Utilising the zebrafish model organism to study the effect of FLCN in early embryonic development

and

Genotyping and platelet phenotyping of cases of rare inherited platelet based bleeding disorders

By Ben Johnson

This project is submitted in partial fulfilment of the requirements for the award of the MRes.

UNIVERSITY OF BIRMINGHAM

Institute of Biomedical Research

School of Clinical and Experimental Medicine

University of Birmingham

May 2013

Utilising the zebrafish model organism to study the effect of FLCN in early embryonic development

By Ben Johnson

This project is submitted in partial fulfilment of the requirements for the award of the MRes.

UNIVERSITY OF BIRMINGHAM

Institute of Biomedical Research

School of Clinical and Experimental Medicine

University of Birmingham

May 2013

Abstract

Birt-Hogg-Dubé syndrome (BHD) (OMIM: #135150) is a rare genodermatosis following a dominant pattern of inheritance. The disease is characterised by the clinical presentation of benign cutaneous fibrofolliculomas and a predisposition to pneumothorax and renal cell carcinomas. The heritable genetic cause of BHD has been determined as mutations within the 3638bp transcript encoding the gene *Folliculin* (*FLCN*). FLCN is thought to mediate a molecular effect through two binding proteins; FNIP1 and FNIP2, and a role in the Akt-mTOR pathway through interaction with AMPK has been elucidated through co-immunoprecipitation assays. A variable phenotype has been observed though in systems modelling loss of function. This study therefore intends to take a novel approach to mutational studies by utilising the zebrafish model organism. The ease of use, application of systems such as morpholinos and optical clarity provided by zebrafish embryos make them ideal for functional assays *in vivo.* We exploit this benefit using imaging analysis of a Fucci transgenic line and *in situ* hybridisation and are able to suggest a possible functional role of FLCN within cell cycle regulation and morphogenesis in the developing brain. In addition, progress is made to determining the embryonic lethality of null homozygous mutants in the zebrafish model organism.

Word count - 200

Contents

List of Illustrations

Introduction

Birt-Hogg-Dubé Syndrome (BHD) (OMIM: #135150) is a rare dominantly inherited genodermatosis with a predisposition to hereditary kidney neoplasms and benign cutaneous hamartomas. In past literature it has been cited as Hornstein-Knickenberg syndrome due to a 1975 study by Hornstein and Knickenberg (2) in which they described a sibling pairing with a presentation of perifollicular fibromas with association of intestinal polyps. In 1977 Drs. Birt, Hogg and Dubé (3) described similar cutaneous symptoms in a more extensive kinship with 15/70 affected family members. Both studies presented cases of what is now referred to as BHD. Birt *et al.* (3) used the extensive family available of the index case and the presence or absence of papular skin lesions to determine an autosomal dominant pattern of inheritance through pedigree studies.

Clinical presentation and current treatment options

As both aforementioned studies have suggested one of the main clinical phenotypes observed within patients is the presence of small, rounded, dome-shaped skin lesions from which the Birt *et al.* paper (3) coined the terminology of fibrofolliculomas. These smooth and skin-coloured papules are often determined as the defining pathology of the disease and are most prevalent on the head, neck and trunk of affected patients. They are benign tumours in the hair follicle due to spatially restricted proliferation of collagen and fibroblasts (4). They occur predominately in patients older than 20 (1) with some papers suggesting onset of skin lesions only occurring in the third to fourth decade of life in affected individuals (5). They are often found in vast numbers and can be present together with lesions of similar morphology including trichodiscomas (both present in 31% of patients) and acrochordons (91%) (5).

The most threatening clinical complication of BHD is a linked predisposition to malignant carcinoma, in particular renal neoplasms. In a study of 124 BHD patients Pavlovich (6) found 27% (34/124) to have a renal cell carcinoma (RCC) of varying histology. Zbar, in his 2002 paper (7), suggested an odds

ratio (OR) of 6.9 that patients with BHD will contract a renal tumour in multivariate age adjusted analysis and that the median age of detection was 51 years. The renal tumours published in patients suffering with BHD vary greatly in histology with four main sub-categories observed; papillary RCC, clear cell RCC, chromophobe RCC and oncocytomas. Tumours are often multifocal and bilateral with chromophobe RCCs being most common but vast familial variation and individual variation occurs with many hybrid cases of sub-categories within one kidney or one tumour being reported (8). 38 months follow up of 10 patients by Pavlovich *et al.,* (6) revealed 80% overall survival with five remaining disease free, three retaining small renal tumours and two patients dying of metastatic RCC.

Another important symptom of BHD is pulmonary cysts and spontaneous pneumothorax. The original link between BHD and lung cysts was first reported by Toro *et al.* (5) where the complication was found in four out of 28 cases using chest radiography. Since then, numerous cases of pulmonary cysts within patients suffering from BHD have been reported with recent studies showing 24% prevalence in BHD patients (9) and an OR 32 times as great, for those with BHD compared to those without, in multivariate analysis (7). In BHD it can be said that the pneumothorax seen cannot be classed as spontaneous but more as a possible secondary complication when lung cysts rupture under pressure (4).

Previous diagnostic criteria relied on the presence of five or more facial papules or at least one papule confirmed as a fibrofolliculoma using specific criterion (5). Menko *et al.* (1) have built on this knowledge and suggested the major and minor diagnostic criteria as shown in Table 1. One of the main clinical problems with BHD is its variation; be that across a panel of patients or within a family with a predicted segregating mutation. Clinical presentation of the symptoms stated previously can occur at a varied time post the second decade of life (1, 10). This presents an altogether different challenge in treating BHD and its symptoms and susceptibilities requiring a personal maintenance approach.

Currently treatment aims to achieve just this and focuses on clinical management of phenotypic presentations and screening for critical complications such as renal cell carcinoma. Laser ablation, with varying results, using systems such as erbium:YAG laser (11), is currently suggested for use for aesthetical treatment of fibrofolliculomas and other presented skin lesions. Management of the most threatening clinical complication, RCC, is achieved through regular monitoring from a relatively early age. It is suggested now that scanning should occur from the age of 20 upwards at yearly intervals (1). Computerised tomography is the preferred method of analysis due to its improved ability to detect smaller masses (<3cm diameter) when compared to ultrasound techniques (12) but it does raise the issue of an increased risk to cumulative radiation across a lifetime. Upon detection of a renal carcinoma mass normal staging and grading occurs followed by appropriate treatment, usually nephron sparing surgery, with promising event free survival, for masses larger than 3cm (13). Tackling other complications of BHD, such as pneumothorax and colonic cysts, occurs currently at presentation of symptoms. The high risk factor of pneumothorax in patients suffering with BHD though does provide evidence for the instigation of CT monitoring of pulmonary cysts in patients at risk (1).

Diagnostic criteria for Birt-Hogg-Dubé syndrome (BHD; patients should fulfill one major or two minor criteria for diagnosis)

Major criteria

- At least five fibrofolliculomas or trichodiscomas, at least one histologically confirmed, of adult onset*
- Pathogenic *FLCN* germline mutation

Minor criteria

- Multiple lung cysts: bilateral basally located lung cysts with no other apparent cause, with or without spontaneous primary pneumothorax
- Renal cancer: early onset (<50 years) or multifocal or bilateral renal cancer, or renal cancer of mixed chromophobe and oncocytic histology
- A first-degree relative with BHD

Figure 1. Diagnostic criteria for diagnosis of BHD upon patient presentation, as described by Menko *et al.* (1).

Folliculin, the causative gene underlying BHD

From the first mention of heritability (3) and the dominant pattern of inheritance observed in pedigree studies it had been clear there was a genetic causation of BHD but it was not until 2001 when Khoo *et al.* (14) first mapped the BHD gene to a locus between 17p12-q11.2. They used microsatellite markers every 10cM and haplotype mapping to determine segregation in affected patients. This work was built upon and confirmed by Schmidt *et al.* (15) who narrowed down the location of the causative gene to 17p11.2 using a high penetrance model of disease and linkage analysis software. In all patient families Schmidt analysed high LOD scores were obtained for a <4cM region between microsatellite markers in the 17p11 region. Cloning of the affected gene was first performed by Nickerson *et al.* (16) and followed a logical progression from narrowing down the 4cM region using additional familial recombination studies to then gene mining the critical region. This method returned two possible alternately spliced transcripts. Finally cDNA extracted from kidney and lung, thought to be a combination of the two alternately spliced transcripts, was sequenced with >4 fold coverage. The resultant product is a sequence of 3638 nucleotides/14 exons, 11 of which are coding, that they named *folliculin (FLCN)* after the fibrofolliculomas in clinical presentation.

Studies by Schmidt *et al.* (17) and Toro (18) have showed prevalence of *FLCN* mutations in 84 and 88% of BHD patients, respectively. As of the 26th of March 2013 there are 153 described variants within the European Birt-Hogg-Dube consortium database (19). Toro (18) reported that 47% (24/51) of familial mutations occur within exon 11, 19 of which occur at a “mutational hotspot” in a mononucleotide tract of eight cytosine residues, $(C)_8$, and this was reported earlier by Schmidt (17). In Toro’s study 74% had an insertion of cytosine at position 1733 in the coding sequence (c.1733insC) and 24% had a deletion (c.1733delC) and a literature review by Toro found 125 mutations from 40 families at this point in the sequence. Mutations in the 2008 Toro study occurred in only translated exons 4, 5, 6, 9, 11, 12 and 13 and consisted of insertions, deletions, missense and nonsense single nucleotide changes and splice site mutations (18). With the majority of the

mutations in the Toro study being one and two nucleotide insertions or deletions the majority of mutations produced a frame shift and a subsequent altered AA sequence which is predicted to truncate the protein product. Several studies have attempted to create a genotype-phenotype link between mutation and the spectrum of clinical complications. Schmidt (17) initially reported an increased prevalence of RCC in patients with c.1733delC compared to c.1733insC, this was again reported by Toro *et al.* (18) but it was found to be not statistically significant (p=0.11). Toro also found no other links between mutation type, mutation vs. no mutation and any of the main phenotypic presentations.

The 3638bp *FLCN* transcript codes for a 579 AA, 64kDa protein. mRNA expression of *FLCN* in humans is seen in tissues of the skin, kidney, lung as well as tissues of the tonsils, lymph nodes, spleen, brain, breast, bladder, testis, prostate, ovary, myometrium, pancreas and parotid gland (20). Expression is also seen in the epidermis of fibrofolliculomas and lung blebs from BHD patients but no expression was noted in renal carcinoma cells of the same patients. Interestingly no expression was seen in the endothelial cells of the colon.

Molecular function of FLCN

To this date there seems to be a level of ambiguity surrounding the molecular function of FLCN. A second mutational hit, either by somatic mutation or loss of heterozygosity (LOH), has been previously reported in tumour samples as evidence of a tumour suppressor role of FLCN. LOH was confirmed in 17% of renal tumours (21) indicating a possible tumour suppressor function of FLCN but in similar studies looking at LOH in fibrofolliculomas (22) no second-hit mutations could be observed suggesting haploinsuffuciency may be the causative effect of tumourogensis in this particular phenotype.

Efforts to determine the function of FLCN have mainly been focused around studying possible binding partners of the full-length transcript protein product and associated pathways where it may play a role. Initial studies in *Drosophila* suggested a role downstream of the JAK/STAT pathway in germline stem cell maintenance in fly testis (23), but it was a 130kDa interacting protein, first

established using co-immunoprecipitation experiments (24), designated FLCN-interacting protein 1 (FNIP1) that elucidated a possible explanation for the phenotypic presentation of BHD. The team from the National Cancer Institute in Frederick confirmed a similar expression pattern of FNIP1 to FLCN and interactions of FNIP1 with the C-termini of FLCN from mutational studies. Using similar co-immunoprecipitation studies the group were also able to prove that FNIP1 can bind to AMP-activated protein kinase (AMPK) independently of FLCN or together to form a complex. FNIP1 can also act as a substrate of AMPK and be phosphorylated in the AMPK pathway and the same can be said for FLCN where its phosphorylation is diminished upon AMPK inhibition. More recently an uncharacterised cDNA sequence, homologous to FNIP1, was discovered using BLAST sequence searches and designated as FNIP2 (25). Both sequences shared 49% identity and 74% sequence similarity with highly conserved regions in the C-termini and the N half of the protein. Like FNIP1, FNIP2 binds to FLCN in *in vitro* studies and co-localises with AMPK in the cytoplasm of the HEK293 kidney cell line. FNIP1 and 2 can also form multimers independent of FLCN. The difference between the possible roles of both FNIP1 and FNIP2 seems to lie in their slightly altered expression patterns, with higher expression of FNIP2 seen in the fat, liver and pancreas. Interestingly when varying histological renal cell carcinoma samples were tested for mRNA expression levels of the two FLCN-interacting proteins there seemed to be some deviation in expression depending on the type of tumour tested (25).

The interactions of FLCN and its two reported binding proteins with AMPK poses an interesting hypothesis to the function of FLCN and how deregulation through mutation may lead to the phenotypic effects seen in BHD. AMPK acts as a sensor of ATP levels in eukaryotic cells and is activated upon elevated AMP due to a depleted ATP environment (26). It is upon this activation that it is believed that AMPK can inhibit the function of mTOR, through phosphorylation, which, in turn, correlates with findings of a reduced mTOR function in low nutrient levels (27). mTOR, a key member of the PI3K-AKT-mTOR signalling pathway, usually functions to increase cellular growth through increased protein synthesis and/or reduced protein degradation as well as number of other

cellular processes (reviewed in (28)). In theory this model of an altered AMPK and mTOR pathway due to *FLCN* mutation could explain the predisposition to uncontrolled cellular growth in fibrofolliculomas and carcinomas in BHD patients. This possible link is further strengthened with the involvement of causative genes from other hamartoma syndromes such as; TSC, Peutz-Jeghers syndrome and PTEN-related hamartoma syndromes playing roles within the AMPK-mTOR signalling pathway (29, 30). The main flaw with this hypothesis currently is that relies on not just an interaction between FLCN and AMPK but a reaction causing a change to how AMPK relays it's downstream signal. This is still yet to be determined.

Current models of BHD to elucidate the role of FLCN further involve *in vivo* cell-specific and whole organism knockouts in mice and *in vitro* tumour isolated cell lines. Through a Cre-recombinase system of mutation Baba *et al.* (31) were able to establish a murine model that produced Cre-mediated deletion of floxed *FLCN* sequences with a significantly lower level of expression compared to normal renal tissue. By week one after birth inactivated-FLCN kidneys began to enlarge and this accumulated in the vast production of cysts within the lumens of ducts and tubules and significant kidney failure by week three, at which point mice were euthanized. This experimental data has been confirmed in a competing paper where death by renal failure was observed at three weeks of age (32). Baba built on these initial findings to report an increase in proliferation and a marked difference in size of proximal ducts and tubules. They also went further to assess the effect on the previously linked Akt-mTOR pathway. They observed elevated mTOR phosphorylation and increased kidney:body weight ratio in *FLCN* knockout mice as well as increased survival of knockout mice when treated with the mTOR inhibitor rapamycin (31).

Hasumi *et al* (33) focused their efforts on producing a homozygous murine null mutant strain for *FLCN*. *FLCN* homozygous null mice ($BHD^{d/d}$) were embryonic lethal with no homozygous embryos observable at E9.5. Interestingly there was a lack of an organised cell layer in early homozygous mutant embryos and images from the visceral endoderm showed an increased cytoplasm and disorganised structure. Kidney cyst development was also noted in $BHD^{d/+}$ mice mimicking the

tumour phenotype within humans and like before (31) proteins from the Akt-mTOR pathway were elevated in comparison to normal kidney samples.

To date work studying the functional effect of FLCN has focused mainly around its most detrimental clinical phenotype of renal cysts and carcinoma. What we propose in the following study is a novel method to study the effects of FLCN in the early stages of development using the zebrafish model organism. A role of FLCN during development has already previously been suggested (33) so a better understanding of FLCN at this early and fundamental step will potentially help elucidate it's mechanism of effect in simple cellular processes such as growth and cellular organisation.

<u>The use of the Fucci transgenic system to study a novel phenotype within the midbrain</u>

Originally developed in 1985 by Summerton *et al.* and published in 1989 (34) morpholinos are an antisense 25-mer oligonucleotide designed around an adapted ribonucleoside backbone. They mediate a temporary effect, up to 5 days (35), by targeted binding with high efficacy to RNA, blocking the translation initiation complex by an RNAse H-independent steric blocking mechanism or, alternatively, splicing machinery. Using morpholino induced temporary knockdown targeted to the splice acceptor site of exon 2 (Sp2) of *FLCN* a novel phenotype during early development was observed. This phenotype consisted of fluid retention within the midbrain: hydrocephalus and is shown in Figure 2. We believe there may be a case for this phenotype being caused by a disorganisation of cells/a lack of cells within the midbrain during early development. Prior work in mouse models has already shown a disorganisation of cells and developmental structures in murine models (33) so this novel phenotype could be a similar effect at a later stage of development. We initially propose that the phenotype observed may be an effect of deregulated cell cycle control leading to an altered cell growth and proliferation. This model will potentially fit with the uncontrolled cellular growth seen in fibrofollicumoas and lung and renal cysts and could be as a result of altered mTOR activity but this will need to be explored further before a real link can be established.

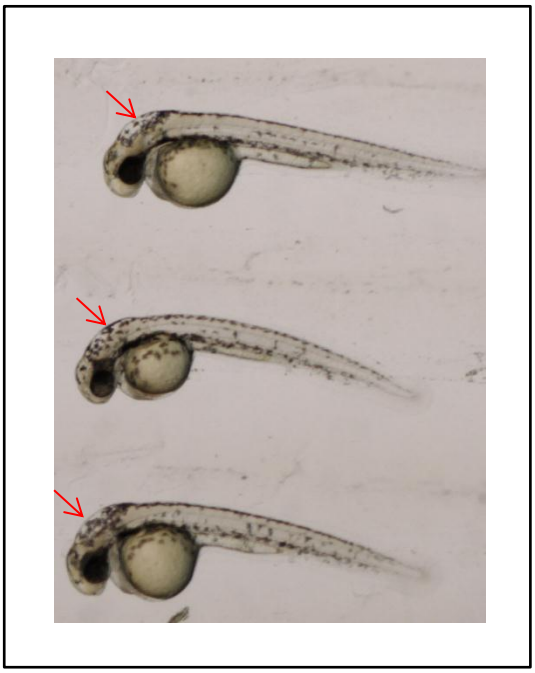


Figure 2 showing fluid retention in the midbrain in Sp2 morpholino injected embryos at 48hpf. Fluid retention is displayed as the clear area around the midbrain indicated by the red arrows. This novel phenotype is observable in all Sp2 injected embryos but some variability in severity is present.

To study the novel phenotype we intend to employ *in vivo* imaging tactics to study cell-cycle control. Originally developed for cell line models to study cell-cycle progression (36) the Fucci (*f*luorescent *u*biquitination-based *c*ell *c*ycle *i*ndicator) transgenic line provides a unique perspective for studying spatial and temporal regulation of the cell cycle. More recently this system has been applied to the zebrafish model organism (37) allowing for live *in vivo* time-lapse imaging facilitated by the translucent nature of zebrafish embryos. The Fucci transgenic line revolves around the conjugation of fluorescent proteins; monomeric Kusabira Orange2 (mKO_2) and monomeric Azami Green (mAG) to the ubiquitination domains of *Ctd1* and *geminin* respectively. Cells control Ctd1 and geminin activity at the protein level by targeting them for destruction using ubiquitination so that Ctd1 levels peak during the G1 phase of the cell cycle and geminin the S/G2/M phase. Transfection of the conjugated proteins allows them to accumulate reciprocally in the nuclei of cells effectively creating a cell cycle marking system by labelling cells in G1 orange/red and those in S/G2/M green. The benefit of this system is that cell-cycle progression can monitored live by the length of each stage of the cell cycle thus creating a platform for mutational comparisons and their effect on the cell cycle. The way in which we hope to utilise this imaging system is by optimising an experimental procedure

that can use the basics of the Fucci transgenic line, different fluorescence corresponding to different stages of the cell cycle, to assess cell cycle control over time in both FLCN temporary knockout and wild type embryos. We hypothesise that by focusing on the phenotypically interesting area of the tectum and midbrain we can establish a role of FLCN in cell-cycle control. Also by direct comparison of mutant embryos to wild type embryos we have the potential to suggest a theory for the phenotypical manifestations seen in BHD as a result of aberrant cell-cycle control.

Zebrafish as a model organism to study *FLCN*

Zebrafish, *Danio rerio,* (Taxonomy ID – 7955) have been used as a model organism for over 30 years (38), they are desirable for studying vertebrate development due to their ease of use, fast reproductive cycle and beneficial anatomical and developmental features such as a translucent embryo and development outside of a womb. They have been utilised for many genetic screens including random mutagenesis by ENU generation of point mutations and as a result there is a large surrounding open database of genomic and developmental information available at www.zfin.org. They provide an ideal intermediate in terms of cost and conserved homology and are favoured in this particular study due to their optical clarity.

FLCN is highly conserved throughout eukaryotes with orthologous sequences in organisms down to *D.melongaster* and *C.elegans.* The zebrafish orthologue of *FLCN*, *zFLCN*, is located on zebrafish chromosome 16 and encodes a 558 AA protein with 68.8% sequence identity when compared to the human sequence. Other members of the mTOR pathway are also conserved within zebrafish including mTOR and the coding subunits of AMPK, as well provisional homologs are suggested for both FNIP1 and FNIP2 proving the adaptability of the zebrafish model to this field. There is little published information regarding zFLCN and its function so the approach I will take in this study is a complete categorisation of the role of FLCN in early development of the zebrafish model organism. Therefore it can be said we have two main aims for the initial body of work excluding the application of the FUCCI transgenic system;

- Determine *FLCN* mRNA expression patterns in early development using whole mount *in-situ* hybridisation.
- Confirm/dispute embryonic lethality previously established in murine models using a homozygous null mutant breeding strain.

Whole mount *in-situ* hybridisation will be used as opposed to tissue *in-situ* hybridisation employed by prior localisation studies as the small size of zebrafish embryos makes it unsuitable for sectioning. The confirmation of embryonic lethality and hence agreement with work published in mice will provide a more substantial level of evidence to a possible function of FLCN in early development. Whether this putative function at early development is the same or a similar manifestation of that observed in adult tissue that can cause a retention of fluid within the mid-hindbrain is yet to be seen.

Overall we aim to achieve a better understanding of the role of FLCN in early embryonic development to help progress studies into BHD and its causation.

Methods

Zebrafish husbandry and care

Adult zebrafish were maintained in a facility utilising a recirculating water system with bead and sand filtration devices. UV lights sterilise post filtrated water. Water temperature is maintained at 28.5°C and between pH 7.2 and 7.5. Fish were reared and cared for according to standard protocols as set out within the 5th edition of The Zebrafish Book, the chapter dedicated to general methods to laboratory care, and home office regulations. Personnel followed home office guidelines for handling of transgenic and wild type fish.

Fish were bred according to standardised protocols where up to four fish were placed in breeding tanks fitted with collection trays. All tanks were maintained at 28.5°C and standardised humidity and were set up overnight to allow the fish to settle and grow a custom to their new surroundings. In the case where the precise time of laying was desired males and females were separated by a plastic screen until desired. In these cases time of laying was noted and embryos from separate breeding tanks were kept separate throughout subsequent experiments. Fish were never set up for breeding more than once every seven days and always returned to the same takes that they were removed from.

All experimental procedures involving zebrafish and care are covered by home office regulations and the personal licenses held within the University of Birmingham.

Transgenic lines

Fucci transgenic line: Tg{[EF1α:mKO2-zCdt1(1/190)][EF1α:mAG-zGem(1/100)]}.

Founder fish were supplied by Maya Sugiyama of the Laboratory for Cell Function and Dynamics, Advanced Technology Development Group, Brain Science Institute, RIKEN, 2-1 Hirosawa, Wako-city, Saitama 351-0198, Japan (37).

Heterozygous embryos for use in time-lapse imaging were generated by outcrossing founding females to wild type males.

Flcn heterozygous fish (Flcn$^{e36/+}$) were generated and supplied by the Sanger zebrafish mutational project (39). The e36 allele contains a point mutation affecting the 12th amino acid within exon 2. The G->T change is a nonsense mutation that causes a truncated non-functional 4 exon protein with a 194 amino acid total. The flanking region around the mutation is as follows;

GCCCTGTGCCACTTTTGT**[G/T]**AGCTCCATGGCCCAC.

The mutation disrupts a Sac1 restriction site.

Morpholinos and Micro-injections

Morpholinos (www.gene-tools.com) were injected into 1 cell embryos post fertilisation. 1.4nl, measured by a graded graticule, of morpholino at a concentration of 1mM, diluted in phenol red, was injected through the chorion into the single cell. Injections were performed using needles pulled from capillary glass tubing, Borosilicate – thin wall with filament (G100TF4) (1mm outer diameter, 0.78mm inner diameter). Embryos were stored at 28.5°C immediately after injection and moved to a 35°C incubator when the rate of development was wished to be increased.

The following morpholino sequences were used during experimentation,

Mismatch (mm) morpholino –

Splice 2 (Sp2) morpholino – CGTTCATCTGGAGGAAACAAACATA.

Localisation of the Splice2 morpholino is shown in Figure 3.

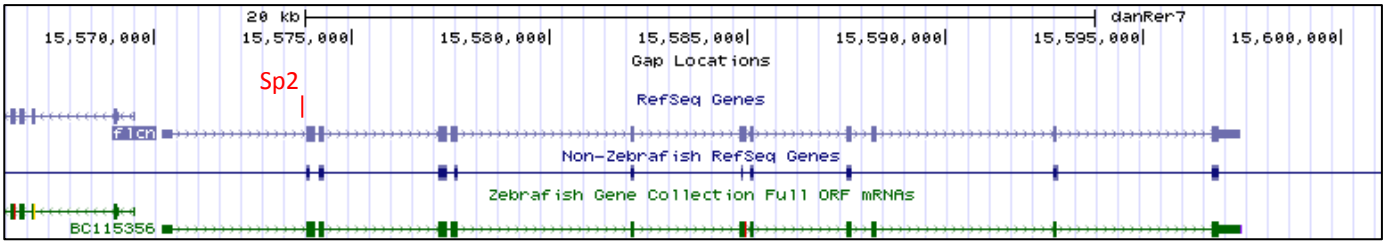


Figure 3.
Shown is the binding region of the splice 2 morpholino in relation to the *FLCN* sequence. The putative effect of the morpholino is a temporary knock down of the gene by abolishment of the function of the full length protein. The morpholino blocks splicing machinery binding at the exon 2 acceptor site causing the retention of the 3392 bases long 1st intron. The mismatch morpholino binds…

***In vivo* time-lapse imaging of heterozygous Fucci embryos**

Time-lapse imaging was performed on live embryos for the duration of 11 hours from 12 hours post fertilisation (bud-stage/1 somite stage) onwards. Preparation was performed by first adding 110µl of 1% agarose (Bioline) in 30% Danieus solution (1x solution: 2% NaCl (50x), 1% KCl, 1% $MgSO_4.7H_20$, 1% $Ca(NO_3)_2$, 1% HEPES, 92% H_20) (diluted in dH_2O) to each well of a 96-well flat bottomed plate (Greiner Bio-One). This was covered by 30µl of absolute ethanol and a brass 96 pin fakir bed template was inserted until the agarose set, upon which the template was wash thoroughly with distilled water. Embryos were manually dechorinated on 1% agarose in 30% Danieus and transferred with 60µl of solution to the prepared 96 well plate, one embryo per well. Embryos were orientated using a Microlance 3 needle with an outer diameter of 0.51mm (BD biosciences). 40µl of mineral oil (Sigma) was applied to the surface of the Danieus to prevent evaporation. A Leica LSI TCS Zoom confocal microscope and bundled software were used to capture the time-lapse video. The final experimental parameters used for time-lapse imaging are described in figure 4.

	LSI parameters		
Laser	Gain	Offset	Frame average
Green	1250	-5	8
Red	1200	-4	4
Brightfield	371	-	-
Format	1024 x 1024		
Optical zoom	1.362		
Step size	9.19µm		
z-volume	358.395		
Time per stack	22m 36s 367ms		
Lens	x5		
Speed	400Hz		
Room temperature	24°C		

Figure 4. The final parameters used for time-lapse videos captured with the confocal LSI microscope from Leica. All parameters were maintained throughout experimental procedures and repeats. Room temperature was maintained using an adjustable room cooling system. Time-lapse was conducted in xyzt mode using the 532 and 488 nm lasers adjusted to capture the wavelength of fluorescence produced by the conjugated Fucci proteins. 40 slices, still images taken along the z-axis, were taken per stack.

Time-lapse images were viewed and subsequently analysed using the Fiji (alternatively known as ImageJ) processing package (40). Overlay colours were added to the images from corresponding lasers and still images of individual frames were taken with the bundled software.

***In-situ* hybridisation**

Whole mount *In-situ* hybridisation (ISH) was performed using previously available probes at a working concentration of 1µg/ml in hybridisation mix. Antisense RNA probes were made in house to hybridise to the full length open reading frame of FLCN mRNA and were inserted into a pC52+ plasmid vector. ISH was performed on previously morpholino injected embryos at seven stages from 50% epiboly to 24 hours post fertilisation. All embryos were initially fixed by removing the 1x E3 medium (5mM NaCl, 0.17mM KCl, 0.33mM CaCl and 0.33mM $MgSO_4$ made up in dH_20) and replacing it with 4% PFA in PBS and stored overnight at 4°C. For embryos older than 20 somites the chorions were removed prior to fixation, those younger the chorions were removed post fixation. All embryos were washed and subsequently stored in 100% methanol at -20°C. Embryos were rehydrated in successive dilutions of PBS/Tween 20 0.1% and placed in hybridisation mix prior to being left overnight at 68°C in hybridisation mix containing FLCN antisense RNA probe. Embryos were then placed in antibody solution overnight at 4°C (PBT/2% goat serum/2mg:ml BSA/1:2500 anti-DIG antibody). The following day embryos were washed and the anti-serum removed. Embryos were placed in the dark in staining solution containing 100mg/ml NBT and 50mg/ml BCIP overnight with slight agitation at 4°C. The staining reaction was stopped and embryos re-fixed for 20 minutes in 4% PFA in PBS to ensure no further staining. Fixed embryos were set up for long-term storage in glycerol by gradual dehydration by methanol and increasing concentrations of glycerol (10, 20, 50 and 100% diluted in dH_2O were necessary).

In-situ hybridised embryos were mounted in a slide and images taken with a mounted HD Canon digital SLR camera at 10x optical zoom.

Polymerase chain reaction (PCR)

PCR was performed on whole euthanized embryos at regular stages from 5 days post fertilisation to 24 dpf. Whole euthanized embryos were added to 30µl of DNA extraction buffer (10mM tris pH8, 2mM EDTA, 0.2% Triton X-100 and 200µg/ml Proteinase K) containing Proteinase K at 200µg/ml and incubated at 55°C for three hours in 8 strip PCR eppendorf tubes. Proteinase K was inactivated at 95°C for 15 minutes. PCR tubes were centrifuged at 4000rpm for 10 minutes and the supernatant transferred to a fresh tube. 5µl of supernatant was visualised by etheidum bromide staining and loaded onto a 1% agarose gel in TAE (40mM Tris, 20mM acetic acid and 1mM EDTA made to a 1x solution in dH_20) along with 2µl of loading buffer to test success of DNA extraction. 3µl of DNA containing supernatant was used per 25µl PCR reaction. The following PCR primers were used for amplification of a 342bp section of *FLCN;*

Forward primer - catataagaatatgtttgtttcctcca

Reverse primer – tcagGGCTGTCGCTCTTTAC

The following cyclic PCR reaction was used;

Figure 5		
	95°C for 5 minutes	
	94°C for 30 seconds 59°C for 20 seconds 72°C for 45 seconds	x35 cycles
	72°C for 5 minutes	

Figure 5. Cyclic PCR parameters

5µl of PCR product was added to 1µl of Sac1 restriction enzyme and 2µl of NEbuffer1.1. Reactions were made up to 20µl with distilled water and incubated at 37°C for 2 hours. Sac1 enzyme recognises the 5bp palindromic sequence 5'…GAGCT ˇC…3'. The PCR reaction amplifies a 342bp sequence and after digestion the products produced are 276 and 66bp fragments. Therefore in heterozygotes because of the disruption by the point mutation to the Sac1 restriction site two bands will be observable within the visualised gel, one indicating a 342bp undigested fragment and the

other the 276bp digested fragment. The 66bp fragment will intentionally not be observable on the gel due to the running time chosen. 5µl of loading buffer (10mM Tris-HCl (ph 7.6), 0.03% bromophenol blue, 0.03% xylene cyanol FF, 60% glycerol and 60mM EDTA) was added to the digested PCR product which was visualised by ethedium bromide staining on a 1% agarose gel in TAE ran for 30 minutes at 120volts. PCR gels were also loaded with 100bp DNA ladder to confirm the presence of bands corresponding to extracted *FLCN* by size and dH_2O as a negative sample. PCR gels were viewed using a standard ultraviolet transilluminator imaging system.

Results

In vivo cell cycle analysis reveals possible cell cycle control deregulation in *FLCN* knock down embryos

Time-lapse imaging was first optimised using both mismatch and splice embryos for reference. The settings previously explained in Figure 4 were the final settings decided on providing optimal frame number and resolution per frame. Figure 6A shows still images from both mismatch and splice embryos. Both the red and green filters are shown to provide a comparison between percentages of cells in G1 to those in G2/S/M phases.

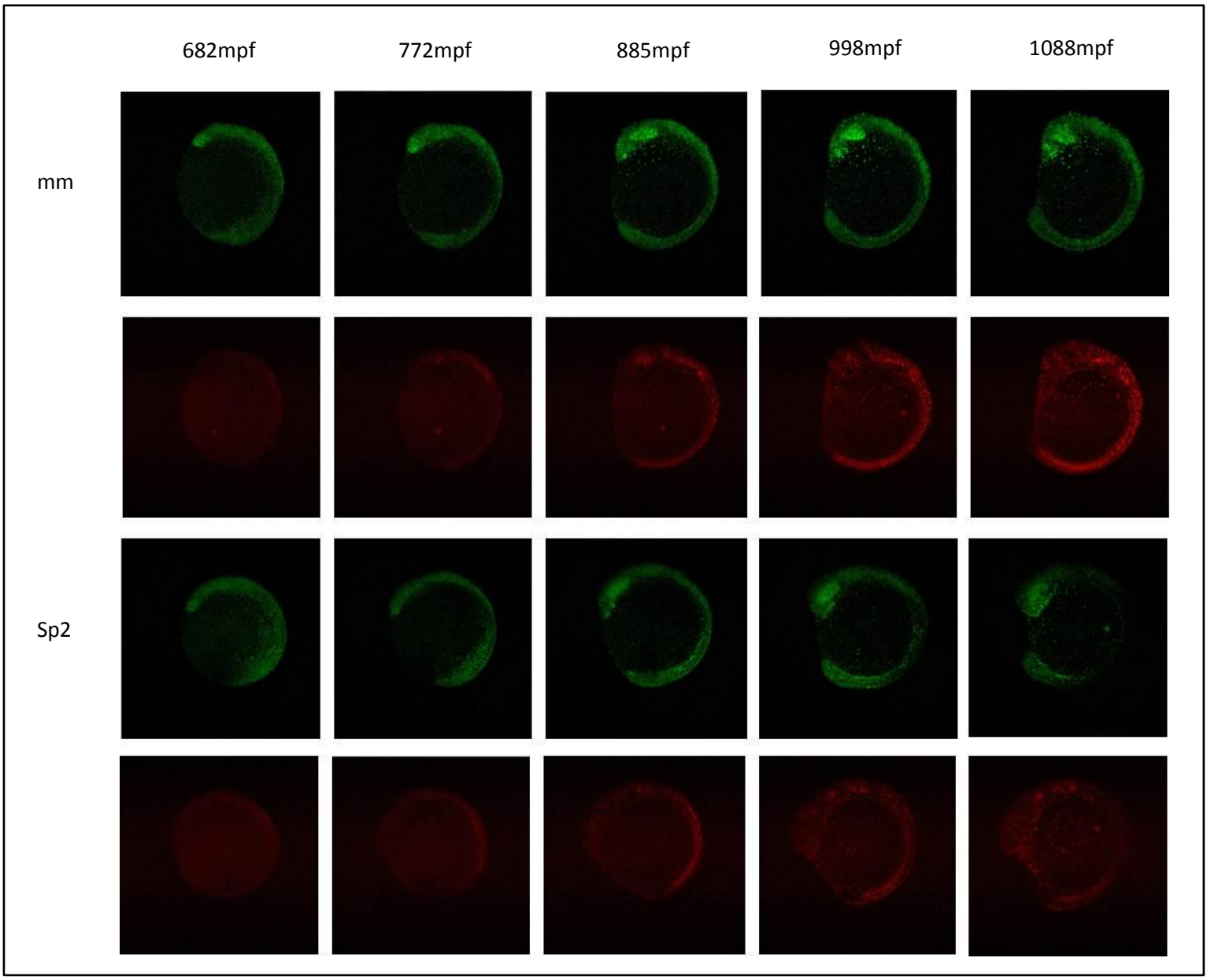
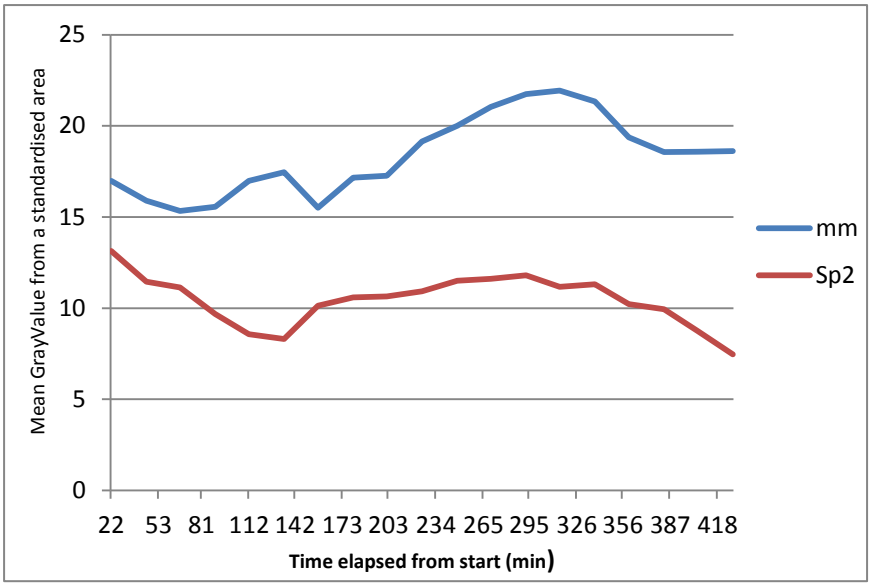


Figure 6A. Shown are still images from both a splice 2 and a mismatch time-lapse video. Still images were taken at the following time points post fertilisation. These correspond to 22, 112, 225, 338 and 428 minutes elapsed since the start of the time-lapse at the bud-1 somite stage.

Slight differences are observable in the fluorescence between the mismatch and splice2 embryos but clarification of these differences is hard to achieve without statistical analysis. We performed our statistical analysis using a protocol developed in house to compare levels of fluorescence. The results from the statistical analysis are shown in Figure 6B.

Green fluorescence (S/G2/M)



Red fluorescence (G1)

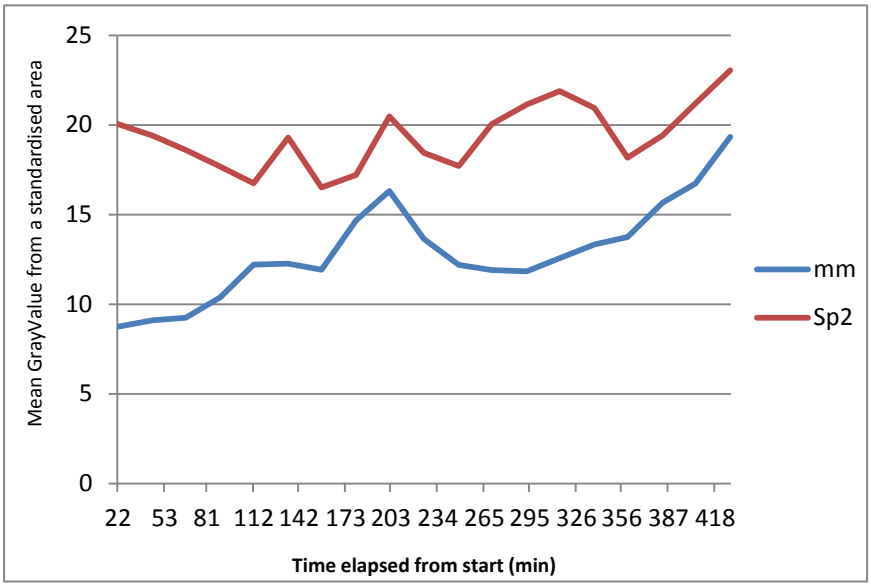


Figure 6B shows the statistical analysis of the time-lapse videos. Analysis was performed by using an internal measurement tool measuring the intensity of fluorescence (GrayValue) over an area of standard size on each frame. The measurement area was placed specifically over the region of the tectum. The value was averaged and plotted for each frame to determine the change in intensity over time for both the red and green fluorescent signal.

The region of the tectum was chosen for statistical analysis due to previous phenotype observed in Sp2 injected embryos. Noticeabl within the graphs is a clear difference in fluorescence intensity between the Sp2 and mm embryos analysed. This indicates an increased number of cells within the Sp2 injected embryo that are within the G1 (red) stage of the cell cycle whereas the cells within the mismatch embryo spend more time in S/G2/M stage shown by an increased intensity of the green signal. This experiment and statistical analysis has been repeated in work prior to this study and has confirmed the same patterns of intensity and differences between the two types of injected embryos.

Analysis of *FLCN* expression suggests tissue specific loss upon morpholino knock-down

In total seven different time points/stages of development between 50% epiboly and 24hpf were fixed for *in situ* hybridisation using the full length *FLCN* probe. The results are shown in the large panel in Figure 7A overleaf. Localisation of *FLCN* mRNA appears uniform and similar in both mm and Sp2 injected embryos as well as wild type uninjected embryos. The point where slight deviation occurs is at 24hpf. The most notable difference from localisation seen within mm and wild type embryos compared to Sp2 injected embryos is a variation in localisation and therefore subsequent expression at the midbrain-hindbrain boundary. This is shown in more detail in Figure 7B. This variation in staining was present across all Sp2 injected embryos at 24hpf that were screened using *in situ* hybridisation (n=8).

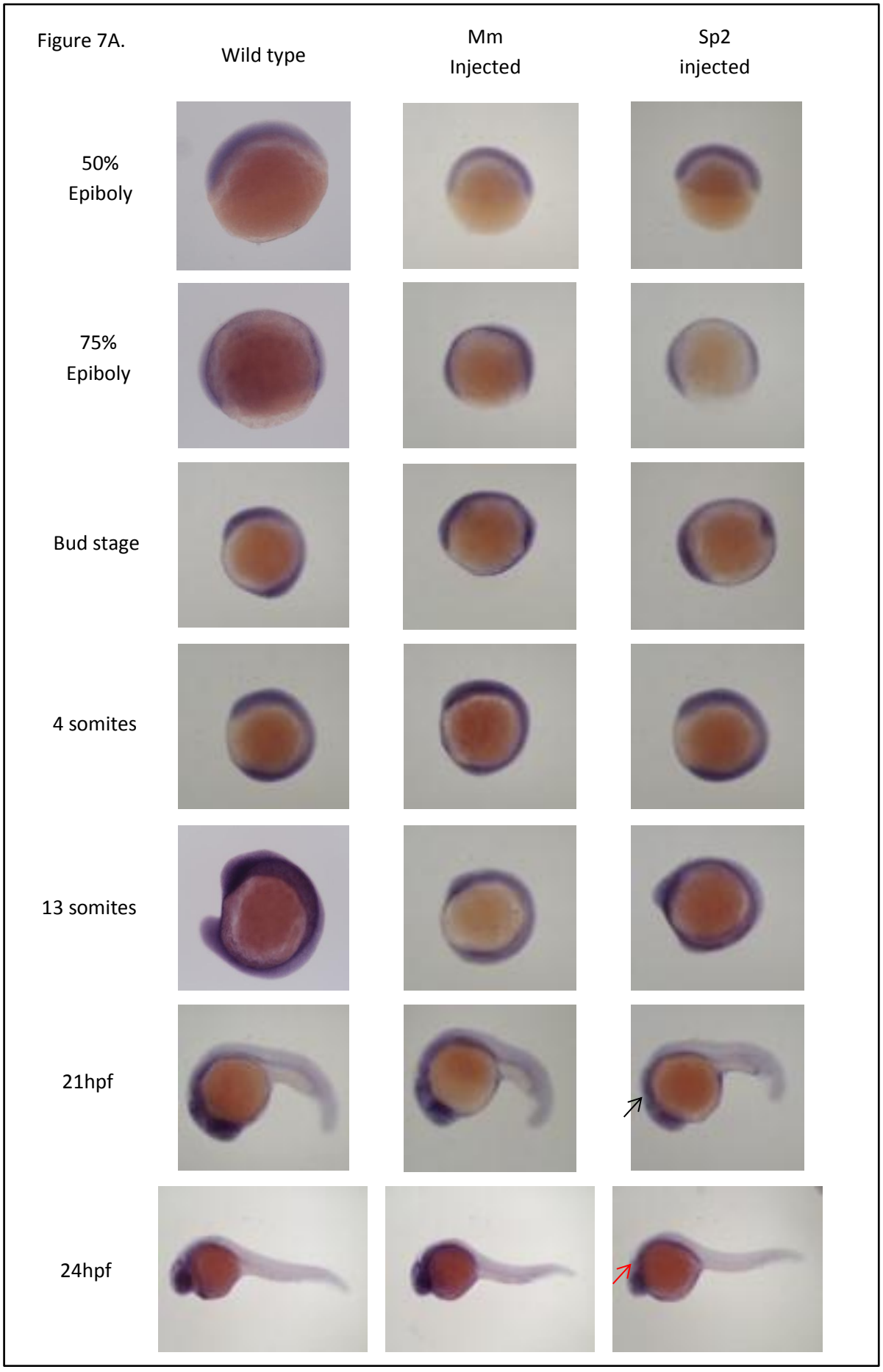


Figure 7A shows wild type, mm and Sp2 injected embryos across seven stages of early development probed with an *in situ* hybridisation probe to the full length *FLCN* mRNA. Little difference is observable between the three type of embryo at the early stages up to and including 13 somites. Noticeable change starts to appear at 21hpf where Sp2 embryos have a slight lack of expression within the developing mid to hindbrain region; this is marked with an arrow. Large variations in localisation and therefore expression are apparent at 24hpf where a loss of expression in the midbrain-hindbrain boundary is observed in Sp2 embryos and is show by the red arrow.

Figure 7B

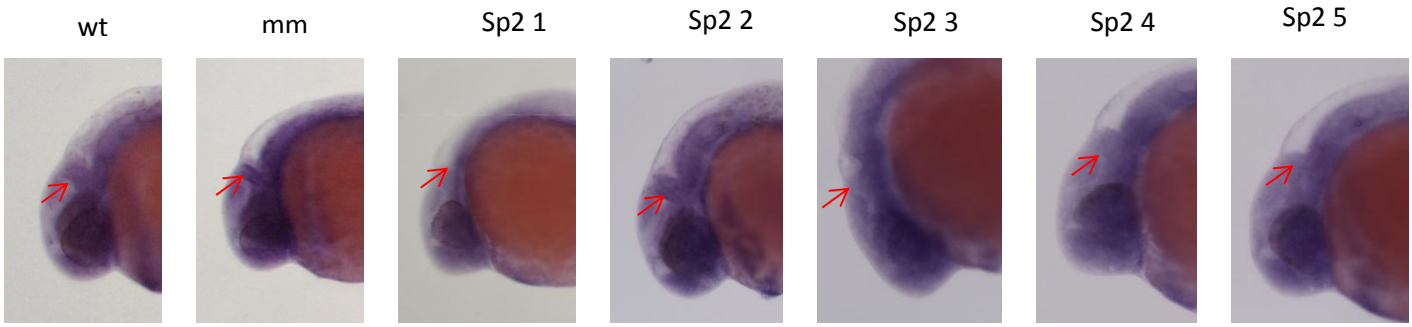


Figure 7B compares the mid to hindbrain regions of uninjected and injected embryos at a higher magnification (15x). The most obvious feature that was present uniformly in wild type and mismatch embryos but not Splice2 morpholino injected embryos was the midbrain-hindbrain boundary indicated by red arrows. In splice2 embryos (n=8) a widely variable staining was present in this region with few cases showing the stereotypical dark band of staining noticeable in the wt and mm embryos. This phonotypical variability is also present in the clear area posterior to the boundary with a number of embryos showing a reduction in the size of this corresponding hindbrain region.

PCR genotyping of *FLCN* mutant embryos reveals non-mendelian segregation rations and suggest that homozygous mutants do not survive past juvenile development

Embryonic digests to be used for PCR were available for the following time points; 5, 6, 7, 8, 9, 12, 15, 21 and 24dpf. At each stage a randomised selection of fish were taken. Fish that had not survived at the point of collection were not tested and were subsequently removed from the tank. The results from the PCR reaction are summarised in Figure 8. The subsequent genotype percentages at each stage post fertilisation are shown in Table 1. The expected ratio from a double heterozygous mutant cross is 1:2:1, homozygous mutant:heterozygous:homozygous wild type. Apart from 8dpf, which seems to be anomalous result, the percentages of the genotypes tends to follow the same predicted ratio. After this point the percentage of homozygous mutant embryos falls from 22% at 9dpf to 10%, 12%, 9% and finally 0 at 24dpf indicating a reduction of embryos homozygous for the mutant allele. In a repeat test performed prior to this work (data not shown) at 24dpf one embryo out of 22 was

determined to be homozygous mutant. This indicates that although the numbers may fall homozygous mutants are still present, but in very low numbers, up to and including 24dpf. Therefore if the mutation is detrimental to the survival of the embryo the effect cannot, in some cases, impact the organism up to and including this stage.

	No of embryos (percentage of total excluding failed samples)			
Days post fertilisation	Homozygous wild type (-/-)	Heterozygous (+/-)	Homozygous mutant (-/-)	Total
5	7 (29%)	9 (38%)	8 (33%)	24 (100%)
6	6 (25%)	11 (46%)	7 (29%)	24 (100%)
7	3 (19%)	10 (62%)	3 (19%)	16 (100%)
8	7 (44%)	9 (56%)	0 (0%)	16 (100%)
9	7 (30%)	11 (48%)	5 (22%)	23 (100%)
12	4 (20%)	14 (70%)	2 (10%)	20 (100%)
15	6 (24%)	16 (64%)	3 (12%)	25 (100%)
21	8 (35%)	13 (56%)	2 (9%)	23 (100%)
24	4 (31%)	9 (69%)	0 (0%)	13 (100%)

Table 1. Percentages of the three genotypes at each stage are shown. All percentages are taken from excluding the failed samples, average 1 per gel, from the total number.

Figure 8.

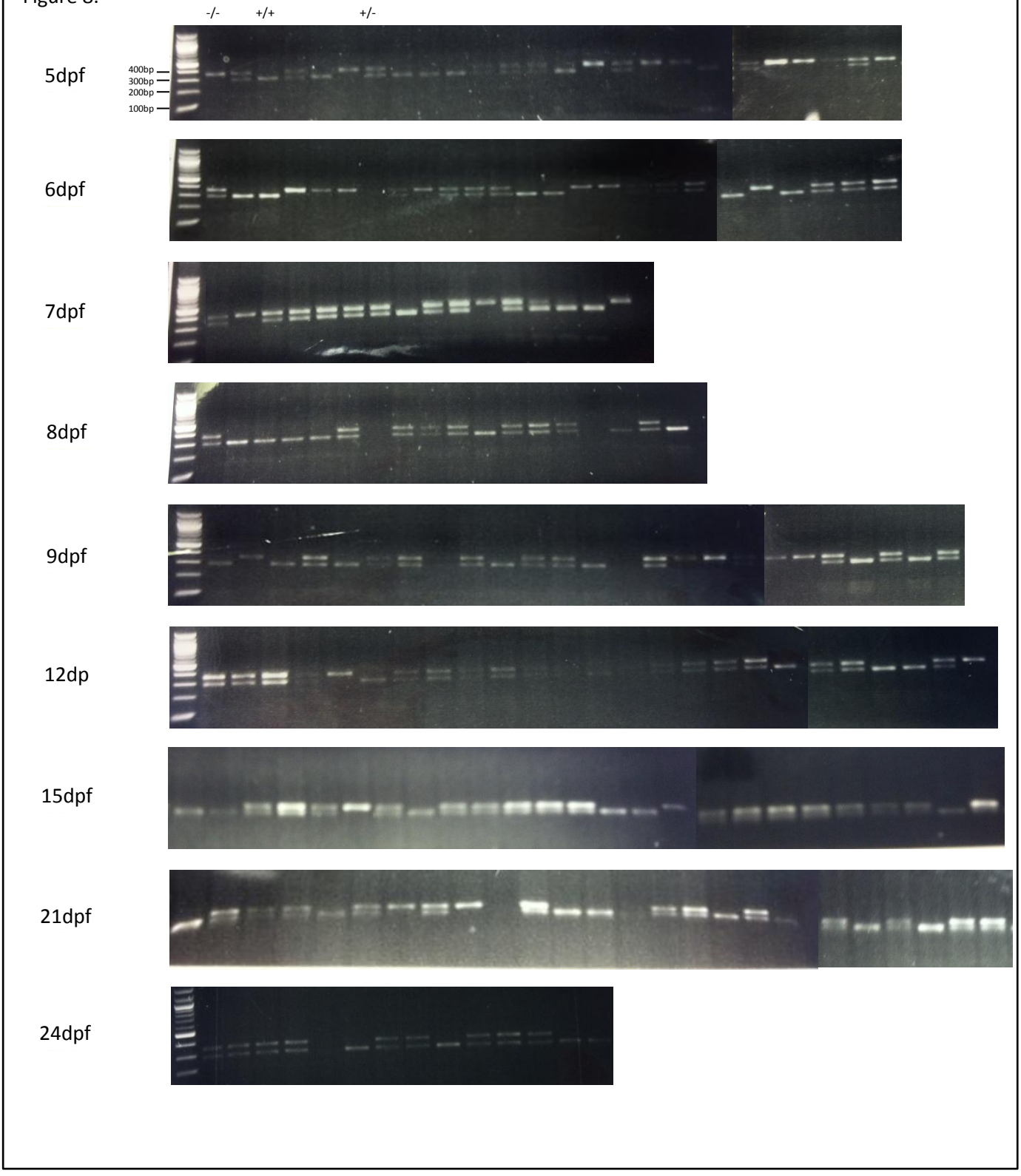


Figure 8. The collection of PCR pictures taken at different time points from 5dpf to 24dpf. Where possible 25 embryos were collected and ran in parallel with a 100bp DNA ladder. 15 and 21dpf are without a 100bp DNA ladder due to degradation of the ladder. The ladder is labelled in the gel from 5dpf for clarity. Homozygous mutant results are represented by a single larger band (342bp), this is due to the e36 point mutation occurring within the Sac1 restriction site. Heterozygotes therefore have two bands, the larger and the smaller band (276bp) due to a successful digest in the unaffected wild type allele. Homozygous wild type embryos will only contain this smaller band as both alleles are wild type and therefore digested by the Sac1 restriction enzyme. Three lanes from the 5dpf gel picture are labelled to show how the three possible genotypes are presented.

Discussion

In vivo time-lapse imaging in Fucci embryos

The use of the Fucci transgenic line provides us with a novel approach of visualising the effect of mutation. It is an easily applicable system to study cell cycle control, cellular and organism growth and therefore proliferation. In theory the system has the ability to be modified and applied to a wide area of study. This could be achieved by utilising the fundamental theories of the method and changing the genes expressed to widen the range or by altering the transgenic constructs to restrict expression in a tissue specific manner, an approach that could improve this particular study. In this study however we have shown this applicability of the already established Fucci system to studying the effects of FLCN in early development without modification. In addition, using statistical analysis, we can suggest a novel effect of the Sp2 morpholino on the development of the zebrafish embryo and in particular the cell cycle. An effect of FLCN on the cell cycle has always been a proposed idea due to the nature of mutations to manifest as uncontrolled cellular growth and the formation of potentially cancerous neoplasms (6, 7). Current knowledge points towards the idea that the disruption to cellular growth is due to the effect on the Akt-mTOR pathway (24). There could however be an alternative effect working in conjunction to produce the phenotypes observed. The reduction in signal intensity of green fluorescence in the Sp2 indicates a reduced number of cells in the S/G2/M stage of the cell cycle. This is confirmed by increased signal intensity, when compared to the mismatch injected embryo, of red fluorescence representative of cells within the G1 stage of the cell cycle. On this evidence a proposal can therefore be made that FLCN could potentially be involved in maintenance and regulation of the time cells spend in S/G2/M. This proposed loss of regulation seen within temporary *FLCN* mutants could help explain the uncontrolled cellular growth that characterises the symptoms of BHD. Recently though a group in Tokyo lead by Kawai (41) have elucidated at a role of FLCN in control of expression of cyclin D1. They found that levels of cyclin D1, encoded by the *CCND1* gene, were markedly increased in *FLCN* knockdown experiments in HeLA cells. They suggest that regulation occurs post-transcriptionally and may be in part due to miRNA or

RNA binding to the 3'-UTR of the *CCND1*. If experimentally reproducible their results provide an interesting idea to the effect of FLCN on cellular growth. Cyclin D1 functions in the G1/S phase transition, therefore an increased expression in cyclin D1 would lead to an increase in cells progressing through the cell cycle and hence an increase in cellular growth. In the context of this study this concept is interesting in the sense that if there was an increase in cyclin D1 expression you would therefore predict that majority of cells would be present in S/G2/M phases due to a shorter time spent in G1. This would be because levels of cyclin D1 would presumably have already reached a threshold needed for the transition and this idea could, in its own right, explain the uncontrolled cellular growth. It is in fact the inverse that is observed within this study though as a higher percentage, and therefore signal intensity, is accounted to cells in G1. The increased intensity from cells in G1 falls in line with the theory that a mutation in *FLCN* would decrease activation of AMPK which would subsequently mean a reduced level of inhibition of mTOR allowing for increased cell growth. This is because cells progressing through G1 can be said to be growing as opposed to proliferating. The complicated nature of FLCNs effect on cell cycle control means that this then is a key area that would benefit from further study to reproduce results and categorise an effect. This could be achieved in numerous ways, one of which would be the use of a quantitative method to study the protein levels of cyclin D1.

The use of morpholino based knock downs has allowed us to temporarily remove the function of FLCN and therefore study its effects. Morpholinos used in this way can be seen as both beneficial and detrimental to the overall study. Although they allow us a temporary control over imitating a mutation they are in essence not functionally the same as an initial heritable mutation, in particular one occurring in the C8 mutational hotspot, and a possible secondary loss of heterozygosity or "second hit". At first two initial morpholinos were designed for use, one of which, the Splice2 site directed morpholino, is used extensively in this study. Although the putative effect of this morpholino is a complete loss of function due to the retention of the 3392bp first intron this has not been fully confirmed and tested as western blots to do so are currently being optimised. FLCN

transduces its effects through the C-terminal binding of its binding proteins so there is an argument that the Splice 2 affected protein could potentially still be translated and function. This is highly improbable as the retention of such a large intron would most likely tag the mRNA for degradation and if translated it would cause a subsequent frameshift to the downstream coding sequence similar to what is observed in most patients. Nevertheless it cannot be yet ruled out so alternatively the use of a second morpholino, targeted to the splice donor site of exon 2, could be employed. The effect of this Splice1 morpholino, ATGACACTCCCCTCTCGCTCACCTC, is the retention of the second intron which contains a premature stop codon. It would therefore be interesting to repeat experiments previously performed with the Splice2 morpholino with the Splice1 morpholino to determine whether the same functional effects are observed.

Effect of *FLCN* mutation on the midbrain-hindbrain boundary

In this study we have managed to elucidate a novel effect of FLCN mutations that leads to variable staining within the mid-hindbrain boundary. Variability ranged from embryos showing a complete lack of staining to a pattern more similar to the staining observed in wild type and mismatch injected embryos (Sp2 2, Figure 7B). Although highly heterogeneous the effect could help explain the fluid retention phenotype observed in Sp2 injected embryos at 48hpf. A deregulation in cellular organisation has already been reported in developing murine embryos (33) and although this work is in a different context to the work published by Hasumi *et al* the idea does match relatively well with the observed fluid retention phenotype and the theorised effect on cell cycle control. Short term further work can help address this issue and determine whether a loss in the mid-hindbrain boundary can result in an increased volume of fluid around the midbrain and tectum. Relatively simple assays could be tested such as the injection of a soluble die or, alternatively, several midbrain-hindbrain markers such as fgf8, pax2 and eng2 (42) exist that could be tested for presence by *in situ* hybridisation in Sp2 injected embryos.

Homozygous lethality of FLCN knockout

Although a small data subset was used and homozygous mutant embryos were present up to and including 24dpf the percentage of $FLCN^{-/-}$ embryos present reduced considerably after 9dpf and strayed away from the expected 1:2:1 ratio. Without additional time points we cannot confirm or exclude an embryonic lethality of the homozygous null mutant. As previously mentioned embryonic lethality has been noted previously in mice embryos from E5.5-6.5 up to the point where no homozygous embryos were present post E9.5 (33). If embryonic lethality is conserved within zebrafish it does not occur before 24dpf, this indicates that there is a vast difference in the time point of lethality across organisms. At E9.5 mouse embryos contain a developing heart, brain as well as other developing organs such as the pancreas and liver. Inversely by 12dpf, the first stage at which a reduction in the number of homozygous mutant embryos is observed, zebrafish have full motility and digestion. Therefore if embryonic lethality is to be observed in both organisms the mutation must be taking effect at a different developmental stage and possibly by an alternative mechanism. It is therefore important to further our time points used and possibly try an alternative mutant allele to determine whether embryonic lethality in zebrafish is consistent with other model organisms.

In conclusion in this study we have shown the applicability of the zebrafish model organism to study the effects of *FLCN* mutations in early development. We have been able to suggest a possible functional role of FLCN within cell cycle regulation as well as regulated morphogenesis in the developing brain. In addition we have made progress to determining the embryonic lethality of a homozygous null mutation and shown the effectiveness and future potential adaptations of zebrafish in studying a complex inherited disorder such as BHD.

Acknowledgements

I would like to acknowledge all members of the Mueller lab for their continued help in and support throughout my project. In particular I would like to thank Harmeet for putting up with me throughout the project, Emma for helping me with all the technical stuff and Ferenc for his help with my write up.

1. Menko, F. H., van Steensel, M. A., Giraud, S., Friis-Hansen, L., Richard, S., Ungari, S., Nordenskjold, M., Hansen, T. V., Solly, J., and Maher, E. R. (2009) Birt-Hogg-Dube syndrome: diagnosis and management. *The lancet oncology* **10**, 1199-1206
2. Hornstein, O. P., and Knickenberg, M. (1975) Perifollicular fibromatosis cutis with polyps of the colon--a cutaneo-intestinal syndrome sui generis. *Archives for dermatological research. Archiv fur dermatologische Forschung* **253**, 161-175
3. Birt, A. R., Hogg, G. R., and Dube, W. J. (1977) Hereditary multiple fibrofolliculomas with trichodiscomas and acrochordons. *Archives of dermatology* **113**, 1674-1677
4. Adley, B. P., Smith, N. D., Nayar, R., and Yang, X. J. (2006) Birt-Hogg-Dube syndrome: clinicopathologic findings and genetic alterations. *Archives of pathology & laboratory medicine* **130**, 1865-1870
5. Toro, J. R., Glenn, G., Duray, P., Darling, T., Weirich, G., Zbar, B., Linehan, M., and Turner, M. L. (1999) Birt-Hogg-Dube syndrome: a novel marker of kidney neoplasia. *Archives of dermatology* **135**, 1195-1202
6. Pavlovich, C. P., Grubb, R. L., 3rd, Hurley, K., Glenn, G. M., Toro, J., Schmidt, L. S., Torres-Cabala, C., Merino, M. J., Zbar, B., Choyke, P., Walther, M. M., and Linehan, W. M. (2005) Evaluation and management of renal tumors in the Birt-Hogg-Dube syndrome. *The Journal of urology* **173**, 1482-1486
7. Zbar, B., Alvord, W. G., Glenn, G., Turner, M., Pavlovich, C. P., Schmidt, L., Walther, M., Choyke, P., Weirich, G., Hewitt, S. M., Duray, P., Gabril, F., Greenberg, C., Merino, M. J., Toro, J., and Linehan, W. M. (2002) Risk of renal and colonic neoplasms and spontaneous pneumothorax in the Birt-Hogg-Dube syndrome. *Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology* **11**, 393-400
8. Roth, J. S., Rabinowitz, A. D., Benson, M., and Grossman, M. E. (1993) Bilateral renal cell carcinoma in the Birt-Hogg-Dube syndrome. *Journal of the American Academy of Dermatology* **29**, 1055-1056
9. Toro, J. R., Pautler, S. E., Stewart, L., Glenn, G. M., Weinreich, M., Toure, O., Wei, M. H., Schmidt, L. S., Davis, L., Zbar, B., Choyke, P., Steinberg, S. M., Nguyen, D. M., and Linehan, W. M. (2007) Lung cysts, spontaneous pneumothorax, and genetic associations in 89 families with Birt-Hogg-Dube syndrome. *American journal of respiratory and critical care medicine* **175**, 1044-1053
10. Khoo, S. K., Giraud, S., Kahnoski, K., Chen, J., Motorna, O., Nickolov, R., Binet, O., Lambert, D., Friedel, J., Levy, R., Ferlicot, S., Wolkenstein, P., Hammel, P., Bergerheim, U., Hedblad, M. A., Bradley, M., Teh, B. T., Nordenskjold, M., and Richard, S. (2002) Clinical and genetic studies of Birt-Hogg-Dube syndrome. *Journal of medical genetics* **39**, 906-912
11. Gambichler, T., Wolter, M., Altmeyer, P., and Hoffman, K. (2000) Treatment of Birt-Hogg-Dube syndrome with erbium:YAG laser. *Journal of the American Academy of Dermatology* **43**, 856-858

12. Jamis-Dow, C. A., Choyke, P. L., Jennings, S. B., Linehan, W. M., Thakore, K. N., and Walther, M. M. (1996) Small (< or = 3-cm) renal masses: detection with CT versus US and pathologic correlation. *Radiology* **198**, 785-788
13. Herring, J. C., Enquist, E. G., Chernoff, A., Linehan, W. M., Choyke, P. L., and Walther, M. M. (2001) Parenchymal sparing surgery in patients with hereditary renal cell carcinoma: 10-year experience. *The Journal of urology* **165**, 777-781
14. Khoo, S. K., Bradley, M., Wong, F. K., Hedblad, M. A., Nordenskjold, M., and Teh, B. T. (2001) Birt-Hogg-Dube syndrome: mapping of a novel hereditary neoplasia gene to chromosome 17p12-q11.2. *Oncogene* **20**, 5239-5242
15. Schmidt, L. S., Warren, M. B., Nickerson, M. L., Weirich, G., Matrosova, V., Toro, J. R., Turner, M. L., Duray, P., Merino, M., Hewitt, S., Pavlovich, C. P., Glenn, G., Greenberg, C. R., Linehan, W. M., and Zbar, B. (2001) Birt-Hogg-Dube syndrome, a genodermatosis associated with spontaneous pneumothorax and kidney neoplasia, maps to chromosome 17p11.2. *American journal of human genetics* **69**, 876-882
16. Nickerson, M. L., Warren, M. B., Toro, J. R., Matrosova, V., Glenn, G., Turner, M. L., Duray, P., Merino, M., Choyke, P., Pavlovich, C. P., Sharma, N., Walther, M., Munroe, D., Hill, R., Maher, E., Greenberg, C., Lerman, M. I., Linehan, W. M., Zbar, B., and Schmidt, L. S. (2002) Mutations in a novel gene lead to kidney tumors, lung wall defects, and benign tumors of the hair follicle in patients with the Birt-Hogg-Dube syndrome. *Cancer cell* **2**, 157-164
17. Schmidt, L. S., Nickerson, M. L., Warren, M. B., Glenn, G. M., Toro, J. R., Merino, M. J., Turner, M. L., Choyke, P. L., Sharma, N., Peterson, J., Morrison, P., Maher, E. R., Walther, M. M., Zbar, B., and Linehan, W. M. (2005) Germline BHD-mutation spectrum and phenotype analysis of a large cohort of families with Birt-Hogg-Dube syndrome. *American journal of human genetics* **76**, 1023-1033
18. Toro, J. R., Wei, M. H., Glenn, G. M., Weinreich, M., Toure, O., Vocke, C., Turner, M., Choyke, P., Merino, M. J., Pinto, P. A., Steinberg, S. M., Schmidt, L. S., and Linehan, W. M. (2008) BHD mutations, clinical and molecular genetic investigations of Birt-Hogg-Dube syndrome: a new series of 50 families and a review of published reports. *Journal of medical genetics* **45**, 321-331
19. Lim, D. H., Rehal, P. K., Nahorski, M. S., Macdonald, F., Claessens, T., Van Geel, M., Gijezen, L., Gille, J. J., Giraud, S., Richard, S., van Steensel, M., Menko, F. H., and Maher, E. R. (2010) A new locus-specific database (LSDB) for mutations in the folliculin (FLCN) gene. *Human mutation* **31**, E1043-1051
20. Warren, M. B., Torres-Cabala, C. A., Turner, M. L., Merino, M. J., Matrosova, V. Y., Nickerson, M. L., Ma, W., Linehan, W. M., Zbar, B., and Schmidt, L. S. (2004) Expression of Birt-Hogg-Dube gene mRNA in normal and neoplastic human tissues. *Modern pathology : an official journal of the United States and Canadian Academy of Pathology, Inc* **17**, 998-1011
21. Vocke, C. D., Yang, Y., Pavlovich, C. P., Schmidt, L. S., Nickerson, M. L., Torres-Cabala, C. A., Merino, M. J., Walther, M. M., Zbar, B., and Linehan, W. M. (2005) High frequency of somatic frameshift BHD gene mutations in Birt-Hogg-Dube-associated renal tumors. *Journal of the National Cancer Institute* **97**, 931-935
22. van Steensel, M. A., Verstraeten, V. L., Frank, J., Kelleners-Smeets, N. W., Poblete-Gutierrez, P., Marcus-Soekarman, D., Bladergroen, R. S., Steijlen, P. M., and van Geel, M. (2007) Novel mutations in the BHD gene and absence of loss of heterozygosity in fibrofolliculomas of Birt-Hogg-Dube patients. *The Journal of investigative dermatology* **127**, 588-593
23. Singh, S. R., Zhen, W., Zheng, Z., Wang, H., Oh, S. W., Liu, W., Zbar, B., Schmidt, L. S., and Hou, S. X. (2006) The Drosophila homolog of the human tumor suppressor gene BHD interacts with the JAK-STAT and Dpp signaling pathways in regulating male germline stem cell maintenance. *Oncogene* **25**, 5933-5941
24. Baba, M., Hong, S. B., Sharma, N., Warren, M. B., Nickerson, M. L., Iwamatsu, A., Esposito, D., Gillette, W. K., Hopkins, R. F., 3rd, Hartley, J. L., Furihata, M., Oishi, S., Zhen, W., Burke, T.

R., Jr., Linehan, W. M., Schmidt, L. S., and Zbar, B. (2006) Folliculin encoded by the BHD gene interacts with a binding protein, FNIP1, and AMPK, and is involved in AMPK and mTOR signaling. *Proceedings of the National Academy of Sciences of the United States of America* **103**, 15552-15557

25. Hasumi, H., Baba, M., Hong, S. B., Hasumi, Y., Huang, Y., Yao, M., Valera, V. A., Linehan, W. M., and Schmidt, L. S. (2008) Identification and characterization of a novel folliculin-interacting protein FNIP2. *Gene* **415**, 60-67
26. Kimura, N., Tokunaga, C., Dalal, S., Richardson, C., Yoshino, K., Hara, K., Kemp, B. E., Witters, L. A., Mimura, O., and Yonezawa, K. (2003) A possible linkage between AMP-activated protein kinase (AMPK) and mammalian target of rapamycin (mTOR) signalling pathway. *Genes to cells : devoted to molecular & cellular mechanisms* **8**, 65-79
27. Shamji, A. F., Nghiem, P., and Schreiber, S. L. (2003) Integration of growth factor and nutrient signaling: implications for cancer biology. *Molecular cell* **12**, 271-280
28. Sarbassov, D. D., Ali, S. M., and Sabatini, D. M. (2005) Growing roles for the mTOR pathway. *Current opinion in cell biology* **17**, 596-603
29. Inoki, K., Corradetti, M. N., and Guan, K. L. (2005) Dysregulation of the TSC-mTOR pathway in human disease. *Nature genetics* **37**, 19-24
30. Shaw, R. J., Bardeesy, N., Manning, B. D., Lopez, L., Kosmatka, M., DePinho, R. A., and Cantley, L. C. (2004) The LKB1 tumor suppressor negatively regulates mTOR signaling. *Cancer cell* **6**, 91-99
31. Baba, M., Furihata, M., Hong, S. B., Tessarollo, L., Haines, D. C., Southon, E., Patel, V., Igarashi, P., Alvord, W. G., Leighty, R., Yao, M., Bernardo, M., Ileva, L., Choyke, P., Warren, M. B., Zbar, B., Linehan, W. M., and Schmidt, L. S. (2008) Kidney-targeted Birt-Hogg-Dube gene inactivation in a mouse model: Erk1/2 and Akt-mTOR activation, cell hyperproliferation, and polycystic kidneys. *Journal of the National Cancer Institute* **100**, 140-154
32. Chen, J., Futami, K., Petillo, D., Peng, J., Wang, P., Knol, J., Li, Y., Khoo, S. K., Huang, D., Qian, C. N., Zhao, P., Dykema, K., Zhang, R., Cao, B., Yang, X. J., Furge, K., Williams, B. O., and Teh, B. T. (2008) Deficiency of FLCN in mouse kidney led to development of polycystic kidneys and renal neoplasia. *PloS one* **3**, e3581
33. Hasumi, Y., Baba, M., Ajima, R., Hasumi, H., Valera, V. A., Klein, M. E., Haines, D. C., Merino, M. J., Hong, S. B., Yamaguchi, T. P., Schmidt, L. S., and Linehan, W. M. (2009) Homozygous loss of BHD causes early embryonic lethality and kidney tumor development with activation of mTORC1 and mTORC2. *Proceedings of the National Academy of Sciences of the United States of America* **106**, 18722-18727
34. Stirchak, E. P., Summerton, J. E., and Weller, D. D. (1989) Uncharged stereoregular nucleic acid analogs: 2. Morpholino nucleoside oligomers with carbamate internucleoside linkages. *Nucleic acids research* **17**, 6129-6141
35. Bill, B. R., Petzold, A. M., Clark, K. J., Schimmenti, L. A., and Ekker, S. C. (2009) A primer for morpholino use in zebrafish. *Zebrafish* **6**, 69-77
36. Sakaue-Sawano, A., Kurokawa, H., Morimura, T., Hanyu, A., Hama, H., Osawa, H., Kashiwagi, S., Fukami, K., Miyata, T., Miyoshi, H., Imamura, T., Ogawa, M., Masai, H., and Miyawaki, A. (2008) Visualizing spatiotemporal dynamics of multicellular cell-cycle progression. *Cell* **132**, 487-498
37. Sugiyama, M., Sakaue-Sawano, A., Iimura, T., Fukami, K., Kitaguchi, T., Kawakami, K., Okamoto, H., Higashijima, S., and Miyawaki, A. (2009) Illuminating cell-cycle progression in the developing zebrafish embryo. *Proceedings of the National Academy of Sciences of the United States of America* **106**, 20812-20817
38. Streisinger, G., Walker, C., Dower, N., Knauber, D., and Singer, F. (1981) Production of clones of homozygous diploid zebra fish (Brachydanio rerio). *Nature* **291**, 293-296

39. Kettleborough, R. N., Busch-Nentwich, E. M., Harvey, S. A., Dooley, C. M., de Bruijn, E., van Eeden, F., Sealy, I., White, R. J., Herd, C., Nijman, I. J., Fenyes, F., Mehroke, S., Scahill, C., Gibbons, R., Wali, N., Carruthers, S., Hall, A., Yen, J., Cuppen, E., and Stemple, D. L. (2013) A systematic genome-wide analysis of zebrafish protein-coding gene function. *Nature* **496**, 494-497
40. Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B., Tinevez, J. Y., White, D. J., Hartenstein, V., Eliceiri, K., Tomancak, P., and Cardona, A. (2012) Fiji: an open-source platform for biological-image analysis. *Nature methods* **9**, 676-682
41. Kawai, A., Kobayashi, T., and Hino, O. (2013) Folliculin regulates cyclin D1 expression through cis-acting elements in the 3' untranslated region of cyclin D1 mRNA. *International journal of oncology* **42**, 1597-1604
42. Rhinn, M., Lun, K., Ahrendt, R., Geffarth, M., and Brand, M. (2009) Zebrafish gbx1 refines the midbrain-hindbrain boundary border and mediates the Wnt8 posteriorization signal. *Neural development* **4**, 12

Genotyping and platelet phenotyping of cases of rare inherited platelet based bleeding disorders

By Ben Johnson

This project is submitted in partial fulfilment of the requirements for the award of the MRes.

UNIVERSITY OF BIRMINGHAM

Institute of Biomedical Research

School of Clinical and Experimental Medicine

University of Birmingham

May 2013

Abstract

Platelet based bleeding disorders are a rare subset of bleeding diathesis that can incorporate thrombocytopenias and platelet function defects. Both thrombocytopenias and platelet function defects present clinically with symptoms relating to hypocoagulability that often manifest as a variation of bleeding episodes. Platelet based bleeding disorders covers a wide variety of individual bleeding disorders with highly heterogeneous severity often due to the genetic component of disease involved. To date there is a large collection of inherited platelet based bleeding disorders incorporating a spectrum of genes with an elucidated role in platelet development or function. This study aims to progress work in this field by using a whole exome sequencing approach following the Genotyping and Platelet Phenotyping (GAPP) protocol as previously described (2). We focus our efforts on two patients families and determine novel candidate variations as potentially disease causing. These candidate variations offer a starting point for subsequent research which can help produce a novel diagnosis and determine new genes with an involvement in platelet functioning and development. In addition we screen a panel of thrombocytopenic patients for variations within the newly discovered 5'-UTR of *ANKRD26* (1) to try and diagnose previously unclassified diseases.

Word count - 193

Contents

List of Illustrations

Introduction

Bleeding diathesis is an unusual susceptibility to haemorrhaging upon injury. Disorders in this spectrum are due to a coagulopathy and in most cases a hypocoagulability. Common symptoms of diseases in this spectrum stem from a prolonged and/or increased rate of bleeding due to a decreased or improper coagulation at the site of injury. Clinically patients can present with complications such as disproportionate bruising, petechial and purpura in early childhood and extensive bleeding from cutaneous membranes, such as; epistaxis, menorrhagia and bleeding gingiva (reviewed in (3)). Hypocoagulability disorders can be split into three main sections; those affecting clotting factors, platelet function disorders (PFDs) or thrombocytopenias. This study will focus only around inherited disorders affecting platelet function and thrombocytopenias. Diagnostically the pattern of bleeding can help distinguish PFDs and thrombocytopenias from coagulation disorders but the different disorders share similar symptoms and phenotypes (2). PFDs and thrombocytopenias can be grouped together to be classed as platelet-based bleeding disorders. Inherited platelet-based bleeding disorders are rare and as such their population statistics are unknown but it is widely believed they are drastically under diagnosed emphasising the need for correct diagnostics to ensure disease management.

Platelets are small anucleate cells derived from megakaryocytes through megakaryopoiesis. In their development megakaryocytes follow a pathway consisting of differentiation from hematopoietic stem cells, endomitosis/nuclear polyploidisation and cytoplasm maturation, all of which prepare the cell for the formation of platelets (reviewed in (4)). Currently platelet biogenesis is thought to occur by the proplatelet model of formation which suggests development of megakaryocyte branches into the sinusoids of bone marrow where platelets are shed into circulation (5). This has been quite well evidentially supported but the fragmentation model, where megakaryocytes travel from the bone marrow to the lung where they are fragmented to form platelets (6), cannot be completely excluded. One of the key players in platelet formation is the growth factor thrombopoetin (TPO) and

its receptor *c-Mpl*. C-Mpl$^{-/-}$ knockout mice have shown an 85% reduction in platelet number (7) but thrombopoetin is only one of several proteins involved in platelet formation that can lead to thrombocytopenia.

Normal platelet count in whole blood is 150-450x10^9/l, such that anything below 150x10^9/l is classed as a thrombocytopenia (3). Platelet count is highly variable between individuals and even within affected family members meaning differing levels of severity can exist within patients. In addition patients with thrombocytopenia are often seen associated with other abnormalities. It is therefore difficult when classifying inherited thrombocytopenia and is important to avoid misdiagnosis as other conditions such as immune thrombocytopenia which may result in altered treatment. As a result various methods of classification currently exist trying to separate inherited thrombocytopenia on a number of criteria including; mode of inheritance, platelet size, genetic mutations and existing abnormalities (8, 9). An Italian based laboratory have also suggested a diagnostic algorithm to help diagnose and distinguish inherited thrombocytopenias (10).

Inherited thrombocytopenias

Current classified inherited thrombocytopenias are outline in Table 1.

Defects in transcription factors affecting early pathways of platelet biogenesis are one of the more frequent causes with mutations being observed in genes such as *GATA1* (11) and *RUNX1* (12). GATA1 and GATA2 are zing-finger transcription factors that bind to a common co-factor FOG1. Both genes are expressed in developing megakaryocytes and are presumed to have an overlapping function (13). GATA1 knockout mice display thrombocytopenia with deregulated megakaryocyte proliferation and cytoplasmic *maturaTtion* (14) and these phenotypic effects are similar to those observed in patients with a *GATA1* mutation. Mutations in *GATA1* play a role in several disorders, possibly due to its downstream effect on protein and DNA binding in both megakaryopoiesis and erythropoiesis. Mutations that affect binding to its co-factor FOG1 cause an X-linked

macrothrombocytopenia with severe bleeding (3) and some papers suggest a dyserythropoietic anaemia dependant on disruption of GATA1's DNA binding function (15).

RUNX1 is a member of the RUNT of family transcription factors. *RUNX1* mutations are most noted in forms of paediatric and adult leukaemia and monoallelic mutations can cause a familial platelet disorder with predisposition to acute myelogenous leukaemia (12). It is believed *RUNX1* mutations cause an effect due to a deregulation of MYH10 silencing meaning a switch from mitosis to endomitosis cannot occur (16). The result is an unaltered platelet size but a reduced number of mature megakaryocytes leading to a reduced platelet count.

Table 1.

Inherited thrombocytopenias
With small platelets (microthrombocytopenia)
Wiskott–Aldrich syndrome
With normal-sized platelets
Familial platelet disorder and predisposition to acute myelogenous leukaemia
Congenital amegakaryocytic thrombocytopenia
Amegakaryocytic thrombocytopenia with radio-ulnar synostosis
Thrombocytopenia with absent radii syndrome
ANKRD26-related thrombocytopenia (THC2)
With large platelets (macrothrombocytopenia)
Bernard–Soulier syndrome
Di-George Syndrome (velocardiofacial syndrome)
Platelet-type von Willebrand disease
Benign mediterranean macrothrombocytopenia
Dyserythropoietic anaemia with thrombocytopenia
X-linked thrombocytopenia with thalassemia
Paris–Trousseau –Jacobsen's syndrome
MYH9-related disease (May–Hegglin anomaly, Sebastian syndrome, Fechtner syndrome, Epstein syndrome)
Gray platelet syndrome
Montreal platelet syndrome
Macrothrombocytopenia with platelet expression of glycophorin A
ITGA2/ITGB3-related thrombocytopenia
TUBB-1 related macrothrombocytopenia
FLNA-related macrothrombocytopenia
Sitosterolaemia

Table 1. A list of inherited thrombocytopenias classified by size.

Other examples of transcription factors mutated in cases of thrombocytopenia include *FLI1* and *NF-E2*. Hemizygous loss of *FLI1* is thought to underlie Paris-Trousseau/Jacobsens syndrome (17). But there has been recent publications to oppose this or even suggest that a possible positional effect could be disease causing (18, 19). NF-E2 is believed to act independently of thrombopoetin and be involved in late platelet formation as null mice are shown to have no circulating platelets in knockout studies (20).

Cases of familial thrombocytopenia that do not affect transcription factors are also observed with a rare prevalence. Three of the most notable disorders are Wiskott-Aldrich syndrome (WAS), MYH-9 related disorders and congenital amegakarycotyic thrombocytopenia (CAMT).

WAS is a rare X-linked recessive disease classified by microthrombocytopenia, eczema and immunodeficiency. Symptoms of WAS were first reported by Wiskott in 1937 (21) and later by Aldrich in 1954 (22). It has an incidence of 4 live male births in 1 million and arises from defects in the *WAS* gene located at Xp11.22, with a high frequency of missense mutations in the two mutational hotspots in exon 2 and intron 6 (affecting a splice acceptor site) (23). Noticeably reduced expression of the *WAS* protein, WASP, in disease (24) causes an aberrant regulation of the cytoskeleton which gives rise to its clinical presentations (25). Complete absence of WASP expression has been noted in 54% of patients in a large multinational cohort. WAS is categorised by a severe thrombocytopenia with a platelet count varying between 5 and $50x10^9$/l and presentation of eczema shortly after birth (3).

MYH-9 related disorders encompass several syndromic and non-syndromic conditions some of which were previously referred to as May–Hegglin anomaly, Sebastian syndrome, Fechtner syndrome and Epstein syndrome. All syndromes now referred to as a MYH-9 related disorder have a deleterious mutation within the *MYH9* gene that encodes the non-muscle myosin II-A heavy chain (NMMHC-IIA) protein (26). A recent study has previously suggested that to date 44 mutations have been reported within the *MYH9* gene, with 79% occurring within six common residues (27), but it is worth noting that the clinical picture is varied both between patients and throughout a patient's life (28). Platelet

counts vary from <10x10^9/l to above 150x10^9/l and patients can present with a macrothrombocytopenia with a variation of other clinical symptoms that seems to follow a genotype-phenotype correlation (27).

CAMT is a rare recessively inherited disorder that presents with severe thrombocytopenia present at birth. This is associated with a complete absence of megakaryocytes and a susceptibility to lead to bone marrow failure and pancytopenia that can occur within the first few months of life (29, 30). Mutations in the thrombopoetin receptor gene *MPL* are the cause of CAMT and most mutations seem to affect the TPO binding domain or lead to a truncated protein abolishing its function (30). CAMT is unique in the sense that its onset is early and can present as a life threatening disorder due to complications affecting all aspects of haematopoiesis.

Platelet function disorders

Once platelets are formed and shed into the circulation they have an average lifespan of 10 days. Upon endothelium damage platelets undergo a well regulated series of responses from adhesion, to spreading and to aggregation and formation of a thrombus. Endothelial damage exposes subendothelial components such as collagen, fibronectin and Von-Willebran factor (VWF). Through receptors present on the platelet surface an initial interaction is made with VWF that aids in slowing the platelets sufficiently so static adhesion and then aggregation can occur. Activation of platelets occurs as platelets are recruited to the site of injury where agonists such as adenosine diphosphate (ADP) and thrombin stimulate internal signalling cascades through transmembrane receptors and G-proteins. These receptors can subsequently activate enzymes within cellular metabolism to alter processes in cytoskeleton organisation to allow for spreading, secretion of α-granules and formation of a procoagulant surface to allow for formation of a thrombus (31).

The majority of defects affecting platelet function occur either in receptors or signalling cascades involved in platelet activation or, alternatively, they are commonly found to affect processes involving granule secretion. However there does exist defects that are classed independently of

these two subgroups, these include; Glanzmann Thrombasthenia (GT), Bernard-Soulier syndrome and Scott Syndrome.

GT is a rare autosomal recessive disorder characterised by a deficiency of platelet GPIIB/IIIA from the fibrinogen receptor integrin αIIbβ3 (32). Mutations occur within the genes encoding GPIIB and GPIIIA, IGTA2B and ITGB3, respectively, and the current published mutations can be found in the mutation database: http://med.mssm.edu/glanzmanndb. Patients have a normal platelet count and can be classed into three sub-groups dependant on the percentage of their surface G-proteins (33). Clinical features have been summarised in a large cohort of 177 patients and consists of purpura, epistaxis, bleeding gingiva with presentation often before the age of five (34). αIIbβ3 is key in platelet aggregation by binding fibrinogen and other adhesive proteins allowing for platelet-platelet adhesion, so mutations that disrupt its function affect a platelets ability to adhere thus causing the bleeding phenotype observed (35).

Bernard-Soulier syndrome (BSS) is an interesting example of both a platelet function disorder and a thrombocytopenia. Clinically it presents as a potentially severe thrombocytopenia with large platelets and tendency for spontaneous bleeding (36). It is characterised by the absence of the major carbohydrate containing G –protein complex GPIb/IX/V. The first mutation was observed in *GP1bα* (37), a gene encoding part of a sub-unit of the complete complex, but since mutations have also been noted in *GP1bβ* and *GP9* affecting their respected translated proteins. BSS can be inherited in both a recessive and dominant pattern with a varied level of severity observed depending on mutation (38). Phenotypic effects arise due to the abolishment of the usual function of the GPIb/IX/V in binding Von-WIllebrand factor (39). Mutations in one protein of the complex result in low surface expression of the overall complex due to a tightly linked association and the need for all proteins to be present for complete biosynthesis (40).

Scott Syndrome differs from previously mentioned disorders due to its involvement with an alternate platelet surface component. First described by Weiss in 1979 (41) Scott Syndrome was recognised as a severe bleeding disorder but where complications seem to be confined to only

bleeding episodes, usually following injury or invasive procedure. The main attribute of Scott Syndrome is a defect in membrane lipid scrambling disrupting the asymmetrical distribution of lipids, including phosphatidylserine (PS), across the plasma membrane (42). PS exposure on the exoplasmic surface of the plasma membrane is hallmark of apoptosis (43) but within platelets it functions to create the prothrombinase complex to allow for conversion of prothrombin to thrombin (44). This subsequently activates and promotes a localised procoagulant activity within platelets. Normal lipid mobility is achieved through the action of a calcium-dependant scramblase (45) and several studies have previously tried to determine the encoding gene of this transport protein. *ABCA1* was first suggested as a candidate (46) but *TMEM16F* is a more recent suggestion thought to now be the causative gene encoding the defected scramblase in Scott Syndrome (45).

Interestingly there exists a syndrome thought to be the inverse of Scott Syndrome named Stormorken Syndrome. First described in 1985 (47) patients have a low platelet count , bleeding and spontaneous platelet activation and aggregation with a full exposure of surface PS (48, 49).

Disorders that can be classed as affecting platelet receptors and signal transduction pathways are an ill-defined group that show inhibition of platelet activation to one or more agonist. Platelet count is usually normal but aggregation to one or more agonists is reduced or deaggregation is observed. Defects have currently been observed in the collagen receptor GPVI, where no response to collagen could be observed (50). As well as in the adenosine diphosphate (ADP) receptors P2Y12 and P2Y1 (50, 51) and the ATP receptor and ligand-gated ion-channel P2X1 (52), both of which affect platelet activation. In addition defects have been noted in other platelet surface receptors such as the prothrombin receptor Thromboxane A2 (TXA_2) (53).

Patients with defects attributed to secretion can be further grouped into two categories; those that involve platelet granule deficiencies and those that can be considered primary secretion defects. It is worth noting that primary secretion defect disorders are a diminishing group of disorders that cannot otherwise be classified. Defects are seen in two types of secretory granules that platelets contain; dense core and α-granules and disorders are associated with a deficiency of either one or

both combined. Both dense core and α-granules contain molecules thought to act in a paracrine way to promote local aggregation of surrounding platelets. Dense core granules contain smaller molecules such as ADP whereas α-granules contain the larger VWF and PF4 (54). A reduction in these granules therefore has a detrimental impact in platelet aggregation and often a secondary wave usually caused by secretion is not noticed in aggregation tests. To date several disorders have been reported affecting both types of granules. Disorders that only affect dense core granules include: Hermansky-Pudlak Syndrome (HPS) (55), which affects *HPS* genes 1-9 and several other loci (56), and Chediak- Hagashi Syndrome (CHS), caused by mutations in *LYST* which is thought to play a role in vesicle trafficking (57). Both syndromes share similar clinical presentations and can be classed together as dense granule storage pool diseases (δ-SPD) but they remain functionally distinct. Alpha granule deficiencies or α-SPD present clinically with differences to those involving dense core granules with symptoms like hypopigmentation and albinism confined to HPS and CHS (58). Alpha granule deficiencies include Grey Platelet syndrome (GPS), a syndrome showing both recessive and dominant inheritance patterns and named due to the complete absence of alpha granules in platelets causing a greyish colour in peripheral blood smears (59, 60). Recent publications have suggested the molecular cause of GPS as a BEACH (Beige and Chédiak-Higashi) domain containing protein involved in vesicle trafficking known as *NBEAL2* (61). Other alpha granule deficiencies include Quebec platelet syndrome (62), affecting factor V and Arthrogryposis-renal dysfunction-cholestasis (ARC) where mutations in the gene *VPS33B* is the known cause (63).

Approach to the diagnosis of platelet based bleeding disorders and the GAPP study

Upon suspicion of a platelet function disorder it is ideal to phenotype the patient to help elucidate a possible cause. To date the "gold standard" of testing platelet activation is *ex vivo* light transmission aggregometry (LTA). First described by Born in 1962 (64) LTA is a fairly inexpensive test utilising patient platelet-rich-plasma (PRP) or wash-platelets to determine aggregation in response to a selection of agonists by monitoring light transmission. LTA has the ability to inform you of details

about aggregation within a patient and can help hint at a deficiency by observation of the effects due to different agonists. Although functionally simple the test is not robust and flaws arise due to a wide variation in controls that can overlap with patients (65) and also due to a non-standard practice across institutions (66).

Many other platelet function tests are also available and beneficial in diagnosis such as flow cytometry (reviewed in (67)), platelet function analyser 100 (PFA-100) (68) and more recently several parallel functional tests have arisen as a more efficient means of testing. One such platform is the 96 well plate aggregation assay: *Optimul* (69). Although providing a high throughput system there seems to be some discrepancy with a reduction in aggregation to arachidonic acid on *Optimul* (70). In combination with a thorough medical and family history platelet function testing can be ideal to diagnose certain sub-sections of platelet based bleeding disorders. Diagnostics becomes more of a challenge however when patients present with a novel or more complex phenotype, for example, a combined thrombocytopenia hinted by a low platelet count and an observable deaggregation to certain agonists.

Correct classification of disease is paramount for two main reasons; genetic counselling and treatment. In conjunction with the consultant counselling can help affected patients understand the inheritance patterns of their disease and susceptibility to a couple's offspring. This additional information from a thorough diagnosis can alter a patient's lifestyle and choices allowing them to cope and manager their disease more effectively. Current treatment for inherited platelet-based bleeding disorders varies widely. A lot of treatment is based around prevention of bleeding episodes through correct management of disease. This can materialise as day to day caution in extremely severe cases to correctly adjusted protocols for invasive routine procedures such as dental treatment and childbirth. The most common and effective form of treatment of symptoms currently is platelet transfusion but complications can arise due to the development of alloimmune antibodies against HLA and αIIβIII (71), however molecular treatments do exist as an alternative and include;

- Antifibrinolytic agents. (E.g. tranexamic acid) Benificial for prevention of subcutaneous bleeds and can be administered orally (3).
- Desmopressin. A synthetic analogue of the hormonal diuretic vasopressin first used for treatment of Von Willebrand disease and haemophilia. It functions by increasing plasma levels of VWF, Factor VIII and tissue plasminogen activator and has been shown to shorten bleeding time in dense granule storage pool diseases and signal transduction disorders and possibly BSS and HPS but with some ambiguity (72).
- Recomibinant Factor VIIa. First used for the successful treatment of epistaxis in a two year old with GT (73) rFVIIa is an alternative to platelet transfusion in patients with GT, although variable success is observed (71).

Also worth noting is the use of haematopoietic stem cell transplantation which has been shown to be effective in GT (74) and BSS (75) but arguably most importantly is the innovation into gene therapy. Preliminary work has suggested the idea that gene therapy may be a feasible form of treatment especially in GT patients with transducible αIIβIII (76, 77). This is a potentially exciting field and fortifies the need for accurate and extensive diagnostics to provide the correct treatment on an individual basis in the near future.

Genotyping of patients is not uncommon and several previous studies have taken a genome wide approach to studying possible causative and mutated genes (78, 79). Many of these GWAS have produced a vast list of genes with SNPs in genes involved, or with putative functions, in platelet functioning or count but many have no credible link to disease. Definitive genotyping in platelets is also complicated by diseases such as the previously mentioned HPS that has several possible disease causing genes.

The Genotyping and Phenotyping Project (GAPP)

The genotype and platelet phenotyping (GAPP) consortium takes a novel approach to, in particular, genotypic diagnostics which is outlined in Figure 1 (2). Through collaboration the consortium aims to achieve phenotypic and genotypic categorisation of patients entered into the patient bank using a standardised set of protocols. Each patient is considered individually and screened using a next generation whole exome sequencing approach which is in addition to more classic platelet phenotyping performed prior. The data from whole exome sequencing provides a platform from which a panel of candidate genes can be produced. Candidate genes can then be tested for segregation and the impact of mutation analysed. This approach is not without its shortcomings though as only the coding regions and a few immediate adjacent bases are sequenced and it cannot determine imprinting defects which may play a role in disease (80). What it can achieve, though, is a complete categorisation of an individual's disease which may potentially show classification of a previously known mutant or the discovery of a novel change. These novel changes then provide us with the opportunity to achieve novel insights into molecular functions within platelets and this has been illustrated recently by the discovery of a patient with a mutation in the PDZ binding domain in the C-terminus of *P2Y12* which shows an inability to recycle the receptor (81).

Figure 1.

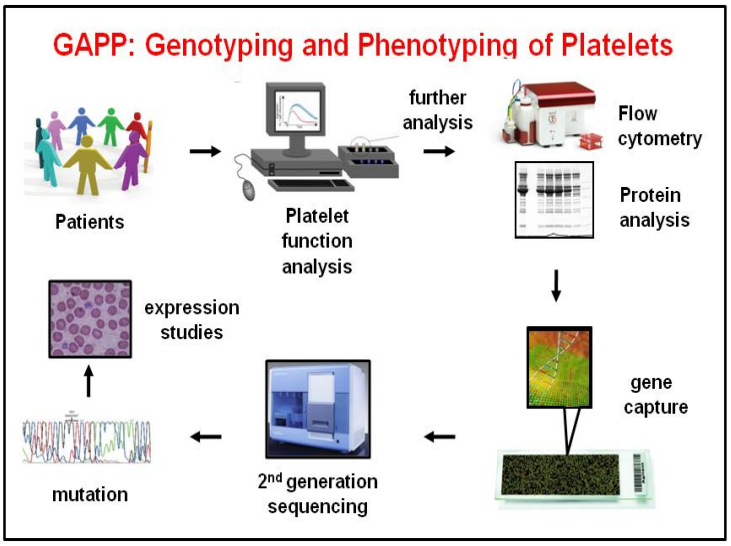


Figure 1. The GAPP work flow approach to diagnosis and prognosis of inherited platelet based bleeding disorders. Areas focused on in this study will be from the 2nd generation sequencing onwards and will incorporate previous results suggested by platelet function analysis.

Variations in the 5'-UTR of *ANKRD26* as the causative mutation underlying Thrombocytopenia 2 (THC2)

Ankirin Repeat domain 26 is a protein of unknown function thought to be involved in protein-protein interaction due to the ankyrin repeat domains. Recently mutations in the 5'-UTR of *ANKRD26* have been reported in patients thought to be suffering from a rare dominant form of thrombocytopenia, THC2 (MIM 188000). This gene maps to 10p12.1, the location of the previously suggested causative genes for THC2 (82, 83), and to date 12 single nucleotide heterozygous variations have been reported in a 21 base stretch of the 5'-UTR from c.-113 to c.-134. So far these mutations have been observed in 78 patients from 21 families presenting with the clinical symptoms of THC2 (1, 84). Partial inactivation of *ANKRD26* in mouse models leads to obesity and insulin resistance whereas platelet count is normal (85) leading to suggest that a gain of function mutation may be responsible for the thrombocytopenia observed. It can be said then that this is possibly due to a deficiency in formation of platelets as megakaryocyte numbers are normal in preliminary data (1). The reported mutations so far in the 5'-UTR of *ANKRD26* in patients suggest a higher incidence and involvement of this gene and THC2 than first expected. It is because of this that we will therefore screen a panel of patients with thrombocytopenia for mutations in the 5'-UTR of *ANKRD26* prior to whole exome sequencing to identify novel disease-causing genes. Patients within the panel have varied levels of thrombocytopenia from mild to severe, consistent with previously reported cases, meaning it is therefore worthwhile to determine the prevalence of *ANKRD26* as a causative gene in these unclassified cases.

Aims of this Project

In this study we plan to adopt the approach of the GAPP consortium and apply it to two sets of patients. One family has a single affected individual and the other is a large kindred with three generations of affected family members. Patients will follow the protocol outlined in (2) with the aim of elucidating a possible causative mutation in both patient sets. Any novel mutations will be

analysed with the aim to determine a functional role of the gene in platelet functioning/formation and the subsequent detrimental effect of the mutation.

In addition we plan to screen a previously undiagnosed thrombocytopenic panel of patients for mutations in the 5'-UTR of *ANKRD26* in the hope of possibly classifying unknown patients with THC2.

Therefore our aims for this study are two-fold;

- Use whole exome sequencing data to determine a causative mutation and defective gene or genes in two unrelated cases of inherited platelet based bleeding disorders.
- Screen a thrombocytopenic panel of patients for mutations in the 5'-UTR of *ANKRD26* to determine frequency of mutations in a previously reported mutational hotspot.

Overall through these outlined aims above we hope to further our understanding in rare, so far, unclassified cases of inherited platelet based bleeding disorders. We therefore aim that through standardised protocols we can elucidate a molecular cause in these patients and when presented with novel genes determine their related molecular functions within platelets to aid in disease management and treatment.

Methods

Patient recruitment

All patients and healthy controls within this study were entered into the GAPP study and are therefore covered under the GAPP study ethical license (NIHR portfolio status in the non-malignant haematology subgroup (ID 9858); REC no 06/MRE07/36). All patients and healthy controls willingly volunteered for participation within this study and have fully consented to the extent of the study and its experimental procedures. All patient personal information has been kept confidential and anonymous throughout the study. More information regarding the ethical approval and licensing of the GAPP study can be requested from Professor Steve Watson, Centre for Cardiovascular Sciences, Institute of Biomedical Research, College of Medical and Dental Sciences, University of Birmingham, Birmingham, B15 2TT, UK.

Preparation of patient blood and DNA extraction

Patient blood was taken in clinic local to the patients in question or at their home. Blood samples were taken following GAPP guidelines which include no samples being taken recently after blood/platelet transfusion. DNA, present within a white cell layer known as the "buffy-coat", and Platelet rich plasma (PRP) were extracted from fresh patient whole blood using an in-house preparation by centrifugation. The white cell rich buffy-coat was stored at -80°C while PRP was tested for concentration using a coulter particle count and size analyser and subsequently used fresh in aggregation assays or adjusted to a platelet concentration of $5x10^8$/ml in 1x SDS containing sample buffer (50mM Tris-HCl [ph 6.8], 2% SDS, 10% glycerol, 1% β-mercaptoethanol, 12.5mM EDTA and 0.02% bromophenol blue) for use in protein studies and western blotting.

DNA was extracted from buffy-coat using the Gentra Puregene blood kit available from Qiagen™ (www.qiagen.com) and the included protocol sheet for DNA preparation from buffy-coat. The method relies on a spin protocol to remove cell debris and DNA is precipitated using isopropanol and

washed with ethanol. The concentration of extracted hydrated DNA was determined using a Nanodrop™ 2000 Spectrophotometer available from Thermo Scientific. 2µl of extracted DNA was tested and concentration (ng/µl), 260/280 and 280/230 values were recorded and the sample stored and logged within the on-site GAPP patient database for further testing.

Polymerase chain reaction (PCR)

PCR was performed a using standardised in house protocol on extracted patient and control DNA at a concentration of 20ng/µl. Where a concentration of 20ng/µl was not available DNA was used neat. Oligonucleotide primers were designed using Primer3 sofware and validated using In-silico PCR software available at http://genome.ucsc.edu/cgi-bin/hgPcr. All oligonucleotides are available from www.sigmaaldrich.com (see attached appendix for specific sequences). PCR followed the standard cyclical protocol of denaturing, hybridisation and annealing as follows;

Figure 2.

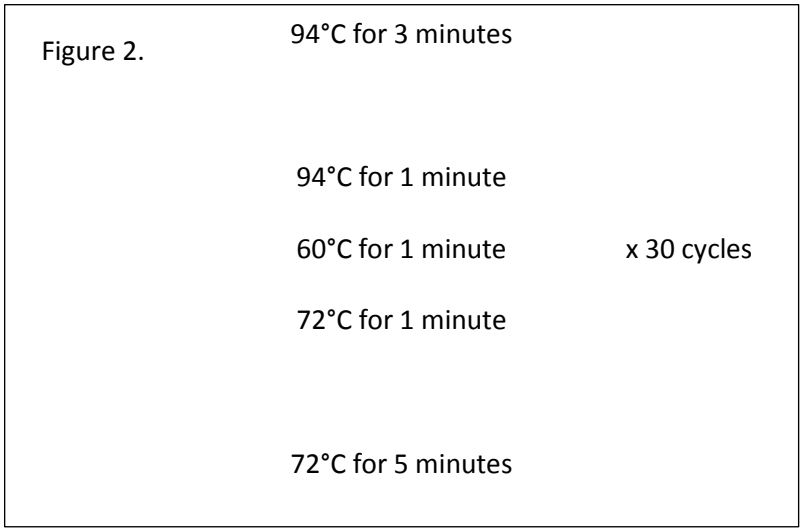


Figure 2. Cyclic PCR parameters.

PCR reactions were performed using RedTaq® Readymix™ PCR reaction mix (20 mM Tris-HCl, pH 8.3, 100 mM KCl, 3 mM MgCl2, 0.002 % gelatin, 0.4 mM dNTP mix (dATP, dCTP, dGTP, TTP), stabilizers and 0.06 unit/mL of Taq DNA Polymerase) also available from Sigma-Aldrich (http://www.sigmaaldrich.com/catalog/product/sigma/r2523?lang=en&region=GB). PCR reactions

were performed in 25µl reactions containing 20ng of genomic DNA, 2x Sigma RedTaq Readymix and 5.0 µmol of each primer. PCR reaction was performed in parallel in 96-well plates and performed in Bio-rad Tetrad thermal cyclers. All primers were tested initially on two unaffected healthy controls, upon failure annealing temperature was varied along with the inclusion of 5µl of GC-rich buffer to determine a successful parameter. All final PCR reactions both for sequencing and other tests were performed with an annealing temperature of 60°C and without GC-rich buffer. All primers were added at a concentration of 10µM and a 0.5µl volume per reaction. PCR products were visualised using ethedium bromide staining on a 1% agarose TAE gel and viewed using a standard ultraviolet transilluminator imaging system. All gels were also loaded with 2µl of 1kb DNA ladder (NEB) and a negative dH_20 control.

Sequencing

Whole exome sequencing was performed externally with collaborators in Imperial College London. The SureSelect human AllExon 50Mb kit (Agilent Technologies) was used and sequencing was performed on the HiSeq 2000 (Illumina) with 100bp paired-end reads. The sequences were aligned to the reference genome (hg19) with Novoalign (Novocraft Technologies) that aligns based upon Needleman-Wunsch algorithm with affine gap penalties. Duplicate reads and reads mapping to multiple locations were excluded from downstream analysis. The SamTools software package and in-house software tools were used to identify and quality filter single nucleotide substitutions and small insertions/deletions. All calls with a read coverage of <4 were excluded. Novelty of variants was determined by comparison to dbSNP137 and 1000 Genomes variant calls and variants previously indentified in >600 in house control exomes sequenced and analysed using the same protocol as described above.

Sanger sequencing was performed in house to verify candidate mutations and determine segregation using the list of PCR primers shown in the appendix. Successful PCR products were cleaned and concentrated using the *micro*CLEAN solution (available from www.microzone.co.uk),

applied in equal volumes and the protocol designed for 96-well eppendorf plates was followed. The purified PCR product was sequenced in both forward and reverse directions using the relevant primers (2pmol/µl) and the BigDye® Terminator v3.1 Cycle Sequencing Kit, available from Applied Biosystems®.

PCR cycling conditions were:-

96°C for 30 seconds

50°C for 15 seconds x30

60°C for 4 minutes

Conditions were used to allow for optimal elongation and random termination to provide all lengths of fragments necessary for good coverage sequencing. The EDTA method of precipitation was used for sequencing reaction clean up and sequencing was performed on an ABI 3730 automated sequencer using a capillary sequencing system. Sequencing results were analysed using MutationSurveyor® software from Soft Genetics®. When no match could be made, of the sequence to one within the enclosed database, sequences were analysed using Chromas sequence chromatogram viewer. A minimum 20 fold coverage was used as a quality cut off. All sequencing was repeated at least once more for clarity and more times in the case of a failed sample.

The same sequencing protocol was applied for a screen of the whole coding region of TMEM30B, using the three primer pairs displayed in the appendix (TMEM30B 1,2 and 3 F and R), in a panel of potential patients.

Results

Characterisation of a novel gene defect in a family with an inherited thrombocytopenia (Family 1)

Clinical features of bleeding history in family 1

A 31 year old female first presented to a haematological clinic with a history of cutaneous bleeding, prolonged bleeding (after minor wounds and dental treatment), menorrhagia, postpartum haemorrhaging and spontaneous muscle haematomas. The patient had previously been treated for their symptoms using platelet and red blood cell transfusions, antifibrinolytics, utering packing after postpartum haemorrhaging and iron therapy. An extensive family history was taken from the index patient, who will hence forth be known as patient 4:3, and a family pedigree was produced indicating 8 affected family members across three generations which can be seen in Figure 3. All patients presented with a variation of bleeding diathesis similar to the symptoms previously described in the index case. Platelet counts, tested within the haematological referring centre, varied between $74x10^9$ and $140x10^9$ indicating a mild to moderate thrombocytopenia throughout the family. Bleeding was disproportionate to platelet count within the index case indicating a possible secondary platelet function effect.

Platelet function testing in family 1

Platelet phenotyping was performed on patient PRP in-house by Dr Marie Lordkipanidzé and Dr Gillian Lowe using LTA and flow-cytometry following the guidelines as set out within the GAPP project (2). Five patients, indicated by green arrows in Figure3, were subjected to platelet phenotyping. The following agonsists were used in LTA; ADP, Collagen, Adrenaline, Arachidonic acid, Thrombin receptor activating proteins (PAR-1 petide and PAR-4 peptide) and risctocetin. All patients shared similar results of deaggregation at 10 and 30µM ADP, primary wave only at 10 and 30µM collagen, deaggregation at PAR-1 100µM and shape change at collagen 1µg/ml. In addition tested patients showed an undetectable level of ATP secretion with 100µM PAR-1 and ATP secretion on the lower levels of normal (0.75nmol/$1x10^8$ platelets) with PAR-4 peptide 500µM. Flow cytometry in

whole blood also supported a defect in PAR-1 or its subsequent signalling. Aggregation traces for ADP 10µM and PAR-1 peptide 100µM in patient 4:1 are shown in Figure 4 as a reference.

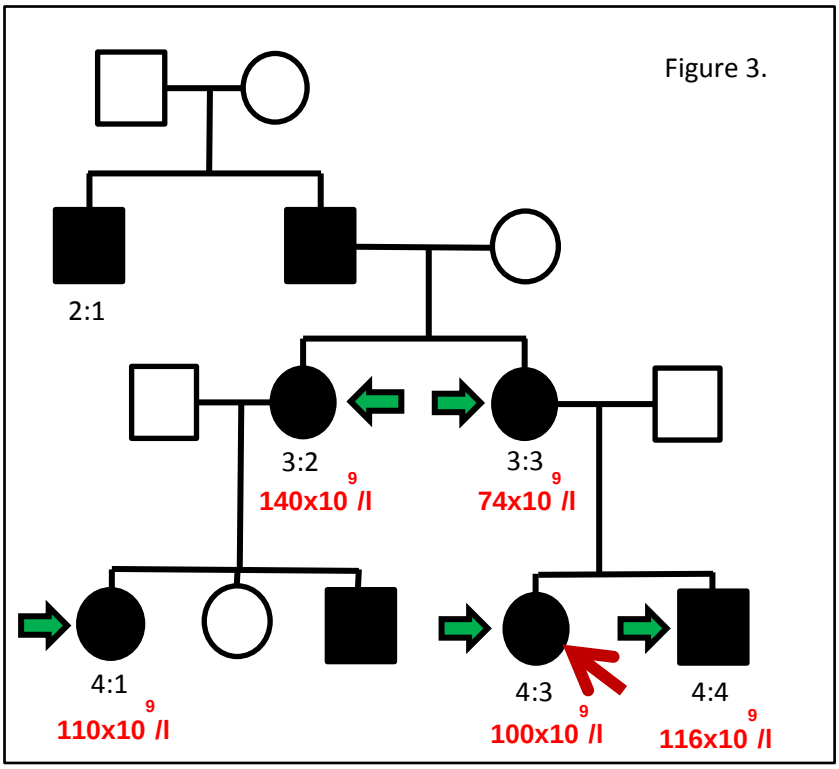


Figure 3 shows the family pedigree produced from the family history in family 1. Males and females shaded in black are affected members of the family and those unshaded are unaffected members. Patient samples were available for the 6 patients labelled. The index case is marked by a red arrow. Platelet counts are shown in red and available for five patients. The same five patients have also received platelet function testing and this is indicated by the green arrows.

Figure 4.

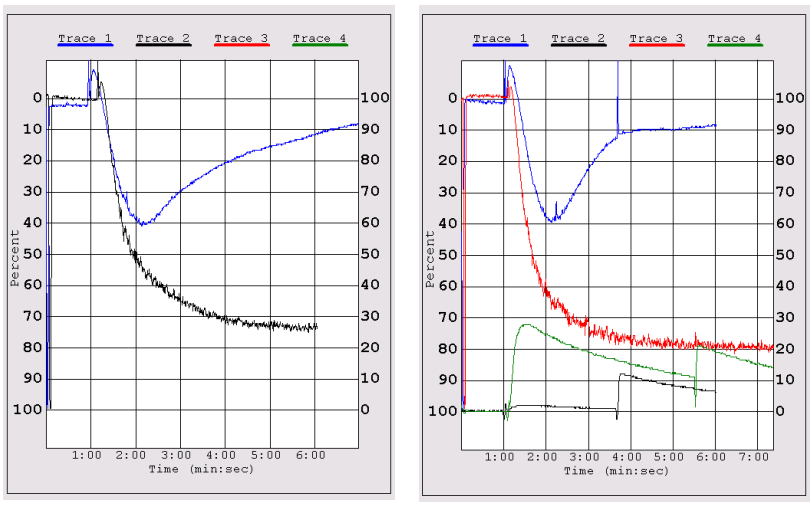


Figure 4 shows two LTA aggregation traces from patient 4:1. In the ADP 10µM trace the black line represents a control (PRP from an unaffected, unrelated individual) trace taken on the same day and the blue line represents the patients trace. As you can see deaggregation occurs at this ADP concentration which is shown by a reduction in percentage (of light transmission) in the blue line after the initial increase. The PAR-1 100µM trace uses the following key; blue line – patient aggregation, red line – control aggregation, black line - patient ATP secretion and green line – control secretion. Once again deaggregation to the agonist is observed and also noteable is the lack of initial secretion within the patient.

Whole exome sequencing in family 1

Whole exome sequencing was performed on three patients from family 1; individuals 4:3, 3:3 and 4:1. 389 novel variations were found within the three patients after analysis and comparison to reference databases; dbSNP137, 1000 Genomes and our in house database of >600 whole exomes. Of these 389 eight novel variants were shared across all three patients, four of which were non-synonymous amino acid changes or insertions/deletions. The eight mutations are shown in table 2. Familial segregation of the four remaining variants was undertaken using traditional Sanger sequencing with all the available affected family members. Primers flanking the region of the variants in the four genes were PCR-amplified and sequenced in all six patients and an unaffected related control, sibling of 4.1. The only variant that could be excluded at this stage was in *TOR2A* on

the basis that the mutation was not observed in all family members screened. The results from the Sanger sequencing and the presence or absence of the novel variant in each gene are summarised in Table 2.

Table 2.

	Patient			Patient 3:3	Patient 4:3	Patient 4:1	Patient 3:2	Patient 4:4	Patient 2:1	Unaffected
Novel variations share between all three patients	**Non-synonymous/insertions/deletions**	**wild type allele**	**mutation**							
CYP39A1: c.G627A: p.E209E										
DST: c.A12063G: p.T4021T										
FURIN: c.A1941T: p.S647S										
DNAJB12: c.G39A: p.R13R										
TOR2A: c.A1C: p.M1L	TOR2A: c.A1C: p.M1L	**A**	**C**	A	A/C	A/C	A	A/C	-	-
TMEM30B: c.778delC: p.P260fs	**TMEM30B: c.778delC: p.P260fs**	**C**	**-**	C/-	C/-	C/-	C/-	C/-	C/-	C/-
NEMF: c.C2884T: p.H962Y	**NEMF: c.C2884T: p.H962Y**	**C**	**T**	C/T	C/T	C/T	C/T	C/T	C/T	C
SLFN14 c.T659A: p.V220D	**SLFN14 c.T659A: p.V220D**	**T**	**A**	T/A	T/A	T/A	T/A	T/A	T/A	T

Table 2 shows a combination of both the whole exome sequencing results and those from Sanger sequencing. The four novel genes shared between all three patients that are either an insertion or deletion or cause a non-synonymous amino acid change are TOR2A, TMEM30B, NEMF and SLFN1. The candidate mutations and the effect to the protein sequence are also presented in the table. Sanger sequencing was used for confirmation of the initially suggested mutation and to determine presence across the affected family. TOR2A can therefore be secluded on these premises as the mutation is not confirmed in 3:3 by Sanger sequencing and it fails to be present in all family members negating it as a possible causative mutation. Therefore the three candidate mutations in TMEM30B, NEMF and SLFN14 cannot be narrowed down further at this stage. TMEM30B is not excluded due to the fact that the unaffected family member had no supplementary platelet phenotyping or any platelet counts. Therefore although the family member is thought to be unaffected we cannot confirm this so for all remaining purposes TMEM30B will continue to be included in further study.

The result of Sanger sequences was the production of three candidate mutations as a possible cause for disease within the affected family.

<u>The three remaining candidate genes</u>

Transmembrane protein 30B (*TMEM30B*), alternatively *CDC50B*, maps to the reverse strand at chromosome 14q23.1. The main protein coding transcript (Ensembl ID - ENST00000555868) encodes a one exon, 4471bp, 351AA sequence expressed in most endogenous tissues according to BioGPS microarray data (http://biogps.org). As of yet no function has been elucidated for TMEM30B but the

discovered mutation causes a frame shift and is therefore predicted to have a high impact on the original protein structure.

Nuclear export mediated factor (*NEMF*) is a protein with predicted function in nuclear export due to the domain contained within the N terminal. NEMF also maps to chromosome 14 but at position q21.3. There are seven predicted protein coding transcripts, the largest of which is a 5038bp transcript encoding a 1076 amino acid protein sequence. The largest transcript has 33 coding exons and the mutation found within family 1 affects exon 28 causing a non-synonymous histidine to tyrosine change.

Schlafen family member 14 (*SLFN14*) is a gene with a transcript length of 2889bp and a translation length of 912 amino acids. It is part of the larger Schlafen family of proteins and has an unknown protein function. A valine to aspartic acid change due to a T to A substitution in exon 1 of 4 is the mutation found within all available affected members of family 1.

All three candidate mutations are expressed and translated within platelets (MOPED and PaxDb).

The three candidate mutations suggested by Sanger sequencing were initially tested using MutationTaster online mutational predicting software (www.mutationtaster.org). Predictions are based upon evolutionary conservation, splice-site changes, loss of protein features and effects to amount of mRNA and a score is produced using a naïve Bayes classifier (86). This approach was utilised as an initial method to decide which mutation to pursue further in the first instance. TMEM30B was the only mutation out of the three predicted to be disease causing with a probability score of 1 so we will therefore look to expand our further work on this gene initially.

Sanger sequencing of a panel of patients with thrombocytopenia for a TMEM30B mutation

Mutation screening was performed using Sanger sequencing and three primer pairs spanning the entire coding region of *TMEM30B.* Unclassified patients from the GAPP study were selected according to a similar phenotype as patients within family 1. 21 patients were screened and no novel variations or patients containing the c.778delC heterozygous mutation within family 1 were found. Two previously reported heterozygous SNPs were observed within the subset, c.272T>C and

c.222C>T, were found in both homozygous and heterozygous states within patients. c.272T>C (rs137950125) is a synonymous variant found in 98% of the population in the database of the 1000 genomes project. c.222C>T is also a synonymous variant found at a frequency of 21% in the 1000 genomes project. As both variations have been previously reported synonymous (do not change the amino acid) and are not thought to occur in the two functional transmembrane domains they can therefore be excluded as being potentially disease causing.

Investigation of a novel gene defect in Family 2

Clinical features and phenotyping of family

A young child was first referred to a haematology centre with a strong bleeding history. He presented with a moderate platelet count, $77x10^9$/l and $99x10^9$/l on two separate occasions, and a history of easy bruising for which he wore protective headwear during early childhood to protect against possible haematoma formation. He was diagnosed at first presentation to have a platelet function disorder with moderate thrombocytopenia. He is the only affected family member within his family although his mother has stated a history of easy bruising but has not been investigated fully. Three individuals were available to study within this family; the mother, the affected index case and an unaffected sibling. The pedigree in figure 5 shows all members of the family and individuals are labelled where samples were available.

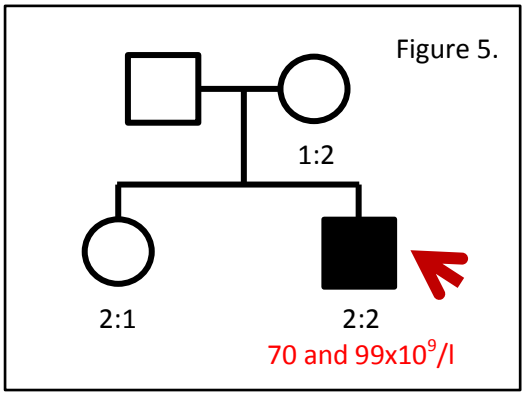


Figure 5 shows the family pedigree for family 2. The index case, 2:2, is indicated with the red arrow. Data is only available for the affected individual, his mother (1:2) and an unaffected sibling (2:1). The affected patients mother did note symptoms of a bleeding diathesis but no investigation had been undertaken indicating a possible de novo or dominant inheritance as the cause of disease.

Platelet function testing of family 2

Platelet function testing was performed using the same GAPP guidelines previously mentioned for family 1 and experimental procedures were performed by Dr Ban Dawood. Deaggregation was observed in the patient at all ADP concentrations (10, 30 and 100µM) but it can be suggested there was no defect in cAMP or cGMP production due to full inhibition in the presence of PGI_2. Full aggregation was observed in response to arachidonic acid negating a possible defect in the TxA_2

pathway. Marked inhibition of aggregation with deaggregation at high concentration was noticed in effect to PAR-1 peptide and the beginning of deaggregation was also observed at the intermediate (250µM) concentration of PAR-4. In addition the level of ATP secretion was markedly reduced in response to ADP, PAR1 and collagen.

Whole exome sequencing of the affected individual of family 2

Whole exome sequencing was performed on the one affected family member; 2.2. Sequencing revealed 108 novel variations (after comparison to databases previously mentioned) that were either insertions or deletions, non-synonymous amino acids changes, stop gains or losses and unknown mutations. Synonymous amino acid changes were excluded for the time being as they do not alter the amino acid sequence. From the 108 novel variations 10 were selected as an initial screen to confirm the mutation and determine familial segregation. These ten included all insertions and deletions, stop gains and a single novel missense variant which was also found to be mutated in family 1. The initial screen was intended to encompass 13 genes that fit the above criteria but three primer pairs failed to work upon control screening. The results from the Sanger sequencing can be found in Table 3.

Table 3.

			Mother	Affected	Unaffected sibling	
Gene	Mutation					
			Patient 1:2	Patient 2:2	Patient 2:1	control
IFIT1	c.497_498insA:p.N166fs,c.590_591insA:p.N197fs.		no ins	het insA	het insA	no ins
CADPS2	c.2935_2937del:p.979_979del,c.2923_2925del:p.975_975del		het delCC	het delCC	het delCC	no del
CHIC1	c.189_191del:p.63_64del,		het delAG	no del	no del	Fail
STAB1	c.3364_3365insC:p.R1122fs,		no ins	het insC	no ins	no ins
CLCNKB	c.118delA:p.R40fs,		no del	no del (F read only)	no del	no del (F read only)
KCNAB3	c.532_534del:p.178_178del,		no del	het del	no del	Fail
WASL	c.862_864del:p.288_288del,		no del	no del	no del	no del
ALS2CL	c.T2207G:p.L736X,c.T248G:p.L83X		het change	het change	het change	no change
C1orf204	c.C37T:p.R13X		homo 7481C>G	het C>T change and homo 7481C>G	homo 7481C>G	homo 7481C>G
SLFN14	c.A652G:p.K218E,		no change	het A>G	Fail	no change

Table 3 shows the Sanger sequencing results from all 3 available individuals within the family as well as an additional control sample. Rows labelled in grey are excluded on the basis that either the mutation is found in all family members, including unaffected, and is therefore not disease causing. White rows are inconclusive due to no confirmation of the initial mutation in the affected patient. This may although indicate a false positive results from the original exome sequencing and this cannot be disregarded. Rows highlighted in blue are, at this stage, possible disease causing mutations. This is indicated through Sanger sequencing confirming in the initial mutation and then no mutations observable in the control and unaffected family members.

From the above sequencing data it can be said that at this point we can exclude three possible mutations on the basis that they are found within unaffected family members so are more than likely non-disease causing variants. Initially we can suggest four possible candidate mutations and genes that, at this point, segregate and are only found within the affected individual. This is primary data and currently excludes a vast body of variations first discovered by exome sequencing but it does provide us with a foundation of from which to continue and progress work.

Mutation screening of the *ANKRD26* 5'-UTR mutations in a panel of patients with thrombocytopenia

In total 39 patients previously determined to be thrombocytopenic (<150x10^0/l platelet count) were entered into a sequencing screen looking in particular at the 5'-UTR region. All patients were presented with an unclassified thrombocytopenia with a possible secondary defect. Patients showed similar platelet counts to those previously identified with an *ANKRD26* mutation. In total three variations were observed across the panel with the results being summarised in Table 4 overleaf. Out of the three variations observed only one falls into the previously reported mutation hotspot of the 5'-UTR of ANKRD26. This single nucleotide change, c.-126T>T/G is found in a heterozygous state in one patient and falls between two previously described mutations in patients with THC2 (1). The location of the mutation is explained in Figure 6 overleaf. The two other variations fall outside of the mutational hotspot and are in fact previously reported common variations. c.59A>A/G and c.59A>G are present in a frequency of 39 and 58%, respectively in this study, and the G allele (rs7897309) is found in 93% of the worldwide population (1000 genomes). c.-140C>C/G is found also as frequently in the 1000 genomes database with a prevalence of 96% (rs41299222).

Table 4.

Patient	Mutations
1	c.59A>G
2	c.59A>G
3	c.59A>G
4	c.59A>A/G
5	Fail
6	c.59A>A/G
7	c.59A>G
8	c.-126T>T/G c.59A>A/G
9	c.59A>A/G
10	c.59A>G
11	c.59A>G
12	Fail
13	c.59A>A/G
14	c.59A>G
15	c.59A>A/G
16	Fail
17	c.59A>G
18	c.59A>G Reverse read only
19	c.59A>G
20	Fail
21	c.59A>G
22	c.59A>G
23	c.59A>G
24	Fail
25	c.59A>A/G
26	c.59A>G
27	c.59A>A/G
28	c.59A>A/G
29	c.59A>A/G
30	wild type
31	c.59A>A/G
32	c.59A>A/G
33	c.-140C>C/G c.59A>G
34	c.-140C>G c.59A>G
35	c.59A>G
36	c.59A>A/G
37	c.59A>A/G
38	c.59A>G
39	Fail

Table 4 shows the results from the screening of the thrombocytopenic panel for mutations in the 5'-UTR region of ANKRD26. Patient families are contained within boxes. 6/39 patients failed to produce any PCR product, in most cases this was due to a limited supply of patient material and therefore DNA to sequence. Three variable bases were noticed across the 39 patients screened, c.59A>A/G, c.-140C>C/G and c.-126T>T/G. c.59A>A/G was found both heterozygously and homozygously in a high frequency of patients with only one patient being wild type at that base. c.-140C>C/G was found in two patients within the same family , further observation reveals these patients are mother and daughter which may explain the presence in both patients, the variation is not observed in any other members of the family. The final variation, c.-126T>T/G, is found a single unrelated patient, patient 8.

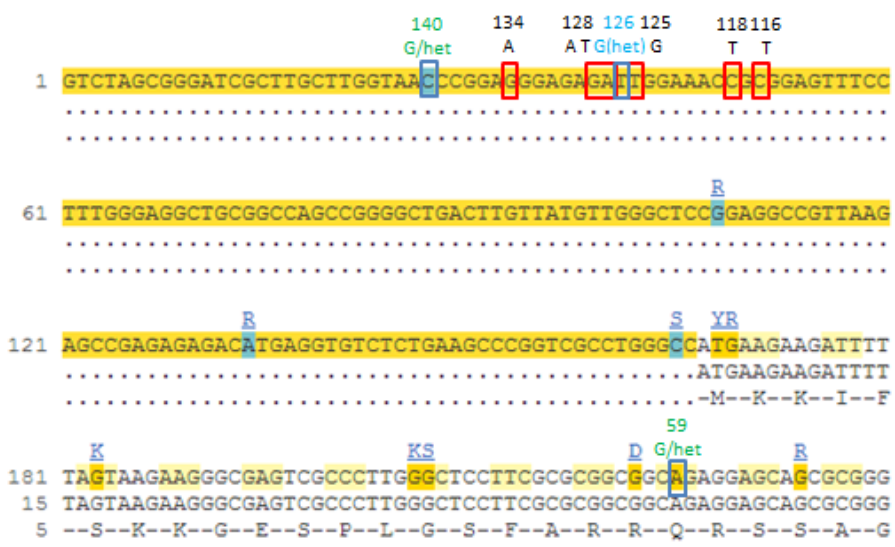


Figure 6 shows an Ensembl search of the 5'-UTR of ANKRD26. Highlighted in red boxes are the six reported mutations in the "mutational hostpot" from -134 to -113 as previously reported by Pippuci *et al.* (1). Within blue boxes are the three variations observed within this study. Worth noting is the c.-126T>T/G change with blue labelling that falls directly between two previously suggested disease causing mutations, these are believed to be the novel cause of THC2.

Discussion

Although no conclusive data has been generated from this study we have managed to reduce the number of possible candidate mutations through a variety of sequencing techniques and have made valuable progress in achieving our aforementioned aims. At this current stage patient reports and causative mutations are incomplete but there are a number of interesting leads that can provide areas for future research and study.

A novel variant in *TMEM30B* as a candidate mutation in family 1

Out of the three candidate variants suggested by whole exome and Sanger sequencing in family 1 *TMEM30* is an interesting possibility for the causative gene. Although the same mutation was present within the unaffected family member it is still not possible to exclude it without a full individual report/platelet function testing. The complex nature of the affected individuals within this family, displaying with both a varied level of thrombocytopenia and a suggested platelet function defect, hints at the possibility of a compound effect involving two or more mutated genes. It is on this basis that therefore a suggestion could be made that although the unaffected individual is classed as phenotypically normal there may be a mild thrombocytopenia present, consistent with other family members, that has not previously been observed due to little or no recent bleeding episodes. The c.778delC variant observed within family 1 occurs within the second transmembrane region of TMEM30B. The effect of this single base deletion is a frame shift which results in the removal of the wild type stop codon and extension of the peptide sequence by 33 amino acids. We predict therefore that the probable effect of this mutation is a degree of loss of function to the protein by removal of the second transmembrane domain externalising the C terminal if the protein does translocate, this mutation is seen in a heterozygous state in all patients though so a complete knock out will not be produced.

Work in 2010 by Bryde *et al.* (87) has tried to elucidate the possible function of TMEM30B within human cell line models and this could potentially provide an explanation to one aspect of the phenotype observed within family 1.

They hypothesise a possible role of TMEM30B, and TMEM30A, in the production of the asymmetrical phospholipid distribution across the plasma membrane. They believe that this effect is mediated through interaction with P4-ATPases and yeast homologues have been shown to be responsible for transport of the phospholipids; phosphatidylserine (PS), phosphatidylethanolamine (PE) and phosphatidylcholine (PC) (88). It was on this notion that Coleman and Molday based their work. They initially determined a cellular localisation of TMEM30B at the plasma membrane through exogenous expression and continued to show interaction of TMEM30B with ATP8B1, B2, B4 and A1 P4-ATPases and markedly increased expression when co-expressed. They also categorised that P4-ATPases fail to translocate from the endoplasmic reticulum to the plasma membrane in the absence of both TMEM30A and TMEM30B and that TMEM30A is required for the conversion of ATP8B1 and B2 to their phosphoenzyme intermediates.

The specific asymmetrical distribution of phospholipids across the endoplasmic and cytoplasmic membrane is important in many cellular processes such as membrane stability, vesicle transport, recognition of apoptosis and blood coagulation which is already been made apparent in the defect in the phospholipid scramblase seen within Scott syndrome (45, 88, 89). Based upon the work previously mentioned by Bryde *et al.* (87) we have two hypotheses as to the effect of the mutation and how it can potentially cause disease. We believe that if TMEM30B does in fact play a role in the function of a P4-ATPase the mutation presented within family 1 has the potential to disrupt either its translocation out of the endoplasmic reticulum or its function involved in transporting the polar phospholipids. If we therefore presume that TMEM30B, and not TMEM30A (TMEM30C is expressed strictly in the testis so can therefore be ruled out in this instance), interacts with a P4-ATPase acting in a flippase manner, that is normally functioning to remove PS and other phospholipids from the exoplasmic to the cytoplasmic leaflet of the plasma membrane, we are able to suggest a two-fold idea that may explain the thrombocytopenia observed. Due to the role of PS in apoptosis and blood coagulation we believe that a possible effect of this particular mutation might be that PS is blocked from being transported from the exoplasmic membrane. It is this retention, on the exoplasmic side,

that may trigger apoptotic like signals and a local coagulation resulting in removal of platelets by coagulation or phagocytosis and therefore a reduced platelet count in what would appear as a Stormorken-like disease. Alternatively if TMEM30B is shown to interact with the opposite of a flippase, a floppase, then the platelet function defect may be equally explained by a reduced exposure of PS leading to reduced coagulation. The mutation in TMEM30B therefore provides quite an interesting development from which to progress. For these hypotheses to be true a number of criteria need to be satisfied, this is why the use of the cloning approach to study TMEM30B localisation and the effect of the mutation is an ideal starting point for further work to commence. In addition an Annexin-V binding assay could be used to determine the presence of PS on the outer membrane and henceforth validate or disprove the proposed hypothesis.

Novel variations in SLFN14 in both family 1 and 2

Although TMEM30B does provide us with an interesting lead we cannot rule out the two other possible candidate mutations without further work. Interestingly the presence of a mutation in the functionally unknown SLFN14 is apparent in both patients from family 1 and those from family 2. Closer investigation into the nature of these mutations reveals that both the mutations are within seven bases of each other and therefore could be affected the same functional domain. When the mutations are analysed using the mutation prediction software Polyphen (Polymorphism Phenotyping version2.1.0)(90) the mutation found within family 1, p.V220D, returns with two alternate prediction scores (HumDiv and HumVar) of 0.773 and 0.665 and is predicted to be "possibly disease causing" (all prediction scores are out of a total of 1). In contrast the variant found in family 2, p.K218E, returns HumDiv and HumVar scores of 0.999 and 0.996 respectively and is predicted to be "probably disease causing". As V220D represents a nonpolar to acidic polar change and K218E represents a basic polar to acidic polar change the difference seems to lie within the conservation of the amino acids. A figure showing the amino acid conservation in a number of organisms is shown in Figure 7. Although the valine at amino acid 220 is not well conserved the presence of both mutations in close juxtaposition to one another provides an interesting prospect in

terms of cause of disease. Patients from both family 1 and family 2 are known to have a compound disease consisting of a level of thrombocytopenia and an additional platelet function disorder. This consistency continues when observing the platelet phenotyping as both patients present with deaggregation to ADP and PAR-1 peptide and an additional ATP secretion defect. Although it is not possible to draw conclusions without the known function of the protein and categorising the mutation this work does provide an insight into the potential of SLFN14 and a suggestion into an area of further research to be completed.

Figure 7.

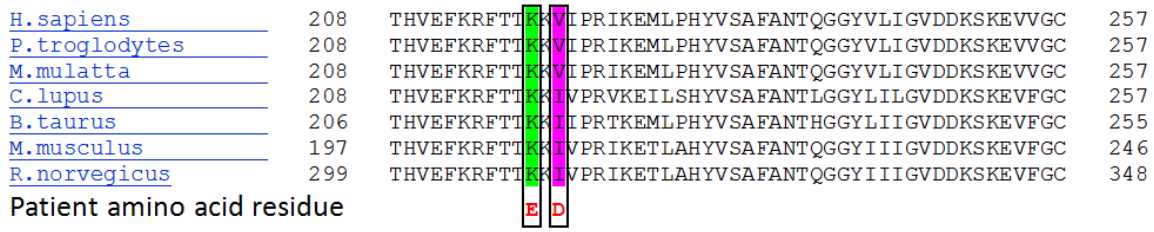


Figure 7 shows the conservation of amino acids in the region flanking the two mutations in SLFN14. As you can see conservation is maintained throughout majority of mammals across the region with extensive conservation at the lysine at position 218 in the *H.sapiens* coding region. The valine at position 220, the mutated amino acid within family 1, is less well conserved. No homologues of SLFN14 have yet been published in lesser order model organisms such as *D.rerio* or *C.elegans.*

It is worth noting that *SLFN14* is one of four possible candidate mutations within family 2 and there are still 98 additional novel variants to screen. In cases where there is little familial background of disease there is more of an argument for a complex multigenic disease in which several genes can be causative in combination. At this stage we cannot rule this idea out and our only method to confirm or deny this theory is to extend the family history and variant segregation.

A novel variation in the 5'-UTR of *ANKRD26* in a patient with thrombocytopenia

One of our initial aims was to determine the prevalence of mutations within the 5'-UTR of ANKRD26 in a panel of thrombocytopenic patients. In total 33 patients were successfully screened using Sanger sequencing and one single nucleotide change was found that falls in the proposed mutational hotspot. This mutation c.-126T>T/G is a novel variant not previously reported within the 1000 genomes database and our internal database. Further work is needed to establish this mutation and whether the patient can therefore be diagnosed with THC2, this could be achieved through the use of a luciferase assay similar to as previously described by Pippuci (1).

Whole exome sequencing in the diagnosis of patients with rare platelet based bleeding disorders

Although whole exome sequencing has been proven to generate novel candidate variations it cannot give us the complete picture of a patient's genome. The method of sequencing used in this study stretches 5 bases from the exons into the untranslated regions allowing coverage of splice acceptor and donor sites. It does not however cover the vast non-coding section of a patient's genome so variations in enhancers, suppressors, non-coding RNAs and other intronic variants will be missed. There is however an option to work a whole genome sequencing approach into the GAPP work flow after whole exome sequencing when no variants are shown. Structuring in this way is a lot more cost effective and efficient but still retains the option for complete coverage if needed.

Acknowledgements

I would like to acknowledge all members of the Watson/Morgan lab, in particular Dr Neil Morgan for his continued help and support throughout the project.

1. 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., and Balduini, C. L. (2011) Mutations in the 5' UTR of ANKRD26, the ankirin repeat domain 26 gene, cause an autosomal-dominant form of inherited thrombocytopenia, THC2. *American journal of human genetics* **88**, 115-120
2. Watson, S. P., Lowe, G. C., Lordkipanidze, M., and Morganon, N. V. (2013) Genotyping and phenotyping of platelet function disorders. *Journal of thrombosis and haemostasis : JTH*
3. Bolton-Maggs, P. H., Chalmers, E. A., Collins, P. W., Harrison, P., Kitchen, S., Liesner, R. J., Minford, A., Mumford, A. D., Parapia, L. A., Perry, D. J., Watson, S. P., Wilde, J. T., and Williams, M. D. (2006) A review of inherited platelet disorders with guidelines for their management on behalf of the UKHCDO. *British journal of haematology* **135**, 603-633
4. Geddis, A. E. (2010) Megakaryopoiesis. *Seminars in hematology* **47**, 212-219
5. Becker, R. P., and De Bruyn, P. P. (1976) The transmural passage of blood cells into myeloid sinusoids and the entry of platelets into the sinusoidal circulation; a scanning electron microscopic investigation. *The American journal of anatomy* **145**, 183-205
6. Sharnoff, J. G., and Scardino, V. (1960) Pulmonary megakaryocytes in human fetuses and premature and full-term infants. *Archives of pathology* **69**, 139-141
7. Gurney, A. L., Carver-Moore, K., de Sauvage, F. J., and Moore, M. W. (1994) Thrombocytopenia in c-mpl-deficient mice. *Science (New York, N.Y.)* **265**, 1445-1447
8. Drachman, J. G. (2004) Inherited thrombocytopenia: when a low platelet count does not mean ITP. *Blood* **103**, 390-398
9. Balduini, C. L., Iolascon, A., and Savoia, A. (2002) Inherited thrombocytopenias: from genes to therapy. *Haematologica* **87**, 860-880
10. Balduini, C. L., Cattaneo, M., Fabris, F., Gresele, P., Iolascon, A., Pulcinelli, F. M., and Savoia, A. (2003) Inherited thrombocytopenias: a proposed diagnostic algorithm from the Italian Gruppo di Studio delle Piastrine. *Haematologica* **88**, 582-592
11. Millikan, P. D., Balamohan, S. M., Raskind, W. H., and Kacena, M. A. (2011) Inherited thrombocytopenia due to GATA-1 mutations. *Seminars in thrombosis and hemostasis* **37**, 682-689
12. Walker, L. C., Stevens, J., Campbell, H., Corbett, R., Spearing, R., Heaton, D., Macdonald, D. H., Morris, C. M., and Ganly, P. (2002) A novel inherited mutation of the transcription factor RUNX1 causes thrombocytopenia and may predispose to acute myeloid leukaemia. *British journal of haematology* **117**, 878-881
13. Chang, A. N., Cantor, A. B., Fujiwara, Y., Lodish, M. B., Droho, S., Crispino, J. D., and Orkin, S. H. (2002) GATA-factor dependence of the multitype zinc-finger protein FOG-1 for its essential role in megakaryopoiesis. *Proceedings of the National Academy of Sciences of the United States of America* **99**, 9237-9242
14. Shivdasani, R. A., Fujiwara, Y., McDevitt, M. A., and Orkin, S. H. (1997) A lineage-selective knockout establishes the critical role of transcription factor GATA-1 in megakaryocyte growth and platelet development. *The EMBO journal* **16**, 3965-3973
15. Nurden, A. T., Freson, K., and Seligsohn, U. (2012) Inherited platelet disorders. *Haemophilia : the official journal of the World Federation of Hemophilia* **18 Suppl 4**, 154-160
16. Lordier, L., Bluteau, D., Jalil, A., Legrand, C., Pan, J., Rameau, P., Jouni, D., Bluteau, O., Mercher, T., Leon, C., Gachet, C., Debili, N., Vainchenker, W., Raslova, H., and Chang, Y. (2012) RUNX1-induced silencing of non-muscle myosin heavy chain IIB contributes to megakaryocyte polyploidization. *Nature communications* **3**, 717
17. Raslova, H., Komura, E., Le Couedic, J. P., Larbret, F., Debili, N., Feunteun, J., Danos, O., Albagli, O., Vainchenker, W., and Favier, R. (2004) FLI1 monoallelic expression combined with its hemizygous loss underlies Paris-Trousseau/Jacobsen thrombopenia. *The Journal of clinical investigation* **114**, 77-84

18. Trkova, M., Becvarova, V., Hynek, M., Hnykova, L., Hlavova, E., Kreckova, G., Kulovany, E., Cutka, D., Zatloukalova, J., Markova, K., Sukova, M., Horacek, J., and Stejskal, D. (2012) SNP array and phenotype correlation shows that FLI1 deletion per se is not responsible for thrombocytopenia development in Jacobsen syndrome. *American journal of medical genetics. Part A* **158A**, 2545-2550
19. Basinko, A., Audebert-Bellanger, S., Douet-Guilbert, N., Le Franc, J., Parent, P., Quemener, S., La Selve, P., Bovo, C., Morel, F., Le Bris, M. J., and De Braekeleer, M. (2011) Subtelomeric monosomy 11q and trisomy 16q in siblings and an unrelated child: molecular characterization of two der(11)t(11;16). *American journal of medical genetics. Part A* **155A**, 2281-2287
20. Shivdasani, R. A., Rosenblatt, M. F., Zucker-Franklin, D., Jackson, C. W., Hunt, P., Saris, C. J., and Orkin, S. H. (1995) Transcription factor NF-E2 is required for platelet formation independent of the actions of thrombopoietin/MGDF in megakaryocyte development. *Cell* **81**, 695-704
21. A., W. (1937) Familiärer, angeborener Morbus Werlhofii? *Monatsschr Kinderheilkd.*, 212-216
22. Aldrich, R. A., Steinberg, A. G., and Campbell, D. C. (1954) Pedigree demonstrating a sex-linked recessive condition characterized by draining ears, eczematoid dermatitis and bloody diarrhea. *Pediatrics* **13**, 133-139
23. Albert, M. H., Bittner, T. C., Nonoyama, S., Notarangelo, L. D., Burns, S., Imai, K., Espanol, T., Fasth, A., Pellier, I., Strauss, G., Morio, T., Gathmann, B., Noordzij, J. G., Fillat, C., Hoenig, M., Nathrath, M., Meindl, A., Pagel, P., Wintergerst, U., Fischer, A., Thrasher, A. J., Belohradsky, B. H., and Ochs, H. D. (2010) X-linked thrombocytopenia (XLT) due to WAS mutations: clinical characteristics, long-term outcome, and treatment options. *Blood* **115**, 3231-3238
24. Derry, J. M., Ochs, H. D., and Francke, U. (1994) Isolation of a novel gene mutated in Wiskott-Aldrich syndrome. *Cell* **78**, 635-644
25. Remold-O'Donnell, E., Rosen, F. S., and Kenney, D. M. (1996) Defects in Wiskott-Aldrich syndrome blood cells. *Blood* **87**, 2621-2631
26. Seri, M., Cusano, R., Gangarossa, S., Caridi, G., Bordo, D., Lo Nigro, C., Ghiggeri, G. M., Ravazzolo, R., Savino, M., Del Vecchio, M., d'Apolito, M., Iolascon, A., Zelante, L. L., Savoia, A., Balduini, C. L., Noris, P., Magrini, U., Belletti, S., Heath, K. E., Babcock, M., Glucksman, M. J., Aliprandis, E., Bizzaro, N., Desnick, R. J., and Martignetti, J. A. (2000) Mutations in MYH9 result in the May-Hegglin anomaly, and Fechtner and Sebastian syndromes. The May-Hegglin/Fechtner Syndrome Consortium. *Nature genetics* **26**, 103-105
27. Balduini, C. L., Pecci, A., and Savoia, A. (2011) Recent advances in the understanding and management of MYH9-related inherited thrombocytopenias. *British journal of haematology* **154**, 161-174
28. Seri, M., Pecci, A., Di Bari, F., Cusano, R., Savino, M., Panza, E., Nigro, A., Noris, P., Gangarossa, S., Rocca, B., Gresele, P., Bizzaro, N., Malatesta, P., Koivisto, P. A., Longo, I., Musso, R., Pecoraro, C., Iolascon, A., Magrini, U., Rodriguez Soriano, J., Renieri, A., Ghiggeri, G. M., Ravazzolo, R., Balduini, C. L., and Savoia, A. (2003) MYH9-related disease: May-Hegglin anomaly, Sebastian syndrome, Fechtner syndrome, and Epstein syndrome are not distinct entities but represent a variable expression of a single illness. *Medicine* **82**, 203-215
29. Stoddart, M. T., Connor, P., Germeshausen, M., Ballmaier, M., and Steward, C. G. (2013) Congenital Amegakaryocytic Thrombocytopenia (CAMT) Presenting as Severe Pancytopenia in the First Month of Life. *Pediatric blood & cancer*
30. van den Oudenrijn, S., Bruin, M., Folman, C. C., Peters, M., Faulkner, L. B., de Haas, M., and von dem Borne, A. E. (2000) Mutations in the thrombopoietin receptor, Mpl, in children with congenital amegakaryocytic thrombocytopenia. *British journal of haematology* **110**, 441-448
31. Jin, J., Daniel, J. L., and Kunapuli, S. P. (1998) Molecular basis for ADP-induced platelet activation. II. The P2Y1 receptor mediates ADP-induced intracellular calcium mobilization and shape change in platelets. *The Journal of biological chemistry* **273**, 2030-2034

32. Nurden, A. T., and Caen, J. P. (1974) An abnormal platelet glycoprotein pattern in three cases of Glanzmann's thrombasthenia. *British journal of haematology* **28**, 253-260
33. Ramasamy, I. (2004) Inherited bleeding disorders: disorders of platelet adhesion and aggregation. *Critical reviews in oncology/hematology* **49**, 1-35
34. George, J. N., Caen, J. P., and Nurden, A. T. (1990) Glanzmann's thrombasthenia: the spectrum of clinical disease. *Blood* **75**, 1383-1395
35. Nurden, A. T., Fiore, M., Nurden, P., and Pillois, X. (2011) Glanzmann thrombasthenia: a review of ITGA2B and ITGB3 defects with emphasis on variants, phenotypic variability, and mouse models. *Blood* **118**, 5996-6005
36. Bernard, J. S., J.P. (1948) Sur une nouvelle variete de dystrophie thrombocytaire-hemorragipare congenitale. *Semaine des Hopitaux de Paris* **244**, 159-160
37. Ware, J., Russell, S. R., Vicente, V., Scharf, R. E., Tomer, A., McMillan, R., and Ruggeri, Z. M. (1990) Nonsense mutation in the glycoprotein Ib alpha coding sequence associated with Bernard-Soulier syndrome. *Proceedings of the National Academy of Sciences of the United States of America* **87**, 2026-2030
38. Vettore, S., Scandellari, R., Moro, S., Lombardi, A. M., Scapin, M., Randi, M. L., and Fabris, F. (2008) Novel point mutation in a leucine-rich repeat of the GPIbalpha chain of the platelet von Willebrand factor receptor, GPIb/IX/V, resulting in an inherited dominant form of Bernard-Soulier syndrome affecting two unrelated families: the N41H variant. *Haematologica* **93**, 1743-1747
39. Nurden, A. T., and Caen, J. P. (1975) Specific roles for platelet surface glycoproteins in platelet function. *Nature* **255**, 720-722
40. Ulsemer, P., Strassel, C., Baas, M. J., Salamero, J., Chasserot-Golaz, S., Cazenave, J. P., De La Salle, C., and Lanza, F. (2001) Biosynthesis and intracellular post-translational processing of normal and mutant platelet glycoprotein GPIb-IX. *The Biochemical journal* **358**, 295-303
41. Weiss, H. J., Vicic, W. J., Lages, B. A., and Rogers, J. (1979) Isolated deficiency of platelet procoagulant activity. *The American journal of medicine* **67**, 206-213
42. Rosing, J., Bevers, E. M., Comfurius, P., Hemker, H. C., van Dieijen, G., Weiss, H. J., and Zwaal, R. F. (1985) Impaired factor X and prothrombin activation associated with decreased phospholipid exposure in platelets from a patient with a bleeding disorder. *Blood* **65**, 1557-1561
43. Fadok, V. A., Voelker, D. R., Campbell, P. A., Cohen, J. J., Bratton, D. L., and Henson, P. M. (1992) Exposure of phosphatidylserine on the surface of apoptotic lymphocytes triggers specific recognition and removal by macrophages. *Journal of immunology (Baltimore, Md. : 1950)* **148**, 2207-2216
44. Majumder, R., Quinn-Allen, M. A., Kane, W. H., and Lentz, B. R. (2008) A phosphatidylserine binding site in factor Va C1 domain regulates both assembly and activity of the prothrombinase complex. *Blood* **112**, 2795-2802
45. Lhermusier, T., Chap, H., and Payrastre, B. (2011) Platelet membrane phospholipid asymmetry: from the characterization of a scramblase activity to the identification of an essential protein mutated in Scott syndrome. *Journal of thrombosis and haemostasis : JTH* **9**, 1883-1891
46. Albrecht, C., McVey, J. H., Elliott, J. I., Sardini, A., Kasza, I., Mumford, A. D., Naoumova, R. P., Tuddenham, E. G., Szabo, K., and Higgins, C. F. (2005) A novel missense mutation in ABCA1 results in altered protein trafficking and reduced phosphatidylserine translocation in a patient with Scott syndrome. *Blood* **106**, 542-549
47. Stormorken, H., Sjaastad, O., Langslet, A., Sulg, I., Egge, K., and Diderichsen, J. (1985) A new syndrome: thrombocytopathia, muscle fatigue, asplenia, miosis, migraine, dyslexia and ichthyosis. *Clinical genetics* **28**, 367-374
48. Stormorken, H. (2002) [Stormorken's syndrome]. *Tidsskrift for den Norske laegeforening : tidsskrift for praktisk medicin, ny raekke* **122**, 2853-2856

49. Stormorken, H., Holmsen, H., Sund, R., Sakariassen, K. S., Hovig, T., Jellum, E., and Solum, O. (1995) Studies on the haemostatic defect in a complicated syndrome. An inverse Scott syndrome platelet membrane abnormality? *Thrombosis and haemostasis* **74**, 1244-1251
50. Moroi, M., Jung, S. M., Okuma, M., and Shinmyozu, K. (1989) A patient with platelets deficient in glycoprotein VI that lack both collagen-induced aggregation and adhesion. *The Journal of clinical investigation* **84**, 1440-1445
51. Oury, C. (1999) Congenital deficiency of the phospholipase C coupled platelet P2Y1 receptor leads to a mild bleeding disorder. *Thrombosis and haemostasis*, 20-21
52. Oury, C., Toth-Zsamboki, E., Van Geet, C., Thys, C., Wei, L., Nilius, B., Vermylen, J., and Hoylaerts, M. F. (2000) A natural dominant negative P2X1 receptor due to deletion of a single amino acid residue. *The Journal of biological chemistry* **275**, 22611-22614
53. Hirata, T., Kakizuka, A., Ushikubi, F., Fuse, I., Okuma, M., and Narumiya, S. (1994) Arg60 to Leu mutation of the human thromboxane A2 receptor in a dominantly inherited bleeding disorder. *The Journal of clinical investigation* **94**, 1662-1667
54. Ren, Q., Ye, S., and Whiteheart, S. W. (2008) The platelet release reaction: just when you thought platelet secretion was simple. *Current opinion in hematology* **15**, 537-541
55. Hermansky, F., and Pudlak, P. (1959) Albinism associated with hemorrhagic diathesis and unusual pigmented reticular cells in the bone marrow: report of two cases with histochemical studies. *Blood* **14**, 162-169
56. Di Pietro, S. M., and Dell'Angelica, E. C. (2005) The cell biology of Hermansky-Pudlak syndrome: recent advances. *Traffic (Copenhagen, Denmark)* **6**, 525-533
57. Tchernev, V. T., Mansfield, T. A., Giot, L., Kumar, A. M., Nandabalan, K., Li, Y., Mishra, V. S., Detter, J. C., Rothberg, J. M., Wallace, M. R., Southwick, F. S., and Kingsmore, S. F. (2002) The Chediak-Higashi protein interacts with SNARE complex and signal transduction proteins. *Molecular medicine (Cambridge, Mass.)* **8**, 56-64
58. Cattaneo, M. (2003) Inherited platelet-based bleeding disorders. *Journal of thrombosis and haemostasis : JTH* **1**, 1628-1636
59. Raccuglia, G. (1971) Gray platelet syndrome. A variety of qualitative platelet disorder. *The American journal of medicine* **51**, 818-828
60. Jantunen, E., Hanninen, A., Naukkarinen, A., Vornanen, M., and Lahtinen, R. (1994) Gray platelet syndrome with splenomegaly and signs of extramedullary hematopoiesis: a case report with review of the literature. *American journal of hematology* **46**, 218-224
61. 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., and Di Paola, J. (2011) Mutations in NBEAL2, encoding a BEACH protein, cause gray platelet syndrome. *Nature genetics* **43**, 738-740
62. Hayward, C. P., Cramer, E. M., Kane, W. H., Zheng, S., Bouchard, M., Masse, J. M., and Rivard, G. E. (1997) Studies of a second family with the Quebec platelet disorder: evidence that the degradation of the alpha-granule membrane and its soluble contents are not secondary to a defect in targeting proteins to alpha-granules. *Blood* **89**, 1243-1253
63. Lo, B., Li, L., Gissen, P., Christensen, H., McKiernan, P. J., Ye, C., Abdelhaleem, M., Hayes, J. A., Williams, M. D., Chitayat, D., and Kahr, W. H. (2005) Requirement of VPS33B, a member of the Sec1/Munc18 protein family, in megakaryocyte and platelet alpha-granule biogenesis. *Blood* **106**, 4159-4166
64. Born, G. V. (1962) Aggregation of blood platelets by adenosine diphosphate and its reversal. *Nature* **194**, 927-929
65. Blais, N., Pharand, C., Lordkipanidze, M., Sia, Y. K., Merhi, Y., and Diodati, J. G. (2009) Response to aspirin in healthy individuals. Cross-comparison of light transmission aggregometry, VerifyNow system, platelet count drop, thromboelastography (TEG) and urinary 11-dehydrothromboxane B(2). *Thrombosis and haemostasis* **102**, 404-411

66. Cattaneo, M., Hayward, C. P., Moffat, K. A., Pugliano, M. T., Liu, Y., and Michelson, A. D. (2009) Results of a worldwide survey on the assessment of platelet function by light transmission aggregometry: a report from the platelet physiology subcommittee of the SSC of the ISTH. *Journal of thrombosis and haemostasis : JTH* **7**, 1029
67. Linden, M. D., Frelinger, A. L., 3rd, Barnard, M. R., Przyklenk, K., Furman, M. I., and Michelson, A. D. (2004) Application of flow cytometry to platelet disorders. *Seminars in thrombosis and hemostasis* **30**, 501-511
68. Jilma, B. (2001) Platelet function analyzer (PFA-100): a tool to quantify congenital or acquired platelet dysfunction. *The Journal of laboratory and clinical medicine* **138**, 152-163
69. Chan, M. V., Armstrong, P. C., Papalia, F., Kirkby, N. S., and Warner, T. D. (2011) Optical multichannel (optimul) platelet aggregometry in 96-well plates as an additional method of platelet reactivity testing. *Platelets* **22**, 485-494
70. Lordkipanidze, M., Chan, M., Lowe, G., Dawood, B., Warner, T., and Watson, S. (2012) Head-to-head comparisons of the Optimul assay (96-well plate aggregometry) with gold standard light transmission aggregometry for detection of acquired and inherited platelet defects. . In *Presented at the 58th Scientific and Standardisation Committee meeting of the ISTH*, Liverpool, United Kingdom
71. Poon, M. C., D'Oiron, R., Von Depka, M., Khair, K., Negrier, C., Karafoulidou, A., Huth-Kuehne, A., and Morfini, M. (2004) Prophylactic and therapeutic recombinant factor VIIa administration to patients with Glanzmann's thrombasthenia: results of an international survey. *Journal of thrombosis and haemostasis : JTH* **2**, 1096-1103
72. Coppola, A., and Di Minno, G. (2008) Desmopressin in inherited disorders of platelet function. *Haemophilia : the official journal of the World Federation of Hemophilia* **14 Suppl 1**, 31-39
73. Tengborn, L., and Petruson, B. (1996) A patient with Glanzmann thrombasthenia and epistaxis successfully treated with recombinant factor VIIa. *Thrombosis and haemostasis* **75**, 981-982
74. McColl, M. D., and Gibson, B. E. (1997) Sibling allogeneic bone marrow transplantation in a patient with type I Glanzmann's thrombasthenia. *British journal of haematology* **99**, 58-60
75. Rieger, C., Rank, A., Fiegl, M., Tischer, J., Schiel, X., Ostermann, H., and Kolb, H. J. (2006) Allogeneic stem cell transplantation as a new treatment option for patients with severe Bernard-Soulier Syndrome. *Thrombosis and haemostasis* **95**, 190-191
76. Wilcox, D. A., Olsen, J. C., Ishizawa, L., Griffith, M., and White, G. C., 2nd (1999) Integrin alphaIIb promoter-targeted expression of gene products in megakaryocytes derived from retrovirus-transduced human hematopoietic cells. *Proceedings of the National Academy of Sciences of the United States of America* **96**, 9654-9659
77. Wilcox, D. A., Olsen, J. C., Ishizawa, L., Bray, P. F., French, D. L., Steeber, D. A., Bell, W. R., Griffith, M., and White, G. C., 2nd (2000) Megakaryocyte-targeted synthesis of the integrin beta(3)-subunit results in the phenotypic correction of Glanzmann thrombasthenia. *Blood* **95**, 3645-3651
78. Guerrero, J. A., Rivera, J., Quiroga, T., Martinez-Perez, A., Anton, A. I., Martinez, C., Panes, O., Vicente, V., Mezzano, D., Soria, J. M., and Corral, J. (2011) Novel loci involved in platelet function and platelet count identified by a genome-wide study performed in children. *Haematologica* **96**, 1335-1343
79. Bunimov, N., Fuller, N., and Hayward, C. P. (2013) Genetic loci associated with platelet traits and platelet disorders. *Seminars in thrombosis and hemostasis* **39**, 291-305
80. Izzi, B., Francois, I., Labarque, V., Thys, C., Wittevrongel, C., Devriendt, K., Legius, E., Van den Bruel, A., D'Hooghe, M., Lambrechts, D., de Zegher, F., Van Geet, C., and Freson, K. (2012) Methylation defect in imprinted genes detected in patients with an Albright's hereditary osteodystrophy like phenotype and platelet Gs hypofunction. *PloS one* **7**, e38579

81. Nisar, S., Daly, M. E., Federici, A. B., Artoni, A., Mumford, A. D., Watson, S. P., and Mundell, S. J. (2011) An intact PDZ motif is essential for correct P2Y12 purinoceptor traffic in human platelets. *Blood* **118**, 5641-5651
82. Punzo, F., Mientjes, E. J., Rohe, C. F., Scianguetta, S., Amendola, G., Oostra, B. A., Bertoli-Avella, A. M., and Perrotta, S. (2010) A mutation in the acyl-coenzyme A binding domain-containing protein 5 gene (ACBD5) identified in autosomal dominant thrombocytopenia. *Journal of thrombosis and haemostasis : JTH* **8**, 2085-2087
83. Gandhi, M. J., Cummings, C. L., and Drachman, J. G. (2003) FLJ14813 missense mutation: a candidate for autosomal dominant thrombocytopenia on human chromosome 10. *Human heredity* **55**, 66-70
84. Noris, P., Perrotta, S., Seri, M., Pecci, A., Gnan, C., Loffredo, G., Pujol-Moix, N., Zecca, M., Scognamiglio, F., De Rocco, D., Punzo, F., Melazzini, F., Scianguetta, S., Casale, M., Marconi, C., Pippucci, T., Amendola, G., Notarangelo, L. D., Klersy, C., Civaschi, E., Balduini, C. L., and Savoia, A. (2011) Mutations in ANKRD26 are responsible for a frequent form of inherited thrombocytopenia: analysis of 78 patients from 21 families. *Blood* **117**, 6673-6680
85. Bera, T. K., Liu, X. F., Yamada, M., Gavrilova, O., Mezey, E., Tessarollo, L., Anver, M., Hahn, Y., Lee, B., and Pastan, I. (2008) A model for obesity and gigantism due to disruption of the Ankrd26 gene. *Proceedings of the National Academy of Sciences of the United States of America* **105**, 270-275
86. Schwarz, J. M., Rodelsperger, C., Schuelke, M., and Seelow, D. (2010) MutationTaster evaluates disease-causing potential of sequence alterations. *Nature methods* **7**, 575-576
87. Bryde, S., Hennrich, H., Verhulst, P. M., Devaux, P. F., Lenoir, G., and Holthuis, J. C. (2010) CDC50 proteins are critical components of the human class-1 P4-ATPase transport machinery. *The Journal of biological chemistry* **285**, 40562-40572
88. Pomorski, T., Lombardi, R., Riezman, H., Devaux, P. F., van Meer, G., and Holthuis, J. C. (2003) Drs2p-related P-type ATPases Dnf1p and Dnf2p are required for phospholipid translocation across the yeast plasma membrane and serve a role in endocytosis. *Molecular biology of the cell* **14**, 1240-1254
89. Balasubramanian, K., and Schroit, A. J. (2003) Aminophospholipid asymmetry: A matter of life and death. *Annual review of physiology* **65**, 701-734
90. Adzhubei, I. A., Schmidt, S., Peshkin, L., Ramensky, V. E., Gerasimova, A., Bork, P., Kondrashov, A. S., and Sunyaev, S. R. (2010) A method and server for predicting damaging missense mutations. *Nature methods* **7**, 248-249

Appendix

PCR primers

Cardiff Family

Primer name	Sequence	Scale (µmol)	Purification
TOR2A_F	CTTTCTGAGCCACTACGATGG	0.025	DST
TOR2A_R	CTCAACGGGTACCCTCCAC	0.025	DST
NEMF_F	GGGAAGAAAGGAAAAACAAAGG	0.025	DST
NEMF_R	TGTATTTCTGGATGACATTCCATT	0.025	DST
SLFN14_F	GGGTTTAGAGCCCAAAGAGG	0.025	DST
SLFN14_R	CAGCAGAAGTGGAATGTAGGC	0.025	DST
TMEM30B_Exon1_1F	CTGAGCTGGCGGGAAAC	0.025	DST
TMEM30B_Exon1_1R	ACTCGTTGACAGGGTGGC	0.025	DST
TMEM30B_Exon1_2F	TGCCCGAGCTCTTCCAG	0.025	DST
TMEM30B_Exon1_2R	GTGCGCATCCACACCAC	0.025	DST
TMEM30B_Exon1_3F	GTCAAGTTCCGCAACCC	0.025	DST
TMEM30B_Exon1_3R	GAGATGCCAAAAGCACCTTG	0.025	DST
TMEM30B_Internal_F	GGCATCAAGGAGCTGGAGTA	0.025	DST
TMEM30B_Internal_R	GCCCCTGGGAACAAATTACT	0.025	DST
Birmingham Family			
IFIT_F	ATGGGCCTTGCTGAAGTG	0.025	DST
IFIT_R	GCCACCTCAAATGTGGGC	0.025	DST
CADPS2_F	TGGCCCATATTTTCCAAATG	0.025	DST
CADPS2_R	CGCAGAACACAGCCCAG	0.025	DST
CHIC1_F	ACCTCGGCAGGTTCAAACTC	0.025	DST
CHIC1_R	CCCTCGATCACTTCTCATTCC	0.025	DST
STAB1_F	CTGGCAGTCTCTCTGTTGGG	0.025	DST
STAB1_R	CTCGCCCTCTATGCCTTTG	0.025	DST
CLCNKB_F	caaaatggagatcgcaacc	0.025	DST
CLCNKB_R	gaaaggaagagcaaggggtg	0.025	DST
KCNAB3_F	GGGAGGAGAGGAAGGAATTG	0.025	DST
KCNAB3_R	AATGTTCCCTACTCCTCTGGC	0.025	DST
WASL_F	AAATAAAATGCAAAGATGAGACCC	0.025	DST
WASL_R	TCCTGCAGTAGTTGGAACCTG	0.025	DST
ALS2CL_F	ACTCCAGTTCTGTGGCAAGG	0.025	DST
ALS2CL_R	CTCTTTCCCCAAGGGAGGTC	0.025	DST
C1orf204_F	CGAAACATCAGAAAGGGATTG	0.025	DST
C1orf204_R	CCCAGCTGCAGACAGAGAG	0.025	DST