Identification and Characterisation of the Underlying Defects in Patients with Inherited Platelet Bleeding Disorders

By:

Maryam Ahmed Aldossary

A thesis submitted in partial fulfilment of the requirements for the degree of Doctor of Philosophy

The University of Sheffield
Faculty of Medicine, Dentistry and Health
Department of Infection, Immunity & Cardiovascular Disease

Submission Date

January 2019

Abstract

The underlying genetic defects remain unknown in about 50% of inherited platelet bleeding disorders (IPDs). This study investigated the use of whole exome sequencing (WES) to identify candidate gene defects in 34 index cases enrolled in the UK Genotyping and Phenotyping of Platelets study with a history of bleeding, whose platelets demonstrated defects in agonist-induced dense granule secretion or Gi-signalling. WES analysis identified a median of 98 candidate disease-causing variants per index case highlighting the complexity of IPDs. Sixteen variants were in genes that had previously been associated with IPDs, two of which were selected for further characterisation. Two novel *FLI1* alterations, predicting p.Arg340Cys/His substitutions in the DNA binding domain of FLI1 were shown to reduce transcriptional activity and nuclear accumulation of FLI1, suggesting that these variants interfere with the regulation of essential megakaryocyte genes by FLI1 and may explain the bleeding tendency in affected patients. Expression of a novel truncated p.Arg430* variant of ETV6 revealed it to be stably expressed, possessing normal repressor activity in HEK 293T cells and a slight reduction in repressor activity in megakaryocytic cells. Further studies are required to confirm the pathogenicity of this variant. To identify novel genes involved in platelet granule biogenesis and secretion, gene expression was examined in megakaryocytic cells before and after knockdown of *FLI1*, defects in which are associated with platelet granule abnormalities. Comparison of the gene expression data with that from platelets from patients with *FLI1* defects and with the results of WES analysis in patients with secretion defects highlighted several genes of interest, including *C18orf32*, *IGFBP2, PLCG2*, *SCFD2, SLC24A3, ST8SIA6* and *ZBTB45* which, between them, harboured ten candidate causative variants among the patients with defects in platelet secretion. Further work is warranted to explore the contribution of these genes to platelet secretory pathways.

Acknowledgements

Firstly, I would like to express my genuine appreciation to my supervisor Prof Martina Daly for her help. Her guidance, support, motivation, enthusiasm and patience helped me at each step during the last four years at both the academic and non-academic level. I could not have imagined having a better advisor and mentor for my PhD study. Also, I would like to thank Prof Michael Makris for his insightful comments and encouragement and Dr Vincenzo Leo for his patience and valuable support, especially during the first year. With a special mention, I also would like to thank Dr Simon Webster for his help, suggestions, and insight, which drove me to think deeply during our discussions in the lab. Also, I would like to thank all the members of the Haemostasis group, including Dr Dan Hampshire, Prof Anne Goodeve and Ian Peake for their comments, questions and support, and for all the laughs and fun that we have had during the last four years. It was fantastic to have had the opportunity to be a member of the Sheffield Haemostasis group.

I would also like to express my appreciation for all those from outside the University of Sheffield who helped in various aspects during my PhD: The UK-GAPP study coordinator Dr Neil V. Morgan from the University of Birmingham, UK, and all members of the UK-GAPP study group; Our collaborators Dr Courtney D. Thornburg and Diane Masser-Frye from Rady Children's Hospital, San Diego, USA, who identified and provided details of the second FLI1 patient; Dr Michael Simpson of King's College London, UK, who carried out the whole exome sequencing. Additionally, Dr Maria Trojanowska from Boston University, USA, for providing us with the FLI1 plasmid and Dr Leila Noetzli from the University of Colorado Anschutz Medical Campus, USA, and Michael Y. Zhang from the University of Washington, USA, for providing the ETV6 plasmids.

Equally, I am grateful for all the support that was provided from within the University of Sheffield. I would like to thank all the core facilities for their support, especially Sheffield Microarray and Next Generation Sequencing Core Facility in the Sheffield Institute for Translational Neuroscience (SITRAN), the Medical School Flow Cytometry Core Facility, Wolfson Light Microscopy Facility, and the Core Genomic Facility. Additionally, special thanks to all the technical staff, the PGR team, the medical IT team, the administrative staff and the members of the medical school. Also, I would like to thank my colleagues in our student office and all other students for their chats, smiles, help and a few µl of reagents.

Additionally, I would like to acknowledge my sponsors, the "Imam Abdurrahman bin Faisal University" and the British Heart Foundation, who initially funded the UK-GAPP study. Also, I would like to thank the Saudi Arabian Cultural Bureau and the Embassy of Saudi Arabia in the UK for their efforts. I am grateful to Dr Layla Bashawri for introducing me to research, to Dr Sarah Alshamari for making this possible and to Dr Faisal Alzahrani for his encouragement in taking this step.

Last but not least, I would like to express my great love, debt and appreciation for my mother Hussah and my father Ahmed, who kept pushing me towards this point. Additionally, I would like to thank my sisters, brothers and aunts and all my family and friends who have supported me along the way and in my life in general. I should also not forget my husband and my two lovely daughters for being there for me.

Dedication

"This is dedicated …

To my grandpa Abdul-Aziz. Although it has been three years since you left us, I still do not know how I will keep visiting that house and seeing your favourite place on that couch empty. I miss you being around, giving your blessing, and being a part of all the special moments in our lives. Of all the times, 9pm every day, the time that we expected you to return from your office, is one of the most difficult times, without hearing the noise of the garage door opening, seeing your car approaching through that large window in the living room, and listening to your short steps before you opened the door with a smile and carrying a bag of hot fresh bread. I really miss you a lot and everything that left with you.

To my beloved mum Hussah, who keeps motivating me throughout my life. To my dear dad Ahmed, who keeps making me dream. Your big heart that gives unconditional love, your ultimate care, attention and support that include even all the minor details in my life, your strictness and passion about my education, you are one of a kind. I could not have any better parents and I wish to give my kids part of what you gave me.

To you.

Because I owe it all to you. Many Thanks"

Table of Contents

List of Figures

List of Tables

Appendices

List of Supplementary Data

Supplementary Data #1. List of unique genes identified by whole exome sequencing in a subgroup of index cases having platelet Gi-signalling defect

Supplementary Data #2. List of unique genes identified by whole exome sequencing in a subgroup of index cases having platelets secretion defect

Supplementary Data #3. List of differentially express coding transcripts in wild-type differentiated Dami cells relative to untreated cells

Supplementary Data #4. List of functional annotation clusters of Gene Ontology terms for genes differentially expressed in wild-type differentiated Dami cells relative to untreated cells

Supplementary Data #5. List of differentially express coding transcripts in *FLI1* knockdown Dami cells relative to wild-type before differentiation

Supplementary Data #6. List of differentially express coding transcripts in *FLI1* knockdown Dami cells relative to wild-type following differentiation

Supplementary Data #7. List of functional annotation clusters of Gene Ontology terms for genes differentially expressed in *FLI1* knockdown Dami cells relative to wild-type

Supplementary Data #8. List of functional annotation clusters of Gene Ontology terms for genes identified by whole exome sequencing which are differentially expressed in FLI1-deficient platelets

Supplementary Data #9. List of functional annotation clusters of Gene Ontology terms for genes differential expression in FLI1-deficient platelets and in response to *FLI1* knockdown

List of abbreviations

ADP	Adenosine diphosphate
ATP	Adenosine triphosphate
BM	Bone marrow
BSA	Bovine serum albumin
CADD	Combined Annotation Dependent Depletion
Cas	CRISPR associated protein
CID	C-terminal inhibitory domain
CRISPR	Clustered regularly interspaced short palindromic repeats
CRP	Collagen-related peptide
DAG	Diacylglycerol
DAPI	4',6-diamidino-2-phenylindole
DAVID	Database for Annotation, Visualization and Integrated Discovery
DMEM	Dulbecco's modified Eagle's medium
DMSO	Dimethyl sulfoxide
DSB	double-stranded break
EGFP	Enhanced green fluorescent protein
ETS	E26 transformation-specific or E-twenty-six
EV	Empty vector
FACS	Fluorescence-activated cell sorting
FBS	Foetal bovine serum
GAPP	Genotyping and Phenotyping of Platelets study
gDNA	Genomic DNA
GFP	Green fluorescent protein
GO	Gene Ontology
GP	Glycoprotein
GPCR	G-protein-coupled receptor
GPS	Grey platelet syndrome
gRNA	Guide RNA
HEK 293T	Human embryonic kidney 293T
HPS	Hermansky-Pudlak syndrome
Hr	Hour
HSC	Haematopoietic stem cells
IP3	Inositol trisphosphate
IPD	Inherited platelet bleeding disorder
Kb	Kilobases
kDa	Kilodalton
LDS	Lithium dodecyl sulphate
LID	Linker inhibitory damper domain
LTA	Light transmission aggregometry
MAF	Minor allele frequency
MAPK	Mitogen-activated protein kinase
Min	Minute
MK	Megakaryocyte
MKRP	MK repopulating progenitor
MPV	Mean platelet volume
NES	Nuclear export signal

NGS	Next-generation sequencing
NLS	Nuclear localisation signal
PAGE	Polyacrylamide gel electrophoresis
PAM	proto-spacer adjacent motif
PAR	Protease-activated receptor
PBS	Phosphate buffered saline
PCR	Polymerase chain reaction
PI3K	Phosphoinositide 3-kinase
PIC	Protease Inhibitor Cocktail
PLC	Phospholipase C
PMA	Phorbol 12-myristate 13-acetate
PNT	Pointed N-terminal domain
PTS	Paris-Trousseau syndrome
qPCR	Quantitative polymerase chain reaction
RIPA	Radioimmunoprecipitation assay buffer
rpm	Revolutions per minute
RQ	Relative quantity
Sec	Second
SEM	The standard error of the mean
SFK	Src family kinase
SNARE	Soluble N-ethylmaleimide-sensitive fusion attachment protein receptor
SNV	Single nucleotide variant
TBA	Tris-borate acetate
TBP	TATA-binding protein
THC5	Thrombocytopenia 5
TPO	Thrombopoietin
TXA2	Thromboxane A2
VWF	von Willebrand factor
WES	Whole exome sequencing
WT	Wild-type

Poster and oral presentations arising from this work

Poster in the University of Sheffield Medical School Research Meeting June 16-17 2016 "Use of whole exome sequencing to investigate patients with unexplained platelet bleeding disorders; Identification and characterisation of a novel *FLI1* variant in patients with dense granule secretion defects**"**.

Poster in the University of Sheffield Medical School Research Meeting June 15-16 2017 "Characterisation of Two Novel FLI1 Variants causing Substitution of Arginine 340 in the ETS Domain of FLI1 in Patients with Platelet Dense Granule Secretion Disorders**"**.

Oral communication in the 2016 British Society for Haemostasis and Thrombosis (BSHT), Anticoagulation in Practice (AiP) & UK Platelet Group Annual Meeting, Leeds, UK, 10-11-16 "Use of whole exome sequencing to investigate patients with unexplained platelet bleeding disorders: Identification and characterization of a novel FLI1 variant in a patient with a dense granule secretion defect**"**.

Oral communication in International Society on Thrombosis and Haemostasis (ISTH) 2017 Congress, Berlin, Germany, 12-7-17 "Characterisation of Two Novel FLI1 Variants causing Substitution of Arginine 340 in the ETS Domain of FLI1 in Patients with Platelet Dense Granule Secretion Disorders**" OC 59.4.**

1 Chapter 1. Introduction

1.1 Introduction

1.1.1 Megakaryopoiesis and platelet production

Platelets are the second most abundant cell in blood, with counts normally ranging from 150 to 350 x 10^9/L in the general population (Kaushansky, 2009). They are small anucleate discoid cells derived from megakaryocytes (MKs) and play a critical role in primary haemostasis by accumulating at sites of vascular injury and mediating formation of a fibrin-rich platelet plug that stems bleeding. MKs account for less than 1% of all cells in the bone marrow (BM) but, through the process of megakaryopoiesis, they produce the 10^{11} platelets required each day to maintain normal platelet counts (Kaushansky, 2009). As shown in Figure 1.1, during conventional megakaryopoiesis, haematopoietic stem cells (HSC) are directed to differentiate sequentially through multipotent progenitors, common myeloid progenitors, MK-erythroid progenitors, MK progenitors, to mature MKs that release platelets. However, non-conventional models of megakaryopoiesis have also been described (shown in Figure 1.1). HSC that exhibit platelet-biased long-term reconstitution capacity (Plt-LT-HSC) are capable of differentiating directly to MKs without passing through the MK-erythroid progenitor stage while maintaining their self-renewal capacity and ability to generate multiple blood lineage cells (Sanjuan-Pla et al., 2013). Additionally, HSC progenitors called MK repopulating progenitors (MKRPs), that are proposed to derive from Plt-LT-HSC and maintain their self-renewal capacity, can directly differentiate into MK progenitors (Yamamoto et al., 2013). While Plt-LT-HSC and MKRPs contribute to MK production during steady-state conditions, another HSC that is thought to be derived from MKRPs called stem-like MK-committed progenitors (SL-MKPs) and which can differentiate directly to MK, are activated in response to stress induced by inflammation, infection or other states of thrombocytopenia (Haas et al., 2015).

The process of megakaryopoiesis is orchestrated by a wide variety of factors, including thrombopoietin (TPO) and its receptor (MPL), many other cytokines and several transcription factors (reviewed in Behrens & Alexander (2018); Deutsch & Tomer (2013); Tijssen & Ghevaert (2013)).

TPO acts as the primary regulator of megakaryopoiesis by binding to its receptor, MPL, and activating several downstream signalling pathways, though the presence of

residual platelets in mice and human with TPO or MPL deficiency (Alexander et al., 1996; de Sauvage et al., 1996; Ihara et al., 1999; Savoia et al., 2007) indicates the contribution of TPO-independent mechanisms to the regulation of megakaryopoiesis (reviewed in Zheng et al. (2008)). Additionally, several cytokines contribute to megakaryopoiesis in both steady-state and emergency megakaryopoiesis such as multi-colony-stimulating factor, interleukin (IL) -1α, IL-3, IL-6, IL-9, IL-11, IL-21, stem cell factor, and interferon-α (Behrens & Alexander, 2018).

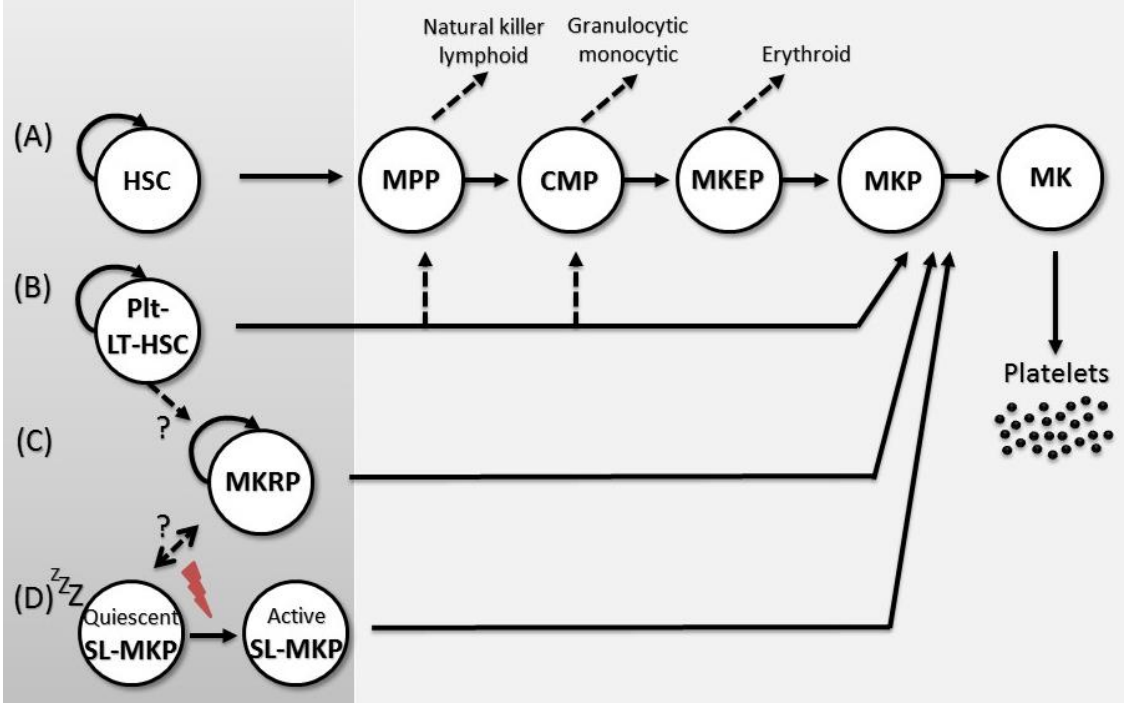


Figure 1.1 Megakaryopoiesis
In conventional megakaryopoiesis (A), haematopoietic stem cells (HSCs) are directed to differentiate sequentially through multipotent progenitors (MPPs), common myeloid progenitors (CMPs), megakaryocyte (MK)-erythroid progenitors (MKEPs), MK progenitors (MKPs), to mature MKs that release platelets. In non-conventional megakaryopoiesis (B-D), platelet-biased long-term reconstitution capacity HSC (Plt-LT-HSC), MK repopulating progenitors (MKRPs), and quiescent stem-like MK-committed progenitors (SL-MKPs) are able to differentiate directly to MK progenitors. Curved arrows indicate self-renewal capacity. Adapted from Behrens & Alexander (2018).

Each stage in the process of megakaryopoiesis requires a specific set of transcription factors. RUNX1 (Runt-related transcription factor 1) is considered to be a core regulator of early and late MK differentiation that interacts with other essential transcription factors to form stage-specific transcription factor complexes. For example, the RUNX1/ETS1 (E26 transformation-specific 1)/CBFβ (Core-binding factor subunit β) complex is required in the early stages of megakaryopoiesis, while the

RUNX1/Friend leukaemia virus integration 1 (FLI1)/GATA1 complex plays an important regulatory role in the late stages of megakaryopoiesis (Okada et al., 2013). Other transcription factors and co-factors that are known to contribute to MK lineage commitment, as well as to platelet production and function, include GATA2, FOG1 (Friend of GATA1), TAL1 (T-cell acute lymphocytic leukaemia 1, SCL), EVI1 (Ecotropic Virus Integration Site 1 Protein Homolog), NFE2 (Nuclear Factor Erythroid 2), SRF (Serum Response Factor), MKL1 (Megakaryoblastic Leukaemia 1) and GFI1B (Growth Factor-Independent 1B).

The process of MK maturation commences with multiple cycles of endomitosis without cytokinesis that are accompanied by an increase in the cytoplasmic volume in proportion to the degree of ploidy, resulting in a single MK with a ploidy of 16N or greater, which has a single poly-lobulated nucleus and a large cytoplasm. Subsequently, MKs develop a demarcation membrane system, becoming filled with platelet-specific proteins, organelles and granules inherited by platelets. MK maturation is followed by thrombopoiesis, the last stage of megakaryopoiesis which results in the release of platelets (reviewed in Kosaki (2005)). Platelet release is believed to be mediated via proplatelet formation, whereby, following MK maturation and migration toward the vascular niche, they develop dynamic cytoplasmic structures that project through junctions in the lining of blood sinuses, forming pseudopod-like structures called proplatelets. After further elongation and branching into multiple platelet-sized swellings at their ends, the platelets are released into the circulation due to the high shear forces that operate in the sinusoidal vessels. During this process, cellular organelles are packaged into the proplatelet before the platelets are released. The alternative platelet territory model for platelet release proposes that, during MK maturation, the developing demarcation membrane forms zones or territories around each platelet and the nascent platelets are then released into the circulation following explosive fragmentation that starts at the MK surface and proceeds to the centre of the cell (Kosaki, 2005).

1.1.2 Platelet ultrastructure

Platelets are small discoid cells, approximately 2.5 µm in diameter, that do not possess a nucleus, but they have a number of distinct structural features (reviewed in Gremmel

et al. (2016); Thon & Italiano (2012)). Although platelets do not contain DNA, they do have MK-derived mRNA allowing the expression of platelet proteins.

The peripheral zone of a platelet contains a glycocalyx and a plasma membrane with internal invaginations that form an open canalicular system. The glycocalyx facilitates the uptake of plasma proteins, and provides an adhesive surface in response to haemostatic demands. The plasma membrane is a phospholipid bilayer in which negatively charged phospholipids are expressed predominantly in the inner leaflet thus maintaining the non-procoagulant status of the platelet surface. The membrane contains Na^+ and Ca^{2+}-ATP pumps that control the intracellular ionic environment. Additionally, the membrane has dynamic, cholesterol- and sphingolipid-rich microdomains called lipid rafts that have an important role in signalling and intracellular trafficking in platelets. The plasma membrane is also densely packed with many receptors, including glycoprotein (GP) and integrin receptors. The open canalicular system acts mainly as a massive membrane reservoir for activated platelets, serving as a gate that allows entry of external elements and release of granule contents, as well as a storage site for some plasma membrane GP receptors.

The plasma membrane is connected to the cytoplasm via an actin cytoskeleton and contractile microtubule ring that play a role in platelet production, maintaining the discoid shape of resting platelets, linking the platelet receptors to the actin cytoskeleton and to cytoplasmic signalling proteins, thereby mediating shape change, spreading and degranulation following platelet activation in response to vascular injury. Additionally, many other proteins have a role in organising the actin cytoskeleton including myosin IIA, actin crosslinking proteins filamin A and α-actinin, as well as several regulator and effector proteins such as cofilin, profilin, Rho GTPases (RhoA, Cdc42 and Rac1), formins and WASp (Wiskott-Aldrich syndrome protein) (reviewed in Poulter & Thomas (2015); Shin et al. (2017)).

The platelet cytoplasm contains many cellular organelles, including mitochondria, endoplasmic reticulum (called dense tubular system), Golgi and several types of granules. Despite their low number, mitochondria provide the platelet's energy requirements. The dense tubular system is a residual endoplasmic reticulum network which acts as a reservoir for calcium that is released upon platelet activation. The Golgi

stack comprises remnants from the MK that are rarely seen in platelets. Recently, Golgi proteins were found within platelets and shown to map to a series of scattered separated structures rather than the typical stacked Golgi structures (Yadav et al., 2017).

Platelets possess three main types of granules, alpha (α)-granules, dense (δ)-granules and lysosomes, each with different cargoes (reviewed in Sharda & Flaumenhaft (2018)). The most abundant of these are the α-granules, which number between 50 and 80 per platelet and contain many proteins that are either synthesised in the MKs (e.g. platelet factor 4 (PF4), β-thromboglobulin, platelet-derived growth factor (PDGF), thrombospondin) or endocytosed from plasma (e.g. factor V, fibrinogen, von Willebrand factor (VWF)). These proteins have diverse roles in platelet adhesion, coagulation, and wound healing. Each platelet also has between 5 and 8 electron-dense granules which store ADP, ATP, GTP, serotonin, histamine, pyrophosphate and divalent cations that are essential for recruiting other platelets to the site of release. Lysosomes are the least abundant granule type, storing proteolytic enzymes which are involved in vascular remodelling.

1.1.3 The role of platelets in primary haemostasis

Platelets have roles in many biological processes, both physiological, such as wound healing, angiogenesis, immunity and inflammation, and pathological, including atherosclerosis, ischaemia, disseminated intravascular coagulation, bleeding, cancer and autoimmune diseases (reviewed in Golebiewska & Poole (2015); Ware et al. (2013)). However, their primary role is to arrest bleeding following vascular injury.

Normally, platelets circulate in the bloodstream in a resting, non-reactive, discoid shape which is maintained through many mechanisms, e.g. the inhibitory effects of endothelial prostacyclin, nitric oxide (NO), and platelet–endothelial cell adhesion molecule-1 (PECAM-1) (Li et al., 2010; Ming et al., 2011). Most platelets do not participate in irreversible adhesion and are removed by macrophages in the spleen and liver after 7 to 10 days in the circulation. However, at sites of vascular injury, platelets are required to change their status to a reactive form to arrest bleeding. They participate in complex processes that involve interactions with the extracellular matrix, other platelets, as well as other cells via receptors present on their surface membranes

to trigger intracellular signalling pathways that lead to the formation of a platelet thrombus at the site of injury. There are generally considered to be five phases of platelet activity at sites of vessel injury: initiation, extension, secretion, consolidation and procoagulant activity.

1.1.3.1 Initiation

Circulating platelets remain in a resting state until they come into contact with immobilised adhesive proteins at sites of exposed subendothelial matrix, whereupon they become activated, and are able to bind to collagen through direct interactions mediated mostly via the platelet receptors α2β1 and GPVI that are important for platelet adhesion and collagen-induced platelet activation respectively (Berndt et al., 2014; Li et al., 2010). GPVI is non-covalently coupled to the Fc receptor γ-chain (FcRγ) and to Src family kinases (SFKs). As shown in Figure 1.2 when GPVI is crosslinked by collagen, the cytoplasmic immunoreceptor tyrosine-based activation motif (ITAM) within the FcRγ chain is tyrosine phosphorylated by SFKs leading to a serious of signalling events that result in activation of phospholipase C gamma (PLCγ), elevation of cytosolic calcium, thromboxane A2 (TXA2) synthesis, granule secretion, and integrin activation (Berndt et al., 2014; Li et al., 2010).

In regions of low shear such as the venous circulation, some platelet integrins, including α2β1, are able to mediate stable binding of platelets to the subendothelial matrix directly (Dumont et al., 2009; Li et al., 2010). In contrast, in areas of high shear, platelets become tethered at sites of injury through the platelet GPIb-IX-V complex, which binds reversibly to immobilised VWF in the sub-endothelium via GPIb, thus allowing platelets to roll across the site of the injury (Varga-Szabo et al., 2008; Yago et al., 2008). However, this interaction is stabilised through direct interactions with collagen, which are mediated via the platelet collagen receptors α2β1 and GPVI. The interaction of VWF with GPIb-IX-V leads to activation of a SFK and phosphoinositide 3-kinase (PI3K) resulting in PLCγ activation. It also triggers signalling through Akt, resulting in activation of the downstream mitogen-activated protein kinases (MAPKs), leading to TXA2 synthesis, and granule secretion (Figure 1.2) (Berndt et al., 2014; Li et al., 2010).

Platelet adhesion triggers signalling events that ultimately lead to an increase in intracellular calcium levels and structural changes that are associated with platelet spreading and the release of intracellular granule contents and integrin activation (Berndt et al., 2014; Li et al., 2010; Varga-Szabo et al., 2009).

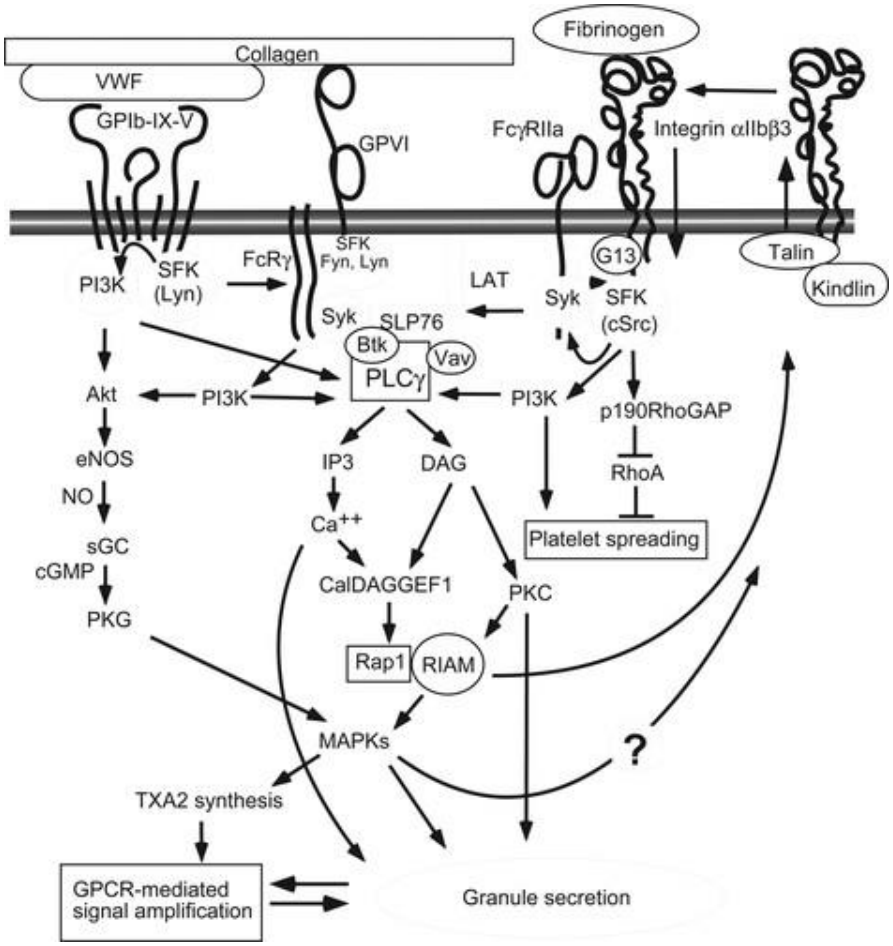


Figure 1.2 Signalling pathways of the major platelet adhesion receptors
Btk; Bruton tyrosine kinase, CalDAG-GEF1; DAG-regulated GEF 1, DAG; diacylglycerol, eNOS; endothelial NO synthase, FcRγ; immunoglobulin epsilon receptor subunit γ, FcγRIIa; immunoglobulin γFc receptor IIa, GP; glycoprotein, IP3; inositol trisphosphate, LAT; transmembrane adapter linker for activated T-cells, MAPKs; mitogen-activated protein kinases, NO; nitric oxide, PI3K; phosphoinositide 3-kinases, PKG; cGMP-dependent protein kinase, PLCγ; phospholipase C gamma, RIAM; Rap1-GTP–interacting adaptor molecule, SFK; Src family kinases, sGC; soluble guanylyl cyclase, SLP76; Src homology 2 domain– containing leukocyte phosphoprotein of 76-kilodalton, TXA2; thromboxane A2, VWF; von Willebrand factor. Adapted, with permission, from Li et al. (2010).

1.1.3.2 Secretion

The release of the contents of the intracellular secretory granules of platelets into the surrounding environment is fundamental to thrombus formation as it mediates further platelet activation and interactions with other cells (reviewed in Joshi & Whiteheart

(2017); Sharda & Flaumenhaft (2018)). The process of exocytosis is typically mediated by the SNARE (soluble N-ethylmaleimide-sensitive fusion attachment protein receptor) complexes, which are formed through protein interactions between the vesicle (v)-SNAREs (also known as vesicle-associated membrane proteins (VAMPs)), which reside on the vesicle or granule membrane, and target (t)-SNAREs present on the cell membrane and are subgrouped into the synaptosomal-associated proteins (SNAPs) and the syntaxins (STXs) (Hong, 2005).

Although platelets contain different types of granules, current knowledge supports a single model for platelet secretion that involves the (v)-SNARE, mainly VAMP8, and the (t)-SNAREs, STX11 and SNAP23 (Golebiewska & Poole, 2015; Heijnen & Van der Sluijs, 2015). In addition to SNAREs other 'granule' machinery and 'plasma membrane' machinery are involved. The 'granule' machinery includes the SNARE-associated proteins UNC13D (Protein unc-13 homolog D; also known as MUNC13-4), the small GTPase Rab27 (Ras-related protein Rab-27) and the Rab effector SLP4 (Synaptotagmin-like protein 4), while the 'plasma membrane' machinery includes STXBP2 (Syntaxin-binding protein 2; also known as Munc18b) and the regulator STXBP5 (Syntaxin-binding protein 5; also known as Tomosyn 1).

Following platelet activation, sequential steps lead to secretion including granule tethering, docking via SNARE engagement, and membrane fusion, then cargo release (see Figure 1.3). Initially, granules tether to the lipid rafts in specific regions of the plasma membrane through the Rab27-UNC13D interaction and SLP4. Tethering facilitates pairing of the SNARE proteins between the vesicles (VAMP8) and the membrane (STX11-SNAP23-STXBP2) that lead to membrane fusion. Finally, interaction of the opposing SNARE is mediated by the Rab27-UNC13D complex on the granule leading to cargo release (Joshi & Whiteheart, 2017; Sharda & Flaumenhaft, 2018). The presence of multiple SNAREs and SNARE-effectors in platelets that are able to contribute to granule exocytosis suggest a restricted spatial and temporal regulation of secretion for granules of different type or content (Golebiewska & Poole, 2015; Heijnen & Van der Sluijs, 2015; Joshi & Whiteheart, 2017; Sharda & Flaumenhaft, 2018).

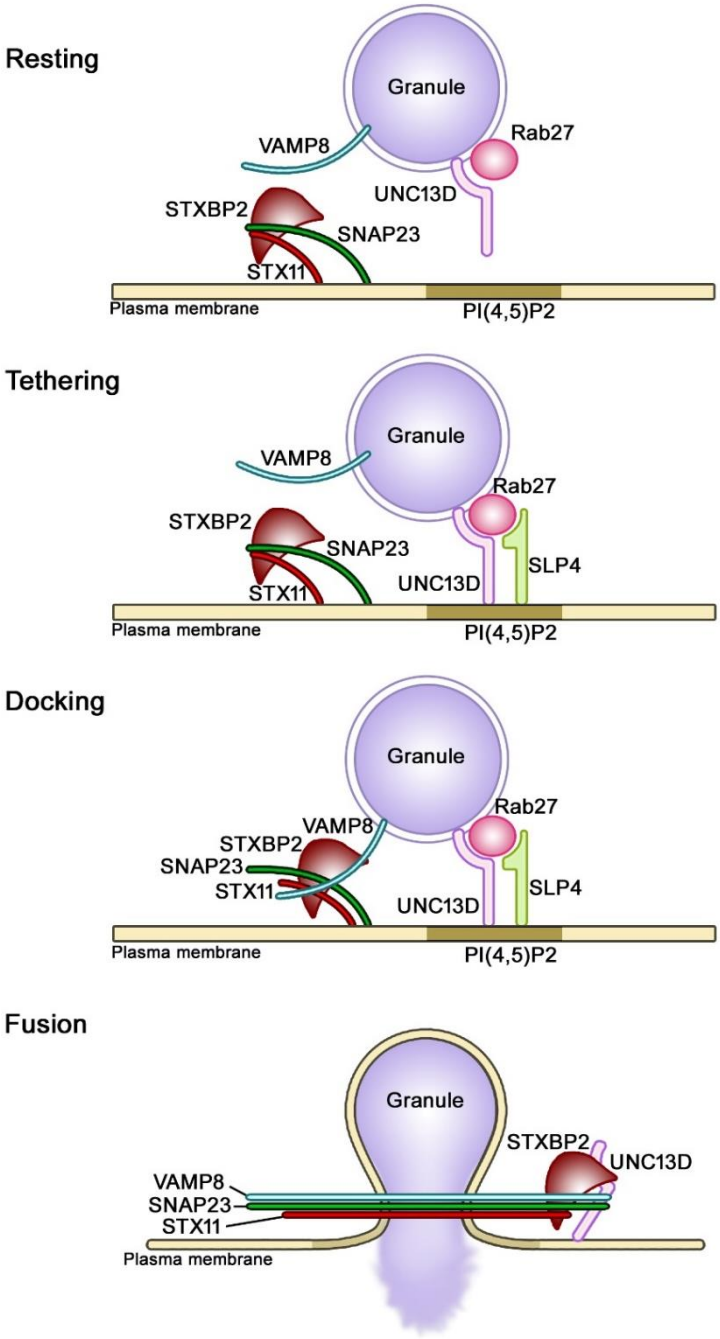


Figure 1.3 The currently accepted model of platelet granule secretion
Three SNARE proteins form the core SNARE complex in platelet secretion; vesicle transmembrane VAMP8, membrane-anchored STX11 and SNAP23. Additionally, the SNARE-associated proteins, UNC13D and STXBP2, the small GTPases, Rab27 and SLP4 are also essential for secretion. Adapted, with permission, from Heijnen & Van der Sluijs (2015).

1.1.3.3 Extension

The molecules released from platelets, which include ADP, epinephrine, TXA2 and serotonin, promote the recruitment and activation of other platelets at the site of injury. Platelets respond to these soluble agonists via specific seven-transmembrane-domain G-protein-coupled receptors (GPCRs) which trigger downstream signalling through

one of four G-proteins each consisting of 3 subunits, α, β and γ. When a ligand binds to its cognate GPCR, the α-subunit of the heterotrimeric G-protein detaches from the receptor and the β/γ complex, triggering downstream signalling events. G-proteins are classified based on the similarity of their α-subunits into four subfamilies: Gq, Gi/Gz, G12/G13 and Gs. Figure 1.4 shows the platelet GPCRs and their ligands (reviewed in Li et al. (2010)).

With the exception of Gs, signalling through the G-proteins in platelets results in platelet activation, while signalling through Gs, which is coupled to the prostacyclin receptor, results in inhibition of platelet activation following stimulation of adenylyl cyclase-dependent cAMP synthesis. G13 signalling activates RhoA which induces myosin light chain-dependent contraction, resulting in platelet shape change and granule secretion (Klages et al., 1999; Kozasa et al., 1998). Signalling through Gi relieves the inhibitory effect of cAMP-dependent protein kinase, activating the PI3K/Akt signalling pathway that induces granule secretion, TXA2 synthesis and activation of the critical mediator of integrin activation; Rap1 (Li et al., 2003b; Lova et al., 2003). Gq transmits cellular signals via the PI3K/Akt signalling pathway, and via activation of PLCβ that catalyses the hydrolysis of phosphatidylinositol 4,5 bisphosphate to inositol trisphosphate (IP3) and diacylglycerol (DAG). IP3 and DAG activate calcium mobilisation and protein kinase C (PKC) respectively, leading to platelet granule secretion. Additionally, they both indirectly activate αIIbβ3 integrin and MAPKs through activation of the DAG-regulated guanine nucleotide exchange factor (GEF) 1 (CalDAG-GEF1) that activates Rap1 (Crittenden et al., 2004; Varga-Szabo et al., 2009).

Despite the considerable redundancy in the signalling pathways in platelets, it is noteworthy that while upstream signalling events vary according to the stimulus and the receptor activated, the downstream signalling pathways that ultimately lead to granule secretion, TXA2 synthesis, shape change and activation of αIIbβ3 are similar.

1.1.3.4 Consolidation / Aggregation

The αIIbβ3 integrin is the most abundant receptor on the platelet surface and it plays a crucial role in mediating platelet aggregation (Li et al., 2010; Wagner et al., 1996). It is maintained in a low-affinity state on resting platelets but following platelet activation, αIIbβ3 undergoes a conformational change to a high-affinity form that allows binding

to its ligands, fibrinogen and VWF, in an “inside-out” signalling process that requires the binding of talin and kindlin to the cytoplasmic domain of β3 (Bouaouina et al., 2008; Ma et al., 2008). Binding of αIIbβ3 to its ligands results in positive feedback and “outside-in” signalling, where Gα13 binds to the cytoplasmic domain of β3 and activates SFKs (Gong et al., 2010) that subsequently inactivate RhoA allowing platelet spreading on fibrinogen (Arthur et al., 2000; Gong et al., 2010) and mediating activation of PLCγ2 (Boylan et al., 2008) leading to further platelet activation, secretion, spreading and platelet-platelet interactions (aggregation), forming the primary platelet plug (Shattil et al., 2010).

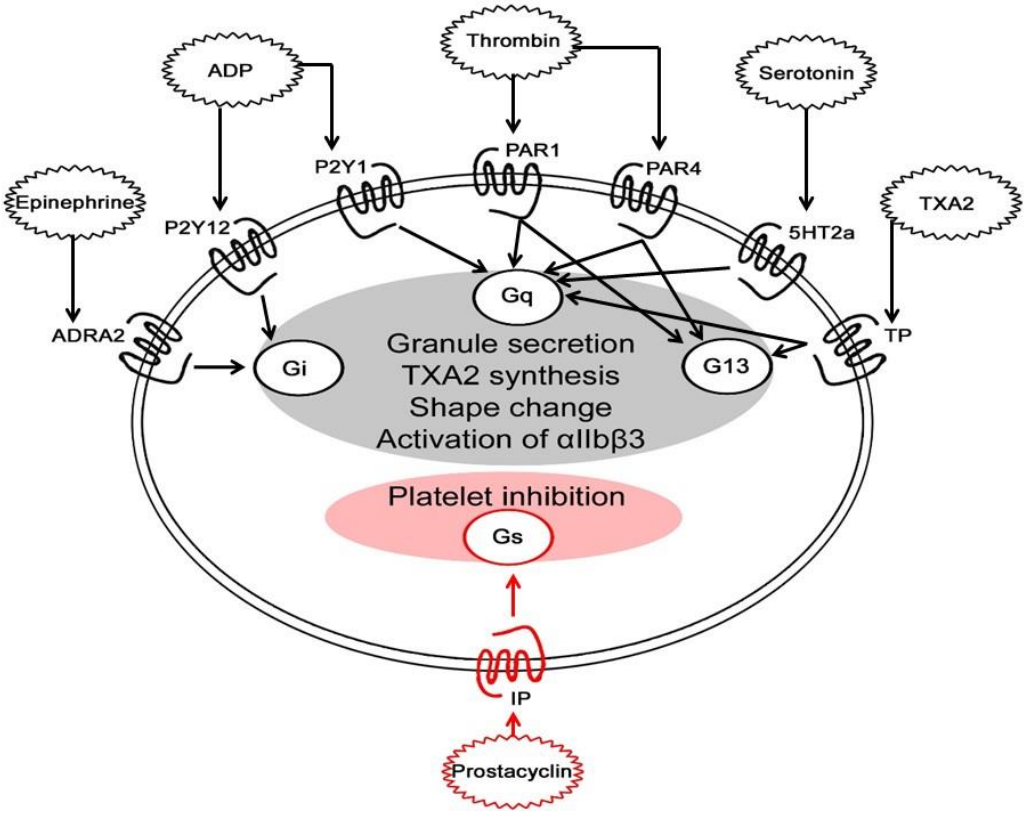


Figure 1.4 Platelet G-protein-coupled receptors and their G-proteins
Platelet agonists bind to their cognate receptors stimulating platelet activation via the G-proteins, Gi, Gq or G13. Thrombin, serotonin and thromboxane A2 (TXA2) mediate platelet activation through the Gq and G13-coupled receptors, protease-activated receptor-1 (PAR-1), PAR-4, 5-hydroxytryptamine (5HT) and TXA2/prostaglandin H2 receptors (TP) respectively. ADP mediates initial platelet activation via the Gq-coupled receptor P2Y1, while sustained platelet aggregation is mediated through the Gi-coupled P2Y12 receptor. The Gi-coupled α_2-adrenergic receptor (ADRA2) activates platelets in response to epinephrine. Stimulation of the Gs-coupled prostacyclin receptor (IP) leads to platelet inhibition. Adapted, with permission, from Li et al. (2010).

1.1.3.5 Platelet Procoagulant Activity

Platelets also have procoagulant activity, as the exposed anionic phospholipids on the activated platelet membrane provide a surface for activation of coagulation factors X and prothrombin (Suzuki et al., 2010). The assembly of the tenase and prothrombinase complexes results in the generation of thrombin, which catalyses the formation of fibrin monomers, which when crosslinked, form a mesh that coats the primary platelet plug to form a stable haemostatic clot at the site of injury.

1.1.4 Inherited platelet bleeding disorders

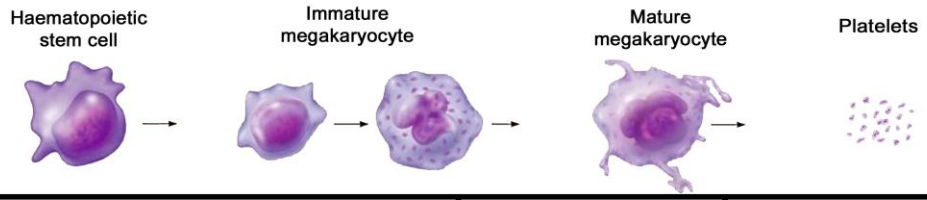
Abnormalities in platelet number or function can be manifested as a predisposition to either thrombosis or bleeding. Symptoms of excessive bleeding can arise from either acquired or inherited causes, both presenting with similar bleeding patterns. However, patients with inherited defects usually present with symptoms early in life and have a family history of bleeding, whereas acquired disorders tend to be sporadic, occurring later in life. Inherited platelet bleeding disorders (IPDs) are a heterogeneous group of conditions arising from defects in genes which have a role in platelet production or function that result in an increased risk of bleeding.

Patients with IPDs present with a broad range of mild to severe bleeding symptoms. Typically, they exhibit spontaneous mucocutaneous bleeding, epistaxis, menorrhagia and gastrointestinal bleeding. Patients may also develop purpura, bruise easily and experience prolonged bleeding after trauma. Some IPDs are associated with additional characteristic clinical signs and symptoms which facilitate their diagnosis. For example, patients with Wiskott-Aldrich syndrome usually suffer from severe immunodeficiency, while patients with Hermansky-Pudlak syndrome commonly present with skin and hair hypopigmentation (Derry et al., 1994; Hermansky & Pudlak, 1959). In contrast, some patients can be asymptomatic and are only at risk of bleeding when exposed to haemostatic challenges such as surgery or childbirth.

1.1.4.1 The molecular basis of inherited platelet disorders

IPDs tend to be classified according to the predominant features observed in affected patients, thus disorders in platelet secretion or aggregation are well established, as are inherited forms of thrombocytopenia. However, considerable heterogeneity in the phenotypic and clinical expression of IPDs mean that the features of different subgroups of IPDs frequently overlap. For example, Quebec platelet syndrome is

classified as a platelet secretion disorder, although it is also characterised by moderate thrombocytopenia (Tracy et al., 1984). As understanding of platelet biology has grown, some of the overlapping features have been explained by the recognition that many of the genes required for platelet development encode proteins that are also involved in mature platelet function (Bariana et al., 2017; Daly, 2017; Pecci & Balduini, 2014; Savoia, 2016). This knowledge has resulted in the grouping of IPDs based on the particular qualitative and/or quantitative defects displayed by the platelets as well as any additional clinical features of affected individuals and specific gene defects identified. To date, defects in at least 69 genes have been identified in patients with IPDs (Figure 1.5, Appendix 1).



Megakaryocyte differentiation and maturation		Proplatelet production and platelet release	Platelet function
ABCG5	*ACBD5*	*ACTN1*	*ANO6*
ABCG8	*ANKRD18A*	*ARPC1B*	*CD36*
HOXA11	*ANKRD26*	*CDC42*	*EPHB2*
MECOM	*CYCS*	*DIAPH1*	*FERMT3*
MPL	*ETV6*	*FLNA*	*GBA*
RBM8A	*FLI1*	*GEN*	*GP6*
THPO	*FYB*	*GP1BA*	*MLPH*
	G6b-B	*GP1BB*	*MUNC18b*
	GATA1	*GP9*	*MYO5A*
	GFI1B	*ITGA2B*	*P2RX1*
	HPS1-9	*ITGB3*	*P2RY12*
	LYST	*KDSR*	*PLA2G4A*
	MASTL	*MYH9*	*RAB27A*
	MUNC13-4	*PRKACG*	*RASGRP2*
	NBEAL2	*PTPRJ*	*SBF2*
	PLAU	*TUBB1*	*SLC35A1*
	RUNX1	*VPS45*	*STX11*
	SALL4	*WAS*	*TBXA2R*
	SLFN14	*WDR1*	*TBXAS1*
	SRC		
	STIM1		
	TREML1		
	VPS33B		
	VIPAS39		

Figure 1.5 Genes associated with inherited platelet bleeding disorders and their role in megakaryocyte maturation, platelet production and function
Many genes that play roles at different stages of megakaryocyte differentiation, proplatelet production and platelet function have been identified in association with inherited platelet bleeding disorders. Defects in these genes can result in thrombocytopenia and/or platelet function abnormalities. Adapted, with permission, from Pecci & Balduini (2014); Savoia (2016) and further updated.

1.1.4.2 Diagnosis of inherited platelet function disorders

Diagnosis of an IPD is usually based on the clinical evaluation of the patient and the phenotype of the patient's platelets. In a minority of cases, diagnosis is supported by DNA analysis to identify the underlying gene defect. Prior to undertaking platelet phenotyping, initial laboratory tests are usually conducted to exclude coagulation disorders, in particular type 1 VWD, the clinical features of which are similar to those of IPDs (Gresele et al., 2015).

Platelet phenotyping

Platelet phenotyping includes measurement of platelet count and size, the use of light microscopy to examine platelet morphology and platelet function testing using light transmission aggregometry (LTA). Additional assays and analyses that are occasionally used to phenotype platelets include assessment of platelet receptor levels by flow cytometry, electron microscopy to examine platelet ultrastructure, aggregometry using antagonists targeting specific platelet receptors, assessment of platelet thrombus formation on collagen at arterial shear rates, serotonin uptake and clot retraction assays (Gresele et al., 2015).

As the gold standard test used to assess platelet function, LTA records the change in light transmission of a sample of platelet-rich plasma that occurs over time in response to specific agonists (Born, 1962). Although the exact response depends on the agonist and the concentration at which it is used, classically the initial platelet response commences with a lag phase, followed by a decrease in light transmission due to changes in platelet shape, then by an increase in transmitted light, known as the primary reversible wave, due to platelet aggregation in the presence of an external agent. A secondary wave is then observed as a result of sustained aggregation following the release of platelet granule contents, which result in a further increase in light transmission. Panels of agonists are used at a range of concentrations. The agonists include ADP, adrenaline, collagen, ristocetin, arachidonic acid, U46619 (a TXA2 analogue), protease-activated receptor-1 (PAR-1) and PAR-4 activation peptides and collagen-related peptide (CRP). LTA has also been adapted to allow assessment of dense granule secretion by the addition of luciferin-luciferase reagent and monitoring ATP release in a lumi-aggregometer (Miller, 1984).

The differential aggregation responses of platelets to a panel of agonists facilitate the identification of the two best characterised IPDs, Glanzmann thrombasthenia and Bernard-Soulier syndrome, since they are associated with characteristic aggregation responses. Furthermore, where the diagnosis of a specific disorder is not possible, the responses allow patients to be subgrouped according to the specific pattern of agonist responses. Thus, patients can be grouped according to whether their platelets display defects in Gi signalling, the TXA2 pathway, dense granule secretion, Gq signalling or GPVI signalling (Dawood et al., 2012). Indeed, most platelet defects identified by LTA were shown to fall into the first three of these subgroups (Dawood et al., 2012). A defect in the Gi signalling pathway appears in LTA as a defect in response to ADP and adrenaline. TXA2 pathway defects are characterised by a marked abnormality in response to arachidonic acid and the TXA2 analogue, U46619, can be used to distinguish between defects in TXA2 synthesis (normal platelet aggregation) and defects in the TXA2 receptor or in its downstream signalling pathway (impaired platelet aggregation). Patients with dense granule secretion defects present with a decrease in ATP secretion in response to all agonists (Dawood et al., 2012).

Genetic analysis in inherited platelet disorders

In addition to being of diagnostic value, genetic analysis in patients with IPDs can complement the results of platelet phenotyping, identify reasons for unexplained bleeding, allow assessment of bleeding risk, and facilitate genetic counselling of affected individuals. Improvements in our understanding of megakaryopoiesis, thrombopoiesis and platelet function in recent years have been driven by the development of molecular and genetic technologies such as next-generation sequencing (NGS), which have in turn led to the association of 34 new genes with IPDs over the last eight years, a number that represents approximately half of all genes known to contribute to IPDs in humans (Figure 1.6).

Characteristic clinical features and platelet phenotypes are helpful in directing genetic analyses to a single or a few genes to confirm the diagnosis, and Sanger sequencing is useful for confirming the diagnosis of IPDs in those cases when phenotypic analysis suggests defects in particular genes such as in Glanzmann Thrombasthenia and Bernard-Soulier syndrome, or in membrane receptors for specific agonists such as the P2Y12 ADP, TXA2 and GPVI receptors. However, in those cases where platelet phenotyping does not direct genetic analysis, Sanger sequencing is impractical. The

development of NGS technologies has enabled large-scale simultaneous genetic analysis of target gene panels, whole exomes, or whole genomes and has facilitated the discovery of many novel genes associated with IPDs, of which the first was *NBEAL2*, defects in which are associated with Grey platelet syndrome (Albers et al., 2011; Gunay-Aygun et al., 2011; Kahr et al., 2011). Since then, several other genes have been implicated in different IPDs using NGS technologies and many novel mutations in known genes associated with IPDs have also been documented. The role of NGS in the investigation of patients with IPDs is discussed more fully in chapter 3 of this thesis.

Approximately half of all patients with IPDs present with clearly associated clinical features that allow their diagnosis by sequencing a panel of genes, one of which is likely to harbour the underlying causative defect (Gresele et al., 2015; Sivapalaratnam et al., 2017). The other 50% of patients remain partially diagnosed, and the underlying genetic defect(s) are not identified. Although considerable advancements in the molecular genetic techniques that facilitate DNA analysis have been made in recent years, genetic diagnosis in these patients remains challenging due to the variable clinical expression, disease heterogeneity and multifactorial basis of the disease, and the possible contribution of other factors, such as infection, drug use, and other acquired disorders, to the diagnosis as well as the lack of knowledge of some aspects of platelet biology (Daly et al., 2014; Watson et al., 2013). These patients are categorised with mucocutaneous bleeding of unknown cause

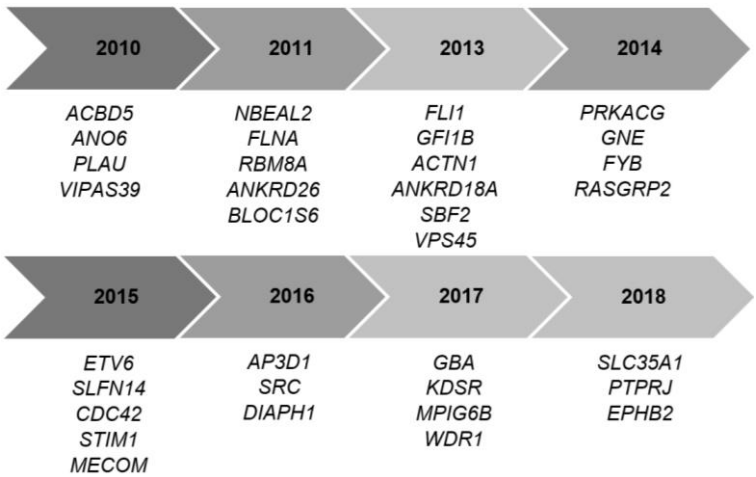


Figure 1.6 Genes recognised to be associated with inherited platelet bleeding disorders since 2010

1.2 The aims of this study

The underlying genetic defect remains unknown in approximately 50% of patients with IPDs. The overarching hypothesis of this study is that IPDs in these patients are due to monogenic or oligogenic inheritance of alterations in genes encoding proteins essential for platelet formation and/or function. Thus, this study had three broad aims, which were to:

i. Identify potential disease-causing genetic variants in patients with IPDs via analysis of whole exome sequence (WES) data using a variety of bioinformatic tools
ii. Investigate the pathogenicity of selected variants using *in vitro* approaches
iii. Identify novel genes associated with IPDs

To achieve these aims, WES analysis was undertaken for 34 individuals who had been clinically diagnosed as having excessive bleeding symptoms and a suspected IPD. All patients were enrolled in the UK Genotyping and Phenotyping of Platelets (UK-GAPP) study and extensive platelet phenotyping had resulted in them being assigned to one of two subgroups according to whether their platelets showed a Gi-signalling defect, or a defect in platelet secretion. WES analysis was undertaken for all patients resulting in the identification of 98 candidate disease-causing variants in each index case (see chapter 3). Two potential disease-causing genetic variants in *FLI1* and *ETV6* that were identified in patients with platelet secretion defects were selected for further characterisation (see chapters 4 and 5). Finally, to identify novel genes involved in platelet granule biogenesis and secretion, gene expression was examined in a megakaryocytic cell line before and after knockout/down of *FLI1*, which encodes a transcription factor required for megakaryopoiesis, defects in which are associated with platelet granule abnormalities. Comparison of the gene expression data with that in platelets from patients with *FLI1* defects and with the results of WES analysis in patients with secretion defects highlighted several novel genes of interest (see chapter 6).

2 Chapter 2. Materials and Methods

2.1 Materials

2.1.1 Plasmids

The SnapGene Viewer software (version 4.1.5) was used to generate schematic diagrams for all plasmids used in this work, which can be viewed in Appendix 2. The expression vector pSG5 was from Agilent Technologies and the enhanced green fluorescent protein (EGFP) plasmids, pEGFP-N2-Empty Vector (EV) and pEGFP-N3-EV, were from Clontech laboratories. The plasmid pCMV6-AC-HA-His-EV (pCMV6) was purchased from Origene (OriGene Technologies), while the pGEM®-T Easy vector was supplied by Promega. The CRISPR plasmid pSpCas9(BB)-2A-EGFP was from Addgene (PX458, plasmid # 48138) and used to generate the CRISPR/Cas9 (clustered regularly interspaced short palindromic repeats / CRISPR associated protein 9) plasmid after ligation with the designed guide RNA (gRNA) sequence (Ran et al., 2013).

The pSG5-FLI1 plasmid containing the full-length wild-type (WT) *FLI1* cDNA was kindly provided by Professor M Trojanowska, Boston University (Watson et al., 1992). Plasmid pEGFP-N2-FLI1 containing the *FLI1* cDNA was generated by Dr V Leo, Sheffield Haemostasis Group prior to this study. The pCMV6-AC-HA-His-ETV6 (pCMV6-ETV6) plasmid containing the full-length *ETV6* cDNA was a gift from Professor J Di Paola, University of Colorado Anschutz Medical Campus (Noetzli et al., 2015).

Three derivatives of the pGL3 reporter plasmid containing promoter fragments from three different genes cloned upstream of the Firefly luciferase cDNA were used. The pGL3-GP6 plasmid containing a fragment corresponding to nucleotides -238 to -1 of the *GP6* (glycoprotein VI) promoter was generated by members of the Haemostasis Research Group prior to this study (Stockley et al., 2013). The pGL3-MMP3 and pGL3-PF4 plasmids were a gift from Dr A Shimamura, Department of Paediatrics, University of Washington (Zhang et al., 2015); pGL3-MMP3 contains a DNA fragment corresponding to nucleotides -388 to -1 of the matrix metallopeptidase 3 *(MMP3)* promoter, whereas pGL3-PF4 contains a fragment corresponding to nucleotides -300 to -1 of the platelet factor 4 (*PF4*) gene. The plasmid pRL-null was purchased from Promega (Promega, Southampton, UK).

2.1.2 Oligonucleotide primers

Oligonucleotide primers were supplied by Eurofins Genomic apart from the TaqMan™ probes which were from Applied Biosystems Life Technologies. Oligonucleotides were handled and stored according to the manufacturers' instructions. Working oligonucleotide primer concentrations for polymerase chain reaction (PCR) and sequencing were 10 µm and 1 µm, respectively. Table 2.1 provides details of all oligonucleotide primers used.

Table 2.1 Oligonucleotide primers used in the study

Name	Sequence (5' - 3')
Sequencing and PCR amplification	
M13_F	TGTAAAACGACGGCCAGT
M13_R	CAGGAAACAGCTATGACC
T7_F	TAATACGACTCACTATAGGG
CMV_F	CGCAAATGGGCGGTAGGCGTG
EGPF-N	CGTCGCCGTCCAGCTCGACCAG
U6_F	CGTAACTTGAAAGTATTTCGATTTCTTGGC
FLI1-exon7_F*	CAGGGCCAGCACATAGTAGA
FLI1-exon7_F#	*tgtaaaacgacggccagt*GAGAAAGCACATCTGTCAAG
FLI1-exon7_R#	*caggaaacagctatgacc*CCTCCATCAGTTGACCATGT
FLI1-exon8_F#*	*tgtaaaacgacggccagt*CTTATGGTTGGTACGGTTGT
FLI1-exon8_R#*	*caggaaacagctatgacc*CAGGTGTCTGGACTTAGGAC
FLI1-exon9_F#	*tgtaaaacgacggccagt*GAACTGGGTTCTGCCTTCTC
FLI1-exon9_R#	*caggaaacagctatgacc*ACATATGTCCTGTTGAGTCC
FLI1-exon9_R*	TTGGGGTTGGGGTAGATTCC
FLI1-PS2_F	ATGACCACCAACGAGAGGAG
FLI1-PS3_F	AGCTGTGGCAATTCCTCCT
FLI1-PS5_R	GGACTTTTGTTGAGGCCAGA
FLI1-PS6_R	GGGAGGGACAAAGTTCACCT
FLI1-PS7_R	GTTGGGGTTGGGGTAGATT
ETV6-exon8-F#	*tgtaaaacgacggccagt*CCAGCTGTATAAGATGATGG
ETV6-exon8-R#	*caggaaacagctatgacc*TAGTTTGTCTAAGGTGCTCC
ETV6-cDNA_1	ATAATCACTGCCCAGCGTCC
Mutagenesis (bold font indicates the targeted nucleotide(s))	
FLI1-R340C_S	TTTATCATAGTAATAAC**A**GAGGGCCCGGCTCAGC
FLI1-R340C_AS	GCTGAGCCGGGCCCTC**T**GTTATTACTATGATAAA
FLI1-R340H_S	GCTGAGCCGGGCCCTCC**A**TTATTACTATGATAAAAA
FLI1-R340H_AS	TTTTTATCATAGTAATAA**T**GGAGGGCCCGGCTCAGC
ETV6-R430X_S	ATGAAATCATGAGTGGC**T**G**A**ACAGACCGTCTGGAG
ETV6-R430X_AS	CTCCAGACGGTCTGT**T**C**A**GCCACTCATGATTTCAT
ETV6-R399C_S	AAATGTCCAGAGCCCTG**T**GCCACTACTACAAACTA
ETV6-R399C_AS	TAGTTTGTAGTAGTGGC**A**CAGGGCTCTGGACATTT
ETV6-R399X_S	GAGAAAATGTCCAGAGCCCTG**T**G**A**CACTACTACAAACTAAACATTA
ETV6-R399X_AS	TAATGTTTAGTTTGTAGTAGTG**T**C**A**CAGGGCTCTGGACATTTTCTC
CRISPR Cloning (phosphorylated oligonucleotides)	
FLI1-exon 7-S	*cacc*GATCGTTTGTGCCCCTCCAA
FLI1-exon 7-AS	*aaac*TTGGAGGGGCACAAACGATC
FLI1-exon 9-S	*caccg*AATGACGGACCCCGATGAGG
FLI1-exon 9-AS	*aaac*CCTCATCGGGGTCCGTCATT*c*
Quantitative polymerase chain reaction (qPCR)	

Taqman FLI1- FAM™	Hs00956711_m1
Taqman B2M- VIC™	Hs99999907_m1

Oligonucleotides were designed for FLI1; ENST00000527786: NM_002017 and ETV6; ENST00000396373: NM_001987. Italic lowercase letters indicate nucleotides added to facilitate cloning or sequencing; # primers have M13 tails to facilitate sequencing; * primers used to screen for CRISPR edits.

2.1.3 Preparation of plasmids and mutagenesis kits

The QuikChange® Lightning Site-Directed mutagenesis kit was purchased from Agilent Technologies. XL10-Gold® Ultracompetent cells were supplied with the mutagenesis kit, while calcium competent *Escherichia coli* (*E. coli*) NM554 cells were prepared in-house and made available for use in the laboratory. Luria-Bertani (LB) broth Miller (BP1426), LB agar Miller (BP1425) and NZY+ broth (BP2465) were purchased from Fisher BioReagents™. QIAprep miniprep, QIAprep maxiprep kits and Qiagen-tip 500 columns were purchased from Qiagen and used for plasmid purification. Ampicillin (A6140) and kanamycin (K-4000) were supplied by Sigma.

2.1.4 Cell lines, tissue culture media and transfection reagent

Human embryonic kidney (HEK) 293T cells were purchased from the American Type Culture Collection (ATCC), while human megakaryocytic Dami cells were kindly provided by Dr N Morgan, University of Birmingham. Dami cells were cultured in Roswell Park Memorial Institute (RPMI) 1640 medium + GlutaMAX™, while HEK 293T were cultured in Dulbecco's modified Eagle's medium (DMEM) + GlutaMAX™. For simplification, these media will be referred to as RPMI and DMEM hereafter. Both were supplemented with 10% foetal bovine serum (FBS) and antibiotic-antimycotic reagent. The media (Cat Nos. 10566-016, 61870-010), 100X antibiotic-antimycotic (Cat No. 15240062) (10,000 units/mL of penicillin, 10,000 µg/mL of streptomycin, and 25 µg/mL of amphotericin B) and FBS (Cat No. 10500064) were purchased from Gibco®. Dimethyl sulfoxide (DMSO; Cat No. D2438) and 1X trypsin-EDTA (0.12% trypsin, 0.02% EDTA in Dulbecco′s phosphate buffered saline] solutions; Cat No. 59430C) were purchased from Sigma.

Cells were transfected using either Lipofectamine® LTX reagent which was purchased from Life Technologies or jetPEI® transfection reagent from Polyplus.

2.1.5 Antibodies

Antibodies were used for western blotting and protein localisation studies. Table 2.2 lists the antibodies, their suppliers and the dilutions at which they were used.

Table 2.2 Antibodies used in the study, their suppliers and working dilutions

Antibody	Dilution		Modification	Company	Catalogue No.
	Microscopy	WB			
Rabbit polyclonal anti-human FLI1	1:250	1:500	Unconjugated	Thermo Fisher Scientific	PA5-29597
Rabbit polyclonal anti-human ETV6	---	1:500	Unconjugated	Abcam	ab185816
Mouse monoclonal anti-human TATA-binding protein (TBP)	---	1:1000	Unconjugated	Abcam	ab51841
Mouse monoclonal anti-human Cyclophilin B	---	1:1000	Unconjugated	Abcam	ab178397
Mouse monoclonal anti-human β-tubulin I antibody	---	1:500	Unconjugated	Sigma	T7816
Goat polyclonal anti-rabbit IgG (H+L) antibodies	1:500	---	Alexa Fluor™ 488 dye	Invitrogen	A-11008
Donkey polyclonal anti-rabbit IgG (H + L) infrared fluorescent dye (IRDye)® 680RD	---	1:10000	IRDye ® 680RD	Li-Cor	926-68073
Donkey polyclonal anti-mouse IgG (H + L) IRDye ® 800CW	---	1:10000	IRDye ® 800CW	Li-Cor	926-32212

2.1.6 Protein analysis reagents

Protein was extracted using the radioimmunoprecipitation assay buffer (RIPA) from Thermo Scientific (89901) or the NE-PER™ nuclear and cytoplasmic extraction reagents which were purchased from Pierce Biotechnology. SIGMAFAST™ Protease Inhibitor Cocktail (PIC) tablets, EDTA-Free (S8830) were from Sigma. The Pierce bicinchoninic acid (BCA) protein assay kit was from Thermo Fisher Scientific. Western blotting was carried out using Novex™ NuPAGE® Gels and XCell SureLock™ Mini-Cell system (Invitrogen, Thermo Fisher Scientific). NuPAGE 4-12% Bis-Tris gels 1.5 mm, NuPAGE lithium dodecyl sulphate (LDS) sample buffer, 10X NuPAGE reducing agent, NuPAGE® antioxidant, NuPAGE™ MOPS SDS running buffer, Invitrolon™ 0.45 µm pore size PVDF membrane and NuPAGE® transfer buffer were supplied by Invitrogen, Thermo Fisher Scientific. Odyssey Blocking buffer and Chameleon™ pre-stained protein ladder were obtained from Li-Cor. Methanol (M/4000/17) was

purchased from Fisher Chemical, while Tween®20 (663684B) was from VWR Life Science.

2.1.7 Enzymes and cloning reagents

Several of the reagents required for cloning were purchased from New England Biolabs including 20,000 unit/mL BbsI-high fidelity (R3539L) restriction enzyme, 10X restriction enzyme (CutSmart®) buffer, 400,000 U/mL T4 DNA ligase (M0202S), 10X T4 ligation buffer and 10 mg/mL bovine serum albumin (BSA). TA cloning was carried out using the pGEM®-T Easy Vector System from Promega (A1360) which, in addition to the vector, includes T4 DNA ligase (3 u/µL) and 2X rapid ligation buffer.

Several Taq polymerase kits were used including the 1.1X Reddymix DNA Polymerase Mix (AB-0575/LD), which contains Taq DNA Polymerase, dNTPs, reaction buffer, 1.5 mM magnesium chloride and a red gel loading dye (Thermo Scientific), GoTaq® G2 DNA polymerase kit (M7845), which contains 5X colourless GoTaq® reaction buffer and GoTaq® G2 DNA polymerase (Promega). The expand long template enzyme mix and 10X expand long template buffer 1, with 17.5 mM $MgCl_2$ supplied with the Expand Long Template PCR system from Roche Diagnostics (11681834001) were also used. A solution of 100 mM deoxy-adenosine triphosphate (dATP) (U1240) was obtained from Promega, and 10 mM dNTPs (KN1011) were purchased from Kapa Biosystems.

2.1.8 Other commercial kits and reagents

Additional reagents and kits not mentioned above, but which were used during this study are listed in Table 2.3.

Table 2.3 Commercial kits and reagents used in this study

Kit / Chemical / Reagent	Catalogue number / Product code	Supplier
4',6-diamidino-2-phenylindole (DAPI)	D1306	Molecular Probes® by Life Technologies
Agarose	BP1356	Fisher BioReagents™
Ammonium chloride	A9434	Sigma
BD Cytofix/Cytoperm™ Fixation buffer	554722	BD Bioscience
BSA lyophilised powder	A7888	Sigma
Clariom™ D Assay, human	902922	Affymetrix; Applied Biosystems™
Ethanol	E/0665DF/17	Fisher Chemical
Ethidium bromide (10 mg/mL in H_2O)	E1510	Sigma
EZ-RNA Total RNA Isolation Kit	20-400-100	Biological industries
GenElute™ mammalian genomic DNA miniprep kits	G1N70	Sigma
Glycerol	G/0650	Fisher
HyperLadder 1kb	BIO-33053	Bioline
HyperLadder IV (100s bp)	BIO-33056	Bioline
Isopropanol (propan-2-ol)	P/7500/17	Fisher Chemical
Phorbol 12-myristate 13-acetate (PMA)	P1585	Sigma
Phosphate buffered saline (PBS)	P4417	Sigma
ProLong® Gold anti-fade reagent	P36930	Molecular Probes by Life Technologies
Promega® Dual Luciferase® Reporter Assay System kit	E1910	Promega
QIAquick gel extraction Kit	28704	Qiagen
QIAquick PCR Purification Kit	28106	Qiagen
QuantiTect® reverse transcription kit	205313	Qiagen
QuickExtract™ DNA extraction solution	QE09050	Lucigen- Epicentric
RT-PCR grade water	AM9935	Ambion
Sodium acetate 3 M	71196	Sigma
TaqMan™ gene expression master mix	4369016	Applied biosystem
Thrombopoietin (TPO), recombinant human protein (10 µg)	PHC9514	Gibco® by Life Technologies
Tris-borate acetate (TBA)	EC-872	National Diagnostics
Triton® X-100	306324N	VWR Life Science

2.1.9 Software and online tools

Several software and online tools were used during this study. These are listed in Appendix 3.

2.2 Methods

2.2.1 Oligonucleotide primer design

Oligonucleotide primers for use in mutagenesis reactions were designed using the QuikChange primer design tool (http://www.genomics.agilent.com/primerDesignProgram.jsp) from Agilent Technologies [accessed between 2015 and 2016]. Forward and reverse primers were complementary to each other and the region of interest, except where mutagenic nucleotides were introduced (indicated in bold font in Table 2.1). Primers for use in PCR were designed using Prime3Web (http://primer3.ut.ee/) [accessed between 2015 and 2018]. Where appropriate, M13 tails were incorporated to facilitate sequencing (indicated by # in Table 2.1).

2.2.2 Nucleic acid quantification

DNA and RNA were quantified in sample solutions by measuring absorbance at 260 nm using the NanoDrop 1000 Microfluid spectrophotometer (Thermo Fisher Scientific UK). The purity of nucleotide samples was assessed using the 260/280 and 260/230 ratios of the absorbance measurements.

2.2.3 DNA sequencing

Sanger sequencing was performed using the Applied Biosystems 3730 DNA Analyser housed in the Core Genomic Facility, University of Sheffield. The sequence traces were analysed using FinchTV software (version 1.4.0) (Geospiza, Inc.), the Standard Nucleotide BLAST tool available at http://blast.ncbi.nlm.nih.gov/Blast.cgi?PROGRAM=blastn&PAGE_TYPE=BlastSearch&LINK_LOC=blasthome [accessed 2015-2018] and CodonCode Aligner (version 6.0.2) (CodonCode Corporation).

2.2.4 Amplification of genomic DNA using polymerase chain reaction

PCR was used to amplify target regions of genomic DNA prior to confirming the presence of sequence variants of interest, or the introduction of CRISPR-mediated edits. PCR was performed using 100 ng of genomic DNA, 10 µM forward primer, 10 µM reverse primer (Table 2.1) and 12.5 µl of 1.1X ReddyMix PCR master mix in a total volume of 25 µl. Reactions were subjected to the following thermal cycling in a GeneAmp® PCR system 9700 (Applied Biosystems®): an initial denaturation step at

95ºC for 2 minutes (mins), followed by 35 cycles of denaturation for 60 seconds (secs) at 95ºC, annealing for 60 secs at 55ºC and extension for 1 min at 72ºC, with a final extension performed at 72ºC for 10 mins. The PCR product was either directly sequenced or subjected first to agarose gel electrophoresis to check the size of the amplified product prior to sequence analysis or other use.

2.2.5 Agarose gel electrophoresis of DNA

Agarose gel electrophoresis was used to separate DNA fragments based on their size, and when analysed alongside an appropriate DNA ladder it allowed estimation of the size of DNA fragments. DNA fragments of 1 kilobase (Kb) or less were separated in 2% agarose, while 1% agarose was used for larger fragments. The gel was prepared by dissolving the appropriate amount of agarose in 50 mL of 1X Tris-borate acetate (TBA) by heating. The mixture was allowed to cool before adding 2 µL ethidium bromide (10 mg/ml) and then pouring it directly into the cassette. Following solidification of the gel, it was submerged in 1X TBA electrophoresis buffer and samples were loaded (7.5–10 µL of DNA ladder or 20–25 µL of PCR product). Samples were electrophoresed at 120 V for 30–40 mins until the dye in the DNA ladder had migrated 50 to 80% of the length of the gel. Images were acquired using a Bio-Rad Gel Doc 2000 ultraviolet (UV) transilluminator using Quantity One imaging software from Bio-Rad laboratories (version 4.6.8).

2.2.6 DNA extraction and purification

Following electrophoresis, a gel slice containing the DNA fragment of interest was excised using a sharp, clean scalpel under a FastGene® Blue LED transilluminator (Geneflow Limited) and transferred into a 1.5 mL Eppendorf tube. The DNA fragment was extracted from the gel slice and purified using the QIAquick gel extraction kit according to the manufacturers' instructions. Briefly, per 1X volume of agarose gel, 3X volumes of QG buffer were added. The sample was incubated at 50ºC for 10 mins with vortexing every 2–3 mins. When the agarose had dissolved, 1X volume of isopropanol was added to the sample, and the homogenous mixture was applied to the QIAquick column. After centrifugation for 1 min and discarding the flow-through, 500 µL of QG buffer was added to the column, and the flow-through was discarded following another centrifugation step. The column was then washed with 750 µL of PE buffer and incubated for 5 mins before centrifugation. After discarding the flow-through from the washing step, residual washing buffer was removed by an additional centrifugation

step. DNA was then eluted by adding 30 µL of EB buffer (10 mM Tris·Cl, pH 8.5) and incubating the sample for 5 mins before collecting DNA in a clean 1.5 mL Eppendorf tube by centrifugation. All centrifugation steps were performed at ~13,000 x *g* for 1 min in a table-top microcentrifuge at room temperature.

In some situations, PCR products were purified directly using the QIAquick PCR purification kit following the manufacturers' instructions. Briefly, 5 volumes of buffer PB were added to 1 volume of the PCR product. The mixture was loaded onto a QIAquick column before centrifugation to allow DNA binding to the membrane. The flow-through was discarded and the membrane washed with 750 µL PE buffer before centrifuging the column again and discarding the flow-through. After a further centrifugation step, 30 µL EB elution buffer was added. The column was allowed to stand for 5 mins and the DNA then collected by centrifugation. All centrifugation steps are carried out at ~13,000 x *g* for 60 secs at room temperature.

2.2.7 Site-directed mutagenesis

Nucleotide alterations were introduced into the cDNA sequence of interest in expression plasmids for either FLI1 or ETV6 using the QuikChange Lightning Site-Directed mutagenesis kit. Mutagenesis was achieved by PCR amplification of 100 ng of plasmid DNA template, diluted with 125 ng of each oligonucleotide primer (Table 2.1), 1 µl dNTP mix, 1.5 µl QuikSolution reagent, 5 µL 10X reaction buffer and 1 µl of QuikChange Lightning Enzyme in a final volume of 50 µl. Thermal cycling conditions were initiated with a denaturation step at 95ºC for 2 mins, followed by 18 cycles of denaturation at 95ºC for 20 secs, annealing at 60ºC for 10 secs and extension at 68ºC for 4 mins (30 secs/Kb of plasmid), with a final extension step at 68ºC for 5 mins. The amplified product was then treated with 2 µl Dpn I for 5 mins at 37ºC to digest the parental, methylated DNA before its introduction into ultracompetent XL10-Gold® cells.

2.2.8 Transformation of competent *E. coli* with plasmid DNA

Dpn I-treated products of mutagenesis reactions were introduced into XL10-Gold® ultracompetent cells, while transformations involving other plasmids were performed using competent NM554 *E. coli*. An aliquot of XL10-Gold® ultracompetent cells or competent NM554 cells was thawed on ice. Two to three µl of the DNA sample was then mixed with 45 µl of cells in a pre-chilled 14-mL BD Falcon polypropylene round-

bottomed tube, before incubation on ice for 20 mins, heat-pulsed at 42ºC for 30 secs and incubated on ice for 5 mins. Following the addition of 500 µL of pre-warmed NZY+ broth, the bacteria were allowed to recover by incubation at 37ºC, with shaking at 200–250 revolutions per minute (rpm) for 60 mins. The bacteria were then plated on LB agar containing 100 µg/mL of either ampicillin or kanamycin depending on the plasmid being prepared and incubated at 37ºC overnight.

2.2.9 Plasmid DNA purification

2.2.9.1 Plasmid DNA purification using the QIAprep spin mini kit

A single colony was used to inoculate 5 mL of LB broth containing 100 µg/mL of either ampicillin or kanamycin. The culture was then incubated overnight at 37ºC, with shaking at 200 rpm. A glycerol stock was prepared from each inoculate and stored at -80°C for later use. Plasmid mini-preparations were isolated using the QIAprep spin miniprep kit according to the manufacturers' instructions. In brief, bacterial cells were harvested by centrifugation at 6,000 x *g* for 3 mins and resuspended in 250 µl P1 resuspension buffer. A 250 µl aliquot of P2 alkaline lysis buffer was added to lyse the cells and denature the DNA for 4 mins before the reaction was terminated by adding 350 µl P3 neutralisation buffer, which allows small plasmid DNA molecules to renature into double-stranded DNA, while efficient renaturation is not possible for large genomic DNA. Cell debris, proteins and genomic DNA were removed by centrifugation for 10 mins at ~13,000 x *g*, and the supernatant was then applied to a QIAprep spin column, allowing DNA to bind to the membrane. The column was then centrifuged at ~13,000 x *g* for 10 mins. The eluate was discarded and the column was washed with 750 µL PE buffer and centrifuged again at ~13,000 x *g* for 1 min, followed by a further centrifugation step to remove any residual buffer. DNA was then eluted from the column by the addition of 50 µl EB elution buffer, followed by centrifugation at 13,000 x *g* for 1 min. Direct sequencing was used to confirm the presence of any desired nucleotide changes in the plasmid before use in further work (see section 2.2.3).

2.2.9.2 Plasmid DNA purification using the Qiagen plasmid maxi kit

A sample from a glycerol stock was used to inoculate 250 mL of LB broth containing 100 µg/mL of antibiotic before incubation overnight at 37ºC with shaking at 200 rpm. Bacterial cells were pelleted by centrifugation at 6,000 x *g* for 15 mins and resuspended in 10 mL P1 resuspension buffer. Cell lysis was achieved by the addition of 10 mL of

P2 and incubation for 5 mins. A 10 mL aliquot of chilled buffer N3 was then added to neutralise P2 and the sample incubated on ice for 20 mins to allow renaturation of plasmid DNA. A Qiagen-tip 500 column was equilibrated by allowing flow-through of 10 mL QBT buffer. Following the removal of the cell debris, proteins and genomic DNA by centrifugation at 20,000 x *g* for 10 mins, the DNA-containing solution was applied to the equilibrated column. The column was washed twice with 15–25 mL QC buffer before eluting the DNA in 15 mL QF eluting buffer. The DNA was then precipitated by adding 10.5 mL of isopropanol and centrifuging the sample at 15,000 x *g* for 30 mins at 4ºC. The supernatant was discarded, and the DNA pellet was washed with 5 mL 70% ethanol before centrifugation at 15,000 x *g* for 10 mins. The supernatant was discarded and the DNA pellet allowed to air-dry for 10 mins before it was resuspended in 500–1,000 µl distilled water. The final concentration of the DNA was determined and its purity assessed using the NanoDrop 1000 spectrophotometer as mentioned in section 2.2.2. Direct sequencing was usually conducted before using the plasmid in further studies (see section 2.2.3).

2.2.10 Immortalised mammalian cell line culture

The HEK 293T and Dami established cell lines were utilised for *in vitro* investigations. The reported doubling times for both cell lines are between 24 and 30 hours (hrs) (https://www.dsmz.de/catalogues/details/culture/ACC-635.html) (Greenberg et al., 1988). The HEK 293 line is an adherent human epithelial embryonic kidney cell line derived by adenoviral transformation of cells from an apparently healthy female foetus. HEK 293T cells are a variant of HEK 293 cells that have the SV40 Large T-antigen, which allows for episomal replication of transfected plasmids containing the SV40 origin of replication.

The Dami suspension cell line was derived from the peripheral blood of a 57-year-old male patient with megakaryoblastic leukaemia (Greenberg et al., 1988), and therefore is a more suitable model for investigating the role of variants identified in patients with platelet bleeding disorders. Additionally, when Dami cells are treated with phorbol 12-myristate 13-acetate (PMA) and thrombopoietin (TPO), they differentiate further along the megakaryocyte (MK) lineage (Greenberg et al., 1988; Lev et al., 2011). Cytogenetically, Dami cells are nearly triploid (+1, +2, +3, +4, +5, +6, +8, +11, +12, +13, +15, +17, +19, +20, +21, +22) (Greenberg et al., 1988).

Cells were grown to sub-confluence in T75 flasks, and all passages and other manipulations of cell lines were conducted in a class II type A biological safety cabinet. Media, phosphate buffered saline (PBS) and trypsin were pre-warmed at 37ºC in a water bath before commencing any procedures. Cell lines were maintained in a humidified tissue culture incubator at 37ºC in the presence of 5% CO_2. Cells were passaged twice a week or whenever the microscopic estimation of cell confluency reached 70–90%.

2.2.10.1 Passaging cells

HEK 293T cells are adherent and were cultured in DMEM medium supplemented with 10% FBS and 1X antibiotic-antimycotic reagent. Cells were passaged by first removing the growth medium by aspiration, then washing the cells with 10 mL PBS. Following removal of the PBS, the cells were incubated with 2 mL trypsin for about 2 mins at 37ºC. Following cell detachment, 8 mL of medium was added to neutralise the trypsin. A 0.5–1 mL aliquot of cells was then added to 10–15 mL fresh medium.

Dami cells were cultured in RPMI 1640 medium supplemented with 10% FBS and 1X antibiotic-antimycotic reagent. Cells were passaged by adding 0.5–1 mL of cells from a confluent flask directly into a T75 flask containing 15 mL fresh medium.

2.2.10.2 Cell counting and seeding

Cells were counted manually using a haemocytometer before being seeded. Briefly, 25 µL of a homogeneous cell suspension was loaded into the haemocytometer chamber. Using the 10X objective lens of a light microscope, the focus was set on the grid lines, and the cells in at least three squares, containing 16 smaller squares, were counted. The number of cells per mL was determined by multiplying the average number of cells in each square by 10,000. The number of cells seeded varied depending on the experiment and the surface area of the culture dish (summarised in Table 2.4).

2.2.10.3 Transfection of HEK 293T and Dami cells

HEK 293T and Dami cells were transfected differently using different plasmid concentrations and transfection reagents. HEK 293T cells were transfected using Lipofectamine LTX reagent, while Dami cells were transfected using jetPEI. Both transfection reagents form complexes with DNA molecules allowing the negatively

charged plasmid DNA to overcome the electrostatic repulsion of the cell membrane and enter the cell. Transfection reagents were used according to the manufacturers' instructions. The transfection conditions varied according to the experimental design (see Table 2.4). HEK 293T cells were seeded and incubated overnight before being transfected. When cells reached 50–70% confluency, DNA was first diluted in serum-free DMEM media before adding the Lipofectamine LTX. Dami cells were seeded and transfected at the same time unless otherwise specified. The DNA and jetPEI were diluted separately in 150 mM NaCl. The diluted solutions were vortexed before adding the diluted transfection reagent to the diluted DNA. The ratio of Lipofectamine LTX to DNA used was 1:1, while the jetPEI to DNA ratio was 2:1 unless otherwise specified.

The transfection mixture was vortexed for 15 secs, and incubated at room temperature for 15 mins in the case of Lipofectamine LTX, or 15–30 mins for jetPEI. The transfection mixtures were then added drop-wise using a pipette on top of the cell culture medium, and the plate was gently rocked to ensure homogeneous distribution. Cells were incubated in the tissue culture incubator following transfection and after 24 hrs, the media was replaced in the case of plates containing HEK 293T cells or topped-up in the case of Dami cells.

2.2.10.4 Cryopreservation and thawing of cells

Cells from a confluent T75 flask were collected into a 50 mL Falcon tube and centrifuged for 5 mins at 1,500 rpm. The medium was discarded, and the cell pellet resuspended in 2 mL of cell freezing medium containing 95% FBS and 5% dimethyl sulfoxide. The cell suspension was transferred in 1 mL aliquots into cryovials that were placed in a polystyrene container and frozen at -80ºC for 24 hrs before long-term storage in liquid nitrogen.

Cells were thawed by placing the cryovial directly in a 37ºC water bath for 2 mins. The freezing media was removed by transferring the contents of the cryovial to a Falcon tube containing 5 mL of pre-warmed media, centrifuging the suspension at 1,500 rpm for 5 mins, and discarding the supernatant. The cell pellet was then resuspended in 5 mL of fresh pre-warmed media. The resuspended cells were transferred to a T75 flask containing 10 mL media and incubated in the tissue culture incubator. Cells were passaged at least twice before use in further studies.

Table 2.4 Transfection conditions for HEK 293T and Dami cells

Investigation	Cells	Plate	Seeding density (cells/mL)	Growth medium (per well)	Plasmid DNA concentration	Transfection reagent	Transfection diluent	Total transfection mixture volume
Dual luciferase reporter assay	HEK 293T	24-well plate	4-5 x 10^5	500 µl	500 ng	1 µl	Serum-free DMEM media	100 µl
	Dami	24-well plate	8 x 10^5	500 µl	1,000 ng	2 µl	150 mM NaCl	100 µl
Cellular localisation	HEK 293T	6-well plate	2.5 x 10^5	2 mL	2,500 ng	5 µl	Serum-free DMEM media	250 µl
	Dami	6-well plate	5 x 10^5	2 mL	3,000 ng	6 µl	150 mM NaCl	200 µl
LDS-PAGE and western blotting	HEK 293T	6-well plate	5 x 10^5	2 mL	2,500 ng	5 µl	Serum-free DMEM media	250 µl

DMEM; Dulbecco's modified Eagle's medium, LDS; lithium dodecyl sulphate, PAGE; polyacrylamide gel electrophoresis.

2.2.11 Dual luciferase reporter assays

Dual luciferase reporter assays were performed to examine the ability of FLI1 variants to transactivate the *GP6* promoter and the ability of ETV6 variants to trans-repress the *MMP3* and *PF4* promoters. HEK 293T and Dami cells were transfected with three plasmids: a transcription factor expression plasmid (pSG5-FLI1 or pCMV-ETV6 bearing WT or mutated cDNAs for FLI1 or ETV6, or the empty vectors pSG5 and pCMV), a pGL3-reporter plasmid (containing a *GP6*, *MMP3* or *PF4* promoter fragment upstream of the Firefly luciferase cDNA), and pRL-null. Use of pRL-null allowed the resulting data to be normalised to correct for transfection efficiency. In some experiments, variant transcription factors were co-expressed with the WT counterpart to mimic the heterozygous situation.

2.2.11.1 Transfection of cells for dual luciferase reporter assays

HEK 293T cells and Dami cells were seeded into 24-well plates and transfected as described earlier (see section 2.2.10.3). For each well of HEK 293T cells, the transfection mixture contained 100 ng of expression vector bearing the WT or mutated transcription factor cDNA or the corresponding EV with 200 ng of pGL3-reporter, 200 ng of pRL-null and 1 µl of Lipofectamine LTX in a total volume of 100 µl serum-free DMEM. To transfect each well of Dami cells, a transfection mixture containing 200 ng of the transcription factor expression vector, 400 ng of the pGL3-reporter of interest and 400 ng of pRL-null diluted with 150 mM NaCl to a final volume of 50 µL was prepared. In a separate tube, 2 µL of jetPEI reagent was diluted with 150 mM NaCl to a final volume of 50 µL and mixed by vortexing. The diluted transfection reagent was added to the diluted DNA, the mixture vortexed and incubated at room temperature for 15–30 mins before being added to the cells.

2.2.11.2 Cell lysis and dual luciferase reporter assays

Forty-eight hrs post-transfection, cells were washed with 200 µl of PBS, and after removing the PBS, cells were lysed by incubation at room temperature for 30 mins with 100 µl 1X passive lysis buffer supplied with the Promega® Dual Luciferase® Reporter Assay kit. Each lysate was then divided into four 20 µL aliquots in the wells of a 96-well clear-bottom white plate. Following the addition of 24 µl luciferase reagent to three wells, Firefly luciferase activity was measured using a microplate reader (Varioskan Flash, Thermo Scientific) and SkanIt RE for VarioSkan Flash software (version 2.4.5).

Quenching of luminescence arising from Firefly luciferase and parallel activation of Renilla luciferase were accomplished by the addition of 24 µl of Stop&Glo® reagent and Renilla luciferase activity was then measured using the same settings.

Following subtraction of background luminescence derived from the sample and reagents, Firefly luciferase results were normalised for transfection efficiency using Renilla luciferase. The normalised ratios were then expressed as the fold change in relation to lysates from cells transfected with the EV. The results were plotted as the mean and standard error of the mean (SEM), and paired t-tests were used for comparison.

2.2.12 Localisation of overexpressed FLI1 in mammalian cell lines

Cloning of the cDNA for green fluorescent protein (GFP) either upstream or downstream of the cDNA for a gene of interest, without an intermediate stop codon, usually allows expression of a fusion protein that retains the properties of the GFP without affecting the protein encoded by the gene of interest. The presence of the GFP tag in the hybrid protein facilitates cellular localisation studies. This approach was adopted to further characterise FLI1 variants as part of this study. Direct detection of FLI1 by immunostaining was also carried out.

2.2.12.1 Transfection of cells for cellular localisation studies

HEK 293T and Dami cells were seeded onto glass coverslips in 6-well plates and transfected with the appropriate transcription factor expression plasmid (pEGFP-N2-FLI1, pSG5-FLI1, or derivatives thereof). To promote attachment and differentiation of Dami cells following seeding, 0.2 µL 1 mM PMA and 2 µL 10 µg/mL TPO were added to achieve final concentrations of 100 nM PMA and 10 ng/mL TPO. The cells were incubated overnight before being transfected as described in section 2.2.10.3.

2.2.12.2 Fixing and staining of cells

Following incubation for 24 hrs at 37ºC in 5% CO_2, cells were washed twice with 1–2 mL PBS before adding 1 mL of BD Cytofix/Cytoperm™ Fixation buffer to each well. The plates were then incubated for 20 mins at room temperature on an orbital shaking platform at 130 rpm. The residual fixative solution was then removed in three washing steps each with 1–2 mL of PBS.

When immunostaining was used to detect FLI1, excess formaldehyde was quenched by adding 1 mL 50 mM NH_4Cl to each well for 10 mins before washing the cells three times with PBS. Cells were permeabilised by adding 1 mL 0.2% Triton X-100 in PBS for 5 mins, followed by four washes with PBS. Protein binding sites were blocked by incubation of the coverslips with 1 mL 1% BSA in PBS for 20 mins at room temperature. Following three washes with PBS to remove excess unbound albumin, coverslips were incubated with polyclonal rabbit anti-human FLI1 antibodies diluted 1:250 in blocking buffer to detect FLI1. Following incubation for 1 hr at room temperature, coverslips were washed three times with the blocking buffer, then twice with PBS alone. They were then incubated in the dark for 1 hr at room temperature with a 1:500 dilution of secondary polyclonal goat anti-rabbit IgG (H+L) antibodies conjugated to Alexa Fluor™ 488. Unbound antibodies were removed by three washes with blocking buffer followed by a further three washes with PBS alone.

Coverslips were then mounted onto glass slides using ~300 µL 4',6-diamidino-2-phenylindole (DAPI)-proLong® Gold anti-fade reagent which was prepared by adding 1 µL 1000X DAPI nuclear counterstain to 1 mL proLong® Gold anti-fade reagent. Slides were stored in the dark at room temperature for 24 hrs to allow the mounting medium to dry. The coverslip edges were then sealed with clear varnish and the slides stored in the dark at 4ºC before examination (usually within 2–3 weeks).

2.2.12.3 Microscopy

The inverted Ti-eclipse Nikon widefield microscope with dual camera system housed in the Wolfson Light Microscopy Facility at the University of Sheffield was used to examine and quantify the distribution of proteins of interest within the cells. Slides were imaged for DAPI (blue; stain maximum excitation 350 nm, maximum emission 470 nm) and either GFP (bright green; maximum excitation 488 nm, maximum emission 510 nm) or Alexa Fluor 488 (bright green: maximum excitation 490 nm, maximum emission 525 nm). Images were acquired using NIS elements advance research software (version 4) with the following specification: 100X oil objective lenses, 395 nm and 470 nm exciting laser using SpectraX LED excitation, with specific emission filter for quad-DAPI and quad-FITC, and detected with Dual Andor Zyla sCMOS, 2560 x 2160, 6.5 µm pixels. Images were taken for 8–30 cells per experiment and for each cell, a series

of Z-stacked images were taken at 0.2–0.3 µm steps from the top to the bottom of the cell.

Images were stored as ND2 files and opened, processed and quantitated using ImageJ software (version 1.5). After opening an image, the Z-stacks were compressed to obtain the maximum intensity projection. The freehand selection tool was used to draw borders around the signals from DAPI and FLI1, and the signal areas were quantified using the programme. The percentage of nuclear FLI1 was equivalent to the proportion of total FLI1 (fluorescent FLI1-EGFP) which co-localised with the nuclear DAPI signal. The results were plotted as the mean and SEM, and t-test was used for comparison.

On a single occasion, super-resolution images were acquired for cells expressing WT-FLI1-EGFP and R340C-FLI1-EGFP by members of the technical team in the facility using the Structured Illumination super-resolution microscope and images were provided as JPEG files.

2.2.13 Analysis of the distribution of FLI1 and ETV6 variants in HEK 293T cells

LDS- polyacrylamide gel electrophoresis (PAGE) and western blotting were carried out to examine the subcellular localisation of FLI1 and ETV6 variants and to assess their ability to translocate to the nucleus in HEK 293T cells.

2.2.13.1 Transfection of HEK 293T cells for protein extraction

HEK 293T cells were seeded into 6-well plates and transfected with the expression plasmid bearing the WT, mutated cDNA of interest (pSG5-FLI1 or pCMV6-ETV6) or EV as described in section 2.2.10.3.

2.2.13.2 Extraction of nuclear and cytoplasmic proteins from HEK 293T cells

Forty-eight hrs post-transfection, the nuclear and cytoplasmic proteins were extracted from cells using the NE-PER™ nuclear and cytoplasmic extraction reagents according to the manufacturers' instructions. In brief, harvested HEK 293T cells were washed with ice-cold PBS, and the supernatant discarded after pelleting the cells by centrifugation at 500 x *g* for 5 mins, leaving the cell pellet as dry as possible. The pelleted cells were resuspended by adding 200 µL ice-cold CER I reagent and 2 µL 100X PIC and vortexing vigorously for 15 secs. The cell suspension was then

incubated for 10 mins on ice, before adding 11 µl ice-cold CER II. The mixture was vortexed vigorously twice for 5 secs with an intervening one-minute incubation on ice. The combined use of CER I and CER II disrupted the outer cell membrane and released the cytoplasmic contents, which were collected by centrifugation at 16,000 x *g* for 5 mins at 4ºC. The supernatant was immediately removed into a clean pre-chilled tube and the remaining insoluble nuclear pellet was then resuspended in 25 µL ice-cold NER and 0.25 µL 100X PIC and incubated on ice. Every 10 mins, the suspension was vortexed vigorously for 15 secs before being returned to the ice. After 40 mins, the mixture was centrifuged at 16,000 x *g* for 10 mins at 4ºC, and the supernatant containing the nuclear proteins was immediately transferred to a clean pre-chilled tube. All extracts were aliquoted before storage at -20ºC.

2.2.13.3 Determination of protein concentration in cellular extracts

Protein concentration was determined using the Pierce bicinchoninic acid (BCA) Protein Assay Kit according to the manufacturers' instructions. In brief, following preparation of the standards and working reagent as instructed, a 96-well plate was placed on ice and the wells loaded with 10 µL of the standards and test samples. A 200 µL ice-cold BCA working reagent was added to each well, and the plate incubated at 37ºC for 30 mins. The plate was then cooled to room temperature on ice before measuring the absorbance of sample wells at 562 nm using a plate reader (Varioskan Flash, Thermo Scientific) and SkanIt RE for VarioSkan Flash software (version 2.4.5). Protein concentrations were interpolated from the standard curves using GraphPad Prism.

2.2.13.4 Sample preparation and lithium dodecyl sulphate- polyacrylamide gel electrophoresis

LDS-PAGE was carried out using a Novex™ NuPAGE® Gel and XCell SureLock™ Mini-Cell system. Samples were prepared for electrophoresis in pre-chilled tubes by diluting 10 µg of protein with 2.5 µL 4X NuPAGE LDS sample buffer, 1 µL 10X NuPAGE reducing agent and water to a final volume of 12.5 µL. Samples were denatured by boiling at 100°C for 10 mins before loading 10 µL aliquots into 1.5 mm NuPAGE 4–12% Bis-Tris gels. A 5–10 µL aliquot of Chameleon duo pre-stained protein ladder was included in every electrophoresis run. NuPAGE™ MOPS SDS running buffer was used for electrophoresis, and 500 µL of NuPAGE® antioxidant was added to the upper buffer

chamber to maintain proteins in a reduced state during electrophoresis. The samples were separated for 50–60 mins at 200 V at room temperature in an ice bath.

2.2.13.5 Electrophoretic transfer of proteins

Western blotting was carried out in a semi-wet XCell II™ blot module using 1X NuPAGE® transfer buffer supplemented with 10% methanol (or 20% for two blots) and 0.1% NuPAGE® antioxidant to enhance protein transfer. An Invitrolon™ 0.45 µm pore size PVDF membrane was activated by soaking the membrane in methanol for 30 secs. After rinsing the membrane in deionised water, it was saturated in the transfer buffer for several mins before assembly of the blot.

Following electrophoresis, the gel-membrane sandwich was assembled from the cathode to the anode in the following order: 2–3 layers of blotting pads, filter paper, the gel which was trimmed to remove the wells and the foot, the transfer membrane, filter paper, and 2–3 further layers of blotting pads. When two gels were blotted at the same time, the layers were repeated in the same order. All the blotting pads, pieces of filter paper and transfer membranes were pre-soaked in transfer buffer and kept saturated during assembly. The blot module was then filled with transfer buffer while the outer buffer chamber was filled with water to dissipate the heat produced during blotting. Protein transfer was carried out overnight at 4ºC at a constant voltage of 15 V or at 18 V if two gels were blotted together.

2.2.13.6 Protein detection on western blots

Following protein transfer, any remaining protein binding sites on the membrane were blocked using Odyssey® PBS blocking buffer for 1 hr. The membrane was then incubated for 1 hr with a 1:500 dilution of polyclonal rabbit anti-human FLI1 or ETV6 diluted in blocking buffer containing 0.2% Tween20. Additionally, monoclonal rabbit anti-human Cyclophilin B, which acts as a cytoplasmic marker, and monoclonal mouse anti-human TATA-binding protein (TBP), which serves as a nuclear marker, were also used, both at a 1:1000 dilution. The membrane was washed 4 times, each time for 5 mins, with 15 mL 0.1% Tween20-PBS, before being incubated for 1 hr with the secondary antibody which was diluted in blocking buffer containing 0.2% Tween20. Two secondary polyclonal antibodies were used, each at a 1:10,000 dilution, donkey anti-rabbit immunoglobulins conjugated to IRDye 680RD and donkey anti-mouse

immunoglobulins conjugated to IRDye 800CW. The membrane was then washed 4 times, each for 5 mins with 15 mL 0.1% Tween20-PBS, then twice, each for 5 mins with 15 mL PBS alone to remove the Tween20. All steps were carried out at room temperature using a rolling mixer for the incubations and a rocking shaker for the washes. Membranes were protected from light as much as possible at all stages of the detection procedure.

Western blots were scanned using the Odyssey® Sa Infrared Imaging System (LI-COR) through the Image studio software (version 2.0.13) set on the following parameters: focus offset 2 mm, intensity 7, and resolution 100 µM. The density of the band in the acquired image was then measured using Image Studio Lite software (version 5.2), where the average intensity of 3 pixels bordering each band was set as the background. The intensity of either the FLI1 or the ETV6 band on the western blot was normalised to the relevant nuclear (TBP) or cytoplasmic (cyclophilin B) housekeeping protein, and the ratio of FLI1 or ETV6 localized in the nucleus to that localized in the cytoplasm was calculated using the following equation:

$$\frac{(\text{Target protein in Nuc} \,/\, \text{TBP in Nuc})}{(\text{Target protein in Cyto}/\, \text{Cyclophilin B in Cyto})}$$

The percentage of nuclear FLI1 was determined from the same data using the following equation:

$$\frac{(\text{FLI1 in Nuc} \,/\, \text{TBP in Nuc})}{(\text{FLI1 in Nuc} \,/\, \text{TBP in Nuc}) + (\text{FLI1 in Cyto}/\, \text{Cyclophilin B in Cyto})} \text{X}100$$

The results were plotted as the mean and SEM, and a Mann-Whitney test was used to compare results.

2.2.14 Use of CRISPR/Cas9 to knockout/down FLI1 in Dami cells

A CRISPR/Cas9 gene editing approach was adopted to generate stable Dami cell lines in which *FLI1* expression was knocked out/down. The procedure was similar to that described by Bauer et al. (2015).

2.2.14.1 CRISPR guide design

CRISPR guides for editing *FLI1* were designed using several online tools ([accessed 2017–2018], links for all tools are found in Appendix 3), commencing with the Zhang

lab tool (Hsu et al., 2013) which suggested several candidate guide sequences targeting *FLI1*. Further tools were then used to facilitate selection between the top 5 candidate guides including, the Efficiency Prediction tool (Housden et al., 2015), the Integrated DNA Technologies (IDT) tool, the CRISPOR tool (Haeussler et al., 2016), the CCTop - CRISPR/Cas9 target online predictor tool (Stemmer et al., 2017) and the ChopChop tool (Labun et al., 2016). The guides that were predicted to be effective by the majority of the design tools, which also avoided the intron-exon junctions, were selected.

To facilitate cloning into the CRISPR plasmid pX458 using BbsI, and expression of the guide from the U6 promoter, additional nucleotides were added to each guide to result in 24- or 25- oligonucleotides. Thus, the forward oligonucleotides had the sequence 5´-CACCg-target sequence-3´ and the reverse complement oligonucleotide had the sequence 5´-AAAC-target sequence-c-3´. However, guanine was not added after the CACC sequence to the forward oligonucleotide and cytosine was not added to the 3' end of the reverse complement oligonucleotide if guanine was the first nucleotide in the guide sequence.

2.2.14.2 Cloning of the CRISPR guides

Forward and reverse CRISPR guides were purchased as phosphorylated oligonucleotides (mentioned in section 2.1.2) and cloned into pX458. Before cloning, guides were annealed to form double-stranded DNA oligonucleotides. Annealing reactions contained 1 µL of 100 µM stock of each guide oligonucleotide and 1 µl 10X T4 ligation buffer in a final volume of 10 µl. Oligonucleotides were heated in a thermocycler at 37ºC for 30 mins, then denatured at 95ºC for 5 mins before being allowed to anneal by reducing the temperature to 25ºC at 5ºC/min. The annealed oligos were ligated into the CRISPR plasmid using the Golden Gate cloning, where digestion and ligation occur in the same reaction. Ligation mixtures contained 100 ng pX458 circular vector, 1 µL 1 µM annealed oligonucleotides, 5 µl 10X restriction enzyme buffer, 20 U BbsI, 10 mM ATP, 5 µg BSA, and 750 U T4 DNA ligase in a final volume of 50 µl. Samples were then subjected to 20 cycles of restriction digestion at 37ºC for 5 mins and ligation at 20ºC for 5 mins, before inactivating the enzymes by heating at 80ºC for 20 mins. The ligated oligonucleotides were then introduced into competent *E. coli* NM554 cells, mini-preparations of transformed colonies were isolated, and the

integrity of the inserted guide sequence was confirmed by sequencing using the U6 promoter primer, before maxi-preparation of the plasmid (see sections 2.2.3, 2.2.8, 2.2.9). The purified plasmid was sequenced again using the same primer before transfection.

2.2.14.3 Transfecting Dami cells with CRISPR plasmids

The conditions for transfection of Dami cells using jetPEI were optimised by transfecting cells with a range of concentrations of pEGFP-N3 (1,000–12,000 ng) at different jetPEI:DNA ratios (4:1, 3:1, 2:1 and 1:1) as described earlier (see section 2.2.10.3). Cell viability and transfection efficiency were assessed 48 hrs post-transfection using fluorescence-activated cell sorting (FACS) (see section 2.2.14.4).

In the optimised protocol, one million Dami cells in 1 mL RPMI 1640 were seeded into each well of a 6-well plate. Cells were transfected with a total of 3,000 ng of CRISPR plasmids (1,500 ng of each plasmid) diluted to 100 µl with 150 mM NaCl, and 9 µl of jetPEI reagent diluted to 100 µl with 150 mM NaCl to achieve a jetPEI:DNA ratio of 3:1. Following overnight incubation, cells were supplemented with 2 mL of fresh media and sorted 48 hrs post-transfection using FACS (see section 2.2.14.4)

2.2.14.4 Fluorescence-activated cell sorting

Forty-eight hrs post-transfection, the viability and efficiency of transfection of cells were assessed using the BD FACSAria IIu (BD Bioscience) housed in the Flow Cytometry Core Facility at the University of Sheffield Medical School. Where appropriate, cells were also sorted using the EGFP signal. Cells were prepared for analysis by passage through a 50 µm filter into a FACS tube and then loaded into the machine. Viable and single cells were gated initially according to side scatter areas (SSC-A) and forward scatter areas (FSC-A), and their distribution on the forward scatter height (FSC-H) and FSC-A plot, respectively. EGFP positive cells were detected using the blue 530/30 filter and gated according to SSC-A. Figure 2.1 illustrates the parameters used to sort EGFP positive cells.

Gated cells were bulk sorted into a collection tube containing 1 mL of media by the cell sorter and then sorted into single cells by limiting dilution into a 96-well plate containing 100 µL of fresh media per well. Limiting dilution was carried out by first counting the

cells using a haemacytometer. The volume of medium containing 30 cells was then diluted to 10 mL to achieve a concentration of 3 cells/mL, before transferring 100 µL of the cell suspension into each well of a 96-well plate. The final cell concentration was 0.3 cells/well, ensuring that some wells received a single cell and minimising the likelihood that any well received more than one cell. Single cells were then allowed to expand undisturbed in the incubator for 14 days.

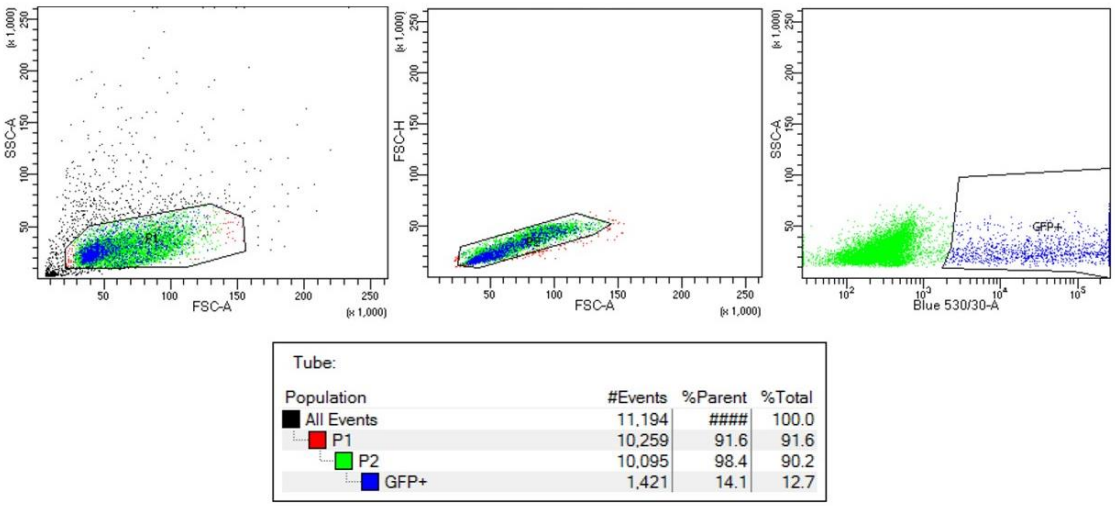


Figure 2.1 Fluorescence-activated cell sorting of cells expressing enhanced green fluorescent protein

2.2.14.5 Genomic DNA extraction from CRISPR-edited Dami cells

Genomic DNA (gDNA) was extracted from CRISPR-edited Dami cells using one of two methods. QuickExtract™ DNA extraction solution was used to isolate PCR-ready gDNA from an aliquot of cells grown in 96-well plates. Cells were pelleted by centrifuging at 1,500 rpm for 5 mins. The media was decanted and cell pellets were each resuspended in 50 µl extraction solution before being transferred to 0.5 mL PCR tubes/plate. DNA was extracted by heating at 65ºC for 6 mins, then 98ºC for 2 mins in a thermocycler before storage of samples at 4ºC prior to analysis.

The GenElute™ mammalian genomic DNA miniprep kit was used to extract higher grade DNA for further characterisation. Cells from a confluent 6-well plate were pelleted by centrifugation for 5 mins at 300 × *g*. After removing the culture medium and washing the cells with PBS, the pellet was resuspended in 200 µL of resuspension

solution. To eliminate RNA and protein, 20 µL of RNase A solution was added, the sample mixed and incubated for 2 mins, before addition of 20 µL of the Proteinase K solution. Cells were lysed by adding 200 µL of Lysis Solution C, vortexing the mixture thoroughly until it was homogenised, then heating at 70ºC for 10 mins. During sample incubation, the binding column was equilibrated by adding 500 µL of the column preparation solution and allowing it to flow-through via centrifugation. The lysate was mixed with 200 µL of 100% ethanol before being loaded onto the equilibrated column using a wide bore pipette tip to reduce DNA shearing. The column was centrifuged, and the flow-through discarded. The column was then washed twice with 500 µL wash solution, discarding the flow-through each time, and centrifuged again for 3 mins at 13,000 × *g* to remove any residual ethanol. DNA was eluted into a new collection tube by adding 200 µL elution solution, incubation for 5 mins and a final centrifugation step. All DNA extraction steps were performed at room temperature, and all centrifugation steps were at 13,000 × *g* for 1 min unless otherwise specified.

To have DNA suspended in water rather than the elution solution, eluted DNA was precipitated by adding 20 µL 3M sodium acetate and 550 µL 100% cold (-20ºC) ethanol. Following incubation for 10 mins to overnight at -20ºC, DNA was recovered by centrifugation at 13,000 ×*g* for 15 mins at 4ºC and washed with 500 µL 70% cold (-20ºC) ethanol. After removing the ethanol, the DNA pellet was air-dried for 10 mins then resuspended in 100–200 µL water.

2.2.14.6 Screening for CRISPR-mediated edits at the DNA level

Following extraction of genomic DNA, preliminary screening for the desired *FLI1* edit was carried out by PCR amplification across the region that was targeted for deletion followed by agarose gel electrophoresis of the amplified product as described in sections 2.2.4 and 2.2.5. Intact alleles were detected in those clones that had been edited by amplification of a DNA fragment from within the target region using a second set of primers. The primers used for screening are indicated by an asterisk in Table 2.1.

2.2.14.7 Protein extraction from CRISPR-edited Dami cells

Cells were harvested from a 6-well plate and washed twice with ice-cold PBS by centrifugation at 1,000-1,500 rpm for 5 mins and discarding the supernatant. The

washed cell pellet was resuspended in 50-200 µL of ice-cold RIPA buffer mixed with 0.5-2 µL 100X PIC depending on the cell density and incubated on ice. Every 10 mins, the suspension was vortexed vigorously for 15 secs and returned to the ice. After 30-40 mins, the mixture was centrifuged at 14,000 ×*g* for 15 mins at 4ºC, and the supernatant, containing the cellular proteins was immediately aliquoted into a clean pre-chilled tube and stored at -20ºC prior to analysis.

2.2.14.8 FLI1 protein quantification in CRISPR-edited Dami cells

FLI1 knockout/down was assessed by measuring FLI1 protein levels by LDS-PAGE and immunoblotting of cellular lysates from CRISPR-edited Dami cells as described in sections 2.2.13.4 to 2.2.13.6. The housekeeping protein was detected by probing the blot with a 1:500 dilution of monoclonal mouse anti-human β-tubulin I antibody.

2.2.14.9 RNA extraction from CRISPR-edited Dami cells

Total RNA was extracted using the EZ-RNA kit which involves disruption of the cells in guanidine thiocyanate/detergent solution and extraction of RNA using phenol and chloroform and its recovery by precipitation with alcohol. The extraction was performed according to the manufacturers' instructions in a class II type A biological safety cabinet and in a pre-PCR room at a workstation reserved for RNA work. Briefly, following harvesting and washing cells with PBS by centrifugation at 1,000–1,500 rpm for 5 mins and discarding the supernatant, 500 µL denaturation solution was added to a PBS-washed cell pellet, followed 5 mins later by 500 µL extraction solution. The suspension was vortexed vigorously for 15 secs, incubated for 10 mins and then centrifuged at 13,000 x *g* for 15 mins, allowing separation of the RNA into the upper aqueous phase, while the DNA and proteins remained in the interphase and the lower organic phase respectively. Carefully, 300–400 µL of the upper phase was transferred into a new 1.5 mL Eppendorf tube, and the RNA precipitated by adding an equal volume of isopropanol. Following incubation for 10 mins, RNA was collected by centrifugation at 14,000 x *g* for 8 mins. The isopropanol was removed and the RNA pellet washed with 500 µL 75% cold (-20ºC) ethanol before being centrifuged at 14,000 x *g* for 5 mins. The alcohol was removed and the RNA pellet was air-dried for 5 mins before being resuspended in 50–100 µL of ice-cold RNase-free water. Following quantitation using the NanoDrop, the RNA was aliquoted into pre-chilled tubes and stored at -20ºC. All

incubation steps were carried out at room temperature and the centrifugation steps were at 4ºC.

2.2.14.10 Quantitative polymerase chain reaction of *FLI1* expression using TaqMan™ probes

A sample of RNA was usually transcribed to cDNA immediately following extraction. Alternatively, an aliquot of RNA was thawed on ice along with the reagents from the QuantiTect® reverse transcription kit prior to conversion to cDNA. Reverse transcription was performed according to the manufacturers' instructions, assembling the reaction mixtures on ice, using RNase-free water. In a 0.5 mL PCR tube, 1,000 ng of RNA was mixed with 2 µL gDNA wipeout buffer in a final volume of 14 µL. The mixture was incubated for 5 mins at 42ºC, then returned to 4ºC, before adding 4 µL 5X Quantiscript RT buffer, 1 µL RT primer mix and 1 µL Quantiscript reverse transcriptase and incubating the sample at 42ºC for 15 mins. The reverse transcriptase was then inactivated by heating at 95ºC for 3 mins. After cooling to 4ºC, and the product was either used directly for quantitative polymerase chain reaction (qPCR), or stored at 4ºC for the short-term or at -20ºC for long-term periods. A GeneAmp® PCR system 9700 (Applied Biosystems®) thermocycler was used for all incubation steps. Two control samples were included in the procedure, the non-template control which contained no RNA, and the no reverse transcriptase control.

The 20 µL cDNA reaction was diluted to 100 µL with RT-PCR grade water. The *FLI1* probe (Hs00956711_m1) was FAM™ labelled, while the probe for the housekeeping beta-2 microglobulin gene (Hs99999907_m1) was VIC™ labelled; both probes spanned exons to minimise detection of any contaminating genomic DNA. Expression of *FLI1* and the housekeeping gene were measured for each test sample in separate reactions in triplicate. A reaction master mix was prepared for each probe which included 5 µL 2X TaqMan™ gene expression master mix and 0.5 µL 20X TaqMan™ gene expression assay per replicate. A 384-well reaction plate was loaded with 6 µL of the master mix, then with 5 µL of the diluted cDNA reaction using an automated pipette dispenser. After sealing with optical adhesive film, the plate was inverted and tapped to mix the contents of the wells, then centrifuged briefly to ensure that the liquid was at the bottom of the wells. The TaqMan™ probes were kept ice-cold and protected from light as much as possible at all stages of the procedure.

Using the SDS software (version 2.4), the plate was loaded into the 7900HT Fast Real-Time PCR system (Applied Biosystems™) and subjected to standard thermal cycling conditions: 95ºC for 10 mins, followed by 50 cycles of denaturation at 95ºC for 15 secs, and annealing and extension at 60ºC for 1 min. When the run was complete, the RQ manager software (version 1.2.1) was used to calculate the relative quantity (RQ) of *FLI1* knockdown in CRISPR-edited Dami cells, following normalisation to the housekeeping gene. The default settings were used and the results were plotted as log10 of the RQ values.

2.2.14.11 Characterisation of CRISPR-mediated edits in Dami cell clones by TA cloning

Double strand breakpoints in the *FLI1* gene that were introduced using CRISPR/Cas9 were mapped by sequencing DNA amplified from selected clones either directly or following TA cloning. DNA was amplified for TA cloning using a high fidelity DNA polymerase. PCRs contained 100 ng of genomic DNA, 10 µM forward primer, 10 µM reverse primer, 2.5 µL 10X PCR Buffer with 17.5 mM $MgCl_2$ (Expand Long Template Buffer 1), and 0.375 µl Expand long template enzyme in a total volume of 25 µl. Following thermal cycling as described earlier (see section 2.2.4), amplicons were purified as described in section 2.2.6. Prior to TA cloning, deoxyadenosine nucleotide overhangs were added to the purified amplicon by mixing 4 µl of the purified DNA with 1 µl 5X colourless GoTaq® reaction buffer, 1 µl 2 mM dATP and 0.2 µl GoTaq® G2 DNA polymerase and incubating the sample at 72ºC for 30 mins in the GeneAmp® PCR system 9700 (Applied Biosystems®).

The product was ligated to the pGEM-T vector using a vector to insert ratio of 1:3 and the manufacturers' recommended protocol. The Biomath Ligations calculator from Promega (see Appendix 3 for link) was used to calculate the amount of insert required for ligation, assuming the vector size and quantity were 3,015 bp and 50 ng respectively. Ligation mixtures containing the required amount of PCR product, 5 µl 2X ligation buffer, 1 µl pGEM-T vector, 1 µl T4 DNA ligase in a final volume of 10 µl were prepared and incubated overnight at 4ºC. Competent NM554 *E. coli* cells were then transformed with 2–3 µL of the ligation mixture as described previously (see section 2.2.8). Mini-preparations of plasmid DNA were purified from individual colonies and sequenced using the T7_F primer (as described in sections 2.2.9.1 and 2.2.3)

2.2.14.12 Differentiation of Dami cells

To investigate the role of FLI1 in late megakaryopoiesis, Dami cells were treated with PMA and TPO both before and following knockdown of *FLI1*. PMA is a potent activator of the protein kinase C signalling pathway, which in Dami cells leads to megakaryocytic maturation, an increase in ploidy and expression of markers of late megakaryopoiesis including glycoproteins IIb/IIIa and Ib (Greenberg et al., 1988; Long et al., 1984; Lumelsky & Schwartz, 1997). TPO is a critical haematopoietic cytokine that acts primarily as a regulator of MK progenitor expansion and differentiation. It is essential for the production of mature α-granules in the cytoplasm of Dami cells treated with PMA, and it drives further MK differentiation (Briquet-Laugier et al., 2004). FLI1 levels in Dami cells that are differentiated using PMA and TPO have been observed to peak twice, at day one and following 5 to 7 days of treatment, remaining moderately elevated after that (Lev et al., 2011). To reflect the maximum effect of FLI1 on the Dami cells, RNA for transcriptome analysis was extracted after 6 days of treatment.

Following an established protocol for Dami differentiation (Lev et al., 2011), Dami cells were seeded into T75 flasks at a density of 2.5 x 10^5 cells/mL in 10 mL of medium. Cells were treated daily for 6 days using 100 nM PMA (1 µL of stock 1mM PMA solution) and 10 ng/mL TPO (10 µL of stock 10 µg/mL TPO solution). The medium was replaced on day 3 before adding the PMA and TPO and likewise on every subsequent day. On day 7, stimulated cells were washed with PBS before they were collected in 500 µL detergent solution for RNA extraction using a cell scraper. RNA was then extracted as previously described (see section 2.2.14.9).

2.2.14.13 Transcriptome analysis of Dami cells

Transcriptome analysis was performed in the Microarray and Next Generation Sequencing Core Facility, housed in the Sheffield Institute for Translational Neuroscience (SITraN) at the University of Sheffield. Prior to analysis, RNA quality, integrity and concentration were assessed using the NanoDrop spectrophotometer and Agilent 2100 bioanalyzer. Transcriptome analysis was carried out using Clariom™ D Assay_human chips (Affymetrix; Applied Biosystems™) with the Affymetrix Gene Chip Microarray, Affymetrix Gene Chip hybridisation and the Illumina HiScan SQ systems. The Clariom™ D Assay_human chip has more than 6,765,500 probes enabling

detection of in excess of 542,500 transcripts from more than 134,700 genes (https://www.thermofisher.com/order/catalog/product/902922).

Results were analysed using the Affymetrix Transcriptome Analysis Console (TAC) software (version 4.0.1.36) and Clariom^{TM} D human (version 2.0). Transcriptome data were compared between the WT clones, both before and following differentiation of the cells. Also, transcriptome data were compared between the WT and *FLI1* knockdown clones, both before and following differentiation of the cells. Transcripts that showed a ≥1.5 or ≤-1.5-fold log change in expression, and $p<0.05$ were highlighted (a screenshot from the programme showing the different settings used is displayed in Appendix 4).

2.2.15 Statistical analysis

GraphPad Prism software (version 7.03) was used for all statistical calculations. A *p*-value less than 0.05 was considered significant. Different statistical tests were used depending on the data being analysed. Normality of the data was assessed using the D'Agostino-Pearson normality test. To compare differences between two groups, paired or unpaired T-tests were used when data showed a Gaussian distribution. Where data were not normally distributed a Mann-Whitney test was used. The results were plotted as mean and standard error of the mean unless otherwise specified.

3 Chapter 3. Identification of the Underlying Genetic Defects in Patients with Unexplained Inherited Platelet Bleeding Disorders Using Whole Exome Sequencing

3.1 Introduction

3.1.1 Use of next-generation sequencing to investigate patients with inherited platelet bleeding disorders

The development of next-generation sequencing (NGS) technologies has enabled large-scale simultaneous genetic analysis of target gene panels, exomes or whole genomes. Whole exome sequencing (WES) has been applied widely, helping to establish the genetic basis for many diseases, better management options and deeper understanding of the physiological and pathological aspects of many syndromes (Bainbridge et al., 2011; Dixon-Salazar et al., 2012; Stark et al., 2016). WES permits analysis of the coding regions and flanking intronic sequences, which account for 1–3% of the whole genome. Since most of the known genetic causes of human disease are located in the exons, it offers an attractive approach for large-scale genetic analysis of patients with unexplained inherited disorders. With recent advancements in this technology, the challenge has shifted from the procedures and methodology involved in sequence analysis to data interpretation, assessing the functional impact of novel gene defects and linking the causative gene defects to particular clinical features (Gilissen et al., 2012). The challenge is to identify one or more genetic variants that are responsible for the disorder among the 20,000-50,000 genetic variants that are recognised in each patient. When candidate causative variants are highlighted, their significance often remains elusive, particularly where they occur in genes encoding proteins with unknown functions (Gilissen et al., 2012). Nonetheless, the development of pipelines for analysis of NGS data, bioinformatic tools for predicting the effects of candidate gene defects, and strategies for prioritising candidate variants are helping to increase the power of this technique (Ku et al., 2012).

WES has facilitated the discovery of many novel genes associated with inherited platelet bleeding disorders (IPDs). The first of these was *NBEAL2* (Neurobeachin Like 2), which was found to be associated with Grey platelet syndrome (GPS) in three different studies (Albers et al., 2011; Gunay-Aygun et al., 2011; Kahr et al., 2011). Although these three studies utilised different methods, they adopted similar approaches, combining the power of genome-wide analysis and homozygosity mapping with WES. All three studies were based on the previous report from Gunay-Aygun and colleagues that the locus at chromosome 3p21.1-22.1 was significantly

associated with GPS, which limited the number of candidate genes to 197 (Gunay-Aygun et al., 2010). Kahr and colleagues relied on sequencing of RNA isolated from the platelets of an individual with GPS, which led directly to the identification of *NBEAL2* as the causative gene (Kahr et al., 2011). Albers et al. (2011) sequenced DNA from four patients with GPS and prioritised candidate genes using the knowledge that GPS is a rare recessive disorder, and that causative variants would therefore be novel, inherited as homozygous or compound heterozygous defects, and that these would be considered deleterious to the function of the encoded protein using bioinformatic prediction tools (Albers et al., 2011). Gunay-Aygun and colleagues sequenced DNA from 15 unrelated GPS patients and identified *NBEAL2* as the causative gene in all patients using similar assumptions (Gunay-Aygun et al., 2011). Subsequently, segregation of *NBEAL2* variants within the families of affected patients was confirmed in both studies (Albers et al., 2011; Gunay-Aygun et al., 2011). Since the identification of *NBEAL2* as the causative gene in GPS, NGS technologies have implicated several other genes in IPDs, including those encoding *ACTN1* (Actinin Alpha 1), *ANKRD18A* (Ankyrin Repeat Domain 18A), *CDC42* (Cell Division Cycle 42), *ETV6* (ETS Variant 6) , *FLI1* (Fli-1 Proto-Oncogene), *FYB* (FYN Binding Protein 1), *GBA* (Glucosylceramidase Beta) , *GFI1B* (Growth Factor Independent 1B Transcriptional Repressor), *GNE* (Glucosamine (UDP-N-Acetyl)-2-Epimerase/N-Acetylmannosamine Kinase), *KDSR* (3-Ketodihydrosphingosine Reductase), *PRKACG* (Protein Kinase CAMP-Activated Catalytic Subunit Gamma), *PTPRJ* (Protein Tyrosine Phosphatase, Receptor Type J), *RASGRP2* (RAS Guanyl Releasing Protein 2) , *RBM8A* (RNA Binding Motif Protein 8A), *SBF2* (SET Binding Factor 2) , *SRC* (SRC Proto-Oncogene), *THPO* (Thrombopoietin) (Abuzenadah et al., 2013; Albers et al., 2012; Canault et al., 2014; Dasouki et al., 2013; Ferreira et al., 2017; Futterer et al., 2018; Hamamy et al., 2014; Izumi et al., 2014; Kunishima et al., 2013; Marconi et al., 2018; Morgan et al., 2013; Motokawa et al., 2018; Noetzli et al., 2015; Noris et al., 2018; Stockley et al., 2013; Takeichi et al., 2017; Takenouchi et al., 2015; Takenouchi et al., 2016; Turro et al., 2016; Wan et al., 2017; Zhang et al., 2015).

WES analysis will normally identify about 25,000 genetic variants in the exome of a single individual (Gilissen et al., 2012). This enormous number of variants necessitates adopting various strategies to filter and prioritise the data for downstream analyses or, indeed, to identify the variants that are most likely to explain the patient's phenotype.

In most cases, prioritisation of WES data to identify a disease-causing variant usually commences with the exclusion of common variants listed in public databases, then focuses on non-synonymous, nonsense and frameshift variants. The removal of common and synonymous single nucleotide variants (SNVs) usually results in a panel of candidate causative variants, and further prioritisation is typically required to reduce the number of candidate variants for further analysis using a variety of approaches (Figure 3.1). There are several strategies: (i) examination of segregation of candidate variants among family members (linkage analysis) requires WES data to be available for other family members, either affected and/or unaffected, and allows variants that are present in both the patient and unaffected members to be excluded, while those shared among only affected family members are retained for further investigation; (ii) homozygosity mapping in those cases where there is a history of consanguinity takes into consideration homozygous variants that are contained within large regions of homozygosity; (iii) when the disease is inherited recessively, considering only homozygous as well as compound heterozygous variants will be helpful, especially in cases when other family members are unavailable for study (double-hit strategy); (iv) data can be compared to those of unrelated patients who have similar phenotypes on the basis that they are likely to have defects in the same or different genes which interact or play a role in the same, or in a related pathway (overlap strategy); (v) a focus on de novo variants (de novo strategy); and (vi) applying bioinformatic tools can predict the possible impact of candidate gene defects on protein structure and function (candidate gene strategy). The overall diagnostic success rate for such approaches in identifying causative variants in Mendelian disorders has been reported to be 30%-60% (Gilissen et al., 2012; Stark et al., 2016). Innovative ways for handling genetic data that allow candidate gene prioritisation continue to be developed. For example, the Human Phenotype Ontology (HPO) project is an international open-source coding system which uses logarithmic computational software to link candidate disease-associated genes identified by WES to patient phenotypes using standardised phenotypic terms (Kohler et al., 2014). This approach has recently proved to be effective in identifying genetic defects in patients with IPDs (Stritt et al., 2016; Turro et al., 2016; Westbury et al., 2015).

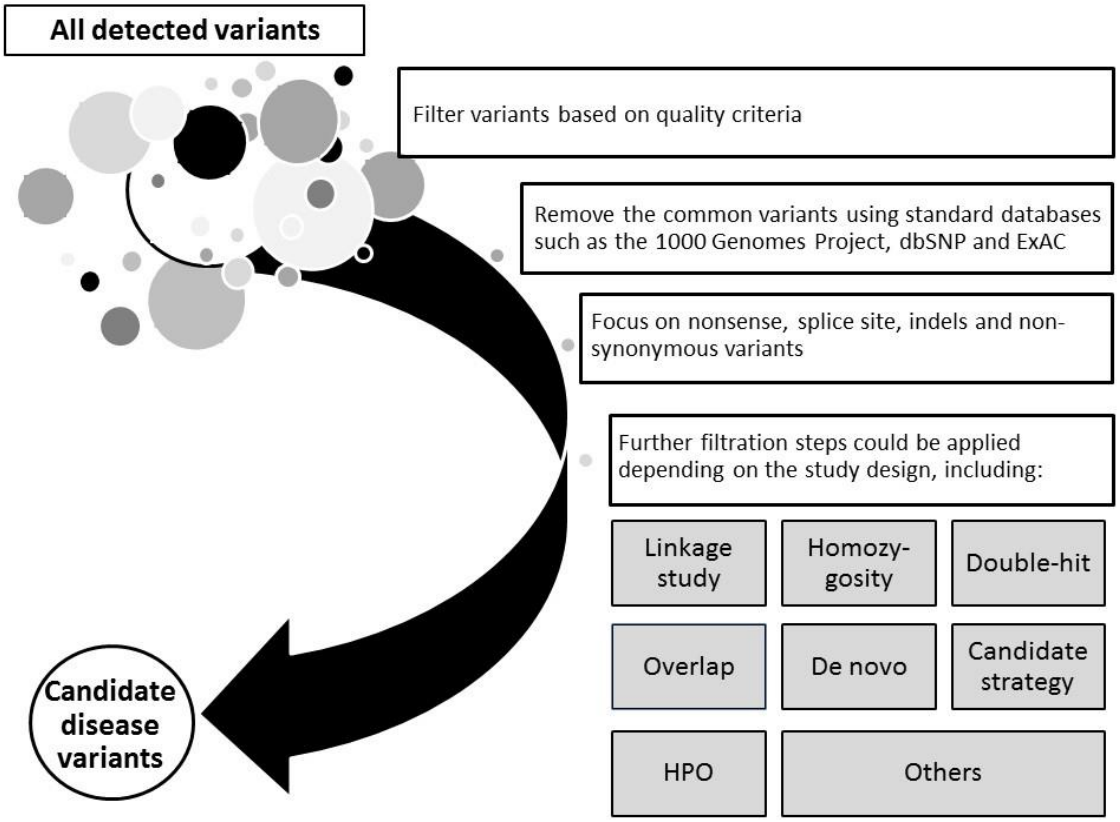


Figure 3.1 Approaches used to identify causative gene defects following whole exome sequencing

Depending on the study design, several approaches can be utilised to prioritise variants to a limited number of single nucleotide variants (SNVs) that are then considered to be candidate causative defects. However, early steps in prioritising variants are almost always shared among different study designs.

3.1.2 The UK Genotyping and Phenotyping of Platelets study

The UK Genotyping and Phenotyping of Platelets study (UK-GAPP) was funded by the British Heart Foundation from 2010 to 2015, and involved researchers from Birmingham, Bristol and Sheffield. The study aimed to undertake extensive genotyping and platelet phenotyping to investigate IPDs in patients recruited through Haemophilia Centres throughout the UK (http://www.birmingham.ac.uk/research/activity/cardiovascular-sciences/research/platelet-group/platelet-gapp/index.aspx). Since the research described in this thesis has arisen from work carried out in Sheffield as part of the UK-GAPP study, a historical overview of the findings of the UK-GAPP study to date is warranted to allow the present study to be placed in context. A brief history of the UK-GAPP study therefore follows.

3.1.2.1 Platelet phenotyping

Although light transmission aggregometry (LTA) is the gold standard method for assessment of platelet function, its use as a tool to identify platelet function defects is influenced by several pre-analytical factors, sampling methods, aspects of the methodology and approaches used to interpret the results. Thus, prior to recruitment of patients with IPDs for assessment of platelet function, considerable work was undertaken to determine the optimal conditions for sampling by performing LTA on samples from control subjects, thus establishing normal ranges for platelet aggregation and secretion against which data from patients could then be compared. Normal platelet aggregation traces in response to a panel of eight different agonists, adenosine diphosphate (ADP), TXA2 mimetic (U46619), adrenaline, arachidonic acid, collagen, collagen-related peptide (CRP), protease-activated receptor (PAR)-1, and PAR-4 peptides, each used at three different concentrations, were established using lumi-aggregometry for approximately 100 healthy volunteers (Dawood et al., 2007). A comprehensive multicentre study involving platelet phenotyping of more than 600 patients clinically diagnosed as having IPDs then commenced. The findings obtained for 111 of these patients were published in 2012 (Dawood et al., 2012). LTA confirmed the presence of a platelet defect in approximately 60% of cases studied, and it was possible to subgroup the cases according to the defects identified. Thus, cases were subgrouped according to whether they had a defect in dense granule secretion, a defect in Gi-receptor signalling, a TXA2 pathway defect, or defects in signalling through the purinergic receptor P2Y12 ADP receptor, the TXA2 receptor, glycoprotein (GP) VI pathway or a Gq signalling defect. It should be noted that the majority of patients had a defect in one of the first three of these categories and that most patients showed a partial impairment, rather than complete abrogation, of the response that was overcome at high agonist concentrations. The use of LTA failed to identify a platelet function defect in the remaining 40% of patients. The possibility that this was due to the limitations of LTA was examined by performing additional phenotypic assays, such as optical multichannel platelet aggregometry or measurement of platelet P-selectin release, but these also failed to demonstrate the presence of a defect (Chan & Warner, 2012; Fox et al., 2009). It is possible that the bleeding symptoms observed in the patients may not have been due to platelet defects, or the assays used to investigate the suspected platelet disorders may not have been sensitive enough to detect a platelet defect. Also, the bleeding symptoms may have been caused by defects in

platelet adhesion or aggregation under flow which would not have been detected with the assays utilised (Watson et al., 2013).

3.1.2.2 DNA analysis in patients recruited to the UK Genotyping and Phenotyping of Platelets study

The decision whether to analyse certain candidate genes in patients recruited to the UK-GAPP study by Sanger sequencing or to undertake WES analysis has been largely driven by the results of platelet phenotyping as described above. When the clinical presentation, family history and results of platelet phenotyping suggested the strong likelihood of a defect in a particular gene or group of genes, Sanger sequencing of the candidate genes was undertaken. This approach led to the identification of a small number of novel defects in the P2Y12 ADP receptor and the TXA2 receptor, which were successfully characterised and shown to recapitulate the platelet phenotype when expressed in a heterologous cell line (Daly et al., 2009; Dawood et al., 2012; Mumford et al., 2010; Mumford et al., 2013; Nisar et al., 2011; Nisar et al., 2014; Patel et al., 2014).

However, in most cases, the results of platelet phenotyping did not suggest a particular candidate gene that could be targeted for Sanger sequencing. The UK-GAPP study group therefore took advantage of rapidly developing NGS technology to undertake simultaneous sequence analyses of a panel of 216 candidate genes that were previously known to be associated with IPDs, human orthologs of genes in which defects were shown to cause platelet defects in animal models, or genes that were relevant to platelet function but not previously associated with IPDs (Jones et al., 2012). They adopted a strategy for mapping and filtering the sequencing output that depended on earlier assumptions that causative SNVs were likely to be rare, that they would occur in the coding region of one of the 216 candidate genes, and that they would be predicted to be pathogenic using various bioinformatic prediction tools. Where possible, they also employed the clinical and laboratory phenotype data of patients to direct the analysis toward certain target genes among the 216 genes sequenced. They analysed DNA from ten patients diagnosed with IPDs in this way but focused their attention on one patient with a clinical diagnosis of Hermansky-Pudlak syndrome (HPS) whose platelets displayed an absence of ATP secretion in response to all agonists. The focus was thus directed to a shorter list of 57 candidate genes

which encoded proteins involved in dense granule formation and secretion. By removing common variants, synonymous variants that were predicted to be non-functional, and heterozygous SNVs, because the mode of inheritance of HPS is recessive, a novel SNV in *HPS4* was highlighted as the causative defect. However, it was not possible to identify the causative SNVs in any of the remaining nine patients who were studied at the same time. This study highlighted the difficulty of identifying causative variants in IPD patients due to the presence of many candidate gene defects even when the focus was on a limited number of genes. When using a sequencing method or approach that yields a large number of variants, as is the case with WES or whole genome sequencing, the identification of causative defects is even more challenging. It is further complicated by the fact that apparently healthy individuals carry hundreds of mildly disadvantageous SNVs without showing any apparent ill effects (Xue et al., 2012).

In 2013, in a follow-on study to the one described above, Stockley et al. (2013) undertook NGS analysis of a panel of 260 platelet genes in a group of 13 patients with IPDs that were characterised by a significant reduction in dense granule secretion with no other features of HPS. Their findings revealed an enrichment of defects in the transcription factor genes *FLI1* and *RUNX1* (Runt-related transcription factor 1) in half of the index cases in this subgroup. Affected members of these families presented with bleeding disorders which were characterised by defects in platelet dense granule secretion, and in most cases, the presence of a *FLI1* or *RUNX1* defect was also associated with mild thrombocytopenia. RUNX1 and FLI1 are transcription factors that have a role in megakaryopoiesis and are known to cooperate in regulating the last stages of platelet production (Huang et al., 2009; Okada et al., 2013; Tijssen et al., 2011; Zang et al., 2016). It is known that 94% of patients with Paris-Trousseau syndrome (PTS), who have a deletion of the long arm of chromosome 11 that includes *FLI1*, have a bleeding tendency that is characterised by thrombocytopenia and platelet dysfunction (Grossfeld et al., 2004). The report by Stockley et al. (2013) was the first to describe the association of platelet secretion defects with abnormalities in *FLI1*. Further characterisation of two of the non-synonymous *FLI1* alterations identified in this study revealed a significant reduction in their transactivational capacity (Stockley et al., 2013).

The difficulty of identifying the underlying causative genes in patients with IPDs was again highlighted in the study by Leo et al. (2015), in which WES analysis was undertaken in two groups of patients, 12 with defects in Gi-signalling and 6 with defects in secretion, before downstream bioinformatic studies focused on an extended panel of 329 platelet genes (Leo et al., 2015). Assuming that unrelated patients with similar phenotypes will have defects in similar or closely related genes or pathways, downstream analysis of the genetic data was conducted separately for the two patient subgroups. Using a combination of bioinformatic prediction tools and functional annotation analysis, 13 genes which harboured deleterious defects were highlighted in the subgroup of patients with Gi-signalling defects, which could potentially explain the condition in 75% of the patients in this group. Similarly, 14 genes harbouring deleterious defects were identified in the subgroup of patients with platelet secretion defects, with at least one of these occurring in each of the patients in this subgroup. Although some defects were found in genes that were previously associated with IPDs, such as *P2RY12* (Purinergic receptor P2Y12)*, LYST* (Lysosomal trafficking regulator) and *STXBP2* (Syntaxin-binding protein 2), most of the variants were in genes that had not previously been associated with IPDs.

A drawback of the studies described above was the possible exclusion of causative defects due to restricting the analysis to certain genes. This was avoided by more recent studies which used WES analysis to elucidate the genetic basis of IPDs. Thus, SNVs in three consecutive codons of the *SLFN14* (Schlafen family member 14) were identified in three unrelated index cases affected by a bleeding disorder which was characterised by thrombocytopenia and a defect in platelet secretion (Fletcher et al., 2015). SFLN14 was later shown to be a ribosome-associated protein involved in ribosomal RNA and mRNA degradation in rabbit reticulocytes (Pisareva et al., 2015). Subsequently, the expression of all variant forms of SFLN14 was shown to be reduced due to posttranslational degradation and misfolding when compared with the wild-type protein (Fletcher et al., 2018). The findings suggested a possible role for SFLN14 in the degradation of RNA during thrombopoiesis. WES also facilitates rapid detection of causative defects in patients with inherited thrombocytopenia. Thus, WES analysis of DNA from 37 patients recruited to the UK-GAPP study with inherited bleeding diatheses characterised by thrombocytopenia of unknown aetiology identified pathogenic variants in 14 genes previously known to associate with inherited

thrombocytopenia in 46% of cases (Johnson et al., 2016). More recently, a combined approach of WES analysis, followed by *in silico* prediction of the likelihood of pathogenicity of novel homozygous variants that mapped to a tightly-linked homozygous region, and review of RNA sequence data for haematopoietic progenitors suggested a *GNE* variant to be the causative defect in a consanguineous family with severe congenital thrombocytopenia (Futterer et al., 2018).

3.2 Hypothesis and aims of this study

Many factors complicate diagnosis of IPDs, including issues relating to the laboratory investigations conducted, such as the lack of standardisation of methods for analysing platelet function between laboratories, the necessity to have the expertise to interpret LTA results and the identification of several plausible candidate gene defects in most affected individuals using WES analysis. Various factors related to the nature of the disease can also complicate the diagnosis, including differences in the severity of the bleeding phenotype among patients, which are dependent on the nature of any vascular challenges faced by patients. The extensive redundancy of the receptors and the complexity of the signalling pathways in platelets pose a further challenge. It is also recognised that different SNVs in one or more genes can result in considerable variability in clinical presentation among patients (Albers et al., 2012; Daly et al., 2009; Pecci et al., 2008). Allelic heterogeneity, in which similar IPDs can be caused by defects in different genes, is also a complicating factor. It is therefore likely that the contribution of multiple genetic loci affecting different aspects of platelet biology and/or other aspects of the haemostatic system may explain the patient phenotype in some cases, supporting the rationale for a multifactorial aetiology in IPDs.

The enhanced power of genetic analysis provided by NGS technologies, such as WES, has helped to overcome some of the challenges described above, facilitating the identification of many novel genes associated with IPDs. It is predicted that WES analysis guided by platelet phenotyping, in addition to standard initial clinical evaluation, will aid the identification of the underlying causes for at least some patients with unexplained bleeding, which allow assessment of bleeding risk in affected patients and facilitate genetic counselling. Also, the outcome of these analyses will expand our

understanding of the relationship between platelet phenotype and genotype, yielding further insights into the molecular mechanisms of platelet physiology, which in turn will further aid the diagnosis of IPDs.

The study summarised in this chapter involved 34 index cases who were recruited to the UK-GAPP study for investigation of an IPD. After extensive platelet phenotyping, cases were subgrouped according to whether their platelets displayed a defect in a Gi-signalling pathway, or in secretion. Based on the overarching hypothesis of the study that IPDs are due to monogenic or oligogenic inheritance of alterations in genes encoding proteins that are essential for platelet formation and/or function, our aim was to identify potential disease-causing genetic variants via WES analysis. Given the interests of the group in the genes regulating the biogenesis and secretion of platelet granules, further downstream studies focused on those patients with secretion defects.

3.3 Materials and methods

3.3.1 Patients

Thirty-four index cases who were recruited to the UK-GAPP study because they were suspected to have an IPD were investigated. Platelet phenotyping, which was undertaken at the time of enrolment, separated them into two broad, but distinct, subgroups. Thus, 12 cases had platelet profiles that were consistent with the presence of a Gi-signalling defect, as they displayed transient aggregation in response to ADP and a reduction or absence of aggregation in response to adrenaline, and an absence of ATP secretion in response to both agonists (Dawood et al., 2012). Platelets from the remaining 22 cases displayed profiles that were consistent with an abnormality in secretion, showing reduced ATP secretion in response to all platelet agonists (Dawood et al., 2012). DNA samples were available from the 34 index cases and from an additional affected family member for six of the index cases. Index cases were recruited based on clinical and laboratory criteria which included (i) a history of excessive bleeding symptoms, which either presented as spontaneous mucocutaneous bleeding or prolonged bleeding following injury; (ii) the absence of any known causes of acquired platelet dysfunction; (iii) exclusion of Glanzmann thrombasthenia, Bernard-Soulier syndrome and HPS; and (iv) exclusion of a

coagulation defect. It should be noted that 22 of the 34 index cases had been the subject of a previous investigation that focused on 329 platelet genes (Leo et al., 2015).

The UK-GAPP study was approved by the National Research Ethics Service Committee West Midlands–Edgbaston (REC reference: 06/MRE07/36). Written informed consent was given by all participants in accordance with the Declaration of Helsinki before providing venous blood samples for analysis, which were collected using 3.2% trisodium citrate tubes (S-Monovette® 0.106 mol L^{-L}; Sarstedt, Leicester, UK).

3.3.2 Genetic analysis: identification of candidate single nucleotide variants

WES was carried out in collaboration with Dr Michael Simpson, Kings College, London. Sequence data for each patient was provided as a large Variant Call Format (VCF) file and examined to identify SNVs of potential clinical relevance using a pipeline that was modified slightly from that used in a previous study (Leo et al., 2015) (Figure 3.2).

SNVs were initially filtered according to their frequency in the general population against three different databases, the 1000 Genomes project (http://www.1000genomes.org/), the NHLBI exome sequencing project (http://evs.gs.washington.edu/EVS/) and an in-house database (Dr M Simpson, Kings College London) consisting of more than 900 exomes [accessed 2015]. Variants previously identified as having a frequency of over 1% were excluded from any further investigations. Furthermore, only SNVs that altered the amino acid sequence of the encoded proteins were retained, while synonymous SNVs were excluded. The potential consequences of the remaining variants were investigated using an online prediction tool, CADD [Combined Annotation Dependent Depletion, version 1.3 (http://cadd.gs.washington.edu/score) (Kircher et al., 2014)]. CADD combines conservation-based metrics with functional metrics using 65 different annotations to generate a single value for each SNV. Based on the CADD inventor recommendation, a cut-off ranked C-score (PHRED) of 20 was used. Long indels that could not be analysed using CADD were considered to be pathogenic and taken forward within the pipeline. Following the removal of variants that were predicted to be benign according to CADD, the remaining SNVs were investigated for gene expression in platelets using PaxDb.4 [(https://pax-db.org/) (Wang et al., 2015)], which provided information based

on two studies (Kim et al., 2014b; Martens et al., 2005), and data from a third study describing the platelet transcriptome (Londin et al., 2014). SNVs in genes which were found to be expressed in platelets according to one or more of these studies were retained. Where affected family members were available, shared SNVs were only considered for further analysis as all included family members presented with similar phenotypes to their corresponding index cases.

WES data from the 34 index cases were analysed similarly for the majority of the pipeline. However, the 34 index cases comprised two subgroups of 12 cases who were diagnosed with Gi-signalling abnormalities as their platelets showed a defect in aggregation and secretion in response to agonists for the Gi-coupled receptors for ADP and adrenaline and 22 cases were diagnosed with dense granule secretion defects as their platelets demonstrated a decrease in ATP secretion to all platelet agonists. Assuming that defects in a single gene would not result in two different disease phenotypes, SNVs that were present in genes shared by the two subgroups were removed from the analysis in the final step of the pipeline.

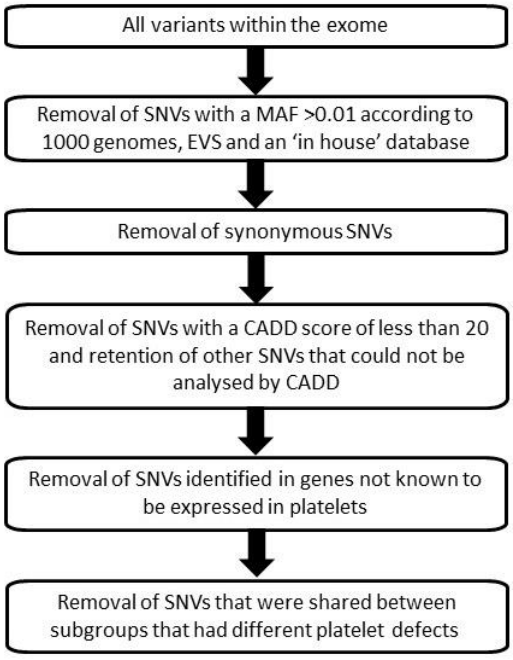


Figure 3.2 Pipeline for analysis of whole exome sequence data
Sequential steps to filter single nucleotide variants (SNVs) were as follows: SNVs having a minor allele frequency (MAF) in the general population of >0.01 and synonymous SNVs were removed; the pathogenicity of remaining SNVs was predicted using Combined Annotation Dependent Depletion (CADD) tool, and those SNVs having scores of 20 or greater or that could not be analysed by CADD were retained. Further prioritisation was achieved by retaining SNVs in genes known to be expressed in platelets and shared among affected family members (when available). Finally, SNVs located in genes that were shared between subgroups that had different platelet defects were removed. EVS; Exome sequence variant (NHLBI exome sequencing project).

3.4 Results

3.4.1 Characteristics of index cases

A total of 34 index cases (F1.1 to F34.1) were selected for investigation. All participants had been evaluated clinically and enrolled in the UK-GAPP study following diagnosis of a suspected platelet function disorder. Extensive platelet phenotyping was undertaken on samples from all subjects at the time of recruitment to the study to examine platelet aggregation and granule secretion in response to a range of agonists using lumi-aggregometry (Dawood et al., 2012; Dawood et al., 2007). The index cases selected for study belonged to one of two subgroups. Thus, 22 cases (F1.1 to F22.1) had been diagnosed with defects in dense granule secretion based on a reduction in ATP secretion from platelets in response to PAR-1 specific peptide (SFLLRN; 100 µmol/L) or thrombin (1U/ml) compared to platelets from healthy control subjects. The remaining 12 cases (F23.1 to F34.1) were classified as having Gi-signalling abnormalities as their platelets displayed reversible aggregation in response to 10 µmol/L ADP and absence of a secondary wave of aggregation in response to adrenaline (10 µmol/L). Six first-degree affected relatives were also included in the study (F1.2, F18.2, F25.2, F29.2, F30.2 and F34.2).

3.4.2 Exome sequencing and prediction of candidate gene defects

WES was performed on DNA from the index cases and affected family members, and analysis of the resulting genetic data was undertaken to identify candidate gene defects that could explain the bleeding tendency. Alignment of sequence data with the human genome resulted in the identification of approximately 25,000 sequence variants in the exome from each index case. Removal of synonymous variants and those variants identified by the 1000 Genomes project, the NHLBI exome sequencing project or an in-house database as having a minor allele frequency (MAF) in the general population of greater than 0.01 reduced the number of variants to a median of 379 SNVs in each index case. The remaining SNVs were then analysed using CADD, a tool that scores the deleteriousness of SNVs in the human genome. Following removal of those variants having a CADD_PHRED score of less than 20, a median of 174 (range 76–296) candidate SNVs remained per index case. Further prioritisation was achieved by limiting the analysis to SNVs located in genes expressed in platelets. This prioritisation was achieved by comparison with the data from three studies, two of

the platelet proteome, and one of the platelet transcriptome, and resulted in a reduction in the median number of SNVs to 103 per index case (range 46–175). Where data were available for an affected family member, focussing only on shared variants further reduced the number of SNVs by approximately 50% (Table 3.1).

Table 3.1 Median number of single nucleotide variants in each exome at different stages of the analysis pipeline

	Median (range)
SNVs identified in exome sequence	24,774 (11,768–27,715)
Non-synonymous SNVs* with MAF ≤ 0.01	379 (209–711)
SNVs with CADD_PHRED score ≥ 20	174 (76–296)
SNVs in genes expressed in platelets	103 (46–175)
SNVs shared between affected family members	98 (31-175)

CADD; Combined Annotation Dependent Depletion, MAF; minor allele frequency, SNV; single nucleotide variant.

In the subgroup of 22 index cases with platelet secretion defects, a total of 2,066 variants were identified in 1,476 candidate genes, while among the 12 index cases with Gi-signalling defects, a total of 1,059 variants were detected in 891 genes (Supplementary Data #1 & #2). Assuming that different platelet phenotypes will be due to causative variants in different gene subsets, those SNVs that occurred in genes common to both subgroups of index cases were filtered out from the data. The number of genes in which candidate SNVs were identified in both subgroups was 346. The removal of SNVs that occurred in these genes reduced the number of candidate genes by approximately 30%, with 1,130 and 545 unique genes remaining in the subgroups of index cases having secretion defects and Gi-signalling defects, respectively (Table 3.2). Table 3.3 shows the number of SNVs in each patient at different stages of the analysis pipeline.

Table 3.2 Numbers of single nucleotide variants and genes in which they occur in each subgroup of index cases

	Secretion defect	**Gi-signalling defect**
Total number of candidate SNVs/genes	2,066/1,476	1,059/891
Total number of unshared SNVs/genes	1,465/1,130	599/545

SNV; single nucleotide variant.

Table 3.3 Number of single nucleotide variants in each index case at different stages of the analysis pipeline

Index case	Platelet phenotype	SNVs identified in exome sequence	Non-synonymous SNVs* with MAF ≤ 0.01	SNVs with CADD_PHRED score ≥ 20	SNVs in genes expressed in platelets	SNVs shared between affected family members	SNVs in genes unique to platelet phenotype
F1.1	Sec	24,740	368	158	96	48	38
F1.2[#]	Sec	26,483	598	254	149		
F2.1	Sec	27,715	711	296	175	---	122
F3.1	Sec	24,658	364	162	90	---	66
F4.1	Sec	24,709	389	187	109	---	72
F5.1	Sec	24,874	366	169	100	---	69
F6.1	Sec	24,657	423	200	116	---	79
F7.1	Sec	25,026	378	172	99	---	71
F8.1	Sec	24,510	393	175	99	---	72
F9.1	Sec	24,940	370	180	114	---	86
F10.1	Sec	25,206	500	221	121	---	74
F11.1	Sec	25,883	424	212	127	---	83
F12.1	Sec	24,700	359	153	86	---	64
F13.1	Sec	12,535	240	99	47	---	31
F14.1	Sec	17,860	304	121	63	---	47
F15.1	Sec	11,768	209	76	46	---	35
F16.1	Sec	25,090	379	81	96	---	72
F17.1	Sec	26,260	454	197	111	---	81
F18.1	Sec	17,350	309	118	72	31	24
F18.2[#]	Sec	15,001	247	99	54		
F19.1	Sec	25,860	455	223	123	---	85
F20.1	Sec	17,462	523	208	102	---	72
F21.1	Sec	15,982	401	157	97	---	77
F22.1	Sec	16,101	277	119	66	---	45
F23.1	Gi	24,439	400	194	107	---	62

F24.1	Gi	25,121	367	171	106	---	71
F25.1	Gi	24,967	467	223	127	72	41
F25.2#	Gi	25,074	493	249	139		
F26.1	Gi	25,007	398	198	100	---	55
F27.1	Gi	24,910	336	147	82	---	48
F28.1	Gi	25,729	543	240	129	---	75
F29.1	Gi	24,704	378	190	113	63	31
F29.2#	Gi	24,632	361	186	112		
F30.1	Gi	25,092	356	182	103	57	30
F30.2#	Gi	25,023	379	163	108		
F31.1	Gi	24,631	332	150	96	---	57
F32.1	Gi	24,713	326	156	103	---	63
F33.1	Gi	25,092	419	175	107	49	29
F33.2#	Gi	24,807	406	167	89		
F34.1	Gi	24,244	363	165	95	---	38
Median		24,774	379	174	103	98	65
Minimum		11,768	209	76	46	31	24
Maximum		27,715	711	296	175	175	122

* After removal of synonymous and intronic variants. # affected relative of index case. CADD; Combined Annotation Dependent Depletion, Gi; Gi-signalling defect, MAF; minor allele frequency, Sec; secretion defect, SNV; single nucleotide variant.

3.4.3 Variants in platelet disorder genes identified among index cases with platelet secretion defects

Given the focus of this study on genes regulating the biogenesis and secretion of platelet granules, further downstream studies focused on those patients with secretion defects, as well as the 1,130 genes which harboured SNVs in this subgroup of index cases.

While the majority of variants identified were present in genes that have not previously been associated with IPDs, several variants occurred in genes that have previously been implicated in IPDs including *ABCG8* (ATP binding cassette subfamily G member 8), *BLOC1S3* (Biogenesis of lysosomal organelles complex 1 subunit 3)*, ETV6*, *FERMT3 (*Fermitin family member 3), *FLNA* (Filamin A), *FLI1*, *GFI1B*, *ITGB3* (Integrin subunit beta 3), *P2RX1* (Purinergic receptor P2X 1), and *RUNX1* (Table 3.4). Thus, candidate defects were identified in at least one of the latter genes in 54% (12 out of 22) of the index cases with platelet secretion defects, with four index cases harbouring defects in more than one of these genes. Four of these genes, *ETV6, FLI1*, *GFI1B* and *RUNX1,* which encode transcription factors that are known to have a role in the regulation of megakaryopoiesis, were shown to harbour candidate defects in six of the index cases with platelet secretion disorders. Querying these SNVs against “The Exome Aggregation Consortium (ExAC)” database (Lek et al., 2016) revealed that, with the exception of a c.322C>G:p.L108V SNV in *BLOC1S3* which had a frequency of 1.16%, all of the SNVs occurred at frequencies of less than 1% in the general population (Exome Aggregation Consortium [ExAC], Cambridge, MA URL: http://exac.broadinstitute.org [accessed October 2018]).

Table 3.4 Variants present in genes previously associated with inherited platelet bleeding disorders that were identified among index cases with platelet secretion defects

Patient	Gene	CADD_ PHRED score	Type of SNV	Alteration	rs number*	ExAC**	Location/Domain of the change in the protein***
F1.1+F1.2	*FLI1*	-	FD	NM_002017:c.992_995del:p.331_332del	---	---	DNA-binding domain (Ets domain)
F4.1	*ETV6*	45	SG	NM_001987:c.1288C>T:p.R430X	---	---	Winged helix-turn-helix DNA-binding domain
F6.1	*RUNX1*	37	SG	NM_001001890:c.236G>A:p.W79X	---	---	DNA-binding domain (Runt domain)
F6.1	*ITGB3*	33	NS	NM_000212:c.349C>T:p.R117W	---	8.28e-06	Integrin beta subunit, N-terminal
F7.1	*RUNX1*	27.1	SP	NM_001001890:c.270+1G>T	---	---	---
F7.1	*BLOC1S3*	23.9	NS	NM_212550:c.322C>G:p.L108V	rs75792246	0.01164	---
F10.1	*FERMT3*	24.4	NS	NM_178443:c.293G>A:p.R98Q	rs140992702	1.653e-05	---
F11.1	*FLI1*	26.9	NS	NM_002017:c.1018C>T:p.R340C	---	---	DNA-binding domain (Ets domain)
F11.1	*ABCG8*	24.9	NS	NM_022437:c.1540C>T:p.P514S	---	---	Non-cytoplasmic domain
F12.1	*BLOC1S3*	23.9	NS	NM_212550:c.322C>G:p.L108V	rs75792246	0.01164	---
F13.1	*FLNA*	27	NS	NM_001456:c.806T>A:p.L269Q	---	2.313e-05	Between calponin homology and immunoglobulin-like fold
F16.1	*BLOC1S3*	23.9	NS	NM_212550:c.322C>G:p.L108V	rs75792246	0.01164	---
F19.1	*BLOC1S3*	23.9	NS	NM_212550:c.322C>G:p.L108V	rs75792246	0.01164	---
F21.1	*GFI1B*	20.5	NS	NM_001135031:c.289G>A:p.D97N	rs145562579	0.005939	---
F21.1	*BLOC1S3*	24.9	NS	NM_212550:c.499C>T:p.L167F	rs572296006	0.0005295	---
F22.1	*P2RX1*	24.8	NS	NM_002558:c.1111G>A:p.A371T	---	8.24e-06	Cytoplasmic domain

*rs number from the dbSNP database [accessed 2018]. **ExAC Browser (Beta) - version 0.3.1 from Exome Aggregation Consortium (http://exac.broadinstitute.org/) [accessed 2018]. ***Location/Domain of amino acid substitution predicted using InterPro. CADD; Combined Annotation Dependent Depletion, FD; frameshift deletion, NS; non-synonymous, SG; stop-gain, SNV; single nucleotide variant, SP; splicing.

3.5 Discussion

In this study, WES was used to investigate the underlying genetic defects in 34 index cases with IPDs who were recruited to the UK-GAPP study. Based on extensive platelet phenotyping at the time of study enrolment, 22 index cases were diagnosed as having defects in platelet secretion, while the remaining 12 had defects in Gi-signalling pathways. The pipeline used to prioritise candidate defects from the WES data initially filtered variants according to their frequency in the population and variant type. Those with a MAF of 0.01 or greater were removed, as were synonymous and intronic (non-splice site) variants. The potential effects of gene variants were then predicted using CADD, a tool that combines the information from many functional annotation tools to derive single scores as measures of the deleteriousness of sequence variants. Variants achieving CADD_PHRED scores of 20 or greater were predicted to be pathogenic. Removal of variants having CADD_PHRED scores of less than 20, and occurring in genes that are not expressed in platelets, reduced the number of candidate variants to 103 in each index case. Where WES analysis was also undertaken on an affected family member, removal of variants that were not shared between the affected family members brought about an approximate 50% further reduction in the number of candidate gene defects in each corresponding index case. Assuming that causative variants would occur within genes that function in the same or related pathways for each subgroup of index cases, those genes that were represented among the SNVs identified in both subgroups were excluded from further analysis, thereby reducing the number of candidate SNVs by approximately 30%. A median of 70 candidate SNVs per index case, and a total of 1,130 possible candidate gene defects were identified across the 22 patients with secretion defects. Similarly, a median of 50 candidate SNVs per index case, and a total of 545 possible candidate gene defects were identified among the 12 patients with defects in Gi-signalling pathways, highlighting the heterogeneity and complexity of IPDs.

Given the focus of this study on genes regulating the biogenesis and secretion of platelet granules, further downstream studies focused on those findings from the patients with defects in platelet secretion.

Heterozygous defects were identified in ten genes previously associated with IPDs, *ABCG8, BLOC1S3*, *ETV6, FERMT3, FLI1*, *FLNA, GFI1B, ITGB3, P2RX1* and *RUNX1*, and defects in at least one of these genes were identified in about half of the index cases. However, it is important to note that several of the IPDs associated with these genes are recessively inherited and that the heterozygous presence of a variant in one of these genes alone is unlikely to fully explain the bleeding symptoms observed. Additionally, in four index cases, more than one candidate variant was identified suggesting that more than one gene may contribute to the pathogenicity of the IPDs. The associations of the ten previously mentioned genes with IPDs are discussed below.

Interestingly, defects in only four of these genes have been previously associated with abnormalities in dense granule secretion, these being *BLOC1S3, ETV6, FLI1, GFI1B* and *RUNX1* (Ferreira et al., 2017; Mao et al., 2017; Marneth et al., 2017; Morgan et al., 2006; Poggi et al., 2017; Saultier et al., 2017; Stockley et al., 2013). In the *ETV6, FLI1, FLNA, GFI1B, ITGB3* and *RUNX1*-related IPDs, α-granule abnormalities have been documented (Aneja et al., 2011; Berrou et al., 2017; Favier et al., 2018; Ferreira et al., 2017; Glembotsky et al., 2014; Mao et al., 2017; Marneth et al., 2017; Noetzli et al., 2015; Stevenson et al., 2013; Stevenson et al., 2015). Although dense and α-granules differ in their cargo, they are understood to have similar granule release machinery (Heijnen & Van der Sluijs, 2015) which suggests that the defects identified in some of these genes could, at least partly, explain the observed phenotypes in the affected cases.

Apart from *BLOC1S3, FLNA* and *ITGB3,* the genes mentioned above which have been associated with platelet granule abnormalities all encode transcription factors that cooperate during megakaryopoiesis to determine several aspects of platelet biology. Current evidence suggests that ETV6 acts as a transcriptional repressor and has a significant role in early haematopoiesis, affecting the development of multiple lineages (Rasighaemi et al., 2015). More recently, several groups have reported its association with familial thrombocytopenia and a predisposition to haematological malignancy. Defects in platelet aggregation, accompanied by platelet hypogranularity, elongated α-granules, and abnormal dense granules have also been described (Melazzini et al., 2016; Moriyama et al., 2015; Noetzli et al., 2015; Poggi et al., 2017; Topka et al., 2015;

Zhang et al., 2015). The results of further work which was conducted to characterise the nonsense c.1288C>T transition in *ETV6* and its potential contribution to the bleeding tendency in index case F4.1 will be described in chapter 5 of this thesis. FLI1 is a key regulator of the differentiation of megakaryocytic/erythroid progenitors into platelets (Bastian et al., 1999; Pang et al., 2006; Vo et al., 2017) through its regulation of several megakaryocytic genes (Bastian et al., 1999; Deveaux et al., 1996; Eisbacher et al., 2003; Gosiengfiao et al., 2007; Hart et al., 2000; Lemarchandel et al., 1993; Moussa et al., 2010; Schwachtgen et al., 1997; Zang et al., 2016; Zhang et al., 1993). In recent years, germline defects in *FLI1* have been identified in patients with IPDs that were characterised by abnormalities in platelet granules and variable thrombocytopenia (Poggi et al., 2015; Saultier et al., 2017; Stevenson et al., 2015; Stockley et al., 2013). Further work was undertaken to characterise the c.1018C>T transition in *FLI1* which was identified in index case F11.1 and the findings are summarised in chapter 4 of this thesis. GFI1B plays a role in the development of both megakaryocytic and erythroid lineages by acting as a transcriptional repressor (Saleque et al., 2002). A number of studies have reported defects in *GFI1B* in association with “Bleeding Disorder, Platelet-Type, 17” which presents as a GPS-like phenotype in addition to red cell anisopoikilocytosis (Ferreira et al., 2017; Marneth et al., 2017; Schulze et al., 2016; 2017; Stevenson et al., 2013). In addition to macrothrombocytopenia and a reduced number of platelet α-granules, the platelets from affected patients were reported to have a reduction in dense granules. However, there was variation in the extent of the associated erythropoiesis defect as well as in the mode of inheritance (Ferreira et al., 2017; Schulze et al., 2016; 2017), suggesting that the location of variants may be important in determining the clinical phenotype. The identification of two GFI1B isoforms that preferentially promote either megakaryocytopoiesis or erythropoiesis would support this hypothesis (McClellan et al., 2017; Polfus et al., 2016; Schulze et al., 2016). RUNX1 is a core regulator of haematopoiesis (North et al., 2002), regulating the expression of many genes involved in megakaryocyte differentiation (Ichikawa et al., 2004), including genes involved in platelet granule biogenesis e.g. *NFE2* (Nuclear Factor, Erythroid 2) (Glembotsky et al., 2014) and *PLDN* (Pallidin) (Mao et al., 2017), granule content e.g. *PF4* (Platelet Factor 4) (Aneja et al., 2011), and signalling and trafficking e.g. *ALOX12* (Arachidonate 12-Lipoxygenase, 12S Type) (Kaur et al., 2010), *PRKCQ* (Protein Kinase C Theta) (Jalagadugula et al., 2011), *RAB27B* and *RAB1B* members of RAS oncogene family

(Glembotsky et al., 2014; Jalagadugula et al., 2018), and *PCTP* (Phosphatidylcholine transfer protein) (Songdej et al., 2016). Interestingly, defects in *RAB27A* and *PLDN* have been shown to cause Griscelli syndrome and HPS respectively, which are both characterised by abnormalities in platelet dense granules (Cullinane et al., 2011; Ménasché et al., 2000).

Defects in *BLOC1S3, FLNA* and *ITGB3* have also been associated with platelet granule abnormalities. *BLOC1S3* encodes the BLOC-1 subunit that is involved in the biogenesis of lysosome-related organelles. Homozygous defects in *BLOC1S3* are known to cause HPS type 8, of which dense granule deficiency is a characteristic feature, as well as oculocutaneous albinism (hypopigmentation) and impaired visual acuity (Morgan et al., 2006). Interestingly, five of the index cases studied were heterozygous for one of two *BLOC1S3* SNVs identified in this study (F7.1, F12.1, F16.1 F19.1 and F21.1). Two of these cases (F7.1 and F21.1) were also heterozygous for defects in *RUNX1* and *GFI1B*. Defects in *FLNA* encoding filamin A, an actin-binding protein, have been associated with macrothrombocytopenia as a result of disturbances in the platelet cytoskeleton that affect proplatelet formation and platelet release (Nurden et al., 2011). The presence of enlarged α-granules in the platelets has also been reported (Nurden et al., 2011). Given the X-linked dominant mode of inheritance of *FLNA* defects, it is likely that the c.806T>A *FLNA* alteration predicting a p.L269Q substitution in filamin A is the causative genetic defect in the female index case F13.1. *ITGB3* encodes β3 integrin, which assembles with αIIb integrin subunit to form the αIIbβ3 fibrinogen receptor in platelets. Homozygous or compound heterozygous defects in *ITGB3* result in the autosomal recessive disorder Glanzmann Thrombasthenia that is characterised by qualitative and/or quantitative deficiency of the αIIbβ3 receptor, which in turn alters the ability of platelets to aggregate. A small number of heterozygous alterations have also been shown to alter proplatelet formation and platelet release, resulting in GT-like thrombocytopenia with giant platelets and abnormal giant α-granules (Favier et al., 2018; Ghevaert et al., 2008). Interestingly, the heterozygous *ITGB3* SNV (c.349C>T:p.R117W) is co-inherited with a nonsense *RUNX1* variant in index case F6.1. It is, therefore, possible that it is the combination of the two variants that results in the bleeding symptoms observed in the index case, though further work would be required to investigate this.

The remaining SNVs that were identified in genes previously associated with IPDs among those index cases with platelet secretion defects occurred in *ABCG8, FERMT3*, and *P2RX1*. However, none of these genes has previously been reported to be associated with platelet granule abnormalities. *ABCG8* encodes sterolin-2, a transporter protein for dietary sterols (Berge et al., 2000). Recessively inherited *ABCG8* alterations result in a condition known as thrombocytopenia associated with sitosterolemia or Mediterranean macrothrombocytopenia (Rees et al., 2005; Su et al., 2006), a metabolic disorder which also affects the early phase of megakaryopoiesis by altering the bone marrow microenvironment. Although electron microscopy did not reveal any structural abnormalities in the platelet granules in affected individuals, aggregometry revealed a consistent defect in aggregation to ristocetin (Rees et al., 2005; Su et al., 2006). Given the recessive pattern of inheritance of this disorder, the novel *ABCG8* variant (c.1540C>T:p.P514S) identified in index case F11.1 is unlikely to explain their bleeding symptoms. Furthermore, the *ABCG8* variant was co-inherited with a non-synonymous SNV in FLI1, the characterisation of which will be described in the following chapter. Homozygous defects in *FERMT3* encoding Kindlin-3, a mediator of integrin activation (Moser et al., 2008), are associated with macrothrombocytopenia along with disturbed 'inside-out' integrin activation in leukocytes and platelets which results in leukocyte adhesion deficiency type III (Kuijpers et al., 2009). Given the recessive mode of inheritance, the *FERMT3* variant (c.293G>A:p.R98Q) identified in index case F10.1, is unlikely to independently explain their bleeding symptoms. *P2RX1* encodes P2X1, the only ATP activated receptor in platelets. The release of ATP at the injury site and activation of the receptor causes a rapid influx of calcium that mediates rapid and reversible shape change, transient granule centralization and transient aggregation mediated by activation of αIIbβ3 integrin (Mahaut-Smith et al., 2011). The only heterozygous *P2RX1* alteration that has been described was associated with symptoms of excessive bleeding in the presence of a normal platelet count, as well as normal platelet size and morphology (Oury et al., 2000). However, selective impairment of ADP-induced platelet aggregation was observed (Oury et al., 2000). The absence of further cases associating *P2RX1* defects with IPDs emphasises the importance of undertaking further work to characterise the *P2RX1* variant identified in index case F22.1 (c.1111G>A:p.A371T) to assess its possible contribution to their bleeding symptoms.

WES analysis has facilitated the identification of candidate gene defects in this study, some of which could potentially explain, or contribute to, the dense granule secretion abnormalities and bleeding symptoms observed in the index cases. However, the use of this approach to identify the underlying genetic causes of unexplained bleeding symptoms does have limitations. Some variants may have been overlooked by focussing only on the exome, as intronic variants could contribute to disease. Also, by excluding those variants having MAFs greater than 1%, genetic variants that could potentially contribute to disease may have been overlooked. Furthermore, while non-synonymous SNVs are more likely to be causative than synonymous SNVs, the latter can also be deleterious, mainly through the alteration of splice site positions but also by altering codon preferences and affecting the rate of protein translation (Plotkin & Kudla, 2011); these have been excluded from the analysis in this study. In the event that no causative variants are identified for an index case, the pipeline used to analyse the genetic data could be altered to include intronic and synonymous SNVs or to relax the stringency of the CADD predictions of pathogenicity. These limitations are not unique to this study but common across similar studies using NGS technology to identify the causative defects for complex and heterogeneous disorders (Gilissen et al., 2012).

Despite the previously mentioned shortcomings, the number of candidate SNVs that could potentially contribute to the IPDs among the cases studied was reduced from ~25,000 to ~100 for each index case. Focussing on the subgroup of 22 cases with platelet secretion abnormalities identified probable candidate gene defects affecting *ETV6* and *FLI1* in three index cases, and further studies were undertaken to characterise two of these (see chapters 4 and 5). As discussed above, the defects identified in *FLNA, GFI1B, ITGB3, P2RX1* and *RUNX1* could potentially contribute to the bleeding tendency in another five index cases, though, further investigation is required to evaluate the pathogenicity and association of these variants with bleeding.

The identification of genetic defects that could potentially contribute to the bleeding tendency in approximately 35% of the index cases was achieved mainly as a result of the previous reported association of these genes with IPDs. This highlights the need to adopt alternative approaches to identify which, if any, of the remaining genes that were found to harbour candidate gene defects could be associated with IPDs in the

other index cases. Given that FLI1 is a master regulator of megakaryopoiesis, and that *FLI1* defects are associated with profound reductions in dense granule secretion, it is feasible that some of the FLI1-regulated genes could harbour defects in patients with platelet secretion abnormalities. This hypothesis was tested by knockdown of *FLI1* in the megakaryocytic Dami cell line and gene expression analysis to identify novel genes involved in platelet granule biogenesis and secretion. Those genes which are differentially expressed after *FLI1* knockdown and harbour SNVs among cases with platelet secretion abnormalities would then be considered to be strong candidates for further investigation. This approach, and the results obtained, will be described fully in chapter 6.

4 Chapter 4. Characterisation of Two Novel *FLI1* Variants causing Substitution of Arginine 340 in the ETS Domain of FLI1 in Patients with Platelet Dense Granule Secretion Disorders

4.1 Introduction

4.1.1 Discovery of FLI1

Friend virus (FV) -A and FV-P are complexes of spleen focus-forming virus (SFFV), either SFFV-A or SFFV-P, and Friend murine leukaemia virus (F-MuLV) (Ben-David & Bernstein, 1991). The FV complex is oncogenic and known to induce erythroleukaemia in mice, which is characterised by rearrangement of the E26 transformation-specific or E-twenty-six (ETS) family member Spi-1 proto-oncogene (*Spi1)* in 95% of cases (Friend, 1957; Moreau-Gachelin et al., 1988). F-MuLV alone does not induce erythroleukaemia in adult mice (Troxler & Scolnick, 1978), although it induces a variety of haematopoietic neoplasms in susceptible newborn mice (Silver & Kozak, 1986). In 1990, *Fli1* was identified as a common proviral insertion site that was rearranged in over 75% of FMuLV-induced erythroleukaemia cells (Ben-David et al., 1990). A year later, murine FLI1 was recognised as a member of the ETS family (Ben-David et al., 1991) and based on extensive amino acid similarity, the human homologue, *FLI1*, was cloned and characterised from a T-cell leukaemia line (Watson et al., 1992). The gene is almost 127 kilobase (Kb) in size and located on chromosome 11 (11q23-24).

The Ensembl genome browser (www.ensembl.org) lists nine *FLI1* transcripts which have been annotated by the Genome Reference Consortium for Human Build 38 (GRCh38). There are four non-coding transcripts and five protein-coding transcripts, of which four have RefSeq IDs from the National Center for Biotechnology Information (NCBI) as well as Consensus Coding Sequence (CCDS) identifiers. The transcript with the most extended open reading frame (NM_002017) spans 4.1 Kb of genomic DNA, comprising nine exons which are transcribed to yield a 51 kilodalton (kDa) protein that has 452 amino acids. The other three protein-coding transcripts, NM_001271010 (4.15 Kb), NM_001271012 (3.43 Kb) and NM_001167681 (3.86 Kb), are translated to generate 386, 259 and 419 amino acid isoforms of FLI1.

FLI1, together with the ETS transcription factors, ERG and FEV, belongs to the ERG subfamily of ETS transcription factors (Hollenhorst et al., 2011). FLI1 and ERG have roles in megakaryocyte (MK) differentiation and angiogenesis (Kruse et al., 2009; Liu et al., 2008; McLaughlin et al., 2001), while FEV plays a role in brain development (Hendricks et al., 2003).

4.1.2 Expression of FLI1

During murine embryogenesis, *Fli1* is expressed before day 8 in the extraembryonic yolk sac of the haemangioblasts (Mélet et al., 1996). At day E8.5 *Fli1* can be detected in all parts of the newly formed mesoderm, primarily in the endothelial and neural crest cells, and from day E11.5, it can be detected in MK-like cells in the foetal liver, endothelial cells, spleen and thymus (Mélet et al., 1996). A similar pattern of expression was observed for the FLI1 homologue in Xenopus embryos where it was localised to angioblasts, endothelial cells and the neural crest (Mager et al., 2004). In zebrafish, *fli1* expression is one of the earliest indicators of haemangioblast formation during embryogenesis, where it is detected in sites of developing vasculature, suggesting a role in the formation of both the endothelium and blood cells (Brown et al., 2000). Interestingly, *fli1* expression overlaps with that of *gata2* in the haemangioblast, but in the later stages of development, *fli1* is found in the developing vasculature, while *gata2* is expressed in haematopoietic cells (Brown et al., 2000).

Fli1 is highly expressed in the thymus, heart, muscle and spleen of adult mice and expressed at lower levels in the brain, kidney, testes and liver (Ben-David et al., 1991; Mélet et al., 1996). In humans, *FLI1* is expressed in both endothelial and haematopoietic cell lineages (Hollenhorst et al., 2004) and found in peripheral blood lymphocytes, thymus, bone marrow (BM), ovary, spleen, heart (Watson et al., 1992) and platelets (Bastian et al., 1999). It has also been detected in transformed cell lines having lymphoid, myeloid and erythroid origins (Hromas et al., 1993; Klemsz et al., 1993; Watson et al., 1992).

4.1.3 Structure of FLI1

As shown in Figure 4.1 like other members of the ETS family, FLI1 possesses the hallmark ETS domain (residues 280-365). In addition, it has a pointed N-terminal domain (PNT; residues 99-198).

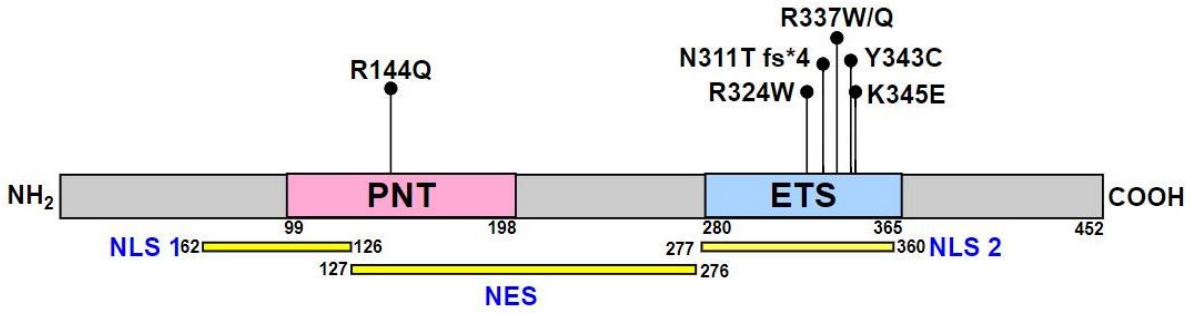


Figure 4.1 Human FLI1 protein domain structure
The domain structure of FLI1 (NP_002008) and the locations of amino acid substitutions that have been reported in patients with inherited platelet bleeding disorders are indicated. p.R144Q (Poggi et al., 2015); p.R324W (Stevenson et al., 2015); p.Asn331Thrfs*4, p.R337W and p.Y343C (Stockley et al., 2013); p.R337Q and p.K345E (Saultier et al., 2017); ETS; ETS DNA binding domain, NES; nuclear export signal, NLS; nuclear localisation signal, PNT; pointed N-terminal domain.

4.1.3.1 The ETS domain of FLI1

As with other ETS family members, the DNA binding capability of FLI1 lies within its ETS domain (Liang et al., 1994). The conserved ETS domain is approximately a 85 amino acid sequence that binds to target DNA sequences having a core (C/A)**GGA**(A/T) motif. The preferred DNA sequence to which FLI1 binds is (A/t)(C/t)(C/a)**GGA**A(G/A)(T/c) (Szymczyna & Arrowsmith, 2000).

FLI1 was the first ETS family member for which a structure of an ETS domain bound to DNA was described (Liang et al., 1994). It comprises three alpha helices, H1 (residues 283-292), H2 (315-323), and H3 (332-344) as well as four short antiparallel stranded beta sheets, B1 (300-303), B2 (308-311), B3 (348-351), and B4 (357-361) (Liang et al., 1994). The H3 helix is the main region that interacts with the major groove in DNA while the loop between the H2 and H3 helices, and the wing between the B3 and B4 beta sheets, interact with the minor groove of DNA (Liang et al., 1994).

In addition to its role in DNA binding, the ETS domain incorporates one of two nuclear localisation signals (NLSs) present in FLI1; nuclear localisation signal 2 (NLS2) (Hu et al., 2005). Some residues required for DNA binding also appear to be vital for nuclear targeting of FLI1. For example, substitution of K325, R337, R340 or K350 with alanine residues causes a reduction in DNA binding capacity as well as nuclear accumulation of FLI1, while FLI1 variants in which R334 or R355 have been substituted with alanine residues still retain their DNA binding capability but fail to show nuclear accumulation

(Hu et al., 2005). Hu et al. (2005) also demonstrated that the affinity of the ETS domain for its target DNA sequence is higher than that for the nucleocytoplasmic shuttling receptors. The ETS domain has also been shown to mediate interactions of FLI1 with itself (Hou & Tsodikov, 2015) and with other proteins, such as GATA1 (GATA binding protein 1, erythroid transcription factor) and KLF1 (Kruppel like factor 1) (Eisbacher et al., 2003; Starck et al., 2003).

4.1.3.2 The pointed N-terminal domain of FLI1

The PNT domain, which is also known as the helix-loop-helix (HLH) or the sterile alpha motif (SAM) domain, comprises approximately 80 residues that are organised as four alpha helices (Hollenhorst et al., 2011). This domain is present in one-third of ETS family members (Hollenhorst et al., 2011). However, PNT of FLI1 has not received as much attention as its ETS domain. As shown in Figure 4.1 it overlaps partly with the nuclear localisation signal 1 (NLS1) and the nuclear export signal (NES) (see below) (Hu et al., 2005) and is reported to be involved in the interaction between FLI1 and ETV6 (Kwiatkowski et al., 1998).

4.1.3.3 The nuclear localisation and export signals of FLI1

Two NLS and one NES have been identified in FLI1 (Figure 4.1) (Hu et al., 2005). NLS1 has been localised to residues 62 to 126, partly overlapping with the PNT domain, while NLS2 is located between residues 277 and 360, coinciding with the ETS domain (Hu et al., 2005). The NES has been mapped to the amino acid sequence that encompasses residues 127 to 276 (Hu et al., 2005). Although each NLS can independently direct FLI1 to the nucleus, two signals are required for normal function as only combined mutations in both signals can entirely abolish nuclear accumulation of FLI1 (Hu et al., 2005).

4.1.4 Functions of FLI1

Extensive *in vitro* and *in vivo* studies have revealed a critical role for FLI1 in vasculogenesis, as well as haematopoietic cell proliferation and differentiation. Several murine models of FLI1 deficiency have been reported and those having phenotypic features related to these aspects of FLI1 function are summarised in Table 4.1. While these roles provide the primary focus for this thesis, FLI1 also has roles in determining cellular adhesion properties, regulation of the extracellular matrix (Hart et al., 2000;

Remy et al., 2002), and apoptosis (Cui et al., 2009; Yi et al., 1997). FLI1 acts mainly as a transcriptional activator (Rao et al., 1993), though it has been shown to repress expression of some target genes (Kubo et al., 2003; Starck et al., 2003; Tamir et al., 1999).

4.1.4.1 Role of FLI1 in vasculogenesis, angiogenesis and blood cell formation

FLI1 is one of the transcription factors considered to be essential for control of haematopoietic stem cell differentiation and one of the regulators of early endothelial development (Kruse et al., 2009; Liu et al., 2008; Wilson et al., 2010). In a recent study, conditional overexpression of 15 transcription factors known to be important for induction of haematopoiesis in human embryonic stem cells showed that only *FLI1* overexpression was able to induce differentiation to the haematopoietic lineage (Zhao et al., 2018).

Two mouse models which fail to express *Fli1* (Hart et al., 2000; Spyropoulos et al., 2000) and a mouse expressing a truncated FLI1 have been described (Moussa et al., 2010). In all three cases, no visible phenotype was observed in heterozygous embryos, and the heterozygous adult mice had peripheral blood cell counts and bleeding times within the normal range (Hart et al., 2000; Moussa et al., 2010; Spyropoulos et al., 2000). In contrast, homozygous null mice presented with fatal bleeding at mid-gestation (E11.5 - E12.5) as a result of aberrant vasculogenesis, haematopoiesis and megakaryopoiesis (Hart et al., 2000; Spyropoulos et al., 2000). Histological examinations revealed endothelial abnormalities, disorganised columnar epithelium and disrupted basement membranes (Spyropoulos et al., 2000). Although normal yolk sac vascularisation was observed in the *Fli1* null embryo, the ability of *Fli1*$^{-/-}$ cells to contribute to the development of the vascular endothelium and the foetal liver was compromised (Hart et al., 2000). Homozygous mice expressing the truncated FLI1 showed a significant reduction in viability and were also thrombocytopenic (Moussa et al., 2010).

Asano et al. (2010) reported a mouse model with conditional knockout of *Fli1* in endothelial cells which displayed severe abnormalities of the vasculature, including a disorganised dermal vascular network characterised by irregular vessel diameters, an increase in vascular permeability and impaired basement membrane development

(Asano et al., 2010). Supporting its role in the development of the vasculature, FLI1 was found to regulate expression of several endothelial genes considered essential for vascular homeostasis and angiogenesis (Abedin et al., 2014; Asano et al., 2010; Deramaudt et al., 1999; Gory et al., 1998; Göttgens et al., 2002; Hart et al., 2000; Landry et al., 2005; Le Bras et al., 2010; Marks-Bluth et al., 2015; Pimanda et al., 2006).

4.1.4.2 Role of FLI1 in erythroid/megakaryocytic differentiation

Platelets and MKs express FLI1 (Bastian et al., 1999), and early evidence supported the vital role of FLI1 in megakaryocytic differentiation. Treatment of human erythroleukaemia K562 cells with phorbol 12-myristate 13-acetate (PMA) to induce MK differentiation is accompanied by overexpression of *FLI1* and a severe reduction in *GATA1* expression (Athanasiou et al., 1996). In contrast, erythropoietin-induced erythroid differentiation of the HB60 cell line resulted in downregulation of *Fli1* while overexpression of *Fli1* inhibited erythroid differentiation (Tamir et al., 1999).

Overexpression of *fli1* in Xenopus embryos and zebrafish resulted in a spectrum of abnormalities including developmental anomalies in the head and heart, alterations in cell adhesion properties and absence of erythrocyte differentiation (Brown et al., 2000; Remy et al., 2002). Interestingly, overexpression of both *FLI1* and *ERG* in human BM erythroblasts resulted in transdifferentiation into MKs that were able to produce functional platelets (Siripin et al., 2015).

Examination of foetal liver from the murine *Fli1* knockout model mentioned above (section 4.1.4.1, Table 4.1), revealed reduced numbers of pronormoblasts and basophilic normoblasts (Spyropoulos et al., 2000), with an elevated number of abnormal undifferentiated MKs, which showed decreasing numbers of α-granules and had disorganised platelet demarcation membranes (Hart et al., 2000). The mice expressing the truncated FLI1 exhibited thrombocytopenia, abnormal platelet activation and aggregation as well as prolonged bleeding time (Moussa et al., 2010). Additionally, some of the genes associated with MK development were found to be downregulated, including those encoding *Cd36* (CD36 molecule), *Gp9* (Glycoprotein (GP) IX), *Itga2b* (Integrin subunit alpha 2b), *Mafg* (MAF BZIP transcription factor G) , *Mpl* (MPL proto-oncogene, Thrombopoietin receptor), *Nfe2* (Nuclear factor erythroid

2), *Pf4* (Platelet factor 4), *Rab27B* (RAB27B, Member of RAS oncogene family) (Moussa et al., 2010). *In vitro* cell culture studies of murine *Fli1* knockout cells at E10.0, prior to haemorrhage, revealed a defect in erythropoiesis, downregulation of *Mpl*, and absence of megakaryopoiesis (Kawada et al., 2001). Assessment of the effect of FLI1 on adult haematopoiesis by induction of *Fli1* deletion in adult mice showed, in addition to a defect in myelopoiesis in multiple lineages, an increase in the number of MK-erythrocyte progenitors, a decrease in large mature MKs, and increased numbers of erythrocytes in the BM (Starck et al., 2010). The peripheral blood revealed mild thrombocytopenia with a normal number of circulating red cells (Starck et al., 2010).

In vivo and *in vitro* studies have demonstrated that FLI1 modulates the expression of both early and late MK-specific genes (Pang et al., 2006). These include *CD36* (Moussa et al., 2010), *GP1BA* (GP Ib platelet subunit alpha) (Eisbacher et al., 2003), *GP9* (Bastian et al., 1999; Eisbacher et al., 2003; Hart et al., 2000; Moussa et al., 2010), *HOXA10* (Homeobox A10) (Gosiengfiao et al., 2007), *HPS4* (Biogenesis of lysosomal organelles complex 3 subunit 2), *ITGA2B* (Lemarchandel et al., 1993; Moussa et al., 2010; Zhang et al., 1993), *MAFG* (Moussa et al., 2010), *MPL* (Deveaux et al., 1996; Moussa et al., 2010), *NFE2* (Moussa et al., 2010), *PF4* (Lemarchandel et al., 1993; Moussa et al., 2010), *RAB27B* (Moussa et al., 2010; Zang et al., 2016), *VWF* (Von Willebrand factor) (Schwachtgen et al., 1997).

Table 4.1 Features of murine FLI1 deficiency models

Model	Phenotype
Substitution of the ETS domain with a lacZ reporter gene [(1)]	• Die at E11.5 • Dysmegakaryopoiesis; undifferentiated MKs that show abnormal ultrastructural features including reduced α-granule numbers, disorganised platelet demarcation membranes (resembles patients with 11q deletions) • Compromised ability of *Fli1*$^{-/-}$ cells to contribute to vascular endothelium and MK lineages • Downregulation of *Tek* (TEK Receptor Tyrosine Kinase) and *Gp9*
Substitution of the ETS domain with a l*oxP*-flanked *Neo* cassette [(2)]	• Bleed at E11.5 and die by E12.5 • Endothelial related abnormalities at the site of haemorrhage at E11.0 • Absence of red cells from the yolk sac vasculature at E11.0 • Reduced progenitor numbers in the foetal livers at E11.0 • *In vitro* cell culture of progenitor cells from *Fli1*$^{-/-}$ embryo yolk sacs at E10.0 showed absence of MK colonies and a moderate loss of erythroid progenitors • *In vitro* cell culture studies of embryonic stem cells and cells from the aorta-gonad-mesonephros region of E10.0 embryos [(3)] confirmed abnormal erythroid development and defective megakaryopoiesis that was partially explained by reduced *Mpl* expression
Truncated Fli1 (amino acids 1 to 384)[(4)]	• Early postnatal lethality (30% survival of homozygotes to adulthood) • Thrombocytopenia • Prolonged bleeding time • Abnormalities in platelet activation and aggregation • Downregulation of MK genes (*Mpl, Itga2b, Cd36, Gp9, Pf4*, *Nfe2*, *Mafg* and *Rab27B*) • Reduced binding of GATA1 to the promoters of some target genes
Inducible *Fli1* gene deletion by targetting exon 9 in adult mice [(5)]	• Mild thrombocytopenia • Bone marrow has an increased number of bipotent MK-erythrocytic progenitors that are unable to generate mature MK colonies • Increased numbers of erythrocytes and natural killer cells, with a decreased number of granulocytic cells
Conditional knockout of *Fli1* in endothelial cells by targetting exons 3 and 4 [(6)]	• Severe abnormalities of the skin vasculature (recapitulates the scleroderma phenotype of dermal blood vessels) • Increased vessel permeability • Impaired pericyte/vascular smooth muscle cell coverage of the blood vessels • Alteration in the expression of some endothelial genes, including *Cdh5* (Cadherin 5), *Col4a1* (Collagen type IV alpha 1 chain), *Mmp9* (Matrix metalloproteinase 9), *Pdgfb* (Platelet-derived growth factor subunit B), *Pecam1* (Platelet and endothelial cell adhesion molecule 1), *S1pr1* (sphingosine-1-phosphate receptor 1), *Tek* (TEK Receptor Tyrosine Kinase)

(1) Hart et al., 2000, (2) Spyropoulos et al., 2000, (3) Kawada et al., 2001, (4) Moussa et al., 2010, (5) Starck et al., 2010, (6) Asano et al., 2010. ETS; ETS DNA binding domain, MK; megakaryocyte.

4.1.5 The role of FLI1 in disease

Aberrant *FLI1* expression, as a result of inherited or acquired genetic alterations, has been implicated in the pathogenesis of haematological and non-haematological malignancies (Bonetti et al., 2013; Del Portillo et al., 2017; Delattre et al., 1992; Kornblau et al., 2011; Lee et al., 2015; Paulo et al., 2012; Scheiber et al., 2014; Song et al., 2015; Torlakovic et al., 2008), autoimmune disorders such as systemic lupus erythematosus (Georgiou et al., 1996; Mathenia et al., 2010; Zhang et al., 1995; Zhang et al., 2004) and systemic sclerosis (Asano et al., 2010; Kubo et al., 2003; Takahashi et al., 2017) and bleeding disorders (Poggi et al., 2015; Raslova et al., 2004; Saultier et al., 2017; Stevenson et al., 2015; Stockley et al., 2013; Vo et al., 2017). More recently, FLI1 deficiency has been implicated in the pathogenesis of pulmonary arterial hypertension (Looney et al., 2017). The broad spectrum of diseases associated with FLI1 defects most likely reflects the disturbed expression of different downstream targets of FLI1. The association between FLI1 and inherited platelet bleeding disorders (IPDs) is discussed further below.

4.1.5.1 Bleeding disorders due to structural chromosomal abnormalities that result in deletion of FLI1: Jacobsen syndrome and Paris-Trousseau syndrome

Jacobsen syndrome and Paris-Trousseau syndrome are caused by partial deletions of chromosome 11 (Breton-Gorius et al., 1995; Jacobsen et al., 1973), which result in loss of up to 16 megabases of genomic DNA that encompasses approximately 300 genes, including *FLI1*.

Also known as distal 11q deletion syndrome, Jacobsen syndrome is a rare congenital disorder that shows variable phenotypic expression depending on the size and location of the deletion affecting the q arm of chromosome 11 (Penny et al., 1995). Most commonly, patients present with cardiac defects, growth and psychomotor retardation, trigonocephaly, dysmorphic facies, digit anomalies, pancytopenia, thrombocytopenia (Penny et al., 1995) and dense granule storage pool deficiency (White, 2007). Jacobsen syndrome is also characterised by the presence of giant α-granules in 3-20% of platelets (White, 2007), that have been shown to arise through fusion of α-granules (Krishnamurti et al., 2001).

Paris-Trousseau syndrome (PTS) is an autosomal dominant disorder resulting from deletion of chromosome 11q23. Patients present with a lifelong mild haemorrhagic tendency, dysmegakaryopoiesis and macrothrombocytopenia (Breton-Gorius et al., 1995; Favier et al., 1993; Favier et al., 2003). The dysmegakaryopoiesis manifests as an increase in the number of micro-MKs in the BM, the presence of large platelets in the peripheral blood, about 15% of which contain giant fused α-granules, and a defect in platelet secretion (Breton-Gorius et al., 1995). The presence of giant α-granules in platelets, but not MKs, of affected patients suggests that fusion occurs within the platelets (Breton-Gorius et al., 1995; Favier et al., 2003). Interestingly, Raslova et al. (2004) described two distinct populations of normal and small immature MKs in the BM of PTS patients which were thought to arise as a result of transient monoallelic expression of FLI1 early during megakaryopoiesis alongside the hemizygous loss of FLI1 (Raslova et al., 2004). However, Vo et al. (2017) did not confirm the presence of two different MK populations, hypothesising that FLI1 haploinsufficiency alone underlies PTS. The severity of bleeding in individuals with PTS does not correlate with their platelet count, indicating an intrinsic platelet defect (Grossfeld et al., 2004). This is supported by the reported correction of the thrombocytopenia within the first two years of life in some patients, while the platelet abnormality remains (Favier et al., 2003; Grossfeld et al., 2004). The similarities in phenotype between patients with PTS and Jacobsen syndrome have led to the suggestion that PTS is a variant of Jacobsen syndrome (Favier et al., 2003; Krishnamurti et al., 2001).

Deletions at 11q23 result in hemizygous expression of many genes, including *ETS1* and *FLI1*, loss of which were suggested to explain the defects observed in the MKs (Breton-Gorius et al., 1995). Considerable evidence supports the loss of *FLI1* over that of *ETS1* as the underlying cause of the dysmegakaryopoiesis observed. Firstly, MKs express 100-fold more *FLI1* than *ETS1* (Raslova et al., 2004). ETS1 is not essential for megakaryopoiesis as MKs were able to differentiate normally in *Ets1* null mice (Bartel et al., 2000), while homozygous *Fli1* null mice presented with dysmegakaryopoiesis similar to that observed in patients with terminal deletions of 11q (Hart et al., 2000; Spyropoulos et al., 2000). Moreover, studies at the single cell level revealed that MK maturation, assessed as the transition from a $CD42a^-$ to a $CD42a^+$ phenotype, is sensitive to FLI1 level (Raslova et al., 2004), and overexpression of *FLI1* in haematopoietic stem cells ($CD34^+$) derived from a PTS patient was shown to rescue

megakaryopoiesis *in vitro* (Raslova et al., 2004). Finally, in a recent study using MKs derived from induced pluripotent stem (iPS) cells from a PTS patient and a human iPS cell line with a targeted heterozygous *FLI1* deletion, the features of derived MKs and platelets were shown to replicate the defects described in PTS and the phenotype was rescued when *FLI1* was overexpressed (Vo et al., 2017). These results support the loss of *FLI1* as the underlying cause of the MK/platelet abnormalities observed in Jacobsen syndrome and PTS, confirming the essential role of FLI1 in megakaryopoiesis, platelet production and function.

4.1.5.2 Bleeding disorders due to intragenic *FLI1* variation

Heterozygous alterations in *FLI1* were first associated with IPDs by the UK Genotyping and Phenotyping of Platelets (UK-GAPP) study group who reported enrichment of both *FLI1* and *RUNX1* defects among 13 index cases with a history of excessive bleeding that was characterised predominantly by a significant reduction in platelet dense granule secretion (Stockley et al., 2013). Thus, next-generation sequencing of a panel of 260 platelet genes identified two index cases with novel heterozygous missense mutations predicting p.Arg337Trp and p.Tyr343Cys substitutions in the ETS domain of FLI1, both of which resulted in a loss in transactivation capacity of FLI1. A third index case was found to be heterozygous for a 4 base pair deletion, p.Asn331Thrfs*4 (Stockley et al., 2013). The p.Tyr343Cys variant was associated with thrombocytopenia in the affected index case and another affected family member, while the p.Asn331Thrfs*4 defect was associated with an increase in mean platelet volume (Stockley et al., 2013). Interestingly, in addition to their association with excessive bleeding, the missense *FLI1* defects were associated with alopecia, eczema or psoriasis, as well as recurrent viral infections in the affected members of the two families identified. While this study was ongoing, four further novel FLI1 missense variants were reported in the literature, all of which were associated with bleeding and platelet granule abnormalities (Poggi et al., 2015; Saultier et al., 2017; Stevenson et al., 2015).

4.2 Hypothesis and Aims

The first part of this study (chapter 3) focused on 34 index cases with unexplained IPDs that were characterised by defects in either dense granule secretion or Gi-signalling,

all of whom were enrolled in the UK-GAPP study. Whole exome sequencing (WES) and bioinformatic analysis of the genetic data obtained for those 22 patients with dense granule secretion defects highlighted novel *FLI1* variants in the index cases from two families; F1 (c.992-995del:p.Asn331Thrfs*4) and F11 (c.1018C>T:p.Arg340Cys). The 4bp deletion identified in F1 was predicted to cause a frameshift and introduce a premature stop codon and no further work was undertaken on this variant as part of this study. The *FLI1* defect in F11 (hereinafter referred to as family A) was of interest since it predicted a p.R340C substitution in a region of the protein that was previously shown to harbour defects (c.1009C>T:p.Arg337Trp and c.1028A>G:p.Tyr343Cys) in members of two other families enrolled in the UK-GAPP study with IPDs, which were characterised predominantly by defects in platelet secretion and other clinical features including mild thrombocytopenia, eczema, alopecia and recurrent viral infection (Stockley et al., 2013). However, in contrast to the defects identified previously, the predominant clinical feature in the patient with the p.Arg340Cys defect was in platelet secretion. It was therefore of interest to study this variant further to examine how it differs in its properties from the other two variants.

During the course of this study, we were approached by Courtney D. Thornburg and Diane Masser-Fryeat from Rady Children's Hospital, San Diego as they had identified another novel *FLI1* variant also affecting codon 340, but which was predicted to result in substitution of arginine by histidine (c.1019G>A:p.Arg340His). Interestingly, the affected members of the family (referred to here as family B) who carried this *FLI1* variant were clinically similar to those cases previously reported by Stockley et al. (2013).

Given the critical role of FLI1 in megakaryopoiesis, we hypothesised that the IPDs which affected the index case in family A, and three members of family B were due to the inherited *FLI1* variants c.1018C>T:p.Arg340Cys and c.1019G>A:p.Arg340His respectively. Since neither of the index cases nor any of their family members were available for investigation, this hypothesis was explored using experimental approaches which aimed to reproduce the patient phenotypes *in vitro*. In addition, further work was undertaken to characterise the p.Arg337Trp and p.Tyr343Cys variants of FLI1 which were identified previously in our laboratory. The aims of the work described in this chapter were therefore to:

i. Predict the DNA binding capability of four FLI1 variants p.R340C, p.R340H, p.R337W and p.Y343C
ii. Assess the transactivation capacity of two novel FLI1 variants in which arginine 340 is substituted by either cysteine or histidine, p.R340C and p.R340H, and further investigate the transactivation capacity of the previously described variants, p.R337W and p.Y343C
iii. Evaluate nuclear localisation of all four FLI1 variants p.R340C, p.R340H, p.R337W and p.Y343C

4.3 Methods

4.3.1 Prediction of the effects of amino acid substitutions in FLI1

The possible effects of the R340C, R340H, R337W and Y343C amino acid substitutions on the structure of FLI1 and on the FLI1-DNA interaction were predicted using the crystal structure of the ETS domain of FLI1 (amino acids 279-371) bound to the double-stranded oligonucleotide GACCGGAAGTG which was deposited in the RCSB Protein Data Bank (PDB ID: 5JVT) (Hou et al., 2016). The structure was visualised, *in silico* mutagenesis was carried out, and polar interactions were predicted using the tools available in Pymol (version 0.99rc6). The most common amino acid side chain orientations were used in the simulations.

4.3.2 Assessment of transcriptional activity of overexpressed FLI1 variants using dual luciferase reporter assays in mammalian cell lines

Dual luciferase reporter assays were performed to examine the ability of FLI1 variants to transactivate the *GP6* (glycoprotein VI) promoter in HEK 293T cells and the megakaryocytic Dami cell line as described in section 2.2.11.

4.3.3 Evaluation of the subcellular localisation of overexpressed FLI1 variants in mammalian cell lines using wide-field microscopy

To evaluate nuclear localisation of FLI1 variants, FLI1 was detected either by immunostaining or by assessing the enhanced green fluorescent protein (EGFP) signal from FLI1-EGFP fusion proteins as described in section 2.2.12.

4.3.4 Evaluation of the subcellular localisation of overexpressed FLI1 variants in HEK 293T cells using lithium dodecyl sulphate polyacrylamide gel electrophoresis and western blotting

Subcellular localisation of FLI1 was evaluated by electrophoresis and immunoblotting of cytoplasmic and nuclear fractions of lysates from HEK 293T cells that transiently overexpressed different FLI1 variants as described in section 2.2.13.

4.4 Results

4.4.1 Identification of two *FLI1* defects which predict substitution of arginine 340 in FLI1

4.4.1.1 Clinical features of index cases with *FLI1* variants

Figure 4.2 shows the pedigrees for families A and B. The index case in family A (F.A.II.1) was recruited to the UK-GAPP study with a history of bleeding and a diagnosis of storage pool disease. Phenotyping of her platelets confirmed a reduction in agonist-induced dense granule ATP secretion. Following the referral of the index case, her daughter, F.A.III.1 and son, F.A.III.2 were both diagnosed with the same condition. Both were reported to have thrombocytopenia at birth, which was confirmed by the medical record in the case of F.A.III.1. However, the platelet counts for all three affected family members were within the normal range on later occasions. While F.A.II.1 and F.A.III.1 both had a history of bleeding, F.A.III.2 had no bleeding symptoms though he developed arthritis at the age of 18. DNA was available only from subject F.A.III.2.

The index case (F.B.I.1) in family B, and her two sons (F.B.II.1, F.B.II.2), all had a history of mucocutaneous bleeding symptoms and a diagnosis of storage pool disease, which was defined by a reduction in dense granule secretion. They also suffered from mild eczema and had a history of thrombocytopenia. The index case (F.B.I.1) had been diagnosed with Ehlers-Danlos syndrome III. Additionally, F.B.I.1 and F.B.II.1 had a history of recurrent infections and F.B.II.2 had a history of mild neutropenia.

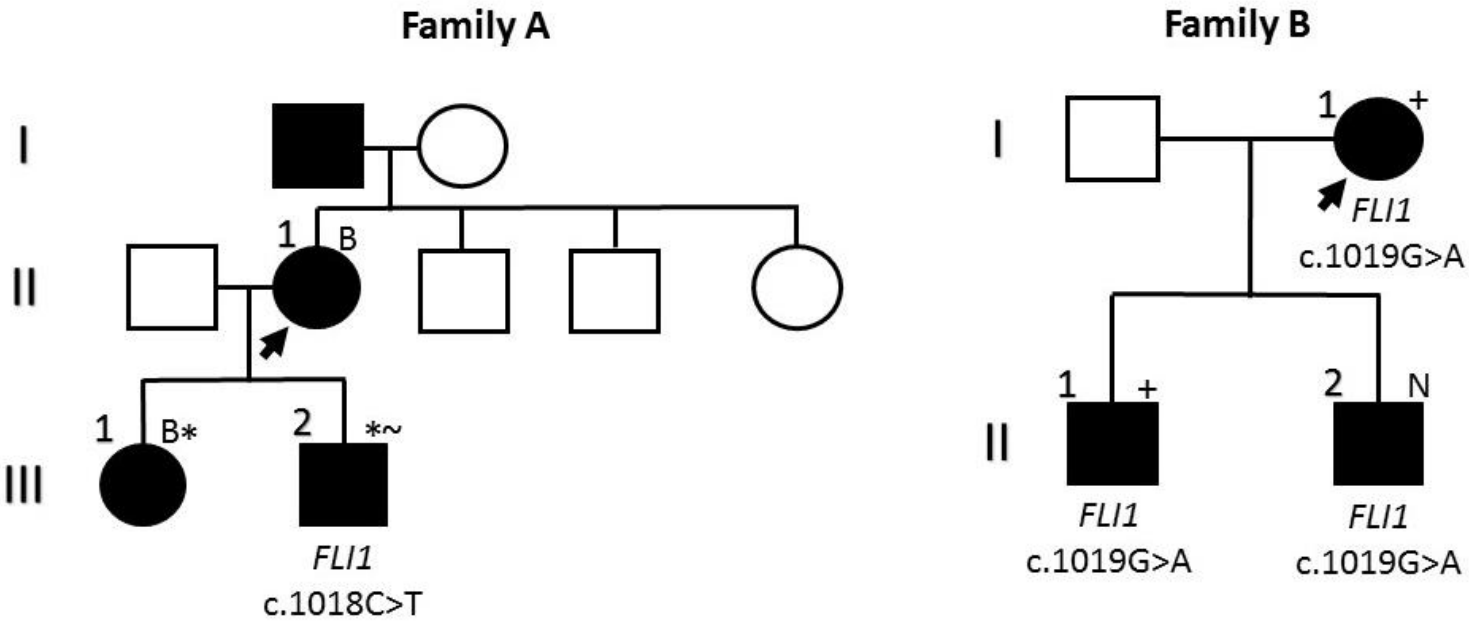


Figure 4.2 Inheritance of dense granule secretion defects and other clinical features in families with *FLI1* variants

Individuals heterozygous for the c.1018C>T and c.1019G>A transitions in *FLI1* are indicated. The index case in each family is indicated by an arrow. Members of family A diagnosed with storage pool disease and having a confirmed dense granule secretion defect are indicated by black filled symbols. The presence of a bleeding history is indicated by a "B". History of thrombocytopenia is indicated by an asterisk, and a history of arthritis is indicated by the "~" symbol. Members of family B diagnosed with storage pool disease and having bleeding symptoms, a history of thrombocytopenia and eczema are indicated by black filled symbols. A "+" symbol indicates a history of recurrent infections while "N" indicates a history of neutropenia.

4.4.1.2 Confirmation of *FLI1* c.1018C>T and c.1019G>A by Sanger sequencing

Analysis of WES data from participants enrolled in the UK-GAPP study identified a novel heterozygous SNV in *FLI1* (c.1018C>T) in one participant, F11.1 (chapter 3). For simplification, F11 will be named family A from here. The presence of the c.1018C>T transition, predicting substitution of arginine by cysteine at amino acid position 340 in FLI1 (p.R340C), was confirmed by Sanger sequencing following PCR amplification of a genomic DNA fragment spanning the candidate defect from F.A.III.2 (Figure 4.3).

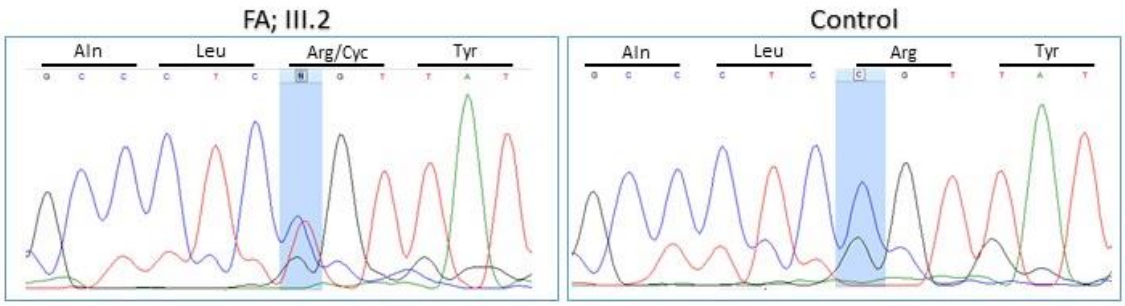


Figure 4.3 Sequencing of a fragment of *FLI1* amplified from the DNA of index case F.A.III.2 and a healthy control subject

Sanger sequencing confirmed the presence of the heterozygous FLI1:c.1018C>T transition (shaded in blue) in F.A.III.2 (previously F11.1).

WES analysis was undertaken for three members of family B resulting in the identification of a novel heterozygous c.1019G>A transition in *FLI1* which predicts substitution of arginine by histidine at amino acid position 340 (p.R340H). The presence of this defect was confirmed by Sanger sequencing in all three family members at the referring centre in San Diego.

4.4.2 Predicted interactions of FLI1 variants with DNA

Assessment of the polar interactions between the ETS domain of FLI1 and a double-stranded DNA oligonucleotide incorporating the FLI1 consensus sequence (GACCGGAAGTG), identified thirteen amino acids in FLI1 which were predicted to interact directly with the DNA. These were Q282, L283, W321, K325, N329, K334, R337, R340, Y341, Y342, Y343, K350, and Y356 (Figure 4.4).

Arginine residues 337 and 340 and tyrosine 343 are located in the ETS-H3 helix (332-344), the primary helix that interacts with DNA. The positively charged guanidinium group of arginine 337 is predicted to have polar interactions with the third guanine in

the DNA sequence GACCG**G**AAGTG, also with tyrosine 341 in FLI1. Similarly, arginine 340 is predicted to have polar interactions with the second guanine in the DNA sequence GACC**G**GAAGTG, also with serine 336 in FLI1. The aromatic ring in tyrosine 343 is predicted to have two polar interactions with the second cytosine in the DNA sequence GAC**C**GGAAGTG (Figure 4.5 A-C, Appendix 5).

In silico substitution of arginine 337 by tryptophan is predicted to maintain the interactions between FLI1 and DNA, while the interaction with tyrosine 341 is lost. Substitution of arginine 340 by cysteine is predicted to lead to loss of the interactions made by the guanidine group with DNA and with serine 336, while substitution with a histidine residue at the same position, which introduces a positively charged imidazole ring, maintains the interaction with the consensus DNA sequence, though the interaction with serine 336 is lost. Similarly, substitution of tyrosine 343 by cysteine was predicted to result in a complete loss of the interaction between FLI1 and DNA. Interestingly, the free thiol group introduced by the Y343C and R340C substitutions is not anticipated to interact with other residues in the vicinity (Figure 4.5 D-G, Appendix 5).

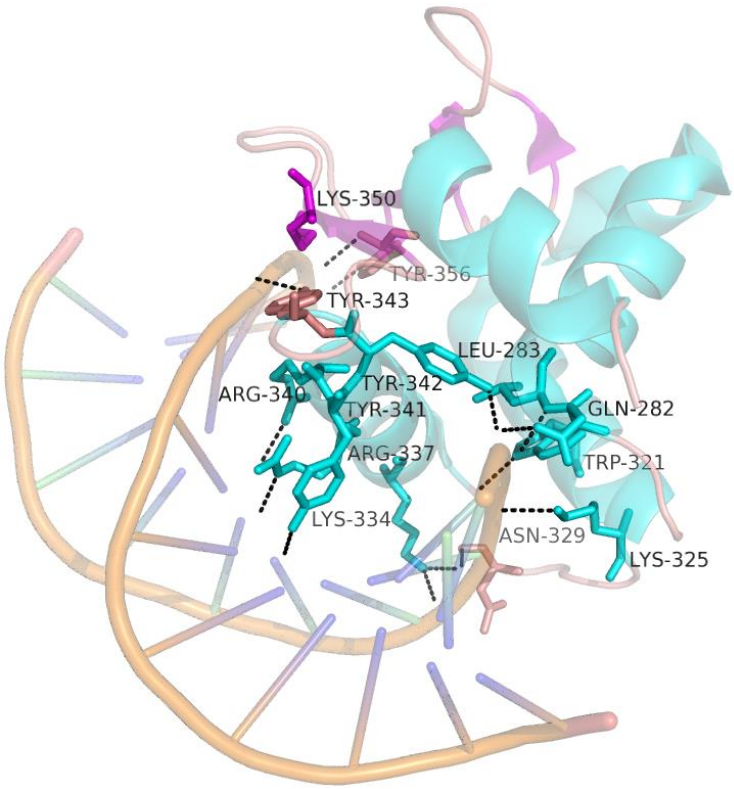


Figure 4.4 Partial structure of FLI1 showing those amino acids that are predicted to interact with the double-stranded DNA fragment GACCGGAAGTG
The amino acids that interact with DNA are shown as sticks. Polar interactions are indicated by dashed lines. Colour coding: cyan; helices, magenta; sheets, loops; pink.

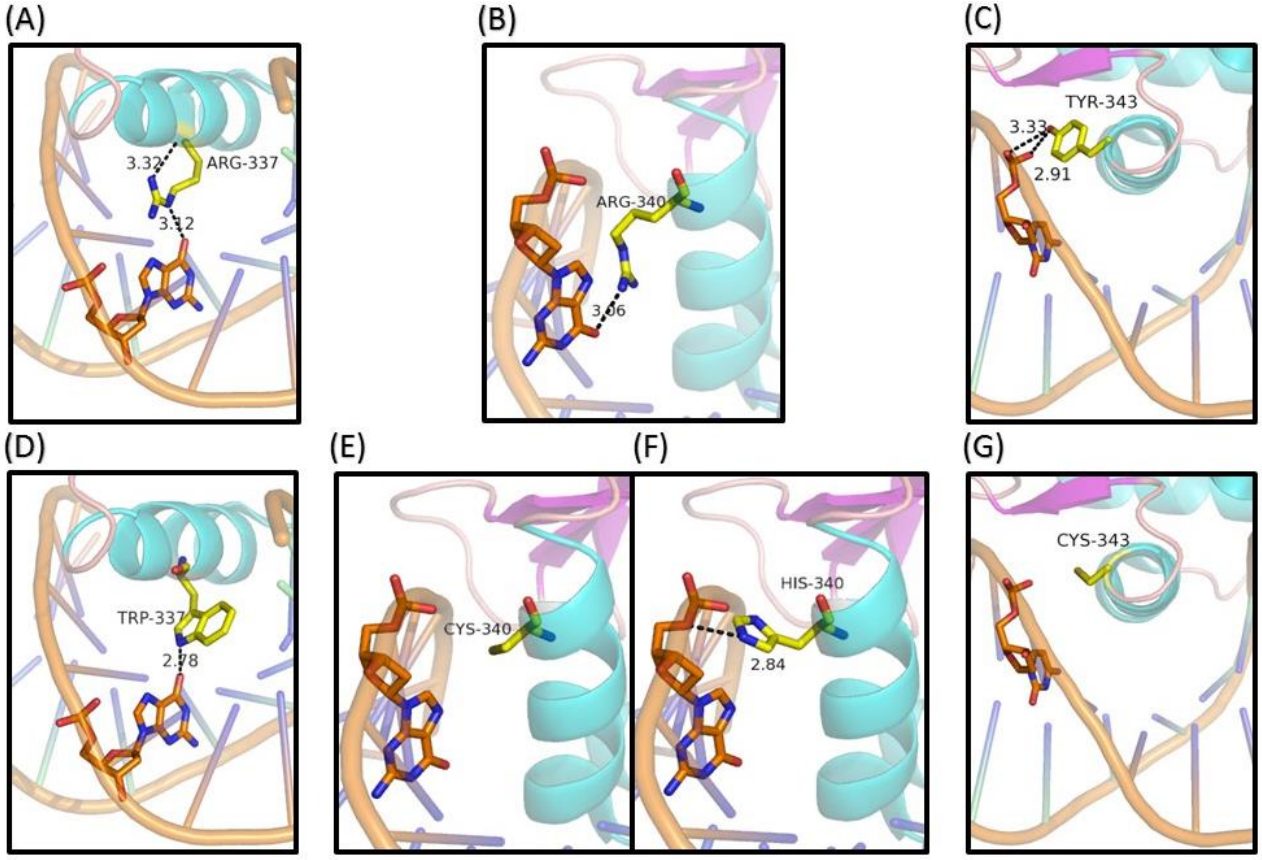


Figure 4.5 Predicted interactions of the residues located at positions 337, 340 and 343 with neighbouring residues and with DNA and the predicted effect of their substitutions found in patients with the FLI1 defect

The predicted interactions of FLI1 with DNA and within FLI1 at (A) R337, (B) R340, (C) Y343 and the effect of their substitutions to (D) R337W (E) R340C, (F) R340H, and (G) Y343C. Dashed lines indicate polar interactions. Colour coding: cyan; helices, magenta; sheets, loops; pink.

4.4.3 Effects of the FLI1 variants on gene transcription

FLI1 promotes the expression of several genes in MKs including *GP1BA*, *GP6*, *GP9* and *ITGA2B*. The FLI1 binding sites in the promoters of these genes have been used in previous studies to evaluate transactivation by different FLI1 variants (Bastian et al., 1999; Hu et al., 2005; Stevenson et al., 2015; Stockley et al., 2013). To examine the effects of the p.R340C and p.R340H substitutions on the transcriptional activity of FLI1, the ability of the corresponding recombinant FLI1 variants to transactivate the *GP6* promoter was assessed in both HEK 293T and Dami cells using a dual luciferase reporter assay. In addition to the R340C and R340H plasmids generated as part of this study, derivatives of pSG5-FLI1 encoding the R337W- and Y343C-FLI1 variants, which were previously generated in our lab and shown to have reduced capacity to transactivate *GP6* in HEK 293T cells, were also used in these experiments (Stockley et al., 2013).

4.4.3.1 FLI1 variants show reduced transactivation of the *GP6* promoter in HEK 293T cells

Compared to HEK 293T cells which were transfected with the empty vector (EV; pSG5), a 5-fold increase in luciferase activity was observed in cells expressing wild-type (WT) FLI1 ($p<0.0001$). There was a significant reduction in luciferase activity in the presence of the R337W (86.58% reduction compared to WT-FLI1; $p=0.0124$) and the Y343C (88.33% reduction compared to WT-FLI1; $p<0.0001$) variants, confirming previous results (Stockley et al., 2013). Similarly, there was a significant reduction in *GP6* promoter activity in the presence of either the R340C- or the R340H-FLI1 variant when compared to cells expressing WT-FLI1. Interestingly, while *GP6* transactivation by the R340C variant was reduced to a similar extent to that observed with the R337W and Y343C variants (83.27% reduction compared to WT-FLI1; $p<0.0001$), there was a less dramatic reduction in *GP6* transactivation in the presence of the R340H variant (64.20% reduction compared to WT-FLI1; $p=0.0007$).

Co-expression of either the R337W or the Y343C variant with WT-FLI1 to mimic heterozygosity resulted in significant reductions in luciferase activity compared to cells expressing WT-FLI1 alone (R337W+WT-FLI1 68.68% reduction, $p=0.009$; Y343C+WT-FLI1 43.19% reduction, $p=0.0284$), confirming previous results (Stockley et al., 2013). Likewise, co-expression of either the R340C or the R340H variant with

WT-FLI1 led to 31.32% (p=0.0348) and 41.83% (p=0.0001) reductions in luciferase activity respectively when compared to cells expressing WT-FLI1 alone (Figure 4.6 A, Table 4.2).

4.4.3.2 FLI1 variants show reduced transactivation of the *GP6* promoter in Dami cells

Having demonstrated a loss of transactivation capacity of the two R340 variants in HEK 293T cells, their ability to induce *GP6* promoter activity was investigated in megakaryocytic Dami cells along with the R337W and Y343C variants, the activity of which had not previously been studied in Dami cells. Luciferase activity was increased 2-fold in Dami cells expressing WT-FLI1 ($p<0.0001$) when compared with cells transfected with the EV. There were significant reductions, between 31 and 57%, in *GP6* promoter activity in the presence of each of the four FLI1 variants studied when compared to cells expressing WT-FLI1 (R337W: 31.28%, $p<0.0001$; Y343C: 56.41%, $p<0.0001$; R340C: 44.10%, $p=0.0002$; R340H: 49.74%, $p<0.0001$). Similarly, compared with cells expressing WT-FLI1 alone, a significant reduction in luciferase activity was observed when each of the variants was co-expressed with WT-FLI1 to mimic heterozygosity (R337W: 26.15%, $p=0.0025$; Y343C: 29.23%, $p=0.0009$; R340C: 22.56%, $p=0.0009$; R340H: 21.54%, $p=0.0006$) (Figure 4.6 B, Table 4.3). Interestingly, the R340C and R340H variants displayed similar transactivation capacity in Dami cells.

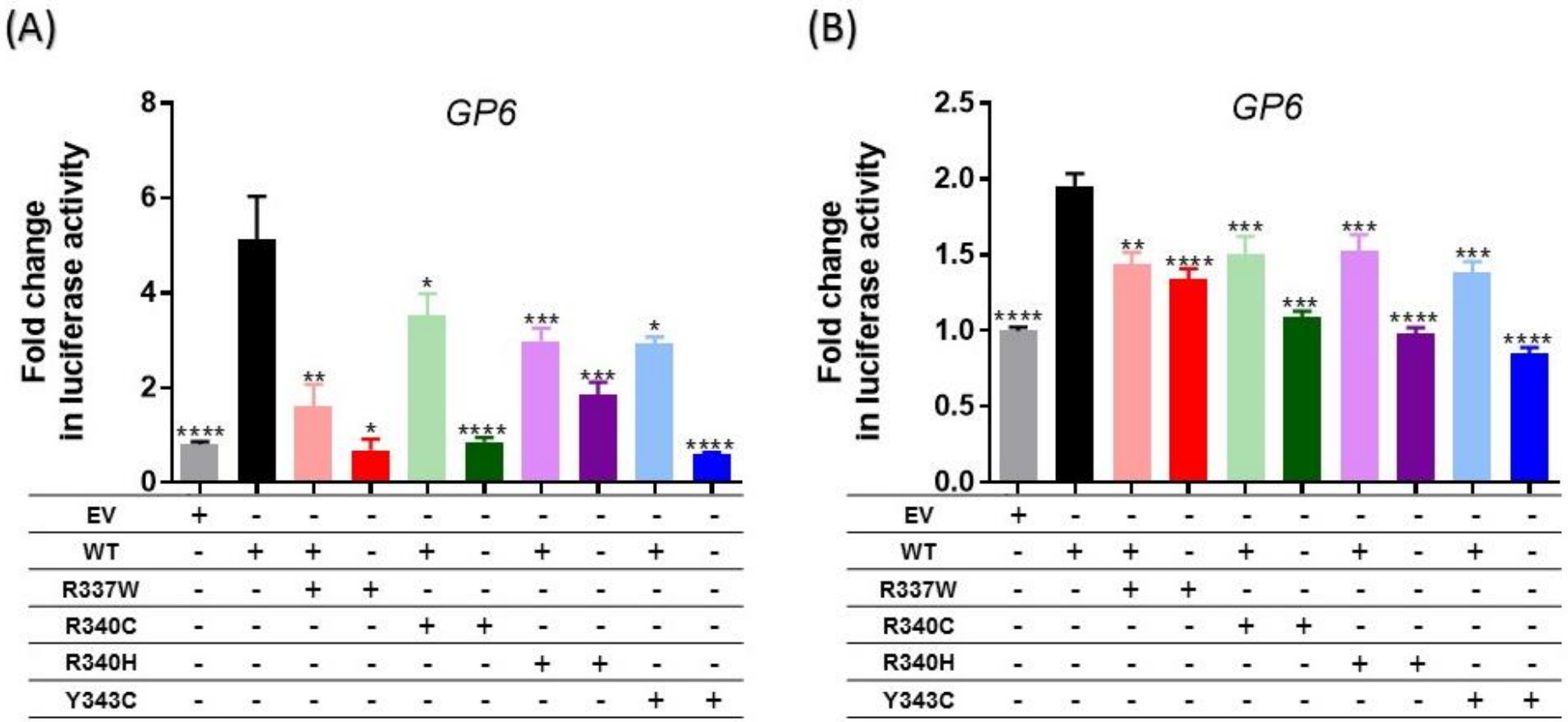


Figure 4.6 Transactivation of the *GP6* promoter by wild-type and mutated FLI1 variants

(A) HEK 293T and (B) Dami cells were transfected with wild-type (WT) or mutated FLI1 constructs, or combinations thereof, or with the empty vector (EV), in addition to pGL3-GP6-luciferase and pRLnull-Renilla reporters as described. Firefly and Renilla luciferase expression were assessed in cell lysates 48 hours later. Transcriptional activity of FLI1 variants was measured by calculating the ratio of signal from the *GP6* promoter (Firefly) compared to a control promoter (Renilla). The data are expressed as fold change in luciferase activity relative to that observed in cells transfected with EV and represent the mean ± standard error of the mean of at least three independent experiments. Paired t-tests were used for comparison, NS $p>0.05$, * $p\leq0.05$, ** $p\leq0.01$, *** $p\leq0.001$, **** $p\leq0.0001$.

Table 4.2 Comparison of the transactivation capacity of FLI1 variants in HEK 293T cells

	WT	EV	R337W+ WT	R337W	R340C+W T	R340C	R340H+W T	R340H	Y343C +WT	Y343C
Mean	5.14	0.80	1.61	0.69	3.53	0.86	2.99	1.84	2.92	0.60
% activity compared to WT-FLI1	100	---	31.32	13.42	68.68	16.73	58.17	35.80	56.81	11.67
% reduction in activity compared to WT-FLI1	---	---	68.68	86.58	31.32	83.27	41.83	64.20	43.19	88.33
p-value	---	<0.0001	0.009	0.0124	0.0348	<0.0001	0.0001	0.0007	0.0284	<0.0001
Number of repeats (each in triplicate)	---	15	3	5	4	5	4	4	3	4

For further details, see the legend to Figure 4.6. EV; empty vector, WT; wild-type.

Table 4.3 Comparison of the transactivation capacity of FLI1 variants in Dami cells

	WT	EV	R337W+ WT	R337W	R340C+W T	R340C	R340H+W T	R340H	Y343C +WT	Y343C
Mean	1.95	1.00	1.44	1.34	1.51	1.09	1.53	0.98	1.38	0.85
% activity compared to WT-FLI1	100		73.85	68.72	77.44	55.90	78.46	50.26	70.77	43.59
% reduction in activity compared to WT-FLI1	---	---	26.15	31.28	22.56	44.10	21.54	49.74	29.23	56.41
p-value	---	<0.0001	0.0025	<0.0001	0.0009	0.0002	0.0006	<0.0001	0.0009	<0.0001
Number of repeats (each in triplicate)	---	10	3	3	4	3	4	3	4	3

For further details, see the legend to Figure 4.6. EV; empty vector, WT; wild-type.

.

4.4.4 Intracellular distribution of wild-type and mutated FLI1-EGFP fusion proteins

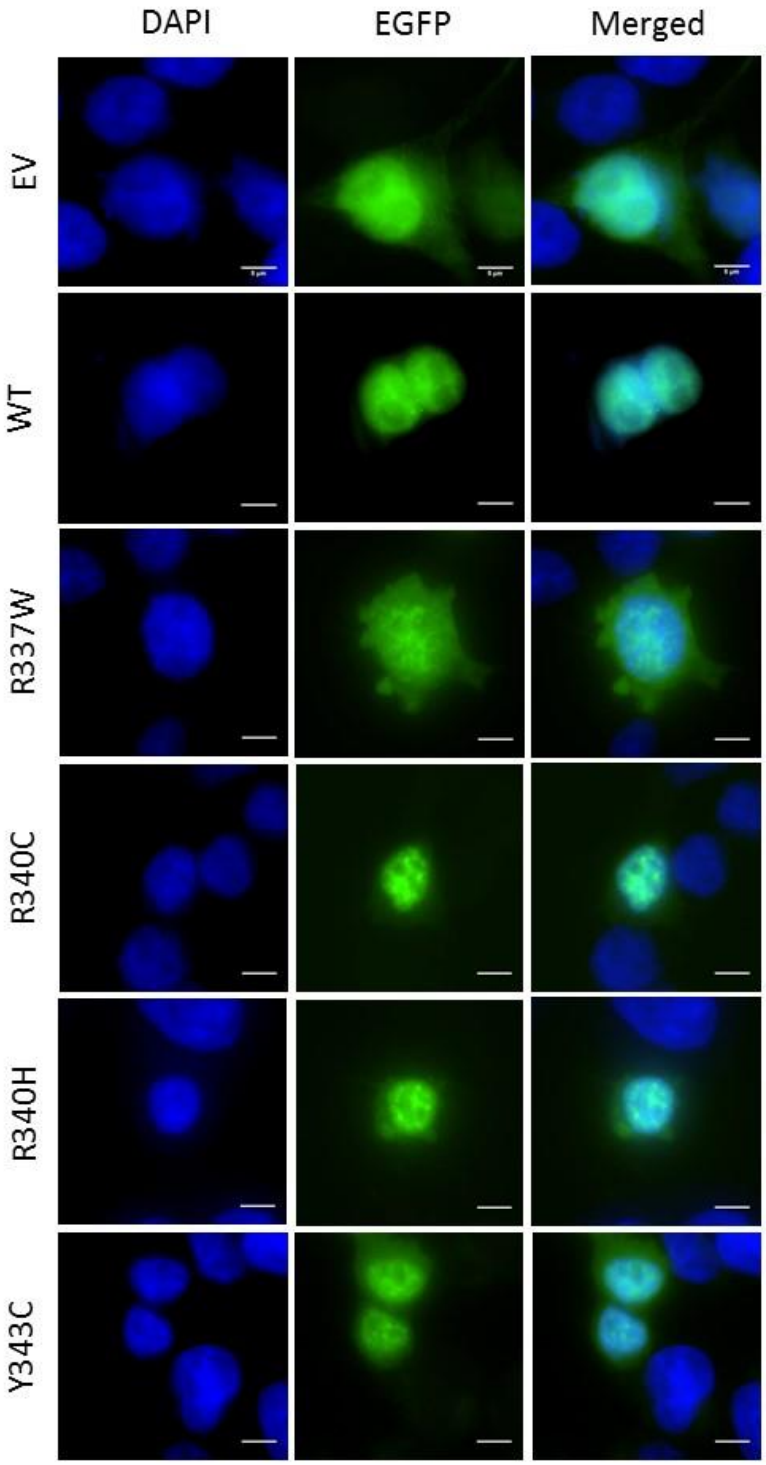
The intracellular expression of a particular gene of interest can be examined following its expression as a recombinant fusion protein formed by cloning the cDNA of the gene of interest in-frame with the cDNA for a fluorescent protein without an intermediate stop codon. This fusion results in expression of a fluorescently tagged protein that can be easily detected by microscopy in transfected mammalian cells. HEK 293T and Dami cells were transfected with pEGFP-N2-WT-FLI1 or derivative constructs encoding each of the four FLI1 variants (R337W, R340C/H, Y343C) being investigated. The localisation of FLI1 within the cells was then assessed using wide-field and super-resolution microscopy, with images being taken for quantification purposes.

4.4.4.1 Reduced nuclear localisation of EGFP-tagged FLI1 variants in HEK 293T cells

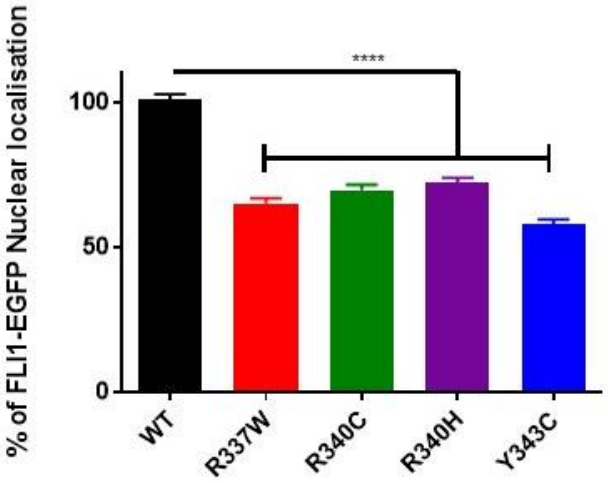
Analysis of the expression of the WT-FLI1-EGFP fusion protein in HEK 293T cells revealed it to be localised entirely to the nucleus. Compared to cells expressing the WT fusion protein, there was a significant reduction in nuclear expression of the R337W-, R340C-, R340H- and Y343C-FLI1-EGFP fusion proteins, with 69.64% of the R340C variant, and 72.17% of the R340H variant locating to the nucleus ($p<0.0001$) (Figure 4.7 A and B, Appendix 6). The R337W and Y343C fusion variants also showed reduced nuclear accumulation, 65% and 58.01% being located to the nucleus respectively ($p< 0.0001$).

In contrast to the WT-FLI1 fusion protein, which appeared to be evenly distributed in the nuclei of HEK 293T cells, the fusion proteins of the four FLI1 variants appeared to condense, forming aggregates within the nucleus, while the fraction of the variants that was located in the cytoplasm was evenly distributed. This pattern of aggregation within the nucleus was confirmed for the R340C-FLI1 variant when the cells were imaged by super-resolution microscopy (Figure 4.7 C).

(A)



(B)



(C)

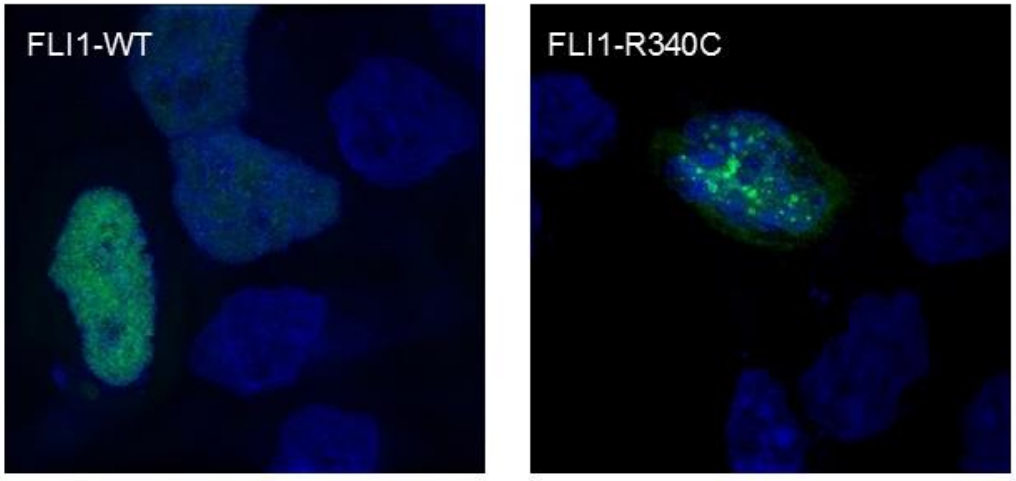


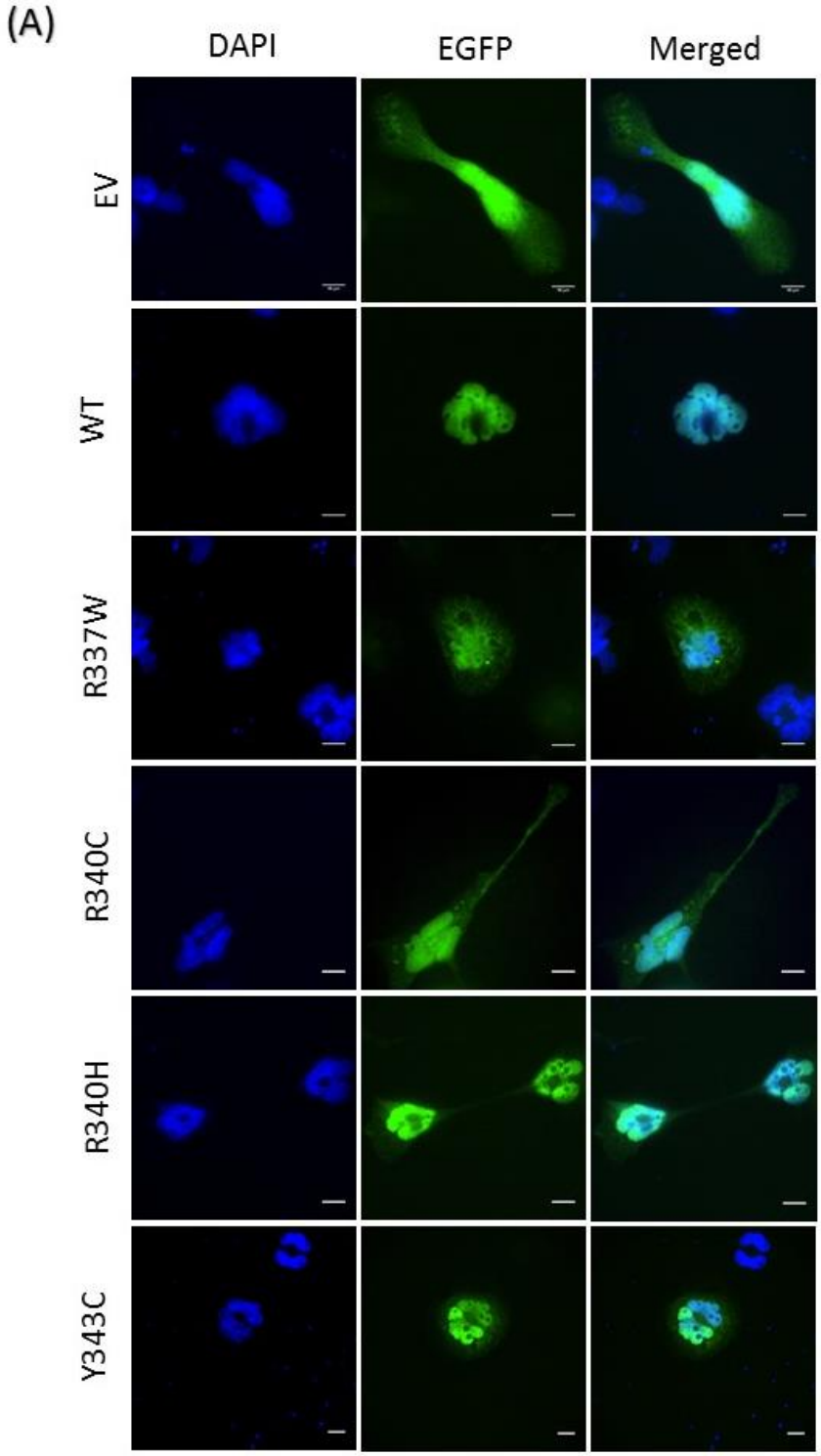
Figure 4.7 Reduced nuclear localisation of EGFP-tagged FLI1 variants in HEK 293T cells

Cells were transfected with pEGFP-N2 constructs encoding fusion proteins of wild-type (WT) or variant forms of FLI1 with enhanced green fluorescent protein (EGFP). (A) Fluorescence images captured by wide-field microscopy of HEK 293T cells transiently expressing EGFP-tagged WT, the indicated FLI1 variants or empty vector (EV). 4',6-diamidino-2-phenylindole (DAPI) was used as a nuclear counterstain and appears blue. Merged images that contain both fluorescent signals are also included. The scale bar represents 5 µm in all images. (B) The percentage of FLI1 located to the nucleus was determined in at least 20 cells per experiment (n=3). The t-test was used to compare results,****p<0.0001. (C) Super-resolution microscopy images of HEK 293T cells expressing EGFP-tagged WT-FLI1 and the R340C-FLI1 variant.

4.4.4.2 Reduced nuclear localisation of EGFP-tagged FLI1 variants in Dami cells

Analysis of the distribution of WT-FLI1-EGFP fusion protein in Dami cells revealed 78.16% of the fluorescent signal to be localised to the nucleus. Compared to cells expressing the WT-FLI1 fusion protein, there was a significant reduction in nuclear expression in cells expressing either the R340C or the R340H-FLI1-EGFP fusion

proteins, with 51.73% of the R340C variant and 45.18% of the R340H variant reaching the nucleus ($p<0.0001$) (Figure 4.8, Appendix 7). The R337W and Y343C variants also showed reduced nuclear accumulation with 44.64% and 46.88% of the signal locating to the nucleus respectively ($p<0.0001$).



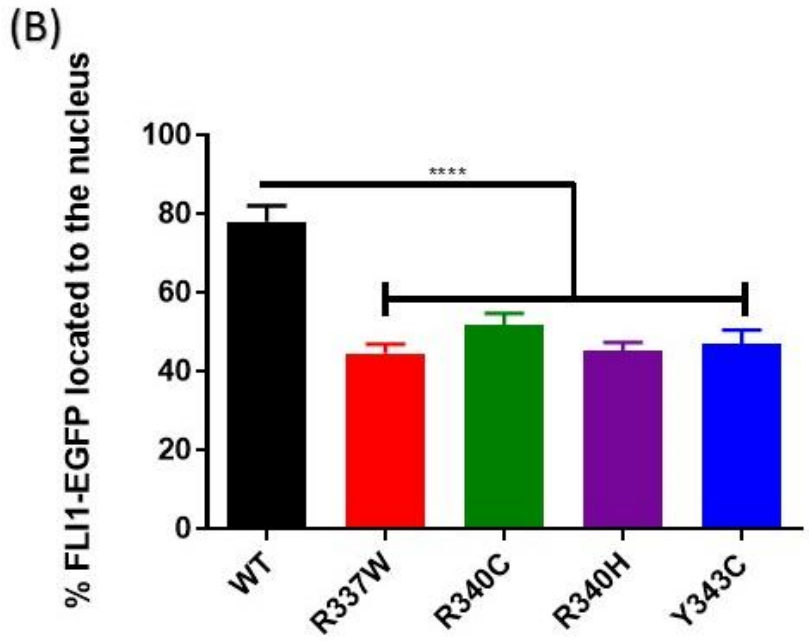


Figure 4.8 Reduced nuclear localisation of EGFP-tagged FLI1 variants in Dami cells

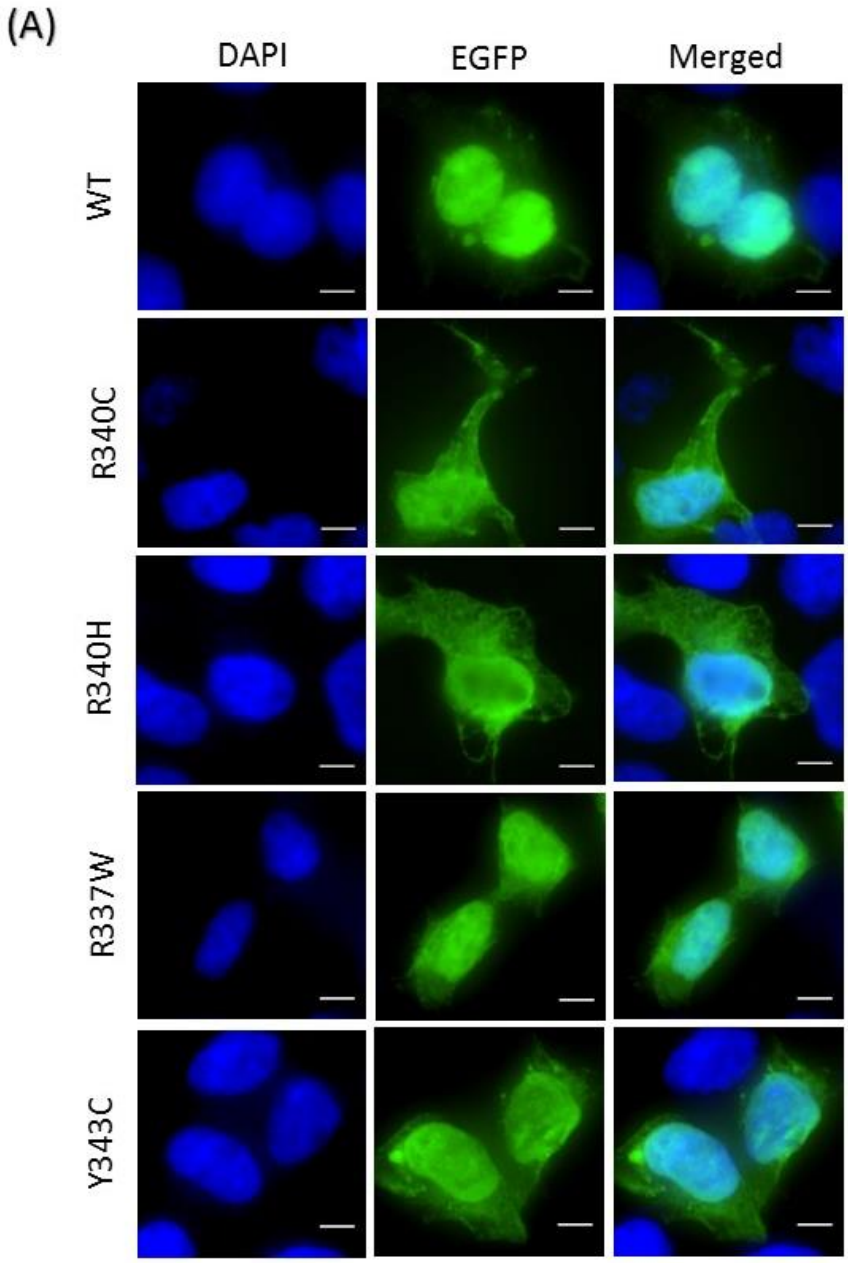
Cells were transfected with pEGFP-N2 constructs encoding fusion proteins of wild-type (WT) or variant forms of FLI1 with enhanced green fluorescent protein (EGFP). (A) Fluorescence images captured by wide-field microscopy of Dami cells transiently expressing EGFP-tagged WT-FLI1, the indicated FLI1 variants or empty vector (EV). 4',6-diamidino-2-phenylindole (DAPI) was used as a nuclear counterstain which appears blue. Merged images that contain both fluorescent signals are also included. The scale bar represents 10 µm in all images. (B) The percentage of FLI1 located to the nucleus was determined in at least 8 cells per experiment (n=3) and the t-test was used to compare results, ****p<0.0001.

4.4.5 Assessment of nuclear localisation of FLI1 variants in HEK 293T after immunological detection of FLI1

The above studies demonstrated a reduction in nuclear localisation of the EGFP-tagged FLI1 variants in HEK 293T and Dami cells. As it was possible that the EGFP tag could affect trafficking of FLI1 in these cells, the intracellular distribution of the FLI1 variants was further investigated by expressing them in HEK 293T cells, which are not known to express endogenous FLI1, and then detecting the expressed FLI1 using an anti-FLI1 antibody.

The microscope images showed that the majority of the WT-FLI1 signal originated in the nucleus, with a weak cytoplasmic signal also being observed. In contrast, all four FLI1 variants showed diffuse patterns of staining, both within the nucleus and the cytoplasm. Almost 70% of WT-FLI1 expression was localised to the nucleus compared to approximately 43% for each of the R340C, R340H and R337W variants and 47.39%

of the Y343C variant ($p<0.0001$ for comparisons between all variants and WT-FLI1) (Figure 4.9, Appendix 8).



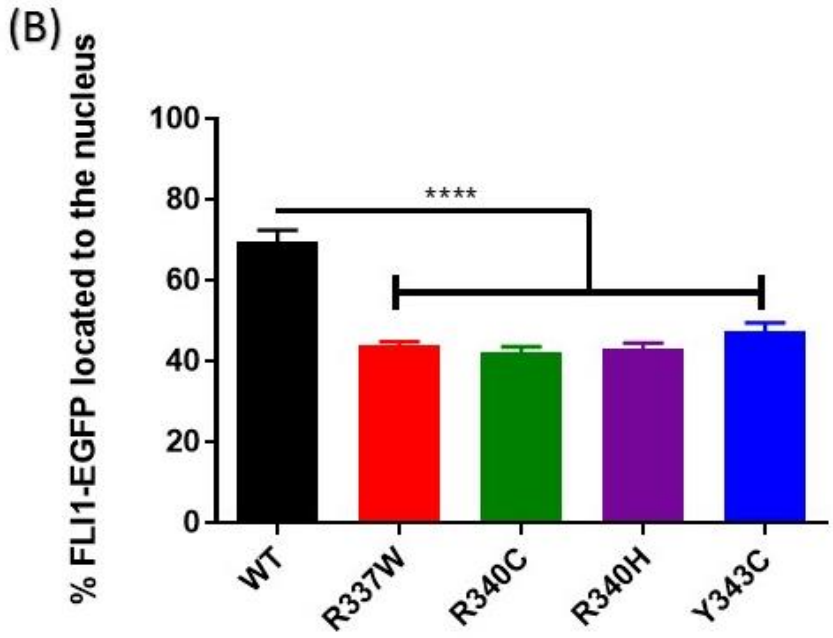


Figure 4.9 Reduced nuclear localisation of the FLI1 variants in HEK 293T cells
Cells were transfected with pSG5-WT-FLI1 or derivatives encoding the identified FLI1 variants. FLI1 was detected using a primary anti-FLI1 antibody and a secondary antibody conjugated to Alexa Fluor™ 488. (A) Fluorescence images using wide-field microscopy of HEK 293T cells transiently expressing wild-type (WT)-FLI1, the indicated FLI1 variants, or the empty vector (EV) appears green. 4',6-diamidino-2-phenylindole (DAPI) was used as a nuclear counterstain and appears blue. Merged images containing both fluorescent signals are also shown. The scale bar represents 5 µm. (B) The percentage of FLI1 located to the nucleus was determined in at least 20 cells per experiment (n=3), and the t-test was used to compare results, ****p<0.0001.

4.4.6 Assessment of the subcellular distribution of FLI1 variants in HEK 293T cells

The effects of the R337W, R340C, R340H and Y343C amino acid substitutions in FLI1 on its subcellular localisation were also investigated by electrophoresis and immunoblotting of cytoplasmic and nuclear fractions of lysates from HEK 293T cells transiently overexpressing the different FLI1 variants.

Electrophoresis and immunoblotting of fractions extracted from cells expressing WT-FLI1 revealed that FLI1 was present in both nuclear and cytoplasmic extracts. Densitometric analysis of the blots and calculation of the normalised nuclear to cytoplasmic ratios of FLI1, using the relevant housekeeping protein, revealed 5.6-fold enrichment of WT-FLI1 in the nucleus when compared with that present in the cytoplasm. All four FLI1 variants were also distributed between both cellular fractions (Figure 4.10 A). However, in contrast to WT-FLI1, the R340C, R340H and R337W-FLI1 variants were located primarily in the cytoplasm and showed significant

reductions in nuclear accumulation (ratio of FLI1 nuclear: FLI1 cytoplasmic 0.8; p=0.0286 for all three variants). The Y343C variant was also confined mostly to the cytoplasm, though it showed a slightly greater capacity to reach the nucleus (ratio of nuclear: cytoplasmic FLI1 1.32, p=0.0286) (Figure 4.10 B, Appendix ***9***).

Quantification of the immunoblotting results revealed 80.61% of WT-FLI1 to be present in the nuclear extract. In contrast, the proportion of FLI1 present in the nuclear fraction was reduced to less than 55% (range between 40.57-54.58%) in lysates from HEK 293T cells expressing each of the four FLI1 variants studied (Figure 4.10 C, Appendix 10).

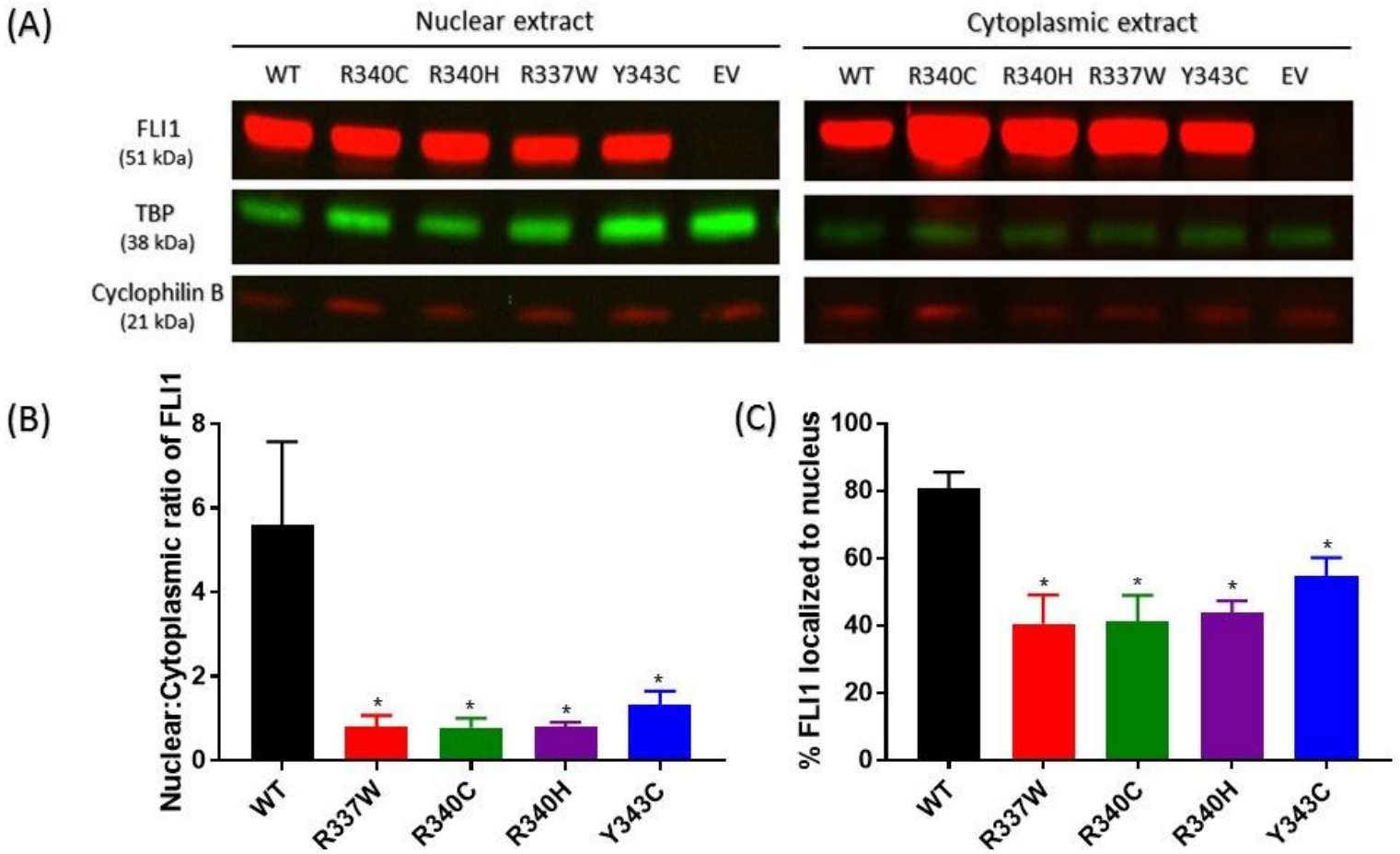


Figure 4.10 Amino acid substitutions in the ETS domain impair nuclear localisation of FLI1

(A) Detection of FLI1 in western blots of cytoplasmic and nuclear fractions from HEK 293T cells transiently overexpressing wild-type (WT)-FLI1 and the R340C-, R340H-, R337W- and Y343C-FLI1 variants. The blot was also probed for TATA-binding protein (TBP) and cyclophilin B which act as markers of nuclear and cytoplasmic fractions respectively. (B) Nuclear: Cytoplasmic ratios of FLI1 and (C) percentage of FLI1 localised to the nucleus derived from densitometric analysis of the western blots. The Mann-Whitney test was used for comparison, * $p \leq 0.05$. Data shown represent a minimum of three independent experiments.

4.5 Discussion

FLI1 belongs to the ETS family of transcription factors and is known to regulate a number of genes that are crucial for the formation of haematopoietic and vascular endothelial cells as well as regulation of megakaryopoiesis. In Jacobsen syndrome and PTS, a partial deletion of the long arm of chromosome 11, which includes *FLI1*, is associated with bleeding, thrombocytopenia, dense granule storage pool deficiency and the presence of characteristic giant α-granules in platelets (Breton-Gorius et al., 1995; Favier et al., 1993; Favier et al., 2003; Grossfeld et al., 2004; Krishnamurti et al., 2001; Penny et al., 1995; White, 2007). More recently, alterations in *FLI1* were found by our group to be enriched among patients with IPDs, which were characterised by excessive bleeding and a failure in dense granule secretion (Stockley et al., 2013). In particular, the R337W and Y343C variants of FLI1 were identified and shown to have a complete loss in transactivation capacity (Stockley et al., 2013). However, the possibility that this could be due to a failure in nuclear translocation was not explored. Since the report by Stockley et al. (2013), several other novel *FLI1* variants have been identified in patients with IPDs (Poggi et al., 2015; Saultier et al., 2017; Stevenson et al., 2015) (Figure 4.1, Table 4.4).

In this chapter, two novel variants of *FLI1*, both predicting substitution of arginine 340 in the ETS domain, but which are associated with different clinical features in the affected patients, were investigated to assess their possible pathogenic effects and association with bleeding symptoms. *In silico* and *in vitro* expression studies were used to predict their DNA binding capability, assess their transactivation capacity and investigate their effects on nuclear translocation of FLI1. The R337W and Y343C variants of FLI1, which had been previously identified by our group, were included in these studies. The c.1018C>T defect in *FLI1* which predicts the p.R340C substitution was identified in a patient enrolled in the UK-GAPP study who had been diagnosed with storage pool disorder, and whose platelets showed a reduction in dense granule secretion. He also had thrombocytopenia at birth, though this later resolved. There was a history of bleeding in the family affecting both the mother and the sister of the patient. Interestingly, while he did not have a bleeding history, he developed arthritis at the age of 18 years. The c.1019G>A transition in *FLI1,* which predicted the p.R340H substitution, was identified in three members of a family referred for investigation of a

bleeding history, all of whom had a diagnosis of storage pool deficiency characterised by a reduction in dense granule secretion. They also suffered from eczema and mild thrombocytopenia. Two of the affected family members had a history of recurrent infections, while the third had a history of mild neutropenia.

Arginine 337, arginine 340 and tyrosine 343 are located within the highly conserved ETS DNA binding domain of FLI1. This region also encompasses one of two NLSs that are known to occur in FLI1. Structural modelling predicted that these three amino acids all directly contact DNA. Furthermore, substitution of either R340 or Y343 by a cysteine residue was predicted to entirely abolish the interaction of FLI1 with DNA and the interactions of these amino acids with other residues in FLI1. In contrast, substitution of R340 by histidine or of R337 by tryptophan was predicted to disrupt intramolecular interactions within FLI1, while maintaining the interaction of these variants with DNA. It is also possible that the loss of intramolecular interactions as a result of the amino acid substitutions at positions 337, 340 and 343 could, in turn, cause conformational changes in the ETS domain that affect its ability to bind to its target DNA sequence. This was supported by previous work which showed that substitution of either arginine 337 or 340 by alanine, lysine, aspartic acid, asparagine, or glutamic acid abolished the ability of FLI1 to bind to DNA when tested using electrophoretic mobility shift assays (Hu et al., 2005; Liang et al., 1994). In addition arginine 324 that is not involved in any direct interactions with DNA, however, it provides structural stability to the H2 helix and aids in positioning the loop between the H2 and H3 helices by forming a salt bridge with aspartate 289 (Hou & Tsodikov, 2015). It is noteworthy that disruption of this salt bridge as a result of the c.970C>T transition in *FLI1* that predicted substitution of arginine 324 by tryptophan was identified in patients who presented with an IPD that mimicked PTS (Stevenson et al., 2015).

The effects of the amino acid substitutions at R337, R340 and Y343 were assessed *in vitro* by examining the transactivation capacity of the corresponding FLI1 variants and their ability to undergo nuclear translocation in two mammalian cell lines. Thus, HEK 293T cells, which do not express FLI1, were used for initial investigations, with further work being undertaken in Dami cells derived from the blood of a patient with megakaryoblastic leukaemia, an established model for MK/platelets (Briquet-Laugier et al., 2004; Greenberg et al., 1988; Lev et al., 2011).

The transactivation capacity of FLI1 variants was assessed by investigating their ability to bind to an ETS binding site in the promoter of a known FLI1 target gene, *GP6*, using a dual luciferase reporter assay. All four variants, R337W, R340C/H and Y343C, showed significant reductions in their ability to transactivate the *GP6* promoter, both in HEK 293T and Dami cells. Furthermore, co-expression of each of the four variants with WT-FLI1 in order to mimic the heterozygous situation in the patients also led to significant reductions in transcriptional activity. These results confirm the earlier findings for the R337W and Y343C variants (Stockley et al., 2013), suggesting that the R340C and R340H variants could similarly contribute to the bleeding tendency observed in patients by causing a reduction in transcription of MK-specific genes. Although the modelling predicts a greater impact of the R340C and Y343C substitutions on the interaction of FLI1 with DNA, all four variants appear to have comparable transactivation capacity. This is perhaps not surprising given the similar reductions in transcription capacity of other FLI1 variants described while this work was ongoing, which also have substitutions within the ETS domain, the activities of which have been studied by luciferase assays using a variety of promoter sequences (Saultier et al., 2017; Stevenson et al., 2015). The failure to distinguish between the variants in terms of their ability to interact with DNA to promote transcription may reflect the blunt nature of the luciferase assay as a tool to distinguish subtle differences in ability to promote transcription, or indeed could be a reflection of the necessity for the integrity of the ETS domain of FLI1. Interestingly, the R324W variant of FLI1 demonstrated a reduced ability to drive the expression of several target genes, despite predictions that the arginine 324 does not directly contact DNA (Hou & Tsodikov, 2015; Stevenson et al., 2015). Overall, however, the results of the reporter gene assays suggest that the R337W, R340C/H and Y343C substitutions disrupt the binding of FLI1 to DNA, therefore the ability of FLI1 to activate MK-specific genes.

One of the factors that could contribute to the observed reduction in transactivation of FLI1 variants, is an impaired ability to translocate to the nucleus. The effects of the amino acid substitutions at R337, R340 and Y343 on nuclear accumulation of FLI1 were therefore assessed in HEK 293T and Dami cells. There was a 30-40% reduction in nuclear accumulation of each of the FLI1 variants following their overexpression as EGFP-tagged proteins in both HEK 293T and Dami cells. Similar results were obtained both by microscopic analysis of the cellular distribution of FLI1 and densitometric

analysis of western blots of nuclear and cytoplasmic extracts of HEK 293T cells expressing the non-GFP-tagged FLI1 variants. Interestingly, in contrast to WT-FLI1, each of the four FLI1 variants displayed a speckled distribution pattern in the nuclei of HEK 293T cells, suggesting that they may be interacting abnormally with other nuclear proteins and/or DNA to form nuclear aggregates. However, it is also possible that the EGFP tag interfered with trafficking of the FLI1 variants in HEK 293T cells. The absence of the speckled pattern of FLI1 distribution following expression of the FLI1 variants without GFP tags in HEK 293T cells would support this possibility. The presence of FLI1 in the nucleus, albeit at reduced levels, can be partly explained by the presence of a second intact NLS (NLS1) that is known to independently direct FLI1 to the nucleus (Hu et al., 2005). These findings also agree with those of Hu et al. (2005) who showed that substitution of either arginine 340 or arginine 337 with alanine inhibited accumulation of FLI1 in the nucleus as a result of a loss of the interaction between NLS2 and the nucleocytoplasmic shuttling receptors (Hu et al., 2005). Similarly, other FLI1 variants having amino acid substitutions in the ETS domain, such as R337Q and K345E, demonstrated an increased cytoplasmic accumulation of FLI1 (Saultier et al., 2017).

Overall, the reduction in transcriptional activity and nuclear accumulation of the R340C and R340H-FLI1 variants support the hypothesis that these variants interfere with normal FLI1 regulation of essential MK genes and are likely to explain the bleeding tendency in affected members of families A and B.

To date, the affected members of families with inherited defects in *FLI1* have all been reported to present with a bleeding tendency and a platelet disorder that was characterised predominantly by a defect in platelet granule secretion, though they have also shown variation in phenotypic expression of the platelet disorder, also in whether or not additional clinical features have been present (Poggi et al., 2015; Saultier et al., 2017; Stevenson et al., 2015; Stockley et al., 2013). Thus, while the predominant abnormality in individuals carrying the R340C and R324W variants was a defect in platelet dense granule secretion, the R340H, R337W and Y343C variants were also associated with other clinical features including recurrent infections, eczema, psoriasis and alopecia. In addition to MKs, *FLI1* is expressed in endothelial cells and other haematopoietic lineages and there is also evidence supporting a role for FLI1 in

leukocyte differentiation (Masuya et al., 2005; Mélet et al., 1996; Starck et al., 2010; Suzuki et al., 2013; Zhang et al., 1995; Zhang et al., 2008). It is therefore possible that the additional clinical features observed in patients with *FLI1* defects are due, at least partly, to abnormal regulation of target genes in these cell types. Similarly, the transient monoallelic expression of *FLI1* during early megakaryopoiesis may explain the variation in platelet phenotype observed among different affected members of the same family (Raslova et al., 2004).

Further work is required to explain the platelet granule abnormalities associated with *FLI1* defects. Electron microscope images of platelets from patients with *FLI1* variants have revealed the presence of giant α-granules that resemble those seen in platelets from patients with PTS (Saultier et al., 2017; Stevenson et al., 2015). They also show a dramatic reduction in dense granules, the accumulation of glycogen in vacuoles in the cytoplasm and the existence of autophagosome-like structures in a subpopulation of platelets (Saultier et al., 2017). Given the role of FLI1 in regulating megakaryocytic genes, and the association between *FLI1* gene defects and platelet granule abnormalities, further investigation of the genes that are regulated by FLI1 in MKs and in platelets from patients with *FLI1* defects may yield insights into the mechanisms governing platelet granule biogenesis and secretion, identifying novel genes defects which may be associated with IPDs. This hypothesis will be the focus of the work described in chapter 6 of this thesis.

Table 4.4 Germline *FLI1* variants associated with inherited platelet bleeding disorders

FLI1 defect	Location within the protein	Mode of inheritance	Bleeding	Platelet count	Platelet size (MPV)	Platelet phenotype	Other clinical feature
p.R144Q[(1)]	PNT	---	No	Low	---	---	---
c.970C>T: p.R324W [(2)]	ETS (Not in contact with DNA)	Hom	Yes	Low	Increase	Defect in dense granule secretion; 4% of platelets displayed fused α-granules	No other features
c.992-995del; p.Asn331Thrfs*4 [(3)]	ETS	Het	Yes	Normal/ Low	Normal	Defect in dense granule secretion	---
c.1009 C>T: p.R337W [(3)]	ETS	Het	Yes	Normal	Normal	Defect in dense granule secretion; Presence of fused α-granules	Eczema; Recurrent viral infections; Psoriasis; Alopecia
c.1010G>A: p.R337Q [(4)]	ETS	Het	No	Low	Increase	Absence of dense granules; 25% of patients' platelets displayed giant α-granules	No other features
c.1018C>T: p.R340C	ETS	Het	Yes/No	Low at birth/ Normal	---	Storage pool disease; Defect in dense granule secretion	Arthritis
c.1019G>A: p.R340H	ETS	Het	Yes	Low	---	Dense granule storage pool disease	Eczema; Recurrent infections
c.1028 A>G: p.Y343C [(3)]	ETS	Het	Yes	Low	Normal	Defect in dense granule secretion; Presence of fused α-granules	Infective endocarditis; Eczema; Colitis; Alopecia
c.1033A>G: p.K345E [(4)]	ETS	Het	Yes	Low	Increase	Absence of dense granules; 29% of patients' platelets displayed giant α-granules	No other features

(1) Poggi et al., 2015; (2) Stevenson et al., 2015; (3) Stockley et al., 2013; (4) Saultier et al., 2017. ETS; ETS DNA binding domain, Het; heterozygous, Hom; homozygous, MPV; mean platelet volume, PNT; pointed N-terminal domain.

5 Chapter 5. Characterisation of a Novel Nonsense Variant in *ETV6* Identified in a Patient with a Platelet Secretion Disorders

5.1 Introduction

5.1.1 Discovery of ETV6

ETV6 was first identified as a novel ETS-like gene on chromosome 12 that was disrupted by a t(5;12) translocation in a subgroup of patients with chronic myelomonocytic leukaemia (Golub et al., 1994). This translocation was found to result in a chimeric transcript that fuses the N-terminal domain of ETV6 to the tyrosine kinase domain of platelet-derived growth factor receptor β (Golub et al., 1994). Subsequently, ETV6 was shown to form part of chimeric proteins that have been identified in many haematological (De Braekeleer et al., 2012) and non-haematological malignancies (Brenca et al., 2016; Knezevich et al., 1998a; Knezevich et al., 1998b; Leeman-Neill et al., 2014; Tognon et al., 2002) and about 30 *ETV6* partner genes have been identified (De Braekeleer et al., 2012).

5.1.2 The *ETV6* gene

ETV6, also known as *TEL* (Translocation Leukaemia) or *TEL1* (ETS-Related Protein Tel1) or *THC5* (Thrombocytopenia 5), was assigned to the E26 transformation-specific or E-twenty-six (ETS) family of transcription factors based on its amino acid homology with other members of this family (Golub et al., 1994) and to the "Yan" subfamily, which is the only ETS subfamily in which all members (ETV6, ETV7 and yan) are repressors (Hollenhorst et al., 2011; Lopez et al., 1999; Poirel et al., 2000). The *ETV6* gene has been localised to chromosome position 12p13.1 and spans 250 kilobase (Kb) of genomic DNA. The *ETV6* transcript (NM_001987) is approximately 6 Kb in size, comprising 8 exons that are translated to yield a 53 kilodalton (kDa) protein comprising 452 amino acids (NP_001978).

5.1.3 The ETV6 protein

ETV6 is expressed in multiple tissues including heart, brain, placenta, lung, liver, skeletal muscle, kidney and pancreas (Golub et al., 1994). In addition to the ETS domain (residues 338-424), ETV6 has three important functional domains: the pointed N-terminal domain (PNT; residues 56-123), the linker inhibitory damper domain (LID; residues 124-337) and the C-terminal inhibitory domain (CID; residues 423-452) (Green et al., 2010) (Figure 5.1).

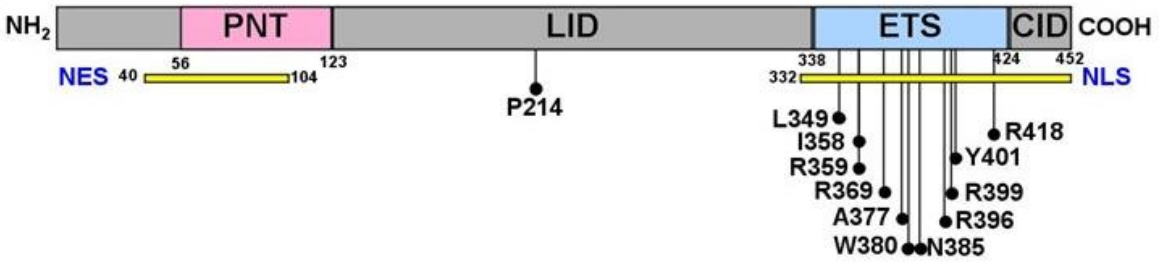


Figure 5.1 Human ETV6 protein domain structure
The domain structure of the ETV6 protein (NP_001978) and the locations of amino acid substitutions reported in patients with inherited platelet bleeding disorders are indicated. p.P214L (Zhang et al., 2015, Noetzli et al., 2015, Melazzini et al., 2016, Poggi et al., 2017); p.L349P (Topka et al., 2015); p.I358M (Poggi et al., 2017); p.R359* (Moriyama et al., 2015); p.R369W (Melazzini et al., 2016); p.R369Q (Zhang et al., 2015); p.A377T (Poggi et al., 2017); p.W380R (Melazzini et al., 2016); p.N385fs (Topka et al., 2015; Melazzini et al., 2016); p.R396G (Poggi et al., 2017); p.R399C (Zhang et al., 2015); p.Y401N (Poggi et al., 2017); p.Y401H (Poggi et al., 2017); p.R418G+p.N385Vfs*7p.N 385_418del (Noetzli et al., 2015, Melazzini et al., 2016). CID; C-terminal inhibitory domain, ETS; ETS DNA binding domain, LID; Linker inhibitory damper domain, NES; nuclear export signal, NLS; nuclear localisation signal, PNT; pointed N-terminal domain.

5.1.3.1 The ETS domain of ETV6

Like other members of the ETS family, ETV6 possesses a conserved ETS domain that preferentially binds T(C/G/T)(A/C)**GGA**AGT sequences in the regulatory elements of target genes (Buijs et al., 2000; Szymczyna & Arrowsmith, 2000). Residues in the ETS domain, as well as the adjacent CID domain, have also been shown to form a nuclear localisation signal (Park et al., 2006) (section 5.1.3.5) and to aid in recruitment of proteins that modulate the transcription repression activity of ETV6 (Nordentoft & Jørgensen, 2003; Roukens et al., 2010).

5.1.3.2 The pointed N-terminal domain of ETV6

The PNT domain of ETV6 acts primarily as a homo- and hetero-oligomerisation domain (Kwiatkowski et al., 1998; Poirel et al., 2000). It also plays a role in determining the subcellular distribution of ETV6 as it contains part of the nuclear export signal (Wood et al., 2003) (section 5.1.3.5).

ETV6 was found to partner with itself through the PNT in normal haematopoietic cell development (Kwiatkowski et al., 1998). However, the functional relevance of ETV6 dimerisation was not realised until monomeric ETV6 was shown to be inactive and unable to repress transcription of the known ETV6 target, *MMP3* (Matrix Metallopeptidase 3) (Wood et al., 2003). Monomeric ETV6 has been shown to bind to

its target DNA sequence with ten-fold lower affinity than the dimeric protein (Green et al., 2010). In addition, ETV6 dimerisation was found to stabilise binding of the ETS domain to its target sequence in DNA by overcoming the autoinhibition effect of the CID (Green et al., 2010).

ETV6 also forms heterodimers through its PNT domain with other ETS proteins including FLI1 and ETV7. While the functional significance of ETV6-ETV7 dimerisation remains to be determined (Poirel et al., 2000; Potter et al., 2000), ETV6-FLI1 dimerisation has been shown to suppress the transactivation capacity of FLI1 (Kwiatkowski et al., 1998). ETV6 has also been reported to heterodimerise through its PNT with other proteins such as UBE2I (Ubiquitin-conjugating enzyme E2 I) (Chakrabarti et al., 1999; Chakrabarti et al., 2000), and SIN3A (SIN3 transcription regulator family member A) (Fenrick et al., 1999) and these interactions modulate the transcriptional activity of ETV6 and its subcellular localisation.

5.1.3.3 The C-terminal inhibitory domain of ETV6

Self-regulatory capacity is a feature of several ETS family members which is well characterised for ETV6 (Coyne et al., 2012; Green et al., 2010). The CID of ETV6 is formed by two helices, H4 and H5, which are in close contact with the three helices that form the ETS domain (Coyne et al., 2012). Deletion of the CID caused a 10-fold increase in affinity of the ETV6 monomer for DNA to a level that was similar to the DNA affinity of the wild-type (WT) ETV6 dimer (Green et al., 2010). In another study, a C-terminal fragment of ETV6 that contained both the ETS and CID domains (without the LID), was found to repress DNA binding by the ETS domain by approximately 50-fold compared to the ETS domain alone (Coyne et al., 2012).

The CID domain regulates ETV6 function through steric hindrance of the ETS domain. The H5 helix is packed anti-parallel to H3 helix, preventing it from entering the DNA major groove (Coyne et al., 2012). Furthermore, H4 and H5 helices are stacked against an N-terminal portion of H1 helix and some residues in the H2-H3 turn, which prevents these helices from contacting the negatively charged phosphodiester backbone of DNA that is essential for the ETS domain to fall into the DNA minor groove (Coyne et al., 2012). Additionally, the dynamic flexibility of the ETS domain, which is necessary for DNA binding, is found to be quenched in ETV6 by the CID (Coyne et al., 2012).

The two amphipathic helices H4 and H5 have been shown to be only marginally stable (Coyne et al., 2012) and a conformational equilibrium between inhibited and uninhibited structures of ETV6, which is determined by the folding and unfolding of H5 helix, is suggested to regulate ETV6 function. When H5 is folded, ETV6 has an inhibited conformation, but when it unfolds, ETV6 adopts a flexible unblocked conformation in which the CID is displaced and the ETS domain is exposed for DNA binding (Green et al., 2010). This model is supported by the finding that both of the negatively charged glutamate residues at positions 431 and 434 in murine ETV6 (equivalent to glutamine residues 435 and 438 in human ETV6) are essential for the conformational equilibrium of the CID domain (Green et al., 2010).

5.1.3.4 The central domain / linker inhibitory damper domain of ETV6

The linker inhibitory damper (LID) domain is located between the PNT and the ETS domains and it interferes with the full inhibitory potential of the CID through its interaction with H5 (Green et al., 2010). ETV6 lacking the LID showed a 50 to 100-fold lower affinity for the target DNA sequence than the full-length ETV6 dimer (Green et al., 2010). The dampening of CID autoinhibition by the LID has been suggested to be due to competition between the CID-LID and CID-ETS interactions (Green et al., 2010). Similar to the PNT, the LID is involved in the recruitment of other co-repressors (Chakrabarti & Nucifora, 1999; Guidez et al., 2000; Wang & Hiebert, 2001) that modulate ETV6 repressor function.

5.1.3.5 The nuclear localisation and export signals of ETV6

ETV6 is mainly located in the nucleus (Poirel et al., 1997), and a nuclear localisation signal (NLS) has been identified between residues 332 and 452 in the C-terminal region of the protein (Park et al., 2006). A small fraction of ETV6 is found in the cytoplasm (Van Rompaey et al., 1999; Wood et al., 2003) and a nuclear export signal (NES) has been located between residues 40 and 104 (Wood et al., 2003). Post-translational modification of specific residues within the PNT domain has been shown to play a role in determining the subcellular distribution and function of ETV6 (Chakrabarti et al., 2000; Wood et al., 2003).

5.1.4 Function of ETV6

ETV6 is one of a small subset of ETS family members that function primarily as transcriptional repressors (Hollenhorst et al., 2011; Lopez et al., 1999). It has also been described as a putative tumour suppressor (Fenrick et al., 2000; Irvin et al., 2003; Van Rompaey et al., 2000).

Animal models have shown that ETV6 is essential for embryonic development, angiogenesis and hematopoiesis. *Etv6*$^{-/-}$ mice are embryonic lethal between E10.5 and E11.5, with embryos at this stage displaying growth retardation, defects in yolk sac angiogenesis and patches of apoptosis in the neural tube and mesenchymal tissues while hematopoiesis appears unaffected (Wang et al., 1997). The importance of Etv6 in haematopoietic stem cell production was demonstrated in Xenopus embryos, where its depletion resulted in the failure of the dorsal aorta to be specified as an artery, subsequently preventing the emergence of the first haematopoietic stem cell (Ciau-Uitz et al., 2010). A mouse model generated from *Etv6*$^{-/-}$ embryonic stem cells highlighted the essential role of ETV6 in the shift of hematopoiesis from foetal liver to bone marrow (Wang et al., 1998). The importance of ETV6 for the survival of adult haematopoietic stem cells was revealed in a murine conditional *Etv6* knockout model (Hock et al., 2004). Thus, no defects were observed for any committed lineage, apart from megakaryopoiesis which showed a five-fold increase in megakaryocyte (MK) colony-forming cells, but an approximate 50% reduction in platelet counts suggesting a defect in the later stages of megakaryopoiesis (Hock et al., 2004). Additionally, morpholino-mediated knockdown of *etv6* in zebrafish highlighted the importance of Etv6 for angiogenesis (Roukens et al., 2010) and for the development of multiple haematopoietic lineages (Rasighaemi et al., 2015). Thus, Etv6 has been shown to be essential for sprouting and branching of endothelial cells, also in haematopoietic cell differentiation and lineage maturation (Rasighaemi et al., 2015; Roukens et al., 2010).

Overexpression of ETV6 in a human leukaemia cell line led to accumulation of transcripts which were specific to erythroid differentiation following erythropoietin treatment, and to a reduction in transcripts which were specific to megakaryocytic differentiation following thrombopoietin treatment (Takahashi et al., 2005). This finding indicates that ETV6 controls the fate of common progenitors of the human erythroid

and MK lineages, promoting differentiation of the erythroid lineage and hindering megakaryocytic differentiation (Takahashi et al., 2005).

5.1.5 Role of ETV6 in disease

ETV6 was initially recognised in patients with chronic myelomonocytic leukaemia characterised by a t(5;12) translocation that resulted in the fusion of the N-terminal region of ETV6 encompassing the PNT and central domains with the tyrosine kinase domain of platelet-derived growth factor receptor β (Golub et al., 1994). Somatic mutations in *ETV6* have been described in many haematological malignancies (De Braekeleer et al., 2012) and are associated with poor overall survival (Bejar et al., 2011). Non-haematological malignancies including congenital fibrosarcoma (Knezevich et al., 1998b), gastrointestinal stromal tumours (Brenca et al., 2016), breast cancer (Tognon et al., 2002), renal neoplasms (Knezevich et al., 1998a) and thyroid cancer (Leeman-Neill et al., 2014) have also been associated with somatic *ETV6* rearrangement. Although the majority of genetic rearrangements that include ETV6 in the oncogenic chimaeras involve the PNT, rearrangements involving the ETS domain have also been described (Buijs et al., 2000).

Early in 2015, Zhang et al. (2015) described three germline *ETV6* mutations in familial thrombocytopenia and haematologic malignancy. Soon after, while the present study was ongoing, further germline *ETV6* defects were reported, representing a new form of autosomal dominant thrombocytopenia called “thrombocytopenia 5 or ETV6-related thrombocytopenia”, which is characterised by a reduced number of normal sized platelets in the circulation, and a predisposition to malignancy, similar in clinical presentation to that of patients with *RUNX1* (Runt Related Transcription Factor 1) or *ANKRD26* (Ankyrin Repeat Domain 26) defects. To date, six different studies describing 18 families with germline *ETV6* mutations in patients with inherited platelet bleeding disorders (IPDs) have been reported (Melazzini et al., 2016; Moriyama et al., 2015; Noetzli et al., 2015; Poggi et al., 2017; Topka et al., 2015; Zhang et al., 2015).

5.2 Hypothesis and Aims

The work described in chapter 3 of this thesis focused on 34 patients who had been diagnosed with IPDs that were characterised phenotypically by defects in either dense granule secretion or in Gi-signalling, all of whom were recruited for investigation by the UK Genotyping and Phenotyping of Platelets (UK-GAPP) study. Whole exome sequencing (WES) and downstream bioinformatic analysis, identified a novel nonsense *ETV6* variant (NM_001987:c.1288C>T:p.Arg430*) in participant F4.1 who had been diagnosed with a platelet secretion disorder. This *ETV6* defect was predicted to be deleterious using the Combined Annotation Dependent Depletion (CADD) tool, and was assigned a PHRED score of 45. It was, therefore, a strong candidate for the causative defect underlying the bleeding disorder in patient F4.1. However, prior to this study, the predominant clinical phenotype of patients with platelet disorders in whom *ETV6* variants had been identified was thrombocytopenia with a predisposition to malignancy (Noetzli et al., 2015; Topka et al., 2015; Zhang et al., 2015), in contrast to patient F4.1 in whom the predominant clinical feature was a reduction in platelet dense granule secretion. Given the atypical phenotypic expression of the *ETV6* variant identified in F4.1, further studies were warranted to examine how it differs in its properties from previously characterised *ETV6* variants.

The work described in this chapter explored the hypothesis that the c.1288C>T:p.Arg430*-ETV6 variant is the pathogenic defect underlying the bleeding tendency in F4.1. As patient F4.1 and their family members were unavailable for further studies, a combination of *in silico* and *in vitro* methods were used to achieve the following aims:

i. Model and predict the potential effects of the c.1288C>T *ETV6* variant
ii. Examine the expression of the p.R430*-ETV6 variant *in vitro* and evaluate its ability to translocate to the nucleus
iii. Assess the transactivation capacity of the p.R430*-ETV6 variant

5.3 Methods

5.3.1 Predicting the effects of the c.1288C>T transition on splicing of the *ETV6* RNA

The impact of the NM_001987:c.1288C>T transition in *ETV6* on RNA splicing was predicted using Human Splicing Finder (Version 3.1) (http://www.umd.be/HSF3/) [accessed 2018].

5.3.2 Modelling the truncated p.R430*-ETV6 variant

The location of R430 in ETV6 was visualised by modelling the C-terminal region of ETV6 (NP_001978) (residues 339-451) using the web-based SWISS-MODEL tool (https://swissmodel.expasy.org/) [accessed 2018], and the structure for *Mus musculus* ETV6 (PDB ID: 2lf8.1) (Coyne et al., 2012) as a template. The structure was visualised using Pymol (version 0.99rc6).

5.3.3 Assessment of the subcellular localisation of overexpressed ETV6 variants in HEK 293T cells using lithium dodecyl sulphate polyacrylamide gel electrophoresis and western blotting

The expression and subcellular localisation of ETV6 variants was evaluated following electrophoresis and immunoblotting of cytoplasmic and nuclear fractions of lysates from HEK 293T cells that were transiently overexpressing the different ETV6 variants as described earlier (section 2.2.13).

5.3.4 Assessment of the transcriptional activity of overexpressed ETV6 variants using dual-luciferase reporter assays in mammalian cell lines

Dual luciferase reporter assays were conducted to examine the ability of ETV6 variants to transrepress the *MMP3* and *PF4* (Platelet factor 4) promoters in HEK 293T and megakaryocytic Dami cells as described earlier (section 2.2.11).

5.4 Results

5.4.1 The clinical phenotype of index case F4.1

Index case F4.1 is a female with a history of excessive cutaneous bruising at exposed sites. At 13 years, she had a bleeding episode that lasted for 10 days after dental extraction, which required packing and suturing. She suffered from menorrhagia after menarche, and experienced post-partum bleeding after at least one of her three deliveries. At 25 years, she had an ovarian cyst removed and the surgeon reported excessive bleeding. Treatments for her bleeding included iron therapy, emergency treatment with blood products and hysterectomy. On recruitment to the UK-GAPP study, her platelet count and mean platelet volume were within the normal range. However, assessment of platelet aggregation and ATP secretion in response to a range of agonists by lumiaggregometry revealed a profound reduction in ATP secretion (Figure 5.2, Data provided by Dr Neil Morgan, University of Birmingham).

5.4.2 Confirmation of a c.1288C>T transition in *ETV6* in index case F4.1

WES analysis identified a novel heterozygous transition in *ETV6* (NM_001987:c.1288C>T) in participant F4.1 (section 3.4.3) that was confirmed by PCR amplification and direct sequencing of a DNA fragment spanning the candidate defect (Figure 5.3).

5.4.3 Predicted effect of the c.1288C>T transition on *ETV6* splicing

The c.1288C>T transition was not predicted to introduce a new splice site that would affect processing of the *ETV6* RNA. However, it was predicted to introduce an exon splice enhancer site, though this was not expected to have an impact on splicing. The output from the Human Splicing Finder tool is shown in Figure 5.4.

5.4.4 Modelling the truncated p.R430*-ETV6 variant

The arginine residue at amino acid position 430 in ETV6 was visualised by modelling the C-terminal region of human ETV6 using the structure for murine ETV6 as a template. Clearly, truncation of ETV6 at residue 430 results in loss of the remaining 22 amino acids of the protein, which encompass all of H5 helix in the CID domain (Figure 5.5).

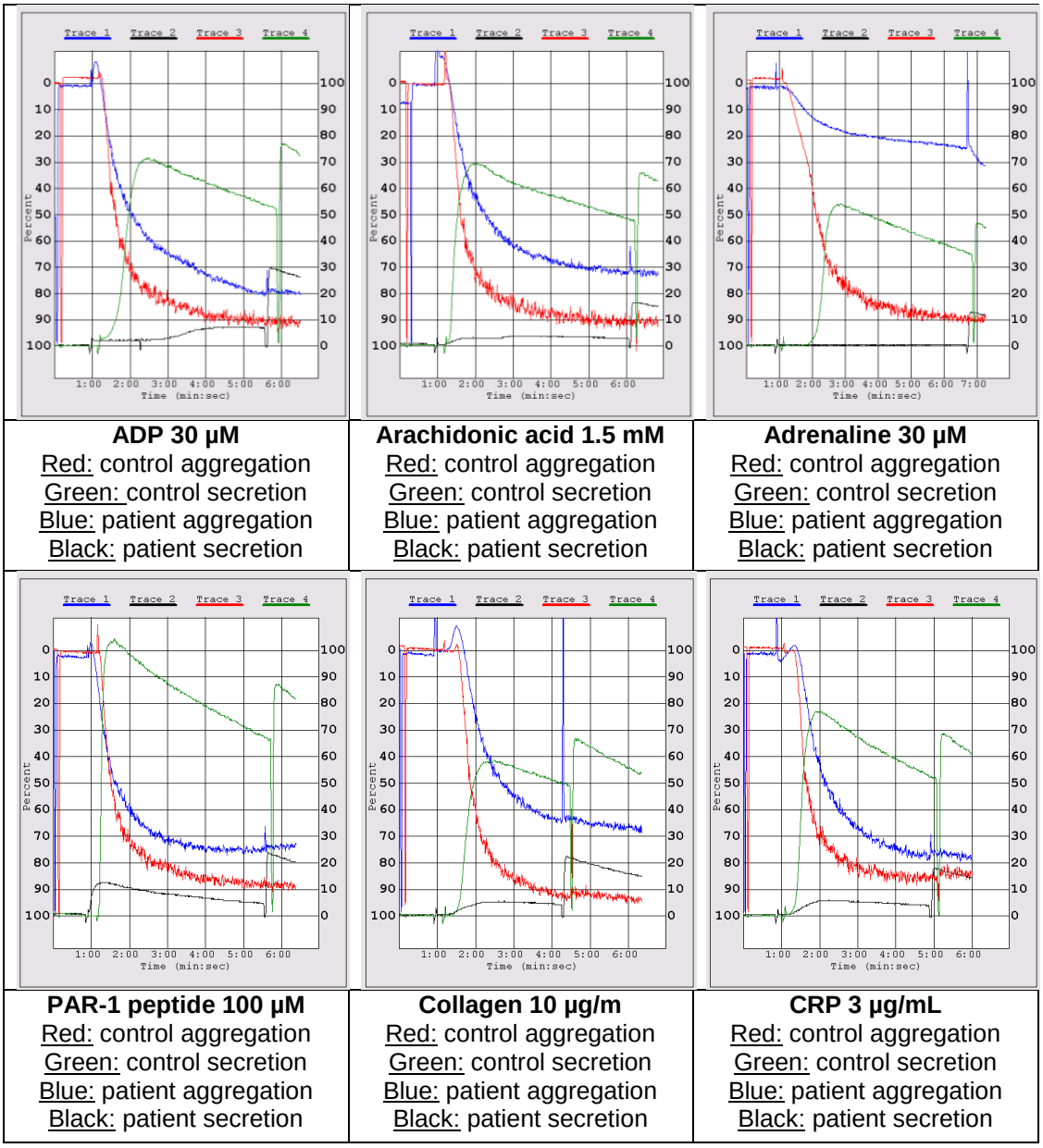


Figure 5.2 Platelet aggregation and ATP secretion in response to a range of agonists in index case F4.1 and a normal subject (control)

ADP; Adenosine diphosphate, CRP; Collagen-related peptide, PAR-1; Protease-activated receptor-1. Data provided by Dr Neil Morgan, University of Birmingham.

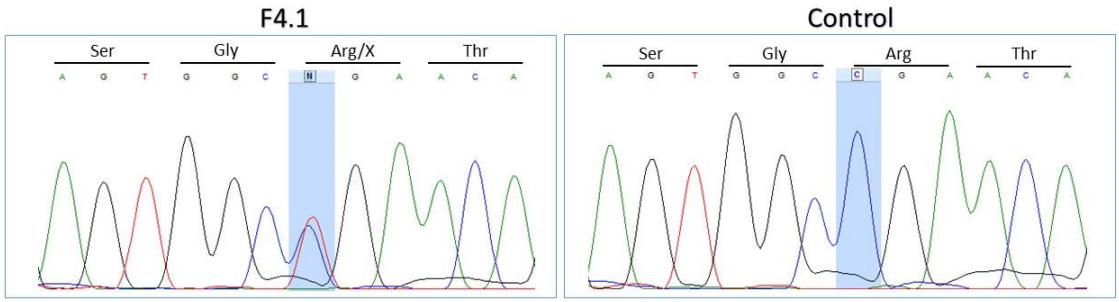


Figure 5.3 Sequencing traces of *ETV6* fragment amplified from the DNA of the index case F4.1 and a healthy control subject

Sanger sequencing confirmed the presence of the heterozygous ETV6:c.1288C>T transition (shaded in blue) in F4.1.

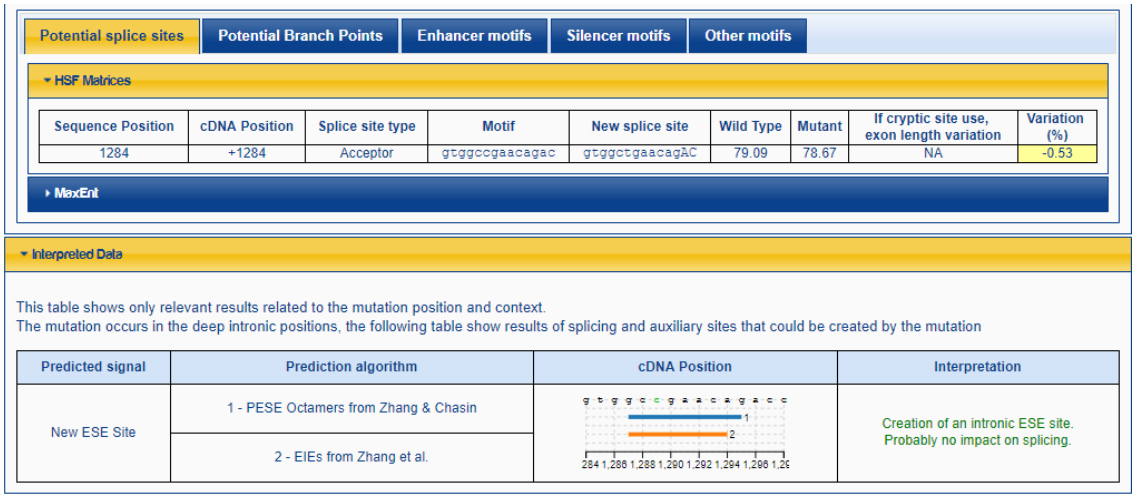


Figure 5.4 The output from Human Splicing Finder tool shows the predicted effect of the c.1288C>T transition on *ETV6* splicing

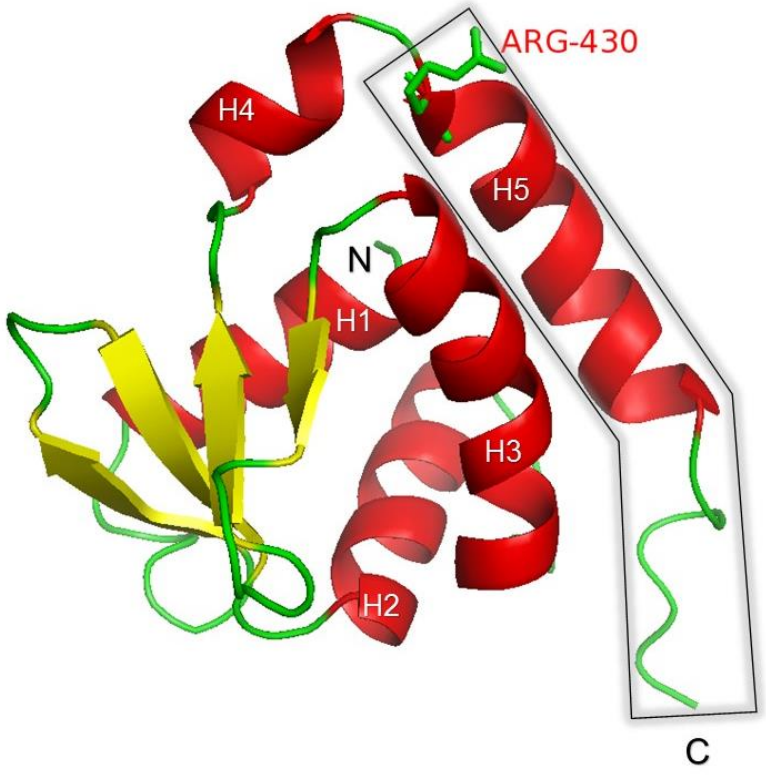


Figure 5.5 Partial structure of ETV6 showing the ETS domain, the C-terminal inhibitory domain and the location of arginine 430
Arginine 430 is shown as a stick. Colour coding: red; helices, yellow; sheets, loops; green.

5.4.5 Expression and subcellular localisation of the p.R430*-ETV6 variant

The c.1288C>T transition in *ETV6* is predicted to result in premature termination of protein translation at residue 430. To investigate whether the truncated ETV6 molecule

is stable *in vitro*, the transition was introduced into the ETV6 expression vector pCMV6-ETV6 and the plasmid was then transfected into HEK 293T cells. While the truncated ETV6 variant is predicted to retain all of the ETS domain, part of the NLS will be removed. For this reason, the effect of the p.R430* defect on subcellular localisation of ETV6 was also determined. Two additional variants were made for inclusion in these experiments. Firstly, R399C-ETV6 was included as an example of a previously characterised variant which has been shown to have reduced nuclear localisation and was therefore used as an internal control. A second truncated variant, R399*-ETV6, was made to allow comparison of the effect of premature termination of protein translation within the ETS domain to that within the CID domain.

Electrophoresis and immunoblotting of cytoplasmic and nuclear fractions extracted from cells transfected with WT and variant forms of ETV6 revealed both WT-ETV6 and the R399C-ETV6 variant migrated as a single protein species with the expected molecular weight of 53 kDa. In contrast, the R430*-ETV6 variant migrated as a protein species, which was estimated using the Protein Molecular Weight tool (https://www.bioinformatics.org/sms/prot_mw.html) [accessed 2016-2017] to have a molecular weight of 50 kDa. As expected, the R399*-ETV6 variant migrated faster than both WT-ETV6 and the R430*-ETV6 variant, having a molecular weight of approximately 45 kDa (Figure 5.6.A).

WT-ETV6 was present mostly in the nuclear, but also in the cytoplasmic, extracts when overexpressed in HEK 293T cells, and densitometric analysis of the blots revealed an 89-fold enrichment of WT-ETV6 in the nucleus when compared with that present in the cytoplasm (Appendix 11). All three of the ETV6 variants studied were also detected in both cellular fractions (Figure 5.6 A). However, in contrast to WT-ETV6, which was predominantly associated with the nucleus, the R399C displayed only five-fold enrichment in the nuclear fraction ($p=0.04$), the R399* variant was equally distributed between the nucleus and the cytoplasm ($p=0.04$), while R430* exhibited approximately 22-fold enrichment in the nucleus compared to the cytoplasm ($p=0.06$) (Figure 5.6 B, Appendix 11). All three ETV6 variants appeared to be present at higher levels in the cytoplasmic fractions than the WT-ETV6, which suggests an impaired ability to traffic to the nucleus (Figure 5.6 C). However, blots of the nuclear fractions revealed reduced

levels of the p.R399C and p.R399* variants and increased levels of the R430* variant when compared to WT-ETV6.

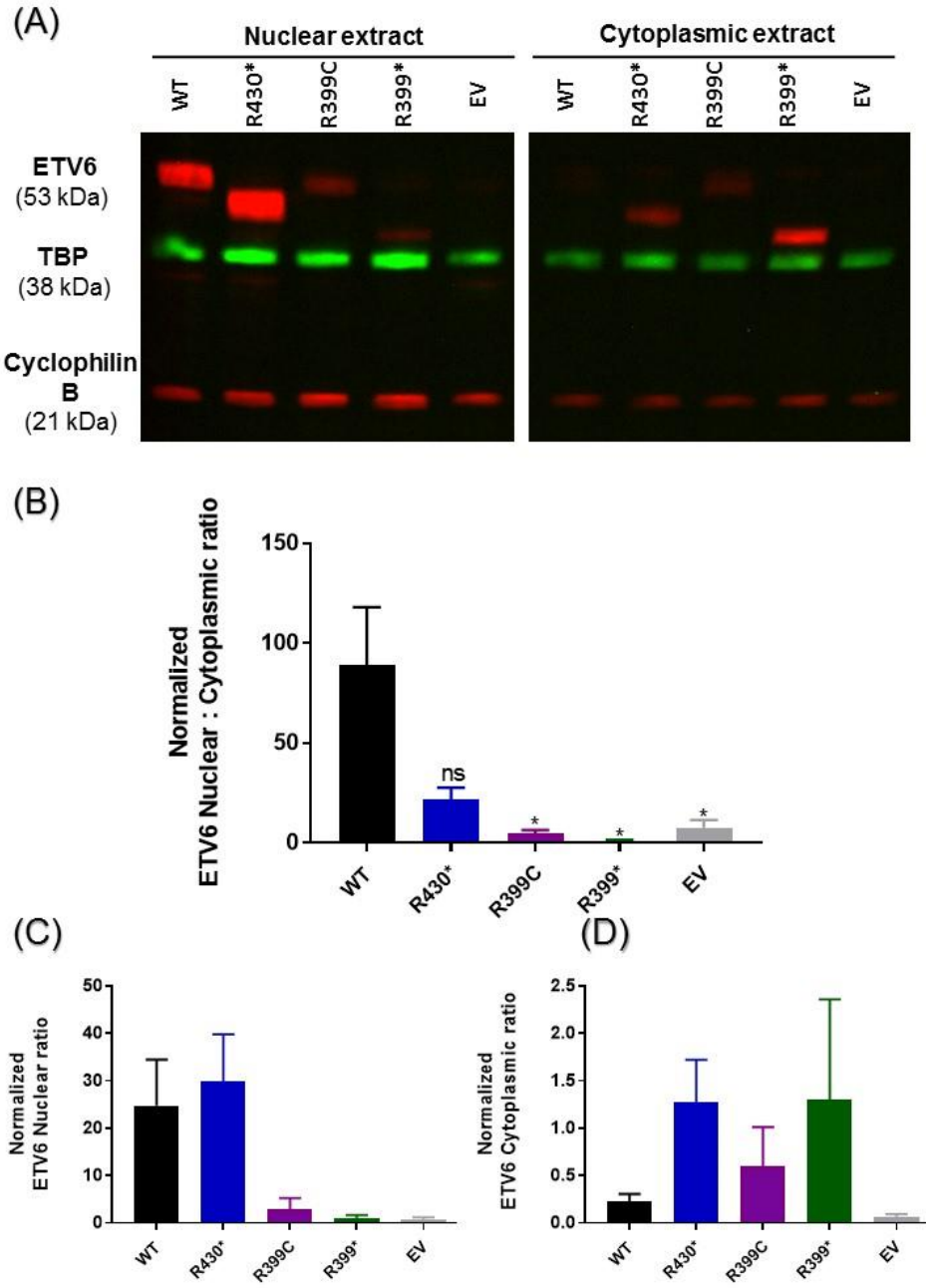


Figure 5.6 Subcellular localisation of wild-type and variant forms of ETV6
(A) Detection of ETV6 following western blotting of cytoplasmic and nuclear fractions from HEK 293T cells transiently overexpressing wild-type (WT)-ETV6 and the R430*, R399C and R399*-ETV6 variants. The blot was also probed for TATA-binding protein TBP (TBP) and cyclophilin B which act as markers of nuclear and cytoplasmic fractions respectively (B) Nuclear:cytoplasmic ratios of ETV6 calculated by densitometric analysis of the western blots (C-D) Normalised ETV6 ratios for nuclear and cytoplasmic fractions. The Mann-Whitney test was used for comparison, NS $p>0.05$, * $p\leq0.05$. Data represent a minimum of three independent experiments.

5.4.6 Effect of the R430*-ETV6 variant on transrepression of ETV6 target genes

ETV6 is known to downregulate *MMP3* and *PF4* expression, with previous studies using these genes in reporter assays to investigate the repression activity of different ETV6 variants (Noetzli et al., 2015; Topka et al., 2015; Zhang et al., 2015). The transrepression capacity of the recombinant R430*-ETV6 variant was therefore compared with that of WT-ETV6 using a dual luciferase reporter assay to measure repression of *MMP3* and *PF4* promoter activity in both HEK 293T and Dami cells. In addition to the R430* variant, derivatives of pCMV6-ETV6 encoding the R399C variant, which had been previously shown to have reduced transrepression capacity, and the R399* variant were used in these experiments.

5.4.6.1 Transrepression ability of ETV6 variants in HEK 293T cells

There was an approximate 50% reduction in *MMP3* promoter activity in the presence of WT-ETV6 compared to that observed in cells transfected with the empty vector (EV=1), confirming the ability of WT-ETV6 to repress *MMP3* expression (0.49-fold, $p \leq 0.0001$). Compared to cells expressing WT-ETV6, there were significant increases in luciferase activity, to almost the same levels as those observed in cells transfected with the EV, in the presence of both R399C (1.03-fold, $p<0.0001$) and R399* (0.70-fold, $p<0.0001$) variants, indicating significant reductions in repression activity. Interestingly, the R430*-ETV6 variant showed similar transrepression activity to that of WT-ETV6 (0.55 fold, $p=0.1392$) (Figure 5.7 A, Table 5.1).

Similar results were obtained using the *PF4* luciferase reporter construct. Compared to cells transfected with the EV, there was a 65% reduction in luciferase activity in the presence of WT-ETV6, confirming repression of *PF4* promoter activity (0.35-fold, $p<0.0001$). Both of R399 variants showed a loss in repression capacity (R399C 1.03-fold, $p<0.0001$; R399* 0.83-fold, $p<0.0001$). The R430* variant was able to repress *PF4* promoter activity, to a level which was comparable with that of WT-ETV6 (0.34-fold, $p=0.7645$) (Figure 5.7 B, Table 5.2)

The possibility that the R430* variant interferes with the transrepressor function of WT-ETV6 through a dominant-negative effect was explored by co-expressing different amounts of the R430* variant and WT-ETV6 in HEK 293T cells. However, co-expression of the R430* variant with WT-ETV6, using different ratios of the two

expression plasmids, resulted in transrepression of the *MMP3* promoter (25% R430* +75% WT-ETV6 0.6-fold, p=0.0104; 50% R430* + 50% WT-ETV6 0.62- fold, p=0.0618; 75% R430* + 25% WT-ETV6 0.48-fold, p=0.2304) and *PF4* (25% R430* + 75% WT-ETV6 0.38-fold, p=0.212; 50% R430* + 50% WT-ETV6 0.34-fold, p=0.4512; 75% R430* + 25% WT-ETV6 0.29-fold, p=0.0514) which was comparable to that observed in cells expressing WT-ETV6 alone (Figure 5.7, Table 5.1, Table 5.2). In contrast, co-expression of equal amounts of either the R399C or the R399* expression plasmid with the WT-ETV6 expression plasmid resulted in a significant increase in luciferase activity compared to cells expressing WT-ETV6 alone (*MMP3:* 50% R399C + 50% WT-ETV6 0.76-fold, p=0.0022; 50% R399* + 50% WT-ETV6 0.57-fold, p=0.0096; *PF4:* 50% R399C + 50% WT-ETV6 0.72-fold, p=0.0054; 50% R399* + 50% WT-ETV6 0.61-fold, p<0.0001). These results confirm those previously reported for the R399C-ETV6 variant (Topka et al., 2015; Zhang et al., 2015).

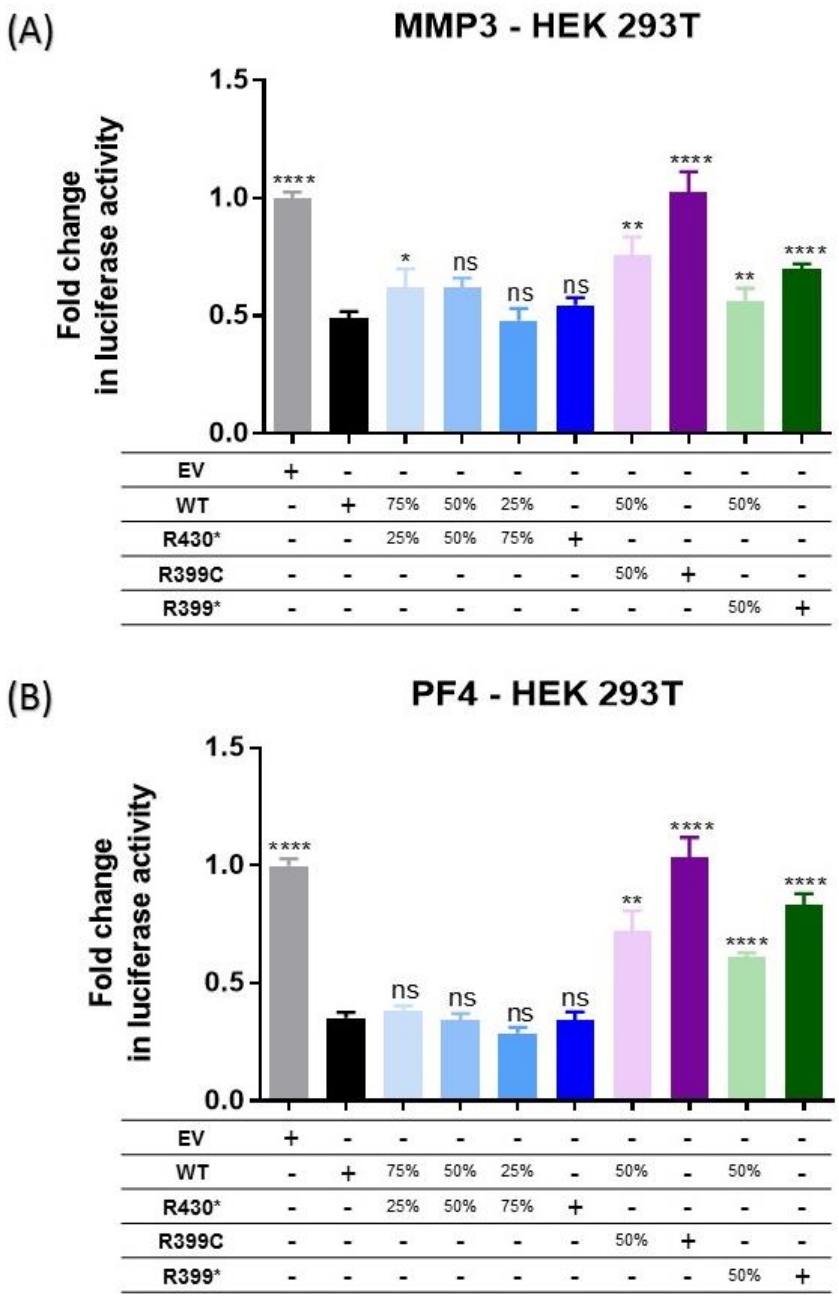


Figure 5.7 Transrepression of *MMP3* and *PF4* promoter activity by wild-type and variant forms of ETV6 in HEK 293T cells

Cells were transfected with wild-type (WT) or mutated ETV6 constructs, or with the empty vector (EV), in addition to either pGL3-MMP3-luciferase (A) or pGL3-PF4-luciferase (B) and pRLnull-Renilla reporters as described. Cells were also co-transfected with different ratios of the WT-ETV6 and ETV6 variant expression plasmids as indicated, maintaining a constant amount of the total DNA introduced into the cells. Firefly and Renilla luciferase expression were assessed in cell lysates 48 hours later. Transcriptional activity of ETV6 variants was measured by calculating the ratio of signal from the *MMP3*/*PF4* promoter (Firefly) compared to a control promoter (Renilla). The data are expressed as fold change in luciferase activity and represent the mean ± standard error of the mean of at least three independent experiments. Paired t-test was used for comparison NS $p>0.05$, * $p\leq0.05$, ** $p\leq0.01$, *** $p\leq0.001$, **** $p\leq0.0001$.

Table 5.1 Comparison of the transrepression capacity of ETV6 variants in HEK 293T cells on the *MMP3* promotor

	EV	WT	25% R430	50% R430	75% R430	R430*	50% R399C	R399C	50% R399*	R399*
Mean	1.00	0.49	0.62	0.62	0.48	0.55	0.76	1.03	0.57	0.70
% activity compared to EV	100	48.77	62.22	61.94	47.90	54.51	75.71	102.60	56.62	69.78
% reduction in activity compare to EV	0	51.23	37.78	38.06	52.10	45.49	24.29	-2.60	43.38	30.22
p-value	<0.0001	---	0.0104	0.0618	0.2304	0.1392	0.0022	<0.0001	0.0096	<0.0001
Number of repeats (each in triplicate)	7	---	3	5	3	6	3	4	3	4

For further details, see the legend to Figure 5.7. EV; empty vector, WT; wild-type.

Table 5.2 Comparison of the transrepression capacity of ETV6 variants in HEK 293T cells on the *PF4* promotor

	EV	WT	25% R430	50% R430	75% R430	R430*	50% R399C	R399C	50% R399*	R399*
Mean	1.00	0.35	0.38	0.34	0.29	0.34	0.72	1.03	0.61	0.83
% activity compared to EV	100.00	35.37	38.43	34.36	28.65	34.34	72.36	103.30	61.20	83.03
% reduction in activity compare to EV	0.00	64.63	61.57	65.64	71.35	65.66	27.64	-3.30	38.80	16.97
p-value	<0.0001	---	0.212	0.4512	0.0514	0.7645	0.0054	<0.0001	<0.0001	<0.0001
Number of repeats (each in triplicate)	8	---	3	7	3	8	3	4	3	4

For further details, see the legend to Figure 5.7. EV; empty vector, WT; wild-type.

5.4.6.2 Transrepression ability of the ETV6 variants in Dami cells

The capacity of the R430*-ETV6 variant to transrepress *PF4* and *MMP3* promoter activity was also investigated in the megakaryocytic Dami cell line. There was a 23.5% reduction in *MMP3* reporter activity in Dami cells in the presence of WT-ETV6 compared to cells transfected with the EV (0.77-fold; p<0.0001), confirming the ability of ETV6 to repress target gene expression. There was a significant increase in luciferase activity in the presence of both the R399C (1.06-fold, p<0.0001) and the R399* (0.98-fold, p=0.0004) variants compared to cells expressing WT-ETV6, which confirm the findings in HEK 293T cells. Interestingly, a small, but significant, increase in luciferase activity was also observed in the presence of the R430*-ETV6 (0.89-fold, p=0.0036) suggesting that it may have impaired transrepression activity in Dami cells (Figure 5.8 A, Table 5.3).

The effects of WT and variant forms of ETV6 were similar when promoter activity was assessed using the *PF4* luciferase reporter in Dami cells. Thus, there was a 46% reduction in *PF4* promoter activity in the presence of WT-ETV6 compared to cells transfected with the EV (0.54-fold, p<0.0001).The two R399-ETV6 variants failed to repress *PF4* promoter activity (R399C 1.20-fold; R399* 1.27-fold; p<0.0001), while the R430* variant led to a small but significant loss in transrepression activity (0.61-fold, p<0.0001) when compared with that observed in the presence of WT-ETV6 (Figure 5.8 B, Table 5.3).

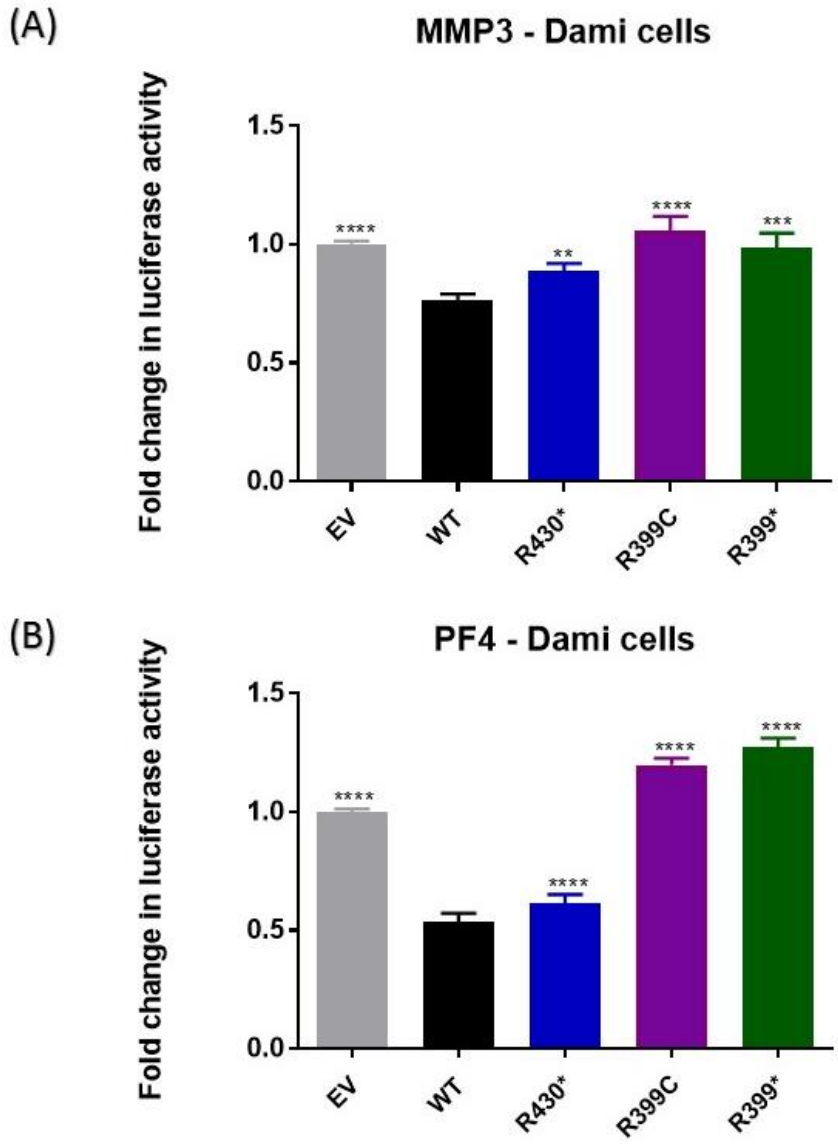


Figure 5.8 Transrepression of *MMP3* and *PF4* promoter activity by wild-type and variant forms of ETV6 in Dami cells

Cells were transfected with wild-type (WT) or mutated ETV6 constructs, or with the empty vector (EV), in addition to either pGL3-MMP3-luciferase (A) or pGL3-PF4-luciferase (B) and pRLnull-Renilla reporters as described. Firefly and Renilla luciferase expression were assessed in cell lysates 48 hours later. Transcriptional activity of ETV6 variants was measured by calculating the ratio of signal from the *MMP3*/*PF4* promoter (Firefly) compared to a control promoter (Renilla). The data are expressed as fold change in luciferase activity and represent the mean ± standard error of the mean of at least three independent experiments. Paired t-test was used for comparison, NS $p>0.05$, * $p\leq0.05$, ** $p\leq0.01$, *** $p\leq0.001$, **** $p\leq0.0001$.

Table 5.3 Transrepression of the *MMP3* and *PF4* promoters by wild-type and variant forms of ETV6 in Dami cells

MMP3					
	EV	**WT**	**R430***	**R399C**	**R399***
Mean	1.00	0.77	0.89	1.06	0.98
% activity compared to EV	100.00	76.50	88.95	105.70	98.27
% reduction in activity compare to EV	0.00	23.50	11.05	-5.70	1.73
p-value	<0.0001	---	0.0036	<0.0001	0.0004
Number of repeats (each in triplicate)	5	---	5	5	5

PF4					
	EV	**WT**	**R430***	**R399C**	**R399***
Mean	1.00	0.54	0.61	1.20	1.27
% activity compared to EV	100.00	53.74	61.34	119.90	127.40
% reduction in activity compare to EV	0.00	46.26	38.66	-19.90	-27.40
p-value	<0.0001	---	<0.0001	<0.0001	<0.0001
Number of repeats (each in triplicate)	5	---	5	5	5

For further details, see the legend to Figure 5.8. EV; empty vector, WT; wild-type.

5.5 Discussion

ETV6 is a transcriptional repressor that belongs to the ETS family of transcription factors (Golub et al., 1994; Lopez et al., 1999), which plays a role in hematopoiesis (Ciau-Uitz et al., 2010; Hock et al., 2004; Rasighaemi et al., 2015; Wang et al., 1998), including in the late stages of megakaryopoiesis (Hock et al., 2004). In 2015, Zhang et al. reported three germline missense variations in *ETV6* that segregated with dominant transmission of thrombocytopenia and haematologic malignancy. Since then, germline defects in *ETV6* have been associated with a new form of IPD called "Thrombocytopenia 5" (THC5).

To date, a total of 15 *ETV6* alterations have been reported to be associated with THC5 (Figure 5.1 and Table 5.4). Patients have presented with a history of mild to moderate bleeding, mild thrombocytopenia usually with normal sized platelets, and a predisposition to various neoplasms (Melazzini et al., 2016; Moriyama et al., 2015; Noetzli et al., 2015; Poggi et al., 2017; Topka et al., 2015; Zhang et al., 2015). Where platelets from affected individuals have been evaluated, they display mild defects in spreading, aggregation, clot retraction velocity and a shortened lifespan (Melazzini et al., 2016; Poggi et al., 2017). Ultrastructural analyses revealed round hypogranular

platelets with elongated α-granules and a disorganised open canalicular system (Poggi et al., 2017). *ETV6* alterations were also associated with MK hyperplasia characterised by defects in MK maturation and in proplatelet production, as indicated by the presence of small hypolobulated MKs, and a decrease in the number and length of proplatelet branches (Poggi et al., 2017). However, the exact role of ETV6 in megakaryopoiesis requires further elucidation, and the targets of ETV6 remain to be fully characterised.

In this chapter, a novel nonsense *ETV6* alteration that was predicted to result in premature termination of protein translation at residue 430 was investigated to assess its possible contribution to the bleeding phenotype observed in the affected patient. The p.R430* defect is different to previously reported *ETV6* variants in that protein translation is predicted to terminate in the CID domain which has not previously been shown to harbour defects related to IPDs. Additionally, the p.R430* variant was associated with a reduction in dense granule secretion and not with thrombocytopenia, as was the case for all other reported *ETV6* variants. *In silico* and *in vitro* expression studies were used to explore the possible effects of the c.1288C>T:p.R430* defect, alongside the previously characterised p.R399C-ETV6 variant. In addition, a truncated variant of ETV6 caused by the introduction of a stop codon at position 399 in the cDNA (p.R399*) was included in the experiments.

Table 5.4 Germline variants in *ETV6* associated with inherited platelet bleeding disorders

***ETV6* defect**	**Domain**	**Platelet count**	**Platelet size (MPV)**	**Bleeding / Platelets / MK phenotype**	**Other haematological features**	**Malignancies / Additional features**
c.641C>T:p.P214L (1, 2, 5, 6)	LID	Low	N/Dec	Mild to moderate bleeding; Normal platelets ultrastructure; Elongated α-granules; Small hypolobulated MK in BM; Decreased TPO plasma level; Reduced platelet half-life; Decrease in proplatelet-bearing MKs	Variable MCV; Abnormal RBC precursors in BM	ALL; MDS; MPAL; Breast fibroadenoma; Meningioma
c.1046T>C:p.L349P (3)	ETS	Low	---	---	Anaemia; Dec/N MCV	ALL; Arthritis; Ankylosing spondylitis; Uveitis; Secondary amenorrhea; Cleft lip/palate
p.I358M(6)	ETS	Low	Inc	Round larger and smaller platelets; Elongated α-granules; Hypolobulated small MKs in BM	---	AML
c.1075C>T:p.R359* (4)	ETS	Low	Dec	---	---	ALL; Mild intellectual disability; Learning disability
c.1105C>T:p.R369W (5)	ETS	Low	Dec	---	---	---
c.1106G>A:p.R369Q (1)	ETS	Low	Dec	Petechia; Epistaxis	---	Skin cancer; Colon cancer; Reading disability; GERD oesophagal stricture
p.A377T(6)	ETS	Low	Dec	Platelet anisocytosis; Hypogranular platelets; Poorly organised open canalicular system	Dyserythropoiesis; Delay granulocyte maturation in BM	---
c.1138T>A:p.W380R (5)	ETS	Low	Dec	Bleeding	N/Inc MCV	ALL; PCV

c.1153-5_1153_1delAACAG:p.N385fs [3]	ETS	Low	---	---	---	ALL; MDS; AML; Craniofacial and musculoskeletal anomalies; Right ear anterior placement; Downward shaped mouth; Joint hypermobility; CNS heterotopias
c.1153-1_1165del:p.N385Vfs*7 [5]	ETS	Low	N/Dec	Mildly reduced platelet aggregation	Inc MCV	ALL; Breast cancer; Breast fibroma
p.R396G[6]	ETS	Low	N	Mild bleeding; Dysmegakaryopoiesis; Almost no mature MK in BM	---	---
c.1195C>T:p.R399C [1]	ETS	Low	Dec	Bruising; Menorrhagia	Refractory anaemia; RAEB-I; Neutropenia	MDS; MM; ALL; Colorectal carcinoma; Myopathy; Gastrointestinal dysmotility; GERD; Developmental delay; Seizures; Dental disease; Delayed puberty
p.Y401N[6]	ETS	Low	Inc	Decreased platelet lifespan; Decrease in proplatelet-bearing MKs	---	---
p.Y401H[6]	ETS	Low	---	Platelet dense storage pool deficiency; Platelet aggregation defects; Abnormal dense granules	---	---
c.1252A>G:p.R418G+p.N385Vfs*7p.N385_418del [2, 5]	ETS	Low	Dec	Severe menorrhagia; Bruising; Nose bleeding	---	---
c.1288C>T:p.R430*	CID	N	N	Bleeding; Granule secretion defect	---	---

(1) Zhang et al., 2015, (2) Noetzli et al., 2015, (3) Topka et al., 2015, (4) Moriyama et al., 2015, (5) Melazzini et al., 2016, (6) Poggi et al., 2017. ALL; acute lymphoblastic leukaemia, AML; acute myeloid leukaemia, BM; bone marrow, CMML; chronic myelomonocytic leukaemia , Dec; decrease, ETS; ETS DNA binding domain, GERD; gastrointestinal oesophageal reflux disease, Inc; increase, LID; linker inhibitory damper domain, MCV; mean corpuscular volume, MDS; myelodysplastic syndrome, MK; megakaryocyte, MM; multiple myeloma, MPAL; mixed-phenotype acute leukaemia, MPV; mean platelet volume, N; normal, PCV; polycythemia vera, RAEB-I; refractory anaemia with excess blasts type I, RBC; red blood cell, TPO; thrombopoietin, WBC; white blood cells.

The c.1288 C>T transition in *ETV6* was identified in index case F4.1, who had enrolled in the UK-GAPP study with a history of bleeding, and whose platelets displayed a defect in dense granule secretion. The transition was predicted to introduce a premature stop codon at position 430 that would result in a truncated protein lacking H5 helix of the CID, a region known to negatively regulate the activity of the ETS domain. ETV6 and the other ETS family member, ETS1, are both described as having two autoinhibitory helices at the C-terminal end of the ETS domains (Coyne et al., 2012). In the case of ETS1, two additional autoinhibitory helices at the N-terminal of the ETS domain have been described and all four helices were found to interfere with binding of DNA to the ETS domain through allosteric mechanisms (Lee et al., 2005). Deletion of the two N-terminal inhibitory helices in ETS1 was found to abrogate the autoinhibitory function allowing ETS1 to bind DNA with higher affinity (Garvie & Wolberger, 2001), while deletion of the two C-terminal inhibitory helices impaired expression and/or solubility of ETS1 (Donaldson et al., 1994; Werner et al., 1997). These studies highlight the importance of the ETS appendix helices in regulating the function, folding and stability of the ETS domain. In ETV6, the CID region comprising residues 424 to 452 also contains the C-terminal part of the NLS, which encompasses amino acids 332 to 452 and includes the entire ETS domain (Park et al., 2006). Based on previous studies, it was predicted that the truncated R430*-ETV6 protein would be less stable than full-length ETV6, show reduced nuclear accumulation and have reduced ability to transrepress target genes. We therefore investigated the consequences of loss of H5 helix on ETV6 function by examining nuclear localisation of the truncated R430*-ETV6 and its ability to transrepress the activity of known ETV6 target genes, *MMP3* and *PF4.*

The c.1288C>T transition in *ETV6* that results in the p.R430* variant occurs in exon 8 of *ETV6,* which is the last exon of the gene. Nonsense variants that occur in the closing exons of genes have been reported to be less likely to result in nonsense-mediated decay (Lejeune & Maquat, 2005). In HEK 293T cells, the R430* variant was expressed as a stable truncated protein, though its subcellular distribution pattern differed from that of WT-ETV6. While the level of nuclear R430*-ETV6 was comparable with that of WT-ETV6, and higher than that of the R399C and R399* ETV6 variants, the level of cytoplasmic ETV6 was elevated in cells expressing R430*-ETV6 or either of the two R399 variants compared with those expressing WT-ETV6. When the level of ETV6

distributed between the nuclear and cytoplasmic fractions was expressed as a ratio, the ratio for the R430*-ETV6 variant was reduced compared with that for WT-ETV6 due to the elevation in the level of cytoplasmic R430*-ETV6 variant, though not significantly. The increase in cytoplasmic levels and a decrease in nuclear localisation for both R399C- and R399*-ETV6 variants were in agreement with previous findings for the R399C variant (Topka et al., 2015; Zhang et al., 2015) and other previously described ETV6 variants (Noetzli et al., 2015; Topka et al., 2015; Zhang et al., 2015).

The ability of the three ETV6 variants to transrepress the *MMP3* and *PF4* promoters was compared with that of WT-ETV6 using dual luciferase reporter assays in HEK 293T cells and in the megakaryocytic Dami cell line. Interestingly, the R430* variant demonstrated repression activity comparable to that of WT-ETV6 in HEK 293T cells. However, it showed a small, but significant, reduction in repression activity in megakaryocytic Dami cells. These findings could be explained by tissue-specific differences between the two cell lines, as the presence of specific cofactors or signal specific post-transcriptional modifications have a direct impact on subcellular localisation and activity of ETV6 (Chakrabarti et al., 1999; Chakrabarti et al., 2000). Given that several previously described ETV6 variants were shown to interact with WT-ETV6 to exert a dominant-negative effect on target gene repression (Noetzli et al., 2015; Poggi et al., 2017; Topka et al., 2015; Zhang et al., 2015), the effect of the R430*-ETV6 variant on target gene repression was also examined in the presence of WT-ETV6. However, co-expression of the R430*-ETV6 variant with WT-ETV6 at different ratios resulted in repression activity comparable to that observed in the presence of WT-ETV6 alone. The R399C variant showed a reduction in ETV6 transrepression activity in both cell lines, which confirms previous reports for this variant (Topka et al., 2015; Zhang et al., 2015). Similarly, the R399* truncation variant also showed a reduction in transrepression activity in both cell lines, consistent with previous findings that the integrity of the ETS domain is required for ETV6 to bind to DNA (Noetzli et al., 2015; Poggi et al., 2017; Topka et al., 2015; Zhang et al., 2015).

The findings reported here do not support the hypothesis that the c.1288C>T:p.R430* *ETV6* variant is the pathogenic defect underlying the bleeding tendency in F4.1. It is possible that the truncated ETV6 is not stably expressed in the patient's platelets. Indeed, a previously described *ETV6* variant that was predicted to result in a truncated

ETV6 molecule (c.1252A>G:p.Asn385Valfs*7) was found to be expressed as a truncated protein in HEK 293T, but not in platelets from the affected patients (Noetzli et al., 2015).

If the truncated R430*-ETV6 is expressed as a stable protein in the patient's platelets, we might expect loss of H5 helix of the CID, which is known to repress ETV6 function by 10-50 fold, to lead to a gain of function for ETV6 (Coyne et al., 2012). Deletion of the CID region of ETV6 was previously shown to cause a 10-fold increase in the affinity of the ETV6 monomer for DNA to a level that was similar to the DNA affinity of the WT dimer (Green et al., 2010). The truncated R430*-ETV6 could therefore act as a more potent repressor of its target genes, or continuously repress target gene activity and/or repress expression of genes that it does not normally target. This abnormal repression activity could, in turn, alter expression of essential MK/platelet genes which could potentially contribute to the bleeding phenotype observed in our patient. While this hypothesis needs to be tested, a gain of function effect was recently described in a similar condition, where a nonsense mutation in *DIAPH1* (Rho-effector diaphanous-related formin 1) leading to loss of the C-terminal autoregulatory domain of DIAPH1 was described in a patient with macrothrombocytopenia and progressive hearing loss (Neuhaus et al., 2017; Stritt et al., 2016). DIAPH1 has a recognisable role in actin remodelling in the platelet cytosol (Higashi et al., 2008), and more recently, was found to play a critical regulatory role in proplatelet formation in MKs (Pan et al., 2014). Loss of the diaphanous autoregulatory domain resulted in constitutive activation of DIAPH1, resulting in reduced proplatelet formation and altered cytoskeletal regulation in platelets (Stritt et al., 2016).

It is possible that the nonsense c.1288C>T transition in *ETV6* does not contribute to the bleeding tendency observed in our patient, in which case, my findings highlight the necessity to exercise caution when interpreting the bioinformatic predictions of the likely pathogenicity of candidate genetic variations, and the importance of conducting functional studies on identified variants. Re-examination of the WES data for index case F4.1 (Appendix 12) revealed the presence of a heterozygous non-frameshift insertion in *ZFPM1*, the gene encoding Zinc Finger Protein, FOG Family Member 1, which could potentially be of interest. ZFPM1 is a transcription factor required for the formation of committed MK and erythroid progenitors (Mancini et al., 2012) and mice

with disrupted ZFPM1 showed Grey platelet syndrome-like macrothrombocytopenia and pleiotropic platelet defects, including abnormal granule formation and a dramatic decrease in P-selectin expression (Wang et al., 2011). *ZFPM1* alterations have not been associated with IPDs and the alteration that occurs in F4.1 (c.1685_1686insGGGCGC:p.A562delinsAGA) was not predicted to disrupt any of the zinc finger or other recognised functional domains of ZFPM1 when it was inspected using the InterPro online tool (https://www.ebi.ac.uk/interpro/) [accessed 2015-2016]. Nonetheless, the possibility that this variant contributes, either alone or in combination with other genetic variants, to the bleeding tendency in the patient warrants further investigation.

Further studies of the c.1288C>T:R430*-ETV6 variant would ideally involve analysis of RNA in platelets from the index case F4.1, as well as other affected and unaffected family members. In particular, quantitative measurement of the *ETV6* transcript by qPCR would determine whether it was stably expressed or removed via nonsense-mediated decay. Failure to detect a stable *ETV6* transcript would provide strong support for the hypothesis that the c.1288 C>T transition is the underlying pathogenic defect in the patient. However, if the variant is stably expressed as a truncated protein, transcriptomic analysis could be undertaken to evaluate the expression of known ETV6 target genes (e.g. *EGR1* and *TRAF1*). This approach also has the potential to identify other genes that may be differentially expressed in the patient's platelets compared to control subjects, possibly as a result of a gain of function of ETV6.

6 Chapter 6. CRISPR-generated *FLI1* knockdown in megakaryocytic Dami cell line as *in vitro* cell model to study FLI1 role in megakaryopoiesis

6.1 Introduction, hypothesis and aims

Whole exome sequencing (WES) has contributed to the establishment of the genetic basis for many disorders, including inherited platelet bleeding disorders (IPDs) (see section 3.1.1). However, despite continuous advancements in the field, identification of disease-causing genetic variants in the exomes from single patients with heterogeneous disorders is still a major challenge. Chapter 3 summarised the findings obtained when WES was applied to the identification of potential disease-causing variants in 34 index cases with uncharacterised IPDs. In a subgroup of 22 index cases whose platelets displayed defects in platelet secretion, WES highlighted 1,465 variants in 1,130 genes, of which only nine variants, affecting *ETV6* (chapter 5), *FLNA*, *FLI1* (chapter 4), *GFI1B*, *ITGB3*, *P2RX1* and *RUNX1*, were considered likely to contribute to the bleeding symptoms in eight patients. The recognition of these variants was primarily due to the previously described association of defects in these genes with IPDs. However, a median of 70 candidate gene defects was identified in each of the remaining 14 index cases. Clearly a systematic exploration of each of these defects to determine its likelihood of being causative would be impractical. An alternative approach was therefore sought to identify novel candidate genes that could contribute to the bleeding disorder in these patients.

Given the essential role of FLI1 as a master regulator of megakaryopoiesis and the association of *FLI1* defects with abnormal platelet secretion, it is likely that the genes which are differentially expressed due to the loss of FLI1 activity will also have a role in platelet granule biogenesis and secretion. Variants that are harboured by these genes and present in patients with platelet secretion defects, would be more likely to be causative, though further work would be required to confirm this. Therefore, the work in this chapter explored the hypothesis that variants in genes that are differentially expressed following knockout/down of *FLI1* in megakaryocytes (MKs) will be enriched among subjects with abnormalities in platelet secretion. This was achieved by (i) knockout/down of *FLI1* in the megakaryocytic Dami cell line; (ii) transcriptome analysis of Dami cells following *FLI1* knockout/down; (iii) identification of differentially expressed genes that also harbour candidate defects in index cases with unexplained IPDs characterised by defects in platelet secretion.

A CRISPR/Cas9 (Clustered regularly interspaced short palindromic repeats / CRISPR associated protein 9) based method was used to knockout/down *FLI1*. The CRISPR/Cas system is a prokaryotic adaptive immune system against phage infection and plasmid transfer that occurs in nature (Barrangou et al., 2007). It comprises a non-specific nuclease (Cas) that mediates a nucleic acid double-strand break (DSB) at the target site when guided by a sequence-specific guide RNA (gRNA). The use of type II CRISPR/Cas9 systems for gene editing in mammalian cells was described in 2013 (Cong et al., 2013; Mali et al., 2013). Since then, the simplicity, reliability, high efficiency and cost-effectiveness of CRISPR/Cas9 technology have led to its widespread use in genome editing applications (Adli, 2018).

6.2 Materials and Methods

CRISPR/Cas9 knockout/down of *FLI1* in Dami cells was executed following the protocol described by Bauer et al. (2015). The workflow for this procedure is illustrated in Figure 6.1, and the detailed methodology is provided in section 2.2.14. The knockout/down relies on the introduction of simultaneous DSBs in DNA by Cas9, guided by a pair of gRNAs that target two different exons which result in loss of the intervening genomic sequence through the mechanism of non-homologous end joining repair (Canver et al., 2014). The CRISPR/Cas9 was delivered using the CRISPR pX458 plasmid (~9.3 kb) that contains a cassette encoding *Streptococcus pyogenes* (Sp)-Cas9 with enhanced green fluorescent protein (EGFP) as a selection marker, and the cloning backbone for a single gRNA.

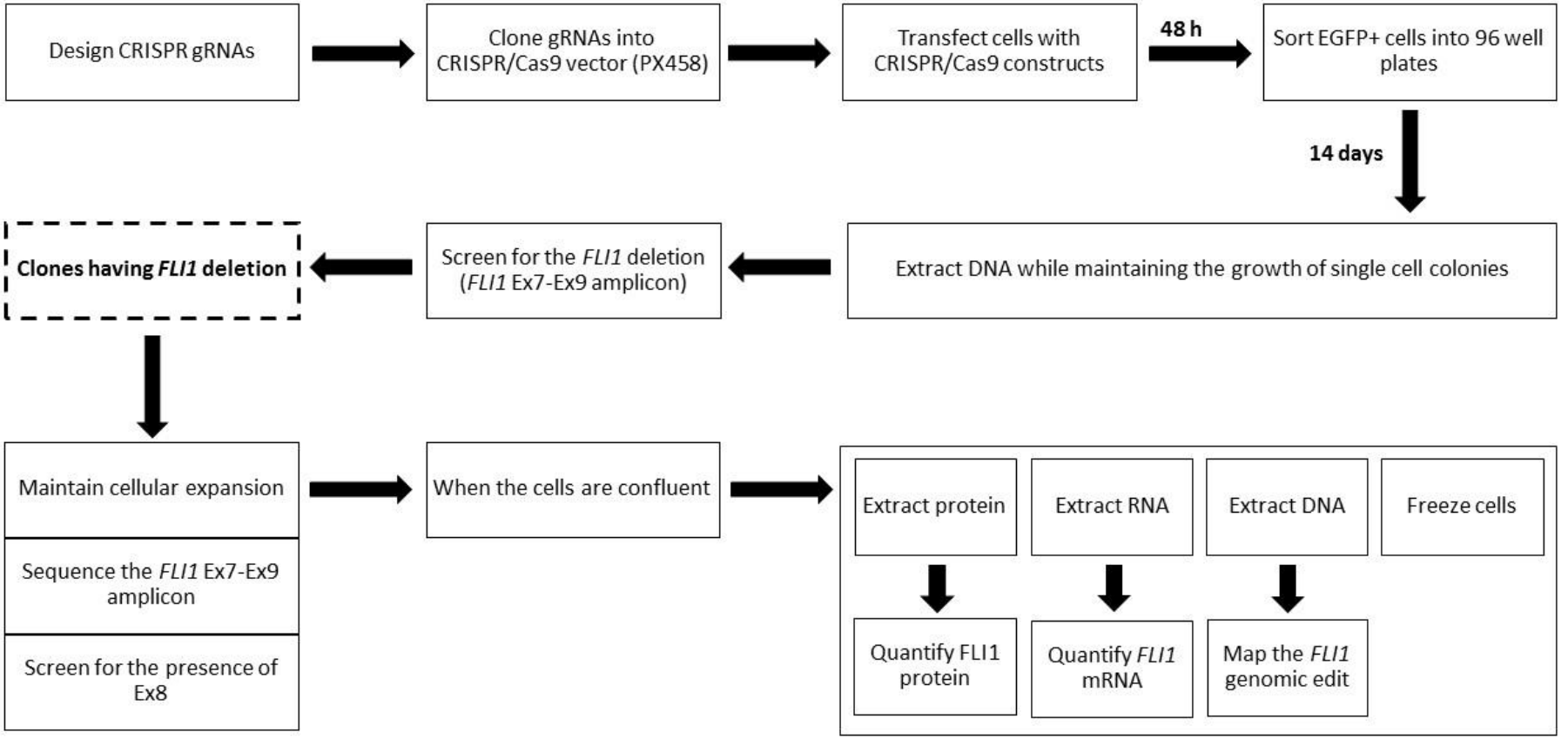


Figure 6.1 The workflow for CRISPR-mediated *FLI1* knockdown/out in Dami cells
EGFP; enhanced green fluorescent protein, gRNA; guide RNA, h; hours.

6.2.1 Design and cloning of guide RNAs

Several factors were considered when designing the CRISPR gRNA molecules for *FLI1* knockout/down. Coding regions that are present in all *FLI1* coding transcripts were targeted to increase the likelihood of gene knockout by introducing a small indel within the exon if the simultaneous DSBs did not occur. Given that the likelihood of introducing a specific deletion is inversely related to its size (Bauer et al., 2015), gRNAs were positioned to achieve a deletion of less than 5 kb, with the highest possible binding specificity and cutting efficiency (Appendix 13). Accordingly, exons 7 and 9 were selected as the locations for the gRNAs (Figure 6.2).

Additional nucleotides were added to each gRNA oligonucleotide to facilitate their cloning and expression (Table 6.1). The guides were then cloned separately into the CRISPR pX458 plasmid using the “Golden Gate assembly” cloning strategy.

Table 6.1 Sequences of CRISPR guides targeting exons 7 and 9 of *FLI1*

	5´-Guide Sequence-3´	5´-Reverse complement-3´	Direction
Exon 7	**CACC**GATCGTTTGTGCCCCTCCAA	**AAAC**TTGGAGGGGCACAAACGATC	-
Exon 9	**CACCg**AATGACGGACCCCGATGAGG	**AAAC**CCTCATCGGGGTCCGTCATT**c**	+

Additional nucleotides are indicated in bold.

6.2.2 Transfection of Dami cells with CRISPR plasmids and fluorescence-activated cell sorting

Dami cells were transfected with 1,500 ng of each pX458 CRISPR plasmid, pX458-Ex7 (targeting exon 7) and pX458-Ex9 (targeting exon 9), using a 3:1 ratio of jetPEI to DNA. They were also transfected with the native pX458 CRISPR plasmid as a control. The transfected cells were allowed to recover for 48 hours, before they were sorted into single cells in 96-well plates and allowed to expand undisturbed for 14 days.

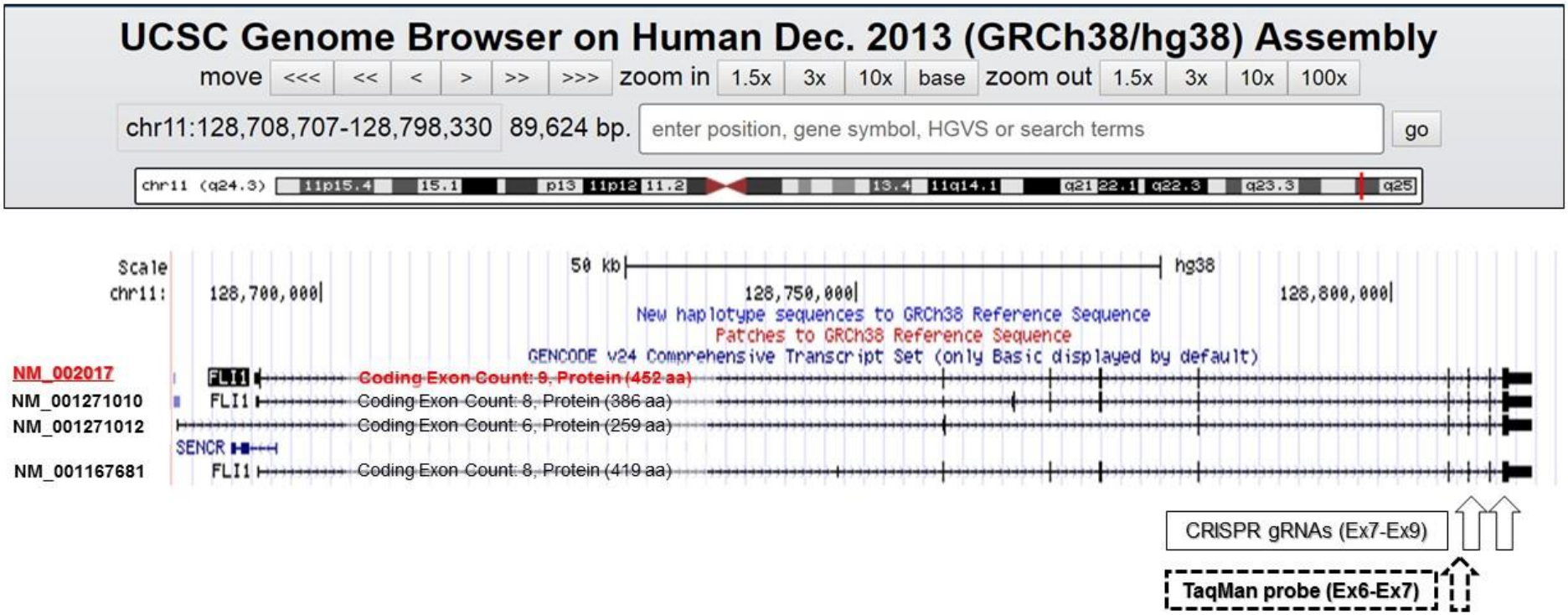


Figure 6.2 Screenshot from UCSC showing the four human *FLI1* transcripts and the location of binding of the CRISPR guide RNAs and *FLI1* TaqMan™ probe

The transcripts were annotated manually with the RefSeq coding, number of exons and the size of the translated protein. CRISPR gRNAs were designed to target exons 7 and 9 in NM_002017 and delete the intervening sequence. The binding location for the *FLI1* TaqMan™ probe is also indicated.

6.2.3 Screening for CRISPR-mediated edits at the DNA level

QuickExtract™ solution was used to extract PCR-ready genomic DNA (gDNA) from an aliquot of cells growing in 96-well plate. Preliminary screening for the desired *FLI1* deletion was carried out by PCR amplification of the extracted gDNA using primers designed to amplify across the deleted region (Ex7_F and Ex9_R), which would typically result in an amplicon of approximately 539 bp in those cells where a deletion had occurred. DNA samples from clones that were positive for the deletion were then screened by PCR using a second set of primers, which were designed to amplify a 244 bp fragment of exon 8 (Ex8_F and Ex8_R), to detect alleles where the desired edit had not occurred. Given that Dami cells are triploid (Greenberg et al., 1988), the smaller 244 bp fragment would fail to be amplified only in those cells where the deletion had occurred on all three *FLI1* alleles.

6.2.4 Characterisation of CRISPR-edited clones

FLI1 knockout/down was evaluated at protein and RNA level in clones that were positive for deletion by DNA analysis. Following extraction of total cellular protein using radioimmunoprecipitation assay (RIPA) buffer, FLI1 protein expression was assessed by lithium dodecyl sulphate - polyacrylamide gel electrophoresis (LDS-PAGE) and immunoblotting using antibodies to FLI1 and β-tubulin. The intensity of the 51 kilodalton (kDa) FLI1 protein detected on blots was normalised to that of the β-tubulin housekeeping protein and calculated as a percentage relative to the FLI1 species detected in wild-type (WT) Dami cells.

FLI1 RNA was quantified by qPCR using TaqMan™ probes for *FLI1* and the housekeeping beta-2 microglobulin RNA. *FLI1* RNA levels in each clone were normalised to those of the housekeeping gene and expressed relative to RNA levels in WT Dami cells. Results were plotted as log 10 of the relative quantity (RQ) from three replicates per sample.

Following extraction of gDNA using the GenElute™ mammalian genomic DNA mini-prep kit, the precise DSB points were determined for selected clones by sequencing amplified DNA spanning the targeted region, either directly, or following TA cloning.

6.2.5 Transcriptome analysis of wild-type Dami cells and cells displaying reduced *FLI1* expression

Selected *FLI1* knockdown and control clones were seeded in parallel into two culture flasks. The cells in one flask were stimulated to differentiate over six days by daily treatment with thrombopoietin (TPO) and phorbol 12-myristate 13-acetate (PMA), while the cells in the second flask were left untreated. RNA was purified from the differentiated, and undifferentiated cells and subjected to transcriptome analysis using the Affymetrix Gene Chip Microarray system and Clariom^{TM} D Assay_human chips (Affymetrix; Applied Biosystems™). Transcriptome data from clones that had been treated similarly was combined prior to analysis and transcripts showing a ≥1.5 or ≤-1.5 fold log change in expression, at a $p<0.05$ level of significance were highlighted.

Platelet transcriptome data which was obtained for two affected members of a family with an inherited bleeding tendency who also carried the *FLI1* c.1028A>G:p.Tyr343Cys variant, and three matched control subjects were kindly made available for use in the comparative analysis as part of this study by Dr Simon Webster, Haemostasis Group, Sheffield.

The Database for Annotation, Visualization and Integrated Discovery (DAVID) functional annotation tool was used to analyse and compare the transcriptomic data. DAVID is a web-accessible programme (see Appendix 3 for link) that allows clustering of genes as defined by the user. The Gene Ontology (GO) options "GOTERM_BP_FAT, GOTERM_CC_FAT and GOTERM_MF_FAT" which generate clusters of genes according to their associated biological process, cellular component and molecular function GO terms respectively, and a classification stringency setting of "Medium" were selected for the analysis. Appendix 14 is a screenshot showing the settings used for functional annotation analysis in DAVID. The group enrichment scores, defined as the geometric mean of the p-values associated with each annotation in a cluster, were used to rank their significance, with an enrichment score greater than 1.3, which is equivalent to a p-value of 0.05, being considered significant.

6.3 Results

6.3.1 Optimising transfection of Dami cells using jetPEI

Transfection of Dami cells using jetPEI was optimised experimentally utilising flow cytometry to evaluate transfection efficiency and cell viability. First, cells were transfected with varying amounts of the pEGFP-N3 plasmid using the recommended 2:1 ratio of jetPEI:DNA (Table 6.2). The results indicated that 3,000 ng was the optimal amount of DNA to use for transfection. Cells were then transfected with 3,000 ng of the same plasmid at varying ratios of jetPEI to DNA, which showed that a 3:1 ratio of jetPEI to DNA resulted in the highest transfection efficiency without compromising cell viability (Table 6.2).

Table 6.2 Optimising transfection of Dami cells with jetPEI

	Amount of pEGFP-N3 plasmid used					
	1 µg	**3 µg**	6 µg	8 µg	10 µg	12 µg
Viability	82%	**75.1%**	58.4%	44.5%	34.4%	20.5%
Transfection efficiency	0.5%	**4.7%**	2.6%	2%	1.7%	1.1%

	jetPEI: DNA ratio			
	3 µg (1:1)	3 µg (2:1)	**3 µg (3:1)**	3 µg (4:1)
Viability	79.4%	75.1%	**73.6%**	70.4%
Transfection efficiency	0.5%	4.7%	**7.7%**	8.2%

6.3.2 Screening for genomic edits in *FLI1*

DNA samples from approximately 80 single cell clones were screened by PCR for the 539 bp fragment that indicated loss of 3.9 Kb of DNA, which was predicted to occur by targeting exons 7 and 9 of *FLI1* by CRISPR/Cas9. The 539 bp fragment was amplified from DNA isolated from nine clones; #1, #8, #12, #18, #25, #29, #44, #60 and #B (Figure 6.3 A). Sanger sequencing of the amplified fragment identified the deletion breakpoints for clones #29 and #44 (Figure 6.3 C). A 244 bp fragment corresponding to exon 8 of *FLI1* was amplified from all nine clones, indicating that none of them had deletions of exons 7 to 9 on all three *FLI1* alleles (Figure 6.3 B).

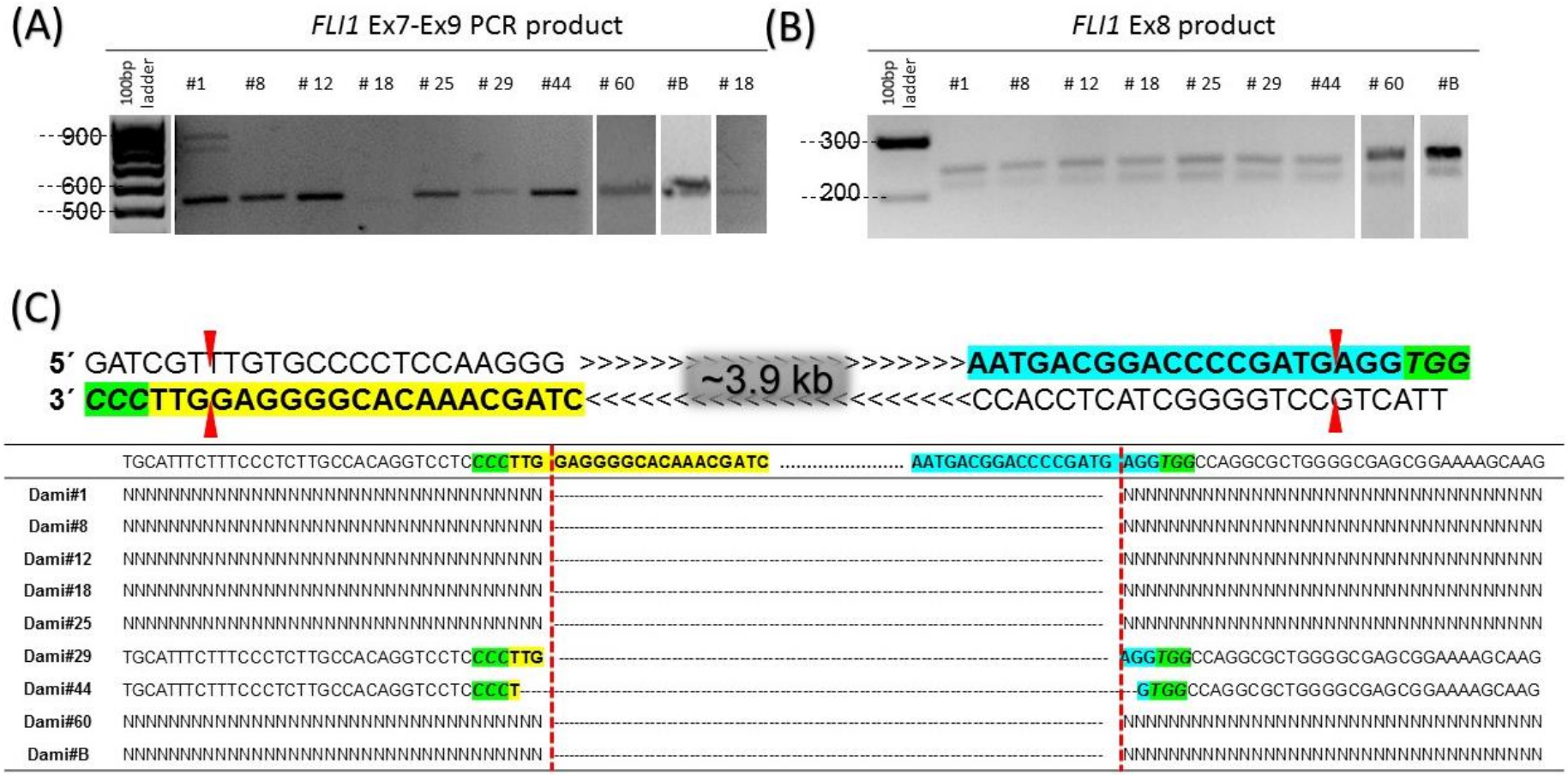


Figure 6.3 Screening of Dami cell clones for genomic edits introduced by targeting exons 7 and 9 of *FLI1* by CRISPR/Cas9
Agarose gel electrophoresis of PCR fragments amplified from Dami cell clones showing (A) the 539 bp deletion-specific fragment amplified using primers located in exons 7 and 9 and (B) the 244 bp fragment of exon 8. (C) The locations of the double-strand break points in clones #29 and #44 were identified by sequencing the corresponding 539 bp deletion-specific fragments. The guide RNA sequence is shown in bold font highlighted in either yellow (exon 7) or blue (exon 9). The protospacer adjacent motif (PAM) sequence is in bold, italic font and highlighted in green. The predicted cutting site is indicated by the red arrows and vertical broken lines.

It is noteworthy that several of the single cell clones that began to expand after sorting became non-viable after 14 days in the absence of any apparent infection. Otherwise, there were no obvious differences in the rate of growth of the control clones (numbered C1 to C6) and the CRISPR-edited clones, with the exception of clones #60 and #B, which took longer to reach confluency (~21 days and ~60 days respectively).

6.3.3 FLI1 protein and RNA levels in CRISPR-edited Dami cell clones

LDS-PAGE and western blotting of FLI1 in cell extracts from native Dami cells and from the CRISPR-edited clones identified above resulted in detection of a protein with the expected molecular weight of FLI1 (~51 kDa) in all but one clone (clone #B). Densitometric analysis showed that compared to the WT cells, clones #1, #12, #25, #29, #60 and #B all had lower levels of the 51 kDa FLI1 species. Additionally, lower molecular weight forms of FLI1 were detected in clones #1, #12, #29, #44, #60 and #B. The polyclonal antibody to FLI1 that was used in this study was raised to the N-terminal 222 amino acids of FLI1 derived from transcript NM_002017 (51 kDa), which are also present in the FLI1 isoforms NM_001167681 (~ 48 kDa) and NM_001271010 (~44 kDa). This suggests that the lower molecular weight proteins detected by western blotting could be other FLI1 isoforms or degraded FLI1 protein (Figure 6.4).

Similarly, qPCR of *FLI1* RNA in the CRISPR-edited clones showed that clones #1, #25, #29, #44, #60 and #B all had lower *FLI1* expression relative to that of polyclonal Dami cells (Figure 6.5). The TaqMan™ probe used for qPCR of *FLI1* RNA binds to exons 6-7 of the coding sequence, and therefore should detect all four *FLI1* transcripts.

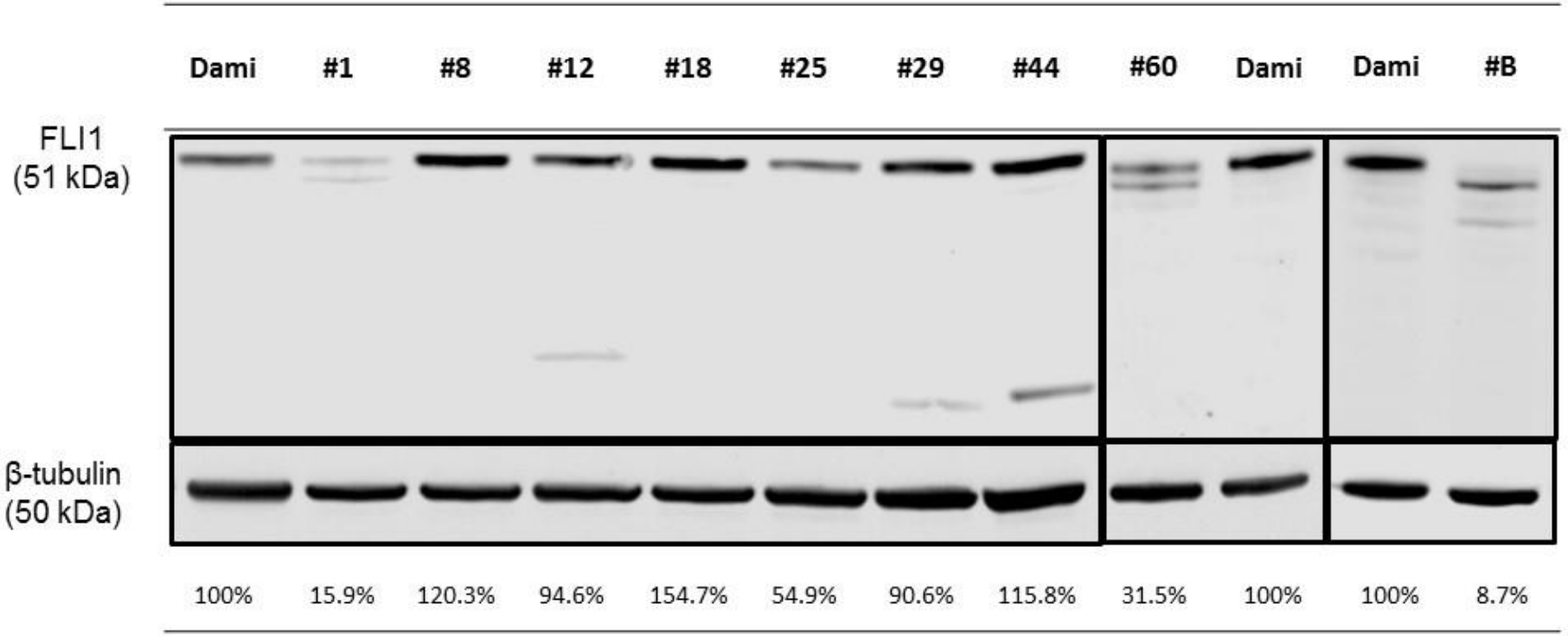


Figure 6.4 Reduced FLI1 protein expression in CRISPR-edited clones relative to wild-type polyclonal Dami cells
Western blotting of protein extracted using radioimmunoprecipitation assay (RIPA) buffer was used to evaluate FLI1 protein expression in CRISPR-edited clones. The blot was also probed for β-tubulin as a housekeeping protein. FLI1 expression by the clones was calculated as a percentage relative to that in wild-type polyclonal Dami cells.

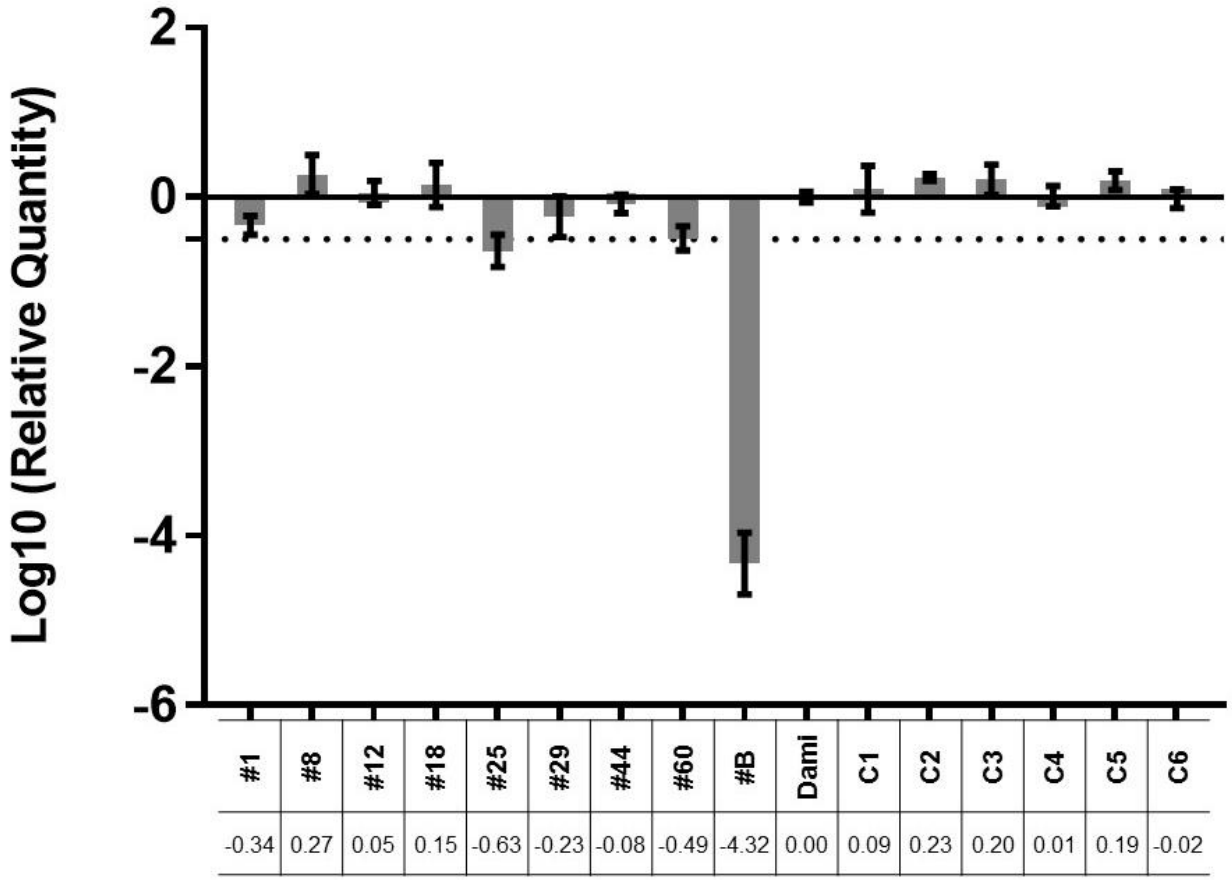


Figure 6.5 Reduced *FLI1* RNA expression in CRISPR-edited clones relative to wild-type polyclonal Dami cells

Quantitative polymerase chain reaction (qPCR) was used to evaluate relative *FLI1* RNA expression in CRISPR-edited clones and six single wild-type Dami clones (C1-C6). *FLI1* RNA expression relative to that of polyclonal Dami cells following normalisation to beta-2 microglobulin gene expression is shown. The data are plotted as relative quantity (RQ) from triplicate samples and the error bar represents the calculated maximum and minimum expression levels at the 99% confidence level.

6.3.4 Mapping of genomic edits in selected clones

Based on the results above, clones #1, #25, #60 and #B were selected for further study as they all expressed approximately 50% or less of FLI1 at protein level, thereby mimicking the situation in platelets from a patient with *FLI1* haploinsufficiency. The deletion breakpoints were mapped for clones #1, #25 and #60 by sequencing the PCR amplified products corresponding to exons 7 and 9, and the product amplified using primers that flanked the deleted region (Figure 6.6), either directly or following TA cloning. Figure 6.7 shows the breakpoints of the deletions in clones #1, #25, #60, while the breakpoints in clone #B were not characterised.

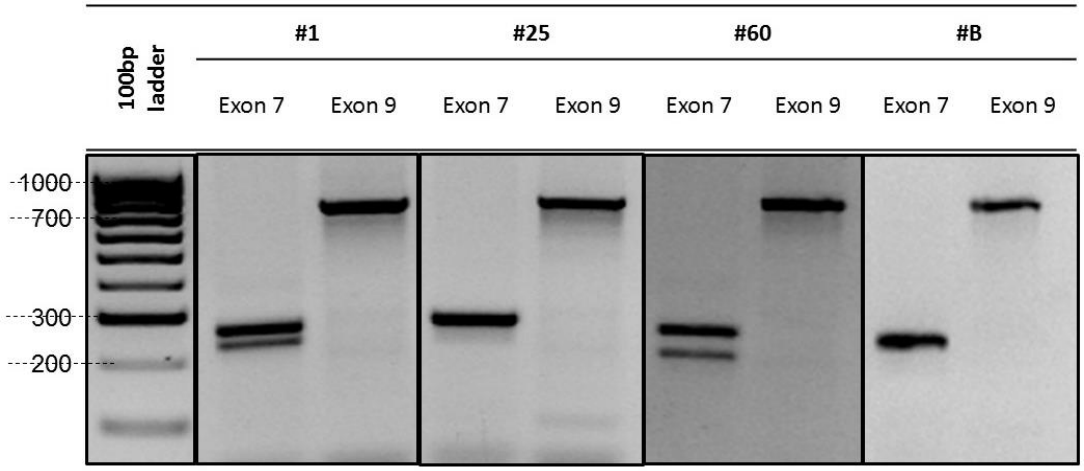


Figure 6.6 Amplification of CRISPR/Cas9 targeted exons in selected *FLI1* knockdown clones

Amplified PCR products corresponding to *FLI1* exons 7 and 9 were separated in 2% agarose. The expected sizes of the exon 7 and exon 9 amplicons are 254 and 711 bp respectively.

Reference sequence			
Ex7-Ex9	3923bp	TTGCCACAG**GTCCTC**CCCTTGGAGGGGCACAAACGATC>>>>>>3350bp>>>>>>AATGACGGACCCCGATGAGGTGGCCAGGCGCTGGGGCGAGCGGA	
EX7	255bp	TTTCCCTCTTGCCACAG**GTCCTC**CCCTTGGAGGGGCACAAACGATCAGTAAGAATACAGAGCAACGGCCCCAGCCAGGTACCTGCC	
EX9	712bp	**TGGGAGGGGACCAACGGGGAGTTCAAAATGACGGACCCCGATGAGGTGGCCAGGCGCTGGGGCGAGCGGAAAAGCAAGCCCAAC	
Clone#1			
Ex7-Ex9	539bp	TTGCCACAG**GTCCTC**CCCTTG----------AGGTGGCCAGGCGCTGGGGCGAGCGGA	
	~800bp	TTGCCACAG**GTCCTC**CCCTT------AGGATCTAACCTCGCTTATTATA>>>290bp>>>AATGACGGACCCC----------CAGGCGCTGGGGCGAGCGGA	
EX7	255bp	TTTCCCTCTTGCCACAG**GTCCTC**CCCTTGGAGGGGCACAAACGATCAGTAAGAATACAGAGCAACGGCCCCAGCCAGGTACCTGCC	WT
	~224bp	TTTCCCTCTTGCCACAG----------TAAGAATACAGAGCAACGGCCCCAGCCAGGTACCTGCC	31bp deletion
EX9	712bp	**TGGGAGGGGACCAACGGGGAGTTCAAAATGACGGACCCCGATGAGGTGGCCAGGCGCTGGGGCGAGCGGAAAAGCAAGCCCAAC	WT
		**TGGGAGGGGACCAACGGGGAGTTCAAAATGACGGACCCC----------CAGGCGCTGGGGCGAGCGGAAAAGCAAGCCCAAC	11bp deletion
Clone#25			
Ex7-Ex9	539bp	TTGCCACAG**GTCCTC**CCCTTG----------AGGTGGCCAGGCGCTGGGGCGAGCGGA	
		TTGCCACAG**GGGAGGACCT**----------GTGGC----------GCTGGGGCGAGCGGA	
EX7	255bp	TTTCCCTCTTGCCACAG**GTCCTC**CCCTTGGAGGGGCACAAACGATCAGTAAGAATACAGAGCAACGGCCCCAGCCAGGTACCTGCC	WT
EX9	712bp	**TGGGAGGGGACCAACGGGGAGTTCAAAATGACGGACCCCGATGAGGTGGCCAGGCGCTGGGGCGAGCGGAAAAGCAAGCCCAAC	WT
Clone#60			
Ex7-Ex9	539bp	TTGCCACAG**GTCCTC**CCCTTGG----------AGGTGGCCAGGCGCTGGGGCGAGCGGA	
EX7	255bp	TTTCCCTCTTGCCACAG**GTCCTC**CCCTTGGAGGGGCACAAACGATCAGTAAGAATACAGAGCAACGGCCCCAGCCAGGTACCTGCC	WT
	~210bp	TTTCCCTCTTGCCACAG**GTCCTC**CCC----------AGCCAGGTACCTGCC	45bp deletion
EX9	712bp	**TGGGAGGGGACCAACGGGGAGTTCAAAATGACGGACCCCGATGAGGTGGCCAGGCGCTGGGGCGAGCGGAAAAGCAAGCCCAAC	WT

Figure 6.7 Mapping of genomic edits in selected *FLI1* knockdown clones
Guide sequences are indicated in bold font and highlighted in either yellow (targeting exon 7) or blue (targeting exon 9). Proto-spacer adjacent motif (PAM) sequences are shown in bold, italic font, with green highlight. Genomic changes are shown in red font. Dashed red lines indicated deleted sequence.

6.3.5 Transcriptome analysis of differentiated and undifferentiated Dami cell clones

Dami cell clones #1, #25, #60 and #B were selected for transcriptome analysis in parallel with two WT Dami single cell clones, C1 and C4, which were selected as they showed minimal differences in *FLI1* expression relative to the polyclonal Dami cell line (Figure 6.5). Transcriptomes were analysed, both before and following differentiation using PMA and TPO. Transcriptomic profiling was usually performed on two samples, derived from consecutive passages, of each clone (*FLI1* knockdown: #1, #25, #60; WT control: C1 and C4), both before and following differentiation for six days. The transcriptome of only one sample of clone #B was analysed without differentiation of the cells.

6.3.5.1 Effect of PMA and TPO on gene expression in wild-type Dami cell clones

Transcriptomic profiling highlighted 19,937 transcripts which were differentially expressed in WT Dami cells (C1 and C4) following treatment with TPO and PMA (≥1.5 or ≤-1.5 fold log change in expression after differentiation and $p<0.05$). These included 1,676 coding transcripts which were significantly upregulated and 5,608 coding transcripts significantly downregulated in differentiated cells relative to untreated cells (Table 6.3, Figure 6.8, Supplementary Data #3).

Table 6.3 Numbers of differentially expressed transcripts following differentiation of Dami cells

	Downregulated after differentiation	**Upregulated after differentiation**
No. of differentially expressed transcripts (1.5-fold change, p< 0.05)	19,937	
	8,898	11,039
No. of differentially expressed coding transcripts (1.5-fold change, p< 0.05)	7,284	
	5,608	1,676

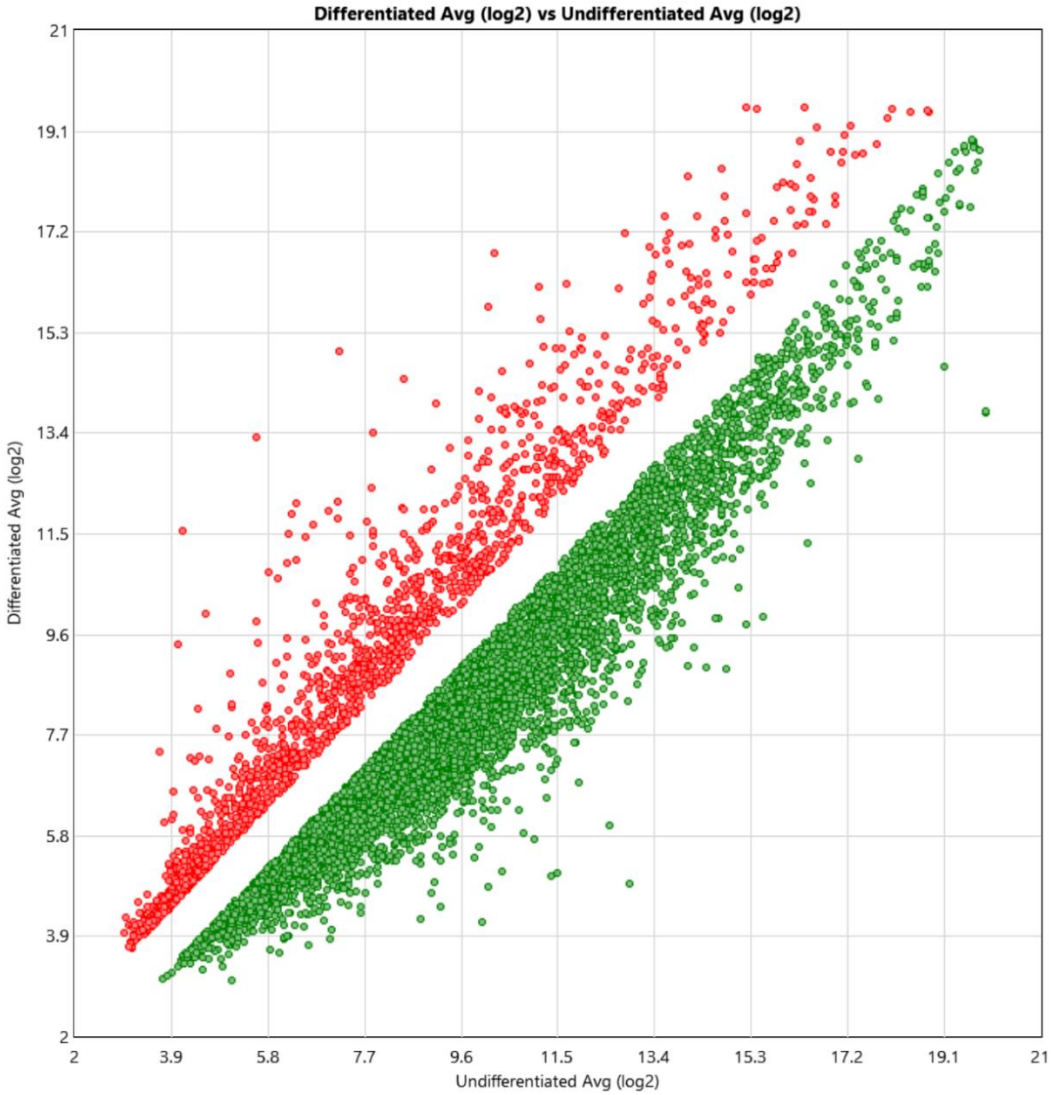


Figure 6.8 Scatter plot of 7,284 coding transcripts that were differentially expressed following differentiation of Dami cells relative to undifferentiated cells
The 1,676 upregulated transcripts are shown in red, and the 5,608 downregulated transcripts are shown in green.

Functional annotation analysis is an approach which, for a given list of genes, allows identification of subsets or clusters of those genes that encode proteins which are related by function, biological pathway or interaction. Using the functional annotation analysis tool in DAVID, this approach was applied to assess the changes in gene expression that occurred due to differentiation in Dami cells. Thus, functional annotation analysis of the 3,000 transcripts showing the greatest change in expression with differentiation (i.e. the 1,500 most upregulated and 1,500 most downregulated genes) identified 109 significant gene clusters, some of which had associated GO terms relating to erythrocyte differentiation, platelet activation and vacuolar transport (Supplementary Data #4).

6.3.5.2 Effect of *FLI1* knockdown on gene expression in Dami cells before and after differentiation

Changes in gene expression after *FLI1* knockdown were investigated in Dami cells both before and following differentiation. Compared to WT Dami cells, 1,385 and 817 genes were differentially expressed after *FLI1* knockdown in untreated cells and following differentiation respectively (Table 6.4, Figure 6.9, Supplementary Data #5 & #6).

Table 6.4 Numbers of differentially expressed transcripts following *FLI1* knockdown in differentiated and undifferentiated Dami cells

	Undifferentiated Dami		Differentiated Dami	
	Down-regulated	**Up-regulated**	**Down-regulated**	**Up-regulated**
No. of differentially expressed transcripts (1.5-fold change, p<0.05)	3,386		2,724	
	2,013	1,373	1,395	1,329
No. of differentially expressed coding transcripts (1.5-fold change, p<0.05)	1,385		817	
	1,163	222	489	328
	2,052*			

*150 transcripts were represented in both conditions.

The 2,052 genes that were differentially expressed following *FLI1* knockdown either before or following differentiation were subjected to functional annotation analysis using DAVID. This identified 62 clusters which had significantly associated GO annotations. Interestingly, the five clusters with the highest enrichment scores had associated GO terms that included ‘haemostasis’, ‘wound healing’, ‘platelet activation’, ‘membrane-bound vesicle’, ‘extracellular vesicle’ and ‘platelet activation’ (Table 6.5, Supplementary Data #7).

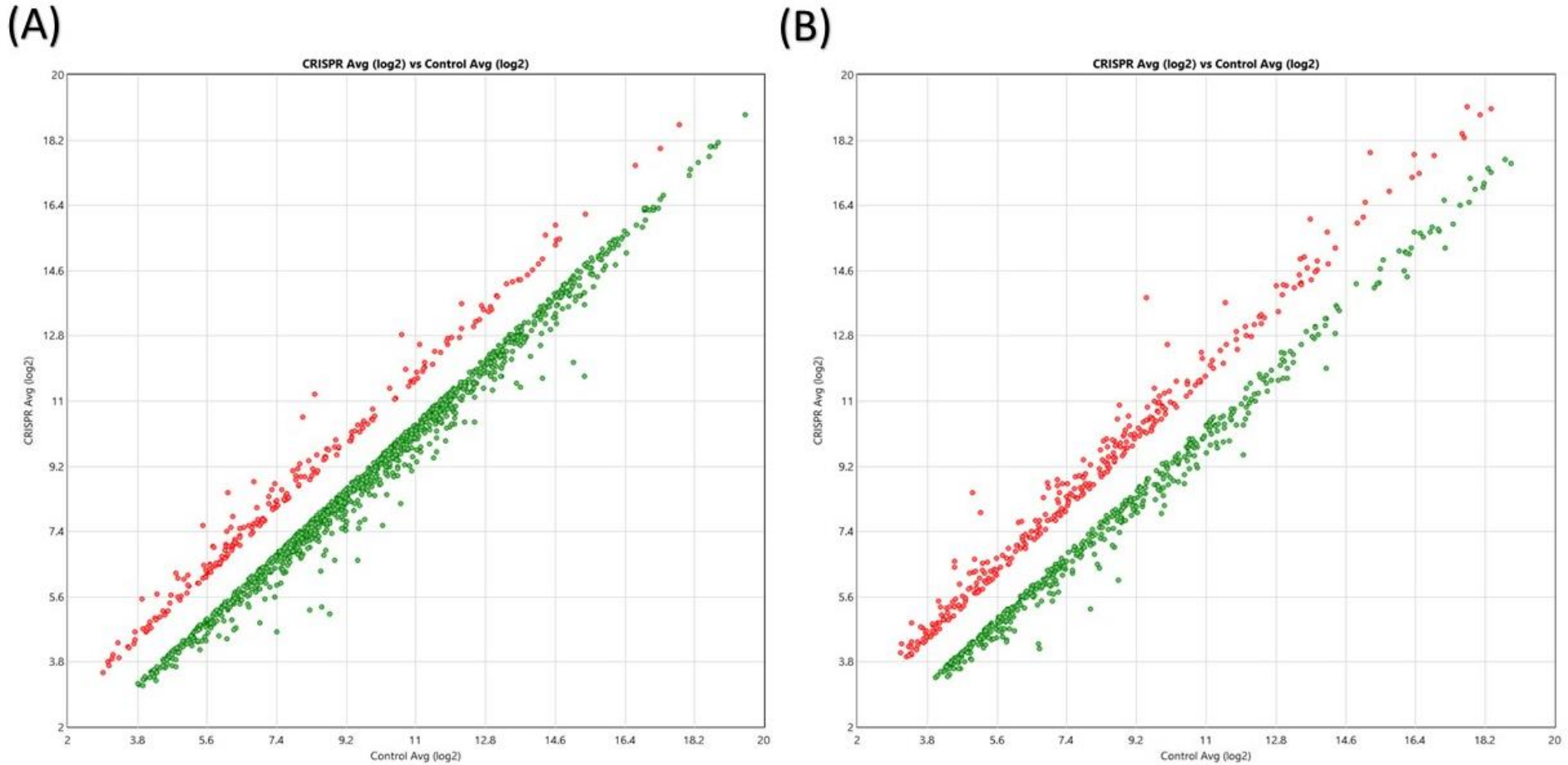


Figure 6.9 Scatter plots of coding transcripts that were differentially expressed in Dami cell clones following *FLI1* knockdown
A total of (A) 1,385 coding transcripts in untreated and (B) 817 coding transcripts in differentiated conditions were differentially expressed in *FLI1* knockdown clones relative to wild-type clones. Upregulated genes are shown in red and downregulated genes are shown in green.

Table 6.5 Top five gene clusters identified by functional annotation analysis of 2,052 genes that were differentially expressed following *FLI1* knockdown in untreated and differentiated Dami cells

Cluster	Enrichment score	Clustered Gene Ontology terms	Gene count
1	6.6	Wound healing / platelet activation / haemostasis / response to wounding / blood coagulation / coagulation / regulation of body fluid levels	119
2	5.24	Nucleosome / protein-DNA complex / nuclear chromosome part / DNA packaging complex / nuclear chromosome / chromosome organisation / nuclear chromatin / chromosome / chromatin / chromosomal part	197
3	4.73	Membrane-bound vesicle / extracellular vesicle / extracellular organelle / extracellular exosome / extracellular region part / extracellular region	544
4	3.83	Platelet activation / platelet aggregation / homotypic cell-cell adhesion	39
5	3.81	Establishment of protein localisation / cellular localisation / protein localisation / protein transport / intracellular transport / macromolecule localisation / establishment of localisation in cell / cellular macromolecule localisation / cellular protein localisation / intracellular protein transport / protein localisation to organelle / single-organism intracellular transport / single-organism cellular localisation / protein targeting / establishment of protein localisation to organelle / protein import	397

6.3.5.3 Platelet gene expression in subjects with a heterozygous *FLI1* defect

While this work was ongoing, platelets became available from two affected members of a previously studied family with an inherited bleeding tendency characterised predominantly by a defect in dense granule secretion, both of whom carried the c.1028A>G:p.Tyr343Cys *FLI1* variant. The platelet RNA isolated from the patients and three matched control subjects was subjected to transcriptome analysis using the Clariom^{TM} D Assay_human chips by Dr Simon Webster, Haemostasis Group, Sheffield. Analysis of coding transcript expression, adopting the same criteria as those used above (≥1.5; ≤-1.5-fold log change, $p<0.05$) highlighted a total of 2,836 differentially expressed coding transcripts (1,834 upregulated; 1,002 downregulated). These data were kindly made available for comparative analysis as part of this study by Dr Simon Webster.

6.3.6 Comparative studies

Comparison of gene expression data for untreated and differentiated Dami cells in which *FLI1* was knocked down with that from platelets expressing a heterozygous *FLI1*

gene defect, and with the WES data from 22 index cases with platelet secretion defects should allow prioritisation of candidate genes for investigation of function.

6.3.6.1 Comparison of candidate genes identified by whole exome sequencing and following *FLI1* knockdown

WES analysis of samples from 22 patients with platelet secretion defects identified variants in 1,130 candidate genes (see chapter 3). Comparison of these genes with the list of 2,052 genes that were differentially expressed following *FLI1* knockdown in Dami cells identified 135 shared genes (Table 6.6). Of these, 62 genes had previously been reported to be regulated by FLI1 or by a transcription factor complex which included FLI1 (GATA1/2, RUNX1, FLI1, and SCL) in MKs (Tijssen et al., 2011).

Table 6.6 Candidate genes identified by whole exome sequencing and also showing differential expression in response to *FLI1* knockdown

112 down-regulated	***ACOT8, ACSF3,*** *ADAMTS9, ALKBH5,* ***AP2A1, ARHGEF18,*** *ATAD3B, ATG4D,* ***ATP7B,*** *BLOC1S3, BLVRB,* ***CABIN1,*** *CARS2, CC2D1A,* ***CCDC51,*** *CCDC88B,* ***CDC16,*** *CDK16, CHURC1-FNTB,* ***CLIC6, COPE, COTL1, CSK,*** *DCHS1, EML3,* ***EPN1,*** *FAM160B2,* ***FBL,*** *FBLN2,* ***FERMT3, FHOD1,*** *FLNA,* ***GATAD2A,*** *GCDH, GMIP,* ***HDAC5, HIP1, HNRNPUL2,*** *IFT80, IGFBP2, IL12RB2, ILVBL,* ***INPP4B,*** *IP6K2, IPO13, IRF3, LAMTOR4,* ***LMNA,*** *MAEA, MAN2B1, MAPK3, MCM5, MED16,* ***MRPS34,*** *MYO1E, NUCB1,* ***P2RX1, PAQR7,*** *PCIF1, PCNXL2,* ***PDLIM7, PFKFB3,*** *PGLS, PHKA2,* ***PLCG2, PLEKHG3,*** *PLEKHG4, PLTP, PRKAR2A, PROSER2,* ***PSRC1,*** *PVRL2,* ***RCOR1,*** *ROMO1, RRBP1, RRP36,* ***RUVBL2, SCFD2, SCNM1,*** *SDSL, SEMA4B, SEMA6C,* ***SIPA1,*** *SLC4A11, SLC7A1, SPOPL, SRM, SSRP1, STEAP3,* ***SUGT1,*** *SULT1A2, TBC1D9B, TC2N,* ***TKT, TLN1,*** *TSC22D1,* ***TTYH3,*** *TUBB8, TYK2,* ***UBR7,*** *UPF1,* ***UROD, VAC14,*** *VPS52, WDR91,* ***WRNIP1, XPNPEP1,*** *YLPM1,* ***ZBTB45,*** *ZNF358, ZNF385A,* ***ZNF592***
23 up-regulated	***AAK1, C18orf32, CAPG, CHST11, CMTM6, ESAM,*** *GPATCH8, KRT79,* ***LAPTM4B, MACF1,*** *MICAL1, MTMR3,* ***NEK6, ORMDL1, PDE3A,*** *PKLR,* ***RALGDS, SLC24A3, SLC8A3, ST8SIA6,*** *WDPCP,* ***ZBTB26,*** *ZNF385D*

Genes regulated by FLI1 (Tijssen et al., 2011) are indicated in bold.

6.3.6.2 Comparison of genes identified by whole exome sequencing and genes differentially expressed in FLI1-deficient platelets

Similarly, comparison of the 2,836 genes which were differentially expressed in FLI1-deficient platelets with the list of 1,130 candidate genes highlighted by WES in patients with defects in platelet secretion identified 215 shared genes, 92 of which had been previously described to be regulated by FLI1 (Tijssen et al., 2011) (Table 6.7). Functional annotation analysis of these 215 genes identified four significant gene clusters (Table 6.8, Supplementary Data #8).

Table 6.7 Candidate genes identified by whole exome sequencing and also showing differential expression in FLI1-deficient platelets

74 down-regulated	*AARS,* ***AFTPH,*** *ALAD, ANK2, ANKRD27,* ***ANO2, ARAP1,*** *BACH1,* ***BCOR,*** *BICD1, C10orf76,* ***CACNB1, CAMTA1,*** *CAPN11,* ***CDKN1A, CREM,*** *DGKI, DYNC1I1,* ***ERGIC1, ETV6,*** *FBXO11, FOCAD,* ***FSTL4, GFI1B,*** *GPATCH8, HADHB, HECTD1, HGD, HIST1H2AH,* ***HIST1H2BD,*** *IGFBP2,* ***KDM5B,*** *KRT18,* ***LAPTM4B,*** *LCORL, LRMP, MCF2L2,* ***MEF2D,*** *NBEA, NID2, ODF3, PAM,* ***PHF14, PLCG2,*** *POTEE,* ***PPHLN1, PPIL3,*** *PPP1R15A,* ***PTGIR,*** *PYGL,* ***RGS3,*** *RIN2,* ***RUNX1, SCFD2,*** *SH3TC2,* ***SHANK2, SLC2A3, SMOX,*** *STARD9,* ***STK39,*** *TAF1D, TCF20, THEM5, TOX,* ***TTC13, UBA7,*** *UBR2, WRN, XYLT1,* ***ZBTB45***, ***ZC3H4,*** *ZNF385D, ZNF462, ZNF506*
141 up-regulated	*ABCB6, ABCC3, ADIPOR1,* ***AKAP7,*** *ALKBH5,* ***AMZ2,*** *API5,* ***ARL6IP5, ATAD2B,*** *ATP11B,* ***B4GALT3, BMP2K,*** *BTBD7,* ***C18orf32,*** *C18orf8,* ***C1orf27,*** *C21orf58,* ***C6orf62,*** *CALCOCO2, CARD6, CCBL2, CCDC175, CCP110, CCT6A, CDK12, CHURC1-FNTB,* ***COTL1,*** *CRTAP, CTBS,* ***CTSZ,*** *DAPK1,* ***DARS, DDOST,*** *DEK,* ***DENND4A,*** *DENR, DZIP1,* ***ELAVL1, ERICH1,*** *EXOC5, FYCO1,* ***GAB1, GATAD2A, GCOM1,*** *GMPS, GOLGB1, GORASP1, GPD1L, GRIPAP1, HADH, HBD, HMGB1, HMGB2, HN1,* ***HNRNPUL2, HOMER2,*** *HYOU1,* ***INPP4B,*** *IRAK4,* ***ITGB3,*** *KIAA0100, KIDINS220,* ***LMNB2,*** *LXN, MORC1, MORF4L1,* ***MPV17, MRVI1,*** *MTR, NAB1, NAGA, NCAPD2,* ***NEK4,*** *NIPA1, NMRK1, NPHP3,* ***NRP2, NSUN2, PABPC1, PAK2,*** *PANK2,* ***PDE4DIP,*** *PEX19,* ***PGM2, PITRM1,*** *PLTP,* ***PLXNB3, POU2F1,*** *PPT2,* ***PTPN12, PTPRJ,*** *RAB3GAP1, RANBP10, RB1, REXO2,* ***RNF6,*** *RNPS1, ROS1, SAV1,* ***SCFD1, SERPINB9, SLC24A3,*** *SLC25A20,* ***SLC30A9, SLC35F5,*** *SLC45A3,* ***SNAP29, SNF8, SNRPB2,*** *SNX31, SPOPL,* ***SPTBN1,*** *SRP14,* ***ST8SIA6,*** *TANGO2, TBC1D13, TBC1D9B, TC2N, TCF25,* ***TEX2,*** *TFDP2, TMEM183A,* ***TRAPPC6B, TREML2,*** *TRIM10, TRPM7,* ***TTYH3,*** *TXNDC16,* ***UBE2Q2, UBQLN1, UBR7,*** *VAPA, VEZT, VNN1,* ***VPS39, WASL,*** *XIRP2,* ***XPOT,*** *YWHAH, ZNF418,* ***ZNF592***

Genes regulated by FLI1 (Tijssen et al., 2011) are indicated in bold.

Table 6.8 Significant gene clusters identified by functional annotation analysis of 215 candidate genes identified by whole exome sequencing and also showing differential expression in FLI1-deficient platelets

Cluster	Enrichment score	Clustered Gene Ontology terms*	Gene count
1	2.23	ER to Golgi vesicle-mediated transport / establishment of vesicle localisation / vesicle localisation / Golgi vesicle transport / vesicle targeting / vesicle organisation / COPII-coated vesicle budding / membrane budding / establishment of organelle localisation / organelle localisation	32
2	1.63	Vesicle organisation / organelle fusion / vesicle fusion / organelle membrane fusion / SNARE binding / membrane fusion / single-organism membrane fusion	15
3	1.49	Negative regulation of platelet-derived growth factor receptor signalling pathway / regulation of platelet-derived growth factor receptor signalling pathway / platelet-derived growth factor receptor signalling pathway / negative regulation of cell proliferation	7
4	1.37	Vacuole fusion / organelle fusion / autophagosome maturation / positive regulation of vacuole organisation / vacuole organisation / regulation of vacuole organisation / autophagosome / macroautophagy / positive regulation of macroautophagy / autophagy	18

* The first 10 Gene Ontology terms are listed for each cluster.

6.3.6.3 Comparison of genes showing differential expression in FLI1-deficient platelets and in response to *FLI1* knockdown

Comparison of the 2,836 differentially expressed genes in FLI1-deficient platelets with the 2,052 genes that were differentially expressed in Dami cells following *FLI1* knockdown identified 390 shared genes. Of the 186 genes that were differentially expressed in the same direction in both platelets and Dami cells, 74 genes had been previously shown to be regulated by FLI1 (Tijssen et al., 2011) (Table 6.9). Functional annotation analysis of the 186 genes that were differentially expressed in the same direction in both *FLI1* deficient Dami cells and platelets identified six significant gene clusters (Table 6.10, Supplementary Data #9).

Table 6.9 Genes showing differential expression in the same direction in FLI1-deficient platelets and in Dami cells following *FLI1* knockdown

110 down-regulated	***ABCA3, ACOT7,*** *ACP6, ACSF2,* ***ADCY6,*** *AP1S3,* ***AP2S1,*** *BSDC1,* ***C1orf116,*** *C1orf21,* ***CABLES1,*** *CARD19, CLN6,* ***COMMD7, COX4I1,*** *CTNS, CTSW, CUL4A,* ***DLST, DNM1,*** *DOK2, DPYD,* ***EIF2B2, ELOF1,*** *ENO2, ERI3, ERI3-IT1,* ***FAH,*** *FAM58A,* ***FCER1G, FRMD3,*** *GRB14,* ***GRIK4,*** *GRTP1, HIST1H2AB, HIST1H2AD, HIST1H2AE, HIST1H2AM, HIST1H2BB, HIST1H2BJ,* ***HIST1H2BL,*** *HIST1H3D, HIST1H3G, HIST1H3H, HIST1H3J, HIST2H2BD, HIST2H2BE, HIST2H2BF, HIST2H4A, HIST2H4A, HIST2H4B, HIST2H4B, HMG20B, HSD17B3, IGFBP2,* ***KDM4B, LGALS12,*** *LOC650293,* ***LYL1, LYST, MAFK, MEA1,*** *MIEF1, MIR4701, MIR6823, MIR7641-2,* ***NDUFA11,*** *PAPSS1, PFKFB4, PHEX,* ***PIH1D1, PIK3C3, PLCG2, PRKCD,*** *PTENP1-AS,* ***RASGRP2, RECK, RGS9,*** *RHCE, RHD, RNU5F-1,* ***SCARB1, SCFD2,*** *SLA2,* ***SLC37A1,*** *SNX2, SPHK1, SRY, SUSD3, SUV39H1,* ***TMCC2,*** *TMEM9,* ***TNFAIP8L1, TPRA1,*** *TRBV26OR9-2, TREML1,* ***TSPAN15,*** *TUBA4A, TWF2, U3, UBE2E3, VKORC1, VN1R110P, VTRNA2-1, WDR75, WFDC1,* ***XPO7, ZBTB45,*** *ZNF317,* ***ZNF672***
76 up-regulated	*AK3, ALS2CR12, APLF,* ***ARCN1, B3GNT2, C18orf32,*** *CCDC117, CCND1,* ***CD36, CDC27,*** *CEP57L1,* ***CITED2,*** *CTCFL, DCBLD2,* ***EHD2,*** *FAM60BP,* ***GCNT1, GFOD1, GTDC1,*** *GYPB, HERC2P7, HLA-E, ID2,* ***IDE,*** *ITGA2, KCTD13,* ***LSM3,*** *MAGT1,* ***MBD5, MED4, MINPP1, NCOA4, NEK1, NOL7, OCIAD1,*** *PABPC1P9, PBX2,* ***PCMT1,*** *PDCD6IPP2,* ***PELI2, PLEKHF2,*** *PLXNA3, RCL1, RHAG, RPL12P32, RPL23AP45, RPL35AP33, RPL6P12,* ***RPS3A,*** *RPS7P8, RSAD2, SC5D,* ***SDHC,*** *SELT, SELT.1, SIAE, SKA2, SKP1P1,* ***SLC24A3, SLC7A5, SNCA,*** *SNORD73A,* ***SNX16, ST8SIA6,*** *STAU2,* ***SURF4, TMEM154,*** *TMEM167B, TNFSF13B,* ***UBLCP1,*** *UQCRB, WHAMMP1, ZBTB44, ZDHHC20, ZDHHC21,* ***ZMYM5***

Genes regulated by FLI1 (Tijssen et al., 2011) are indicated in bold.

Table 6.10 Significant gene clusters identified by functional annotation analysis of 186 genes that were differentially expressed in the same direction in FLI1-deficient platelets and in Dami cells following *FLI1* knockdown

Cluster	Enrichment score	Clustered Gene Ontology terms*	Gene count
1	5.25	Blood coagulation / coagulation / haemostasis / wound healing / regulation of body fluid levels / response to wounding	20
2	2.70	Nucleosome / DNA packaging complex / protein-DNA complex / nuclear nucleosome / chromatin organisation / nucleosome assembly / chromatin silencing / protein heterodimerisation activity / chromatin assembly / nuclear chromatin	106
3	1.92	Membrane-bounded vesicle / extracellular exosome / extracellular vesicle / extracellular organelle / extracellular region part / extracellular region /	57
4	1.62	Nucleoside metabolic process / purine ribonucleoside metabolic process / purine nucleoside metabolic process / glycosyl compound metabolic process / ribonucleoside metabolic process / purine ribonucleoside triphosphate metabolic process / ribonucleoside triphosphate metabolic process / purine nucleoside / triphosphate metabolic process / purine ribonucleoside monophosphate metabolic process	50
5	1.58	Nitrogen utilisation / ammonium transmembrane transport / organic cation transport / ammonium transmembrane transporter activity / ammonium transport / cation transmembrane transport / nitrogen compound transport / cation transmembrane transporter activity	13
6	1.37	Vacuolar membrane / endosome membrane / endosomal part / vacuole / vacuolar part / late endosome / endosome / early endosome membrane / early endosome	19

* The first 10 Gene Ontology terms are listed for each cluster.

6.3.6.4 Candidate genes identified through whole exome sequencing, and differentially expressed in both FLI1-deficient platelets and following *FLI1* knockdown in Dami cells

Finally, comparison of the candidate genes identified through WES, which were also differentially expressed in FLI1-deficient platelets and following *FLI1* knockdown in Dami cells highlighted 23 genes, 14 of which are known to be regulated by FLI1 (Tijssen et al., 2011) (Table 6.11, Figure 6.10). Of these 23 genes, only seven were found to exhibit differential expression in the same direction in both the Dami cell model, and in FLI1-deficient platelets, three were upregulated (*C18orf32*, *SLC24A3*, and *ST8SIA6*) and four were downregulated (*IGFBP2*, *PLCG2*, *SCFD2,* and *ZBTB45*).

Table 6.11 Candidate genes identified through whole exome sequencing, and differentially expressed in FLI1-deficient platelets and following *FLI1* knockdown in Dami cells

ZNF385D, ***SCFD2,*** ***LAPTM4B,*** *GPATCH8, IGFBP2,* ***PLCG2,*** ***ZBTB45,*** ***UBR7,*** *ALKBH5, PLTP,* ***COTL1,*** ***SLC24A3,*** ***GATAD2A,*** *CHURC1-FNTB,* ***HNRNPUL2,*** *SPOPL,* ***ST8SIA6,*** *TBC1D9B,* ***ZNF592,*** ***C18orf32,*** ***INPP4B,*** *TC2N,* ***TTYH3***

Genes regulated by FLI1 (Tijssen et al., 2011) are indicated in bold. Genes differentially expressed in the same direction in FLI1-deficient platelets and following *FLI1* knockdown in Dami cells are underlined.

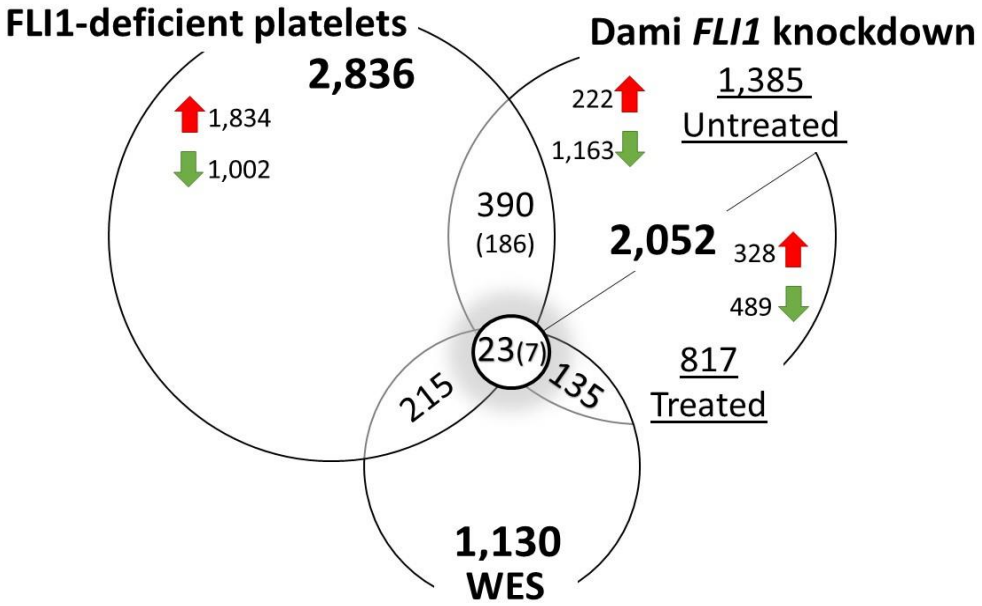


Figure 6.10 Numbers of candidate genes identified by whole exome sequencing, and by gene expression analysis in FLI1-deficient platelets and Dami cells following *FLI1* knockdown

The numbers in parentheses indicate genes that are differentially expressed in the same direction. Red arrows signify upregulated genes, while green arrows represent downregulated genes.

6.3.7 Assessment of off-target effects

One of the drawbacks of using CRISPR/Cas9 as a gene editing tool is the potential for off-target effects. In this study, high gRNA specificity was a criterion at the design stage. Despite this, the Zhang lab tool highlighted 14 other coding regions as potential off-target sites for the gRNAs, five with the gRNA targeting exon 7 and nine with the gRNA targeting exon 9. One of these potential off-target sites had three mismatches with the gRNA sequence, while the remainder had four mismatches (Table 6.12).

Although DNA sequencing of these loci would directly confirm or exclude off-target effects, the absence of a significant reduction in the expression of these genes in CRISPR-edited clones when compared to WT cells is an indirect way of exploring

whether off-target effects have occurred. However, it should be noted that a significant reduction in expression of an alternative target could indicate that it has been edited as a result of an off-target effect or that it has been downregulated as a result of *FLI1* knockdown.

Of the possible off-target genes, only the *CALCR* (Calcitonin receptor) gene showed a significant reduction in expression in the undifferentiated CRISPR-edited clones, while *KLHL38* (Kelch Like Family Member 38) and *TNS2* (Tensin 2) both showed reduced expression following differentiation (Table 6.12).

Table 6.12 Predicted exonic off-target sites for the guide RNAs used to knockdown *FLI1* highlighted using the Zhang lab tool

Off-target site "Gene"	Number of mismatches	FLI1 Untreated conditions		FLI1 Differentiated conditions	
		Fold change	p-value	Fold change	p-value
Guide RNA targeting exon 7 GATCGTTTGTGCCCCTCCAAGGG Guiding quality score 83					
GCAT	4	-1.06	0.5943	-1.05	0.4958
EVC2	4	1.07	0.5793	-1.14	0.4222
SDK1	4	-1.01	0.4493	1.16	0.2151
KLHL38	4	-1.12	0.3616	**-1.3**	**0.0362**
COL14A1	4	1.01	0.4567	1.18	0.3488
Guide RNA targeting exon 9 AATGACGGACCCCGATGAGGTGG Guiding quality score 92					
FBXL19-AS1	4	-1.23	0.2458	-1.07	0.8212
ERG	3	1.37	0.9156	-1.98	0.3314
CADPS	4	1.1	0.6547	1.17	0.3215
STAT5A	4	-1.13	0.0517	1.07	0.8877
TNS2	4	-1.2	0.5820	**-1.71**	**0.0073**
SPPL2B	4	-1.22	0.0824	-1.17	0.4387
FNDC3B	4	-1.23	0.3294	-1.67	0.7012
CALCR	4	**-1.43**	**0.0016**	1.23	0.3090
KMT2D	4	-1.1	0.2768	1.02	0.6258

Proto-spacer adjacent motif (PAM) sequences are underlined.

6.4 Discussion

The study described in chapter 3 of this thesis highlighted the shortcomings of using WES as a genetic diagnosis tool to identify the underlying causative defects in unrelated patients with heterogeneous IPDs. Thus, while likely candidate defects were identified in a small number of cases with IPDs that were characterised by defects in platelet secretion, a median of 70 candidate variants was detected in the majority of cases. Rather than systematically investigating each of these candidate defects in turn, an alternative approach was required to identify novel candidate genes that could contribute to the bleeding disorder in these cases. Given the essential role that FLI1 plays in megakaryopoiesis and the association of *FLI1* defects with abnormal platelet secretion, it was hypothesised that knockout/down of *FLI1* would lead to changes in expression of genes having a role in platelet granule biogenesis and secretion, furthermore, that defects in these genes would be represented among patients with IPDs that were characterised by defects in platelet secretion.

A CRISPR/Cas9 approach, which involved introducing two DSBs into *FLI1* and deletion of the intervening sequence by non-homologous end joining, was therefore used to knockout/down *FLI1* expression in the megakaryocytic Dami cell line prior to transcriptome analysis to identify changes in expression of genes that may be important in platelet granule biogenesis and secretion. CRISPR/Cas9 editing in mammalian cells using this approach has previously been reported to be effective for introducing deletions ranging from 1.3 Kb to over 1000 Kb in size (Canver et al., 2014). Although the use of two gRNAs increases the potential for off-target effects, it also has advantages. Firstly, it allows cheap, quick and straightforward screening for clones with the desired edit by conventional PCR amplification across the deletion breakpoints, which can then be analysed by agarose gel electrophoresis. Secondly, the use of two gRNAs increases the likelihood of target gene knockout by introducing the intended genomic deletion or other small indels within coding regions.

Following transfection of Dami cells with CRISPR plasmids and enrichment of the successfully transfected cells, single Dami cell clones were initially screened for alterations in *FLI1* by PCR amplification of a DNA fragment that spanned the region targeted for deletion. The degree of *FLI1* knockdown was then evaluated in edited

clones at both RNA and protein levels, and for some clones, the locations of the DSB points were also defined. The use of two gRNAs for CRISPR-mediated targeting of coding regions was reported to be highly efficient, as the intended genomic deletions were found to occur at a greater than expected frequency in edited clones (Bauer et al., 2015; Canver et al., 2014). Thus, one study reported successful mono- or bi-allelic deletion of between 2.8 and 4.5 Kb of DNA in 12-61% edited clones (Canver et al., 2014). In this study, 11% of the transfected clones had the desired deletion of approximately 3.9 Kb of *FLI1*. This lower editing rate could be explained by a reduced efficiency of the gRNAs used to introduce DSBs, or due to the essential role of FLI1 in the cells, hence lower tolerance of genetic alterations. The failure of several single cell clones to survive following expansion after sorting would support the latter suggestion, as does the observed lower proliferative capacity and/or increased apoptosis of embryonic stem cells in an *in vivo Fli1* knockout model (Hart et al., 2000). Previous studies involving targeting of essential gene transcripts in *in vitro* and *in vivo* models have reported the development of host resistance, presumably as a physiological necessity to cope with the hostile perturbation (Ajiro et al., 2015; Gu et al., 2011; Tang et al., 2006). Dosage compensation (increased transcription of a target gene from a single allele to result in the same expression level) and genetic compensation (changes in RNA or protein levels of another gene(s) that can functionally compensate for the loss of function) have been also described in response to gene knockout (reviewed in El-Brolosy & Stainier (2017)).

Studies have shown that gRNA sequence mismatches are tolerated by the CRISPR/Cas9 system, raising the potential for undesired off-target effects when CRISPR-based approaches are used for gene editing (Hsu et al., 2013). Comparison of the transcript levels for the predicted off-target loci in *FLI1* edited clones with those in WT clones revealed significant reductions in *CALCR*, *KLHL38* and *TNS2* expression, despite the presence of four mismatches between the gRNA sequences and the complementary regions in these genes (Table 6.12). Given that four mismatches are less likely to be tolerated by the CRISPR/Cas9 system (Hsu et al., 2013), and the absence of a consistent reduction in expression of these genes both before and following differentiation, the likelihood of these being true off-target effects is low, though it would be of interest to sequence these three genes to confirm this.

Transcriptomic profiling was undertaken for four Dami cell clones which showed *FLI1* knockdown due to genomic editing and expressed approximately 50% or less of the levels of FLI1 observed in WT Dami cell clones, and two WT single cell Dami clones. The transcriptomes were examined both before and following treatment with PMA and TPO to induce megakaryocytic differentiation.

The transcriptomic profile of WT Dami cells identified 7,284 coding transcripts that were differentially regulated in WT clones in response to differentiation. Functional annotation analysis of the 3,000 most up- or downregulated genes, revealed, among others, enrichment for GO terms relating to erythrocyte differentiation, platelet activation and vacuolar transport, which is not surprising given that Dami cells are derived from the blood of a patient with megakaryoblastic leukaemia, as well as their ability to differentiate to form proplatelet-like extensions and platelet-like particles containing α- and dense granules (Briquet-Laugier et al., 2004; Greenberg et al., 1988; Lev et al., 2011).

Similarly, comparison of the transcriptomic profile of Dami cells showing *FLI1* knockdown with that of WT cells, before and following differentiation, identified a total of 2,052 genes that were differentially expressed. Functional annotation analysis of these genes identified 62 significant clusters, of which those showing the highest enrichment scores had associated GO terms that included ‘haemostasis’, ‘wound healing’, ‘platelet activation’, ‘nucleosome’, ‘membrane-bound vesicle’, ‘extracellular vesicle’ and ‘platelet activation’, reflecting once again the regulatory role of FLI1 in multiple aspects of MK and platelet biology.

In addition to the gene expression data for Dami cells showing *FLI1* knockdown, and the WES findings for patients with platelet secretion disorders, platelet transcriptome profiles were available for two members of a family affected by a bleeding disorder, whose platelets showed a profound loss in dense granule secretion, both of whom also carried a *FLI1* defect. This allowed comparative analysis to identify and prioritise novel FLI1-regulated genes that could be implicated in platelet granule biogenesis and secretion. This approach highlighted several genes of interest. An exhaustive discussion of the potential role of each of these in platelet biology is not possible here. However, in the remainder of this chapter, I have highlighted the genes which are

known to be regulated by FLI1 (Tijssen et al., 2011) and have previously been associated with IPDs, along with those for which there is evidence supporting a role in platelet biology, as well as those where the evidence is more tenuous (see tables below). The final part of this chapter focuses on those genes that are shared across the three data sets that were used for the comparative analysis.

Comparison of those genes which were differentially expressed after *FLI1* knockdown in Dami cells with those genes that harboured defects in patients with platelet secretion disorders highlighted 135 genes that were shared between the two groups, of which, 62 were known to be regulated by FLI1. As shown in Table 6.13, of these, two genes *FERMT3* and *P2RX1* had been previously associated with IPDs, while several others had been implicated in platelet secretion (*CSK*, *FHOD1*, *VAC14*) and other aspects of platelet function (*PDE3A, PDLIM7*, *PLCG2*, *SLC8A3, TLN1*). The two datasets also shared a subset of genes, which have not been implicated in IPDs or directly with platelet formation or function, and a selection of these, which may be worthy of further investigation for potential roles in platelet biology has been included (see Table 6.13).

Table 6.13 Selection of FLI1-regulated genes which were differentially expressed after *FLI1* knockdown in Dami cells and which harboured defects in patients with platelet secretion disorders*

	Gene	Role of the corresponding gene product	Reference
IPDs and/or platelet secretion defects	*FERMT3* (Fermitin Family Member 3/ Kindlin-3)	Encodes kindlin-3, which mediates linkage of the cytoskeleton to αIIbβ3 triggering platelet activation and aggregation	(Kuijpers et al., 2009; Moser et al., 2008)
	P2RX1 (Purinergic Receptor P2X 1)	Acts as an ATP-gated ion channel	(Oury et al., 2000)
Important in platelet secretion/ granule formation	*CSK* (C-Terminal Src Kinase)	Negatively regulates SFKs inhibiting integrin activation in resting platelets	(Obergfell et al., 2002)
	FHOD1 (Formin Homology 2 Domain Containing 1)	Acts as a key regulator of platelet stress fibre formation	(Thomas et al., 2011)
	VAC14 (Vac14, PIKFYVE Complex Component)	Interacts with NBEAL2 in human MK. Some *NBEAL2* variants causing Grey platelet syndrome disrupt binding to Vac14	(Mayer et al., 2018)
Other aspects of platelet biology	*PDE3A* (Phosphodiesterase 3A)	Mediates platelet aggregation by blocking the inhibitory effect of cAMP	(Feijge et al., 2004)
	PDLIM7 (PDZ And LIM Domain 7)	Involved in organising the actin cytoskeleton by regulating cycling between the GTP/GDP-bound states of Arf6	(Urban et al., 2016)
	PLCG2 (Phospholipase C Gamma 2)	A transmembrane signalling enzyme that, through the production of second messengers, transmits signals from multiple platelet receptors across the cell membrane to activate platelets	(Li et al., 2010)
	SLC8A3 (Solute Carrier Family 8 Member A3)	Encodes K+ -independent Na^+/Ca^{2+} exchanger, NCX3, which promotes transient calcium influx to increase platelet cytosolic calcium during collagen activation	(Roberts et al., 2012)
	TLN1 (Talin 1)	A cytoskeletal protein that binds to the cytoplasmic domain of the β3 subunit and is required for inside-out activation of platelet integrin	(Nieswandt et al., 2007; Tadokoro et al., 2003)
Other genes of potential interest	*AAK1* (AP2 Associated Kinase 1)	During endocytosis, it phosphorylates the AP2 complex to enhance its affinity for membrane protein sorting signals	(Ricotta et al., 2002)
	AP2A1 (Adaptor-Related Protein Complex 2 Subunit Alpha 1)	Subunit of AP2 adaptor complex that facilitates clathrin-mediated endocytosis	(Jackson et al., 2010)
	ARHGEF18 (Rho/Rac Guanine Nucleotide Exchange Factor 18)	Regulates actin and myosin distribution through RhoA signalling at the junctional complex which regulates tight junction assembly and epithelial morphogenesis	(Terry et al., 2011)
	CABIN1 (Calcineurin Binding Protein 1)	Regulates synaptic vesicle endocytosis	(Lai et al., 2000)
	CAPG (Capping Actin Protein, Gelsolin Like)	Involved in regulating actin-based movement in macrophages which coincides with calcium oscillations in the formation of membrane protrusions (ruffling), phagocytosis, and vesicle movement within the cytoplasm (rocketing)	(Witke et al., 2001)
	CLIC6 (Chloride Intracellular Channel 6)	CLIC4 has been found to regulate apical exocytosis through retromer- and actin-mediated endocytic trafficking in renal cells	(Chou et al., 2016)
	COPE (Coatomer Protein Complex Subunit Epsilon)	One of the proteins of the COPI vesicular coat complex that mediates transport from the Golgi apparatus to the ER and within the Golgi stacks	(Béthune & Wieland, 2018)

	COTL1 (Coactosin Like F-Actin Binding Protein 1)	Regulates actin dynamics of T-cells to promote lamellipodial protrusion at the immune synapse towards antigen-presenting cells or target cells	(Kim et al., 2014a)
	EPN1 (Epsin 1)	Adaptor protein for synaptic vesicle endocytosis at neuronal synapses	(Kyung et al., 2016)
	HIP1 (Huntingtin-interacting protein 1)	Involved in clathrin-mediated endocytosis	(Gottfried et al., 2010)
	INPP4B (Inositol polyphosphate 4-phosphatase type II)	Regulates signalling associated with endocytic trafficking	(Chew et al., 2016)
	LAPTM4B (Lysosomal Protein Transmembrane 4 Beta)	An endosomal transmembrane protein that regulates lysosomal sorting and degradation	(Tan et al., 2015)
	MACF1 (Microtubule-Actin Crosslinking Factor 1)	Cross-links actin to other cytoskeletal proteins and binds to microtubules to mediate cell migration, focal adhesions, signalling and vesicle transport from the trans-Golgi network to the cell periphery	(Hu et al., 2016)
	PSRC1 (Proline And Serine Rich Coiled-Coil 1)	A microtubule-associated protein that orchestrates microtubule dynamics and directional cell migration	(Zhang et al., 2013)
	*RALGDS (*Ral Guanine Nucleotide Dissociation Stimulator)	Responsible for Ral-dependent exocytosis of WPBs in endothelial cells	(Rondaij et al., 2008)
	SCFD2 (Sec1 Family Domain Containing 2)	Member of the Sec1/Munc18 family of proteins that cooperate with SNARE complexes in membrane fusion events through their interactions with syntaxins	(Halachmi & Lev, 1996; Li et al., 2013)
	SLC24A3 (Solute Carrier Family 24 Member 3)	Belongs to the SLC24–Na^+/(Ca^{2+}–K^+) exchanger family of proteins known to play an important role in intracellular calcium homeostasis	(Schnetkamp, 2013)

*Genes highlighted in this table are derived from a manual search of the GeneCards website (https://www.genecards.org/) [accessed 2018]. AP2; adaptor protein 2, Arf6; ADP-ribosylation factor 6, COPII; coat protein complex II, GPCR; G-protein-coupled receptor, IPDs: inherited platelet bleeding disorders, MK; megakaryocyte, SFKs; Src family kinases, SNARE; soluble N-ethylmaleimide-sensitive factor attachment receptor, WPBs; Weibel-Palade bodies.

Of the genes that were differentially expressed in platelets from patients with a *FLI1* defect, 215 were also represented among the genes that harboured defects in patients with platelet secretion disorders. Out of these, 92 genes were known to be regulated by FLI1 (Table 6.7). As shown in Table 6.14, these included five genes which have previously been associated with IPDs (*ETV6, GFI1B, ITGB3*, *PTPRJ, RUNX1*) and others which encode proteins involved in platelet secretion (*SLC2A3, SNAP29*), and other aspects of platelet function *(CDKN1A*, *GAB1*, *MRVI1*, *PAK2, PLCG2, PTGIR, WASL*). There were also several genes that could potentially have a role in regulating the platelet secretory pathways, though further investigation is required to explore this possibility (Table 6.14).

There were 186 genes that were differentially expressed in the same direction in platelets from patients with a *FLI1* defect and following *FLI1* knockdown in Dami cells. Further inspection of the subset of 74 genes known to be regulated by FLI1 (Table 6.9), highlighted genes associated with IPDs (*CD36, LYST*, *RASGRP2, SCARB1*), genes encoding proteins that have a role in platelet secretion *(PIK3C3, PRKCD*) and other aspects of platelet biology (*ADCY6, CABLES1, COMMD7*, *PLCG2, FCER1G**,*** *SNCA*) as well as genes encoding proteins that could potentially have a role in platelet function (Table 6.15). Interestingly, inspection of the 186 genes that were differentially regulated (Table 6.9) identified several groups of related genes (e.g. *DOK2, DOK3* / *SUSD3, SUSD6* / *TMEM108**,*** *TMEM9*, *TMEM154, TMEM167B* / *ZDHHC20, ZDHHC21* / *ZNF317, ZNF672**,*** *ZNF98*), which could reflect direct or indirect regulation of several genes in similar biological pathways by FLI1. Of particular interest, was the cluster of 14 histone genes. Although the exact role of histones in platelets remains to be fully elucidated, initial studies have shown that histones H1, H2, H3 and H4 induce platelet activation (Carestia et al., 2013) via a mechanism that is regulated by albumin (Lam et al., 2013).

Table 6.14 Selection of FLI1-regulated genes which were differentially expressed in platelets from patients with *FLI1* defects and which harboured defects in patients with platelet secretion disorders*

	Gene	**Role of the corresponding gene product**	**Reference**
IPDs and/or platelet secretion defects	*ETV6* (ETS Variant 6)	Transcriptional repressor that has a role in embryonic development, angiogenesis, haematopoiesis and megakaryopoiesis	(Zhang et al., 2015)
	GFI1B (Growth Factor-Independent 1B Transcriptional Repressor)	Transcription factor which has a pivotal role in haematopoiesis (erythropoiesis and megakaryocyte development)	(Stevenson et al., 2013)
	ITGB3 (Integrin Subunit Beta 3)	Subunit of the platelet αIIbβ3 receptor which mediates platelet adhesion and aggregation and triggers "outside-in" signalling. αIIbβ3 plays a role in proplatelet formation by regulating actin remodelling; Subunit of the αvβ3 receptor that has a role in platelet adhesion, and aggregation, and endothelial cell adhesion migration and angiogenesis	(Ghevaert et al., 2008; Nurden & Caen, 1975)
	PTPRJ (Protein Tyrosine Phosphatase, Receptor Type J)	A receptor-like protein tyrosine phosphatase that is critical for initiating GPVI signalling in platelets through activation of Src family kinases	(Marconi et al., 2018)
	RUNX1 (Runt Related Transcription Factor 1)	Transcription factor required for haematopoietic stem cell generation; It is also essential for the maturation of T and B lymphocytes; In megakaryopoiesis, it acts as a core regulator of early and late MK differentiation	(Heller et al., 2005)
Important in platelet secretion/ granule formation	*SLC2A3* (Solute Carrier Family 2 Member 3)	Mediates glucose utilization and glycogenolysis in platelets, promotes α-granule release, platelet activation and postactivation functions	(Fidler et al., 2017)
	SNAP29 (Synaptosome Associated Protein 29)	Contributes to the regulation of platelet α-granule secretion and thrombus stability	(Williams et al., 2016)
Other aspects of platelet biology	*CDKN1A* (Cyclin-Dependent Kinase Inhibitor 1A)	Regulates MK differentiation	(Rubinstein et al., 2012)
	GAB1 (GRB2 Associated Binding Protein 1)	Signalling molecule in collagen-stimulated PI3K signalling pathway	(Moraes et al., 2010)
	MRVI1 (Murine Retrovirus Integration Site 1 Homolog)	Plays a central role in NO/cGMP dependent inhibition of platelet aggregation and thrombus formation	(Antl et al., 2007)
	PAK2 (P21 (RAC1) Activated Kinase 2)	A critical effector of Rho GTPases (CDC42 and RAC1) that plays a role during spreading of platelet lamellipodia (early shape change) and shedding of platelet microvesicles	(Crespin et al., 2009; Vidal et al., 2002)
	PLCG2 (Phospholipase C Gamma 2)	See Table 6.13	
	PTGIR (Prostaglandin I2 Receptor)	A GPCR that plays a role in inhibiting platelet activation	(Li et al., 2010)
	WASL (Wiskott-Aldrich Syndrome Like)	Major mediator of early rapid actin cytoskeleton responses, including filopodia formation; Positively regulates demarcation membrane system development and proplatelet formation being a direct target of CDC42	(Palazzo et al., 2016; Shcherbina et al., 2001)
Other genes of potential interest	*AFTPH* (Aftiphilin)	Binds to clathrin AP-1 and 2; Involved in the response of WPBs to secretagogues and release of their contents; Involved in clathrin-mediated trafficking in neurons	(Burman et al., 2005; Lui-Roberts et al., 2008)
	ANO2 (Anoctamin 2)	Transmembrane calcium-activated chloride channel that facilitates the scrambling of phospholipids between leaflets of the membrane bilayer leading to the release of extracellular vesicles	(Whitlock & Hartzell, 2017)

	ARAP1 (ArfGAP With RhoGAP Domain, Ankyrin Repeat And PH Domain 1)	Regulates Arf-, Rho-, and Cdc42-dependent actin cytoskeleton-related activities including cell shape change, filopodia and stress fibre formation	(Miura et al., 2002)
	CACNB1 (Calcium Voltage-Gated Channel Auxiliary Subunit Beta 1)	Auxiliary- cytosolic- subunit of the voltage-gated calcium channel that controls trafficking of the main α1-subunit to the plasma membrane, its regulation and its gating properties; Interacts with a number of proteins involved in diverse aspects of cell signalling; Other family members interact with synaptic proteins	(Rima et al., 2016)
	C1orf27 (*ODR4* (Odr-4 GPCR Localization Factor Homolog)	Transmembrane protein in the endoplasmic reticulum that has a role in the maturation, trafficking or localisation of a subset of GPCRs to sensory cilia	(Chen et al., 2014; Dwyer et al., 1998)
	COTL1 (Coactosin Like F-Actin Binding Protein 1)	See Table 6.13	
	CTSZ (Cathepsin Z)	Cleaves regulatory motifs of several integrins (e.g. β2, β3 subunits) affecting their function (cell adhesion, phagocytosis, maturation, proliferation, activation, and cytoskeletal rearrangment) in macrophages, T lymphocytes, dendritic cells, and neuronal cells	(Kos et al., 2009)
	DENND4A (DENN Domain Containing 4A)	DENND4 ortholog regulates the polarised secretion of basement membrane components; DENND4A-C proteins have GEF activity towards Rab10, a Rab involved in the regulation of basolateral trafficking in polarised cells	(Marat et al., 2011)
	ERGIC1 (Endoplasmic Reticulum-Golgi Intermediate Compartment 1)	Cycling membrane protein that may have a role in transport between the ER and Golgi	(Breuza et al., 2004)
	FSTL4 (Follistatin Like 4)	Regulates synaptic plasticity in neuronal cells	(Suzuki et al., 2018)
	HOMER2 (Homer Scaffold Protein 2)	Interaction with CDC42 influences actin cytoskeleton organization and cell morphology; modifies calcium signalling through a GPCR to regulate the frequency of calcium oscillations in pancreatic acini	(Shin et al., 2003; Shiraishi-Yamaguchi et al., 2009)
	INPP4B (inositol polyphosphate 4-phosphatase type II)	See Table 6.13	
	PDE4DIP (Phosphodiesterase 4D Interacting Protein)	Involved in tethering non-centrosomal microtubules to Golgi membranes	(Wu et al., 2016)
	LAPTM4B (Lysosomal Protein Transmembrane 4 Beta)	See Table 6.13	
	PLXNB3 (Plexin B3)	Acts as a transmembrane receptor for semaphorin 5A that suppresses human glioma cell motility and morphology through Rac1 and the actin cytoskeleton by mediating disassembly of F-actin stress fibres, and disruption of focal adhesions	(Li et al., 2012)
	PTPN12 (Protein Tyrosine Phosphatase, Non-Receptor Type 12)	Involved in immunity, vascular development, adhesion, cell migration and embryonic viability being a key regulator of signalling pathways in cell-cell and cell-extracellular matrix interactions; Acts as binding partner for Filamin A	(Duval et al., 2015; Rhee et al., 2014; Souza et al., 2012)
	RGS3 (Regulator Of G Protein Signalling 3)	Negatively regulates intracellular calcium release via inactivation of Gα; Other family members (RGS1, 2, 13, 14, 16, 18) have roles in haematopoiesis, megakaryopoiesis and platelet function	(Freisinger et al., 2010; Louwette et al., 2012)

	SCFD1 (Sec1 Family Domain Containing 1)	Member of the Sec1/Munc18 family that cooperate with SNARE complexes in membrane fusion events through their interactions with syntaxins; Involved in vesicle transport from ER to Golgi; Binds to ER-localized Syntaxin5 where it is possibly involved in stabilizing the open conformation of the SNARE, a prerequisite for Q-SNARE complex assembly and sorting into COPII-coated vesicles	(Adolf et al., 2018; Dascher & Balch, 1996; Li et al., 2013)
	SCFD2 (Sec1 Family Domain Containing 2)	See Table 6.13	
	SHANK2 (SH3 And Multiple Ankyrin Repeat Domains 2)	Acts as binding partner for dynamin and cortactin	(Okamoto et al., 2001)
	SLC24A3 (Solute Carrier Family 24 Member 3)	See Table 6.13	
	SNF8 (SNF8, ESCRT-II Complex Subunit)	Component of the ESCRT-II complex that is required for endocytosis and lysosomal degradation of transmembrane proteins; SNF8 acts as a regulator of the calcium-permeable cation channels TRPC6 (Transient receptor potential canonical channel)	(Babst et al., 2002; Carrasquillo et al., 2012)
	TRAPPC6B (Trafficking Protein Particle Complex 6B)	A component of the TRAPP tethering complexes that function in specific stages of inter-organelle traffic and also can activate the GTPase Rab1; *TRAPPC6A* could be involved in hypopigmentation	(Brunet & Sacher, 2014)
	TREML2 (Triggering Receptor Expressed On Myeloid Cells Like 2)	Another family member (TREML1) is known to regulate granule construction in platelets and is associated with IPDs due to platelet secretion defect	(Nurden et al., 2008)
	VPS39 (VPS39, HOPS Complex Subunit)	A subunit of the HOPS-tethering complex that acts as a Rab7-binding subunit; the HOPS complex promotes clustering, tethering and fusion of late endosomes with lysosomes and vacuoles by binding and stabilising SNARE complexes preventing their dissociation	(Kleine Balderhaar & Ungermann, 2013; Starai et al., 2008)

*Genes highlighted in this table are derived from a manual search of the GeneCards website (https://www.genecards.org/) [accessed 2018]. AP; adaptor protein , COPII; coat protein complex II , ER; endoplasmic reticulum, ESCRT; endosomal sorting complex required for transport, GEF; guanine-nucleotide exchange factor, GP; glycoprotein, GPCR; G-protein-coupled receptor, HOPS; homotypic fusion and protein sorting, IPDs: inherited platelet bleeding disorders, MK; megakaryocyte, NO; nitric oxide PI3K; phosphatidylinositol 3-kinase, SNARE; soluble N-ethylmaleimide-sensitive factor attachment receptor, TRAPP; transport protein particle, WPBs; Weibel-Palade bodies.

Table 6.15 Selection of FLI1-regulated genes which are differentially expressed in platelets from patients with a *FLI1* defect, and following *FLI1* knockdown in Dami cells*

	Gene	**Role of the corresponding gene product**	**Reference**
IPDs and/or platelet secretion defects	*CD36* (CD36 Molecule)	Multi-ligand receptor that induces platelet adhesion and activation	(Kashiwagi et al., 1994)
	LYST (Lysosomal Trafficking Regulator)	Scaffolding protein that facilitates membrane events, including both fission and fusion	(Nagle et al., 1996)
	RASGRP2 (RAS Guanyl Releasing Protein 2)	Guanine nucleotide exchange factor that is critical for activation of small GTPases, including RAP1	(Canault et al., 2014)
	SCARB1 (Scavenger Receptor Class B Member 1)	Plays a role in megakaryopoiesis and platelet production and in regulating platelet activation and aggregation	(Vergeer et al., 2011)
Important in platelet secretion/ granule formation	*PIK3C3* (Phosphatidylinositol 3-Kinase Catalytic Subunit Type 3)	Lipid kinase that catalyses the conversion of phosphatidylinositol into PI3P; Controls granule biogenesis, intracellular trafficking, migration and platelet production in MK; Regulates platelet secretion and thrombus growth	(Valet et al., 2017)
	PRKCD (Protein Kinase C Delta)	Signalling protein that regulates platelet functional responses including dense granule secretion and TXA2 generation downstream of PARs and GPVI receptors	(Chari et al., 2009)
Other aspects of platelet biology	*ADCY6* (Adenylate Cyclase 6)	Encodes a protein that belongs to the adenylyl cyclase family, which is required for the synthesis of cAMP downstream of GPCR-Gs subunits that are activated when prostacyclin binds to its receptor	(Smolenski, 2012)
	CABLES1 (Cdk5 And Abl Enzyme Substrate 1)	Regulates HSCs and the process of megakaryopoiesis	(He, 2018)
	COMMD7 (COMM Domain Containing 7)	Positive regulator of thrombus formation in zebrafish after laser injury; Acts as a repressor of transcription of the *NFKB1*	(Burstein et al., 2005; Vermeersch et al., 2018)
	FCER1G (Fc Fragment Of IgE Receptor Ig)	Immunoglobulin receptor that non-covalently binds to GPVI and mediates platelet activation by collagen; Plays a role in mediating signalling via the platelet GPIb-IX-V complex	(Li et al., 2010)
	PLCG2 (Phospholipase C Gamma 2)	See Table 6.13	
	SNCA (Synuclein Alpha)	Inhibits thrombin-induced platelet α-granule release in vitro; Negatively regulates dopamine neurotransmission	(Abeliovich et al., 2000; Park et al., 2002)
Other genes of potential interest	*AP2S1* (Adaptor-Related Protein Complex 2 Subunit Sigma 1)	Component of the AP2 complex that acts in clathrin-mediated endocytosis of the plasma membrane; Plays a role in extracellular-calcium homeostasis	(Nesbit et al., 2013; Ohno, 2006)
	ARCN1 (Archain 1)	Encodes the coatomer subunit delta of COPI that is required for vesicle budding in the early secretory pathway	(Beck et al., 2009)
	ABCA3 (ATP Binding Cassette Subfamily A Member 3)	A lipid transporter involved in the biogenesis of intracellular multi-lamellar vesicles in alveolar epithelial cells	(Yamano et al., 2001)
	EHD2 (EH Domain Containing 2)	Links clathrin-mediated endocytosis to the actin cytoskeleton	(Guilherme et al., 2004)
	DNM1 (Dynamin 1)	Plays a role in regulating fusion pore geometry and kinetics of endo- and exocytotic vesicles; Mutations in DNM2 have been associated with	(Lasič et al., 2017; Züchner et al., 2005)

		subclinically low counts of neutrophils, lymphocytes, erythrocytes and platelets in Charcot-Marie-Tooth disease	
	LYL1 (LYL1, Basic Helix-Loop-Helix Family Member)	Required for survival of adult haematopoietic stem and progenitor cells	(Souroullas et al., 2009)
	MAFK (MAF BZIP Transcription Factor K)	Plays a role in erythropoiesis, megakaryopoiesis and platelet production	(Onodera et al., 2000)
	PLEKHF2 (Pleckstrin Homology And FYVE Domain Containing 2)	Modulates the structure and function of endosomes by a Rab5-dependent mechanism	(Lin et al., 2010)
	SCFD2 (Sec1 Family Domain Containing 2)	See Table 6.13	
	SLC24A3 (Solute Carrier Family 24 Member 3)	See Table 6.13	
	SNX16 (Sorting Nexin 16)	Regulates traffic between early and late endosomal compartments; Regulates recycling and trafficking of E-cadherin	(Hanson & Hong, 2003; Xu et al., 2017)
	SURF4 (Surfeit 4)	Required to maintain the architecture of the ER-Golgi intermediate compartment and Golgi apparatus by controlling COPI recruitment	(Mitrovic et al., 2008)

*Genes highlighted in this table are derived from a manual search in GeneCards website (https://www.genecards.org/) [accessed 2018]. AP; adaptor protein, COPI; coat protein complex I, ER; endoplasmic reticulum, GP; glycoprotein, GPCR; G-protein-coupled receptor, IPDs: inherited platelet bleeding disorders, MK; megakaryocyte, NFKB1; Nuclear Factor Kappa B Subunit 1, PAR; protease-activated receptor, TXA2; thromboxane A2.

Of the genes that were differentially expressed in *FLI1* deficient Dami cells and platelets, twenty-three also harboured defects in patients with platelet secretion disorders. Of these, seven genes were differentially expressed in the same direction in both the Dami cells and platelets (Table 6.11). Thus, *C18orf32* (Chromosome 18 Open Reading Frame 32), *SLC24A3* (Solute Carrier Family 24 Member 3), and *ST8SIA6* (ST8 Alpha-N-Acetyl-Neuraminide Alpha-2,8-Sialyltransferase 6) were all upregulated, while *IGFBP2* (Insulin-Like Growth Factor Binding Protein 2), *PLCG2* (Phospholipase C Gamma 2), *SCFD2* (Sec1 Family Domain Containing 2), and *ZBTB45* (Zinc Finger And BTB Domain Containing 45) were downregulated. Each of these genes will be briefly described below, along with the evidence where it exists, for them having a role in platelet biology.

C18orf32 is a 76 amino acid protein that activates the NFκB (Nuclear factor kappa B) signalling pathway (Matsuda et al., 2003), which is known to have a dual regulatory role in platelet function (Fuentes et al., 2016). Through activation of NFκB signalling, it leads to phosphorylation of SNAP23, which enhances soluble N-ethylmaleimide-sensitive factor attachment receptor (SNARE) complex formation, resulting in membrane fusion and granule release (Karim et al., 2013). C18orf32 is also involved in regulating the posttranslational modification process that anchors glycosylphosphatidylinositol (GPI) to proteins (Liu et al., 2018) in the endoplasmic reticulum. The GPI in the anchored proteins then acts as a sorting signal to transport the proteins through the Golgi apparatus via vesicles to the plasma membrane (Muñiz & Riezman, 2016). C18orf32 has also been shown to interact with glucagon-like peptide 1 receptor (GLP1R) (Huang et al., 2013). GLP1R is a guanine nucleotide-binding G-protein-coupled receptor (GPCR) that signals through a Gs-protein complex and cAMP on ligand binding and attenuates platelet aggregation and thrombosis (Cameron-Vendrig et al., 2016). However, the physiological relevance of the C18orf32-GLP1R interaction remains to be determined. Despite the interesting role of C18orf32 in platelets, the missense defect identified in the corresponding gene c.83T>C:p.V28A in index case F11 was inherited along with a defect in *FLI1* (NM_002017:c.1018C>T:p.R340C), that was more likely to explain the bleeding phenotype (Table 6.16).

SLC24A3 encodes the 644 amino acid NCKX3 protein ($Na^+/Ca^{2+}/K^+$ exchanger 3) that belongs to the SLC24-$Na^+/(Ca^{2+}\text{-}K^+)$ exchanger family of proteins known to play an important role in intracellular calcium homeostasis (reviewed in Schnetkamp (2013)). Members of this family are widely expressed and known to have critical roles in sensory cells (retinal and olfactory), throughout the brain, epidermal melanocytes and in pigment-producing cells. Although very little is known about the physiological role of SLC24A3, a number of variants have been associated with pathophysiological changes in systolic blood pressure (Citterio et al., 2011). In platelets, *SLC24A3* was found to be upregulated in response to aspirin (Voora et al., 2013) and downregulated in platelets from a patient with a *RUNX1* mutation (Sun et al., 2007). Interestingly, *SLC24A3* expression was found to increase with *FLI1* deficiency, which suggests it is normally repressed by FLI1 to regulate intracellular calcium homeostasis in platelets. The missense defect in *SLC24A3* that was identified in index case F7 (c.650C>G:p.S217C) predicted an amino acid substitution within one of the Na+/Ca^{2+} exchanger membrane region domains. However, the index case also carried a candidate genetic defect in *RUNX1* (NM_001001890:c.270+1G>T) that was more likely explain the bleeding phenotype (Table 6.16).

ST8SIA6 encodes the 398 amino acid ST8 α-N-Acetyl-Neuraminide α-2,8-sialyltransferase protein which belongs to the α2,8-sialyltransferase (ST8Sia) family of enzymes and catalyses the transfer of α-2,8-linked disialic acid to glycoconjugates (reviewed in Huang et al. (2017)). ST8SIA6 has been shown to be responsible for modifying the natural killer inhibitory receptor, sialic acid binding Ig-like lectin 7 (Siglec-7) (Avril et al., 2006). Interestingly, Siglec-7 is expressed on the membranes of α-granules and colocalises with CD62P in platelets (Nguyen et al., 2014). Furthermore, activation of platelets causes an increase in Siglec-7 expression, which correlates closely with the increase in CD62P expression (Nguyen et al., 2014). Moreover, when Siglec-7 was cross-linked with its ligand, ganglioside, it was found to promote platelet apoptosis without affecting any other platelet functions (Nguyen et al., 2014). In this study, *ST8SIA6* was found to be upregulated with *FLI1* deficiency. This could potentially lead to an increase in modification of SIGLEC7, which, when cross-linked with its ligand, will drive platelets to apoptosis, contributing to the mild thrombocytopenia associated with *FLI1* defects. The non-frameshift deletion (c.40_42del:p.14_14del) that was identified in *ST8SIA6* in index case F9.1 was not co-

inherited with any other obvious causative gene defects (Table 6.16), supporting the need for further investigation of the role of this gene in platelet granule biogenesis and secretion, and a possible contribution of *ST8SIA6* to the pathogenesis of IPDs.

The 325 amino acid protein encoded by *IGFBP2* is one of six similar IGFBPs (Insulin-like growth factor binding proteins) which bind and modulate the biologic effects of Insulin-like growth factors I and II (IGFI and IGFII). They also have IGF-independent bioactivity mediated through their interaction with cell surface receptors (e.g. α5β1 and αvβ3) or the intracellular actions of IGFBP2 (reviewed in Bach (2018)). IGFBP2, similar to other IGFBPs (IGFBP3-6) has been found to increase intracellular calcium levels (Seurin et al., 2013). Knockdown of *igfbp2* in zebrafish embryos disrupted cardiovascular development and resulted in specific angiogenic abnormalities (Wood et al., 2005), while *Igfbp2* knockout mice were phenotypically normal apart from minor gender-specific changes in bone structure and in the weights of spleen and liver (DeMambro et al., 2008; Wood et al., 2000). The normal phenotype in the knockout mice was explained by an elevation in the level of other IGFBPs to compensate for the loss of *Igfbp2* (Wood et al., 2000). IGFBP2 has been found to support expansion/proliferation, survival and cycling of hematopoietic stem cells (Huynh et al., 2008; Huynh et al., 2011), and migration of human haematopoietic stem and progenitor cells (Bartling et al., 2010). More recently, under turbulent flow, IGFBP2 was found to be released from MKs and to facilitate platelet shedding *ex vivo* (Ito et al., 2018). It could be hypothesised that IGFBP2 plays a role in platelet secretion via IGF-independent mechanisms, that are likely to involve interactions with integrins or other cell surface receptors or which could be mediated through the intracellular action of IGFBP2. Interestingly, this study identified three different non-frameshift deletions in *IGFBP2* in three index cases, F1.1, F12.1 and F17.1. However, the *IGFBP2* defect that was identified in index case F1.1 and their affected relative, F1.2, was co-inherited with a 4 bp deletion in *FLI1* (NM_002017:c.992_995del:p.331_332del) that is likely to explain the bleeding phenotype (Table 6.16). Nevertheless, the presence of *IGFBP2* defects in three index cases with defects in platelet secretion is interesting and warrants further investigation of the potential contribution of *IGFBP2* to platelet granule biogenesis and secretion.

The 1,265 amino acid protein encoded by *PLCG2* (Phospholipase C Gamma 2), PLCγ2, has a well-established role in platelet activation. Activation of PLCγ2 occurs downstream of signalling through the GPIb-IX-V, GPVI and αIIbβ3 receptors and is regulated by phosphatidylinositol 3-kinase (PI3K) (Li et al., 2010). Activation of PLC is critical for efficient platelet activation, as it catalyses the formation of inositol-1,4,5-trisphosphate (IP3) which promotes intracellular calcium release and diacylglycerol (DAG) required for activation of protein kinase C (PKC) (Li et al., 2010). *Plcg2*$^{-/-}$ mice are viable, but exhibit increased perinatal lethality, internal bleeding, reduced B-cell numbers and major functional defects in B-cells, platelets, mast cells and natural killer cells (Wang et al., 2000). However, in addition to PLCγ2, platelets also express the PLCγ1 isoform, which is able to support activation downstream of GPVI in PLCγ2-deficient murine platelets (Suzuki-Inoue et al., 2003). Heterozygous variants of *PLCG2* have been associated with auto-inflammation, antibody deficiency and an immune dysregulation syndrome known as familial cold autoinflammatory syndrome 3 (Ombrello et al., 2012; Zhou et al., 2012). Interestingly, despite the role of *PLCG2* in platelet activation, defects in platelet function or bleeding have not been reported, though affected patients have not been assessed for a bleeding tendency. This study identified non-synonymous heterozygous *PLCG2* variants (c.1712A>G:p.N571S and c.2032G>A:p.D678N) in two index cases (F9.1 and F21.1), both of which predict substitutions of residues located within the Src homology 2 (SH2) domains involved in signal transduction. In the case of F21.1 however, the *PLCG2* defect was co-inherited with a heterozygous *GFI1B* defect (NM_004188: c.289G>A:p.D97N) which, given the association of *GFI1B* with IPDs, is more likely to explain the bleeding phenotype (Table 6.16). The presence of a platelet defect in the knockout mouse model would support a contribution from the two heterozygous *PLCG2* variants to the bleeding phenotype, which requires further investigation.

SCFD2 is a 684 amino acid protein that belongs to the Sec1/Munc18 protein superfamily (Li et al., 2013). By interacting with SNARE proteins, members of this superfamily, which includes the Sec1 family (STXBP1, STXBP2 and STXBP3), Vps45 family (VPS45), Vps33 family (Vps33A and Vps33B), Sly1 family (SCFD1), and MIP3/SCFD2 family (SCFD2), act as indispensable regulators of vesicle fusion in eukaryotic cells (Halachmi & Lev, 1996; Lobingier & Merz, 2012). Defects in members of the Sec1/Munc18 protein superfamily have been found to be associated with IPDs

characterised by abnormalities in platelet granules, platelet secretion and, in some cases, thrombocytopenia. Familial haemophagocytic lymphohistiocytosis type 5 is associated with autoimmune disease, anaemia, thrombocytopenia, abnormalities in all platelet granules, and impaired cytotoxic granule exocytosis in the patient's natural killer cells due to homozygous, compound heterozygous or heterozygous defects in *STXBP2* (Al Hawas et al., 2012; Cote et al., 2009). A homozygous variant of VPS45 was associated with a significant bleeding tendency due to life-threatening thrombocytopenia and platelet dysfunction, life-threatening infections, congenital neutropenia, lack of lysosomes in patients' fibroblasts, primary myelofibrosis, extramedullary haematopoiesis and progressive bone marrow failure (Stepensky et al., 2013). Platelets from affected patients displayed defective platelet aggregation, reduced α-granules and a distorted open canalicular system (Stepensky et al., 2013). A mouse with a homozygous missense mutation in *Vps33a* (vacuolar protein sorting-associated protein 33 A), *buff* mouse, exhibited hypopigmentation and decreased platelet activity, resembling the phenotype of Hermansky-Pudlak syndrome (HPS) in humans (Suzuki et al., 2003). More recently, homozygous *VPS33A* variants were described in a human mucopolysaccharidosis-like condition (lysosomal storage disease), where the characteristic symptoms of HPS, including cutaneous albinism and bleeding diathesis, were not observed (Dursun et al., 2017; Kondo et al., 2016). However, patients developed haematopoietic disorders including anaemia, thrombocytopenia, leukocytopenia as well as recurrent infections and bone marrow hypoplasia (Dursun et al., 2017; Kondo et al., 2016). Mutations in *VPS33B* are associated with the autosomal recessive arthrogryposis-renal dysfunction cholestasis (ARC) syndrome, where enlarged platelets with absent or deficient platelet α-granules are documented (Gissen et al., 2004; Kim et al., 2010). Studies in a murine model of VPS33B deficiency highlighted that VPS33B regulates protein sorting into α-granule destined organelles during megakaryopoiesis (Bem et al., 2015). The role, if any, of SCFD2 in intracellular vesicle fusion remains to be clarified. Interestingly, *SCFD2* downregulation was observed in platelets from a patient with a heterozygous *RUNX1* mutation (Sun et al., 2007). Similarly, in this study, *FLI1* deficiency was associated with reduced *SCFD2* expression, which suggests that *SCFD2* is upregulated by RUNX1 and FLI1, either individually or as a complex, and is required for vesicle formation and/or trafficking. The identification of a heterozygous *SCFD2* variant (c.469C>T:p.P157S) in index case F2.1, in the absence of other obvious likely

candidate gene defects, supports a contribution from this variant to the defect in platelet secretion and bleeding symptoms observed in the index case, which merits further investigation (Table 6.16).

Very little is known about the role of the Zinc finger and BTB domain containing 45 protein encoded by *ZBTB45*. It is proposed to be a regulator for glial differentiation (Södersten et al., 2010) and has also been found to induce cell cycle arrest and to inhibit cell proliferation by activating genes of the p53 pathway (Kim et al., 2011). The missense *ZBTB45* variant (c.976G>A:p.G326R) identified in F2.1 was not co-inherited with candidate defects in other established IPD genes, which increases the likelihood of it being a causative defect. Further work is therefore warranted to investigate the role, if any, of ZBTB45 in platelets and its possible association with IPDs.

The defects in the seven genes (*C18orf32*, *SLC24A3*, *ST8SIA6*, *IGFBP2*, *PLCG2, SCFD2,* and *ZBTB45*) highlighted above were co-inherited in some cases with defects in other genes, for which there is strong evidence of an association with bleeding. Although this reduces the likelihood of them being causative, given the known heterogeneity of IPDs, and the variations in bleeding severity observed among patients carrying the same or similar gene defects, it is possible that they still contribute to the phenotypic expression of the bleeding tendency. Given that this study focused primarily on single unrelated index cases, family studies would help in directing further investigation of the above genes. Alternatively, validation of the association of these genes with platelet secretion defects, which could commence with characterisation of the *PLCG2* variants, could utilise CRISPR/Cas9 knockout/down of the candidate genes, followed by functional characterisation of the resulting clones to determine their role in platelet granule biogenesis and secretion.

In addition to further investigation of the seven genes discussed above, further studies of genes highlighted by other comparisons could be undertaken, particularly those genes that are differentially expressed in platelets from patients with a *FLI1* defect and in Dami cells following *FLI1* knockdown. Transcriptomic profiling of platelets from additional patients carrying *FLI1* variants would increase the power of this approach to detect novel genes involved in platelet secretion by excluding transcriptomic variation that is unrelated to the *FLI1* defect. Interestingly, a very recent publication identified a

novel form of inherited thrombocytopenia which was due to a defect in *PTPRJ* (Protein Tyrosine Phosphatase, Receptor Type J) (Marconi et al., 2018), a gene that was identified as harbouring a candidate defect in the WES analysis of an index case in this study (F11.1), which also showed differential expression in FLI1-deficient platelets, providing support for this approach.

Table 6.16 Defects identified by whole exome sequencing analysis in patients with platelet secretion disorders, which occurred in genes that were differentially expressed in *FLI1* deficient Dami cells and platelets

Gene	Patient	Type of variant	Zygosity	Change	Domain*	CADD _PHRED score	ExAC**	rs number***	Other IPD genes
C18orf32	F11.1+F11.2	NS	Het	NM_001035005:c.83T>C:p.V28A	---	22.8	---	---	*FLI1*, *ABCG8 PTPRJ*
SLC24A3	F7.1	NS	Het	NM_020689: c.650C>G:p.S217C	N-terminal Na^+/Ca^{2+} exchanger membrane region	24.9	0.002447	rs147680736	*RUNX1*, *BLOC1S3*
ST8SIA6	F9.1	NFD	Het	NM_001004470:c.40_42del:p.14_14del	---	---	---	---	---
IGFBP2	F1.1+F1.2	NFD	Het	NM_000597:c.41_42insCCCGCCGCT:p.P14delinsPPPL	---	---	---	---	*FLI1*
IGFBP2	F12.1	NFD	Het	NM_000597:c.61_62insGGCCGCTGC:p.L21delinsRPLL	---	---	---	---	*BLOC1S3*
IGFBP2	F17.1	NFD	Het	NM_000597:c.46_47insAGCTGCTGC:p.P16delinsQLLP	---	---	---	---	---
PLCG2	F9.1	NS	Het	NM_002661:c.1712A>G:p.N571S	N-SH2 domain	22.9	0.006559	rs75472618	---
PLCG2	F21.1	NS	Het	NM_002661: c.2032G>A:p.D678N	C-SH2 domain	22.3	0.0003306	rs541071022	*GFI1B, BLOC1S3*
SCFD2	F2.1	NS	Het	NM_152540: c.469C>T:p.P157S	---	23.5	0.002365	rs144687608	---
ZBTB45	F2.1	NS	Het	NM_032792:c.976G>A:p.G326R	---	25.5	0.002012	rs140831088	---

*Using InterPro - version 7 [accessed October-2018]. **ExAC Browser (Beta) (version 0.3.1) from Exome Aggregation Consortium (http://exac.broadinstitute.org/) [accessed 2018]. ***rs number from the dbSNP database [accessed 2018]. CADD: Combined Annotation Dependent Depletion; Het: heterozygous; IPDs: inherited platelet bleeding disorders; NFD: non-frameshift deletion; NS: non-synonymous single nucleotide variants.

7 Chapter 7. General discussion, final summary and future work

Inherited platelet bleeding disorders (IPDs) are a heterogeneous group of conditions that arise from defects in genes which have a role in platelet production or function, and result in an increased risk of bleeding. This study commenced with a cohort of 34 individuals who had enrolled in the UK Genotyping and Phenotyping of Platelets (UK-GAPP) study, with a clinical diagnosis of unexplained excessive bleeding symptoms that was suspected to be due to a platelet disorder. Prior to studies undertaken in this thesis, extensive platelet phenotyping had been undertaken for all patients, which had resulted in them being categorised into two subgroups. Thus, 12 patients were diagnosed as having a Gi-signalling defect, while platelets from the remaining 22 patients displayed defects in platelet secretion. The study aimed to identify the underlying genetic defects in these patients and to highlight novel genes associated with IPDs. Ultimately, results of this work should, in the future, provide valuable insights into the pathogenesis of IPDs, contribute to our understanding of platelet physiology and allow identification of novel targets for potential therapeutic intervention.

Whole exome sequencing (WES) analysis was undertaken in order to identify candidate genetic defects underlying the IPDs in all patients studied. The bioinformatic pipeline, devised to prioritise candidate defects for further investigation, filtered variants according to their frequency in the population, the variant type, its predicted effect and whether or not it occurred in a gene that is expressed in platelets. These steps reduced the number of candidate variants from an median of 24,774 to approximately 100 for each index case. Further analysis, which grouped patients according to their platelet phenotype and assumed that causative variants in each group would occur within subsets of genes that function in different pathways, which would be specific to each phenotype, resulted in a further reduction in the number of candidate single nucleotide variants (SNVs) per patient. Thus, a median of 70 plausible candidate variants, affecting a total of 1,130 genes, were identified for each of the 22 patients with secretion defects, while a median of 45 candidate variants affecting a total of 545 genes were detected in each of the patients with the Gi-signalling defects. The large number of candidate gene defects in each group highlighted the heterogeneity and complexity of the IPDs and the difficulty in achieving a genetic diagnosis, particularly for single index cases.

Given the interests of the Sheffield Haemostasis group in platelet granule biogenesis and secretion and related disorders, subsequent investigations focused on those 22

patients whose platelets displayed defects in secretion. Of 1,465 potential disease-causing variants that may have contributed to the bleeding disorder, only 16 occurred in genes that have previously been implicated in IPDs and, based on the existing knowledge about their IPD-related conditions, nine of these, affecting *ETV6*, *FLNA*, *FLI1*, *GFI1B*, *ITGB3*, *P2RX1* and *RUNX1* in 8 index cases, were considered likely to contribute to the bleeding tendency.

Interestingly, six of the nine highlighted variants occurred in genes encoding transcription factors. Recent studies have associated defects in the haematopoietic transcription factors *RUNX1*, *GFI1B*, *GATA1*, *FLI1* and *ETV6* with both quantitative and qualitative platelet disorders and variable bleeding symptoms in affected patients (reviewed in Daly (2017)). In this study, novel variants affecting *FLI1* and *ETV6* that were identified by WES in patients with platelet secretion disorders were selected for further characterisation. FLI1 and ETV6 are E26 transformation-specific or E-twenty-six (ETS) transcription factors that are known to have a role in megakaryopoiesis (Hart et al., 2000; Hock et al., 2004; Kawada et al., 2001; Moussa et al., 2010; Starck et al., 2010; Takahashi et al., 2005). FLI1 acts mainly as a transcriptional activator (Rao et al., 1993), while ETV6 is known to act as a transcriptional repressor (Lopez et al., 1999). The association between these two ETS family members and IPDs was established re latively recently (Stockley et al., 2013; Zhang et al., 2015), and while this work was ongoing, several germline variants were identified in both genes in patients with IPDs. All of the *FLI1* variants that have been described to date have been associated with a profound defect in platelet secretion in the affected patients (Poggi et al., 2015; Saultier et al., 2017; Stevenson et al., 2015), while the germline *ETV6* defects have mainly been associated with thrombocytopenia and a predisposition to haematologic malignancy (Melazzini et al., 2016; Moriyama et al., 2015; Noetzli et al., 2015; Poggi et al., 2017; Topka et al., 2015; Zhang et al., 2015).

The association between platelet secretion defects and variants of *FLI1* was first described by the UK-GAPP study group (Stockley et al., 2013). Members of affected families carrying *FLI1* variants (c.1009C>T:p.Arg337Trp and c.1028A>G:p.Tyr343Cys) presented with bleeding disorders as a result of defects in platelet dense granule secretion which, in the majority of cases, were accompanied by mild thrombocytopenia and other immune disorders (Stockley et al., 2013). Initial investigation of these variants revealed them both to have reduced transactivation capacity (Stockley et al.,

2013). In this study, the identification of a third *FLI1* variant (c.1018C>T:p.Arg340Cys) in index case F11.1 that also predicted an amino acid substitution in the DNA binding domain of FLI1 was of interest. In contrast to the two previously described variants, it was associated with a defect in platelet secretion in the absence of thrombocytopenia or any accompanying immune features. During the course of this study, we collaborated with colleagues to investigate a fourth novel *FLI1* variant that also affected codon 340, but which was predicted to result in substitution of arginine by histidine (c.1019G>A:p.Arg340His). Interestingly, the affected members of the family who carried this *FLI1* variant were clinically similar to the cases previously reported by Stockley et al. (2013). Characterisation of the two novel FLI1 variants identified in this study (p.Arg340Cys and p.Arg340His) revealed them both to have reduced transcriptional activity and nuclear accumulation, leading to the conclusion that these variants interfere with the regulation of essential megakaryocyte-specific genes by FLI1 and are likely to explain the bleeding tendency in the affected patients. In addition to megakaryocytes, FLI1 is expressed in endothelial cells and in many haematopoietic lineages (Bastian et al., 1999; Masuya et al., 2005; Mélet et al., 1996; Starck et al., 2010; Suzuki et al., 2013; Zhang et al., 1995; Zhang et al., 2008). This possibly explains the association of *FLI1* defects with other clinical features, including alopecia, eczema, psoriasis and recurrent viral infections in some families. It is possible that the observed heterogeneity of the clinical phenotypes is a reflection of the presence of four coding *FLI1* transcripts which may have tissue-specific differences in expression and activity, about which little is known. It would therefore be interesting to compare the expression of different *FLI1* transcripts and functional activity of FLI1 in endothelial cells, B-lymphocytes, T-lymphocytes, granulocytes, monocytes and other cells. In addition, evaluation of the effects of *FLI1* variants in these cells, for example by comparing the transcriptomic profiles in affected patients with those of control subjects, could help in understanding tissue-specific differences in the role of FLI1.

The novel *ETV6* defect identified in index case F4.1 (c.1288C>T:p.Arg430*), who had been diagnosed as having a platelet secretion defect, was a strong candidate for causing the underlying bleeding disorder, particularly given the recently described association of *ETV6* variants with IPDs and the predicted highly deleterious effects of the SNV compared to all other SNVs identified by WES analysis across the 22 index cases with platelet secretion defects. It was therefore selected for further

characterisation. Interestingly, the identified *ETV6* SNV occurred in a region of the gene that had not previously been associated with IPDs and the predominant clinical feature was a reduction in platelet dense granule secretion rather than the thrombocytopenia associated with previously reported *ETV6* variants (Melazzini et al., 2016; Moriyama et al., 2015; Noetzli et al., 2015; Poggi et al., 2017; Topka et al., 2015; Zhang et al., 2015). The studies reported in chapter 5 showed that the truncated R430*-ETV6 variant was expressed as a stable protein in the nucleus and the cytoplasm of human embryonic kidney (HEK) 293T cells. Also, while no significant difference in the ability of the R430*-ETV6 variant to repress transcription was observed in HEK 293T cells, a small, but significant, reduction in repressor activity was observed in the presence of the R430*-ETV6 variant in Dami cells. Thus, there was insufficient evidence from the *in vitro* studies to support an association of the R430*-ETV6 variant with the bleeding symptoms in the index case with the c.1288C>T *ETV6* variant, and further studies would be required to confirm the pathogenicity of this nonsense *ETV6* variant. The experimental approach used in this study to investigate the *ETV6* variant had some limitations. In particular, overexpression of a cDNA encoding the truncated variant did not allow assessment of the effect of the *ETV6* variant on RNA stability. As the results of the luciferase reporter assays indicated a defect in transrepression of the truncated ETV6 variant in Dami cells, but not in HEK 293T, further studies in Dami cells, particularly to assess expression of an N-terminal tagged truncated variant, would be of interest. More importantly, *ETV6* transcript analysis should be carried out following isolation of platelet RNA from the affected index case to determine whether it is stably expressed *in vivo*. Notwithstanding, the findings of this study emphasise the necessity to exercise caution when interpreting the possible effects of nonsense variants and to undertake appropriate functional studies to confirm bioinformatic predictions where possible.

In this study, recognition of potential disease-causing variants in 8 out of 22 index cases with platelet secretion defects was primarily due to the previously described association of defects in these genes with IPDs. This outcome highlights the limitations of using WES as a genetic diagnosis tool to identify the underlying causative defects in unrelated patients with heterogeneous IPDs. However, the ultimate aim of this study was to identify novel genes that have a role in megakaryopoiesis and platelet function

(mainly secretion), impairment of which could be associated with a bleeding tendency and thus to improve diagnosis of the IPDs in affected patients.

Platelet granule biogenesis occurs within the MKs and results in three main types of platelet granule that differ in size, number and cargo. Additionally, heterogeneity of each granule type, as a result of differences in their size, composition and localisation that are then reflected by differences in transport, release time and ultimately function, has been described (Jena et al., 2017; Peters et al., 2012). Degranulation is the process by which platelets, in response to specific stimuli, release cargo from their storage granules to influence the surrounding microenvironment, with the strength of the stimulation (concentration and potency) controlling the rate and extent of platelet secretion (Chatterjee et al., 2011; Jonnalagadda et al., 2012). The concept of agonist-dependent patterns of released cargo remains controversial (Jonnalagadda et al., 2012; van Holten et al., 2014). Granule biogenesis and secretion is a complex, tightly regulated process that has some degree of redundancy and is mediated by the formation of SNARE (soluble N-ethylmaleimide-sensitive fusion attachment protein receptor) complexes and regulated by a number of SNARE regulators (reviewed in Golebiewska & Poole (2015); Heijnen & Van der Sluijs (2015); Sharda & Flaumenhaft (2018)). Despite the different granule types, their reported heterogeneity and the presence of multiple SNAREs and SNARE regulators in platelets, a single model of platelet granule secretion has been proposed (Golebiewska & Poole, 2015; Heijnen & Van der Sluijs, 2015; Sharda & Flaumenhaft, 2018) (Figure 1.3). In contrast to this model, some studies report that certain SNAREs or SNARE regulators are associated with specific granule types or have differential effects on different types of granule. For example, Rab4 is crucial for the exocytosis of α-granules, while Rab27b is a key regulator of dense granule biogenesis and exocytosis (Shirakawa et al., 2000; Tolmachova et al., 2007). Additionally, in a VAMP8 knock-out mouse model, while mild agonist-evoked release was impaired for all three granule types, release from lysosomes and α-granules was affected to a greater degree than that from dense granules (Graham et al., 2009).

Very little is known about the role of transcription factor defects in platelet granule formation or secretion, though it is likely that their effects are mediated through the differential expression of genes that are normally regulated by these transcription factors. For example, dysregulation of the gene encoding Palladin (*PLDN*), a subunit of BLOC-1 which is involved in granule

biogenesis, was recently found to be just one possible explanation for the platelet dense granule deficiency observed in a patient with *RUNX1* haploinsufficiency (Mao et al., 2017). Indeed, several studies have highlighted genes that were downregulated in a patient with a *RUNX1* defect which could explain the associated platelet phenotype (Aneja et al., 2011; Jalagadugula et al., 2018; Jalagadugula et al., 2011; Jalagadugula et al., 2010; Kaur et al., 2010). Since transcription factors regulate multiple genes, defects in transcription factor genes in patients with IPDs are likely to be associated with a combination of qualitative and quantitative defects in platelets and may also affect other blood cells. Thus, abnormalities in platelet dense granule secretion could represent just one aspect of a bleeding disorder caused by a transcription factor defect. Identification of the genes that are regulated by these transcription factors would improve our understanding of the molecular mechanisms that govern platelet granule biogenesis and secretion and ultimately aid diagnosis of such conditions.

Much of what is already known about platelet granule biogenesis and secretion has originated from the detailed investigation of patients with IPDs that are characterised by granule abnormalities, such as Hermansky-Pudlak syndrome (Ammann et al., 2016; Anikster et al., 2001; Cullinane et al., 2011; Li et al., 2003a; Morgan et al., 2006; Oh et al., 1996; Shotelersuk et al., 2000; Suzuki et al., 2002; Zhang et al., 2003) or Grey platelet syndrome (Albers et al., 2011; Gunay-Aygun et al., 2011; Kahr et al., 2011). Similarly, in this study, the FLI1-related platelet granule secretion defect was utilised to highlight a number of FLI1 effector genes involved in platelet granule function (see chapter 6). *FLI1* knockdown in Dami cells highlighted differentially expressed coding transcripts of 2,052 genes that showed enrichment for Gene Ontology terms relating to ‘haemostasis’, ‘wound healing’, ‘platelet activation’, ‘membrane-bound vesicle’, ‘extracellular vesicle’ and ‘platelet activation’. Similarly, platelet transcriptome data for two members of a family who carried the c.1028A>G *FLI1* variant predicting the p.Tyr343Cys substitution in the DNA binding domain of FLI1, made available by Dr Simon Webster, Sheffield, highlighted differentially expressed coding transcripts of 2,836 genes. Comparison of the platelet transcriptome data with that from *FLI1* deficient Dami cells identified 186 genes which were differentially expressed in the same direction in both datasets. Interestingly, of those genes known to be regulated by FLI1, four were previously associated with IPDs (*CD36*, *LYST*, *RASGRP2, SCARB1*), while other genes were associated with platelet secretion and other aspects of platelet biology (A*DCY6*, *CABLES1*, *COMMD7*, *FCER1G*, *PIK3C3*, *PLCG2*,

PRKCD, *SNCA*), though none of these had been previously associated with IPDs (see table Table 6.15 for references).

Comparison of those genes that were shown by WES analysis to harbour deleterious defects in patients with IPDs (from chapter 3) with those which were differentially expressed either in *FLI1*-deficient Dami cells or platelets (from chapter 6) identified 135 and 215 genes respectively that were shared between the datasets. Among those genes known to be regulated by FLI1, several had previously been associated with IPDs, including *ETV6, FERMT3*, *GFI1B, ITGB3*, *P2RX1, PTPRJ* and *RUNX1*. Others, which were known to be regulated by FLI1 but had not previously been associated with IPDs, played a role in platelet secretion or other aspects of platelet biology (*CDKN1A*, *CSK*, *FHOD1*, *GAB1*, *MRVI1*, *PAK2*, *PDE3A*, *PDLIM7*, *PLCG2*, *PTGIR*, *SLC2A3*, *SLC8A3*, *SNAP29*, *TLN1*, *VAC14*, *WASL*) (see tables Table 6.13 and Table 6.14 for references). Only 23 genes were shared by all three datasets, of which seven were differentially expressed in the same direction in both FLI1-deficient Dami cells and platelets: *C18orf32*, *SLC24A3*, *ST8SIA6, IGFBP2*, *PLCG2*, *SCFD2* and *ZBTB45*. It would be interesting to explore the roles of these genes in platelet granule biogenesis and secretion, as well as their possible contribution to the bleeding symptoms among the affected index cases.

There were several limitations to the work described in this thesis, not least being the lack of availability of index cases and their family members for follow up studies. As mentioned earlier, analysis of platelet RNA from index case F4.1 to determine whether the nonsense *ETV6* variant that they had inherited was stably expressed would have facilitated further investigation of the pathogenicity of this variant. Additionally, the availability of DNA samples from other affected and unaffected family members would have allowed association of candidate gene defects identified by WES with platelet secretion defects. Indeed, WES analysis in DNA samples from other first-degree relatives of index cases would greatly increase the likelihood of genetic diagnosis by WES alone.

The use of *in vitro* cell-based models for functional studies of candidate gene defects was also a limitation. Generally, all established cell lines are induced by genetic alterations to be immortal and able to proliferate indefinitely, which frequently causes them to behave differently to the primary cells from which they are derived. Both HEK

293T cells and megakaryocytic Dami cells were used in the studies described in this thesis. While HEK 293T cells are readily transfected and widely used for *in vitro* investigations, they are derived from kidney cells that express neither *ETV6* nor *FLI1*. On the other hand, Dami cells display many of the morphologic and biochemical features of the megakaryocytic lineage and express both *ETV6* and *FLI1*. However, they require treatment with TPO and PMA to induce differentiation and may not faithfully reflect the differences in the expression of FLI1 and other transcription factors that occur during megakaryopoiesis and platelet production. In addition to the limitations associated with the mammalian cell lines, assessments of the transactivation capacity and cellular localisation of FLI1 and ETV6 variants were performed following overexpression of the variants using naked plasmid DNA, which does not reflect the normal *in vivo* situation.

There were also limitations to the transcriptomic profiling and its related analysis. For instance, analysis of platelet transcriptomes from two related members of a family, both with the same *FLI1* defect, could detect changes in gene expression that were due to the subjects being related, rather than the *FLI1* defect. Similarly, the use of the Dami cell line as a model to evaluate the effect of FLI1 knockdown/out could highlight genetic variations within the cells that do not accurately reflect FLI1 deficiency *in vivo*. In addition, the transcriptome data were derived from *FLI1* knockdown clones and a complete knockout might be more useful for identifying FLI1-regulated genes involved in granule biogenesis and secretion, though a complete knockout may not be viable. Furthermore, as with any CRISPR/Cas9 gene editing approach, the possibility of off-target effects exists and should be considered by sequencing the likely off-target loci in relevant clones. Despite the strong correlation between gene expression profiling data obtained using microarrays and RNA sequence analysis, the latter would overcome many of the technical issues inherent to microarray probe performance, as well as the need for prior knowledge of the sequence. RNA sequence data are also amenable to re-interrogation when new knowledge becomes available.

Appendices

Appendix 1. Genetic basis of inherited platelet defects

Disease	Features		MOI	Protein: Gene (Chromosomal location)
	Platelet and MK related	Others		
ACTN1 -related thrombocytopenia = α-actinin- related disease (Kunishima et al., 2013)	Macrothrombocytopenia	---	AD	Alpha-actinin-1: *ACTN1* (14q24)
ANKRD18A -related thrombocytopenia (Morgan et al., 2013)	Severe thrombocytopenia; Marked impairment of platelet activation by a range of agonists	---	AR	Ankyrin repeat domain-containing protein 18A: *ANKRD18A* (9p13.1)
ANKRD26 -related thrombocytopenia = Autosomal dominant thrombocytopenia = Thrombocytopenia 2 (Gandhi et al., 2003; Pippucci et al., 2011; Punzo et al., 2010)	Many small, hypo-lobulated, dystrophic MKs in BM; Normal or small platelet size; MKs and platelets have particulate cytoplasmic structures, consisting of an accumulation of proteasome complexes and polyubiquitinated proteins that might contribute to mild to moderate thrombocytopenia; Platelets are deficient in GPIa and α-granules	Leukocytosis; Increased haemoglobin levels; Potential association with haematological malignancies, particularly acute myeloid leukaemia	AD	Ankyrin repeat domain-containing protein 26: *ANKRD26* (10p12.1); Serine/threonine-protein kinase greatwall: *MASTL* (10p12.1); Acyl-CoA-binding domain-containing protein 5: *ACBD5* (10p12.1)
ARPC1B -related thrombocytopenia (Kahr et al., 2017)	Microthrombocytopenia; Dense granule deficiency; Defective platelet spreading	Recurrent infections; Eosinophilia; Cutaneous vasculitis; Predisposition to inflammatory diseases	AR	Actin-related protein 2/3 complex subunit 1B: *ARPC1B* (7q22.1)
Arthrogryposis-renal dysfunction cholestasis syndrome (ARC) (Cullinane et al., 2010; Gissen et al., 2004)	Normal platelet count; Enlarged platelets; Absent/deficient platelet α-granules	Arthrogryposis; Renal dysfunction; Cholestasis; High mortality rate in the first year after birth due to severe multisystem defects	AR	Vacuolar protein sorting-associated protein 33B: VPS33B (15q26.1); Spermatogenesis-defective protein 39 homolog: VIPAS39 (14q24.3)
Bernard-Soulier syndrome (BSS) (Berndt et al., 1983; Clemetson et al., 1982; Kunishima et al., 2001; Savoia et al., 2001)	Moderate macrothrombocytopenia; GPIb-IX-V complex severely reduced; Failure of ristocetin-induced aggregation	---	AR/AD	Platelet glycoprotein Ib alpha chain: *GP1BA* (17p13.2); Platelet glycoprotein Ib beta chain:*GP1BB* (22q11.21); Platelet glycoprotein IX: *GP9* (3q21.3)
Bleeding disorder, platelet-type, 17 = GFI1B -related thrombocytopenia = Monoallelic Grey platelet syndrome (Stevenson et al., 2013)	Dysmorphic MKs; Macrothrombocytopenia; α-granule and dense granule deficiency	Red blood cells anisocytosis; Mild myelofibrosis	AD/AR	Zinc finger protein Gfi1b: *GFI1B* (9q34-13)

Chediak-Higashi syndrome (Nagle et al., 1996)	Normal platelet count; Absence of platelet dense granules; Reduced/absent ATP release from platelet granule	Giant inclusions in granulocytes and their precursors; Severe immunodeficiency; Defective phagocytosis; Neutropenia; Increased susceptibility to infection and lymphoma; Progressive neurological dysfunction; Skin and hair hypopigmentation	AR	Lysosomal trafficking regulator: *LYST* (CHS1) (1q42.3)
Congenital amegakaryocytic thrombocytopenia (Ihara et al., 1999)	Absence of MKs in BM; Very severe thrombocytopenia; Normal size platelets	Increased plasma TPO levels; Pancytopenia; Progress to BM aplasia; Neurological abnormalities	AR	Thrombopoietin receptor: *MPL* (1p34)
CYCS -related thrombocytopenia = Thrombocytopenia Cargeeg (Morison et al., 2008)	Mild thrombocytopenia; Normal size platelets	---	AD	Cytochrome c: *CYCS* (7p15.3)
Cytosolic phospholipase A2 deficiency (Adler et al., 2008)	Normal platelet count; Decreased platelet aggregation with ADP and collagen	Small intestinal ulceration	AR	Cytosolic phospholipase A2: *PLA2G4A* (1q25)
Di Georges / Velocardiofacial syndrome = 22q11.2 deletion syndrome (Latger-Cannard et al., 2004)	Macrothrombocytopenia (sometimes)	Immunodeficiency; Neuro-psychomotor delay; Speech delay; Seizures; Congenital heart defect; Psychiatric disorder	AD	Chromosomal deletion (22q11.2)
DIAPH1 -related thrombocytopenia (Stritt et al., 2016)	Macrothrombocytopenia; Abnormal bleeding symptoms; Platelets show heterogeneity in size, shape, and granule abnormality	Progressive loss of hearing; Mild neutropenia	AD	Protein diaphanous homolog 1: *DIAPH1* (5q31.3)
EPHB2-related defect (Berrou et al., 2018)	Excessive recurrent bleeding; Normal platelet counts; Defect in platelet aggregation, αIIbβ3 activation, granule secretion and thrombus formation on collagen under flow	---	AR	Ephrin type-B receptor 2 *EPHB2* (1p36.12)
ETV6 -related thrombocytopenia = Thrombocytopenia 5 (Zhang et al., 2015)	Increased number of immature small and hypo-lobulated MKs in BM; Thrombocytopenia; Normal size platelets; Elongated α-granules	Red cell macrocytosis; Abnormal erythrocyte precursors; Association with different cancers and malignancies especially leukaemia and myelodysplastic syndromes	AD	Transcription factor ETV6: *ETV6* (12p13)
Familial haemophagocytic types 3-5 (Cote et al., 2009; Feldmann et al., 2003; Spessott et al., 2015; zur Stadt et al., 2005)	Thrombocytopenia; Abnormal platelet aggregation due to secretion defects in all granules; Normal granule cargo	Autoimmune disease; Reduced red cells number; Immune dysregulation	AR/AD	Protein unc-13 homolog D: *UNC13D* / *Munc13–4* (17q25.1); Syntaxin-11: *STX11* (6q24.2); Syntaxin-binding protein 2: *STXBP2* / *Munc18-2* (19p13.2)

Familial platelet disorder with propensity to develop acute myeloid Leukaemia (FPD/AML) (Song et al., 1999)	MKs are small but increased in number; Mild thrombocytopenia; Normal size platelets; Reduced α-granules; Reduced platelet ATP secretion in response to all agonists, dense granule deficiency	Hallmark is that 40% of affected individuals are at risk of haematological malignancies; Growth retardation; Malformations; Dysmorphic features; Mental retardation	AD	Runt-related transcription factor 1: *RUNX1=AML1= CBFA2* (12q22.12)
FLI1 dysfunction (Stockley et al., 2013)	Some have thrombocytopenia; Normal/large size platelets; Fused α-granules in 1-5% of circulating platelets; Dense granule secretion defect	Eczema; Recurrent infection; Alopecia; Neutropenia	AD	Friend leukaemia integration transcription factor: *FLI1* (11q23)
FLNA -related thrombocytopenia = Filamin A disorder (Nurden et al., 2011)	Macrothrombocytopenia; Enlarged α-granules; Present of abnormal fragmentation of the platelet cytoplasm	Neurological, heart, skeletal and muscular developmental disease including periventricular nodular heterotopia	XD	Filamin A: *FLNA* (Xq28)
FYB -related thrombocytopenia (Hamamy et al., 2014)	Normal number of MKs in BM; Decreased number of mature multilobulated MKs in BM; Microthrombocytopenia; Reduced formation of platelet filopodia; Increased platelet clearance from circulation	Eczema during infancy	AR	FYN-binding protein 1: *FYB1* (5p13.1)
G6bB -related defect (Hofmann et al., 2017)	Macrothrombocytopenia; Distinctive pattern of BM reticulin fibrosis centred around clusters of atypical MKs; Mild to moderate bleeding symptoms	Myelofibrosis; Mild anaemia; Mild leukocytosis	AR	Megakaryocyte and platelet inhibitory receptor G6bB: *MPIG6B* (6p21.33)
GATA1 -related thrombocytopenia: X-linked thrombocytopenia with thalassemia (XLTT) / Dyserythropoietic anaemia with thrombocytopenia (Nichols et al., 2000)	Dysmegakaryopoiesis; Variable degrees of macrothrombocytopenia (milder in XLTT); Decrease in platelet α-granules	Dyserythropoiesis; Variable degrees of anaemia	XR	Erythroid transcription factor: *GATA1* (X p11.23)
Gaucher disease (Wan et al., 2017)	Thrombocytopenia	Splenomegaly; Anaemia; Low RBC count; Low β-glucocerebrosidase activity in leukocytes; Gaucher cells in BM; Gastric cancer	AR	Glucosylceramidase: *GBA* (1q22)
Glanzmann thrombasthenia (GT) (Nurden & Caen, 1975)	Normal platelet count; Quantitative or qualitative deficiency of the integrin αIIbβ3 (GPIIb/IIIa); Selective ability of platelets to aggregate in the presence of ristocetin	---	AR	Integrin alphaIIb: *ITGA2B* (7q21.11); Integrin beta3: *ITGB3* (17q21.31)
GPIV thrombospondin receptor defect (Kashiwagi et al., 1994)	Normal platelet count; Deficiency of the GPIV receptor (CD36)	Metabolic syndrome; Atherosclerotic cardiovascular disease; Cardiomyopathy	AR	Platelet glycoprotein IV: *CD36* (7q21.11)
GPVI collagen receptor defect (Dumont et al., 2009)	Normal platelet count; Absent/reduced platelets aggregation with collagen	---	AR	Platelet glycoprotein VI: *GP6* (19q13.42)
Grey platelet syndrome (GPS)	Macrothrombocytopenia; Grey platelets in blood film; Defective α-granule formation; Shortened platelet	Myelofibrosis; Splenomegaly; High serum vitamin B12	AR/AD	Neurobeachin-like protein 2: *NBEAL2* (3 p21.31); Trem-

(Gunay-Aygun et al., 2011; Nurden et al., 2008)	lifespan; Defects in collagen and/or thrombin-induced platelet aggregation; Platelets deficient in GPVI (some patients)			like transcript 1 protein: *TREML1* (6p21.1)
Griscelli syndrome type1-3 (Ménasché et al., 2003; Ménasché et al., 2000; Pastural et al., 1997)	Normal (type 1,3) or low (type 2) platelet count; Deficiency or abnormality in platelet dense granules	Partial albinism; Neurological defects; Severe immunodeficiency with a defect in cytotoxic lymphocyte activity	AR	Type1: Unconventional myosin Va: *MYO5A* (15q21.2) ; Type2: Ras-related protein Rab-27A: *RAB27A* (15q21.3); Type3: Melanophilin: *MLPH* (2q37.3)
GT-like thrombocytopenia = ITGA2/ ITGB3- related thrombocytopenia (Ghevaert et al., 2008; Kunishima et al., 2011)	Macrothrombocytopenia; Spontaneous partial activation of αIIbβ3 integrin; Enlarged α-granules	---	AD	Integrin alpha-IIb: *ITGA2B* (7q21.11); Integrin beta-3: *ITGB3* (17q21.31)
Hermansky-Pudlak syndrome (HPS) (Ammann et al., 2016; Anikster et al., 2001; Cullinane et al., 2011; Li et al., 2003a; Morgan et al., 2006; Oh et al., 1996; Shotelersuk et al., 2000; Suzuki et al., 2002; Zhang et al., 2003)	Normal platelet count; Absence of platelet dense granules; Reduced/absent ATP release from platelet granule	Skin, eye and hair hypopigmentation; Pulmonary fibrosis; Ceroid accumulation; Colitis	AR	Hermansky-Pudlak syndrome 1 protein: *HPS1* (10q24.2); AP-3 complex subunit β-1 (HPS2): *AP3B1* (5q14.1); HPS 3 protein: *HPS3* (3q24); HPS 4 protein: *HPS4* (22q12.1); HPS 5 protein: *HPS5* (11p15.1); HPS 6 protein: *HPS6* (10q24.32); Dysbindin (HPS 7 protein): *DTNBP1* / *HPS7* (6p22.3); Biogenesis of lysosome-related organelles complex 1 subunit 3 (HPS 8 protein): *BLOC1S3* / *HPS8* (19q13.32); Biogenesis of lysosome-related organelles complex 1 subunit 6 (HPS 9 protein / Paladin): *BLOC1S6* / *HPS9* / *PLDN* (15q21.1); AP-3 complex subunit delta-1 (HPS 10 protein): *AP3D1* / *HPS10* (19p13.3)
IVIC syndrome (Paradisi & Arias, 2007)	Mild thrombocytopenia	Upper limb anomalies; Extraocular motor disturbances; Congenital bilateral hearing loss; Heart problems; Leukocytosis; Hypoplasia; Kidney malrotation	AD	Sal-like protein 4: *SALL4* (20q13.2)

KDSR -related defect (Takeichi et al., 2017)	Thrombocytopenia; Impaired platelet function; Normal to increased number of MKs	Hyperkeratosis, ichthyosis; Symmetric erythrokeratoderma	AR	3-Ketodihydrosphingosine reductase: *KDSR* (18q21.33)
Leukocyte adhesion deficiency type III syndrome (Kuijpers et al., 2009)	Macrothrombocytopenia; Loss of 'inside-out' integrin activation in platelets	Loss of 'inside-out' integrin activation in white blood cells, and endothelial cells; Increased susceptibility to infections without pus formation; Poor wound healing; Late detachment of the umbilical cord	AR	Fermitin family homolog 3 (Kindlin-3): *FERMT3* (11q13.1)
MYH9 -related diseases: May-hegglin anomaly, Sebastin syndrome, Fechtner syndrome, Epstein syndrome (Seri et al., 2000)	Macrothrombocytopenia	Döhle-like inclusions in neutrophils; Glomerulonephritis; Renal failure; Deafness; Presenile cataracts; Elevated liver enzyme	AD	Myosin-9: *MYH9* (22q12.3)
P2RY12 ADP receptor Defect (Hollopeter et al., 2001)	Normal platelet count; Reduced transient aggregation to ADP	---	AR	P2Y purinoceptor 12: *P2RY12* (17p13.2)
P2X1 receptor (Oury et al., 2000)	Normal platelet count; Decreased platelet ADP-dependent response	---	AD	P2X purinoceptor 1: *P2RX1* (17p13.3)
Paris-Trousseau syndrome (Breton-Gorius et al., 1995)	Increased number of small immature, hypo-lobulated, dystrophic MKs in BM; Severe macrothrombocytopenia in babies, but often increases with age to near normal levels; Giant α-granules with a deficiency in dense granule; Reduced platelet ATP secretion is in response to all agonists	Mental retardation; Cardiac, facial and neurological defects; Developmental delay	AD	Chromosomal deletion **(**11q23), Between 170 to more than 340 genes deleted
Platelet-type von Willebrand disease = Monoallelic BSS (Miller et al., 1991)	Mild macrothrombocytopenia; Platelets agglutinate spontaneously in response to low ristocetin concentrations	Reduction of plasma von Willebrand factor	AD	Platelet glycoprotein Ib alpha chain: *GP1BA* (17q13.2)
PRKACG and GNE -related disease (Izumi et al., 2014; Manchev et al., 2014)	Macrothrombocytopenia; Platelet dysfunction due to defective aggregation; FLNA is expressed at low-level; Complete abolition of CD62P (P-selectin) expression; Variable CD42b and CD41 expression level	Myopathy with rimmed vacuoles or sialuria; Neurological symptoms; Developmental delay; Skull abnormalities; Severe body myopathy	AR	cAMP-dependent protein kinase catalytic subunit gamma: *PRKACG* (9q13); Bifunctional UDP-N-acetylglucosamine 2-epimerase/N-acetylmannosamine kinase: *GNE* (9p13)
PTPRJ-related thrombocytopenia (Marconi et al., 2018)	Defect in late MK maturation; Reduction in the ability of MK to migrate toward BM extracellular matrix; Defective platelet production; Thrombocytopenia	---	AR	Receptor-type tyrosine-protein phosphatase eta: *PTPRJ* (11p11.2)

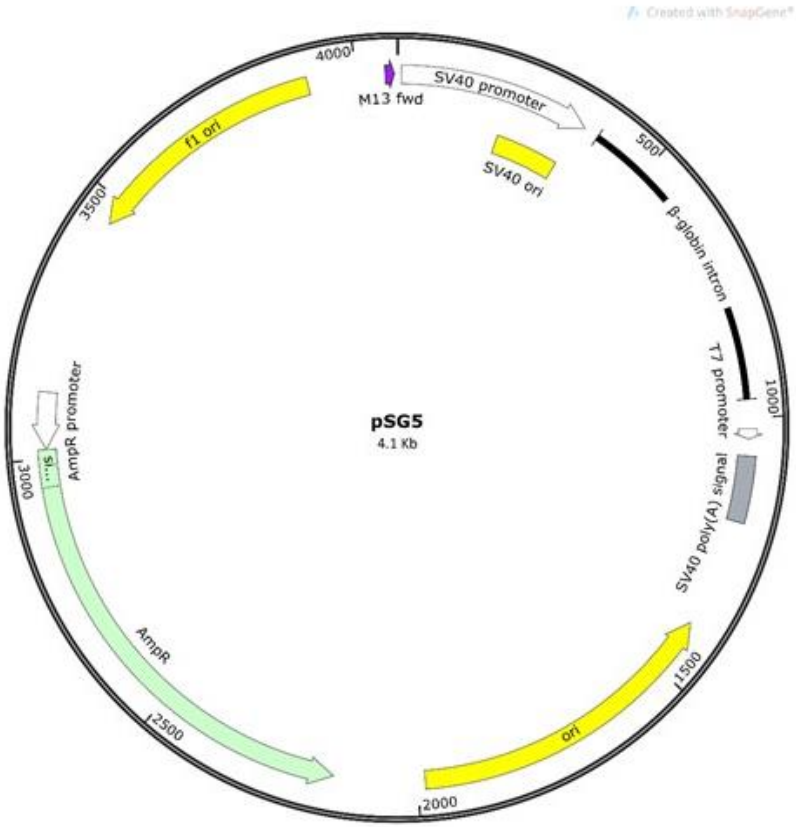
Quebec platelet disorder (Paterson et al., 2010)	Moderate thrombocytopenia; Gain-of-function defect in fibrinolysis; Increased urokinase-type plasminogen activator storage in platelets; Degradation of α-granule proteins despite normal ultrastructure; Lack of platelet aggregation in response to epinephrine	---	AD	Urokinase-type plasminogen activator: *PLAU* (10q22.2)
Radioulnar synostosis with amegakaryocytic thrombocytopenia (RUSAT) = Congenital thrombocytopenia with radioulnar synostosis (Niihori et al., 2015; Thompson & Nguyen, 2000)	Severe reduction of MKs in BM; Thrombocytopenia; Normal size platelets	Skeletal defects mainly a proximal fusion of the ulna and radius; Sensorineural deafness; Potential development of aplastic anaemia	AD	Homeobox protein HOX-A11:*HOXA11* (7p15.2); MDS1 and EVI1 complex locus protein: *MECOM* (3q26.2); Remains unidentified for many patients
RASGRP2 -related defect (Canault et al., 2014)	Severe bleeding; Normal platelet count; Normal granule content with slightly reduced α-granule secretion; Heterozygous have normal bleeding platelet aggregation, but their platelets failed to undergo normal adhesion under flow and spreading	---	AR	RAS guanyl releasing protein 2: *RASGRP2* (11q13)
SBF2 -related thrombocytopenia (Either a new variant form of Griscelli syndrome or a variant Charcot-Marie-Tooth type 4 disease) (Abuzenadah et al., 2013)	Severe thrombocytopenia; Normal MKs; Mild bleeding	Fair coloured hair and skin; Lymphocytosis; No neurologic or other immunologic abnormalities	AR	Myotubularin-related protein 13: *MTMR13* / *SBF2* (11p15)
SCARB1 (Vergeer et al., 2011)	Normal platelet count; Increased unesterified cholesterol content in platelets; Diminished platelet-aggregation response; Increased P-selectin expression (activated platelets); increased adhesion and spreading when exposed to immobilized fibrinogen	High HDL cholesterol levels; Reduced capacity for efflux of cholesterol from macrophages; Decreased adrenal steroidogenesis	AD	Scavenger receptor class B member 1: *SCARB1* (12q24.31)
Scott syndrome (Suzuki et al., 2010)	Mild bleeding; Impaired surface exposure of procoagulant phosphatidylserine on platelets and a severe decrease in the production of microparticles following platelet activation; Reduced thrombin generation and impaired clot formation	-	AR	Anoctamin-6: *ANO6* / *TMEM16F* (12q12)
SLC35A1 -related thrombocytopenia (Kauskot et al., 2018)	Moderate macrothrombocytopenia; Increased immature MKs in BM	Delay psychomotor development; Epilepsy; Ataxia; Microcephaly; Choreiform movements	AR	CMP-sialic acid transporter: *SLC35A1* (6q15)
SLFN14 -related thrombocytopenia (Fletcher et al., 2015)	Moderate thrombocytopenia; Excessive bleeding; Platelet secretion defects; Reduced aggregation and ATP secretion in response to ADP, collagen, and PAR-1 peptide (Gi signalling pathway)	---	AD	Protein SLFN14: *SLFN14* (17q12)
SRC -related thrombocytopenia (Turro et al., 2016)	Hypercellular BM dysplasia with increased numbers of MKs; Dysmegakaryopoiesis; Alteration in MK actin organisation and reduction in proplatelet formation;	Early onset myelofibrosis; Bone pathologies	AD	Proto-oncogene tyrosine-protein kinase Src: *SRC* (20q11.23)

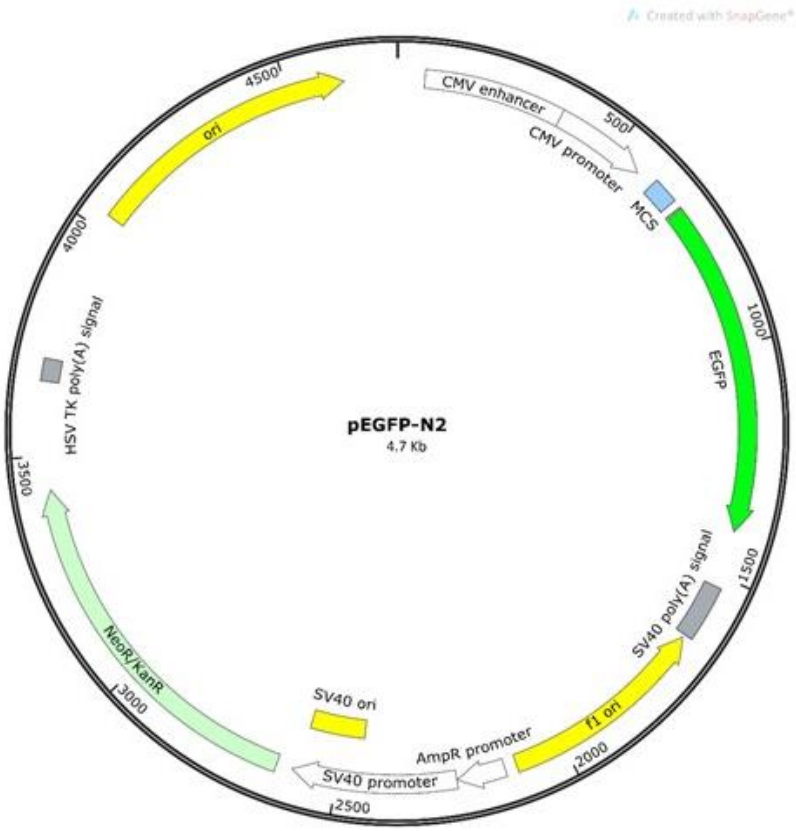
	Thrombocytopenia; Dysmorphic platelet with highly variable in size; A paucity of α-granules; Elevated plasma TPO levels			
Stormorken York platelet syndrome (Markello et al., 2015)	Mild macrothrombocytopenia; Presence of giant aberrant opaque organelles (abnormal lysosomes) in MKs and circulating platelets; Abnormalities in platelet ultrastructure; α-granules abnormality	Mitochondrial myopathy	AD	Stromal interaction molecule 1: *STIM1* (11p15.4)
Takenouchi-Kosaki syndrome (Takenouchi et al., 2015)	Macrothrombocytopenia	Developmental delay; Dysmorphic facial features; Mild eczema, Sensorineural hearing loss; Camptodactyly; Lymphedema; Congenital hypothyroidism; Immunological disturbance	AD	Cell division control protein 42 homolog: *CDC42* (1p36)
Thrombocytopenia associated with sitosterolemia = Mediterranean stomatocytosis/macrothrombocytopenia (MSMT) (Rees et al., 2005)	Normal or increased MK number in BM; Macrothrombocytopenia; Platelet in hyper-activated status	Sterol storage disorder; Stomatocytic haemolysis; Splenomegaly; Xanthomas; Premature atherosclerosis; Arthritis	AR	ATP-binding cassette sub-family G member 5 (Sterolin 1): *ABCG5* (2p21); ATP-binding cassette sub-family G member 8 (Sterolin 2): *ABCG8* (2p21)
Thrombocytopenia in association with absent radii (TAR) (Albers et al., 2011)	Severely reduced MKs in BM; Thrombocytopenia and a severe bleeding tendency in the first years of life, however, platelet count tends to rise, often reaching normal values in adult life; Normal-size platelets	Skeletal defects (Hallmark of the disease is bilateral radial aplasia); Facial dysmorphism; Shortness; Macrocephaly; Renal and cardiac defects; Capillary haemangiomata; Gastroenteritis and cow's milk intolerance; Predisposition to acute myeloid and lymphoid leukaemias	AR	RNA-binding protein 8A: *RBM8A* (1q21.1)
Thrombopoietin-related thrombocytopenia (Savoia et al., 2007)	Mild thrombocytopenia in the heterozygous state	Normal plasma TPO level; Aplastic anaemia in the homozygous state	AR	Thrombopoietin: *THPO* (3q27.1)
Thromboxane A synthase (Ghosal syndrome) (Genevieve et al., 2008)	Thrombocytopenia; Defective platelet aggregation with arachidonic acid	Increased bone density	AR	Thromboxane A synthase: *TBXAS1* (7q34)
Thromboxane A2 receptor (Hirata et al., 1994)	Normal platelet count; Reduced platelet aggregation with arachidonic acid and thromboxane A2 in heterozygous and absent in homozygous conditions	---	AD	Thromboxane A2 receptor: *TBXA2R* (19p13.3)
TUBB1 -related thrombocytopenia (Kunishima et al., 2009)	Moderate macrothrombocytopenia; Normal platelet aggregation	---	AD	Tubulin beta-1 chain: *TUBB1* (20q13.32)
VPS45 -related thrombocytopenia (Stepensky et al., 2013)	Life-threatening thrombocytopenia; Significant bleeding tendency; Platelet dysfunction; Decreased α-granule; Distorted open-channel system	Life-threating infections; Congenital neutropenia; Primary myelofibrosis with extramedullary	AR	Vacuolar protein sorting-associated protein 45: *VPS45* (1q21.2)

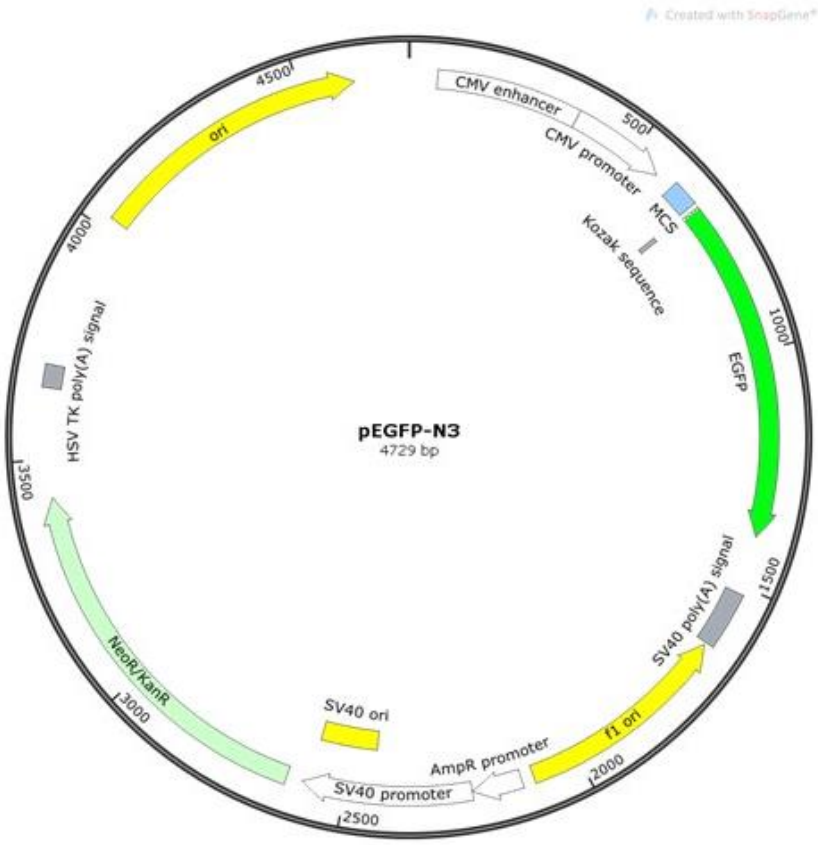
		haematopoiesis; Progressive BM failure		
WDR1 -related thrombocytopenia (Standing et al., 2017)	Thrombocytopenia; Small, atypical MKs with no nucleus and failure of the demarcating membrane system to develop resulting in a large peripheral zone devoid of organelles and granules	Life threating immunodeficiency; Autoinflammatory periodic fever	AR	WD repeat-containing protein 1*: WDR1* (4p16.1)
Wiskott–Aldrich syndrome (WAS) / X-link thrombocytopenia (XLT) (Derry et al., 1994)	Microthrombocytopenia; Reduced number of dense and α-granules; Reduced platelet life span	XLT has a milder clinical presentation than WAS; Severe immune deficiency; Susceptibility to infections, eczema, and autoimmune phenomena; At risk of lymphoproliferative disorders	XR	Wiskott-Aldrich syndrome protein: *WAS* (Xp11.23)

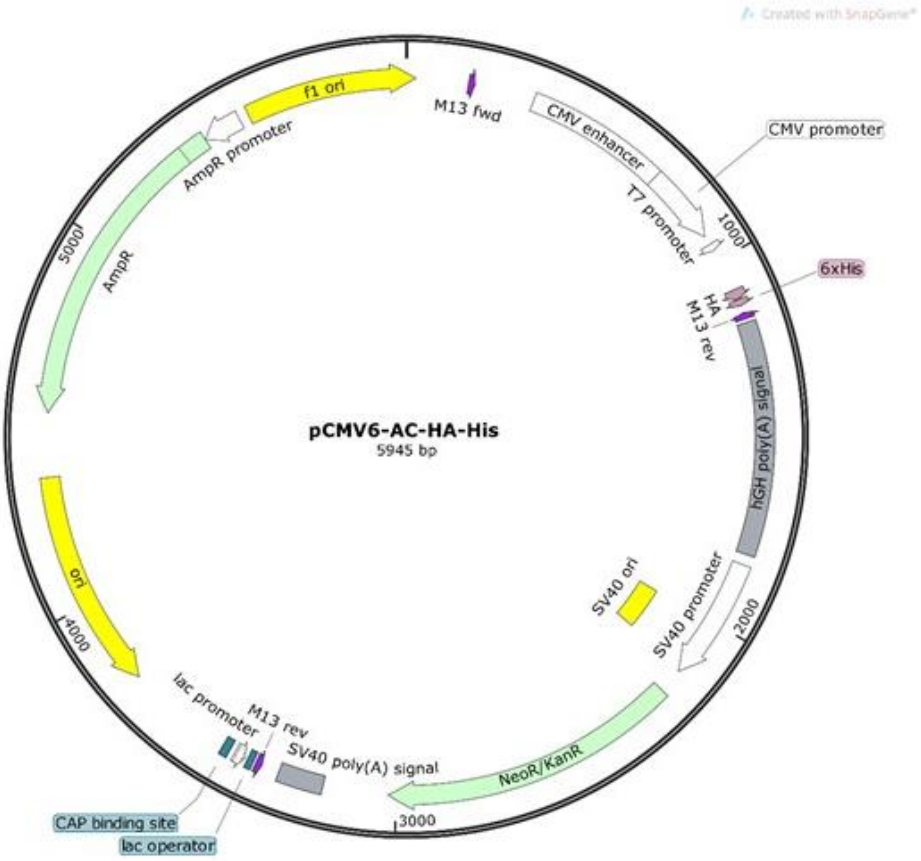
Compiled from Bunimov et al. 2013; Savoia 2015; Nurden & Nurden 2014; Noetzli et al. 2015; Nurden & Nurden 2011; Pecci & Balduini 2014; Favier & Raslova 2015 and further updated. In all cases, the first report to link a gene to an IPD is referenced. AD; autosomal dominant, AR; autosomal recessive, BM; bone marrow, GP; glycoprotein, MK; megakaryocyte, MOI; mode of inheritance, PAR-1; protease-activated receptor 1, TPO; thrombopoietin, XD; X-linked dominant, XR; X-linked recessive.

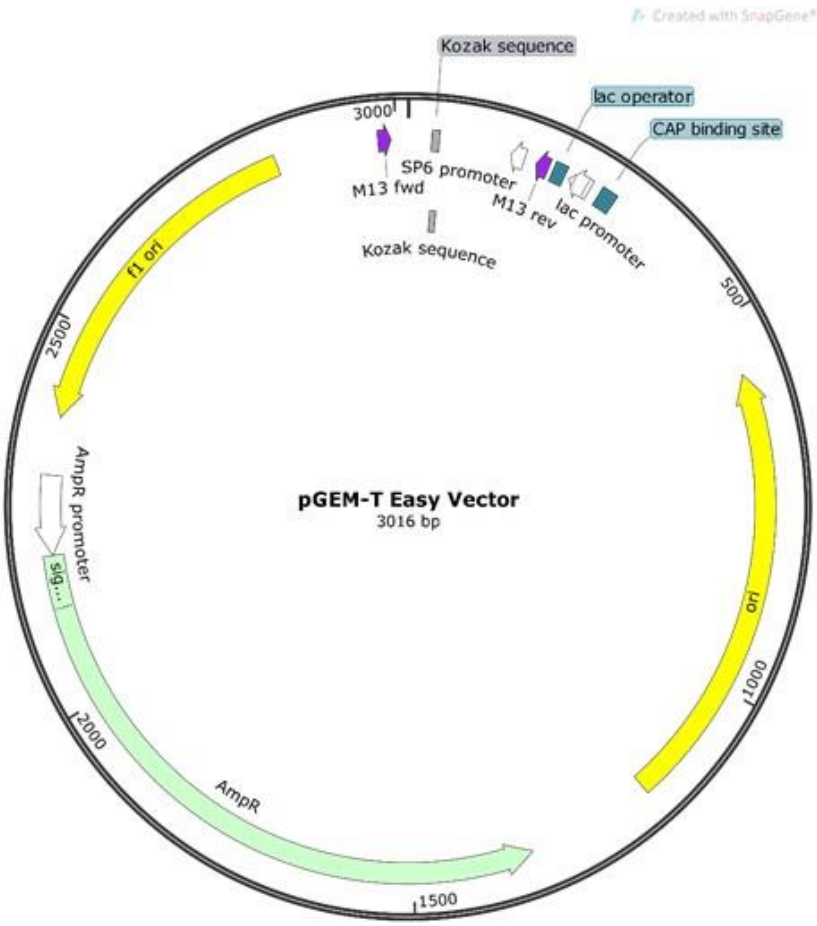
Appendix 2. Schematic representations of plasmids used in this study

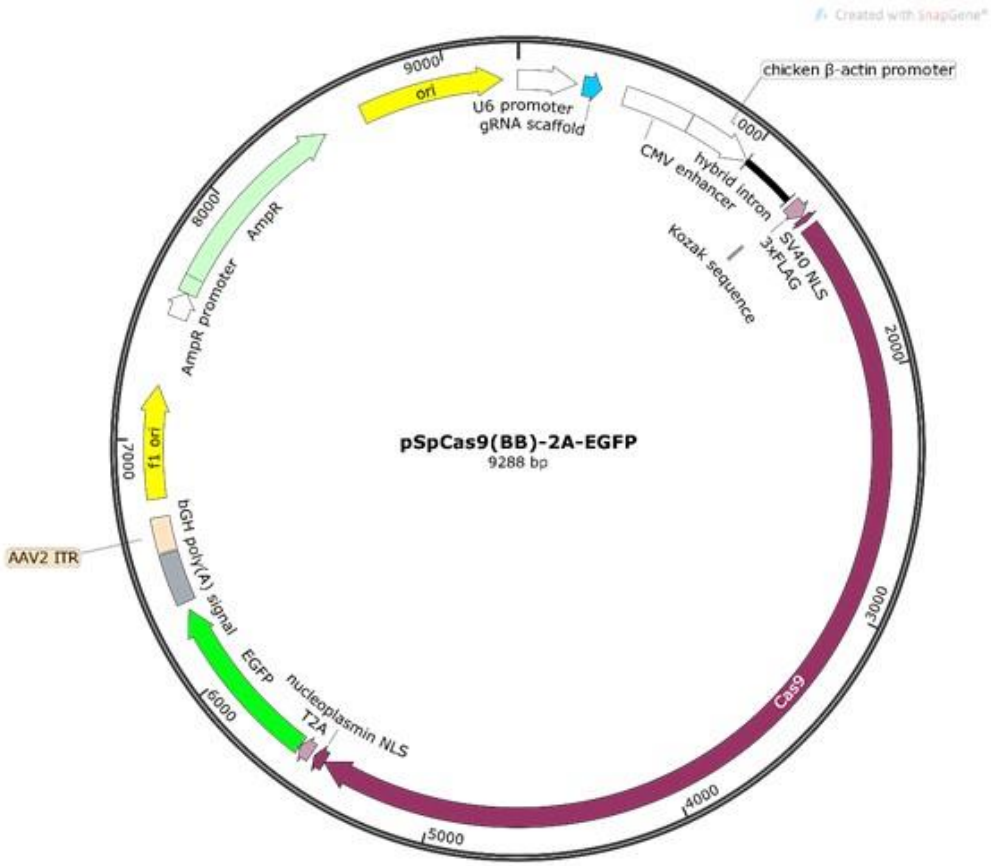


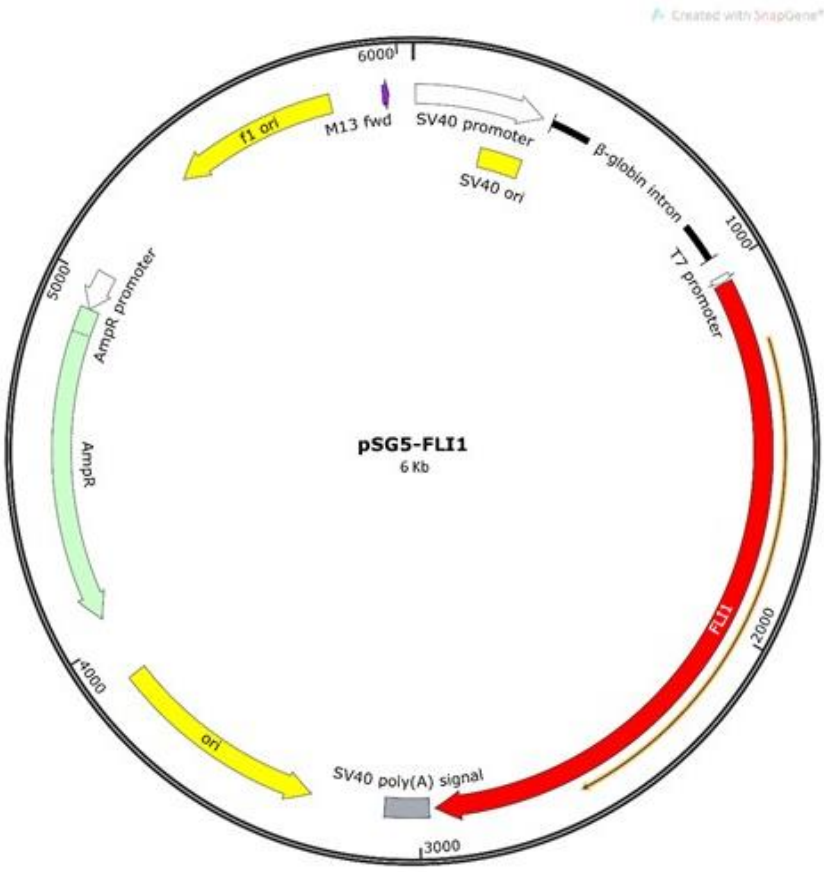


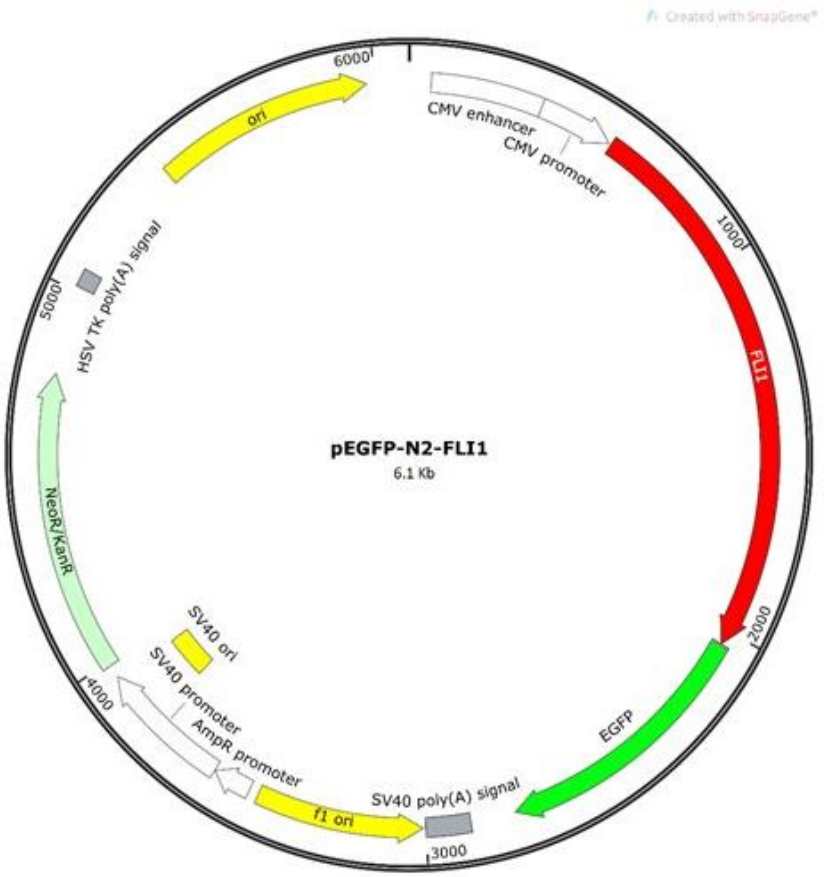

Created with SnapGene®
pEGFP-N3
4729 bp
ori
CMV enhancer
CMV promoter
MCS
Kozak sequence
EGFP
SV40 poly(A) signal
f1 ori
AmpR promoter
SV40 promoter
SV40 ori
NeoR/KanR
HSV TK poly(A) signal
500
1000
1500
2000
2500
3000
3500
4000
4500

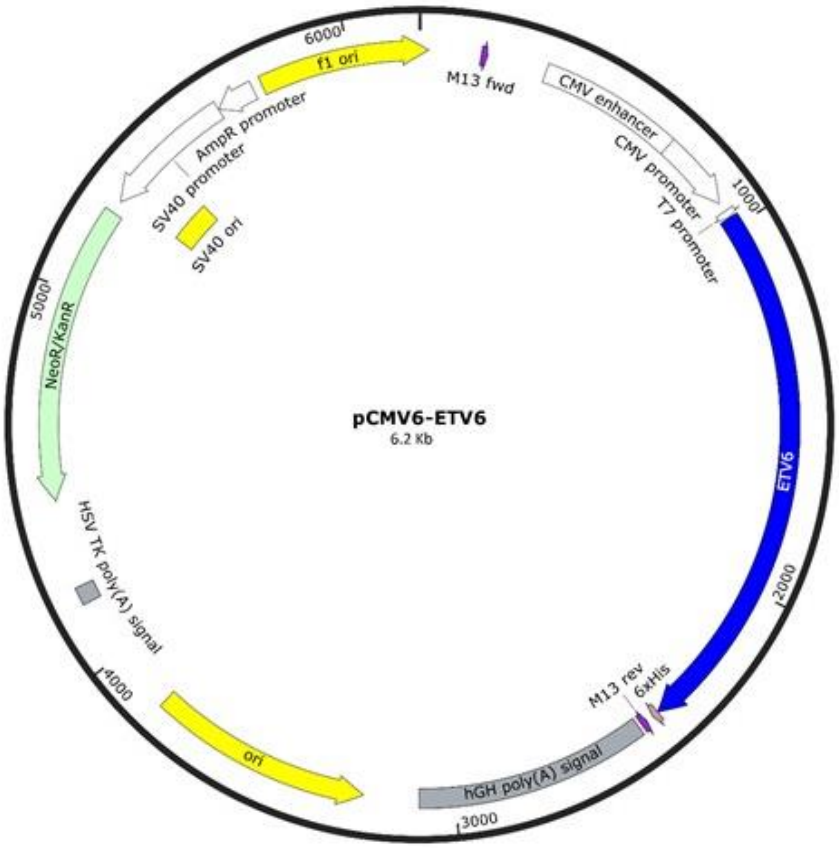

Created with SnapGene®
pCMV6-AC-HA-His
5945 bp
f1 ori
M13 fwd
CMV enhancer
CMV promoter
T7 promoter
6xHis
HA
M13 rev
hGH poly(A) signal
SV40 ori
SV40 promoter
SV40 poly(A) signal
NeoR/KanR
ori
lac promoter
M13 rev
CAP binding site
lac operator
AmpR promoter
AmpR
1000
2000
3000
4000
5000

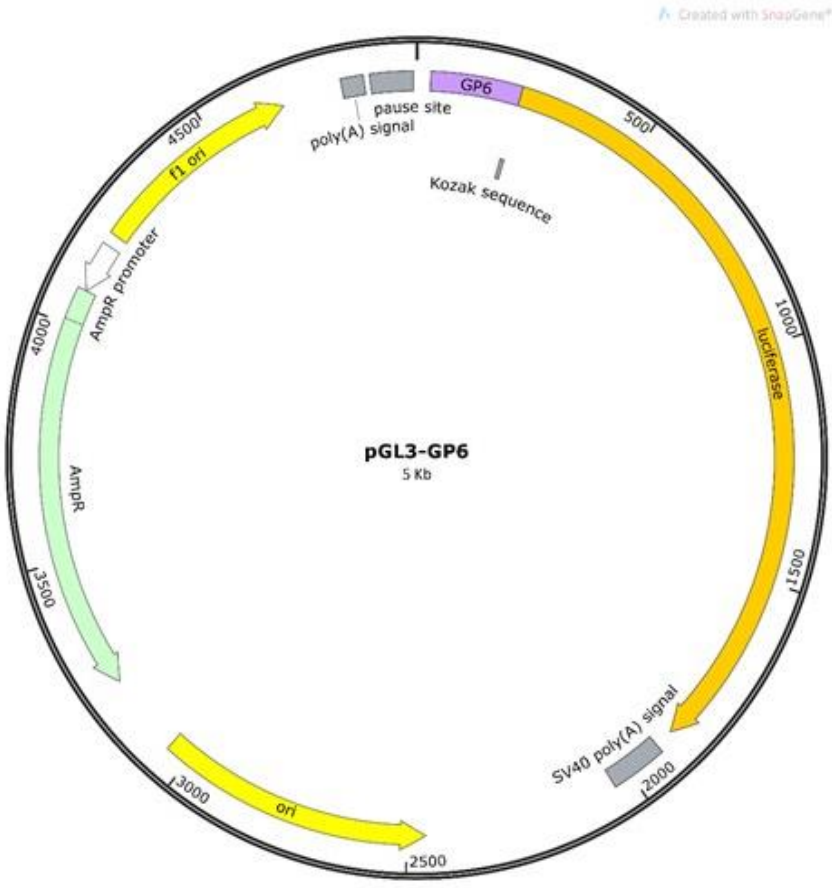

Created with SnapGene®
Kozak sequence
lac operator
CAP binding site
3000
SP6 promoter
M13 fwd
M13 rev
lac promoter
Kozak sequence
500
f1 ori
2500
AmpR promoter
sig...
pGEM-T Easy Vector
3016 bp
ori
1000
2000
AmpR
1500

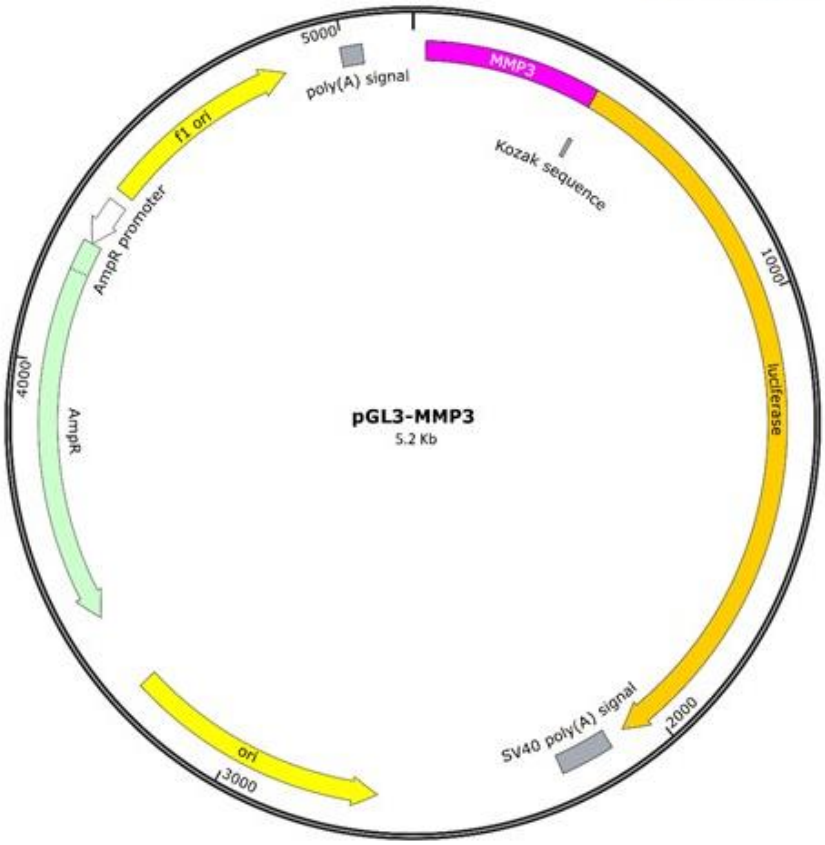

Created with SnapGene®
9000
ori
U6 promoter
gRNA scaffold
chicken β-actin promoter
1000
CMV enhancer
hybrid intron
SV40 NLS
3xFLAG
Kozak sequence
2000
8000
AmpR
AmpR promoter
pSpCas9(BB)-2A-EGFP
9288 bp
7000
f1 ori
AAV2 ITR
bGH poly(A) signal
3000
Cas9
EGFP
nucleoplasmin NLS
6000
T2A
4000
5000

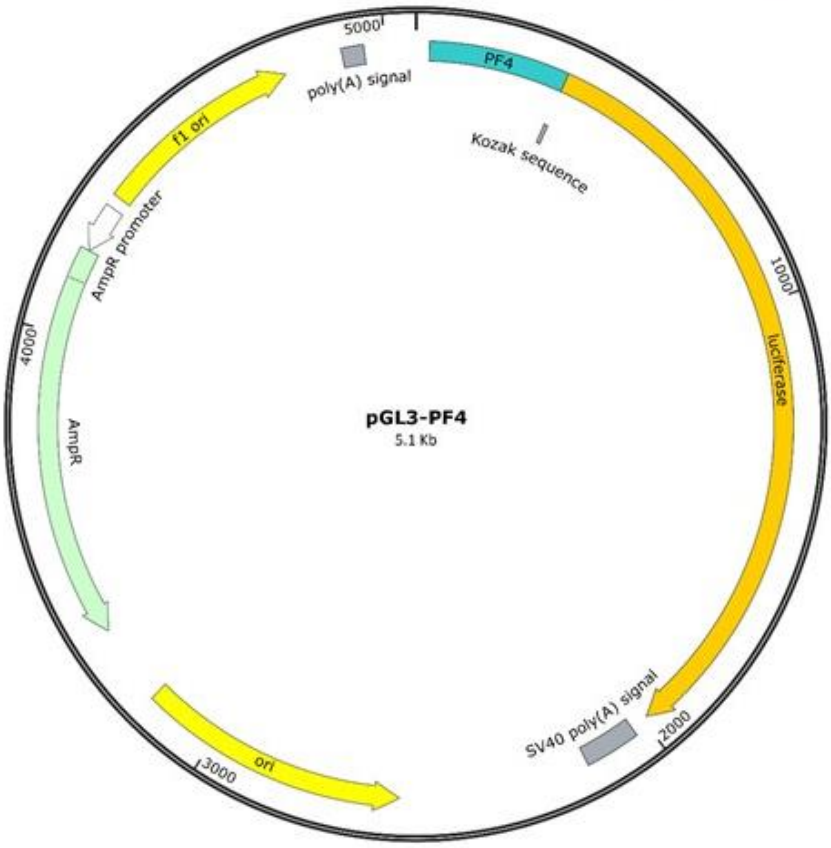

Created with SnapGene®
6000
f1 ori
M13 fwd
SV40 promoter
SV40 ori
β-globin intron
1000
T7 promoter
5000
AmpR promoter
AmpR
pSG5-FLI1
6 Kb
FLI1
2000
4000
ori
SV40 poly(A) signal
3000

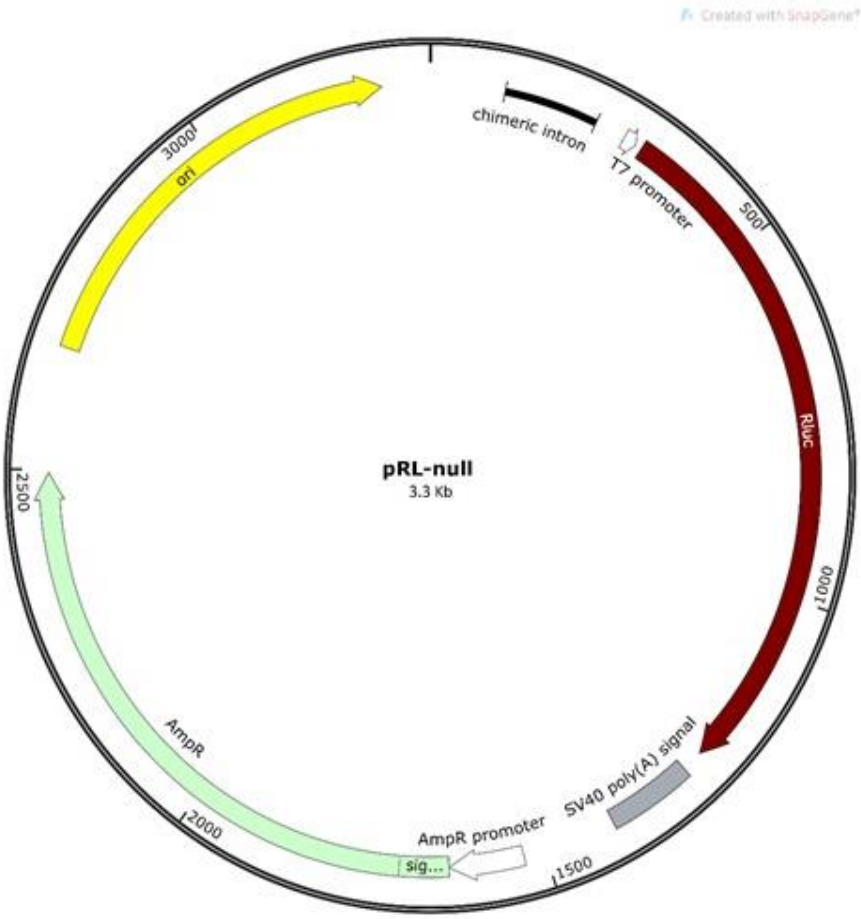

Created with SnapGene®
6000
ori
CMV enhancer
CMV promoter
1000
5000
HSV TK poly(A) signal
FLI1
pEGFP-N2-FLI1
6.1 Kb
NeoR/KanR
SV40 ori
2000
SV40 promoter
AmpR promoter
4000
EGFP
SV40 poly(A) signal
f1 ori
3000


Created with SnapGene®
6000
f1 ori
M13 fwd
CMV enhancer
CMV promoter
T7 promoter
1000
AmpR promoter
SV40 promoter
SV40 ori
5000
NeoR/KanR
pCMV6-ETV6
6.2 Kb
ETV6
2000
HSV TK poly(A) signal
4000
M13 rev
6xHis
ori
hGH poly(A) signal
3000


Created with SnapGene®
GP6
pause site
poly(A) signal
4500
500
f1 ori
Kozak sequence
AmpR promoter
4000
1000
luciferase
pGL3-GP6
5 Kb
AmpR
3500
1500
SV40 poly(A) signal
3000
ori
2000
2500


Created with SnapGene®
5000
poly(A) signal
MMP3
f1 ori
Kozak sequence
AmpR promoter
1000
luciferase
4000
AmpR
pGL3-MMP3
5.2 Kb
SV40 poly(A) signal
2000
ori
3000


Created with SnapGene®
5000
poly(A) signal
PF4
f1 ori
Kozak sequence
AmpR promoter
1000
luciferase
4000
AmpR
pGL3-PF4
5.1 Kb
SV40 poly(A) signal
2000
ori
3000


Created with SnapGene®
chimeric intron
T7 promoter
500
Rluc
1000
SV40 poly(A) signal
1500
AmpR promoter
sig...
2000
AmpR
2500
3000
ori
pRL-null
3.3 Kb

Appendix 3. Software and online tools used during this study

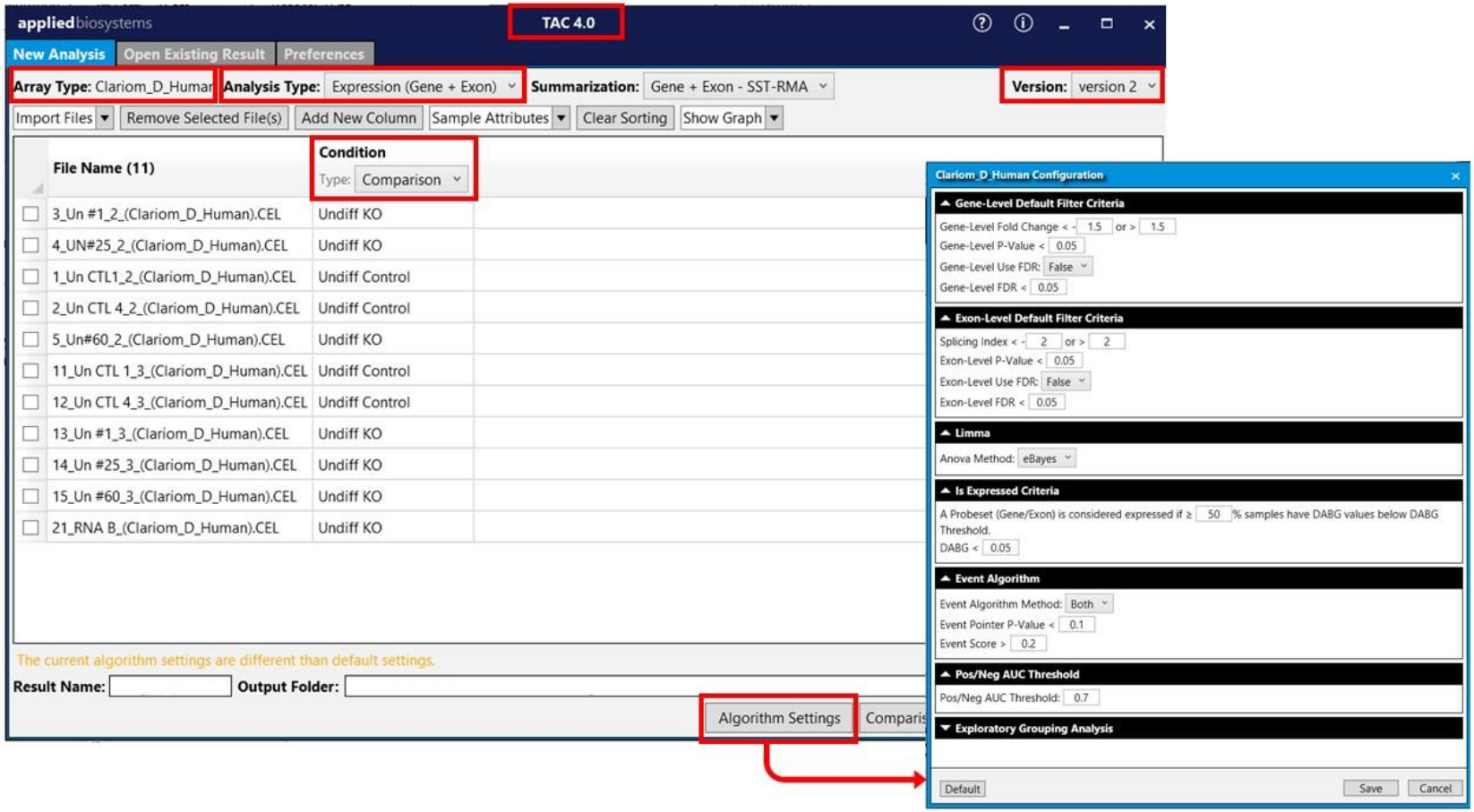
Software	Purpose	Developer	Version
CodonCode Aligner	Viewing and analysing the trace data from Sanger DNA Sequencing	CodonCode Corporation	6.0.2
FinchTV	Viewing and analysing the trace data from Sanger DNA Sequencing	Geospiza, Inc.	1.4.0
GraphPad Prism	Statistical analysis	GraphPad Software	7.03
Image studio	Software of the Odyssey® Sa infrared imaging system for imaging of blots	LI-COR	2.0.13
Image studio lite	Processing blot images and measuring band intensities	LI-COR	5.2
ImageJ	Opening, processing and quantifying ND2 image files acquired by the inverted Ti-eclipse Nikon widefield microscope	National Institutes of Health	1.5
NIS elements advance research software	Software of the inverted Ti-eclipse Nikon widefield microscope used to acquire images	Nikon	4
Pymol	Visualising 3D protein structures and introducing *in silico* mutagenesis	DeLano Scientific LLC	0.99rc6
Quantity One imaging	Acquisition of images taken using the Bio-Rad Gel Doc 2000 ultraviolet (UV) transilluminator	Bio-Rad	4.6.8
RQ manager	Analysis of results from the 7900HT Fast Real-Time PCR system	Applied Biosystems™	1.2.1
SDS	Operating the 7900HT Fast Real-Time PCR system	Applied Biosystems™	2.4
SkanIt RE for VarioSkan Flash software	The software of the Varioskan Flash plate reader	Thermo Scientific	2.4.5
SnapGene Viewer	Generating vector maps	GSL Biotech LLC	4.1.5
Transcriptome Analysis Console (TAC)	Microarray analysis	Affymetrix ™	4.0.1

Online tool	Purpose	Website	Accessed
CCTop - CRISPR/Cas9 target online predictor (Stemmer et al., 2017)	Designing gRNAs for a targeted sequence	https://crispr.cos.uni-heidelberg.de/	2017-2018
ChopChop (Labun et al., 2016; Montague et al., 2014)	Designing gRNAs for a targeted sequence	http://chopchop.cbu.uib.no/index.php	2017-2018
Combined Annotation Dependent Depletion (CADD) (Kircher et al., 2014)	Predicting and scoring the deleteriousness of SNVs and small indels in the human genome	http://cadd.gs.washington.edu/score	2015-2016 Version 1.3
CRISPOR programme (Haeussler et al., 2016)	Designing gRNAs for a targeted sequence	http://crispor.tefor.net/	2017-2018
CRISPR design Zhang lab MIT web-based tool (Hsu et al., 2013)	Designing gRNAs for a targeted sequence	http://crispr.mit.edu/	2017-2018
The Database for Annotation, Visualization and Integrated Discovery (DAVID) (Huang da et al., 2009)	Functional annotation tool	https://david.ncifcrf.gov/summary.jsp	2018 Version 6.8

Efficiency Prediction Tool (Housden et al., 2015)	Calculating efficiency of pre-designed gRNAs and identifying any U6 terminator sequences	http://www.flyrnai.org/evaluateCrispr/input	2017-2018
Ensembl genome browser	Genome browser	www.ensembl.org	GRCh38
Exome Aggregation Consortium "ExAC" database	Database used to find the allele frequencies of certain SNV	http://exac.broadinstitute.org	2015 -2016 Version 0.3
Human Splicing Finder	Predicts the effects of intronic and exonic SNV on splicing	http://www.umd.be/HSF3/	2018 Version 3.1
Integrated and technologies (IDT) tool	Designing gRNAs for a targeted sequence	https://eu.idtdna.com/site/order/designtool/index/CRISPR_CUSTOM	2017-2018
InterPro: protein sequence analysis & classification	Database of protein families, domains and functional sites used to determine the location of variations within a protein	https://www.ebi.ac.uk/interpro/	2015-2016
PaxDb.4	Proteomic database used to access the human platelet proteomic	https://pax-db.org/	2015-2016 Version 4
Prime3Web	Designing PCR primers	http://primer3.ut.ee/	2015-2018
Promega Biomath Ligations calculator	Calculating the amount of insert required for TA cloning	https://www.promega.co.uk/resources/tools/biomath-calculators/	2018
Protein Molecular Weight tool	Protein molecular weight estimation	https://www.bioinformatics.org/sms/prot_mw.html	2016-2017
QuikChange primer design tool	Designing mutagenesis primers	http://www.genomics.agilent.com/primerDesignProgram.jsp	2015-2016
RCSB Protein Data Bank	Database that contains information about experimentally-determined three-dimensional structures of proteins, nucleic acids, and complex assemblies	https://www.rcsb.org/	2015-2018
Standard Nucleotide BLAST tool	Viewing and analysing trace data from Sanger DNA Sequencing	http://blast.ncbi.nlm.nih.gov/Blast.cgi?PROGRAM=blastn&PAGE_TYPE=BlastSearch&LINK_LOC=blasthome	2015-2018
SWISS-MODEL web-based tool	Automated protein structure homology-modelling server	https://swissmodel.expasy.org/	2018
UCSC Genome Browser	Genome browser	http://genome-euro.ucsc.edu/cgi-bin/hgGateway?redirect=manual	2015-2018

gRNAs; guide RNA, PCR; polymerase chain reaction, SNVs; single nucleotide variants.

Appendix 4. Screenshot from the Affymetrix transcriptome analysis console software showing the settings used



Appendix 5. In silico prediction of alteration in the interaction of FLI1 with DNA for R337W, R340C/H and Y343C substitutions

R340	R340C	R340H	R337	R337W	Y343	Y343C
	Mutation 62.6% (highest)	**Mutation 35.7% (highest)**		**Mutation 26.9% (highest)**		**Mutation 68.3% (highest)**
3.06 Å→ guanidino → 5th G in the DNA	---	2.84 Å → imidazole ring → 5th G in the DNA	**3.12 Å → guanidino → 6th G in the DNA**	2.78 Å → imidazole ring → 6th G in the DNA	**3.33 Å → aromatic ring →4th C**	---
2.72 Å → guanidino → S336 in FLI1	---	---	**3.56 Å → guanidino → Y341 in FLI1**	---	**2.91 Å → aromatic ring →4th C**	---

Appendix 6. Percentage of wild-type FLI1 and FLI1 variants showing nuclear accumulation in HEK 293T cells assessed using FLI1-EGFP fusion protein

	WT	R337W	R340C	R340H	Y343C
No. of cells evaluated	68	102	64	98	87
% FLI1 localised to the nucleus	100	65	69.64	72.17	58.01
% Reduction		35	30.36	27.83	41.99
Standard error of the mean	1.747	1.891	1.93	1.764	1.538
p-value	---	<0.0001	<0.0001	<0.0001	<0.0001

WT; wild-type.

Appendix 7. Percentage of wild-type FLI1 and FLI1 variants showing nuclear accumulation in Dami cells assessed using FLI1-EGFP fusion protein

	WT	R337W	R340C	R340H	Y343C
No. of cells evaluated	33	38	27	39	26
% FLI1 localised to the nucleus	78.16	44.64	51.73	45.18	46.88
% FLI1 localised to nucleus (normalised)	100	57.11	66.18	57.80	59.98
% Reduction	---	42.89	33.82	42.20	40.02
Standard error of the mean	3.951	2.34	3.026	2.184	3.672
p-value	---	<0.0001	<0.0001	<0.0001	<0.0001

WT; wild-type.

Appendix 8. Percentage of wild-type FLI1 and FLI1 variants showing nuclear accumulation in HEK 293T cells assessed using anti-FLI1 antibodies

	WT	R337W	R340C	R340H	Y343C
No. of cells evaluated	74	125	73	83	78
% FLI1 localised to the nucleus	69.60	43.99	42.32	42.95	47.39
% FLI1 localised to nucleus (normalised)	100	63.20	60.80	61.71	68.09
% Reduction	---	36.80	39.20	38.29	31.91
Standard error of the mean	2.88	0.89	1.31	1.60	2.17
p-value	---	<0.0001	<0.0001	<0.0001	<0.0001

WT; wild-type.

Appendix 9. The ratio of FLI1 distributed between the nuclear and cytoplasmic fractions of HEK 293T cells expressing wild-type and variant forms of FLI1 as determined by immunoblotting

	WT	R337W	R340C	R340H	Y343C
No. of experiments	4	4	4	4	4
Mean ratio of FLI1 (normalised)	5.60	0.80	0.79	0.80	1.32
Standard error of the mean	1.98	0.27	0.22	0.11	0.33
p-value	---	0.0286	0.0286	0.0286	0.0286

WT; wild-type.

Appendix 10. The percentage of wild-type FLI1 and FLI1 variants showing nuclear localisation as determined by immunoblotting

	WT	R337W	R340C	R340H	Y343C
No. of experiments	4	4	4	4	4
% FLI1 in nuclear fraction	80.61	40.57	41.21	43.87	54.58
% FLI1 localised to nucleus (normalised)	100	50.33	51.12	54.42	67.71
% Reduction	---	49.67	48.88	45.58	32.29
Standard error of the mean	4.935	8.539	7.716	3.491	5.545
p-value	---	0.0286	0.0286	0.0286	0.0286

WT; wild-type.

Appendix 11. The ratio of ETV6 distributed between the nuclear and cytoplasmic fractions of HEK 293T cells expressing wild-type and variant forms of ETV6 as determined by immunoblotting

	WT	R430*	R399C	R399*	EV
No. of experiments	5	4	3	3	3
Mean ratio of ETV6 (normalised)	89.22	21.55	4.93	0.91	7.73
Standard error of the mean	28.67	6.13	1.46	0.55	3.72
p-value	---	0.06	0.04	0.04	0.04

WT; wild-type.

Appendix 12. Candidate genes in the F4.1 participant that were identified using whole exome sequencing and the used pipeline

Gene	CADD_ PHRED	Gene	CADD_ PHRED	Gene	CADD_ PHRED
ETV6	**45**	*PDLIM7*	26.2	*MYO6*	23.4
MSH6	35	*PTH1R*	25.5	*KIF14*	23.1
GOLGA4	35	*RNF6*	25.2	*ZNF236*	23.1
ABCB1	35	*HLA-C*	25	*NT5C3A*	23
SRI	35	*SPTAN1*	24.9	*CACNA1C*	22.9

STRADA	34	*USP32*	24.8	*ATG4B*	22.7
ADAM17	34	*ARHGAP10*	24.5	*AAGAB*	22.6
MTMR3	34	*PNLIPRP2*	24.4	*SULT1A2*	22.6
CIAO1	33	*PPFIA2*	24.3	*HIST1H2AH*	22.5
ELOVL7	33	*PNPLA6*	24.3	*GAPVD1*	22
FSTL4	33	*NRP2*	24.2	*CCDC141*	21.2
KIAA1841	32	*SRCAP*	24	*AFTPH*	20.5
NAGA	32	*SUGP2*	24	*CCDC88B*	Unpredicted
WDR60	32	*ITSN2*	24	*RNF103*	Unpredicted
ACER2	32	*NUP210L*	23.9	*AMPD1*	Unpredicted
GMPS	29.3	*MPEG1*	23.9	*RAI1*	Unpredicted
LNPEP	29	*TXNDC16*	23.8	*ATF7IP2*	Unpredicted
ALDH2	28.7	*MED16*	23.8	*UPF1*	Unpredicted
ACSF3	27.8	*SPTBN1*	23.8	*SRP14*	Unpredicted
SFI1	27.5	*DND1*	23.7	*ZFPM1 (FOG1)*	Unpredicted
DNAJC21	27.5	*RAB11FIP3*	23.6	*REC8*	Unpredicted
KCNK6	27.1	*IPO5*	23.5	*MCC*	Unpredicted
COL12A1	27	*TXNDC17*	23.4	*GAS8-AS1 (C16orf3)*	Unpredicted
KMT2C	26.4	*MBOAT7*	23.4	*SUPT20HL1*	Unpredicted

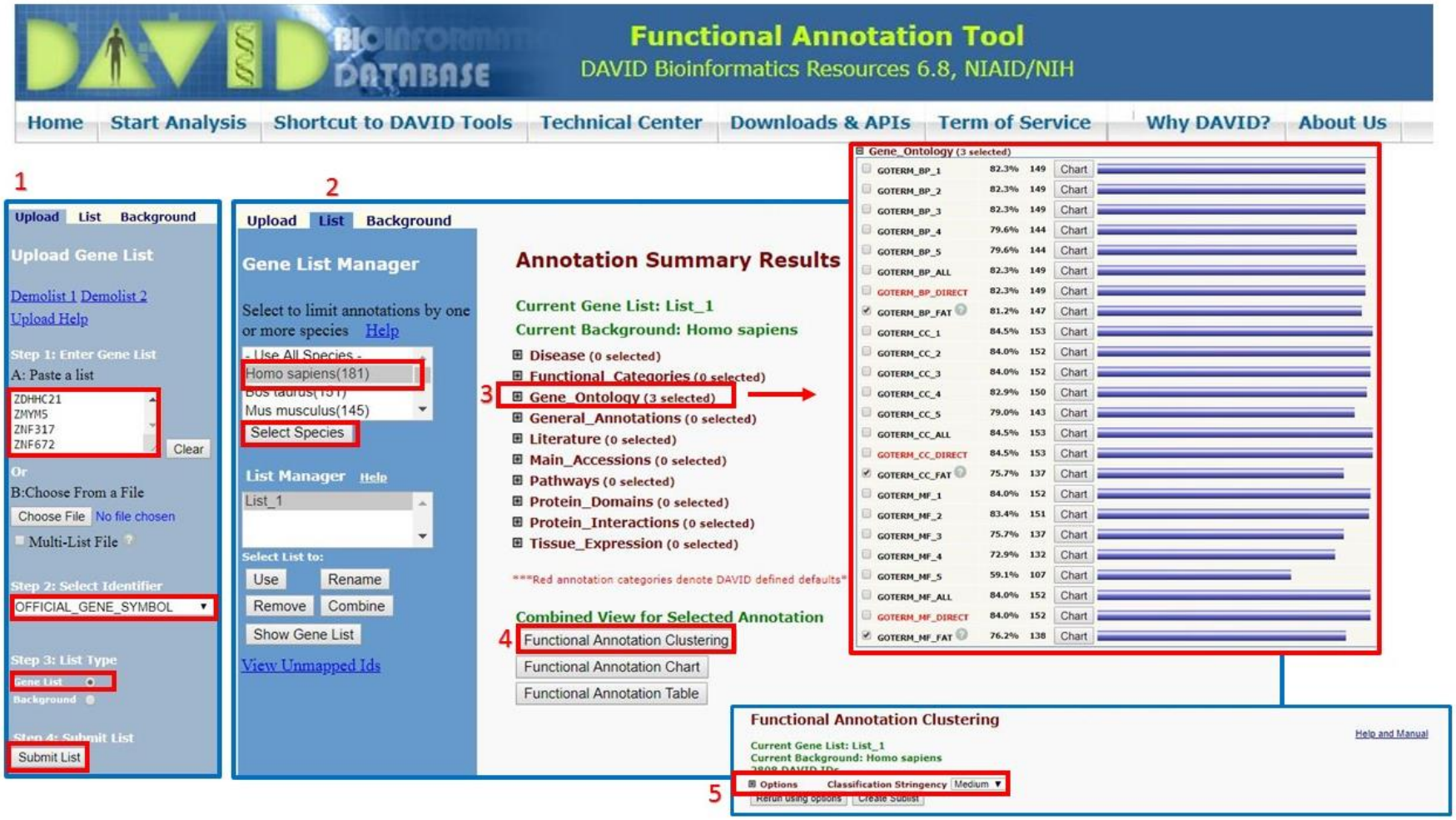
CADD; Combined Annotation Dependent Depletion

Appendix 13. The top five designed guides for FLI1 exons 6, 7, 9 and 8, their specificity, and predicted cutting efficiency

	Guide sequence and proto-spacer adjacent motif (PAM)*	Guiding quality score*	Efficiency score**
Exon 6			
1	CTTCTGACTGAGTCATAAGAAGG	65	3.47
2	TGGGGCAATAACATGAATTCTGG	63	7.46
3	TCTTATGACTCAGTCAGAAGAGG	57	7.28
4	TTCTGACTGAGTCATAAGAAGGG	56	4.90
5	CTCAGTCAGAAGAGGAGCTTGGG	56	6.01
Exon 7			
1#	GATCGTTTGTGCCCCTCCAAGGG	83	5.23
2	TGATCGTTTGTGCCCCTCCAAGG	82	3.61
3	ATCGTTTGTGCCCCTCCAAGGGG	79	3.79
4	GTTTGTGCCCCTCCAAGGGGAGG	66	6.57
5	TCAGTAAGAATACAGAGCAACGG	42	3.90
Exon 8			
1	AGGGTTGGCTAGGCGACTGCTGG	86	6.63
2	TGGCTAGGCGACTGCTGGTCGGG	78	5.22
3	GTCGGGCCCAGGATCTGATACGG	77	6.03
4	TTGGCTAGGCGACTGCTGGTCGG	75	9.35
5	GGCGACTGCTGGTCGGGCCCAGG	74	4.14
Exon 9			
1	CAAAATGACGGACCCCGATGAGG	94	4.27
2#	AATGACGGACCCCGATGAGGTGG	92	6.44
3	CATAGTAATAACGGAGGGCCCGG	87	5.70
4	ACAGCTGGCGTTGGCGCTGTCGG	83	6.20
5	CCAGGTGATACAGCTGGCGTTGG	83	4.64

* Data were generated from CRISPR design Zhang lab tool. ** Data were generated from Efficiency Prediction Tool. # The selected guide.

Appendix 14. Screenshot showing the settings used for functional annotation analysis in Database for Annotation, Visualization and Integrated Discovery (DAVID)



Bibliography

Abedin, M. J., Nguyen, A., Jiang, N., Perry, C. E., Shelton, J. M., Watson, D. K. & Ferdous, A. (2014) Fli1 acts downstream of Etv2 to govern cell survival and vascular homeostasis via positive autoregulation. *Circulation research*, CIRCRESAHA. 113.303145.

Abeliovich, A., Schmitz, Y., Fariñas, I., Choi-Lundberg, D., Ho, W.-H., Castillo, P. E., Shinsky, N., Verdugo, J. M. G., Armanini, M. & Ryan, A. (2000) Mice lacking α-synuclein display functional deficits in the nigrostriatal dopamine system. *Neuron*, 25(1), 239-252.

Abuzenadah, A. M., Zaher, G. F., Dallol, A., Damanhouri, G. A., Chaudhary, A. G., Al-Sayes, F., Gari, M. A., AlZahrani, M., Hindawi, S. & Al-Qahtani, M. H. (2013) Identification of a novel SBF2 missense mutation associated with a rare case of thrombocytopenia using whole-exome sequencing. *J Thromb Thrombolysis*, 36(4), 501-6.

Adler, D. H., Cogan, J. D., Phillips, J. A., 3rd, Schnetz-Boutaud, N., Milne, G. L., Iverson, T., Stein, J. A., Brenner, D. A., Morrow, J. D., Boutaud, O. & Oates, J. A. (2008) Inherited human cPLA(2alpha) deficiency is associated with impaired eicosanoid biosynthesis, small intestinal ulceration, and platelet dysfunction. *J Clin Invest*, 118(6), 2121-31.

Adli, M. (2018) The CRISPR tool kit for genome editing and beyond. *Nature communications*, 9(1), 1911.

Adolf, F., Rhiel, M., Hessling, B., Hellwig, A. & Wieland, F. T. (2018) Proteomic Profiling of Mammalian COPII Vesicles. *bioRxiv*, 253294.

Ajiro, M., Jia, R., Wang, R.-H., Deng, C.-X. & Zheng, Z.-M. (2015) Adapted Resistance to the Knockdown Effect of shRNA-Derived Srsf3 siRNAs in Mouse Littermates. *International journal of biological sciences*, 11(11), 1248.

Al Hawas, R., Ren, Q., Ye, S., Karim, Z. A., Filipovich, A. H. & Whiteheart, S. W. (2012) Munc18b/STXBP2 is required for platelet secretion. *Blood*, 120(12), 2493-500.

Albers, C. A., Cvejic, A., Favier, R., Bouwmans, E. E., Alessi, M. C., Bertone, P., Jordan, G., Kettleborough, R. N., Kiddle, G., Kostadima, M., Read, R. J., Sipos, B., Sivapalaratnam, S., Smethurst, P. A., Stephens, J., Voss, K., Nurden, A., Rendon, A., Nurden, P. & Ouwehand, W. H. (2011) Exome sequencing identifies NBEAL2 as the causative gene for gray platelet syndrome. *Nat Genet*, 43(8), 735-7.

Albers, C. A., Paul, D. S., Schulze, H., Freson, K., Stephens, J. C., Smethurst, P. A., Jolley, J. D., Cvejic, A., Kostadima, M., Bertone, P., Breuning, M. H., Debili, N., Deloukas, P., Favier, R., Fiedler, J., Hobbs, C. M., Huang, N., Hurles, M. E., Kiddle, G., Krapels, I., Nurden, P., Ruivenkamp, C. A., Sambrook, J. G., Smith, K., Stemple, D. L., Strauss, G., Thys, C., van Geet, C., Newbury-Ecob, R., Ouwehand, W. H. & Ghevaert, C. (2012) Compound inheritance of a low-frequency regulatory SNP and a rare null mutation in exon-junction complex subunit RBM8A causes TAR syndrome. *Nat Genet*, 44(4), 435-9, S1-2.

Alexander, W., Roberts, A., Nicola, N., Li, R. & Metcalf, D. (1996) Deficiencies in progenitor cells of multiple hematopoietic lineages and defective megakaryocytopoiesis in mice lacking the thrombopoietic receptor c-Mpl. *Blood*, 87(6), 2162-2170.

Ammann, S., Schulz, A., Krageloh-Mann, I., Dieckmann, N. M., Niethammer, K., Fuchs, S., Eckl, K. M., Plank, R., Werner, R., Altmuller, J., Thiele, H., Nurnberg, P., Bank, J., Strauss, A., von Bernuth, H., Zur Stadt, U., Grieve, S., Griffiths, G. M., Lehmberg, K., Hennies, H. C. & Ehl, S. (2016) Mutations in AP3D1 associated with immunodeficiency and seizures define a new type of Hermansky-Pudlak syndrome. *Blood*, 127(8), 997-1006.

Aneja, K., Jalagadugula, G., Mao, G., Singh, A. & Rao, A. K. (2011) Mechanism of platelet factor 4 (PF4) deficiency with RUNX1 haplodeficiency: RUNX1 is a transcriptional regulator of PF4. *J Thromb Haemost*, 9(2), 383-91.

Anikster, Y., Huizing, M., White, J., Shevchenko, Y. O., Fitzpatrick, D. L., Touchman, J. W., Compton, J. G., Bale, S. J., Swank, R. T., Gahl, W. A. & Toro, J. R. (2001) Mutation of a new gene causes a unique form of Hermansky-Pudlak syndrome in a genetic isolate of central Puerto Rico. *Nat Genet*, 28(4), 376-80.

Antl, M., von Brühl, M.-L., Eiglsperger, C., Werner, M., Konrad, I., Kocher, T., Wilm, M., Hofmann, F., Massberg, S. & Schlossmann, J. (2007) IRAG mediates NO/cGMP-dependent inhibition of platelet aggregation and thrombus formation. *Blood*, 109(2), 552-559.

Arthur, W. T., Petch, L. A. & Burridge, K. (2000) Integrin engagement suppresses RhoA activity via a c-Src-dependent mechanism. *Current Biology*, 10(12), 719-722.

Asano, Y., Stawski, L., Hant, F., Highland, K., Silver, R., Szalai, G., Watson, D. K. & Trojanowska, M. (2010) Endothelial Fli1 deficiency impairs vascular homeostasis: a role in scleroderma vasculopathy. *Am J Pathol*, 176(4), 1983-98.

Athanasiou, M., Clausen, P. A., Mavrothalassitis, G. J., Zhang, X.-K., Watson, D. K. & Blair, D. G. (1996) Increased expression of the ETS-related transcription factor FLI-1/ERGB correlates with and can induce the megakaryocytic phenotype. *Cell Growth and Differentiation-Publication American Association for Cancer Research*, 7(11), 1525-1534.

Avril, T., North, S. J., Haslam, S. M., Willison, H. J. & Crocker, P. R. (2006) Probing the cis interactions of the inhibitory receptor Siglec-7 with α2, 8-disialylated ligands on natural killer cells and other leukocytes using glycan-specific antibodies and by analysis of α2, 8-sialyltransferase gene expression. *Journal of leukocyte biology*, 80(4), 787-796.

Babst, M., Katzmann, D. J., Snyder, W. B., Wendland, B. & Emr, S. D. (2002) Endosome-associated complex, ESCRT-II, recruits transport machinery for protein sorting at the multivesicular body. *Developmental cell*, 3(2), 283-289.

Bach, L. A. (2018) 40 YEARS OF IGF1: IGF-binding proteins. *Journal of molecular endocrinology*, 61(1), T11-T28.

Bainbridge, M. N., Wiszniewski, W., Murdock, D. R., Friedman, J., Gonzaga-Jauregui, C., Newsham, I., Reid, J. G., Fink, J. K., Morgan, M. B., Gingras, M. C., Muzny, D. M., Hoang, L. D., Yousaf, S., Lupski, J. R. & Gibbs, R. A. (2011) Whole-genome sequencing for optimized patient management. *Sci Transl Med*, 3(87), 87re3.

Bariana, T. K., Ouwehand, W. H., Guerrero, J. A., Gomez, K., BRIDGE Bleeding, T., Disorders, P. & Consortia, T. (2017) Dawning of the age of genomics for platelet granule disorders: improving insight, diagnosis and management. *British journal of haematology*, 176(5), 705-720.

Barrangou, R., Fremaux, C., Deveau, H., Richards, M., Boyaval, P., Moineau, S., Romero, D. A. & Horvath, P. (2007) CRISPR provides acquired resistance against viruses in prokaryotes. *Science*, 315(5819), 1709-1712.

Bartel, F. O., Higuchi, T. & Spyropoulos, D. D. (2000) Mouse models in the study of the Ets family of transcription factors. *Oncogene*, 19(55), 6443.

Bartling, B., Koch, A., Simm, A., Scheubel, R., Silber, R.-E. & Santos, A. N. (2010) Insulin-like growth factor binding proteins-2 and-4 enhance the migration of human CD34−/CD133+ hematopoietic stem and progenitor cells. *International journal of molecular medicine*, 25(1), 89-96.

Bastian, L. S., Kwiatkowski, B. A., Breininger, J., Danner, S. & Roth, G. (1999) Regulation of the megakaryocytic glycoprotein IX promoter by the oncogenic Ets transcription factor Fli-1. *Blood*, 93(8), 2637-2644.

Bauer, D. E., Canver, M. C. & Orkin, S. H. (2015) Generation of genomic deletions in mammalian cell lines via CRISPR/Cas9. *Journal of visualized experiments: JoVE*(95).

Beck, R., Ravet, M., Wieland, F. & Cassel, D. (2009) The COPI system: molecular mechanisms and function. *FEBS letters*, 583(17), 2701-2709.
Behrens, K. & Alexander, W. S. (2018) Cytokine control of megakaryopoiesis. *Growth Factors*, 1-15.
Bejar, R., Stevenson, K., Abdel-Wahab, O., Galili, N., Nilsson, B., Garcia-Manero, G., Kantarjian, H., Raza, A., Levine, R. L. & Neuberg, D. (2011) Clinical effect of point mutations in myelodysplastic syndromes. *New England Journal of Medicine*, 364(26), 2496-2506.
Bem, D., Smith, H., Banushi, B., Burden, J. J., White, I. J., Hanley, J., Jeremiah, N., Rieux-Laucat, F., Bettels, R., Ariceta, G., Mumford, A. D., Thomas, S. G., Watson, S. P. & Gissen, P. (2015) VPS33B regulates protein sorting into and maturation of α-granule progenitor organelles in mouse megakaryocytes. *Blood*, blood-2014-12-614677.
Ben-David, Y. & Bernstein, A. (1991) Friend virus-induced erythroleukemia and the multistage nature of cancer. *Cell*, 66(5), 831-834.
Ben-David, Y., Giddens, E., Letwin, K. & Bernstein, A. (1991) Erythroleukemia induction by Friend murine leukemia virus: insertional activation of a new member of the ets gene family, Fli-1, closely linked to c-ets-1. *Genes & development*, 5(6), 908-918.
Ben-David, Y., Giddens, E. B. & Bernstein, A. (1990) Identification and mapping of a common proviral integration site Fli-1 in erythroleukemia cells induced by Friend murine leukemia virus. *Proceedings of the National Academy of Sciences*, 87(4), 1332-1336.
Berge, K. E., Tian, H., Graf, G. A., Yu, L., Grishin, N. V., Schultz, J., Kwiterovich, P., Shan, B., Barnes, R. & Hobbs, H. H. (2000) Accumulation of dietary cholesterol in sitosterolemia caused by mutations in adjacent ABC transporters. *Science*, 290(5497), 1771-1775.
Berndt, M., Metharom, P. & Andrews, R. (2014) Primary haemostasis: newer insights. *Haemophilia*, 20, 15-22.
Berndt, M. C., Gregory, C., Chong, B. H., Zola, H. & Castaldi, P. A. (1983) Additional glycoprotein defects in Bernard-Soulier's syndrome: confirmation of genetic basis by parental analysis. *Blood*, 62(4), 800-807.
Berrou, E., Adam, F., Lebret, M., Planche, V., Fergelot, P., Issertial, O., Coupry, I., Bordet, J.-C., Nurden, P. & Bonneau, D. (2017) Gain-of-function mutation in filamin A potentiates platelet integrin αIIbβ3 activation. *Arteriosclerosis, thrombosis, and vascular biology*, 37(6), 1087-1097.
Berrou, E., Soukaseum, C., Favier, R., Adam, F., Elaib, Z., Kauskot, A., Bordet, J.-C., Ballerini, P., Loyau, S. & Feng, M. (2018) A mutation of the human EPHB2 gene leads to a major platelet functional defect. *Blood*, 132(19), 2067-2077.
Béthune, J. & Wieland, F. T. (2018) Assembly of COPI and COPII vesicular coat proteins on membranes. *Annual review of biophysics*.
Bonetti, P., Testoni, M., Scandurra, M., Ponzoni, M., Piva, R., Mensah, A. A., Rinaldi, A., Kwee, I., Tibiletti, M. G. & Iqbal, J. (2013) Deregulation of ETS1 and FLI1 contributes to the pathogenesis of diffuse large B-cell lymphoma. *Blood*, 122(13), 2233-2241.
Born, G. V. R. (1962) Aggregation of blood platelets by adenosine diphosphate and its reversal. *Nature*, 194(4832), 927-929.
Bouaouina, M., Lad, Y. & Calderwood, D. A. (2008) The N-terminal domains of talin cooperate with the phosphotyrosine binding-like domain to activate β1 and β3 integrins. *Journal of Biological Chemistry*, 283(10), 6118-6125.
Boylan, B., Gao, C., Rathore, V., Gill, J. C., Newman, D. K. & Newman, P. J. (2008) Identification of FcγRIIa as the ITAM-bearing receptor mediating αIIbβ3 outside-in integrin signaling in human platelets. *Blood*, 112(7), 2780-2786.
Brenca, M., Rossi, S., Polano, M., Gasparotto, D., Zanatta, L., Racanelli, D., Valori, L., Lamon, S., Dei Tos, A. P. & Maestro, R. (2016) Transcriptome sequencing identifies ETV6-NTRK3 as a gene fusion involved in GIST. *J Pathol*, 238(4), 543-9.
Breton-Gorius, J., Favier, R., Guichard, J., Cherif, D., Berger, R., Debili, N., Vainchenker, W. & Douay, L. (1995) A new congenital dysmegakaryopoietic thrombocytopenia (Paris-Trousseau) associated with giant platelet alpha-granules and chromosome 11 deletion at 11q23 [see comments]. *Blood*, 85(7), 1805-1814.
Breuza, L., Halbeisen, R., Jenö, P., Otte, S., Barlowe, C., Hong, W. & Hauri, H.-P. (2004) Proteomics of endoplasmic reticulum-Golgi intermediate compartment (ERGIC) membranes from brefeldin A-treated HepG2 cells identifies ERGIC-32, a new cycling protein that interacts with human Erv46. *Journal of Biological Chemistry*, 279(45), 47242-47253.
Briquet-Laugier, V., El. Golli, N., Nurden, P., Lavenu-Bombled, C., Dubart-Kupperschmitt, A., Nurden, A. & Rosa, J.-P. (2004) Thrombopoietin-induced Dami cells as a model for α-granule biogenesis. *Platelets*, 15(6), 341-344.
Brown, L. A., Rodaway, A. R., Schilling, T. F., Jowett, T., Ingham, P. W., Patient, R. K. & Sharrocks, A. D. (2000) Insights into early vasculogenesis revealed by expression of the ETS-domain transcription factor Fli-1 in wild-type and mutant zebrafish embryos. *Mechanisms of development*, 90(2), 237-252.
Brunet, S. & Sacher, M. (2014) In sickness and in health: the role of TRAPP and associated proteins in disease. *Traffic*, 15(8), 803-818.
Buijs, A., van Rompaey, L., Molijn, A. C., Davis, J. N., Vertegaal, A. C., Potter, M. D., Adams, C., van Baal, S., Zwarthoff, E. C. & Roussel, M. F. (2000) The MN1-TEL fusion protein, encoded by the translocation (12; 22)(p13; q11) in myeloid leukemia, is a transcription factor with transforming activity. *Molecular and cellular biology*, 20(24), 9281-9293.
Burman, J. L., Wasiak, S., Ritter, B., de Heuvel, E. & McPherson, P. S. (2005) Aftiphilin is a component of the clathrin machinery in neurons. *FEBS letters*, 579(10), 2177-2184.
Burstein, E., Hoberg, J. E., Wilkinson, A. S., Rumble, J. M., Csomos, R. A., Komarck, C. M., Maine, G. N., Wilkinson, J. C., Mayo, M. W. & Duckett, C. S. (2005) COMMD proteins: a novel family of structural and functional homologs of MURR1. *Journal of Biological Chemistry*.
Cameron-Vendrig, A., Reheman, A., Siraj, M. A., Xu, X. R., Wang, Y., Lei, X., Afroze, T., Shikatani, E., El-Mounayri, O. & Noyan, H. (2016) Glucagon-like peptide-1 receptor activation attenuates platelet aggregation and thrombosis. *Diabetes*, db151141.
Canault, M., Ghalloussi, D., Grosdidier, C., Guinier, M., Perret, C., Chelghoum, N., Germain, M., Raslova, H., Peiretti, F., Morange, P. E., Saut, N., Pillois, X., Nurden, A. T., Cambien, F., Pierres, A., van den Berg, T. K., Kuijpers, T. W., Alessi, M. C. & Tregouet, D. A. (2014) Human CalDAG-GEFI gene (RASGRP2) mutation affects platelet function and causes severe bleeding. *J Exp Med*, 211(7), 1349-62.
Canver, M. C., Bauer, D. E., Dass, A., Yien, Y. Y., Chung, J., Masuda, T., Maeda, T., Paw, B. H. & Orkin, S. H. (2014) Characterization of genomic deletion efficiency mediated by clustered regularly interspaced palindromic repeats (CRISPR)/Cas9 nuclease system in mammalian cells. *Journal of Biological Chemistry*, 289(31), 21312-21324.
Carestia, A., Rivadeneyra, L., Romaniuk, M. A., Fondevila, C., Negrotto, S. & Schattner, M. (2013) Functional responses and molecular mechanisms involved in histone-mediated platelet activation. *Thrombosis and haemostasis*, 110(05), 1035-1045.
Carrasquillo, R., Tian, D., Krishna, S., Pollak, M. R., Greka, A. & Schlöndorff, J. (2012) SNF8, a member of the ESCRT-II complex, interacts with TRPC6 and enhances its channel activity. *BMC cell biology*, 13(1), 33.
Chakrabarti, S. R. & Nucifora, G. (1999) The leukemia-associated gene TEL encodes a transcription repressor which associates with SMRT and mSin3A. *Biochemical and biophysical research communications*, 264(3), 871-877.

Chakrabarti, S. R., Sood, R., Ganguly, S., Bohlander, S., Shen, Z. & Nucifora, G. (1999) Modulation of TEL transcription activity by interaction with the ubiquitin-conjugating enzyme UBC9. *Proceedings of the National Academy of Sciences*, 96(13), 7467-7472.
Chakrabarti, S. R., Sood, R., Nandi, S. & Nucifora, G. (2000) Posttranslational modification of TEL and TEL/AML1 by SUMO-1 and cell-cycle-dependent assembly into nuclear bodies. *Proceedings of the National Academy of Sciences*, 97(24), 13281-13285.
Chan, M. V. & Warner, T. D. (2012) Standardised optical multichannel (optimul) platelet aggregometry using high-speed shaking and fixed time point readings. *Platelets*, 23(5), 404-8.
Chari, R., Getz, T., Nagy Jr, B., Bhavaraju, K., Mao, Y., Bynagari, Y. S., Murugappan, S., Nakayama, K. & Kunapuli, S. P. (2009) Protein kinase Cδ differentially regulates platelet functional responses. *Arteriosclerosis, thrombosis, and vascular biology*, 29(5), 699-705.
Chatterjee, M., Huang, Z., Zhang, W., Jiang, L., Hultenby, K., Zhu, L., Hu, H., Nilsson, G. P. & Li, N. (2011) Distinct platelet packaging, release, and surface expression of proangiogenic and antiangiogenic factors on different platelet stimuli. *Blood*, 117(14), 3907-3911.
Chen, C., Itakura, E., Weber, K. P., Hegde, R. S. & de Bono, M. (2014) An ER complex of ODR-4 and ODR-8/Ufm1 specific protease 2 promotes GPCR maturation by a Ufm1-independent mechanism. *PLoS genetics*, 10(3), e1004082.
Chew, C. L., Chen, M. & Pandolfi, P. P. (2016) Endosome and INPP4B. *Oncotarget*, 7(1), 5.
Chou, S.-Y., Hsu, K.-S., Otsu, W., Hsu, Y.-C., Luo, Y.-C., Yeh, C., Shehab, S. S., Chen, J., Shieh, V. & He, G.-a. (2016) CLIC4 regulates apical exocytosis and renal tube luminogenesis through retromer-and actin-mediated endocytic trafficking. *Nature communications*, 7, 10412.
Ciau-Uitz, A., Pinheiro, P., Gupta, R., Enver, T. & Patient, R. (2010) Tel1/ETV6 specifies blood stem cells through the agency of VEGF signaling. *Developmental cell*, 18(4), 569-578.
Citterio, L., Simonini, M., Zagato, L., Salvi, E., Carpini, S. D., Lanzani, C., Messaggio, E., Casamassima, N., Frau, F. & D'Avila, F. (2011) Genes involved in vasoconstriction and vasodilation system affect salt-sensitive hypertension. *PLoS One*, 6(5), e19620.
Clemetson, K. J., McGregor, J. L., James, E., Dechavanne, M. & Lüscher, E. (1982) Characterization of the platelet membrane glycoprotein abnormalities in Bernard-Soulier syndrome and comparison with normal by surface-labeling techniques and high-resolution two-dimensional gel electrophoresis. *The Journal of clinical investigation*, 70(2), 304-311.
Cong, L., Ran, F. A., Cox, D., Lin, S., Barretto, R., Habib, N., Hsu, P. D., Wu, X., Jiang, W. & Marraffini, L. (2013) Multiplex genome engineering using CRISPR/Cas systems. *Science*, 1231143.
Cote, M., Menager, M. M., Burgess, A., Mahlaoui, N., Picard, C., Schaffner, C., Al-Manjomi, F., Al-Harbi, M., Alangari, A., Le Deist, F., Gennery, A. R., Prince, N., Cariou, A., Nitschke, P., Blank, U., El-Ghazali, G., Menasche, G., Latour, S., Fischer, A. & de Saint Basile, G. (2009) Munc18-2 deficiency causes familial hemophagocytic lymphohistiocytosis type 5 and impairs cytotoxic granule exocytosis in patient NK cells. *J Clin Invest*, 119(12), 3765-73.
Coyne, H. J., 3rd, De, S., Okon, M., Green, S. M., Bhachech, N., Graves, B. J. & McIntosh, L. P. (2012) Autoinhibition of ETV6 (TEL) DNA binding: appended helices sterically block the ETS domain. *J Mol Biol*, 421(1), 67-84.
Crespin, M., Vidal, C., Picard, F., Lacombe, C. & Fontenay, M. (2009) Activation of PAK1/2 during the shedding of platelet microvesicles. *Blood Coagulation & Fibrinolysis*, 20(1), 63-70.
Crittenden, J. R., Bergmeier, W., Zhang, Y., Piffath, C. L., Liang, Y., Wagner, D. D., Housman, D. E. & Graybiel, A. M. (2004) CalDAG-GEFI integrates signaling for platelet aggregation and thrombus formation. *Nature medicine*, 10(9), 982.
Cui, J. W., Vecchiarelli-Federico, L. M., Li, Y. J., Wang, G. J. & Ben-David, Y. (2009) Continuous Fli-1 expression plays an essential role in the proliferation and survival of F-MuLV-induced erythroleukemia and human erythroleukemia. *Leukemia*, 23(7), 1311-9.
Cullinane, A. R., Curry, J. A., Carmona-Rivera, C., Summers, C. G., Ciccone, C., Cardillo, N. D., Dorward, H., Hess, R. A., White, J. G., Adams, D., Huizing, M. & Gahl, W. A. (2011) A BLOC-1 mutation screen reveals that PLDN is mutated in Hermansky-Pudlak Syndrome type 9. *Am J Hum Genet*, 88(6), 778-787.
Cullinane, A. R., Straatman-Iwanowska, A., Zaucker, A., Wakabayashi, Y., Bruce, C. K., Luo, G., Rahman, F., Gürakan, F., Utine, E. & Özkan, T. B. (2010) Mutations in VIPAR cause an arthrogryposis, renal dysfunction and cholestasis syndrome phenotype with defects in epithelial polarization. *Nature genetics*, 42(4), 303.
Daly, M. E. (2017) Transcription factor defects causing platelet disorders. *Blood Rev*, 31(1), 1-10.
Daly, M. E., Dawood, B. B., Lester, W. A., Peake, I. R., Rodeghiero, F., Goodeve, A. C., Makris, M., Wilde, J. T., Mumford, A. D., Watson, S. P. & Mundell, S. J. (2009) Identification and characterization of a novel P2Y 12 variant in a patient diagnosed with type 1 von Willebrand disease in the European MCMDM-1VWD study. *Blood*, 113(17), 4110-3.
Daly, M. E., Leo, V. C., Lowe, G. C., Watson, S. P. & Morgan, N. V. (2014) What is the role of genetic testing in the investigation of patients with suspected platelet function disorders? *Br J Haematol*, 165(2), 193-203.
Dascher, C. & Balch, W. E. (1996) Mammalian Sly1 regulates syntaxin 5 function in endoplasmic reticulum to Golgi transport. *Journal of Biological Chemistry*, 271(27), 15866-15869.
Dasouki, M. J., Rafi, S. K., Olm-Shipman, A. J., Wilson, N. R., Abhyankar, S., Ganter, B., Furness, L. M., Fang, J., Calado, R. T. & Saadi, I. (2013) Exome sequencing reveals a thrombopoietin ligand mutation in a Micronesian family with autosomal recessive aplastic anemia. *Blood*, 122(20), 3440-9.
Dawood, B. B., Lowe, G. C., Lordkipanidze, 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*, 120(25), 5041-9.
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 A(2) pathways. *Platelets*, 18(5), 329-45.
De Braekeleer, E., Douet-Guilbert, N., Morel, F., Le Bris, M.-J., Basinko, A. & De Braekeleer, M. (2012) ETV6 fusion genes in hematological malignancies: a review. *Leukemia research*, 36(8), 945-961.
de Sauvage, F. J., Carver-Moore, K., Luoh, S.-m., Ryan, A., Dowd, M., Eaton, D. L. & Moore, M. W. (1996) Physiological regulation of early and late stages of megakaryocytopoiesis by thrombopoietin. *Journal of Experimental Medicine*, 183(2), 651-656.
Del Portillo, A., Komissarova, E. V., de Gonzalez, A. K., Bokhari, A., Remotti, H., Sepulveda, J. & Sepulveda, A. (2017) Functional role of Friend Leukemia Integration-1 (FLI1) in gastric carcinogenesis. AACR.
Delattre, O., Zucman, J., Plougastel, B., Desmaze, C., Melot, T., Peter, M., Kovar, H., Joubert, I., de Jong, P. & Rouleau, G. (1992) Gene fusion with an ETS DNA-binding domain caused by chromosome translocation in human tumours. *Nature*, 359(6391), 162.
DeMambro, V. E., Clemmons, D., Horton, L., Bouxsein, M., Wood, T., Beamer, W., Canalis, E. & Rosen, C. J. (2008) Gender-specific changes in bone turnover and skeletal architecture in igfbp-2-null mice. *Endocrinology*, 149(5), 2051-2061.
Deramaudt, B. M. M., Remy, P. & Abraham, N. (1999) Upregulation of human heme oxygenase gene expression by Ets-family proteins. *Journal of cellular biochemistry*, 72(3), 311-321.

Derry, J. M., Ochs, H. D. & Francke, U. (1994) Isolation of a novel gene mutated in Wiskott-Aldrich syndrome. *Cell*, 78(4), 635-644.
Deutsch, V. R. & Tomer, A. (2013) Advances in megakaryocytopoiesis and thrombopoiesis: from bench to bedside. *British journal of haematology*, 161(6), 778-793.
Deveaux, S., Filipe, A., Lemarchandel, V., Ghysdael, J., Romeo, P.-H. & Mignotte, V. (1996) Analysis of the thrombopoietin receptor (MPL) promoter implicates GATA and Ets proteins in the coregulation of megakaryocyte-specific genes. *Blood*, 87(11), 4678-4685.
Dixon-Salazar, T. J., Silhavy, J. L., Udpa, N., Schroth, J., Bielas, S., Schaffer, A. E., Olvera, J., Bafna, V., Zaki, M. S., Abdel-Salam, G. H., Mansour, L. A., Selim, L., Abdel-Hadi, S., Marzouki, N., Ben-Omran, T., Al-Saana, N. A., Sonmez, F. M., Celep, F., Azam, M., Hill, K. J., Collazo, A., Fenstermaker, A. G., Novarino, G., Akizu, N., Garimella, K. V., Sougnez, C., Russ, C., Gabriel, S. B. & Gleeson, J. G. (2012) Exome sequencing can improve diagnosis and alter patient management. *Sci Transl Med*, 4(138), 138ra78.
Donaldson, L. W., Petersen, J. M., Graves, B. J. & McIntosh, L. P. (1994) Secondary structure of the ETS domain places murine Ets-1 in the superfamily of winged helix-turn-helix DNA-binding proteins. *Biochemistry*, 33(46), 13509-13516.
Dumont, B., Lasne, D., Rothschild, C., Bouabdelli, M., Ollivier, V., Oudin, C., Ajzenberg, N., Grandchamp, B. & Jandrot-Perrus, M. (2009) Absence of collagen-induced platelet activation caused by compound heterozygous GPVI mutations. *Blood*, 114(9), 1900-3.
Dursun, A., Yalnizoglu, D., Gerdan, O. F., Yucel-Yilmaz, D., Sagiroglu, M. S., Yuksel, B., Gucer, S., Sivri, S. & Ozgul, R. K. (2017) A probable new syndrome with the storage disease phenotype caused by the VPS33A gene mutation. *Clinical dysmorphology*, 26(1), 1-12.
Duval, D., Labbé, P., Bureau, L., Tourneau, T. L., Norris, R. A., Markwald, R. R., Levine, R., Schott, J.-J. & Mérot, J. (2015) MVP-associated filamin a mutations affect FlnA-PTPN12 (PTP-PEST) interactions. *Journal of cardiovascular development and disease*, 2(3), 233-247.
Dwyer, N. D., Troemel, E. R., Sengupta, P. & Bargmann, C. I. (1998) Odorant receptor localization to olfactory cilia is mediated by ODR-4, a novel membrane-associated protein. *Cell*, 93(3), 455-466.
Eisbacher, M., Holmes, M. L., Newton, A., Hogg, P. J., Khachigian, L. M., Crossley, M. & Chong, B. H. (2003) Protein-protein interaction between Fli-1 and GATA-1 mediates synergistic expression of megakaryocyte-specific genes through cooperative DNA binding. *Molecular and Cellular Biology*, 23(10), 3427-3441.
El-Brolosy, M. A. & Stainier, D. Y. (2017) Genetic compensation: A phenomenon in search of mechanisms. *PLoS genetics*, 13(7), e1006780.
Favier, M., Bordet, J. C., Favier, R., Gkalea, V., Pillois, X., Rameau, P., Debili, N., Alessi, M. C., Nurden, P. & Raslova, H. (2018) Mutations of the integrin αIIb/β3 intracytoplasmic salt bridge cause macrothrombocytopenia and enlarged platelet α-granules. *American journal of hematology*, 93(2), 195-204.
Favier, R., Douay, L., Esteva, B., Portnoi, M., Gaulard, P., Lecompte, T., Perot, C., Adam, M., Lecrubier, C. & den Akker Van, J. (1993) A novel genetic thrombocytopenia (Paris-Trousseau) associated with platelet inclusions, dysmegakaryopoiesis and chromosome deletion AT 11q23. *Comptes rendus de l'Academie des sciences. Serie III, Sciences de la vie*, 316(7), 698-701.
Favier, R., Jondeau, K., Boutard, P., Grossfeld, P., Reinert, P., Jones, C., Bertoni, F. & Cramer, E. M. (2003) Paris-Trousseau syndrome : clinical, hematological, molecular data of ten new cases. *Thromb Haemost*, 90(5), 893-7.
Feijge, M. A., Ansink, K., Vanschoonbeek, K. & Heemskerk, J. W. (2004) Control of platelet activation by cyclic AMP turnover and cyclic nucleotide phosphodiesterase type-3. *Biochemical pharmacology*, 67(8), 1559-1567.
Feldmann, J., Callebaut, I., Raposo, G., Certain, S., Bacq, D., Dumont, C., Lambert, N., Ouachée-Chardin, M., Chedeville, G. & Tamary, H. (2003) Munc13-4 is essential for cytolytic granules fusion and is mutated in a form of familial hemophagocytic lymphohistiocytosis (FHL3). *Cell*, 115(4), 461-473.
Fenrick, R., Amann, J. M., Lutterbach, B., Wang, L., Westendorf, J. J., Downing, J. R. & Hiebert, S. W. (1999) Both TEL and AML-1 contribute repression domains to the t (12; 21) fusion protein. *Molecular and Cellular Biology*, 19(10), 6566-6574.
Fenrick, R., Wang, L., Nip, J., Amann, J. M., Rooney, R. J., Walker-Daniels, J., Crawford, H. C., Hulboy, D. L., Kinch, M. S. & Matrisian, L. M. (2000) TEL, a putative tumor suppressor, modulates cell growth and cell morphology of ras-transformed cells while repressing the transcription of stromelysin-1. *Molecular and cellular biology*, 20(16), 5828-5839.
Ferreira, C. R., Chen, D., Abraham, S. M., Adams, D. R., Simon, K. L., Malicdan, M. C., Markello, T. C., Gunay-Aygun, M. & Gahl, W. A. (2017) Combined alpha-delta platelet storage pool deficiency is associated with mutations in GFI1B. *Molecular genetics and metabolism*, 120(3), 288-294.
Fidler, T. P., Middleton, E., Rowley, J. W., Boudreau, L. H., Campbell, R. A., Souvenir, R., Funari, T., Tessandier, N., Boilard, E. & Weyrich, A. S. (2017) Glucose transporter 3 potentiates degranulation and is required for platelet activation. *Arteriosclerosis, thrombosis, and vascular biology*, ATVBAHA. 117.309184.
Fletcher, S. J., Johnson, B., Lowe, G. C., Bem, D., Drake, S., Lordkipanidze, M., Guiu, I. S., Dawood, B., Rivera, J., Simpson, M. A., Daly, M. E., Motwani, J., Collins, P. W., Watson, S. P., Morgan, N. V., Genotyping, U. K. & Phenotyping of Platelets study, g. (2015) SLFN14 mutations underlie thrombocytopenia with excessive bleeding and platelet secretion defects. *J Clin Invest*, 125(9), 3600-5.
Fletcher, S. J., Pisareva, V. P., Khan, A., Tcherepanov, A., Morgan, N. V. & Pisarev, A. V. (2018) Role of the novel endoribonuclease SLFN14 and its disease causing mutations in ribosomal degradation. *RNA*, rna. 066415.118.
Fox, S. C., May, J. A., Shah, A., Neubert, U. & Heptinstall, S. (2009) Measurement of platelet P-selectin for remote testing of platelet function during treatment with clopidogrel and/or aspirin. *Platelets*, 20(4), 250-9.
Freisinger, C. M., Fisher, R. A. & Slusarski, D. C. (2010) Regulator of g protein signaling 3 modulates wnt5b calcium dynamics and somite patterning. *PLoS genetics*, 6(7), e1001020.
Friend, C. (1957) Cell-free transmission in adult Swiss mice of a disease having the character of a leukemia. *Journal of Experimental Medicine*, 105(4), 307-318.
Fuentes, E., Rojas, A. & Palomo, I. (2016) NF-κB signaling pathway as target for antiplatelet activity. *Blood reviews*, 30(4), 309-315.
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 a severe form of congenital thrombocytopenia. *Blood*, blood-2018-04-847798.
Gandhi, M. J., Cummings, C. L. & Drachman, J. G. (2003) FLJ14813 missense mutation: a candidate for autosomal dominant thrombocytopenia on human chromosome 10. *Human heredity*, 55(1), 66-70.
Garvie, C. W. & Wolberger, C. (2001) Recognition of specific DNA sequences. *Molecular cell*, 8(5), 937-946.
Genevieve, D., Proulle, V., Isidor, B., Bellais, S., Serre, V., Djouadi, F., Picard, C., Vignon-Savoye, C., Bader-Meunier, B., Blanche, S., de Vernejoul, M. C., Legeai-Mallet, L., Fischer, A. M., Le Merrer, M., Dreyfus, M., Gaussem, P., Munnich, A. & Cormier-Daire, V. (2008) Thromboxane synthase mutations in an increased bone density disorder (Ghosal syndrome). *Nat Genet*, 40(3), 284-6.

Georgiou, P., Maroulakou, I., Green, J., Dantis, P., Romanospica, V., Kottaridis, S., Lautenberger, J., Watson, D., Papas, T. & Fischinger, P. (1996) Expression of ets family of genes in systemic lupus erythematosus and Sjogren's syndrome. *International journal of oncology*, 9(1), 9-18.
Ghevaert, C., Salsmann, A., Watkins, N. A., Schaffner-Reckinger, E., Rankin, A., Garner, S. F., Stephens, J., Smith, G. A., Debili, N., Vainchenker, W., de Groot, P. G., Huntington, J. A., Laffan, M., Kieffer, N. & Ouwehand, W. H. (2008) A nonsynonymous SNP in the ITGB3 gene disrupts the conserved membrane-proximal cytoplasmic salt bridge in the alphaIIbbeta3 integrin and cosegregates dominantly with abnormal proplatelet formation and macrothrombocytopenia. *Blood*, 111(7), 3407-14.
Gilissen, C., Hoischen, A., Brunner, H. G. & Veltman, J. A. (2012) Disease gene identification strategies for exome sequencing. *Eur J Hum Genet*, 20(5), 490-7.
Gissen, P., Johnson, C. A., Morgan, N. V., Stapelbroek, J. M., Forshew, T., Cooper, W. N., McKiernan, P. J., Klomp, L. W., Morris, A. A., Wraith, J. E., McClean, P., Lynch, S. A., Thompson, R. J., Lo, B., Quarrell, O. W., Di Rocco, M., Trembath, R. C., Mandel, H., Wali, S., Karet, F. E., Knisely, A. S., Houwen, R. H., Kelly, D. A. & Maher, E. R. (2004) Mutations in VPS33B, encoding a regulator of SNARE-dependent membrane fusion, cause arthrogryposis-renal dysfunction-cholestasis (ARC) syndrome. *Nat Genet*, 36(4), 400-4.
Glembotsky, A. C., Bluteau, D., Espasandin, Y. R., Goette, N. P., Marta, R. F., Marin Oyarzun, C. P., Korin, L., Lev, P. R., Laguens, R. P., Molinas, F. C., Raslova, H. & Heller, P. G. (2014) Mechanisms underlying platelet function defect in a pedigree with familial platelet disorder with a predisposition to acute myelogenous leukemia: potential role for candidate RUNX1 targets. *J Thromb Haemost*, 12(5), 761-72.
Golebiewska, E. M. & Poole, A. W. (2015) Platelet secretion: From haemostasis to wound healing and beyond. *Blood Rev*, 29(3), 153-62.
Golub, T. R., Barker, G. F., Lovett, M. & Gilliland, D. G. (1994) Fusion of PDGF receptor β to a novel ets-like gene, tel, in chronic myelomonocytic leukemia with t (5; 12) chromosomal translocation. *Cell*, 77(2), 307-316.
Gong, H., Shen, B., Flevaris, P., Chow, C., Lam, S. C.-T., Voyno-Yasenetskaya, T. A., Kozasa, T. & Du, X. (2010) G protein subunit Gα13 binds to integrin αIIbβ3 and mediates integrin "outside-in" signaling. *Science*, 327(5963), 340-343.
Gory, S., Dalmon, J., Prandini, M.-H., Kortulewski, T., de Launoit, Y. & Huber, P. (1998) Requirement of a GT box (Sp1 site) and two Ets binding sites for vascular endothelial cadherin gene transcription. *Journal of Biological Chemistry*, 273(12), 6750-6755.
Gosiengfiao, Y., Horvat, R. & Thompson, A. (2007) Transcription factors GATA-1 and Fli-1 regulate human HOXA10 expression in megakaryocytic cells. *DNA and cell biology*, 26(8), 577-587.
Gottfried, I., Ehrlich, M. & Ashery, U. (2010) The Sla2p/HIP1/HIP1R family: similar structure, similar function in endocytosis? : Portland Press Limited.
Göttgens, B., Nastos, A., Kinston, S., Piltz, S., Delabesse, E. C., Stanley, M., Sanchez, M. J., Ciau-Uitz, A., Patient, R. & Green, A. R. (2002) Establishing the transcriptional programme for blood: the SCL stem cell enhancer is regulated by a multiprotein complex containing Ets and GATA factors. *The EMBO journal*, 21(12), 3039-3050.
Graham, G. J., Ren, Q., Dilks, J. R., Blair, P., Whiteheart, S. W. & Flaumenhaft, R. (2009) Endobrevin/VAMP-8-dependent dense granule release mediates thrombus formation in vivo. *Blood*, 114(5), 1083-90.
Green, S. M., Coyne, H. J., 3rd, McIntosh, L. P. & Graves, B. J. (2010) DNA binding by the ETS protein TEL (ETV6) is regulated by autoinhibition and self-association. *J Biol Chem*, 285(24), 18496-504.
Greenberg, S. M., Rosenthal, D. S., Greeley, T. A., Tantravahi, R. & Handin, R. I. (1988) Characterization of a new megakaryocytic cell line: the Dami cell. *Blood*, 72(6), 1968-1977.
Gremmel, T., Frelinger III, A. L. & Michelson, A. D. (2016) Platelet physiology, *Seminars in thrombosis and hemostasis*. Thieme Medical Publishers.
Gresele, P., Physiology, S. o. P., Harrison, P., Gachet, C., Hayward, C., Kenny, D., Mezzano, D., Mumford, A., Nugent, D. & Nurden, A. (2015) Diagnosis of inherited platelet function disorders: guidance from the SSC of the ISTH. *Journal of Thrombosis and Haemostasis*, 13(2), 314-322.
Grossfeld, P. D., Mattina, T., Lai, Z., Favier, R., Jones, K. L., Cotter, F. & Jones, C. (2004) The 11q terminal deletion disorder: a prospective study of 110 cases. *American journal of medical genetics Part A*, 129(1), 51-61.
Gu, W., Payne, E., Sun, S., Burgess, M. & McMillan, N. A. (2011) Inhibition of cervical cancer cell growth in vitro and in vivo with dual shRNAs. *Cancer gene therapy*, 18(3), 219.
Guidez, F., Petrie, K., Ford, A. M., Lu, H., Bennett, C. A., MacGregor, A., Hannemann, J., Ito, Y., Ghysdael, J. & Greaves, M. (2000) Recruitment of the nuclear receptor corepressor N-CoR by the TEL moiety of the childhood leukemia–associated TEL-AML1 oncoprotein. *Blood*, 96(7), 2557-2561.
Guilherme, A., Soriano, N. A., Bose, S., Holik, J., Bose, A., Pomerleau, D. P., Furcinitti, P., Leszyk, J., Corvera, S. & Czech, M. P. (2004) EHD2 and the novel EH domain binding protein EHBP1 couple endocytosis to the actin cytoskeleton. *Journal of Biological Chemistry*, 279(11), 10593-10605.
Gunay-Aygun, M., Falik-Zaccai, T. C., Vilboux, T., Zivony-Elboum, Y., Gumruk, F., Cetin, M., Khayat, M., Boerkoel, C. F., Kfir, N., Huang, Y., Maynard, D., Dorward, H., Berger, K., Kleta, R., Anikster, Y., Arat, M., Freiberg, A. S., Kehrel, B. E., Jurk, K., Cruz, P., Mullikin, J. C., White, J. G., Huizing, M. & Gahl, W. A. (2011) NBEAL2 is mutated in gray platelet syndrome and is required for biogenesis of platelet alpha-granules. *Nat Genet*, 43(8), 732-4.
Gunay-Aygun, M., Zivony-Elboum, Y., Gumruk, F., Geiger, D., Cetin, M., Khayat, M., Kleta, R., Kfir, N., Anikster, Y., Chezar, J., Arcos-Burgos, M., Shalata, A., Stanescu, H., Manaster, J., Arat, M., Edwards, H., Freiberg, A. S., Hart, P. S., Riney, L. C., Patzel, K., Tanpaiboon, P., Markello, T., Huizing, M., Maric, I., Horne, M., Kehrel, B. E., Jurk, K., Hansen, N. F., Cherukuri, P. F., Jones, M., Cruz, P., Mullikin, J. C., Nurden, A., White, J. G., Gahl, W. A. & Falik-Zaccai, T. (2010) Gray platelet syndrome: natural history of a large patient cohort and locus assignment to chromosome 3p. *Blood*, 116(23), 4990-5001.
Haas, S., Hansson, J., Klimmeck, D., Loeffler, D., Velten, L., Uckelmann, H., Wurzer, S., Prendergast, Á. M., Schnell, A. & Hexel, K. (2015) Inflammation-induced emergency megakaryopoiesis driven by hematopoietic stem cell-like megakaryocyte progenitors. *Cell stem cell*, 17(4), 422-434.
Haeussler, M., Schönig, K., Eckert, H., Eschstruth, A., Mianné, J., Renaud, J.-B., Schneider-Maunoury, S., Shkumatava, A., Teboul, L. & Kent, J. (2016) Evaluation of off-target and on-target scoring algorithms and integration into the guide RNA selection tool CRISPOR. *Genome biology*, 17(1), 148.
Halachmi, N. & Lev, Z. (1996) The Sec1 family: a novel family of proteins involved in synaptic transmission and general secretion. *Journal of neurochemistry*, 66(3), 889-897.
Hamamy, H., Makrythanasis, P., Al-Allawi, N., Muhsin, A. A. & Antonarakis, S. E. (2014) Recessive thrombocytopenia likely due to a homozygous pathogenic variant in the FYB gene: case report. *BMC medical genetics*, 15(1), 135.
Hanson, B. J. & Hong, W. (2003) Evidence for a role of SNX16 in regulating traffic between the early and later endosomal compartments. *Journal of Biological Chemistry*, 278(36), 34617-34630.

Hart, A., Melet, F., Grossfeld, P., Chien, K., Jones, C., Tunnacliffe, A., Favier, R. & Bernstein, A. (2000) Fli-1 is required for murine vascular and megakaryocytic development and is hemizygously deleted in patients with thrombocytopenia. *Immunity*, 13(2), 167-177.
He, L. (2018) *Multiple Functions of Cables1 in Hematopoiesis*Paris Saclay.
Heijnen, H. & Van der Sluijs, P. (2015) Platelet secretory behaviour: as diverse as the granules… or not? *Journal of thrombosis and haemostasis*, 13(12), 2141-2151.
Heller, P. G., Glembotsky, A. C., Gandhi, M. J., Cummings, C. L., Pirola, C. J., Marta, R. F., Kornblihtt, L. I., Drachman, J. G. & Molinas, F. C. (2005) Low Mpl receptor expression in a pedigree with familial platelet disorder with predisposition to acute myelogenous leukemia and a novel AML1 mutation. *Blood*, 105(12), 4664-70.
Hendricks, T. J., Fyodorov, D. V., Wegman, L. J., Lelutiu, N. B., Pehek, E. A., Yamamoto, B., Silver, J., Weeber, E. J., Sweatt, J. D. & Deneris, E. S. (2003) Pet-1 ETS gene plays a critical role in 5-HT neuron development and is required for normal anxiety-like and aggressive behavior. *Neuron*, 37(2), 233-247.
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(2), 162-169.
Higashi, T., Ikeda, T., Shirakawa, R., Kondo, H., Kawato, M., Horiguchi, M., Okuda, T., Okawa, K., Fukai, S. & Nureki, O. (2008) Biochemical characterization of the Rho GTPase-regulated actin assembly by diaphanous-related formins, mDia1 and Daam1, in platelets. *Journal of Biological Chemistry*, 283(13), 8746-8755.
Hirata, T., Kakizuka, A., Ushikubi, F., Fuse, I., Okuma, M. & Narumiya, S. (1994) Arg60 to Leu mutation of the human thromboxane A2 receptor in a dominantly inherited bleeding disorder. *Journal of Clinical Investigation*, 94(4), 1662.
Hock, H., Meade, E., Medeiros, S., Schindler, J. W., Valk, P. J., Fujiwara, Y. & Orkin, S. H. (2004) Tel/Etv6 is an essential and selective regulator of adult hematopoietic stem cell survival. *Genes Dev*, 18(19), 2336-41.
Hofmann, I., Crispin, A., Campagna, D., Calicchio, M., Schmitz-Abe, K., Obeng, E. A., Markianos, K., Senis, Y. A. & Fleming, M. D. (2017) Congenital Thrombocytopenia and Myelofibrosis Due to Germline Mutations in G6b-B—a Megakaryocyte-Specific Immunoreceptor Tyrosine-Based Inhibitory Motif (ITIM) Receptor. Am Soc Hematology.
Hollenhorst, P. C., Jones, D. A. & Graves, B. J. (2004) Expression profiles frame the promoter specificity dilemma of the ETS family of transcription factors. *Nucleic acids research*, 32(18), 5693-5702.
Hollenhorst, P. C., McIntosh, L. P. & Graves, B. J. (2011) Genomic and biochemical insights into the specificity of ETS transcription factors. *Annual review of biochemistry*, 80, 437-471.
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(6817), 202-207.
Hong, W. (2005) SNAREs and traffic. *Biochimica et Biophysica Acta (BBA)-Molecular Cell Research*, 1744(2), 120-144.
Hou, C. & Tsodikov, O. V. (2015) Structural Basis for Dimerization and DNA Binding of Transcription Factor FLI1. *Biochemistry*, 54(50), 7365-74.
Hou, C., Weidenbach, S., Cano, K. E., Wang, Z., Mitra, P., Ivanov, D. N., Rohr, J. & Tsodikov, O. V. (2016) Structures of mithramycin analogues bound to DNA and implications for targeting transcription factor FLI1. *Nucleic Acids Res*, 44(18), 8990-9004.
Housden, B. E., Valvezan, A. J., Kelley, C., Sopko, R., Hu, Y., Roesel, C., Lin, S., Buckner, M., Tao, R. & Yilmazel, B. (2015) Identification of potential drug targets for tuberous sclerosis complex by synthetic screens combining CRISPR-based knockouts with RNAi. *Sci. Signal.*, 8(393), rs9-rs9.
Hromas, R., May, W., Denny, C., Raskind, W., Moore, J., Maki, R. A., Beck, E. & Klemsz, M. J. (1993) Human FLI-1 localizes to chromosome 11Q24 and has an aberrant transcript in neuroepithelioma. *Biochimica et Biophysica Acta (BBA)-Gene Structure and Expression*, 1172(1-2), 155-158.
Hsu, P. D., Scott, D. A., Weinstein, J. A., Ran, F. A., Konermann, S., Agarwala, V., Li, Y., Fine, E. J., Wu, X. & Shalem, O. (2013) DNA targeting specificity of RNA-guided Cas9 nucleases. *Nature biotechnology*, 31(9), 827.
Hu, L., Su, P., Li, R., Yin, C., Zhang, Y., Shang, P., Yang, T. & Qian, A. (2016) Isoforms, structures, and functions of versatile spectraplakin MACF1. *BMB reports*, 49(1), 37.
Hu, W., Philips, A. S., Kwok, J. C., Eisbacher, M. & Chong, B. H. (2005) Identification of nuclear import and export signals within Fli-1: roles of the nuclear import signals in Fli-1-dependent activation of megakaryocyte-specific promoters. *Molecular and cellular biology*, 25(8), 3087-3108.
Huang da, W., Sherman, B. T. & Lempicki, R. A. (2009) Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Res*, 37(1), 1-13.
Huang, H., Yu, M., Akie, T. E., Moran, T. B., Woo, A. J., Tu, N., Waldon, Z., Lin, Y. Y., Steen, H. & Cantor, A. B. (2009) Differentiation-dependent interactions between RUNX-1 and FLI-1 during megakaryocyte development. *Molecular and cellular biology*, 29(15), 4103-4115.
Huang, R.-B., Cheng, D., Liao, S.-M., Lu, B., Wang, Q.-Y., Xie, N.-Z., A Troy II, F. & Zhou, G.-P. (2017) The intrinsic relationship between structure and function of the sialyltransferase ST8Sia family members. *Current topics in medicinal chemistry*, 17(21), 2359-2369.
Huang, X., Dai, F., Gaisano, G., Giglou, K., Han, J., Zhang, M., Kittanakom, S., Wong, V., Wei, L. & Showalter, A. (2013) The identification of novel proteins that interact with the GLP-1 receptor and restrain its activity. *Molecular Endocrinology*, 27(9), 1550-1563.
Huynh, H., Iizuka, S., Kaba, M., Kirak, O., Zheng, J., Lodish, H. F. & Zhang, C. C. (2008) Insulin-like growth factor-binding protein 2 secreted by a tumorigenic cell line supports ex vivo expansion of mouse hematopoietic stem cells. *Stem Cells*, 26(6), 1628-1635.
Huynh, H., Zheng, J., Umikawa, M., Zhang, C., Silvany, R., Iizuka, S., Holzenberger, M., Zhang, W. & Zhang, C. C. (2011) IGF binding protein 2 supports the survival and cycling of hematopoietic stem cells. *Blood*, 118(12), 3236-3243.
Ichikawa, M., Asai, T., Saito, T., Yamamoto, G., Seo, S., Yamazaki, I., Yamagata, T., Mitani, K., Chiba, S. & Hirai, H. (2004) AML-1 is required for megakaryocytic maturation and lymphocytic differentiation, but not for maintenance of hematopoietic stem cells in adult hematopoiesis. *Nature medicine*, 10(3), 299.
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. *Proceedings of the National Academy of Sciences*, 96(6), 3132-3136.
Irvin, B. J., Wood, L. D., Wang, L., Fenrick, R., Sansam, C. G., Packham, G., Kinch, M., Yang, E. & Hiebert, S. W. (2003) TEL, a putative tumor suppressor, induces apoptosis and represses transcription of Bcl-XL. *Journal of Biological Chemistry*, 278(47), 46378-46386.
Ito, Y., Nakamura, S., Sugimoto, N., Shigemori, T., Kato, Y., Ohno, M., Sakuma, S., Ito, K., Kumon, H. & Hirose, H. (2018) Turbulence activates platelet biogenesis to enable clinical scale ex vivo production. *Cell*, 174(3), 636-648. e18.
Izumi, R., Niihori, T., Suzuki, N., Sasahara, Y., Rikiishi, T., Nishiyama, A., Nishiyama, S., Endo, K., Kato, M., Warita, H., Konno, H., Takahashi, T., Tateyama, M., Nagashima, T., Funayama, R., Nakayama, K., Kure, S., Matsubara, Y., Aoki, Y. & Aoki, M.

(2014) GNE myopathy associated with congenital thrombocytopenia: a report of two siblings. *Neuromuscul Disord*, 24(12), 1068-72.
Jackson, L. P., Kelly, B. T., McCoy, A. J., Gaffry, T., James, L. C., Collins, B. M., Höning, S., Evans, P. R. & Owen, D. J. (2010) A large-scale conformational change couples membrane recruitment to cargo binding in the AP2 clathrin adaptor complex. *Cell*, 141(7), 1220-1229.
Jacobsen, P., Hauge, M., Henningsen, K., Hobolth, N., Mikkelsen, M. & Philip, J. (1973) An (11; 21) translocation in four generations with chromosome 11 abnormalities in the offspring. *Human heredity*, 23(6), 568-585.
Jalagadugula, G., Goldfinger, L. E., Mao, G., Lambert, M. P. & Rao, A. K. (2018) Defective RAB1B-related megakaryocytic ER-to-Golgi transport in RUNX1 haplodeficiency: impact on von Willebrand factor. *Blood advances*, 2(7), 797-806.
Jalagadugula, G., Mao, G., Kaur, G., Dhanasekaran, D. N. & Rao, A. K. (2011) Platelet protein kinase C-theta deficiency with human RUNX1 mutation: PRKCQ is a transcriptional target of RUNX1. *Arterioscler Thromb Vasc Biol*, 31(4), 921-7.
Jalagadugula, G., Mao, G., Kaur, G., Goldfinger, L. E., Dhanasekaran, D. N. & Rao, A. K. (2010) Regulation of platelet myosin light chain (MYL9) by RUNX1: implications for thrombocytopenia and platelet dysfunction in RUNX1 haplodeficiency. *Blood*, 116(26), 6037-45.
Jena, B. P., Stemmer, P. M., Wang, S., Mao, G., Lewis, K. T. & Walz, D. A. (2017) Human Platelet Vesicles Exhibit Distinct Size and Proteome. *Journal of proteome research*, 16(7), 2333-2338.
Johnson, B., Lowe, G. C., Futterer, J., MacDonald, D., Simpson, M. A., Guiu, I. S., Drake, S., Bem, D., Leo, V. & Fletcher, S. J. (2016) Whole exome sequencing identifies genetic variants in inherited thrombocytopenia with secondary qualitative function defects. *haematologica*, haematol. 2016.146316.
Jones, M. L., Murden, S. L., Bem, D., Mundell, S. J., Gissen, P., Daly, M. E., Watson, S. P., Mumford, A. D. & group, U. G. s. (2012) Rapid genetic diagnosis of heritable platelet function disorders with next-generation sequencing: proof-of-principle with Hermansky-Pudlak syndrome. *J Thromb Haemost*, 10(2), 306-9.
Jonnalagadda, D., Izu, L. T. & Whiteheart, S. W. (2012) Platelet secretion is kinetically heterogeneous in an agonist-responsive manner. *Blood*, blood-2012-07-445080.
Joshi, S. & Whiteheart, S. W. (2017) The nuts and bolts of the platelet release reaction. *Platelets*, 28(2), 129-137.
Kahr, W. H., 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. *Nat Genet*, 43(8), 738-40.
Kahr, W. H., Pluthero, F. G., Elkadri, A., Warner, N., Drobac, M., Chen, C. H., Lo, R. W., Li, L., Li, R. & Li, Q. (2017) Loss of the Arp2/3 complex component ARPC1B causes platelet abnormalities and predisposes to inflammatory disease. *Nature communications*, 8, 14816.
Karim, Z. A., Zhang, J., Banerjee, M., Chicka, M. C., Al Hawas, R., Hamilton, T. R., Roche, P. A. & Whiteheart, S. W. (2013) IκB kinase (IKK) phosphorylation of SNAP-23 controls platelet secretion. *Blood*, blood-2012-11-470468.
Kashiwagi, H., Tomiyama, Y., Kosugi, S., Shiraga, M., Lipsky, R., Kanayama, Y., Kurata, Y. & Matsuzawa, Y. (1994) Identification of molecular defects in a subject with type I CD36 deficiency. *Blood*, 83(12), 3545-3552.
Kaur, G., Jalagadugula, G., Mao, G. & Rao, A. K. (2010) RUNX1/core binding factor A2 regulates platelet 12-lipoxygenase gene (ALOX12): studies in human RUNX1 haplodeficiency. *Blood*, 115(15), 3128-35.
Kaushansky, K. (2009) Determinants of platelet number and regulation of thrombopoiesis. *ASH Education Program Book*, 2009(1), 147-152.
Kauskot, A., Pascreau, T., Adam, F., Bruneel, A., Reperant, C., Lourenco-Rodrigues, M.-D., Rosa, J.-P., Petermann, R., Maurey, H. & Auditeau, C. (2018) A mutation in the gene coding for the sialic acid transporter SLC35A1 is required for platelet life span but not proplatelet formation. *Haematologica*, haematol. 2018.198028.
Kawada, H., Ito, T., Pharr, P. N., Spyropoulos, D. D., Watson, D. K. & Ogawa, M. (2001) Defective megakaryopoiesis and abnormal erythroid development in Fli-1 gene-targeted mice. *International journal of hematology*, 73(4), 463.
Kim, J., Shapiro, M. J., Bamidele, A. O., Gurel, P., Thapa, P., Higgs, H. N., Hedin, K. E., Shapiro, V. S. & Billadeau, D. D. (2014a) Coactosin-like 1 antagonizes cofilin to promote lamellipodial protrusion at the immune synapse. *PLoS One*, 9(1), e85090.
Kim, M.-Y., Lee, K.-M. & Hur, M.-W. (2011) A novel POK protein, ZNF499, induces BAX gene expression with apoptosis. Federation of American Societies for Experimental Biology.
Kim, M. S., Pinto, S. M., Getnet, D., Nirujogi, R. S., Manda, S. S., Chaerkady, R., Madugundu, A. K., Kelkar, D. S., Isserlin, R., Jain, S., Thomas, J. K., Muthusamy, B., Leal-Rojas, P., Kumar, P., Sahasrabuddhe, N. A., Balakrishnan, L., Advani, J., George, B., Renuse, S., Selvan, L. D., Patil, A. H., Nanjappa, V., Radhakrishnan, A., Prasad, S., Subbannayya, T., Raju, R., Kumar, M., Sreenivasamurthy, S. K., Marimuthu, A., Sathe, G. J., Chavan, S., Datta, K. K., Subbannayya, Y., Sahu, A., Yelamanchi, S. D., Jayaram, S., Rajagopalan, P., Sharma, J., Murthy, K. R., Syed, N., Goel, R., Khan, A. A., Ahmad, S., Dey, G., Mudgal, K., Chatterjee, A., Huang, T. C., Zhong, J., Wu, X., Shaw, P. G., Freed, D., Zahari, M. S., Mukherjee, K. K., Shankar, S., Mahadevan, A., Lam, H., Mitchell, C. J., Shankar, S. K., Satishchandra, P., Schroeder, J. T., Sirdeshmukh, R., Maitra, A., Leach, S. D., Drake, C. G., Halushka, M. K., Prasad, T. S., Hruban, R. H., Kerr, C. L., Bader, G. D., Iacobuzio-Donahue, C. A., Gowda, H. & Pandey, A. (2014b) A draft map of the human proteome. *Nature*, 509(7502), 575-81.
Kim, S. M., Chang, H. K., Song, J. W., Koh, H., Han, S. J. & Group, S. P. L. D. R. (2010) Agranular platelets as a cardinal feature of ARC syndrome. *Journal of pediatric hematology/oncology*, 32(4), 253-258.
Kircher, M., Witten, D. M., Jain, P., O'Roak, B. J., Cooper, G. M. & Shendure, J. (2014) A general framework for estimating the relative pathogenicity of human genetic variants. *Nat Genet*, 46(3), 310-5.
Klages, B., Brandt, U., Simon, M. I., Schultz, G. & Offermanns, S. (1999) Activation of G12/G13 results in shape change and Rho/Rho-kinase–mediated myosin light chain phosphorylation in mouse platelets. *The Journal of cell biology*, 144(4), 745-754.
Kleine Balderhaar, H. J. & Ungermann, C. (2013) CORVET and HOPS tethering complexes–coordinators of endosome and lysosome fusion. *J Cell Sci*, 126(6), 1307-1316.
Klemsz, M. J., Maki, R. A., Papayannopoulou, T., Moore, J. & Hromas, R. (1993) Characterization of the ets oncogene family member, fli-1. *Journal of Biological Chemistry*, 268(8), 5769-5773.
Knezevich, S. R., Garnett, M. J., Pysher, T. J., Beckwith, J. B., Grundy, P. E. & Sorensen, P. H. (1998a) ETV6-NTRK3 gene fusions and trisomy 11 establish a histogenetic link between mesoblastic nephroma and congenital fibrosarcoma. *Cancer Research*, 58(22), 5046-5048.
Knezevich, S. R., McFadden, D. E., Tao, W., Lim, J. F. & Sorensen, P. H. (1998b) A novel ETV6-NTRK3 gene fusion in congenital fibrosarcoma. *Nature genetics*, 18(2), 184-187.
Kohler, S., Doelken, S. C., Mungall, C. J., Bauer, S., Firth, H. V., Bailleul-Forestier, I., Black, G. C., Brown, D. L., Brudno, M., Campbell, J., FitzPatrick, D. R., Eppig, J. T., Jackson, A. P., Freson, K., Girdea, M., Helbig, I., Hurst, J. A., Jahn, J., Jackson, L. G., Kelly, A. M., Ledbetter, D. H., Mansour, S., Martin, C. L., Moss, C., Mumford, A., Ouwehand, W. H., Park, S. M., Riggs, E. R., Scott, R. H., Sisodiya, S., Van Vooren, S., Wapner, R. J., Wilkie, A. O., Wright, C. F., Vulto-van Silfhout, A. T., de Leeuw, N., de

Vries, B. B., Washingthon, N. L., Smith, C. L., Westerfield, M., Schofield, P., Ruef, B. J., Gkoutos, G. V., Haendel, M., Smedley, D., Lewis, S. E. & Robinson, P. N. (2014) The Human Phenotype Ontology project: linking molecular biology and disease through phenotype data. *Nucleic Acids Res*, 42(Database issue), D966-74.
Kondo, H., Maksimova, N., Otomo, T., Kato, H., Imai, A., Asano, Y., Kobayashi, K., Nojima, S., Nakaya, A. & Hamada, Y. (2016) Mutation in VPS33A affects metabolism of glycosaminoglycans: a new type of mucopolysaccharidosis with severe systemic symptoms. *Human molecular genetics*, 26(1), 173-183.
Kornblau, S. M., Qiu, Y. H., Zhang, N., Singh, N., Faderl, S., Ferrajoli, A., York, H., Qutub, A. A., Coombes, K. R. & Watson, D. K. (2011) Abnormal expression of FLI1 protein is an adverse prognostic factor in acute myeloid leukemia. *Blood*, 118(20), 5604-5612.
Kos, J., Jevnikar, Z. & Obermajer, N. (2009) The role of cathepsin X in cell signaling. *Cell adhesion & migration*, 3(2), 164-166.
Kosaki, G. (2005) In vivo platelet production from mature megakaryocytes: does platelet release occur via proplatelets? *International journal of hematology*, 81(3), 208-219.
Kozasa, T., Jiang, X., Hart, M. J., Sternweis, P. M., Singer, W. D., Gilman, A. G., Bollag, G. & Sternweis, P. C. (1998) p115 RhoGEF, a GTPase activating protein for Gα12 and Gα13. *Science*, 280(5372), 2109-2111.
Krishnamurti, L., Neglia, J., Nagarajan, R., Berry, S., Lohr, J., Hirsch, B. & White, J. (2001) Paris-Trousseau syndrome platelets in a child with Jacobsen's syndrome. *American journal of hematology*, 66(4), 295-299.
Kruse, E. A., Loughran, S. J., Baldwin, T. M., Josefsson, E. C., Ellis, S., Watson, D. K., Nurden, P., Metcalf, D., Hilton, D. J. & Alexander, W. S. (2009) Dual requirement for the ETS transcription factors Fli-1 and Erg in hematopoietic stem cells and the megakaryocyte lineage. *Proceedings of the National Academy of Sciences*, 106(33), 13814-13819.
Ku, C. S., Cooper, D. N., Polychronakos, C., Naidoo, N., Wu, M. & Soong, R. (2012) Exome sequencing: dual role as a discovery and diagnostic tool. *Ann Neurol*, 71(1), 5-14.
Kubo, M., Czuwara-Ladykowska, J., Moussa, O., Markiewicz, M., Smith, E., Silver, R. M., Jablonska, S., Blaszczyk, M., Watson, D. K. & Trojanowska, M. (2003) Persistent down-regulation of Fli1, a suppressor of collagen transcription, in fibrotic scleroderma skin. *The American journal of pathology*, 163(2), 571-581.
Kuijpers, T. W., van de Vijver, E., Weterman, M. A., de Boer, M., Tool, A. T., van den Berg, T. K., Moser, M., Jakobs, M. E., Seeger, K., Sanal, O., Unal, S., Cetin, M., Roos, D., Verhoeven, A. J. & Baas, F. (2009) LAD-1/variant syndrome is caused by mutations in FERMT3. *Blood*, 113(19), 4740-6.
Kunishima, S., Kashiwagi, H., Otsu, M., Takayama, N., Eto, K., Onodera, M., Miyajima, Y., Takamatsu, Y., Suzumiya, J., Matsubara, K., Tomiyama, Y. & Saito, H. (2011) Heterozygous ITGA2B R995W mutation inducing constitutive activation of the alphaIIbbeta3 receptor affects proplatelet formation and causes congenital macrothrombocytopenia. *Blood*, 117(20), 5479-84.
Kunishima, S., Kobayashi, R., Itoh, T. J., Hamaguchi, M. & Saito, H. (2009) Mutation of the beta1-tubulin gene associated with congenital macrothrombocytopenia affecting microtubule assembly. *Blood*, 113(2), 458-61.
Kunishima, S., Naoe, T., Kamiya, T. & Saito, H. (2001) Novel heterozygous missense mutation in the platelet glycoprotein Ibβ gene associated with isolated giant platelet disorder. *American journal of hematology*, 68(4), 249-255.
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. *Am J Hum Genet*, 92(3), 431-8.
Kwiatkowski, B. A., Bastian, L. S., Bauer, T. R., Tsai, S., Zielinska-Kwiatkowska, A. G. & Hickstein, D. D. (1998) The ets family member Tel binds to the Fli-1 oncoprotein and inhibits its transcriptional activity. *Journal of Biological Chemistry*, 273(28), 17525-17530.
Kyung, J. W., Bae, J. R., Kim, D.-H., Song, W. K. & Kim, S. H. (2016) Epsin1 modulates synaptic vesicle retrieval capacity at CNS synapses. *Scientific Reports*, 6, 31997.
Labun, K., Montague, T. G., Gagnon, J. A., Thyme, S. B. & Valen, E. (2016) CHOPCHOP v2: a web tool for the next generation of CRISPR genome engineering. *Nucleic acids research*, 44(W1), W272-W276.
Lai, M. M., Luo, H. R., Burnett, P. E., Hong, J. J. & Snyder, S. H. (2000) The calcineurin-binding protein cain is a negative regulator of synaptic vesicle endocytosis. *Journal of Biological Chemistry*, 275(44), 34017-34020.
Lam, F. W., Cruz, M. A., Leung, H.-C. E., Parikh, K. S., Smith, C. W. & Rumbaut, R. E. (2013) Histone induced platelet aggregation is inhibited by normal albumin. *Thrombosis research*, 132(1), 69-76.
Landry, J.-R., Kinston, S., Knezevic, K., Donaldson, I. J., Green, A. R. & Göttgens, B. (2005) Fli1, Elf1, and Ets1 regulate the proximal promoter of the LMO2 gene in endothelial cells. *Blood*, 106(8), 2680-2687.
Lasič, E., Stenovec, M., Kreft, M., Robinson, P. J. & Zorec, R. (2017) Dynamin regulates the fusion pore of endo-and exocytotic vesicles as revealed by membrane capacitance measurements. *Biochimica et Biophysica Acta (BBA)-General Subjects*, 1861(9), 2293-2303.
Latger-Cannard, V., Bensoussan, D., Gregoire, M.-J., Marcon, F., Cloez, J.-L., Leheup, B., Jonveaux, P., Lecompte, T. & Bordigoni, P. (2004) Frequency of thrombocytopenia and large platelets correlates neither with conotruncal cardiac anomalies nor immunological features in the chromosome 22q11. 2 deletion syndrome. *European journal of pediatrics*, 163(6), 327-328.
Le Bras, A., Samson, C., Trentini, M., Caetano, B., Lelievre, E., Mattot, V., Beermann, F. & Soncin, F. (2010) VE-statin/egfl7 expression in endothelial cells is regulated by a distal enhancer and a proximal promoter under the direct control of Erg and GATA-2. *PLoS One*, 5(8), e12156.
Lee, G. M., Donaldson, L. W., Pufall, M. A., Kang, H. S., Pot, I., Graves, B. J. & McIntosh, L. P. (2005) The structural and dynamic basis of Ets-1 DNA binding autoinhibition. *J Biol Chem*, 280(8), 7088-99.
Lee, W. J., Lim, J. S. & Lee, Y. S. (2015) Multiple cerebral hemorrhages associated with Friend leukemia integration 1 (FLI1) positive cardiac angiosarcoma and left atrial thrombi. *International Journal of Stroke*, 10(2).
Leeman-Neill, R. J., Kelly, L. M., Liu, P., Brenner, A. V., Little, M. P., Bogdanova, T. I., Evdokimova, V. N., Hatch, M., Zurnadzy, L. Y. & Nikiforova, M. N. (2014) ETV6-NTRK3 is a common chromosomal rearrangement in radiation-associated thyroid cancer. *Cancer*, 120(6), 799-807.
Lejeune, F. & Maquat, L. E. (2005) Mechanistic links between nonsense-mediated mRNA decay and pre-mRNA splicing in mammalian cells. *Current opinion in cell biology*, 17(3), 309-315.
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(7616), 285.
Lemarchandel, V., Ghysdael, J., Mignotte, V., Rahuel, C. & Romeo, P. (1993) GATA and Ets cis-acting sequences mediate megakaryocyte-specific expression. *Molecular and Cellular Biology*, 13(1), 668-676.
Leo, V. C., Morgan, N. V., Bem, D., Jones, M. L., Lowe, G. C., Lordkipanidze, M., Drake, S., Simpson, M. A., Gissen, P., Mumford, A., Watson, S. P., Daly, M. E. & Group, U. G. S. (2015) Use of next-generation sequencing and candidate gene analysis to identify underlying defects in patients with inherited platelet function disorders. *J Thromb Haemost*, 13(4), 643-50.

Lev, P., Goette, N., Glembotsky, A., Laguens, R., Meckert, P. C., Salim, J., Heller, P., Pozner, R., Marta, R. & Molinas, F. (2011) Production of functional platelet-like particles by the megakaryoblastic DAMI cell line provides a model for platelet biogenesis. *Platelets*, 22(1), 26-36.
Li, L., Shimada, T., Takahashi, H., Koumoto, Y., Shirakawa, M., Takagi, J., Zhao, X., Tu, B., Jin, H. & Shen, Z. (2013) MAG 2 and three MAG 2-INTERACTING PROTEIN s form an ER-localized complex to facilitate storage protein transport in Arabidopsis thaliana. *The Plant Journal*, 76(5), 781-791.
Li, W., Zhang, Q., Oiso, N., Novak, E. K., Gautam, R., O'Brien, E. P., Tinsley, C. L., Blake, D. J., Spritz, R. A., Copeland, N. G., Jenkins, N. A., Amato, D., Roe, B. A., Starcevic, M., Dell'Angelica, E. C., Elliott, R. W., Mishra, V., Kingsmore, S. F., Paylor, R. E. & Swank, R. T. (2003a) Hermansky-Pudlak syndrome type 7 (HPS-7) results from mutant dysbindin, a member of the biogenesis of lysosome-related organelles complex 1 (BLOC-1). *Nat Genet*, 35(1), 84-9.
Li, X., Law, J. & Lee, A. (2012) Semaphorin 5A and plexin-B3 regulate human glioma cell motility and morphology through Rac1 and the actin cytoskeleton. *Oncogene*, 31(5), 595.
Li, Z., Delaney, M. K., O'Brien, K. A. & Du, X. (2010) Signaling during platelet adhesion and activation. *Arterioscler Thromb Vasc Biol*, 30(12), 2341-9.
Li, Z., Zhang, G., Le Breton, G. C., Gao, X., Malik, A. B. & Du, X. (2003b) Two waves of platelet secretion induced by thromboxane A2 receptor and a critical role for phosphoinositide 3-kinases. *Journal of Biological Chemistry*, 278(33), 30725-30731.
Liang, H., Mao, X., Olejniczak, E. T., Nettesheim, D. G., Yu, L., Meadows, R. P., Thompson, C. B. & Fesik, S. W. (1994) Solution structure of the ets domain of Fli-1 when bound to DNA. *Nature Structural & Molecular Biology*, 1(12), 871-876.
Lin, W.-J., Yang, C.-Y., Lin, Y.-C., Tsai, M.-C., Yang, C.-W., Tung, C.-Y., Ho, P.-Y., Kao, F.-J. & Lin, C.-H. (2010) Phafin2 modulates the structure and function of endosomes by a Rab5-dependent mechanism. *Biochemical and biophysical research communications*, 391(1), 1043-1048.
Liu, F., Walmsley, M., Rodaway, A. & Patient, R. (2008) Fli1 acts at the top of the transcriptional network driving blood and endothelial development. *Curr Biol*, 18(16), 1234-40.
Liu, Y.-S., Guo, X.-Y., Hirata, T., Rong, Y., Motooka, D., Kitajima, T., Murakami, Y., Gao, X.-D., Nakamura, S. & Kinoshita, T. (2018) N-Glycan–dependent protein folding and endoplasmic reticulum retention regulate GPI-anchor processing. *J Cell Biol*, 217(2), 585-599.
Lobingier, B. T. & Merz, A. J. (2012) Sec1/Munc18 protein Vps33 binds to SNARE domains and the quaternary SNARE complex. *Molecular biology of the cell*, 23(23), 4611-4622.
Londin, E. R., Hatzimichael, E., Loher, P., Edelstein, L., Shaw, C., Delgrosso, K., Fortina, P., Bray, P. F., McKenzie, S. E. & Rigoutsos, I. (2014) The human platelet: strong transcriptome correlations among individuals associate weakly with the platelet proteome. *Biology direct*, 9(1), 3.
Long, M. W., Smolen, J. E., Szczepanski, P. & Boxer, L. (1984) Role of phorbol diesters in in vitro murine megakaryocyte colony formation. *The Journal of clinical investigation*, 74(5), 1686-1692.
Looney, A. P., Han, R., Stawski, L., Marden, G., Iwamoto, M. & Trojanowska, M. (2017) Synergistic role of endothelial ERG and FLI1 in mediating pulmonary vascular homeostasis. *American journal of respiratory cell and molecular biology*, 57(1), 121-131.
Lopez, R. G., Carron, C., Oury, C., Gardellin, P., Bernard, O. & Ghysdael, J. (1999) TEL is a sequence-specific transcriptional repressor. *Journal of Biological Chemistry*, 274(42), 30132-30138.
Louwette, S., Van Geet, C. & Freson, K. (2012) Regulators of G protein signaling: role in hematopoiesis, megakaryopoiesis and platelet function. *Journal of Thrombosis and Haemostasis*, 10(11), 2215-2222.
Lova, P., Paganini, S., Hirsch, E., Barberis, L., Wymann, M., Sinigaglia, F., Balduini, C. & Torti, M. (2003) A selective role for phosphatidylinositol 3, 4, 5-trisphosphate in the Gi-dependent activation of platelet Rap1B. *Journal of Biological Chemistry*, 278(1), 131-138.
Lui-Roberts, W. W., Ferraro, F., Nightingale, T. D. & Cutler, D. F. (2008) Aftiphilin and γ-synergin are required for secretagogue sensitivity of Weibel-Palade bodies in endothelial cells. *Molecular biology of the cell*, 19(12), 5072-5081.
Lumelsky, N. L. & Schwartz, B. S. (1997) Protein kinase C in erythroid and megakaryocytic differentiation: possible role in lineage determination. *Biochimica et Biophysica Acta (BBA)-Molecular Cell Research*, 1358(1), 79-92.
Ma, Y.-Q., Qin, J., Wu, C. & Plow, E. F. (2008) Kindlin-2 (Mig-2): a co-activator of β3 integrins. *The Journal of cell biology*, 181(3), 439-446.
Mager, A., Grapin-Botton, A., Ladjali, K., Meyer, D., Wolff, C., Stiegler, P., Bonnin, M. & Remy, P. (2004) The avian fli gene is specifically expressed during embryogenesis in a subset of neural crest cells giving rise to mesenchyme. *International Journal of Developmental Biology*, 42(4), 561-572.
Mahaut-Smith, M. P., Jones, S. & Evans, R. J. (2011) The P2X1 receptor and platelet function. *Purinergic signalling*, 7(3), 341-356.
Mali, P., Yang, L., Esvelt, K. M., Aach, J., Guell, M., DiCarlo, J. E., Norville, J. E. & Church, G. M. (2013) RNA-guided human genome engineering via Cas9. *Science*, 339(6121), 823-826.
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. *Blood*, 124(16), 2554-63.
Mancini, E., Sanjuan-Pla, A., Luciani, L., Moore, S., Grover, A., Zay, A., Rasmussen, K. D., Luc, S., Bilbao, D. & O'carroll, D. (2012) FOG-1 and GATA-1 act sequentially to specify definitive megakaryocytic and erythroid progenitors. *The EMBO journal*, 31(2), 351-365.
Mao, G., Goldfinger, L., Fan, D., Lambert, M., Jalagadugula, G., Freishtat, R. & Rao, A. (2017) Dysregulation of PLDN (pallidin) is a mechanism for platelet dense granule deficiency in RUNX 1 haplodeficiency. *Journal of Thrombosis and Haemostasis*, 15(4), 792-801.
Marat, A. L., Dokainish, H. & McPherson, P. S. (2011) DENN domain proteins: regulators of Rab GTPases. *Journal of Biological Chemistry*, jbc. R110. 217067.
Marconi, C., Di Buduo, C. A., LeVine, K., Barozzi, S., Faleschini, M., Bozzi, V., Palombo, F., McKinstry, S., Lassandro, G. & Giordano, P. (2018) A new form of inherited thrombocytopenia caused by loss-of-function mutations in PTPRJ. *Blood*, blood-2018-07-859496.
Markello, T., Chen, D., Kwan, J. Y., Horkayne-Szakaly, I., Morrison, A., Simakova, O., Maric, I., Lozier, J., Cullinane, A. R. & Kilo, T. (2015) York platelet syndrome is a CRAC channelopathy due to gain-of-function mutations in STIM1. *Molecular genetics and metabolism*, 114(3), 474-482.
Marks-Bluth, J., Khanna, A., Chandrakanthan, V., Thoms, J., Bee, T., Eich, C., Kang, Y. C., Knezevic, K., Qiao, Q. & Fitch, S. (2015) SMAD1 and SMAD5 expression is coordinately regulated by FLI1 and GATA2 during endothelial development. *Molecular and cellular biology*, 35(12), 2165-2172.

Marneth, A. E., Van Heerde, W. L., Hebeda, K. M., Laros-van Gorkom, B. A., Barteling, W., Willemsen, B., de Graaf, A. O., Simons, A., Jansen, J. H. & Preijers, F. (2017) Platelet CD34 expression and α/δ-granule abnormalities in GFI1B-and RUNX1-related familial bleeding disorders. *Blood*, 129(12), 1733-1736.
Martens, L., Van Damme, P., Van Damme, J., Staes, A., Timmerman, E., Ghesquiere, B., Thomas, G. R., Vandekerckhove, J. & Gevaert, K. (2005) The human platelet proteome mapped by peptide-centric proteomics: a functional protein profile. *Proteomics*, 5(12), 3193-204.
Masuya, M., Moussa, O., Abe, T., Deguchi, T., Higuchi, T., Ebihara, Y., Spyropoulos, D. D., Watson, D. K. & Ogawa, M. (2005) Dysregulation of granulocyte, erythrocyte, and NK cell lineages in Fli-1 gene–targeted mice. *Blood*, 105(1), 95-102.
Mathenia, J., Reyes-Cortes, E., Williams, S., Molano, I., Ruiz, P., Watson, D., Gilkeson, G. & Zhang, X. (2010) Impact of Fli-1 transcription factor on autoantibody and lupus nephritis in NZM2410 mice. *Clinical & Experimental Immunology*, 162(2), 362-371.
Matsuda, A., Suzuki, Y., Honda, G., Muramatsu, S., Matsuzaki, O., Nagano, Y., Doi, T., Shimotohno, K., Harada, T. & Nishida, E. (2003) Large-scale identification and characterization of human genes that activate NF-κB and MAPK signaling pathways. *Oncogene*, 22(21), 3307.
Mayer, L., Jasztal, M., Pardo, M., de Haro, S. A., Collins, J., Bariana, T. K., Smethurst, P. A., Grassi, L., Petersen, R. & Nurden, P. (2018) Nbeal2 interacts with Dock7, Sec16a, and Vac14. *Blood*, 131(9), 1000-1011.
McClellan, D., Singer, J., Velinder, M., Bareyan, D., Casey, M., Lucente, H., Weston, A., Sharma, S. & Engel, M. (2017) GFI1B Splice Variants Differentially Regulate Transcription and Megakaryocyte/Erythrocyte Fate in Hematopoiesis. Am Soc Hematology.
McLaughlin, F., Ludbrook, V. J., Cox, J., von Carlowitz, I., Brown, S. & Randi, A. M. (2001) Combined genomic and antisense analysis reveals that the transcription factor Erg is implicated in endothelial cell differentiation. *Blood*, 98(12), 3332-3339.
Melazzini, F., Palombo, F., Balduini, A., De Rocco, D., Marconi, C., Noris, P., Gnan, C., Pippucci, T., Bozzi, V., Faleschini, M., Barozzi, S., Doubek, M., Di Buduo, C. A., Kozubik, K. S., Radova, L., Loffredo, G., Pospisilova, S., Alfano, C., Seri, M., Balduini, C. L., Pecci, A. & Savoia, A. (2016) Clinical and pathogenic features of ETV6-related thrombocytopenia with predisposition to acute lymphoblastic leukemia. *Haematologica*, 101(11), 1333-1342.
Mélet, F., Motro, B., Rossi, D. J., Zhang, L. & Bernstein, A. (1996) Generation of a novel Fli-1 protein by gene targeting leads to a defect in thymus development and a delay in Friend virus-induced erythroleukemia. *Molecular and cellular biology*, 16(6), 2708-2718.
Ménasché, G., Ho, C. H., Sanal, O., Feldmann, J., Tezcan, I., Ersoy, F., Houdusse, A., Fischer, A. & Basile, G. d. S. (2003) Griscelli syndrome restricted to hypopigmentation results from a melanophilin defect (GS3) or a MYO5A F-exon deletion (GS1). *J Clin Invest*, 112(3), 450-6.
Ménasché, G., Pastural, E., Feldmann, J., Certain, S., Ersoy, F., Dupuis, S., Wulffraat, N., Bianchi, D., Fischer, A., Le Deist, F. & de Saint Basile, G. (2000) Mutations in RAB27A cause Griscelli syndrome associated with haemophagocytic syndrome. *Nature Genetics*, 25, 173.
Miller, J. L. (1984) Platelet function testing: an improved approach utilizing lumi-aggregation and an interactive computer system. *American journal of clinical pathology*, 81(4), 471-476.
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(11), 4761-4765.
Ming, Z., Hu, Y., Xiang, J., Polewski, P., Newman, P. J. & Newman, D. K. (2011) Lyn and PECAM-1 function as interdependent inhibitors of platelet aggregation. *Blood*, 117(14), 3903-6.
Mitrovic, S., Ben-Tekaya, H., Koegler, E., Gruenberg, J. & Hauri, H.-P. (2008) The cargo receptors Surf4, endoplasmic reticulum-Golgi intermediate compartment (ERGIC)-53, and p25 are required to maintain the architecture of ERGIC and Golgi. *Molecular biology of the cell*, 19(5), 1976-1990.
Miura, K., Jacques, K. M., Stauffer, S., Kubosaki, A., Zhu, K., Hirsch, D. S., Resau, J., Zheng, Y. & Randazzo, P. A. (2002) ARAP1: a point of convergence for Arf and Rho signaling. *Molecular cell*, 9(1), 109-119.
Montague, T. G., Cruz, J. M., Gagnon, J. A., Church, G. M. & Valen, E. (2014) CHOPCHOP: a CRISPR/Cas9 and TALEN web tool for genome editing. *Nucleic acids research*, 42(W1), W401-W407.
Moraes, L. A., Barrett, N. E., Jones, C., Holbrook, L., Spyridon, M., Sage, T., Newman, D. & Gibbins, J. M. (2010) Platelet endothelial cell adhesion molecule-1 regulates collagen-stimulated platelet function by modulating the association of phosphatidylinositol 3-kinase with Grb-2-associated binding protein-1 and linker for activation of T cells. *Journal of Thrombosis and Haemostasis*, 8(11), 2530-2541.
Moreau-Gachelin, F., Tavitian, A. & Tambourin, P. (1988) Spi-1 is a putative oncogene in virally induced murine erythroleukaemias. *Nature*, 331(6153), 277.
Morgan, N. V., Pasha, S., Johnson, C. A., Ainsworth, J. R., Eady, R. A., Dawood, B., McKeown, C., Trembath, R. C., Wilde, J., Watson, S. P. & Maher, E. R. (2006) A germline mutation in BLOC1S3/reduced pigmentation causes a novel variant of Hermansky-Pudlak syndrome (HPS8). *Am J Hum Genet*, 78(1), 160-6.
Morgan, V., Lowe, G., Motwani, J., Williams, M., Simpson, M., Kirby, G., Maher, E. & Watson, S. (2013) A mutation in Ankrd18a is associated with a severe congenital thrombocytopenia. *Journal of Thrombosis and Haemostasis*, 11, 42.
Morison, I. M., Cramer Borde, E. M., Cheesman, E. J., Cheong, P. L., Holyoake, A. J., Fichelson, S., Weeks, R. J., Lo, A., Davies, S. M., Wilbanks, S. M., Fagerlund, R. D., Ludgate, M. W., da Silva Tatley, F. M., Coker, M. S., Bockett, N. A., Hughes, G., Pippig, D. A., Smith, M. P., Capron, C. & Ledgerwood, E. C. (2008) A mutation of human cytochrome c enhances the intrinsic apoptotic pathway but causes only thrombocytopenia. *Nat Genet*, 40(4), 387-9.
Moriyama, T., Metzger, M. L., Wu, G., Nishii, R., Qian, M., Devidas, M., Yang, W., Cheng, C., Cao, X. & Quinn, E. (2015) Germline genetic variation in ETV6 and risk of childhood acute lymphoblastic leukaemia: a systematic genetic study. *The Lancet Oncology*, 16(16), 1659-66.
Moser, M., Nieswandt, B., Ussar, S., Pozgajova, M. & Fässler, R. (2008) Kindlin-3 is essential for integrin activation and platelet aggregation. *Nature medicine*, 14(3), 325.
Motokawa, M., Watanabe, S., Nakatomi, A., Kondoh, T., Matsumoto, T., Morifuji, K., Sawada, H., Nishimura, T., Nunoi, H. & Yoshiura, K.-i. (2018) A hot-spot mutation in CDC42 (p. Tyr64Cys) and novel phenotypes in the third patient with Takenouchi-Kosaki syndrome. *Journal of human genetics*, 63(3), 387.
Moussa, O., LaRue, A. C., Abangan, R. S., Williams, C. R., Zhang, X. K., Masuya, M., Gong, Y. Z., Spyropoulos, D. D., Ogawa, M. & Gilkeson, G. (2010) Thrombocytopenia in mice lacking the carboxy-terminal regulatory domain of the Ets transcription factor Fli1. *Molecular and cellular biology*, 30(21), 5194-5206.
Mumford, A. D., Dawood, B. B., Daly, M. E., Murden, S. L., Williams, M. D., Protty, M. B., Spalton, J. C., Wheatley, M., Mundell, S. J. & Watson, S. P. (2010) A novel thromboxane A2 receptor D304N variant that abrogates ligand binding in a patient with a bleeding diathesis. *Blood*, 115(2), 363-9.
Mumford, A. D., Nisar, S., Darnige, L., Jones, M. L., Bachelot-Loza, C., Gandrille, S., Zinzindohoue, F., Fischer, A. M., Mundell, S. J., Gaussem, P. & Group, U. G. S. (2013) Platelet dysfunction associated with the novel Trp29Cys thromboxane A(2) receptor variant. *J Thromb Haemost*, 11(3), 547-54.

Muñiz, M. & Riezman, H. (2016) Trafficking of glycosylphosphatidylinositol anchored proteins from the endoplasmic reticulum to the cell surface. *Journal of lipid research*, 57(3), 352-360.
Nagle, D. L., Karim, M. A., Woolf, E. A., Holmgren, L., Bork, P., Misumi, D. J., McGrail, S. H., Dussault, B. J., Perou, C. M. & Boissy, R. E. (1996) Identification and mutation analysis of the complete gene for Chediak–Higashi syndrome. *Nature genetics*, 14(3), 307-311.
Nesbit, M. A., Hannan, F. M., Howles, S. A., Reed, A. A., Cranston, T., Thakker, C. E., Gregory, L., Rimmer, A. J., Rust, N. & Graham, U. (2013) Mutations in AP2S1 cause familial hypocalciuric hypercalcemia type 3. *Nature genetics*, 45(1), 93.
Neuhaus, C., Lang-Roth, R., Zimmermann, U., Heller, R., Eisenberger, T., Weikert, M., Markus, S., Knipper, M. & Bolz, H. (2017) Extension of the clinical and molecular phenotype of DIAPH1-associated autosomal dominant hearing loss (DFNA1). *Clinical genetics*, 91(6), 892-901.
Nguyen, K. A., Hamzeh-Cognasse, H., Palle, S., Anselme-Bertrand, I., Arthaud, C.-A., Chavarin, P., Pozzetto, B., Garraud, O. & Cognasse, F. (2014) Role of siglec-7 in apoptosis in human platelets. *PLoS One*, 9(9), e106239.
Nichols, K. E., Crispino, J. D., Poncz, M., White, J. G., Orkin, S. H., Maris, J. M. & Weiss, M. J. (2000) Familial dyserythropoietic anaemia and thrombocytopenia due to an inherited mutation in GATA1. *Nature genetics*, 24(3), 266-270.
Nieswandt, B., Moser, M., Pleines, I., Varga-Szabo, D., Monkley, S., Critchley, D. & Fässler, R. (2007) Loss of talin1 in platelets abrogates integrin activation, platelet aggregation, and thrombus formation in vitro and in vivo. *Journal of Experimental Medicine*, 204(13), 3113-3118.
Niihori, T., Ouchi-Uchiyama, M., Sasahara, Y., Kaneko, T., Hashii, Y., Irie, M., Sato, A., Saito-Nanjo, Y., Funayama, R. & Nagashima, T. (2015) Mutations in MECOM, encoding oncoprotein EVI1, cause radioulnar synostosis with amegakaryocytic thrombocytopenia. *The American Journal of Human Genetics*, 97(6), 848-854.
Nisar, S., Daly, M. E., Federici, A. B., Artoni, A., Mumford, A. D., Watson, S. P. & Mundell, S. J. (2011) An intact PDZ motif is essential for correct P2Y12 purinoceptor traffic in human platelets. *Blood*, 118(20), 5641-51.
Nisar, S. P., Lordkipanidze, M., Jones, M. L., Dawood, B., Murden, S., Cunningham, M. R., Mumford, A. D., Wilde, J. T., Watson, S. P., Mundell, S. J., Lowe, G. C. & group, U. G. s. (2014) A novel thromboxane A2 receptor N42S variant results in reduced surface expression and platelet dysfunction. *Thromb Haemost*, 111(5), 923-32.
Noetzli, L., Lo, R. W., Lee-Sherick, A. B., Callaghan, M., Noris, P., Savoia, A., Rajpurkar, M., Jones, K., Gowan, K., Balduini, C., 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. *Nat Genet*, 47(5), 535-538.
Nordentoft, I. & Jørgensen, P. (2003) The acetyltransferase 60 kDa trans-acting regulatory protein of HIV type 1-interacting protein (Tip60) interacts with the translocation E26 transforming-specific leukaemia gene (TEL) and functions as a transcriptional co-repressor. *Biochemical Journal*, 374(1), 165-173.
Noris, P., Marconi, C., De Rocco, D., Melazzini, F., Pippucci, T., Loffredo, G., Giangregorio, T., Pecci, A., Seri, M. & Savoia, A. (2018) A new form of inherited thrombocytopenia due to monoallelic loss of function mutation in the thrombopoietin gene. *British journal of haematology*, 181(5), 698-701.
North, T. E., De Bruijn, M. F., Stacy, T., Talebian, L., Lind, E., Robin, C., Binder, M., Dzierzak, E. & Speck, N. A. (2002) Runx1 expression marks long-term repopulating hematopoietic stem cells in the midgestation mouse embryo. *Immunity*, 16(5), 661-672.
Nurden, A. & Caen, J. (1975) Specific roles for platelet surface glycoproteins in platelet function. *Nature*, 255(5511), 720.
Nurden, A. T., Nurden, P., Bermejo, E., Combrie, R., McVicar, D. W. & Washington, A. V. (2008) Phenotypic heterogeneity in the Gray platelet syndrome extends to the expression of TREM family member, TLT-1. *Thromb Haemost*, 100(1), 45-51.
Nurden, P., Debili, N., Coupry, I., Bryckaert, M., Youlyouz-Marfak, I., Sole, G., Pons, A. C., Berrou, E., Adam, F., Kauskot, A., Lamaziere, 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. *Blood*, 118(22), 5928-37.
Obergfell, A., Eto, K., Mocsai, A., Buensuceso, C., Moores, S. L., Brugge, J. S., Lowell, C. A. & Shattil, S. J. (2002) Coordinate interactions of Csk, Src, and Syk kinases with αIIbβ3 initiate integrin signaling to the cytoskeleton. *The Journal of cell biology*, 157(2), 265-275.
Oh, J., Bailin, T., Fukai, K., Feng, G. H., Ho, L., Mao, J.-i., Frenk, E., Tamura, N. & Spritz, R. A. (1996) Positional cloning of a gene for Hermansky–Pudlak syndrome, a disorder of cytoplasmic organelles. *Nature genetics*, 14(3), 300-306.
Ohno, H. (2006) Physiological roles of clathrin adaptor AP complexes: lessons from mutant animals. *Journal of biochemistry*, 139(6), 943-948.
Okada, Y., Watanabe, M., Nakai, T., Kamikawa, Y., Shimizu, M., Fukuhara, Y., Yonekura, M., Matsuura, E., Hoshika, Y., Nagai, R., Aird, W. C. & Doi, T. (2013) RUNX1, but not its familial platelet disorder mutants, synergistically activates PF4 gene expression in combination with ETS family proteins. *J Thromb Haemost*, 11(9), 1742-50.
Okamoto, P. M., Gamby, C., Wells, D., Fallon, J. & Vallee, R. B. (2001) Dynamin isoform-specific interaction with the shank/ProSAP scaffolding proteins of the postsynaptic density and actin cytoskeleton. *Journal of Biological Chemistry*, 276(51), 48458-48465.
Ombrello, M. J., Remmers, E. F., Sun, G., Freeman, A. F., Datta, S., Torabi-Parizi, P., Subramanian, N., Bunney, T. D., Baxendale, R. W. & Martins, M. S. (2012) Cold urticaria, immunodeficiency, and autoimmunity related to PLCG2 deletions. *New England Journal of Medicine*, 366(4), 330-338.
Onodera, K., Shavit, J. A., Motohashi, H., Yamamoto, M. & Engel, J. D. (2000) Perinatal synthetic lethality and hematopoietic defects in compound mafG:: mafK mutant mice. *The EMBO journal*, 19(6), 1335-1345.
Oury, C., Toth-Zsamboki, E., Van Geet, C., Thys, C., Wei, L., Nilius, B., Vermylen, J. & Hoylaerts, M. F. (2000) A natural dominant negative P2X1 receptor due to deletion of a single amino acid residue. *J Biol Chem*, 275(30), 22611-4.
Palazzo, A., Bluteau, O., Messaoudi, K., Marangoni, F., Chang, Y., Souquere, S., Pierron, G., Lapierre, V., Zheng, Y. & Vainchenker, W. (2016) The cell division control protein 42–Src family kinase–neural Wiskott–Aldrich syndrome protein pathway regulates human proplatelet formation. *Journal of Thrombosis and Haemostasis*, 14(12), 2524-2535.
Pan, J., Lordier, L., Meyran, D., Rameau, P., Lecluse, Y., Kitchen-Goosen, S., Badirou, I., Mokrani, H., Narumiya, S. & Alberts, A. S. (2014) The formin DIAPH1 (mDia1) regulates megakaryocyte proplatelet formation by remodeling the actin and microtubule cytoskeletons. *Blood*, blood-2013-12-544924.
Pang, L., Xue, H.-H., Szalai, G., Wang, X., Wang, Y., Watson, D. K., Leonard, W. J., Blobel, G. A. & Poncz, M. (2006) Maturation stage–specific regulation of megakaryopoiesis by pointed-domain Ets proteins. *Blood*, 108(7), 2198-206.
Paradisi, I. & Arias, S. (2007) IVIC syndrome is caused by a c.2607delA mutation in the SALL4 locus. *Am J Med Genet A*, 143(4), 326-32.
Park, H., Seo, Y., Kim, J. I., Kim, W.-j. & Choe, S. Y. (2006) Identification of the nuclear localization motif in the ETV6 (TEL) protein. *Cancer Genetics and Cytogenetics*, 167(2), 117-121.

Park, S. M., Jung, H. Y., Kim, H. O., Rhim, H., Paik, S. R., Chung, K. C., Park, J. H. & Kim, J. (2002) Evidence that α-synuclein functions as a negative regulator of Ca++-dependent α-granule release from human platelets. *Blood*, 100(7), 2506-2514.
Pastural, E., Barrat, F. J., Dufourcq-Lagelouse, R., Certain, S., Sanal, O., Jabado, N., Seger, R., Griscelli, C., Fischer, A. & de Saint Basile, G. (1997) Griscelli disease maps to chromosome 15q21 and is associated with mutations in the myosin-Va gene. *Nature genetics*, 16(3), 289-292.
Patel, Y. M., Lordkipanidze, M., Lowe, G. C., Nisar, S. P., Garner, K., Stockley, J., Daly, M. E., Mitchell, M., Watson, S. P., Austin, S. K. & Mundell, S. J. (2014) A novel mutation in the P2Y12 receptor and a function-reducing polymorphism in protease-activated receptor 1 in a patient with chronic bleeding. *J Thromb Haemost*, 12(5), 716-25.
Paterson, A. D., Rommens, J. M., Bharaj, B., Blavignac, J., Wong, I., Diamandis, M., Waye, J. S., Rivard, G. E. & Hayward, C. P. (2010) Persons with Quebec platelet disorder have a tandem duplication of PLAU, the urokinase plasminogen activator gene. *Blood*, 115(6), 1264-6.
Paulo, P., Barros-Silva, J. D., Ribeiro, F. R., Ramalho-Carvalho, J., Jerónimo, C., Henrique, R., Lind, G. E., Skotheim, R. I., Lothe, R. A. & Teixeira, M. R. (2012) FLI1 is a novel ETS transcription factor involved in gene fusions in prostate cancer. *Genes, Chromosomes and Cancer*, 51(3), 240-249.
Pecci, A. & Balduini, C. L. (2014) Lessons in platelet production from inherited thrombocytopenias. *Br J Haematol*, 165(2), 179-92.
Pecci, A., Panza, E., Pujol-Moix, N., Klersy, C., Di Bari, F., Bozzi, V., Gresele, P., Lethagen, S., Fabris, F. & Dufour, C. (2008) Position of nonmuscle myosin heavy chain IIA (NMMHC-IIA) mutations predicts the natural history of MYH9-related disease. *Human mutation*, 29(3), 409-417.
Penny, L. A., Dell'Aquila, M., Jones, M. C., Bergoffen, J., Cunniff, C., Fryns, J.-P., Grace, E., Graham, J. M., Kousseff, B. & Mattina, T. (1995) Clinical and molecular characterization of patients with distal 11q deletions. *American journal of human genetics*, 56(3), 676.
Peters, C. G., Michelson, A. D. & Flaumenhaft, R. (2012) Granule exocytosis is required for platelet spreading: differential sorting of α-granules expressing VAMP-7. *Blood*, blood-2011-10-389247.
Pimanda, J. E., Chan, W. I., Donaldson, I. J., Bowen, M., Green, A. R. & Göttgens, B. (2006) Endoglin expression in the endothelium is regulated by Fli-1, Erg, and Elf-1 acting on the promoter and a–8-kb enhancer. *Blood*, 107(12), 4737-4745.
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. *Am J Hum Genet*, 88(1), 115-20.
Pisareva, V. P., Muslimov, I. A., Tcherepanov, A. & Pisarev, A. V. (2015) Characterization of novel ribosome-associated endoribonuclease SLFN14 from rabbit reticulocytes. *Biochemistry*, 54(21), 3286-3301.
Plotkin, J. B. & Kudla, G. (2011) Synonymous but not the same: the causes and consequences of codon bias. *Nat Rev Genet*, 12(1), 32-42.
Poggi, M., Canault, M., Favier, M., Turro, E., Saultier, P., Ghalloussi, D., Baccini, V., Vidal, L., Mezzapesa, A., Chelghoum, N., Mohand-Oumoussa, B., Falaise, C., Favier, R., Ouwehand, W. H., Fiore, M., Peiretti, F., Morange, P. E., Saut, N., Bernot, D., Greinacher, A., BioResource, N., Nurden, A. T., Nurden, P., Freson, K., Tregouet, D. A., Raslova, H. & Alessi, M. C. (2017) Germline variants in ETV6 underlie reduced platelet formation, platelet dysfunction and increased levels of circulating CD34+ progenitors. *Haematologica*, 102(2), 282-294.
Poggi, M., Canault, M., Lucca, P., Pouymayou, C., Ghalloussi, D., Saut, N., Morange, P., Tregouet, D. & Alessi, M. (2015) Rock1 OR FLI1 mutations in patients with mild thrombocytopenia. *Journal of Thrombosis and Haemostasis*, 13, 421.
Poirel, H., Lopez, R. G., Lacronique, V., Valle, V. D., Mauchauffe, M., Berger, R., Ghysdael, J. & Bernard, O. A. (2000) Characterization of a novel ETS gene, TELB, encoding a protein structurally and functionally related to TEL. *Oncogene*, 19(41), 4802-4806.
Poirel, H., Oury, C., Carron, C., Duprez, E., Laabi, Y., Tsapis, A., Romana, S. P., Mauchauffe, M., Le Coniat, M. & Berger, R. (1997) The TEL gene products: nuclear phosphoproteins with DNA binding properties. *Oncogene*, 14(3).
Polfus, L. M., Khajuria, R. K., Schick, U. M., Pankratz, N., Pazoki, R., Brody, J. A., Chen, M.-H., Auer, P. L., Floyd, J. S. & Huang, J. (2016) Whole-exome sequencing identifies loci associated with blood cell traits and reveals a role for alternative GFI1B splice variants in human hematopoiesis. *The American Journal of Human Genetics*, 99(2), 481-488.
Potter, M. D., Buijs, A., Kreider, B., van Rompaey, L. & Grosveld, G. C. (2000) Identification and characterization of a new human ETS-family transcription factor, TEL2, that is expressed in hematopoietic tissues and can associate with TEL1/ETV6. *Blood*, 95(11), 3341-3348.
Poulter, N. S. & Thomas, S. G. (2015) Cytoskeletal regulation of platelet formation: Coordination of F-actin and microtubules. *The international journal of biochemistry & cell biology*, 66, 69-74.
Punzo, F., Mientjes, E. J., Rohe, C. F., Scianguetta, S., Amendola, G., Oostra, B. A., Bertoli-Avella, A. M. & Perrotta, S. (2010) A mutation in the acyl-coenzyme A binding domain-containing protein 5 gene (ACBD5) identified in autosomal dominant thrombocytopenia. *J Thromb Haemost*, 8(9), 2085-7.
Ran, F. A., Hsu, P. D., Wright, J., Agarwala, V., Scott, D. A. & Zhang, F. (2013) Genome engineering using the CRISPR-Cas9 system. *Nature protocols*, 8(11), 2281.
Rao, V. N., Ohno, T., Prasad, D., Bhattacharya, G. & Reddy, E. (1993) Analysis of the DNA-binding and transcriptional activation functions of human Fli-1 protein. *Oncogene*, 8(8), 2167-2173.
Rasighaemi, P., Onnebo, S. M., Liongue, C. & Ward, A. C. (2015) ETV6 (TEL1) regulates embryonic hematopoiesis in zebrafish. *Haematologica*, 100(1), 23-31.
Raslova, H., Komura, E., Le Couedic, J. P., Larbret, F., Debili, N., Feunteun, J., Danos, O., Albagli, O., Vainchenker, W. & Favier, R. (2004) FLI1 monoallelic expression combined with its hemizygous loss underlies Paris-Trousseau/Jacobsen thrombopenia. *J Clin Invest*, 114(1), 77-84.
Rees, D. C., Iolascon, A., Carella, M., O'Marcaigh A, S., Kendra, J. R., Jowitt, S. N., Wales, J. K., Vora, A., Makris, M., Manning, N., Nicolaou, A., Fisher, J., Mann, A., Machin, S. J., Clayton, P. T., Gasparini, P. & Stewart, G. W. (2005) Stomatocytic haemolysis and macrothrombocytopenia (Mediterranean stomatocytosis/macrothrombocytopenia) is the haematological presentation of phytosterolaemia. *Br J Haematol*, 130(2), 297-309.
Remy, P., Senan, F., Meyer, D., Mager, A. & Hindelang, C. (2002) Overexpression of the Xenopus Xl-fli gene during early embryogenesis leads to anomalies in head and heart development and erythroid differentiation. *International Journal of Developmental Biology*, 40(3), 577-589.
Rhee, I., Zhong, M.-C., Reizis, B., Cheong, C. & Veillette, A. (2014) Control of dendritic cell migration, T cell-dependent immunity, and autoimmunity by protein tyrosine phosphatase PTPN12 expressed in dendritic cells. *Molecular and cellular biology*, 34(5), 888-899.

Ricotta, D., Conner, S. D., Schmid, S. L., von Figura, K. & Höning, S. (2002) Phosphorylation of the AP2 μ subunit by AAK1 mediates high affinity binding to membrane protein sorting signals. *The Journal of cell biology*, 156(5), 791-795.
Rima, M., Daghsni, M., Fajloun, Z., M'rad, R., Brusés, J. L., Ronjat, M. & De Waard, M. (2016) Protein partners of the calcium channel β subunit highlight new cellular functions. *Biochemical Journal*, 473(13), 1831-1844.
Roberts, D. E., Matsuda, T. & Bose, R. (2012) Molecular and functional characterization of the human platelet Na+/Ca2+ exchangers. *British journal of pharmacology*, 165(4), 922-936.
Rondaij, M. G., Bierings, R., van Agtmaal, E. L., Gijzen, K. A., Sellink, E., Kragt, A., Ferguson, S. S., Mertens, K., Hannah, M. J. & van Mourik, J. A. (2008) Guanine exchange factor RalGDS mediates exocytosis of Weibel-Palade bodies from endothelial cells. *Blood*, 112(1), 56-63.
Roukens, M. G., Alloul-Ramdhani, M., Baan, B., Kobayashi, K., Peterson-Maduro, J., Van Dam, H., Schulte-Merker, S. & Baker, D. A. (2010) Control of endothelial sprouting by a Tel-CtBP complex. *Nature cell biology*, 12(10), 933-942.
Rubinstein, J. D., Elagib, K. E. & Goldfarb, A. N. (2012) Cyclic AMP signaling inhibits megakaryocytic differentiation by targeting a transcription factor 3 (E2A)-cyclic dependent kinase inhibitor 1A (CDKN1A) transcriptional axis. *Journal of Biological Chemistry*, jbc. M112. 366476.
Saleque, S., Cameron, S. & Orkin, S. H. (2002) The zinc-finger proto-oncogene Gfi-1b is essential for development of the erythroid and megakaryocytic lineages. *Genes & development*, 16(3), 301-306.
Sanjuan-Pla, A., Macaulay, I. C., Jensen, C. T., Woll, P. S., Luis, T. C., Mead, A., Moore, S., Carella, C., Matsuoka, S. & Jones, T. B. (2013) Platelet-biased stem cells reside at the apex of the haematopoietic stem-cell hierarchy. *Nature*, 502(7470), 232.
Saultier, P., Vidal, L., Canault, M., Bernot, D., Falaise, C., Pouymayou, C., Bordet, J. C., Saut, N., Rostan, A., Baccini, V., Peiretti, F., Favier, M., Lucca, P., Deleuze, J. F., Olaso, R., Boland, A., Morange, P. E., Gachet, C., Malergue, F., Faure, S., Eckly, A., Tregouet, D. A., Poggi, M. & Alessi, M. C. (2017) Macrothrombocytopenia and dense granule deficiency associated with FLI1 variants: ultrastructural and pathogenic features. *Haematologica*, 102(6), 1006-1016.
Savoia, A. (2016) Molecular basis of inherited thrombocytopenias. *Clin Genet*, 89(2), 154-62.
Savoia, A., Balduini, C. L., Savino, M., Noris, P., Del Vecchio, M., Perrotta, S., Belletti, S., Poggi, V. & Iolascon, A. (2001) Autosomal dominant macrothrombocytopenia in Italy is most frequently a type of heterozygous Bernard-Soulier syndrome. *Blood*, 97(5), 1330-1335.
Savoia, A., Dufour, C., Locatelli, F., Noris, P., Ambaglio, C., Rosti, V., Zecca, M., Ferrari, S., di Bari, F., Corcione, A., Di Stazio, M., Seri, M. & Balduini, C. L. (2007) Congenital amegakaryocytic thrombocytopenia: clinical and biological consequences of five novel mutations. *Haematologica*, 92(9), 1186-93.
Scheiber, M. N., Watson, P. M., Rumboldt, T., Stanley, C., Wilson, R. C., Findlay, V. J., Anderson, P. E. & Watson, D. K. (2014) FLI1 expression is correlated with breast cancer cellular growth, migration, and invasion and altered gene expression. *Neoplasia*, 16(10), 801-813.
Schnetkamp, P. P. (2013) The SLC24 gene family of Na+/Ca2+–K+ exchangers: From sight and smell to memory consolidation and skin pigmentation. *Molecular aspects of medicine*, 34(2-3), 455-464.
Schulze, H., Schlagenhauf, A., Manukjan, G., Beham-Schmid, C., Andres, O., Klopocki, E., König, E.-M., Haidl, H., Panzer, S. & Althaus, K. (2016) An autosomal-recessive GFI1B mutation defines the splice isoform p37 as essential for biogenesis of functional human platelets, but dispensable for erythropoiesis. Am Soc Hematology.
Schulze, H., Schlagenhauf, A., Manukjan, G., Beham-Schmid, C., Andres, O., Klopocki, E., König, E.-M., Haidl, H., Panzer, S. & Althaus, K. (2017) Recessive grey platelet-like syndrome with unaffected erythropoiesis in the absence of the splice isoform GFI1B-p37. *haematologica*, 102(9), e375-e378.
Schwachtgen, J.-L., Janel, N., Barek, L., Duterque-Coquillaud, M., Ghysdael, J., Meyer, D. & Kerbiriou-Nabias, D. (1997) Ets transcription factors bind and transactivate the core promoter of the von Willebrand factor gene. *Oncogene*, 15(25), 3091.
Seri, M., Cusano, R., Gangarossa, S., Caridi, G., Bordo, D., Lo, N. C., Ghiggeri, G. M., Ravazzolo, R., Savino, M. & Del Vecchio, M. (2000) Mutations in MYH9 result in the May-Hegglin anomaly, and Fechtner and Sebastian syndromes. The May-Hegglin/Fechtner Syndrome Consortium. *Nature genetics*, 26(1), 103-105.
Seurin, D., Lombet, A., Babajko, S., Godeau, F. & Ricort, J.-M. (2013) Insulin-like growth factor binding proteins increase intracellular calcium levels in two different cell lines. *PloS one*, 8(3), e59323.
Sharda, A. & Flaumenhaft, R. (2018) The life cycle of platelet granules. *F1000Research*, 7.
Shattil, S. J., Kim, C. & Ginsberg, M. H. (2010) The final steps of integrin activation: the end game. *Nat Rev Mol Cell Biol*, 11(4), 288-300.
Shcherbina, A., Miki, H., Kenney, D. M., Rosen, F. S., Takenawa, T. & Remold-O'Donnell, E. (2001) WASP and N-WASP in human platelets differ in sensitivity to protease calpain. *Blood*, 98(10), 2988-2991.
Shin, D. M., Dehoff, M., Luo, X., Kang, S. H., Tu, J., Nayak, S. K., Ross, E. M., Worley, P. F. & Muallem, S. (2003) Homer 2 tunes G protein–coupled receptors stimulus intensity by regulating RGS proteins and PLCβ GAP activities. *J Cell Biol*, 162(2), 293-303.
Shin, E.-K., Park, H., Noh, J.-Y., Lim, K.-M. & Chung, J.-H. (2017) Platelet Shape Changes and Cytoskeleton Dynamics as Novel Therapeutic Targets for Anti-Thrombotic Drugs. *Biomolecules & therapeutics*, 25(3), 223.
Shiraishi-Yamaguchi, Y., Sato, Y., Sakai, R., Mizutani, A., Knöpfel, T., Mori, N., Mikoshiba, K. & Furuichi, T. (2009) Interaction of Cupidin/Homer2 with two actin cytoskeletal regulators, Cdc42 small GTPase and Drebrin, in dendritic spines. *BMC neuroscience*, 10(1), 25.
Shirakawa, R., Yoshioka, A., Horiuchi, H., Nishioka, H., Tabuchi, A. & Kita, T. (2000) Small GTPase Rab4 regulates Ca2+-induced α-granule secretion in platelets. *Journal of Biological Chemistry*, 275(43), 33844-33849.
Shotelersuk, V., Dell'Angelica, E. C., Hartnell, L., Bonifacino, J. S. & Gahl, W. A. (2000) A new variant of Hermansky-Pudlak syndrome due to mutations in a gene responsible for vesicle formation. *The American journal of medicine*, 108(5), 423-427.
Silver, J. & Kozak, C. (1986) Common proviral integration region on mouse chromosome 7 in lymphomas and myelogenous leukemias induced by Friend murine leukemia virus. *Journal of virology*, 57(2), 526-533.
Siripin, D., Kheolamai, P., Yaowalak, U., Supokawej, A., Wattanapanitch, M., Klincumhom, N., Laowtammathron, C. & Issaragrisil, S. (2015) Transdifferentiation of erythroblasts to megakaryocytes using FLI1 and ERG transcription factors. *Thrombosis and haemostasis*, 114(03), 593-602.
Sivapalaratnam, S., Collins, J. & Gomez, K. (2017) Diagnosis of inherited bleeding disorders in the genomic era. *British journal of haematology*, 179(3), 363-376.
Smolenski, A. (2012) Novel roles of cAMP/cGMP-dependent signaling in platelets. *Journal of Thrombosis and Haemostasis*, 10(2), 167-176.
Södersten, E., Lilja, T. & Hermanson, O. (2010) The novel BTB/POZ and zinc finger factor Zbtb45 is essential for proper glial differentiation of neural and oligodendrocyte progenitor cells. *Cell Cycle*, 9(24), 4866-4875.
Song, W.-J., Sullivan, M. G., Legare, R. D., Hutchings, S., Tan, X., Kufrin, D., Ratajczak, J., Resende, I. C., Haworth, C. & Hock, R. (1999) Haploinsufficiency of CBFA2 causes familial thrombocytopenia with propensity to develop acute myelogenous leukaemia. *Nature genetics*, 23(2), 166-175.

Song, W., Zhang, T., Li, W., Mu, R., Zhang, L., Li, Y., Jin, B., Wang, N., Li, A. & Cui, J. (2015) Overexpression of Fli-1 is associated with adverse prognosis of endometrial cancer. *Cancer investigation*, 33(9), 469-475.
Songdej, N., Mao, G., Voora, D., Goldfinger, L. E., Del Carpio-Cano, F., Myers, R. & Rao, A. K. (2016) PCTP (phosphatidylcholine transfer protein) is regulated by RUNX1 in platelets/megakaryocytes and is associated with adverse cardiovascular events. Am Soc Hematology.
Souroullas, G. P., Salmon, J. M., Sablitzky, F., Curtis, D. J. & Goodell, M. A. (2009) Adult hematopoietic stem and progenitor cells require either Lyl1 or Scl for survival. *Cell stem cell*, 4(2), 180-186.
Souza, C. M., Davidson, D., Rhee, I., Gratton, J.-P., Davis, E. C. & Veillette, A. (2012) The phosphatase PTP-PEST/PTPN12 regulates endothelial cell migration and adhesion, but not permeability, and controls vascular development and embryonic viability. *Journal of Biological Chemistry*, jbc. M112. 387456.
Spessott, W. A., Sanmillan, M. L., McCormick, M. E., Patel, N., Villanueva, J., Zhang, K., Nichols, K. E. & Giraudo, C. G. (2015) Hemophagocytic lymphohistiocytosis caused by dominant negative mutations in STXBP2 that inhibit SNARE-mediated membrane fusion. *Blood*, blood-2014-11-610816.
Spyropoulos, D. D., Pharr, P. N., Lavenburg, K. R., Jackers, P., Papas, T. S., Ogawa, M. & Watson, D. K. (2000) Hemorrhage, Impaired Hematopoiesis, and Lethality in Mouse Embryos Carrying a Targeted Disruption of the Fli1Transcription Factor. *Molecular and cellular biology*, 20(15), 5643-5652.
Standing, A. S., Malinova, D., Hong, Y., Record, J., Moulding, D., Blundell, M. P., Nowak, K., Jones, H., Omoyinmi, E. & Gilmour, K. C. (2017) Autoinflammatory periodic fever, immunodeficiency, and thrombocytopenia (PFIT) caused by mutation in actin-regulatory gene WDR1. *Journal of Experimental Medicine*, 214(1), 59-71.
Starai, V. J., Hickey, C. M. & Wickner, W. (2008) HOPS proofreads the trans-SNARE complex for yeast vacuole fusion. *Molecular biology of the cell*, 19(6), 2500-2508.
Starck, J., Cohet, N., Gonnet, C., Sarrazin, S., Doubeikovskaia, Z., Doubeikovski, A., Verger, A., Duterque-Coquillaud, M. & Morle, F. (2003) Functional cross-antagonism between transcription factors FLI-1 and EKLF. *Molecular and Cellular Biology*, 23(4), 1390-1402.
Starck, J., Weiss-Gayet, M., Gonnet, C., Guyot, B., Vicat, J.-M. & Morlé, F. (2010) Inducible Fli-1 gene deletion in adult mice modifies several myeloid lineage commitment decisions and accelerates proliferation arrest and terminal erythrocytic differentiation. *Blood*, 116(23), 4795-4805.
Stark, Z., Tan, T. Y., Chong, B., Brett, G. R., Yap, P., Walsh, M., Yeung, A., Peters, H., Mordaunt, D. & Cowie, S. (2016) A prospective evaluation of whole-exome sequencing as a first-tier molecular test in infants with suspected monogenic disorders. *Genetics in Medicine*, 18(11), 1090.
Stemmer, M., Thumberger, T., del Sol Keyer, M., Wittbrodt, J. & Mateo, J. L. (2017) Correction: CCTop: An Intuitive, Flexible and Reliable CRISPR/Cas9 Target Prediction Tool. *PloS one*, 12(4), e0176619.
Stepensky, P., Saada, A., Cowan, M., Tabib, A., Fischer, U., Berkun, Y., Saleh, H., Simanovsky, N., Kogot-Levin, A. & Weintraub, M. (2013) The Thr224Asn mutation in the VPS45 gene is associated with congenital neutropenia and primary myelofibrosis of infancy. *Blood*, blood-2012-12-475566.
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. *J Thromb Haemost*, 11(11), 2039-47.
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(17), 2027-30.
Stockley, J., Morgan, N. V., Bem, D., Lowe, G. C., Lordkipanidze, M., Dawood, B., Simpson, M. A., Macfarlane, K., Horner, K., Leo, V. C., Talks, K., Motwani, J., Wilde, J. T., Collins, P. W., Makris, M., Watson, S. P., Daly, M. E., Genotyping, U. K. & Phenotyping of Platelets Study, G. (2013) Enrichment of FLI1 and RUNX1 mutations in families with excessive bleeding and platelet dense granule secretion defects. *Blood*, 122(25), 4090-3.
Stritt, S., Nurden, P., Turro, E., Greene, D., Jansen, S. B., Westbury, S. K., Petersen, R., Astle, W. J., Marlin, S., Bariana, T. K., Kostadima, M., Lentaigne, C., Maiwald, S., Papadia, S., Kelly, A. M., Stephens, J. C., Penkett, C. J., Ashford, S., Tuna, S., Austin, S., Bakchoul, T., Collins, P., Favier, R., Lambert, M. P., Mathias, M., Millar, C. M., Mapeta, R., Perry, D. J., Schulman, S., Simeoni, I., Thys, C., Consortium, B.-B., Gomez, K., Erber, W. N., Stirrups, K., Rendon, A., Bradley, J. R., van Geet, C., Raymond, F. L., Laffan, M. A., Nurden, A. T., Nieswandt, B., Richardson, S., Freson, K., Ouwehand, W. H. & Mumford, A. D. (2016) A gain-of-function variant in DIAPH1 causes dominant macrothrombocytopenia and hearing loss. *Blood*, 127(23), 2903-14.
Su, Y., Wang, Z., Yang, H., Cao, L., Liu, F., Bai, X. & Ruan, C. (2006) Clinical and molecular genetic analysis of a family with sitosterolemia and co-existing erythrocyte and platelet abnormalities. *Haematologica*, 91(10), 1392-1395.
Sun, L., Gorospe, J., Hoffman, E. & Rao, A. (2007) Decreased platelet expression of myosin regulatory light chain polypeptide (MYL9) and other genes with platelet dysfunction and CBFA2/RUNX1 mutation: insights from platelet expression profiling. *Journal of Thrombosis and Haemostasis*, 5(1), 146-154.
Suzuki-Inoue, K., Inoue, O., Frampton, J. & Watson, S. P. (2003) Murine GPVI stimulates weak integrin activation in PLCγ2–/– platelets: involvement of PLCγ1 and PI3-kinase. *Blood*, 102(4), 1367-1373.
Suzuki, E., Williams, S., Sato, S., Gilkeson, G., Watson, D. K. & Zhang, X. K. (2013) The transcription factor Fli-1 regulates monocyte, macrophage and dendritic cell development in mice. *Immunology*, 139(3), 318-327.
Suzuki, J., Umeda, M., Sims, P. J. & Nagata, S. (2010) Calcium-dependent phospholipid scrambling by TMEM16F. *Nature*, 468(7325), 834-8.
Suzuki, R., Fujikawa, A., Komatsu, Y., Kuboyama, K., Tanga, N. & Noda, M. (2018) Enhanced extinction of aversive memories in mice lacking SPARC-related protein containing immunoglobulin domains 1 (SPIG1/FSTL4). *Neurobiology of learning and memory*, 152, 61-70.
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., Fong, C. T., Castellan, C., Miterski, B., Swank, R. T. & Spritz, R. A. (2002) Hermansky-Pudlak syndrome is caused by mutations in HPS4, the human homolog of the mouse light-ear gene. *Nat Genet*, 30(3), 321-4.
Suzuki, T., Oiso, N., Gautam, R., Novak, E. K., Panthier, J.-J., Suprabha, P., Vida, T., Swank, R. T. & Spritz, R. A. (2003) The mouse organellar biogenesis mutant buff results from a mutation in Vps33a, a homologue of yeast vps33 and Drosophila carnation. *Proceedings of the National Academy of Sciences*, 100(3), 1146-1150.
Szymczyna, B. R. & Arrowsmith, C. H. (2000) DNA-binding specificity studies of 4 ETS proteins supports an" indirect read-out" mechanism of protein-DNA recognition. *Journal of Biological Chemistry*.
Tadokoro, S., Shattil, S. J., Eto, K., Tai, V., Liddington, R. C., de Pereda, J. M., Ginsberg, M. H. & Calderwood, D. A. (2003) Talin binding to integrin ß tails: a final common step in integrin activation. *Science*, 302(5642), 103-106.

Takahashi, T., Asano, Y., Sugawara, K., Yamashita, T., Nakamura, K., Saigusa, R., Ichimura, Y., Toyama, T., Taniguchi, T. & Akamata, K. (2017) Epithelial Fli1 deficiency drives systemic autoimmunity and fibrosis: Possible roles in scleroderma. *Journal of Experimental Medicine*, 214(4), 1129-1151.
Takahashi, W., Sasaki, K., Kvomatsu, N. & Mitani, K. (2005) TEL/ETV6 accelerates erythroid differentiation and inhibits megakaryocytic maturation in a human leukemia cell line UT-7/GM. *Cancer Sci*, 96(6), 340-8.
Takeichi, T., Torrelo, A., Lee, J. Y., Ohno, Y., Lozano, M. L., Kihara, A., Liu, L., Yasuda, Y., Ishikawa, J. & Murase, T. (2017) Biallelic mutations in KDSR disrupt ceramide synthesis and result in a spectrum of keratinization disorders associated with thrombocytopenia. *Journal of Investigative Dermatology*, 137(11), 2344-2353.
Takenouchi, T., Kosaki, R., Niizuma, T., Hata, K. & Kosaki, K. (2015) Macrothrombocytopenia and developmental delay with a de novo CDC42 mutation: Yet another locus for thrombocytopenia and developmental delay. *American Journal of Medical Genetics Part A*, 167(11), 2822-2825.
Takenouchi, T., Okamoto, N., Ida, S., Uehara, T. & Kosaki, K. (2016) Further evidence of a mutation in CDC42 as a cause of a recognizable syndromic form of thrombocytopenia. *American Journal of Medical Genetics Part A*, 170(4), 852-855.
Tamir, A., Howard, J., Higgins, R. R., Li, Y.-J., Berger, L., Zacksenhaus, E., Reis, M. & Ben-David, Y. (1999) Fli-1, an Ets-related transcription factor, regulates erythropoietin-induced erythroid proliferation and differentiation: evidence for direct transcriptional repression of the Rb gene during differentiation. *Molecular and Cellular Biology*, 19(6), 4452-4464.
Tan, X., Sun, Y., Thapa, N., Liao, Y., Hedman, A. C. & Anderson, R. A. (2015) LAPTM4B is a PtdIns (4, 5) P2 effector that regulates EGFR signaling, lysosomal sorting, and degradation. *The EMBO journal*, e201489425.
Tang, S., Tao, M., McCoy Jr, J. P. & Zheng, Z.-M. (2006) Short-term induction and long-term suppression of HPV16 oncogene silencing by RNA interference in cervical cancer cells. *Oncogene*, 25(14), 2094.
Terry, S. J., Zihni, C., Elbediwy, A., Vitiello, E., San, I. V. L. C., Balda, M. S. & Matter, K. (2011) Spatially restricted activation of RhoA signalling at epithelial junctions by p114RhoGEF drives junction formation and morphogenesis. *Nature cell biology*, 13(2), 159.
Thomas, S., Calaminus, S., Machesky, L., Alberts, A. & Watson, S. (2011) G-protein coupled and ITAM receptor regulation of the formin FHOD1 through Rho Kinase in platelets. *Journal of Thrombosis and Haemostasis*, 9(8), 1648-1651.
Thompson, A. A. & Nguyen, L. T. (2000) Amegakaryocytic thrombocytopenia and radio-ulnar synostosis are associated with HOXA11 mutation. *Nature genetics*, 26(4), 397-398.
Thon, J. N. & Italiano, J. E. (2012) Platelets: production, morphology and ultrastructure, *Antiplatelet Agents*Springer, 3-22.
Tijssen, M. R., Cvejic, A., Joshi, A., Hannah, R. L., Ferreira, R., Forrai, A., Bellissimo, D. C., Oram, S. H., Smethurst, P. A., Wilson, N. K., Wang, X., Ottersbach, K., Stemple, D. L., Green, A. R., Ouwehand, W. H. & Gottgens, B. (2011) Genome-wide analysis of simultaneous GATA1/2, RUNX1, FLI1, and SCL binding in megakaryocytes identifies hematopoietic regulators. *Dev Cell*, 20(5), 597-609.
Tijssen, M. R. & Ghevaert, C. (2013) Transcription factors in late megakaryopoiesis and related platelet disorders. *J Thromb Haemost*, 11(4), 593-604.
Tognon, C., Knezevich, S. R., Huntsman, D., Roskelley, C. D., Melnyk, N., Mathers, J. A., Becker, L., Carneiro, F., MacPherson, N. & Horsman, D. (2002) Expression of the ETV6-NTRK3 gene fusion as a primary event in human secretory breast carcinoma. *Cancer cell*, 2(5), 367-376.
Tolmachova, T., Åbrink, M., Futter, C. E., Authi, K. S. & Seabra, M. C. (2007) Rab27b regulates number and secretion of platelet dense granules. *Proceedings of the National Academy of Sciences*, 104(14), 5872-5877.
Topka, S., Vijai, J., Walsh, M. F., Jacobs, L., Maria, A., Villano, D., Gaddam, P., Wu, G., McGee, R. B., Quinn, E., Inaba, H., Hartford, C., Pui, C. H., Pappo, A., Edmonson, M., Zhang, M. Y., Stepensky, P., Steinherz, P., Schrader, K., Lincoln, A., Bussel, J., Lipkin, S. M., Goldgur, Y., Harit, M., Stadler, Z. K., Mullighan, C., Weintraub, M., Shimamura, A., Zhang, J., Downing, J. R., Nichols, K. E. & Offit, K. (2015) Germline ETV6 Mutations Confer Susceptibility to Acute Lymphoblastic Leukemia and Thrombocytopenia. *PLoS Genet*, 11(6), e1005262.
Torlakovic, E. E., Slipicevic, A., Florenes, V. A., Chibbar, R., DeCoteau, J. F. & Bilalovic, N. (2008) Fli-1 expression in malignant melanoma. *Histology and histopathology*, 23(10), 1309.
Tracy, P. B., Giles, A. R., Mann, K. G., Eide, L. L., Hoogendoorn, H. & Rivard, G. E. (1984) Factor V (Quebec): a bleeding diathesis associated with a qualitative platelet Factor V deficiency. *The Journal of clinical investigation*, 74(4), 1221-1228.
Troxler, D. H. & Scolnick, E. M. (1978) Rapid leukemia induced by cloned Friend strain of replicating murine type-C virus Association with induction of xenotropic-related RNA sequences contained in spleen focus-forming virus. *Virology*, 85(1), 17-27.
Turro, E., Greene, D., Wijgaerts, A., Thys, C., Lentaigne, C., Bariana, T. K., Westbury, S. K., Kelly, A. M., Selleslag, D. & Stephens, J. C. (2016) A dominant gain-of-function mutation in universal tyrosine kinase SRC causes thrombocytopenia, myelofibrosis, bleeding, and bone pathologies. *Science translational medicine*, 8(328), 328ra30-328ra30.
Urban, A. E., Quick, E. O., Miller, K. P., Krcmery, J. & Simon, H.-G. (2016) Pdlim7 regulates Arf6-dependent actin dynamics and is required for platelet-mediated thrombosis in mice. *PloS one*, 11(10), e0164042.
Valet, C., Levade, M., Chicanne, G., Bilanges, B., Cabou, C., Viaud, J., Gratacap, M.-P., Gaits-Iacovoni, F., Vanhaesebroeck, B. & Payrastre, B. (2017) A dual role for the class III PI3K, Vps34, in platelet production and thrombus growth. *Blood*, blood-2017-04-781641.
van Holten, T. C., Bleijerveld, O. B., Wijten, P., de Groot, P. G., Heck, A. J., Barendrecht, A. D., Merkx, T. H., Scholten, A. & Roest, M. (2014) Quantitative proteomics analysis reveals similar release profiles following specific PAR-1 or PAR-4 stimulation of platelets. *Cardiovascular research*, 103(1), 140-146.
Van Rompaey, L., Dou, W., Buijs, A. & Grosveld, G. (1999) Tel, a frequent target of leukemic translocations, induces cellular aggregation and influences expression of extracellular matrix components. *Neoplasia*, 1(6), 526-536.
Van Rompaey, L., Potter, M., Adams, C. & Grosveld, G. (2000) Tel induces a G1 arrest and suppresses Ras-induced transformation. *Oncogene*, 19(46), 5244.
Varga-Szabo, D., Braun, A. & Nieswandt, B. (2009) Calcium signaling in platelets. *J Thromb Haemost*, 7(7), 1057-66.
Varga-Szabo, D., Pleines, I. & Nieswandt, B. (2008) Cell adhesion mechanisms in platelets. *Arterioscler Thromb Vasc Biol*, 28(3), 403-12.
Vergeer, M., Korporaal, S. J., Franssen, R., Meurs, I., Out, R., Hovingh, G. K., Hoekstra, M., Sierts, J. A., Dallinga-Thie, G. M. & Motazacker, M. M. (2011) Genetic variant of the scavenger receptor BI in humans. *New England Journal of Medicine*, 364(2), 136-145.
Vermeersch, E., Nuyttens, B. P., Tersteeg, C., Broos, K., De Meyer, S. F., Vanhoorelbeke, K. & Deckmyn, H. (2018) Functional Genomics for the Identification of Modulators of Platelet-Dependent Thrombus Formation. *TH Open*, 2(03), e272-e279.
Vidal, C., Geny, B., Melle, J., Jandrot-Perrus, M. & Fontenay-Roupie, M. (2002) Cdc42/Rac1-dependent activation of the p21-activated kinase (PAK) regulates human platelet lamellipodia spreading: implication of the cortical-actin binding protein cortactin. *Blood*, 100(13), 4462-4469.

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*, 129(26), 3486-3494.
Voora, D., Cyr, D., Lucas, J., Chi, J.-T., Dungan, J., McCaffrey, T. A., Katz, R., Newby, L. K., Kraus, W. E. & Becker, R. C. (2013) Aspirin exposure reveals novel genes associated with platelet function and cardiovascular events. *Journal of the American College of Cardiology*, 62(14), 1267-1276.
Wagner, C. L., Mascelli, M., Neblock, D., Weisman, H., Coller, B. & Jordan, R. (1996) Analysis of GPIIb/IIIa receptor number by quantification of 7E3 binding to human platelets. *Blood*, 88(3), 907-914.
Wan, L., Wu, H., Xie, F. & Nie, Y. (2017) Thrombocytopenia and GBA gene mutation in a patient with adult type 1 Gaucher disease. *Platelets*, 28(8), 829-831.
Wang, D., Feng, J., Wen, R., Marine, J.-C., Sangster, M. Y., Parganas, E., Hoffmeyer, A., Jackson, C. W., Cleveland, J. L. & Murray, P. J. (2000) Phospholipase Cγ2 is essential in the functions of B cell and several Fc receptors. *Immunity*, 13(1), 25-35.
Wang, L. & Hiebert, S. W. (2001) TEL contacts multiple co-repressors and specifically associates with histone deacetylase-3. *Oncogene*, 20(28).
Wang, L. C., Kuo, F., Fujiwara, Y., Gilliland, D. G., Golub, T. R. & Orkin, S. H. (1997) Yolk sac angiogenic defect and intra-embryonic apoptosis in mice lacking the Ets-related factor TEL. *The EMBO journal*, 16(14), 4374-4383.
Wang, L. C., Swat, W., Fujiwara, Y., Davidson, L., Visvader, J., Kuo, F., Alt, F. W., Gilliland, D. G., Golub, T. R. & Orkin, S. H. (1998) The TEL/ETV6 gene is required specifically for hematopoiesis in the bone marrow. *Genes & development*, 12(15), 2392-2402.
Wang, M., Herrmann, C. J., Simonovic, M., Szklarczyk, D. & von Mering, C. (2015) Version 4.0 of PaxDb: Protein abundance data, integrated across model organisms, tissues, and cell-lines. *Proteomics*, 15(18), 3163-8.
Wang, Y., Meng, R., Hayes, V., Fuentes, R., Yu, X., Abrams, C. S., Heijnen, H. F., Blobel, G. A., Marks, M. S. & Poncz, M. (2011) Pleiotropic platelet defects in mice with disrupted FOG1-NuRD interaction. *Blood*, blood-2011-06-363580.
Ware, J., Corken, A. & Khetpal, R. (2013) Platelet function beyond hemostasis and thrombosis. *Current opinion in hematology*, 20(5).
Watson, D. K., Smyth, F. E., Thompson, D. M., Cheng, J. Q., Testa, J. R., Papas, T. S. & Seth, A. (1992) The ERGB/Fli-1 gene: isolation and characterization of a new member of the family of human ETS transcription factors. *Cell growth and differentiation*, 3, 705-705.
Watson, S. P., Lowe, G. C., Lordkipanidze, M., Morgan, N. V. & consortium, G. (2013) Genotyping and phenotyping of platelet function disorders. *J Thromb Haemost*, 11 Suppl 1, 351-63.
Werner, M. H., Clore, G. M., Fisher, C. L., Fisher, R. J., Trinh, L., Shiloach, J. & Gronenborn, A. M. (1997) Correction of the NMR structure of the ETS1/DNA complex. *Journal of biomolecular NMR*, 10(4), 317-328.
Westbury, S. K., Turro, E., Greene, D., Lentaigne, C., Kelly, A. M., Bariana, T. K., Simeoni, I., Pillois, X., Attwood, A., Austin, S., Jansen, S. B., Bakchoul, T., Crisp-Hihn, A., Erber, W. N., Favier, R., Foad, N., Gattens, M., Jolley, J. D., Liesner, R., Meacham, S., Millar, C. M., Nurden, A. T., Peerlinck, K., Perry, D. J., Poudel, P., Schulman, S., Schulze, H., Stephens, J. C., Furie, B., Robinson, P. N., van Geet, C., Rendon, A., Gomez, K., Laffan, M. A., Lambert, M. P., Nurden, P., Ouwehand, W. H., Richardson, S., Mumford, A. D., Freson, K. & Consortium, B.-B. (2015) Human phenotype ontology annotation and cluster analysis to unravel genetic defects in 707 cases with unexplained bleeding and platelet disorders. *Genome Med*, 7(1), 36.
White, J. G. (2007) Platelet storage pool deficiency in Jacobsen syndrome. *Platelets*, 18(7), 522-527.
Whitlock, J. M. & Hartzell, H. C. (2017) Anoctamins/TMEM16 proteins: chloride channels flirting with lipids and extracellular vesicles. *Annual review of physiology*, 79, 119-143.
Williams, C., Savage, J., Harper, M., Moore, S., Hers, I. & Poole, A. (2016) Identification of roles for the SNARE-associated protein, SNAP29, in mouse platelets. *Platelets*, 27(4), 286-294.
Wilson, N. K., Foster, S. D., Wang, X., Knezevic, K., Schutte, J., Kaimakis, P., Chilarska, P. M., Kinston, S., Ouwehand, W. H., Dzierzak, E., Pimanda, J. E., de Bruijn, M. F. & Gottgens, B. (2010) Combinatorial transcriptional control in blood stem/progenitor cells: genome-wide analysis of ten major transcriptional regulators. *Cell Stem Cell*, 7(4), 532-44.
Witke, W., Li, W., Kwiatkowski, D. J. & Southwick, F. S. (2001) Comparisons of CapG and gelsolin-null macrophages: demonstration of a unique role for CapG in receptor-mediated ruffling, phagocytosis, and vesicle rocketing. *J Cell Biol*, 154(4), 775-784.
Wood, A. W., Schlueter, P. J. & Duan, C. (2005) Targeted knockdown of insulin-like growth factor binding protein-2 disrupts cardiovascular development in zebrafish embryos. *Molecular endocrinology*, 19(4), 1024-1034.
Wood, L. D., Irvin, B. J., Nucifora, G., Luce, K. S. & Hiebert, S. W. (2003) Small ubiquitin-like modifier conjugation regulates nuclear export of TEL, a putative tumor suppressor. *Proceedings of the National Academy of Sciences*, 100(6), 3257-3262.
Wood, T. L., Rogler, L. E., Czick, M. E., Schuller, A. G. & Pintar, J. E. (2000) Selective alterations in organ sizes in mice with a targeted disruption of the insulin-like growth factor binding protein-2 gene. *Molecular Endocrinology*, 14(9), 1472-1482.
Wu, J., De Heus, C., Liu, Q., Bouchet, B. P., Noordstra, I., Jiang, K., Hua, S., Martin, M., Yang, C. & Grigoriev, I. (2016) Molecular pathway of microtubule organization at the Golgi apparatus. *Developmental cell*, 39(1), 44-60.
Xu, J., Zhang, L., Ye, Y., Shan, Y., Wan, C., Wang, J., Pei, D., Shu, X. & Liu, J. (2017) SNX16 regulates the recycling of E-cadherin through a unique mechanism of coordinated membrane and cargo binding. *Structure*, 25(8), 1251-1263. e5.
Xue, Y., Chen, Y., Ayub, Q., Huang, N., Ball, E. V., Mort, M., Phillips, A. D., Shaw, K., Stenson, P. D., Cooper, D. N., Tyler-Smith, C. & Genomes Project, C. (2012) Deleterious- and disease-allele prevalence in healthy individuals: insights from current predictions, mutation databases, and population-scale resequencing. *Am J Hum Genet*, 91(6), 1022-32.
Yadav, S., Williamson, J. K., Aronova, M. A., Prince, A. A., Pokrovskaya, I. D., Leapman, R. D. & Storrie, B. (2017) Golgi proteins in circulating human platelets are distributed across non-stacked, scattered structures. *Platelets*, 28(4), 400-408.
Yago, T., Lou, J., Wu, T., Yang, J., Miner, J. J., Coburn, L., Lopez, J. A., Cruz, M. A., Dong, J. F., McIntire, L. V., McEver, R. P. & Zhu, C. (2008) Platelet glycoprotein Ibalpha forms catch bonds with human WT vWF but not with type 2B von Willebrand disease vWF. *J Clin Invest*, 118(9), 3195-207.
Yamamoto, R., Morita, Y., Ooehara, J., Hamanaka, S., Onodera, M., Rudolph, K. L., Ema, H. & Nakauchi, H. (2013) Clonal analysis unveils self-renewing lineage-restricted progenitors generated directly from hematopoietic stem cells. *Cell*, 154(5), 1112-1126.
Yamano, G., Funahashi, H., Kawanami, O., Zhao, L.-X., Ban, N., Uchida, Y., Morohoshi, T., Ogawa, J., Shioda, S. & Inagaki, N. (2001) ABCA3 is a lamellar body membrane protein in human lung alveolar type II cells. *FEBS letters*, 508(2), 221-225.
Yi, H.-k., Fujimura, Y., Ouchida, M., Prasad, D., Rao, V. N. & Reddy, E. S. P. (1997) Inhibition of apoptosis by normal and aberrant Fli-1 and erg proteins involved in human solid tumors and leukemias. *Oncogene*, 14(11), 1259.
Zang, C., Luyten, A., Chen, J., Liu, X. S. & Shivdasani, R. A. (2016) NF-E2, FLI1 and RUNX1 collaborate at areas of dynamic chromatin to activate transcription in mature mouse megakaryocytes. *Sci Rep*, 6, 30255.

Zhang, L., Eddy, A., Teng, Y.-T., Fritzler, M., Kluppel, M., Melet, F. & Bernstein, A. (1995) An immunological renal disease in transgenic mice that overexpress Fli-1, a member of the ets family of transcription factor genes. *Molecular and cellular biology*, 15(12), 6961-6970.
Zhang, L., Lemarchandel, V., Romeo, P.-H., Ben-David, Y., Greer, P. & Bernstein, A. (1993) The Fli-1 proto-oncogene, involved in erythroleukemia and Ewing's sarcoma, encodes a transcriptional activator with DNA-binding specificities distinct from other Ets family members. *Oncogene*, 8(6), 1621-1630.
Zhang, L., Shao, H., Zhu, T., Xia, P., Wang, Z., Liu, L., Yan, M., Hill, D. L., Fang, G. & Chen, Z. (2013) DDA3 associates with microtubule plus ends and orchestrates microtubule dynamics and directional cell migration. *Scientific reports*, 3, 1681.
Zhang, M. Y., Churpek, J. E., Keel, S. B., Walsh, T., Lee, M. K., Loeb, K. R., Gulsuner, S., Pritchard, C. C., Sanchez-Bonilla, M., Delrow, J. J., Basom, R. S., Forouhar, M., Gyurkocza, B., Schwartz, B. S., Neistadt, B., Marquez, R., Mariani, C. J., Coats, S. A., Hofmann, I., Lindsley, R. C., Williams, D. A., Abkowitz, J. L., Horwitz, M. S., King, M. C., Godley, L. A. & Shimamura, A. (2015) Germline ETV6 mutations in familial thrombocytopenia and hematologic malignancy. *Nat Genet*, 47(2), 180-5.
Zhang, Q., Zhao, B., Li, W., Oiso, N., Novak, E. K., Rusiniak, M. E., Gautam, R., Chintala, S., O'Brien, E. P., Zhang, Y., Roe, B. A., Elliott, R. W., Eicher, E. M., Liang, P., Kratz, C., Legius, E., Spritz, R. A., O'Sullivan, T. N., Copeland, N. G., Jenkins, N. A. & Swank, R. T. (2003) Ru2 and Ru encode mouse orthologs of the genes mutated in human Hermansky-Pudlak syndrome types 5 and 6. *Nat Genet*, 33(2), 145-53.
Zhang, X. K., Gallant, S., Molano, I., Moussa, O. M., Ruiz, P., Spyropoulos, D. D., Watson, D. K. & Gilkeson, G. (2004) Decreased expression of the Ets family transcription factor Fli-1 markedly prolongs survival and significantly reduces renal disease in MRL/lpr mice. *The Journal of Immunology*, 173(10), 6481-6489.
Zhang, X. K., Moussa, O., LaRue, A., Bradshaw, S., Molano, I., Spyropoulos, D. D., Gilkeson, G. S. & Watson, D. K. (2008) The transcription factor Fli-1 modulates marginal zone and follicular B cell development in mice. *The Journal of Immunology*, 181(3), 1644-1654.
Zhao, H., Zhao, Y., Li, Z., Ouyang, Q., Sun, Y., Zhou, D., Xie, P., Zeng, S., Dong, L. & Wen, H. (2018) FLI1 and PKC co-activation promote highly efficient differentiation of human embryonic stem cells into endothelial-like cells. *Cell death & disease*, 9(2), 131.
Zheng, C., Yang, R., Han, Z., Zhou, B., Liang, L. & Lu, M. (2008) TPO-independent megakaryocytopoiesis. *Critical reviews in oncology/hematology*, 65(3), 212-222.
Zhou, Q., Lee, G.-S., Brady, J., Datta, S., Katan, M., Sheikh, A., Martins, M. S., Bunney, T. D., Santich, B. H. & Moir, S. (2012) A hypermorphic missense mutation in PLCG2, encoding phospholipase Cγ2, causes a dominantly inherited autoinflammatory disease with immunodeficiency. *The American Journal of Human Genetics*, 91(4), 713-720.
Züchner, S., Noureddine, M., Kennerson, M., Verhoeven, K., Claeys, K., De Jonghe, P., Merory, J., Oliveira, S. A., Speer, M. C. & Stenger, J. E. (2005) Mutations in the pleckstrin homology domain of dynamin 2 cause dominant intermediate Charcot-Marie-Tooth disease. *Nature genetics*, 37(3), 289.
zur Stadt, U., Schmidt, S., Kasper, B., Beutel, K., Diler, A. S., Henter, J. I., Kabisch, H., Schneppenheim, R., Nurnberg, P., Janka, G. & Hennies, H. C. (2005) Linkage of familial hemophagocytic lymphohistiocytosis (FHL) type-4 to chromosome 6q24 and identification of mutations in syntaxin 11. *Hum Mol Genet*, 14(6), 827-34.