Identification and functional investigation of genes in patients with inherited bleeding disorders

by

Ibrahim Abdullah Almazni

A thesis submitted to the University of Birmingham for the degree of DOCTOR OF PHILOSOPHY

Institute of Cardiovascular Sciences
College of Medical and Dental Sciences
The University of Birmingham
February 2021

Abstract

Inherited bleeding disorders (IBDs) comprise a heterogeneous group of diseases that reflect abnormalities in blood vessels, coagulation proteins and platelets. The diagnosis of patients with IBDs is challenging and approximately 50-60% of patients recruited to the UK-GAPP study have no known cause of their bleeding despite a strongly indicative inherited component. Whole Exome Sequencing (WES) analysis using a bioinformatics pipeline workflow was carried out in 6 patients including 4 isolated individuals and two related individuals with bleeding episodes and low platelet counts. A stop gained variant within *CD36*: c.975T>G; p. Tyr325* was identified in related patients 5.1 and 5.2 and a number of other plausible candidate variants were also found for the other isolated individuals. Functional studies were conducted on stop gained variants within *CD36* which showed the effect of this variant on the expression of CD36 protein using Western blot and flow cytometry techniques. This finding was confirmed through the activation of NFAT-luciferase reporter assay by the WT CD36, but not by the mutant constructs. In addition, a comprehensive bioinformatic analysis of 126 patients with suspected platelet disorders was carried out using the Congenica diagnostic software to identify both structural and sequence genetic variants. 28.2% of patients were noted with platelet function defects and 19.6% of patients represented thrombocytopenia only. A total of 135 variants in genes implicated in bleeding disorders were identified across all patients. 22 patients were classified as pathogenic and 26 patients as likely pathogenic, while 87 patients had uncertain pathogenicity. The Congenica software has also identified potential CNVs in 3 patients. In conclusion, this thesis has demonstrated that the application of combined phenotyping and genotyping coupled with WES technology and bioinformatic tools are efficient and effective approaches for refining and identifying new sequence variants in patients with suspected IBDs.

Publications arising from this work

1. Khan AO, Stapley R, Pike JA, Wijesinghe SN, Reyat JS, **Almazni I,** Machlus KR, Morgan NV (2021) Novel gene variants in patients with platelet-based bleeding using combined exome sequencing and RNAseq murine expression data. *J Thromb Haemost* 19:262-268.
2. **Almazni I**, Stapley RJ, Khan AO, Morgan NV (2020) A comprehensive bioinformatic analysis of 126 patients with an inherited platelet disorder to identify both sequence and copy number genetic variants. *Human Mutation* 41: 1848-1865.
3. **Almazni I**, Stapley R, Morgan NV (2019) Inherited thrombocytopenia: Update on genes and genetic variants which may be associated with bleeding. *Front Cardiovasc Med* 6: 80.

4. **Almazni I**, Chudakou P, Dawson-Meadows A, Downes K, Freson K, Mason J, Page P, Reay K, Myers B, Morgan NV (2021) A novel *RUNX1* exon 3 - 7 deletion causing a familial platelet disorder. *Platelets In press.*

Acknowledgements

First, I would like to sincerely thank my main supervisor Professor Neil Morgan for all the help, guidance and support he offered during my PhD and subsequent writing of this thesis. He guided me during these years in a successful way to achieve my goals. I would like to thank my second supervisor Prof Timothy Barrett for his supervision during my PhD. None of this work would have been possible without their supervision. A special thank you goes to Dr Annabel Maclachlan, Dr Alexandre Slater, Connie Koo, Ying Di, Rachel Stapley, Jack Yule, for their assistance during my research and for teaching me various laboratory techniques. I would like to thank my supervisor Dr Neil Morgan, Dr Michael Tomlinson, Dr Abdullah Khan, Connie Koo, Rachel Stapley for being very supportive whenever needed and for their helpful comments on the thesis. I would like to thank everyone in the Birmingham Platelet Group, past and present, for their help, chats and support during my PhD. Special thanks go to my friends Fawaz Alanazi and Maan Harbi for their help, comments and support. I would like to acknowledge the Saudi Government (Ministry of Higher Education) especially Najran University for providing generous funding for this research. Finally, I would like to extend my eternal gratitude and thanks to my father and mother, brothers, sisters and friends thanks a lot for your support, pray and encouragement throughout my PhD journey. My special thanks goes to my incredibly lovely wife for her care, understanding and patience - thank you from the bottom of my heart for everything you have done for me. Thank you very much to my daughters Toleen and Mila, just looking into your eyes every morning gives me the strength to work hard in order to give you an amazing future.

Contents

List of figures

List of tables

List of abbreviations

AA	Arachidoinc acid
ACD	Acid-citrate-dextrose
ACMG	American College of Medical Genetics and Genomics
ADP	Adenosine diphosphate
AD	Autosomal dominant
AMP	Association for Molecular Pathology
AR	Autosomal recessive
ATP	Adenosine triphosphate
BAK	Pro-apoptotic protein
BAT	Bleeding Assessment Tool
BDGP	Berkeley Drosophila Genome Project
BDP	Bleeding and platelet disorder
BM	Bone marrow
BPD	Bleeding platelet disorder
BT	Bleeding time
BSA	Bovin serum albumin
BSS	Bernard-Soulier syndrome
cAMP	Cyclin adenosine monophosphate
CAMT	Congenital amegakaryocytic thrombocytopenia
CHS	Chediak-Higashi syndrome
CLEC-2	C-type lectin receptor 2
CLP	Common lymphoid progenitor
CMM	Common myeloid progenitor
CNV	Copy number variant
COX	cyclooxygenase
CRISPR	Clustered regularly interspaced short palindromic repeats
DDAVP	1-desamino-8-d-arginine vasopressin
DMS	Demarcation membrane system
DNA	Deoxyribonucleic acid
DS	Down syndrome
DTT	Dithiothreitol
EDTA	Ethylenediaminetetraacetic acid
EVS	Exon Variant Server
ExAC	Exome Aggregation Consortium
FcRγ	Fc receptor gamma chain
FITC	Foundation for International Technological Cooperation
FLI1	Friend leukaemia virus integrin 1
FLNA	Filamin A
FLNB	Filamin B
FNAIT	Fetal/neonatal alloimmune thrombocytopenia
FOG1	Friend of GATA-1
FPD-AML	Familial platelet disorder with propensity to acute myelogenous leukemia
GAPP	Genotyping and Phenotyping in platelets
GATA-1	GATA-binding factor 1
GDP	Guanosine diphosphate
GMP	Granulocyte macrophage progenitor
GTP	Guanosine triphosphate

GP	Glycoprotein
GPO	Glycine-proline-hydroxyproline
GPCR	G-protein coupled receptor
GPS	Grey platelet syndrome
GT	Glanzman thrombasthenia
HD	Huntington's disease
HEK	Human Embryonic Kidney
HGDM	Human Gene Mutation Database
HGP	Human Genome Project
HLA	Human leukocyte antigen
HMW	High molecular weight
HPA	Human platelet antigens
HPO	Human phenotype ontology
HPR	Horseradish Peroxidase
HPS	Hermansky-Pudlak syndrome
HSC	Haematopoietic stem cell
HTS	High-throughput sequencing
IBD	Inherited bleeding disorders
ICH	Intracranial haemorrhage
IMS	Invaginated membrane system
IPD	Inherited platelet disorder
IPF	Immature platelet fraction
iPSC	Inducible pluripotent stem cell
IR	Interpretation Request
ISTH	International Society of Thrombosis and Hemostasis
IT	Inherited thrombocytopenia
ITP	Idiopathic thrombocytopenic purpura
JAK2	Janus kinase 2
JNK-2	Jun N-terminal kinase
LOVD	Leiden Open Variation Database
LB	Lymphoblast
LTA	Light transmission aggregometry
MAF	Minor allele frequency
MAPK	Mitogen activated protein kinase
MCV	Mean platelet volume
MEA	Multiple electrode aggregometry
MEP	Megakaryocyte erythroid progenitor
MK	Megakaryocyte
MKB	Megakaryoblast
MPP	Multipotent progenitors
MPV	Mean platelet volume
NCBI	National Center for Biotechnology Information
NFAT	Nuclear factor of activated T-cells
NF-E2	Nuclear factor erythroid 2
NGS	Next-generation sequencing
NO	Nitric oxide
NPF	Nucleation-promoting factor
NSAID	Non-steroidal anti-inflammatory drug
OMIM	Online Medelian Inheritance in Man
oxLDL	oxidised low-density lipoprotein

PBS	Phosphate-buffered saline
PCR	Polymerase chain reaction
PFA	Platelet Function Analyser
PFD	Platelet function disorders
PFT	Platelet function testing
PGI_2	Prostacyclin
PI3K	Phosphoinositol-3 kinase
PLA_2	Phospholipase A 2
PLT	Platelet
PMA	Phorbol myristate acetate
PPP	Platelet poor plasma
PROVEAN	Protein Variant Effect Analyser
PRP	Platelet rich plasma
PS	Phosphatidylserine
PT	Prothrombin time
PTK	Protein tyrosine kinase
PT-VWF	Platelet type–von Willebrand disease
RBC	Red blood cell
RD	Rare disease
RGD	Rare genetic disease
RNA	Ribonucleic acid
RUNX1	Runt-related transcription factor 1
SDM	Site-Directed Mutagenesis
SEM	Subendothelial matrix
SFK	Src family kinase
S1P	Sphingosine 1 phosphate
SIFT	Sorting Intolerant from Tolerant
SNP	Single nucleotide polymorphism
SNV	Single nucleotide variant
STAT5	Signal transducer and activator of transcription 5
TAE	Tris acetate-EDTA
TAR	Thrombocytopenia absent radii
TBS-T	Tris-buffered saline with Tween
THC2	Thrombocytopenia type-2
TPO	Thrombopoietin
TSP	Thrombospondin
T-TAS	Total Thrombus Formation Analysis System
TXA_2	Thromboxane synthase
TXA_2R	Thromboxane synthase receptor
UCSC	University of California Santa Cruz
uPA	Urokinase plasminogen activator
UK_GAPP	United Kingdom Genotyping and Phenotyping of Platelets
UTR	Untranslated region
VWD	von Willebrand disease
VWF	von Willebrand factor
WAS	Wiskott-Aldrich syndrome
WASP	Wiskott-Aldrich syndrome protein
WES	Whole exome sequencing
WGS	Whole genome sequencing
WT	Wild-type

XLT	X-linked thrombocytopenia
XLTT	X-linked thrombocytopenia with thalassemia

Chapter.1
Introduction

1.1 Human genetic disorders and Mendelian inheritance principles overview

A genetic disorder is defined as a disease that is caused by pathogenic changes in an affected individual's DNA. Protein coding regions constitute ~1-1.5% (Hatje et al., 2019, Venter et al., 2001) of the human genome which involves approximately (20,000 - 25,000 genes) distributed across ~233,785 exons (Sakharkar et al., 2004, Consortium, 2004).Generally, genetic disorders can be divided into three major classes which include: single gene disorders (monogenic), chromosomal abnormality disorders and multifactorial (complex) disorders (Strachan and Read, 2019). These genetic disorders can result from defective genes or environmental and lifestyle factors or both which indicate that the genes and genetic-environmental interactions are the main parameters associated with genetic diseases (Iourov et al., 2019, Heng, 2009). Multifactorial diseases such as coronary heart disease are caused by multiple genetic defects or environmental factors which involve drug use, air pollution, infection from the mother during pregnancy or proximity to harmful waste sites (Khan et al., 2015, Poulter, 1999, Ritz et al., 2002).

Chromosomal disorders are a group of disorders that arise from alterations in either the number or structure of the chromosomes. These numerical abnormalities result from the presence of more or less copies of the normal diploid set. The loss or gain of chromosomal material during chromosome rearrangements cause structural abnormalities. Down syndrome (DS), the most common chromosomal disorder, results from the presence of three copies (trisomy) of chromosome 21 (Patil et al., 2014, Luthardt and Keitges, 2001, Kirsch-Volders et al., 2002).

Monogenic disorders are characterised by genetic changes in the DNA sequence for one gene which subsequently produce an altered or faulty protein product. For instance, the inherited genetic disorders involving production of abnormal haemoglobin (Sickle cell anaemia) are the most common monogenic diseases caused by a single gene in humans (Weatherall, 2000, Lonergan et al., 2001). Monogenic disorders can be classified based on Mendelian heredity principles into four main categories: autosomal dominant disorders (AD), autosomal recessive disorders (AR), X-linked disorders and Y -linked disorders. In AD, only one affected allele of the relevant gene is affected and cause the disorder. The chance of transmitting the defective allele from affected parents to their offspring is about 50% (Mahdieh and Rabbani, 2013). As an example, Huntington's disease (HD), an autosomal dominant progressive neurodegenerative disorder, caused by a genetic defect in the *HTT* gene (4p16.3) due to expansion repeats of CAG sequence (Apolinário et al., 2017). In addition, a heterozygous mutation in the *FGFR3* gene causes Achondroplasia which is considered as the most commonly inherited form of dwarfism, and also follows an AD inheritance pattern (Su et al., 2010). AR diseases result from the presence of two abnormal copies of the mutated gene. One copy comes from the father and the other from the mother. The mutated allele is transmitted from heterozygous carrier parents, which make the affected offspring homozygous or compound heterozygous for the mutated alleles (Mahdieh and Rabbani, 2013). Chances of inheriting an AR trait for having an affected offspring is 25% for heterozygous parents. For example, homozygous or compound heterozygous mutations in the *CFTR* gene can cause cystic fibrosis, characterised by secretion of thick and sticky mucus in the lungs and other organs (Cutting, 2015) resulting in breathing difficulties and respiratory distress. Mutations in genes located on the X chromosome lead to X-linked disorders. Both

males and females can be affected by X-linked dominant and recessive disorders but males are more affected because they only have one copy of the X chromosome (hemizygous)(Khan et al., 2015, Mahdieh and Rabbani, 2013).Haemophilia B, also known as factor IX deficiency, is an X-linked recessive disorder which is caused by mutations in the *F9* gene and subsequently prevents blood clotting (Wang et al., 2016).

1.2 Consanguinity and increased risk of genetic malformation

Consanguinity in marriages is a profoundly social habit that has existed since the time of the earliest modern humans(Tadmouri et al., 2009). Consanguinity in marriage is defined as the union between two related individuals as first or second cousins who shared a common ancestor (Hamamy, 2012). Consanguinity in marriage is relatively common and estimated to total one billion of the world's population, where they live in communities that prefer consanguinity in marriage (Bittles and Black, 2010a, Modell and Darr, 2002). High rates of consanguinity marriage are seen in most populations in North Africa, Middle East and West Asia, where the familial unions are expected to reach 20-50% of all marriages (Figure 1.1) (Bittles, 2001, Hamamy et al., 2011). Furthermore, the migration of populations from these regions lead to increase the consanguinity marriage rate in North America and Europe (Teeuw et al., 2014). Although the consanguineous marriage is thought to strengthen the social relationships between families and reduce the financial living costs, the risk of transmitted genetic defects into offspring are increased(Hamamy, 2012). It has been shown that the rate of genetic defects at birth is estimated to be 1.7-2.8% in first cousin marriages mostly contributable to AR disorders (Bennett et al., 2002). In addition, the chance of having AR disorders is increased with consanguinity; however, the risk of expression of AD disorders in progeny of consanguineous marriages is not increased

particularly when only one of the parents is affected (Bittles and Black, 2010b, Hamamy et al., 2011, Hamamy et al., 2007). As a result, offspring of consanguineous unions are more susceptible to AR disorders due to the inheritance of AR gene mutations from a common ancestor (Jaber et al., 1998, Zlotogora, 2002). The probability of passing one or more identical mutant recessive genes to an offspring is increased when parents are more closely biologically related. For illustration, it is estimated that first cousin marriages share 12.5% (1/8) of their genes. Thus, on average, 6.25% (1/6) of gene loci will be homozygous in their offspring (Bennett et al., 2002). The genetic studies of consanguineous families with clinical phenotypes are favourable and useful for identification of novel disease-causing genes specially when they share similar phenotypes (Erzurumluoglu et al., 2016). In chapter 3, WES and bioinformatic analysis was conducted in one large consanguineous family with an inherited bleeding disorder.

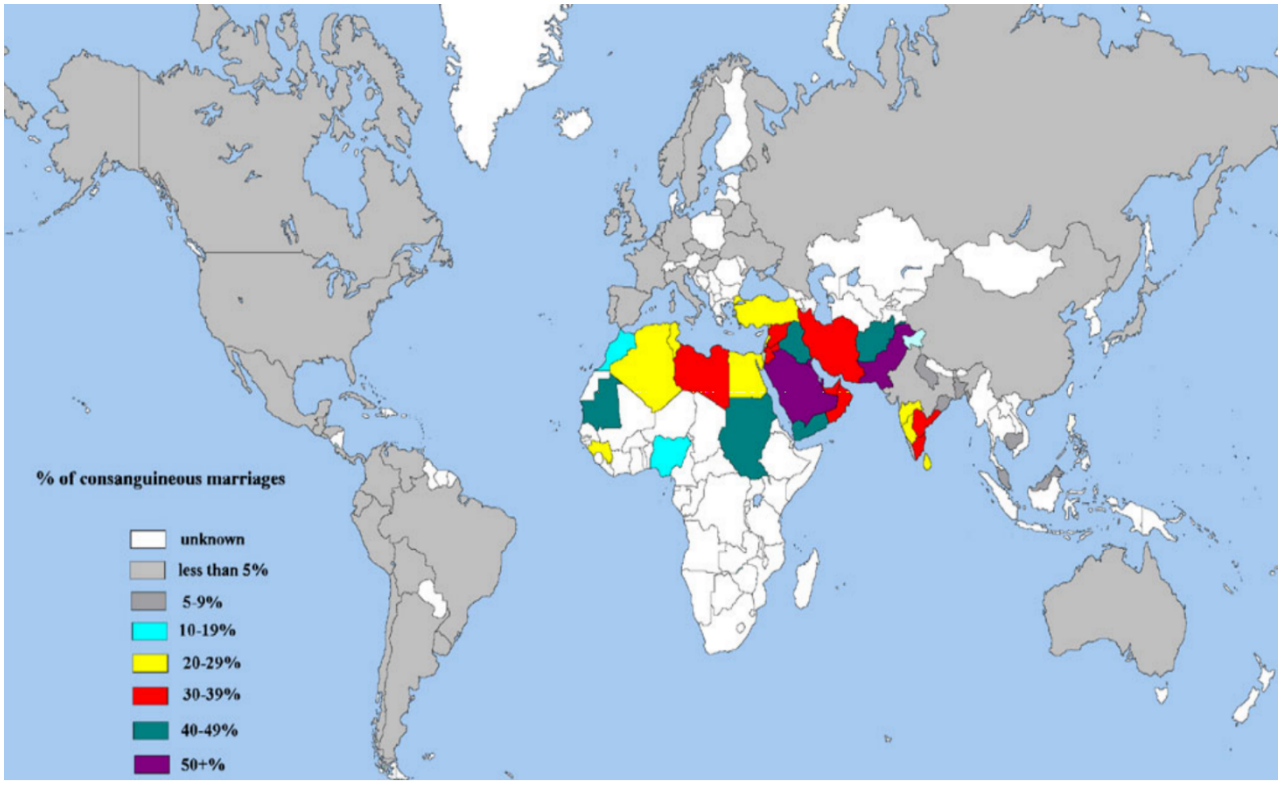


Figure 1.1. Global consanguinity rate. This figure shows the global distribution of consanguineous marriages for first or second cousins. The colours represents the percentages of consanguineous marriages as seen in the left (Bittles and Black, 2010a, Hamamy, 2012, ALHENC-GELAS et al., 2010).

1.3 Rare disease

Rare diseases (RDs) are numerous and a heterogeneous group of diseases that are disparate geographically. Although the minority of these diseases can be prevented and cured, the majority of them are chronic and can subsequently lead to early death (Wakap et al., 2020). A disease is classified as 'rare' when it affects less than 200,000 individuals in the United States (Whicher et al., 2018) or fewer than 1 in 2,000 people in Europe (Lopes and Oliveira, 2013) at any time. It is predicted that ~80% of rare diseases result from genetic defects which suggests the presence of strong phenotype-genotype association in these disorders (Lopes and Oliveira, 2013). The precise number of rare genetic diseases (RGDs), also called Mendelian or monogenic diseases, is difficult to be estimated due to the phenotypes of many that yet need to be defined (Boycott et al., 2017). It is estimated that the number of RGDs is over 7,000 according to the information provided by rare diseases databases (Boycott et al., 2013, Boycott et al., 2017). In addition, approximately 75% of these disorders affect mainly children which could induce high impact on the well-being of children and their families (Beaulieu et al., 2014). Over 3,500 RGDs have been investigated and their aetiology determined by genetic linkage analyses, candidate gene analysis and NGS technologies that are sometimes costly and time-consuming (Boycott et al., 2013). Despite the progressive nature of RGDs, early diagnosis of patients with these disorders, such as new-born screening, could help with the management of disease and enhance the well-being as well as reduce the long-term complications in some diseases (Boycott et al., 2017). Due to the rarity of RGDs and the heterogeneity of phenotypes and genotypes information, their diagnostic investigation can be difficult sometimes requiring long and expensive diagnostic assessment (Boycott et al., 2013). Moreover, clinical and genetic information for some RGDs are limited which makes it

difficult for clinicians to distinguish between similar conditions leading to difficulty in genetic counselling and suitable treatment. Clinicians and healthcare workers are advised to exploit all available resources such as online databases in order to help with diagnosis, management and treatment of patients with RGDs (Pogue et al., 2018). Two international recognised databases provide valuable clinical and genetic information for some RGDs: Online Mendelian Inheritance In Man (OMIM) (https://www.omim.org) and orphanet portal for rare diseases and orphan drugs (https://www.orpha.net). As many rare diseases are genetically inherited, they manifest with varying degrees of severity which require proper diagnosis and evaluation (Figure 1.2) (Pogue et al., 2018). The genetic and phenotypic variation of RGDs is gradually being recognised; yet, an intensive understating of RGDs is still unclear. Despite distinct advances in understanding the molecular aetiology of RGDs, the definitive molecular diagnosis of some RGDs remain undetermined (Boycott et al., 2017, Pabinger et al., 2014). Molecular diagnosis of RGDs has rapidly progressed and become less expensive since the employment and wider acceptance of next-generation sequencing (NGS). For example, Whole exome sequencing (WES), a feasible and cost-effective technique, provides effective diagnosis of Mendelian disorders. Subsequently, identifying disease causing genes will provide better disease management, genetic counselling and treatment (Pogue et al., 2018).

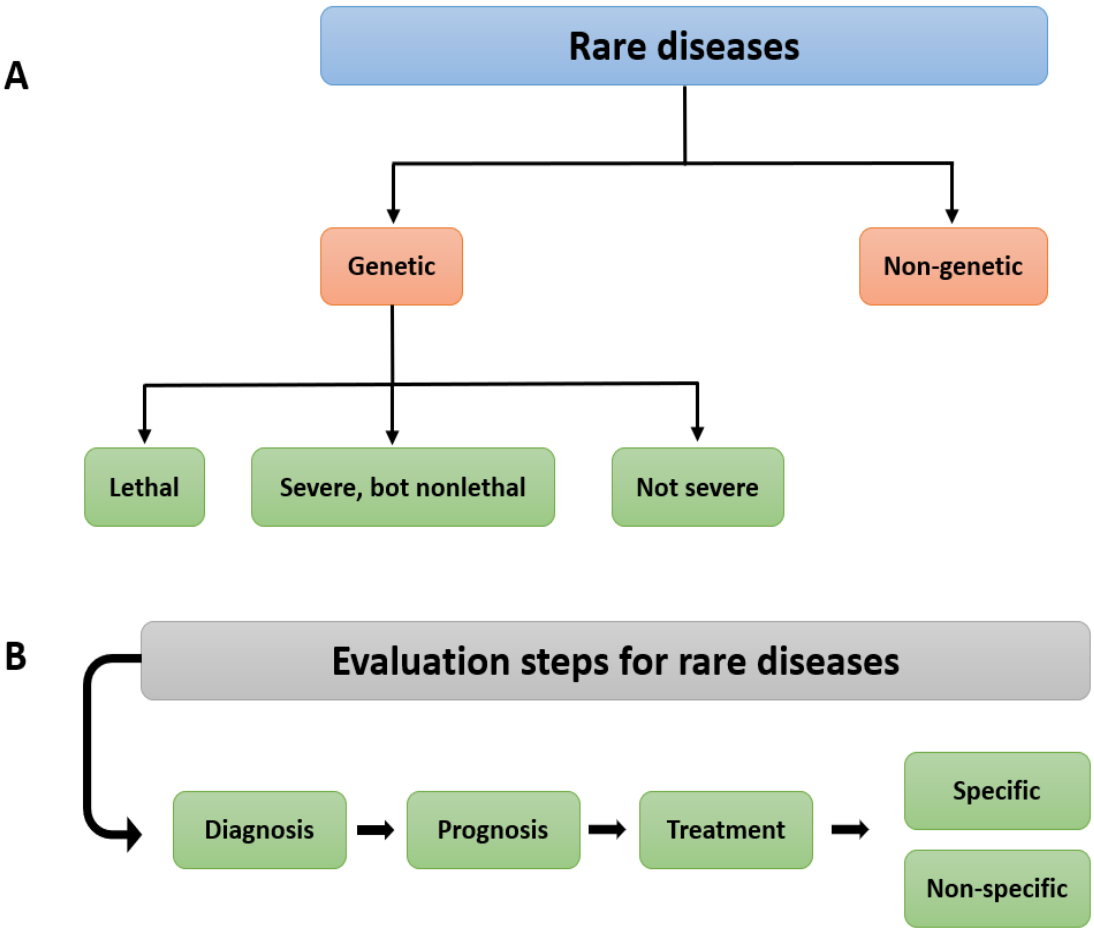


Figure 1.2. Classification and evaluation steps for rare disease. (A) rare diseases can be classified into genetic and non-genetic based -diseases, which manifest with different degrees of severity. (B) Proper diagnosis and evaluation of rare diseases provide better disease management, prognosis, and treatment. Figure amended and based on (Pogue et al., 2018, ALHENC-GELAS et al., 2010).

1.4 Inherited bleeding disorders

Inherited bleeding disorders (IBDs) are a heterogenous group of diseases that arise from abnormalities in blood vessels, coagulation proteins and platelets. They are considered as life-long bleeding diseases which clinically manifest from a series of bleeding episodes. Severe bleeding disorders can affect both males and females but may be more pronounced females due to menstruation and following childbirth (Chi and Kadir, 2012, Lowe et al., 2019). They usually present after birth or in childhood, but mild disorders may not be recognised until later in life after haematological assessment and testing most often after excessive bleeding is experienced during surgery (McDonald and Austin, 2017, Blanchette et al., 1991). The frequencies of IBDs range from 1 in 1000 live births for von Willebrand disease (VWD), the most common bleeding disorder, to 8 cases reported globally for alpha-2-antiplasmin deficiency (Favier et al., 2001, Palla et al., 2015). The molecular mechanisms behind these disorders can be as a result of a genetic defect in any of the genes implicated in the haemostasis pathway (Sivapalaratnam et al., 2017a). IBDs can result from quantitative or qualitative deficiency/defect of von Willebrand factor (VWF), coagulation factors, platelets or fibrinolysis (Al-Rahal, 2018, McDonald and Austin, 2017). In some complex cases, patients could present with combined coagulation protein abnormalities alone or accompanied by platelet disorders (Blanchette et al., 1991). Bleeding is not the only feature of IBDs; patients may also present other non-haematological abnormalities, which make the accurate diagnosis of the disorder difficult. Some IBDs can be easily diagnosed with careful clinical assessments and a combination of laboratory assays. These assays can vary in complexity but nevertheless, such conventional approaches are inadequate for diagnosis of many IBDs (Sivapalaratnam et al., 2017a). Although many genetic causes of bleeding disorders are known as

being primarily caused by genes associated with VWD disease and/or coagulation factor diseases , the genetic causes of rare platelet bleeding disorders are in fact less understood (Bolton-Maggs et al., 2006, Swystun and James, 2017). To investigate the molecular mechanism behind these disorders, it is often best to address all genes implicated in platelet bleeding disorders and how their variants can disrupt the gene function. In this thesis, the research will focus on rare platelet-based bleeding disorders specifically those with low platelet counts (thrombocytopenia).

1.5 Platelet overview

Platelets are small anucleate blood cells that play a significant biological role to maintain normal haemostasis after vascular injury. Platelets were firstly discovered by Max Schultze in 1865 as a normal constituent of the blood when he studied white blood cells (Brewer, 2006). Later in 1882, an Italian pathologist Giulio Bizzozero demonstrated platelets in a comprehensive study when he observed them under the microscope in a blood sample from vessels and in the blood of living animals. Accordingly, Bizzozero described platelets as the third morphological particles of the blood components, unrelated to erythrocytes and leukocytes which play vital roles in both haemostasis and thrombosis. (Ribatti and Crivellato, 2007, Coller, 1984). They are produced from the cytoplasm of large mature cells called Megakaryocytes (MKs) in the blood vessel lumen with an average production rate of approximately $100X10^9$ platelets per day, to maintain the normal circulating range from 150,000 to 450,000 platelets per microliter of blood (Drachman, 2004a, Ghoshal and Bhattacharyya, 2014). The normal average number of blood platelets in the human adult circulation is nearly one trillion platelets (Thon and Italiano, 2010). Furthermore, platelets circulate in the blood to maintain the integrity of blood vessels and play a vital role in preventing excessive bleeding at the site of injury, forming a clot plug to prevent further damage

to the vessel wall and reduce blood loss (Ruggeri, 2002). Upon injury of the blood vessel, proteins of the subendothelial matrix (SEM) are exposed, platelets start rapid functional activation to prevent excessive bleeding and to seal the wound. This activation is essential in a haemostasis, most particularly in the high shear conditions of arteries and venous systems. However, excessive uncontrolled activation of platelets at the site of the rupture may result in impaired haemostasis, leading to health complications. A typical example is thrombosis in diseased arteries, which may lead to thrombotic occlusion and suppression of blood flow, resulting in subsequent myocardial infarction or stroke whereby the tissue is starved of oxygen (Para et al., 2011).

1.6 Haematopoiesis and megakaryopoiesis

Megakaryopoiesis and subsequent thrombopoiesis involves series of complicated biological steps. Megakaryocytes, like all blood cells, are derived from hematopoietic stem cells (HSCs) in the bone marrow. During the commitment of different cell lineages of HSCs, this complex process requires multiple cytokines and transcription factors in order to regulate the differentiation process (Schulze and Shivdasani, 2005, Guo et al., 2015a). The detailed process of haematopoiesis is summarised in (Figure 1.3). HSCs typically reside in a unique microenvironment inside the bone marrow (BM), called endosteal zones, which are considered the primary site of haematopoiesis (Nombela-Arrieta et al., 2013). Schofield initially described these unique microenvironments as "stem cell niches" when he studied the localisation of HSCs (Schofield, 1978). There are two types of microenvironment niches within the BM where HSCs reside including: osteoblastic niche and the vascular niche (Calvi et al., 2003, Kiel et al., 2005). HSCs are mainly maintained by the endosteal osteoblastic niche by providing a quiescent microenvironment for HSCs, while the vascular niche

promotes the self-renewal of HSCs and regulates the initial differentiation and mobilisation of HSCs into all blood lineages (Taichman, 2005, Mikkola and Orkin, 2006, Yoon et al., 2012). The osteoblastic niche is lined with osteoblastic cells, adjacent to the endosteum bone surface area. It has been revealed that an increase in the number of osteoblastic cells lining the endosteum bone area leads to maintain the quiescent stem cell population (Arai et al., 2004, Zhang et al., 2003). In the vascular niche microenvironment, HSCs become more exposed to the vascular endothelial cells in the BM which promotes the operation of self-renewal and differentiation into the various types of blood cells dependent on lineage (Chute et al., 2007, Kopp et al., 2005, Li et al., 2010). The first step for HSCs differentiation into megakaryocytes involves the development of multipotent progenitors (MPPs) which are no longer able to self-renew (Seita and Weissman, 2010, Morrison and Weissman, 1994). Further downstream, MPPs undergo developmental process to produce two oligopotent progenitors: the common myeloid progenitors (CMPs) or common lymphoid progenitors (CLPs) respectively. The CLPs then have the propensity to produce B and T lymphocyte cells as well as natural killer cells (Velten et al., 2017). The CMPs produce megakaryocyte erythroid progenitors (MEPs) or granulocyte macrophage progenitors (GMPs). The latter can progress to generate granulocyte cells and monocytes along with macrophages. MEPs can subsequently differentiate into erythrocytes and megakaryocytes depending on the transcription factor stimulation (Akashi et al., 2000). Megakaryocytes undergo a maturation process that results in mature polyploid megakaryocytes, and followed by the formation of pro-platelets (Schulze and Shivdasani, 2005).

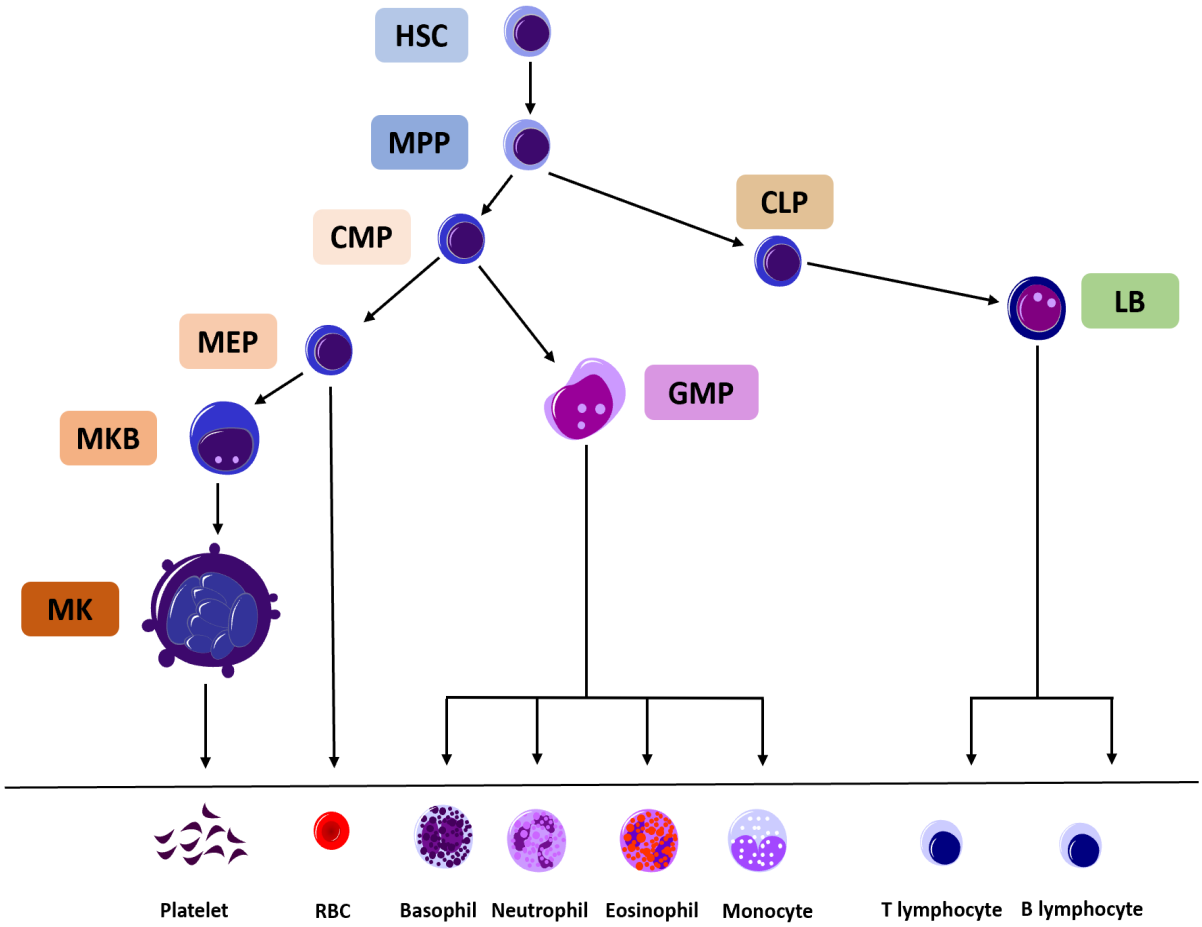


Figure 1.3. Schematic diagram of the hematopoietic hierarchy. The hierarchy starts with hematopoietic stem cells that has the capacity for self-renewal and the potential ability to give rise into all hematopoietic cell types (multipotent). HSC: haematopoietic stem cell, MPP: multipotent progenitor, CMP: common myeloid progenitor, CLP: common lymphoid progenitor, LB: lymphoblast, MEP: megakaryocyte erythroid progenitor, GMP: granulocyte macrophage progenitor, MKB: megakaryoblast, MK: megakaryocyte, RBC: red blood cell. Figure reproduced from (Cheng et al., 2020).

1.6.1 The role of transcription factors and cytokines inmegakaryopoiesis

The process of megakaryocyte differentiation involves many transcription factors as well as cytokine signalling which regulate the production of megakaryocytes and platelets (Tijssen and Ghevaert, 2013a). Thrombopoietin (TPO), a haematopoietic chemokine, is the main driver and regulator of the formation of megakaryocytes and subsequently platelets. It is synthesised mainly in the liver parenchymal cells and to a lesser extent in the kidneys and bone marrow (Jelkmann, 2001, Tijssen and Ghevaert, 2013b). It is required for megakaryocytes to fulfil their developmental proliferation and for the subsequent maturation of platelets (Kaushansky, 1998). Although the major role of TPO is to control the platelet number, it also has maturation and proliferation properties (Murone et al., 1998). TPO acts on its cognate c-Mpl receptor, also known as the myeloproliferative leukemia protein or CD110, on the surface of megakaryocytes and platelets, as well as on HSCs. It is encoded by the c-Mpl proto-oncogene that is homologous with members of the hematopoietic receptor superfamily (Drachman et al., 1995, Kaushansky et al., 1994). TPO binds to the inactive form of c-Mpl receptor which becomes active after a series of conformational changes, which bring the cytoplasmic tails of the receptor closer (Geddis, 2010, Kuter, 2007). The c-Mpl receptor lacks the kinase activity that can be gained by association with the intracytoplasmic tyrosine kinase, known as Janus Kinase 2 (JAK2). The cognate association of c-Mpl with JAK2 regulates the signal transduction of TPO (Besancenot et al., 2014). The auto-phosphorylation of JAK2 subsequently induces phosphorylation downstream, which leads to a triggering of multiple signalling pathways, including phosphoinositol-3 kinase (PI3K), mitogen-activated protein kinase (MAPK) and signal transducers and activators of transcription 5 (STAT5) Megakaryocyte differentiation,

as well as the repression and induction of gene expression, results from the activation of these pathways (Avecilla et al., 2004). Transcription factors also play a vital role in haematopoietic cell lineage differentiation. There are different transcription factors involved in megakaryocyte differentiation which include the zinc finger protein GATA-binding factor 1(GATA-1), the friend of GATA-1(FOG1), the nuclear factor erythroid-2 (NF-E2) and transcription factor friend leukaemia virus integration 1 (FLI-1) (Shivdasani, 2001). GATA-1 is highly expressed in the erythroid and megakaryocytic lineage, which plays a vital role in the maturation, growth and differentiation of megakaryocytes. It has been established that MKs are delayed in their maturation and increase their cell proliferation if they lack GATA-1, which indicates the role GATA-1 plays in proliferation (Shivdasani et al., 1998). A study has demonstrated that FOG-1 modifies GATA-1 activity by binding directly to its N-terminal zinc finger. Consequently, experimental studies revealed that FOG-1 directly enhances the activity of GATA-1 on the promoter of integrin alpha-IIb (Wang et al., 2002). NF-E2 is a transcription factor belonging to a basic leucine zipper family of dimeric transcription factors, which plays an important role in the regulation of erythrocyte and platelet production. The expression of GATA-1 and NF-E2 are increased during the megakaryocyte-erythrocyte lineage development (Andrews, 1998, Akashi et al., 2000). It has been shown that the NF-E2 plays a critical role in platelet production and β1 tubulin expression (Lecine et al., 2000). In addition, Runt-related transcription factor 1 (RUNX1) is another major transcription factor involved in the regulation of haematopoiesis and particularly megakaryopoiesis. It has the capacity to maintain, proliferate and differentiate the HSCs (Schlegelberger and Heller, 2017, North et al., 2004). The RUNX-1 and CBF-β complex has been implicated in the maturation of MKs. It plays a pivotal role in mediating DNA binding. Defects in RUNX-1 disrupt this

complex function because of their negative effect on the binding with CBF-β which subsequently impairs the regulation of genes involved in the HSCs differentiation and maturation (Morgan and Daly, 2017). Moreover, FLI1 also plays a critical role in the regulation of megakaryocyte differentiation (Vo et al., 2017). Additional transcription factors such as ETV6 and GFI1B also bind to critical promotor regions of megakaryocyte expressed genes in order to ensure maturation of megakaryocytes (Tijssen and Ghevaert, 2013a, Foudi et al., 2014). Therefore, genetic defects in these transcription factors can disrupt the megakaryopoiesis process and result in low platelet count or impaired platelet function (Johnson et al., 2016a).

1.6.2 Endomitosis and polyploidisation

The production of platelets then involves two further stages before release into the circulation: endomitosis and formation of proplatelets. Committed megakaryoblasts undergo the endomitosis. This involves progressive development of megakaryoblasts in which they become larger and polyploid promegakaryocytes, and ultimately megakaryocytes that have a propensity to produce platelets (Bluteau et al., 2009). Megakaryocytes follow the normal cycle of cell division and successfully pass through stages G1, S and G2 followed by a short-interrupted M phase. This cycle is then repeated which leads to many DNA replications (Zimmet and Ravid, 2000). Normally, non-muscular myosin IIB heavy chain (MYH10) and filamentous actin work together to produce the contractile ring required for normal mitosis. In this altered form of mitosis, megakaryocytes have shown defective cytokinesis which contributes to impaired cleavage and contractile ring formation (Bluteau et al., 2009, Lordier et al., 2008). In the contractile ring, RUNX1 works to downregulate MYH10 at the contractile ring which leads to subsequent endomitosis (Lordier et al., 2012). Consequently,

megakaryocytes are allowed to produce multi-lobulated polyploid nuclei after geometric distribution from 4N to 64N which subsequently increases the genomic content (Zimmet and Ravid, 2000). Megakaryocyte polyploidisation results in functional amplification of the genetic contents and is likely to enhance protein synthesis parallel to cell enlargement (Raslova et al., 2003). This step is essential for the maturation of megakaryocyte cytoplasm and form the invaginated membrane system (IMS) which is also known as the demarcation membrane system (DMS) important in the latter stage of proplatelet release into the circulation (Machlus and Italiano, 2013).

1.6.3 Proplatelet formation and platelet release

After megakaryocytes have completely matured, they branch and develop long cytoplasmic extensions, which then extend into the sinusoidal blood vessel of the bone marrow. The IMS is an extensive network of cisterns and tubules in the cytoplasm of the megakaryocytes and is continuous with the plasma membrane (Radley and Haller, 1982). The IMS is considered the membrane reservoir necessary for the rapid surface growth and formation of proplatelets (Schulze et al., 2006, Bluteau et al., 2009). Italiano et al. showed that proplatelets are made in the shape of tip swellings at the end of expanded megakaryocyte microtubules as shown in Figure 1.4 (Italiano Jr et al., 1999). Initially, the IMS incorporates polyphosphatidylinositol lipid (PI-4,5-P2) with the association of microtubules and actin filaments (Schulze et al., 2006). Polymerisation of the microtubule β1-tubulin supports the growing tip swellings; however, sliding of microtubules is facilitated by the motor protein dynein which is considered as the main driver for proplatelet elongation (Patel et al., 2005). Although the microtubules support the proplatelet elongation, they also help with the

transportation of granules and other vesicular organelles into platelets, as well as transportation of spliceosomes necessary for protein synthesis in platelets (Richardson et al., 2005, Denis et al., 2005). In addition, Filamentous-actin (F-actin) and spectrin also play important roles in the proplatelet branching and bending respectively which regulate the production of platelets (Bluteau et al., 2009, Italiano Jr et al., 1999). When the proplatelets are formed, they are directed to the lumen of bone marrow sinusoids as a result of an increasing gradient of concentration of sphingosine 1 phosphate (S1P) in the blood (Zhang et al., 2012). Platelets are released into the circulating blood after the activation of megakaryocyte S1P receptors (S1pr1) through downstream S1P signalling (Zhang et al., 2012).

The mechanism of entering proplatelet extensions into the vascular lumen requires penetration of the basement membrane which is mediated by Podosomes. It has been shown that both actin polymerisation and the WASp-Arp2/3 complex pathway are essential for podosome formation and stability (Schachtner et al., 2013). Podosomes degrade the extracellular matrix proteins (ECM), enabling proplatelets to reach into the vascular lumen and as a result of blood shear stress, fragments break off the tip swellings (Junt et al., 2007). The fragments known as preplatelets are generally larger than mature platelets and are heterogenous in size. The preplatelets areanucleate discoid particles with a diameter of 2-10μm and are often seen as barbell-shaped (Thon et al., 2012). Finally, microtubule-based forces drive the separation of platelets through polar located microtubule coils, which are released into the blood stream (Machlus and Italiano, 2013, Thon et al., 2012)(Figure 1.4). The cytoskeletal regulation of platelet production assures that platelets are shaped in uniform size although there is evidence of genetic implications which may affect platelet size (Gieger et al., 2011). After proplatelets release preplatelets , the remaining senescent megakaryocyte

nucleus is destroyed by programmed cell death (apoptosis) and phagocytosis (Gordge, 2005).

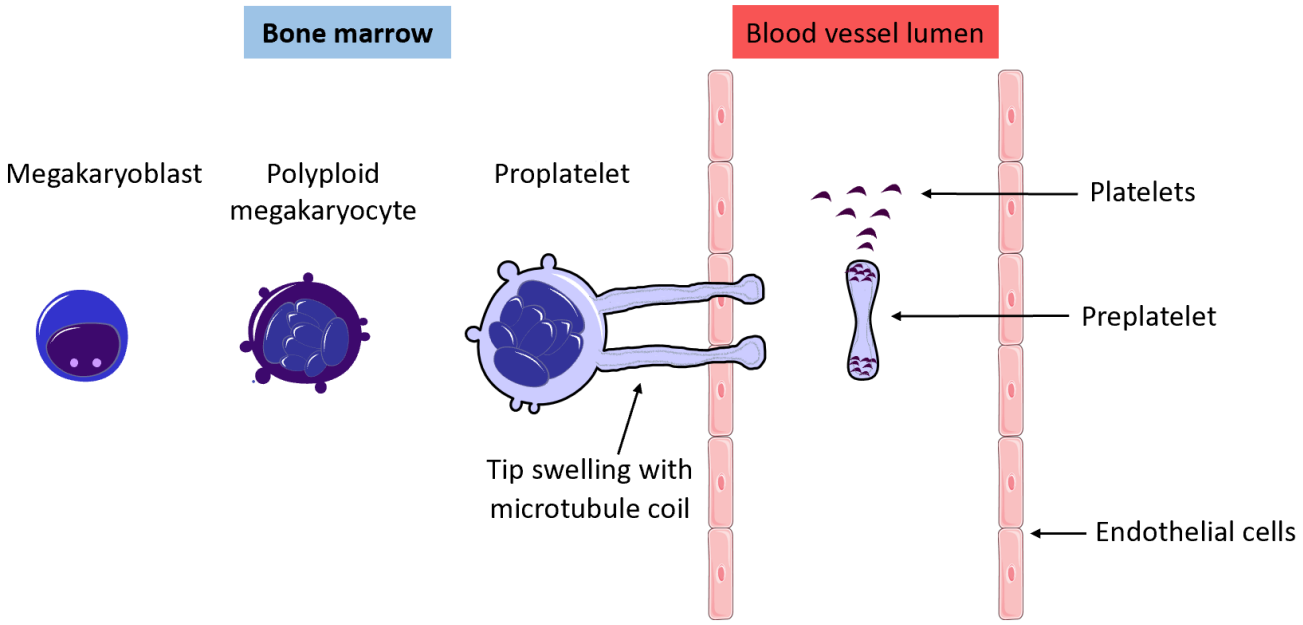


Figure 1.4. Proplatelet formation and platelet release. After the complete maturation of megakaryocytes , they start branching and produce tip swellings lined with microtubules, which protrudes through the endothelial cells lining the blood vessel. As result of blood shear stress, proplatelets are fragmented and released from the tip swellings into the vascular lumen to form preplatelets which then split to produce mature platelets. Figure amended based on (Avanzi and Mitchell, 2014, ALHENC-GELAS et al., 2010).

1.7 Platelet structure

Platelets, also called thrombocytes, are anucleate small fragments of blood cells, discoid in morphology with a diameter of 2-3 μm and 0.5 μm in depth. Although platelets do not have a nucleus, their cytoplasm contains some organelles such as Golgi apparatus, mitochondria, endoplasmic reticulum and lysosomes. The cytoplasmic regions contain an extensive cytoskeleton which is mostly comprised of complex actin network along with microtubules (Hartwig and DeSisto, 1991). The platelet cytoskeleton supports and maintains the discoid shape of resting platelets as well as playing a major role in the shape change of platelets during activation and spreading (Diagouraga et al., 2014, White and Rao, 1998). The coiled microtubules support the spectrin mesh of platelet cytoskeleton and subsequently maintain their discoid shape (Hartwig and DeSisto, 1991, Schwer et al., 2001). The spectrin fibres are bound by actin filaments which are cross-linked by filamin A (FLNA), filamin B (FLNB) and α-actin (Tablin et al., 1988). These strong framework connections support the cytoskeleton (Hartwig and DeSisto, 1991). In addition, there are two important storage granules inside cytoplasm of platelets including the alpha (α) and the dense granules which play vital roles in platelet function(Holinstat, 2017, Marcus et al., 1966). The most abundant are α-granules and previously reported that each platelet can contain up to 80 alpha granules (Blair and Flaumenhaft, 2009). α-granules contain proteins that are secreted into the circulation or bind onto the surface membrane of platelets following platelet activation. These include von Willebrand factor (VWF), GPIb-IV-V, αIIbβ3, fibrinogen and membrane P-selectin which all play important roles during platelet adhesion and aggregation (Blair and Flaumenhaft, 2009, Harrison and Cramer, 1993). Dense granules, named because of the presence of dense core contents under the electron microscope, occur at roughly 4 - 6 granules per platelet.

These store more than 200 small molecules such as adenosine diphosphate (ADP), adenosine triphosphate (ATP), guanosine triphosphate (GTP), guanosine diphosphate (GDP), calcium and serotonin (Youssefian et al., 1997, Ruiz et al., 2004). Moreover, platelets also contain other types of granules called lysosomes that play a major role in degrading proteins (Holinstat, 2017).

1.8 Functional role of platelets

The role of platelets in haemostasis is important to mainly suppress blood loss after vessel injury. After injury, platelets are rapidly recruited to the site of injury and subsequently accumulated to prevent the blood loss. This process involves multiple steps including adhesion, activation, aggregation and ultimately thrombus formation and fibrinolysis (Figure 1.5) (Rodgers, 1999, Arnout et al., 2006). In addition, platelets are also involved with other pathological and physiological processes such as thrombosis, inflammation, tumour growth and host defence (Harrison, 2005a, Jurasz et al., 2004, Zarbock et al., 2007, Yeaman, 2014). The haemostasis process functions at both high shear (arterial) and low shear (venous) rates (Watson, 2009). However, in the instance of errors or mutations which can affect platelet function this process may become erratic and ultimately result in impaired haemostasis such as bleeding (haemorrhage) or blood vessel blockage (thrombosis) (Watson et al., 2013a, Ross et al., 1988). Endothelial cells which line blood vessels maintain the fluidity of the blood flow by preventing the attachment of any of the cellular components and proteins in the circulating blood to the vessel wall (Watson, 2009). Upon vessel injury, the endothelial cells are disrupted, which results in extracellular matrix proteins and subendothelial membranes being exposed to the blood cells. Platelets start the first step in primary haemostasis by rapid adhesion to the extracellular matrix (Ruggeri and Mendolicchio, 2007). Two major molecules primarily mediate platelet tethering and

adhesion, namely, the VWF and the platelet GPIb-IX-V complex. VWF predominantly circulates in human blood plasma and upon blood vessel injurybinds to collagen (type I and III). In the instance of high shear stress, the GPIb-IX-V receptor complex on the surface of the platelet interacts specifically with the A1 domain of VWF to form a bridge with exposed collagen leading to an initial adhesion of platelets at the site of injury (Clemetson, 2012, Yun et al., 2016). Although this binding interaction is insufficient to form stable adhesion, it does promote the platelet tethering, especially at high shear rates (Savage et al., 1996).

One of the major components of the extracellular matrix (ECM) is collagen, which plays an important role in platelet adhesion, aggregation and activation (Nieswandt and Watson, 2003). The immunoglobulin-like surface receptor GPVI is recognised as an important platelet surface receptor, which couples with a signalling adaptor, the Fc receptor gamma chain (FcRγ), and induces tyrosine phosphorylation. This leads to the interaction of GPVI with exposed collagen and subsequently induces integrin activation through the inside-out signalling pathway (Chen et al., 2002, Nieswandt and Watson, 2003). Furthermore, GPVI dimers establish platelet adhesion by binding directly to specific glycine-proline-hydroxyproline (GPO) sequences on exposed collagen. Subsequently, this triggers intracellular signalling and leads to activation, releasing platelet dense granule contents including ADP and thromboxane A_2 (TxA_2) (Nieswandt et al., 2009). Platelets have a major class of surface receptors, known as integrins, which mediate the stable adhesion at high shear by binding to the ECM. They are heterodimeric proteins which consist of α and β subunits and bind tightly to their ligands when changing state from low-affinity to high-affinity during platelet activation(Nieswandt et al., 2009, Nieswandt and Watson, 2003). The most important surface membrane integrins are α2β1 and αIIbβ3 (Watson, 2009). The interaction of

α2β1 with collagen and αIIbβ3 with VWF and fibrinogen/fibronectin mediates platelet activation and promotes stable adhesion and aggregation (Clemetson, 2012, Watson, 2009). Platelet activation is stimulated by the binding of platelet secretion components and tissue factor (TF). The C-type lectin-like receptor 2 (CLEC-2) expressed in platelets plays an important role in platelet activation and thrombus formation (Yun et al., 2016). It binds to its ligand podoplanin and initiates platelet activation signalling through phosphorylation of Src family kinase (SFKs) and protein tyrosine kinase (Syk) in a similar way to GPVI (Lorenz et al., 2015) . The binding of ADP with its receptors (P2Y1 and P2Y12) and TxA_2 (TPs) give rise to further platelet activation (Brass et al., 2005, Furie and Furie, 2008). After full adherence of platelets, they undergo cytoskeletal conformational changes which lead to spreading. When αIIbβ3 binds to fibrinogen, this allows filamin and α-actin to extend filopodia and lamellipodia which increase the surface area and enhance the interaction with other recruited platelets and form a stable layer adhering to the injury site (Durrant et al., 2017).

Activated platelets are also involved in the overall coagulation process. The exposure of procoagulant phospholipids, known as phosphatidylserine (PS), on the activated platelet surface and the surface of endothelial cells provides a surface for attachment of clotting factors. In addition, the release of coagulation factors from α-granules during platelet activation promotes the coagulation cascade forming a fibrin mesh and subsequent thrombus formation (Furie and Furie, 2008, Heemskerk et al., 2002). The increased thrombin formation at the site of injury forms a positive feedback process which reinforces further platelet activation (Gibbins, 2004).

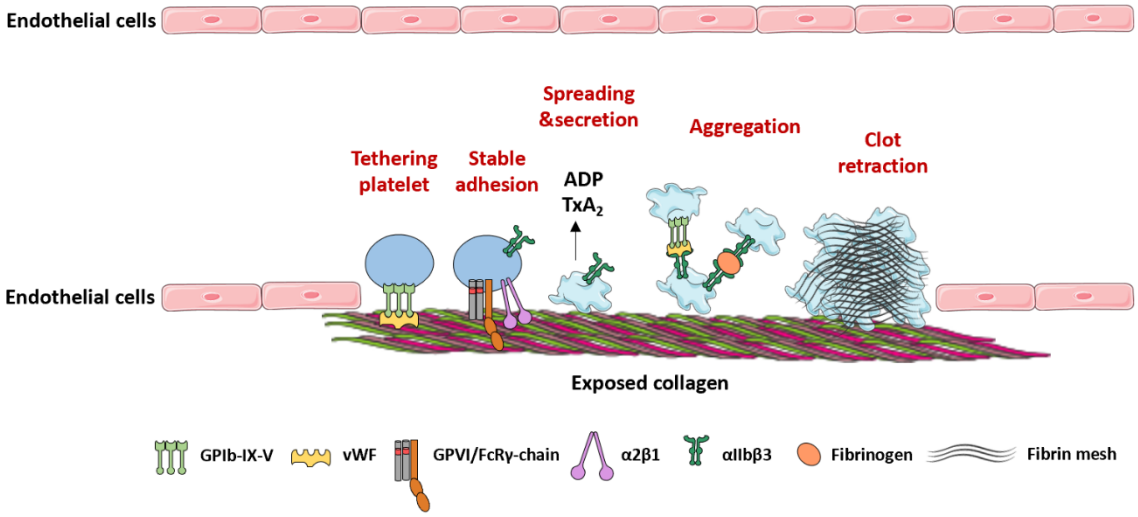


Figure 1.5. Stages of thrombus formation. Subendothelial collagen become exposed upon vascular injury. Platelet tethering is initiated after the binding of platelet surface receptor GPIb-IX-V with VWF. Binding of GPVI with collagen initiates rapid and sustained activation which causes integrin conformational changes from low affinity to high affinity andthese interactions strengthen the adhesion to exposed collagen and VWF. Platelets then spread and release their secretion contents such as ADP and TxA_2 which cause platelet aggregation and subsequent thrombus formation. Blue circle represent resting platelets. Curly blue shape indicates activated platelets. Figure amended based on (Jackson, 2011, Wei et al., 2009).

1.9 Platelet lifespan

Non-activated platelets remain in the blood circulation for 7-10 days due to a fine balance between platelet production and clearance which differs from individual to individual. Platelets are usually cleared from the circulation by phagocytosis in the liver and spleen (Lebois and Josefsson, 2016). The regulation of platelet production and clearance is important to prevent excessive platelet production which may cause spontaneous thrombosis or too much platelet clearance leading to excessive bleeding. Platelets, like megakaryocytes, express the c-Mpl receptors and regulate the TPO level in circulation either by binding or removing it from the circulation. In the case of low platelet counts, less TPO binds to platelets and this leads to the stimulation of the thrombopoiesis process to boost the level of platelets in the circulation (Folman et al., 2000, Hobisch-Hagen et al., 2000). Platelet lifespan can be regulated by an intrinsic balance between apoptotic and survival mechanisms. Pro-survival protein Bcl-X_L (encoded by BCL2L1) maintains platelet survival by preventing the pro-apoptotic protein Bak activity (Mason et al., 2007). However, when platelets age, the level of Bcl-X_L is diminished as a result of degradation of the inherited protein from the megakaryocytes (Bertino et al., 2003). Consequently, the low level of Bcl-X_L is insufficient to supress the pro-apoptotic proteins Bak and Bax which leads to platelet death (Mason et al., 2007).

1.10 Platelet function testing

Platelet function can be disrupted which leads to impaired haemostasis upon vascular injury. Platelet function testing (PFT) is important for the diagnosis, treatment and management of bleeding episodes in affected individuals which can have a significant impact on their life. There are several functional assays that can investigate platelet

function in patients with suspected bleeding disorders. These assays test different parameters such as platelet adhesion, aggregation, secretion and expression of cell surface markers upon platelet activation, in order to produce quantitative and qualitative results regarding overall function. Although many of these tests are mainly used in clinical practice, PFT is continually evolving to produce more rapid, efficient and accurate tests that can be used in both clinical and research settings. The first reported platelet function test was in 1910, when Duke developed the bleeding time test (BT). The test was designed to evaluate the extent of bleeding from a small skin wound by estimating the time required for platelets to form a plug (Duke, 1910). As the Duke method has poor reproducibility and sensitivity, this BT test is now used less frequently.

1.10.1 Light transmission aggregometry

The BT test has been replaced by a platelet aggregation assay using platelet rich plasma (PRP), which is known as light transmission aggregometry (LTA). It was developed by Born in 1962 where he measured the ability of platelets to aggregate to each other using the addition of aggregating agonists, such as adenosine-diphosphate (ADP) and arachidonic acid (AA) (Paniccia et al., 2015, Born, 1962). LTA is still the gold standard platelet function assay and is widely used to investigate the function of platelets by monitoring their response to variable concentrations of agonists. It simply measures the light transmitted through PRP or washed platelet samples in real time. When agonists are added to samples, normal platelets are activated and start aggregating. Hence, the aggregation allows more light to be transmitted through samples as the solution becomes clearer (Kottke-Marchant and Corcoran, 2002). Moreover, LTA can also be coupled with lumiaggregometry in order to simultaneously measure the secretion of ATP from platelet granules. This can be achieved by adding

luciferin-luciferase reagents into platelet mixtures where luciferase enzyme oxidise luciferin in the presence of ATP and consequently produce a fluorescent signal that can be quantified (Holmsen et al., 1972). An increase in fluorescent signals indicates the high concentration of ATP secretion in a sample. Aggregometry assays have also been developed by introducing a 96 well platelet aggregation assay in 1990 (Fratantoni and Poindexter, 1990). The principle is based on LTA with some modifications. PRP samples are added to 96 well plate containing variable concentrations of agonists and the percentage of aggregation is calculated based on the light absorbance of samples and controls. Multiple electrode aggregometry (MEA) has also been used to measure platelet aggregation in whole blood (Toth et al., 2006).

1.10.2 Flow cytometry

A flow cytometry assay is widely used to measure the platelet secretion and the expression of platelet cell surface markers. The principle of flow cytometry is based on the passage of single cells through a laser, where they can be detected, counted and sorted. When fluorescent labelled cells are excited by the laser, they emit light with different wavelengths which can be detected by the detector. Flow cytometry can also be used to evaluate platelet activation by measuring the platelet response to agonists from platelet granule secretion. Here fluorescently labelled P-selectin is detected as an indication of platelet activation and α granule relase. In addition, it can measure the quantity of platelet glycoproteins in a tested sample which can explain the causes behind some platelet function and aggregation disorders (Rubak et al., 2016, Gordon et al., 1995, Michelson, 1996).

1.10.3 Platelet adhesion assays

The capability of platelets to adhere and aggregate to a coated surface under shear stress can be assessed using platelet adhesion assays which mimics the physiological condition. Platelet adherence can be determined using a microfluidic system and surfaces coated with a variety of adhesive molecules, predominantly collagen and fibrinogen. However, there are some commercial available systems such as the Platelet Function Analyser PFA-100™ and Total Thrombus Formation Analysis System (T-TAS) (Kundu et al., 1995, Hosokawa et al., 2012). The aggregation and adhesion of platelets under high shear stress can be tested using the PFA-100™ which provides a quantitative measurement of platelet function in anticoagulated whole blood. It uses a disposable test cartridges containing a membrane coated with collagen or epinephrine and measures the time required to form a full plug on the membrane aperture (Kundu et al., 1995, Harrison, 2005b). The quantitative analysis of thrombus formation under various flow conditions can be evaluated by using (T-TAS), which uses a microchip based flow chamber coated with collagen to measure the blood flow pressure in real time (Hosokawa et al., 2012). The formation of thrombi in the flow chamber is monitored and recorded by a video microscope which tracks changes in blood flow pressure.

1.11 Inherited platelet disorders-symptoms and treatments

There are many disorders that can affect the function (qualitative disorders) and/or number (quantitative disorders) of platelets which lead to impaired haemostasis. These disorders can be inherited or acquired (Pai and Hayward, 2009). Acquired platelet disorders are much more common than inherited and can result typically from drug intake and/or dietary factors but most often due to medical conditions particularly immune disorders (Hassan and Kroll, 2005). However, inherited platelet disorders

(IPDs) are a large group of heterogeneous rare diseases that result from an inherited defect of platelet function and/or number. They are often associated with variable bleeding tendency ranging from severe bleeding, which can be recognized within a few weeks after birth, to mild bleeding that may remain undetected in adulthood (Drachman, 2004a, Israels et al., 2011). Symptoms of these disorders include: epistaxis, petechiae, cutaneous bruising, bleeding from mouth and gums, bleeding after tooth extraction or surgery, menorrhagia and post-partum haemorrhage. IPDs can be classified into two categories: Platelet function disorders (PFDs) or disorders of platelet number (thrombocytopenia or thrombocythemia) (Lambert, 2019, Maclachlan et al., 2017). When the platelet count is less than $150x10^9$/L (normal range 150-400 $x10^9$/L), it is considered to be thrombocytopenia which may result in variable bleeding episodes. Platelet function disorders result from various underlying causes as a result of the defective components of the haemostatic pathway such as receptors of adhesion (GPIb-IX-V), aggregation (Integrin αIIbβ3, GPIIb-IIIa), soluble agonists, the secretion of granules and the activation pathway (Nurden et al., 2012).Some disorders can present with only thrombocytopenia or accompanied by secondary platelet function defect (Cattaneo, 2003). IPDs are clinically rare and difficult to diagnose because they may be misdiagnosed with acquired disorders. Thus, careful assessment of patient's history of bleeding episodes and medical care is required when diagnosing patients with a bleeding disorder (D'Andrea et al., 2009). The majority of patients with IPDs have normal platelet counts (150-$400x10^9$/L) and mild bleeding symptoms, but they are at risk of severe bleeding at times of surgery or trauma when platelet function is challenged (Watson et al., 2013b). Although many affected individuals do not require treatment for their bleeding condition, careful management and monitoring are required to prevent the possibility of bleeding during

pregnancy and surgical operations. To understand the pathophysiology of these disorders, it is often best to address all genes involved in megakaryocyte differentiation and platelet formation and how they are functionally disrupted (Leo et al., 2015, Rabbani et al., 2014).

1.11.1 Treatment of inherited platelet disorders

Careful clinical management of patients with bleeding disorders is necessary to prevent bleeding risk. This can be accomplished in consultation with the patient's primary physician and haematologist prior to any planned procedure, in order to provide appropriate management and treatment (Patton and Ship, 1994). There are some therapies which can be used to treat individuals with bleeding episodes most often given to patients prior to a planned procedure or after exposure to trauma. One of the most common therapies is Desmospressin (DDAVP) which is mainly used to prevent bleeding particularly in patients with mild haemophilia A and von Willebrand disease. It is a synthetic analog of the antidiuretic hormone vasopressin which increases the level of factor VIII and von Willebrand factor in plasma and can subsequently prevent bleeding (Cattaneo, 2002). Antifibrinolytic medications such as tranexamic acid can also be used orally or intravenous (IV) in cases of mild bleeding, including epistaxis and menorrhagia. This treatment prevents the degradation of the fibrin clot by blocking the lysine binding site on plasminogen to interact with fibrin (Wellington and Wagstaff, 2003). In cases of severe bleeding or surgical operation with known high risk of bleeding, platelet transfusion is recommended in order to avoid this (Bolton-Maggs et al., 2006). However, platelet transfusion from multiple donors can lead to risk of developing alloantibodies. To prevent undesirable alloimmunisation, human leukocyte antigen (HLA) matched single donor platelets or leukocyte depleted blood components are used. The recombinant factor VIIa (FVIIa) is also an alternative

therapy that can be used particularly in patients with Glanzmann thrombocytopenia (GT) when there is a lack of response to platelet transfusion due to alloimmunisation (Poon et al., 2004). Furthermore, there has been a significant advancement in the disease therapies including the use of TPO mimetics, haematopoietic stem cell transplantation and gene therapy, which can be used for patients with inherited platelet-based bleeding disorders (Locatelli et al., 2003, Boztug et al., 2010, Pecci et al., 2010).

1.11.2 Platelet function disorders

Platelet function disorders (PFDs) compromise a large group of heterogeneous disorders that are characterised with variable bleeding tendencies (Table 1.1). The majority of hereditary PFDs with severe bleeding are well characterised by using conventional platelet function tests such as LTA and flow cytometry. The genes associated with PFDs are shown in Figure 1.6.

1.11.2.1 Disorders of platelet adhesion and aggregation

Bernard-Soulier Syndrome (BSS; OMIM ID #231200) is a rare inherited platelet disorder that results from genetic defects in the genes of *GPIBA*, *GPIBB* or *GP9* while genetic defect in the gene of *GP5* is not associate with BSS. The molecular defects in these genes leads to absent or reduced expression of glycoproteins (GP) GPIbα, GPIbβ or GPIX on the surface of platelets, which affect the function of GPIb-IX-V receptor complex on the membrane surface of platelets. As a result, the binding of GPIb-IX-V receptor complex with VWF on the exposed subendothelium matrix is impaired which could prevent platelet adhesion (Nurden et al., 2012, Cattaneo, 2003). Most affected individuals with BSS who present with severe bleeding phenotypes are inherited in an AR manner. However, there are a few individuals in the literature

presented with asymptomatic or mild bleeding phenotype which are transmitted in an AD manner (Othman and Emsley, 2017). Most affected individuals with BSS present with variable thrombocytopenia and typically large platelet size (macro-thrombocytopenia).

Glanzmann Thrombasthenia (GT; OMIM ID #273800 and #619267) is a group of rare AR syndromes, whichaffects the megakaryocyte lineage and cause defective platelet aggregation. GT is caused by mutations in the genes encoding ITGA2B and ITGB3 that lead to a reduction in the number and abnormal function of subunits of the alpha-IIb/beta-3 integrin receptor (Nurden, 2006). This receptor mediates the platelet interaction with fibrinogen at low shear and VWF at high shear, to form platelet aggregates at the site of the injury and finally formation of thrombi. However, in GT platelets fail to interact with the adhesive proteins that attach to the aggregated platelets (Cattaneo, 2003, Nurden et al., 2012). Patients with GT present with normal platelet size and counts as well as variable bleeding episodes ranging from mild to severe with mainly mucocutaneous bleeding (Hardisty et al., 1992).

Platelet type–von Willebrand disease (PT-VWF) is a rare autosomal dominant bleeding disorder that results from a gain of function mutation in the platelet GPIbα protein encoded by the *GP1BA* gene. This leads to an excessive binding of GPIbα with its ligand VWF (Miller et al., 1991). As a result, platelets have an affinity to bind with the high molecular weight (HMW) multimers of VWF which are subsequently removed from the blood circulation. Patients with PT-VWF present with mild to moderate mucocutaneous bleeding and a variable level of thrombocytopenia (Othman et al., 2016).

Scott syndrome (SS; OMIM ID #262890) is known as a very rare AR congenital disorder, which is caused by a defect in the platelet phospholipid membrane. As a result, it causes a lack of procoagulant phosphatidylserine (PS), which subsequently leads to impaired platelet coagulant activity upon platelet activation (Toti and Freyssinet, 2005, Zwaal et al., 2004). It is caused by mutations in *ANO6* encoding phospholipid scramblase protein TMEM16F (Millington-Burgess and Harper, 2020).

1.11.2.2 Disorders of platelet secretion and signalling

Platelet secretion disorders result from deficiency of platelet granule secretion upon platelet activation, which can involve platelet α or dense granules, but rarely in both at the same time (Simon et al., 2008). Gray Platelet Syndrome (GPS; OMIM ID #139090) is characterized by a deficiency in α granule content which results in a platelet function defect. It was named GPS because of platelets appearing gray in blood smears as result of the absence of α granules (Raccuglia, 1971). It is associated with enlarged platelets and mild thrombocytopenia with moderate to severe bleeding as a result of biallelic mutations in *NBEAL2*, the gene encoding the neurobeachin- like-2 protein (Kahr et al., 2011).

Hermansky-Pudlak syndrome (HPS) is a group of 10 AR multi-system disorders caused by mutations in the *HPS* gene family (*HPS1-HPS10*) and associated with oculocutaneous albinism, bleeding diathesis and, in some cases, pulmonary fibrosis (Hermansky and Pudlak, 1959, Huizing et al., 2020, Suzuki et al., 2002). This is caused by deficiency in dense granules, in which patient platelets lack their contents which consequently leads to impaired platelet aggregation and secretion.

Chediak-Higashi syndrome (CHS; OMIM ID #214500), is a rare inherited AR disorder characterised by variable oculocutaneous albinism, recurrent infection and mild

bleeding phenotypes also as a result of a dense granule deficiency. It is caused by mutations in the *CHS1* (*LYST*) gene that is associated with lysosomal trafficking (Zarzour et al., 2005).

Moreover, platelet activation is stimulated through interaction between platelet membrane receptors called G-protein-coupled receptors (GPCRs) with their secreted agonists. Deficiency or dysfunction of these receptors can cause impaired platelet function with variable bleeding tendency (Hollopeter et al., 2001). P2Y1 and P2Y12 are two distinct receptors that expressed on the surface of human platelets which bind to ADP upon platelet activation (Cattaneo, 2011). Congenital defects or deficiency in P2Y12 receptors has been identified to cause inherited bleeding disorder platelet-type 8in an AR manner which is characterised with excessive bleeding as a result of impaired ADP-induced aggregation (Cattaneo et al., 1992, Nurden et al., 1995).

Table 1.1. The main inherited platelet function disorders. The table includes a detailed description of the disorders, including the name of the disorders, the causative genes and their phenotypes and any other findings related with the disorders. All the disorders can be found in detail in (Bolton-Maggs et al., 2006).

Platelet function disorders	**Disorders**	**Gene**	**Phenotype and other findings**
Disorders of platelet adhesion and aggregation	Bernard Soulier syndrome	*GP9* *GP1BA* *GP1BB*	Thrombocytopenia, large platelets Anomalies in components of the GPIb-IX-V complex
	Glanzmann thrombasthenia	*ITGA2B* *ITGB3*	Normal platelet count, Absent aggregation with all agonists
	$P2Y_{12}$ receptor defects	*P2Y12*	Normal platelet count, Transient aggregation
	Scott syndrome	*ANO6*	Normal platelet count, impaired platelet coagulant activity upon platelet activation
	Thromboxane A_2	*TBXA2R*	Normal platelet count, Aggregation reduced in response to arachidonic acid and U46619
	von Willebrand disease	*VWF*	Thrombocytopenia, Reduced or absent VWF multimers
Disorders of platelet secretion and signalling	Chediak-Higashi syndrome	*LYST*	Normal platelet count, Skin and hair hypopigmentation Immunodeficiency Reduced α-granules
	Grey Platelet syndrome	*NBEAL2*	Thrombocytopenia, large platelets Grey platelet colour on blood film Absent α-granules
	Hermansky-Pudlak syndrome	*HPS1-9*	Normal platelet count Decreased/absent dense granules

			Oculocutaneous albinism
	Jacobsen/ Paris-Trousseau syndrome	*FLI1*	Thrombocytopenia, large platelets Developmental delay, facial abnormalities
	Quebec platelet disorder	*PLAU*	Thrombocytopenia Urokinase plasminogen activator storage in platelets

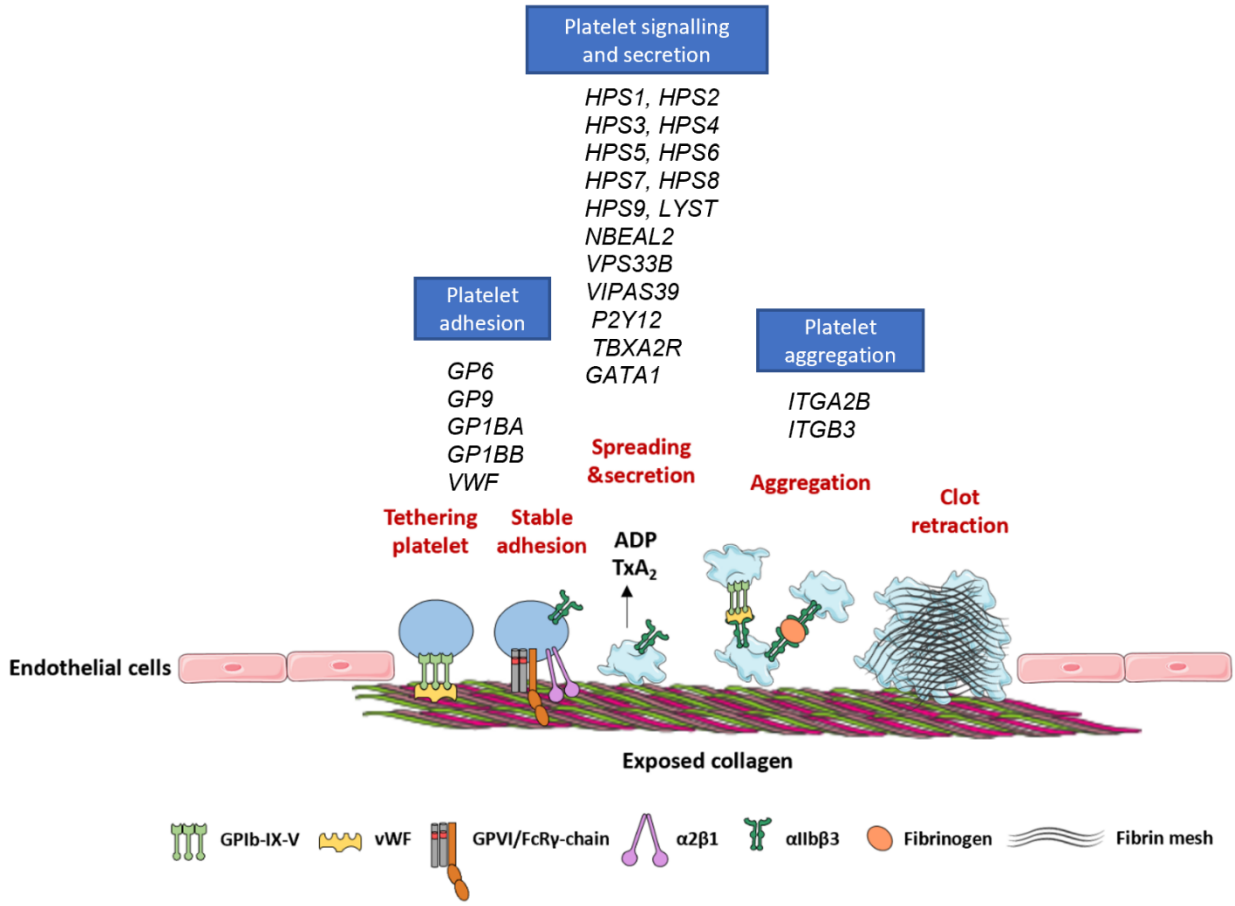


Figure 1.6. Platelet function associated genes. The carton depicts the genes associated with platelet function including platelet adhesion, signalling and secretion as well as aggregation.

1.11.3 Inherited thrombocytopenia

Inherited Thrombocytopenia (IT) is comprised of heterogeneous group of disorders characterized by low platelet counts and variable bleeding diathesis as the main features, often with abnormal platelet function resulting in impaired haemostasis (Balduini et al., 2018). In 1948, thrombocytopenia was firstly identified as a cause of bleeding particularly in immune system disorders. Since then, the Mendelian inheritance disease pattern of IT was discovered initially in Bernard-Soulier syndrome (BSS) (Noris and Pecci, 2017). Advancement in clinical and scientific research has led to an increased understanding of the molecular defects in patients with IT. These defects are variable in severity, ranging from severe bleeding, which can be recognized within a few weeks after birth, to mild bleeding that may remain undiagnosed until incidental recognition during routine blood testing in adulthood (Drachman, 2004b). They manifest with different symptoms including: epistaxis, easy bruising, petechiae, prolonged bleeding from cuts, gum bleeding, excessive bleeding after surgery, hematuria and menorrhagia in women (Johnson et al., 2018, Johnson et al., 2016b). As bleeding is considered the main clinical complication for patients with IT, some patients with common ITs have the propensity to develop other disorders such as hematological malignancies and kidney failure (Noris and Pecci, 2017, Morgan and Daly, 2017). Although there are other causes of thrombocytopenia, such as infections and immune disorders, IT is primarily caused by mutations in genes involved in megakaryocyte differentiation, maturation and platelet release (Balduini et al., 2017). Since the last decade, NGS technologies, namely Whole Exome Sequencing (WES) and Whole Genome Sequencing (WGS) coupled with conventional Sanger sequencing and *in-silico* bioinformatic tools have

been used in conjuction to uncover novel genes with a pivotal role in megakaryocyte biology and platelet biogenesis (Greinacher and Eekels, 2019b, Savoia, 2016).To date, more than 40 genes have been reported to cause different forms of IT, which reflects the immense difficulty in identifying a single causative gene (Table 1.2; Figure 1.6) (Johnson et al., 2016a, Nurden and Nurden, 2020). These genetic variants have various clinical manifestations, phenotypic presentations and sometimes associated with secondary qualitative defects in platelet function (Johnson et al., 2016b). Diverse platelet phenotypes mean there are several approaches in which they can be characterized. One such way is to classify genes based on their influence on megakaryocyte differentiation, platelet production, and removal, and will be discussed below (Johnson et al., 2016a). However, despite these advancements, nearly 50% of patients with IT of unknown genetic etiology still remain undiagnosed (Johnson et al., 2018, Savoia, 2016).

Table 1.2. Genetic causes of inherited thrombocytopenia and their associated syndromes.

Platelet abnormality	Gene	disorders	Inheritance	Other syndromic and features	Reference
Megakaryopoiesis	*ANKRD26*	*ANKRD26*-related thrombocytopenia	Autosomal dominant	Predisposition to leukaemia. Reduction of platelet α-granules. Normal in vitro platelet aggregation and mean platelet volume. Some patients have high level of haemoglobin and leukocyte.	(Bluteau et al., 2014, Pippucci et al., 2011)
	ETV6	*ETV6*-related thrombocytopenia	Autosomal dominant	Leukaemia predisposition. High erythrocyte mean corpuscular volume (MCV). Some patients have elevated red cell MCV.	(Noetzli et al., 2015)
	FLI1	Paris-Trousseau thrombocytopenia /Jacobsen syndrome	Autosomal Dominant/ recessive	Abnormal development of heart and face. Intellectual disabilities. Large α-granules. Abnormal MKs morphology. Normal RBCs and WBCs counts. Moderate thrombocytopenia.	(Stevenson et al., 2015)
	FYB	*FYB*-related thrombocytopenia	Autosomal recessive	Small platelets. Reduction of mature MKs in BM. Significant bleeding tendency. Normal WBCs count. Low mean platelet volume MPV. Mild iron deficiency anaemia.	(Koren et al., 2015)
	GATA1	GATA1-related disease: X-linked thrombocytopenia (XLT) and X-linked thrombocytopenia with thalassemia (XLTT)	X-linked recessive	Dyserythropoietic anaemia. Macrothrombocytopenia. Beta-thalassemia. Congenital erythropoietic porphyria. Erythrocyte abnormalities. Splenomegaly	(Freson et al., 2017)

	GFI1B	Macrothrombocytopenia and platelet function defects	Autosomal dominant	Macrothrombocytopenia. Red cell anisopoikilocytosis. platelet dysfunction. Reduction of platelet α-granules.	(Stevenson et al., 2013)
	HOXA11	Amegakaryocytic thrombocytopenia with radio-ulnar synostosis	Autosomal dominant	Bilateral radioulnar synostosis. Severe bone marrow failure. Cardiac and renal malformations. B-cell deficiency. Hearing loss. Clinodactyly. Some patients show skeletal anomalies. Some patients have developed pancytopenia.	(Horvat-Switzer and Thompson, 2006)
	MECOM	Congenital amegakaryocytic thrombocytopenia and radioulnar synostosis	Autosomal dominant		(Germeshausen et al., 2018)
	MPL	Congenital amegakaryocytic thrombocytopenia (CAMT)	Autosomal recessive	Absence or reduced of MKs in BM. No physical anomalies. Development to BM aplasia in infancy.	(Ihara et al., 1999)
	NBEAL2	Grey Platelet Syndrome	Autosomal recessive	Impaired platelet function. Severe reduction of platelet α-granules contents. Large platelets. Development of myelofibrosis and splenomegaly in some patients. Abnormalities in megakaryocyte development.	(Pluthero et al., 2018)
	RBM8A	Thrombocytopenia-absent radius syndrome	Autosomal recessive	Bilateral radial aplasia. Elevated haemoglobin level in patients with 5'UTR SNP. Normal WBCs count and some patients have leucocytosis and eosinophilia. Anaemia. Skeletal, urogenital, kidney and heart defects. Reduced MKs in BM.	(Manukjan et al., 2017)

	RUNX1	Familial platelet disorder with propensity to acute myelogenous leukemia (FPD/AML)	Autosomal dominant	Platelet defects. Variable platelet counts. Reduction in dense granule secretion observed in secondary qualitative abnormality. Myelodysplasia. Reduced response to several platelet agonists.	(Morgan and Daly, 2017)
	SLFN14	*SLFN14*-related thrombocytopenia	Autosomal dominant	Giant platelets. Decreased ATP secretion. Reduced number of dense granules.	(Simpson et al., 2015)
	SRC	SRC-related thrombocytopenia	Autosomal dominant	Myelofibrosis, bleeding, and bone pathologies. Hypercellular bone marrow with trilineage dysplasia. Platelets are dysmorphic and variable in size. Splenomegaly and congenital facial dysmorphism.	(Turro et al., 2016)
	THPO	Inherited thrombocytopenia from monoallelic *THPO* mutation	Autosomal dominant	Bone marrow aplasia. Normal or enlarged platelet morphology.	(Marconi et al., 2017)
	PTPRJ	Inherited thrombocytopenia	Autosomal recessive	syndromic thrombocytopenia characterized by spontaneous bleeding, small-sized platelets. Impaired platelet function.	(Wen and Wang, 2019)
	GALE	Inherited thrombocytopenia	-	Dysplastic megakaryocytes. Some patients have mild anemia and febrile neutropenia. Big and pale platelets. Galactosemia, hypotonia, seizures, jaundice, galactosuria and hepatomegaly.	(Seo et al., 2019)
	IKZF5	Inherited thrombocytopenia	Autosomal dominant	-	(Lentaigne et al., 2019)
	MASTL	Inherited thrombocytopenia	Autosomal dominant	-	(Hurtado et al., 2018)

	NF-E2	Inherited thrombocytopenia	-	-	(Luk et al., 2020)
	ACTIN1	*ACTN1*-related thrombocytopenia	Autosomal dominant	Congenital macrothrombocytopaenia. Anisocytosis. Absent or mild bleeding diathesis.	(Kunishima et al., 2013)
Platelet production	*CYCS*	CYCS-related thrombocytopenia	Autosomal dominant	Normal platelet size and volume.	(Ong et al., 2017)
	GNE	*GNE* myopathy with congenital thrombocytopenia .	Autosomal recessive	Rimmed vacuoles. Hematological complications are rare. Proteinuria and hematuria in some patients. Membranoproliferative glomerulonephritis. Platelets size are normal to large.	(Futterer et al., 2018, Revel-Vilk et al., 2018)
	GP1BA *GPIBB* *GP9*	Bernard-Soulier Syndrome (BSS)+ platelet type von-Willebrand disease (PTvWD)	Autosomal Dominant/ recessive	Macrothrombocytopaenia. Severe bleeding tendency with platelet function defect. Platelet anisocytosis.	(Othman and Emsley, 2017) (Sivapalaratnam et al., 2017b) (Wright et al., 1993)
	MYH9	*MYH9*-related disease (MYH9-RD)	Autosomal dominant	Congenital macrothrombocytopaenia. Mild bleeding tendency. Development of kidney dysfunction, deafness, cataracts and Döhle-like bodies. Elevated liver enzymes.	(Balduini et al., 2011)
	ITGA2B *ITGB3*	Glanzmann thrombasthenia	Autosomal recessive	Impaired platelet function.	(Burk et al., 1991) (Nurden et al., 2013)
	PRKACG	*PRKACG*-related thrombocytopenia	Autosomal recessive	Giant platelet. Impaired platelet function.	(Manchev et al., 2014)

	TRPM7	TRPM7-related thrombocytopenia	Autosomal dominant	Macrothrombocytopaenia. Atrial fibrillation.	(Stritt et al., 2016a)
	TPM4	Tropomyosin 4-related thrombocytopenia	-	Macrothrombocytopaenia. All other blood cell counts are normal. Mild effect on platelet function.	(Pleines et al., 2017)
	TUBB1	*TUBB1*-related thrombocytopenia	Autosomal dominant	Congenital macrothrombocytopaenia.	(Kunishima et al., 2009)
	WAS	Wiskott-Aldrich syndrome, X-linked thrombocytopenia (XLT)	X-linked recessive	Mild or severe immunodeficiency, hematopoietic malignancies and eczema. Thrombocytopenia with small platelets. Autoimmune haemolytic anaemia.	(Massaad et al., 2013)
	FLNA	Filaminopathies A	X-linked recessive	X-linked dominant form of periventricular nodular heterotopia (FLNA-PVNH) and the otopalatodigital syndrome spectrum of disorders. Haemorrhage and coagulopathy. Abnormal platelet morphology.	(Nurden et al., 2011)
	DIAPH1	Macrothrombocytopenia (MTP) and hearing loss	Autosomal dominant	_	(Stritt et al., 2016b)
	ABCG5 ABCG8	Macrothrombocytopenia associated with Sitosterolemia.	Autosomal recessive	Xanthomas and premature coronary atherosclerosis due to hypercholesterolemia. Hematologic abnormality.	(Bastida et al., 2017) (Bardawil et al., 2017)
Platelet clearance/ other	*STIM1*	Stormorken Syndrome and York platelet syndrome	Autosomal dominant	Tubular myopathy and congenital miosis. Severe immune dysfunction.	(Shahrizaila et al., 2014, Markello et al., 2015)

	ORAI1	Stormorken syndrome	Autosomal dominant	CRAC channelopathy. Severe combined immunodeficiency, autoimmunity, muscular hypotonia, and ectodermal dysplasia.	
	VWF	Von Willebrand disease type IIB	Autosomal Dominant/ recessive	_	(Cooney et al., 1991)
	ADAMTS13	Thrombotic thrombocytopenia purpura	Autosomal recessive	Upshaw-Schulman syndrome. Anaemia.	(Levy et al., 2001)

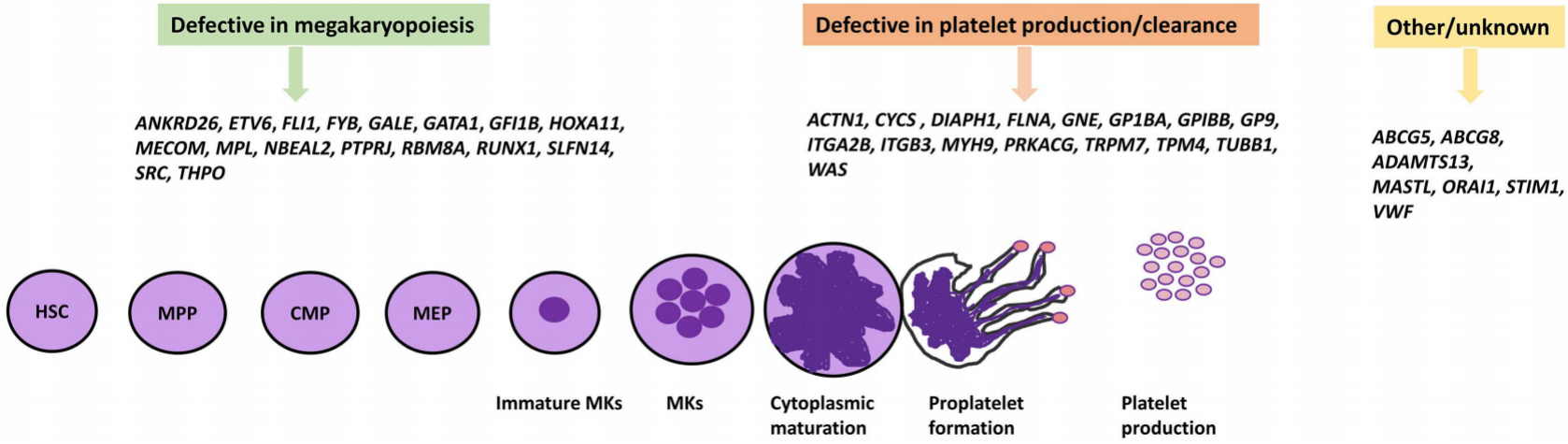


Figure 1.7. Inherited thrombocytopenia causative genes involved in megakaryopoiesis, platelet formation, and others. The differentiation of platelets from HSCs proceeds by multiple different cell lineages which involve many genes encoding a number of transcription factors and proteins. Genetic defects in these genes have been shown to cause IT. HSC, Hematopoietic stem cell; MPP, Multi-Potent Progenitor; CMP, Common myeloid progenitor; MEP, Megakaryocyte-erythroid progenitor; MK, Megakaryocyte (Almazni et al., 2019, ALHENC-GELAS et al., 2010).

1.11.3.1 Disorders of megakaryopoiesis

As described earlier,the process of megakaryopoiesis and thrombopoiesis involves a complicated biological series of events that are highly regulated. There are several disorders associated with megakaryopoiesis and proplatelet formation and release. The process of megakaryopoiesis and thrombopoiesis involves multiple genes and transcriptions factors which play important roles in megakaryocyte differentiation, platelet formation, and release (Johnson et al., 2016a, Schulze and Shivdasani, 2005). IT can result from defects in these genes which present with variable phenotypic display and clinical presentation. Some ITs result from defective differentiation of haemopoietic progenitor cells to MKs, leading to reduction or absence in the number of bone marrow MKs and insufficient numbers of platelets.

1.11.3.1.1 Defects in megakaryocyte differentiation and maturation

THPO-related thrombocytopenia is an AD form of IT that is caused by a monoallelic mutation in the *THPO* gene characterized by normal or slightly large platelet size but no bleeding tendency (Marconi et al., 2017). The main defective mechanism in several forms of IT is a change in MK maturation which therefore leads to the production of immature and dysfunctional MKs. However, the differentiation and maturation of MKs is regulated by several transcription factors; one of which is *GATA1*. *GATA1* is highly expressed in the erythroid and megakaryocytic lineages, and plays a vital role in the maturation and development of erythroid cells and megakaryocytes (Geddis, 2010). X-linked thrombocytopenia with thalassemia and X-linked thrombocytopenia with dyserythropoietic anemia are both caused by mutations in *GATA1,* resulting in impaired MK and erythroid cell maturation (Freson et al., 2017). As a consequence, *GATA1* mutated patients present with large platelets and reduced α- granule contents. They also display a variable degree of anemia and abnormal

morphology in red blood cells (Millikan et al., 2011). Additional transcription factors known to be involved in the maturation of megakaryocytes are *RUNX1*, *ETV6, ANKRD26*, *FLI-1***,** and the transcriptional repressor *GFI1B* acting by binding to promoter regions in MK expressed genes. Thus, multiple mechanisms in MK and platelet maturation are affected as result of alterations in these genes (Tijssen and Ghevaert, 2013a, Foudi et al., 2014). A previous study identified a point mutation in the third helix of *HOXA11* homeodomain causing an inherited syndrome of congenital amegakaryocytic thrombocytopenia and radio-ulnar synostosis (Horvat-Switzer and Thompson, 2006).

Thrombocytopenia absent radii (TAR) syndrome results from a combination of a microdeletion on Chromosome 1 including the *RBM8A* gene, alongside a low frequency non-coding single nucleotide polymorphism (SNP) within the regulatory region of *RBM8A* (Manukjan et al., 2017). As a consequence, hematopoietic progenitors from patients with TAR syndrome fail to differentiate into MKs *in vitro* (Letestu et al., 2000). Variants in the 5'UTR of *ANKRD26* cause familial thrombocytopenia type-2 (THC2) with propensity to leukemia, which results in loss of *RUNX1* and *FLI1* binding and prevents gene silencing (Bluteau et al., 2014). Moreover, heterozygous variants specifically located in the promoter region between c.-134G and c.-113 region highly affects gene expression. Patients with THC2 are characterized by small MKs with hypolobulated nuclei as a result of dysmegakaryopoiesis (Loffredo et al., 2011). A mutation in the *FYB1* gene has recently been identified to cause IT and although the exact mechanism of the mutation is still ambiguous, it has been suggested that thrombocytopenia arises from a reduction of mature MKs in the bone marrow and synthesis of small platelets (Koren et al., 2015). A novel transcriptional factor gene *IKZF5* was recently identified

to regulate megakaryopoiesis. Genetic analysis of unrelated individuals with isolated thrombocytopenia has identified novel rare missense variants in *IKZF5* which cause hereditary thrombocytopenia without any syndromic abnormality. In addition, a novel homozygous frameshift mutation (c.952delA) was identified within nuclear factor erythroid 2 (*NF-E2*) gene in new born patients who developed multiple petechiae and thrombocytopenia two days after birth (Luk et al., 2020). *NF-E2* is considered as heterodimeric transcription factor that is involved in megakaryocyte maturation, pro-platelet formation and release (Motohashi et al., 2010, Lecine et al., 2000, Tiwari et al., 2003).

1.11.3.1.2 Defects in platelet release and survival

After megakaryopoesis, mature MKs undergo essential processing by extending long branches called proplatelets via the bone marrow sinusoids, and subsequently release platelets into the blood circulation (Machlus and Italiano, 2013, Bender et al., 2015). These processes are underpinned by cytoskeletal changes and cellular signaling where in this case mutations disrupt this pathway reducing the circulating platelet count . In mature polyploid MKs, the cytoplasm extends with long beaded cytoplasmic protrusions, due to microtubule sliding. The dimerisation of β1-tubulin with α-tubulin polymerizes into long microtubule bundles inside the MK cortex. A mixed polarity of microtubule bundles runs throughout the extension of proplatelets which are thought to provide fundamental force for microtubule sliding and proplatelet elongation (Bender et al., 2015). *TUBB1* encodes for β1-tubulin and mutations within *TUBB1* are associated with an autosomal dominant form of IT known as a congenital macrothrombocytopenia (Kunishima et al., 2009). WASp is a multidomain protein belonging to a family of actin nucleation-promoting factors (NPFs) which are specifically expressed in hematopoietic cells. WASp plays an important role in actin

polymerization by transmission of surface signals via the actin-related protein (Arp)2/3 complex (Massaad et al., 2013, Snapper and Rosen, 1999). Mutations have been identified in the *WAS* gene which cause a rare X-linked disorder called Wiskott-Aldrich syndrome (WAS). Patients are characterized by micro-thrombocytopenia and immunodeficiency with predisposition to malignancies (Kirchhausen and Rosen, 1996).

1.11.3.2 Importance of Diagnosing IPDs

IPDs can result from impaired platelet function (thrombocytopathy), reduced platelet count (thrombocytopenia) or both. Identification of the genetic causes in patients with IPDs is challenging particularly with ITs because many patients may be misdiagnosed with acquired thrombocytopenia such as immune thrombocytopenia (ITP). As result, many patients with thrombocytopenic IPDs are subjected to splenectomy and sometimes undergo unnecessary treatment with steroids (Balduini et al., 2013). In addition, some patients with *ANKRD26*-related thrombocytopenia (*ANKRD26*-RD) are subjected to chemotherapy because they have been misdiagnosed with myelodysplastic syndrome (Zaninetti et al., 2017). Platelets are involved in other biological roles beyond haemostasis, such as immunity and inflammation (Kim et al., 2018, Swinkels et al., 2018) therefore, mutations in platelet specific genes may cause functional disruption in haemostasis and the inflammatory response. Based on this, inherited bleeding disorders can be classified into three categories including disorders that (i) affect only platelets, (ii) disorders that are associated with syndromic or non-syndromic phenotypes and (iii) disorders with increased risk of haematologic malignancies. This classification can be used for both diagnostic and prognostic purposes (Noris and Pecci, 2017, Greinacher and Eekels, 2019b). Taken together, genetic counselling and prenatal diagnostic tests would be beneficial to identify genetic causes behind IPDs. Subsequently, this will help with the management and treatment of diseases associated with some syndromic forms. For example, patients with *MYH9*-RD can be managed and treated with renin-angiotensin system blockage agent early to prevent subsequent renal failure (Pecci et al., 2008, Pecci et al., 2014).

1.11.3.3 Syndromic disorders associated with IT

The number of IT forms identified has increased over the last few years since the implementation of NGS. Consequently, it has been shown that the bleeding is not the only clinical phenotype with IT, but patients with some IT forms have propensity to develop more syndromic disorders as result of molecular defects in genes responsible for thrombocytopenia. For instance, haematological malignancies, bone marrow aplasia, skeletal malformation, liver and kidney malfunction and deafness as shown in Figure 1.7. The development of these diseases can be more severe for patients than the bleeding itself (Noris and Pecci, 2017) however, it is still important to recognise if these symptoms are present in relatives. Some syndromic phenotypes associated with ITs are variable between family members or can arise later in life. For example, *MYH9* has an important role in the platelet cytoskeleton and has been found expressed in kidney and inner ear cilia (Marigo et al., 2004). Patients with *MYH9*-reltaed disease (*MYH9*-RD) present with giant platelets, thrombocytopenia and variable bleeding tendency. Although macro-thrombocytopenia is the most distinct presentation during life, many patients with *MYH9*-RD can develop deafness, kidney malfunction and/or cataracts (Pecci et al., 2014). Familial platelet disorder with predisposition to acute myeloid leukaemia (FPD/AML) is characterised by platelet function defects and thrombocytopenia as result of genetic mutations within *RUNX1* gene. The bleeding phenotype is not the only characterisation of disorder, but some patients have a propensity to develop myelodysplasia or leukaemia (Morgan and Daly, 2017).

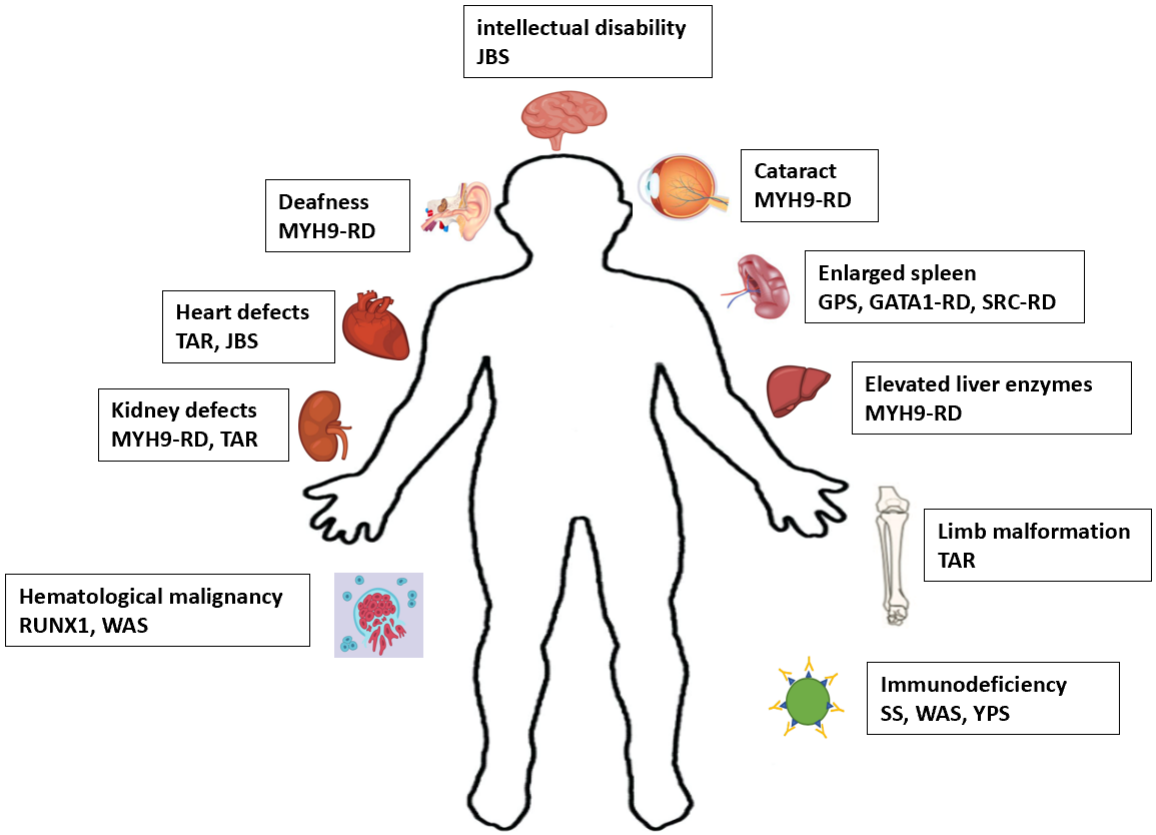


Figure 1.8. Schematic representation of the defects associated with IT. JBS: Jacobsen syndrome, RD: related disease, TAR: Thrombocytopenia-absent radius, GPS: Grey platelet syndrome, SS: Stormorken syndrome, WAS: Wiskott-Aldrich syndrome, YPS: York platelet syndrome (Almazni et al., 2019).

1.11.3.4 Initial diagnostic tools for IPDs

There are several different diagnostic approaches that can be used to diagnose patients with suspected IPDs. The first diagnostic approaches involved a careful historical assessment and physical examination using the Bleeding Assessment Tool scores (BAT-scores) in order to assess the bleeding symptoms in patients with suspected IPDs (Lowe et al., 2013, Perez Botero et al., 2017). Further approaches can be used including complete blood cell counts to assess the platelet number, size and immature platelet fraction in addition to other blood cell types in the case of ITP or anaemia related to MEP differentiation dysfunction (Noris et al., 2016, Ferreira et al., 2017). Peripheral blood films are beneficial to investigate the presence of large platelet size, gray platelets and inclusion bodies within leukocytes (Monteferrario et al., 2014, Ferreira et al., 2017).The second diagnostic tools include platelet function testing such as the gold standard LTA, lumi-aggregometry, functional flow cytometry and flow adhesion assays (Cattaneo et al., 2013). These can be used to assess platelet function defects however, are not suitable in all instances. LTA requires fresh blood samples and at least 80,000 platelets which is often difficult to obtain in patients with ITs. Although this problem can be overcome with whole blood aggregometry in some cases, it is still not suitable for paediatric patients (Harrison et al., 2011, Moenen et al., 2019). The third tier incudes the use of electron-microscopy and western blotting to investigate the expression of platelet proteins (Gresele et al., 2014). Despite the beneficial outcomes of these approaches, the diagnosis of IPDs is still challenging and can be overcome with employing genetic testing as mentioned in the next section (Gresele et al., 2019, Greinacher and Eekels, 2019a).

1.11.3.5 Genetic diagnosis of IPDs

Genetic diagnosis is vital in order to provide patients with a definitive diagnosis for their bleeding symptoms as well as targeted treatments for their disorder. Patients with genetic mutations in *RUNX1* have a predisposition to develop haematological malignancies where the genetic information can be used to monitor the patients' haematological parameters very closely. This emphasises the importance and need for definitive genetic diagnostic tools to provide quick and cost-effective diagnosis for screening patients with IPDs (Johnson et al., 2018, Morgan and Daly, 2017). The molecular basis of IPDs have been elucidated since the adoption of Sanger sequencing and linkage analysis in the 1990s. Recently, Sanger sequencing is considered a low throughput and time-consuming approach which can be used initially as a standard tool to investigate patients in further detail based on clinical description of symptoms and phenotype (Noris and Pecci, 2017). A targeted thrombocytopenia gene panel is a useful approach which can be used as initial screening prior to WES. This targeted panel encompasses all known genes associated with IT which provides a quick and cost effective screening for patients with IT prior to WES and this allowing for a quick genetic diagnosis. The aim of using an IT gene specific panel prior to WES is to filter out patients based on variants in known IT causative genes and subsequently reserving WES for patients with unknown genetic aetiology (Johnson et al., 2018). This could achieve an increase in the efficiency of genetic diagnosis in addition to the reduction of the overall cost. The ThromboGenomics project provided a multi-gene high-throughput sequencing platform (HTS) for the diagnosis of heritable bleeding disorders (Simeoni et al., 2016). The HTS platform covers approximately 96 genes associated with inherited bleeding, thrombotic, coagulation and platelet disorders. The panel showed high sensitivity in detecting causative variants in patients

who had not been previously investigated at the molecular level. It has a high sensitivity to detect variants in the exonic region as well as many of exonic-intronic boundaries and untranslated regions (UTRs) (Simeoni et al., 2016).

1.11.3.6 Next generation sequencing (NGS)

Targeted NGS platforms can be applied to determine the causative genes of IT. As the molecular cause of IT remains unknown in many patients, WGS or WES may be required to uncover more detail of ITs at the molecular level. Several national and international consortia have adopted these approaches to identify disease-causing genes associated with IT. The genes *SLFN14, FYB, STIM1, GFI1b* and *ETV6* are some examples of causative genes detected by these approaches (Noetzli et al., 2015, Koren et al., 2015, Simpson et al., 2015, Shahrizaila et al., 2014, Monteferrario et al., 2014). The results obtained by HTS improves the understanding of the functional role in some causative genes, whose function in platelet production was previously unknown. These techniques will bring substantial benefit to improve our understanding of the molecular mechanisms in megakaryocyte and platelet biogenesis. However, distinguishing pathogenic variants from non-pathogenic variants often requires complex functional and cell lineage studies to prove causality (Johnson et al., 2016b).

1.12 The Genotyping and Phenotyping of Platelets study

In 2009, the United Kingdom Genotyping and Phenotyping of Platelets (UK-GAPP) study was established in Birmingham in collaboration with the Universities of Bristol, Nottingham and Sheffield (https://www.birmingham.ac.uk/research/cardiovascular-sciences/research/platelet-group/platelet-gapp/index.aspx) (Watson et al., 2013b). The study aimed to investigate patients with suspected inherited bleeding disorders of unknown cause. Over 1000 patients with suspected IBDs have been recruited to the UK-GAPP study from over 25 collaborating

Haemophilia Comprehensive Care Centres across the UK as shown in (Figure 1.8). The study used a combination of phenotyping and genotyping approaches such as LTA, flow cytometry and WES in orders to investigate the molecular mechanism behind IBDs and subsequently identifying the candidate variants underlying bleeding disorders (Jones et al., 2012, Watson et al., 2013b). Its ultimate purpose is to identify novel gene defects that could disrupt the normal platelet physiology which eventually increase our understanding of disease pathogenesis. The UK-GAPP study work flow can be seen in (Figure 1.9). The GAPP study work flow was employed since its establishment to investigate hundreds of patients (Johnson et al., 2018, Almazni et al., 2020). Several causative bleeding variants have been identified through the UK-GAPP methodology including the novel variants in *SLFN14* as well as those in *RUNX1* and *FLI1* which shows effective diagnosis for IBDs (Stockley et al., 2013, Fletcher et al., 2015). However, the study is not exempt from the challenges faced in diagnosing these IBDs as the GAPP study involved some patients without platelet function defects and some with normal platelet counts. These required further investigation through WES coupled with functional characterisation to identify the molecular and functional mechanisms behind their disease.

Figure 1.9. Map of the GAPP study network. Red pins show all Haemophilia Comprehensive Care Centres from which patients are referred to the main laboratory at the University of Birmingham shown by the blue star. The blue diamond represent the University of Sheffield and the University of Bristol.

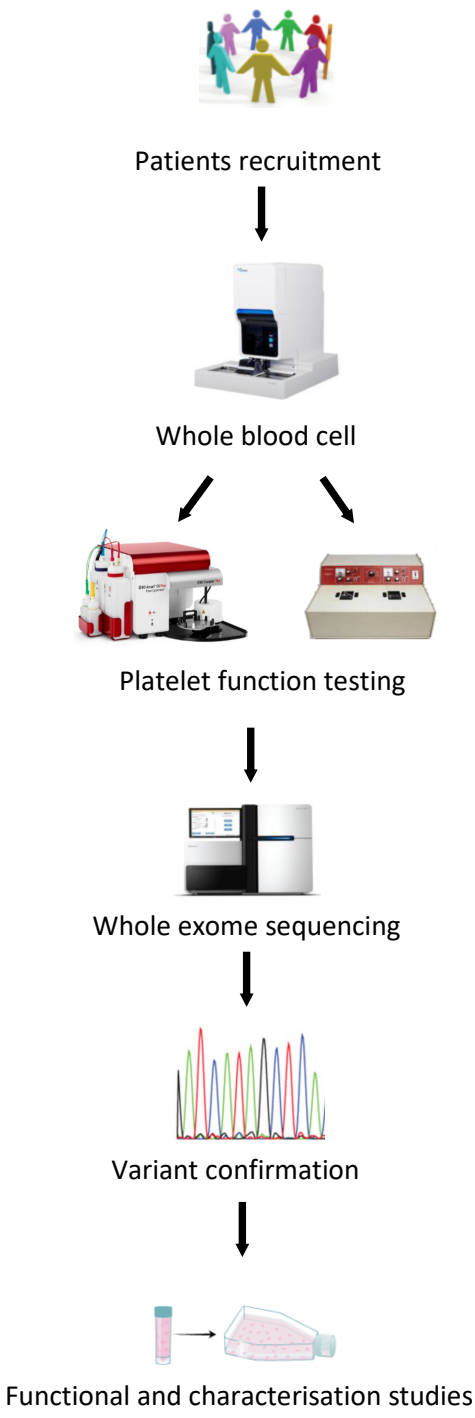


Figure 1.10. The GAPP study workflow. Patients suitable for the study are recruited. Blood cell counts using the Sysmex**.** Depending on the platelet count lumi-aggregometry or flow cytometry is carried out. If diagnosis is still not possible, DNA is selected to go forward for WES. Genetic variants are then confirmed by Sanger sequencing and functional work may need to be carried out.

1.13 Thesis hypothesis and aims

I hypothesise that patients with a variable bleeding tendency and undiagnosed previously with inherited bleeding are caused by genetic variants in both known and novel gene variants. We will investigate this hypothesis using a phenotyping and genotyping approach combined with WES and the Sapientia bioinformatics tool in order to identify the underlying causative variants in a cohort of patients with unknown cause. This will then be followed by functional analysis to investigate the variant's influence on protein function.

The overarching aim in this project is to elucidate the molecular genetic behind the IPDs and investigate novel causative mutations, which can be approached by:

1. Using whole exome sequencing data from individuals with a shared platelet phenotype coupled with Sapientia and bioinformatics tools to identify rare candidate genetic variants.
2. Filtering variants in Sapientia by prioritising them with multiple filtering factors.
3. Confirmation of plausible candidate variants by Sanger sequencing and segregation analysis.
4. Investigate the functional consequence of the candidate variant with appropriate functional assays.

Chapter.2
Materials and methods

2.1 Materials

All general lab consumable materials were purchased from StarLab, UK. Unless otherwise stated all materials and reagents were available from Sigma Aldrich (Sigma-Aldrich™, UK). All antibodies, software, databases and primers used in this thesis are listed in Table 2.1., 2.2 and 2.3 respectively.

Table 2.1. Lists of antibodies used in thesis's experiments. WB: Western blot, FS: Flow cytometry.

Antibodies	Host Species	Source	Working concentration/dilution
Primary antibodies			
Anti-CD36 (GTX112891)	Rabbit Polyclonal	GeneTex, USA	WB: 1/1000
Anti-CD36 (D8L9T)	Rabbit Monoclonal	Sell Signaling Technology®, USA, #14347	WB: 1/4000
Anti-CD36 (FA6-152)	Mouse Monoclonal	Hycult®Biotech, USA	FS: 10 µg/ml
Anti-GAPDH	Rabbit polyclonal	Abcam, UK, #ab9485	WB: 1:2500
Secondary antibodies			
Donkey anti-Rabbit IgG	Donkey, HRP-conjugated	GE Healthcare, UK, #NA934V	WB: 1/10,000
Anti-Mouse (whole molecule) $F(ab')_2$ fragment–FITC antibody	Sheep, FITC conjugate	Sigma-Aldrich™, UK, #F2883	FS: 1/128

Table 2.2. Software used throughout this project.

Software	Supplier
Chromas	Technelysium, Australia
ExonPrimer	Helmholtz Center Munich, Germany
Mutation Taster	Charité, Germany
Polymorphism Phenotyping v2 (PolyPhen)	BWH- Harvard Medical School, USA
Sorting Intolerant from Tolerant (SIFT)	J. Craig Venter Institute, USA
Protein Variation effect analyser	J. Craig Venter Institute, USA
Primer3	Whitehead Institute for Biomedical Research, USA
Congenica diagnostic decision support platform	Congenica®, UK
NEBaseChanger®	New England Biolabs, UK
GraphPad Prism	GraphPad Software, USA
SnapGene software	Insightful Science, USA

Table 2.3. Databases used throughout this project.

Database	URL
1000 Genomes Project	http://www.internationalgenome.org/1000-genomes-browsers/
BDGP	http://www.fruitfly.org/
Blueprint Progenitors	https://blueprint.haem.cam.ac.uk/admin/query_genes
ClinVar	https://www.ncbi.nlm.nih.gov/clinvar/
Ensembl	https://www.ensembl.org/index.html
EVS	http://evs.gs.washington.edu/EVS/
ExAC	http://exac.broadinstitute.org/
gnomAD	https://gnomad.broadinstitute.org/
GAPP Database	https://collaborate.bham.ac.uk/mds/gapp/SitePages/Home.aspx
HGMD	http://www.hgmd.cf.ac.uk/ac/index.php
UCSC Genome Browser	https://genome.ucsc.edu/
ClinVar Miner	https://clinvarminer.genetics.utah.edu/
dbSNP	https://www.ncbi.nlm.nih.gov/projects/SNP/snp_summary.cgi
DECIPHER	https://www.deciphergenomics.org/

2.2 Study approval

The UK-GAPP study was approved by the National Research Ethics Service Committee of West Midlands–Edgbaston (REC reference: 06/MRE07/36) and participants gave written informed consent in accordance with the Declaration of Helsinki. This study was registered at ISRCTN (www.isrctn.com/) as #ISRCTN 77951167. The GAPP study is included in the National Institute of Health Research Non-Malignant Haematology study portfolio (ID9858).

2.3 Patients recruitment to the UK-GAPP Study

Over 1000 patients have been successfully recruited to the UK-GAPP Study since the study has established. Patients were recruited with a suspected inherited bleeding disorder of unknown cause, where known (obvious) causes of inherited bleeding had been eliminated. All patients were required to meet certain inclusion criteria however, any who met one or more exclusion points were excluded from the study. Patients who fulfilled all criteria were consented in writing by signing the consent form as shown in (Appendix, Figure A).

Inclusion criteria:

a. Aged 0-85 years
b. Bleeding episodes in line with a platelet function disorder
c. Normal coagulation factor screen results
d. Absence of evidence suggesting acquired platelet dysfunction

Exclusion criteria:

a. Recently undergone major surgery
b. Suffer from chronic renal failure requiring dialysis
c. Severely anaemic (Hgb <8 g/dl)

d. Recently received any blood products
e. Receiving any known platelet function affecting drugs

Before recruitment to the study, all patients underwent routine clinical screening tests which consisted of full blood cell counts, peripheral blood smear, prothrombin time and activated partial thromboplastic time (APPT). The subsequent findings from these tests can exclude the most common causes of bleeding. Platelets from peripheral blood smears can be visualised under a light microscope. Large platelet size could be indicative Grey Platelet Syndrome or MYH9-related disorder, the presence of leukocyte inclusions which can be followed up with genetic tests, if available and then excluded from the study. In addition, known platelet conditions such as idiopathic thrombocytopenic purpura (ITP) and other non-platelet disorders, including von Willebrand disease and inherited coagulation factor deficiencies were excluded from the study.

2.4 Healthy volunteers

Healthy donor volunteers with ages over 18 years and deemed to be physically well at the time of donation were recruited to the study. All the healthy donors did not take any known platelet inhibiting drugs such as aspirin or other non-steroidal anti-inflammatory drugs (NSAIDs) in the last two weeks before donation. Donors were unable to donate blood to the NHS blood and transplant for up to two weeks before they took part in this study.

2.5 Sample collection

Peripheral blood samples were collected from recruited patients and familial healthy donor volunteers, if available, at the same time in the referring centre. If familial healthy donors were not available, a control blood sample was taken from any healthy

volunteer at the referring centre. It was important that the blood samples from patient and healthy control were collected at the same time to ensure samples experienced same transportation conditions. The blood samples were transported from the referring centre to the location of analysis in Birmingham via courier transport. A previous study conducted by a member of the UK-GAPP study found that the blood samples collected locally or transported via courier were highly comparable.The study showed that PRP samples incubated for up to 6 hours at room temperature maintained the platelets responses to all agonists (Dawood et al., 2007). This meant, samples could be collected nationwide and the recruitment of patients was not limited to Birmingham.

For adult patients, approximately 55 ml blood was collected in 3.2% (w/v) trisodium citrate via venepuncture. Approximately 20 ml blood, depending on the age and weight of patient, was collected from paediatric patients under 18 years also into 3.2% (w/v) trisodium citrate by venepuncture. In some cases, blood samples were collected into BD Vacutainer® 2.7 ml citrate blood collection tubes (BD, UK, #363083), however the collection tube used did occasionally vary depending on the protocol used in each referring haematological centre. Control blood samples collected in house at the site of analysis were collected by venepuncture into 1:9 ratio of 4% (v/v) citrate concentrated solution (Sigma-Aldrich™, UK, #S5770) to blood volume.

A further 2 - 4 ml of ethylenediaminetetraacetic acid (EDTA) anticoagulated blood was taken from all patients and healthy donor controls to perform whole blood cell counts. Blood was collected into BD Vacutainer® plastic whole blood tubes with spray-coated KDEDTA (BD, UK, #367835) by venepuncture with a concentration of 1.8mg of EDTA per millilitre of blood.

2.6 Platelet counts in whole EDTA-anticoagulant blood

Whole blood cell counts were assessed in EDTA-anticoagulant blood using the Sysmex XN-1000™ haematology analyser (Sysmex, UK) post January 2014. Previous to this whole blood cell counts were determined at the referring haematological centres. Any repeated patient samples after January 2014 were measured using Sysmex XN-1000™. Blood cell counts on Sysmex XN-1000™ were performed in manual operating mode for both sealed and open tubes. The measurement of whole blood cell counts by The Sysmex XN-1000™ analyser is considered highly accurate due to specific fluorescent labelling using a flow cytometry based analysis strategy. The machine consists of five different channels that read various blood components such as platelet, white blood cells (WBC), red blood cells (RBC) and haemoglobin. The most important channels for this project were the white cell nucleated (WNR) channel that measure the components of (WBC), basophils and nucleated red blood cell as well as the PLT channel that exclusively measured different parameters of platelets. The PLT channel used three different methods to measure the platelet counts including: platelet fluorescence (PLT-F), platelet impedance (PLT-I) and platelet optical (PLT-O). PLT-F channel used fluorescence RNA staining dye called oxazine which is considered the most accurate reading channel due to its ability to distinguish between platelets and apoptotic white cells which was not previously possible with other analysers. The measurement of platelet production (immature platelet fraction - IPF) and platelet size (mean platelet volume – MPV) was detected by using the PLT-F channel. The overall platelet count was obtained after comparison of platelet count between all the three PLT channels in order to avoid any significant differences in platelet counts. A daily quality control was performed using control blood (XN-CHECK [Sysmex, UK]) to ensure accurate machine performance throughout the

study. The normal range record was obtained by analysing blood samples from 40 healthy controls in collaboration with Sysmex. All blood samples from patients and controls recruited to the UK-GAPP study were analysed in the same way. Figure 2.1 demonstrates an example of the whole blood cell count report from the Sysmex of a single GAPP recruited patient.

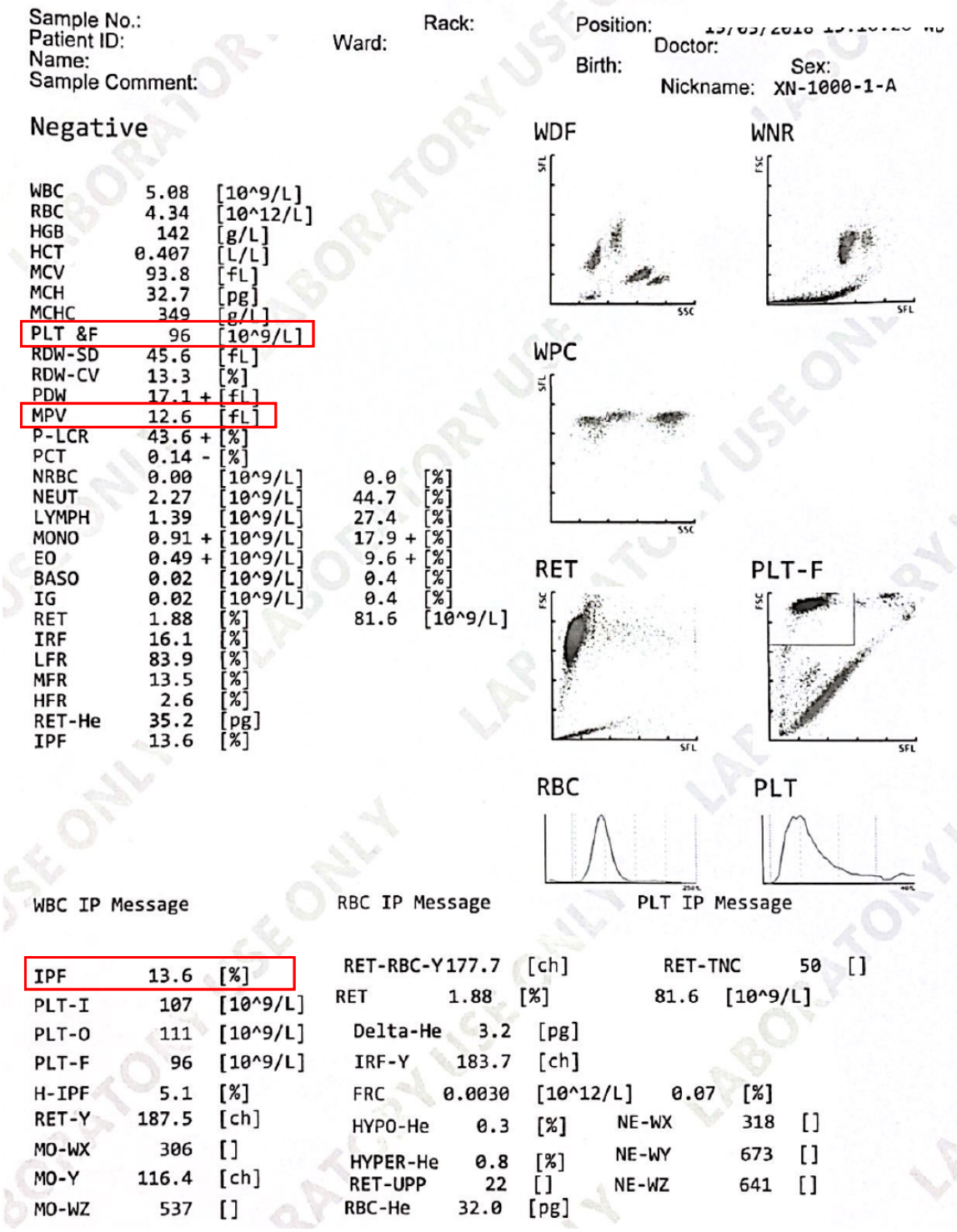


Figure 2.1. Full blood cell counts using the Sysmex XN-1000™ Haematology analyser. The report shows all blood cell counts with their units. The parameters highlighted in the red box are the most important to check for all patients recruited to the GAPP study. PLT: platelet count, MPV: mean platelet volume, IPF: immature platelet fraction.

2.7 DNA extraction

2.7.1 Cell lysis and protein precipitation

DNA was extracted from whole blood or buffy coats by using the Qiagen Gentra Puregene Blood DNA extraction Kit (Qiagen, UK, #158389) according to the standard protocol provided. 1:3 (v/v) ratio was used by adding 3ml of blood to 9ml of RBC lysis solution and mixed by inversion 10 times. This was then incubated for 5 minutes at room temperature, inverting once during incubation. This was then centrifuged for 5 minutes at 2000g to pellet the WBCs using Heraeus™Megafuge™16R centrifuge (Thermo Fisher scientific™, USA). The supernatant was discarded and the tube vortexed to resuspend the pellet in residual liquid. A 3ml cell lysis solution was added and vortexed vigorously for 10 seconds and then incubated at 37^{o}C until the solution became homogeneous. 15 µl RNase A (Sigma-Aldrich™, UK, #SLBN4278V) was then added and mixed 25 times, then the tube was incubated for 15 minutes at 37^{o}C, and then put on ice for 3 minutes to cool the sample. 3ml protein precipitation solution was added and vortexed vigorously for 20 seconds at high speed; the tube was then centrifuged for 5 minutes at 2000g to pellet the precipitated protein.

2.7.2 DNA precipitation and hydration

The supernatant containing the DNA was transferred into a clean, labelled 15 ml Falcon™ tube containing 3ml 100 % (v/v) isopropanol (Fisher Scientific, USA, #389710025) and inverted gently 50 times until the DNA threads were seen. The tube was centrifuged at 2000g for 3 minutes, the supernatant was discarded and the tube drained on a clean piece of paper towel. 3ml 70% (v/v) ethanol (EtOH) (VWR®, France, #20821.330) was added and inverted several times to wash the DNA pellet and then centrifuged for 1 minute at 2000g. The supernatant was discarded and the tube

drained on clean paper for 5-10 minutes. 100-300µl DNA hydration solution was added based on the size of DNA pellet and vortexed at medium speed for 10 seconds. DNA samples were incubated at 55^{o}C overnight to dissolve the DNA pellet. DNA was stored at -80°C until needed.

2.7.3 DNA quantification and analysis

The DNA concentration was measured using either a NanoPhotometer® P360 (IMPLEN, Germany) or the Qubit®3.0 Fluorometer (Life technologies™, USA). The NanoPhotometer® P-Class Submicrolitre Cell was placed into the cell holder. The well was calibrated and blanked with 2µl fresh water. The well was cleaned by tissue paper before applying 1-2µl of DNA sample to be measured. This step was repeated three times and the average of DNA concentration was recorded. The quality of sample was then determined by calculating the 260/280 ratio.

The Qubit® 3.0 Fluorometer (Life Technologies™, USA) was used to measure the DNA concentration for WES following the standard protocol and using the Qubit™ reagents. The Qubit dsDNA Broad Range (BR) and the Qubit dsDNA High Sensitivity (HS) Assay kits (Thermo Fisher Scientific, UK, #Q32850,#Q32851) were used to make up the Qubit® working solution. The Qubit® working solution was prepared by diluting the Qubit® reagent 1:200 in Qubit® buffer. Two standards were used for calibrating the Qubit® Fluorometer. One assay for each individual user sample was set up. For standard assays, 10 µl of each standard were added to 190 µl of working solution and for user samples, 3 µl of DNA samples were added to 197 µl of working solution to give a total volume of 200 µl in 500 µl thin-walled polypropylene Qubit® Assay Tubes (Invitrogen, UK, #Q32856). Samples were then incubated for 2 minutes at room temperature before been run and recorded.

2.8 Whole Exome Sequencing Analysis

Whole exome sequencing (WES) is one of the most common next generation sequencing methods used in our research field which can determine the causative variants of many genetic diseases. The coding regions (exons) of the genome makes up approximately 1.5% of the human genome. As it has been estimated that 85% of disease-causing variants are located within the coding regions of the genome, employing exome sequencing will be a very effective technique to identify these variants (Kellis et al., 2014, Botstein and Risch, 2003, Majewski et al., 2011). As we are assuming that the majority of patients recruited to the UK-GAPP study will have disease-causing variants in their exome's, WES was used in the UK-GAPP study. WES of patients in this project was outsourced by a collaborating group directed by Professor Michael Simpson in King's College London until 2016. Following that, the WES was carried out in Birmingham by a collaboration with Dr Yvonne Wallis from the West Midlands Regional Genetics lab, Birmingham Women's Hospital. Whole exome sequencing was performed on the genomic DNA of patients in this study as previously described (Johnson et al., 2016b). The WES required a minimum of 1µg DNA from each recruited individual. The DNA samples were transported to the sequencing centre in sealed 1.5 ml Eppendof tubes at room temperature. Briefly following enrichment of coding regions and intron/exon boundaries with the SureSelect human AllExon 50Mb kit (Agilent Technologies, USA), captured libraries were sequenced on the Illumina HiSeq® 2500 sequencing system (Illumina®, USA) with 100bp paired-end reads. Sequencing reads were then aligned to the human reference genome (hg19) by applying the Needleman-Wunsch algorithm using NovoAlign (Novocraft Technologies, Malaysia). SAMtools software (Wellcome Trust Sanger Institute, Cambridge) was used to manipulate alignments remove duplicated reads, variant

calling for SNP and small indels, amend the alignment information and view the alignments. All duplicated reads and reads that mapped to multiple read regions were excluded. All read coverages less than 4 reads were excluded from post bioinformatic analysis.

The remaining variants were then filtered for novelty by comparison to known variants present in multiple databases including: dbSNP139, the 1000 Genomes Project, Exon Variant Server (EVS) and the variants from the GAPP database of over 800 whole exome sequences. Post alignment and variant called sequencing bioinformatics analysis was employed in chapter 3.

2.8.1 Prediction of pathogenicity of genetic variants

The potential pathogenicity effect of each of the variants was assessed by utilising five different *in silico* pathogenicity prediction tools which could estimate the effect of candidate variants on the functionality of the protein. The pathogenicity prediction tools employed were as follows:

1- MutationTaster2 is an online free resource that uses a naive Bayes classifier to predict the effect of a DNA sequence change (variant) on the gene product (protein) and its disease-causing potential. Variants were then predicted to be either disease causing or polymorphism. Information on known disease-causing variants and common polymorphisms from various databases is integrated into the analysis tool and is listed in the prediction output (Schwarz et al., 2014).

2- PolyPhen-2 (Polymorphism Phenotyping version 2) is a comparative online tool that can predict the potential effect of amino-acid change on the structure and function of protein. This prediction is based on the sequence, phylogenetic and

structural information characterising the substitution change (Adzhubei et al., 2013). The pathogenicity scores were predicted independently by using two different datasets. For this study, HumVar database was used which includes all human disease causing mutations from UniProtKB as positive control along with common human SNPs that are not associated in diseases .

3- Sorting Intolerant From Tolerant (SIFT) predicts tolerated from deleterious amino acid residue sequences based on sequence homology and the physical properties in multiple aligned sequences. SIFT assumes that any changes in all conserved regions within protein sequences throughout evolution will affect the function of proteins (Kumar et al., 2009). This allows us to predict whether amino acid substitutions will be tolerated or deleterious to the protein function based on the sequence homology. Positions with a normalised probability of less than 0.05 are predicted to be deleterious, those greater than or equal to 0.05 are tolerated.

4- Protein Variation effect Analyzer (PROVEAN) is an online software that provides prediction about the impact of non-synonymous amino acid substitutions or indels on the biological function of protein by using delt alignment-based score. The score is calculated based on the difference between the sequence of variant and the reference sequence homology. The threshold is set at -2.5 for binary classification. Variant scoring below -2.5 are deemed to be deleterious and those above -2.5 are deemed to be neutral or have no effect (Choi et al., 2012).

5- Splice site prediction software was performed using the Berkeley Drosophila Genome Project that can predict the effect of variants on the donor and acceptor splice sites (Consortium, 1999).

In addition, expression level of mRNA for candidate genes in different haematopoietic cells progenitors was examined using the blueprint consortium progenitor software provided by the University of Cambridge as well as RNA-seq data from the study of human and mouse platelet transcriptomes (Rowley et al., 2011).

PhastCons and PhyloP scores provided from MutationTaster can be used to determine the conservation at the site of variant (Pollard et al., 2010). PhastCons values vary between 0 and 1 to represent the likelihood that a nucleotide belongs to a conserved region based on multiple alignment sequence of over 46 species. The closer the value to 1, the more probable the variant to be conserved. PhyloP values ranged from -14 to 6 and positive scores are allocated to sites that are predicted to be conserved, while negative scores are assigned to locations that are predicted to be fast evolved.

2.8.2 Variant classification

The American College of Medical Genetics and Genomics Guidelines (ACMG) and the Association for Molecular Pathology (AMP) have established scientific standard guidelines for geneticists to describe and interpret the sequence variants (Richards et al., 2015). They use standard terminology to describe and classify the potential pathogenicity of variants identified in genes causing disease. These include ("pathogenic", "likely pathogenic" "uncertain significance", "likely benign" and "benign"). The candidate variants from WES analysis of the UK-GAPP study patients in this thesis were classified based on all these categories when predicting the pathogenicity of each one. This report recommended using this terminology for the

interpretation of the wide variety of genetic tests used in the clinical laboratory such as genomes, exomes, genotyping and single gene or panel sequenced genes (Kearney et al., 2011).

2.9 Preparation of washed platelets

Washed platelets were prepared for platelet protein extraction in this thesis. 22.5 ml of water was added to stock 10x Tyrode's concentrate to make up total volume of 25ml. Tyrode's buffer contains (134 mM NaCl, 2.9 mM KCl, 10 mM HEPES, 12 mM NaHCO3, 0.34 mM NA2HPO4, 1 mM $MgCl_2$, [all reagents available from Sigma-Aldrich™, UK]). Tyrode's buffer was warmed in a 37°C water bath along with 10ml acid-citrate-dextrose (ACD) (39mM citric acid, 75 mM sodium citrate, 135 mM dextrose [all reagents available from Sigma-Aldrich™, UK]). 22.5mg of Glucose was added to make 25 ml modified Tyrode's buffer at pH 7.3 at 30 ºC. The blood samples in anticoagulant tubes , 4% (v/v) citrate concentrated solution (Sigma-Aldrich™, UK, #S5770) were then transferred to polypropylene tubes and centrifuged at 200g for 20 minutes at room temperature using the Heraeus™Megafuge™ 16R centrifuge (Thermo Fisher Scientific™, USA). Platelet rich plasma (PRP) was pipetted into a 50ml Falcon™ tube. 10µl prostaglandin I2 (PGI_2) (10µl of stock solution at 1 mg/ml [Cayman Chemical Company, USA, #18220]) per 10ml PRP was added and mixed gently. The PRP tube was then centrifuged at 1000g for 10 minutes. The supernatant was discarded and the pellet was resuspended with 2ml warmed modified Tyrode's buffer. 3ml of ACD, 10µl of PGI2 and other 20ml of modified Tyrode's buffer were added and the tubes centrifuged at 1000g for 10 minutes to wash and pellet the platelets. This step was repeated once more. The supernatant was then discarded and the pellet was resuspended in 2ml modified Tyrode's buffer. 5µl of washed platelets was added to 10ml COULTER® ISOTON® II

Diluent (VWR, USA, #8448011) and platelets counted three times using the Z™2 Series COULTER COUNTER® (Beckman Coulter®, USA). The average of the three platelet readings was taken and platelets were diluted to $3x10^8$ /ml before starting the protein extraction. The platelets were diluted with modified Tyrode's buffer to obtain a final concentration of $3x10^8$ in order to extract platelet protein.

2.9.1 Sample storage - platelet protein and buffy coat

To prepare platelet protein, PRP was adjusted to a volume containing $3x10^8$ platelets. This was transferred into new 1.5ml Eppendorf tube. 1µl of PGI_2 (1 mg/ml [Cayman Chemical Company, USA, #18220]) was added to the PRP sample and mixed gently. The PRP sample was centrifuged at 1000g for 10 minutes using the Heraeus™ Pico™ 21 centrifuge (Thermo Fisher Scientific™, USA). The supernatant was carefully aspirated to avoid disrupting the platelet pellets. The platelet pellets were resuspended in 1ml phosphate buffer saline (PBS, [Sigma-Aldrich™, UK, #P4417]). 1µl of PGI_2 was then added to the resuspended pellets , inverted gently to mix, and centrifuged at 2000g for 2 minutes. The supernatant was then aspirated and the protein pellets were resuspended with 600µl 1x sample buffer (60 mM Tris-HCl pH 6.8 [Sigma-Aldrich™, UK, #T6455] , 2% (w/v) sodium dodecyl sulphate [SDS, Sigma-Aldrich™, UK, #L3771], 10% (v/v) glycerol [Sigma-Aldrich™, UK, #G5516], 0.002% (w/v) Bromophenol Blue, 5% (v/v) β-mercaptoethanol [Sigma-Aldrich™, UK, #M6250]). The platelet protein was stored at -80 ºC until needed.

After PRP was removed from the blood sample, the blood sample was centrifuged again at 1000g for 10 minutes. Avisible white layer known as a buffy coat is formed at the top of the red blood cell layer. The buffy coat was then removed by pipetting and transferred into 1.5ml Eppendorf tube and labelled with patient's name. The buffy coat

contains white blood cells which can be used later in order to extract DNA. The buffy coat is stored at -80 ºC until needed.

2.10 Polymerase Chain Reaction (PCR)

Polymerase Chain reaction (PCR) was applied in order to amplify the genomic DNA template for subsequent Sanger sequencing and confirmation of NGS data and genetic variants detected. All PCR reactions were performed based on a standard laboratory protocol. Each single reaction contained a total volume of 25µl reaction mix including: 12.5 µl Sigma-Aldrich® REDTaq® ReadyMix™ DNA polymerase, 0.5µl forward primer, 0.5µl reverse primer both at concentration 10µM, 6.5µl dH_2O and 5µl DNA sample at a concentration 20 ng/µl. MyTaq™ Red Mix DNA polymerase was used as an alternative DNA polymerase mix in some PCR reactions that were unsuccessful with Sigma RedTaq. A PCR reaction with water was used in lieu of template DNA as negative control. All the primer pairs were designed using the ExonPrimer database from the University of California Santa Cruz (UCSC) genome browser (http://genome.ucsc.edu/). All primers were purchased from Sigma-Aldrich®. The primer pair sequences are available in appendix Table A. The PCR reaction was carried out using the DNA Engine Tetrad®2, Peltier Thermal Cycler, (BIO RAD) under the universal cycling conditions of denaturation, annealing and elongation as follows:

- 94 ºC for 3 minutes
  - 94 ºC for 1 minute
  - 60 ºC for 1 minute
  - 72 ºC for 1 minute

  } X 30 cycles
- 72 ºC for 5 minutes

The primer pairs' annealing temperatures for some candidate genes were optimised if annealing was unsuccessful at 60ºC.

2.10.1 Agarose gel electrophoresis

The PCR products were run on agarose gel electrophoresis for visualisation. 1% (w/v) of Ethidium bromide stain was added into 1.5% (w/v) Ultrapure™ Agarose (Invitrogen™, UK, #16500-500) and Tris acetate-EDTA buffer (TAE) containing (40 mM Tris, 20mM acetic acid and 1 mM EDTA). Later in the project, SYBR® Safe DNA gel stain (10,000X concentrated in DMSO) (Invitrogen, USA, #1771543) was used instead of Ethidium bromide. All gels were loaded with 5µl of 1kb hyper DNA ladder (New England Biolabs, UK, #N3232) in order to measure the size of PCR products alongside 10µl of PCR product. The gels were run at 120 V for 20-60 minutes using suitable power pack. The gels were then viewed under an ultraviolet transilluminator (Syngene, Gene Genius Bio Imaging System, UK) to visualize the PCR products and GeneSnap (Synoptics,UK) software was used to capture an image of the gel.

2.11 Sanger sequencing

PCR products were purified and cleaned before Sanger sequencing was carried out. 2.4µl of PCR product was transferred into a new 96-plate and loaded twice for both forward and reverse sequences. 2.4 µl of microCLEAN solution (Microzone, UK, #2MCL) was added into each single well using the Multipette® Stream multi-pipette (Eppendorf, Germany). The plate was covered with a plastic lid and centrifuged at 1900g for 40 minutes using the Hettich Universal 320R plate centrifuge (Hettich®, Germany). The plate was then placed upside-down on a clean paper towel and spun at 30g for 30 seconds to remove the supernatant and any residual liquid.

A PCR-based sequencing reaction was conducted to amplify the targeted DNA sequence by using a single primer. The reaction mixture consisted of a total volume of 10µl containing: 0.5µl Big Dye® Ready Reaction Mix, 2µl Big Dye® Terminator 5x Sequencing Buffer (BigDye™ Terminator v3.1 Cycle Sequencing Kit, Applied Biosystems™, USA, #4337455), 2µl (2pmol/µl) forward or reverse primer specific for the gene of interest and 5.5µl dH_2O. The plate was then vortexed briefly to resuspend reaction contents. The sequencing reaction was conducted under the universal cycling conditions as follows:

- 96 ºC for 30 seconds
    - 96 ºC for 30 seconds
    - 50 ºC for 15 seconds
    - 60 ºC for 4 minutes

  X 30 cycles
- 4ºC for ∞

The sequencing reaction was cleaned and purified by using two ethanol washes. Firstly, 2µl of 0.125M EDTA (Sigma-Aldrich™, UK, #E7889) coupled with 30µl of 100% (v/v) ethanol absolute (VWR, France, #20821.330) were added to each single reaction and centrifuged at 500g for 20 minutes. This wash was removed by placing the plate upside-down and centrifuged at 30g for 30 seconds. The second wash was carried out by adding 90µl of freshly made 70% (v/v) ethanol diluted in dH_2O to each individual well, replaced with the cover lid and centrifuged at 500g for 20 minutes. The plate was then centrifuged at 30g for 30 seconds to remove the wash. The plate was left for 5-10 minutes to let the DNA pellets/sequencing reactions to air dry.

10µl of Hi-Di™ Formamide (Thermo Fisher Scientific™, USA, #4311320) was added to each single reaction and mixed by pipetting to resuspend the DNA pellet. 10µl of water was added to each empty well and then covered with a new adhesive plastic cover. The plate was placed in the PCR machine to denature the DNA at 94ºC for 2 minutes. The plate was then placed immediately on ice to snap chill and keep the DNA denatured and prevent re-annealing.

The plate was then sequenced and DNA fragments were separated by capillary electrophoresis using the ABI 3730XL Automated Sequencer (Applied Biosystems™, USA) operated by the functional genomics lab at the School of Biosciences , University of Birmingham. The sequence data was then analyzed and viewed using Chromas Version 2.4 sequence chromatogram (Technelysium, Australia). Occasionally, unsuccessful Sanger sequencing presented with lack of sequence data, poor or weak trace signal, overlapping peaks which required repeating to obtain clean and readable sequence traces.

2.12 NFAT-luciferase reporter assay

Jurkat T cells were grown in RPMI 1640 medium (SIGMA -ALDRICH®, UK, #RNBJ1179) supplemented with (10 % (v/v) FBS, 100 units/ml penicillin, 100 µg/ml streptomycin and 2 mM L-glutamine). Cells were split to $3.5x10^5$ cells/ml for 20 ml per transfection condition. Cells were incubated overnight at 37ºC in 5% CO_2. Cells were counted under microscope by using Trypan Blue solution (SIGMA -ALDRICH®, UK, #RNBF9999) in order to obtain $1x10^7$ cells per transfection. Cells were then centrifuged at 250gfor 5 minutes and the supernatant poured off. Cells were washed with unsupplemented RPMI media and centrifuged at 250g for 5 minutes and then subsequently removed the supernatant. Cells were then resuspended in 0.4ml per transfection of unsupplemented RPMI media per transfection, to avoid potential binding to stimuli. 20 µg NFAT-luciferase construct was added to each transfection and mixed by pipetting then transferred into sterilised electroporation cuvette (Cell Projects, UK). 12.5 µg of tested constructs were added to each cuvette and mixed by Pasteur pipette. Samples were incubated at room temperature for 10 minutes before transfected by electroporation. Samples were transfected by electroporation (BIO-RAD Gene Pulser, USA) at 250V and 960µF and then incubated at room temperature for 10 minutes. Samples were transferred into 6-well plates (Corning Incorporated COSTAR®, USA) containing 8ml of complete RPMI media in each well and incubated overnight at 37ºC in 5% CO_2. Next day, cells were counted and harvested in 15ml tube at 250g for 5 minutes. Cells were resuspended at $2x10^6$ cells/ml in complete RPMI media. 250µl of harvested cells were transferred into two Eppendorf tubes for subsequent FACS and western blot and stored on ice. The NFAT-luciferase assay was carried out in triplicate for stimulated and unstimulated conditions. 50µl of transfected cells were transferred into all wells followed by 50µl of complete RPMI

media in unstimulated condition wells and 50µl of PMA/Ionomycin in stimulated condition wells. 100ng/ml PMA and 2µM ionomycin were diluted in complete RPMI media. All the wells were mixed and incubated at 37 ºC in 5 % CO_2 for 6 hours.

The luciferase assay was carried out by harvesting the cells in 11µl of luciferase harvest buffer containing (1 M potassium phosphate buffer ph 7.8, 12.5% (v/v) Triton X-100 and 1M DTT). Cells were mixed by stirring until the media turned a slight yellow colour . Cells were incubated at room temperature for 5 minutes. 100µl of luciferase assay buffer containing (1 M potassium phosphate buffer ph 7.8, 0.1M $MgCl_2$, 0.1M ATP and 5ml dH_2O) was transferred into an opaque 96-well plate (CELLSTAR®, Greiner Bio-One, Germany) and 100µl of samples were added and mixed by pipetting. The plate was inserted into luminometer machine after injection of 50µl of 1mM luciferin (Berthold Technologies, Germany). The luciferase activity was measured using a luminometer software – MICROWIN 2000 and the measurement data was then exported and analysed using the Excel program.

2.12.1 Cell counting using haemocytometer

To count cultured cells, the cells were poured into a sterilised 50ml Falcon™ tube. 10µl of cultured cells was added to clean Eppendorf tube and mixed with 10µl Trypan Blue solution (SIGMA -ALDRICH®, UK, #RNBF9999) and mixed. The mixture was loaded into Neubauer Counting Chamber Neubauer (Hecht Glaswarenfabrik GmbH&Co KG, Assistent®, Germany) and inserted under the microscope for counting.

2.13 Preparation of lysates from cell culture

Cell lysates were prepared from cell culture for Western blot analysis and flow cytometry. The 250µl of harvested cells prepared in Section 2.12 underwent the subsequent steps.

For Western blot analysis, the Eppendorf tubes containing harvested cells were centrifuged at 2000g for 3 minutes at 4ºC and the supernatants were aspirated. 50µl of cold lysis buffer containing (300mM NaCl, 20mM Tris, 2mM EDTA, 2mM EGTA and 1% (v/v) Nonidet p-40 [NP-40, Sigma-Aldrich™, UK, #I8896]) with protease inhibitors (250µl sodium orthovanadate [Sigma-Aldrich™, UK, #450243], 100µl leupeptin [Enzo Life Sciences, USA, #260-009-M100], 10µl aprotinin [Sigma-Aldrich™, UK, #A1153], 4µl pepstatin [Sigma-Aldrich™, UK, #P5318] and 100µl AEBSF [Calbiochem, UK, #101500]) was added for each sample. The samples were mixed and incubated on ice for 30 minutes at room temperature. The cell suspensions were then centrifuged at 16000g for 10 minutes at 4ºC. The supernatants were removed and placed into new tubes containing 40µl 2x sample reducing buffer and stored at -20ºC until use.

For flow cytometry, the Eppendorf tubes containing harvested cells were centrifuged at 250g1 for 5 minutes at 4ºC to pellet the cells and the supernatants were aspirated. 50µl of primary antibody diluted in FACS buffer (0.2% (w/v) BSA [First Link UK Ltd, UK, #41-00-450] and 0.02% (w/v) sodium azide [Sigma-Aldrich®, UK, #S2002]) in PBS was added to each sample and mixed. Samples were incubated on ice for 30 minutes. 1ml of FACS buffer was added to each sample and vortexed. Samples were centrifuged at 250g1 for 5 minutes at 4ºC to pellet the cells and the supernatants were aspirated. 50µl of secondary antibody diluted in FACS buffer (1:100) was added to each sample and mixed. Samples were incubated on ice for 30 minutes in a dark box. 1ml of FACS buffer was added to wash samples and then vortexed. Samples were centrifuged at 250g1 for 5 minutes at 4ºC and the supernatants were aspirated. 500µl of FACS buffer was added to each tested sample before analysing on an Accuri™ C6 flow cytometer (BD,UK). Samples were ran using antibody colour channels in collecting 10,000 ungated events.

2.14 Western blot analysis

Western blotting was used to investigate the protein expression of candidate genes of interest in this thesis. Western blots were conducted on stored platelet protein and platelet lysates from cell culture.

Samples firstly were boiled at 105 ºC for 5 minutes in order to denature the protein. Samples were run on Precast polyacrylamide 10 and 12 well Bolt™ Bis-Tris Plus gel (Invitrogen,UK, #NW04120BOX and #NW04122BOX) using Bolt™ Mini Gel Tank (Invitrogen, UK). 20x Bolt™ MES SDS Running Buffer (Thermo Fisher Scientific, USA, #NP0002) was diluted to 1x in dH_2O and then added to the gel tank. 5µl of Coloured Prestained Protein Standard, Broad Range (10-250 KDa) (New England BioLabs, USA, #P7719) was loaded into the first well in order to verify the expected protein size. 20µl of denatured samples were loaded into the remaining wells of the gel and saple buffer alone in any spare wells to ensure equal gel running. The gel was run at 60V for 10 minutes followed by 120V for one to two hours depending on the size of protein samples until the standard ladder and samples reached the bottom of the gel. To transfer protein bands from polyacrylamide gel to low fluorescence polyvinylidene difluoride (LV PVDF) membrane, the Trans-Blot® Turbo™ Transfer System (Bio-Rad, UK) and Trans-Blot® Turbo™ RTA mini transfer kit (Bio-Rad, UK, #1704272) were used. PVDF membrane was activated in 100% (v/v) methanol (VWR, France, #20847.307) for two minutes prior transfer protein bands. Blot absorbent filter paper was pre-socked in 1x Trans-Bolt® Turbo™ Transfer buffer (Bio-Rad, UK, #10026938). Filter paper, PVDF membrane, polyacrylamide gel and final filter paper were stacked together and rolled gently to remove air bubbles. They were inserted into the Trans-Bolt® Turbo™ transfer system Cassette (Bio-Rad,UK) and run at 1.5V for 10 minutes. The membrane was blocked in 4% (w/v) bovin serum albumin (BSA) (First Link UK

Ltd, UK, #41-00-450) and 0.1% (w/v) sodium azide (Sigma-Aldrich®, UK, #S2002) in tris-buffered saline with Tween® (TBS-T, [200 nM Trizma base, 1.37 M NaCl, 0.1% (v/v) Tween-20]) for one hour at room temperature. Primary antibodies were diluted as recommended in 4% (w/v) BSA and 0.1% (w/v) sodium azide in TBS-T. PVDF membranes were incubated in primary antibody at 4ºC overnight on a tube roller. Next day, the membrane was washed with TBS-T 3x for 5 minutes.

Horseradish peroxidase-conjugated (HRP) secondary antibodies were diluted in TBS-T as recommended. Membranes were then incubated with secondary antibody for 1 hour at room temperature on a tube roller. After incubation, the membrane was washed 5x in TBS-T. An enhanced chemiluminescence system with the Pierce™ ECL Wester Blotting Substrate Kit (Fisher Scientific, USA, #32106) was used in order to stain proteins on the membrane for developing. The membrane was placed protein side down in a mixture of 1ml Detection Reagent 1 Peroxide Solution and 1ml Detection Reagent 2 Luminol Enhancer Solution for 5 minutes. Membranes were blotted dry and mounted into a film cassette wrapped in clear cling film wrap. Chemiluminescence on the membrane was visualised after exposure of membrane into Amersham Hyperfilm ECL (Fisher Scientific, UK, #GZ28906837) using the XOGRAPH Compact X4 Film Processor (Xograph, UK).

Bound antibodies were stripped from the membrane by incubating the membrane in stripping buffer (TBS-T containing 2% (w/v) SDS) and 1% (v/v) β-mercaptoethanol (Sigma-Aldrch™, UK, #M6250) for 20 minutes at 80ºC. Membranes were then stripped again in stripping buffer without β-mercaptoethanol for 20 minutes at 80 ºC. The membrane was blocked again in 4% (w/v) BSA and 0.1% (w/v) sodium azide in TBS-T for one hour at room temperature before adding primary antibody and repeating.

2.15 Site-directed mutagenesis to introduce *CD36* candidate variant

2.15.1 Site directed mutagenesis kit and mutagenesis primer design

Q5 Site-Directed Mutagenesis (SDM) Kit (NEB®, USA, #E0554S) was used to introduce the CD36 heterozygous stop gain variant (c.975T>G; p. Tyr325Ter) into the human CD36 cDNA wild-type that has been subcloned into mammalian expression pEF-BOS vector by using *BstXI* adapter or *XbaI* linker (Table 2.4) (Thorne et al., 2000). The mutagenic primer sequences were designed by using the NEBase changer tool (http://nebasechanger.neb.com). The entire CD36 cDNA sequence was inserted into the tool. For substitution change, the start (5') and end (3') location of nucleotide to be substituted (Thymine T) was selected followed by selection of desired nucleotide sequence change (Guanine G). The substitution sequence change was incorporated in the centre of the forward mutagenic primer. The 5' end of reverse primer was designed to anneal at the nucleotide adjacent to 5' end of the forward mutagenic primer on the reverse strand. For deletion, the start (5') and end (3') position of regions to be deleted was selected. Forward and reverse primers were designed to flank the deleted region. The 5' end of forward primer was designed to anneal at the nucleotide adjacent to 3' end of deleted region while the 5' end of reverse primer was designed to anneal at the nucleotide adjacent 5' end of the deleted region on the reverse strand.

All primers designed for the substitution and deletion mutagenesis are detailed in (Table 2.5). Figure 2.2 illustrated the primer design for substitution and deletion mutagenesis.

Table 2.4. The structure of pEF-BOS mammalian expression vector and constructed plasmids.

Plasmid vector	**Replication origin**	**Promotor**	**Restriction sites for cDNA insertion**	**Bacterial resistance**	**Size**	**Plasmid type**	**Insert**
Empty pEF-BOS	SV40	Human EF-1alpha	BstXI or XbaI	Ampicillin	5.8 kb	Mammalian Expression	No
pEF-BOS (WT CD36)	SV40	Human EF-1alpha	BstXI or XbaI	Ampicillin	5.8 kb	Mammalian Expression	WT CD36 cDNA
pEF-BOS (Deleted CD36)	SV40	Human EF-1alpha	BstXI or XbaI	Ampicillin	5.3 kb	Mammalian Expression	CD36 cDNA including (a.a1 - a.a324)
pEF-BOS (Substitution CD36)	SV40	Human EF-1alpha	BstXI or XbaI	Ampicillin	5.8 kb	Mammalian Expression	WT CD36 cDNA including c.975T>G change

Table 2.5. Site directed mutagenesis primers to introduce CD36 mutants.

Primer	Oligo
SDM_CD36 F_Deletion	5'- GTATGTACCAAAAAATATTGCTTC
SDM_CD36 R_Deletion	5'- CTATGATGTACAATTTTTTGAGATAATTTTTTC
SDM_CD36 F_Substitution	5'- GTACATCATAGGGTGTGCTAG
SDM_CD36 R_Substitution	5'- AATTTTTTGAGATAATTTTTTCTGTG

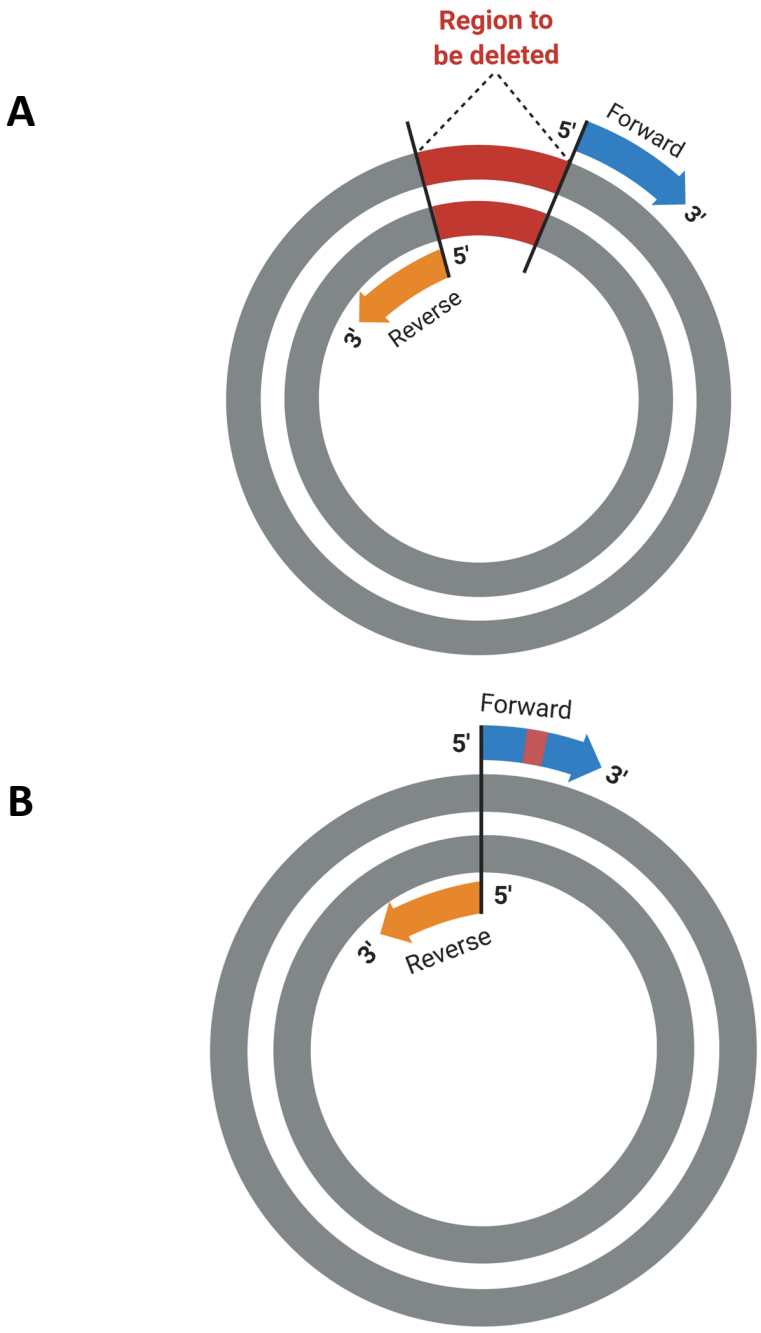


Figure 2.2. Primer design for the site-directed mutagenesis. NEBase Changer® (http://nebasechanger.neb.com) was used to design primers for substitution and deletion mutagenic change within WT CD36 cDNA. Figure A shows the primer design for deletion mutant. Forward (blue) and reverse (orange) primers were designed to anneal to nucleotides adjacent to the deleted region. Figure B illustrates the primer design for substitution change. The substitution change (red) was incorporated in the middle of forward primer (blue). The 5' end of reverse primer was designed to anneal at the nucleotide adjacent to 5' end of the forward mutagenic primer on the reverse strand.

2.15.2 SDM exponential amplification (PCR)

The PCR reaction of SDM was carried out by adding the following reagents in a thin-walled PCR tube to make a total of 25 µl (Table 2.6). All reagents were mixed and the tubes were inserted into the thermal cycling machine (DNA Engine Tetrad®2, Peltier Thermal Cycler, (BIO-RAD, UK) using the conditions in (Table 2.7).

Table 2.6. SDM PCR amplification reagents.

Reagents	Volume	Final concentration
Q5 Hot Start High-Fidelity 2X master Mix	12.5 µl	1X
10 µM forward primer	1.25 µl	0.5 µM
10 µM reverse primer	1.25 µl	0.5 µM
Template DNA (1-25ng/µl)	1 µl	1-25 ng
Nuclease-free water	9.0 µl	
Total	25 µl	

Table 2.7. Thermocycling conditions for SDM PCR amplification.

Step	**Temperature**	**Time**	
Initial denaturation	98 ºC	30 seconds	
Denaturation	98 ºC	10 seconds	25 Cycles
Primer annealing	60 ºC	30 seconds	
Extension	72 ºC	30 seconds	
Final Extension	72 ºC	2 minutes	
Hold	4 ºC	∞	

2.15.3 Kinase, Ligase & DpnI (KLD) treatment and transformation

Following the SDM-PCR reaction, a unique enzyme mix including kinase, ligase and DpnI was used in order to ligate and circulate the amplified linear plasmid products and remove the DNA template. This was carried out by assembling the following reagents in a clean PCR tube to make a 10µl final volume reaction: 1µl PCR product, 5µl of 2X KLD Reaction Buffer (1X final concentration), 1µl of 10X KLD Enzyme Mix (1X final concentration) and 3µl nuclease-free water. The reagents were mixed gently by pipetting and incubated at room temperature for 5 minutes.

The final step transformed the circulated plasmids into a high-efficiency NEB 5-alpha Competent *E.coli* cells. A tube of NEB 5-alpha Competent *E.coli* cells was thawed on ice following by adding 5µl of the KLD reaction mixture prepared from the previous step into the tube cells. The tube was mixed gently by flicking and then incubated on ice for 30 minutes. The cells were then heat shocked at 42 ºC for 30 second follwed by incubation on ice for 5 minutes. Afterwards, 950µl of pre-warmed Super Optimal broth with Catabolite repression (SOC) medium was added into the cells following with gently shaking at 37ºC for one hour. Finally, the cells were mixed and 40-100µl of the cell mixture was spread onto multiple Luria broth agar Miller (LB) selection plates (Sigma-Aldrich®, USA, #MKCG6791) with 1:1000 Ampicillin antibiotic. The plates were then incubated at 37ºC overnight.

2.15.4 Verification of SDM

To check if the *CD36* constructed mutants were successfully incorporated within *CD36* cDNA, 10 colonies were selected and individually grown in 5ml Luria Broth (LB) media (Sigma-Aldrich, USA, #BCCC0494) with 5µl Ampicillin. Cultured colonies were incubated under shaking conditions using(New Brunswick™ Innova® 43/43R Shaker,

Germany) at 37ºC overnight. Plasmid DNA was then extracted using the small-scale DNA extraction method Section 2.15.5 . The mutant constructs were then Sanger sequenced to confirm if the constructed mutants were successfully incorporated in the CD36 cDNA.

2.15.5 Small-scale plasmid DNA extraction and purification using Miniprep kit

Small-scale (up to 20µg) plasmid DNA was extracted from bacterial cell cultures prepared in Section 2.15.4 by using (QIAGEN® Plasmid Mini Kit, Germany , #12123) following the manufacturer's protocol. Bacterial cells cultured in 5ml LB broth were harvested by centrifugation at 6000 x *g* for 15 minutes at 4ºC using Centrifuge 5810R (Eppendorf®, UK) followed by completely decanting the medium. The bacterial pellet was resuspended in 300µl Resuspension Buffer (P1) including RNase A ensuring sufficient mixing. The cells were then lysed by adding 300µl Lysis Buffer (P2) followed by vigorously inverting the tubes 4-6 times and incubating for 5 minutes at room temperature. Lysis buffer solubilises the cell membrane of bacterial cells and subsequently releases DNA. Afterwards, 300µl of prechilled Neutralisation Buffer (P3) was added to the cells mixture and mixed thoroughly by vigorously inverting 4-6 times and then incubated for 5 minutes on ice. Buffer P3 precipitated the protein, cell debris and genomic DNA. The tubes were then centrifuged at 16,000 x *g* for 10 minutes at 4ºC. Meanwhile, QIAGEN-tips 20 were equilibrated by applying 1ml of Buffer QBT and columns emptied by gravity flow. When the high speed centrifugation finished, supernatants were poured into the equilibrated QIAGEN-tips 20 and allowed to enter through the resin by gravity flow. The QIAGEN-tips 20 were washed twice by applying 2ml Buffer QC allowing the QC washing buffer to pass through QIAGEN-tips 20 by

gravity flow. The QIAGEN tips were placed in clean 2ml polypropylene collection tubes followed by adding 800µl of Buffer QF in order to elute the DNA. The eluted DNA was precipitated by adding 560µl fresh absolute isopropanol and mixed by inversion. The DNA mixtures were centrifuged at 15,000 x *g* for 30 minutes at 4ºC followed by careful removal the supernatant. 1ml of freshly prepared 70% (v/v) ethanol was added to wash the DNA pellets followed by a further centrifugation step at 15,000 x *g* for 10 minutes at 4ºC. The supernatant was carefully decanted and DNA pellets were air-dried for 10 minutes. Finally, the DNA pellets were redissolved in suitable volume of TE buffer (pH 8.0). DNA samples were stored at -20 ºC until use.

2.15.6 Large-scale plasmid DNA extraction and purification using Maxiprep kit

The (QIAGEN Plasmid Maxi Kit, QIAGEN, Germany, #12162) was used in order to extract a high DNA yield up to 500µg from bacterial cells. After confirmation of mutant constructs by sequencing, a bacterial glycerol stock of verified colony (labelled previously) was scraped with a sterile cell scraper and cultured in 5ml LB broth with 5µl ampicillin and incubated under shaking conditions at 37ºC for 3-4 hours. The bacterial culture was then cultured in large volume of LB broth medium (150 ml LB + 150 µl ampicillin) and incubated on shaker at 37ºC overnight. The following day, the QIAGEN Plasmid Maxi Kit was used to extract the plasmid DNA from bacterial culture. The protocol of plasmid DNA extraction for QIAGEN Plasmid Maxi Kit is similar to what mentioned previously in Section 2.15.5. The differences between large-scale and small-scale plasmid DNA extraction and purification kits are the volume of solutions to be added, size of QIAGEN-tips (500) and incubation times. The following Table 2.8 illustrates these differences.

Table 2.8. The differences between QIAGEN Mini Kit and QIAGEN Maxi Kit.

Steps	QIAGEN Plasmid Mini Kit	QIAGEN Plasmid Maxi Kit
1st centrifugation	6000 x g 15 minutes at 4 ºC	
Buffer P1	300 µl	10 ml
Buffer P2	300 µl	10 ml
Buffer P3	300 µl, incubation 5 minutes on ice	10 ml, incubation 20 minutes on ice
2nd centrifuge	16,000 x *g* for 10 minutes at 4 ºC	20,000 x *g* for 30 minutes at 4 ºC
QIAGEN tip	20	500
Buffer QBT	1 ml	10 ml
Buffer QC	Two times, 2 ml	Two times, 30 ml
Buffer QF	800 µl	15 ml
Isopropanol	560 µl	10.5 ml
3rd centrifuge	15,000 x *g* for 30 minutes at 4	
70% (v/v) ethanol	1 ml	5 ml
4th centrifuge	15,000 x *g* for 10 minutes at 4 ºC	

2.15.7 Preparation of LB agar plates and broth medium

LB agar plates were prepared dissolving 37g of LB agar Miller (Sigma-Aldrich®, USA, #MKCG6791) powder in 1 litre distilled water and mixed until completely dissolved. The mixture was autoclaved and cooled before adding the appropriate ampicillin (1:1000) under flame. 200µl of ampicillin was added to 200ml of agar mixture and mixed. The agar mixture was incubated for 5 minutes at room temperature before pouring approximately 20ml of the agar mixture into sterilized 10cm diameter Perti dish (avoiding air bubbles). The plates were then incubated at room temperature to cool for 30 minutes before storing (inverted) at 4ºC until used.

The LB liquid broth medium was made by dissolving 10g of LB Broth powder (Sigma-Aldrich®, USA, #BCCC0494) in 500ml distilled water. The broth was autoclaved followed by cooling and then stored at room temperature until used.

2.16 Processing WES data using Congenica Software

Firstly an Interpretation Request (IR) was completed which included information about the proband and any other family members and related clinical data including HPO terms (abnormal bleeding HP:0001892 and/or thrombocytopenia HP:0001873) for affected individuals. Relevant gene panels (Inherited Bleeding Disorder; High Evidence_Green, Medium Evidence Amber and Low Evidence Red, gene lists) containing 119 genes from Panel app https://panelapp.genomicsengland.co.uk/) were applied in the project and deemed suitable for research purposes. However, of this gene panel only 88 genes from the Genomics England website (R90) are considered as suitable for clinical use at this present time. The WES data (either BAM or FASTQ files) of patients were then transferred to the Congenica SFTP server for processing.

The Congenica pipeline could then be used for sequence alignment and variant calling of SNVs, small insertion/deletion (indels), and CNVs (Figure 2.3). The analytical pipeline for the detection of CNVs in genes involved in the IBDs panel was employed using the Exome Depth coverage approach. The exome read depth of the target patient's sample was compared against the read depth of a reference panel (up to 10 WES samples of each gender) to detect regions with different coverage which could represent a CNV event. Using the Congenica software, the lower limit that the Exome Depth calling software uses for CNV calling is ≥20 sequence reads. This ensures that Exome Depth does not consider low quality reads when comparing the reference samples to the target patient.

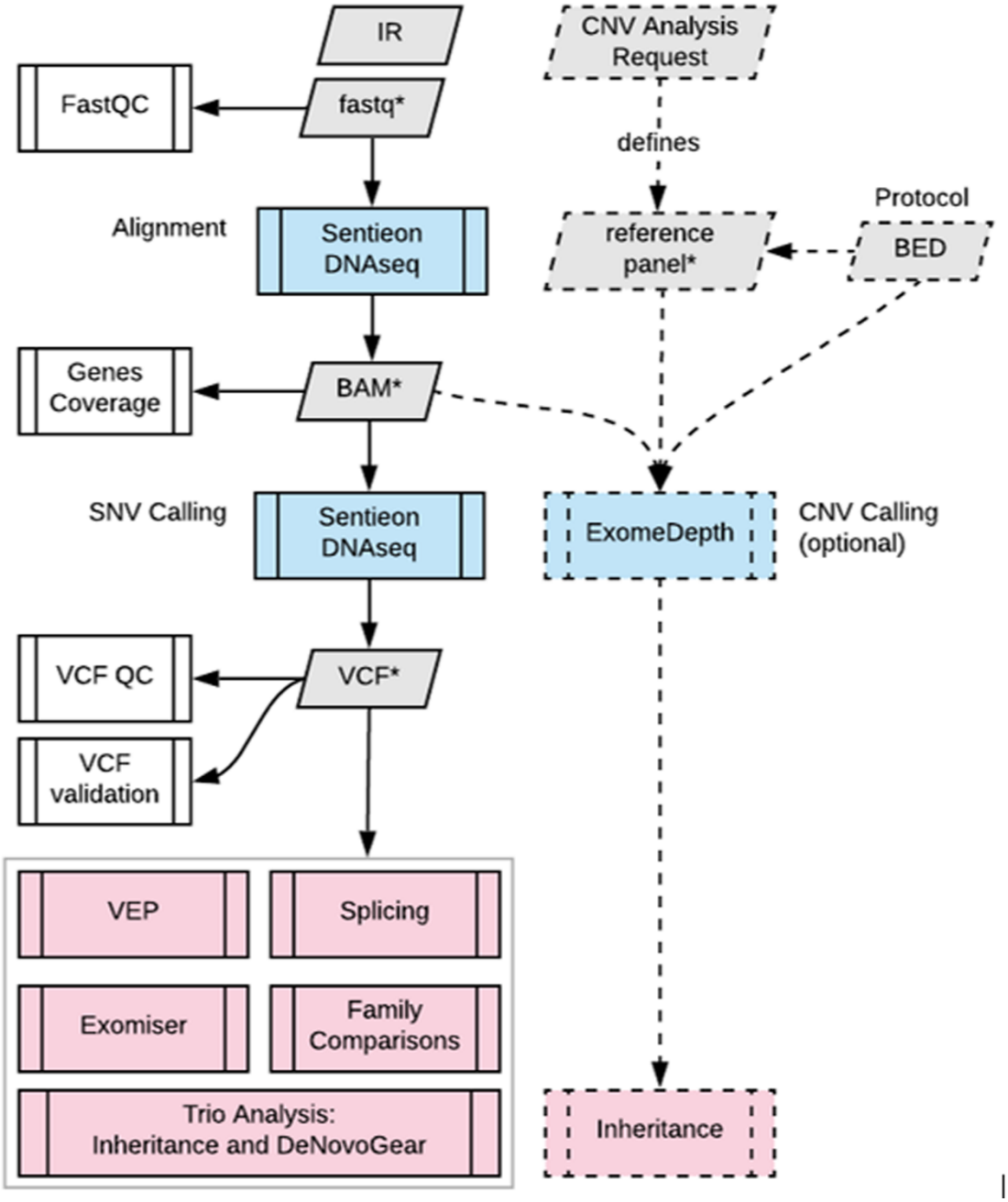


Figure 2.3. Congenica pipeline overview for processing of WES data. Adapted from (https://www.congenica.com/).

Chapter.3

Patient genotyping by whole exome sequencing

3.1 Summary of background to this research

Inherited bleeding and platelet disorders (BPDs) are an extremely heterogenous group of diseases that can affect approximately ~300 individuals per million births. Although some of these disorders such as haemophilia and von Willebrand diseases can be easily diagnosed with a combination of conventional laboratory assays of varying degrees of complexity and clinical assessments, the molecular analysis of patients with other BPDs is often limited. The significant advances in phenotyping assays have enhanced the molecular diagnosis, but the genotyping approach now provides the most accurate and complete diagnosis tool for some of these disorders (Simeoni et al., 2016, Sivapalaratnam et al., 2017a).

Molecular genetic sequencing has rapidly progressed since the invention of Sanger sequencing technology and subsequently completion of the human genome project (Consortium, 2004, Sanger et al., 1977). Nowadays, Sanger sequencing has evolved away from using radiography and gel electrophoresis into a more contemporary version that utilises fluorescence and capillary array electrophoresis, thereby it is commonly used in all genetic diagnostic services. Although it improves the efficiency and efficacy of outputs, there is still some limitations in sensitivity (Ansorge, 2009, Hall, 2007). Since the development of NGS technologies, there has been a fast progress in DNA sequencing with a substantial reduction in costs and increasing throughput, sensitivity and accuracy.. This progression provides vast amount of genome data and subsequently leads to a big shift in usage from small-scale applications of genetic sequencing into large-scale studies such as whole exome and genome sequencing (Liu et al., 2012). NGS involves multiple sequencing applications that provide massive parallel sequencing of the entire genome using (WGS) or (WES) (Goodwin et al., 2016). In addition, this improvement in sequencing technologies has led to rapid and

more accurate diagnosis of genetic diseases, thereby they have been widely adopted in clinical and research areas (Bamshad et al., 2011, Ng et al., 2010). The rate of disease-causing genes discovery has increased since the first WES proof-of-concept study in 2009 (Ng et al., 2009) (Figure 3.1). Improvements and the evolution of multiple NGS applications have helped researchers to understand the molecular mechanism of BPDs and elucidating the molecular aetiology behind many well-known disorders (Maclachlan et al., 2017). Therefore the question arises whether the identification of the molecular mechanisms behind rare bleeding disorders in single pedigrees with strong shared phenotype would benefit from using direct WES and whether this will identify a novel causative gene not previously associated with disease.

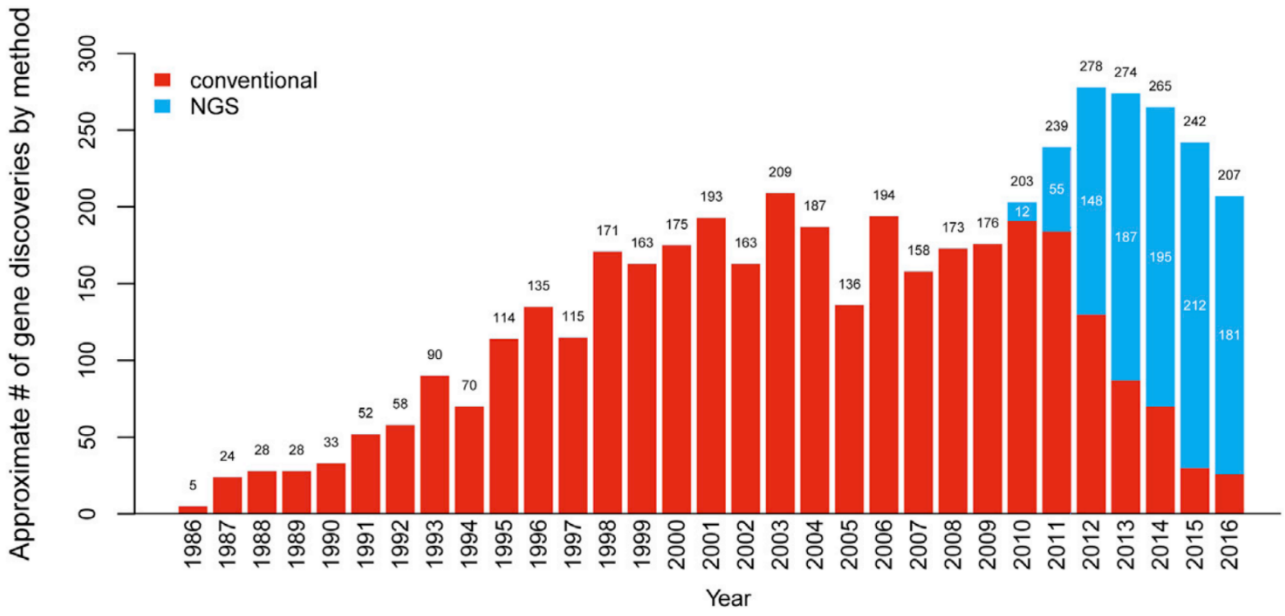


Figure 3.1. Approximate number of genes identified by WES and WGS versus conventional methods . Conventional methods were popularly used to identify rare disease-causing genes since 1986. Since the employment of NGS applications including WES and WGS (Blue) in 2010, the rate of discovering genes associated with rare genetic diseases has increased. Since 2013, the number of genes discovered by WGS and WES techniques were increased . Graph adapted from (Boycott et al., 2017, Silverstein and Febbraio, 2009).

3.2 Aims of this chapter

The aim for this chapter was to genotype two related family members and four unrelated individuals with a suspected IBD of unknown cause. These patients and family members were recruited to the UK-GAPP study. All patients were analysed using WES data coupled with a bioinformatic pipeline that has been developed by the GAPP study team in order to determine plausible candidate variants from WES data. The results will provide insight into the molecular mechanism of disease and help patients with an overall diagnosis as well as helping with treatment and disease management. It will provide insight into potential novel genetic variants which can be pursued further by functional investigation.

3.3 Patient characterisation

During the course of this research, four unrelated affected patients and two related affected patients with bleeding histories and low platelet counts were recruited to the UK-GAPP study and included in the WES analysis of this chapter. Written consent was taken from the patients to take part in the GAPP study and genetic research (Consent form is shown in Figure A in the appendices). The average age of the patients was 12 years and included three female and two male participants in addition to patient 4.I whose sex was not recorded. The platelet counts for all patients were lower than the normal range which averaged from 19-132 $x10^9$/L. Haematological parameters are shown in (Table 3.1). Mean platelet volume (MPV) wasmeasured to examine platelet size and values were found to be high in two patients and could not be detected by the Sysmex analyser in others. In addition, the immature platelet fraction (IPF) was measured and found to be high among all patients, ranging from 12.1-36.8%. Therefore, all patients were noted to have macrothrombocytopenia.

Table 3.1. Patient's details and haematological parameters. The normal range of platelet count (150-450 x10^9/L), MPV (7.5-11.5fL) and IPF (1.3-10.8%). M: male, F: female, IPF: immature platelet fraction, MPV: mean platelet volume, IT: inherited thrombocytopenia. The age of patients at the time of recruitment to the UK-GAPP study ranged between 12-16 years old.

Family/ Patient	Age range	Gender	Platelet count (10^9/L)	IPF (%)	MPV (fL)	Diagnosis
Family 1 / I	8-12	M	19	18	Undetectable	IT
Family 2 / I	12-16	F	68	13.2	13.8	IT
Family 3 / I	12-16	F	132	12.1	12.8	IT
Family 4 / I	12-16	-	45	16	Undetectable	IT
Family 5 / I	12-16	M	79	24.4	Undetectable	IT
Family 5 / II	12-16	F	87	36.8	Undetectable	IT

3.4 Whole exome sequencing

Whole exome sequencing was carried out on the genomic DNA of the 6 patients as previously outlined in Chapter 2, Section 2.8. All WES data of the 6 patients were analysed the using adapted bioinformatic pipeline method which is detailed in the next section. Patients 5.I and 5.II are discussed with further functional studies in Chapter 4.

3.4.1 Analysis of WES data by using an adapted bioinformatic pipeline

A bioinformatic pipeline developed by the GAPP study team was used to analyse WES data of the patients and subsequently identify plausible candidate variants (Figure 3.2). Sequence alignment against the reference genome, annotation and variant calling were performed as well as determination of the variants' novelty and frequency across multiple populations as described in Chapter 2, Section 2.8. WES data was formatted in variant call format (VCF) files which were then accessed using Microsoft Excel for filtering analysis. WES analysis was conducted by two different methods.

The first method used a candidate gene panel which compared variants including all genes implicated in platelet formation, count, function, lifespan or death and to all genes previously known to cause IT. The database of genes was developed by members of the UK-GAPP study which was first used by (Stockley et al., 2013) and then an extended version was created and employed. A wide range of variants, approximately 25,000 to 40,000 variants, were identified per single patient. A database of 358 known platelet-related genes which have been implicated in platelet formation, count, function, lifespan or death, were used initially to compare with the patient's genes in order to narrow down the number of candidate variants (Table B within the

appendix). These variants were then selected to be filtered based on the exclusion criteria created by the GAPP study as follows:

(i) A traditional rare variant cut off value of 0.01 or 1% of the ExAC, gnomAD and GAPP databases was used to filter out variants with minor allele frequency of >0.01 to priorities the remainder for further analysis.

(ii) Variants not known to change the amino acid or those that do not have a potential effect on protein, such as synonymous variants and intron variants were excluded.

(iii) Splice site variants occurring >4 base pair away from the exome were also excluded.

(iv) Comparisons with other affected and unaffected family members on the database were made to select candidate variants where they were shared between affected individuals and if not shared with other affected individuals or if shared with unaffected individuals, they were excluded

.

The second method of analysis was used to determine novel variants; where all non-novel variants were excluded. The novelty was determined by comparison of all variants with a set of databases. These included: Exome Variant Server (EVS), 1000 Genome database, Single Nucleotide Polymorphism database (dbSNP135) and Exome Aggregation Consortium (ExAC). Also, synonymous variants, not changing the amino acids sequence as well as splice site variants occurring >4 base pairs away from exon-intron boundaries were excluded. The remaining variants were compared with variants in other affected family members to exclude variants that are not shared with affected individuals. Taken together, all plausible candidate variants from both methods subsequently underwent in silico pathogenicity prediction in order to classify variants based on the ACMG guidelines.

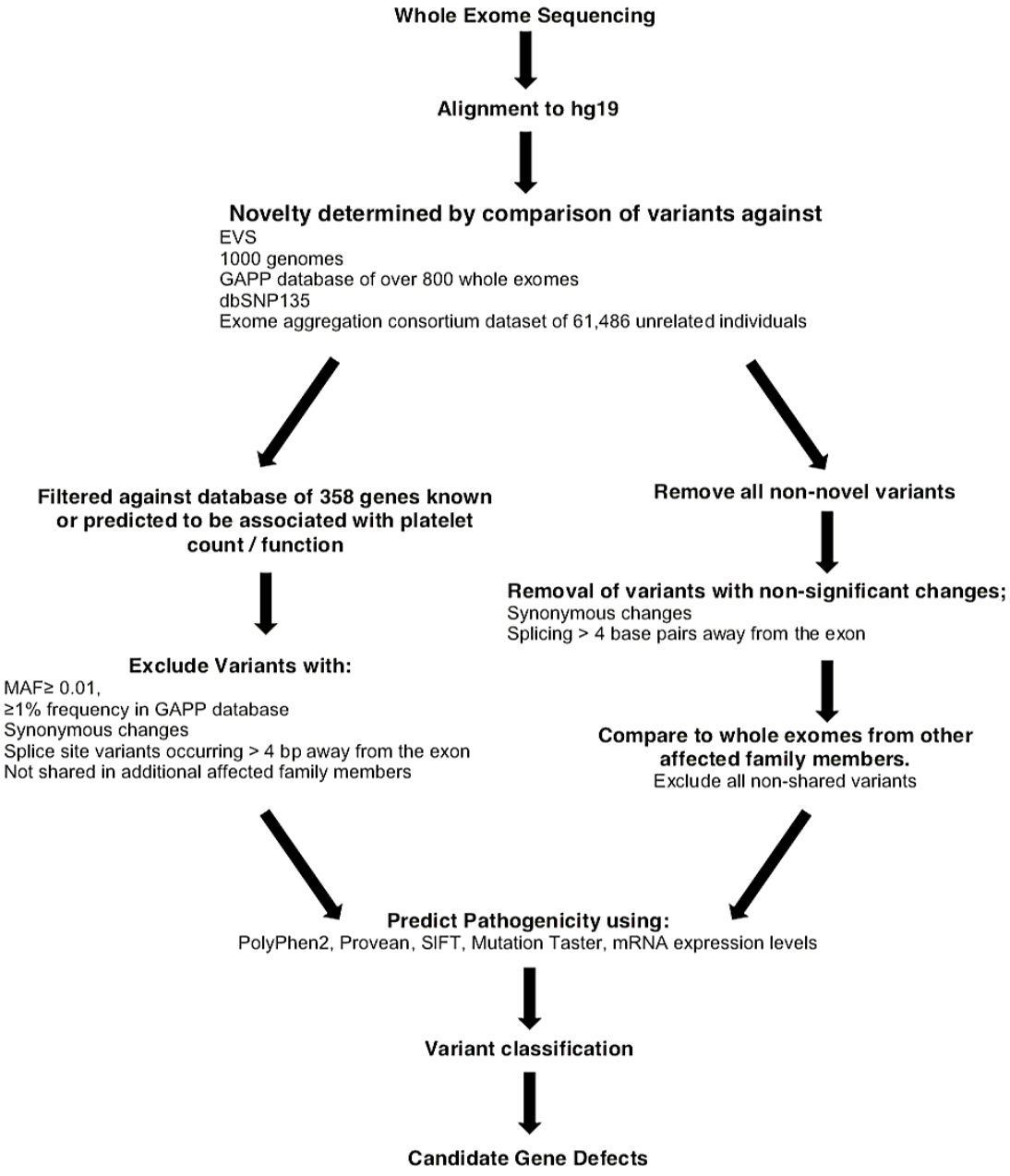


Figure 3.2. Bioinformatic pipeline for analysing WES data. Adapted from (Johnson et al., 2016b).

3.4.1.1 *In silico* pathogenicity prediction

The pathogenicity predictions of variants were assessed using five main analytical tools as outlined and described in Chapter 2, Section 2.8.1. These predicted tools could estimate the effect of candidate variants on the function and structure of the protein. Due to disparity in prediction tools outcomes, the pathogenicity prediction of some variants was not always in agreement with all the prediction tools and this resulted in difficulty of variant classification based on pathogenicity criteria. Therefore variants predicted to be damaging or deleterious by two or more prediction tools were taken forward for further analysis.

3.4.2 Development of a platelet count and function related gene panel

The database of 358 genes associated with platelet count/function used in the bioinformatic pipeline was developed by members of the UK-GAPP study that had been extended from the first version used by Stockley et al (Stockley et al., 2013). The primary list of genes created included genes known to be associated with platelet disorders such as those implicated in thrombocytopenia, platelet formation and lifespan, and plausible candidate genes from *in vitro* studies and animal models (Johnson et al., 2016a). The list has been updated with genes in previously published literature and encompasses genes from an N-ethyl-N nitrosourea (ENU) based mutagenic screen into platelets as performed by the group of Professor Ben Kile (Alexander et al., 2006).

3.4.3 Analysis of candidate variants

Following the bioinformatic pipeline analysis, plausible candidate variants from different patients were classified based on the ACMG guidelines as outlined above in Section 2.8.2. These were then selected for consideration and further functional analysis. The pathogenicity prediction of variants is considered as the most complex and challenging task in the genetic analysis, particularly for novel variants with limited evidence and in isolated cases where the true pathogenicity is unknown. The ACMG guidelines were followed as recommended which provided interpretive categories for assessment and subsequent variant interpretation (Richards et al., 2015). A proposed criteria for interpretation of sequence variants evaluate the evidences based on population, disease-specific and sequence databases as well as computational *in-silico* predictive programs. The 28 criteria that assess evidences are assigned by a specific code and all the codes are weighted (stand-alone, very strong, strong, moderate or supporting) and assigned significance (benign or pathogenic) based on the strength of evidence. Variant classification can be produced by combining these evidence codes to classify variants into categories named: pathogenic, likely pathogenic, benign, likely benign or uncertain significance. Variants that do not have enough evidence to be classified or have contradictory factors were assigned 'variant of uncertain significant' status which is most commonly seen with novel variants.

3.5 Results of WES in patients

3.5.1 Study 1: Candidate gene panel analysis.

WES analysis was carried out using the bioinformatic pipeline workflow which identified an average of 38,812 variants per individuals sample. The bioinformatic pipeline was used to refine the variants using series of filtering steps as displayed in (Figure 3.2). The variants were filtered against a panel of 358 genes known or predicted to be associated with platelet count, function or lifespan. On average, 222 variants from the gene panel were identified per individual. These variants where then filtered, excluding all synonymous and intronic variants followed by excluding all variants with a MAF of more than 0.01. On average, 16 sequence variants with a MAF of ≤ 0.01 were noted among patients in the panel of 358 platelet related genes. Pathogenicity of the variants was predicted by utilizing the prediction tools (Mutation Taster, PolyPhen-2, SIFT, Provean) and the variants were classified based on the ACMG guidelines. This gave an average of 9 sequence variants per patient. The resulting sequence variants are shown in (Table 3.2). Plausible candidate variants in each patient were then selected based on the pathogenicity prediction as seen in (Table 3.3). The ACMG consensus guidelines, including supporting evidence, are detailed in table E and F in appendices based on (Richards et al., 2015).

Table 3.2. Sequence variants filtered by bioinformatics pipeline. Multiple filtering steps were carried out to narrow down the number of sequence variants and elucidate potential plausible candidate variants. Each patient was analysed individually.

Patients	**1.I**	**2.I**	**3.I**	**4.I**	**5.I**	**5.II**
Total number of variants identified by WES	36,040	36,106	35,990	36,971	43,774	43,995
Total number of variants within 358 platelet genes panel	220	213	219	208	233	243
Total number of variants (excluding synonymous and intronic variants)	115	122	120	119	149	135
Total number of variants with MAF ≤ 0.01	15	10	14	8	26	24
Total number of variants after pathogenicity prediction	5	6	4	4	19	17
Total shared variants	NA	NA	NA	NA	11	11
Plausible candidate variants	3	2	3	4	1	1

Table 3.3. Table of candidate genes after bioinformatic pipeline analysis. The table shows the nucleotide change and predicted protein effect as well as multiple pathogenicity prediction tools for assigning the variant classification based on ACMG criteria. SIFT: Sorting intolerance from tolerance. Hom: homozygous. Het: heterozygous. ACMG: The American College of Medical Genetics and Genomics. NA: Not available. Allele frequency obtained from ExAC database. Genes in bold are from the inherited bleeding disorders related gene panel. The remaining genes are from the database of 358 genes developed by members of the UK-GAPP study. All ACMG criteria are described in Table E and F in appendices.

Patient	Gene/ gene transcript	Nucleotide change	Predicted protein effect	Variation type/ (zygosity)	Allele frequency	Mutation taster	PolyPhen-2	SIFT	Provean	ACMG criteria	Classification
1.I	*FCER1G* NG_029043.1	c.173A>G	p. Tyr58Cys	Missense (Het)	0.000008237	Disease causing	Possibly damaging	Tolerated	Deleterious	PP3,PP2	Uncertain significance
	FHOD1 NG_029672.1	c.1525C>G	p. Arg509Gly	Missense (Het)	0.00001817	Disease causing	Benign	Tolerated	Neutral	PP2	Uncertain significance
	FMNL1 NA	c.1807_1809delCCG	p.Pro612del	Inframe deletion (Het)	Novel	NA	NA	NA	Neutral	PM2, PM4	Uncertain significance
2.I	***NBEAL2*** NG_031914.1	c.7369C>T	p. Arg2457Trp	Missense (Het)	0.00005808	Disease causing	Possibly damaging	Deleterious	Deleterious	PP3,PP2	Uncertain significance
	ITGA2B NG_008331.1	c.3076C>T	p.Arg1026Trp	Missense (Het)	Novel	Disease causing	Probably damaging	Deleterious	Deleterious	PM1,PM2, PP2,PP3, PP5,PS4	Pathogenic
3.I	*DNM3* NA	c.1072G>A	p. Gly358Ser	Missense (Het)	Novel	Polymorphism	Probably damaging	Deleterious	Deleterious	PP3,PP2	Uncertain significance
	ARHGEF3 NA	c.1127G>A	p.Arg376His	Missense (Het)	0.0002224	Disease causing	Possibly damaging	Deleterious	Deleterious	PP3	Uncertain significance

	ABCG8 NG_008884.2	c.1629G>T	p.Arg543Ser	Missense (Het)	0.0002308	Disease causing	Probably damaging	Deleterious	Neutral	PS1,PM1, PM2,PP2, PP5,BP4	Pathogenic
4.I	*RAI1* NG_007101.2	c.311A>G	p. Asp104Gly	Missense (Het)	0.00001823	Disease causing	Possibly damaging	Deleterious	Neutral	PP3,PP2	Uncertain significance
	ACVRLA1 NA	c.1445C>T	p.Ala482Val	Missense (Het)	0.001673	Disease causing	Probably damaging	Deleterious	Deleterious	PP3,PP2	Uncertain significance
	FLNA NG_011506.2	c.6143G>A	p.Ser2048Asn	Missense (Het)	Novel	Disease causing	Benign	Deleterious	Deleterious	PP3,PP2, PM2,PM5	Likely pathogenic
			p.Pro48_Pro49del								
	ORAI1 NG_007500.1	c.127_132delCCGCCA		Inframe deletion (Hom)	Novel	Disease causing	NA	NA	NA	PM2,PM4	Uncertain significance
5.I	*CD36* NG_008192.1	c.975T>G	p. Tyr325Ter	Stop gain (Het)	0.007914	Polymorphism	NA	NA	NA	PVS1,PS3, PP1,PP5	Pathogenic
5.II	*CD36* NG_008192.1	c.975T>G	p. Tyr325Ter	Stop gain (Het)	0.007914	Polymorphism	NA	NA	NA	PVS1,PS3, PP1,PP5	Pathogenic

Following patient analysis of WES data using the bioinformatic pipeline, there were 13 potential disease causing genetic variants in genes: *FCER1G, FHOD1, FMNL1, NBEAL2, DNM3, ARHGEF3, RAI1, ACVRLA1, ORAI1* which were classified as variants of uncertain significance. One likely pathogenic variant within *FLNA* was found in patient 4.I. Three pathogenic variants within ITGA2B, *ABCG8* and *CD36* were observed in patients 2.I, 3.I, 5.I and 5.II respectively. Three patients including 2.I, 3.I and 4.I were noted with variants in genes previously implicated in IT. These were as follows: Patient 2.I harboured two missense variants in *NBEAL2* (c.7369C>T; p. Arg2457Trp) which classified as uncertain significance and a novel *ITGA2B* variant (c.3076C>T; p.Arg1026Trp) classified as pathogenic. Patient 3.I presented with pathogenic variant in *ABCG8* (c.1629G>T; p.Arg543Ser). Patient 4.I was noted to have a likely pathogenic missense variant within *FLNA* (c.6143G>A; p.Ser2048Asn), which was novel.

Three patients including 1.I, 5.I and 5.II were observed with variants in genes from the database of 358 gene panel described previously in Section 3.4.2. Patient 1.I was identified with novel inframe deletion variant within *FMNL1* (c.1807_1809delCCG; p.Pro612del). Patient 5.I and also related family member patient 5.II were noted to have a stop gain variant within *CD36*: c.975T>G; p. Tyr325*, which is predicted to truncate the coded protein sequence at amino acid 325.

3.5.2 Study 2: Screening for unknown novel genes.

3.5.2.1 Background and patient details

The family studied here (Family A) is a large family of Pakistani origin with a history of cousin-cousin marriages which were recruited to the GAPP study (Figure 3.3, A). Two phenotypes are presented within the family members: excessive bleeding and oculocutaneous albinism (OCA) which are mainly associated with Hermansky-Pudlak syndrome. However, these two phenotypes do not always segregate together in the same individual where many individuals present only one phenotype (Figure 3.3). This feature suggests that the bleeding and albinism may be caused by separate gene defects. Homozygous or compound heterozygous mutations in the *TYR* gene were identified previously via WES analysis, which cause oculocutaneous albinism 1 (OCA1). A study has been conducted on 17 patients with non-syndromic OCA. The study was aimed to screen patients for whole exons and flanking regions of *TYR* gene. A total of 12 mutations were identified in 10 patients. Of these, two patients were identified with homozygous mutations while eight patients carried compound heterozygous and none of them presented with bleeding episodes (Sun et al., 2018). The genetic cause of the bleeding phenotype in these patients remains unresolved. This section will analyse the WES of some affected individuals from Family A as seen in Figure 3.3 to identify the genetic cause of bleeding and by employing segregation analysis on all other individuals will help to determine the shared genetic variants among affected members which could explain the cause of bleeding tendency.

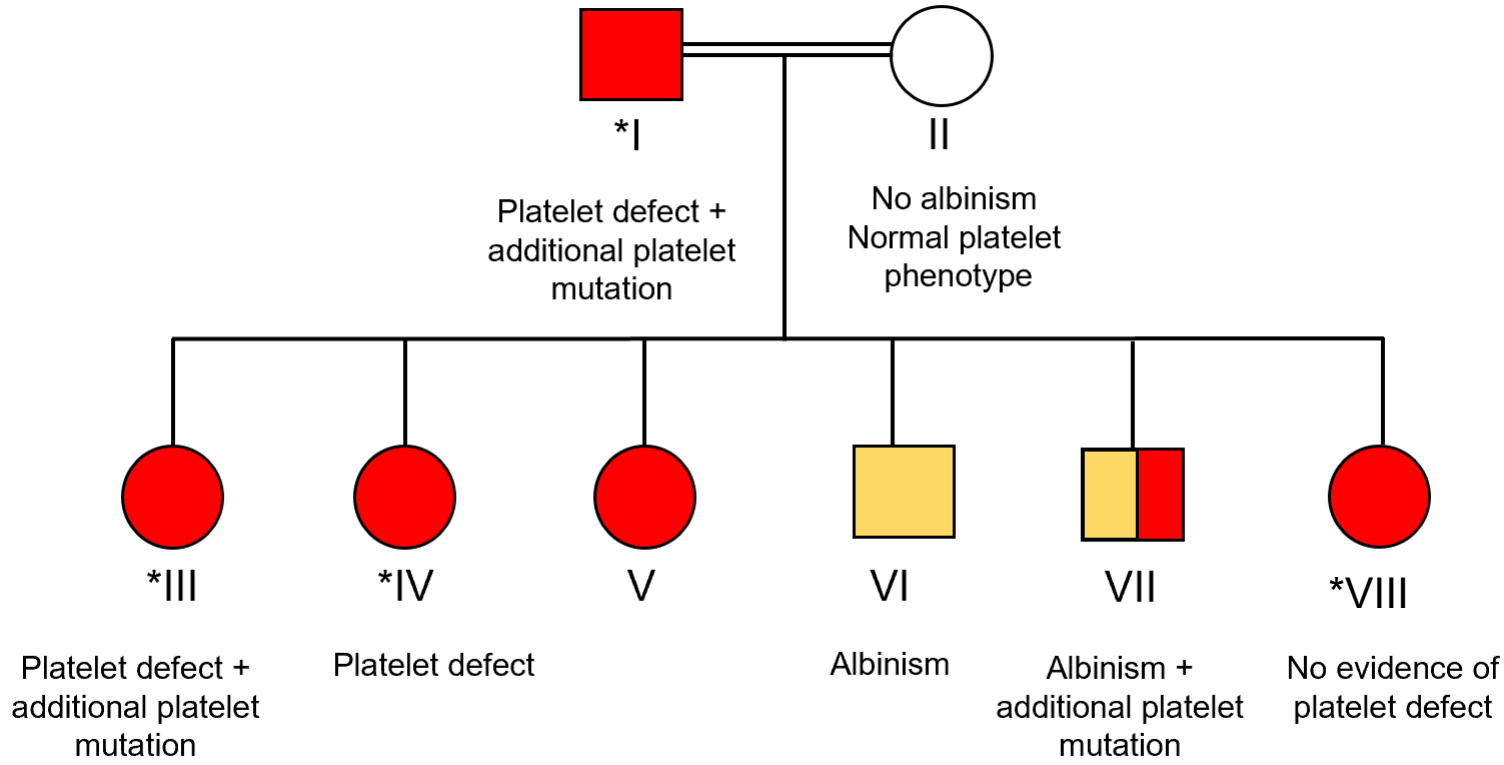


Figure 3.3. Pedigree of consanguineous Pakistani origin family. Pedigree of the family members with excessive bleeding phenotype and albinism. Highlighted red color indicating bleeding tendency, white color indicating unaffected family member and light yellow represent albinism. Patients underwent WES analysis are indicating with a star symbol.

3.5.2.2 Bleeding phenotype and BAT score

Patients I, III, IV, V, VII and VIII have had a history of variable bleeding episodes however the mother, II, and the eldest son, VI reported with no bleeding symptoms. The International Society of Haemostasis and Thrombosis bleeding assessment tool (BAT) was used at the referral centre which provided bleeding scores for each patient based on the patients evaluation and clinical history and as seen in (Table 3.4) (Rodeghiero et al., 2010).

Table 3.4. Patient details, corresponding phenotypes, BAT score and platelet defects. All patients underwent platelet function testing at the time of recruitment except patient VI. COX: cyclo-oxygenase deficiency, Gi: Gi receptor signalling Age is given in years at the time of recruitment, NA: Not available, M: male, F: female, BAT: bleeding assessment tool. See figure B in appendices as an example of the patient questionnaire and scoring system.

Patient	Age	Gender	Bleeding phenotypes	BAT Score	Platelet defect
I	44	M	Epistaxis Bleeding of minor wounds Oral cavity bleeding Bleeding after surgery Excessive bleeding on venepuncture	8	Gi
II	44	F	None	0	Normal
III	22	F	Oral cavity bleeding Bleeding postpartum	2	Gi
IV	24	F	Bleeding of minor wounds Oral cavity bleeding Menorrhagia Excessive bleeding on venepuncture	7	COX
V	21	F	Bleeding after tooth extraction Excessive bleeding on venepuncture	4	Not confirmed
VI	NA	M	None	0	Not tested
VII	16	M	Epistaxis Bleeding of minor wounds Oral cavity bleeding	3	Normal
VIII	14	F	Reported excessive bleeding	NA	Not confirmed

3.5.2.3 Platelet function testing

Platelet function testing was carried out on patients I, II, III, IV, V, VII and VIII while patient VI was not tested. Patient II, an unaffected family member, showed no platelet defect after aggregometry testing as expected. However, patient VII showed normal response to all agonists except for low dose collagen which showed slight disaggregation at 1µg/ml. The remaining affected patients showed reduced aggregation particularly in response to ADP and adrenaline. In addition, patient VII showed reduced aggregation at low doses with arachidonic acid, collagen and PAR-1. Patient I and III showed disaggregation at 30µM of PAR-1 peptide as well as reduced primary wave and absent secondary wave of aggregation to adrenaline which are compatible with a Gi-like defect. Patient IV showed reduced or absent aggregation response to all agonists particularly 10µM ADP and absent aggregation with arachidonic acid which is considered to be most likely relatable to COX (Cyclooxygenase) defects. Patient V showed absent or reduced response to all agonists but the specific platelet defect is still unclear. There is not enough evidence to suggest a platelet defect in patient VIII because all aggregation results showed low response at low doses while high doses showed normal responses. In summary, platelet function testing was carried out on all patients except patient VI who was not tested. The mother, II , showed normal platelet function as expected and patient VII normal response to all agonists except low dose of collagen. Therefore, patient VII is considered to have normal platelet function and thus bleeding symptoms must be due to a possible collagen defect as the patient showed reduced aggregation at low dose of collagen . However, a clear platelet function defect was identified in the other affected family members with varying degrees of impaired Gi signalling in patients I and III and most likely COX defect in patient IV.

3.5.2.4 Whole exome sequencing

Patients I, III, IV and VIII were selected for whole exome sequencing. WES analysis has been previously done using the panel of 358 known platelet related genes (left side of pipeline in Figure 3.2 but no plausible candidate variants were identified. However, this study used the right hand analysis in Figure 3.2 for identification of novel variants. 25,205 variants were identified by variant calling in patient I while 24,884 variants were identified in patient III. These variants were subsequently filtered out using the established filtering pipeline analysis for novel variants. The bioinformatic pipeline was used to narrow down the total variants using a series of filtering steps. All non-novel variants where excluded and an average of 2594 variants (novel and not found in population frequency databases) were identified (Table 3.5). An average of 553 variants were identified after exclusion of all synonymous variants and splicing variants >5bp away from the exon. All variants within genes that have low mRNA expression in progenitor populations of the haematopoietic lineage were excluded. In total, 147 variants were shared between patients I and III. When these variants were compared with other affected family members IV and VIII, a total of 58 variants were shared between all four individuals. Of 147 variants, there were no homozygous variants identified in *HPS* gene family that are known to cause Hermansky-Pudlak syndrome with a bleeding phenotype. In addition, shared homozygous plausible candidate variants were not found among the four patients which suggests that a heterozygous variant or pseudo-dominant inheritance might cause the bleeding disorders. The pathogenicity prediction of the variants were determined by utilizing the prediction tools (Mutation Taster, PolyPhen-2, SIFT, Provean) and the variants predicted to be non-pathogenic in at least one bioinformatic tool were excluded. Three plausible candidate variants shared between all affected family members were then selected based on the pathogenicity prediction in (Table 3.6).

Table 3.5. Sequence variants filtered by bioinformatics pipeline. Multiple filtering steps were carried out to narrow down the number of sequence variants and elucidate potential plausible candidate variants.

Patients	**I**	**III**
Total number of variants identified by WES	25,205	24,884
Removal all non-novel variants	2158	3030
Total number of variants (excluding synonymous and intronic variants)	454	652
Total number of variants with high mRNA expression in MSC lineage	325	336
Total shared variants	147	147
Total shared variants among 4 patients	58	
Exclusion of all predicted to be non-pathogenic in at least 1 bioinformatic tool	11	
Plausible candidate variants	3	

3.5.2.5 WES data analysis to determine candidate variants

This large Pakistani family has been investigated for disease causative variants. Three novel candidate variants in genes not known previously to cause bleeding disorders were identified in the four affected family members I, III, IV and VIII (Table 3.6). These include: inframe insertion variant c.1331-1336dupAGGAGG; p.Glu444-Glu445dup in *NADK* gene, a splice variant c.1881+3A>G in *CEBPZ* and missense variant c.1285G>T; p.Ala429Ser in *NET1*. The allele frequencies of these variants were zero and have never been found in ExAC and GnomAD databases. Segregation analysis was conducted on all affected and unaffected family members to determine if these candidate variants are co-segregated in affected individuals. This analysis identified only a splice variant c.1881+3A>G in *CEBPZ* segregated in affected family members while patient II and VI were wild type as expected due to their absence of bleeding episodes (Figure 3.4). However, platelet function testing in patient VII was normal and thus bleeding symptoms must be due to a collagen defect as the patient showed reduced aggregation at low dose collagen and will be investigated in future work. The splice variant is located upstream of the intron between exon 3 and 4 which could interrupt the splicing process. This would result in subsequent skipping of the exon or create a cryptic splice site, or reduced/absent expression of the mutant allele through nonsense-mediated RNA decay (Bruno et al., 2011). Furthermore, *in silico* expression of NADK, CEBPZ and NET1 in cells of the haematopoietic cell lineage were determined by using the Blueprint epigenome data (Chen et al., 2014). RNAseq data showed high expression of CEBPZ and NET1 in all cells of the haematopoietic lineage as opposed to low expression of NADK (Figure 3.5).

Table 3.6. Summary of candidate genes identified by WES using right hand analysis of bioinformatics pipeline. The table shows the genomic variation and protein effect as well as two pathogenicity prediction tools for assigning the variant classification based on ACMG criteria. SIFT: Sorting Intolerance From Tolerance. Het: heterozygous. NA: not available.

Gene/gene transcript	Zygosity	Variation type	Nucleotide change variation	Predicted protein effect	Allele frequency	PolyPhen-2	SIFT	ACMG criteria	Classification
NADK / NA	Het	Inframe insertion	c.1331_1336dupAGGAGG	p.Glu444_Glu445dup	0	NA	NA	PM2	Uncertain significance
CEBPZ / NG_050962.1	Het	Splicing	c.1881+3A>G	NA	0	NA	NA	PM2	Uncertain Significance
NET1 / NG_050652.1	Het	Missense	c.1285G>T	p.Ala429Ser	0	Possibly damaging	Deleterious	PM2 PP3	Uncertain significance

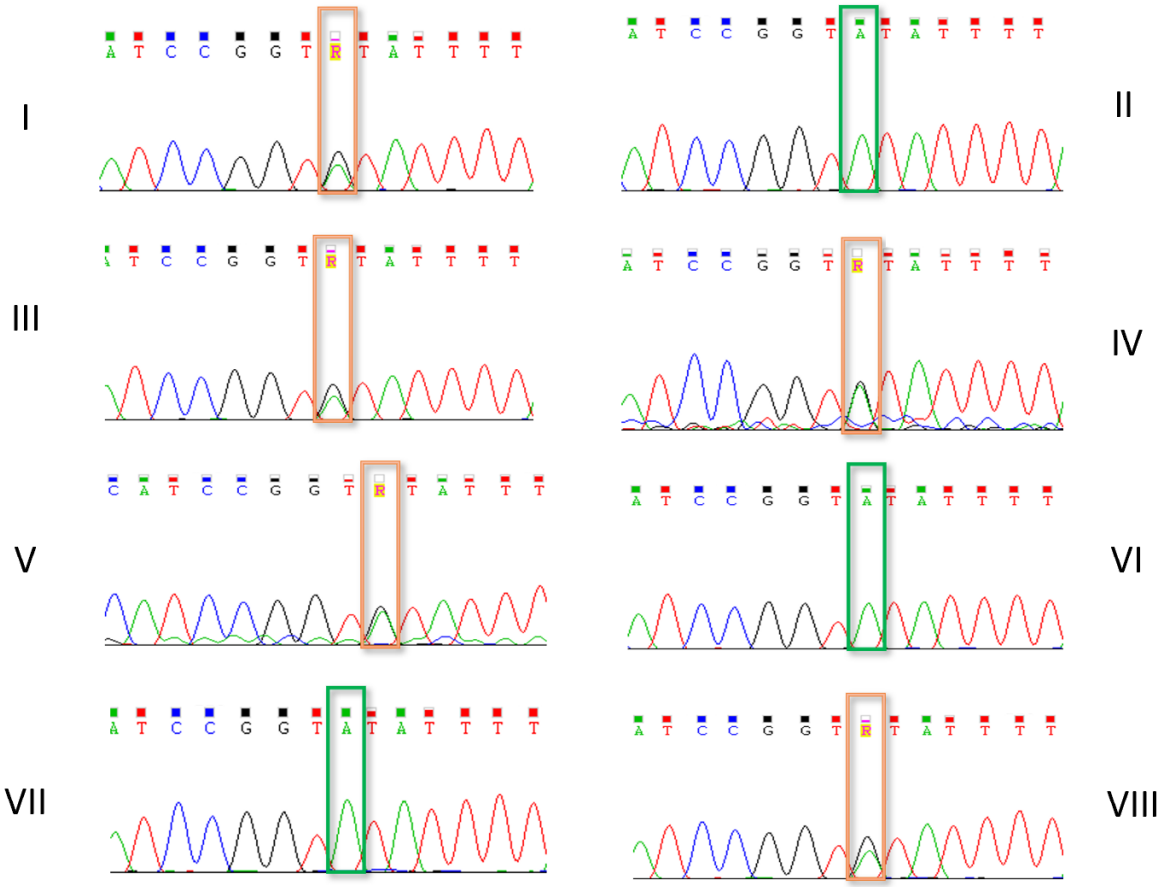


Figure 3.4. Confirmation of *CEBPZ* variant by Sanger sequencing. The results show the presence of heterozygous A>G variant with highlighted orange box in all affected family members except VII who may harbour a causative variant in non-platelet origin. The unaffected family members II and VI show A>A homozygous as wild type as highlighted with green box. Red: T, Blue: C, Black: G, Green: A.

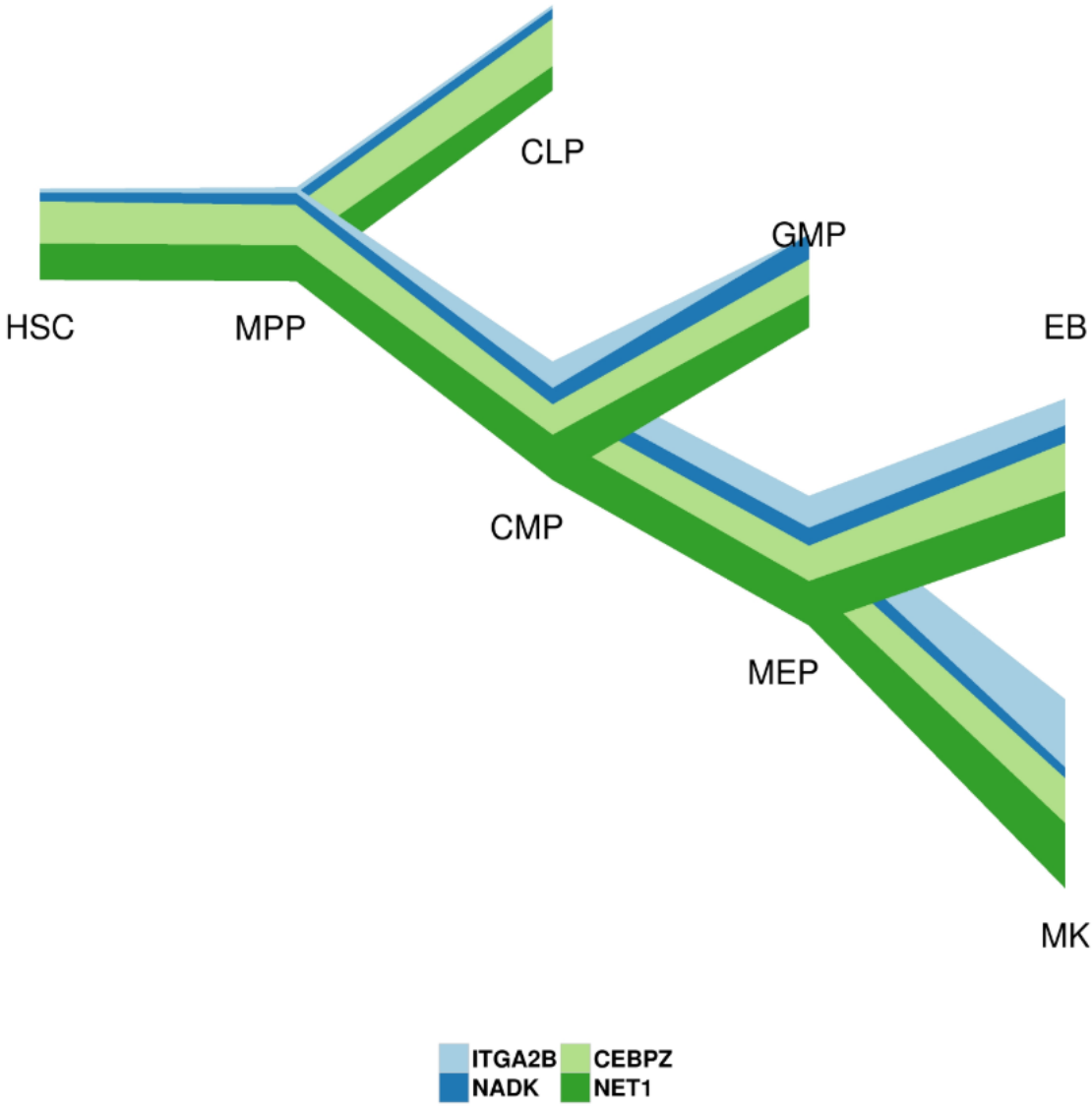


Figure 3.5. In silico expression analysis of *NADK*, *CEBPZ* and *NET1* in cells of the haematopoietic cell lineage. High expression of *CEBPZ* and *NET1* were noted in all cells of the haematopoietic lineage as opposed to low expression of *NADK*. HSC: Haematopoietic stem cell, MPP: Mulitpotent progenitor, CMP: Common myeloid progenitor, MEP: Myelo-erythroid progenitor, EB: Erythroblast, and MK: Megakaryocyte. The blot shows the expression level of mRNA for candidate genes in different haematopoietic cells progenitors using the blueprint consortium progenitor software provided by the University of Cambridge as well as RNA-seq data from the study of human and mouse platelet transcriptomes as mentioned in Section 2.8.1.

3.6 Discussion

There is no doubt that the emergence of NGS technologies such as WGS and WES have increasingly resulted in a massive advancement in the understanding of the molecular basis of rare genetic diseases. Although many rare genetic diseases are still without a feasible treatment, understanding the molecular basis of these diseases can provide an accurate diagnosis and great help for patients and their families (Fernandez-Marmiesse et al., 2018). WES approaches have increasingly been used over the last decade in the molecular diagnosis of IBDs. This chapter aimed to use WES data of patients recruited to the UK-GAPP study with IPDs of unknown causes. Genotype and phenotype data coupled with Sanger sequencing approaches can be used to diagnose patients with bleeding disorders in order to identify the molecular causes behind their disorders. A novel pipeline workflow for analysing WES data was developed by members of the UK-GAPP study which can be used to determine candidate variants. The pipeline was designed to be used in two different ways. Firstly, to analyse variants within genes associated with platelet formation, count, function and lifespan while the second strategy was to assess all novel variants in patients where there are not likely to be any candidate gene variants present. The bioinformatic pipeline uses a traditional rare “cut-off” value of prevalence (0.01) within the population databases. All splice site variants which exceeded 4 base pairs from the intron-exon boundary were excluded. It has been shown that the majority of splice site variants are believed to locate at positions ±1 and 2 from the intron-exon boundary (Krawczak et al., 1992, Anna and Monika, 2018). The presence of splice variants at these positions could affect the length of mRNA and subsequently the protein function by exon skipping or intron inclusion (Richards et al., 2015). All synonymous variants with no significant effect on the function of the protein were also excluded to narrow down

the variants. A panel of 358 genes associated with platelet formation, count, function and lifespan has been developed within the UK-GAPP study which includes data from published literature and animal models of platelet disease.

In study one, 6 patients including 4 isolated individuals and 2 related individuals with bleeding episodes and low platelet counts were recruited to the UK-GAPP study with the average patient age at 12 years old. Platelet counts for all patients were lower the normal range which averaged 19-132 $x10^9$/L. All patients presented with high MPV and IPF. The bioinformatic pipeline workflow was used to analyse WES data. The left side assessing 358 platelet associated genes was used to filter out variants. WES analysis identified 13 plausible candidate variants in genes *FCER1G, FHOD1, FMNL1, NBEAL2, DNM3, ARHGEF3, RAI1, ACVRLA1, ORAI1* which were classified as variants of uncertain significance. One likely pathogenic variant in *FLNA* and three pathogenic in ITGA2B, ABCG8 and CD36 were found. Patient 1.I was identified with a novel in-frame deletion variant within *FMNL1* (c.1807_1809delCCG; p.Pro612del). The variant was classified as a variant of uncertain significance. Patient 2.I harbours a novel missense variant in a gene previously implicated in IT (*ITGA2B:* c.3076C>T; p.Arg1026Trp) which is classified as pathogenic based on the ACMG guidelines. This variant has been reported as pathogenic variant in some databases such as NCBI and LOVD. The platelet surface receptor αIIbβ3, the receptor for fibrinogen, is encoded by *ITGA2B* and *ITGB3*. Genetic variants in these two genes cause qualitative deficiencies in the αIIbβ3 receptor which subsequently cause Glanzmann thrombocytopenia (GT) (Nurden et al., 2013). Patients with GT are often presented with variable bleeding phenotypes and the platelet number and size are often unaffected. However, some patients with genetic variants in *ITGA2B* and *ITGB3* were reported to have reduced platelet count and marked increase of platelet size (Ghevaert

et al., 2008, Khoriaty et al., 2019). This is applicable with what was found in patient 2.I as the patient was presented with macrothrombocytopenia. Patient 3.I presented with a pathogenic missense variant in *ABCG8* (c.1629G>T; p.Arg543Ser). Genetic mutations in *ABCG5* and *ABCG8* , two adjacent ATP-binding cassette transport genes, have been reported to cause a rare autosomal recessive disorder called Sitosterolemia which characterised by increasing of plasma sterols (Bastida et al., 2020, Desai et al., 2021). These genes regulate the absorption and excretion of sterols from circulation. When mutated, the accumulation of circulating plasma sterols lead to platelets death and subsequent thrombocytopenia (Hubacek et al., 2001). It has been reported that 30% of patients with Sitosterolemia are associated with macrothrombocytopenia (Bastida et al., 2020). The presentation of reduced platelet count in patient 3.I could result from the pathogenic missense variant in ABCG8. Patient 4.I was noted with a likely pathogenic missense variant within *FLNA* (c.6143G>A; p.Ser2048Asn) which was novel in the population frequency databases. As mutations within *FLNA* gene have been reported in patients with variable platelet size and low platelet count, this could be applicable with what was found in patient 4.I (Nurden et al., 2011). Patient 5.I and 5.II were noted with stop gained variants within *CD36*: c.975T>G; p. Tyr325*, which is predicted to truncate the coded protein sequence at amino acid 325. These patients will be discussed in more detail in Chapter 4. All these candidate variants have not been reported yet to the relevant clinician which need further investigation to confirm if they are associated with bleeding disorders in these patients. Recruitment of more affected family individuals with similar phenotypes will help in the diagnosis of these patients in order to confirm if these variants segregate in family members. Determining the cause of the bleeding disorder

could help in patient management and to provide patients with appropriate treatment to prevent excessive bleeding.

In study two, the right hand side of the bioinformatic pipeline was used to assess variants within novel genes because the left side was used previously in our lab on this data set and no likely candidate variants were identified. It has been shown that the chance of having AR disorders is increased with consanguineous marriages due to the inheritance of AR gene mutations from a common ancestor (Zlotogora, 2002, Hamamy et al., 2011). However, a study was undertaken on patients to analyse the effect of consanguinity on different types of genetic diseases. The study showed that consanguineous marriage was recorded in 51.5% of cases with autosomal dominant diseases (Shawky et al., 2013). This could explain the presence of heterozygous variants shared among the patients.

Some genetic disorders are genetically heterogenous and caused by mutations that result in different phenotypes where the condition is syndromic. Hermansky-Pudlak syndrome is one example in which patients with this disorder presented with different clinical manifestations, such as oculocutaneous albinism, prolonged bleeding and congenital neutropenia. It is mostly caused by mutations in the *HPS* gene family (Suzuki et al., 2002, Wei, 2006). This indicates that some genetic mutations can cause genetic disorders with different phenotypes which make the genetic diagnosis more difficult. However, variants in *TYR* gene are less likely to be causative of the bleeding phenotype because mutations in the *TYR* gene were identified to cause oculocutaneous albinism 1 (OCA1) without any bleeding episodes (Sun et al., 2018). This suggests that the cause of bleeding is not associated with *TYR* gene and it is more likely to be other genes. Following WES analysis, three novel candidate variants were found to be shared between family members I, III, IV, V and VIII in a large

consanguineous family of Pakistani origin who have a history of excessive bleeding and albinism. Segregation analysis was carried out on all affected and unaffected family members to determine which variants segregate with the bleeding phenotype. This identified only a heterozygous *CEBPZ*; c.1881+3A>G variant, which was shared among affected family members. The splice variant is located upstream of the intron between exon 3 and 4 which could interrupt the splicing process and subsequently lead to exon skipping. *CEBPZ* , CCAAT/enhancer binding protein, is a member of a family of transcription factors of the basic leucine-zipper class that plays an important role in haematopoietic differentiation especially myeloid differentiation through the binding to CCAAT/ enhancer box within promoter regions of genes (Musialik et al., 2014, Pulido-Salgado et al., 2015, Barbieri et al., 2017). Exome sequencing was carried out on 34 AML patients. Two patients were identified with mutations in *CEBPZ* which suggest that *CEBPZ* is a novel mutated gene in AML (Herold et al., 2014). As *CEBPZ* is involved in the myeloid differentiation, it may play an important role in megakaryocyte differentiation and subsequent platelet formation. However, its role in megakaryocyte differentiation and platelet formation is still unclear which needs further studies. Furthermore, *in silico* expression of *CEBPZ* in cells of the haematopoietic cell lineage that were obtained from RNAseq data of the Blueprint epigenome showed high expression of *CEBPZ* in all cells of the haematopoietic lineage. Thus, this needs further functional studies to investigate this *CEBPZ* variant and how likely it is associated with the bleeding disorder and platelet function. A heterozygous *CEBPZ* nonsense mutation leading to truncation of the protein product could result in loss of the mutated allele through nonsense-mediated decay (NMD) leading to haploinsufficiency as a mutational mechanism. Alternatively, the mutation could lead

to a dominant negative result where the loss of the mutant allele affects the normal function of the protein product. The summary of family A is shown in Table 3.7.

Table 3.7. Summary of Family A in study 2. Patient details, corresponding phenotypes, BAT score and platelet defects. The last column indicates the zygosity of the segregation result conducted on patients. COX: cyclo-oxygenase deficiency, Gi: Gi receptor signalling. Age is given in years, NA: Not available, M: male, F: female, BAT: bleeding assessment tool. Hom: homozygous, Het: heterozygous and WT: wild type.

Patient	Age	Gender	Bleeding phenotypes	BAT Score	Platelet defect	Zygosity of *CEBPZ* varaint
I	44	M	Epistaxis Bleeding of minor wounds Oral cavity bleeding Bleeding after surgery Excessive bleeding on venepuncture	8	Gi	Het
II	44	F	None	0	Normal	WT
III	22	F	Oral cavity bleeding Bleeding postpartum	2	Gi	Het
IV	24	F	Bleeding of minor wounds Oral cavity bleeding Menorrhagia Excessive bleeding on venepuncture	7	COX	Het
V	21	F	Bleeding after tooth extraction Excessive bleeding on venepuncture	4	Not confirmed	Het
VI	NA	M	None	0	Not tested	WT
VII	16	M	Epistaxis Bleeding of minor wounds Oral cavity bleeding	3	Normal	WT
VIII	14	F	Reported excessive bleeding	NA	Not confirmed	Het

In summary, genotype and phenotype data from the GAPP study provided valuable information about the genes involved in platelet production regulation, the platelet physiological process and their involvement in haemostasis. This data can be applied to investigate patients with platelet function disorders and subsequent bleeding history. Although the bioinformatic pipeline used here has some limitations such as being time-consuming and difficult to identify copy number variants (CNVs), it is still a useful tool to analyse small cohort of patients. Further automated bioinformatic tools are becoming available such as the Congenica which is explored in chapter 5. This chapter has demonstrated that the application of combined phenotyping and genotyping coupled with WES technology and bioinformatic tools are an efficient and effective approach for refining and identifying new sequence variants in patients with suspected IBDs. They can provide invaluable information about the genes involved in platelet production and platelet physiology subsequently involved in haemostasis.

3.7 Future work

The effectiveness of molecular next generation sequencing analysis relies on subsequent functional studies to confirm the effect of the genetic variants on gene and protein function. Future work would be to carry out functional assays to investigate these variants and how likely they are associated with diseases. Pathogenicity *in silico* prediction tools can be used to classify variants but without *in vitro or in vivo* studies, the role of these variants in the mechanism of diseases are often ambiguous and unclear.

Future work needs to be done to investigate the effect of the *CEBPZ* splice variant. To evaluate this variant further, an end point and/or quantitative reverse transcription polymerase chain reaction (qRT-PCR) can be performed using specific *CEBPZ* primers to investigate if this splice variant results in differential expression of the gene specifically in platelets. Moreover, inducible pluripotent stem cells (iPSCs) are an ideal option as they self-renew and have been widely used recently. They can be used to produce a reasonable number of megakaryocytes combined with CRISPR-mediated genomic editing to introduce the candidate variant, in order to study *CEBPZ*, where primary patient material is difficult to source.

3.8 Key findings

- In study one, WES analysis identified 13 potential disease causing genetic variants in genes: *FCER1G, FHOD1, FMNL1, NBEAL2, DNM3, ARHGEF3, RAI1, ACVRLA1, ORAI1* which were classified as variants of uncertain significance and one likely pathogenic variant in *FLNA, and* three pathogenic in ITGA2B, ABCG8 and CD36.

- In study two, three novel candidate variants were found to be shared between four related family members (I, III, IV and VIII) in a large consanguineous family of Pakistani origin who have a history of excessive bleeding and albinism. Segregation analysis with all affected and unaffected family members found that only the splice variant within *CEBPZ*; c.1881+3A>G were segregated among the affected family members.

- Combination of WES data with platelet function testing and bleeding phenotypes are useful approaches to investigate patients with inherited bleeding disorders with unknown cause. In addition, analysing WES data using the bioinformatic pipeline workflow is a highly sensitive and valuable tool which can be mainly used for the determination of novel variants.

Chapter.4

Investigating the functional effects of a CD36 variant

4.1 Summary of background to this research

The utilisation of next-generation sequencing techniques is widely used in clinical and research fields and is readily available. Consequently, the number of identified genetic variants associated with diseases and syndromes has increased over recent years as a result of sequencing more affected individuals (Leo et al., 2015, Ng et al., 2009). However, the biggest challenge facing researchers is the investigation of candidate genetic variants in order to determine their functional effects in understanding their molecular mechanisms especially with more variants of uncertain significance that have been identified through NGS (Starita et al., 2017).

In Chapter 3, the analysis of WES data by using a bioinformatic pipelines identified a number of potential disease causing genetic variants in multiple affected individuals. Of these, some variants are classified as pathogenic and likely pathogenic using ACMG guidelines, which are most likely to be causative of the bleeding phenotypes in affected individuals. Computational prediction tools with a combination of population and disease-specific databases were used for interpretation of these variants and subsequently classifying their pathogenicity based on different criteria as stated in Chapter 3. Although these prediction tools and databases provide valuable information about the potential consequences of sequence variants, the classification of variants should be taken with caution. These *in-silico* prediction tools should not be used solely to classify variants because variants may rarely classify incorrectly. This could cause a detrimental effect to the patient's health and subsequently lead to problems from the perspective of genetic counselling. Further supporting evidence will provide more insight about the consequence of a variant which helps to confirm whether the resulting sequence is the cause of pathologic phenotype (Tchernitchko et al., 2004).

From a clinical perspective, a final variant classification report of any patient should be checked and reviewed by a highly trained and qualified clinical scientific team based on ACMG guidelines (Downes et al., 2019). As many of the variants identified through this project are predicted to be pathogenic or likely pathogenic, functional studies are required to elucidate the molecular mechanism behind these variants, as well as investigating the resulting functional consequences which eventually correlate with phenotypic presentation in related patients. It has been shown that some disorders, or even some variants in the same gene, can present different phenotypes which make diagnosis difficult. For example, MYH9-related disease is an autosomal-dominant disorder caused by mutations in *MYH9*. Patients with MYH9-related disease show variable degrees of thrombocytopenia as well as variable bleeding tendencies. Some patients develop additional clinical phenotypes such as deafness, renal problems and cataracts (Pecci et al., 2008, Pecci et al., 2014).

Combining genetic studies with functional investigation could provide a more accurate diagnosis and link a genetic variant to a specific phenotype. This could also help in the future to investigate patients who present with slightly different phenotypes for the same established disorders. Functional characterisation of a variants can be employed by using multiple functional techniques and assays in order to investigate the functional consequences of genetic variants at molecular and protein levels. This will provide insights into the effect of a variant on the gene and protein expression in human patients using *in vitro* and *in vivo* approaches, such as cell culture or animal models if applicable. The information gained from such functional studies will provide more evidence as to what extent a variant will alter gene and/or protein function subsequently causing health problems in human patients.

4.2 Aims of this chapter

This results chapter will follow on from the previous work in Chapter 3, where WES analysis has identified a stop gain variant within the *CD36* gene in two related siblings with inherited bleeding and low platelet counts (patients 5.I and 5.II). This chapter will aim to investigate the functional consequences of this genetic variant using appropriate functional assays.

4.3 CD36 glycoprotein

CD36 (also known as GPIV) is an integral membrane glycoprotein expressed at the cell surface of various cells, including macrophages, monocytes, platelets, adipocytes and microvascular endothelial cells (Park, 2014, Silverstein and Febbraio, 2009). CD36 was firstly identified in platelets which was then termed glycoprotein IV because it was the fourth band observed in the gel of platelet membrane (Clemetson et al., 1977). It belongs to the class B scavenger receptor family, which play an important role as adhesive molecules for many different ligands such as thrombospondin (TSP), oxidized phospholipids, oxidised low-density lipoprotein (oxLDL), long-chain fatty acids and *plasmodium falciparum*-infected erythrocytes (Baruch et al., 1996, Silverstein and Febbraio, 2009, Glatz et al., 2010). Additionally, it has been shown that CD36 is involved in multiple physiological and pathological processes such as thrombosis, haemostasis, atherogenesis (formation of fatty deposits in the arteries), angiogenesis the innate immune response and diabetes (Silverstein et al., 2010, Cabrera et al., 2014, Ni, 2012, Zhu et al., 2012, Silverstein and Febbraio, 2009). Despite its abundant presence, its function in platelets is still unclear (NERGIZ-UNAL et al., 2011). In the past, CD36 was considered as a receptor for collagen (Tandon et al., 1989, Diaz-Ricart et al., 1993), but other studies have shown that platelets from

CD36-deficient patients displayed a normal response to collagen (Yamamoto et al., 1992, Kehrel et al., 1998, Daniel et al., 1994).

The expression level of CD36 in monocytes, platelets and adipocytes is high based on transcriptome expression data from Proteomics Data Bank database (Samaras et al., 2020). The data in (Figure 4.1) shows the medians of protein expression in different tissues and cell lines across multiple tested samples. The median CD36 expression in adipocytes, monocytes and platelets was 7.8, 6.5 and 6.4 respectively which showed the abundancy of CD36 expression in these cells. Thrombospondin 1 (TSP-1), encoded by the *THBS1* gene, is a receptor on platelets which has been shown to bind to activated platelet surfaces during platelet stimulations which form a specific complex with fibrinogen (Leung, 1984). As the platelet membrane glycoprotein αIIbβ3 serves as receptor for fibrinogen during platelet aggregation, the binding of TSP-1 with CD36 reinforces the molecular bridge formed between fibrinogen and αIIbβ3 (Leung and Nachman, 1986, Leung, 1984). This is supported by the study which revealed that antibodies to TSP-1 inhibited platelet aggregation (Lawler, 1986). In addition, cross linking studies have shown that CD36 is in close proximity to αIIbβ3 in platelets thus strengthening this interaction between CD36, αIIbβ3 and fibrinogen (Dorahy et al., 1996). A study using immunoelectron microscopy has demonstrated that TSP-1, fibrinogen and αIIbβ3 were colocalised on activated platelets and the association of CD36 with αIIbβ3 and its binding with TSP-1 may explain the activation of platelets and subsequent aggregation (Asch et al., 1985). A further study has also shown that CD36 is associated with CD9 and integrins on human platelets which may form specialised domains for platelet adhesion and aggregation. By immunofluorescence confocal microscopy, they have shown CD36, CD9 and α6 were located on intact platelet membranes (Miao et al., 2001). Fixed human platelets were incubated with

mouse anti-CD36 and rabbit anti-α6 antibodies. The experiments showed CD36 protein and α6 integrin co-localised on platelet membranes. Similarly, platelets were incubated with rabbit anti-CD36 and mouse anti-CD9 antibodies and the data showed that CD36 and CD9 proteins were co-localised on platelet membranes also . Taken together, the co-localisation of CD36, CD9 and integrins on platelet membranes may form a protein complex domain that enhances platelet adhesion and aggregation.

In addition, it has been shown that CD36 is involved in macrophage foam cell formation which provokes the signalling cascade upon binding to oxLDL (Platt and Gordon, 2001, Silverstein and Febbraio, 2000). The binding of oxLDL with CD36 on macrophages induces activation of Lyn, Vav family guanine nucleotide exchange factors and c-Jun N-terminal kinase (JNK)-2 (Rahaman et al., 2006, Wilkinson et al., 2006, Park et al., 2009). The activation of these pathways is required for oxLDL internalisation and foam cell formation (Huang et al., 2011). Moreover, these pathways additionally activate platelets through CD36 in a ligand dependent manner which lead to platelet hyperactivity (Chen et al., 2008, Chen et al., 2011). A study conducted by Huang et al investigated the mechanism by which CD36 signals (Huang et al., 2011). The study analysed mouse peritoneal macrophage lysates by immunoprecipitation. They performed mass spectrometry analysis by using a monoclonal anti-CD36 IgA. The results showed that CD36 interacted with tetraspanin CD9 (Tspan) which was confirmed by western blot. CD36 from mouse macrophages was precipitated by an anti-CD9 monoclonal antibody and CD9 was likewise precipitated by anti-CD36 monoclonal antibody. The tetraspanin CD9 has been shown to be expressed in many different cell types including platelets, macrophages and vascular endothelial cells (Zhang et al., 2009). Thus, the localisation of CD36 with CD9 on platelets from Miao

et al study is compatible with results from Huang et al study where CD36 and CD9 co-localise on the macrophage cell membrane (Miao et al., 2001, Huang et al., 2011).

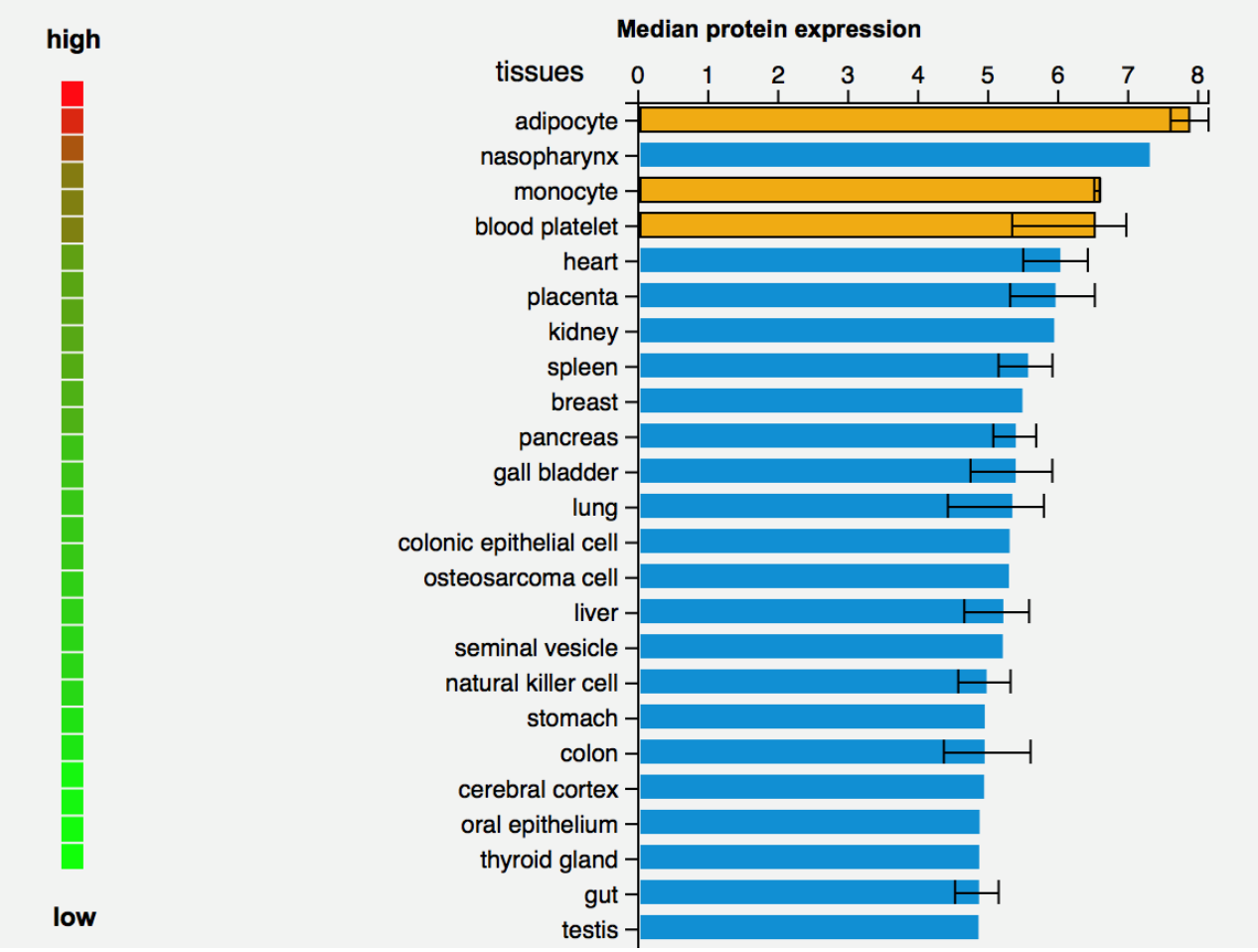


Figure 4.1. The expression level of CD36 across different tissues. Expression level data is shown as median protein expression among multiple tested samples. The median is given as 0 being low expression which is represented as a green colour on the left bar that illustrates the intensity of median value. The median value of 8 is the highest expression level which is represented as red colour on the scale bar on the left. Data obtained from (https://www.proteomicsdb.org/).

4.3.1 CD36 glycoprotein structure

The gene encoding CD36 transmembrane glycoprotein is located on chromosome *7q11.2* and is encoded by 15 exons. CD36 is comprised of 427 amino acids in length, one extracellular domain, two transmembrane fragments and two cytoplasmic fragments with C and N-terminals as seen in Figure 4.2 (Rać et al., 2007, Armesilla and Vega, 1994).The molecular mass of the functional receptor is 78 - 88kDa which differs based on the cell type and post-translational glycosylation (Oquendo et al., 1989, Greenwalt et al., 1992). It has been shown that the sequence of CD36 protein is highly conserved between human and rodent proteins (Febbraio et al., 2001). Most of the exons located in the extracellular membrane domain are highly glycosylated (exons 4 - 13 and part of exon 14) which could be the result of involvement of CD36 in multiple interactions (Armesilla and Vega, 1994). There are six centrally clustered cysteine residues located in the middle of extracellular domain including (Cys 243, Cys 272, Cys 311, Cys 313, Cys 322 and Cys 333). Theses cysteines residues form three disulfide bridges as can be seen in (Figure 4.2) (Armesilla and Vega, 1994). Recombinant and synthetic peptide studies revealed that the sequence region between amino acids 93 and 120 is the most likely site for the protein binding of the thrombospondin type I repeat (TSR) (Pearce et al., 1995, Leung et al., 1992). Interestingly, it has been shown that the binding of TSP-1 to platelets is regulated by the ectodomain phosphorylation of threonine at position 92 (Asch et al., 1993). It has been shown that the amino acid sequence extending from residues 97 - 110 interacts with malaria *plasmodium falciparum* (Baruch et al., 1999). The binding site of oxLDL within the CD36 extracellular domain remains incompletely characterised, but recent studies from Podrez's group and others have suggested that the region extending

between amino acids 157 - 171 contains the major binding site (Kar et al., 2008, Pearce et al., 1998).

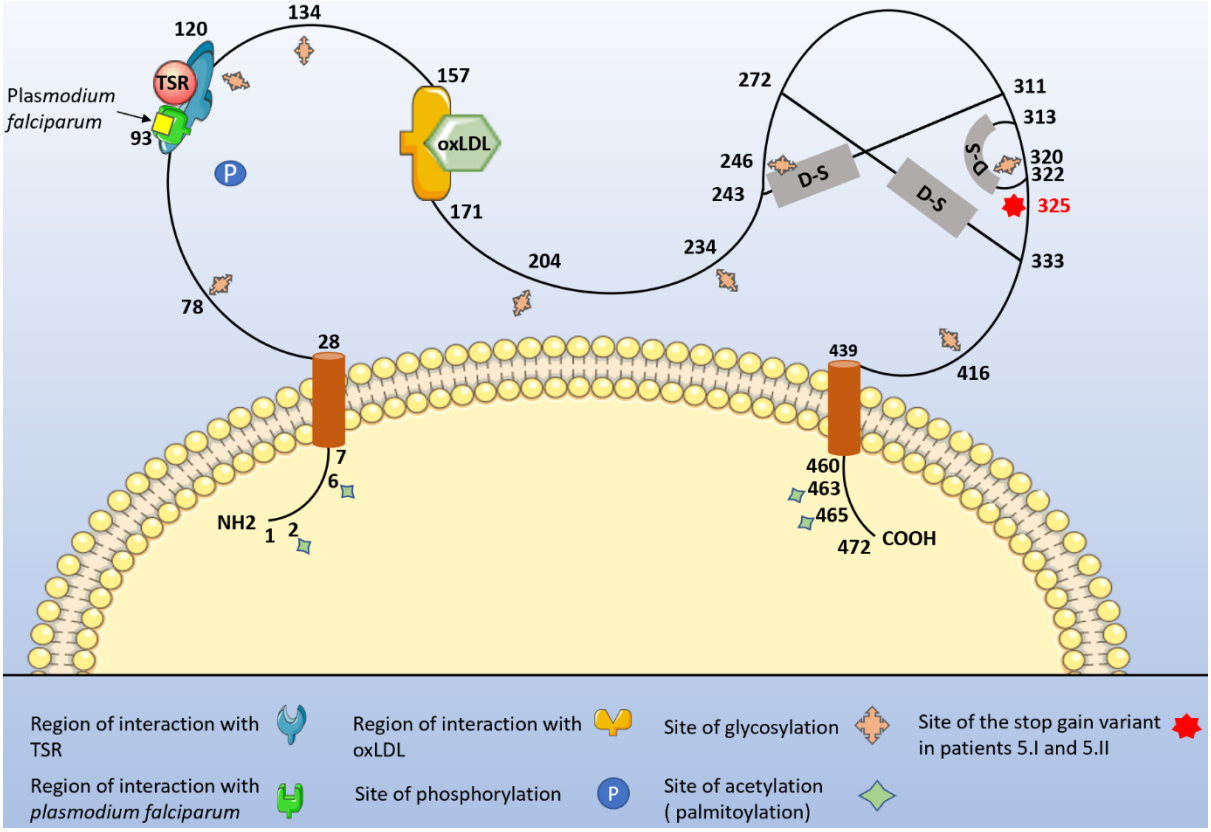


Figure 4.2. Diagram of CD36 structure. CD36 has two short cytoplasmic domains representing the C-terminal and N-terminal parts of the molecule, two transmembrane domains and two large extracellular domains. TSR: thrombospondin type I repeat, oxLDL: oxidised low-density lipoprotein, D-S: disulfide bonds. This diagram was re-designed based on (Rać et al., 2007, ALHENC-GELAS et al., 2010, Silverstein and Febbraio, 2009)

4.3.2 CD36 deficiency

CD36 deficiency can be classified into two types: type I is the most common type characterised by the absence of CD36 on both monocytes and platelets. In type II deficiency, monocytes express CD36 but platelets do not . Type II deficiency is more frequent in Asians and Afro-Americans (about 3% - 4% of the population) however it is very rare in Caucasians, only occurring in approximately 0.3% of the population (Take et al., 1993, Yamamoto et al., 1994). To date, more than 20 mutations have been identified in the coding sequence of *CD36* gene which lead to type I CD36 deficiency, while the molecular mechanisms underlying type II CD36 deficiency is still unclear (Rać et al., 2007, Take et al., 1993, Yamamoto et al., 1994).In African and Asian populations there is an exceptionally high frequency of *CD36* mutations (3-6%) which lead to CD36 deficiency and increased susceptibility to malaria (Curtis and Aster, 1996, Yamashita et al., 2007, Aitman et al., 2000). Sequence analysis of five affected African-American patients with CD36 deficiency has revealed the presence of homozygous or compound heterozygous frameshift or nonsense mutations within *CD36* and one of these patients also presented with thrombocytopenia. In addition, 80% of the mutant alleles detected in all five subjects were homozygous or compound heterozygous for a stop mutation in exon 10 (T1264G; ENST00000447544.7) which is compatible with the variant in patients 5.1 and 5.2 (c.975T>G; p. Tyr325Ter) (Aitman et al., 2000). In the Aitman paper, the genomic sequence was used as 1264 which is identical to the coding sequence 975 of the CD36.

Moreover, it has been shown that anti-CD36 antibodies develop in individuals with CD36 type I deficiency during pregnancy or blood transfusion. Consequently, this phenomenon could cause a number of immune-mediated platelet disorders such as fetal/neonatal alloimmune thrombocytopenia (FNAIT) and platelet transfusion

refractoriness (PTR) (Lin et al., 2018, Fujino et al., 2001). FNAIT is considered one of the most common causes of severe thrombocytopenia and intracranial haemorrhage (ICH) in foetus and newborns, which is estimated to occur in 1/800 - 1/1000 live births (Kamphuis et al., 2014, Turner et al., 2005). It is caused by maternal alloimmunisation that produces antibodies against paternal-derived fetal human platelet antigens (HPA) , which subsequently lead to the destruction of platelets(Kaplan, 2002, Xu and Santoso, 2018). It has been reported that more than 75% of FNAIT cases among Caucasian populations are caused by alloantibodies against human platelet antigen HPA-1a, whereas anti-HPA-4b alloantibodies inducing FNAIT are more frequent in Japanese patients (Mueller-Eckhardt et al., 1989, Peterson et al., 2013, Ohto et al., 2004). However, both alloantibodies have not yet been reported in the Chinese population. In contrast, CD36 type I deficient mothers in the Chinese population were frequently reported to develop anti-CD36 isoantibodies which cause FNAIT (Xia et al., 2014, Wu et al., 2017). It has been shown that anti-CD36 isoantibodies develop in type I CD36 deficient mothers and appear to cause FNAIT more frequently than other HPA alloantibodies in the Asian population (Xu et al., 2018). Individuals with anti-CD36 mediated FNAIT can present heterogenous clinical phenotypes including widespread petechial haemorrhages, gastrointestinal bleeding, severe thrombocytopenia, anaemia and hydrops fetalis (Kankirawatana et al., 2001, Taketani et al., 2008, Okajima et al., 2006, Curtis et al., 2002). Nowadays, intravenous immunoglobulin (IVIG) and corticosteroids are commonly used to treat individuals with high risk cases of FNAIT caused by anti-HPA-1a, however, little is known about the benefit of IVIG to treat cases with FNAIT caused by anti-CD36 isoantibodies (Bussel and Primiani, 2008, Xu et al., 2018).

4.3.3 Expression level of CD36 in platelets and monocytes

CD36 is known as a highly polymorphic gene, however, there are extremely limited studies investigating the link between CD36 genetic polymorphisms and the protein expression level (Ghosh et al., 2011). *CD36* tag SNPs were evaluated across a cohort of unrelated African-American subjects in order to investigate their effect on the metabolic syndrome (MetS). One of the studied SNPs was rs3211938 (c.975T>G; p. Tyr325Ter) which was identified in patients 5.I and 5.II and has been reported to associate with susceptibility to malaria (Aitman et al., 2000, Love-Gregory et al., 2008). Cell surface expression of CD36 on monocytes and platelets was evaluated by flow cytometry and showed that individuals homozygous for this allele lack CD36 expression on both cell types, while heterozygous individuals had reduced CD36 expression level on both monocytes and platelets. There was a positive correlation between CD36 SNPs and the level of CD36 in peripheral blood monocyte and platelets (Love-Gregory et al., 2008, Love-Gregory et al., 2011). A study conducted by Ghosh et al initially analysed CD36 platelet expression in 32 healthy donors which involved subjects from Asian population(Ghosh et al., 2011). The data showed that CD36 platelet expression ranged from 0 to 14000 molecule per platelet. They then analysed platelets from a group of 500 subjects and they found variability in platelet CD36 expression level with a mean of 7876 ±1924 molecules per platelet and there was no substantial difference in the CD36 expression level between male and female donors. Then, the functional consequence of variability of CD36 platelet expression was assessed through the investigation of effect of oxLDL on platelet aggregations. Aggregation test in response to low-dose ADP was conducted on PRP from different donors with variable CD36 platelet expression. The results showed that oxLDL induced significant increase in aggregation response within individuals with high CD36

expression level but not from CD36 null donors. Taken together, these studies reveal the level of platelet CD36 expression is highly variable among subjects from different populations which could affect the functional response of platelets to oxLDL.

4.3.3.1 *In silico* expression of CD36 in the haematological cell lineages.

The blueprint research project (https://www.blueprint-epigenome.eu/) conducted between October 2011 to September 2016 has provided valuable data about the epigenetic and transcriptomic profiles of normal and malignant haematopoietic stem cells. RNA from 6 human progenitors cells and two precursor populations representing the classical myeloid and lymphoid stages of the haematopoiesis have been sequenced. Subsequently, the RNAseq data were then aligned with the human reference transcriptome and genome (Chen et al., 2014). The Blueprint RNAseq data were used to determine the expression of CD36 through haematopoietic cell lineages. RNAseq data showed high expression of CD36 in all progenitor cell populations of the haematopoietic lineage that lead to megakaryocytes particularly within erythroblast and megakaryocyte cell populations (Figure 4.3).

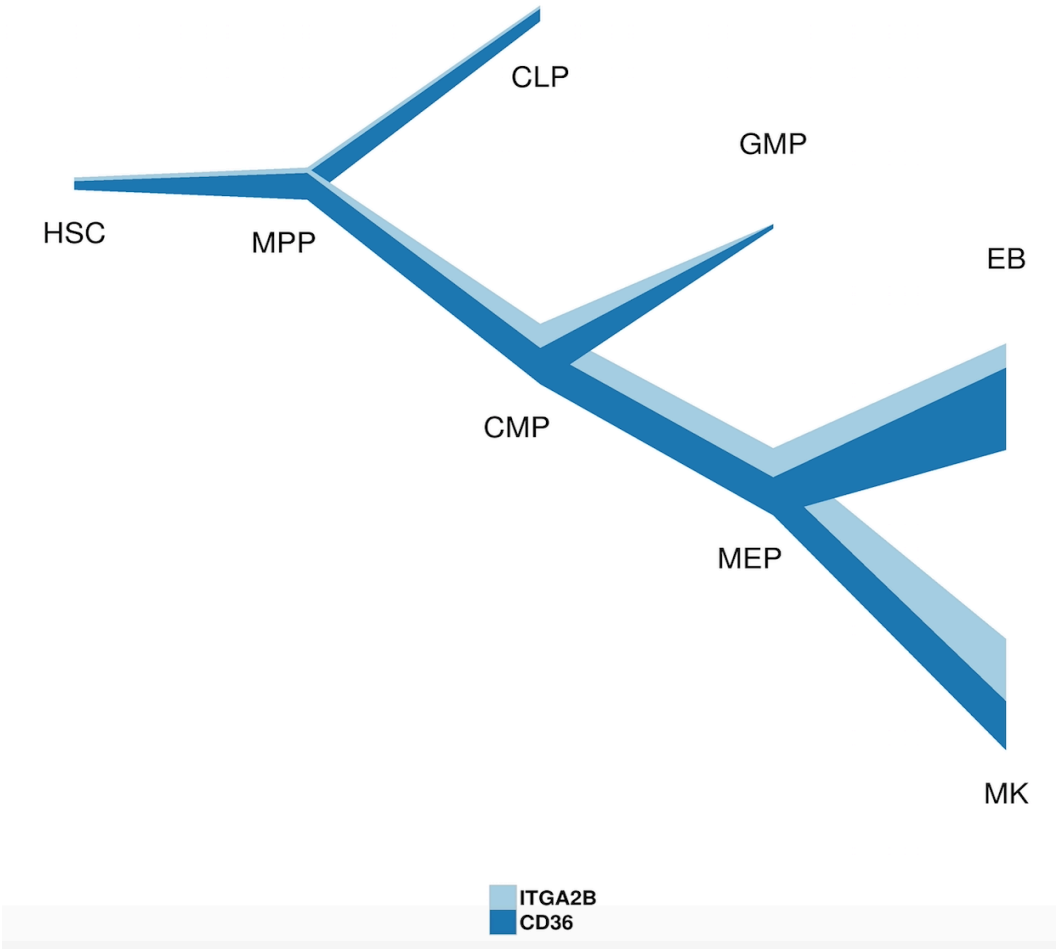


Figure 4.3. *In silico* Blueprint river plot for visualising CD36 expression in cells of the haematopoietic cell lineage. High CD36 expression in all cell lineages of the haematopoietic lineage. ITGA2B was used for comparison due to its known high expression in all the haematopoietic stem cell lineages. HSC: Haematopoietic stem cell, MPP: Mulitpotent progenitor, CLP: Common lymphoid progenitor, CMP: Common myeloid progenitor, GMP: Granulocyte-monocyte progenitor, MEP: Myelo-erythroid progenitor, EB: Erythroblast, and MK: Megakaryocyte.

4.4 Patient characteristics

Patients 5.I and 5.II were recruited to the UK-GAPP study as mentioned in Section 2.3. (pedigree in Figure 4.4). At the time of recruitment, patients were fully evaluated and their clinical history showed patients had experienced bleeding episodes. Whole blood cells counts were taken which revealed the presence of low platelet counts and high MPV and IPF values as seen in Table 4.1. Previous results from the analysis of WES data of the patients 5.I and 5.II identified a heterozygous stop gain variant in exon 10 of the *CD36* gene (c.975T>G; p. Tyr325Ter) which was reported in dbSNP as rs3211938. This variant is predicted to encode a shorter CD36 protein which lacks the C-terminal transmembrane domain. Therefore, this variant may disrupt the normal function of CD36 protein or adversely affect the normal gene product such as in a dominant negative manner. The tyrosine residue at position 325 which is changed to a stop codon is highly conserved across multiple species as seen in Figure 4.5. This variant was classified as pathogenic based on the ACMG guidelines (Table 4.2).

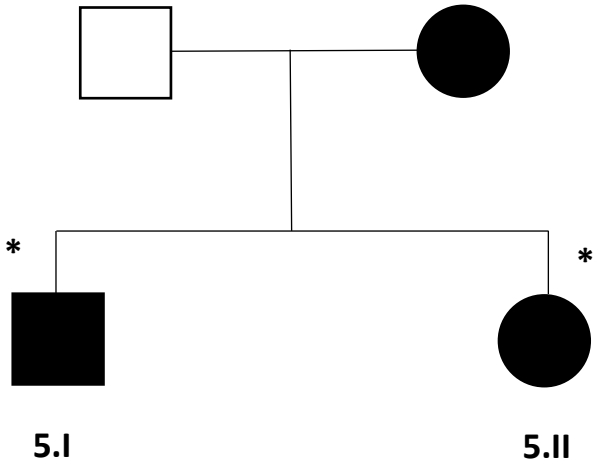


Figure 4.4. Family pedigree of the affected patients 5.I and 5.II. Affected individuals are highlighted in black. The mother was reported to have bleeding episodes and a low platelet counts. Asterisks (*) indicate patients whose whole exomes were sequenced.

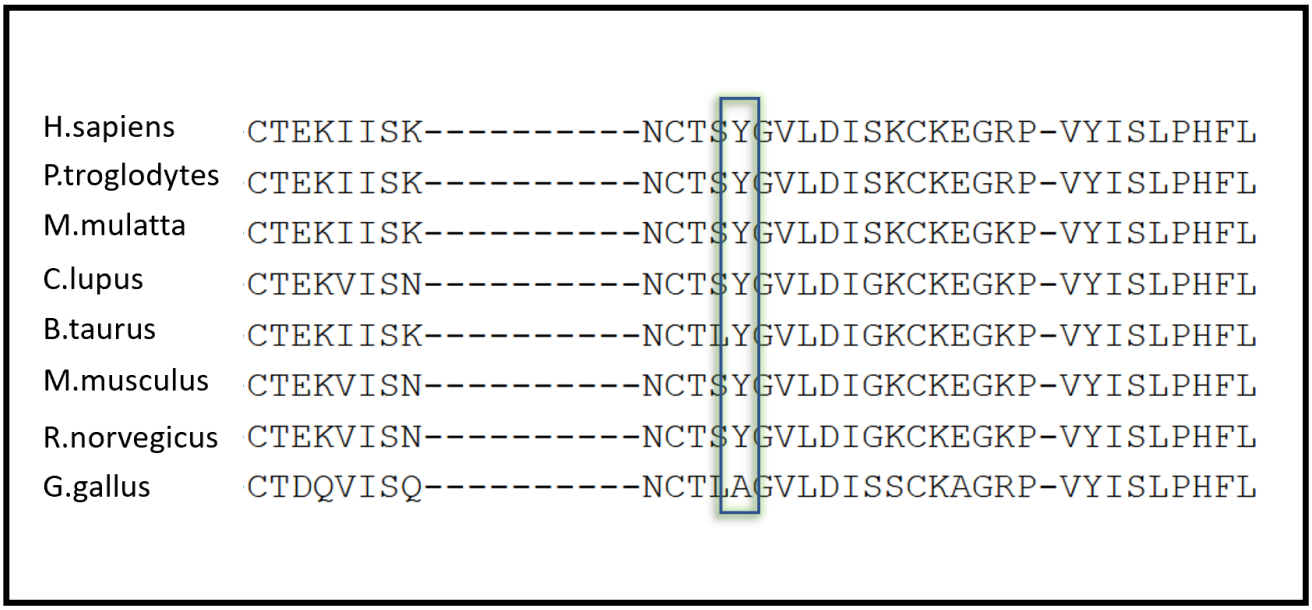


Figure 4.5. The conservation of the tyrosine 325 residue across multiple species. The nonsense variant c.975T>G; p. Tyr325Ter results in change of tyrosine residue at position 325 to stop codon which is predicted to truncate the full length of 472 amino acids. The location of tyrosine residue is shown by the green box. The name of each CD36 species is shown on the left hand side.

Table 4.1. Patient's details and haematological parameters. Haematological parameters of patients 5.I and 5.II. Patients age is recorded as an age range at the time of recruitment to the study. M: male, F: female, WBC: white blood cell, RBC: red blood cell, Mono: monocyte, IPF: immature platelet fraction, MPV: mean platelet volume. MPV values were shown in the table as large because platelets with large volume (> than the value of the normal range) are undetectable by the Sysmex analyser. The Sysmex blood cell analyser shows the haematological parameters normal ranges which were taken from (Pekelharing et al., 2010).

Parameters	5.I	5.II	Unit	Normal range
Age	12-16	12-16	—	—
Gender	M	F	—	—
WBC	3.55	5.01	$x10^9$/L	M 3.91 – 10.90 F 4.49 – 12.68
RBC	4.73	3.87	$x10^{12}$/L	M 4.44 – 5.61 F 3.92 – 5.08
Mono	0.28	0.32	$x10^9$/L	M 0.29 – 0.95 F 0.25 – 0.84
Platelet	79	87	$x10^9$/L	M 166 - 308 F 173 - 390
IPF	24.4	36.8	%	M 0.8 – 6.3 F 0.8 – 6.2
MPV	Large	Large	fL	M 9.3 – 12.1 F 9.1 – 11.9

Table 4.2. *In silico* analysis of the CD36 variant (NM_001001548). The table shows the nucleotide change, predicted protein effect, frequency in population database as well as the variant classification based on ACMG guidelines.

Gene	Nucleotide change	Predicted protein effect	Variant type	Allele frequency	ACMG parameters	Classification
CD36	c.975T>G	p. Tyr325Ter	stop gain	0.007914	PVS1,PS3,PP1 ,PP5	Pathogenic

4.5 Sanger sequencing confirmation

DNA from both patients 5.I and 5.II were analysed. Segregation studies were carried out to confirm if the CD36 nonsense variant c.975T>G; p. Tyr325Ter segregated in both patients and to remove potentially false positive results. Sanger sequencing confirmed that this variant was present in both patients in a heterozygous state and thus suggested an autosomal dominant inheritance pattern. The mother was reported to have bleeding episodes and low platelets counts. The confirmation of the variant in patients 5.I and 5.II is shown in (Figure 4.6).

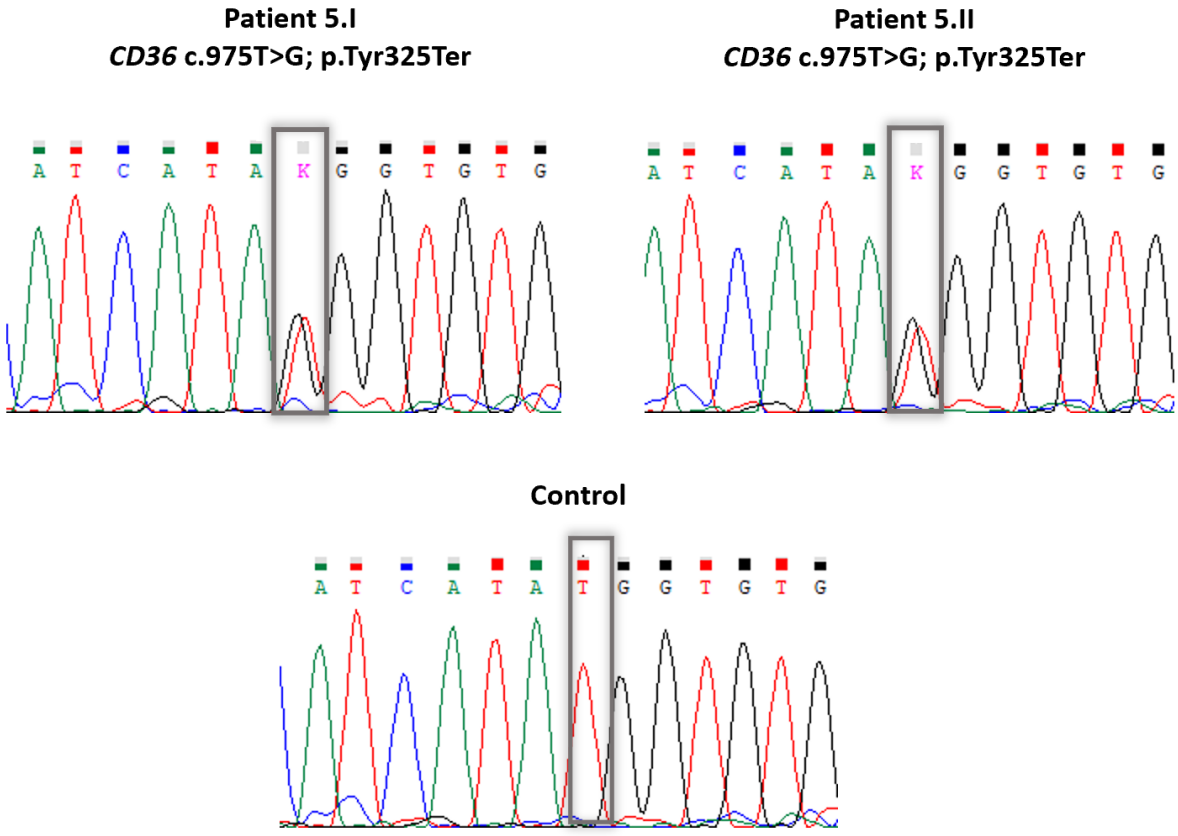


Figure 4.6. The confirmation of the CD36 variant c.975T>G; p. Tyr325Ter in patients 5.I and 5.II The grey box indicates the location of variant change in both patients compared with the normal WRT sequence in the control.

4.6 CD36 expression in protein lysates from patients

CD36 protein expression in patients 5.I and 5.II was investigated using western blot technique. Stored platelet protein lysates were available for the two patients from when recruited to the study. The protein samples from the patients were run in parallel with 6 healthy volunteer control platelet protein samples and prepared in the same way Section 2.14. The CD36 protein was detected by using anti-CD36 antibody (GTX112891, GenTex) (a table of antibodies used in this thesis can be seen in (Table 2.1). The CD36 protein blot is illustrated in (Figure 4.7). There appeared to be difference in expression of CD36 protein in patients 5.I and 5.II compared to the healthy controls. Both patients 5.I and 5.II showed two protein bands while the 6 healthy control expressed only one band. In both patients, one band was expressed at ~79 kDa which is within the normal expected size of the CD36 protein (~78-88 kDa). The other protein band was expressed at ~58kDa which is more likely to represent the truncated part of the mature CD36 protein. However, all 6 control samples expressed the same larger band at ~79 kDa which is compatible to the mature CD36 protein expected size. GAPDH was used as a loading control which showed that the expression among all the samples highly similar. (Figure 4.7).

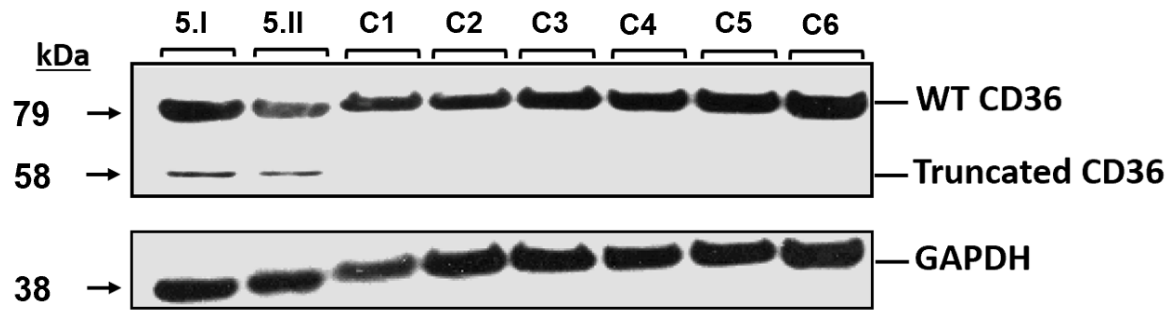


Figure 4.7. Protein expression of CD36-wild type and CD36-truncated protein from the patients platelet lysates. C: healthy control samples. SDS-PAGE immunoblot expression analysis of samples probed with rabbit polyclonal IgG antibody anti-CD36 and anti-GAPDH as loading control. The molecular weight size is indicated on the left.

4.7 CD36 mutagenesis

Site-directed mutagenesis is considered as a powerful method that can be widely used in molecular biology to introduce mutants into plasmid vectors and study the effect of amino acid changes on the structure and function of proteins (Fisher and Pei, 1997). Human cDNA for CD36 was sub-cloned into the expression vector pEF-BOS by using XbaI restriction enzyme (Thorne et al., 2000). The pEF-BOS vector is 5.8kb which includes the human polypeptide chain elongation factor 1α (EF-1α). This promotor has been shown to have very efficient stimulation of transcription *in vitro* (Uetsuki et al., 1989). The structure of pEF-BOS expression vector is shown in (Figure 4.8) (Mizushima and Nagata, 1990). The Tomlinson research group at the University of Birmingham previously performed a CD36 functional assay by over-expressing CD36 in an immortalized human T lymphocyte cell line () and subsequently used a nuclear factor of activated T-cells (NFAT) transcriptional luciferase reporter assay (personal communication)Jurkat cells were transfected with WT CD36 using different DNA concentrations. The optimal concentration for WT CD36 to be activated by the NFAT reporter was 12.5μg. Two experimental conditions were used including unstimulated condition and stimulated condition by using oxLDL.The data showed that wild-type (WT) CD36 activated the NFAT reporter and the activation was slightly increased with the stimulation of CD36 with its ligand oxLDL (Figure 4.9). Initially, two primers were designed for the upstream and downstream regions of CD36 cDNA boundaries with the pEF-BOS vector in order to confirm the insertion of WT CD36 cDNA into the vector by Sanger sequencing (Figure 4.10) (appendices, Table A). To mimic this variant within the cloned WT CD36 cDNA, Q5 site direct mutagenesis kit (New England Biolabs, USA) was applied on WT CD36 cDNA to introduce two types of mutant construct that mimic the nonsense variant in the patients (Figure 4.11) including:

1- Stop codon TAG was inserted at the equivalent of c.973_975 which subsequently deleted all amino acids following the nonsense variant (Tyr325Ter) to the stop codon of the mature protein at amino acid residue 472.
2- Substitution mutation that changes T of tyrosine residue to G (c.975T>G, p.Tyr325Ter)

Thus, the two type of constructs were created. Firstly to ensure the protein was truncated correctly at the specified site and secondly to replicate the substitution mutation which may not have truncated correctly when expressed in Jurkat T cells.

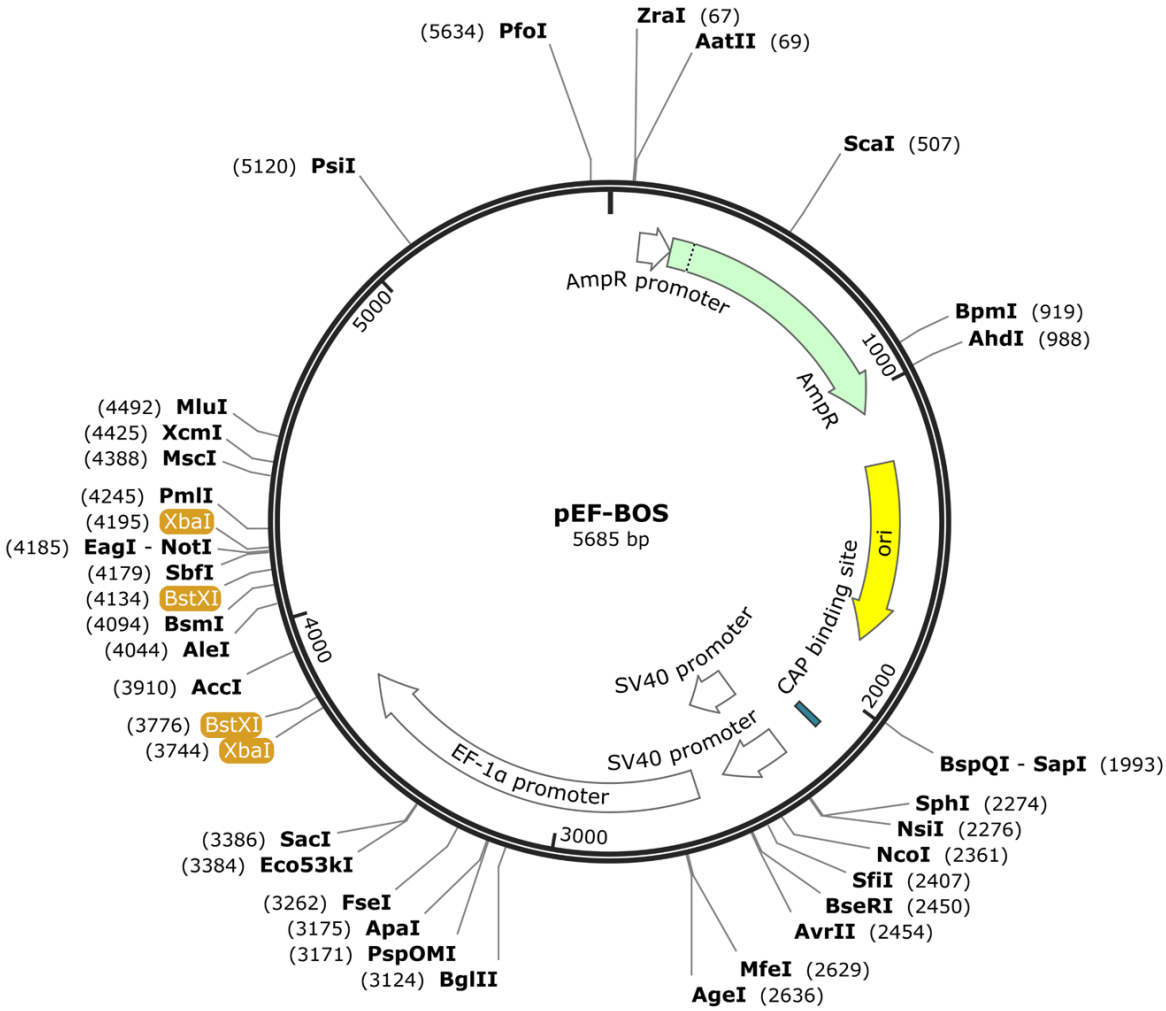


Figure 4.8. The structure of the pEF-BOS vector. The map shows the SV40 replication origin, the human EF-1α promotor, origin of replication (yellow) restriction enzyme sites and ampicillin resistance regions (green). A cDNA to be expressed can be inserted using *BstXI* adapter or *XbaI* linker which are shown with highlighted brown colour. The map was generated using SnapGene software (https://www.snapgene.com/).

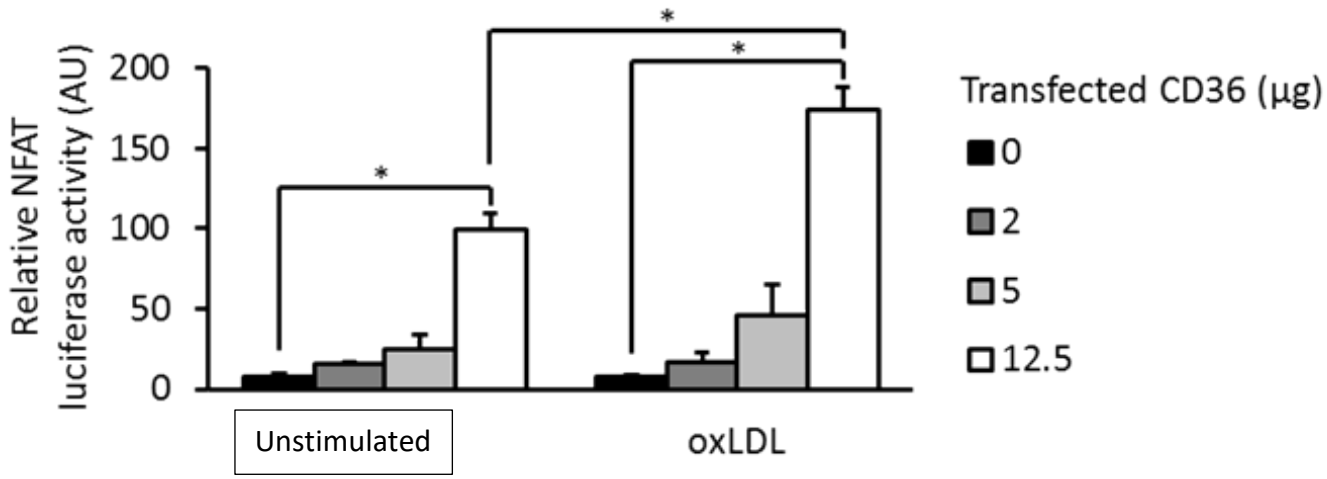


Figure 4.9. Activation of NFAT luciferase by WT CD36. Jurkat cells were transfected with WT CD36 using different DNA concentrations. WT CD36 activated the NFAT reporter with optimal concentration of 12.5µg. The WT CD36 was stimulated with oxLDL which showed slightly increased activation compared with unstimulated condition. The error bars are standard error of the mean (*P<0.05). The data were analysed by using one way ANOVA (n=3) . (This work and data has been previously done and analysed by the Tomlinson research group, manuscript under revision).

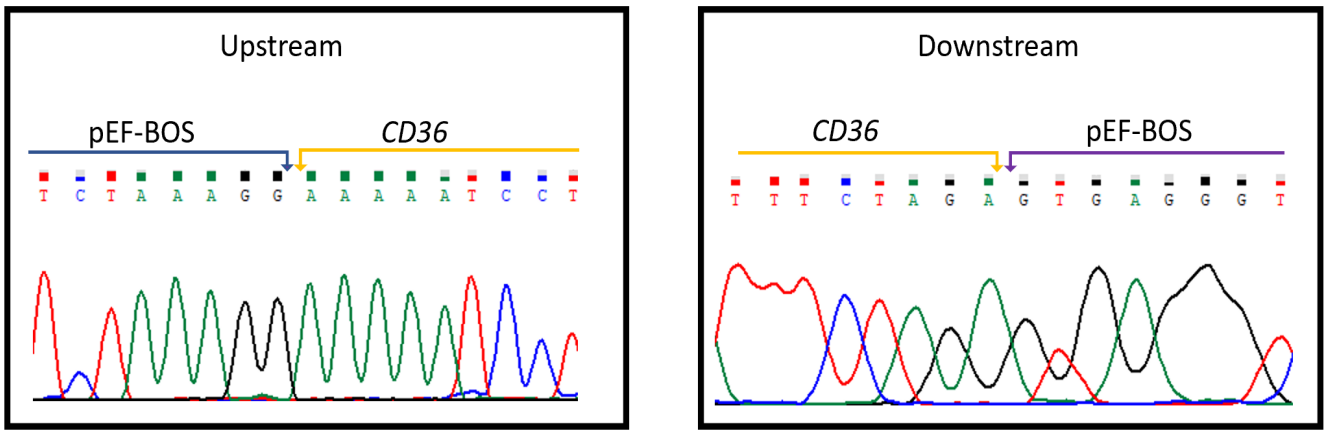


Figure 4.10. Sequencing the CD36 cDNA and the plasmid vector pEF-BOS boundaries. The yellow arrow represents the CD36 cDNA region and the purple arrow represents the pEF-BOS plasmid vector region. The figure on the left shows the sequence boundary of the plasmid vector and the upstream region of CD36 and the figure on the right shows the sequence boundary of the plasmid vector and the downstream region of CD36.

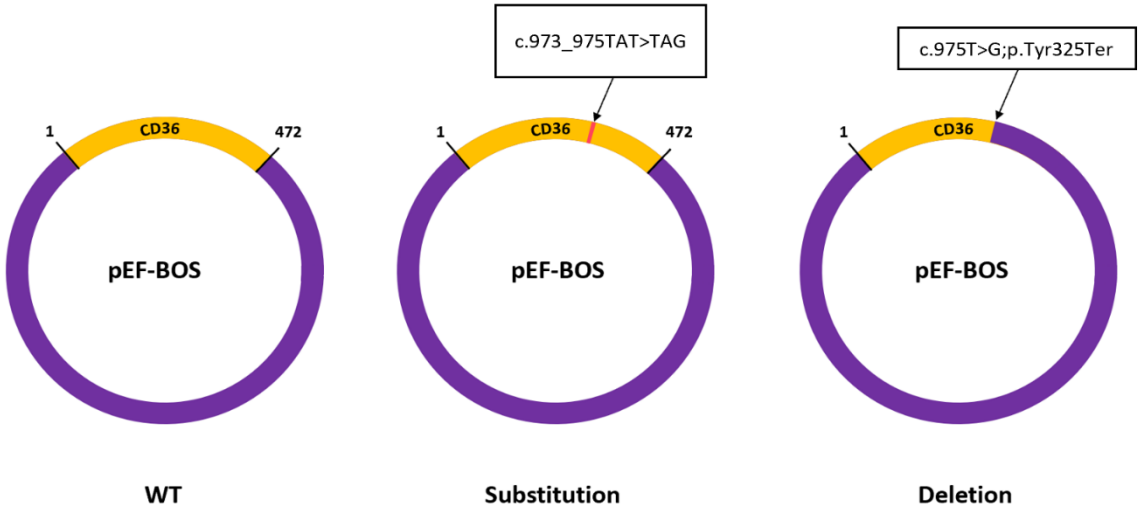


Figure 4.11. Diagram showing the type of mutagenesis carried out on human CD36 constructs. Yellow region represents the CD36 cDNA and the purple colour represents the pEF-BOS vector. The grey tag in substitution construct indicates the location of the nucleotide substitution change c.973_975 TAT>TAG at amino acid 325. The deletion construct includes CD36 cDNA which involved amino acid 1-325.

4.7.1 Transformation of mutagenesis constructs

The WT CD36 cDNA and the two versions of the mutant constructs were transformed into chemically competent *E.coli* cells following growth on LB agar at 37°C overnight as mentioned previously in Section 2.15.3. Multiple colonies were retrieved from each mutant construct in order to purify the plasmid DNAs using small scale and large scale plasmid DNA extraction kit (Qiagen kit) (see Section 2.15. 4-6).To confirm successful mutagenesis and transformation of the mutant constructs, DNA samples from WT and mutant constructs were amplified by PCR followed by Sanger sequencing. All the constructs were confirmed by Sanger sequencing as shown in (Figure 4.12).

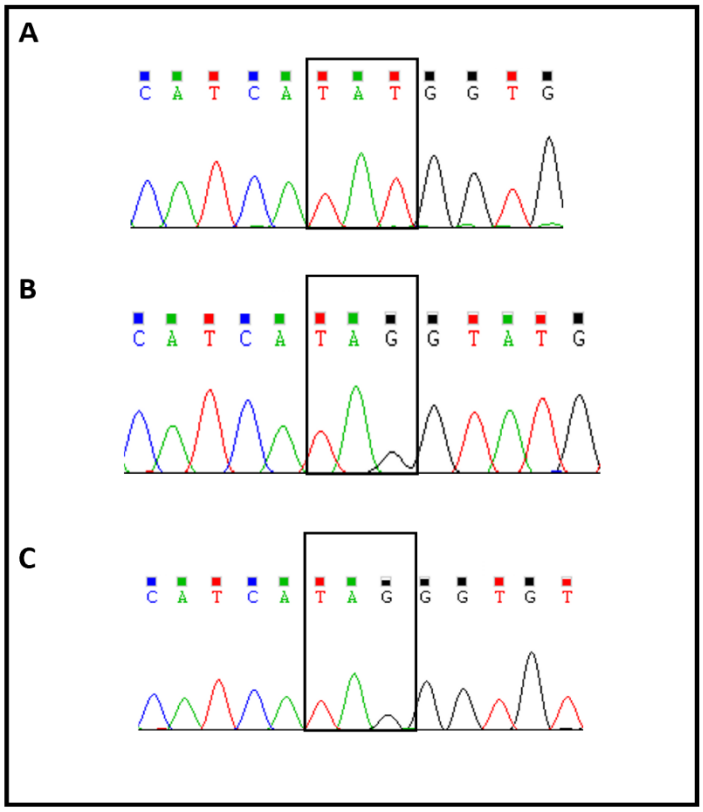


Figure 4.12. Confirmation of CD36 transformation and subsequent mutagenesis. A: represents the confirmation of WT-CD36 transformation. B: represents the confirmation of mutagenesis deletion in CD36 construct. C: represents the confirmation of mutagenesis substitution in CD36 construct.

4.8 NFAT-luciferase reporter assays

The activation of NFAT is initiated by the stimulation of antigen receptors on T cells via the calcium/ calcineurin signaling which results in the rapid translocation of NFAT transcription factors from cytoplasm to the nucleus. This activation leads to the expression of a variety of genes encoding cytokinase and membrane proteins that are involved in the immune system. NFAT-luciferase reporter is considered as sensitive readout which is widely utilized for studying the signaling pathway through the immune receptor tyrosine based activation motif (ITAM) containing B-cell and T-cell receptors (Tomlinson et al., 2007). Here, the NFAT reporter assay was used to investigate the effect of the CD36 nonsense variant and to determine whether the mutant constructs transfected into Jurkat T cells can activate the NFAT-luciferase reporter. Jurkat T cells were transfected with an empty vector as negative control, CD36 WT and CD36 mutant constructs in addition to NFAT-luciferase reporter (see Section 2.12). After transfection, the samples were incubated overnight at 37°C. Samples were then stimulated with PMA and ionomycin for 6hrs. The luciferase activity was measured as mean ± SEM for a total of three experiments . The data showed that the stimulation of WT CD36 activated the NFAT-luciferase reporter with up to 10% activity. However, the activation of NFAT-luciferase reporter in mutant constructs was low (0.2%) which was similar to the empty vector (Figure 4.13).

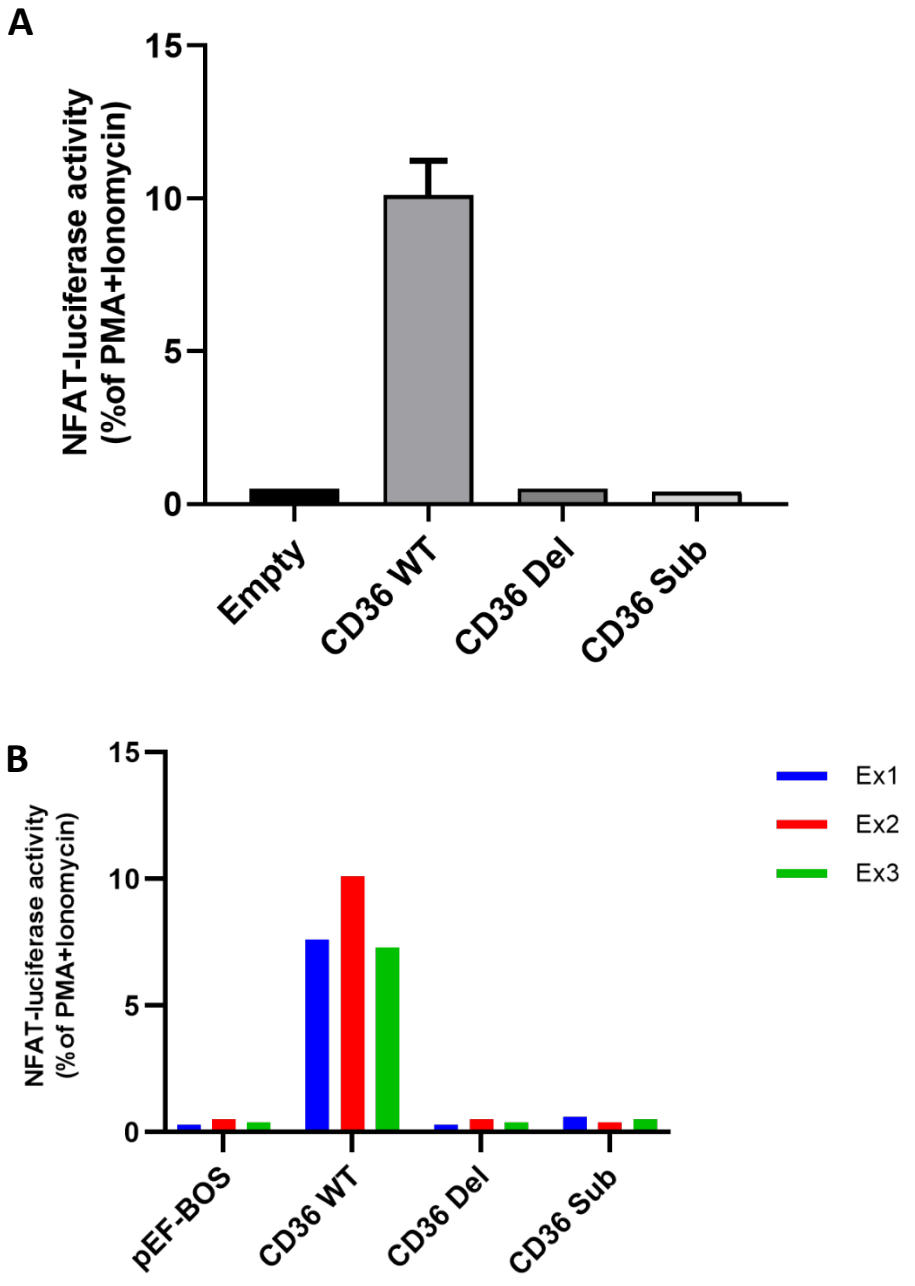


Figure 4.13. Activity of NFAT-luciferase reporter assay in CD36 constructs . The effect of mutated CD36 constructs was monitored using an NFAT-luciferase reporter assay. A: shows the NFAT-luciferase activity as (%) after normalisation of the stimulated and unstimulated conditions and the luciferase activity was measured as mean ± SEM for a total of three experiments. B: indicates the activity of NFAT-luciferase reporter assay for a total of three experiments (n=3). Only CD36 WT was activated by the NFAT-luciferase while the mutated constructs were similar to the empty vector. WT: wild type, Del: deletion, Sub: substitution.

4.9 Expression of mutant CD36 constructs in transfected cells

To characterise the effect of CD36 nonsense variant *in vitro*, an overexpression model was used. To investigate the expression of WT CD36 and mutant constructs on the cell surface, the mammalian vector, pEF-BOS, containing WT or mutant CD36 was used as outlined in Section 2.15.1. Accordingly, flow cytometry and western blot were used to assess this expression. Jurkat-T cells grown to ~ 80% confluence were transfected with WT CD36, mutant constructs and an empty pEF-BOS vector in separate wells as described in Section 2.12. 24 hours post transfection, cells were harvested and purified as previously described in Section 2.13. 250µl of transfected cells per condition were used in subsequent experiments. In flow cytometry, mouse monoclonal antibody FA6-152 (Hycult®Biotech) for human CD36 was used at concentration 10µg/ml, which recognises epitopes at amino acids 155-183. This can bind to the mutant constructs. After running the empty vector sample in flow cytometry, cells were gated and baseline applied using the empty vector (negative control) as readout for successful transfection and expression. Only WT CD36, but not the mutant constructs, were detected by flow cytometry and expressed on the cell surface as shown in (Figure 4.14,A). The percentage of expression was analysed by using one away-ANOVA for a total of three experiments as shown in Figure 4.14,C. Analysis revealed that there is a potential difference between the WT CD36 and the negative control (pEF-BOS empty vector) but the mutant constructs were similar to the empty vector.

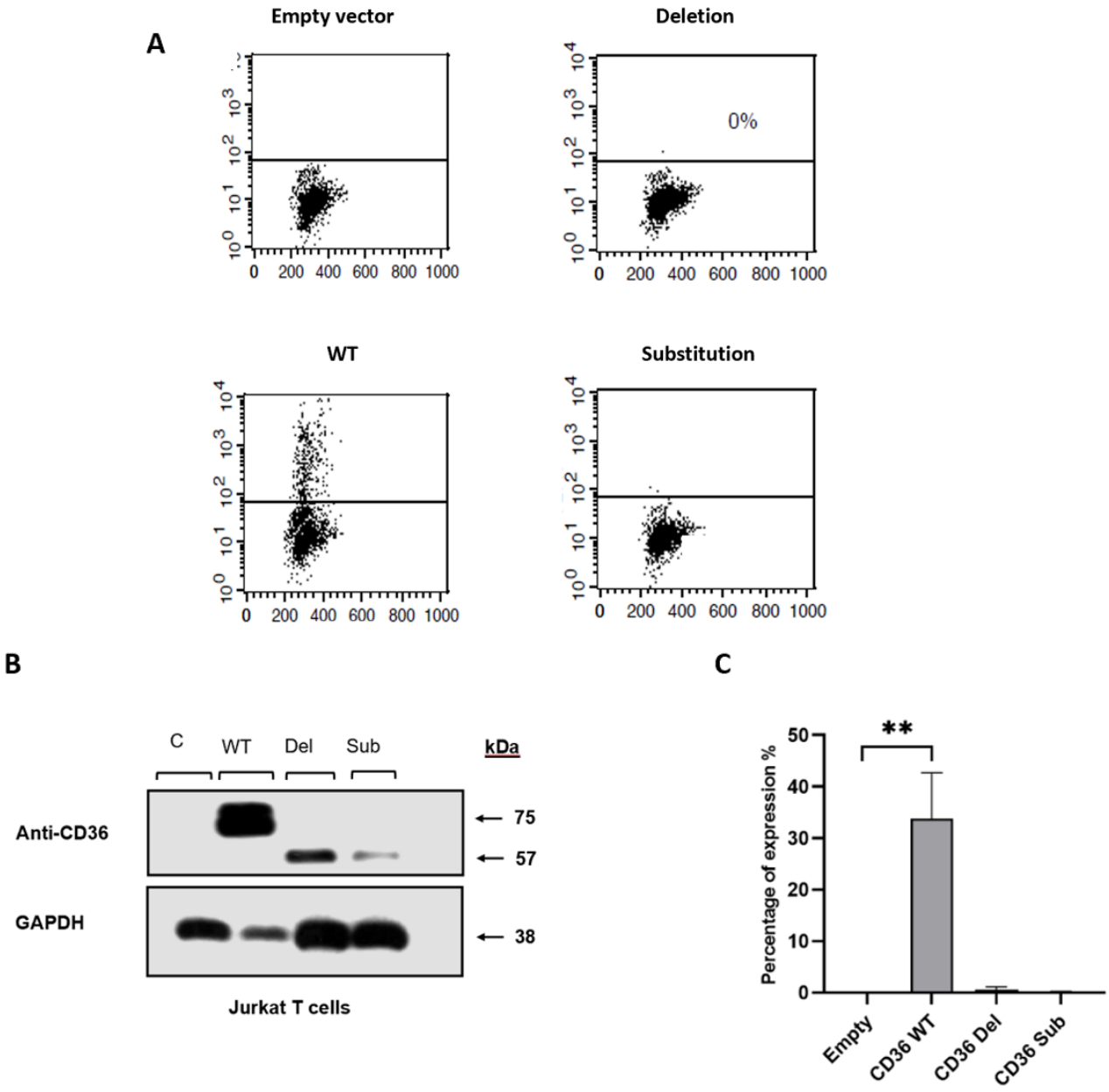


Figure 4.14. Expression of CD36 in Jurkat T cells transfected with wild type and mutant CD36 plasmids. A: indicates the expression of CD36 constructs in Jurkat T cells using FACS (n=3). B: indicates the expression of CD36 constructs in Jurkat T cells using Western blotting. C: showed the percentage expression of CD36 in Jurkat T cells as mean ± SEM of three experiments. The empty vector was used as negative control and baseline for the expression. pEF6 empty vector; WT: pEF6/CD36 wild type; Del: pEF6/CD36 deleted mutant; Sub: pEF6/CD36 substitution mutant. SDS-PAGE immunoblot expression analysis of samples probed with anti-CD36 and anti-GAPDH. Expected sizes of the samples are indicated on the right.

Furthermore, western blot analysis was used to measure CD36 protein expression of the constructs transfected in Jurkat-T cells . 20µl of cell lysates from transfected cells were run on the gel and subsequently probed with anti-CD36 at a diluted concentration of 1:4000 (D8L9T) (Cell Signaling Technology®). This antibody recognises epitopes at amino acid residue 112 and the wild type CD36 was expressed at the expected size ~ 78-88 kDa. The mutant constructs truncated at amino acid 325 are expected to result in a deletion of ~ 18 kDa which were expressed at ~ 57kDa. The western blot result showed thatonly WT CD36 was expressed at the expected size of ~78-88 kDa. The deleted mutant (truncated at amino acid residue 325, corresponding to ~ 18 kDa deletion) was expressed at expected size of ~57 kDa. The blot of expression is shown in (Figure 4.14,B). Therefore, a crystal structure analysis of the WT CD36 was performed and based on the published CD36 structure by Hsieh to determine the effect of the nonsense variant on the structure of the CD36 as shown in Figure 4.15 (Hsieh et al., 2016). This was conducted using the SWISS-MODEL tool (https://swissmodel.expasy.org/) to generate the structure which provides an annotated 3D model for proteins. This provides automated homology modelling tools for proteins based on relevant model organisms and sequence information obtained from UniProtKB. Figure 4.15 shows part of the CD36 protein which could be deleted as result of the truncation of the protein, potentially disrupting overall protein function and signalling.

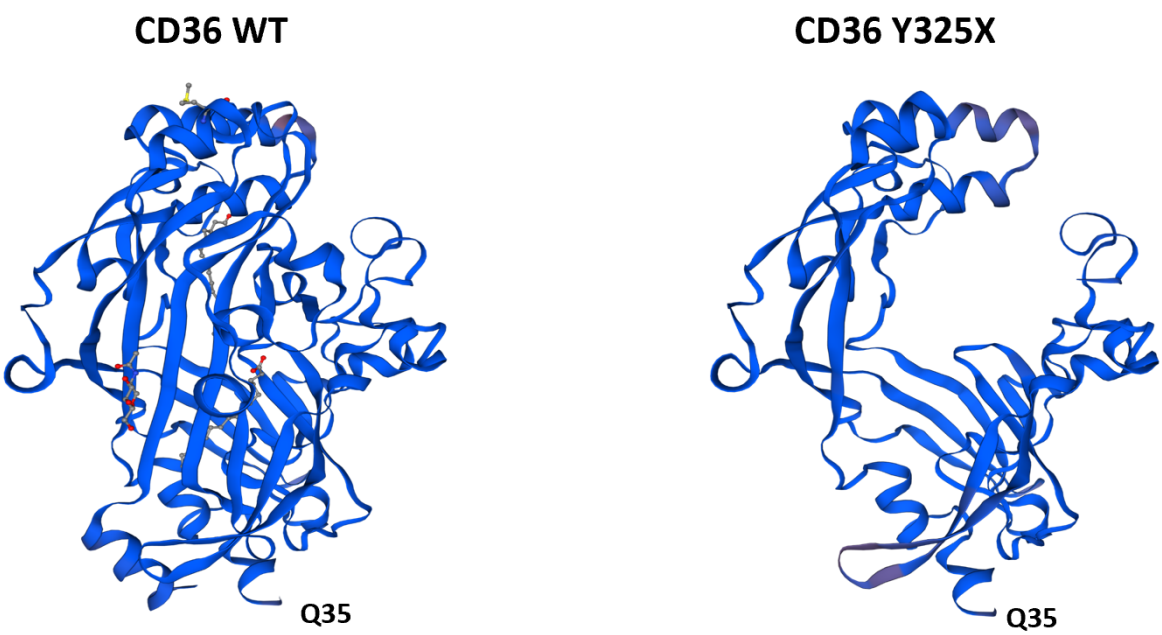


Figure 4.15. Modelled structure of CD36 ectodomain. This crystal structure was done by Connie Koo which shows the result of CD36 nonsense variant on the structure of protein. The crystal structure on the left show the WT CD36 and the one on the right shows the deleted part of CD36 as result of truncation. This was conducted using the SWISS-MODEL tool (https://swissmodel.expasy.org/) based on the modelled structure of *CD36* ectodomain study done by (Hsieh et al., 2016).

4.10 Discussion

Functional studies were carried out on patients 5.I and 5.II as previously identified in chapter 3 following the identification of a pathogenic variant (c.975T>G) in the CD36 gene, which was predicted to lead to a truncation of the mature protein with a termination codon at amino acid 325 (p.Tyr325Ter). As a result, the heterozygous *CD36* nonsense variant leading to truncation of the protein product could result in loss of the mutated allele through NMD mechanism which eliminates mRNA containing the premature translation-termination codon (PTC) (Brogna and Wen, 2009). This could result in low concentration of CD36 mRNA inside cells which subsequently could affect the normal function of the protein. Both patients presented with similar symptoms involving a history of bleeding and low platelet counts. Their mother reported to have had bleeding episodes which provided supporting evidence that the bleeding phenotype segregated in the family in an autosomal dominant pattern. OMIM608404 is an AR disorder caused by homozygous or compound heterozygous mutations in the *CD36* gene which lead to CD36 deficiency. However, the heterozygous CD36 nonsense variant could be dominantly inherited which might result from the absence of substantial activity of one allele in which a product of the other allele is not adequate for normal function and this mechanism is known as haploinsufficiency (Aitman et al., 2000). On the other hand, patients 5.I and 5.II may harbour a second variant within *CD36* that is missed by exome coverage during the exome sequencing or the presence of a CNV on the other allele which could coincide with AR inheritance. All these are possibilities that would need to be investigated further.

It has been reported that a patient with deficiency in CD36 (glycoprotein IV) presented with macrothrombocytopenia (Yufu et al., 1990). It has been shown that mutations in *CD36* lead to CD36 deficiency and subsequent reduced expression (Aitman et al.,

2000). This could have an effect on the platelet production or destruction but the mechanism needs to be studied further. In addition, the heterozygous variant in the patients predicted to truncate the full length of the protein which could act as a dominant negative on the wild type of the protein. As result, the gene product of mutated *CD36* may adversely affects the normal wild type gene product within the same cell. This could result to loss the activity and signalling pathway at the C-terminal of the CD36 protein which could subsequently affect the normal function of the protein.

Due to the COVID-19 pandemic and time constraints, some functional assays were unfinished or remain outstanding in order to fully determine the effect of the nonsense variant on the function of the protein. It has been shown previously that CD36 is an integral membrane glycoprotein expressed at the cell surface of various types of cells, including macrophages, monocytes, platelets, adipocytes and microvascular endothelial cells (Park, 2014, Silverstein and Febbraio, 2009). Despite these findings, its function in platelets is still unclear (NERGIZ-UNAL et al., 2011).

The functional investigation of this variant in platelets began by determining the expression of CD36 in patient protein lysates. The western blot analysis of CD36 expression in patients 5.I and 5.II identified two bands at sizes ~79 and ~58 kDa. This result suggests the band expressed at ~58 kDa represents the truncated part of the mature protein compared to the platelet lysates from controls which expressed only one band at expected size of mature CD36 (~78-88 kDa). This finding was confirmed by transfecting Jurkat T cells with CD36 WT and mutant constructs. The western blot result showed that only WT CD36 expressed one band at the expected size of normal CD36 ~78-88 kDa. However, the mutant constructs expressed one band at expected size of ~57 kDa which corresponds with ~18 kDa deletion. In addition, the expression of CD36 on the cell surface was assessed by flow cytometry. The expression analysis

showed that only the WT CD36 had cell surface expression in transfected Jurkat T cells while the mutant constructs were not expressed and in fact similar to the empty vector. These findings indicate that the nonsense variant may affect CD36 localisation on the cell surface membrane. As part of the external domain, one transmembrane fragment and one cytoplasmic fragment with C-terminal part of the molecule are involved in the truncated section, this may have an effect on the function of the protein arising from the cell surface and affecting downstream signalling processes in different cell types (Rać et al., 2007). Moreover, the activation of NFAT-luciferase reporter assay by the WT CD36, but not by the mutant constructs provides supportive evidence on the subsequent effect of the nonsense variant on the function the CD36 protein. This is compatible with the findings from collaborators in the Tomlinson group whereby the transfected Jurkat T cells with WT CD36 can activate the NFAT reporter and is indeed slightly increased when stimulated with oxLDL. Therefore, the modelled structure of the WT CD36 ectodomain in Figure 4.15 illustrates that the nonsense variant could alter the structure of the overall protein and the deleted part may be involved with protein signalling. It has been shown that CD36 provokes the signalling cascade which results in platelet hyperactivity as well as the formation of macrophage foam cells (Platt and Gordon, 2001, Silverstein and Febbraio, 2000). This signalling induces activation of fyn, Lyn Src family kinases, Vav family guanine nucleotide exchange factors and c-Jun N-terminal kinase (JNK)-2 (Rahaman et al., 2006, Wilkinson et al., 2006, Park et al., 2009, Silverstein et al., 2010). This signalling cascade and subsequent activation occurs within the intracytoplasmic C-terminal domain (Silverstein et al., 2010). As the truncated part of CD36 involves the C-terminal domain, this could alter the signalling cascade of the protein. Taken together, the exact mechanism of the truncated version needs to be explored further to investigate if the

mutant does in fact have an effect on the WT protein such as in a dominant negative manner which could affect the function of normal protein.

4.11 Future work

Future work should be focussed on using functional approaches to investigate the effect of this variant on the total CD36 protein function and how likely it is associated with bleeding disorders and thrombocytopenia. Re-designing the myc tag and inserting into the N-terminal domain of CD36 protein would aid this followed by with the NFAT-luciferase reporter assay to determine if the constructs could activate the NFAT-luciferase reporter. In addition, investigation into the expression of these constructs by using western blotting and flow cytometry approaches would be useful. As CD36 has been found to interact with CD9 tetraspanin on the surface of platelets, co-immunoprecipitation could be used to investigate CD36 and CD9 interactions, co-localisation on the platelet surface and subsequent signalling and cell function properties. Co-expression of WT and mutant constructs is a useful way for investigating a dominate negative effect of the variant on the WT of the protein.

4.12 Key findings

1. Identification of the CD36 nonsense variant in patients 5.I and 5. after analysing WES data by adapted bioinformatic pipeline.
2. The wild type CD36 was expressed at the expected size ~78-88 kDa and the mutant constructs truncated at amino acids 325 are expected to delete ~18 kDa which is shown at 57 kDa.
3. The NFAT-luciferase reporter assay showed that the WT CD36 can activate the NFAT-luciferase reporter up to 10% while the mutant constructs showed similar activation to the empty vector at only 0.2%.
4. The expression of CD36 by flow cytometry showed that only WT CD36, but not the mutant constructs, was expressed on the surface.

Chapter.5

A comprehensive bioinformatic analysis of patients with a suspected platelet disorder to identify both Structural and Sequence genetic variants

5.1 Summary of background to this research

The adapted bioinformatic pipeline method used in Chapter 3 was used to analyse the WES data of recruited patients. This is considered an initial analytical tool to analyse WES data. However, this method is time-consuming and not suitable for analysing large numbers of patients.

Targeted next-generation sequencing (NGS) panels were used to highlight platelet-specific genes that have been previously implicated in bleeding disorders (Lek et al., 2016). Targeted NGS panels can be used in order to interpret DNA sequence changes with particular use in a clinical diagnostic setting or in pre-screening, filtering out patients with variants in known genes, and subsequently employing WES for those who may harbour variants in novel genes (Johnson et al., 2018, Simeoni et al., 2016, Lek et al., 2016). These approaches were applied in the UK-GAPP study where patients with known mutations in haemophilia A and B, coagulation mediated genes and platelet associated genes known to cause bleeding were eliminated. However, many of these panels do not search for copy number variations (CNVs), and indeed we, and others have not found definitively causative variants in approximately 40%–50% patients despite a strongly indicative inherited component for their bleeding (Johnson et al., 2016b, Johnson et al., 2018, Bastida et al., 2018, Leinøe et al., 2017, Lentaigne et al., 2016). In this study, we address this by applying a newly established, comprehensive genetic analysis software that detects both single-nucleotide variations (SNVs) and CNVs. The Congenica platform is primarily applied for genetic diagnostics in clinical laboratories. The question arises whether it is also suitable to interrogate a large cohort of patients recruited to the UK-GAPP research study and whether this will help to perform a robust and comprehensive analysis in finding known

and novel genetic variants associated with disease, including rare CNVs not previously detected.

5.2 Aim of this chapter

Previously, the UK-GAPP study has used Whole Exome Sequencing in combination with deep platelet phenotyping to identify pathogenic genetic variants in both known and novel genes in approximately 40% of the patients. The remaining patients currently have no known cause of bleeding. To further interrogate this population, this results chapter will aim to use an IBD specific gene panel of 119 genes using the Congenica Clinical Interpretation Platform. This software was used to analyse and rapidly interpret the WES data of 126 patients recruited to the UK-GAPP study to detect both SNVs and CNVs.

5.3 Overview of the Congenica platform/software

The Congenica diagnostic platform is a newly established, comprehensive genetic analysis software, used as an automated clinical decision support platform which enables users to interrogate the human genome and identify disease-causing variants (https://www.congenica.com/). It was first established in Cambridge, UK in 2016 with a footprint in the USA and China. It was born out as result of a research collaboration between the Wellcome Trust Sanger Institute and the National Health Services (NHS) And utilised in many research projects, such as Genomics England 100,000 Genome Project and China's 100K Wellness Pioneer Project. The Congenica platform enables users to diagnose cases more efficiently, which leads to an overall improvement of the genetic and clinical diagnosis and management of the patient. Users are able to prioritize and review genetic variants, as well as assign pathogenicity, after which the

software calculates overall pathogenicity based on the American College of Medical Genetics and Genomics (ACMG) guidelines (Richards et al., 2015). It collates all essential information to make an informed and robust decision for the identification of causal genetic variants. The Congenica platform can be used to rapidly interpret and analyse NGS data and is primarily applied for genetic diagnostics in clinical laboratories for variant validation and reporting. The software improves workflow efficiency and increases diagnostic yield starting from sequence data to finally generating reports. This results from an integration of multiple comparison databases and variant prioritisation tools in addition to the interactive genome browser to visualise many different genomic tracks as shown in (Figure 5.1). These provide access to the most accurate databases such as ExAC or GnomAD for up to date frequencies of the genetic variants within populations, and Decipher and ClinVar which help users to access these databases quickly and easily. The Decipher database contains the data from 39,083 patients with rare genetic diseases which provides information about phenotypes and genotypes of these disorders which can help with genetic diagnosis (https://www.deciphergenomics.org/). In addition, users are able to check and review the quality control of the inputted patient NGS data. This includes the number of SNP alleles for each base change, number of variants in each chromosome, type of variant and zygosity. The gene coverage with different read depths can be also accessed (Figure 5.2). After analysing variants and assigning the pathogenicity scores of each variant, users can make a decision on each variant and create a report which will include all the essential information about the reported variant as shown in (Figure 5.3).

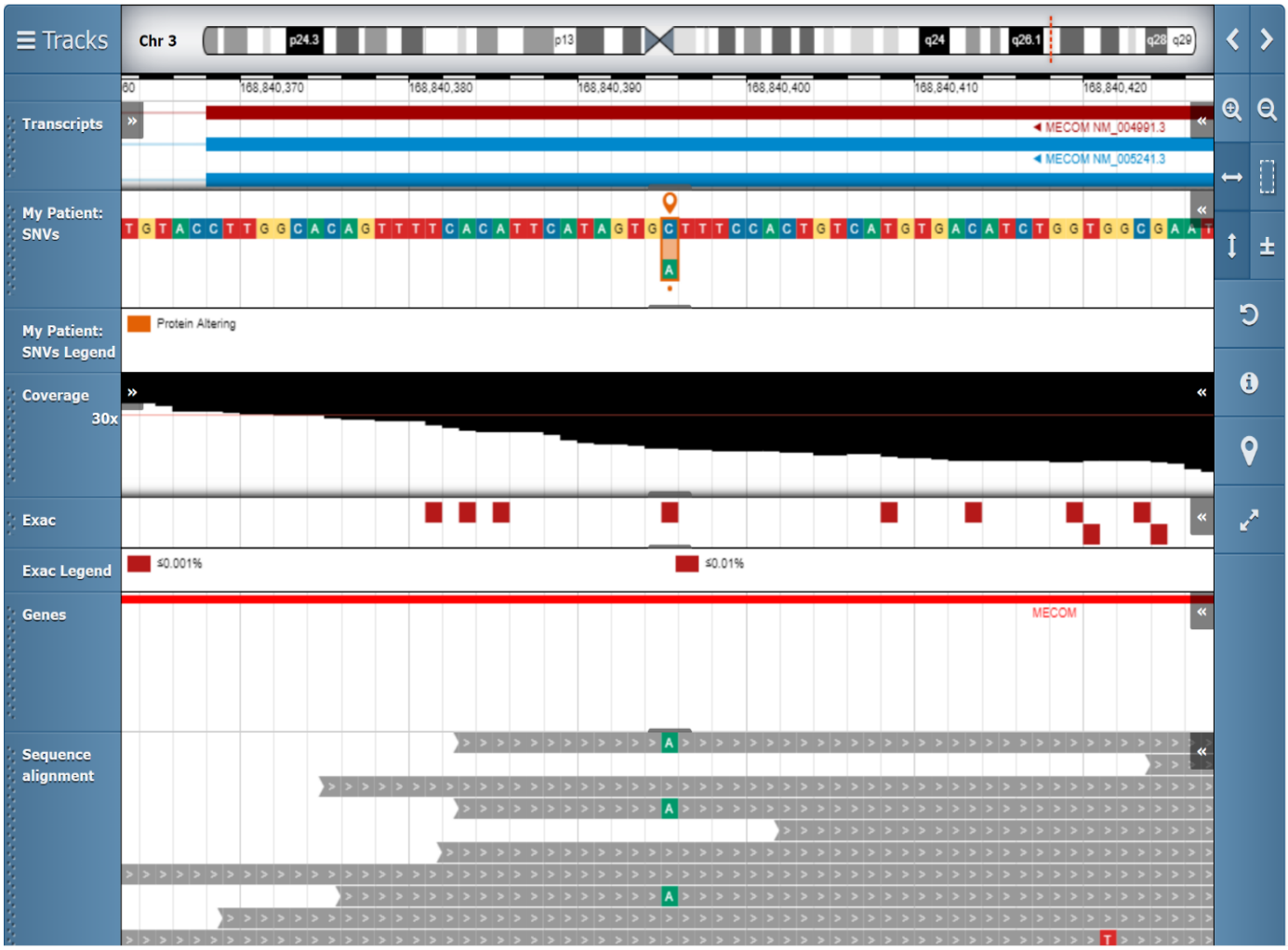
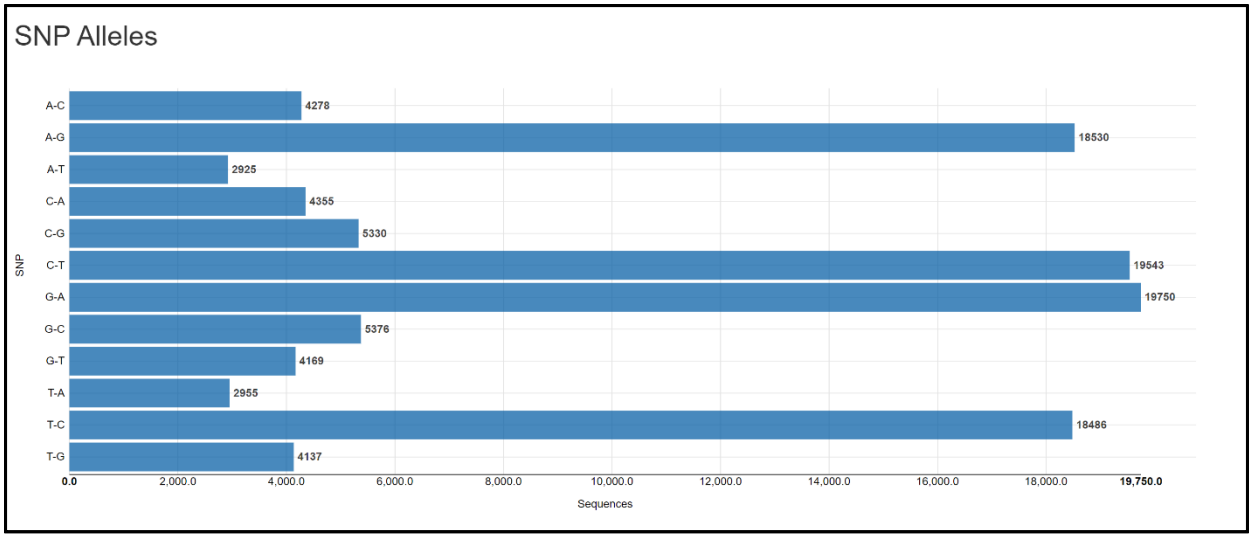
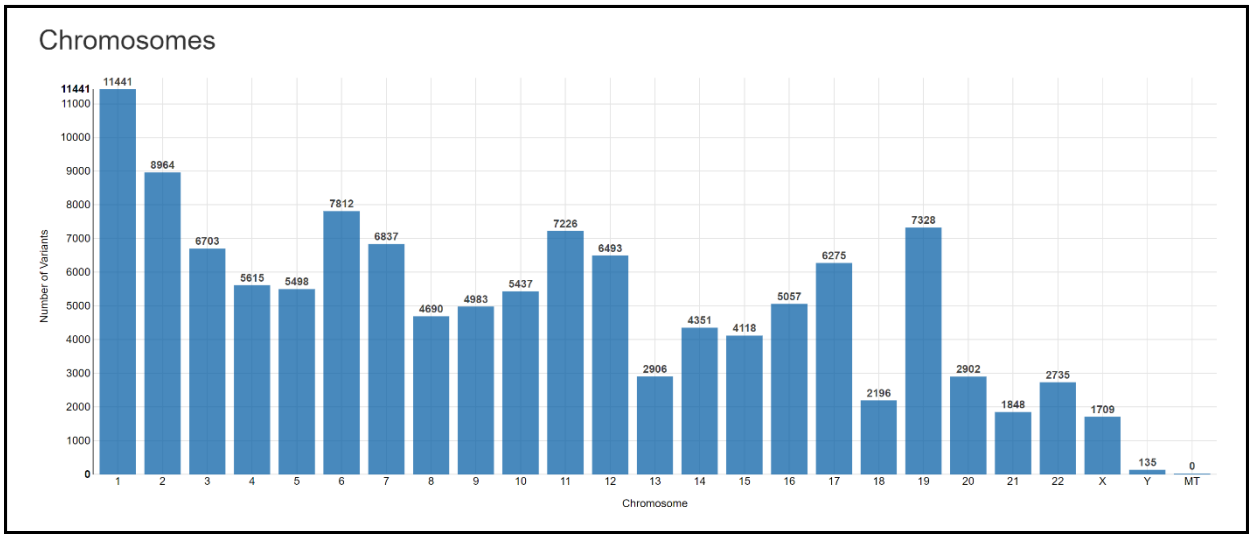


Figure 5.1. Interactive genome browser in the Congenica platform . This is used to visualise many different genomic tracks that include information about the variant in question. Users can access multiple databases that help and support with decision making and classifying variants.

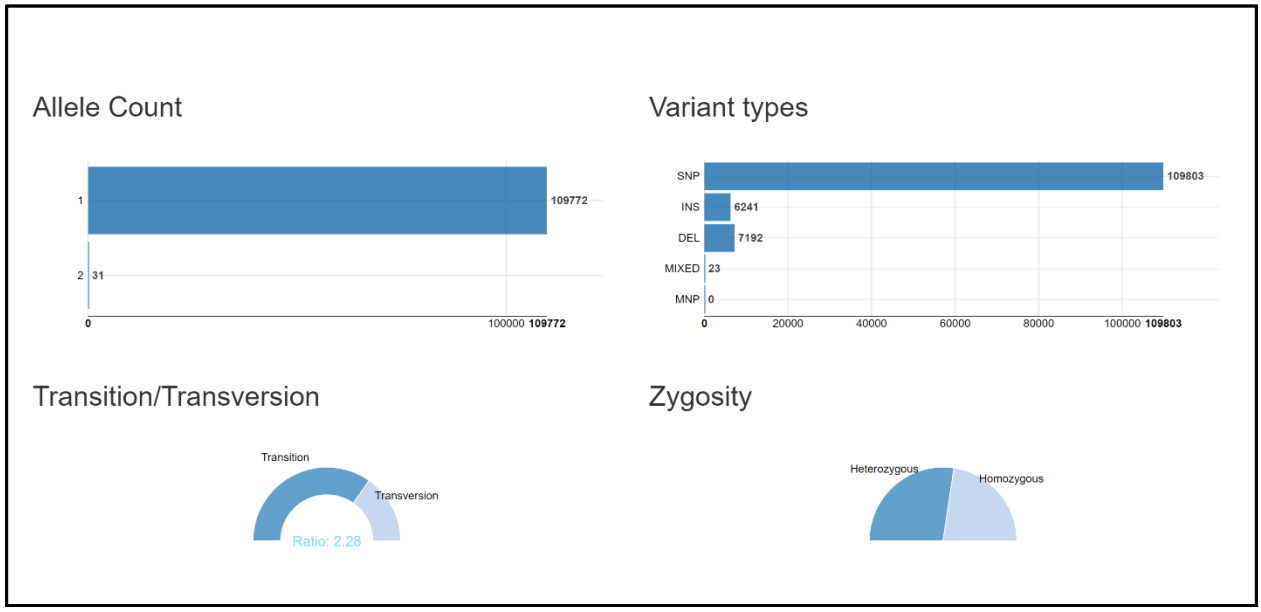
A


SNP Alleles
SNP
A-C
A-G
A-T
C-A
C-G
C-T
G-A
G-C
G-T
T-A
T-C
T-G
4278
18530
2925
4355
5330
19543
19750
5376
4169
2955
18486
4137
0.0
2,000.0
4,000.0
6,000.0
8,000.0
10,000.0
12,000.0
14,000.0
16,000.0
18,000.0
19,750.0
Sequences

B


Chromosomes
Number of Variants
11441
11000
10000
9000
8000
7000
6000
5000
4000
3000
2000
1000
0
11441
8964
6703
5615
5498
7812
6837
4690
4983
5437
7226
6493
2906
4351
4118
5057
6275
2196
7328
2902
1848
2735
1709
135
0
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
X
Y
MT
Chromosome

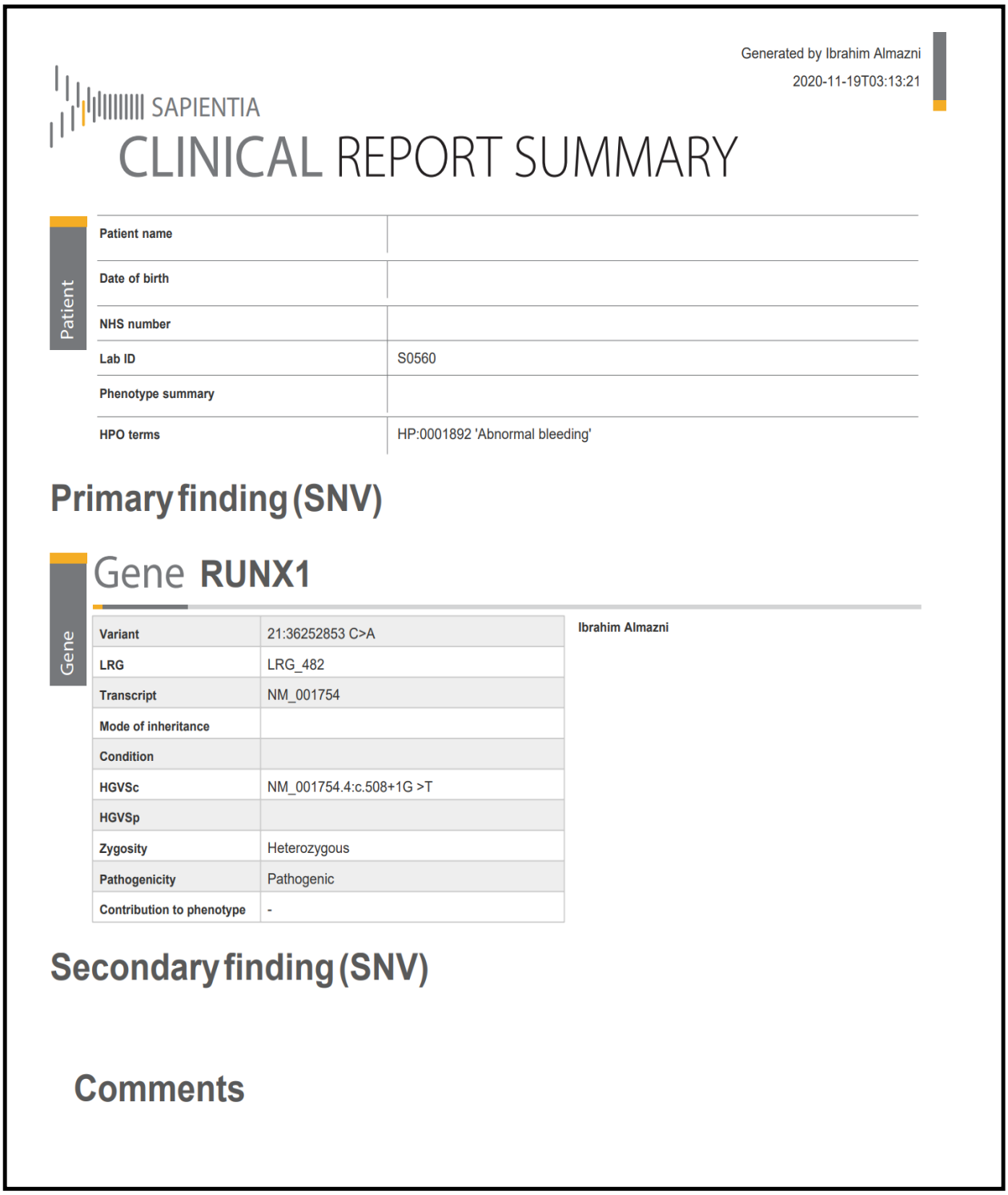
C



D

Gene	Transcript	Description	> 1x	> 15x	> 30x	> 50x
ABCG5	NM_022436	ATP binding cassette subfamily G member 5	100.00%	100.00%	100.00%	98.89%
ABCG8	NM_022437	ATP binding cassette subfamily G member 8	100.00%	100.00%	100.00%	96.73%
ACTN1	NM_001102	actinin alpha 1	100.00%	98.85%	84.45%	31.83%
ACVRL1	NM_001077401	activin A receptor like type 1	100.00%	100.00%	98.05%	90.14%
ADAMTS13	NM_139025	ADAM metallopeptidase with thrombospondin type 1 motif 13	100.00%	96.79%	95.38%	90.10%
ANKRD26	NM_001256053	ankyrin repeat domain 26	100.00%	100.00%	100.00%	98.80%
ANO6	ENST00000435642	anoctamin 6	100.00%	99.26%	95.82%	92.83%
AP3B1	NM_003664	adaptor related protein complex 3 subunit beta 1	100.00%	100.00%	100.00%	99.86%
AP3D1	NM_003938	adaptor related protein complex 3 subunit delta 1	100.00%	99.79%	88.08%	57.64%

Figure 5.2. Visualising the quality control of NGS data and gene coverage for each patients. A: shows the number of SNP alleles for each base change. B: indicates the number of sequence variants in each chromosome. C: represents the ratio of transitions or transversions and level of zygosity. D: shows the gene coverage with different read depth.



Generated by Ibrahim Almazni
2020-11-19T03:13:21

SAPIENTIA

CLINICAL REPORT SUMMARY

Patient

Patient name	
Date of birth	
NHS number	
Lab ID	S0560
Phenotype summary	
HPO terms	HP:0001892 'Abnormal bleeding'

Primary finding (SNV)

Gene

Gene RUNX1

Variant	21:36252853 C>A
LRG	LRG_482
Transcript	NM_001754
Mode of inheritance	
Condition	
HGVSc	NM_001754.4:c.508+1G >T
HGVSp	
Zygosity	Heterozygous
Pathogenicity	Pathogenic
Contribution to phenotype	-

Ibrahim Almazni

Secondary finding (SNV)

Comments

Figure 5.3.The Congenica platform can be used for reporting of variant interpretation. Users are able to create variant interpretation reports for each patient which includes all the essential information about the variant as well as the pathogenicity classification.

5.4 Patients characteristics of exomes included in this chapter

126 patients were recruited to the UK-GAPP study with variable phenotypic presentation and platelet counts. This cohort included 126 recruits in 89 families. Five different known genetic variants were identified previously by WES in 9 patients with a suspected inherited platelet disorder (IPD). These patients were just used for validating the WES analysis by the Congenica software. The remaining 117 patients were the main subjected patients for WES analysis in this chapter.

Age at recruitment was available for 106 patients, while data was unavailable in the patient reports and referrals from referring clinicians for the remaining 20 patients. The cohort included both paediatric and adult patients. An average age of 30 years was noted among patients, ranging from a baby of 39 weeks to the oldest man at 82 years. Information on sex (male/female)was available for the majority of patients which included 53.9% (n=68) female, 31.7% (n=40) male and the remaining 14.2% (n=18) were unknown (Figure 5.4, A). The clinical findings and bleeding phenotypes along with other haematological parameters results can be seen in (Table 5.1). The bleeding phenotypes of the patients were provided by the referring centres for 93 of 126 patients. The most common bleeding phenotype reported among patients were easy bruising with 68% of patients (n=64). Additionally, bleeding phenotypes were noted in a number of patients which include menorrhagia in 37% patients (n=35), epistaxis in 36% patients (n=34), bleeding after tooth extraction in 32% patients (n=30), bleeding from minor wounds in 32% patients (n=30) and bleeding after surgery in 26% patients (n=25) (Figure 5.4, B).

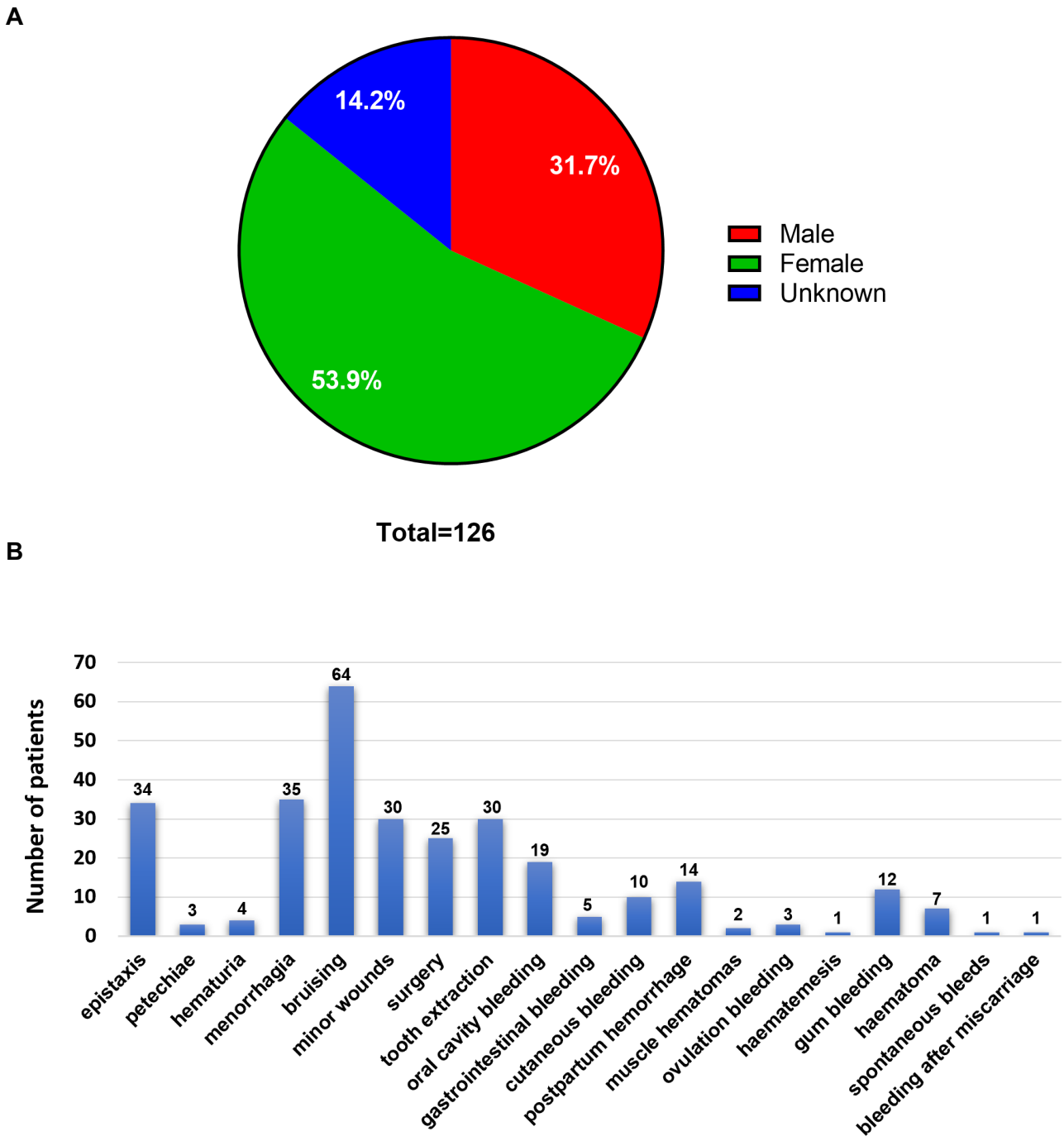


Figure 5.4. Demographic of 126 patients recruited to the UK-GAPP study. A: represents the gender of patients with percentage of male, female and unknown gender. B: indicates the number of different bleeding phenotypes noted among all patients.

5.5 Patient haematological parameters

Whole blood cell counts including platelet count, mean platelet volume (MPV) and immature platelet fraction (IPF) were assessed using the Sysmex XN-1000™ Haematology Analyser (Sysmex, UK) for most patients as mentioned in Section 2.6. However, platelet count data for patients who were recruited before the acquisition of the Sysmex machine were assessed in platelet rich plasma (PRP). Patients and their family members with a bleeding defect were recruited, irrespective of platelet count. After preliminary blood counting in whole blood, platelet counts were classified as normal (150-400x10^9/L) or thrombocytopenic (<150x10^9/L). The platelet counts in whole blood were available for 62/117 patients and platelet counts were measured in PRP on the day of recruitment for 39/117 patients, while platelet counts for the remaining 16 patients were unavailable in the referring reports (Figure 5.5, A). Of the 62 patients with platelet count in whole blood, 12 (19.6%) patients were within the normal range and 48 (77%) patients were classified as thrombocytopenic, with the majority of patients displaying both a reduced platelet count and platelet function defect. Two patients were borderline the normal range 145-155 x10^9/L (Figure 5.5, B). Of the 39 patients with platelet count in PRP, 35 (89%) patients were within normal range compared to healthy controls and 4 (10%) were classified as thrombocytopenic (Figure 5.5, C). Platelet size was examined by measuring MPV values and they were available for 88 of 117 patients. Two patients were more likely to have high MPV values because they were reported to have undetectable MPV as the analyser is on occasion unable to measure the MPV in samples with particularly large platelet size and low platelet counts (Abe et al., 2006). Of the 117 recruited patients, 19 (16%) were deemed to have a macrothrombocytopenia based on the MPV values and the referring reports (Table 5.1). In addition, IPF was undetectable or not measured in 104 of patients, however in those which were assessed, it ranged from 7.1 – 56.4%. Both of these

factors, which often coincide, highlight the variability observed in platelet characteristics and multifactorial considerations and interpretations needed in characterising platelet function and bleeding disorders.

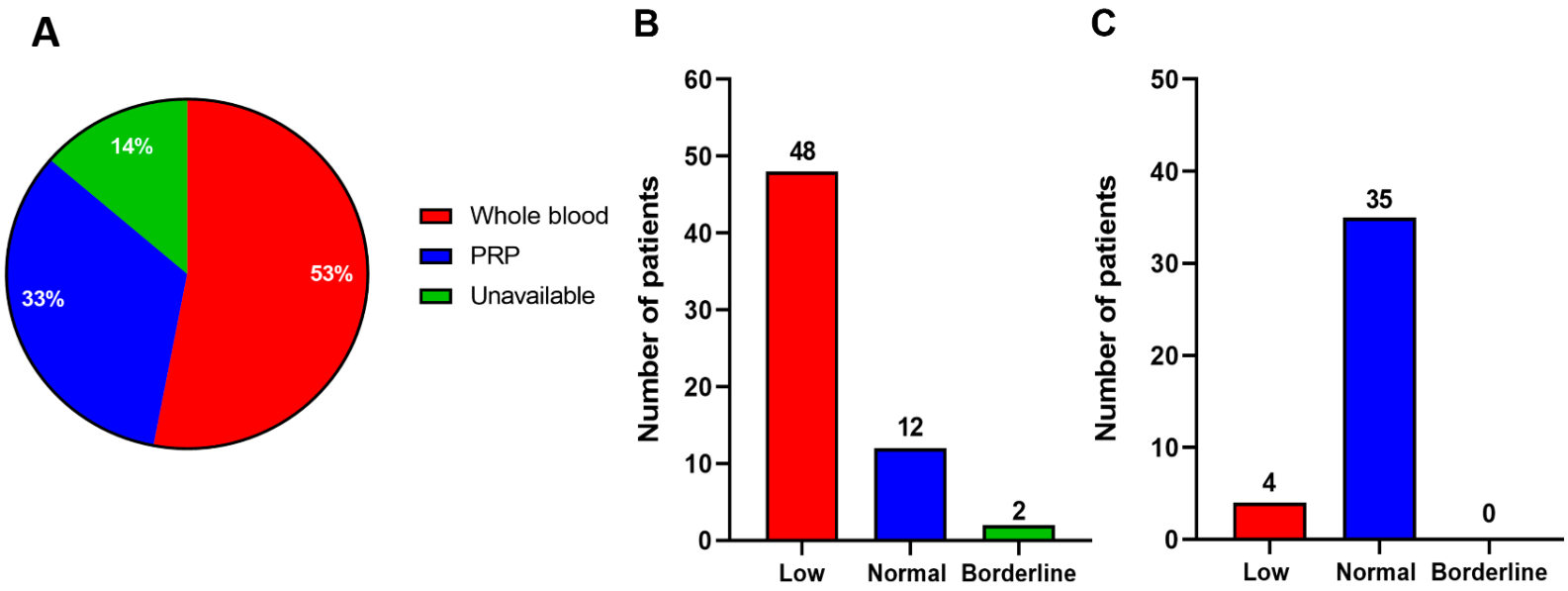


Figure 5.5. Platelet counts of 126 patients. (A) Methods employed for performing platelet counts. (B) Summary of platelet counts in whole blood samples. (C) Summary of platelet counts obtained in PRP of patient samples.

5.6 Platelet function testing

Platelet function testing was carried out on patients and healthy controls and tests were selected based on the platelet count in PRP. Patients and controls with platelet counts in PRP >1 $x10^8$/ml were assessed by light transmission aggregometry(LTA) as described previously (Dawood et al., 2012, Dawood et al., 2007). Whereas patients and controls with platelet counts in PRP <1 $x10^8$/ml were analysed by custom designed flow cytometry as described previously (Johnson et al., 2016b).PRP samples from patients and controls were assessed using agonist panels as previously described. The reason for using flow cytometry to assess platelet function is that LTA was insensitive for patients with thrombocytopenia as described in previous work (Dawood et al., 2007). However in this study, platelet function testing identified qualitative or possible qualitative defects in 68/117 patients and the presence of thrombocytopenia in 58/117 patients (Table 5.1). Platelet secretion defects were found in 26 patients, while Gi and Gq signalling pathway defects were found in 25 and one patient respectively, 12 patients displayed a cyclooxygenase (COX) defect. 13 patients showed possible qualitative defects based on platelet function testing from the referring centres. The remaining 15 patients were grouped without observable platelet defects or low platelet count (Figure 5.6). Platelet function testing was unavailable for nine patients and this is likely due to an insufficient platelet count in PRP to perform functional analysis . Overall, platelet function studies revealed the presence of a combination of qualitative defects secondary to the reduction in platelet count in 30.7% (36/117) of patients.

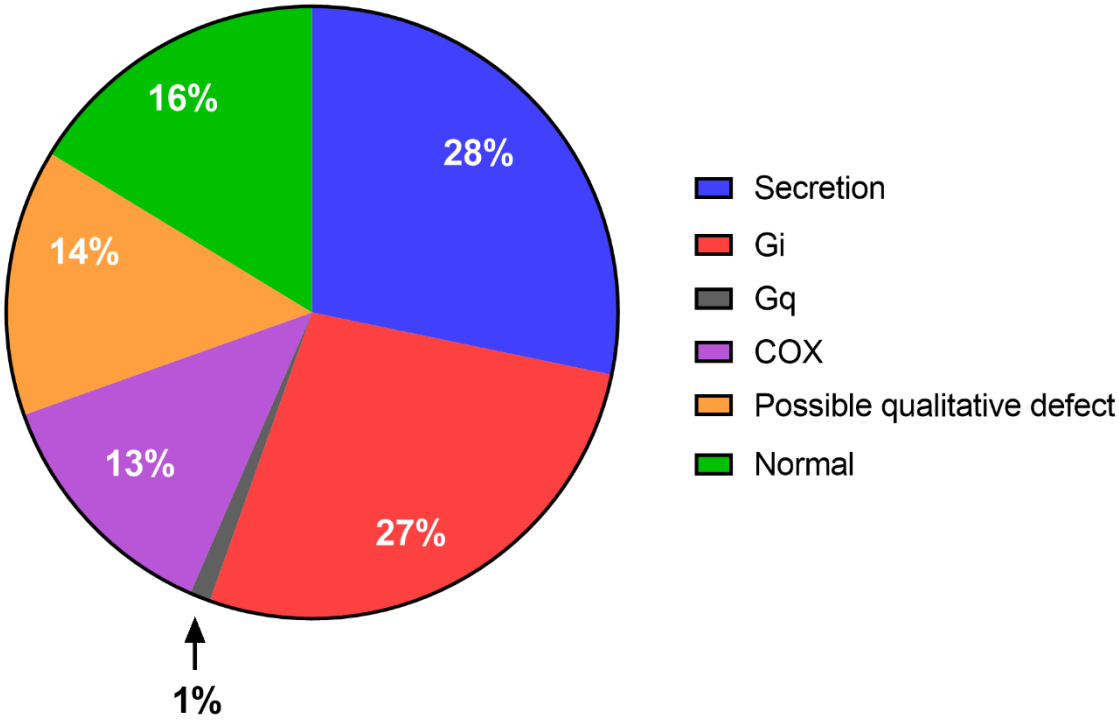


Figure 5.6. Gross classification of platelet function analysis of 117 patients recruited to the GAPP study based on lumi-aggregometry and/or flow cytometry. COX: cyclo-oxygenase deficiency, Gi: Gi receptor signalling, Gq: Gq receptor signalling.

Table 5.1. A summary of test results and the bleeding phenotype for the 126 affected recruits from 89 families. The white and grey colour were used to separate each family individually: male; F: female, W: week, Mean platelet count (normal range :150-400 x10 ^{9}/L), Thrombocytopenia is defined as a platelet count < 150 x10 ^{9}/L, (x10^{8}/ml: platelet count in PRP), Mean platelet volume (normal range : 9-12 fL), COX: cyclo-oxygenase deficiency, Gi: Gi receptor signalling, Gq: Gq receptor, NA: not available, ITP: immune thrombocytopenia purpura. 1: epistaxis; 2: petechiae; 3: hematuria; 4: menorrhagia; 5: Bruising; 6: bleeding from minor wounds; 7: bleeding after surgery; 8: bleeding after tooth extraction; 9: oral cavity bleeding;10: gastrointestinal bleeding; 11: cutaneous bleeding; 12: postpartum haemorrhage; 13: muscle hematomas; 14: ovulation bleeding; 15: haematemesis; 16: gum bleeding; 17:haematoma; 18: spontaneous bleeds; 19: bleeding after miscarriage.

Patient	Age	Gender	Platelet count (x10^{9}/L)	Mean platelet volume MPV (fL)	Platelet defect	Bleeding phenotype
1.1	22	F	NA	8.6	Gi, secretion	5, 9, 19
1.2	44	M	NA	8.9	Gi, secretion	1, 5, 6, 7
1.3	21	F	190	12.8	Gi, secretion	1, 8, 11
1.4	14	F	273	11.2	Gi, secretion	1, 4
2.1	33	M	130-140	7.1	Secretion, Gi, thrombocytopenia	1, 9, 11
2.2	6	M	70-110	7.5	Secretion, Gi, Gq, thrombocytopenia	1, 5, 9
3.1	NA	Unknown	NA	NA	Thrombocytopenia	1, 5, 9
3.2	NA	Unknown	110	9.3	COX, thrombocytopenia, PAR-1, Gi	NA
4.1	48	F	142	11.4	Secretion, Gi, thrombocytopenia	1, 4, 5, 6, 8, 9, ITP in pregnancy
4.2	19	F	157	11.8	Secretion	5, 6, 8, 9, 15

5	25	F	NA	8.5	Secretion	4, 5, 6, 10, 16
6.1	NA	Unknown	NA	NA	Secretion	NA
6.2	NA	Unknown	NA	NA	Gi	NA
7	NA	F	NA	8.4	Secretion, Gi	4, 5, 7, 8, 12
8	37	F	NA	8.6	Secretion	NA
9	82	M	447	10.1	Gi	5, 7
10.1	41	F	NA	7.7	Gi	4, 5, 6, 7, 8, 9, 13
10.2	11	F	NA	7.7	Gi	4, 5, 6, 8, 16
11	6	F	NA	9.3	Gi	5, 17
12	30	Unknown	NA	9.8	COX	1, 4, 6, 8
13	48	F	NA	NA	Gi	1, 6, 7, 12
14.1	40	F	NA	8.5	Gi	7, 17
14.2	22	F	NA	8.9	Gi	7, 8
15.1	40	F	NA	8.9	Gi	8
15.2	14	F	240	10.4	COX	1, 5, 7, 8, 16, 4
16	20	M	100	8	Thrombocytopenia, secretion	6
17	16	F	139	8	Thrombocytopenia, secretion	1, 17
18	11	Unknown	225	NA	Secretion	5, 8
19	51	F	NA	NA	Secretion	5, 6, 8, 4

20	17	M	NA	NA	Gi	8
21	42	F	NA	9.1	Gi	NA
22.1	8	F	NA	8.3	Gi	5
22.2	NA	Unknown	NA	NA	NA	5
23	28	Unknown	NA	NA	Gi	1, 5
24	66	F	130	9.7	Thrombocytopenia	1, 4, 6, 8, 7, 11
25	3	M	90	7.6	Thrombocytopenia, secretion	5
26	15	M	NA	8.1	Mild thrombocytopenia, secretion	5, 7, 17
27	16	F	119	8.8	Mild thrombocytopenia, secretion	5, 7, 8
28	5	F	50-60	8.6	Thrombocytopenia, COX	5
29.1	51	F	104-124	9.6	Mild thrombocytopenia, possible GPIV	4, 12
29.2	79	M	NA	8.6	Mild thrombocytopenia, possible GPIV&COX	5
30	16	M	25-40	8.5	Thrombocytopenia, possible qualitative defect	18
31	16	F	15-30	9.4	Thrombocytopenia, possible qualitative defect	17
32.1	35	F	80-90	10.3	Thrombocytopenia, possible COX	4, 5, 7, 8, 16
32.2	33	F	50	12	Thrombocytopenia, possible COX	4, 5, 12, 16
33	70	M	11	13.4	Severe thrombocytopenia, Possible qualitative defect	5, 7, 8
34	43	F	43	14	Thrombocytopenia	4, 5, 6, 16

35	29	F	100	10.3	Moderate thrombocytopenia, secretion	1, 4, 5, 6, 8, 9
36	30	F	50-60	10.4	Thrombocytopenia, qualitative defect	4, 5, 12
37.1	22	F	60-80	10.7	Thrombocytopenia	4, 5, 6, 12, 14, 16
37.2	29	F	60-80	10.1	Thrombocytopenia	1, 3, 4, 5, 6, 12, 16
37.3	29	F	30-40	11.4	Macro-thrombocytopenia	1, 4, 5, 6, 8, 16
38.1	16	M	100	9	Mild thrombocytopenia	5, 6
38.2	37	F	100	9.2	Mild thrombocytopenia	5, 6
38.3	38	F	80	7.9	Thrombocytopenia, possible qualitative defect	1, 4, 5, 6, 8, 10
38.4	35	F	NA	8.6	Mild thrombocytopenia	3, 4, 5, 6, 16
38.5	63	M	70	8.2	Mild thrombocytopenia, possible qualitative defect	5, 8
39	51	F	60	9.5	Macro-thrombocytopenia	NA
40	NA	Unknown	NA	NA	Secretion	NA
41	35	F	73	9.6	Thrombocytopenia, possible qualitative defect	4, 6, 7, 8, 9, 11, 12
42	30	F	20	9.7	Severe thrombocytopenia, possible qualitative defect	1, 5, 9
43	17	M	NA	NA	Thrombocytopenia, possible qualitative defect, VWF	5, 6, 9, VWF
44.1	46	F	NA	0.1	Normal	1, 4, 5, 8, 16, 17
44.2	17	M	NA	11.9	Severe thrombocytopenia	1, 5

45	67	M	100-120	9.5	Mild thrombocytopenia, possible qualitative defect	1, 7, 11, ITP
46	18	M	66	8.1	Mild thrombocytopenia, qualitative defect P-selection	NA
47	NA	Unknown	NA	NA	Thrombocytopenia	NA
48	39 W	M	NA	NA	Thrombocytopenia, RUNX1	NA
49	3	M	17	Undetectable	Severe thrombocytopenia, secretion	5, 7
50.1	4	F	NA	8.6	Normal	1, 5, 6, 7
50.2	6	F	NA	8.6	Normal	5, 6
51.1	17	F	NA	8.9	COX	4, 5, 7, 8
51.2	9	M	NA	8.9	Gi	1, 5
51.3	7	F	NA	8.8	Normal	5, 6
52.1	49	F	NA	8	Normal	NA
52.2	29	M	NA	10	Normal	NA
53.1	44	M	NA	NA	Normal	7, 8
53.2	12	M	NA	NA	Normal	1, 5, 6, 13
54.1	41	F	NA	8.6	Normal	4, 5
54.2	19	F	NA	9.4	Normal	4, 5, 8
55	38	F	NA	9.2	Normal	4, 5, 9, 12
56	7	M	123	13	Thrombocytopenia, secretion, COX	NA
57	29	F	147	12.7	Gi	4, 5, 8, 9, 12, 17

58.1	9	M	NA	8.1	Normal	2, 5
58.2	14	F	NA	7.8	Normal	1, 4, 7
59	41	F	NA	9.2	Normal	NA
60	47	M	262	8.7	Possible COX & secretion	16
61	33	F	310	10	Secretion	4, 5, 7, 9, 12,
62	37	F	179	11.1	Possible Gi & COX	1, 4, 5, 8, 10, 13
63.1	39	F	221	13.9	Lifespan & maturation, ITP	NA
63.2	54	M	107	NA	Thrombocytopenia, ITP, increased platelet clearance, decreased platelet maturation	NA
63.3	28	M	85	NA	Macro-thrombocytopenia, possible secondary defect	NA
64	29	F	53	11.7	Thrombocytopenia, possible qualitative defect	3, 5, 9, 11
65.1	28	F	130	13.8	Macro-thrombocytopenia	1, 4, 12
65.2	6	F	112	14	Macro-thrombocytopenia	1, 5
66.1	10	F	71	Undetected	Macro-thrombocytopenia	NA
66.2	NA	M	NA	NA	Macro-thrombocytopenia	NA
67	14	F	133	11.7	Thrombocytopenia, secondary defect	NA
68	41	M	92	8.9	Thrombocytopenia	5, 10
69	32	M	57	11.5	Thrombocytopenia, possible defect in the glycoprotein IIb/IIIa signalling complex,	NA

70.1	NA	Unknown	NA	NA	NA	NA
70.2	NA	Unknown	NA	NA	NA	NA
71	30	M	59	13	Thrombocytopenia, secondary defect	1, 6, 7, 9
72	76	M	120	14.6	Macro-thrombocytopenia, possible COX defect	7
73	NA	Unknown	NA	NA	NA	NA
74	NA	M	140	11.6	Thrombocytopenia	NA
75	NA	Unknown	NA	NA	NA	NA
76	NA	M	157	11.4	Thrombocytopenia, possible platelet defect	2, 5
77	78	M	91	NA	Thrombocytopenia	9, 11
78	60	F	175	13	Normal	1, 3, 4, 6, 7, 8, 9, 12
79	NA	F	35	12.2	Severe thrombocytopenia, increased platelet clearance or decreased platelet maturation	2, 4, 5, 11
80	32	F	74	14.3	Thrombocytopenia, delayed secretion defect	1, 3, 4, 6, 8, 9, 11, 12, 13, 14
81	39	F	152	12.6	Borderline thrombocytopenia, mild platelet defect	1, 4, 7, 9, 10, 11, 12,14
82	NA	Unknown	NA	NA	NA	NA
83	NA	Unknown	NA	NA	NA	NA
84.1	NA	Unknown	NA	NA	NA	NA
84.2	NA	Unknown	NA	NA	NA	NA
85.1	39	F	NA	NA	Thrombocytopenia, secretion	1, 4, 5, 7

85.2	17	M	NA	NA	Possible secretion	5, 6, 9
85.3	19	F	NA	NA	Thrombocytopenia, secretion	NA
86.1	54	M	107	NA	Thrombocytopenia, increased platelet clearance or decreased platelet maturation, possible secondary defect	ITP
86.2	28	M	85	NA	Macrothrombocytopenia, secondary defect	NA
86.3	39	F	221	13.9	Thrombocytopenia, possible secondary defect, decreased platelet lifespan/platelet maturation	ITP
87	7	F	20	10.4	Thrombocytopenia, possible qualitative defect	1, 5, 6, 17
88	31	F	100	11.1	Thrombocytopenia, Gi, PAR-1, secretion	4, 5, 6, 8, 13
89	28	M	NA	10	Normal	NA

5.7 Validation of WES analysis with known variants

Validation of the WES analysis in the GAPP study was performed using Congenica software. Five different known genetic variants were identified previously by WES in 9 patients with a suspected inherited platelet disorder (IPD) (Table 5.2). We used two trios (1 affected parent and 2 affected children) and three unrelated affected individuals, all with known or likely pathogenic variants in platelet or megakaryocyte related genes (Figure 5.7). This analysis was conducted in a blind manner to assess the reliability and robustness of the software in correctly highlighting all known genetic variants in these patients using panels of genes implicated in IBDs (Appendices, Table C). The first trio (Family A) including exomes of patients (85.1, 85.2 and 85.3) were found to share the same splicing sequence variant in *RUNX1*; c.98-1G>A. The second trio (Family B) including exomes of (86.1, 86.2 and 86.3) all shared a variant c.503G>T; p.Cys168Phe in *GFI1B*. Patient 87 displayed a homozygous missense mutation c.1246G>A; p.Gly416Arg in the *GNE* gene. In patient 88 a missense variant c.659T>A; p.Val220Asp in *SLFN14* and finally in patient 89 a stop gain mutation c.1611C>A; p.Cys537Ter located within *THBD* was identified. All known variants were found in the patients as expected and thus successfully verified Congenica software against our previously analysed WES data.

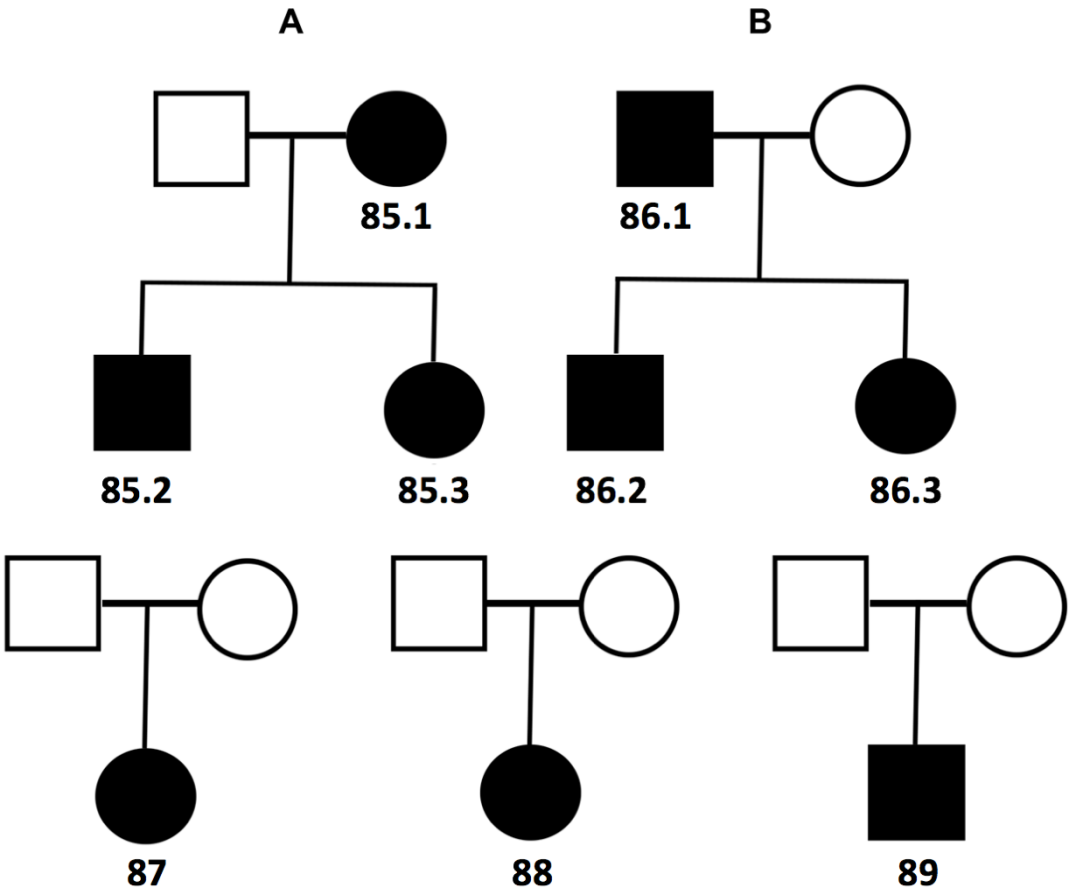


Figure 5.7. Pedigrees of initial 9 patients with a suspected IPD used for software validation. Families A & B represent two separate trios which include one affected parent and two affected children. 87, 88 and 89 represent three unrelated affected individuals.

Table 5.2. Nine patients used in trial analysis with their five known candidate variants used for validation of the Congenica software. The references indicate the publications where these variants have been observed.

Patient	Gene	Variation	Type
(Family A) 85.1	*RUNX1* (Zhang et al., 2018)	c.98-1G>A	Splice acceptor
(Family A) 85.2	*RUNX1*	c.98-1G>A	Splice acceptor
(Family A) 85.3	*RUNX1*	c.98-1G>A	Splice acceptor
(Family B) 86.1	*GFI1B* (Johnson et al., 2018)	c.503G>T; p.Cys168Phe	Missense
(Family B) 86.2	*GFI1B*	c.503G>T; p.Cys168Phe	Missense
(Family B) 86.3	*GFI1B*	c.503G>T; p.Cys168Phe	Missense
87	*GNE* (Futterer et al., 2018)	c.1246G>A; p.Gly416Arg	Missense
88	*SLFN14* (Fletcher et al., 2015)	c.659T>A; p.Val220Asp	Missense
89	*THBD* (Rabbolini et al., 2017)	c.1611C>A; p.Cys537Ter	Stop gain

5.8 Whole exome sequencing analysis to identify new SNVs and CNVs using the Congenica platform

WES data of all 117 patients was analysed using the Congenica platform based on the exome depth coverage following the phenotyping and platelet function studies workflow. The Congenica pipeline was used for exome sequence alignment and variant calling of SNVs, InDels and CNVs in order to determine plausible candidate variants. The read depth sequencing coverage is variable among patients which ranged from 10 to 1500 (Table D in Appendices). The Exome Depth integrated tool was used to determine CNVs based on read coverage difference between selected patients and reference controls as outlined in Section 2.16. Following exclusion of variants based on these criteria, an average of 2 - 6 variants (SNVs, small indels and splice site) were noted per patient. *In silico* pathogenicity prediction tools that have been integrated to the Congenica software were employed for further analysis including PolyPhen and SIFT to identify the most plausible candidate variants. A total of 135 variants in genes implicated in bleeding disorders were identified across all the 117 patients and all variants were observed in a heterozygous state (Table 5.3). In total, 22 variants were classified as pathogenic and 26 variants were likely pathogenic when considering the ACMG consensus guidelines. The remaining 87 variants were classified as of uncertain significance. The graphical illustration of this summary is shown in (Figure 5.8, A). There was a marked difference in the number of reported variants between the 4 classes of variants in patient groups: platelet count (35); platelet function (43); combined platelet count and function (59); and normal count (17) (Figure 5.8, B).

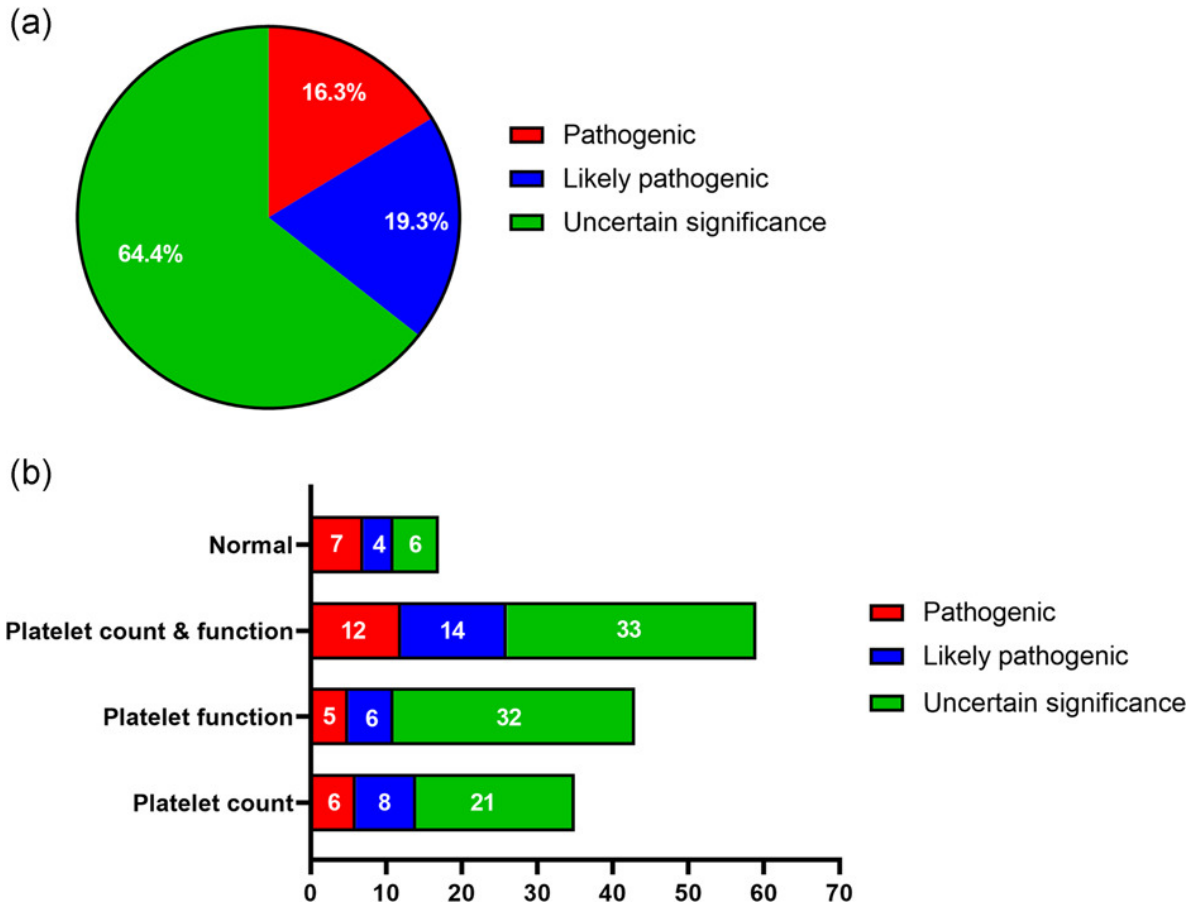


Figure 5.8. Summary of the WES analysis of 117 individuals using Congenica software. (A) Classification of 135variants observed across all 117 patients based on the pathogenicity prediction analysis which include 22 pathogenic variants, 26 likely pathogenic variants and 87 uncertain significance varaints. (B) Number of reported variants based on the pathogenicity prediction in each different class of platelet phenotype observed among the patients. All these reported variants were identified when applying the IBD gene panel.

5.8.1 Candidate variants identified in patient's cohort

Following the analysis of WES data by Congenica platform, 63% of patients (74/117) were observed to harbour variants within genes previously implicated with inherited thrombocytopenia (IT). A total of 48/135 (35.5%) variants with MAF of 0, unless otherwise stated, were identified across the 117 patients (Table 5.3). In total, 14/48 (29.1%) variants have been published previously. The number of variants found to be shared in the same affected family members were 21. Plausible candidate variants were present within the following genes (*RUNX1, SLFN14, FLI1, ETV6, HPS3, F10, P2RY12, SMAD4, TUBB1, GP1BA, GBA, CYCS, VWF, THBD, LYST, ADAMTS13, GFI1B, ITGA2B, NBEAL2, MECOM and MYH9*). Two rare variants were noted between five related affected family members including; *RUNX1:* c.611G>A; p.Arg204Gln in patients 38.1 and 38.2 and *WAS:* c.1456G>A; p.Glu486Lys in patients 38.3, 38.4 and 38.5. A stop gain variant within *ADAMTS13:* c.1315G>T; p.Glu439Ter was shared between two related affected individuals 53.1 and 53.2. Two related affected individuals with macrothrombocytopenia shared a variant within the newly discovered gene (involved in platelet disorders) *MECOM:* c.951G>T; p.Lys317Asn (Germeshausen et al., 2018). A novel stop gained variant within *ETV6:* c.1288C>T; p.Arg430Ter was noted in patient 7, which was subsequently classified as pathogenic. Patient 26 presented with a novel pathogenic splice donor variant c.904+1_904+2insGCCTGTTCACAA and novel splice region variant c.904+3A>G in *SMAD4* gene which could affect the normal splicing process.

Table 5.3. Variants identified by analysis of whole exome sequencing of 117 patients with suspected inherited bleeding disorders using the Inherited bleeding gene panel. VEP: Variant effect predictor, SIFT: Sorting intolerance from tolerance, NA: Not available, NVD: No variant detected. Variants previously reported in the literature ae indicated. The grey colour is just used to separate each family from others.

Patient	Gene(s)	VEP	MAX AF	Nucleotide change	Predicted protein effect	PolyPhen	SIFT	Pathogenicity
1.1	*NVD*							
1.2	*NVD*							
1.3	*NVD*							
1.4	*NVD*							
2.1 2.2	*RUNX1* (Stockley et al., 2013)	Splice donor	0	c.508+1G>T	NA	NA	NA	Pathogenic
	RUNX1 (Stockley et al., 2013)	Splice donor	0	c.508+1G>T	NA	NA	NA	Pathogenic
3.1 3.2	*SLFN14* (Johnson et al., 2016b)	Missense	0	c.659T>A	p.Val220Asp	Possibly_damaging	Deleterious	Pathogenic
	SLFN14 (Johnson et al., 2016b)	Missense	0	c.659T>A	p.Val220Asp	Possibly_damaging	Deleterious	Pathogenic

4.1 4.2	*FLI1* (Johnson et al., 2016b) *FLI1* (Johnson et al., 2016b)	Frameshift Frameshift	0 0	c.992_995del c.992_995del	p.Asn331ThrfsTer4 p.Asn331ThrfsTer4	NA NA	NA NA	Pathogenic Pathogenic
5	*FGA*	Missense	0.00951	c.1366A>G	p.Thr456Ala	Possibly_damaging	Deleterious	Uncertain significance
6.1 6.2	*NVD* *MPIG6B* *VPS33B*	Missense Missense	0 < 0.0001	c.132G>C c.434T>C	p.Trp44Cys p.Leu145Ser	Probably_damaging Probably_damaging	Deleterious Deleterious	Uncertain significance Uncertain significance
7	*ETV6*	Stop gained	0	c.1288C>T	p.Arg430Ter	NA	NA	Pathogenic
8	*VWF* *ANKRD26*	Frameshift Missense	< 0.0001 0.00324	c.2516del c.3004G>A	p.Gly839GlufsTer4 p.Glu1002Lys	NA Possibly_damaging	NA Deleterious	Pathogenic Uncertain significance
9	*SLC45A2*	Missense	< 0.0001	c.1471G>A	p.Gly491Arg	Probably_damaging	Deleterious	Uncertain significance
10.1 10.2	*HPS3* *HPS3*	Missense Missense	0 0	c.479G>A c.479G>A	p.Ser160Asn p.Ser160Asn	Benign Benign	Tolerated Tolerated	Uncertain significance Uncertain significance
11	*LYST* *AP3D1*	Missense Missense	0.0005 0.000116	c.8960C>G c.1246G>A	p.Pro2987Arg p.Glu416Lys	Probably_damaging Possibly_damaging	Deleterious Deleterious	Uncertain significance Uncertain significance
12	*COL5A2*	Missense	0.00264	c.4067A>G	p.Asp1356Gly	Benign	Deleterious	Uncertain significance
13	*F7*	Missense	< 0.0001	c.857C>T	p.Ala286Val	Benign	Deleterious	Uncertain significance
14.1	*F10* *NBEAL2*	Missense Missense	0 0.0066	c.1325G>A c.6866G>A	p.Gly442Asp p.Arg2289Gln	Probably_damaging Possibly_damaging	Deleterious Deleterious	Likely Pathogenic Uncertain significance

	GBA	Missense	0.00363	c.1226A>G	p.Asn409Ser	Benign	Deleterious	Uncertain significance
14.2	*F10*	Missense	0	c.1325G>A	p.Gly442Asp	Probably_damaging	Deleterious	Likely Pathogenic
	NBEAL2	Missense	0.0066	c.6866G>A	p.Arg2289Gln	Possibly_damaging	Deleterious	Uncertain significance
	GBA	Missense	0.00363	c.1226A>G	p.Asn409Ser	Benign	Deleterious	Uncertain significance
15.1	*VWF*	Missense	0.00558	c.2561G>A	p.Arg854Gln	Possibly_damaging	Deleterious	Likely Pathogenic
	NBEA	5 prime UTR	0	c.-161C>T	NA	NA	NA	Uncertain significance
15.2	*VWF*	Missense	0.00558	c.2561G>A	p.Arg854Gln	Possibly_damaging	Deleterious	Likely Pathogenic
	NBEA	5 prime UTR	0	c.-161C>T	NA	NA	NA	Uncertain significance
16	*ACVRL1*	Missense	< 0.0001	c.653G>A	p.Arg218Gln	Benign	Deleterious	Likely Pathogenic
	RUNX1 (Stockley et al., 2013)	Stop gained	0	c.317G>A	p.Trp106Ter	NA	NA	Pathogenic
	ITGB3 (Johnson et al., 2016b)	Missense	< 0.0001	c.349C>T	p.Arg117Trp	Possibly_damaging	Deleterious	Likely Pathogenic
17	*RUNX1* (Stockley et al., 2013)	Splice donor	0	c.351+1G>T	NA	NA	NA	Pathogenic
	F11	Stop gained	0.00127	c.403G>T	p.Glu135Ter	NA	NA	Pathogenic
	SERPINC1	Missense	0.00276	c.1246G>T	p.Ala416Ser	Probably_damaging	Deleterious	Uncertain significance
	F13A1	Missense	0.000192	c.1149G>T	p.Arg383Ser	Probably_damaging	Deleterious	Uncertain significance
18	*PTPN11*	Missense	< 0.0001	c.922A>G	p.Asn308Asp	Benign	Deleterious	Likely Pathogenic

19	*FGB*	Missense	0.00674	c.794C>T	p.Pro265Leu	Probably_damaging	Deleterious	Uncertain significance
20	*GP6*	Missense	0.00209	c.172C>T	p.Arg58Cys	Possibly_damaging	Deleterious	Uncertain significance
	THBD	Missense	0.00528	c.1502C>T	p.Pro501Leu	Possibly_damaging	Deleterious	Uncertain significance
21	*PLG*	Missense	0.00407	c.1469G>A	p.Arg490Gln	Probably_damaging	Deleterious	Uncertain significance
	ARPC1B	Missense	< 0.0001	c.308G>A	p.Arg103His	Benign	Deleterious	Uncertain significance
22.1 22.2	*P2RY12* (Leo et al., 2015)	Missense	0.00015	c.365G>A	p.Arg122His	Probably_damaging	Deleterious	Likely Pathogenic
	P2RY12 (Leo et al., 2015)	Missense	0.00015	c.365G>A	p.Arg122His	Probably_damaging	Deleterious	Likely Pathogenic
23	*P2RY12* (Leo et al., 2015)	Missense	0.0001	c.772C>A	p.Pro258Thr	Probably_damaging	Deleterious	Pathogenic
	MCFD2	Missense	0.00027	c.416C>T	p.Ala139Val	Benign	Tolerated	Uncertain significance
24	*VPS33B*	Missense	0.00267	c.1274G>A	p.Ser425Asn	Probably_damaging	Deleterious	Uncertain significance
	ITGB3	Missense	0.00528	c.197T>G	p.Leu66Arg	Probably_damaging	Deleterious	Uncertain significance
	LYST	Missense	0.00264	c.9017A>G	p.Lys3006Arg	Probably_damaging	Tolerated	Uncertain significance
25	*RUNX1*	Missense	0	c.403G>A	p.Gly135Ser	Probably_damaging	Deleterious	Likely Pathogenic
	VWF	Missense	0.00276	c.7988G>C	p.Arg2663Pro	Benign	Tolerated	Uncertain significance
26	*RUNX1*	Missense	0	c.593A>T	p.Asp198Val	Possibly_damaging	Deleterious	Likely Pathogenic
	SMAD4	Splice donor	0	c.904+1_904+2insGCCTGTTCACAA	NA	NA	NA	Pathogenic
		Splice region	0		NA	NA	NA	Uncertain significance
	SMAD4	Missense	0	c.904+3A>G	p.Arg671His	Benign	Deleterious	Uncertain significance

	GGCX			c.2012G>A				
27	*TUBB1* *TPM4*	Missense Missense	0.0053 0	c.13G>A c.440C>T	p.Val5Ile p.Ala147Val	Probably_damaging Benign	Tolerated Deleterious	Uncertain significance Uncertain significance
28	*PLAT*	Missense	0.00163	c.928C>T	p.Arg310Cys	Possibly_damaging	Deleterious	Uncertain significance
29.1 29.2	*TUBB1* (Johnson et al., 2016b) *TUBB1* (Johnson et al., 2016b)	Missense Missense	0.0008 0.0008	c.721C>T c.721C>T	p.Arg241Trp p.Arg241Trp	Probably_damaging Probably_damaging	Deleterious Deleterious	Uncertain significance Uncertain significance
30	*GP1BA* *GP1BA* *GBA*	Frameshift Frameshift Stop gained	0 0.00235 0	c.1274_1275del c.1277_1313deL c.653G>A	p.Glu425AlafsTer72 p.Pro426ArgfsTer34 p.Trp218Ter	NA NA NA	NA NA NA	Likely Pathogenic Uncertain significance Pathogenic
31	*MPIG6B* *HPS6*	Splice region Inframe insertion	0.00022 0	c.621G>T c.256_264dup	NA p.Trp86_Ala88dup	NA NA	NA NA	Uncertain significance Uncertain significance
32.1 32.2	*CYCS* (Johnson et al., 2016b) *GGCX* *CYCS* (Johnson et al., 2016b) *GGCX*	Missense Missense Missense	0 0.0014 0	c.155C>T c.1217G>A c.155C>T	p.Ala52Val p.Arg406His p.Ala52Val	Benign Probably_damaging Benign	Tolerated Tolerated Tolerated	Likely Pathogenic Uncertain significance Likely Pathogenic

		Missense	0.0014	c.1217G>A	p.Arg406His	Probably_damaging	Tolerated	Uncertain significance
33	*COL5A1*	Missense	0	c.1715C>A	p.Pro572His	Probably_damaging	Deleterious	Uncertain significance
	TUBB1 (Johnson et al., 2016b)	Frameshift	0	c.1080dup	p.Leu361AlafsTer19	NA	NA	Likely Pathogenic
34	*KLKB1*	Missense	0.00132	c.772C>T	p.Leu258Phe	Benign	Deleterious	Uncertain significance
	NBEAL2	Missense	0.00407	c.2375G>A	p.Arg792Gln	Benign	Deleterious	Uncertain significance
35	*COL5A1*	Missense	0.00027	c.145C>T	p.His49Tyr	Probably_damaging	Deleterious	Uncertain significance
	TBXAS1	Missense	0.00162	c.1523A>T	p.Glu508Val	Possibly_damaging	NA	Uncertain significance
	THBD	Missense	0.001	c.407T>G	p.Leu136Trp	Probably_damaging	Deleterious	Uncertain significance
36	*RASGRP2*	Missense	0.0008	c.281C>T	p.Pro94Leu	Benign	Tolerated	Uncertain significance
	VWF	Missense	0.0024	c.6424C>T	p.Leu2142Phe	Probably_damaging	Deleterious	Uncertain significance
	VWF	Missense	0.00212	c.3365C>T	p.Thr1122Met	Possibly_damaging	Deleterious	Uncertain significance
	GP1BA (Johnson et al., 2016b)	Missense	0.00417	c.1761A>C	p.Gln587His	Unknown	Deleterious	Uncertain significance
37.1	*GP1BA* (Johnson et al., 2016b)	Missense	0	c.413G>T	p.Gly138Val	Probably_damaging	Deleterious	Likely Pathogenic
37.2	*GP1BA* (Johnson et al., 2016b)	Missense	0	c.413G>T	p.Gly138Val	Probably_damaging	Deleterious	Likely Pathogenic

37.3	*MYH9 (Johnson et al., 2016b)*	Missense	0	c.3493C>T	p.Arg1165Cys	Probably_damaging	Deleterious	Pathogenic
38.1	*RUNX1*	Missense	0	c.611G>A	p.Arg204Gln	Possibly_damaging	Deleterious	Likely Pathogenic
38.2	*RUNX1*	Missense	0	c.611G>A	p.Arg204Gln	Possibly_damaging	Deleterious	Likely Pathogenic
38.3	*WAS* (Johnson et al., 2018)	Missense	0	c.1456G>A	p.Glu486Lys	Probably_damaging	Deleterious	Uncertain significance
38.4	*WAS* (Johnson et al., 2018)	Missense	0	c.1456G>A	p.Glu486Lys	Probably_damaging	Deleterious	Uncertain significance
38.5	*WAS* (Johnson et al., 2018)	Missense	0	c.1456G>A	p.Glu486Lys	Probably_damaging	Deleterious	Uncertain significance
39	*NVD*							
40	*NVD*							
41	*ACTN1*	Missense	0	c.2662G>C	p.Gly888Arg	Probably_damaging	Deleterious	Uncertain significance
	PLAT	Missense	0.00139	c.1481G>C	p.Gly494Ala	Probably_damaging	Tolerated	Uncertain significance
42	*ABCG8*	Missense	0.00157	c.1924G>A	p.Ala642Thr	Benign	Tolerated	Uncertain significance
43	*RUNX1*	Splice acceptor	0	c.98-1G>A	NA	NA	NA	Pathogenic
		Missense	0.001	c.7390C>T	p.Arg2464Cys	Probably_damaging	Deleterious	Pathogenic

	VWF (Lester et al., 2007) *GBA* *PROS1* (ALHENC-GELAS et al., 2010)	Frameshift Missense	< 0.0001 0.000572	c.26_27del c.284G>A	p.Glu9GlyfsTer8 p.Gly95Glu	NA Possibly_damaging	NA Deleterious	Likely Pathogenic Likely Pathogenic
44.1 44.2	*RUNX1* (Stockley et al., 2013) *RUNX1* (Stockley et al., 2013)	Splice donor Splice donor	0 0	c.351+1G>T c.351+1G>T	NA NA	NA NA	NA NA	Pathogenic Pathogenic
45	*HRG*	Missense	0.00162	c.1379G>A	p.Arg460Gln	Benign	Tolerated	Likely benign
46	*FLNA* (Johnson et al., 2018)	Missense	0.0073	c.5948C>T	p.Ser1983Leu	Probably_damaging	Deleterious	Likely benign
47	*GGCX*	Missense	0.0014	c.1217G>A	p.Arg406His	Probably_damaging	Tolerated	Uncertain significance
48	*RUNX1* (Lamolda et al., 2019) *SLFN14* *SLFN14* *SLFN14* *SLFN14*	Missense Missense Missense Missense	0 0 0 0	c.586A>G c.2686T>C c.1481A>G c.859A>G	p.Thr196Ala p.Ser896Pro p.Gln494Arg p.Lys287Glu	Possibly_damaging Probably_damaging Benign Probably_damaging	Deleterious Deleterious Deleterious Tolerated	Likely Pathogenic Uncertain significance Uncertain significance Uncertain significance

		Frameshift	0	c.3_4insCTAGTCGACTATA	p.Glu2LeufsTer10	NA	NA	Pathogenic
49	*ABCG5*	Missense	< 0.0001	c.692T>C	p.Ile231Thr	Probably_damaging	Deleterious	Uncertain significance
50.1 50.2	*STXBP2* *NVD*	Missense	0.00458	c.1586G>C	p.Arg529Pro	Probably_damaging	Deleterious	Uncertain significance
51.1	*ADAMTS13*	Splice region	< 0.0001	c.3568+7T>G	NA	NA	NA	Uncertain significance
51.2	*ADAMTS13*	Splice region	< 0.0001	c.3568+7T>G	NA	NA	NA	Uncertain significance
51.3	*ADAMTS1*	Splice region	< 0.0001	c.3568+7T>G	NA	NA	NA	Uncertain significance
52.1 52.2	*LYST* *THBD* (Dargaud et al., 2015) *LYST* *THBD* (Dargaud et al., 2015)	Stop gained Stop gained Stop gained Stop gained	0 0 0 0	c.4288C>T c.1611C>A c.4288C>T c.1611C>A	p.Arg1430Ter p.Cys537Ter p.Arg1430Ter p.Cys537Ter	NA NA NA NA	NA NA NA NA	Likely Pathogenic Pathogenic Likely Pathogenic Pathogenic
53.1	*ADAMTS13*	Stop gained	0	c.1315G>T	p.Glu439Ter	NA	NA	Pathogenic
53.2	*ADAMTS13*	Stop gained	0	c.1315G>T	p.Glu439Ter	NA	NA	Pathogenic
54.1 54.2	*NVD* *NVD*							
55	*F10* *MPL*	Missense Missense	0.0007 0.000297	c.1222G>A c.712G>T	p.Asp408Asn p.Gly238Cys	Benign Possibly_damaging	Deleterious Deleterious	Uncertain significance Uncertain significance

56	*GFI1B* (Johnson et al., 2016b)	Splice donor	0	c.814+1G>A	NA	NA	NA	Pathogenic
57	*ACVRL1*	Missense	0.00432	c.1445C>T	p.Ala482Val	Probably_damaging	Deleterious	Uncertain significance
	ITGA2B (Johnson et al., 2018)	Stop gained	0	c.2176A>T	p.Lys726Ter	NA	NA	Pathogenic
	THPO	Frameshift	0.0001	c.610dup	p.Glu204GlyfsTer123	NA	NA	Uncertain significance
	HOXA11	Missense	< 0.0001	c.248A>G	p.Tyr83Cys	Benign	Deleterious	Uncertain significance
58.1	*NBEAL2*	Splice donor	0	c.6801+1G>C	NA	NA	NA	Pathogenic
58.2	*NBEAL2*	Splice donor	0	c.6801+1G>C	NA	NA	NA	Pathogenic
59	*MECOM*	Missense	0.000175	c.580T>G	p.Tyr194Asp	NA	NA	Uncertain significance
	F8	Missense	0	c.5441A>T	p.Asp1814Val	Benign	Deleterious	Likely Pathogenic
60	*MPL*	Missense	0.00472	c.1063A>G	p.Lys355Glu	Benign	Tolerated	Uncertain significance
	NBEAL2	Missense	0.000184	c.5866G>A	p.Val1956Met	Possibly_damaging	Tolerated	Uncertain significance
	PROZ	Missense	0.00346	c.647C>T	p.Thr216Ile	Probably_damaging	Deleterious	Uncertain significance
61	*COL5A1*	Missense	0.000124	c.2146G>A	p.Glu716Lys	Possibly_damaging	Deleterious	Uncertain significance
62	*NVD*							
63.1 63.2 63.3	*GFI1B* (Rabbolini et al., 2017)	Missense	0.00438	c.503G>T	p.Cys168Phe	Probably_damaging	Deleterious	Likely Pathogenic

	GFI1B (Rabbolini et al., 2017)	Missense	0.00438	c.503G>T	p.Cys168Phe	Probably_damaging	Deleterious	Likely Pathogenic
	GFI1B (Rabbolini et al., 2017)	Missense	0.00438	c.503G>T	p.Cys168Phe	Probably_damaging	Deleterious	Likely Pathogenic
64	*NVD*							
65.1	*MECOM*	Missense	0.0001	c.951G>T	p.Lys317Asn	NA	NA	Uncertain significance
	SLFN14	Missense	0.0014	c.916G>C	p.Asp306His	Benign	Tolerated	Uncertain significance
65.2	*MECOM*	Missense	0.0001	c.951G>T	p.Lys317Asn	NA	NA	Uncertain significance
	SLFN14	Missense	0.0014	c.916G>C	p.Asp306His	Benign	Tolerated	Uncertain significance
66.1 66.2	*MYH9* (Savoia and Pecci, 2015)	Stop gained	0.0001	c.5797C>T	p.Arg1933Ter	NA	NA	Pathogenic
	MYH9 (Savoia and Pecci, 2015)	Stop gained	0.0001	c.5797C>T	p.Arg1933Ter	NA	NA	Pathogenic
67	*NVD*							
68	*RASGRP2*	Missense	0.000102	c.1159C>T	p.Arg387Cys	Probably_damaging	Deleterious	Uncertain significance
69	*F10*	Missense	0.000547	c.1406G>A	p.Arg469Lys	Benign	Tolerated	Uncertain significance
70.1 70.2	*ETV6* *NVD*	Stop gained	0	c.313C>T	p.Arg105Ter	NA	NA	Likely Pathogenic

71	*RUNX1* *RUNX1*	Missense Missense	0 0	c.1256T>G c.1270T>C	P.Val419Gly p.Ser424Pro	Benign Possibly_damaging	Deleterious Deleterious	Likely Pathogenic Likely Pathogenic
72	*FGG* *STXBP2* *TUBB1* (Johnson et al., 2016b)	Missense Missense Missense	0.00792 0.000539 0.0008	c.323C>G c.499C>T c.721C>T	p.Ala108Gly p.Arg167Trp p.Arg241Trp	Benign Possibly_damaging Probably_damaging	Tolerated Deleterious Deleterious	Likely Pathogenic Uncertain significance Uncertain significance
73	*NVD*							
74	*RUNX1* *COL5A1*	Missense Missense	0 0.000121	c.1265A>C c.5411C>A	p.Glu422Ala p.Thr1804Asn	Benign Benign	Deleterious Deleterious	Uncertain significance Uncertain significance
75	*GFI1B* (Rabbolini et al., 2017)	Missense	0.00438	c.503G>T	p.Cys168Phe	Probably_damaging	Deleterious	Likely Pathogenic
76	*THBD*	Missense	0	c.716C>T	p.Ala239Val	Benign	Tolerated	Uncertain significance
77	*THBD* *COL5A2*	Missense Missense	0 0.0001	c.752G>A c.2786C>T	p.Gly251Asp p.Ala929Val	Probably_damaging Probably_damaging	Deleterious Tolerated	Likely Pathogenic Uncertain significance
78	*STXBP2* *MCFD2*	Missense Missense	0.000201 0	c.911C>T c.364G>A	p.Thr304Met p.Asp122Asn	Probably_damaging Probably_damaging	Deleterious Deleterious	Uncertain significance Likely Pathogenic
79	*NBEAL2* *AP3D1*	Missense Missense	0.000128 0.0014	c.3184G>A c.1363G>A	p.Val1062Ile p.Ala455Thr	Possibly_damaging Possibly_damaging	Deleterious Deleterious	Uncertain significance Uncertain significance
80	*RUNX1* *MPL*	Missense Missense	0 0	c.1270T>G c.305G>A	p.Ser424Ala p.Arg102His	Possibly_damaging Probably_damaging	Deleterious Deleterious	Likely Pathogenic Uncertain significance

81	*AP3B1*	Missense	0.000809	c.2188C>T	p.Arg730Trp	Benign	Deleterious	Uncertain significance
82	*TUBB1*	Missense	0.0001	c.4C>T	p.Arg2Cys	Probably_damaging	Deleterious	Uncertain significance
	TUBB1	Missense	0.0002	c.68T>C	p.Met23Thr	Benign	Tolerated	Uncertain significance
83	*LPA*	Stop gained	0.001	c.5081C>G	p.Ser1694Ter	NA	NA	Uncertain significance
84.1	*F5*	Missense	0.00806	c.5245C>G	p.Leu1749Val	Possibly_damaging	Tolerated	Uncertain significance
	F5	Missense	0.00806	c.5054C>G	p.Thr1685Ser	Benign	Tolerated	Uncertain significance
84.2	*NBEAL2*	Missense	0.00276	c.5021G>A	p.Arg1674His	Benign	Deleterious	Uncertain significance
	HPS5	Missense	0.00593	c.345G>A	p.Met115Ile	Benign	Tolerated	Uncertain significance

5.8.2 Copy number variation found in the patient cohort

The analytical pipeline for detection of CNVs in genes included in the IBD panel was employed using the ExomeDepth coverage approach. The exome read depth of the target patient's sample was compared against the read depth of a panel of reference samples to detect regions with different coverage which could represent a CNV event. CNVs occurring within one of the genes associated with IBDs was identified in 15 patients (Table 5.4). Overall, the CNV analysis revealed an average of four CNVs per exome. Thirteen patients were observed to have common CNVs and only three patients deemed to have plausible CNVs with low frequency. There were three rare structural variants covering large regions on chromosomes 11 and 17 encompassing numerous genes, including candidate genes within the IBD gene panel. First, a rare 604-kbp CNV loss was noted on chromosome 11q24.3 in patient 35 which covered nine genes including *FLI1*. A further rare deletion was found on chromosome 11q24.3 in patient 71 which covered 31 genes including *FLI1* (Figure 5.9, A and B). Following exome depth alignment with controls, the reads ratio was 0.5 which indicates hemizygosity, as observed in (Table 5.4). In addition, a rare CNV gain was noted in patient 45 within *TBXA2R* on chromosome 19p13.3 and containing four genes in total (Figure 5.9, C). The CNV reads ratio was 2.72 which is indicative of a heterozygous insertion. As patient 26 was observed with two novel pathogenic SNVs within SMAD4, the patient observed also had a rare CNV gain in *SMAD4*. A SNV was not detected in patient 67; however, three CNVs gains in genes associated with IBD were detected in this patient including *PIGA*, *F9* and *FLNA*.

5.9 Oligogenic findings in patient cohort

Within the patient cohort there were several examples of potential oligogenic inheritance involving either two or more gene variations from the IBD gene panel.

Of particular interest was patient 16 who demonstrated an apparent pathogenic missense variant in RUNX1 and a likely pathogenic variant in *ITGB3*, both of which are plausible candidate variants to explain the thrombocytopenia and bleeding history observed. Patient 20 harbored two heterozygous missense variants within *GP6* and *THBD* in which the patient had a platelet function disorder and episodes of bleeding. In patient 30, likely pathogenic and pathogenic variations were found in *GP1BA* and *GBA* respectively. Again this patient had a low platelet count and a history of bleeding.

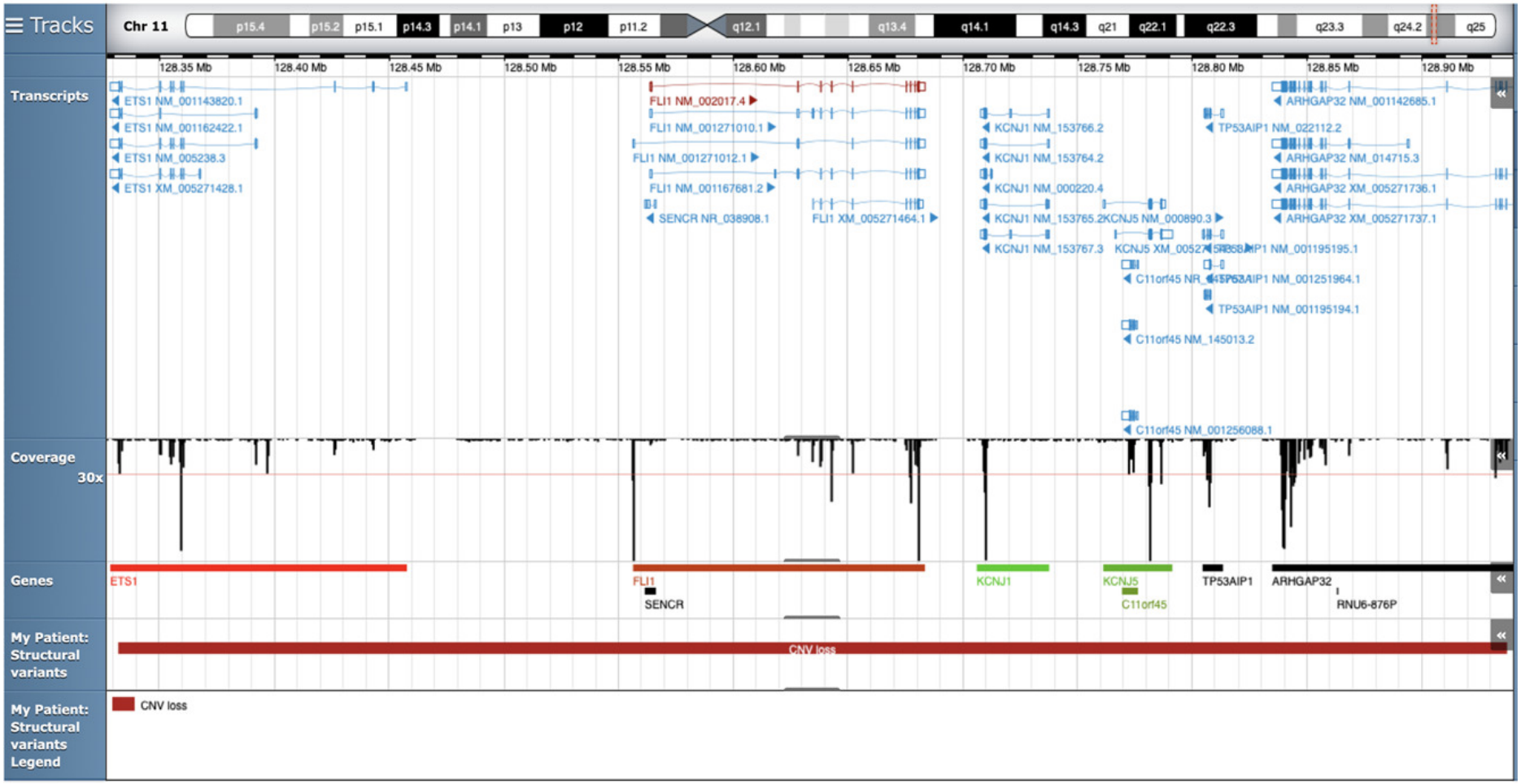
Table 5.4. Copy number variations detected in 15 patients by using Exome Depth calling approach. All genes selected are IBD related genes. Reads expected shows the read depth of patients in the reference panel. Reads observed shows the read depth of the target patient sample. Reads ratio is calculated as (reads observed/reads expected). CNV: represents the copy number variant which is calculated as (reads observed/reads expected)*2. Bayes factor: a measure of confidence of this CNV call being a real event where the higher that number, the more confident one can be about the presence of a CNV. Patients overlap indicates how many patients in the project have the same CNV. The grey colour is just used to separate each family from others.

Patient	Gene	Band	Location	Size	Type	Reads Expected	Reads Observed	Reads Ratio	CNV	Bayes Factor	Patients overlap
2.1	*ANKRD26*	10p12.1	10:27280843-27389421	10.8kbp	Loss	254	157	0.62	1.236	8.26	1
3.1	*ANKRD26*	10p12.1	10:27280843-27389421	10.8kbp	Gain	271	369	1.36	2.72	6.13	1
21	*PIGA*	Xp22.31-p21.3	X:15337573-15353676	21.0Mbp	Loss	187776	105800	0.56	1.126	3050	16
	GATA1	Xp21.1-q13.3	X:48644962-48652716	41.5Mbp	Loss	458560	256748	0.56	1.12	6580	27
	WAS	Xp21.1-q13.3	X:48534985-48549818	41.5Mbp	Loss	458560	256748	0.56	1.12	6580	27
	F9	Xq25-q27.2	X:138612917-138645617	19.0Mbp	Loss	153271	86478	0.56	1.128	2360	24
	F8	Xq27.3-q28	X:154064063-154255215	8.3Mbp	Loss	165645	90968	0.55	1.098	2410	31
	FLNA	Xq27.3-q28	X:153576892-153603006	8.3Mbp	Loss	165645	90968	0.55	1.098	2410	31
26	*SMAD4*	18q21.2	18:48494410-48611415	2.0kbp	Gain	355	495	1.39	2.78	10.8	0
35	*FLI1*	11q24.3	11:128556430-128683162	604.6kbp	Loss	9862	5420	0.55	1.1	422	1
45	*SLFN14*	17q12	17:33875144-33885117	5.2kbp	Loss	833	402	0.48	0.966	6.35	4

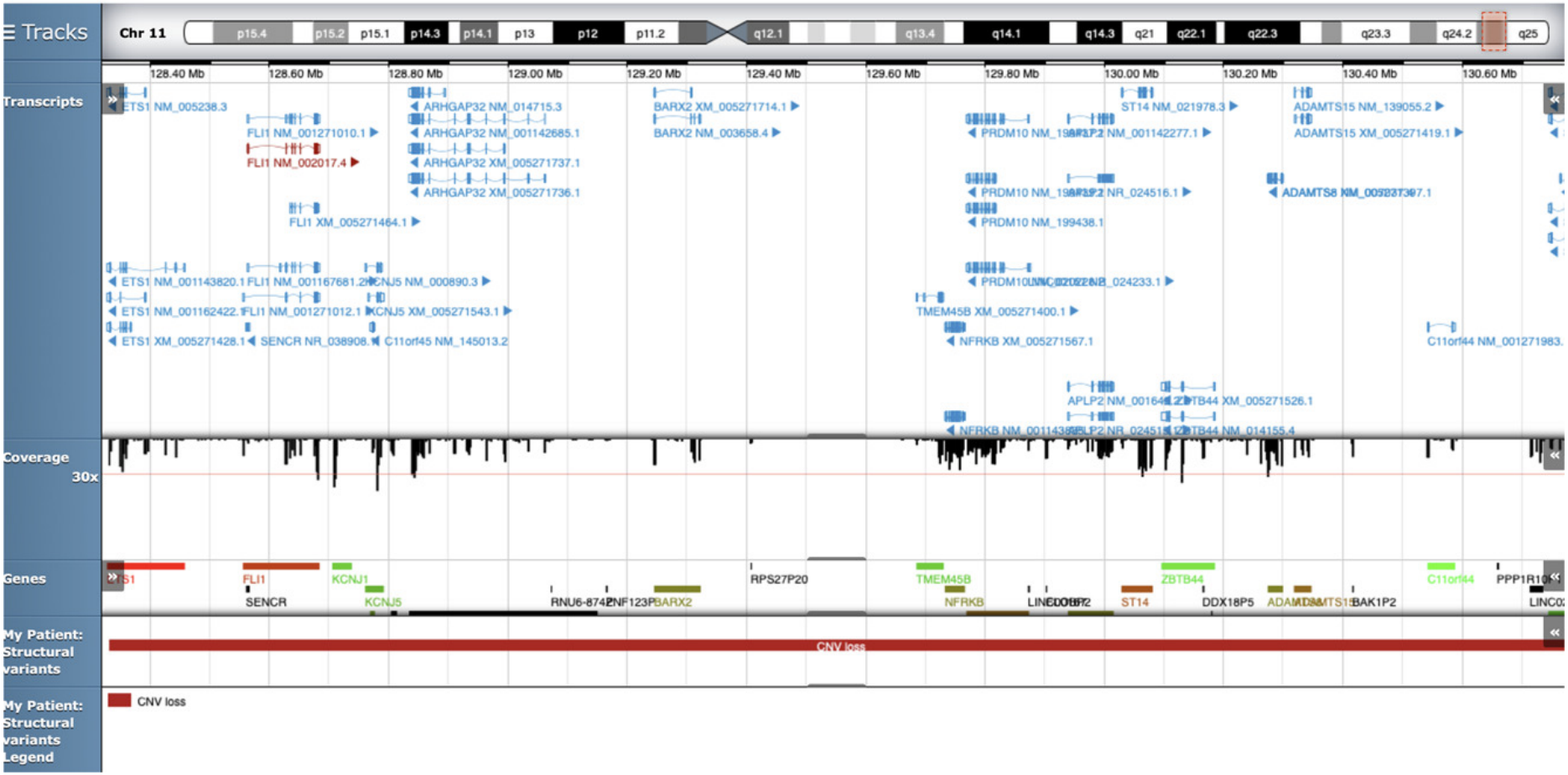
	TBXA2R	19p13.3	19:3594504-3606838	43.5kbp	Gain	1467	1993	1.36	2.72	12.2	0
	GP6	19q13.42	19:55525073-55549632	77.3kbp	Loss	1667	843	0.51	1.012	39.2	0
46	*FYB1*	5p13.1	5:39105338-39274630	145.7kbp	Gain	2146	2630	1.23	2.46	12.6	2
	NBEA	13q13.3	13:35516424-36247159	98.7kbp	Gain	527	742	1.41	2.82	6.86	0
	SLFN14	17q12	17:33875144-33885117	9.9kbp	Loss	7036	3693	0.52	1.05	15.1	4
47	*SLFN14*	17q12	17:33875144-33885117	9.9kbp	Loss	4964	2886	0.58	1.162	6.03	4
	GP1BB	22q11.21	22:19710468-19712294	37.8kbp	Loss	421	224	0.53	1.064	8.47	0
48	*SLFN14*	17q12	17:33875144-33885117	9.9kbp	Gain	4090	11381	2.78	5.56	13.3	4
52.1	*PIGA*	Xp22.2-p22.13	X:15337573-15353676	3.6Mbp	Loss	49781	32738	0.66	1.316	275	3
	GATA1	Xp11.3-p11.22	X:48644962-48652716	5.3Mbp	Loss	154142	94557	0.61	1.226	756	15
	WAS	Xp11.3-p11.22	X:48534985-48549818	5.3Mbp	Loss	154142	94557	0.61	1.226	756	15
	F9	Xq27.1	X:138612917-138645617	476.4kbp	Loss	11569	7853	0.68	1.358	53.5	1
	F8	Xq28	X:154064063-154255215	4.4Mbp	Loss	148620	91530	0.62	1.232	730	15
	FLNA	Xq28	X:154064063-154255215	4.4Mbp	Loss	148620	91530	0.62	1.232	730	15
52.2	*PIGA*	Xp22.33-p21.3	X:15337573-15353676	23.5Mbp	Gain	205946	298796	1.45	2.9	2890	21
	GATA1	Xp11.3-q13.3	X:48644962-48652716	29.3Mbp	Gain	361570	519986	1.44	2.88	4150	25
	WAS	Xp11.3-q13.3	X:48534985-48549818	29.3Mbp	Gain	361570	519986	1.44	2.88	4150	25
	F9	Xq25-q27.3	X:138612917-138645617	14.4Mbp	Gain	134092	194039	1.45	2.9	1690	21
	F8	Xq27.3-q28	X:154064063-154255215	9.9Mbp	Gain	153704	221043	1.44	2.88	1830	38
57	*COL5A1*	9q34.3	9:137533620-137736686	190bp	Loss	158	108	0.68	1.368	5.27	0

67	*PIGA*	Xp22.2-p22.11	X:15337573-15353676	9.8Mbp	Gain	56608	80502	1.42	2.84	1070	10
	F9	Xq26.3-q27.2	X:138612917-138645617	5.4Mbp	Gain	22574	32057	1.42	2.84	386	21
	FLNA	Xq28	X:153576892-153603006	6.0Mbp	Gain	58278	86619	1.49	2.98	1350	39
71	*FLI1*	11q24.3	11:128556430-128683162	2.5Mbp	Loss	20510	10805	0.53	1.054	912	2
72	*F9*	Xq26.3-q27.2	X:138612917-138645617	5.4Mbp	Gain	23065	32667	1.42	2.84	469	21
	FLNA	Xq28	X:153576892-153603006	260.1kbp	Gain	15952	22830	1.43	2.86	322	7

(a)


Tracks
Chr 11
p15.4 p15.2 p15.1 p14.3 p14.1 p13 p12 p11.2 q12.1 q13.4 q14.1 q14.3 q21 q22.1 q22.3 q23.3 q24.2 q25
128.35 Mb 128.40 Mb 128.45 Mb 128.50 Mb 128.55 Mb 128.60 Mb 128.65 Mb 128.70 Mb 128.75 Mb 128.80 Mb 128.85 Mb 128.90 Mb
Transcripts
ETS1 NM_001143820.1
ETS1 NM_001162422.1
ETS1 NM_005238.3
ETS1 XM_005271428.1
FLI1 NM_002017.4
FLI1 NM_001271010.1
FLI1 NM_001271012.1
FLI1 NM_001167681.2
SENCR NR_038908.1
FLI1 XM_005271464.1
KCNJ1 NM_153766.2
KCNJ1 NM_153764.2
KCNJ1 NM_000220.4
KCNJ1 NM_153765.2
KCNJ5 NM_000890.3
KCNJ1 NM_153767.3
TP53AIP1 NM_022112.2
ARHGAP32 NM_001142685.1
ARHGAP32 NM_014715.3
ARHGAP32 XM_005271736.1
ARHGAP32 XM_005271737.1
TP53AIP1 NM_001195195.1
TP53AIP1 NM_001251964.1
TP53AIP1 NM_001195194.1
C11orf45 NM_145013.2
C11orf45 NM_001256088.1
Coverage
30x
Genes
ETS1
FLI1
SENCR
KCNJ1
KCNJ5
C11orf45
TP53AIP1
ARHGAP32
RNU6-876P
My Patient: Structural variants
CNV loss
My Patient: Structural variants Legend
CNV loss

(b)


Tracks
Chr 11
p15.4 p15.2 p15.1 p14.3 p14.1 p13 p12 p11.2 q12.1 q13.4 q14.1 q14.3 q21 q22.1 q22.3 q23.3 q24.2 q25
128.40 Mb 128.60 Mb 128.80 Mb 129.00 Mb 129.20 Mb 129.40 Mb 129.60 Mb 129.80 Mb 130.00 Mb 130.20 Mb 130.40 Mb 130.60 Mb
Transcripts
ETS1 NM_005238.3
FLI1 NM_001271010.1
FLI1 NM_002017.4
ARHGAP32 NM_014715.3
ARHGAP32 NM_001142685.1
ARHGAP32 XM_005271737.1
ARHGAP32 XM_005271736.1
FLI1 XM_005271464.1
BARX2 XM_005271714.1
BARX2 NM_003658.4
ST14 NM_021978.3
ADAMTS15 NM_139055.2
ADAMTS15 XM_005271419.1
PRDM10 NM_001142277.1
PRDM10 NM_024516.1
PRDM10 NM_199438.1
ETS1 NM_001143820.1
FLI1 NM_001167681.2
KCNJ5 NM_000890.3
ETS1 NM_001162422.1
FLI1 NM_001271012.1
KCNJ5 XM_005271543.1
ETS1 XM_005271428.1
SENCR NR_038908.1
C11orf45 NM_145013.2
TMEM45B XM_005271400.1
NFRKB XM_005271567.1
C11orf44 NM_001271983.
APLP2
ZBTB44 XM_005271526.1
ZBTB44 NM_014155.4
Coverage
30x
Genes
ETS1
FLI1
KCNJ1
SENCR
KCNJ5
RPS27P20
BARX2
TMEM45B
NFRKB
ZBTB44
ST14
DDX18P5
BAK1P2
C11orf44
PPP1R10
My Patient: Structural variants
CNV loss
My Patient: Structural variants Legend
CNV loss

(c)

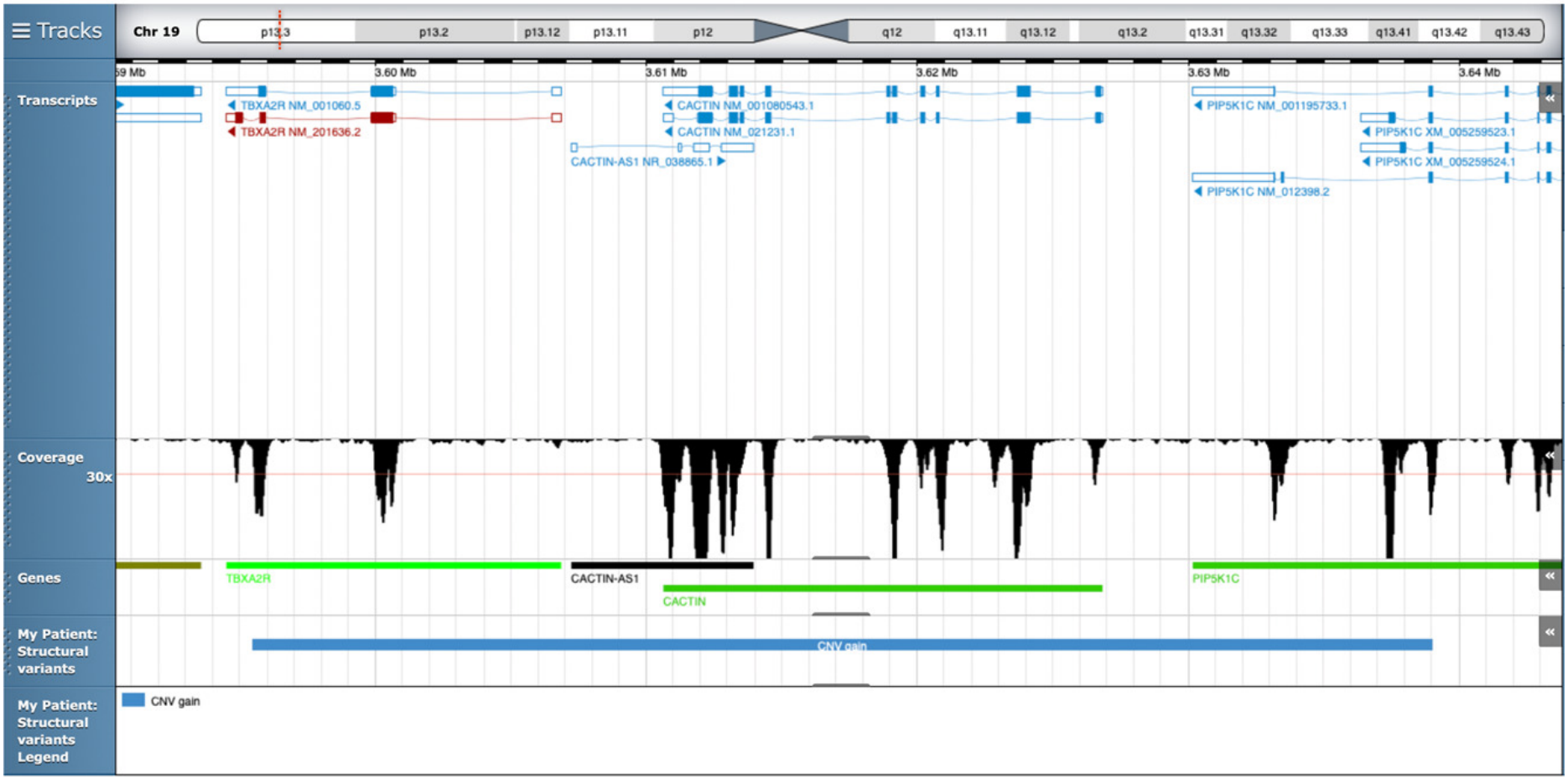


Figure 5.9. Copy number variants found in cohort of GAPP patients. (A)CNV loss was noted on chromosome 11q24.3 in patient 35 which covered 9 genes including *FLI1*. (B) CNV loss was noted on chromosome 11q24.3 in patient 71 which covered 31 genes including *FLI1*. (C) CNV gain was noted on chromosome 19p13.3 in patient 45 which covered 4 genes including *TBXA2R*.

5.10 Discussion

NGS approaches have increasingly been used over the last decade in the molecular diagnosis of IBDs. Here, a large-scale application of WES analysis was carried out by using a robust molecular diagnostics platform for diagnosis of 117 patients recruited to the UK-GAPP study. The aim was to assess the ability of Congenica software to analyse WES data of the patients for both sequence and structural variants by targeting the known IBD genes. Subsequently, patients with variants in known bleeding disorder genes can be eliminated by a series of filtration steps and WES data targeted for those with undetected variants who may harbour a variant in novel genes. A total of 119 genes were included and applied to patient cohort for filtering. However, it is important to note that currently only 88 of these genes are considered clinical-grade genes according to Genomics England. The remainder of the genes may become clinical grade once more variants are identified in patients and deemed pathogenic over time. The UK-GAPP study has used WES data in combination with deep platelet phenotyping to identify the disease-causing variants in both known and novel genes in approximately 40% of patients. However, the remaining patients are still without a clear molecular and genetic diagnosis for their bleeding disorder. 126 patients were analysed by the Congenica software for both SNVs and CNVs following targeting the known IBD gene panel of 119 genes.

Phenotypic presentation and platelet counts varied considerably among the recruited patients, which is consistent with the variability of clinical presentation between patients with suspected IBDs. However, the majority of patients 33/117 (28.2%) were noted with a platelet function defect and 23/117(19.6%) patients represented thrombocytopenia only. Of the 117 recruited patients, 15 (12.8%) patients were deemed to have a macrothrombocytopenia. Platelet function studies revealed the

presence of a combination of platelet defects in addition to thrombocytopenia in 36/117 (30%) of patients. The majority of patients with platelet defects displayed both secretion and Gi defects. However, a previous study has shown that some patients with normal lumi-aggregeometry results have platelet spreading defects, indicating the difficulties faced when diagnosing patients with IBDs and the multitude of assays required for platelet phenotype disorders to be diagnosed (Khan et al., 2020).

Overall, a total of 135 variants in genes implicated in bleeding disorders were identified across all 117 patients and all variants were observed in a heterozygous state, implicating dominant inheritance patterns. This project has shown that the majority of plausible candidate variants were associated with IT genes which explain the association of thrombocytopenia with platelet defects in the majority of patients. When considering pathogenicity prediction, 22 patients were classified as pathogenic and 26 patients as likely pathogenic, while 87 patients had uncertain pathogenicity and therefore classified as uncertain significance. A targeted WES analysis was previously carried out on some patients which identified genetic variants in IT with or without secondary qualitative defects (Johnson et al., 2018, Johnson et al., 2016b). This study has conclusively identified these genetic variants, which indicates the ability of the Congenica platform to analyse and provides suitable validation of WES data in these patients.

As the majority of the patients have been previously analysed in other studies that used the bioinformatic pipeline described in Chapter 3, the performance of the Congenica tool for detecting SNVs and CNVs was compared to these studies (Johnson et al., 2018, Johnson et al., 2016b). 25 variants were identified by the Congenica software as well as the other bioinformatic pipeline and the majority of them were either pathogenic or likely pathogenic. A further 24 variants thatwere classified

as pathogenic or likely pathogenic were identified by the Congenica software only. Therefore, this showed that the Congenica software is a more robust tool to analyze WES as it provides a higher variant detection rate and more suitable for analysing large cohort of patients compared with other bioinformatic tools. It is also important to note that no variants of uncertain significance were included here in this evaluation as it is difficult to assign causality but they are still plausible pathogenic variants. The Number of patients reported with SNVs in IBD related genes is shown in Figure 5.10.

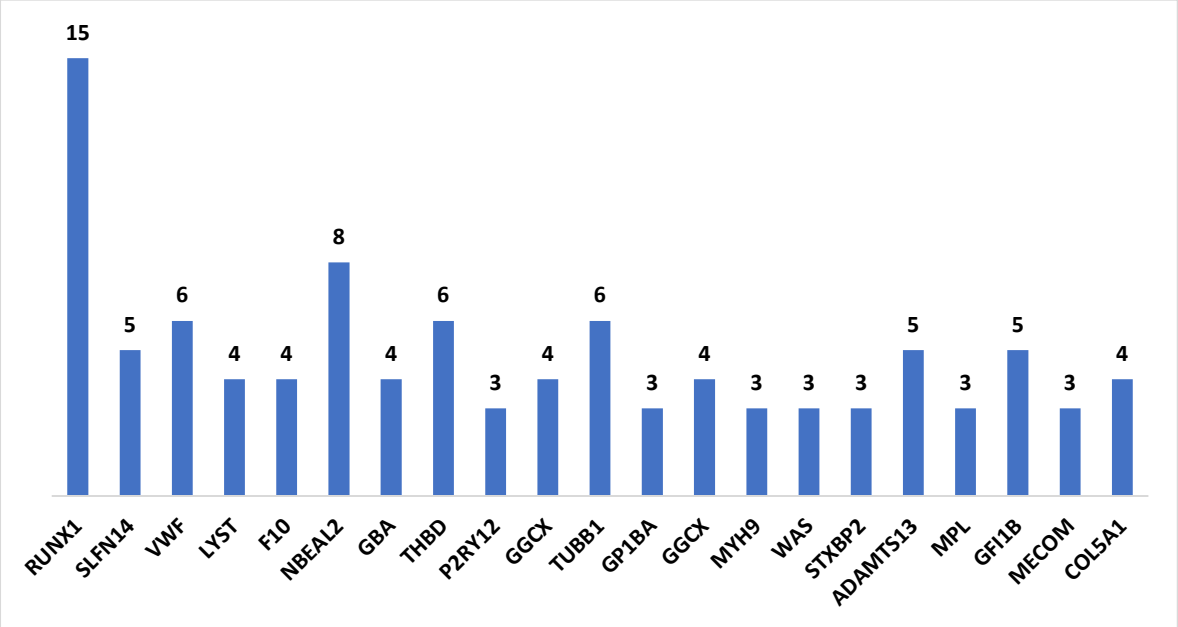


Figure 5.10. Number of patients reported with SNVs in IBD related genes. The number of patients identified with variants in each gene is shown above each related gene bar. This graph includes only genes that have been identified in more than or equal to 3 patients. The remaining genes were reported in less than or equal to 2 patients.

Congenica software also has the added benefit of detecting CNVs, a process which is notoriously difficult yet valuable in identifying rare causative variants in heterogeneous diseases. Congenica utilizes the integrated ExomeDepth tool to compare a target with reference and here, identified rare CNVs in this population. Congenica software alongside targeted gene panel searching, allows for efficient and accessible detection of variants and with some clinical interpretation will be a valuable tool when analyzing large datasets (Nowakowska, 2017, Valsesia et al., 2012).

Paris-Trousseau syndrome, characterised by a bleeding defect with large α-granules and abnormal megakaryocyte morphology is well documented, which is caused by a dominant inheritance of q23 deletion on chromosome 11 (Stevenson et al., 2015). Patients with this disorder have variable size of chromosomal deletion associated with different components of the syndrome. A hemizygous deletion of *FLI1* was attributed to the platelet defect in 2 cases of our cohort. These CNVs noted in patients 35 and 71 cover the deletion region in *FLI1* and are also surrounded by several flanking genes. Both patients presented with thrombocytopenia and a secretion defect which suggest the platelet phenotype and the CNVs in *FLI1* to be associated with their disorders. Thromboxane receptor deficiency is an autosomal recessive or dominant disorder characterized by bleeding symptoms associated with quantitative or qualitative defects within the thromboxane receptor (Mundell and Mumford, 2018). Although there were no plausible candidate SNVs in the 119 candidate genes or the thromboxane receptor in patient 45, there was a rare CNV duplication in the *TBXA2R* gene and thus deduce that either alone or in combination with variants in *GP6* and *SLFN14* which were also detected, could be causative of the patient's thrombocytopenia and bleeding. As patient 26 was observed with two novel heterozygous pathogenic SNVs within *SMAD4*, the patient also had a rare CNV gain

in *SMAD4*. This could indicate that the *SMAD4* gene is more likely to be associated with the bleeding disorders in this patient. It has been shown that mutations within *SMAD4* gene can cause two distinct clinically autosomal dominant diseases called hereditary haemorrhagic telangiectasia (HTT) and juvenile polyposis (JP). Patients with HTT present with a heterogenous phenotype including bleeding, anaemia and stroke (Schwenter et al., 2012). As the two SNVs in *SMAD4* are heterozygous, this is compatible with the HTT inheritance pattern. An SNV was not detected in patient 67; however, three CNVs gain in genes associated with IBD were detected in this patient including *PIGA*, F9 and *FLNA* which may be associated with its bleeding disorder.

Patients without a plausible SNV or CNV candidate might harbour a novel gene that is associated with their bleeding disorder. In the future, it would be interesting to investigate these CNVs further to determine the extent of the contiguous deletions or insertions by long-range polymerase chain reaction and sequencing to determine the breakpoints and mechanisms of the variant, as well as confirming these regions using multiplex ligation-dependent probe amplification, should kits be available for these genomic regions.

In summary, this chapter has shown validation and a practical approach of a robust diagnostic platform that can be employed for WES analysis. This study has made use of data from a cohort of patients with suspected IBDs; a broad category of diseases, well acknowledged in the haematology field as difficult to classify and associate to single causative genetic abnormalities. This study has shown the ability of the software to detect CNVs with high efficiency with the use of targeted gene panels as a replacement of traditional methods for detecting CNVs. To conclude, this data reveals use of a highly sensitive and valuable tool which can be used for detecting SNVs and CNVs based on WES data. This is one of the first studies, albeit in a

research setting, to implement this software for both SNV and CNV analysis. This is a leap forward in the ability to classify hugely complex disorders with a high degree of heterogeneity within the wider scientific community providing concise and definitive diagnosis for patients.

5.11 Future work

The effectiveness of molecular next generation sequencing analysis relies on subsequent functional studies to confirm the effect of the genetic variants on gene and protein function. Future work should focus on functional approaches to investigate these variants and how likely they associate with diseases. Pathogenicity *in-silico* prediction tools can be used to classify variants but without *in vitro or in vivo* studies, the role of these variants in the mechanism of diseases are often ambiguous and uncertain.

Further studies are needed to investigate the molecular mechanism of CNVs found in patients 35, 45 and 71. For example, long amplification PCR can be carried out to determine the breakpoints in these structural variants.

5.12 Key findings

- Platelet function studies identified qualitative or possible qualitative defects in 68/117 patients and 36/117 patients were found to harbour both qualitative defects and low platelet counts.
- All 135 variants found among patients were in heterozygous state. Of these variants, 22 variants were classified as pathogenic and 26 variants were classified as likely pathogenic, while the remaining 87 variants were classified as of uncertain significance.
- 63% of patients (74/117) were observed to harbour variants within genes previously implicated with IT.
- The number of variants found to be shared in the same affected family members were 21 variants.
- 24 variants classified as pathogenic or likely pathogenic identified by the Congenica software but not with other bioinformatic analysis tools.
- Identification of three rare structural variants covering large regions on chromosomes 11 and 17 and encompassing numerous genes, including candidate genes within the IBD gene panel.
- Combination of WES data with platelet function testing and bleeding phenotypes are useful approach to investigate patients with inherited bleeding disorders with unknown causes. In addition, analysing WES data by using a robust molecular diagnostic platform is a highly sensitive and valuable approach which can be used for detecting SNVs and CNVs.

Chapter.6
Overall discussion and conclusion

6.1 Overview of inherited platelet and bleeding disorders

Inherited bleeding disorders (IBDs) are a heterogenous group of diseases that result from defective abnormalities in blood vessels, coagulation proteins and platelets. The most common bleeding disorders are VWD diseases and/or coagulation factor diseases; however, the frequency of inherited platelet disorders (IPDs) is low and less understood (Sivapalaratnam et al., 2017a, Bolton-Maggs et al., 2006). As platelets express numerous proteins, IPDs can present with wide ranging bleeding symptoms which makes diagnosis difficult. Patients recruited to the UK-GAPP study both overall and within this thesis display bleeding symptoms that are extremely heterogeneous in terms of their clinical presentation and demographic information. These include patient's platelet characteristics in platelet size and number, as well as the phenotypic bleeding presentation which ranged from mild to severe bleeding. This is consistent with the variability of clinical presentation and laboratory phenotype between patients with suspected IBDs as mentioned previously in chapters 3 and 5 of this thesis.

6.2 Patient genotyping by whole exome sequencing (Chapter 3)

The application of NGS technologies was firstly employed for diagnosis of patients with IBDs through the GAPP study and the BRIDGE - Bleeding and Platelet Disorders (BPD) consortium (Westbury et al., 2015, Nurden and Nurden, 2018, Johnson et al., 2018, Johnson et al., 2016b, Stockley et al., 2013). The later study initially used WES and later WGS to identify the disease-causing genes behind the BPDs. The UK-GAPP study used a WES approach to identify the causative of IPDs in patients with unknown aetiology (Watson et al., 2013b, Maclachlan et al., 2017), in combination with phenotyping to identify the likely causative genes in recruited patients. In chapter 3, exome data was performed in 6 patients which included 4 isolated individuals and 2 related individuals with bleeding symptoms and thrombocytopenia. They were

analysed by a custom designed bioinformatic pipeline established by the members of the UK-GAPP study. This was used to narrow down the number of candidate variants and classify potential disease-causing variants based on the ACMG guidelines. Two pathogenic variants within *ITGA2B* and *ABCG8* genes were identified in isolated patients 2 and 3. In addition, one pathogenic variant in *CD36* was shared between two affected patients 5.1 and 5.2. Patient 4 was identified with a likely pathogenic in the *FLNA* gene. All patients were presented with low platelet counts and high MPV and IPF values. *ITGA2B*, *ABCG8* and *FLNA* have previously been associated with IT which is compatible with patient's findings (Almazni et al., 2019). However, the bioinformatic pipeline used in chapter 3 is not without its limitation. It is time-consuming and not suitable for analysing large numbers of patients. Moreover, identifying potential causative variants and/or novel genes in isolated affected patients is challenging.

However, recruitment of additional family members particularly those who share similar phenotypes could help to identify the plausible causative genetic variants behind their disease, by performing enhanced genetic segregation analysis. This was illustrated in the large family of Pakistani origin Section 3.5.2 where a novel splice variant within *CEBPZ (*c.1881+3A>G) segregated between 5 affected family members. This variant could be predicted to result in subsequent exon skipping or creation of a cryptic splice site leading to reduced/absent expression of the mutant allele through nonsense-mediated RNA decay. In summary, this chapter has demonstrated that the combination of phenotyping and genotyping coupled with WES technology and bioinformatic tools are an efficient and effective approach for refining and identifying new sequence variants in patients with suspected IPDs. The advent of the 100,000 Genome Project will provide a considerably larger volume of genetic data which over

time will aid the investigation of a wide spectrum of rare genetic diseases (Turnbull et al., 2018).

6.3 Functional characterisation of a *CD36* variant in a family with IT

The ability to identify the genetic causes behind IPDs is often difficult due to the heterogeneity of disease as well as the possibility that identified variants may be benign. Functional studies of candidate variants is therefore required in order to determine the functional effect of variants on the protein and whether the candidate variants can be excluded or not. Functional work from the initial genetic findings in chapter 3 were carried out after identifying a stop gain variant within *CD36* (c.975T>G; p.Tyr325*) by WES in two related affected patients 5.1 and 5.2. The affected patients presented with similar bleeding episodes and low platelet counts in addition to high MPV and IPF values. Functional characterisation was conducted to investigate the functional effect of the *CD36* variant and investigate the potential role of this protein on platelet function which is understudied in the literature to date.

The affected patients 5.1 and 5.2 appear to have large platelet size as their MPV values were above the normal range. It has been shown that many forms of IT are associated with large platelet size which is compatible with patient's findings (Noris et al., 2013). It has also been suggested that the value of IPF is influenced by platelet size which could be a potential parameter for screening macrothrombocytopenia (Miyazaki et al., 2015). CD36 expression in protein lysates from patients' platelets showed a truncation of the mature CD36 protein. This finding was confirmed by the expression of mutant CD36 constructs using flow cytometry, which showed that only WT CD36 was expressed at the cell surface while the mutant constructs were not. Moreover, there was activation of NFAT-luciferase reporter assay by the WT CD36, but not by the mutant (deletion) constructs providing supportive evidence on the

subsequent effect of the nonsense variant on the function the CD36 protein. Furthermore examining the modelled protein structure of the WT CD36 ectodomain demonstrated that the CD36 mutant variant could change the structure of the overall protein, where the deleted part may be involved with and perturb normal protein signalling. It has been shown that CD36 causes platelet hyperactivity through its signalling cascade which occurs within the intracytoplasmic C-terminal domain (Silverstein et al., 2010, Silverstein and Febbraio, 2000). As the truncated part of CD36 involves the C-terminal domain, this could alter the signalling cascade of the protein pathway involved. A study conducted by (Aitman et al., 2000) has revealed the presence of homozygous or compound heterozygous frameshift or nonsense mutations within *CD36* in five affected African-American patients with CD36 deficiency. One of these patients presented with thrombocytopenia which could explain the presence of thrombocytopenia in the patients studied in this thesis. Further functional work is required to fully elucidate the molecular mechanism of CD36 causing the bleeding disorder observed in the patients studied in this thesis, and whether the variant is definitively disease causing. In this study, follow-up functional characterisation is ongoing on the CD36 mutant variant which will be mentioned in the future work below.

6.4 Analysing WES in IBD patients by the Congenica platform

As the adapted bioinformatic workflow used in chapter 3 is time consuming and not suitable to analyse large cohort of patients, the Congenica platform was used to analyse WES data of large cohort of patients recruited to the UK-GAPP study. The UK-GAPP study has used WES data in combination with deep platelet phenotyping to identify the causative variants in both known and novel genes in approximately 40% of patients. However, the remaining patients are still without a clear molecular and

genetic diagnosis for their bleeding disorder. Until recently CNVs and other structural variations were not considered. Indeed finding such lesions is challenging especially for heterozygous CNVs where the WT allele would mask the mutant allele using conventional WES or Sanger sequencing following PCR amplification. In Chapter 5 Congenica software was used to search for both SNVs and CNVs in these patients. 126 patients were analysed by the Congenica software following targeting the known IBD gene panel of 119 genes which contains the most likely mutated genes implicated in IPD to date. 9/126 patients harbour five different known genetic variants that identified previously by WES. These pateints were used for validation of WES analysis by the Congenica software. The remaining 117 patients were subjected for further analysis in this chapter including phenotyping and genotyping. Patients analysed presented with heterogeneous bleeding phenotypes where the easy bruising was the most reported phenotype among patients (68%). The majority of patients with platelet counts in whole blood were thrombocytopenic (77%), while patients with normal platelet count were (19.6%). Although the majority of patients displayed both reduced platelet count and platelet function defects, there are some patients who presented with abnormal platelet function despite their normal platelet counts. It has been reported that a significant number of IPDs have platelet function defects and normal platelet counts with moderate to severe bleeding diathesis (Nurden et al., 2021). These findings highlight the variability of IPDs which indeed require a careful assessment for characterising platelet function and bleeding disorders. A total of 135 SNVs in genes implicated in bleeding disorders were identified across all the 117 patients and all variants were observed in a heterozygous state. After considering the ACMG consensus guidelines, 22 variants were classified as pathogenic and 26 variants were likely pathogenic and the remaining 87 variants were classified as of

'uncertain significance'. The majority of reported variants were within patients with abnormal platelet count and function. 63% of patients were observed to harbour variants within genes previously implicated with IT and 35.5% of these variants were novel. In addition, ExomeDepth coverage was used to identify CNVs by comparing the exome read depth of the target patient's sample against the read depth of a reference panel to detect regions with different coverage, which could represent a CNV event. Three rare CNVs were identified within patients 35, 45 and 71, which illustrate the ability of the Congenica software to detect CNVs with high efficiency when using targeted gene panels as a replacement of traditional methods for detecting CNVs. In addition, 25 variants were identified by the Congenica software as well as other bioinformatic tools used in other studies (Johnson et al., 2018, Johnson et al., 2016b). However, a further 24 variants were classified as pathogenic or likely pathogenic that were detected by the Congenica software only. Therefore, the study has shown that the Congenica software is a more robust tool to analyze WES data as it provides a higher variant detection rate compared with other bioinformatic tools (Johnson et al., 2018, Johnson et al., 2016b, Almazni et al., 2020).

6.5 Final conclusion

There are some limitations of using WES such as the inability of WES to assess the non-coding regions, which may include potential variants in regulatory regions (Nurden and Nurden, 2020, Rabbani et al., 2014, Zhang and Lupski, 2015). This is more likely to be compatible with patients where a causative gene defect was not detected. WES analysis is still time consuming and the coverage of regions of interest is not always complete or as robust as desired. In addition the detection of CNVs by WES is still difficult but this can be reached through using advanced bioinformatic analytical software such as the Congenica software (Almazni et al., 2020). The Congenica

platform shows its ability to detect CNVs by using exome depth as identified in patients 35, 45 and 71 (chapter 5). WGS is a promising method which can potentially detect novel genes and their variants in coding and non-coding regions, that might associated with patients with an unexplained bleeding disorder (Nurden and Nurden, 2018, Turro et al., 2020). WGS method can be combined with other selective approaches such as comparing clinical phenotypes of affected patients with mouse models or gene expression databases. These approaches helped to identify the *TRPM7* and *DIAPH1* genes from an existing mouse model and subsequently searching of genes causing deafness and associated platelet (syndromic) defects respectively. The improvement of bioinformatic tools would help to analyse WGS data. A collaboration among scientists, researchers, clinicians and industrial vendors either nationally or worldwide is needed for the next phase of NGS and genomic medicine, in order to search for novel genes associated with heritable platelet disorders (Desai and Jere, 2012).

6.6 Future work arising from this thesis

Sometimes attributing a disease-causing genetic variant to a disease such as IPD is only the first step in diagnosis and working out the precise aetiology and functional consequence of these variants is important, albeit in a research context. Below there is suggested future work beyond this thesis which should take place to strengthen the research findings.

- The novel splice variant c.1881+3A>G in *CEBPZ* segregated in patients 6.1, 6.3, 6.4, 6.5 and 6.8, and would result in subsequent skipping of exon or create a cryptic splice site or reduced/absent expression of the mutant allele through nonsense-mediated RNA decay. Future works need to be done to investigate

the effect of the *CEBPZ* splice variant. To evaluate this variant further, an end point and/or quantitative reverse transcription polymerase chain reaction (qRT-PCR) can be performed using specific *CEBPZ* primers to investigate if this splice variant results in differential expression/spllicing of the gene specifically in platelets/megakaryocytes. Moreover, inducible pluripotent stem cells (iPSCs) are an ideal option as they self-renewing and widely used recently. They can be used to produce a reasonable number of megakaryocytes combined with CRISPR-mediated genomic editing to introduce the candidate variant, in order to study *CEBPZ*, where primary patient material is difficult to source.

- The plausible candidate variants in patients 1, 2, 3 and 4 could be causative of bleeding. Recruitment of additional affected and unaffected family members to determine if these candidate variants segregate within affected family members and strengthen the genetic evidence. Further functional characterisation to investigate the effect of these variants through using *in-vivo* and/or *in-vitro* approaches would be helpful to understand the effect of these variants on their proteins function.
- Patients with no plausible candidate variants after WES analysis could harbour variants in novel genes or within non-coding regions of genes. Recruitment of additional affected and unaffected family members would be useful to identify such candidate variants through using segregation analysis. Application of WGS would be applicable to identify potential variants in regulatory regions.
- The heterozygous stop gain variant in exon 10 of the *CD36* gene (c.975T>G; p.Tyr325Ter) in patients 5.1 and 5.2 is likely the causative variant of the bleeding disorder and IT. As the mother is reported to have bleeding episodes, it is recommended to obtain a blood sample in order to determine if the mother

has the same variant and thus confirming maternal dominant inheritance. Further functional studies are ongoing, such as co-immunoprecipitation to investigate the CD36 and CD9 localisation on the platelet surface. In addition, the localisation of CD36 and CD9 can be investigated by using the immunofluorescence confocal microscopy.

- The CNVs identified in patients 35, 45 and 71 are plausible structural variants associated with their bleeding disorder, which were identified using read coverage (ExomeDepth) from WES data. Long amplification PCR can be carried out to determine the breakpoints in these structural variants. Furthermore Multiplex ligation-dependent probe amplification (MLPA) can be used to confirm these regions, but only once such commercial kits containing the candidate genes of interest are available.

References

2009.
Abe, Y., Wada, H., Tomatsu, H., Sakaguchi, A., Nishioka, J., Yabu, Y., Onishi, K., Nakatani, K., Morishita, Y. & Oguni, S. 2006. A simple technique to determine thrombopoiesis level using immature platelet fraction (IPF). *Thrombosis research,* 118**,** 463-469.
Adzhubei, I., Jordan, D. M. & Sunyaev, S. R. 2013. Predicting functional effect of human missense mutations using PolyPhen-2. *Current protocols in human genetics,* 76**,** 7.20. 1-7.20. 41.
Aitman, T. J., Cooper, L. D., Norsworthy, P. J., Wahid, F. N., Gray, J. K., Curtis, B. R., McKeigue, P. M., Kwiatkowski, D., Greenwood, B. M. & Snow, R. W. 2000. Malaria susceptibility and CD36 mutation. *Nature,* 405**,** 1015-1016.
Akashi, K., Traver, D., Miyamoto, T. & Weissman, I. L. 2000. A clonogenic common myeloid progenitor that gives rise to all myeloid lineages. *Nature,* 404**,** 193-197.
Al-Rahal, N. K. 2018. Inherited Bleeding Disorders in Iraq and Consanguineous Marriage. *International Journal of Hematology-Oncology and Stem Cell Research,* 12**,** 273.
Alexander, W. S., Viney, E. M., Zhang, J.-G., Metcalf, D., Kauppi, M., Hyland, C. D., Carpinelli, M. R., Stevenson, W., Croker, B. A. & Hilton, A. A. 2006. Thrombocytopenia and kidney disease in mice with a mutation in the C1galt1 gene. *Proceedings of the National Academy of Sciences,* 103**,** 16442-16447.
ALHENC-GELAS, M., Canonico, M., Morange, P.-E., Emmerich, J. & GROUP, G. G. T. 2010. Protein S inherited qualitative deficiency: novel mutations and phenotypic influence. *Journal of Thrombosis and Haemostasis,* 8**,** 2718-2726.
Almazni, I., Stapley, R. & Morgan, N. V. 2019. Inherited thrombocytopenia: update on genes and genetic variants which may be associated with bleeding. *Frontiers in cardiovascular medicine,* 6**,** 80.
Almazni, I., Stapley, R. J., Khan, A. O. & Morgan, N. V. 2020. A comprehensive bioinformatic analysis of 126 patients with an inherited platelet disorder to identify both sequence and copy number genetic variants. *Human Mutation,* 41**,** 1848-1865.
Andrews, N. C. 1998. Molecules in focus The NF-E2 transcription factor. *The international journal of biochemistry & cell biology,* 30**,** 429-432.
Anna, A. & Monika, G. 2018. Splicing mutations in human genetic disorders: examples, detection, and confirmation. *Journal of applied genetics,* 59**,** 253-268.
Ansorge, W. J. 2009. Next-generation DNA sequencing techniques. *New biotechnology,* 25**,** 195-203.
Apolinário, T., Paiva, C. & Agostinho, L. 2017. Intermediate alleles of Huntington's disease HTT gene in different populations worldwide: a systematic review. *Genet. Mol. Res,* 16**,** 1-11.
Arai, F., Hirao, A., Ohmura, M., Sato, H., Matsuoka, S., Takubo, K., Ito, K., Koh, G. Y. & Suda, T. 2004. Tie2/angiopoietin-1 signaling regulates hematopoietic stem cell quiescence in the bone marrow niche. *Cell,* 118**,** 149-161.
Armesilla, A. L. & Vega, M. A. 1994. Structural organization of the gene for human CD36 glycoprotein. *Journal of Biological Chemistry,* 269**,** 18985-18991.
Arnout, J., Hoylaerts, M. & Lijnen, H. 2006. Haemostasis. *The Vascular Endothelium II.* Springer.
Asch, A. S., Leung, L., Polley, M. J. & Nachman, R. L. 1985. Platelet membrane topography: colocalization of thrombospondin and fibrinogen with the glycoprotein IIb-IIIa complex. *Blood* 66**,** 926–934.
Asch, A. S., Liu, I., Briccetti, F. M., Barnwell, J. W., Kwakye-Berko, F., Dokun, A., Goldberger, J. & Pernambuco, M. 1993. Analysis of CD36 binding domains: ligand specificity controlled by dephosphorylation of an ectodomain. *Science,* 262**,** 1436-1440.
Avanzi, M. P. & Mitchell, W. B. 2014. Ex vivo production of platelets from stem cells. *British journal of haematology,* 165**,** 237-247.

Avecilla, S. T., Hattori, K., Heissig, B., Tejada, R., Liao, F., Shido, K., Jin, D. K., Dias, S., Zhang, F. & Hartman, T. E. 2004. Chemokine-mediated interaction of hematopoietic progenitors with the bone marrow vascular niche is required for thrombopoiesis. *Nature medicine,* 10**,** 64.

Balduini, A., Raslova, H., Di Buduo, C. A., Donada, A., Ballmaier, M., Germeshausen, M. & Balduini, C. L. 2018. Clinic, pathogenic mechanisms and drug testing of two inherited thrombocytopenias, ANKRD26-and MYH9-related diseases. *European journal of medical genetics*.

Balduini, C., Savoia, A. & Seri, M. 2013. Inherited thrombocytopenias frequently diagnosed in adults. *Journal of Thrombosis and Haemostasis,* 11**,** 1006-1019.

Balduini, C. L., Melazzini, F. & Pecci, A. 2017. Inherited thrombocytopenias—recent advances in clinical and molecular aspects.

Balduini, C. L., Pecci, A. & Savoia, A. 2011. Recent advances in the understanding and management of MYH9-related inherited thrombocytopenias. *British Journal of Haematology,* 154**,** 161-174.

Bamshad, M. J., Ng, S. B., Bigham, A. W., Tabor, H. K., Emond, M. J., Nickerson, D. A. & Shendure, J. 2011. Exome sequencing as a tool for Mendelian disease gene discovery. *Nature Reviews Genetics,* 12**,** 745-755.

Barbieri, I., Tzelepis, K., Pandolfini, L., Shi, J., Millán-Zambrano, G., Robson, S. C., Aspris, D., Migliori, V., Bannister, A. J. & Han, N. 2017. Promoter-bound METTL3 maintains myeloid leukaemia by m 6 A-dependent translation control. *Nature,* 552**,** 126-131.

Bardawil, T., Rebeiz, A., Chaabouni, M., El Halabi, J., Kambris, Z., Abbas, O., Abou Hassan, O., Hamie, L., Bitar, F. & Kibbi, A. G. 2017. Mutations in the ABCG8 gene are associated with sitosterolaemia in the homozygous form and xanthelasmas in the heterozygous form. *European Journal of Dermatology,* 27**,** 519-523.

Baruch, D., Gormely, J., Ma, C., Howard, R. & Pasloske, B. 1996. Plasmodium falciparum erythrocyte membrane protein 1 is a parasitized erythrocyte receptor for adherence to CD36, thrombospondin, and intercellular adhesion molecule 1. *Proceedings of the National Academy of Sciences,* 93**,** 3497-3502.

Baruch, D. I., Ma, X. C., Pasloske, B., Howard, R. J. & Miller, L. H. 1999. CD36 peptides that block cytoadherence define the CD36 binding region for Plasmodium falciparum-infected erythrocytes. *Blood, The Journal of the American Society of Hematology,* 94**,** 2121-2127.

Bastida, J. M., Benito, R., González-Porras, J. R. & Rivera, J. 2020. ABCG5 and ABCG8 gene variations associated with sitosterolemia and platelet dysfunction. *Platelets***,** 1-5.

Bastida, J. M., Benito, R., Janusz, K., Díez-Campelo, M., Hernández-Sánchez, J. M., Marcellini, S., Girós, M., Rivera, J., Lozano, M. L., Hortal, A., Hernández-Rivas, J. M. & González-Porras, J. R. 2017. Two novel variants of the ABCG5 gene cause xanthelasmas and macrothrombocytopenia: a brief review of hematologic abnormalities of sitosterolemia. *Journal of Thrombosis and Haemostasis,* 15**,** 1859-1866.

Bastida, J. M., Lozano, M. L., Benito, R., Janusz, K., Palma-Barqueros, V., Del Rey, M., Hernández-Sánchez, J. M., Riesco, S., Bermejo, N. & González-García, H. 2018. Introducing high-throughput sequencing into mainstream genetic diagnosis practice in inherited platelet disorders. *haematologica,* 103**,** 148-162.

Beaulieu, C. L., Majewski, J., Schwartzentruber, J., Samuels, M. E., Fernandez, B. A., Bernier, F. P., Brudno, M., Knoppers, B., Marcadier, J. & Dyment, D. 2014. FORGE Canada Consortium: outcomes of a 2-year national rare-disease gene-discovery project. *The American Journal of Human Genetics,* 94**,** 809-817.

Bender, M., Thon, J. N., Ehrlicher, A. J., Wu, S., Mazutis, L., Deschmann, E., Sola-Visner, M., Italiano, J. E. & Hartwig, J. H. 2015. Microtubule sliding drives proplatelet elongation and is dependent on cytoplasmic dynein.

Bennett, R. L., Motulsky, A. G., Bittles, A., Hudgins, L., Uhrich, S., Doyle, D. L., Silvey, K., Scott, C. R., Cheng, E. & McGillivray, B. 2002. Genetic counseling and screening of consanguineous

couples and their offspring: recommendations of the National Society of Genetic Counselors. *Journal of genetic counseling,* 11**,** 97-119.
Bertino, A., Qi, X., Li, J., Xia, Y. & Kuter, D. 2003. Apoptotic markers are increased in platelets stored at 37 C. *Transfusion,* 43**,** 857-866.
Besancenot, R., Roos-Weil, D., Tonetti, C., Abdelouahab, H., Lacout, C., Pasquier, F., Willekens, C., Rameau, P., Lecluse, Y. & Micol, J.-B. 2014. JAK2 and MPL protein levels determine TPO-induced megakaryocyte proliferation vs differentiation. *Blood,* 124**,** 2104-2115.
Bittles, A. H. 2001. Consanguinity and its relevance to clinical genetics. *Clinical genetics,* 60**,** 89-98.
Bittles, A. H. & Black, M. L. 2010a. Consanguinity, human evolution, and complex diseases. *Proceedings of the National Academy of Sciences,* 107**,** 1779-1786.
Bittles, A. H. & Black, M. L. 2010b. The impact of consanguinity on neonatal and infant health. *Early human development,* 86**,** 737-741.
Bizzozero, J. 1882. Ueber einen neuen Formbestandtheil des Blutes und dessen Rolle bei der Thrombose und der Blutgerinnung. *Archiv für pathologische Anatomie und Physiologie und für klinische Medicin,* 90**,** 261-332.
Bizzozero, J. 1982. *On a new blood particle and its role in thrombosis and blood coagulation*, H. Huber.
Blair, P. & Flaumenhaft, R. 2009. Platelet α-granules: basic biology and clinical correlates. *Blood reviews,* 23**,** 177-189.
Blanchette, V. S., Sparling, C. & Turner, C. 1991. 2 Inherited bleeding disorders. *Bailliere's clinical haematology,* 4**,** 291-332.
Bluteau, D., Balduini, A., Balayn, N., Currao, M., Nurden, P., Deswarte, C., Leverger, G., Noris, P., Perrotta, S., Solary, E., Vainchenker, W., Debili, N., Favier, R. & Raslova, H. 2014. Thrombocytopenia-associated mutations in the ANKRD26 regulatory region induce MAPK hyperactivation. *Journal of Clinical Investigation,* 124**,** 580-591.
Bluteau, D., Lordier, L., Di Stefano, A., Chang, Y., Raslova, H., Debili, N. & Vainchenker, W. 2009. Regulation of megakaryocyte maturation and platelet formation. *Journal of Thrombosis and Haemostasis,* 7**,** 227-234.
Bolton-Maggs, P. H., Chalmers, E. A., Collins, P. W., Harrison, P., Kitchen, S., Liesner, R. J., Minford, A., Mumford, A. D., Parapia, L. A. & Perry, D. J. 2006. A review of inherited platelet disorders with guidelines for their management on behalf of the UKHCDO. *British journal of haematology,* 135**,** 603-633.
Born, G. V. R. 1962. Aggregation of blood platelets by adenosine diphosphate and its reversal. *Nature,* 194**,** 927-929.
Botstein, D. & Risch, N. 2003. Discovering genotypes underlying human phenotypes: past successes for mendelian disease, future approaches for complex disease. *Nature genetics,* 33**,** 228-237.
Boycott, K. M., Rath, A., Chong, J. X., Hartley, T., Alkuraya, F. S., Baynam, G., Brookes, A. J., Brudno, M., Carracedo, A. & den Dunnen, J. T. 2017. International cooperation to enable the diagnosis of all rare genetic diseases. *The American Journal of Human Genetics,* 100**,** 695-705.
Boycott, K. M., Vanstone, M. R., Bulman, D. E. & MacKenzie, A. E. 2013. Rare-disease genetics in the era of next-generation sequencing: discovery to translation. *Nature Reviews Genetics,* 14**,** 681-691.
Boztug, K., Schmidt, M., Schwarzer, A., Banerjee, P. P., Díez, I. A., Dewey, R. A., Böhm, M., Nowrouzi, A., Ball, C. R. & Glimm, H. 2010. Stem-cell gene therapy for the Wiskott–Aldrich syndrome. *New England Journal of Medicine,* 363**,** 1918-1927.
Brass, L. F., Zhu, L. & Stalker, T. J. 2005. Minding the gaps to promote thrombus growth and stability. *The Journal of clinical investigation,* 115**,** 3385-3392.
Brewer, D. B. 2006. Max Schultze (1865), G. Bizzozero (1882) and the discovery of the platelet. *British journal of haematology,* 133**,** 251-258.

Brogna, S. & Wen, J. 2009. Nonsense-mediated mRNA decay (NMD) mechanisms. *Nature structural & molecular biology,* 16**,** 107.

Bruno, I. G., Karam, R., Huang, L., Bhardwaj, A., Lou, C. H., Shum, E. Y., Song, H.-W., Corbett, M. A., Gifford, W. D. & Gecz, J. 2011. Identification of a microRNA that activates gene expression by repressing nonsense-mediated RNA decay. *Molecular cell,* 42**,** 500-510.

Burk, C., Newman, P., Lyman, S., Gill, J., Coller, B. & Poncz, M. 1991. A deletion in the gene for glycoprotein IIb associated with Glanzmann's thrombasthenia. *The Journal of clinical investigation,* 87**,** 270-276.

Bussel, J. B. & Primiani, A. 2008. Fetal and neonatal alloimmune thrombocytopenia: progress and ongoing debates. *Blood reviews,* 22**,** 33-52.

Cabrera, A., Neculai, D. & Kain, K. C. 2014. CD36 and malaria: friends or foes? A decade of data provides some answers. *Trends in parasitology,* 30**,** 436-444.

Calvi, L., Adams, G., Weibrecht, K., Weber, J., Olson, D., Knight, M., Martin, R., Schipani, E., Divieti, P. & Bringhurst, F. R. 2003. Osteoblastic cells regulate the haematopoietic stem cell niche. *Nature,* 425**,** 841-846.

Cattaneo, M. 2002. Desmopressin in the treatment of patients with defects of platelet function. *Haematologica,* 87**,** 1122-1124.

Cattaneo, M. 2003. Inherited platelet-based bleeding disorders. *Journal of Thrombosis and Haemostasis,* 1**,** 1628-1636.

Cattaneo, M. 2011. Molecular defects of the platelet P2 receptors. *Purinergic signalling,* 7**,** 333-339.

Cattaneo, M., Cerletti, C., Harrison, P., Hayward, C., Kenny, D., Nugent, D., Nurden, P., Rao, A., Schmaier, A. & Watson, S. 2013. Recommendations for the standardization of light transmission aggregometry: a consensus of the working party from the platelet physiology subcommittee of SSC/ISTH. *Journal of Thrombosis and Haemostasis,* 11**,** 1183-1189.

Cattaneo, M., Lecchi, A., Randi, A. M., McGregor, J. L. & Mannucci, P. M. 1992. Identification of a new congenital defect of platelet function characterized by severe impairment of platelet responses to adenosine diphosphate.

Chen, H., Locke, D., Liu, Y., Liu, C. & Kahn, M. L. 2002. The platelet receptor GPVI mediates both adhesion and signaling responses to collagen in a receptor density-dependent fashion. *Journal of Biological Chemistry,* 277**,** 3011-3019.

Chen, K., Febbraio, M., Li, W. & Silverstein, R. L. 2008. A specific CD36-dependent signaling pathway is required for platelet activation by oxidized low-density lipoprotein. *Circulation research,* 102**,** 1512-1519.

Chen, K., Li, W., Major, J., Rahaman, S. O., Febbraio, M. & Silverstein, R. L. 2011. Vav guanine nucleotide exchange factors link hyperlipidemia and a prothrombotic state. *Blood, The Journal of the American Society of Hematology,* 117**,** 5744-5750.

Chen, L., Kostadima, M., Martens, J. H., Canu, G., Garcia, S. P., Turro, E., Downes, K., Macaulay, I. C., Bielczyk-Maczynska, E. & Coe, S. 2014. Transcriptional diversity during lineage commitment of human blood progenitors. *Science,* 345.

Cheng, H., Zheng, Z. & Cheng, T. 2020. New paradigms on hematopoietic stem cell differentiation. *Protein & cell,* 11**,** 34-44.

Chi, C. & Kadir, R. A. 2012. Inherited bleeding disorders in pregnancy. *Best practice & research Clinical obstetrics & gynaecology,* 26**,** 103-117.

Choi, Y., Sims, G. E., Murphy, S., Miller, J. R. & Chan, A. P. 2012. Predicting the functional effect of amino acid substitutions and indels. *PloS one,* 7**,** e46688.

Chute, J. P., Muramoto, G. G., Salter, A. B., Meadows, S. K., Rickman, D. W., Chen, B., Himburg, H. A. & Chao, N. J. 2007. Transplantation of vascular endothelial cells mediates the hematopoietic recovery and survival of lethally irradiated mice. *Blood,* 109**,** 2365-2372.

Clemetson, K. J. 2012. Platelets and primary haemostasis. *Thrombosis research,* 129**,** 220-224.

Clemetson, K. J., Pfueller, S. L., Luscher, E. F. & Jenkins, C. S. 1977. Isolation of the membrane glycoproteins of human blood platelets by lectin affinity chromatography. *Biochimica et Biophysica Acta (BBA)-Biomembranes,* 464**,** 493-508.

Coller, B. 1984. Bizzozero and the discovery of the blood platelet. *The Lancet,* 323**,** 804.

Consortium, F. 1999. The FlyBase database of the Drosophila Genome Projects and community literature. *Nucleic Acids Research,* 27**,** 85-88.

Consortium, I. H. G. S. 2004. Finishing the euchromatic sequence of the human genome. *Nature,* 431**,** 931.

Cooney, K. A., Nichols, W. C., Bruck, M. E., Bahou, W. F., Shapiro, A. D., Bowie, E. J. W., Gralnick, H. R. & Ginsburg, D. 1991. The Molecular Defect in Type IIB von Willebrand Disease. *J. Clin. Invest.***,** 1227-1233.

Curtis, B. & Aster, R. 1996. Incidence of the Nak (a)-negative platelet phenotype in African Americans is similar to that of Asians. *Transfusion,* 36**,** 331-334.

Curtis, B. R., Ali, S., Glazier, A. M., Ebert, D. D., Aitman, T. J. & Aster, R. H. 2002. Isoimmunization against CD36 (glycoprotein IV): description of four cases of neonatal isoimmune thrombocytopenia and brief review of the literature. *Transfusion,* 42**,** 1173-1179.

Cutting, G. R. 2015. Cystic fibrosis genetics: from molecular understanding to clinical application. *Nature Reviews Genetics,* 16**,** 45-56.

D'Andrea, G., Chetta, M. & Margaglione, M. 2009. Inherited platelet disorders: thrombocytopenias and thrombocytopathies. *Blood Transfusion,* 7**,** 278.

Daniel, J. L., Dangelmaier, C., Strouse, R. & Smith, J. B. 1994. Collagen induces normal signal transduction in platelets deficient in CD36 (platelet glycoprotein IV). *Thrombosis and haemostasis,* 72**,** 353-356.

Dargaud, Y., Scoazec, J. Y., Wielders, S. J., Trzeciak, C., Hackeng, T. M., Négrier, C., Hemker, H. C., Lindhout, T. & Castoldi, E. 2015. Characterization of an autosomal dominant bleeding disorder caused by a thrombomodulin mutation. *Blood, The Journal of the American Society of Hematology,* 125**,** 1497-1501.

Dawood, B. B., Lowe, G. C., Lordkipanidzé, M., Bem, D., Daly, M. E., Makris, M., Mumford, A., Wilde, J. T. & Watson, S. P. 2012. Evaluation of participants with suspected heritable platelet function disorders including recommendation and validation of a streamlined agonist panel. *Blood, The Journal of the American Society of Hematology,* 120**,** 5041-5049.

Dawood, B. B., Wilde, J. & Watson, S. P. 2007. Reference curves for aggregation and ATP secretion to aid diagnose of platelet-based bleeding disorders: Effect of inhibition of ADP and thromboxane A2 pathways. *Platelets,* 18**,** 329-345.

Denis, M. M., Tolley, N. D., Bunting, M., Schwertz, H., Jiang, H., Lindemann, S., Yost, C. C., Rubner, F. J., Albertine, K. H. & Swoboda, K. J. 2005. Escaping the nuclear confines: signal-dependent pre-mRNA splicing in anucleate platelets. *Cell,* 122**,** 379-391.

Desai, A. N. & Jere, A. 2012. Next-generation sequencing: ready for the clinics? *Clinical genetics,* 81**,** 503-510.

Desai, S. R., Korula, A., Kulkarni, U. P., Menon, A. A., Ramachandran, S. V., Sindhuvi, E., Nellickal, A. J., Nair, S. C. & George, B. 2021. Sitosterolemia: Four Cases of an Uncommon Cause of Hemolytic Anemia (Mediterranean Stomatocytosis with Macrothrombocytopenia). *Indian Journal of Hematology and Blood Transfusion,* 37**,** 157-161.

Diagouraga, B., Grichine, A., Fertin, A., Wang, J., Khochbin, S. & Sadoul, K. 2014. Motor-driven marginal band coiling promotes cell shape change during platelet activation. *Journal of Cell Biology,* 204**,** 177-185.

Diaz-Ricart, M., Tandon, N., Carretero, M., Ordinas, A., Bastida, E. & Jamieson, G. 1993. Platelets lacking functional CD36 (glycoprotein IV) show reduced adhesion to collagen in flowing whole blood.

Dorahy, D. J., Berndt, M. C., Shafren, D. R. & Burns, G. F. 1996. CD36 is spatially associated with glycoprotein IIb-IIIa (αIIbβ3) on the surface of resting platelets. *Biochemical and biophysical research communications,* 218**,** 575-581.

Downes, K., Megy, K., Duarte, D., Vries, M., Gebhart, J., Hofer, S., Shamardina, O., Deevi, S. V., Stephens, J. & Mapeta, R. 2019. Diagnostic high-throughput sequencing of 2396 patients with bleeding, thrombotic, and platelet disorders. *Blood,* 134**,** 2082-2091.

Drachman, J. G. 2004a. Inherited thrombocytopenia: when a low platelet count does not mean ITP. *Blood,* 103**,** 390-398.

Drachman, J. G. 2004b. Review in translational hematology Inherited thrombocytopenia : when a low platelet count does not mean ITP. *Blood,* 103**,** 390-398.

Drachman, J. G., Griffin, J. D. & Kaushansky, K. 1995. The c-Mpl ligand (thrombopoietin) stimulates tyrosine phosphorylation of Jak2, Shc, and c-Mpl. *Journal of Biological Chemistry,* 270**,** 4979-4982.

Duke, W. 1910. The relation of blood platelets to hemorrhagic disease: description of a method for determining the bleeding time and coagulation time and report of three cases of hemorrhagic disease relieved by transfusion. *Journal of the American Medical Association,* 55**,** 1185-1192.

Durrant, T. N., van den Bosch, M. T. & Hers, I. 2017. Integrin αIIbβ3 outside-in signaling. *Blood,* 130**,** 1607-1619.

Erzurumluoglu, A. M., Shihab, H. A., Rodriguez, S., Gaunt, T. R. & Day, I. N. 2016. Importance of genetic studies in consanguineous populations for the characterization of novel human gene functions. *Annals of human genetics,* 80**,** 187-196.

Favier, R., Aoki, N. & De Moerloose, P. 2001. Congenital α2-plasmin inhibitor deficiencies: a review. *British journal of haematology,* 114**,** 4-10.

Febbraio, M., Hajjar, D. P. & Silverstein, R. L. 2001. CD36: a class B scavenger receptor involved in angiogenesis, atherosclerosis, inflammation, and lipid metabolism. *The Journal of clinical investigation,* 108**,** 785-791.

Fernandez-Marmiesse, A., Gouveia, S. & Couce, M. L. 2018. NGS technologies as a turning point in rare disease research, diagnosis and treatment. *Current Medicinal Chemistry,* 25**,** 404-432.

Ferreira, F., Colella, M., Medina, S., Costa-Lima, C., Fiusa, M., Costa, L., Orsi, F., Annichino-Bizzacchi, J., Fertrin, K. & Gilberti, M. 2017. Evaluation of the immature platelet fraction contribute to the differential diagnosis of hereditary, immune and other acquired thrombocytopenias. *Scientific reports,* 7**,** 1-8.

Fisher, C. L. & Pei, G. K. 1997. Modification of a PCRBased Site-Directed Mutagenesis Method. *Biotechniques,* 23**,** 570-574.

Fletcher, S. J., Johnson, B., Lowe, G. C., Bem, D., Drake, S., Lordkipanidzé, M., Guiú, I. S., Dawood, B., Rivera, J. & Simpson, M. A. 2015. SLFN14 mutations underlie thrombocytopenia with excessive bleeding and platelet secretion defects. *The Journal of clinical investigation,* 125**,** 3600-3605.

Folman, C. C., Linthorst, G. E., van Mourik, J., van Willigen, G., de Jonge, E., Levi, M., de Haas, M. & von dem Borne, A. E. 2000. Platelets release thrombopoietin (Tpo) upon activation: another regulatory loop in thrombocytopoiesis? *Thrombosis and haemostasis,* 83**,** 923-930.

Foudi, A., Kramer, D. J., Qin, J., Ye, D., Behlich, A.-S., Mordecai, S., Preffer, F. I., Amzallag, A., Ramaswamy, S. & Hochedlinger, K. 2014. Distinct, strict requirements for Gfi-1b in adult bone marrow red cell and platelet generation. *Journal of Experimental Medicine,* 211**,** 909-927.

Fratantoni, J. C. & Poindexter, B. J. 1990. Measuring platelet aggregation with microplate reader: a new technical approach to platelet aggregation studies. *American journal of clinical pathology,* 94**,** 613-617.

Freson, K., Wijgaerts, A. & Geet, C. V. 2017. GATA1 gene variants associated with thrombocytopenia and anemia.

Fujino, H., Ohta, K., Taniue, A., Nagao, N., Hino, M., Yamane, T., Koh, K. R., Takeoka, Y., Hirose, A. & Aoyama, Y. 2001. Primary refractoriness to platelet transfusion caused by Naka antibody alone. *Vox sanguinis,* 81**,** 42-44.

Furie, B. & Furie, B. C. 2008. Mechanisms of thrombus formation. *New England Journal of Medicine,* 359**,** 938-949.

Futterer, J., Dalby, A., Lowe, G. C., Johnson, B., Simpson, M. A., Motwani, J., Williams, M., Watson, S. P. & Morgan, N. V. 2018. Mutation in GNE is associated with severe congenital thrombocytopenia. *Blood,* 132**,** 1855-1858.

Geddis, A. E. 2010. Megakaryopoiesis. *Seminars in Hematology,* 47**,** 212-219.

Germeshausen, M., Ancliff, P., Estrada, J., Metzler, M., Ponstingl, E., Schwabe, D., Scott, R. H., Unal, S., Wawer, A., Zeller, B. & Ballmaier, M. 2018. MECOM-associated syndrome: a heterogeneous inherited bone marrow failure syndrome with amegakaryocytic thrombocytopenia.

Ghevaert, C., Salsmann, A., Watkins, N. A., Schaffner-Reckinger, E., Rankin, A., Garner, S. F., Stephens, J., Smith, G. A., Debili, N. & Vainchenker, W. 2008. A nonsynonymous SNP in the ITGB3 gene disrupts the conserved membrane-proximal cytoplasmic salt bridge in the αIIbβ3 integrin and cosegregates dominantly with abnormal proplatelet formation and macrothrombocytopenia. *Blood,* 111**,** 3407-3414.

Ghosh, A., Murugesan, G., Chen, K., Zhang, L., Wang, Q., Febbraio, M., Anselmo, R. M., Marchant, K., Barnard, J. & Silverstein, R. L. 2011. Platelet CD36 surface expression levels affect functional responses to oxidized LDL and are associated with inheritance of specific genetic polymorphisms. *Blood,* 117**,** 6355-6366.

Ghoshal, K. & Bhattacharyya, M. 2014. Overview of platelet physiology: its hemostatic and nonhemostatic role in disease pathogenesis. *The Scientific World Journal,* 2014.

Gibbins, J. M. 2004. Platelet adhesion signalling and the regulation of thrombus formation. *Journal of cell science,* 117**,** 3415-3425.

Gieger, C., Radhakrishnan, A., Cvejic, A., Tang, W., Porcu, E., Pistis, G., Serbanovic-Canic, J., Elling, U., Goodall, A. H. & Labrune, Y. 2011. New gene functions in megakaryopoiesis and platelet formation. *Nature,* 480**,** 201-208.

Glatz, J. F., Luiken, J. J. & Bonen, A. 2010. Membrane fatty acid transporters as regulators of lipid metabolism: implications for metabolic disease. *Physiological reviews,* 90**,** 367–417.

Goodwin, S., McPherson, J. D. & McCombie, W. R. 2016. Coming of age: ten years of next-generation sequencing technologies. *Nature Reviews Genetics,* 17**,** 333.

Gordge, M. P. 2005. Megakaryocyte apoptosis: sorting out the signals. *British journal of pharmacology,* 145**,** 271.

Gordon, N., Thom, J., Cole, C. & Baker, R. 1995. Rapid detection of hereditary and acquired platelet storage pool deficiency by flow cytometry. *British journal of haematology,* 89**,** 117-123.

Greenwalt, D. E., Lipsky, R. H., Ockenhouse, C. F., Ikeda, H., Tandon, N. N. & Jamieson, G. 1992. Membrane glycoprotein CD36: a review of its roles in adherence, signal transduction, and transfusion medicine. *Blood,* 80**,** 1105-1115.

Greinacher, A. & Eekels, J. J. 2019a. Simplifying the diagnosis of inherited platelet disorders? The new tools do not make it any easier. *Blood,* 133**,** 2478-2483.

Greinacher, A. & Eekels, J. J. M. 2019b. Diagnosis of hereditary platelet disorders in the era of next-generation sequencing: “primum non nocere”. *Journal of Thrombosis and Haemostasis,* 17**,** 551-554.

Gresele, P., Bury, L., Mezzasoma, A. M. & Falcinelli, E. 2019. Platelet function assays in diagnosis: an update. *Expert review of hematology,* 12**,** 29-46.

Gresele, P., Harrison, P., Bury, L., Falcinelli, E., Gachet, C., Hayward, C., Kenny, D., Mezzano, D., Mumford, A. & Nugent, D. 2014. Diagnosis of suspected inherited platelet function disorders: results of a worldwide survey. *Journal of Thrombosis and Haemostasis,* 12**,** 1562-1569.

Guo, T., Wang, X., Qu, Y., Yin, Y., Jing, T. & Zhang, Q. 2015a. Megakaryopoiesis and platelet production: insight into hematopoietic stem cell proliferation and differentiation. *Stem cell investigation,* 2**,** 3-3.

Guo, T., Wang, X., Qu, Y., Yin, Y., Jing, T. & Zhang, Q. 2015b. Megakaryopoiesis and platelet production: insight into hematopoietic stem cell proliferation and differentiation. *stem cell investigation,* 2.

Hall, N. 2007. Advanced sequencing technologies and their wider impact in microbiology. *Journal of Experimental Biology,* 210**,** 1518-1525.

Hamamy, H. 2012. Consanguineous marriages. *Journal of community genetics,* 3**,** 185-192.

Hamamy, H., Antonarakis, S. E., Cavalli-Sforza, L. L., Temtamy, S., Romeo, G., Ten Kate, L. P., Bennett, R. L., Shaw, A., Megarbane, A. & van Duijn, C. 2011. Consanguineous marriages, pearls and perils: Geneva international consanguinity workshop report. *Genetics in Medicine,* 13**,** 841-847.

Hamamy, H. A., Masri, A. T., Al-Hadidy, A. M. & Ajlouni, K. M. 2007. Consanguinity and genetic disorders. Profile from Jordan. *Saudi medical journal,* 28**,** 1015-1017.

Hardisty, R., Pidard, D., Cox, A., Nokes, T., Legrand, C., Bouillot, C., Pannocchia, A., Heilmann, E., Hourdille, P. & Bellucci, S. 1992. A defect of platelet aggregation associated with an abnormal distribution of glycoprotein IIb-IIIa complexes within the platelet: the cause of a lifelong bleeding disorder.

Harrison, P. 2005a. Platelet function analysis. *Blood reviews,* 19**,** 111-123.

Harrison, P. 2005b. The role of PFA-100® testing in the investigation and management of haemostatic defects in children and adults. *British journal of haematology,* 130**,** 3-10.

Harrison, P. & Cramer, E. M. 1993. Platelet α-granules. *Blood reviews,* 7**,** 52-62.

Harrison, P., Mackie, I., Mumford, A., Briggs, C. & Liesner, R. 2011. winter M., Machin S.; British Committee for Standards in Haematology. Guidelines for the laboratory investigation of heritable disorders of platelet function. *Br J Haematol,* 155**,** 30-44.

Hartwig, J. H. & DeSisto, M. 1991. The cytoskeleton of the resting human blood platelet: structure of the membrane skeleton and its attachment to actin filaments. *The Journal of cell biology,* 112**,** 407-425.

Hassan, A. A. & Kroll, M. H. 2005. Acquired disorders of platelet function. *Hematology,* 2005**,** 403-408.

Hatje, K., Mühlhausen, S., Simm, D. & Kollmar, M. 2019. The Protein-Coding Human Genome: Annotating High-Hanging Fruits. *BioEssays,* 41**,** 1900066.

Heemskerk, J. W., Bevers, E. M. & Lindhout, T. 2002. Platelet activation and blood coagulation. *Thrombosis and haemostasis,* 88**,** 186-193.

Heng, H. H. 2009. The genome-centric concept: resynthesis of evolutionary theory. *Bioessays,* 31**,** 512-525.

Hermansky, F. & Pudlak, P. 1959. Albinism associated with hemorrhagic diathesis and unusual pigmented reticular cells in the bone marrow: report of two cases with histochemical studies. *Blood,* 14**,** 162-169.

Herold, T., Metzeler, K. H., Vosberg, S., Hartmann, L., Röllig, C., Stölzel, F., Schneider, S., Hubmann, M., Zellmeier, E. & Ksienzyk, B. 2014. Isolated trisomy 13 defines a homogeneous AML subgroup with high frequency of mutations in spliceosome genes and poor prognosis. *Blood, The Journal of the American Society of Hematology,* 124**,** 1304-1311.

Hobisch-Hagen, P., Jelkmann, W., Mayr, A., Wiedermann, F., Fries, D., Herold, M., Klingler, A. & Schobersberger, W. 2000. Low platelet count and elevated serum thrombopoietin after severe trauma. *European journal of haematology,* 64**,** 157-163.

Holinstat, M. 2017. Normal platelet function. *Cancer and Metastasis Reviews,* 36**,** 195-198.

Hollopeter, G., Jantzen, H.-M., Vincent, D., Li, G., England, L., Ramakrishnan, V., Yang, R.-B., Nurden, P., Nurden, A. & Julius, D. 2001. Identification of the platelet ADP receptor targeted by antithrombotic drugs. *Nature,* 409**,** 202-207.

Holmsen, H., Storm, E. & Day, H. J. 1972. Determination of ATP and ADP in blood platelets: a modification of the firefly luciferase assay for plasma. *Analytical Biochemistry,* 46**,** 489-501.
Horvat-Switzer, R. D. & Thompson, A. A. 2006. HOXA11 mutation in amegakaryocytic thrombocytopenia with radio-ulnar synostosis syndrome inhibits megakaryocytic differentiation in vitro. *Blood Cells, Molecules, and Diseases,* 37**,** 55-63.
Hosokawa, K., Ohnishi, T., Fukasawa, M., Kondo, T., Sameshima, H., Koide, T., Tanaka, K. A. & Maruyama, I. 2012. A microchip flow-chamber system for quantitative assessment of the platelet thrombus formation process. *Microvascular research,* 83**,** 154-161.
Hsieh, F.-L., Turner, L., Bolla, J. R., Robinson, C. V., Lavstsen, T. & Higgins, M. K. 2016. The structural basis for CD36 binding by the malaria parasite. *Nature communications,* 7**,** 1-11.
Huang, W., Febbraio, M. & Silverstein, R. L. 2011. CD9 tetraspanin interacts with CD36 on the surface of macrophages: a possible regulatory influence on uptake of oxidized low density lipoprotein. *PLoS One,* 6**,** e29092.
Hubacek, J. A., Berge, K. E., Cohen, J. C. & Hobbs, H. H. 2001. Mutations in ATP-cassette binding proteins G5 (ABCG5) and G8 (ABCG8) causing sitosterolemia. *Human mutation,* 18**,** 359-360.
Huizing, M., Malicdan, M. C., Wang, J. A., Pri-Chen, H., Hess, R. A., Fischer, R., O'Brien, K. J., Merideth, M. A., Gahl, W. A. & Gochuico, B. R. 2020. Hermansky–Pudlak syndrome: mutation update. *Human mutation,* 41**,** 543-580.
Hurtado, B., Trakala, M., Ximénez-Embún, P., Bakkali, A. E., Partida, D., Sanz-Castillo, B., Álvarez-Fernández, M., Maroto, M., Sánchez-Martínez, R., Martínez, L., Muñoz, J., de Frutos, P. G. & Malumbres, M. 2018. Thrombocytopenia-associated mutations in Ser/Thr kinase MASTL deregulate actin cytoskeletal dynamics in platelets. *Journal of Clinical Investigation,* 128**,** 5351-5367.
Ihara, K., Ishii, E., Eguchi, M., Takada, H., Suminoe, A., Good ¶, R. A. & Hara, T. 1999. Identification of mutations in the c-mpl gene in congenital amegakaryocytic thrombocytopenia.
Iourov, I. Y., Vorsanova, S. G. & Yurov, Y. B. 2019. Pathway-based classification of genetic diseases. *Molecular cytogenetics,* 12**,** 1-5.
Israels, S. J., Kahr, W. H., Blanchette, V. S., Luban, N. L., Rivard, G. E. & Rand, M. L. 2011. Platelet disorders in children: a diagnostic approach. *Pediatric blood & cancer,* 56**,** 975-983.
Italiano Jr, J. E., Lecine, P., Shivdasani, R. A. & Hartwig, J. H. 1999. Blood platelets are assembled principally at the ends of proplatelet processes produced by differentiated megakaryocytes. *The Journal of cell biology,* 147**,** 1299-1312.
Jaber, L., Halpern, G. J. & Shohat, M. 1998. The impact of consanguinity worldwide. *Public Health Genomics,* 1**,** 12-17.
Jackson, S. P. 2007. The growing complexity of platelet aggregation. *Blood,* 109**,** 5087-5095.
Jackson, S. P. 2011. Arterial thrombosis—insidious, unpredictable and deadly. *Nature medicine,* 17**,** 1423-1436.
Jelkmann, W. 2001. The role of the liver in the production of thrombopoietin compared with erythropoietin. *European journal of gastroenterology & hepatology,* 13**,** 791-801.
Johnson, B., Doak, R., Allsup, D., Astwood MbchB, E., Evans MbchB, G., Grimley, C., James MbchB, B., Myers MbchB, B., Stokley MbchB, S., Thachil MbchB, J., Wilde MbchB, J., Williams, M., Makris, M. M. A., Lowe Mrcp, G. C., Wallis, Y., Daly, M. E. & Morgan, N. V. 2018. A comprehensive targeted next-generation sequencing panel for genetic diagnosis of patients with suspected inherited thrombocytopenia.
Johnson, B., Fletcher, S. J. & Morgan, N. V. 2016a. Inherited thrombocytopenia: novel insights into megakaryocyte maturation, proplatelet formation and platelet lifespan. *Platelets,* 27**,** 519-525.
Johnson, B., Lowe, G. C., Futterer, J., Lordkipanidzé, M., Macdonald, D., Simpson, M. A., Sanchez-Guiú, I., Drake, S., Bem, D., Leo, V., Fletcher, S. J., Dawood, B., Rivera, J., Allsup, D., Biss, T., Bolton-Maggs, P. H., Collins, P., Curry, N., Grimley, C., James, B., Makris, M., Motwani, J., Pavord, S., Talks, K., Thachil, J., Wilde, J., Williams, M., Harrison, P., Gissen, P., Mundell, S.,

Mumford, A., Daly, M. E., Watson, S. P. & Morgan, N. V. 2016b. Whole exome sequencing identifies genetic variants in inherited thrombocytopenia with secondary qualitative function defects. *Haematologica,* 101**,** 1170-1179.

Jones, M., Murden, S., Bem, D., Mundell, S., Gissen, P., Daly, M., Watson, S. & Mumford, A. 2012. Rapid genetic diagnosis of heritable platelet function disorders with next-generation sequencing: proof-of-principle with Hermansky–Pudlak syndrome. *Journal of Thrombosis and Haemostasis,* 10**,** 306.

Junt, T., Schulze, H., Chen, Z., Massberg, S., Goerge, T., Krueger, A., Wagner, D. D., Graf, T., Italiano, J. E. & Shivdasani, R. A. 2007. Dynamic visualization of thrombopoiesis within bone marrow. *Science,* 317**,** 1767-1770.

Jurasz, P., Alonso-Escolano, D. & Radomski, M. W. 2004. Platelet–cancer interactions: mechanisms and pharmacology of tumour cell-induced platelet aggregation. *British journal of pharmacology,* 143**,** 819.

Kahr, W. H. A., Hinckley, J., Li, L., Schwertz, H., Christensen, H., Rowley, J. W., Pluthero, F. G., Urban, D., Fabbro, S., Nixon, B., Gadzinski, R., Storck, M., Wang, K., Ryu, G. Y., Jobe, S. M., Schutte, B. C., Moseley, J., Loughran, N. B., Parkinson, J., Weyrich, A. S. & Di Paola, J. 2011. Mutations in NBEAL2, encoding a BEACH protein, cause gray platelet syndrome. *Nature Genetics,* 43**,** 738-740.

Kamphuis, M. M., Paridaans, N. P., Porcelijn, L., Lopriore, E. & Oepkes, D. 2014. Incidence and consequences of neonatal alloimmune thrombocytopenia: a systematic review. *Pediatrics,* 133**,** 715-721.

Kankirawatana, S., Kupatawintu, P., Juji, T., Veerakul, G., Ngerncham, S., Chongkolwatana, V. & O'Charoen, R. 2001. Neonatal alloimmune thrombocytopeniadue to anti-Naka. *Transfusion,* 41**,** 375-377.

Kaplan, C. 2002. Alloimmune thrombocytopenia of the fetus and the newborn. *Blood reviews,* 16**,** 69-72.

Kar, N. S., Ashraf, M. Z., Valiyaveettil, M. & Podrez, E. A. 2008. Mapping and characterization of the binding site for specific oxidized phospholipids and oxidized low density lipoprotein of scavenger receptor CD36. *Journal of Biological Chemistry,* 283**,** 8765-8771.

Kaushansky, K. 1998. Thrombopoietin. *New England Journal of Medicine,* 339**,** 746-754.

Kaushansky, K., Lok, S., Holly, R. D., Broudy, V. C., Lin, N., Bailey, M. C., Forstrom, J. W., Buddle, M. M., Oort, P. J. & Hagen, F. S. 1994. Promotion of megakaryocyte progenitor expansion and differentiation by the c-Mpl ligand thrombopoietin. *Nature,* 369**,** 568.

Kearney, H. M., Thorland, E. C., Brown, K. K., Quintero-Rivera, F. & South, S. T. 2011. American College of Medical Genetics standards and guidelines for interpretation and reporting of postnatal constitutional copy number variants. *Genetics in Medicine,* 13**,** 680.

Kehrel, B., Wierwille, S., Clemetson, K. J., Anders, O., Steiner, M., Graham Knight, C., Farndale, R. W., Okuma, M. & Barnes, M. J. 1998. Glycoprotein VI is a major collagen receptor for platelet activation: it recognizes the platelet-activating quaternary structure of collagen, whereas CD36, glycoprotein IIb/IIIa, and von Willebrand factor do not. *Blood, The Journal of the American Society of Hematology,* 91**,** 491-499.

Kellis, M., Wold, B., Snyder, M. P., Bernstein, B. E., Kundaje, A., Marinov, G. K., Ward, L. D., Birney, E., Crawford, G. E. & Dekker, J. 2014. Defining functional DNA elements in the human genome. *Proceedings of the National Academy of Sciences,* 111**,** 6131-6138.

Khan, A., Khan, I., Zahid, K. & Nabi, G. A Comprehensive Review on Various Aspects of Genetic Disorders. *Editorial Team***,** 110.

Khan, A., Khan, I., Zahid, K. & Nabi, G. 2015. A Comprehensive Review on Various Aspects of Genetic Disorders. *Editorial Team***,** 110.

Khan, A. O., Maclachlan, A., Lowe, G. C., Nicolson, P. L., Al Ghaithi, R., Thomas, S. G., Watson, S. P., Pike, J. A. & Morgan, N. V. 2020. High-throughput platelet spreading analysis: a tool for the diagnosis of platelet-based bleeding disorders. *haematologica,* 105**,** e124.

Khoriaty, R., Ozel, A. B., Ramdas, S., Ross, C., Desch, K., Shavit, J. A., Everett, L., Siemieniak, D., Li, J. Z. & Ginsburg, D. 2019. Genome-wide linkage analysis and whole-exome sequencing identifies an ITGA 2B mutation in a family with thrombocytopenia. *British journal of haematology,* 186**,** 574-579.

Kiel, M. J., Yilmaz, Ö. H., Iwashita, T., Yilmaz, O. H., Terhorst, C. & Morrison, S. J. 2005. SLAM family receptors distinguish hematopoietic stem and progenitor cells and reveal endothelial niches for stem cells. *cell,* 121**,** 1109-1121.

Kim, S.-J., Davis, R. & Jenne, C. 2018. Platelets as Modulators of Inflammation. *Seminars in Thrombosis and Hemostasis,* 44**,** 091-101.

Kirchhausen, T. & Rosen, F. S. 1996. Disease mechanism: Unravelling Wiskott-Aldrich syndrome. *Current Biology,* 6**,** 676-678.

Kirsch-Volders, M., Vanhauwaert, A., De Boeck, M. & Decordier, I. 2002. Importance of detecting numerical versus structural chromosome aberrations. *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis,* 504**,** 137-148.

Kopp, H.-G., Avecilla, S. T., Hooper, A. T. & Rafii, S. 2005. The bone marrow vascular niche: home of HSC differentiation and mobilization. *Physiology,* 20**,** 349-356.

Koren, A., Khayat, M., Hauschner, H., Pretorius, E., Rosenberg, N., Zalman, L., Elpeleg, O., Salama, I., Shalev, S., Shenkman, B. & Levin, C. 2015. Deleterious mutation in the FYB gene is associated with congenital autosomal recessive small-platelet thrombocytopenia. *Journal of Thrombosis and Haemostasis,* 13**,** 1285-1292.

Kottke-Marchant, K. & Corcoran, G. 2002. The laboratory diagnosis of platelet disorders: an algorithmic approach. *Archives of pathology & laboratory medicine,* 126**,** 133-146.

Krawczak, M., Reiss, J. & Cooper, D. N. 1992. The mutational spectrum of single base-pair substitutions in mRNA splice junctions of human genes: causes and consequences. *Human genetics,* 90**,** 41-54.

Kumar, P., Henikoff, S. & Ng, P. C. 2009. Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. *Nature protocols,* 4**,** 1073.

Kundu, S. K., Heilmann, E. J., Sio, R., Garcia, C., Davidson, R. M. & Ostgaard, R. A. Description of an in vitro platelet function analyzer-PFA-100™. Seminars in thrombosis and hemostasis, 1995. Thieme Medical Publishers, 106-112.

Kunishima, S., Kobayashi, R., Itoh, T. J., Hamaguchi, M. & Saito, H. 2009. Mutation of the 1-tubulin gene associated with congenital macrothrombocytopenia affecting microtubule assembly.

Kunishima, S., Okuno, Y., Yoshida, K., Shiraishi, Y., Sanada, M., Muramatsu, H., Chiba, K., Tanaka, H., Miyazaki, K., Sakai, M., Ohtake, M., Kobayashi, R., Iguchi, A., Niimi, G., Otsu, M., Takahashi, Y., Miyano, S., Saito, H., Kojima, S. & Ogawa, S. 2013. ACTN1 mutations cause congenital macrothrombocytopenia. *American Journal of Human Genetics,* 92**,** 431-438.

Kuter, D. J. 2007. New thrombopoietic growth factors. *Blood, The Journal of the American Society of Hematology,* 109**,** 4607-4616.

Lacruz, R. S. & Feske, S. 2015. Diseases caused by mutations in ORAI1 and STIM1.

Lambert, M. P. 2019. Inherited platelet disorders: a modern approach to evaluation and treatment. *Hematology/Oncology Clinics,* 33**,** 471-487.

Lamolda, M., Montes, R., Simón, I., Perales, S., Martínez-Navajas, G., Lopez-Onieva, L., Ríos-Pelegrina, R., del Moral, R. G., Griñan-Lison, C. & Marchal, J. A. 2019. GENYOi005-A: An induced pluripotent stem cells (iPSCs) line generated from a patient with Familial Platelet Disorder with associated Myeloid Malignancy (FPDMM) carrying a p. Thr196Ala variant. *Stem cell research,* 41**,** 101603.

Lawler, J. 1986. The structural and functional properties of thrombospondin.

Lebois, M. & Josefsson, E. C. 2016. Regulation of platelet lifespan by apoptosis. *Platelets,* 27**,** 497-504.

Lecine, P., Italiano Jr, J. E., Kim, S.-W., Villeval, J.-L. & Shivdasani, R. A. 2000. Hematopoietic-specific β1 tubulin participates in a pathway of platelet biogenesis dependent on the transcription factor NF-E2. *Blood, The Journal of the American Society of Hematology,* 96**,** 1366-1373.

Leinøe, E., Zetterberg, E., Kinalis, S., Østrup, O., Kampmann, P., Norström, E., Andersson, N., Klintman, J., Qvortrup, K. & Nielsen, F. C. 2017. Application of whole-exome sequencing to direct the specific functional testing and diagnosis of rare inherited bleeding disorders in patients from the Öresund Region, Scandinavia. *British journal of haematology,* 179**,** 308-322.

Lek, M., Karczewski, K. J., Minikel, E. V., Samocha, K. E., Banks, E., Fennell, T., O'Donnell-Luria, A. H., Ware, J. S., Hill, A. J. & Cummings, B. B. 2016. Analysis of protein-coding genetic variation in 60,706 humans. *Nature,* 536**,** 285-291.

Lentaigne, C., Freson, K., Laffan, M. A., Turro, E. & Ouwehand, W. H. 2016. Inherited platelet disorders: toward DNA-based diagnosis. *Blood, The Journal of the American Society of Hematology,* 127**,** 2814-2823.

Lentaigne, C., Greene, D., Sivapalaratnam, S., Favier, R., Seyres, D., Thys, C., Grassi, L., Mangles, S., Sibson, K. & Stubbs, M. 2019. Germline mutations in the transcription factor IKZF5 cause thrombocytopenia. *Blood, The Journal of the American Society of Hematology,* 134**,** 2070-2081.

Leo, V., Morgan, N., Bem, D., Jones, M., Lowe, G., Lordkipanidzé, M., Drake, S., Simpson, M., Gissen, P. & Mumford, A. 2015. Use of next-generation sequencing and candidate gene analysis to identify underlying defects in patients with inherited platelet function disorders. *Journal of Thrombosis and Haemostasis,* 13**,** 643-650.

Lester, W. A., Guilliatt, A. M., Enayat, M. S., Rose, P. & Hill, F. G. 2007. The R2464C missense mutation in the von Willebrand factor gene causes a novel abnormality of multimer electrophoretic mobility and falls into the subgroup of type 2 von Willebrand disease'unclassified'. *Thrombosis and haemostasis,* 97**,** 159-160.

Letestu, R., Vitrat, N., Massé, A., Le Couedic, J. P., Lazar, V., Rameau, P., Wendling, F., Vuillier, J., Boutard, P., Plouvier, E., Plasse, M., Favier, R., Vainchenker, W. & Debili, N. 2000. Existence of a differentiation blockage at the stage of a megakaryocyte precursor in the thrombocytopenia and absent radii (TAR) syndrome. *Blood,* 95**,** 1633-1641.

Leung, L. 1984. Role of thrombospondin in platelet aggregation. *The Journal of clinical investigation,* 74**,** 1764-1772.

Leung, L., Li, W., McGregor, J., Albrecht, G. & Howard, R. 1992. CD36 peptides enhance or inhibit CD36-thrombospondin binding. A two-step process of ligand-receptor interaction. *Journal of Biological Chemistry,* 267**,** 18244-18250.

Leung, L. & Nachman, R. 1986. Molecular mechanisms of platelet aggregation. *Annual review of medicine,* 37**,** 179-186.

Levy, G. G., Nichols[2], W. C., Lian[3], E. C., Foroud, T., McClintick, J. N., McGee, B. M., Yang, A. Y., Siemieniak, D. R., Stark, K. R., Gruppok, R., Sarode, R., Shurin, S. B., Chandrasekarani, V., Stabler, S. P., Sabio[22], H., Bouhassira[33], E. E., Upshaw, J. D., Ginsburg, D. & Tsaikk, H.-M. 2001. Mutations in a member of the ADAMTS gene family cause thrombotic thrombocytopenic purpura.

Li, B., Bailey, A. S., Jiang, S., Liu, B., Goldman, D. C. & Fleming, W. H. 2010. Endothelial cells mediate the regeneration of hematopoietic stem cells. *Stem cell research,* 4**,** 17-24.

Lin, M., Xu, X., Lee, H. L., Liang, D. C. & Santoso, S. 2018. Fetal/neonatal alloimmune thrombocytopenia due to anti-CD36 antibodies: antibody evaluations by CD36-transfected cell lines. *Transfusion,* 58**,** 189-195.

Liu, L., Li, Y., Li, S., Hu, N., He, Y., Pong, R., Lin, D., Lu, L. & Law, M. 2012. Comparison of next-generation sequencing systems. *BioMed research international,* 2012.

Locatelli, F., Rossi, G. & Balduini, C. 2003. Hematopoietic stem-cell transplantation for the Bernard–Soulier syndrome. *Annals of internal medicine,* 138**,** 79.

Loffredo, G., Balduini, C. L., Notarangelo, L. D., Casale, M., Perrotta, S., Pujol-Moix, N., Amendola, G., Pecci, A., Noris, P., Scianguetta, S., Melazzini, F., Zecca, M., De Rocco, D., Marconi, C., Klersy, C., Punzo, F., Seri, M., Pippucci, T., Scognamiglio, F., Gnan, C., Savoia, A. & Civaschi, E. 2011. Mutations in ANKRD26 are responsible for a frequent form of inherited thrombocytopenia: analysis of 78 patients from 21 families. *Blood,* 117**,** 6673-6680.

Lonergan, G. J., Cline, D. B. & Abbondanzo, S. L. 2001. Sickle cell anemia. *Radiographics,* 21**,** 971-994.

Lopes, P. & Oliveira, J. L. 2013. An innovative portal for rare genetic diseases research: the semantic Diseasecard. *Journal of biomedical informatics,* 46**,** 1108-1115.

Lordier, L., Bluteau, D., Jalil, A., Legrand, C., Pan, J., Rameau, P., Jouni, D., Bluteau, O., Mercher, T. & Leon, C. 2012. RUNX1-induced silencing of non-muscle myosin heavy chain IIB contributes to megakaryocyte polyploidization. *Nature communications,* 3**,** 1-10.

Lordier, L., Jalil, A., Aurade, F., Larbret, F., Larghero, J., Debili, N., Vainchenker, W. & Chang, Y. 2008. Megakaryocyte endomitosis is a failure of late cytokinesis related to defects in the contractile ring and Rho/Rock signaling. *Blood, The Journal of the American Society of Hematology,* 112**,** 3164-3174.

Lorenz, V., Stegner, D., Stritt, S., Vögtle, T., Kiefer, F., Witke, W., Schymeinsky, J., Watson, S. P., Walzog, B. & Nieswandt, B. 2015. Targeted downregulation of platelet CLEC-2 occurs through Syk-independent internalization. *Blood,* 125**,** 4069-4077.

Love-Gregory, L., Sherva, R., Schappe, T., Qi, J.-S., McCrea, J., Klein, S., Connelly, M. A. & Abumrad, N. A. 2011. Common CD36 SNPs reduce protein expression and may contribute to a protective atherogenic profile. *Human molecular genetics,* 20**,** 193-201.

Love-Gregory, L., Sherva, R., Sun, L., Wasson, J., Schappe, T., Doria, A., Rao, D. C., Hunt, S. C., Klein, S. & Neuman, R. J. 2008. Variants in the CD36 gene associate with the metabolic syndrome and high-density lipoprotein cholesterol. *Human molecular genetics,* 17**,** 1695-1704.

Lowe, G. C., Fickowska, R., Al Ghaithi, R., Maclachlan, A., Harrison, P., Lester, W., Watson, S. P., Myers, B., Clark, J. & Morgan, N. V. 2019. Investigation of the contribution of an underlying platelet defect in women with unexplained heavy menstrual bleeding. *Platelets,* 30**,** 56-65.

Lowe, G. C., Lordkipanidzé, M., Watson, S. P. & group, U. G. s. 2013. Utility of the ISTH bleeding assessment tool in predicting platelet defects in participants with suspected inherited platelet function disorders. *Journal of Thrombosis and Haemostasis,* 11**,** 1663-1668.

Luk, A. D. W., Yang, X., Alcasabas, A. P., Hao, R. C., Chan, K. W., Lee, P. P., Yang, J., Chan, G. C. F., So, J. C. C. & Yang, W. 2020. NF-E2 mutation as a novel cause for inherited thrombocytopenia. *British Journal of Haematology,* 189**,** e41.

Luthardt, F. W. & Keitges, E. 2001. Chromosomal syndromes and genetic disease. *e LS*.

Machlus, K. R. & Italiano, J. E. 2013. The incredible journey: From megakaryocyte development to platelet formation. *Journal of Cell Biology,* 201**,** 785-796.

Maclachlan, A., Watson, S. P. & Morgan, N. V. 2017. Inherited platelet disorders: Insight from platelet genomics using next-generation sequencing. *Platelets,* 28**,** 14-19.

Mahdieh, N. & Rabbani, B. 2013. An overview of mutation detection methods in genetic disorders. *Iranian journal of pediatrics,* 23**,** 375.

Majewski, J., Schwartzentruber, J., Lalonde, E., Montpetit, A. & Jabado, N. 2011. What can exome sequencing do for you? *Journal of medical genetics,* 48**,** 580-589.

Manchev, V. T., Hilpert, M., Berrou, E., Elaib, Z., Aouba, A., Boukour, S., Souquere, S., Pierron, G., Rameau, P., Andrews, R., Lanza, F., Bobe, R., Vainchenker, W., Rosa, J.-P., Bryckaert, M., Debili, N., Favier, R. & Raslova, H. 2014. A new form of macrothrombocytopenia induced by a germ-line mutation in the PRKACG gene Key Points.

Manukjan, G., Bösing, H., Schmugge, M., Strauß, G. & Schulze, H. 2017. Impact of genetic variants on haematopoiesis in patients with thrombocytopenia absent radii (TAR) syndrome. *British Journal of Haematology,* 179**,** 606-617.

Marconi, C., De Rocco, D., Pecci, A., Noris, P., Pippucci, T., Melazzini, F., Savoia, A., Loffredo, G., Giangregorio, T. & Seri, M. 2017. A new form of inherited thrombocytopenia due to

monoallelic loss of function mutation in the thrombopoietin gene. *British Journal of Haematology,* 181**,** 698-701.
Marcus, A. J., Zucker-Franklin, D., Safier, L. B. & Ullman, H. 1966. Studies on human platelet granules and membranes. *The Journal of clinical investigation,* 45**,** 14-28.
Marigo, V., Nigro, A., Pecci, A., Montanaro, D., Stazio, M. D., Balduini, C. L. & Savoia, A. 2004. Correlation between the clinical phenotype of MYH9-related disease and tissue distribution of class II nonmuscle myosin heavy chains.
Markello, T., Chen, D., Kwan, J. Y., Horkayne-Szakaly, I., Morrison, A., Simakova, O., Maric, I., Lozier, J., Cullinane, A. R., Kilo, T., Meister, L., Pakzad, K., Bone, W., Chainani, S., Lee, E., Links, A., Boerkoel, C., Fischer, R., Toro, C., White, J. G., Gahl, W. A. & Gunay-Aygun, M. 2015. York platelet syndrome is a CRAC channelopathy due to gain-of-function mutations in STIM1 HHS Public Access. *Mol Genet Metab,* 114**,** 474-482.
Mason, K. D., Carpinelli, M. R., Fletcher, J. I., Collinge, J. E., Hilton, A. A., Ellis, S., Kelly, P. N., Ekert, P. G., Metcalf, D. & Roberts, A. W. 2007. Programmed anuclear cell death delimits platelet life span. *Cell,* 128**,** 1173-1186.
Massaad, M. J., Ramesh, N. & Geha, R. S. 2013. Wiskott-Aldrich syndrome: A comprehensive review. *Annals of the New York Academy of Sciences,* 1285**,** 26-43.
McDonald, V. & Austin, S. K. 2017. Inherited bleeding disorders. *Medicine,* 45**,** 229-232.
Miao, W.-M., Vasile, E., Lane, W. S. & Lawler, J. 2001. CD36 associates with CD9 and integrins on human blood platelets. *Blood, The Journal of the American Society of Hematology,* 97**,** 1689-1696.
Michelson, A. D. 1996. Flow cytometry: a clinical test of platelet function. *Open Access Articles***,** 290.
Mikkola, H. K. & Orkin, S. H. 2006. The journey of developing hematopoietic stem cells. *Development,* 133**,** 3733-3744.
Miller, J. L., Cunningham, D., Lyle, V. A. & Finch, C. N. 1991. Mutation in the gene encoding the alpha chain of platelet glycoprotein Ib in platelet-type von Willebrand disease. *Proceedings of the National Academy of Sciences,* 88**,** 4761-4765.
Millikan, P. D., Balamohan, S. M., Raskind, W. H. & Kacena, M. A. 2011. Inherited thrombocytopenia due to GATA-1 mutations. *Seminars in thrombosis and hemostasis,* 37**,** 682-9.
Millington-Burgess, S. L. & Harper, M. T. 2020. Gene of the issue: ANO6 and Scott Syndrome. *Platelets,* 31**,** 964-967.
Miyazaki, K., Koike, Y., Kunishima, S., Ishii, R., Danbara, M., Horie, R., Yatomi, Y. & Higashihara, M. 2015. Immature platelet fraction measurement is influenced by platelet size and is a useful parameter for discrimination of macrothrombocytopenia. *Hematology,* 20**,** 587-592.
Mizushima, S. & Nagata, S. 1990. pEF-BOS, a powerful mammalian expression vector. *Nucleic acids research,* 18**,** 5322.
Modell, B. & Darr, A. 2002. Genetic counselling and customary consanguineous marriage. *Nature Reviews Genetics,* 3**,** 225-229.
Moenen, F. C., Vries, M. J., Nelemans, P. J., van Rooy, K. J., Vranken, J. R., Verhezen, P. W., Wetzels, R. J., Ten Cate, H., Schouten, H. C. & Beckers, E. A. 2019. Screening for platelet function disorders with Multiplate and platelet function analyzer. *Platelets,* 30**,** 81-87.
Monteferrario, D., Bolar, N. A., Marneth, A. E., Hebeda, K. M., Bergevoet, S. M., Veenstra, H., Laros-van Gorkom, B. A. P., MacKenzie, M. A., Khandanpour, C., Botezatu, L., Fransen, E., Van Camp, G., Duijnhouwer, A. L., Salemink, S., Willemsen, B., Huls, G., Preijers, F., Van Heerde, W., Jansen, J. H., Kempers, M. J. E., Loeys, B. L., Van Laer, L., Van der Reijden, B. A., Baplg, H. & Genetics, H. S. S. 2014. A Dominant-Negative GFI1B Mutation in the Gray Platelet Syndrome and the Departments of Pathology (K. *n engl j med,* 370**,** 245-53.
Morgan, N. V. & Daly, M. E. 2017. Gene of the issue: RUNX1 mutations and inherited bleeding. *Platelets,* 28**,** 208-210.
Morrison, S. J. & Weissman, I. L. 1994. The long-term repopulating subset of hematopoietic stem cells is deterministic and isolatable by phenotype. *Immunity,* 1**,** 661-673.

Motohashi, H., Kimura, M., Fujita, R., Inoue, A., Pan, X., Takayama, M., Katsuoka, F., Aburatani, H., Bresnick, E. H. & Yamamoto, M. 2010. NF-E2 domination over Nrf2 promotes ROS accumulation and megakaryocytic maturation. *Blood, The Journal of the American Society of Hematology,* 115**,** 677-686.

Mueller-Eckhardt, C., Grubert, A., Weisheit, M., Mueller-Eckhardt, G., Kiefel, V., Kroll, H., Schmidt, S. & Santoso, S. 1989. 348 cases of suspected neonatal alloimmune thrombocytopenia. *The Lancet,* 333**,** 363-366.

Mundell, S. J. & Mumford, A. 2018. TBXA2R gene variants associated with bleeding. *Platelets,* 29**,** 739-742.

Murone, M., Carpenter, D. A. & de Sauvage, F. J. 1998. Hematopoietic Deficiencies in c-mpl and TPO Knockout Mice. *Stem Cells,* 16**,** 1-6.

Musialik, E., Bujko, M., Kober, P., Grygorowicz, M. A., Libura, M., Przestrzelska, M., Juszczyński, P., Borg, K., Florek, I. & Jakóbczyk, M. 2014. Comparison of promoter DNA methylation and expression levels of genes encoding CCAAT/enhancer binding proteins in AML patients. *Leukemia research,* 38**,** 850-856.

NERGIZ-UNAL, R., Lamers, M. M., van Kruchten, R., Luiken, J., Cosemans, J. M., Glatz, J. F., Kuijpers, M. J. & Heemskerk, J. W. 2011. Signaling role of CD36 in platelet activation and thrombus formation on immobilized thrombospondin or oxidized low-density lipoprotein. *Journal of Thrombosis and Haemostasis,* 9**,** 1835-1846.

Ng, S. B., Buckingham, K. J., Lee, C., Bigham, A. W., Tabor, H. K., Dent, K. M., Huff, C. D., Shannon, P. T., Jabs, E. W. & Nickerson, D. A. 2010. Exome sequencing identifies the cause of a mendelian disorder. *Nature genetics,* 42**,** 30-35.

Ng, S. B., Turner, E. H., Robertson, P. D., Flygare, S. D., Bigham, A. W., Lee, C., Shaffer, T., Wong, M., Bhattacharjee, A. & Eichler, E. E. 2009. Targeted capture and massively parallel sequencing of 12 human exomes. *Nature,* 461**,** 272-276.

Ni, H. 2012. The platelet "sugar high" in diabetes. *Blood, The Journal of the American Society of Hematology,* 119**,** 5949-5951.

Nieswandt, B., Varga-Szabo, D. & Elvers, M. 2009. Integrins in platelet activation. *Journal of Thrombosis and Haemostasis,* 7**,** 206-209.

Nieswandt, B. & Watson, S. P. 2003. Platelet-collagen interaction: is GPVI the central receptor? *Blood,* 102**,** 449-461.

Noetzli, L., Lo, R. W., Lee-Sherick, A. B., Callaghan, M., Noris, P., Savoia, A., Rajpurkar, M., Jones, K., Gowan, K., Balduini, C. L., Pecci, A., Gnan, C., De Rocco, D., Doubek, M., Li, L., Lu, L., Leung, R., Landolt-Marticorena, C., Hunger, S., Heller, P., Gutierrez-Hartmann, A., Xiayuan, L., Pluthero, F. G., Rowley, J. W., Weyrich, A. S., Kahr, W. H. A., Porter, C. C. & Di Paola, J. 2015. Germline mutations in ETV6 are associated with thrombocytopenia, red cell macrocytosis and predisposition to lymphoblastic leukemia. *Nature Genetics,* 47**,** 535-538.

Nombela-Arrieta, C., Pivarnik, G., Winkel, B., Canty, K. J., Harley, B., Mahoney, J. E., Park, S.-Y., Lu, J., Protopopov, A. & Silberstein, L. E. 2013. Quantitative imaging of haematopoietic stem and progenitor cell localization and hypoxic status in the bone marrow microenvironment. *Nature cell biology,* 15**,** 533-543.

Noris, P., Klersy, C., Gresele, P., Giona, F., Giordano, P., Minuz, P., Loffredo, G., Pecci, A., Melazzini, F. & Civaschi, E. 2013. Platelet size for distinguishing between inherited thrombocytopenias and immune thrombocytopenia: a multicentric, real life study. *British journal of haematology,* 162**,** 112-119.

Noris, P., Melazzini, F. & Balduini, C. L. 2016. New roles for mean platelet volume measurement in the clinical practice? *Platelets,* 27**,** 607-612.

Noris, P. & Pecci, A. 2017. Hereditary thrombocytopenias: A growing list of disorders. *Hematology,* 2017**,** 385-399.

North, T. E., Stacy, T., Matheny, C. J., Speck, N. A. & de Bruijn, M. F. 2004. Runx1 is expressed in adult mouse hematopoietic stem cells and differentiating myeloid and lymphoid cells, but not in maturing erythroid cells. *Stem cells,* 22**,** 158-168.

Nowakowska, B. 2017. Clinical interpretation of copy number variants in the human genome. *Journal of applied genetics,* 58**,** 449-457.

Nurden, A., Freson, K. & Seligsohn, U. 2012. Inherited platelet disorders. *Haemophilia,* 18**,** 154-160.

Nurden, A. T. 2006. Glanzmann thrombasthenia. *Orphanet journal of rare diseases,* 1**,** 10.

Nurden, A. T. & Nurden, P. 2018. High-throughput sequencing for rapid diagnosis of inherited platelet disorders: a case for a European consensus. *haematologica,* 103**,** 6.

Nurden, A. T. & Nurden, P. 2020. Inherited thrombocytopenias: history, advances and perspectives. *Haematologica*.

Nurden, A. T., Pillois, X. & Wilcox, D. A. 2013. Glanzmann Thrombasthenia: State of the Art and Future Directions. *Seminars in thrombosis and hemostasis,* 39**,** 642-642.

Nurden, P., Debili, N., Coupry, I., Bryckaert, M., Youlyouz-Marfak, I., Solé, G., cile Pons, A.-C., Berrou, E., ric Adam, F., Kauskot, A., Daniel Lamaziè re, J.-M., Rameau, P., Fergelot, P., Rooryck, C., Cailley, D., Arveiler, B., Lacombe, D., Vainchenker, W., Nurden, A. & Goizet, C. 2011. Thrombocytopenia resulting from mutations in filamin A can be expressed as an isolated syndrome.

Nurden, P., Savi, P., Heilmann, E., Bihour, C., Herbert, J.-M., Maffrand, J.-P. & Nurden, A. 1995. An inherited bleeding disorder linked to a defective interaction between ADP and its receptor on platelets. Its influence on glycoprotein IIb-IIIa complex function. *The Journal of clinical investigation,* 95**,** 1612-1622.

Nurden, P., Stritt, S., Favier, R. & Nurden, A. T. 2021. Inherited platelet diseases with normal platelet count: phenotypes, genotypes and diagnostic strategy. *haematologica,* 106**,** 337.

Ohto, H., Miura, S., Ariga, H., Ishii, T., Fujimori, K. & Morita, S. 2004. The natural history of maternal immunization against foetal platelet alloantigens. *Transfusion Medicine,* 14**,** 399-408.

Okajima, S., Cho, K., Chiba, H., Azuma, H., Mochizuki, T., Yamaguchi, M., Sato, S.-i., Ikeda, H., Yamada, H. & Minakami, H. 2006. Two sibling cases of hydrops fetalis due to alloimmune anti-CD36 (Naka) antibody. *Thrombosis and haemostasis,* 95**,** 267-271.

Ong, L., Morison, I. M. & Ledgerwood, E. C. 2017. Megakaryocytes from CYCS mutation-associated thrombocytopenia release platelets by both proplatelet-dependent and -independent processes. *British Journal of Haematology,* 176**,** 268-279.

Oquendo, P., Hundt, E., Lawler, J. & Seed, B. 1989. CD36 directly mediates cytoadherence of Plasmodium falciparum parasitized erythrocytes. *Cell,* 58**,** 95-101.

Othman, M. & Emsley, J. 2017. Gene of the issue: GP1BA gene mutations associated with bleeding. *Platelets,* 28**,** 832-836.

Othman, M., Kaur, H., Favaloro, E., Lillicrap, D., Di Paola, J., Harrison, P., Gresele, P., Disease, S. o. v. W. & Physiology, P. 2016. Platelet type von Willebrand disease and registry report: communication from the SSC of the ISTH. *Journal of Thrombosis and Haemostasis,* 14**,** 411-414.

Pabinger, S., Dander, A., Fischer, M., Snajder, R., Sperk, M., Efremova, M., Krabichler, B., Speicher, M. R., Zschocke, J. & Trajanoski, Z. 2014. A survey of tools for variant analysis of next-generation genome sequencing data. *Briefings in bioinformatics,* 15**,** 256-278.

Pai, M. & Hayward, C. P. Diagnostic assessment of platelet disorders: what are the challenges to standardization? Seminars in thrombosis and hemostasis, 2009. © Thieme Medical Publishers, 131-138.

Palla, R., Peyvandi, F. & Shapiro, A. D. 2015. Rare bleeding disorders: diagnosis and treatment. *Blood, The Journal of the American Society of Hematology,* 125**,** 2052-2061.

Paniccia, R., Priora, R., Liotta, A. A. & Abbate, R. 2015. Platelet function tests: a comparative review. *Vascular health and risk management,* 11**,** 133.

Para, A., Bark, D., Lin, A. & Ku, D. 2011. Rapid platelet accumulation leading to thrombotic occlusion. *Annals of biomedical engineering,* 39**,** 1961-1971.

Park, Y. M. 2014. CD36, a scavenger receptor implicated in atherosclerosis. *Experimental & molecular medicine,* 46**,** e99.

Park, Y. M., Febbraio, M. & Silverstein, R. L. 2009. CD36 modulates migration of mouse and human macrophages in response to oxidized LDL and may contribute to macrophage trapping in the arterial intima. *The Journal of clinical investigation,* 119**,** 136-145.

Patel, S. R., Richardson, J. L., Schulze, H., Kahle, E., Galjart, N., Drabek, K., Shivdasani, R. A., Hartwig, J. H. & Italiano Jr, J. E. 2005. Differential roles of microtubule assembly and sliding in proplatelet formation by megakaryocytes. *Blood,* 106**,** 4076-4085.

Patil, S., Rao, R. S. & Majumdar, B. 2014. Chromosomal and multifactorial genetic disorders with oral manifestations. *Journal of international oral health: JIOH,* 6**,** 118.

Patton, L. L. & Ship, J. A. 1994. Treatment of patients with bleeding disorders. *Dental Clinics of North America,* 38**,** 465-482.

Pearce, S. F. A., Roy, P., Nicholson, A. C., Hajjar, D. P., Febbraio, M. & Silverstein, R. L. 1998. Recombinant glutathione S-transferase/CD36 fusion proteins define an oxidized low density lipoprotein-binding domain. *Journal of Biological Chemistry,* 273**,** 34875-34881.

Pearce, S. F. A., Wu, J. & Silverstein, R. L. 1995. Recombinant GST/CD36 fusion proteins define a thrombospondin binding domain evidence for a single CALCIUM-DEPENDENT binding site ON CD36. *Journal of Biological Chemistry,* 270**,** 2981-2986.

Pecci, A., Granata, A., Fiore, C. E. & Balduini, C. L. 2008. Renin–angiotensin system blockade is effective in reducing proteinuria of patients with progressive nephropathy caused by MYH9 mutations (Fechtner–Epstein syndrome). *Nephrology Dialysis Transplantation,* 23**,** 2690-2692.

Pecci, A., Gresele, P., Klersy, C., Savoia, A., Noris, P., Fierro, T., Bozzi, V., Mezzasoma, A. M., Melazzini, F. & Balduini, C. L. 2010. Eltrombopag for the treatment of the inherited thrombocytopenia deriving from MYH9 mutations. *Blood, The Journal of the American Society of Hematology,* 116**,** 5832-5837.

Pecci, A., Klersy, C., Gresele, P., Lee, K. J. D., De Rocco, D., Bozzi, V., Russo, G., Heller, P. G., Loffredo, G., Ballmaier, M., Fabris, F., Beggiato, E., Kahr, W. H., Pujol-Moix, N., Platokouki, H., Van Geet, C., Noris, P., Yerram, P., Hermans, C., Gerber, B., Economou, M., De Groot, M., Zieger, B., De Candia, E., Fraticelli, V., Kersseboom, R., Piccoli, G. B., Zimmermann, S., Fierro, T., Glembotsky, A. C., Vianello, F., Zaninetti, C., Nicchia, E., Güthner, C., Baronci, C., Seri, M., Knight, P. J., Balduini, C. L. & Savoia, A. 2014. MYH9-related disease: A novel prognostic model to predict the clinical evolution of the disease based on genotype-phenotype correlations. *Human Mutation,* 35**,** 236-247.

Pekelharing, J., Hauss, O., De Jonge, R., Lokhoff, J., Sodikromo, J., Spaans, M., Brouwer, R., De Lathouder, S. & Hinzmann, R. 2010. Haematology reference intervals for established and novel parameters in healthy adults. *Sysmex Journal International,* 20**,** 1-9.

Perez Botero, J., Warad, D. M., He, R., Uhl, C. B., Tian, S., Otteson, G. E., Barness, R. L., Olson, M. C., Gossman, S. C. & Charlesworth, J. E. 2017. Comprehensive platelet phenotypic laboratory testing and bleeding history scoring for diagnosis of suspected hereditary platelet disorders: a single-institution experience. *American Journal of Clinical Pathology,* 148**,** 23-32.

Peterson, J. A., McFarland, J. G., Curtis, B. R. & Aster, R. H. 2013. Neonatal alloimmune thrombocytopenia: pathogenesis, diagnosis and management. *British journal of haematology,* 161**,** 3-14.

Pippucci, T., Savoia, A., Perrotta, S., Pujol-Moix, N., Noris, P., Castegnaro, G., Pecci, A., Gnan, C., Punzo, F., Marconi, C., Gherardi, S., Loffredo, G., De Rocco, D., Scianguetta, S., Barozzi, S., Magini, P., Bozzi, V., Dezzani, L., Di Stazio, M., Ferraro, M., Perini, G., Seri, M. & Balduini, C. L. 2011. Mutations in the 5' UTR of ANKRD26, the ankirin repeat domain 26 gene, cause an

autosomal-dominant form of inherited thrombocytopenia, THC2. *American Journal of Human Genetics,* 88**,** 115-120.

Platt, N. & Gordon, S. 2001. Is the class A macrophage scavenger receptor (SR-A) multifunctional?—The mouse's tale. *The Journal of clinical investigation,* 108**,** 649-654.

Pleines, I., Woods, J., Chappaz, S., Kew, V., Foad, N., Ballester-Beltrán, J., Aurbach, K., Lincetto, C., Lane, R. M., Schevzov, G., Alexander, W. S., Hilton, D. J., Astle, W. J., Downes, K., Nurden, P., Westbury, S. K., Mumford, A. D., Obaji, S. G., Collins, P. W., Delerue, F., Ittner, L. M., Bryce, N. S., Holliday, M., Lucas, C. A., Hardeman, E. C., Ouwehand, W. H., Gunning, P. W., Turro, E., Tijssen, M. R. & Kile, B. T. 2017. Mutations in tropomyosin 4 underlie a rare form of human macrothrombocytopenia. *Journal of Clinical Investigation,* 127**,** 814-829.

Pluthero, F. G., Di Paola, J., Carcao, M. D. & Kahr, W. H. A. 2018. NBEAL2 mutations and bleeding in patients with gray platelet syndrome.

Pogue, R. E., Cavalcanti, D. P., Shanker, S., Andrade, R. V., Aguiar, L. R., de Carvalho, J. L. & Costa, F. F. 2018. Rare genetic diseases: update on diagnosis, treatment and online resources. *Drug discovery today,* 23**,** 187-195.

Pollard, K. S., Hubisz, M. J., Rosenbloom, K. R. & Siepel, A. 2010. Detection of nonneutral substitution rates on mammalian phylogenies. *Genome research,* 20**,** 110-121.

Poon, M. C., d'Oiron, R., Von Depka, M., Khair, K., Negrier, C., Karafoulidou, A., Huth-Kuehne, A., Morfini, M., VIIa, m. o. t. I. D. C. o. r. F. & Group, C. P. D. S. 2004. Prophylactic and therapeutic recombinant factor VIIa administration to patients with Glanzmann's thrombasthenia: results of an international survey. *Journal of Thrombosis and Haemostasis,* 2**,** 1096-1103.

Poulter, N. 1999. Coronary heart disease is a multifactorial disease. *American Journal of Hypertension,* 12**,** 92S-95S.

Pulido-Salgado, M., Vidal-Taboada, J. M. & Saura, J. 2015. C/EBPβ and C/EBPδ transcription factors: basic biology and roles in the CNS. *Progress in Neurobiology,* 132**,** 1-33.

Rabbani, B., Tekin, M. & Mahdieh, N. 2014. The promise of whole-exome sequencing in medical genetics. *Journal of human genetics,* 59**,** 5-15.

Rabbolini, D., Morel-Kopp, M. C., Chen, Q., Gabrielli, S., Dunlop, L., Chew, L., Blair, N., Brighton, T., Singh, N. & Ng, A. 2017. Thrombocytopenia and CD 34 expression is decoupled from α-granule deficiency with mutation of the first growth factor-independent 1B zinc finger. *Journal of Thrombosis and Haemostasis,* 15**,** 2245-2258.

Rać, M. E., Safranow, K. & Poncyljusz, W. 2007. Molecular basis of human CD36 gene mutations. *Molecular Medicine,* 13**,** 288-296.

Raccuglia, G. 1971. Gray platelet syndrome: a variety of qualitative platelet disorder. *The American journal of medicine,* 51**,** 818-828.

Radley, J. & Haller, C. 1982. The demarcation membrane system of the megakaryocyte: a misnomer?

Rahaman, S. O., Lennon, D. J., Febbraio, M., Podrez, E. A., Hazen, S. L. & Silverstein, R. L. 2006. A CD36-dependent signaling cascade is necessary for macrophage foam cell formation. *Cell metabolism,* 4**,** 211-221.

Raslova, H., Roy, L., Vourc'h, C., Le Couedic, J. P., Brison, O., Metivier, D., Feunteun, J., Kroemer, G., Debili, N. & Vainchenker, W. 2003. Megakaryocyte polyploidization is associated with a functional gene amplification. *Blood,* 101**,** 541-544.

Revel-Vilk, S., Shai, E., Turro, E., Jahshan, N., Hi-Am, E., Spectre, G., Daum, H., Kalish, Y., Althaus, K., Greinacher, A., Kaplinsky, C., Izraeli, S., Mapeta, R., Deevi, S. V. V., Jarocha, D., Ouwehand, W. H., Downes, K., Poncz, M., Varon, D. & Lambert, M. P. 2018. GNE variants causing autosomal recessive macrothrombocytopenia without associated muscle wasting. American Society of Hematology.

Ribatti, D. & Crivellato, E. 2007. Giulio Bizzozero and the discovery of platelets. *Leukemia research,* 31**,** 1339-1341.

Richards, S., Aziz, N., Bale, S., Bick, D., Das, S., Gastier-Foster, J., Grody, W. W., Hegde, M., Lyon, E. & Spector, E. 2015. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genetics in medicine,* 17**,** 405-423.

Richardson, J. L., Shivdasani, R. A., Boers, C., Hartwig, J. H. & Italiano Jr, J. E. 2005. Mechanisms of organelle transport and capture along proplatelets during platelet production. *Blood,* 106**,** 4066-4075.

Ritz, B., Yu, F., Fruin, S., Chapa, G., Shaw, G. M. & Harris, J. A. 2002. Ambient air pollution and risk of birth defects in Southern California. *American journal of epidemiology,* 155**,** 17-25.

Rodeghiero, F., Tosetto, A., Abshire, T., Arnold, D., Coller, B., James, P., Neunert, C., Lillicrap, D., VWF, I. S. j. & Group, P. P. H. S. W. 2010. ISTH/SSC bleeding assessment tool: a standardized questionnaire and a proposal for a new bleeding score for inherited bleeding disorders. *Journal of Thrombosis and Haemostasis,* 8**,** 2063-2065.

Rodgers, G. M. 1999. Overview of platelet physiology and laboratory evaluation of platelet function. *Clinical obstetrics and gynecology,* 42**,** 349-359.

Ross, D. W., Ayscue, L. H., Watson, J. & Bentley, S. A. 1988. Stability of hematologic parameters in healthy subjects: Intraindividual versus interindividual variation. *American journal of clinical pathology,* 90**,** 262-267.

Rowley, J. W., Oler, A. J., Tolley, N. D., Hunter, B. N., Low, E. N., Nix, D. A., Yost, C. C., Zimmerman, G. A. & Weyrich, A. S. 2011. Genome-wide RNA-seq analysis of human and mouse platelet transcriptomes. *Blood,* 118**,** e101-e111.

Rubak, P., Nissen, P. H., Kristensen, S. D. & Hvas, A.-M. 2016. Investigation of platelet function and platelet disorders using flow cytometry. *Platelets,* 27**,** 66-74.

Ruggeri, Z. M. 2002. Platelets in atherothrombosis. *Nature medicine,* 8**,** 1227.

Ruggeri, Z. M. & Mendolicchio, G. L. 2007. Adhesion mechanisms in platelet function. *Circulation research,* 100**,** 1673-1685.

Ruiz, F. A., Lea, C. R., Oldfield, E. & Docampo, R. 2004. Human platelet dense granules contain polyphosphate and are similar to acidocalcisomes of bacteria and unicellular eukaryotes. *Journal of Biological Chemistry,* 279**,** 44250-44257.

Sakharkar, M. K., Chow, V. T. & Kangueane, P. 2004. Distributions of exons and introns in the human genome. *In silico biology,* 4**,** 387-393.

Samaras, P., Schmidt, T., Frejno, M., Gessulat, S., Reinecke, M., Jarzab, A., Zecha, J., Mergner, J., Giansanti, P. & Ehrlich, H.-C. 2020. ProteomicsDB: a multi-omics and multi-organism resource for life science research. *Nucleic acids research,* 48**,** D1153-D1163.

Sanger, F., Nicklen, S. & Coulson, A. R. 1977. DNA sequencing with chain-terminating inhibitors. *Proceedings of the national academy of sciences,* 74**,** 5463-5467.

Savage, B., Saldívar, E. & Ruggeri, Z. M. 1996. Initiation of platelet adhesion by arrest onto fibrinogen or translocation on von Willebrand factor. *Cell,* 84**,** 289-297.

Savoia, A. 2016. Molecular basis of inherited thrombocytopenias. *Clinical genetics,* 89**,** 154-162.

Savoia, A. & Pecci, A. 2015. MYH9-related disorders. *GeneReviews®[Internet].* University of Washington, Seattle.

Schachtner, H., Calaminus, S. D., Sinclair, A., Monypenny, J., Blundell, M. P., Leon, C., Holyoake, T. L., Thrasher, A. J., Michie, A. M. & Vukovic, M. 2013. Megakaryocytes assemble podosomes that degrade matrix and protrude through basement membrane. *Blood, The Journal of the American Society of Hematology,* 121**,** 2542-2552.

Schlegelberger, B. & Heller, P. G. RUNX1 deficiency (familial platelet disorder with predisposition to myeloid leukemia, FPDMM). Seminars in Hematology, 2017. Elsevier, 75-80.

Schofield, R. 1978. The relationship between the spleen colony-forming cell and the haemopoietic stem cell. *Blood cells,* 4**,** 7.

Schulze, H., Korpal, M., Hurov, J., Kim, S.-W., Zhang, J., Cantley, L. C., Graf, T. & Shivdasani, R. A. 2006. Characterization of the megakaryocyte demarcation membrane system and its role in thrombopoiesis. *Blood,* 107**,** 3868-3875.

Schulze, H. & Shivdasani, R. 2005. Mechanisms of thrombopoiesis. *Journal of Thrombosis and Haemostasis,* 3**,** 1717-1724.

Schwarz, J. M., Cooper, D. N., Schuelke, M. & Seelow, D. 2014. MutationTaster2: mutation prediction for the deep-sequencing age. *Nature methods,* 11**,** 361-362.

Schwenter, F., Faughnan, M. E., Gradinger, A. B., Berk, T., Gryfe, R., Pollett, A., Cohen, Z., Gallinger, S. & Durno, C. 2012. Juvenile polyposis, hereditary hemorrhagic telangiectasia, and early onset colorectal cancer in patients with SMAD4 mutation. *Journal of gastroenterology,* 47**,** 795-804.

Schwer, H. D., Lecine, P., Tiwari, S., Italiano Jr, J. E., Hartwig, J. H. & Shivdasani, R. A. 2001. A lineage-restricted and divergent β-tubulin isoform is essential for the biogenesis, structure and function of blood platelets. *Current Biology,* 11**,** 579-586.

Seita, J. & Weissman, I. L. 2010. Hematopoietic stem cell: self-renewal versus differentiation. *Wiley Interdisciplinary Reviews: Systems Biology and Medicine,* 2**,** 640-653.

Seo, A., Gulsuner, S., Pierce, S., Ben-Harosh, M., Shalev, H., Walsh, T., Krasnov, T., Dgany, O., Doulatov, S. & Tamary, H. 2019. Inherited thrombocytopenia associated with mutation of UDP-galactose-4-epimerase (GALE). *Human molecular genetics,* 28**,** 133-142.

Shahrizaila, N., Gaffney, P. M., Nesin, V., Ong, E. C., Tsiokas, L., Nicholl, D. J., Kousi, M., Lehmann, T., Wiley, G., Suri, M., Katsanis, N. & Wierenga, K. J. 2014. Activating mutations in STIM1 and ORAI1 cause overlapping syndromes of tubular myopathy and congenital miosis. *Proceedings of the National Academy of Sciences,* 111**,** 4197-4202.

Shawky, R. M., Elsayed, S. M., Zaki, M. E., El-Din, S. M. N. & Kamal, F. M. 2013. Consanguinity and its relevance to clinical genetics. *Egyptian Journal of Medical Human Genetics,* 14.

Shivdasani, R. A. 2001. Molecular and transcriptional regulation of megakaryocyte differentiation. *Stem cells,* 19**,** 397-407.

Shivdasani, R. A., Mcdevitt, M. A., Fujiwara, Y. & Orkin, S. H. 1998. Transcription factor GATA-1 in megakaryocyte development. *Stem Cells,* 16**,** 79-83.

Silverstein, R. L. & Febbraio, M. 2000. CD36 and atherosclerosis. *Current opinion in lipidology,* 11**,** 483-491.

Silverstein, R. L. & Febbraio, M. 2009. CD36, a scavenger receptor involved in immunity, metabolism, angiogenesis, and behavior. *Science signaling,* 2**,** re3-re3.

Silverstein, R. L., Li, W., Park, Y. M. & Rahaman, S. O. 2010. Mechanisms of cell signaling by the scavenger receptor CD36: implications in atherosclerosis and thrombosis. *Transactions of the American Clinical and Climatological Association,* 121**,** 206.

Simeoni, I., Stephens, J. C., Hu, F., Deevi, S. V. V., Megy, K., Bariana, T. K., Lentaigne, C., Schulman, S., Sivapalaratnam, S., Vries, M. J. A., Westbury, S. K., Greene, D., Papadia, S., Alessi, M. C., Attwood, A. P., Ballmaier, M., Baynam, G., Bermejo, E., Bertoli, M., Bray, P. F., Bury, L., Cattaneo, M., Collins, P., Daugherty, L. C., Favier, R., French, D. L., Furie, B., Gattens, M., Germeshausen, M., Ghevaert, C., Goodeve, A. C., Guerrero, J. A., Hampshire, D. J., Hart, D. P., Heemskerk, J. W. M., Henskens, Y. M. C., Hill, M., Hogg, N., Jolley, J. D., Kahr, W. H., Kelly, A. M., Kerr, R., Kostadima, M., Kunishima, S., Lambert, M. P., Liesner, R., López, J. A., Mapeta, R. P., Mathias, M., Millar, C. M., Nathwani, A., Neerman-Arbez, M., Nurden, A. T., Nurden, P., Othman, M., Peerlinck, K., Perry, D. J., Poudel, P., Reitsma, P., Rondina, M. T., Smethurst, P. A., Stevenson, W., Szkotak, A., Tuna, S., Van Geet, C., Whitehorn, D., Wilcox, D. A., Zhang, B., Revel-Vilk, S., Gresele, P., Bellissimo, D. B., Penkett, C. J., Laffan, M. A., Mumford, A. D., Rendon, A., Gomez, K., Freson, K., Ouwehand, W. H. & Turro, E. 2016. A high-throughput sequencing test for diagnosing inherited bleeding, thrombotic, and platelet disorders. *Blood,* 127**,** 2791-2803.

Simon, D., Kunicki, T. & Nugent, D. 2008. Platelet function defects. *Haemophilia,* 14**,** 1240-1249.

Simpson, M. A., Daly, M. E., Morgan, N. V., Watson, S. P., Collins, P. W., Lordkipanidzé, M., Motwani, J., Fletcher, S. J., Dawood, B., Guiú, I. S., Rivera, J., Johnson, B., Drake, S., Lowe, G. C. & Bem, D. 2015. SLFN14 mutations underlie thrombocytopenia with excessive bleeding and platelet secretion defects. *Journal of Clinical Investigation,* 125**,** 3600-3605.

Sivapalaratnam, S., Collins, J. & Gomez, K. 2017a. Diagnosis of inherited bleeding disorders in the genomic era. *British journal of haematology,* 179**,** 363-376.

Sivapalaratnam, S., Westbury, S. K., Stephens, J. C., Greene, D., Downes, K., Kelly, A. M., Lentaigne, C., Astle, W. J., Huizinga, E. G., Nurden, P., Papadia, S., Peerlinck, K., Penkett, C. J., Perry, D. J., Roughley, C., Simeoni, I., Stirrups, K., Hart, D. P., Campbell Tait, R., Mumford, A. D., BioResource, N., Laffan, M. A., Freson, K., Ouwehand, W. H., Kunishima, S. & Turro, E. 2017b. Brief Report Rare variants in GP1BB are responsible for autosomal dominant macrothrombocytopenia.

Snapper, S. B. & Rosen, F. S. 1999. THE WISKOTT-ALDRICH SYNDROME PROTEIN (WASP): Roles in Signaling and Cytoskeletal Organization.

Starita, L. M., Ahituv, N., Dunham, M. J., Kitzman, J. O., Roth, F. P., Seelig, G., Shendure, J. & Fowler, D. M. 2017. Variant interpretation: functional assays to the rescue. *The American Journal of Human Genetics,* 101**,** 315-325.

Stevenson, W. S., Morel-Kopp, M. C., Chen, Q., Liang, H. P., Bromhead, C. J., Wright, S., Turakulov, R., Ng, A. P., Roberts, A. W., Bahlo, M. & Ward, C. M. 2013. GFI1B mutation causes a bleeding disorder with abnormal platelet function. *Journal of Thrombosis and Haemostasis,* 11**,** 2039-2047.

Stevenson, W. S., Rabbolini, D. J., Beutler, L., Chen, Q., Gabrielli, S., Mackay, J. P., Brighton, T. A., Ward, C. M. & Morel-Kopp, M.-C. 2015. Paris-Trousseau thrombocytopenia is phenocopied by the autosomal recessive inheritance of a DNA-binding domain mutation in FLI1. *Blood,* 126**,** 2027-2030.

Stockley, J., Morgan, N. V., Bem, D., Lowe, G. C., Lordkipanidzé, M., Dawood, B., Simpson, M. A., Macfarlane, K., Horner, K. & Leo, V. C. 2013. Enrichment of FLI1 and RUNX1 mutations in families with excessive bleeding and platelet dense granule secretion defects. *Blood,* 122**,** 4090-4093.

Strachan, T. & Read, A. P. 2019. *Human molecular genetics*.

Stritt, S., Nurden, P., Favier, R., Favier, M., Ferioli, S., Gotru, S. K., M van Eeuwijk, J. M., Schulze, H., Nurden, A. T., Lambert, M. P., Turro, E., Burger-Stritt, S., Matsushita, M., Mittermeier, L., Ballerini, P., Zierler, S., Laffan, M. A., Chubanov, V., Gudermann, T., Nieswandt, B. & Braun, A. 2016a. Defects in TRPM7 channel function deregulate thrombopoiesis through altered cellular Mg2+ homeostasis and cytoskeletal architecture. *Nature Communications,* 7.

Stritt, S., Nurden, P., Turro, E., Greene, D., Jansen, S. B., Westbury, S. K., Petersen, R., Astle, W. J., Marlin, S. & Bariana, T. K. 2016b. A gain-of-function variant in DIAPH1 causes dominant macrothrombocytopenia and hearing loss. *Blood,* 127**,** 2903-2914.

Su, N., Sun, Q., Li, C., Lu, X., Qi, H., Chen, S., Yang, J., Du, X., Zhao, L. & He, Q. 2010. Gain-of-function mutation in FGFR3 in mice leads to decreased bone mass by affecting both osteoblastogenesis and osteoclastogenesis. *Human molecular genetics,* 19**,** 1199-1210.

Sun, W., Shen, Y., Shan, S., Han, L., Li, Y., Zhou, Z., Zhong, Z. & Chen, J. 2018. Identification of TYR mutations in patients with oculocutaneous albinism. *Molecular medicine reports,* 17**,** 8409-8413.

Suzuki, T., Li, W., Zhang, Q., Karim, A., Novak, E. K., Sviderskaya, E. V., Hill, S. P., Bennett, D. C., Levin, A. V. & Nieuwenhuis, H. K. 2002. Hermansky-Pudlak syndrome is caused by mutations in HPS4, the human homolog of the mouse light-ear gene. *Nature genetics,* 30**,** 321-324.

Swinkels, M., Rijkers, M., Voorberg, J., Vidarsson, G., Leebeek, F. W. G. & Jansen, A. J. G. 2018. Emerging concepts in immune thrombocytopenia.

Swystun, L. L. & James, P. D. 2017. Genetic diagnosis in hemophilia and von Willebrand disease. *Blood reviews,* 31**,** 47-56.

Tablin, F., Reeber, M. J. & Nachmias, V. T. 1988. Platelets contain a 210K microtubule-associated protein related to a similar protein in HeLa cells. *Journal of Cell Science,* 90**,** 317-324.

Tadmouri, G. O., Nair, P., Obeid, T., Al Ali, M. T., Al Khaja, N. & Hamamy, H. A. 2009. Consanguinity and reproductive health among Arabs. *Reproductive health,* 6**,** 17.

Taichman, R. S. 2005. Blood and bone: two tissues whose fates are intertwined to create the hematopoietic stem-cell niche. *Blood,* 105**,** 2631-2639.

Take, H., Kashiwagi, H., Tomiyama, Y., Honda, S., Honda, Y., Mizutani, H., Furubayashi, T., Karasuno, T., Nishiura, T. & Kanayama, Y. 1993. Expression of GPIV and Naka antigen on monocytes in Naka-negative subjects whose platelets lack GPIV. *British journal of haematology,* 84**,** 387-391.

Taketani, T., Ito, K., Mishima, S., Kanai, R., Uchiyama, A., Hirata, Y., Kumakura, S., Ishikura, H. & Yamaguchi, S. 2008. Neonatal isoimmune thrombocytopenia caused by type I CD36 deficiency having novel splicing isoforms of the CD36 gene. *European journal of haematology,* 81**,** 70-74.

Tandon, N. N., Kralisz, U. & Jamieson, G. 1989. Identification of glycoprotein IV (CD36) as a primary receptor for platelet-collagen adhesion. *Journal of Biological Chemistry,* 264**,** 7576-7583.

Tchernitchko, D., Goossens, M. & Wajcman, H. 2004. In silico prediction of the deleterious effect of a mutation: proceed with caution in clinical genetics. *Clinical chemistry,* 50**,** 1974-1978.

Teeuw, M. E., Loukili, G., Bartels, E. A., Ten Kate, L. P., Cornel, M. C. & Henneman, L. 2014. Consanguineous marriage and reproductive risk: attitudes and understanding of ethnic groups practising consanguinity in Western society. *European Journal of Human Genetics,* 22**,** 452-457.

Thon, J. N., Devine, M. T., Jurak Begonja, A., Tibbitts, J. & Italiano Jr, J. E. 2012. High-content live-cell imaging assay used to establish mechanism of trastuzumab emtansine (T-DM1)–mediated inhibition of platelet production. *Blood, The Journal of the American Society of Hematology,* 120**,** 1975-1984.

Thon, J. N. & Italiano, J. E. Platelet formation. Seminars in hematology, 2010. Elsevier, 220-226.

Thorne, R. F., Marshall, J. F., Shafren, D. R., Gibson, P. G., Hart, I. R. & Burns, G. F. 2000. The Integrins α3β1 and α6β1 Physically and Functionally Associate with CD36 in Human Melanoma Cells REQUIREMENT FOR THE EXTRACELLULAR DOMAIN OF CD36. *Journal of Biological Chemistry,* 275**,** 35264-35275.

Tijssen, M. & Ghevaert, C. 2013a. Transcription factors in late megakaryopoiesis and related platelet disorders. *Journal of Thrombosis and Haemostasis,* 11**,** 593-604.

Tijssen, M. R. & Ghevaert, C. 2013b. Transcription factors in late megakaryopoiesis and related platelet disorders.

Tiwari, S., Italiano Jr, J. E., Barral, D. C., Mules, E. H., Novak, E. K., Swank, R. T., Seabra, M. C. & Shivdasani, R. A. 2003. A role for Rab27b in NF-E2-dependent pathways of platelet formation. *Blood,* 102**,** 3970-3979.

Tomlinson, M., Calaminus, S., Berlanga, O., Auger, J., Bori-Sanz, T., Meyaard, L. & Watson, S. 2007. Collagen promotes sustained glycoprotein VI signaling in platelets and cell lines. *Journal of Thrombosis and Haemostasis,* 5**,** 2274-2283.

Toth, O., Calatzis, A., Penz, S., Losonczy, H. & Siess, W. 2006. Multiple electrode aggregometry: a new device to measure platelet aggregation in whole blood. *Thrombosis and haemostasis,* 96**,** 781-788.

Toti, F. & Freyssinet, J.-M. 2005. An ABC for Scott syndrome? *Blood,* 106**,** 396-397.

Turnbull, C., Scott, R. H., Thomas, E., Jones, L., Murugaesu, N., Pretty, F. B., Halai, D., Baple, E., Craig, C. & Hamblin, A. 2018. The 100 000 Genomes Project: bringing whole genome sequencing to the NHS. *Bmj,* 361.

Turner, M. L., Bessos, H., Fagge, T., Harkness, M., Rentoul, F., Seymour, J., Wilson, D., Gray, I., Ahya, R. & Cairns, J. 2005. Prospective epidemiologic study of the outcome and cost-effectiveness

of antenatal screening to detect neonatal alloimmune thrombocytopenia due to anti-HPA-1a. *Transfusion,* 45**,** 1945-1956.
Turro, E., Astle, W. J., Megy, K., Gräf, S., Greene, D., Shamardina, O., Allen, H. L., Sanchis-Juan, A., Frontini, M. & Thys, C. 2020. Whole-genome sequencing of patients with rare diseases in a national health system. *Nature,* 583**,** 96-102.
Turro, E., Greene, D., Wijgaerts, A., Thys, C., Lentaigne, C., Bariana, T. K., Westbury, S. K., Kelly, A. M., Selleslag, D., Stephens, J. C., Papadia, S., Simeoni, I., Penkett, C. J., Ashford, S., Attwood, A., Austin, S., Bakchoul, T., Collins, P., Deevi, S. V. V., Favier, R., Kostadima, M., Lambert, M. P., Mathias, M., Millar, C. M., Peerlinck, K., Perry, D. J., Schulman, S., Whitehorn, D., Wittevrongel, C., De Maeyer, M., Rendon, A., Gomez, K., Erber, W. N., Mumford, A. D., Nurden, P., Stirrups, K., Bradley, J. R., Raymond, F. L., Laffan, M. A., Van Geet, C., Richardson, S., Freson, K. & Ouwehand, H. 2016. A dominant gain-of-function mutation in universal tyrosine kinase SRC causes thrombocytopenia, myelofibrosis, bleeding, and bone pathologies. *Science Translational Medicine,* 8**,** 328-358.
Uetsuki, T., Naito, A., Nagata, S. & Kaziro, Y. 1989. Isolation and characterization of the human chromosomal gene for polypeptide chain elongation factor-1 alpha. *Journal of Biological Chemistry,* 264**,** 5791-5798.
Valsesia, A., Stevenson, B. J., Waterworth, D., Mooser, V., Vollenweider, P., Waeber, G., Jongeneel, C. V., Beckmann, J. S., Kutalik, Z. & Bergmann, S. 2012. Identification and validation of copy number variants using SNP genotyping arrays from a large clinical cohort. *BMC genomics,* 13**,** 1-15.
Velten, L., Haas, S. F., Raffel, S., Blaszkiewicz, S., Islam, S., Hennig, B. P., Hirche, C., Lutz, C., Buss, E. C. & Nowak, D. 2017. Human haematopoietic stem cell lineage commitment is a continuous process. *Nature cell biology,* 19**,** 271-281.
Venter, J. C., Adams, M. D., Myers, E. W., Li, P. W., Mural, R. J., Sutton, G. G., Smith, H. O., Yandell, M., Evans, C. A. & Holt, R. A. 2001. The sequence of the human genome. *science,* 291**,** 1304-1351.
Vo, K. K., Jarocha, D. J., Lyde, R. B., Hayes, V., Thom, C. S., Sullivan, S. K., French, D. L. & Poncz, M. 2017. FLI1 level during megakaryopoiesis affects thrombopoiesis and platelet biology. *Blood, The Journal of the American Society of Hematology,* 129**,** 3486-3494.
Wakap, S. N., Lambert, D. M., Olry, A., Rodwell, C., Gueydan, C., Lanneau, V., Murphy, D., Le Cam, Y. & Rath, A. 2020. Estimating cumulative point prevalence of rare diseases: analysis of the Orphanet database. *European Journal of Human Genetics,* 28**,** 165-173.
Wang, Q.-Y., Hu, B., Liu, H., Tang, L., Zeng, W., Wu, Y.-Y., Cheng, Z.-P. & Hu, Y. 2016. A genetic analysis of 23 Chinese patients with hemophilia B. *Scientific reports,* 6**,** 25024.
Wang, X., Crispino, J. D., Letting, D. L., Nakazawa, M., Poncz, M. & Blobel, G. A. 2002. Control of megakaryocyte-specific gene expression by GATA-1 and FOG-1: role of Ets transcription factors. *The EMBO journal,* 21**,** 5225-5234.
Watson, S., Lowe, G., Lordkipanidze, M. & Morgan, N. 2013a. Genotyping and phenotyping of platelet function disorders. *Journal of Thrombosis and Haemostasis,* 11**,** 351-363.
Watson, S., Lowe, G., Lordkipanidze, M., Morgan, N. & consortium, G. 2013b. Genotyping and phenotyping of platelet function disorders. *Journal of Thrombosis and Haemostasis,* 11**,** 351-363.
Watson, S. P. 2009. Platelet activation by extracellular matrix proteins in haemostasis and thrombosis. *Current pharmaceutical design,* 15**,** 1358-1372.
Weatherall, D. J. 2000. Single gene disorders or complex traits: lessons from the thalassaemias and other monogenic diseases. *Bmj,* 321**,** 1117.
Wei, A. H., Schoenwaelder, S. M., Andrews, R. K. & Jackson, S. P. 2009. New insights into the haemostatic function of platelets. *British journal of haematology,* 147**,** 415-430.
Wei, M. L. 2006. Hermansky–Pudlak syndrome: a disease of protein trafficking and organelle function. *Pigment cell research,* 19**,** 19-42.

Wellington, K. & Wagstaff, A. J. 2003. Tranexamic acid. *Drugs,* 63**,** 1417-1433.
Wen, R. & Wang, D. 2019. PTPRJ: a novel inherited thrombocytopenia gene. *Blood,* 133**,** 1272-1274.
Westbury, S. K., Turro, E., Greene, D., Lentaigne, C., Kelly, A. M., Bariana, T. K., Simeoni, I., Pillois, X., Attwood, A. & Austin, S. 2015. Human phenotype ontology annotation and cluster analysis to unravel genetic defects in 707 cases with unexplained bleeding and platelet disorders. *Genome medicine,* 7**,** 1-15.
Whicher, D., Philbin, S. & Aronson, N. 2018. An overview of the impact of rare disease characteristics on research methodology. *Orphanet journal of rare diseases,* 13**,** 14.
White, J. G. & Rao, G. 1998. Microtubule coils versus the surface membrane cytoskeleton in maintenance and restoration of platelet discoid shape. *The American journal of pathology,* 152**,** 597.
Wilkinson, B., Koenigsknecht-Talboo, J., Grommes, C., Lee, C. D. & Landreth, G. 2006. Fibrillar β-amyloid-stimulated intracellular signaling cascades require Vav for induction of respiratory burst and phagocytosis in monocytes and microglia. *Journal of Biological Chemistry,* 281**,** 20842-20850.
Wright, S. D., Michaelides, K., Johnson, D., West, N. C. & Tuddenham, E. 1993. Double heterozygosity for mutations in the platelet glycoprotein IX gene in three siblings with Bernard-Soulier syndrome.
Wu, G., Zhou, Y., Li, L., Zhong, Z., Li, H., Li, H., Yu, M., Shen, W. & Ni, H. 2017. Platelet immunology in China: research and clinical applications. *Transfusion medicine reviews,* 31**,** 118-125.
Xia, W., Ye, X., Xu, X., Ding, H., Liu, J., Deng, J., Chen, Y., Shao, Y., Wang, J. & Li, H. 2014. Two cases of platelet transfusion refractoriness and one case of possible FNAIT caused by antibodies against CD36 in China. *Transfusion Medicine,* 24**,** 254-256.
Xu, X., Li, L., Xia, W., Ding, H., Chen, D., Liu, J., Deng, J., Chen, Y., He, Z. & Wang, J. 2018. Successful management of a hydropic fetus with severe anemia and thrombocytopenia caused by anti-CD36 antibody. *International Journal of Hematology,* 107**,** 251-256.
Xu, X. & Santoso, S. 2018. Role of CD 36 in immune-mediated thrombocytopenia in Asian populations. *ISBT Science Series,* 13**,** 317-322.
Yamamoto, N., Akamatsu, N., Sakuraba, H., Yamazaki, H. & Tanoue, K. 1994. Platelet glycoprotein IV (CD36) deficiency is associated with the absence (type I) or the presence (type II) of glycoprotein IV on monocytes. *blood,* 83**,** 392–397.
Yamamoto, N., Akamatsu, N., Yamazaki, H. & Tanoue, K. 1992. Normal aggregations of glycoprotein IV (CD36)-deficient platelets from seven healthy Japanese donors. *British journal of haematology,* 81**,** 86-92.
Yamashita, S., Hirano, K.-I., Kuwasako, T., Janabi, M., Toyama, Y., Ishigami, M. & Sakai, N. 2007. Physiological and pathological roles of a multi-ligand receptor CD36 in atherogenesis; insights from CD36-deficient patients. *Molecular and cellular biochemistry,* 299**,** 19-22.
Yeaman, M. R. 2014. Platelets: at the nexus of antimicrobial defence. *Nature Reviews Microbiology,* 12**,** 426-437.
Yoon, K.-A., Cho, H.-S., Shin, H.-I. & Cho, J.-Y. 2012. Differential regulation of CXCL5 by FGF2 in osteoblastic and endothelial niche cells supports hematopoietic stem cell migration. *Stem cells and development,* 21**,** 3391-3402.
Youssefian, T., Massé, J.-M., Rendu, F., Guichard, J. & Cramer, E. M. 1997. Platelet and megakaryocyte dense granules contain glycoproteins Ib and IIb-IIIa. *Blood, The Journal of the American Society of Hematology,* 89**,** 4047-4057.
Yufu, Y., Ideguchi, H., Narishige, T., Suematsu, E., Toyoda, K., Nishimura, J., Nawata, H. & Oda, S. 1990. Familial macrothrombocytopenia associated with decreased glycosylation of platelet membrane glycoprotein IV. *American journal of hematology,* 33**,** 271-273.
Yun, S.-H., Sim, E.-H., Goh, R.-Y., Park, J.-I. & Han, J.-Y. 2016. Platelet activation: the mechanisms and potential biomarkers. *BioMed research international,* 2016**,** 5.

Zaninetti, C., Santini, V., Tiniakou, M., Barozzi, S., Savoia, A. & Pecci, A. 2017. Inherited thrombocytopenia caused by ANKRD 26 mutations misdiagnosed and treated as myelodysplastic syndrome: report on two cases. *Journal of Thrombosis and Haemostasis,* 15**,** 2388-2392.

Zarbock, A., Polanowska-Grabowska, R. K. & Ley, K. 2007. Platelet-neutrophil-interactions: linking hemostasis and inflammation. *Blood reviews,* 21**,** 99-111.

Zarzour, W., Kleta, R., Frangoul, H., Suwannarat, P., Jeong, A., Kim, S. Y., Wayne, A. S., Gunay-Aygun, M., White, J. & Filipovich, A. H. 2005. Two novel CHS1 (LYST) mutations: Clinical correlations in an infant with Chediak–Higashi syndrome. *Molecular genetics and metabolism,* 85**,** 125-132.

Zhang, F., Kotha, J., Jennings, L. K. & Zhang, X. A. 2009. Tetraspanins and vascular functions. *Cardiovascular Research,* 83**,** 7-15.

Zhang, F. & Lupski, J. R. 2015. Non-coding genetic variants in human disease. *Human molecular genetics,* 24**,** R102-R110.

Zhang, J., Niu, C., Ye, L., Huang, H., He, X., Tong, W.-G., Ross, J., Haug, J., Johnson, T. & Feng, J. Q. 2003. Identification of the haematopoietic stem cell niche and control of the niche size. *Nature,* 425**,** 836-841.

Zhang, L., Orban, M., Lorenz, M., Barocke, V., Braun, D., Urtz, N., Schulz, C., Von Brühl, M.-L., Tirniceriu, A. & Gaertner, F. 2012. A novel role of sphingosine 1-phosphate receptor S1pr1 in mouse thrombopoiesis. *Journal of Experimental Medicine,* 209**,** 2165-2181.

Zhang, Y., Wang, F., Chen, X., Zhang, Y., Wang, M., Liu, H., Cao, P., Ma, X., Wang, T. & Zhang, J. 2018. CSF3R Mutations are frequently associated with abnormalities of RUNX1, CBFB, CEBPA, and NPM1 genes in acute myeloid leukemia. *Cancer,* 124**,** 3329-3338.

Zhu, W., Li, W. & Silverstein, R. L. 2012. Advanced glycation end products induce a prothrombotic phenotype in mice via interaction with platelet CD36. *Blood, The Journal of the American Society of Hematology,* 119**,** 6136-6144.

Zimmet, J. & Ravid, K. 2000. Polyploidy: occurrence in nature, mechanisms, and significance for the megakaryocyte-platelet system. *Experimental hematology,* 28**,** 3-16.

Zlotogora, J. 2002. What is the birth defect risk associated with consanguineous marriages? *American journal of medical genetics,* 109**,** 70.

Zwaal, R. F., Comfurius, P. & Bevers, E. M. 2004. Scott syndrome, a bleeding disorder caused by defective scrambling of membrane phospholipids. *Biochimica et Biophysica Acta (BBA)-Molecular and Cell Biology of Lipids,* 1636**,** 119-128.

Appendicies

Figure A. Consent form for recruitment to the GAPP study.

UNIVERSITY OF BIRMINGHAM

Birmingham Platelet Group

INFORMED CONSENT FOR PATIENT / FAMILY MEMBER

Thank you for reading the information about our research project. If you would like to take part, please read and sign this form.

Patient Identification Code for this study:

Title of Project: Mild bleeding disorders caused by platelet defects

Contact details for research team:

Your referring doctor / research nurse ……..

or

The study team: Mrs Gayle Halford (general enquiries), Dr Marie Lordkipanidzé and Dr Gillian Lowe

Birmingham Platelet Group, Institute for Biomedical Research, School of Clinical and Experimental Medicine, College of Medical and Dental Sciences, University of Birmingham, Birmingham B15 2TT, ☎ 0121 415 8680. Please leave a message if no-one is in the office and we will get back to you.

Please initial all boxes

1. I have read the attached information sheet on this project, and have been given a copy to keep. I have been able to ask questions about the project and I understand why the research is being done and any risks involved. ☐

2. I agree to give a sample of blood for research in this project. I understand how the sample will be collected, that giving a sample for this research is voluntary and that I am free to withdraw my approval for the use of the sample at any time without giving a reason and without my medical treatment or legal rights being affected. ☐

3. I give permission for someone from the local research team to look at my medical records to get information on my bleeding history. I understand that the information will be kept confidential. ☐

4. I agree to answering some questions about my medical history including those needed to complete a bleeding assessment by questionnaire. ☐

5. I understand that my referring doctor and I will be informed if any of the results of the medical tests done as a part of the research are important for my health. ☐

6. I understand that I will not benefit financially if this research leads to the development of a new treatment or medical test. ☐

7. I know how to contact the research team if I need to, and how to get information about the results of the research. ☐

8. Consent for storage of sample

I agree that the sample I have given and the information gathered about me can be stored in the Medical Schools of the Universities of Birmingham, Bristol and Sheffield for the purpose of analysing platelets, bleeding tendency and detection of the gene change responsible for the platelet disorder in myself or a family member, for a maximum of 10 years, after which it will be destroyed.

If I wish I may request the return of the sample after analysis has been completed. ☐

9. Consent for genetic research.

I give permission for genetic analysis to be carried out on the sample I give, as part of this project. I have received written information about this test and I understand what the result could mean to me and/or members of my family. ☐

10. Consent for stem cell studies

I give permission for the stem cells in my blood sample to be used to understand how my platelets are produced.

YES ☐

NO ☐

11. Consent for RUNX1 mutation testing **(ONLY FOR PARTICIPANTS WITH LOW PLATELET COUNTS).**

I give permission for genetic analysis to be carried out on the sample I give to specifically look at the RUNX1 gene. I understand that this mutation is associated with low platelet counts and can in some cases predispose to certain types of leukaemia.

YES ☐

NO ☐

12. I want to be told the results of the tests undertaken in this study. I understand I can change my mind about this later.

YES ☐

NO ☐

.....	.....	.....
Name of patient (BLOCK CAPITALS)	Date	Signature
.....	.....	.....
Name of person taking consent (if different from researcher)	Date	Signature
.....	.....	.....
Name of researcher	Date	Signature

Thank you for agreeing to participate in this research

1 copy for patient, 1 for hospital notes, 1 for researcher

Figure B. The International Society of Thrombosis and Haemostasis Bleeding Assessment Tool, used to assist the diagnosis of bleeding disorders.

Annexe 1: ISTH/SSC Bleeding Assessment Tool

The clinical appreciation of the presence and severity of bleeding symptoms is a fundamental step in the evaluation of patients referred for a possible bleeding disorder. In an attempt to improve the collection and reproducibility of the bleeding history, several Bleeding Assessment Tools (BAT) have been proposed and used. Currently available BAT have some limitations, particularly regarding the lack of pediatric-specific symptoms in some of them and the predominance of the severity of bleeding symptoms over other potentially clinically important features, such as the frequency of symptoms.

To overcome the above-mentioned limitations and to promote the standardization of the available BATs, a Working Group was established within the framework of the ISTH/SSC Subcommittees on VWF and on Perinatal/Pediatric Hemostasis (ISTH/SSC-BAT) during the 53rd SSC Annual Meeting held in Geneva in 2007. Members of the group first met in Toronto on January 2008 and then regularly at each subsequent SSC meeting. This paper presents a structured questionnaire and its clinical use agreed on by the ISTH/SSC-BAT together with a proposal for a new BS system to undergo validity and reliability testing in future studies.This new BAT is intended for inherited bleeding disorders in children and adults. The questionnaire should be collected by a physician or another adequately trained health-professional. Only symptoms and related treatments, if any, before and/or at diagnosis should be reported. Refer to the full text for additional instructions.

Minimal criteria defining a significant bleeding

For each specific bleeding symptom, the ISTH/SSC joint working group proposed minimal criteria in order to classify a symptom as significant and thus receive a score of 1 or more (see also Table 1):

1. Epistaxis: Any nosebleed, especially occuring after puberty, that causes patient concern (e.g., interference or distress with daily or social activities) is considered significant. In general, epistaxis should not be considered significant when it lasts less than 10 minutes, has a frequency of < 5 episodes/year, has a seasonal occurence, or is associated with infections of the upper respiratory tract or other identifiable cause (e.g., dusty dry air).

2. Cutaneous bleeding: Bruises are considered significant when 5 or more (> 1cm) in exposed areas; petechiae when adequately described by the patient or relatives; or hematomas when occurring without trauma.

3. Minor cutaneous wound: Any bleeding episode caused by superficial cuts (e.g., by shaving razor, knife, or scissors) or that requires frequent bandage changes is considered significant. Insignificant bleeding from wounds includes those of duration < 10 minutes and lesions that usually require stitches in normal subjects (e.g., under the chin). Symptoms should also be manifest on more than one occasion to be considered significant.

4. Oral cavity bleeding: Gum bleeding should be considered significant when it causes frankly bloody sputum and lasts for 10 minutes or longer on more than one occasion. Tooth eruption or spontaneous tooth loss bleeding should be considered significant when it requires assistance or supervision by a physician, or lasts at least 10 minutes (bleeding associated with tooth extraction is considered separately). Bleeding occurring after bites to lips, cheek, and tongue should be considered significant when it lasts at least 10 minutes or causes a swollen tongue or mouth.

5. Hematemesis, melena, and hematochezia: Any gastrointestinal bleeding that is not explained by the presence of a specific disease should be considered significant.

6. Hematuria: Only macroscopic hematuria (from red to pale-pink urine) that is not explained by the presence of a specific urologic disease should be considered significant.

7. Tooth extraction: Any bleeding occurring after leaving the dentist's office and requiring a new, unscheduled visit or prolonged bleeding at the dentist's office causing a delay in the procedure or discharge should be considered significant.

8. Surgical bleeding: Any bleeding judged by the surgeon to be abnormally prolonged, that causes a delay in discharge, or requires some supportive treatment is considered significant.

9. Menorrhagia: Any bleeding that interferes with daily activities such as work, housework, exercise or social activities during most menstrual periods should be considered significant. Criteria for significant bleeding may include any of the following: changing pads more frequently than every 2 hours; menstrual bleeding lasting 7 or more days; and the presence of clots > 1 cm combined with a history of flooding. If a patient has previously made a record of her menstrual loss using a pictorial blood loss assessment chart (PBAC), a PBAC score higher than 100 also qualifies for a score of 1.

10. Post-partum bleeding. Vaginal bleeding or uterine discharge (lochia) that lasts for more than 6 weeks. Any bleeding of lesser duration that is judged by the obstetrician as abnormally heavy or prolonged, that causes a delay in discharge, requires some supportive treatment, requires changing pads or tampons more frequently than every 2 hours, or causes progressive anemia is also considered significant.

11. Muscle hematomas or hemarthrosis. Any spontaneous joint / muscle bleeding (not related to traumatic injuries) is considered significant.

12. CNS bleeding. Any subdural or intracerebral hemorrhage requiring diagnostic or therapeutic intervention is scored 3 or 4, respectively.

13. Other bleeding symptoms. When these bleeding symptoms occur during infancy, they are scored 1 or more. Their presence when reported by either the patient or a family member should always prompt detailed laboratory investigation.

Only symptoms and treatment BEFORE and AT diagnosis should be considered

1.	**Epistaxis**

1.1	Have you ever had spontaneous epistaxis?	□ Yes	□ No or trivial (skip to 2)
1.2	Have the symptom ever required medical attention ?	□ Yes	□ No (resolve spontaneously; skip to 1.6)
1.3	If answer to 1.2 is yes, please specify	□ Consultation only □ Cauterization/ Packing □ Treatment with desmopressin / antifibrinolytics/ iron therapy □ Treatment with plasma, platelet or factor concentrates □ Blood (RBC) transfusion	
1.4	How many times in your life did you receive any of the above treatments (# 1.3)?	□ 1 - 2 □ 3 to 5 □ 6 to 10 □ more than 10	
1.5	At what age did you first have symptoms?	□ Before 1 year □ Between 1-5 years of age □ Between 6-12 years of age □ Between 13-25 years of age □ After 25 years of age	
1.6	Approximate number of episodes NOT requiring medical attention	□ less than 1 per year □ 1 per year □ 1-5 every six month □ 1-3 every month □ 1 every week	
1.7	Duration of average single episode (min.) NOT requiring medical attention	□ 1 minute or less □ 1 - 10 minutes □ more than 10 minutes	

app

2.	Cutaneous bleeding (Bruising, ecchymoses, purpura, subcutanueos hematomas)		
2.1	Have you ever had any of the above cutaneous bleeding?	□ Yes	□ No or trivial skip to 3
2.2	Have the symptom ever required medical attention?	□ Yes	No □ skip to 2.6
2.3	If answer to 2.2 is yes, please specify	□ Consultation only □ Treatment with desmopressin □ Treatment with plasma, platelets or factor concentrates □ Blood (RBC) transfusion	
2.4	How many times in your life did you receive any of the above treatments (# 2.3)?	□ 1 - 2 □ 3 to 5 □ 6 to 10 □ more than 10	
2.5	At what age did you first have symptoms?	□ Before 1 year □ Between 1-5 years of age □ Between 6-12 years of age □ Between 13-25 years of age □ After 25 years of age	
2.6	Approximate number of episodes NOT requiring medical attention	□ less than 1 per year □ 1 per year □ 1-5 every six month □ 1-3 every month □ 1 every week	
2.7	Type of bleeding	□ Petechiae □ Bruises □ Hematomas	
2.8	Location	□ Exposed sites □ Unexposed sites □ Both	
2.9	Common size	□ ≤ 1 cm □ >1 cm □ Extensive (palm sized or larger)	
2.10	How many bruises >1 cm in exposed areas in the most severe manifestation?	□ ≤ 5 □ > 5	
2.11	Location of petechiae	□ Limited to lower limbs □ Diffuse	

3.	**Bleeding from minor wounds (not requiring stitches in the average patient)**

3.1	Have you ever had prolonged bleeding from minor wounds?	□ Yes	□ No or trivial skip to 4
3.2	Have the symptom ever required medical attention ?	□ Yes	□ No skip to 3.6
3.3	If answer to 3.2 is yes, please specify	□ Consultation only □ Surgical hemostasis □ Treatment with desmopressin □ Treatment with plasma, platelet or factor concentrates □ Blood (RBC) transfusion	
3.4	How many times in your life did you received any of the above treatments (# 3.3)?	□ 1 - 2 □ 3 to 5 □ 6 to 10 □ more than 10	
3.5	At what age did you first have symptoms?	□ Before 1 year □ Between 1-5 years of age □ Between 6-12 years of age □ Between 13-25 years of age □ After 25 years of age	
3.6	Approximate number of episodes NOT requiring medical attention	□ less than 1 per year □ 1 per year □ 1-5 every six month □ 1-3 every month □ 1 every week	
3.7	Duration of average single episode (min.)	□ 1 to 10 minutes □ more than 10 minutes	

4.	Hematuria

4.1	Have you ever had hematuria ?	□ Yes	□ No skip to 5
4.2	If answer to 4.1 is yes, please specify		
	Presence of associated urologic disease	Yes □ (skip to 5) Specify: □ Stones □ Infection □ Kidney/ bladder disease	No □

Please answer the following questions only for SPONTANEOUS symptoms (answer No to 4.1)

4.3	Have the symptom ever required medical attention ?	Yes □	No □ skip to 4.7
4.4	If answer to 4.3 is yes, please specify	□ Consultation only □ Surgery □ Treatment with desmopressin □ Treatment with plasma, platelet or factor concentrates □ Blood (RBC) transfusion	
4.5	How many times in your life did you received any of the above treatments (# 4.4)?	□ 1 - 2 □ 3 to 5 □ 6 to 10 □ more than 10	
4.6	At what age did you first have symptoms?	□ Before 1 year □ Between 1-5 years of age □ Between 6-12 years of age □ Between 13-25 years of age □ After 25 years of age	
4.7	Approximate number of episodes NOT requiring medical attention	□ less than 1 per year □ 1 per year □ 1-5 every six month □ 1-3 every month □ 1 every week	

5. Gastrointestinal bleeding (Hematemesis, Melena, Hematochezia)

5.1	Have you ever had gastrointestinal bleeding ?	□ Yes	□ No skip to 6
5.2	If answer to 5.1 is yes, please specify		
	Type of bleeding	□ Hematemesis □ Melena □ Hematochezia	
	Presence of associated GI disease	Yes □ Specify: □ Ulcer □ Portal hypertension □ Angiodysplasia	No □

Please answer to the following questions only for SPONTANEOUS symptoms

5.3	Have the symptom ever required medical attention ?	Yes □	No □ skip to 5.7
5.4	If answer to 5.3 is yes, please specify	□ Consultation only □ Surgical haemostasis □ Treatment with desmopressin □ Treatment with plasma, platelet or factor concentrates □ Blood (RBC) transfusion	
5.5	How many times in your life did you received any of the above treatments (# 5.4)?	□ 1 - 2 □ 3 to 5 □ 6 to 10 □ more than 10	
5.6	At what age did you first have symptoms?	□ Before 1 year □ Between 1-5 years of age □ Between 6-12 years of age □ Between 13-25 years of age □ After 25 years of age	
5.7	Approximate number of episodes NOT requiring medical attention	□ less than 1 per year □ 1 per year □ 1-5 every six month □ 1-3 every month □ 1 every week	

6. Oral cavity bleeding (Tooth eruption, spontaneous or after brushing/flossing, gum bleeding, bleeding after bites to lip & tongue)

6.1	Have you ever had oral cavity bleeding ?	□ Yes	□ No or trivial skip to 7
6.2	Have the symptom ever required medical attention ?	Yes □	No □ skip to 6.6
6.3	If answer to 6.2 is yes, please specify	□ Consultation only □ Surgery (dental packing, suture, cauterization) □ Treatment with desmopressin / iron therapy □ Treatment with plasma, platelet or factor concentrates □ Blood (RBC) transfusion	
6.4	How many times in your life did you received any of the above treatments (# 6.3)?	□ 1 - 2 □ 3 to 5 □ 6 to 10 □ more than 10	
6.5	At what age did you first have symptoms?	□ Before 1 year □ Between 1-5 years of age □ Between 6-12 years of age □ Between 13-25 years of age □ After 25 years of age	
6.6	Approximate number of episodes NOT requiring medical attention	□ less than 1 per year □ 1 per year □ 1-5 every six month □ 1-3 every month □ 1 every week	
6.7	Duration of average single episode (min.)	□ 1 to 10 minutes □ more than 10 minutes	

7.	**Bleeding after Tooth/ Teeth extraction**

7.1	Have you ever had bleeding after tooth (teeth) extraction ?	□ Yes	□ No skip to 8
7.2	If answer to 7.1 is yes, please specify		
	Number of extractions	□□	

*Please fill in one of the following forms for **each** tooth extraction*

Age at extraction	□□	Type of extraction	□ Deciduous □ Permanent □ Molar
Actions taken to prevent bleeding		□ None □ Antifibrinolytics □ Desmopressin □ Plasma or clotting factor concentrates □ Platelet infusion	
Bleeding after extraction?		Yes □	No □
Actions taken to control bleeding		□ Resuturing □ Packing □ Antifibrinolytics □ Desmopressin □ Plasma or clotting factor concentrates □ Platelet infusion □ Blood (RBC) transfusion	

8. Bleeding after Surgery or Major Trauma

8.1 Have you ever had bleeding after surgery or major trauma ? ☐ Yes ☐ No, skip to 9

8.2 If answer to 8.1 is yes, please specify

Number of interventions ☐☐

*Please fill in one of the following forms for **each** surgery or major trauma episode*

Age at intervention/trauma ☐☐

Type of surgery

☐ Tonsillectomy/Adenoids
☐ Pharynx/Nose
☐ Major-abdominal
☐ Major-thoracic
☐ Major-gynecology
☐ Other

Actions taken to prevent bleeding

☐ None
☐ Antifibrinolytics
☐ Desmopressin
☐ Plasma or clotting factor concentrates
☐ Platelet infusion

Bleeding after intervention? Yes ☐ No ☐

Actions taken to control bleeding

☐ None
☐ Resuturing
☐ Packing
☐ Antifibrinolytics
☐ Desmopressin
☐ Plasma or clotting factor concentrates
☐ Platelet infusion
☐ Blood (RBC) transfusion

9. Menorrhagia

9.1 Have you ever had very heavy menstrual bleeding (menorrhagia)? □ Yes □ No or trivial skip to 10

If answer to 9.1 is yes, please specify

□ Changing pads/tampons more frequently than every 2 hours

□ Bleeding more than 7 days

□ Clot and flooding

Impairment of daily activities (work, housework, exercise, social activities): □ Never or rarely □ Most menses

9.2 Have the symptom ever required medical attention ? □ Yes □ No skip to 9.6

9.3 If answer to 9.2 is yes, please specify

a □ Consultation only

b □ Pictorial Bleeding Assessment Score _____

c □ Antifibrinolytic therapy

d □ Iron therapy

e □ Hormonal therapy

f □ Combined antifibrinolytics & Hormonal therapy

g □ Hysterectomy / endometrial ablation / D & C

h □ Treatment with desmopressin, plasma or factor concentrates, platelet transfusion

i □ Blood (RBC) transfusion

l □ Hospital admission and emergency treatment

9.4 How many times in your life did you received any of the above treatments (# 9.3 a-l)?

□ 1 - 2
□ 3 to 5
□ 6 to 10
□ more than 10

9.5 At what age did you first have symptoms?

□ At menarche
□ Between 14-25 years of age
□ After 25 years of age

9.6 Have you had time off work/school for menorrhagia?

□ < twice a year
□ > twice a year

9.7 Duration of menorrhagia

□ Since menarche
□ > 12 months
□ < 12 months

9.8 Have you had acute menorrhagia requiring emergency treatment/hospital admission

□ Yes □ No
How many times: ____

10.	Post-partum hemorrhage

10.1	Number of successful pregnancies (live births)	☐☐
10.2	Have you ever had post-partum haemorrhage?	☐ Yes ☐ No or trivial skip to 11
10.3	Did it occur	☐ In the first 24 hours after delivery (Primary) ☐ Between 24 hours and 6 weeks postpartum (Secondary) ☐ Both Primary and Secondary
10.4	How long did vaginal discharge (lochia) last?	☐ < 6 weeks ☐ > 6 weeks
10.5	Did it require changing pads/tampons more frequently than every 2 hours?	☐ Yes ☐ No
10.6	Did this bleeding cause delay of hospital discharge/ readmission to hospital?	☐ Yes ☐ No
10.7	Have the symptom ever required medical treatment?	☐ Yes ☐ No
10.8	If answer to 10.7 is yes, please specify	☐ Consultation only /oxytocin i.v. infusion ☐ Additional uterotonic medications ☐ Iron therapy ☐ Antifibrinolytic therapy ☐ Treatment with plasma or factor concentrates, platelet transfusion ☐ Blood (RBC) transfusion ☐ Any procedure requiring examination under anaesthesia ☐ Uterine balloon/package to tamponade the uterus ☐ Any procedure requiring critical care or surgical intervention (includes: hysterectomy, internal iliac artery legation, uterine artery embolization, uterine brace sutures)
10.9	Number of deliveries that required any of the above treatments (# 10.8)?	☐☐

11. Muscle hematomas or hemarthrosis (spontaneous)

11.1	Have you ever had muscle hematomas or hemarthrosis ?	□ Yes	□ No or trivial skip to 12
11.2	Have the symptom ever required medical attention ?	□ Yes	□ No skip to 11.6
11.3	If answer to 11.2 is yes, please specify	□ Consultation only □ Surgical draining □ Treatment with desmopressin □ Treatment with plasma, platelet or factor concentrates □ Blood transfusion	
11.4	How many times in your life did you receive any of the above treatments (# 11.3)?	□ 1 - 2 □ 3 to 5 □ 6 to 10 □ more than 10	
11.5	At what age did you first have symptoms?	□ Before 1 year □ Between 1-5 years of age □ Between 6-12 years of age □ Between 13-25 years of age □ After 25 years of age	
11.6	Approximate number of episodes NOT requiring medical attention	□ less than 1 per year □ 1 per year □ 1-5 every six month □ 1-3 every month □ 1 every week	

12	Other bleedings

12.1 Have you ever had one of the following?

Excessive umbilical stump bleeding	□ Yes	□ No
Cephalohematoma	□ Yes	□ No
Bleeding at circumcision	□ Yes	□ No
Venipuncture bleeding	□ Yes	□ No
Suction Bleeding	□ Yes	□ No
Ovulation bleeding(in women)	□ Yes	□ No

12.2 Have one of these symptoms ever required medical attention? □ Yes □ No

12.3 If answer to 12.2 is yes, please specify

- □ Consultation only
- □ Surgery
- □ Treatment with desmopressin
- □ Treatment with plasma, platelet or factor concentrates
- □ Blood (RBC) transfusion

12.4 How many times in your life did you receive any of the above treatments (# 12.3) for this symptom?

- □ 1 - 2
- □ 3 to 5
- □ 6 to 10
- □ more than 10

Table 1. Bleeding score

SYMPTOMS (up to the time of diagnosis)	**1.1.1.1** SCORE				
1.1.2	0§	1§	2	3	4
Epistaxis	No/trivial	- > 5/year or - more than 10 minutes	Consultation only*	Packing or cauterization or antifibrinolytic	Blood transfusion or replacement therapy (use of hemostatic blood components and rFVIIa) or desmopressin
Cutaneous	No/trivial	For bruises 5 or more (> 1cm) in exposed areas	Consultation only*	Extensive	Spontaneous hematoma requiring blood transfusion
Bleeding from minor wounds	No/trivial	- > 5/year or - more than 10 minutes	Consultation only*	Surgical hemostasis	Blood transfusion, replacement therapy, or desmopressin
Oral cavity	No/trivial	Present	Consultation only*	Surgical hemostasis or antifibrinolytic	Blood transfusion, replacement therapy or desmopressin

GI bleeding	No/trivial	Present (not associated with ulcer, portal hypertension, hemorrhoids, angiodysplasia)	Consultation only*	Surgical hemostasis, antifibrinolytic	Blood transfusion, replacement therapy or desmopressin
Hematuria	No/trivial	Present (macroscopic)	Consultation only*	Surgical hemostasis, iron therapy	Blood transfusion, replacement therapy or desmopressin
Tooth extraction	No/trivial or none done	Reported in ≤25% of all procedures, no intervention**	Reported in >25% of all procedures, no intervention**	Resuturing or packing	Blood transfusion, replacement therapy or desmopressin
Surgery	No/trivial or none done	Reported in ≤25% of all procedures, no intervention**	Reported in >25% of all procedures, no intervention**	Surgical hemostasis or antifibrinolytic	Blood transfusion, replacement therapy or desmopressin
Menorrhagia	No/trivial	Consultation only* or - Changing pads more frequently than every 2 hours or	- Time off work/school > 2/year or - Requiring antifibrinolytics or	- Requiring combined treatment with antifibrinolytics and hormonal therapy or	- Acute menorrhagia requiring hospital admission and emergency treatment or - Requiring blood transfusion, Replacement therapy,

		- Clot and flooding or - PBAC score>100[#]	hormonal or iron therapy	- Present since menarche and > 12 months	Desmopressin, or - Requiring dilatation & curretage or endometrial ablation or hysterectomy)
Post-partum hemorrhage	No/trivial or no deliveries	Consultation only* or - Use of syntocin or - Lochia > 6 weeks	- Iron therapy or - Antifibrinolytics	- Requiring blood transfusion, replacement therapy, desmopressin or - Requiring examination under anaesthesia and/or the use of uterin balloon/package to tamponade the uterus	- Any procedure requiring critical care or surgical intervention (e.g. hysterectomy, internal iliac artery legation, uterine artery embolization, uterine brace sutures)
Muscle hematomas	Never	Post trauma, no therapy	Spontaneous, no therapy	Spontaneous or traumatic, requiring desmopressin or replacement therapy	Spontaneous or traumatic, requiring surgical intervention or blood transfusion
Hemarthrosis	Never	Post trauma, no therapy	Spontaneous, no therapy	Spontaneous or traumatic, requiring desmopressin or replacement therapy	Spontaneous or traumatic, requiring surgical intervention or blood transfusion
CNS bleeding	Never	-	-	Subdural, any intervention	Intracerebral, any intervention

Other bleedings^	No/trivial	Present	Consultation only*	Surgical hemostasis, antifibrinolytics	Blood transfusion or replacement therapy or desmopressin

In addition to the guidance offered by the table, it is mandatory to refer to the text for more detailed instructions.

§ Distinction between 0 and 1 is of critical importance. Score 1 means that the symptom is judged as present in the patient's history by the interviewer but does not qualify for a score 2 or more

* Consultation only: the patient sought medical evaluation and was either referred to a specialist or offered detailed laboratory investigation

** Example: 1 extraction/surgery resulting in bleeding (100%): the score to be assigned is 2; 2 extractions/surgeries, 1 resulting in bleeding (50%): the score to be assigned is 2; 3 extractions/surgeries, 1 resulting in bleeding (33%): the score to be assigned is 2; 4 extractions/surgeries, 1 resulting in bleeding (25%): the score to be assigned is 1

If already available at the time of collection

^ Include: umbilical stump bleeding, cephalohematoma, cheek hematoma caused by sucking during breast/bottle feeding, conjunctival hemorrhage or excessive bleeding following circumcision or venipuncture. Their presence in infancy requires detailed investigation independently from the overall score.

Table A. The primer pair sequences used in this thesis. SDM: site direct mutagenesis, F: forward, R: reverse.

Primer	Oligo
CD36 - F	5'- TTTAGTATGTGTTAAAATTTCCCAATC
CD36 - R	5'- TGGTTGCTAAAGGATTATGGTAAAG
Vector - insertion boundary - F	5'- CTCAGTGTTGGTGTGGTGATG
Vector - insertion boundary - R	5'- ACCTCCAAACACAGCCAGGACAG
SDM_CD36 F_Deletion	5'- GTATGTACCAAAAAATATTGCTTC
SDM_CD36 R_Deletion	5'- CTATGATGTACAATTTTTTGAGATAATTTTTTC
SDM_CD36 F_Substitution	5'- GTACATCATAGGGTGTGCTAG
SDM_CD36 R_Substitution	5'- AATTTTTTGAGATAATTTTTTCTGTG
SDM_ confirmation_F	5'- AGATGCAGCCTCATTTCCAC
SDM_ confirmation_R	5'- GCAATATTTTTTGGTACATAC
NADK - F	5'- GGAACGTCCGGAAGAAGCAA
NADK - R	5'- GAGGGCAGCGGAAGGATTC
CEBPZ - F	5'- CCGGGTGAAGGCTTTTGTGA
CEBPZ - R	5'- GAGCAACATGCCCACCTAGT
NET1 - F	5'- GAATTCGCTTCCATGACCCC
NET1 - R	5'- AATGGCCGCTCGAATACAGT

Table B. Database of 358 platelet related disease genes.

Genes
ABCA12
ABCB4
ABCC4
ABCG5
ABCG8
ACSL4
ACTN1
ACVRL1
ADAMTS13
ADCY3
ADCY6
ADCY7
ADORA2B
ADRA2A
ADRA2B
ADRBK1
AK3
AKT1
AKT2
ALOX12
ANKRD12
ANKRD18A
ANKRD18B
ANKRD26
ANKRD33
AP3B1
AP3D1
AP3M1
AP3S1
APC
ARHGAP1
ARHGAP17
ARHGAP32
ARHGAP6
ARHGDIA
ARHGDIB
ARHGEF12
ARHGEF3
ARRB1
ASPN
BAK1
BCL2L1
BCOR
BET1L
BLOC1S1
BLOC1S2

BLOC1S3
BLOC1S4
BLOC1S5
BLOC1S6
BMP4
BTBD9
BTK
C14orf133
C19orf55
C20orf42
C6orf25
CD226
CD36
CHD3
CLEC1B
CLEC4F
CNO
Cprin1
CSK
CTTN
CYCS
DAAM1
DIAPH1
DIAPH2
DIAPH3
DNAH11
DNM1L
DNM2
DNM3
DTNBP1
EFNB1
EPHA4
EPHB1
ERG
ETS1
ETV6
EVI1
EXOC1
F2R
F2RL3
FARP2
FCER1G
FCGR2A
FERMT3
FGD3
FGR
FHOD1
FLI1

FLII
FLNA
FMNL1
FMNL3
FYN
GATA1
GDI2
GFI1
GFI1B
GNA12
GNA13
GNAI1
GNAI2
GNAQ
GNAZ
GNB2
GNB3
GNG11
GNG12
GNG13
GNG5
GP1BA
GP1BB
GP5
GP6
GP9
GRAP2
GRB2
GRK5
GRK6
GUCY1A3
GUCY1B3
HBB
HOOK3
HOXA11
HPS1
HPS4
HTR2A
INPP5D
ITGA2
ITGA2B
ITGA5
ITGB1
ITGB3
ITPR1
JAK2
JMJD1C
KIAA1109

KIAA2018
LAIR1
LAT
LCP2
LPAR1
LTBP1
LY6G6F
LYN
LYST
MAP2K2
MAP2K4
MAPK1
MAPK13
MAPK14
MAPK8
MDS1
MKL1
MLK1
MLPH
MMP17
MNX1
MPL
MRPS34
MUC16
MUC2
MUTED
MYB
MYH10
MYH13
MYH9
MYL9
MYLK
MYLK2
MYO18B
MYO3A
MYO5A
MYO5B
NAPA
NAPG
NBEA
NBEAL2
NFE2
NIPSNAP3A
NOTCH1
NOX1
NRG3
NSF
NXF1

ORAI1
P2RX1
P2RY1
P2RY12
P2RY13
PDE2A
PDE3A
PDE4D
PDE5A
PDPK1
PDZD3
PDZK1
PEAR1
PECAM1
PGM3
PHOX2A
PIK3CA
PIK3CB
PIK3CD
PIK3CG
PIK3R1
PIK3R3
PIK3R5
PLA2G4A
PLA2G4C
PLCB2
PLCB3
PLCG2
PLDN
PPP1CA
PPP1CB
PPP1CC
PPP1R12A
PPP1R12C
PPP1R14A
PPP1R2
PRKACA
PRKACB
PRKACG
PRKAR1A
PRKAR2A
PRKCA
PRKCB
PRKCD
PRKCQ
PRKD1
PRKG1
PRKG2

PTEN
PTGIR
PTGS1
PTK2
PTPN1
PTPN11
PTPN12
PTPN18
PTPN2
PTPN6
PTPN7
PTPN9
PTPRA
PTPRC
PTPRJ
RAB27A
RAB27B
RAB38
RAB4A
RABGGTA
RAC1
RAF1
RAI1
RAP1B
RAP1GAP
RAP1GAP2
RAP1GDS1
RASGRP2
RBM8A
RGS10
RGS18
RGS19
RGS20
RGS9
RHOA
RHOC
RHOF
ROCK1
ROCK2
RUNX1
SCAMP2
SCAMP5
SCFD1
SELP
SERPINE2
SH2B3
SIRPA
SLC35D3

SLC9A3R1
SLC9A3R2
SLFN14
SMAD1
SMAD6
SNAP23
SNAP25
SNAP29
SNAPIN
SNX1
SRA1
SRC
SRF
STIM1
STOM
STX11
STX12
STX2
STX4
STX6
STX7
STXBP1
STXBP2
STXBP3
STXBP4
STXBP5L
STXBP6
SUZ12
SYK
SYTL3
SYTL4
TAL1
TAOK1
TBXA2R
TEC
TGFBR3
THPO
TLN1
TLR2
TMCC2
TPM1
TPM4
TRAF4
TREML1
TRPM7
TTC37
TUBA3C
TUBB1

UNC13A
UNC13B
VAMP2
VAMP3
VAMP7
VAMP8
VAV1
VAV2
VAV3
VPS11
VPS16
VPS18
VPS33A
VPS33B
VPS39
VPS41
VPS4B
VPS52
VPS8
VWF
WAS
WDR66
ZFPMI

Table C. Candidate known inherited bleeding genes used to filter plausible candidate gene variants from the WES data of patients analysed within the Congenica software.

ABCG5
ABCG8
ACTB
ACTN1
ACVRL1
ADAMTS13
ANKRD26
ANO6
AP3B1
AP3D1
ARPC1B
BLOC1S3
BLOC1S6
CDC42
CHST14
COL1A1
COL5A1
COL5A2
CYCS
DIAPH1
DTNBP1
ENG
ETV6
F10
F11
F12
F13A1
F13B
F2
F5
F7
F8
F9
FERMT3
FGA
FGB
FGG
FLI1
FLNA
FYB1
GATA1
GBA
GFI1B
GGCX

GNE
GP1BA
GP1BB
GP6
GP9
HOXA11
HPS1
HPS3
HPS4
HPS5
HPS6
HRG
IKZF5
ITGA2B
ITGB3
KDSR
KLKB1
KNG1
LMAN1
LYST
MCFD2
MECOM
MPIG6B
MPL
MYH9
NBEA
NBEAL2
P2RY12
PLA2G4A
PLAT
PLAU
PLG
PROC
PROS1
PTPN11
RASGRP2
RBM8A
RNU4ATAC
RUNX1
SERPINC1
SERPIND1
SERPINE1
SERPINF2
SLC45A2
SLFN14
SMAD4
SRC
STIM1

STXBP2
TBXA2R
TBXAS1
THBD
THPO
TUBB1
VIPAS39
VKORC1
VPS33B
VWF
WAS
LAT
ORAI1
PTPRJ
SLC35A1
TPM4
APOH
COL2A1
HABP2
IFNAR2
LPA
MTHFR
PIGA
PROCR
PROZ
SERPINA10
TFPI

Table D. Sequencing parameters of the patients found with SNVs using Whole exome (WES) data and processed with Congenica software. Software predicts Grantham score based on amino acid substitution, Haploinsufficiency score (HI) and GERP scores as a measure of conservation. Read depth sequencing coverage including Reads split and total depth of each patient from WES data, which was observed at the site of each variation across all DNA samples analysed. Subsequently the Exome Depth integrated tool within Congenica was used to determine copy number variation based on read coverage.

Patient	Gene(s)	Grantham	HI	GERP	Reads split	Depth
2.1	*RUNX1*	NA	0.748	5.31	45/29	74
2.2	*RUNX1*	NA	0.748	5.31	40/38	78
3.1	*SLFN14*	152 (Radical)	NA	4.63	159/160	319
3.2	*SLFN14*	152 (Radical)	NA	4.63	109/105	214
4.1	*FLI1*	NA	0.551	5.2	92/59	151
4.2	*FLI1*	NA	0.551	5.2	60/68	128
5	*FGA*	58 (Moderately conservative)	0.179	3.47	111/119	230
6.2	*MPIG6B*	215 (Radical)	0.064	4.79	40/56	96
	VPS33B	145 (Moderately radical)	0.052	5.54	55/45	100
7	*ETV6*	NA	0.928	5.78	78/50	128
8	*VWF*	NA	0.237	3.84	51/56	107
	ANKRD26	56 (Moderately conservative)	0.25	5.64	127/123	250
9	*SLC45A2*	125 (Moderately radical)	0.093	5.81	88/81	169
10.1	*HPS3*	46 (Conservative)	0.481	2.33	71/82	153
10.2	*HPS3*	46 (Conservative)	0.481	2.33	87/82	169
11	*LYST*	103 (Moderately radical)	0.162	3.86	76/76	152
	AP3D1	56 (Moderately conservative)	0.27	5.98	15/18	33
12	*COL5A2*	94 (Moderately conservative)	0.408	5.98	15/18	33
13	*F7*	64 (Moderately conservative)	0.087	1.73	41/33	74
14.1	*F10*	94 (Moderately conservative)	0.118	5.38	58/66	124
	NBEAL2	43 (Conservative)	NA	3.76	95/83	178
	GBA	46 (Conservative)	0.08	3.53	125/122	247
14.2	*F10*	94 (Moderately conservative)	0.118	5.38	41/48	89
	NBEAL2	43 (Conservative)	NA	3.76	83/77	160
	GBA	46 (Conservative)	0.08	3.53	107/111	218
15.1	*VWF*	43 (Conservative)	0.237	5.56	64/69	133
	NBEA	NA	0.437	3.12	7/5	12
15.2	*VWF*	43 (Conservative)	0.237	5.56	86/80	166

	NBEA	NA	0.437	3.12	5/4	9
16	*ACVRL1*	43 (Conservative)	0.935	4.79	28/17	45
	RUNX1	NA	0.748	3.8	21/27	48
	ITGB3	101 (Moderately radical)	0.495	4.9	16/9	28
17	*RUNX1*	NA	0.748	4.44	22/21	43
	F11	NA	0.041	4.92	82/63	145
	SERPINC1	99 (Moderately conservative)	0.357	5.74	75/47	122
	F13A1	110 (Moderately radical)	0.315	1.71	66/70	136
18	*PTPN11*	23 (Conservative)	0.225	4.5	69/42	111
19	*FGB*	98 (Moderately conservative)	0.647	5.8	52/47	99
20	*GP6*	180 (Radical)	0.045	3.04	27/26	53
	THBD	98 (Moderately conservative)	0.21	4.18	27/28	55
21	*PLG*	43 (Conservative)	0.185	3.6	34/26	60
	ARPC1B	29 (Conservative)	0.375	4.51	45/48	93
22.1	*P2RY12*	29 (Conservative)	0.059	5.41	52/48	100
22.2	*P2RY12*	29 (Conservative)	0.059	5.41	48/63	111
23	*P2RY12*	38 (Conservative)	0.059	5.48	48/68	116
	MCFD2	64 (Moderately conservative)	0.143	5.52	32/35	67
24	*VPS33B*	46 (Conservative)	0.052	5.54	23/26	49
	ITGB3	102 (Moderately radical)	0.495	5.88	43/51	94
	LYST	26 (Conservative)	0.162	5.64	18/17	35
25	*RUNX1*	56 (Moderately conservative)	0.748	5.31	11/9	20
	VWF	103 (Moderately radical)	0.237	202	47/42	89
26	*RUNX1*	152 (Radical)	0.748	5.12	104/81	185
	SMAD4	NA	1	6.17	68/7	75
	SMAD4	NA	1	6.17	66/2	68
	GGCX	29 (Conservative)	0.122	5.78	86/69	155
27	*TUBB1*	29 (Conservative)	0.212	4.56	71/57	128
	TPM4	64 (Moderately conservative)	0.198	4.76	36/27	63
28	*PLAT*	180 (Radical)	0.189	3.54	82/53	135
29.1	*TUBB1*	101 (Moderately radical)	0.212	5.19	208/209	417
29.2	*TUBB1*	101 (Moderately radical)	0.212	5.19	204/229	433
30	*GP1BA*	NA	0.104	-6.29	--	34
	GP1BA	NA	0.104	-0.47	37/17	54
	GBA	NA	0.08	3.66	57/68	125
31	*MPIG6B*	NA	0.064	4.11	86/81	167

	HPS6	NA	0.085	2.88	11/8	19
32.1	*CYCS*	64 (Moderately conservative)	0.341	5.07	25/15	40
	GGCX	29 (Conservative)	0.122	5.54	103/120	223
32.1	*CYCS*	64 (Moderately conservative)	0.341	5.07	20/18	38
	GGCX	29 (Conservative)	0.122	5.54	134/137	271
33	*COL5A1*	77 (Moderately conservative)	0.627	4.47	114/101	215
	TUBB1	NA	0.212	4.58	47/33	80
34	*KLKB1*	22 (Conservative)	0.375	5.1	41/35	76
	NBEAL2	43 (Conservative)	NA	1.5	49/44	93
35	*COL5A1*	83 (Moderately conservative)	0.627	4.68	25/24	49
	TBXAS1	121 (Moderately radical)	0.103	3.74	27/35	62
	THBD	61 (Moderately conservative)	0.21	0.0337	20/14	34
36	*RASGRP2*	98 (Moderately conservative)	0.324	3.38	13/16	29
	VWF	22 (Conservative)	0.237	4.41	27/22	49
	VWF	81 (Moderately conservative)	0.237	4.21	30/29	59
	GP1BA	24 (Conservative)	0.104	-1.62	133/132	265
37.1	*GP1BA*	109 (Moderately radical)	0.104	3.7	119/120	239
37.2	*GP1BA*	109 (Moderately radical)	0.104	3.7	121/146	267
37.3	*MYH9*	180 (Radical)	0.259	4.92	28/48	76
38.1	*RUNX1*	43 (Conservative)	0.748	5.12	87/86	173
38.2	*RUNX1*	43 (Conservative)	0.748	5.12	74/65	139
38.3	*WAS*	56 (Moderately conservative)	0.855	4.7	36/38	74
38.4	*WAS*	56 (Moderately conservative)	0.855	4.7	23/31	54
38.5	*WAS*	56 (Moderately conservative)	0.855	4.7	0/37	37
41	*ACTN1*	125 (Moderately radical)	0.635	5	19/19	38
	PLAT	60 (Moderately conservative)	0.189	5.5	13/9	23
42	*ABCG8*	58 (Moderately conservative)	0.113	1.7	35/27	35/27
43	*RUNX1*	NA	0.748	5.04	56/72	128
	VWF	180 (Radical)	0.237	5.19	17/14	31
	GBA	NA	0.08	3.17	83/64	147
	PROS1	NA	0.097	3.42	35/22	57
44.1	*RUNX1*	NA	0.748	4.44	19/27	46
44.2	*RUNX1*	NA	0.748	4.44	29/20	49
45	*HRG*	43 (Conservative)	0.075	2.54	54/42	96
46	*FLNA*	145 (Moderately radical)	0.752	5.69	0/113	113
47	*GGCX*	29 (Conservative)	0.122	5.54	68/75	143

48	*RUNX1*	58 (Moderately conservative)	0.748	5.12	92/94	186
	SLFN14	74 (Moderately conservative)	NA	5.38	500/67	567
	SLFN14	43 (Conservative)	NA	0.449	1181/328	1509
	SLFN14	56 (Moderately conservative)	NA	4.71	971/198	1169
	SLFN14	NA	NA	3.03	362/174	536
49	*ABCG5*	89 (Moderately conservative)	0.088	5.77	53/49	102
50.1	*STXBP2*	103 (Moderately radical)	0.495	4.21	74/43	117
51.1	*ADAMTS13*	NA	0.059	4.04	8/4	12
51.2	*ADAMTS13*	NA	0.059	4.04	15/16	31
51.3	*ADAMTS13*	NA	0.059	4.04	12/10	22
52.1	*LYST*	NA	NA	3.61	235/247	482
	THBD	NA	0.21	0.572	100/115	215
52.2	*LYST*	NA	0.162	3.61	225/222	447
	THBD	NA	0.21	0.572	92/90	182
53.1	*ADAMTS13*	NA	0.059	1.75	23/29	52
53.2	*ADAMTS13*	NA	0.059	1.75	22/28	50
55	*F10*	23 (Conservative)	0.118	1.73	241/207	448
	MPL	159 (Radical)	0.66	1.07	85/81	166
56	*GFI1B*	NA	0.481	4.65	184/179	363
57	*ACVRL1*	64 (Moderately conservative)	0.935	3.52	48/33	81
	ITGA2B	NA	0.333	4.51	34/29	63
	THPO	NA	0.158	4.1	34/29	198
	HOXA11	194 (Radical)	0.321	4.1	74/86	160
58.1	*NBEAL2*	NA	NA	5.01	98/113	211
58.2	*NBEAL2*	NA	NA	5.01	122/103	225
59	*MECOM*	160 (Radical)	0.867	4.38	73/56	129
	F8	152 (Radical)	0.086	3.78	54/52	106
60	*MPL*	56 (Moderately conservative)	0.66	2.41	110/59	169
	NBEAL2	21 (Conservative)	NA	4.09	139/114	253
	PROZ	89 (Moderately conservative)	0.092	4.02	18/18	36
61	*COL5A1*	56 (Moderately conservative)	0.627	5.18	80/84	164
63.1	*GFI1B*	205 (Radical)	0.481	5.08	77/71	148
63.2	*GFI1B*	205 (Radical)	0.481	5.08	81/63	144
63.3	*GFI1B*	205 (Radical)	0.481	5.08	88/76	164
65.1	MECOM	94 (Moderately conservative)	0.867	4.81	22/22	44
	SLFN14	81 (Moderately conservative)		3.39	31/14	45

65.2	MECOM	94 (Moderately conservative)	0.867	4.81	27/23	50
	SLFN14	81 (Moderately conservative)		3.39	30/33	63
66.1	MYH9		0.259	4.85	23/14	37
66.2	MYH9		0.259	4.85	14/15	29
68	RASGRP2	180 (Radical)	0.324	4.83	51/49	100
69	F10	26 (Conservative)	0.118	-2.49	13/13	26
70.1	ETV6		0.928	4.73	13/15	28
71	RUNX1	109 (Moderately radical)	0.748	5.08	24/5	29
	RUNX1					
72	FGG	60 (Moderately conservative)	0.335	5.07	55/60	115
	STXBP2	101 (Moderately radical)	0.495	4.3	38/31	69
	TUBB1	101 (Moderately radical)	0.212	5.19	48/48	96
74	RUNX1	107 (Moderately radical)	0.748	4.97	16/7	23
	COL5A1	65 (Moderately conservative)	0.627	4.72	58/80	138
75	GFI1B	205 (Radical)	0.481	5.08	22/19	41
76	THBD	64 (Moderately conservative)	0.21	-0.851	15/11	26
77	THBD	94 (Moderately conservative)	0.21	5.06	13/22	35
	COL5A2	64 (Moderately conservative)	0.408	5.38	24/30	54
78	STXBP2	81 (Moderately conservative)	0.495	4.4	33/22	55
	MCFD2	23 (Conservative)	0.143	5.52	65/47	112
79	NBEAL2	29 (Conservative)		3.51	21/32	53
	AP3D1	58 (Moderately conservative)	0.27	4.57	56/41	97
80	RUNX1	99 (Moderately conservative)	0.748	4.97	9/10	19
	MPL	29 (Conservative)	0.66	5.56	51/48	99
81	AP3B1	101 (Moderately radical)	0.335	2.08	47/43	90
82	TUBB1	180 (Radical)	0.212	3.97	59/42	101
	TUBB1	81 (Moderately conservative)	0.212	4.36	45/50	95
83	LPA		0.066	1.65	58/34	92
84.1	F5	32 (Conservative)	0.147	1.34	79/44	123
	F5	58 (Moderately conservative)	0.147	2.24	47/39	86
84.2	NBEAL2	29 (Conservative)		4.49	6/6	12
	HPS5	10 (Conservative)	0.093	3.05	21/33	54

Table E. Criteria for classifying pathogenic variants.

Evidence of pathogenicity	Category
Very strong	**PVS1** null variant (nonsense, frameshift, canonical ±1 or 2 splice sites, initiation codon, single or multiexon deletion) in a gene where LOF is a known mechanism of disease Caveats: ▪ Beware of genes where LOF is not a known disease mechanism (e.g., GFAP, MYH7). ▪ Use caution interpreting LOF variants at the extreme 3′ end of a gene. ▪ Use caution with splice variants that are predicted to lead to exon skipping but leave the remainder of the protein intact. ▪ Use caution in the presence of multiple transcripts.
Strong	**PS1** Same amino acid change as a previously established pathogenic variant regardless of nucleotide change. Example: Val→Leu caused by either G>C or G>T in the same codon. Caveat: Beware of changes that impact splicing rather than at the amino acid/protein level. **PS2** De novo (both maternity and paternity confirmed) in a patient with the disease and no family history Note: Confirmation of paternity only is insufficient. Egg donation, surrogate motherhood, errors in embryo transfer, and so on, can contribute to nonmaternity. **PS3** Well-established in vitro or in vivo functional studies supportive of a damaging effect on the gene or gene product. Note: Functional studies that have been validated and shown to be reproducible and robust in a clinical diagnostic laboratory setting are considered the most well established. **PS4** The prevalence of the variant in affected individuals is significantly increased compared with the prevalence in controls. Note 1: Relative risk or OR, as obtained from case–control studies, is >5.0, and the confidence interval around the estimate of relative risk or OR does not include 1.0. See the article for detailed guidance.

	Note 2: In instances of very rare variants where case–control studies may not reach statistical significance, the prior observation of the variant in multiple unrelated patients with the same phenotype, and its absence in controls, may be used as moderate level of evidence.
Moderate	**PM1** Located in a mutational hot spot and/or critical and well-established functional domain (e.g., active site of an enzyme) without benign variation. **PM2** Absent from controls (or at extremely low frequency if recessive) in Exome Sequencing Project, 1000 Genomes Project, or Exome Aggregation Consortium. Caveat: Population data for insertions/deletions may be poorly called by next-generation sequencing. **PM3** For recessive disorders, detected in trans with a pathogenic variant. Note: This requires testing of parents (or offspring) to determine phase. **PM4** Protein length changes as a result of in-frame deletions/insertions in a nonrepeat region or stop-loss variants. **PM5** Novel missense change at an amino acid residue where a different missense change determined to be pathogenic has been seen before. Example: Arg156His is pathogenic; now you observe Arg156Cys. Caveat: Beware of changes that impact splicing rather than at the amino acid/protein level. **PM6** Assumed de novo, but without confirmation of paternity and maternity.
Supporting	**PP1** Cosegregation with disease in multiple affected family members in a gene definitively known to cause the disease. Note: May be used as stronger evidence with increasing segregation data. **PP2** Missense variant in a gene that has a low rate of benign missense variation and in which missense variants are a common mechanism of disease. **PP3** Multiple lines of computational evidence support a deleterious effect on the gene or gene product (conservation, evolutionary, splicing impact, etc.).

	Caveat: Because many in silico algorithms use the same or very similar input for their predictions, each algorithm should not be counted as an independent criterion. PP3 can be used only once in any evaluation of a variant. **PP4** Patient's phenotype or family history is highly specific for a disease with a single genetic etiology. **PP5** Reputable source recently reports variant as pathogenic, but the evidence is not available to the laboratory to perform an independent evaluation.

Table F. Criteria for classifying benign variants.

Evidence of benign impact	**Category**
Stand-alone	**BA1** Allele frequency is >5% in Exome Sequencing Project, 1000 Genomes Project, or Exome Aggregation Consortium.
Strong	**BS1** Allele frequency is greater than expected for disorder. **BS2** Observed in a healthy adult individual for a recessive (homozygous), dominant (heterozygous), or X-linked (hemizygous) disorder, with full penetrance expected at an early age. **BS3** Well-established in vitro or in vivo functional studies show no damaging effect on protein function or splicing. **BS4** Lack of segregation in affected members of a family. Caveat: The presence of phenocopies for common phenotypes (i.e., cancer, epilepsy) can mimic lack of segregation among affected individuals. Also, families may have more than one pathogenic variant contributing to an autosomal dominant disorder, further confounding an apparent lack of segregation.
Supporting	**BP1** Missense variant in a gene for which primarily truncating variants are known to cause disease. **BP2** Observed in trans with a pathogenic variant for a fully penetrant dominant gene/disorder or observed in cis with a pathogenic variant in any inheritance pattern. **BP3** In-frame deletions/insertions in a repetitive region without a known function. **BP4** Multiple lines of computational evidence suggest no impact on gene or gene product (conservation, evolutionary, splicing impact, etc.).

	Caveat: Because many in silico algorithms use the same or very similar input for their predictions, each algorithm cannot be counted as an independent criterion. BP4 can be used only once in any evaluation of a variant. **BP5** Variant found in a case with an alternate molecular basis for disease. **BP6** Reputable source recently reports variant as benign, but the evidence is not available to the laboratory to perform an independent evaluation. **BP7** A synonymous (silent) variant for which splicing prediction algorithms predict no impact to the splice consensus sequence nor the creation of a new splice site AND the nucleotide is not highly conserved.