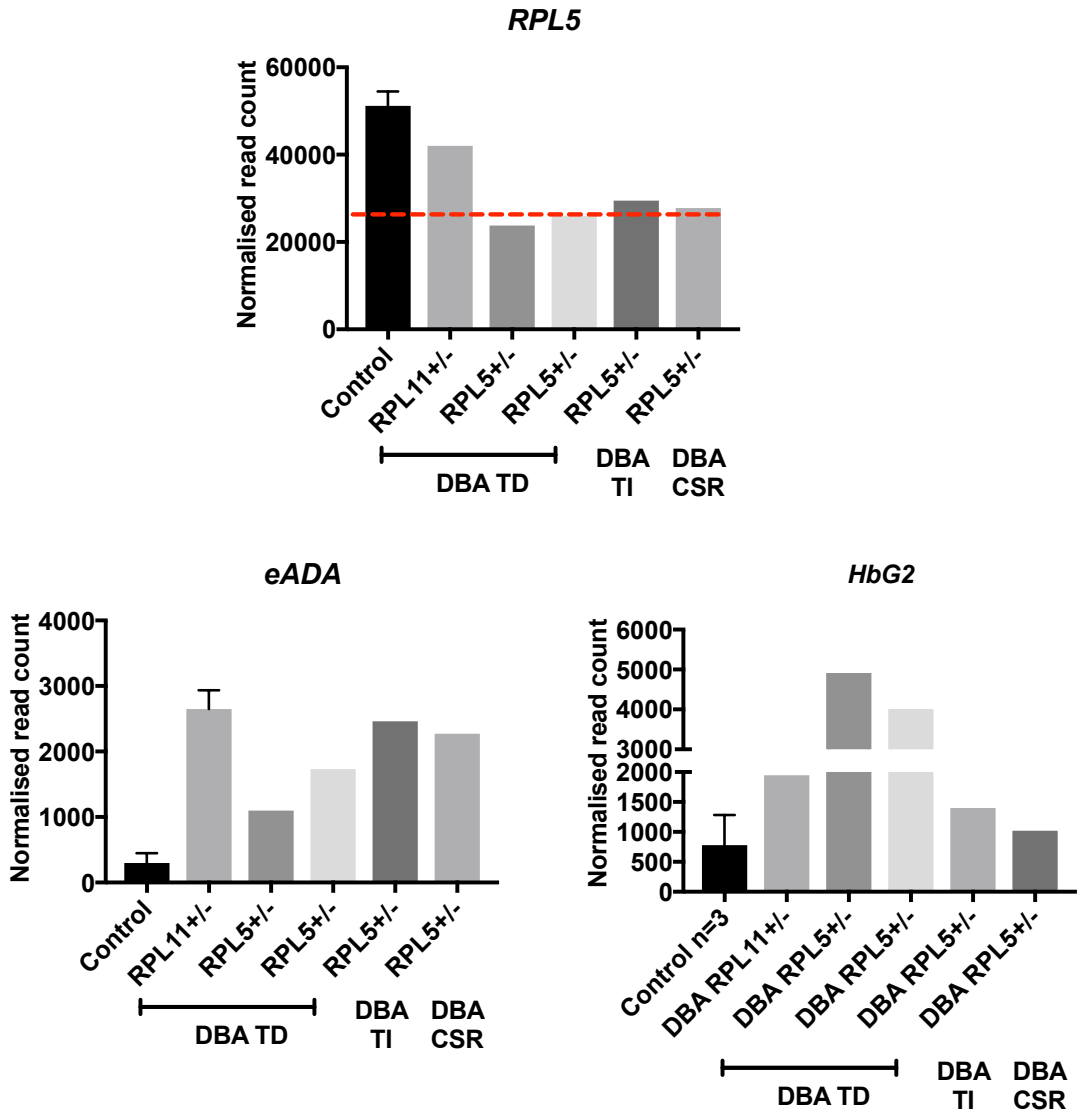
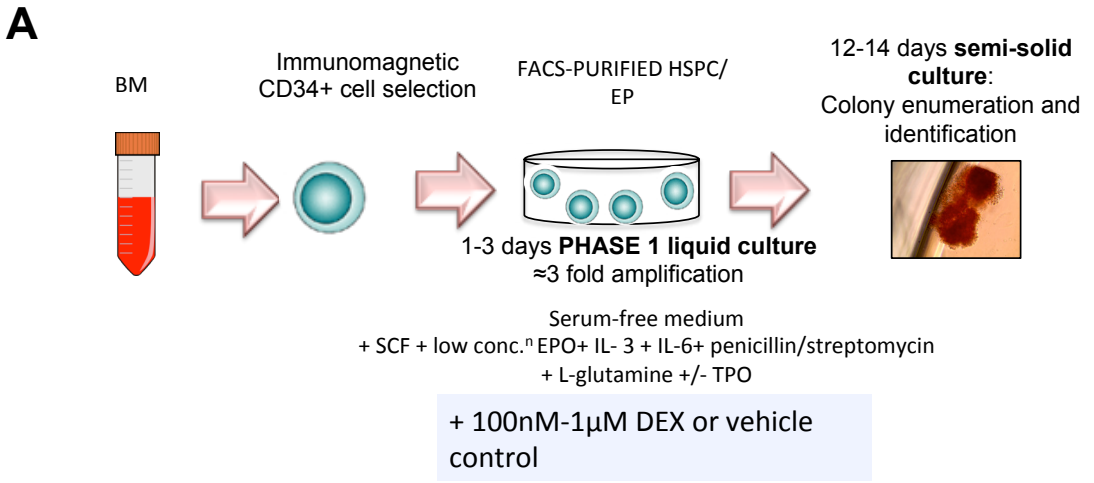
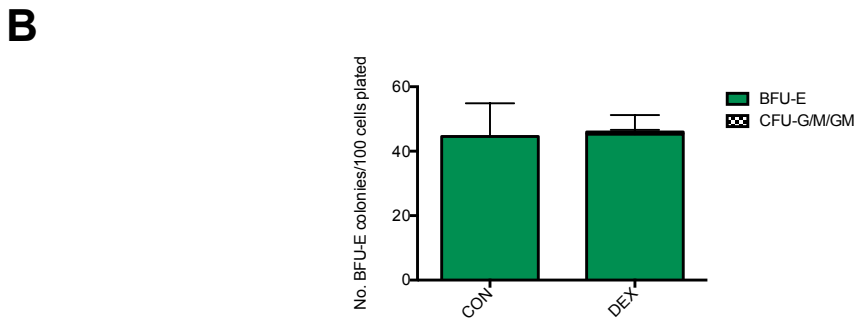


Figure 5:19 Gene expression by RNA-seq in FACS-purified EB from a patient with DBA, pre- and post CS. The patient has a mutation in *RPL5* and was TI prior to the commencement of CS. Gene expression is represented by normalised read counts generated by RNA-seq. Average values were calculated for technical replicates. Red dashed line highlights 50% gene expression relative to CON. *RPL5* remains reduced and *eADA* elevated after CS treatment. *HBG2* expression is also little changed by CS treatment.

5.3.4.2 GC delay normal erythroid differentiation in vitro

My next step was to explore possible underlying mechanisms for the above observation that CS increase EP frequency and function *in vivo*. Given the scarcity of BM samples from CS-dependent patients, and the limited data on GC in normal human erythropoiesis, I first tested the effect of the synthetic GC, DEX on normal BM cells cultured *in vitro.* In these experiments a dose of either 100nM or 1μM was used (as indicated), as both concentrations have been reported to enhance erythropoiesis *in vitro* (Flygare et al. 2011; Ebert 2005). First, 0.5-1x10^2 EEP were FACS-purified from normal BM and cultured (Protocol A with IL-6, lipids and L-glutamine; Table 2.1) for 1-3d in the presence of 1μM DEX or vehicle control. Cells were then washed and plated in methylcellulose (Figure 5:20A). Haematopoietic colonies were scored on d12-14. In 2 independent experiments, I found that EEP proliferation in liquid culture (measured by cell number) was not affected by DEX treatment for 72 hours (D0-D3 fold increase in 2 experiments of of 2.5 and 5 in CON vs 2.8 and 5 in DEX-containing media). Additionally, the number of BFU-E colonies generated by DEX-treated EEP was equivalent to CON (Figure 5:20B). The only observed difference was that prior treatment with DEX appeared to delay EB differentiation during colony formation, as reflected by larger, less compact, poor haemoglobinsed colonies (Figure 5:20C), even after only 24h prior DEX exposure.





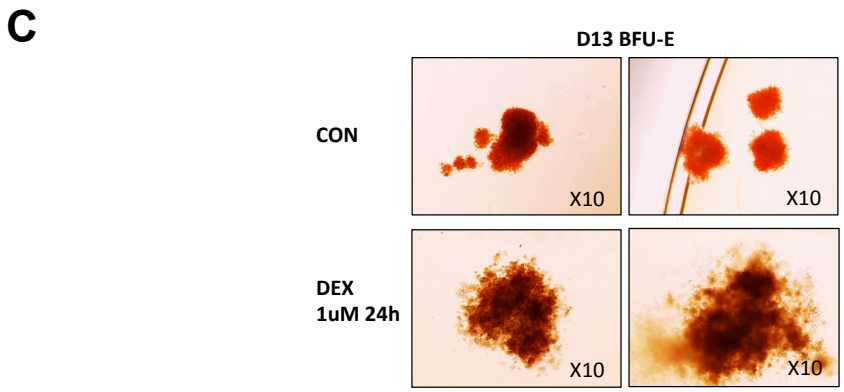


Figure 5:20 EEP exposure to GC delays their differentiation *in vitro*. A) CD34+ cells were positively selected from BMMNC and cryopreserved. On thawing, cells were stained with a cocktail of antibodies to allow FACS-purification of 50-100 EEP into liquid culture in duplicate. EEP were cultured in serum-free medium supplemented with cytokines to support erythroid cell growth. 1µM DEX or vehicle control was added. After 24 hours cells were washed and replated in methylcellulose supportive of myeloid or erythroid development**. B)** Number of haematopoietic colonies scored

on D12-14. Both CON- and DEX- treated EEP gave rise to equivalent numbers of BFU-E. Data from 3 independent experiments are shown. **C)** Examples of colony morphology on D14 semi-solid culture of FACS-purified EEP, pre-exposed to CON or DEX in 3 days liquid culture. EEP exposed to DEX generated larger, less compact and darker colonies that did not fully haemoglobinise, even after 21 days.

As the number of EEP is limiting for detailed cellular assays, next I assessed this apparent DEX-induced differentiation delay using a previously described 2-phase *in vitro* liquid culture system for erythroid cell generation from UCB- derived CD34+ cells (Freyssinier et al. 1999). Figure 5:21 exemplifies the cell differentiation pattern and morphology observed at different time points of the culture. Importantly on D5, cells are CD34+CD36+ i.e., equivalent to late committed EP/early EB.

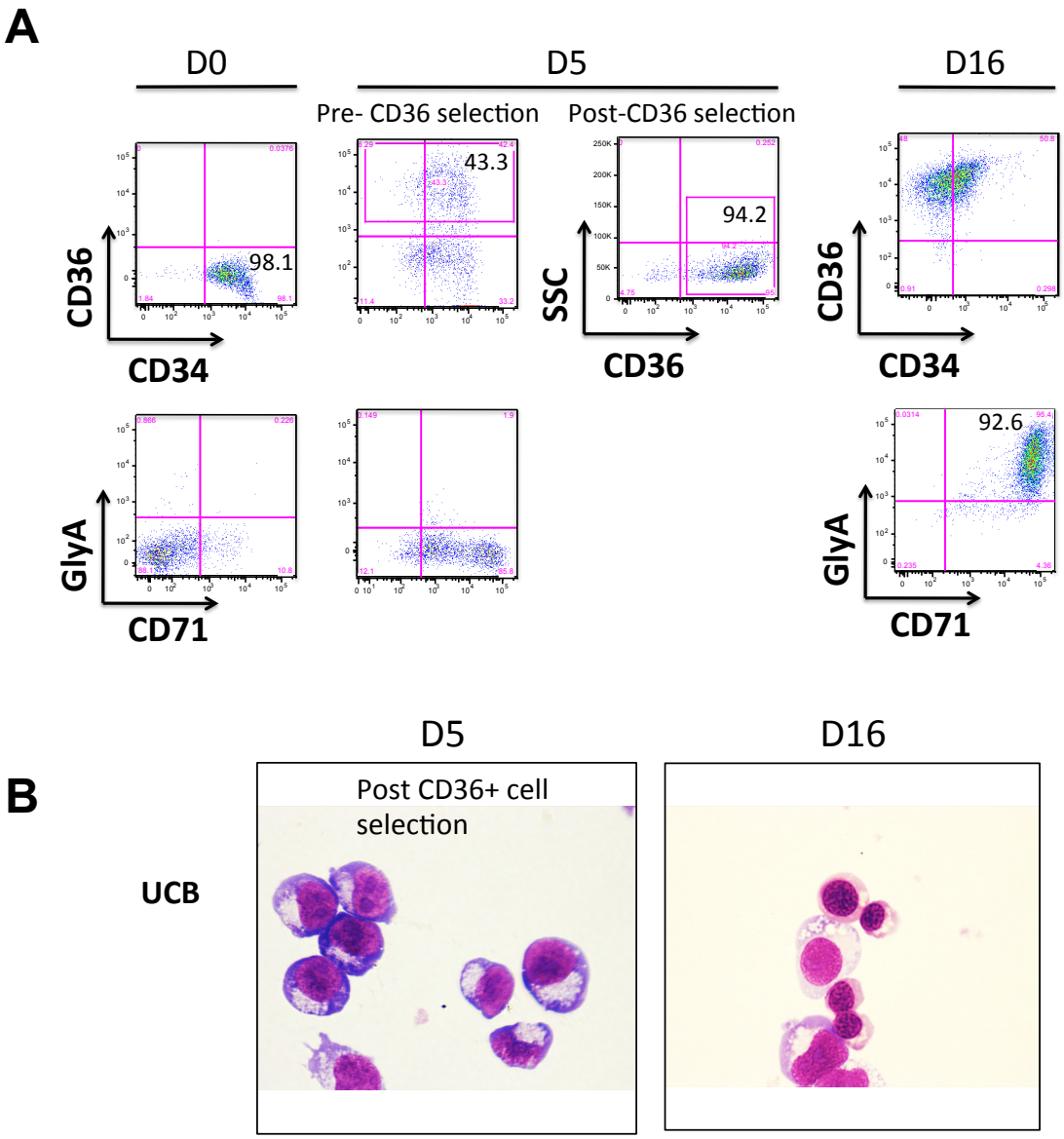


Figure 5:21 Two-phase system for generation of highly enriched erythroid cells. Purified UCB- derived CD34+ cells were first cultured for 5 days in serum-free medium containing SCF (100ng/ml), IL-3 (10ng/ml) and a low concentration of EPO (0.5u/ml), to maximise survival of erythroid-primed progenitors. Data are representative of 7 experiments using UCB **A)** On day 5, ~50% of cells express CD36, >80% are CD71+, and <10% are GlyA+. In the experiment shown, $2x10^4$ CD34+ cells plated on day 0 increased to 1.6 $x10^5$ cells on day 5. CD36+ cells are selected by immunomagnetic beads and re-established in phase 2 of liquid culture with SCF, IL-3 and 4U/ml EPO to enrich, amplify and mature large numbers of terminally

differentiated erythroid precursors on day 14-16, that are all CD36+CD71+GlyA+ and negative for myeloid (CD11b and CD14) and megakaryocytic (CD61) markers. The plots shown are gated on total live cells in the culture at each time point. Quadrants are determined using unstained controls **B)** MGG stained cytospins from normal human BM, showing that on day 5 CD36+ cells are a pure population of LEP/ProEB while on day 14, a mixture of early, intermediate, late erythroblasts and occasional reticulocytes are seen.

Treatment of these cells with 100nM DEX, from D5-D16 induced a block in differentiation with an increase (p=0.018) in the percentage of immature EBs (i.e., expressing CD71 but not GlyA) at d14-16, compared with normal controls where a pure population of mature EBs (expressing both CD71 and GlyA) was observed by this stage (Figure 5:22A&B). The underlying mechanism for this differentiation delay was shown on day 14 to be cell cycle arrest in phase $G_{0/1}$ (Figure 5:22C). This delay in differentiation is in keeping with previous reports (M von Lindern et al. 2001) (Dolznig et al. 2006) (Sjögren et al. 2015) that DEX stimulates erythropoiesis by favouring self-renewal rather than differentiation of erythroid cells.

However, contrary to published data from other groups (M von Lindern et al. 1999; Ebert 2005), and in line with my EEP experiments (Figure 5:20), despite inducing a block in differentiation, DEX-treatment did not induce a higher overall number of erythroid cells in culture and it was not possible to recapitulate *in vitro* the ability of immature EB to undergo final maturation in the presence of DEX *in vivo*. Importantly, previous groups had exposed total CD34+ cells or total BM to DEX to infer an effect on EP, rather than using purified human EP. I therefore explored whether in human BM GC may exert an effect on more immature populations than EEP, unlike in mice FL where a clear direct proliferative effect on FL BFU-E has been shown.

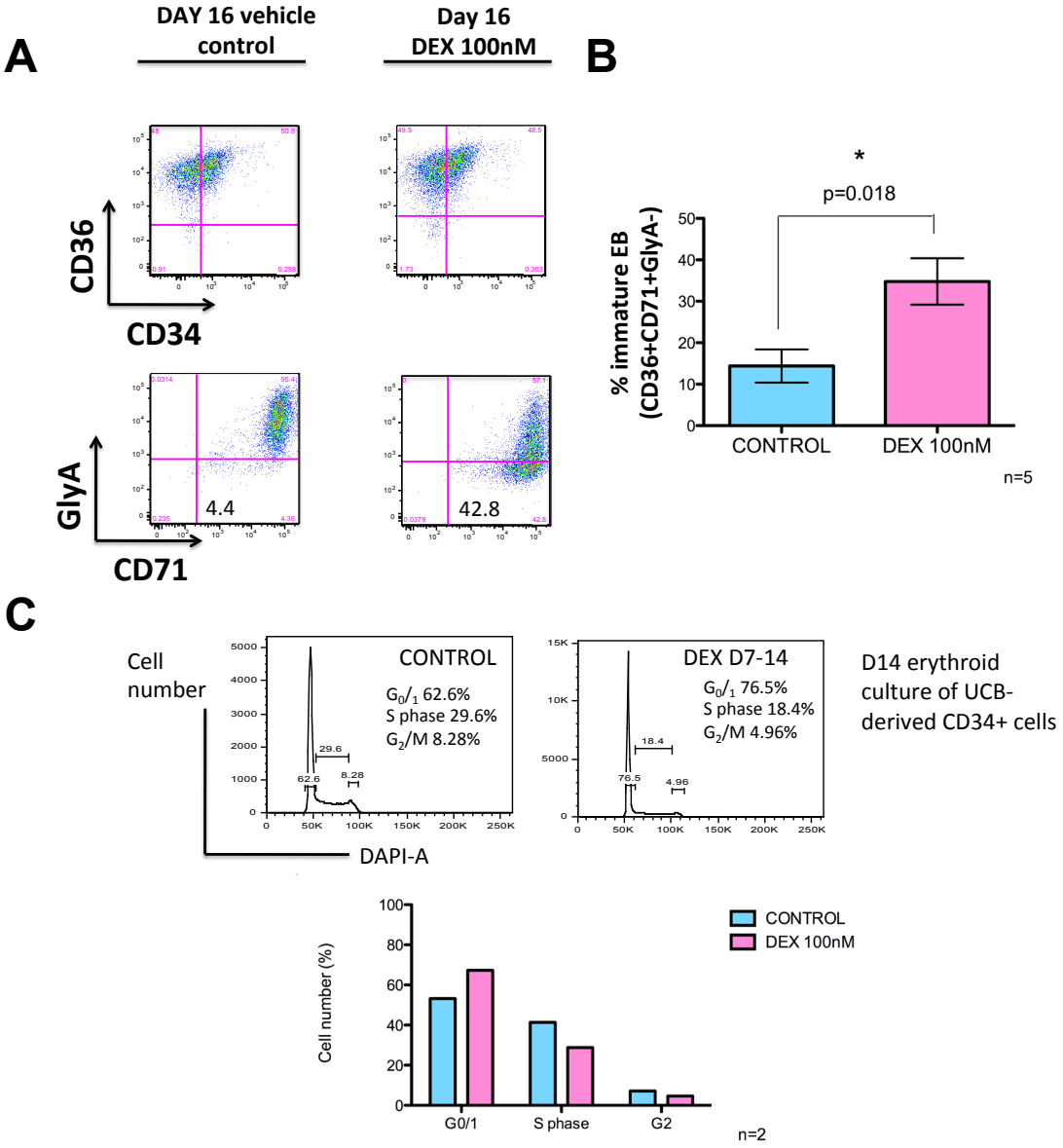


Figure 5:22 GC maintain the immature EB pool. Day 5 CD36+ EBs generated from UCB as outlined in **Figure 5:21** were cultured with 100nM DEX or vehicle till D14-16. At this point the total number of cells in culture was enumerated, erythroid differentiation was analysed by flow cytometry and cell cycle analysis was performed by fixation followed by DAPI staining. **A)** Dexamethasone induces a differentiation arrest in EBs at the CD34-/loCD36+CD71+GlyA- stage, while control cells are virtually all CD34-/loCD36+CD71+GlyA+ by day 16. Plots are representative of 5 experiments and are gated on the total live cells in the culture. Quadrants are determined using unstained controls. **B)** The increase in immature EBs in the presence of DEX in 5

independent experiments. By d14-16 the mean number of immature EBs was $5.8x10^5$ without DEX and $1.4x10^6$ with DEX (from a mean number of $7X10^4$ CD36+ cells on D5). **C)** DEX induces a cell cycle arrest of EBs in $G_{0/1}$, accompanied by a relatively lower proportion of cells in S and G_2/M phases, as shown in the histograms, which are representative of 2 experiments.

5.3.4.3 GC act upstream of MEP in normal BM to stimulate erythropoiesis in vitro

To test whether GC stimulate erythropoiesis by acting on haematopoietic cells in addition to EEP, I first performed an erythroid liquid culture of bulk BM-derived CD34+Lin- cells. Equal numbers of cells were treated with either 1μM DEX or control from D0 to D11 (Protocol A, with IL-6, Table 2.1). Cell enumeration, flow cytometric analysis and cytospins were performed on D11. DEX showed a trend towards an increased number of erythroid cells generated (Figure 5:23A) by increasing the relative proportion of CD71+ cells (Figure 5:23B). Consistent with the data above, I found that DEX also delayed the acquisition of GlyA (Figure 5:23B) hence the ratio of mature: immature E cells was lower in DEX-treated cultures (Figure 5:23C). Differentiation progressed by D17 after DEX removal by washing cells on D11, but had still not reached the same level as CON cells (Figure 5:23B). In conclusion, these data suggest that *in vitro* GC may act upstream of committed EEP to increase generation of CD71+ E cells but reduce the relative proportion of fully formed erythroid cells.

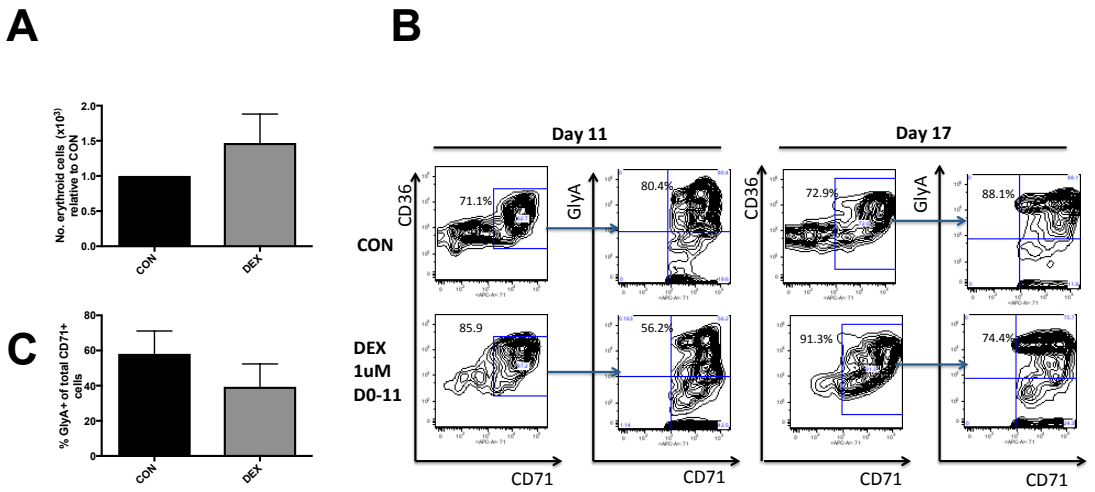
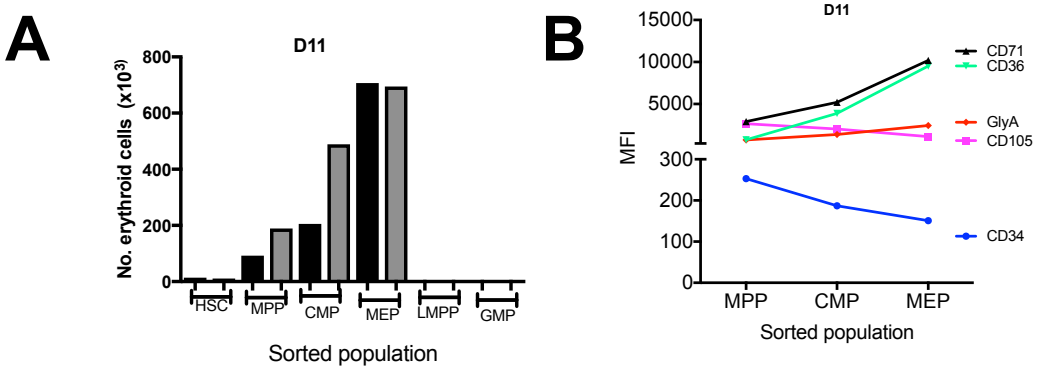
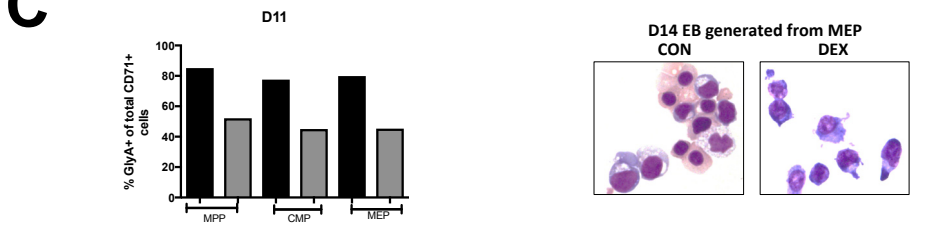


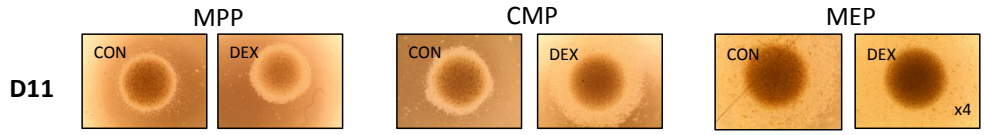
Figure 5:23 Effects of DEX on E culture of CD34+Lin- cells from normal BM A) Total number of E cells on D11 generated from normal BM CD34+Lin- cells in E culture (n=3). Cell number on D11 is normalised to that on D0 because different initial numbers of CD34+Lin- cells were FACS-purifed in the 3 different experiments. The difference in cell number is NS(P=0.2) **B)** Expression of CD71, CD36 and GlyA on D11 and D17 of CON and DEX-treated cells. DEX increased the proportion of CD71+ cells. By gating on CD71+, it can be observed that a lower proportion express GlyA. Removal of DEX on D11 partially overcame this differentiation block as more cells gained GlyA but differentiation lagged behind CON cells throughout the culture **C)** % of GlyA+ cells within CD71+ fraction, showing that DEX delays E maturation (n=3).

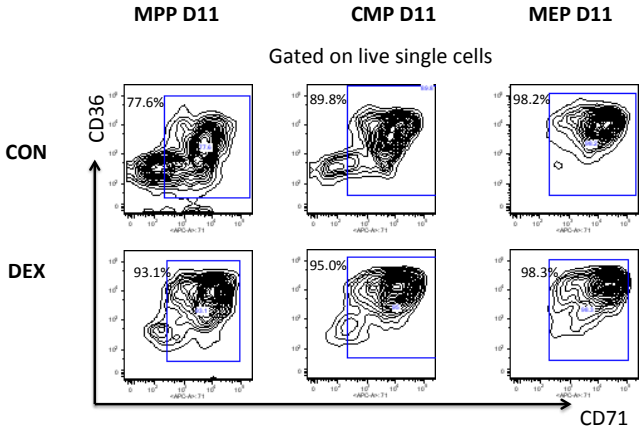
The effect of DEX on CD34+Lin- cells was modest, which may represent the fact that the CD34+Lin- fraction of BM comprises a heterogeneous population of HSPC. I therefore performed 1 experiment in which I assessed the impact of DEX on specific FACS-purified HSPC populations (as defined in Table 1.6). GMP and LMPP populations failed to proliferate in an E culture, confirming that they lack erythroid potential (Figure 5:24A). DEX induced cell death in the HSC culture, which may reflect their need for a unique and specialised microenvironment to facilitate their growth *in vitro* (Breems et al. 1997). The intensity of expression of cell surface markers of E differentiation on D11 confirmed the hierarchical relationships between

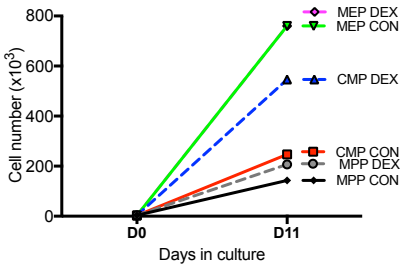
HSPC populations as MEP were more differentiated (marked by loss of CD34/CD105 and gain of CD71/CD36/GlyA) than CMP, which were more mature than MPP (Figure 5:24B). MEP displayed delayed differentiation in response to DEX (Figure 5:24C) but did not increase in number (Figure 5:24A&D). However in the case of MPP and CMP, by D11 the number of E cells increased two- fold in the presence of DEX (Figure 5:24A&D), due to an overall increase in cell number by D11 (Figure 5:24F), as well as a higher proportion of CD71+ cells by D11 (Figure 5:24E) As with MEP, E cells generated on D11 from all of the above HSPC populations were less mature, with a lower proportion expressing GlyA (Figure 5:24C) and a higher proportion in G_0/G_1 cell cycle arrest (Figure 5:24G). In summary, preliminary data suggests that DEX increases E output from HSPC upstream of MEP and maintains a pool of immature CD34-CD71+GlyA- EB by inducing cell cycle arrest. What remains elusive is how EB overcome this arrest *in vivo* in the presence of GC, to allow their full maturation and an increase in Hb.


A
D11
No. erythroid cells (x10³)
800
600
400
200
0
HSC
MPP
CMP
MEP
LMPP
GMP
Sorted population
B
D11
15000
10000
5000
300
200
100
0
MFI
CD71
CD36
GlyA
CD105
CD34
MPP
CMP
MEP
Sorted population


C
D11
% GlyA+ of total CD71+ cells
100
80
60
40
20
0
MPP
CMP
MEP
D14 EB generated from MEP
CON
DEX


D
MPP
CMP
MEP
D11
CON
DEX
CON
DEX
CON
DEX
x4


E
MPP D11
CMP D11
MEP D11
Gated on live single cells
CON
DEX
CD36
77.6%
89.8%
98.2%
93.1%
95.0%
98.3%
CD71


F
800
600
400
200
0
Cell number (x10³)
MEP DEX
MEP CON
CMP DEX
CMP CON
MPP DEX
MPP CON
D0
D11
Days in culture

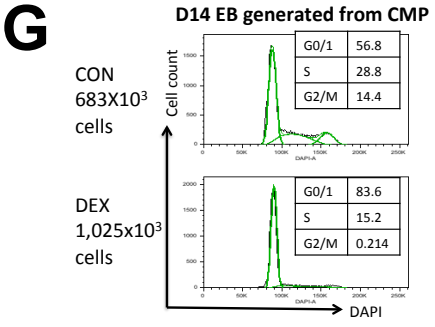

G
D14 EB generated from CMP
CON
683X10³
cells
DEX
1,025x10³
cells
Cell count
G0/1
56.8
S
28.8
G2/M
14.4
G0/1
83.6
S
15.2
G2/M
0.214
DAPI

Figure 5:24 Effects of DEX on E culture of purified HSPC from normal BM. HSPC populations from a normal BM were FACS-isolated and cultured *in vitro* in serum-free medium supplemented with cytokines to support erythroid development, with 1μM DEX or vehicle control. **A)** On day 11, the number of erythroid cells were determined by the total cell number multiplied by the % of CD71+Lin- cells. HSC, LMPP and GMP failed to proliferate in E culture. DEX (grey bars) increased E cell number from MPP and CMP, but not MEP. **B)** Mean Fluorescence Intensity (MFI) of CD34, CD105, CD71, CD36 and GlyA on D11 EB generated from MPP, CMP and MEP in the absence of DEX. From MPP to CMP to MEP, CD34 and CD105 expression declines whereas CD71, CD36 and GlyA expression increases, suggesting that MPP lie at the apex of the hierarchy. **C)** Proportion of GlyA+ (versus GlyA-) cells within the CD71+ EB fraction on D11. DEX consistently retards acquisition of GlyA. Cytospins show delayed acquisition of haemoglobin and delayed maturation of DEX-treated EB **D)** Photographic images (x4 magnification) of 96-well round bottom culture plate, showing cell densities on D11. Each image is a single well. Higher cell density in the presence of DEX in MPP and CMP cultures (but not MEP) is reflected by the increased white halo of cells. **E)** DEX increase the % of CD71+ cells in MPP and CMP cultures. **F)** DEX increases absolute cell numbers from MPP and CMP in E culture. **G)** DEX induces EB cell cycle arrest in G_0/G_1. Shown are CMP-derived cells though the same effect was observed in MPP and MEP cultures.

5.4 Discussion

5.4.1 Relative preservation of the erythroid hierarchy in RPL gene-associated DBA

The factors determining clinical heterogeneity in DBA are not well understood therefore in this chapter I have evaluated the impact of the genotype and CS therapy on the cellular landscape in DBA BM.

First, I tested the hypothesis that RPL and RPS genotypes are associated with distinct haematological phenotypes. My data show that LEP and EB are quantitatively increased in RPL versus RPS TD patients. The fact that CD71+MPP, MEP and EEP are not significantly increased in *RPL* vs *RPS* TD DBA likely reflects the presence of a partial block at the EEP to LEP transition. Importantly, *RPL* TD sample numbers are relatively low (n=6) and, due to limited BM availability, I have grouped together patients with a range of different RP gene mutations. Pathogenic variants in distinct RP genes may in fact exert different effects on erythroid development *in vivo.* For example, clinical data show differences between RPL5 and RPL11 gene-associated DBA but I have grouped them together in the current analysis. Data from a patient with deletion of *RPL35a* is also included in the RPL group although there is insufficient evidence to determine whether *RPL35a* segregates with *RPL5/11* or *RPS* genotypes in terms of clinical features.

Despite these caveats, my data are made more convincing by the concordance between immunophenotypic findings, BM morphology and clinical data. It is now important to examine the underlying mechanisms for these differences by functional characterisation of EP populations. Future work should include clonogenic assays and liquid culture assays to compare the function of EEP and LEP of *RPL* TD, *RPS* TD and control samples. I hypothesise distinct proliferation and/or differentiation profiles of *RPL* and *RPS* EP *ex vivo*, albeit that both are likely to be inferior to normal. In addition to the cellular assays outlined above, further work could entail a comparison between the transcriptome of EEP derived from RPL and RPS patients, as well as normal controls. This may provide insight into the divergent molecular mechanisms underpinning the heterogeneous cellular phenotypes observed. Although RP have a shared role in making up the ribosomal subunits, they also have distinct extra-ribosomal roles as discussed in Chapter 1; for example *RPL5/11* are specifically involved in p53 stabilisation by HDM2 (Dai & Lu 2004) and, unlike many of the RPS (and other RPL) proteins disrupted in DBA, RPL5/11 proteins are necessary for the maintenance of nucleolar structure (Nicolas et al. 2016). Moreover,

over the last few years a new concept in ribosome biology has emerged, that of ribosome specialisation whereby ribosomes are structurally and functionally dynamic across different tissues and within a tissue over time (Slavov et al. 2015) (Mills & Green 2017). It is therefore possible that E lineage cells have differential requirements for particular RP at different stages of development, accounting for the varying haematological phenotypes according to genotype.

Notably, a previous publication has shown a more *severe* erythroid defect of PBMCs isolated from *RPL5/RPL11* compared with *RPS19* mutated patients. (Moniz et al. 2012). While *RPL5/RPL11*-deficient PBMCs exhibited reduced erythroid proliferation, delayed differentiation, and markedly increased apoptosis *in vitro*, *RPS19*-deficient cells exhibited reduced proliferation without inducing significant apoptosis. To explain these differences, the same group later showed that *RPL5/11* but not *RPS19*-deficient erythroid precursors (generated *in vitro* from patient PBMCs) expressed lower levels of GATA-1 protein due to increased proteasomal degradation of HSP70, which itself protects GATA-1 from degradation (Gastou et al. 2017). This paper contradicts a previous study suggesting that selective reduction in translation of GATA-1 mRNA is the unifying mechanism of selective failure of the erythroid lineage in both *RPS and RPL* haploinsufficient cellular models of DBA (Ludwig et al. 2014). It is important to note that the systems used above (either knockdowns or primary cells manipulated by cytokine/growth factor exposure in *vitro*) may not faithfully recapitulate conditions *in vivo* hence the utility of my approach to address erythroid failure mechanisms using primary BM cells, as will be discussed below.

5.4.2 Characterisation and prospective isolation of distinct stages of EB

In order to study the late erythroid defect in patients with DBA due to pathogenic variants in *RPL* genes, I first set out to prospectively isolate distinct stages of EB. This was an important prerequisite for comparing gene expression in equivalent EB populations in normal and DBA. I used a combination of CD105 and GLYA expression

within the LIn-CD34-CD71+ fraction of total BMMNC, to resolve immunophenotypically and morphologically distinct EB stages. Concurrent with my resultant immunophenotypes, a recent publication showed that GLYA expression commences at the proEB stage and CD105 decreases at the polyEB stage (García-Vela et al. 2017). Another group outlined a strategy using integrin α4/CD49d and band 3. As an anti-band 3 antibody is not commercially available I did not adopt this strategy; however I verified that, as suggested by their data, CD49d expression is high in my E3/basophilic EB population.

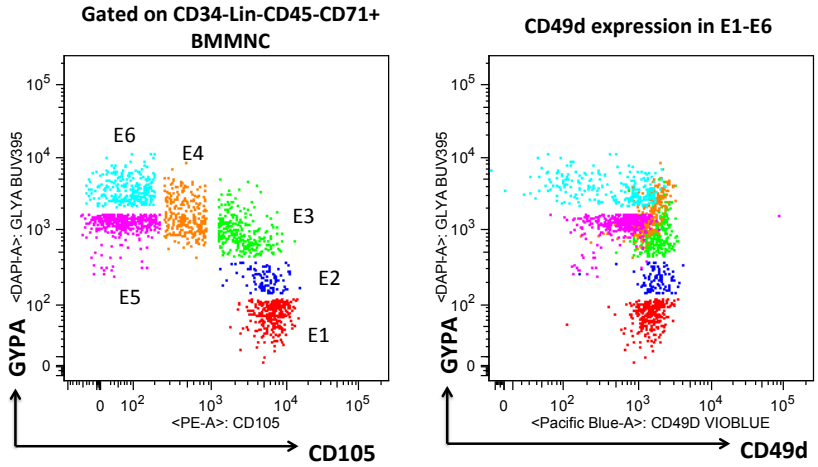


Figure 5:25 CD49d/integrin α 4 expression in EB populations E1-E6, by flow cytometry. E1 to E6 were delineated by CD105/GLYA expression, as outlined in **Figure 5:6** (left plot). An overlay plot was generated showing CD49d expression in each of the E1-E6 populations (right plot). CD49d is highly expressed from ProEB (E1/2) to BasoEB (E3), then decreases.

5.4.3 Mechanisms of late erythroid failure in DBA determined by RNA-seq

Current evidence suggests that defective ribosome biogenesis arising from RP gene haploinsufficiency results in pro-apoptotic erythropoiesis and erythroid development failure, possibly through activation of p53-dependent apoptotic pathways (Sjögren & Flygare 2012). Consistent with this concept, my RNA-seq data show that upstream p53 signalling is one of the most striking molecular signatures in RPL EB (Figure 5.12). It is therefore possible that EP in RPL patients retain some ability to proliferate and

differentiate into EB but are then rendered susceptible to p53-mediated cell death, resulting in anaemia. Indeed, in a recent murine model of RPL11 haploinsufficiency, BFU-E and CFU-E progenitor numbers were *increased* relative to normal mice, with a corresponding decrease in Ter119+ (equivalent to GLYA+) mature E cells. Transcriptome profiling of BFU-E and CFU-E progenitors showed enrichment for the p53 pathway (Sieff et al. 1982). The mechanisms by which p53 is activated in RPL patients remain unknown and warrant further investigation, as this may shed light not only on the pathogenesis of the anaemia in DBA but also on malignant transformation in DBA, to both leukaemia and solid cancer.

In addition to the known pathway of p53-mediated erythroid failure in DBA, these data also illuminate potential new disease mechanisms. Most strikingly, immunoproteasome genes and genes regulated by IFNϒ, IFNα2 and TNFα signalling are highly dysregulated in DBA EB. This provides support for findings from RBC cytosolic proteome analysis in patients with DBA where key upstream regulators of the DBA proteome included mitogen activated kinase 1, IFNγ, TNFα (Pesciotta et al. 2015). It is not yet clear whether inflammation occurs incidentally, for example secondary to p53 activation, which is known to stimulate TNFα in the context of RPS19 haploinsufficiency (Bibikova et al. 2014). Alternatively, one could speculate that impaired biogenesis of ribosomal rRNA, results in free cellular RNA that can aberrantly activate intracellular receptors involved in innate immunity, stimulating inflammatory pathways within erythroid cells. Another possibility is that activation of inflammation in E lineage cells results from a microenvironmental effect such as aberrant cross-talk between EB and their centrally placed macrophages in the BM EB islands. In line with this hypothesis, I found significant downregulation of *GAS6* in DBA EB, a molecule required for normal EB/macrophage interactions and normal E development (Hom et al. 2015). Further experiments are required to confirm or refute the role of macrophages in DBA pathogenesis. In turn, this may shed light on the therapeutic effects of GC in DBA as recent evidence shows that, in addition to

their direct effects on E cells. GC can enhance erythropoiesis via indirect stimulation of EB island macrophages (Falchi et al. 2015).

Another question generated by these data is: what are the ramifications of this inflammatory signature?. Inflammatory cytokines are known to suppress E differentiation (Chapter 1) and in one study this effect occurred via TNFα-mediated inhibition of GATA-1 translation (Bibikova et al. 2014). However, my data do not show evidence for reduced GATA-1 protein or impairment of its function. In fact I found that the expression levels of the majority of GATA-1 target genes are normal in EB from patients with mutations in *RPL5/RPL11*. This finding plays into a current contentious debate on the role of GATA-1 in DBA pathogenesis (Ludwig et al. 2014; O'Brien et al. 2017; Gastou et al. 2017; Ulirsch et al. 2017; Farrar et al. 2017; A. R. Migliaccio & Varricchio 2017). Given its importance, I investigated GATA-1 protein expression in EB directly, by performing GATA-1 IHC on FFPE BM sections from patients with DBA. These experiments suggest that GATA-1 expression is normal, relative to control BM, in DBA BM from both RPS and RPL gene-linked cases, although more sensitive techniques would be required to confirm this. As reduced GATA-1 translation remains an attractive hypothesis given the erythroid tropism in DBA which still remains a major source of biological intrigue.

Interestingly, RNA- sequencing data from in RPL EB showed evidence for the deregulation of other TF. Specifically, *GATA-2* was highly upregulated at the mRNA level. Given that overexpression of *GATA-2* in erythroid cell lines suppresses differentiation (Guo et al. 2015), this observation raises the interesting new concept that failure to downregulate GATA-2 may contribute to E failure in DBA.

One overarching unknown is how the observed changes in gene expression are linked to the underlying ribosomal defect in DBA. Cell number limitations preclude integration of the transcriptome with the proteome in these cells, which would help to elucidate the interplay between translational and transcriptional abnormalities in

DBA. In the absence of these data, we can only hypothesise that the translation of specific mRNAs is selectively impaired in DBA, leading to perturbed protein synthesis and in turn deregulation of genes regulated by these proteins.

5.4.4 Restoration of EP frequency and function in CSR patients with DBA

In this chapter I demonstrate restoration of EEP and LEP defects in CSR patients with the RPS genotype, another novel finding that has not been described previously. Although DEX has been shown to increase BFU-E formation from human CD34+ cells, these were either normal or manipulated by *si*-RNA medicated knockdown of *RPS19.* Furthermore, CFU-GM colony forming capacity was also increased by DEX treatment, hence its effect was not erythroid specific in this model (Ebert 2005). A study performed in the 1980s showed that BM from a CSR patient with DBA produced increased BFU-E and CFU-E colonies post- compared to pre-treatment, whereas E colonies were absent pre- and post-treatment in a CS resistant patient (Chan, Saunders & Freedman 1982a; Chan, Saunders & Freedman 1982b). By contrast, in a more recent study performed using PBMCs from DBA patients and normal donors, CS enhanced *in vitro* erythropoiesis in both steroid-responsive and –resistant DBA patients (Ohene-Abuakwa et al. 2005). This work was limited by its use of unfractionated PB and furthermore, only 5/33 patients had their diagnosis of DBA confirmed by genetic means. By contrast, I was able to use highly purified, well-characterised BM progenitors from a genetically defined cohort of patients where there was a clear correlation between clinical response to steroids and EP frequency Notably, in CSR patients with the RPL genotype, CS did not significantly alter EP number or function. It is possible that CS have a distinct mechanism of action in this subgroup of patients. However, it is also important to note that, for both RPS and RPL patients, different individuals are being compared on and off CS, with the exception of 1 patient with an *RPL11* mutation. In this particular case, BM EB increased from 5% to 18% of total nucleated cells after commencement of CS (Figure

5:18) suggesting that if the same individuals were compared pre- and post-treatment, an improvent by CS in *RPL* patients might be observed.

To explore the mechanisms by which GC enhance DBA erythropoiesis, I demonstrate that CS treatment does not increase GATA-1 protein expression in DBA EB. I also show in 1 patient, by RNA-seq, that CS do not upregulate expression of the mutated RP gene in E lineage cells (Figure 5:11). Finally, I examine the effects of DEX on normal haematopoietic cells *in vitro.* While GC have previously been thought to primarily target BFU-E progenitors directly (Flygare et al. 2011) (Narla et al. 2011), my data suggest, for the first time, that GC exert pleiotropic effects at multiple points of the haematopoietic hierarchy: increasing the generation of CD71+ E cells from MPP and CMP, while retarding differentiation of EB at the point of acquisition of GlyA.

Further experiments, for example using a single cell approach, are needed to confirm these effects and to determine whether GC increase E output from HSPC by increasing self-renewal of committed EP and/or by driving multi-/ bi-potent progenitors towards the E lineage versus other lineages. Recent data show that growth factors can drive lineage fate, for example EPO can increase multipotent progenitor commitment to the E lineage (Grover et al. 2014). It remains to be seen whether the mechanisms of action of GC in DBA are analogous to normal erythropoiesis. Given that I have demonstrated improved EP function in CSR patients with DBA, my future aim is to investigate the molecular mechanisms of GC action in EP +/- HSPC populations sorted from CON, DBA TD and DBA CSR BM. Integration of transcriptomic and chromatin accessibility data, using RNAseq and Assay for Transposase Accessible Chromatin (ATAC)-seq respectively, should help to identify the genes and gene networks directly modulated by GR in DBA EP. In turn this may lead to the development of novel therapeutic modalities.

6 SUMMARY, CONCLUSIONS AND FUTURE WORK

Although DBA was first described almost 80 years ago, the field of DBA research continues to grapple with many unanswered questions (Figure 6:1). Several models have been developed to study DBA; however, I set out to address some of these questions using primary BM samples from patients with the disease.

UNANSWERED QUESTIONS IN DBA

- What genetic lesions underpin the 30-40% cases where a known RP gene is not identified?
- Are there any associations between genotype and haematological phenotype or disease outcomes?
- What is the unifying mechanism that links distinct mutations in different RP genes?
- What are the reasons for variable penetrance in DBA, even in patients with the same genotype?
- Why does DBA usually present after birth and not in fetal life?
- What are the cellular and molecular features of erythroid failure in DBA?
- Why is there a particular tropism for the erythroid lineage in DBA?
- How do steroids restore erythropoiesis in DBA?
- What are the molecular determinants of primary and secondary steroid refractoriness in DBA?
- What are the mechanisms for steroid-induced and spontaneous remission in DBA?
- What is the mechanism of neoplastic transformation in DBA and why is there a predilection for certain malignancies?

Figure 6:1 Unanswered questions in DBA research; 2017

In this chapter, I will summarise my salient findings and discuss their significance in the context of existing knowledge of both erythropoiesis and DBA. I will also

highlight the limitations of my work and the questions that remain unanswered, which inform the directions for future research.

My first aim was to evaluate the clinical and genetic spectrum of DBA in the UK (Chapter 3). This was a prerequisite to the work as the last report from the UK registry was in 2004 and there were no residual registry records. I therefore collected and analysed data from 123 patients in the UK. To my knowledge, this constitutes one of the largest international cohorts (superseded only by USA, Germany and France) and the only one where the vast majority of patients have been assessed at a single centre.

The first important finding is that the UK cohort has much in common with the other international registries of patients with DBA. Specifically, in terms of epidemiology, DBA has an incidence of approximately 4 in 1 million live births and shows no ethnic predilection. *RPS19* is the most common gene affected and a wide spectrum of pathogenic genetic variants, from missense mutations to large deletions, are implicated. The majority of cases occur *de novo* and 30-40% of cases remain genetically undiagnosed, despite comprehensive NGS screening of 80 RP genes.

Future work in this area will involve the validation of novel mutations in known RP genes as well as mutations in new RP genes not previously implicated in DBA, as well as identification of mutations in new non-RP genes that phenocopy DBA. The latter will hopefully be achieved by analysis of the data collected as part of the 100 000 Genomes Project. Notably, traditional family genetic linkage studies are of limited utility in DBA, given the variable penetrance of the disease and the fact that individuals can carry an RP gene defect in the absence of a significant anaemia. Therefore, ascribing pathogenicity to new genetic variants will require the upkeep of accurate mutation databases and collaboration with other centres, to identify variants that recur in pedigrees worldwide, as well as collaboration with biologists able to generate cellular and animal models to test the effects of specific variants.

In terms of clinical features, commensurate with the data from other cohorts (Gazda et al. 2008) (Quarello et al. 2010), I found that: i) RPL genotype is associated with multiple as opposed to no or single congenital abnormalities (P=0.02), ii) *RPL5* is associated with cleft lip/palate (P<0.0001), and iii) *RPL11* is associated with thumb abnormalities (P<0.0001). These findings suggest that the data acquired from the UK cohort are likely to be relevant to DBA patients worldwide.

My analysis also provided several original epidemiological and biological insights. First, I observed that gender appears to be an important modifying factor in that patients with RPL5/11 genotype are much more likely to be female (P<0.0001) and the majority (9/10) of familial cases are transmitted from the mother, rather than the father (P=0.003). The reasons for these underlying gender differences remain elusive and warrant further investigation.

Second, this is the first work to record the prevalence and spectrum of skeletal abnormalities, hernias, neuropsychiatric and gastrointestinal disorders in DBA. In addition, a possible association was identified between RPS genotype and hernias (P=0.06). Prospective multicentre studies are now needed to better characterise these clinical features. In the future, laboratory and clinical phenotypes could be recorded using Human Phenotype Ontology (HPO) terms (Robinson et al. 2008) to improve the consistency of documentation and to facilitate statistical analysis.

Third, several novel associations were identified between RPS versus RPL genotypes and clinical phenotype (Figure 6:2). RPL is associated with a more severe non-haematological phenotype, including birth complications such as prematurity, IUGR and SGA (P=0.04). These are harbingers of short stature in later life, which is also more prevalent in *RPL* patients (P=0.03). Additionally, there is an elevated risk of reduced bone mineral density (P=0.01) and hearing difficulties (P=0.01). By contrast, comparing *RPL11* (and to a lesser extent *RPL5*) with *RPS19*, the former is associated

with a less severe haematological phenotype, manifest as a later age of presentation with anaemia (P=0.002), higher initial response rates to CS (P=0.01) and longer duration of treatment-free or CS-dependent periods as opposed to transfusion-dependence (P=0.02). This has important ramifications for patients as the highest morbidity and mortality in DBA is associated with transfusion-dependence (Lipton et al. 2006).

Another important implication of the above work was that the DBA population from whom I collected BM samples had undergone comprehensive clinical and genetic annotation, allowing subsequent results to be grouped according to genotype and treatment status, which revealed important biological insights, as described below.

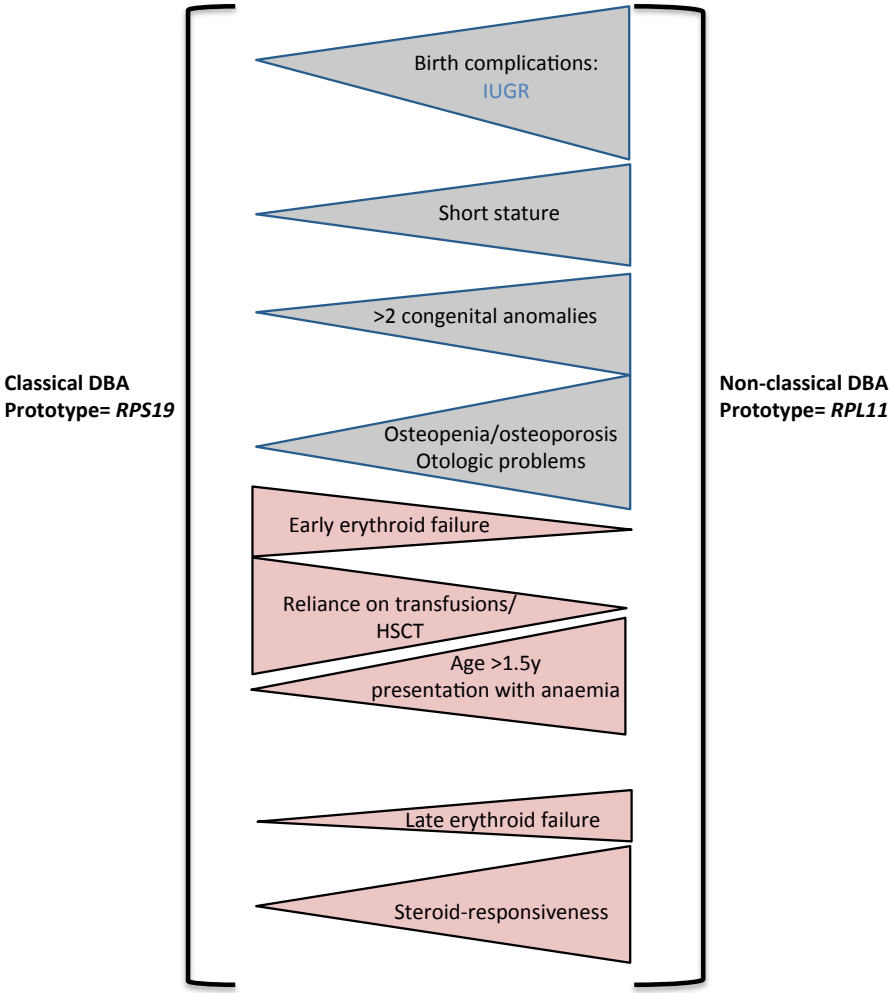


Figure 6:2: A new disease paradigm in DBA. *RPS19* and *RPL11* are prototypes of classical and non-classical DBA, respectively. Grey shading shows non-haematological parameters while pink shading shows haematological parameters. Decrescendo or crescendo shapes reflect whether the incidence of these parameters decreases or increases in non-classical DBA.

My next aim was to investigate the cellular architecture of haematopoiesis in normal and DBA paediatric BM (Chapter 4). I used this approach because the immunophenotypic characterisation of HSPC compartments has advanced considerably over the last decade and provided mechanistic insights into other haematological diseases, such as trisomy 21-associated leukaemias (A. Roy et al. 2012) and MDS (Pang et al. 2013), however to date the cellular hierarchy is poorly characterised in DBA. Specifically, I hypothesised that in DBA, quantitative and/or qualitative defects in MEP and EP populations underpin the erythroid defect.

One obvious challenge of this work was the limited availability of BM samples as the procedure was only done when clinically indicated, both for healthy paediatric BM donors and for patients with DBA. Low volume samples were available and low cell number in DBA was also compounded by 3 factors:

1) hypocellular BM, which is prevalent
2) the majority of experiments required isolation of CD34+ cells (0.5-5% of total cells) from BM, and then analysing sub-populations within this fraction
3) the fact that certain sub-populations, e.g. EP, were under-represented in DBA versus CON BM

It was therefore necessary to optimise all experiments for the lower anticipated cell number expected from the disease sample. Additionally, it was not possible to perform technical replicates and the ideal controls, for all experiments and all samples, for example for clonogenic assays and qRT-PCR. This resulted in low 'n' numbers for some experiments or equivocal results, for examples in cell cycle assays,

which require millions of cells. Despite these caveats, there was remarkable consistency within both the DBA and control groups. This reflects the careful clinical annotation of samples. Above all it suggests that the experimental approaches used are powerful enough to address my hypotheses above. Another strength of this work is that samples were age-matched, given that differences in the frequencies and function of HSPC populations with age have been demonstrated (Pang et al. 2017).

Prior to my wok there were few data defining erythropoiesis in human and in particular paediatric BM, therefore, before I could characterise the defects in DBA BM, it was necessary to investigate the normal erythropoietic hierarchy. To do this, I used a range of known E and MK markers to characterise the bi-potent MEP population in human BM samples. Functional studies confirmed that I was able to use these markers to enumerate and prospectively isolate EP enriched in BFU-E and CFU-E activity in human BM. Notably, this strategy achieved high purities and high clonogenic efficiencies of 67% and 58% respectively (Iskander et al. 2015). Later work from our group confirmed the E lineage specificity of my candidate EP phenotype at a single cell transcriptional level (Psaila et al. 2016). I therefore developed a novel approach allowing the prospective identification, enumeration and isolation of human BM EP that give rise directly and selectively to BFU-E and CFU-E.

Concurrently another group also published a strategy for the identification of human BFU-E and CFU-E progenitors (Li et al. 2014). These 2 approaches have some similarities and some differences, which most likely reflect the tissue and system used. A unique strength of my work is that the data were generated from un-manipulated *ex vivo* samples as opposed to *in vitro* generated erythroid progenitors and precursors. I also validated my approach by demonstrating its utility in elucidating the EP defect in a disease context, DBA.

Although I originally focussed on identifying committed EP within the MEP population *a priori*, I later found that the CMP could be stratified into erythroid, myeloid and megakaryocyte subpopulations in precisely the same way as the MEP, calling into question the historical distinction made between CMP and MEP. These findings provided justification for including CD123- EP within in my sorting strategy for subsequent experiments. While I did not specifically pursue the question of the origins of E commitment in normal haematopoiesis, there is now mounting evidence that MK lineage commitment arises upstream of MEP (Sanjuan-Pla et al. 2013) (Miyawaki et al. 2017) and my data suggest that the same may be true for erythroid cells, although further work is needed to draw firm conclusions.

In addition to identifying EP populations, in Chapter 5 I also present a new strategy for immunophenotyping of more mature EB populations. In so doing I have characterised in detail, using *ex vivo* BM samples, the onset and loss of expression of 6 cell surface markers during human erythropoiesis, from HSC to RBC (Figure 6:3). These data constitute another original contribution to the field as they have the potential to be applied to a range of disorders of erythropoiesis.

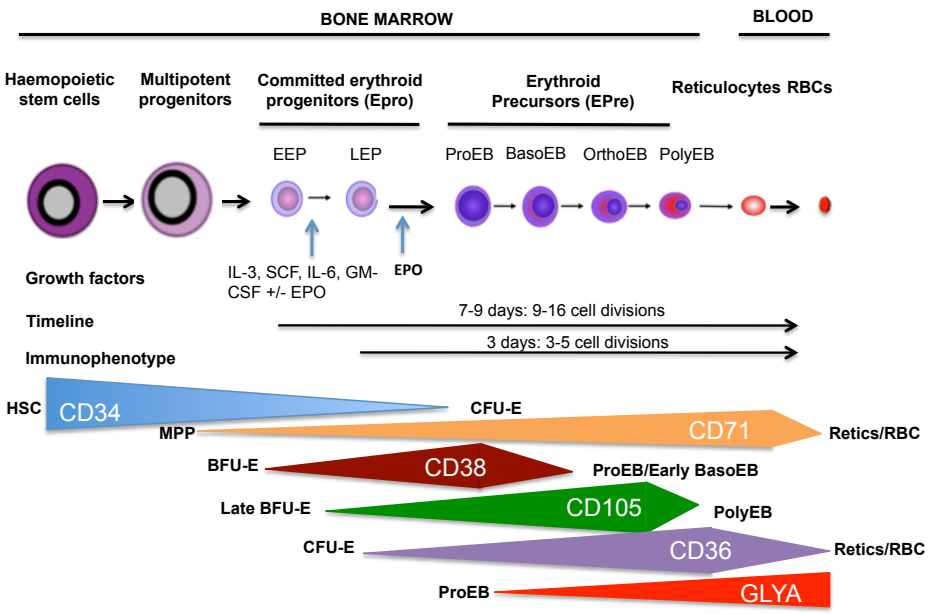


Figure 6:3 Immunophenotypic characterisation of erythropoiesis, from HSC to RBC. By combining FACS-isolation of BM populations *ex vivo* with morphological, cellular

and transcriptional studies, I identified the precise stages of gain and loss of expression of CD34, CD71, CD38, CD105, CD36 and GLYA during normal human erythropoiesis.

As well as providing insight into normal haematopoiesis, another aim of my work was to dissect the cellular architecture of early haematopoiesis and erythropoiesis in DBA. I present novel data showing that HSC MPP, CMP, GMP and LMPP are quantitatively normal in BM from *RPS* TD patients with DBA, compared with age-matched controls, while MEP are selectively reduced. Notably, many patients with DBA have overall BM hypocellularity therefore where HSPC relative frequencies are normal, absolute frequencies may still be reduced. A detailed functional analysis of the CD38- fraction, that includes HSC, MPP and LMPP, was not performed given limited cell number; therefore, further work could involve an assessment of the function of primitive HSPC populations. Qualitative abnormalities in these compartments may help to explain multi-lineage cytopenias in patients. Data from 2 patients suggest that evolution to BMF +/- MDS in DBA is accompanied by a striking loss of HSC, MPP, CMP and MEP, though genetic diagnosis of these patients is required to confirm this observation.

Focussing on the erythroid lineage in DBA, I applied the novel gating strategy described above, to evaluate functionally homogeneous populations of EP. I found that quantitative and qualitative defects in CD71+ EEP and LEP underpin the observed abnormalities in MEP isolated from DBA BM. Although it is well established that EP are impaired in DBA (Lipton et al. 1986; Nathan, Clarke, et al. 1978a; Ohene-Abuakwa et al. 2005), prior to my work the precise nature of these defects had not been elucidated. Using unmanipulated cells *ex vivo*, I clearly demonstrated that erythroid failure in DBA is a product of reduced frequencies of EP, compounded by their failure to proliferate and differentiate into mature EB. Indeed, a three-fold reduction in EEP is followed by a four-fold reduction in LEP and a twenty-fold

reduction in EB. Finally, the finding that EEP are reduced in frequency suggests that the erythroid defect in DBA has its origins upstream of these EP populations. Commensurate with this, I demonstrate a three-fold reduction in the CD71+ MPP subpopulation within the primitive CD38- fraction of DBA BM. If confirmed by more detailed studies, this would suggest that impaired erythropoiesis in DBA originates at the apex of the haematopoietic hierarchy. To confirm this hypothesis, transcriptomic, epigenomic and proteomic profiling of CD34+ CD38- cells could be used to investigate whether primitive HSPC display the molecular hallmarks of DBA.

Consistent with the clinical differences observed between RPS and RPL genotypes (Chapter 3), I also found evidence for 2 distinct cellular phenotypes in TD patients with DBA (Chapter 5). While *RPS* patients exhibit evidence of an *early* EP defect, as described above, EP and EB populations are relatively preserved in *RPL* patients. This may help to reconcile the variable haematological phenotypes observed in previous studies using *in vitro* erythroid culture systems (Ohene-Abuakwa et al. 2005), where genotyping was not available for all patients. In order to elucidate the mechanisms underpinning these differences in early erythropoiesis, future work should investigate the cellular and molecular properties (by cell culture assays and RNA-seq, respectively) of *ex vivo* sorted EP, comparing CON, TD *RPS* DBA and TD *RPL* DBA patients, with the aim of identifying shared and distinct mechanisms of erythroid failure.

To shed light on my observation that *RPL* patients do not have an absolute differentiation arrest at the EEP/LEP stage, yet they are reliant on chronic blood transfusion, I next investigated whether the EB present in RPL gene-associated DBA display evidence of a cell intrinsic *late* erythroid defect. To do this I performed transcriptome profiling by RNA-seq of stage-matched EB, purified from normal and DBA BM. PCA and hierarchical clustering demonstrate a clear distinction between DBA and CON samples. Gene expression data show that RPL EB are functionally abnormal, with a molecular signature in keeping with activation of upstream p53

signalling. This is compatible with the prevailing concept that p53–mediated cell death is an important mechanism in DBA pathogenesis (Dutt et al. 2011; Sieff et al. 2010; Ball 2011).

Importantly, I found that both the expression levels of most GATA-1 target genes (by RNA-seq) and the expression of GATA-1 itself (by IHC on BM trephines) in EB from patients with mutations in *RPL5/RPL11* are comparable with normal controls. This challenges a current hypothesis in the field that selective reduction of GATA-1 translation is the unifying mechanism linking ribosomal biogenesis defects with erythroid failure (Ludwig et al. 2014). Further work is now warranted to investigate GATA-1 expression in progenitor populations upstream of EB, and in patients with the *RPL* or *RPS* genotype.

Finally the RNA-seq data provide the first evidence for a striking pro-inflammatory signature within DBA EB, in support of findings from mature RBC cytosolic proteome analysis in patients with DBA (Pesciotta et al. 2015). Specifically, immunoproteasome genes and genes downstream of IFNγ, IFNα2 and TNFα signalling are highly unregulated in DBA EB. Inflammatory signals are known to impede E differentiation (Chapter 1) however it is not yet clear how inflammation arises within E lineage cells in DBA, and whether this is a primary or secondary mechanism of erythroid failure. These data also raise the tantalising possibility that targeting the microenvironment e.g. by modulating the interactions between EB and macrophages in the EB island, may be a new therapeutic avenue in DBA, at least in RPL-gene associated cases where EB are present.

A natural extension of the work above would be to validate the mechanisms of erythroid failure suggested by transcriptomic analysis of DBA EB, using a model system such as the recently developed erythroid cell line human umbilical cord blood- derived erythroid progenitor (HUDEP) cell lines (Kurita et al. 2013). HUDEP-2 in particular offers improved enucleation rates and higher expression of adult

haemoglobins relative to other erythroid cellular models. It could therefore be used to knock out/down and knock in/overexpress candidate genes of interest to test whether these recapitulate or rescue the DBA phenotype. The high cell numbers required for the above mechanistic studies precludes the use of primary cells.

The final aim of my work was to determine the cellular and molecular mechanisms of GC responsiveness in normal and DBA erythropoiesis. First, I confirmed that EEP, LEP and EB populations are improved in number and function in CSR patients with RPS gene pathogenic variants receiving CS therapy at the time of BM sampling. Notably, the cellular phenotype of CSR patients lies at an intermediate stage between TD DBA and CON, compatible with the clinical observation that patients on CS still demonstrate key signs of the underlying erythropoietic defect, such as macrocytosis, raised HbF and eADA (Vlachos & E. Muir 2010).

Next I investigated whether GC exert a direct effect on BFU-E progenitors, as established in murine studies (Flygare et al. 2011; L. Zhang et al. 2013). Surprisingly I found that addition of a synthetic GC, DEX, delays the differentiation of EEP but fails to increase EEP proliferation *in vitro*. Potentially this could still be a mechanism by which GC enhance erythropoiesis *in vivo,* by promoting survival of the EP pool and retarding their maturation. Indeed this corroborates recent murine data showing that GR-induced genes, such as *Zfp36l2,* maintain normal BFU-E in a self-renewal state, blocking their terminal differentiation (L. Zhang et al. 2013). GC have also been shown to retard accelerated E differentiation observed in *RPS19*$^{+/-}$ murine FL cells (Sjögren et al. 2015). Future work will therefore involve analysing the EP transcriptome in CSR versus TD DBA patients.

However, as the impact of GC on human BM BFU-E progenitors (EEP) was less striking that that reported in murine FL, I also hypothesised that CS may exert their proliferative effect *upstream* of the committed EP that lie within the MEP subfraction of BM. By examining the cellular effects of GC in an *in vitro* erythroid

culture of FACS-purified HSPC from normal BM, I present preliminary data that CS have a stimulatory effect on MPP and CMP, increasing the yield of erythroid cells in culture. Though not yet formally demonstrated, I hypothesise that GC achieve this by either i) increasing the commitment of multi- or bi-potent progenitors towards the E lineage, ii) increasing self-renewal of committed EP that lie upstream of MEP or iii) indirectly promoting erythropoiesis by driving EB island macrophages towards an anti-inflammatory phenotype (Falchi et al. 2015). To reach definitive conclusions regarding the effect of GC on normal haematopoiesis, it is necessary to repeat these experiments, ideally using a single cell approach. Future work could also involve investigation of the effect of GC on DBA HSPC populations *in vitro,* to test my hypothesis that, as in normal HSPC, CS restore the EP defect in DBA but exerting multiple effects on multiple levels of the haematopoietic hierarchy. The fact that CD71+ cells are increased in the CD38 negative primitive HSPC fraction of BM from CSR patients with RPS-gene associated DBA offers the first indication that this may indeed be the case. The above cellular studies could be complemented by genome-wide transcriptomic and epigenomic analysis (by RNA- and ATAC-seq respectively) of the changes induced by GC in human HSPC.

Together these experiments are likely to lead to the identification of novel genes or gene networks involved in DBA and impacted upon by CS. Designing treatments that target these directly may circumvent CS-associated side effects and lead to haematological responses in a higher proportion of patients. Notably, elucidation of the specific molecular determinants of CS-responsiveness versus primary or secondary refractoriness, and of spontaneous or CS-induced remission, was beyond the scope of my current work and remains a source of intrigue in the field.

In conclusion, I have provided the first comprehensive analysis of the haematopoietic hierarchy in DBA, compared with normal paediatric BM. Specifically I have shown that the relative proportions of non-erythroid lineages are preserved in DBA. At the same time, there are quantitative *and* qualitative defects in MEP;

specifically in EP that generate BFU-E and CFU-E colonies *in vitro*. I also show novel data concordant with 2 distinct haematological phenotypes in DBA, characterised by early progenitor and late precursor erythroid defects, which segregate with RPS versus RPL genotypes, respectively. Remarkably, these differences also correlate with differences observed in clinical phenotype as patients with the most common RPL genotype, *RPL11,* have a milder haematological picture and are more likely to respond to CS than patients with the most common RPS genotype, *RPS19*. If confirmed by further studies, these data will have important clinical implications as the identified differences may help to streamline genetic testing to prioritise searching for *RPS* versus *RPL* variants, and, most importantly, allow for prognostication.

Prior to my work there were few, if any, data characterizing haematopoiesis in BM from healthy children. In addition, I confirmed that my novel sorting strategies for different stages of the erythropoietic hierarchy are applicable to adult BM. The significance of the current work therefore extends beyond the field of DBA and can be applied to the investigation of erythropoiesis in several other physiological and pathological contexts.

7 APPENDICES

7.1 Appendix A: Ethics documentation

Imperial College Healthcare NHS
NHS Trust

Centre Number:
Study Number:
Patient Identification Number for this study:

Consent Form for Parents/ Carers of child with bone marrow failure disease

Version 2.0, July 2013

Project title: **Molecular and cellular pathogenesis of bone marrow failure syndromes**

Name of Researchers: Professor Anastasios Karadimitris, Professor Irene Roberts, Dr Josu De La Fuente, Dr Deena Iskander, Department of Haematology, Imperial College Healthcare Trust, Du Cane Road, London, W12 0NN.

Please initial box

1. I confirm that I have read and understood the information sheet, version 1 dated January 2012, for the above study, I have had the opportunity to consider the information, ask questions and have had these questions answered satisfactorily. ☐
2. I understand that my child's participation is voluntary and that I am free to withdraw my child at any time, without giving any reason, without their medical care or legal rights being affected. ☐
3. I understand that sections of my child's medical notes and data collected during the study may be looked at by authorised individuals from Imperial College London, from regulatory agencies or from Imperial College Healthcare NHS Trust, where it is relevant to my child taking part in the research. I give permission to these individuals to have access to my child's records. ☐
4. I agree to my child's GP being informed of my child's participation in this study. ☐
5. I understand, and agree to, data relating to my child's donated sample being stored electronically. Patient identifiable details will be stored on password protected computers at Imperial College London. ☐
6. I would like/ not like to be informed of any new findings relating to bone marrow development or bone marrow failure diseases, which may be discovered through experimental research on my child's blood/ bone marrow sample (delete as appropriate). ☐
7. I agree to genetic tests being conducted on my child sample. ☐
8. I agree for my child's sample to be stored for future use in ethically approved projects. ☐
9. I agree for my sample to be stored/analysed at another laboratory in the UK ☐
10. I agree to…….. taking part in this study. ☐
    Name of child

Name of Parent/Guardian	Date	Signature
Name of Child (if would like to sign)	Date	Signature
Name of person taking consent	Date	Signature
Name of researcher	Date	Signature

Participant's Special Requirements (e.g. other language/communication method)

A copy of this form and the accompanying Information Sheet have been given to the participant

1 copy for patient/participant; 1 copy for Principal Investigator; 1 copy (original) to be kept in medical notes.

1

7.2 Appendix B: Constituents of reagents and buffers

MethoCult™ GF H4034 Optimum (concentrations not disclosed by manufacturer)

Methylcellulose in Iscove's MDM

Fetal bovine serum

Bovine serum albumin

2-Mercaptoethanol

Recombinant human stem cell factor (SCF)

Recombinant human interleukin 3 (IL-3)

Recombinant human erythropoietin (EPO)

Recombinant human granulocyte colony-stimulating factor (G-CSF)

Recombinant human granulocyte-macrophage colony-stimulating factor (GM-CSF)

Supplements

DAPI x1 solution

To 500 ml dd H2O add:

8.5g NaCl (final conc. = 146 mM)

1.2 g Tris Base (final conc. = 10 mM)

Adjust pH to 7.4 with HCl.

Add:

4 ml of 500 mM CaCl2 solution (final conc. = 2 mM)

44 ml of 500 mM MgCl2 solution (final conc. = 22 mM)

50 mg (0.05g) BSA

1 ml nonident P-40 detergent (final conc. = 0.1%)

10 mg DAPI (4,6-diamidino-2-phenylindole*) powder

(final conc. = 10 µg/ml)

100 ml DMSO (final conc. = 10%)

Add dd H20 to final volume of 1 L.

Store in dark at 2-6°C.

7.3 Appendix C: Flow cytometry panels and monoclonal antibodies

Multicolour panels of anti-human fluorophore-conjugated monoclonal antibodies are detailed below. Panels were used for both flow cytometric analysis and fluorescence-activated cell sorting.

HSPC panel: 1, 20, 22, 23, 24, 26, 29

EP panel: 1, 4,10, 11, 12, 19, 21, 23, 25, 29

HSPC +EP panel: 1, 4,10, 11, 12, 19, 21, 23, 25, 27, 29

Erythroid panel: 1, 4, 5, 7, 9, 12, 13-18, 23

Colony panel: 2, 5, 6, 8, 28

+ DAPI (for analysis using LSR Fortessa) or BV510 Aqua live dead fixable stain (for FACS-sorting using BD Fusion or FACSAria II)

Cell death panel: 31+ DAPI

ID	Surface marker	Fluorophore	Clone	Manufacturer
1	CD34	PeCy7	4H11	eBioscience
2	CD34	PerCpCy5.5	581	Biolegend
3	CD36 (thrombospondin receptor)	APC	CB38	BD Pharmingen
4	CD36	PerCPCy5.5	CB38	eBioscience
5	CD235a (Glycophorin A)	Pac Blue	HIR2	Biolegend
6	CD11b (adhesion molecule)	PeCy7	ICRF44	eBioscience
7	CD14 (glycoprotein)	FITC	M5E2	BD
8	CD14	PeCy7	61D3	eBioscience
9	CD71 (transferrin receptor)	APC	OKT9	eBioscience
10	CD71	FITC	OKT9	eBioscience

11	CD41a (GPIIb)	AF405	HIP8	eBioscience
12	CD105 (endoglin)	PE	43A3	eBioscience
13	CD15 (Lewis X)	FITC	H198	eBioscience
14	CD16 (Fb gammaRIII)	FITC	CB16	eBioscience
15	CD56 (NCAM)	FITC	CB56	eBioscience
16	CD10 (glycoprotein)	FITC	CB-CALLA	eBioscience
17	CD19 (glycoprotein)	FITC	HIB19	eBioscience
18	CD3 (subunit of the TCR complex)	FITC	OKT3	eBioscience
19	Lineage cocktail (CD 2/3/56/14/16/19/235a)	APC	RPA-2.10/ OKT3/ CB56/ 61D3/ CB16/ HIB19 /HIT2	eBioscience
20	Lineage cocktail (CD 2/3/56/14/16/19/235a)	FITC	RPA-2.10/ OKT3/ CB56/ 61D3/ CB16/ HIB19 /HIT2	eBioscience

21	CD38 (type II transmembrane protein)	AlexaFlour 700	HIT2	eBioscience
22	CD38	Pac Blue	HIT2	Exbio
23	CD45RA	APC-eFluor 780	HI100	eBioscience
24	CD123 (IL-3Rα)	PE	7G3	BD Biosciences
25	CD123	BV605	7G3	BD Horizon
26	CD90 (Thy-1)	PerCP-eFluor 710	THY1/5E10	eBioscience
27	CD90	BV711	eBio5E10	BD Biosciences
28	CD61 (GPIIIa)	FITC	VI-PL2	BD Pharmingen
29	CD10 (glycoprotein)	APC	5N5C	eBioscience
30	Fixable Viability Stain 510	BV510	N/A	BD Horizon
31	Annexin V	APC	N/A	eBioscience
32	CD45	AlexaFlour 700	30-F11	ebioscience

7.4 Appendix D. RT-PCR Taqman Gene Expression ID

Gene expression by RT-PCR was performed using Taqman Gene Expression Assays (Life Technologies) containing validated probes, as detailed below.

Table 7.1 Taqman Gene Expression Arrays

Gene name	**Probe name/Catalog no.**
GAPDH	Hs02758991_g1
β-actin (ACTB)	4333762T
CD36	Hs01567185_m1
GATA-1	Hs01085823_m1
GATA-2	Hs00231119_m1

8 REFERENCES

Abkowitz, J.L. et al., 2002. Response of Diamond-Blackfan anemia to metoclopramide: evidence for a role for prolactin in erythropoiesis. *Blood*, 100(8), pp.2687–2691.

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

Aerbajinai, W. et al., 2009. SCF induces gamma-globin gene expression by regulating downstream transcription factor COUP-TFII. *Blood*, 114(1), pp.187–194.

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

Alessandri, A.J. et al., 2000. Diamond-blackfan anemia and cyclosporine therapy revisited. *Journal of pediatric hematology/oncology*, 22(2), pp.176–179.

Alkhunaizi, E. et al., 2017. Novel 3q27.2-qter deletion in a patient with Diamond-Blackfan anemia and immunodeficiency: Case report and review of literature. *American journal of medical genetics. Part A*, 173(6), pp.1514–1520.

Alter, B.P. & Young, N.S., 2003. *Inherited bone marrow failure syndromes.* In D. Nathan et al., eds. *Nathan and Oski's Hematology of Infancy and Childhood*. Philadelphia, pp. 280–365.

Alter, B.P. et al., 2009. Cancer in dyskeratosis congenita. *Blood*, 113(26), pp.6549–6557.

Alter, B.P. et al., 1992. Effect of stem cell factor on in vitro erythropoiesis in patients with bone marrow failure syndromes. *Blood*, 80(12), pp.3000–3008.

Alvarez-Dominguez, J.R. et al., 2014. Global discovery of erythroid long noncoding RNAs reveals novel regulators of red cell maturation. *Blood*, 123(4), pp.570–581.

Alvarez-Dominguez, J.R., Zhang, X. & Hu, W., 2017. Widespread and dynamic translational control of red blood cell development. *Blood*, 129(5), pp.619–629.

Angelillo-Scherrer, A. et al., 2008. Role of Gas6 in erythropoiesis and anemia in mice. *The Journal of clinical investigation*, 118(2), pp.583–596.

Angelini, M. et al., 2007. Missense mutations associated with Diamond–Blackfan anemia affect the assembly of ribosomal protein S19 into the ribosome. *Human Molecular Genetics*, 16(14), pp.1720–1727.

Anur, P., Nemecek, E.R. & Kurre, P., 2009. The evolving spectrum of “non-classical” Diamond-Blackfan anaemia--a case of eADA positive pancytopenia in a young adult. *British Journal of Haematology*, 145(3), pp.428–430.

Arbiv, O.A. et al., 2017. Molecular analysis and genotype-phenotype correlation of Diamond-Blackfan anemia. *Clinical Genetics*.

Aspesi, A. et al., 2017. Lymphoblastoid cell lines from Diamond Blackfan anaemia patients exhibit a full ribosomal stress phenotype that is rescued by gene therapy. *Scientific reports*, 7(1), p.12010.

August, C. & King, E., 2005. Establishment of Erythropoiesis Following Bone Marrow Transplantation in a Patient with Congenital Hypoplastic Anemia. pp.1–9.

Bagnara, G.P. et al., 1991. In vitro growth and regulation of bone marrow enriched CD34+ hematopoietic progenitors in Diamond-Blackfan anemia. *Blood*, 78(9), pp.2203–2210.

Ball, S., 2011. Diamond Blackfan anemia. *Hematology / the Education Program of the American Society of Hematology. American Society of Hematology. Education Program*, 2011(1), pp.487–491.

Ball, S.E. et al., 1996. Diamond-Blackfan anaemia in the U.K.: analysis of 80 cases from a 20-year birth cohort. *British Journal of Haematology*, 94(4), pp.645–653.

Ball, S.E. et al., 1995. Is there a role for interleukin-3 in Diamond-Blackfan anaemia? Results of a European multicentre study. *British Journal of Haematology*, 91(2), pp.313–318.

Bauer, A. et al., 1999. The glucocorticoid receptor is required for stress erythropoiesis. *Genes & Development*, 13(22), pp.2996–3002.

Berdoukas, V. et al., 2013. Tissue iron evaluation in chronically transfused children shows significant levels of iron loading at a very young age. *American Journal of Hematology*, pp.n/a–n/a.

Bernini, J.C., Carrillo, J.M. & Buchanan, G.R., 1995. High-dose intravenous methylprednisolone therapy for patients with Diamond-Blackfan anemia refractory to conventional doses of prednisone. *The Journal of pediatrics*, 127(4), pp.654–659.

Bibikova, E. et al., 2014. TNF-mediated inflammation represses GATA1 and activates p38 MAP kinase in RPS19-deficient hematopoietic progenitors. *Blood*, 124(25), pp.3791–3798.

Biddie, S.C. & Hager, G.L., 2009. Glucocorticoid receptor dynamics and gene regulation. *Stress (Amsterdam, Netherlands)*, 12(3), pp.193–205.

Biddie, S.C. et al., 2011. Transcription factor AP1 potentiates chromatin accessibility and glucocorticoid receptor binding. *Molecular Cell*, 43(1), pp.145–155.

Blau, C.A. et al., 1993. Fetal hemoglobin in acute and chronic states of erythroid expansion. *Blood*, 81(1), pp.227–233.

Boehm, D. et al., 2013. Caspase-3 Is Involved in the Signalling in Erythroid Differentiation by Targeting Late Progenitors Z. Ivanovic, ed. *PloS one*, 8(5), p.e62303.

Boria, I. et al., 2008. A new database for ribosomal protein genes which are mutated in Diamond-Blackfan Anemia. M. Greenblatt & S. Tavtigian, eds. *Human Mutation*, 29(11), pp.E263–70.

Boria, I., et al., 2010. The ribosomal basis of Diamond-Blackfan Anemia: mutation and database update. *Human Mutation*, 31(12), pp.1269–1279.

Boultwood, J. et al., 2013. Activation of the mTOR pathway by the amino acid (L)-leucine in the 5q- syndrome and other ribosomopathies. *Advances in biological regulation*, 53(1), pp.8–17.

Breems, D.A., Blokland, E.A. & Ploemacher, R.E., 1997. Stroma-conditioned media improve expansion of human primitive hematopoietic stem cells and progenitor cells. *Leukemia*, 11(1), pp.142–150.

Calvert, R.J. & Robson, T., 1956. Cortisone therapy in erythrogenesis imperfecta. *Archives of disease in childhood*, 31(157), pp.177–181.

Campagnoli, M.F. et al., 2004. Molecular basis of Diamond-Blackfan anemia: new findings from the Italian registry and a review of the literature. *Haematologica*, 89(4), pp.480–489.

Cantor, A.B. & Orkin, S.H., 2002. Transcriptional regulation of erythropoiesis: an affair involving multiple partners. *Oncogene*, 21(21), pp.3368–3376.

Cario, H. et al., 1999. A microdeletion syndrome due to a 3-Mb deletion on 19q13.2 - Diamond-Blackfan anemia associated with macrocephaly, hypotonia, and psychomotor retardation. *Clinical Genetics*, 55(6), pp.487–492.

Carlston, C.M. et al., 2017. Variable expressivity and incomplete penetrance in a large family with non-classical Diamond-Blackfan anemia associated with ribosomal protein L11splicing variant. *American journal of medical genetics. Part A*, 173(10), pp.2622–2627.

Casadevall, N. et al., 1994. Age-related alterations in erythroid and granulopoietic progenitors in Diamond-Blackfan anaemia. *British Journal of Haematology*, 87(2), pp.369–375.

Cathie, I.A.B., 1950. Erythrogenesis imperfecta. *Archives of disease in childhood*, 25(124), pp.313–324.

Chan, H.S., Saunders, E.F. & Freedman, M.H., 1982a. Diamond-Blackfan syndrome. I. Erythropoiesis in prednisone responsive and resistant disease. *Pediatric research*, 16(6), pp.474–476.

Chan, H.S., Saunders, E.F. & Freedman, M.H., 1982b. Diamond-Blackfan syndrome. II. In vitro corticosteroid effect on erythropoiesis. *Pediatric research*, 16(6), pp.477–478.

Chang, T.J. et al., 1993. Inhibition of mouse GATA-1 function by the glucocorticoid receptor: possible mechanism of steroid inhibition of erythroleukemia cell differentiation. *Molecular endocrinology (Baltimore, Md.)*, 7(4), pp.528–542.

Charbord, P. et al., 1996. The purification of CD34 cells from human cord blood: comparison of separation techniques and cytokine requirements for optimal growth of clonogenic progenitors. *British Journal of Haematology*, 94(3), pp.449–454.

Chatr-Aryamontri, A. et al., 2004. Nonsense-mediated and nonstop decay of ribosomal protein S19 mRNA in Diamond-Blackfan anemia. *Human Mutation*, 24(6), pp.526–533.

Chen, J.-J., 2014. Translational control by heme-regulated eIF2α kinase during erythropoiesis. *Current Opinion in Hematology*, 21(3), pp.172–178.

Chiocchetti, A. et al., 2005. Interactions between RPS19, mutated in Diamond-Blackfan anemia, and the PIM-1 oncoprotein. *Haematologica*, 90(11), pp.1453–1462.

Clinton, C. & Gazda, H., 2009. *Diamond-Blackfan Anemia. 2009 Jun 25 [Updated 2016 Apr 7]. In: Adam MP, Ardinger HH, Pagon RA, et al., editors. GeneReviews® [Internet]. Seattle (WA): University of Washington, Seattle; 1993-2017. Available from: https://www.ncbi.nlm.nih.gov/books/NBK7047/*,

Cmejla, R. et al., 2007. Ribosomal protein S17 gene (RPS17) is mutated in Diamond-Blackfan anemia. *Human Mutation*, 28(12), pp.1178–1182.

Cmejla, R. et al., 2000. Ribosomal Protein S19 Gene Mutations in Patients with Diamond-Blackfan Anemia and Identification of Ribosomal Protein S19 Pseudogenes. *Blood Cells, Molecules, and Diseases*, 26(2), pp.124–132.

Cmejlova, J. et al., 2006. Translational efficiency in patients with Diamond-Blackfan anemia. *Haematologica*, 91(11), pp.1456–1464.

CrÃ tien, A. et al., 2010. Genetic variants in the noncoding region of RPS19gene in Diamond-Blackfan anemia: Potential implications for phenotypic heterogeneity. *American Journal of Hematology*, 56, pp.NA–NA.

Da Costa, L., 2013. Defects in DBA: more than meets the eye. *Blood*, 122(6), pp.856–857.

Da Costa, L., 2003. Nucleolar localization of RPS19 protein in normal cells and mislocalization due to mutations in the nucleolar localization signals in 2 Diamond-Blackfan anemia patients: potential insights into pathophysiology. *Blood*, 101(>12), pp.5039–5045.

Da Costa, L. et al., 2010. Diamond-Blackfan anemia, ribosome and erythropoiesis. *Transfusion clinique et biologique*, 17(3), pp.112–119.

Dai, M.-S. & Lu, H., 2004. Inhibition of MDM2-mediated p53 ubiquitination and degradation by ribosomal protein L5. *The Journal of biological chemistry*, 279(43), pp.44475–44482.

Danilova, N. & Gazda, H.T., 2015. Ribosomopathies: how a common root can cause a tree of pathologies. *Disease models & mechanisms*, 8(9), pp.1013–1026.

Danilova, N., Sakamoto, K.M. & Lin, S., 2011. Ribosomal protein L11 mutation in zebrafish leads to haematopoietic and metabolic defects. *British Journal of Haematology*, 152(2), pp.217–228.

Danilova, N., Sakamoto, K.M. & Lin, S., 2008. Ribosomal protein S19 deficiency in zebrafish leads to developmental abnormalities and defective erythropoiesis through activation of p53 protein family. *Blood*, 112(13), pp.5228–5237.

De Keersmaecker, K. et al., 2012. Exome sequencing identifies mutation in CNOT3 and ribosomal genes RPL5 and RPL10 in T-cell acute lymphoblastic leukemia. *Nature genetics*, 45(2), pp.186–190.

De Maria, R., Testa, U., et al., 1999a. Apoptotic role of Fas/Fas ligand system in the regulation of erythropoiesis. *Blood*, 93(3), pp.796–803.

De Maria, R., Zeuner, A., et al., 1999b. Negative regulation of erythropoiesis by caspase-mediated cleavage of GATA-1. *Nature*, 401(6752), pp.489–493.

Debnath, S. et al., 2017. Lentiviral Vectors with Cellular Promoters Correct Anemia and Lethal Bone Marrow Failure in a Mouse Model for Diamond-Blackfan Anemia. *Molecular therapy : the journal of the American Society of Gene Therapy*, 25(8), pp.1805–1814.

Devlin, E.E. et al., 2010. A transgenic mouse model demonstrates a dominant negative effect of a point mutation in the RPS19 gene associated with Diamond-Blackfan anemia. *Blood*, 116(15), pp.2826–2835.

Diamond, L.K. & Blackfan, K.D., 1938. Hypoplastic Anemia. *American Journal of Diseases of Children*, 56, pp.464–467.

Dianzani, I. & Loreni, F., 2008. Diamond-Blackfan anemia: a ribosomal puzzle. *Haematologica*, 93(11), pp.1601–1604.

Dobin, A. et al., 2012. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics*, 29(1), pp.15–21.

Doherty, L. et al., 2010. Ribosomal protein genes RPS10 and RPS26 are commonly mutated in Diamond-Blackfan anemia. *American journal of human genetics*, 86(2), pp.222–228.

Dokal, I. & Vulliamy, T., 2010. Inherited bone marrow failure syndromes. *Haematologica*, 95(8), pp.1236–1240.

Dolznig, H. et al., 2002. Apoptosis protection by the Epo target Bcl-X(L) allows factor-independent differentiation of primary erythroblasts. *Current biology : CB*, 12(13), pp.1076–1085.

Dolznig, H. et al., 2006. Erythroid progenitor renewal versus differentiation: genetic evidence for cell autonomous, essential functions of EpoR, Stat5 and the GR. *Oncogene*, 25(20), pp.2890–2900.

Doré, L.C. & Crispino, J.D., 2011. Transcription factor networks in erythroid cell and megakaryocyte development. *Blood*, 118(2), pp.231–239.

Douay, L. & Giarratana, M.-C., 2009. Ex vivo generation of human red blood cells: a new advance in stem cell engineering. *Methods in molecular biology (Clifton, N.J.)*, 482(Chapter 8), pp.127–140.

Doulatov, S. et al., 2017. Drug discovery for Diamond-Blackfan anemia using reprogrammed hematopoietic progenitors. *Science translational medicine*, 9(376), p.eaah5645.

Doulatov, S. et al., 2012. Hematopoiesis: a human perspective. *Cell stem cell*, 10(2), pp.120–136.

Draptchinskaia, N. et al., 1999. The gene encoding ribosomal protein S19 is mutated in Diamond-Blackfan anaemia. *Nature genetics*, 21(2), pp.169–175.

Dunbar, A.E., Moore, S.L. & Hinson, R.M., 2003. Fetal Diamond-Blackfan anemia associated with hydrops fetalis. *American journal of perinatology*, 20(7), pp.391–394.

Dutt, S. et al., 2011. Haploinsufficiency for ribosomal protein genes causes selective activation of p53 in human erythroid progenitor cells. *Blood*, 117(9), pp.2567–2576.

Dzierzak, E. & Philipsen, S., 2013. Erythropoiesis: development and differentiation.

Cold Spring Harbor perspectives in medicine, 3(4), pp.a011601–a011601.

Ear, J. et al., 2015. RAP-011 improves erythropoiesis in zebrafish model of Diamond-Blackfan anemia through antagonizing lefty1. *Blood*, 126(7), pp.880–890.

Ebert, B.L., 2005. An RNA interference model of RPS19 deficiency in Diamond-Blackfan anemia recapitulates defective hematopoiesis and rescue by dexamethasone: identification of dexamethasone-responsive genes by microarray. *Blood*, 105(12), pp.4620–4626.

Ebert, B.L. et al., 2008. Identification of RPS14 as a 5q- syndrome gene by RNA interference screen. *Nature*, 451(7176), pp.335–339.

Edvardsson, L., Dykes, J. & Olofsson, T., 2006. Isolation and characterization of human myeloid progenitor populations--TpoR as discriminator between common myeloid and megakaryocyte/erythroid progenitors. *Experimental Hematology*, 34(5), pp.599–609.

Fagioli, F. et al., 2014. Haematopoietic stem cell transplantation for Diamond Blackfan anaemia: a report from the Italian Association of Paediatric Haematology and Oncology Registry. *British Journal of Haematology*, 165(5), pp.673–681.

Falchi, M. et al., 2015. Dexamethasone targeted directly to macrophages induces macrophage niches that promote erythroid expansion. *Haematologica*, 100(2), pp.178–187.

Fargo, J.H. et al., 2012. Erythrocyte adenosine deaminase: diagnostic value for Diamond-Blackfan anaemia. *British Journal of Haematology*, 160(4), pp.547–554.

Farrar, J.E. & Dahl, N., 2011. Untangling the Phenotypic Heterogeneity of Diamond Blackfan Anemia. *YSHEM*, 48(2), pp.124–135.

Farrar, J.E. et al., 2008. Abnormalities of the large ribosomal subunit protein, Rpl35a, in Diamond-Blackfan anemia. *Blood*, 112(5), pp.1582–1592.

Farrar, J.E. et al., 2017. Response: Making “perfect” the enemy of good. *Blood*, 130(9), pp.1168–1169.

Farrar, J.E. et al., 2011. Ribosomal protein gene deletions in Diamond-Blackfan anemia. *Blood*, 118(26), pp.6943–6951.

Ferreira, R. et al., 2005. GATA1 function, a paradigm for transcription factors in hematopoiesis. *Molecular and cellular biology*, 25(4), pp.1215–1227.

Ferretti, M.B. et al., 2017. Rps26 directs mRNA-specific translation by recognition of Kozak sequence elements. *Nature structural & molecular biology*, 24(9), pp.700–

707.

Flores Ballester, E. et al., 2015. Adult-onset Diamond-Blackfan anemia with a novel mutation in the exon 5 of RPL11: too late and too rare. *Clinical Case Reports*, 3(6), pp.392–395.

Flygare, J., 2005. Deficiency of ribosomal protein S19 in CD34+ cells generated by siRNA blocks erythroid development and mimics defects seen in Diamond-Blackfan anemia. *Blood*, 105(12), pp.4627–4634.

Flygare, J. et al., 2011. HIF1 synergizes with glucocorticoids to promote BFU-E progenitor self-renewal. *Blood*, 117(12), pp.3435–3444.

Frede, S. et al., 1997. Erythropoietin gene expression is suppressed after lipopolysaccharide or interleukin-1 beta injections in rats. *The American journal of physiology*, 273(3 Pt 2), pp.R1067–71.

Freedman, M.H. & Saunders, E.F., 1978. Diamond-Blackfan syndrome: evidence against cell-mediated erythropoietic suppression. *Blood*, 51(6), pp.1125–1128.

Freedman, M.H., Amato, D. & Saunders, E.F., 1976. Erythroid colony growth in congenital hypoplastic anemia. *The Journal of clinical investigation*, 57(3), pp.673–677.

Freyssinier, J.M. et al., 1999. Purification, amplification and characterization of a population of human erythroid progenitors. *British Journal of Haematology*, 106(4), pp.912–922.

Frischmeyer, P.A. et al., 2002. An mRNA surveillance mechanism that eliminates transcripts lacking termination codons. *Science*, 295(5563), pp.2258–2261.

Frontelo, P. et al., 2007. Novel role for EKLF in megakaryocyte lineage commitment. *Blood*, 110(12), pp.3871–3880.

Fumagalli, S. et al., 2009. Absence of nucleolar disruption after impairment of 40S ribosome biogenesis reveals an rpL11-translation-dependent mechanism of p53 induction. *Nature cell biology*, 11(4), pp.501–508.

Gabbianelli, M., 2002. HbF reactivation in sibling BFU-E colonies: synergistic interaction of kit ligand with low-dose dexamethasone. *Blood*, 101(7), pp.2826–2832.

Gagne, K.E. et al., 2014. Pearson marrow pancreas syndrome in patients suspected to have Diamond-Blackfan anemia. *Blood*, 124(3), pp.437–440.

Gao, X. et al., 2017. Thyroid hormone receptor beta and NCOA4 regulate terminal erythrocyte differentiation. *Proceedings of the National Academy of Sciences of*

the United States of America, 114(38), pp.10107–10112.

García-Vela, J.A. et al., 2017. CD105 expression in early erythroid precursors. *American Journal of Hematology*, 92(8), pp.E155–E156.

Garçon, L. et al., 2013. Ribosomal and hematopoietic defects in induced pluripotent stem cells derived from Diamond Blackfan anemia patients. *Blood*, 122(6), pp.912–921.

Gastou, M. et al., 2017. The severe phenotype of Diamond-Blackfan anemia is modulated by heat shock protein 70. *Blood Advances*, 1(22), pp.1959–1976.

Gazda, H.T. et al., 2006. Defective ribosomal protein gene expression alters transcription, translation, apoptosis, and oncogenic pathways in Diamond-Blackfan anemia. *Stem cells (Dayton, Ohio)*, 24(9), pp.2034–2044.

Gazda, H.T. et al., 2012. Frameshift mutation in p53 regulator RPL26 is associated with multiple physical abnormalities and a specific pre-ribosomal RNA processing defect in diamond-blackfan anemia. *Human Mutation*, 33(7), pp.1037–1044.

Gazda, H.T. et al., 2008. Ribosomal protein L5 and L11 mutations are associated with cleft palate and abnormal thumbs in Diamond-Blackfan anemia patients. *American journal of human genetics*, 83(6), pp.769–780.

Gazda, H.T. et al., 2004. RNA and protein evidence for haplo-insufficiency in Diamond-Blackfan anaemia patients with RPS19 mutations. *British Journal of Haematology*, 127(1), pp.105–113.

Ge, J. et al., 2015. Dysregulation of the Transforming Growth Factor β Pathway in Induced Pluripotent Stem Cells Generated from Patients with Diamond Blackfan Anemia J. Freeman, ed. *PloS one*, 10(8), p.e0134878.

George, A.A., Schiltz, R.L. & Hager, G.L., 2009. Dynamic access of the glucocorticoid receptor to response elements in chromatin. *The international journal of biochemistry & cell biology*, 41(1), pp.214–224.

Gerrard, G. et al., 2013. Target enrichment and high-throughput sequencing of 80 ribosomal protein genes to identify mutations associated with Diamond-Blackfan anaemia. *British Journal of Haematology*, 162(4), pp.530–536.

Gillio, A.P. et al., 1993. Treatment of Diamond-Blackfan anemia with recombinant human interleukin-3. *Blood*, 82(3), pp.744–751.

Giri, N. et al., 2000. Clinical and laboratory evidence for a trilineage haematopoietic defect in patients with refractory Diamond-Blackfan anaemia. *British Journal of Haematology*, 108(1), pp.167–175.

Glader, B.E., Backer, K. & Diamond, L.K., 1983. Elevated erythrocyte adenosine deaminase activity in congenital hypoplastic anemia. *The New England journal of medicine*, 309(24), pp.1486–1490.

Gregory, C.J. & Eaves, A.C., 1977. Human marrow cells capable of erythropoietic differentiation in vitro: definition of three erythroid colony responses. *Blood*, 49(6), pp.855–864.

Gregory, C.J. & Eaves, A.C., 1978. Three stages of erythropoietic progenitor cell differentiation distinguished by a number of physical and biologic properties. *Blood*, 51(3), pp.527–537.

Griffin, B.H., Chitty, L.S. & Bitner-Glindzicz, M., 2017. The 100 000 Genomes Project: What it means for paediatrics. *Archives of disease in childhood. Education and practice edition*, 102(2), pp.105–107.

Griffiths, R.E. et al., 2012. Maturing reticulocytes internalize plasma membrane in glycophorin A-containing vesicles that fuse with autophagosomes before exocytosis. *Blood*, 119(26), pp.6296–6306.

Gripp, K.W. et al., 2014. Diamond-Blackfan anemia with mandibulofacial dystostosis is heterogeneous, including the novel DBA genes TSR2 and RPS28. *American journal of medical genetics. Part A*, 164A(9), pp.2240–2249.

Grover, A. et al., 2014. Erythropoietin guides multipotent hematopoietic progenitor cells toward an erythroid fate. *The Journal of experimental medicine*, 211(2), pp.181–188.

Guo, Y. et al., 2015. Histone demethylase LSD1-mediated repression of GATA-2 is critical for erythroid differentiation. *Drug design, development and therapy*, 9, pp.3153–3162.

Gustavsson, P. et al., 1998. Identification of Microdeletions Spanning the Diamond-Blackfan Anemia Locus on 19q13 and Evidence for Genetic Heterogeneity. *The American Journal of Human Genetics*, 63(5), pp.1388–1395.

Haase, V.H., 2013. Regulation of erythropoiesis by hypoxia-inducible factors. *Blood reviews*, 27(1), pp.41–53.

Haddad, R. et al., 2004. Molecular characterization of early human T/NK and B-lymphoid progenitor cells in umbilical cord blood. *Blood*, 104(13), pp.3918–3926.

Halperin, D.S., Estrov, Z. & Freedman, M.H., 1989. Diamond-Blackfan anemia: promotion of marrow erythropoiesis in vitro by recombinant interleukin-3. *Blood*, 73(5), pp.1168–1174.

Hamaguchi, I., 2003. Proliferation deficiency of multipotent hematopoietic

progenitors in ribosomal protein S19 (RPS19)-deficient diamond–Blackfan anemia improves following RPS19 gene transfer. *Molecular Therapy*, 7(5), pp.613–622.

Hamaguchi, I. et al., 2002. Gene transfer improves erythroid development in ribosomal protein S19-deficient Diamond-Blackfan anemia. *Blood*, 100(8), pp.2724–2731.

Hanspal, M., 1997. Importance of cell-cell interactions in regulation of erythropoiesis. *Current Opinion in Hematology*, 4(2), pp.142–147.

Hattangadi, S.M. et al., 2011. From stem cell to red cell: regulation of erythropoiesis at multiple levels by multiple proteins, RNAs, and chromatin modifications. *Blood*, 118(24), pp.6258–6268.

Hoffman, R. et al., 1976. Diamond-Blackfan syndrome: lymphocyte-mediated suppression of erythropoiesis. *Science*, 193(4256), pp.899–900.

Hofman, I.J.F. et al., 2017. Low frequency mutations in ribosomal proteins RPL10 and RPL5 in multiple myeloma. *Haematologica*, 102(8), pp.e317–e320.

Hollanda, L.M. et al., 2006. An inherited mutation leading to production of only the short isoform of GATA-1 is associated with impaired erythropoiesis. *Nature genetics*, 38(7), pp.807–812.

Hom, J. et al., 2015. The erythroblastic island as an emerging paradigm in the anemia of inflammation. *Immunologic research*, 63(1-3), pp.75–89.

Horos, R. & Lindern, M, 2012. Molecular mechanisms of pathology and treatment in Diamond Blackfan Anaemia. *British Journal of Haematology*.

Huang, X. et al., 2010. Erythropoietin receptor signaling regulates both erythropoiesis and megakaryopoiesis in vivo. *Blood Cells, Molecules, and Diseases*, 44(1), pp.1–6.

Huyhn, A. et al., 1995. Characterization of hematopoietic progenitors from human yolk sacs and embryos. *Blood*, 86(12), pp.4474–4485.

Ichimura, T. et al., 2016. Diagnostic challenge of Diamond–Blackfan anemia in mothers and children by whole-exome sequencing. *International Journal of Hematology*, 105(4), pp.515–520.

Iskander, D., 2012. Patients with Diamond-Blackfan Anaemia have Abnormalities of Cellular and Humoral Immunity. In American Society of Haematology.

Iskander, D. et al., 2015. Elucidation of the EP defect in Diamond-Blackfan anemia by characterization and prospective isolation of human EPs. *Blood*, 125(16),

pp.2553–2557.

Iudicibus, S.D., 2011. Molecular mechanism of glucocorticoid resistance in inflammatory bowel disease. *World Journal of Gastroenterology*, 17(9), p.1095.

Jaako, P. et al., 2012. Dietary L-leucine improves the anemia in a mouse model for Diamond-Blackfan anemia. *Blood*, 120(11), pp.2225–2228.

Jaako, P. et al., 2011. Mice with ribosomal protein S19 deficiency develop bone marrow failure and symptoms like patients with Diamond-Blackfan anemia. *Blood*, 118(23), pp.6087–6096.

Jabr, F.I. & Taher, A., 2006. Diamond-Blackfan anemia in remission for 2 years on valproic acid. *Haematologica*, 91(6 Suppl), p.ELT05.

John, S. et al., 2011. Chromatin accessibility pre-determines glucocorticoid receptor binding patterns. *Nature genetics*, 43(3), pp.264–268.

Juli, G. et al., 2016. Depletion of ribosomal protein S19 causes a reduction of rRNA synthesis. *Scientific reports*, 6(1), p.35026.

Kaspers, G.J.L. et al., 2009. Glucocorticoid Resistance in Childhood Leukemia. *Leukemia & Lymphoma*, 13(3-4), pp.187–201.

Klimchenko, O. et al., 2009. A common bipotent progenitor generates the erythroid and megakaryocyte lineages in embryonic stem cell-derived primitive hematopoiesis. *Blood*, 114(8), pp.1506–1517.

Kolbus, A., 2003. Cooperative signaling between cytokine receptors and the glucocorticoid receptor in the expansion of erythroid progenitors: molecular analysis by expression profiling. *Blood*, 102(9), pp.3136–3146.

Kongsuwan, K. et al., 1985. A Drosophila Minute gene encodes a ribosomal protein. *Nature*, 317(6037), pp.555–558.

Kurita, R. et al., 2013. Establishment of Immortalized Human Erythroid Progenitor Cell Lines Able to Produce Enucleated Red Blood Cells D. T. Covas, ed. *PloS one*, 8(3), p.e59890.

Lahoti, A. et al., 2016. Endocrine Dysfunction in Diamond-Blackfan Anemia (DBA): A Report from the DBA Registry (DBAR). *Pediatric Blood & Cancer*, 63(2), pp.306–312.

Landau, D.A. et al., 2015. Mutations driving CLL and their evolution in progression and relapse. *Nature*, 526(7574), pp.525–530.

Landowski, M. et al., 2013. Novel deletion of RPL15 identified by array-comparative

genomic hybridization in Diamond-Blackfan anemia. *Human genetics*, 132(11), pp.1265–1274.

Leberbauer, C., 2005. Different steroids co-regulate long-term expansion versus terminal differentiation in primary human erythroid progenitors. *Blood*, 105(1), pp.85–94.

Leblanc, T.M. et al., 2007. Metoclopramide treatment in DBA patients: no complete response in a French prospective study. *Blood*, 109(5), pp.2266–2267.

Lee, H.-Y. et al., 2015. PPAR-α and glucocorticoid receptor synergize to promote erythroid progenitor self-renewal. *Nature*, 522(7557), pp.474–477.

Li, J. et al., 2014. Isolation and transcriptome analyses of human erythroid progenitors: BFU-E and CFU-E. *Blood*, 124(24), pp.3636–3645.

Liao, Y., Smyth, G.K. & Shi, W., 2013. The Subread aligner: fast, accurate and scalable read mapping by seed-and-vote. *Nucleic acids research*, 41(10), pp.e108–e108.

Libregts, S.F. et al., 2011. Chronic IFN-γ production in mice induces anemia by reducing erythrocyte life span and inhibiting erythropoiesis through an IRF-1/PU.1 axis. *Blood*, 118(9), pp.2578–2588.

Lim, H.-W. et al., 2015. Genomic redistribution of GR monomers and dimers mediates transcriptional response to exogenous glucocorticoid in vivo. *Genome Research*, 25(6), pp.836–844.

Lindern, von, M et al., 2001. Leukemic transformation of normal murine erythroid progenitors: v- and c-ErbB act through signaling pathways activated by the EpoR and c-Kit in stress erythropoiesis. *Oncogene*, 20(28), pp.3651–3664.

Lindern, von, M et al., 1999. The glucocorticoid receptor cooperates with the erythropoietin receptor and c-Kit to enhance and sustain proliferation of erythroid progenitors in vitro. *Blood*, 94(2), pp.550–559.

Lipton, J.M. et al., 1986. Defective erythroid progenitor differentiation system in congenital hypoplastic (Diamond-Blackfan) anemia. *Blood*, 67(4), pp.962–968.

Lipton, J.M. et al., 2006. Improving clinical care and elucidating the pathophysiology of Diamond Blackfan anemia: An update from the Diamond Blackfan Anemia Registry. *Pediatric Blood & Cancer*, 46(5), pp.558–564.

Liu, Y. et al., 2006. Suppression of Fas-FasL coexpression by erythropoietin mediates erythroblast expansion during the erythropoietic stress response in vivo. *Blood*, 108(1), pp.123–133.

Love, M.I., Huber, W. & Anders, S., 2014. Moderated estimation of fold change and

dispersion for RNA-seq data with DESeq2. *Genome biology*, 15(12), p.31.

Lucarelli, G., 2008. Bone marrow transplanataion in patients with thalassemia. *The New England journal of medicine*, pp.1–5.

Ludwig, L.S. et al., 2014. Altered translation of GATA1 in Diamond-Blackfan anemia. *Nature Medicine*, pp.1–9.

Luzzatto, L. & Karadimitris, A., 1998. Dyskeratosis and ribosomal rebellion. *Nature genetics*, 19(1), pp.6–7.

Luzzatto, L. & Karadimitris, A., 2010. The Molecular Basis of Anemia. In *Molecular Hematology*. Oxford, UK: Wiley-Blackwell, pp. 140–164.

Macdougall, I.C. & Cooper, A.C., 2005. Hyporesponsiveness to erythropoietic therapy due to chronic inflammation. *European journal of clinical investigation*, 35 Suppl 3(s3), pp.32–35.

Majeti, R., Park, C.Y. & Weissman, I.L., 2007. Identification of a hierarchy of multipotent hematopoietic progenitors in human cord blood. *Cell stem cell*, 1(6), pp.635–645.

Manz, M.G. et al., 2002. Prospective isolation of human clonogenic common myeloid progenitors. *Proceedings of the National Academy of Sciences*, 99(18), pp.11872–11877.

Marygold, S.J. et al., 2007. The ribosomal protein genes and Minute loci of Drosophila melanogaster. *Genome biology*, 8(10), p.R216.

Matsson, H. et al., 2004. Targeted disruption of the ribosomal protein S19 gene is lethal prior to implantation. *Molecular and cellular biology*, 24(9), pp.4032–4037.

McGowan, K.A. & Mason, P.J., 2011. Animal Models of Diamond Blackfan Anemia. *Seminars in Hematology*, 48(2), pp.106–116.

McGowan, K.A. et al., 2008. Ribosomal mutations cause p53-mediated dark skin and pleiotropic effects. *Nature genetics*, 40(8), pp.963–970.

McLennan, A.C. et al., 1996. Prenatal diagnosis of Blackfan-Diamond syndrome: case report and review of the literature. *Prenatal diagnosis*, 16(4), pp.349–353.

Medvinsky, A. & Dzierzak, E., 1996. Definitive hematopoiesis is autonomously initiated by the AGM region. *Cell*, 86(6), pp.897–906.

Merryweather-Clarke, A.T. et al., 2011. Global gene expression analysis of human erythroid progenitors. *Blood*, 117(13), pp.e96–108.

Metcalf, D., 2008. Hematopoietic cytokines. *Blood*, 111(2), pp.485–491.

Migliaccio, A.R. & Varricchio, L., 2017. Advanced Cell Culture Models for Diamond Blackfan Anemia and other Erythroid Disorders. *Stem cells (Dayton, Ohio)*.

Migliaccio, G. et al., 2011. Under HEMA conditions, self-replication of human erythroblasts is limited by autophagic death. *Blood Cells, Molecules, and Diseases*, 47(3), pp.182–197.

Mills, E.W. & Green, R., 2017. Ribosomopathies: There's strength in numbers. *Science*, 358(6363), p.eaan2755.

Mirabello, L. et al., 2017. Novel and known ribosomal causes of Diamond-Blackfan anaemia identified through comprehensive genomic characterisation. *Journal of Medical Genetics*, pp.jmedgenet–2016–104346.

Mirabello, L. et al., 2014. Whole-exome sequencing and functional studies identify RPS29 as a novel gene mutated in multicase Diamond-Blackfan anemia families. *Blood*, 124(1), pp.24–32.

Miyawaki, K. et al., 2017. Identification of unipotent megakaryocyte progenitors in human hematopoiesis. *Blood*, 129(25), pp.3332–3343.

Moniz, H. et al., 2012. Primary hematopoietic cells from DBA patients with mutations in RPL11 and RPS19 genes exhibit distinct erythroid phenotype in vitro. *Cell death & disease*, 3(7), p.e356.

Moore, J.B. et al., 2010. Distinct ribosome maturation defects in yeast models of Diamond-Blackfan anemia and Shwachman-Diamond syndrome. *Haematologica*, 95(1), pp.57–64.

Mori, Y. et al., 2015. Prospective isolation of human erythroid lineage-committed progenitors. *Proceedings of the National Academy of Sciences of the United States of America*, 112(31), pp.9638–9643.

Muir, C., Dodds, A. & Samaras, K., 2017. Mid-life extra-haematopoetic manifestations of Diamond-Blackfan anaemia. *Endocrinology, diabetes & metabolism case reports*, 2017.

Muta, K. et al., 1995. Stem cell factor retards differentiation of normal human erythroid progenitor cells while stimulating proliferation. *Blood*, 86(2), pp.572–580.

Nandakumar, S.K., Ulirsch, J.C. & Sankaran, V.G., 2016. Advances in understanding erythropoiesis: evolving perspectives. *British Journal of Haematology*, 173(2), pp.206–218.

Narla, A. & Ebert, B.L., 2010. Ribosomopathies: human disorders of ribosome dysfunction. *Blood*.

Narla, A. et al., 2011. Dexamethasone and lenalidomide have distinct functional effects on erythropoiesis. *Blood*, 118(8), pp.2296–2304.

Narla, A. et al., 2016. Erythrocyte adenosine deaminase levels are elevated in Diamond Blackfan anemia but not in the 5q− syndrome. *American Journal of Hematology*, 91(12), pp.E501–E502.

Nathan, D.G., Clarke, B.J., et al., 1978a. Erythroid precursors in congenital hypoplastic (Diamond-Blackfan) anemia. *The Journal of clinical investigation*, 61(2), pp.489–498.

Nathan, D.G., Hillman, D.G., et al., 1978b. Normal erythropoietic helper T cells in congenital hypoplastic (Diamond-Blackfan) anemia. *The New England journal of medicine*, 298(19), pp.1049–1051.

Nicolas, E. et al., 2016. Involvement of human ribosomal proteins in nucleolar structure and p53-dependent nucleolar stress. *Nature communications*, 7, p.11390.

Niemeyer, C.M. et al., 1991. Treatment trial with recombinant human erythropoietin in children with congenital hypoplastic anemia. *Contributions to nephrology*, 88, pp.276–80– discussion 281.

Notta, F. et al., 2016. Distinct routes of lineage development reshape the human blood hierarchy across ontogeny. *Science*, 351(6269), pp.aab2116–aab2116.

O 039 Boyle, F. et al., 2016. Haemopoietic Stem Cell Transplantation for Diamond Blackfan Anaemia Leads to Early and Sustained Engraftment with Good Long-Term Outcomes, but Has an Increased Risk of Gut Toxicity and Lung GvHD. *Blood*, 128(22), p.2679.

O'Brien, K.A. et al., 2017. Molecular convergence in ex vivo models of Diamond-Blackfan anemia. *Blood*, 129(23), pp.3111–3120.

Ohene-Abuakwa, Y. et al., 2005. Two-phase culture in Diamond Blackfan anemia: localization of erythroid defect. *Blood*, 105(2), pp.838–846.

Olivieri, N.F. et al., 1994. Failure of recombinant human interleukin-3 therapy to induce erythropoiesis in patients with refractory Diamond-Blackfan anemia. *Blood*, 83(9), pp.2444–2450.

Olopade, O.I. et al., 1992. Clinical, morphologic, and cytogenetic characteristics of 26 patients with acute erythroblastic leukemia. *Blood*, 80(11), pp.2873–2882.

Orkin, S.H., 2000. Diversification of haematopoietic stem cells to specific lineages. *Nature Reviews Genetics*, 1(1), pp.57–64.

Ottersbach, K. et al., 2010. Ontogeny of haematopoiesis: recent advances and open questions. *British Journal of Haematology*, 148(3), pp.343–355.

Ozden, F.O. et al., 2011. Oral and dental manifestations of diamond-blackfan anemia: case reports. *European journal of dentistry*, 5(3), pp.344–348.

Palis, J., 2017. Interaction of the Macrophage and Primitive Erythroid Lineages in the Mammalian Embryo. *Frontiers in Immunology*, 7, p.444.

Pang, W.W. et al., 2013. Hematopoietic stem cell and progenitor cell mechanisms in myelodysplastic syndromes. *Proceedings of the National Academy of Sciences of the United States of America*, 110(8), pp.3011–3016.

Pang, W.W., Schrier, S.L. & Weissman, I.L., 2017. Age-associated changes in human hematopoietic stem cells. *Seminars in Hematology*, 54(1), pp.39–42.

Paolini, N.A. et al., 2017. A Ribosomopathy Reveals Decoding Defective Ribosomes Driving Human Dysmorphism. *American journal of human genetics*, 100(3), pp.506–522.

Parrella, S. et al., 2014. Loss of GATA-1 full length as a cause of Diamond-Blackfan anemia phenotype. *Pediatric Blood & Cancer*, 61(7), pp.1319–1321.

Paugh, S.W. et al., 2015. NALP3 inflammasome upregulation and CASP1 cleavage of the glucocorticoid receptor cause glucocorticoid resistance in leukemia cells. *Nature genetics*, 47(6), pp.607–614.

Payne, E.M. et al., 2012. L-leucine improves the anemia and developmental defects associated with Diamond-Blackfan anemia and del(5q) MDS by activating the mTOR pathway. *Blood*, 120(11), pp.2214–2224.

Perdahl, E.B. et al., 1994. Erythroid failure in Diamond-Blackfan anemia is characterized by apoptosis. *Blood*, 83(3), pp.645–650.

Pereira, C., Clarke, E. & Jackie, D.J., 2007. *Stem Cell Assays* M. Vemuri, ed., Totowa, NJ.

Pesciotta, E.N. et al., 2014. Dysferlin and other non-red cell proteins accumulate in the red cell membrane of Diamond-Blackfan Anemia patients. S. R. Ellis, ed. *PloS one*, 9(1), p.e85504.

Pesciotta, E.N. et al., 2015. In-Depth, Label-Free Analysis of the Erythrocyte Cytoplasmic Proteome in Diamond Blackfan Anemia Identifies a Unique Inflammatory Signature. T. Rouault, ed. *PloS one*, 10(10), p.e0140036.

Pishesha, N. et al., 2014. Transcriptional divergence and conservation of human and mouse erythropoiesis. *Proceedings of the National Academy of Sciences of the United States of America*, 111(11), pp.4103–4108.

Porter, J.B. et al., 2014. Mechanisms of plasma non-transferrin bound iron generation: insights from comparing transfused diamond blackfan anaemia with sickle cell and thalassaemia patients. *British Journal of Haematology*, 167(5), pp.692–696.

Pospisilova, D. et al., 2012. The Czech National Diamond-Blackfan Anemia Registry: clinical data and ribosomal protein mutations update. *Blood Cells, Molecules, and Diseases*, 48(4), pp.209–218.

Psaila, B. et al., 2016. Single-cell profiling of human megakaryocyte-erythroid progenitors identifies distinct megakaryocyte and erythroid differentiation pathways. *Genome biology*, 17(1), p.83.

Quarello, P. et al., 2010. Diamond-Blackfan anemia: genotype-phenotype correlations in Italian patients with RPL5 and RPL11 mutations. *Haematologica*, 95(2), pp.206–213.

Quarello, P. et al., 2016. Ribosomal RNA analysis in the diagnosis of Diamond-Blackfan Anaemia. *British Journal of Haematology*, 172(5), pp.782–785.

Quin, J.E. et al., 2014. Targeting the nucleolus for cancer intervention. *Biochimica et biophysica acta*, 1842(6), pp.802–816.

Ramenghi, U. et al., 1999. Diamond-Blackfan anaemia in the Italian population. *British Journal of Haematology*, 104(4), pp.841–848.

Rey, M.A. et al., 2008. Enhanced alternative splicing of the FLVCR1 gene in Diamond Blackfan anemia disrupts FLVCR1 expression and function that are critical for erythropoiesis. *Haematologica*, 93(11), pp.1617–1626.

Rhodes, M.M. et al., 2008. Adherence to macrophages in erythroblastic islands enhances erythroblast proliferation and increases erythrocyte production by a different mechanism than erythropoietin. *Blood*, 111(3), pp.1700–1708.

Ribeil, J.-A. et al., 2007. Hsp70 regulates erythropoiesis by preventing caspase-3-mediated cleavage of GATA-1. *Nature*, 445(7123), pp.102–105.

Robinson, P.N. et al., 2008. The Human Phenotype Ontology: a tool for annotating and analyzing human hereditary disease. *American journal of human genetics*, 83(5), pp.610–615.

Roy, A. et al., 2012. Perturbation of fetal liver hematopoietic stem and progenitor cell development by trisomy 21. *Proceedings of the National Academy of*

Sciences of the United States of America, 109(43), pp.17579–17584.

Roy, N.B.A. et al., 2016. A novel 33-Gene targeted resequencing panel provides accurate, clinical-grade diagnosis and improves patient management for rare inherited anaemias. *British Journal of Haematology*, 175(2), pp.318–330.

Sanada, C. et al., 2016. Adult human megakaryocyte-erythroid progenitors are in the CD34+CD38mid fraction. *Blood*, 128(7), pp.923–933.

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

Sankaran, V.G. et al., 2012. Exome sequencing identifies GATA1 mutations resulting in Diamond-Blackfan anemia. *The Journal of clinical investigation*, 122(7), pp.2439–2443.

Santucci, M.A. et al., 1999. Long-term bone marrow cultures in Diamond-Blackfan anemia reveal a defect of both granulomacrophage and erythroid progenitors. *Experimental Hematology*, 27(1), pp.9–18.

Sato, T. et al., 2000. Erythroid progenitors differentiate and mature in response to endogenous erythropoietin. *The Journal of clinical investigation*, 106(2), pp.263–270.

Schmittgen, T.D. & Livak, K.J., 2008. Analyzing real-time PCR data by the comparative C(T) method. *Nature protocols*, 3(6), pp.1101–1108.

Schwarz, J.M. et al., 2014. MutationTaster2: mutation prediction for the deep-sequencing age. *Nature Methods*, 11(4), pp.361–362.

Shibuya, Y. et al., 2001. Identification of receptor-binding sites of monocyte chemotactic S19 ribosomal protein dimer. *The American journal of pathology*, 159(6), pp.2293–2301.

Siatecka, M. & Bieker, J.J., 2011. The multifunctional role of EKLF/KLF1 during erythropoiesis. *Blood*, 118(8), pp.2044–2054.

Sieff, C. et al., 1982. Changes in cell surface antigen expression during hemopoietic differentiation. *Blood*, 60(3), pp.703–713.

Sieff, C.A. et al., 2010. Pathogenesis of the erythroid failure in Diamond Blackfan anaemia. *British Journal of Haematology*, 148(4), pp.611–622.

Singh, S.A. et al., 2014. p53-Independent cell cycle and erythroid differentiation defects in murine embryonic stem cells haploinsufficient for Diamond Blackfan anemia-proteins: RPS19 versus RPL5. T. Papayannopoulou, ed. *PloS one*, 9(2), p.e89098.

Sjögren, S.E. & Flygare, J., 2012. Progress towards Mechanism-Based Treatment for Diamond-Blackfan Anemia. *The Scientific World Journal*, 2012(2), pp.1–8.

Sjögren, S.E. et al., 2015. Glucocorticoids improve erythroid progenitor maintenance and dampen Trp53response in a mouse model of Diamond-Blackfan anaemia. *British Journal of Haematology*, 171(4), pp.517–529.

Slavov, N. et al., 2015. Differential Stoichiometry among Core Ribosomal Proteins. *Cell reports*, 13(5), pp.865–873.

So, A.Y.-L. et al., 2007. Determinants of cell- and gene-specific transcriptional regulation by the glucocorticoid receptor. *PLoS genetics*, 3(6), p.e94.

Song, M.-J. et al., 2010. A novel initiation codon mutation in the ribosomal protein S17 gene (RPS17) in a patient with Diamond-Blackfan anemia. *Pediatric Blood & Cancer*, 54(4), pp.629–631.

Soni, S. et al., 2006. Absence of erythroblast macrophage protein (Emp) leads to failure of erythroblast nuclear extrusion. *The Journal of biological chemistry*, 281(29), pp.20181–20189.

Soulet, F. et al., 2001. Fibroblast Growth Factor-2 Interacts with Free Ribosomal Protein S19. *Biochemical and Biophysical Research Communications*, 289(2), pp.591–596.

Steinberg-Shemer, O. et al., 2016. Diamond Blackfan Anemia: A Nonclassical Patient With Diagnosis Assisted by Genomic Analysis. *Journal of pediatric hematology/oncology*, 38(7), pp.e260–2.

Stellacci, E. et al., 2009. Interaction between the glucocorticoid and erythropoietin receptors in human erythroid cells. *Experimental Hematology*, 37(5), pp.559–572.

Storch, H., Krüger, W. & Rotzsch, W., 1981. Adenosine Deaminase Activity in Plasma and Blood Cells of Patients with Haematological and Autoimmune Diseases. *Acta haematologica*, 65(3), pp.183–188.

Strauss, L.C. et al., 1986. Antigenic analysis of hematopoiesis. V. Characterization of My-10 antigen expression by normal lymphohematopoietic progenitor cells. *Experimental Hematology*, 14(9), pp.878–886.

Suragani, R.N.V.S. et al., 2014. Transforming growth factor-β superfamily ligand trap ACE-536 corrects anemia by promoting late-stage erythropoiesis. *Nature Medicine*, 20(4), pp.408–414.

Suzuki, N., 2003. Identification and characterization of 2 types of erythroid progenitors that express GATA-1 at distinct levels. *Blood*, 102(10), pp.3575–

3583.

Szvetnik, E.A. et al., 2017. Diamond-Blackfan Anemia Phenotype Caused By Deficiency of Adenosine Deaminase 2. *Blood*, 130(Suppl 1), p.874.

Tavian, M. et al., 1999. Emergence of the haematopoietic system in the human embryo and foetus. *Haematologica*, 84 Suppl EHA-4, pp.1–3.

Terszowski, G., 2005. Prospective isolation and global gene expression analysis of the erythrocyte colony-forming unit (CFU-E). *Blood*, 105(5), pp.1937–1945.

Ulirsch, J.C. et al., 2017. Confounding in ex vivo models of Diamond-Blackfan anemia. *Blood*, 130(9), pp.1165–1168.

Utsugisawa, T. et al., 2016. Erythrocyte glutathione is a novel biomarker of Diamond-Blackfan anemia. *Blood Cells, Molecules, and Diseases*, 59(C), pp.31–36.

Vandekerckhove, J. et al., 2009. Regulation of erythropoiesis. In pp. 1–44.

Varricchio, L. & Migliaccio, A.R., 2014. The role of glucocorticoid receptor (GR) polymorphisms in human erythropoiesis. *American journal of blood research*, 4(2), pp.53–72.

Varricchio, L. et al., 2011. Increased frequency of the glucocorticoid receptor A3669G (rs6198) polymorphism in patients with Diamond-Blackfan anemia. *Blood*, 118(2), pp.473–474.

Vlachos, A. & Muir, E., 2010. How I treat Diamond-Blackfan anemia. *Blood*, 116(19), pp.3715–3723.

Vlachos, A. et al., 2013. Clinical utility gene card for: Diamond - Blackfan Anemia - update 2013. *European journal of human genetics : EJHG*, 21(10).

Vlachos, A. et al., 2008. Diagnosing and treating Diamond Blackfan anaemia: results of an international clinical consensus conference. In British journal of haematology. Blackwell Publishing Ltd, pp. 859–876.

Vlachos, A. et al., 2001. Hematopoietic stem cell transplantation for Diamond Blackfan anemia: a report from the Diamond Blackfan Anemia Registry. *Bone Marrow Transplantation*, 27(4), pp.381–386.

Vlachos, A. et al., 2012. Incidence of neoplasia in Diamond Blackfan anemia: a report from the Diamond Blackfan Anemia Registry. *Blood*, 119(16), pp.3815–3819.

Vlachos, A. et al., 2016. Myelodysplastic Syndrome and Gastrointestinal Carcinomas Characterize the Cancer Risk in Diamond Blackfan Anemia: A Report from the Diamond Blackfan Anemia Registry. In American Society of Haematology.

Vlachos, A., Blanc, L. & Lipton, J.M., 2014. Diamond Blackfan anemia: a model for the translational approach to understanding human disease. *Expert review of hematology*, 7(3), pp.359–372.

Wallis, Y., 2013. Practice Guidelines for the Evaluation of Pathogenicity and the Reporting of Sequence Variants in Clinical Molecular Genetics., pp.1–16. Available at: http://www.acgs.uk.com/media/774853/evaluation_and_reporting_of_sequence_variants_bpgs_june_2013_-_finalpdf.pdf [Accessed January 31, 2018].

Walrafen, P. et al., 2005. Both proteasomes and lysosomes degrade the activated erythropoietin receptor. *Blood*, 105(2), pp.600–608.

Wang, Q. et al., 2008. MicroRNA miR-24 inhibits erythropoiesis by targeting activin type I receptor ALK4. *Blood*, 111(2), pp.588–595.

Wang, R. et al., 2015. Loss of function mutations in RPL27 and RPS27 identified by whole-exome sequencing in Diamond-Blackfan anaemia. *British Journal of Haematology*, 168(6), pp.854–864.

Wangen, J.R. et al., 2014. Assessment of normal erythropoiesis by flow cytometry: important considerations for specimen preparation. *International journal of laboratory hematology*, 36(2), pp.184–196.

Warren, A.J., 2017. Decoding erythropoiesis. *Blood*, 129(5), pp.544–545.

Wiernik, P.H. et al., 2012. *Neoplastic Diseases of the Blood*, Springer Science & Business Media.

Willig, T.N. et al., 1998. High adenosine deaminase level among healthy probands of Diamond Blackfan anemia (DBA) cosegregates with the DBA gene region on chromosome 19q13. The DBA Working Group of Société d'Immunologie Pédiatrique (SHIP). *Blood*, 92(11), pp.4422–4427.

Willig, T.N., Draptchinskaia, N., et al., 1999a. Mutations in ribosomal protein S19 gene and diamond blackfan anemia: wide variations in phenotypic expression. *Blood*, 94(12), pp.4294–4306.

Willig, T.N., Niemeyer, C.M., et al., 1999b. Identification of new prognosis factors from the clinical and epidemiologic analysis of a registry of 229 Diamond-Blackfan anemia patients. DBA group of Société d"Hématologie et d"Immunologie Pédiatrique (SHIP), Gesellshaft für Pädiatrische Onkologie und Hämatologie (GPOH), and the European Society for Pediatric Hematology and Immunology (ESPHI). *Pediatric research*, 46(5), pp.553–561.

Winkler, T. et al., 2016. Successful Treatment of a Diamond-Blackfan Anemia (DBA)

Patient with Eltrombopag. *Blood*, 128 (22), p.2682.

Yoshida, H. et al., 2005. Phosphatidylserine-dependent engulfment by macrophages of nuclei from erythroid precursor cells. *Nature*, 437(7059), pp.754–758.

Yuan, H. et al., 2016. A de novo 1.6Mb microdeletion at 19q13.2 in a boy with Diamond-Blackfan anemia, global developmental delay and multiple congenital anomalies. *Molecular Cytogenetics*, pp.1–6.

Zermati, Y. et al., 2001. Caspase activation is required for terminal erythroid differentiation. *Journal of Experimental Medicine*, 193(2), pp.247–254.

Zhang, E.G. et al., 2011. Managing the difficult case of fetal anemia. *The Journal of Maternal-Fetal & Neonatal Medicine*, 24(12), pp.1498–1503.

Zhang, J. et al., 2009. Mitochondrial clearance is regulated by Atg7-dependent and -independent mechanisms during reticulocyte maturation. *Blood*, 114(1), pp.157–164.

Zhang, L. et al., 2013. ZFP36L2 is required for self-renewal of early burst-forming unit erythroid progenitors. *Nature*, 499(7456), pp.92–96.

Zhang, Y. et al., 2003. Ribosomal protein L11 negatively regulates oncoprotein MDM2 and mediates a p53-dependent ribosomal-stress checkpoint pathway. *Molecular and cellular biology*, 23(23), pp.8902–8912.

Zhou, J. & Cidlowski, J.A., 2005. The human glucocorticoid receptor: one gene, multiple proteins and diverse responses. *steroids*, 70(5-7), pp.407–417.