

We are IntechOpen, the world's leading publisher of Open Access books Built by scientists, for scientists

6,900

Open access books available

186,000

International authors and editors

200M

Downloads

154

Countries delivered to

TOP 1%

most cited scientists

12.2%

Contributors from top 500 universities



WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.

For more information visit www.intechopen.com



DNA-based Diagnosis of Uncharacterized Inherited Macrothrombocytopenias Using Next-generation Sequencing Technology with a Candidate Gene Array

David J. Rabbolini, Marie-Christine Morel Kopp, Sara Gabrielli, Qiang Chen, William S. Stevenson and Christopher M. Ward

Additional information is available at the end of the chapter

<http://dx.doi.org/10.5772/61777>

Abstract

Inherited macrothrombocytopenias comprise a heterogeneous group of inherited platelet disorders that are characterized by large platelets, thrombocytopenia and bleeding tendencies in affected individuals. Diagnostic platforms have traditionally involved a battery of complex phenotypic tests that often fail to reach a diagnosis. Next-generation sequencing lacks the pre-analytical and analytical shortcoming of these tests and provides an attractive alternate diagnostic approach. Our group has developed a candidate gene array targeting genes known to affect platelet function and tested it in a large cohort of Australasian patients with presumed platelet function disorders, particularly macrothrombocytopenia. This array identified causative variants in a significant portion of patients with uncharacterized platelet disorders, including transcription factor mutations that cannot easily be diagnosed with standard platelet phenotyping procedures. We propose that targeted genotypic screening can identify the genetic basis of platelet function defects and has the potential to be developed into a powerful clinical platform to help clinicians diagnose these rare disorders.

Keywords: Inherited macrothrombocytopenia, next-generation sequencing, candidate gene array

1. Introduction

Platelets are essential for clot formation after tissue trauma. Initiation of the platelet plug occurs by adhesion of platelets to the damaged vascular endothelium mediated by interactions of glycoprotein Ib/IX/V complexes with von Willebrand factor (vWF), and GPVI and integrin

$\alpha 2\beta 1$ with collagen [1]. Extension of the platelet plug requires activation of $\alpha IIb\beta 3$ through an “inside-out” signaling cascade which enables receptor cross-linking with fibrinogen and vWF and activation of “outside-in” signaling events [1, 2].

Primary hemostasis relies on both adequate function and number of platelets. Abnormalities in platelet function and/ or number may be acquired (liver disease, chronic kidney disease) or inherited (inherited platelet function disorders, IPFDs or inherited platelet number disorders, IPNDs). The group of inherited macrothrombocytopenias is included in the heterogeneous IPNDs and are characterized by large platelets, thrombocytopenia and bleeding tendencies in affected individuals (Figure 1A, Figure 1B, Figure 1C and Figure 1D) [3].

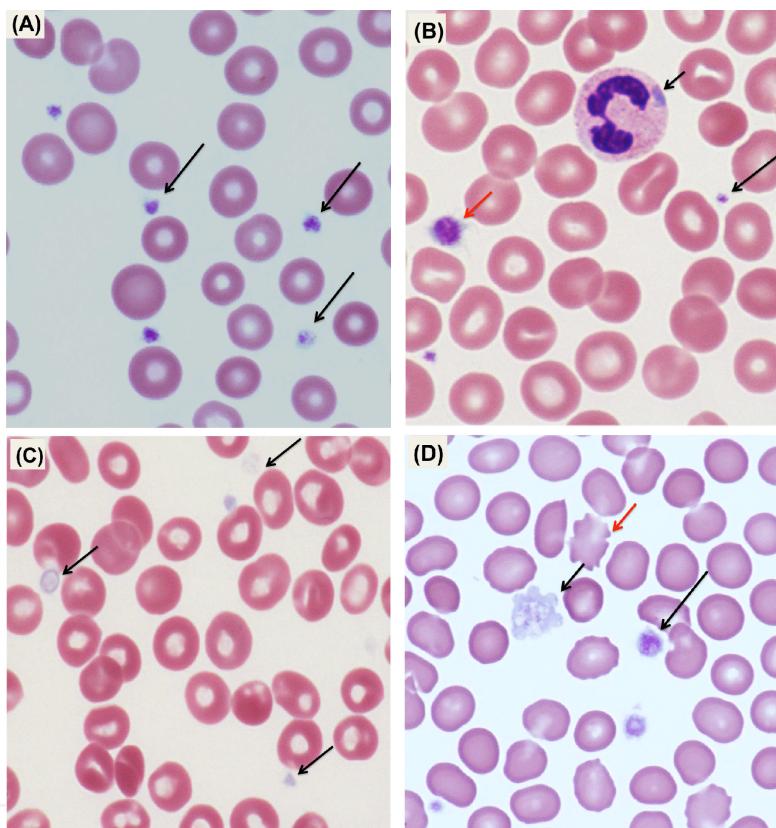


Figure 1. A normal blood film and three blood films demonstrating macrothrombocytopenia associated with mutations in different genes (*MYH9*, *NBEAL2* and *GFI1B*, respectively). (A) A blood film with platelets of normal appearance (black arrows). (B) *MYH9*-related disorder with characteristic inclusion bodies in the neutrophils (small black arrow) and large platelets (red arrow). Normal-sized platelets are also seen (long black arrow). (C) Gray platelet syndrome showing distinctive pale or gray platelets (black arrows). (D) *GFI1B*-related thrombocytopenia (c.880-881insC mutation) resulting in red cells with atypical shapes and sizes (red arrow) and thrombocytopenia with platelets that appear large with normal granulation (long black arrow) as well as hypogranular or gray (short black arrows).

Unfortunately, inherited macrothrombocytopenia is under-recognized with the presence of large platelets on blood film examination often leading to a misdiagnosis of immune thrombocytopenic purpura (ITP), resulting in subsequent inappropriate treatment with steroids or in some cases removal of the spleen [4]. Diagnostic algorithms have traditionally been based around biological laboratory tests examining functional properties and activation pathways

in isolated platelets [3, 5–7]. This phenotypic approach is poorly standardized, technically difficult and not easily reproducible [6–11]. In addition, numerous pre-analytical variables may affect phenotypic test results. These variables include the effect of food (garlic), alcohol, drugs (herbal remedies, non-steroidal anti-inflammatory drugs, anti-platelet medications) and stimulants (smoking and caffeine) on platelet function, activation of platelet samples during venipuncture and transport necessitating careful sample handling as well as the relatively large volume of blood needed (which becomes a major problem when assessing pediatric samples) [12–14]. Despite these complex phenotypic tests, many cases remain without a definitive diagnosis.

Genetic technology may overcome many of the problems surrounding phenotypic testing for thrombocytopenia as DNA is stable, can easily be transported long distances and is not affected by diet or drugs. Moreover, genetic-based tests have provided opportunities to reduce redundancy and heterogeneity of diagnostic algorithms and have shifted our ability to describe inherited platelet disorders from a level of the defective platelet pathway involved, to a molecular level.

The Sanger sequencing method [15] has long been considered the “gold standard” technology to rapidly analyze small regions across a limited number of samples, but it is not suited to screening large numbers of genes in multiple patients [16]. The emergence of next-generation sequencing (NGS) technologies as a diagnostic approach has been able to generate more test sequence increasing the number of gene targets and decreasing the costs [17, 18]. Human whole genome sequencing (WGS) or whole exome sequencing (WES) [19, 20] have proven to be clinically appropriate and practical modalities in describing new genetic mutations in families and identifying known pathogenic mutations in individuals formerly without a diagnosis [17].

Testing approaches may vary depending on whether a novel genetic mutation is likely. WGS and WES are powerful platforms in discovering novel causal variants in individuals with rare penetrant monogenic disorders [21], whilst a candidate gene approach allows assessment of known mutations in genes causing clinical phenotypes.

Whole genome approaches incorporating NGS have recently reported novel mutations in an essential platelet transcription factor GFI1B [22, 23], and a WES approach followed by targeted Sanger sequencing was used successfully to describe mutations in *ACTN1* causing macrothrombocytopenia [24, 25]. Acknowledging these advancements, we employed a targeted candidate gene approach to explore cases of suspected inherited macrothrombocytopenia that remained uncharacterized despite phenotypic testing and hypothesized this to be an effective approach to diagnose inherited macrothrombocytopenia.

2. Materials and methods

2.1. Patients

Diagnostic assessment of patients with uncharacterized thrombocytopenia was performed as part of a human research ethics committee approved study conducted in accordance with the Declaration of Helsinki.

Following informed written consent, 20 ml of blood was taken from an antecubital vein and collected into EDTA tubes. This blood was easily transported, in some cases, over 1,000 km between diagnostic sites in Australia.

A total of 95 patient DNA samples were analyzed. This included two internal controls for which DNA-based diagnosis had previously been established by Sanger sequencing.

32 male patients (mean age 37.4 years, range 18–92 years) and 44 female patients (mean age 38.7 years, range 18–79 years) were included in the NGS assay. The mean age of the cohort was 38.1 years (range 18–92 years). Sixteen de-identified DNA samples were received from referring institutions for which no additional laboratory data were available.

Phenotypic testing data were available for 59 (62.1%) individuals. This included platelet functional analysis (PFA) ($n = 25$, 26.0% of the cohort), light transmission aggregometry / whole blood impedance aggregometry (LTA/WBIA) ($n = 39$, 41.3% of the cohort), flow cytometry ($n = 45$, 47.8% of the cohort) and electron microscopy ($n = 12$, 13% of the cohort). These phenotypic test results suggested a diagnosis to a “pathway level”, that is, a description to the level of the suspected defective biochemical pathway, in only 11 cases. Pathway orientated defects included, storage pool disorders ($n = 3$), platelet glycoprotein deficiency ($n = 3$), platelet signaling defects ($n = 2$), platelet secretion defects ($n = 2$) as well as α -granule disorder ($n = 1$).

2.2. DNA preparation

Genomic DNA (gDNA) was isolated from peripheral blood leukocytes using the Wizard® Genomic DNA purification kit (Promega, Alexandria, NSW, Australia). DNA quality and concentration were assessed using the Nanodrop™ 1000 spectrophotometer (Thermo Scientific, Scoresby, Vic, Australia) that measures the purity of DNA by the ratio of absorbance of molecules at 260 and 280 nm. Samples with ratios between 1.8 and 2.0 were accepted for analysis whilst ratios lower than this may represent the presence of contaminants and these samples were not processed further [26]. At least, 250 ng of input gDNA was prepared per sample.

2.3. Candidate gene identification and gene panel design

An extensive literature search using public databases was performed to assemble an initial candidate gene list of all genes reasonably hypothesized to have an impact on platelet number and size ($n = 173$). A final list of candidate genes ($n = 19$) was derived by including those genes in which mutations were known to be definitively associated with IPNDs (predominantly, macrothrombocytopenia) and by excluding genes, which although known to result in thrombocytopenia, could easily be identified by conventional and clinical methods characterized by distinct clinical phenotypes.

A TruSeq custom amplicon (TruSeq® Custom Amplicon Kit, Illumina Inc., Scoresby, Vic, Australia) specific for the target regions of the selected 19 genes (Table 1, *ACTN1*, *CD36*, *ETS1*, *F2R*, *FLI1*, *GATA1*, *GFI1B*, *GP1BA*, *GP1BB*, *GP6*, *GP9*, *ITGA2*, *ITGA2B*, *ITGB1*, *ITGB3*, *MYH9*, *NBEAL2*, *P2RY12*, *RUNX1*, *TUBB1*) was designed as an entire custom pool using the web-based software tool, Illumina Design Studio (Illumina Inc.). This generated 201 gene targets that were either exons or gene regions that were split into 632 amplicons, each of approximately

250 base pairs (bps). There were no undesignable targets and a total coverage of 91% was predicted for the panel.

Gene	Description (OMIM)	Inheritance	Disorder (abbreviation in this paper, OMIM entry)
<i>ACTN1</i>	Alpha-Actinin-1	AD	α actinin-related thrombocytopenia (α actinin-RT, 615193)
<i>CD36 (GPIV)</i>	Thrombospondin receptor (Glycoprotein IV)	AD	Familial thrombocytopenia with GPIV deficiency (nd, 608404)
<i>ETS1</i>	V-Ets avian erythroblastosis virus E26 oncogene homolog 1	nd	nd
<i>F2R</i>	Coagulation factor II (thrombin) receptor	nd	nd
<i>FLI1</i>	Friend leukaemia virus integration 1	AD	Paris-Trousseau syndrome / Jacobsen syndrome (TCPT/ JBS, 188025, 600588)
<i>GATA1</i>	GATA-binding protein 1	XL	GATA1-related disorders (GATA1-RD, 300367, 314050)
<i>GFI1B</i>	Growth factor-independent 1B	AD	GFI1B-related thrombocytopenia (GFI1B-RT, 187900)
<i>GP1BA</i>	Glycoprotein 1b-alpha polypeptide	AR	Bernard Soulier syndrome (BSS, 231200)
		AD	Platelet type-von Willebrand disease (PT-VWD, 177820)
		AD	Velocardiofacial syndrome (VCFS, 192430)
		AD	Mediterranean thrombocytopenia (nd, 153670)
<i>GP1BB</i>	Glycoprotein 1b-beta polypeptide	AR	Bernard Soulier syndrome (BSS, 231200)
<i>GP6</i>	Glycoprotein VI	AR*	Bleeding disorder, platelet type 11 (614201)
<i>GP9</i>	Glycoprotein IX	AR	Bernard Soulier syndrome (BSS, 231200)
<i>ITGA2</i>	Integrin, alpha-2	AR	GPIa/IIa deficiency (giant platelets and mitral valve insufficiency) (nd,nd)
<i>ITGA2B</i>	Integrin, alpha-2B	AD	Monoallelic ITGA2B/ITGB3-related thrombocytopenia (ITGA2B/ITGB3-RT, 187800)
<i>ITGB1</i>	Integrin, beta-1	AR	GPIa/IIa deficiency (giant platelets and mitral valve insufficiency) (nd,nd)
<i>ITGB3</i>	Integrin, beta-3	AD	Monoallelic ITGA2B/ITGB3-related thrombocytopenia (ITGA2B/ITGB3-RT, 187800)
<i>MYH9</i>	Myosin heavy-chain 9	AD	MYH9-related disease (MYH9-RD, 155100)
<i>NBEAL2</i>	Neurobeachin-like 2	AR	Gray platelet syndrome (GPS, 139090)
<i>P2RY12</i>	Purinergic receptor P2Y, G protein-coupled 12	AR*	Bleeding disorder, platelet type 8 (609821)
<i>RUNX1</i>	Runt-related transcription factor 1	AD	Platelet disorder, familial, with associated myeloid malignancy (FDP/AML, 601399)
<i>TUBB1</i>	Tubulin, beta-1	AD	β 1 Tubulin-related thrombocytopenia (β 1 tubulin-RT, 613112)

Table 1. Candidate gene list. OMIM, online Mendelian inheritance in man; AR, autosomal recessive; AD, autosomal dominant; XL, X-linked; nd, not defined, *In progress (OMIM)

2.4. Next-generation sequencing

The Truseq custom amplicon library preparation kit and the MiSeq Illumina sequencer platform (Illumina Inc.) were used to create the sequencing library and perform resequencing respectively. All steps were performed in-house according to the manufacturer's instructions [27, 28].

Library preparation was performed by enrichment of the target regions using an amplicon-based multiplex polymerase chain reaction (PCR) method. Here, a custom amplicon tube (CAT) containing upstream and downstream oligonucleotides specific for the target regions was hybridized to the unfragmented gDNA samples in a 96-well plate. Unbound oligonucleotides were then removed by a series of wash steps using manufacturer supplied reagents. A proprietary extension–ligation mix containing DNA polymerase and ligase (Illumina Inc.) extended and ligated the upstream bound oligonucleotide through the targeted region to the 5' end of the downstream oligonucleotide. The resulting extension–ligation products containing the targeted genomic region flanked by common sequences required for amplification were then amplified by standard PCR on a thermal cycler. The amplicon size (250 bps), the number of amplicons in the CAT (632 amplicons) and the type of input DNA (high quality) determined the number of PCR cycles ($n = 24$). The PCR reaction incorporated two unique, sample-specific, multiplexing index sequences (barcoding) that would later be used by the alignment software (MiSeq reporter) to identify individual samples following library pooling, and common adapters required for cluster generation. PCR products were purified by AMPure XP beads (Beckman Coulter, Lane Cove, NSW, Australia) and the quantity of each library was normalized by an integrated bead-based method. Equal volumes of the normalized libraries were then combined, diluted in hybridization buffer (Illumina Inc.) and heat denatured.

The MiSeq Illumina instrument was used to resequence the pooled library by paired-end sequencing. The DNA library was immobilized to the single-use glass-based MiSeq flow cell through the adapter sequences. Bridge PCR amplification then generated clusters of clonal copies of each DNA molecule. These templates were then sequenced using platform-specific reversible dye terminator sequencing-by-synthesis chemistry. Sequence alignment to the reference genome (GRCh37/hg19) was performed using on-instrument software (MiSeq reporter software, Illumina Inc.) that aligned the reads in BAM format and outputted variant calls in.vcf files. Variant calls were generated using ANNOVAR software (<http://www.open-bioinformatics.org/annovar>) [29] with an acceptance threshold Q-score of 30, corresponding to a 1:1000 error rate and genomic datasets were viewed using the Integrative Genomics viewer (IGV) (www.broadinstitute.org/igv/) [30]. Sanger sequencing was performed to provide data for bases with insufficient coverage and validate variants of clinical significance.

2.5. Data analysis

The University of California, Santa Cruz (UCSC), genome browser (<http://genome.ucsc.edu>) was used for variant analysis and variants were cross-checked against databases including the NHLBI-Extended Sequencing Project (ESP), 1000 Genomes Project Database [31] and the Database of Single-Nucleotide Polymorphisms (dbSNP, <http://www.ncbi.nlm.nih.gov/SNP/>). Bioinformatic tools, Sorting Intolerant From Tolerant (SIFT, <http://sift.jcvi.org/>) [32], Polymorphism Phenotyping-2 (PolyPhen-2, <http://genetics.bwh.harvard.edu/pph2/>) [33] and Mutation

taster (<http://www.mutationtaster.org/>) [34] were used to predict variant effects on protein structure and function in the cases of variants lacking published literature.

2.6. Nomenclature and descriptions for variant reporting

All variants identified were annotated according to Human Genome Variation Society (HGVS) nomenclature for clinical reporting (<http://www.hgvs.org>). The variant elements included gene name, zygosity, cDNA nomenclature, protein nomenclature, exon number and clinical assertion.

Descriptions of sequence variations were adapted from the American College of Medical Genetics and Genomics (ACMG) recommendations for standards for interpretation and reporting of sequence variations and are listed below [35]:

Pathogenic: The sequence variation has been reported in the literature and is a recognized cause of the disorder.

Likely pathogenic: The sequence variation is previously unreported and is of the type that is expected to cause the disorder.

Variant of uncertain significance (VUS): The sequence variation is previously unreported and is of the type which may or may not be causative of the disorder.

Likely non-pathogenic: The sequence variation is previously unreported and is probably not causative of disease.

Non-pathogenic: The sequence variation is previously reported and is a recognized neutral variant.

3. Results

3.1. Next-generation sequencing platform performance

Next-generation sequencing on the Illumina platform produced 13 690 589 (96.74%) reads that passed initial filtering. This process removes any clusters demonstrating excessive intensity corresponding to bases other than the called base. Only reads that passed the quality filter were assigned a quality score. A quality score of Q30 was accepted in the run predictive of an error probability of $\leq 0.1\%$. One sample was excluded from analysis due to poor DNA quality that generated poor-quality scores across all genomic regions.

Overall coverage across all genomic targets was 92.3%. This was consistent with the initial software prediction.

3.2. Candidate gene panel results

A total of 703 non-synonymous variants were detected; 75 of these variants were novel and had not been reported in the dbSNP database. An average of eight non-synonymous variants was detected per patient.

Two individuals with known mutations in *GFI1B*, *GP1BA* and *GP9* by Sanger sequencing were included as controls. NGS successfully called the first, *GFI1B* c.880-881insC, but failed to detect the second, a patient with a phenotype consistent with the inherited macrothrombocytopenia Bernard-Soulier syndrome (BSS). This patient's genotype had previously been confirmed by Sanger sequencing and included mutations in both the *GPIBA* (*GPIBA* c.2217C>T) and the *GP9* genes (c.1829A>G and c.1859T>G). Failure to detect these mutations may have been caused by sequencing errors introduced by GC-rich motifs in these regions [36, 37].

Pathogenic mutations were detected in 16 individuals (17.4% of the cohort) whilst 36 individuals (39.1%) had VUS and 40 individuals (43.0%) were without identifiable pathogenic mutations (Table 2, Table 3).

Genes	Number of individuals with pathogenic mutations	Number of mutations detected of uncertain significance
<i>ACTN1</i>	0	8
<i>GP1BA</i>	1**	2
<i>GP1BB</i>	0	2
<i>GP9</i>	0	1
<i>MYH9</i>	6	3
<i>TUBB1</i>	0	3
<i>NBEAL2</i>	1	7
<i>FLI1</i>	0	1
<i>GATA1</i>	0	3
<i>GFI1B</i>	3	2
<i>RUNX1</i>	2**	0
<i>CD36</i>	0	13
<i>F2R</i>	0	0
<i>GP6</i>	0	5
<i>ITGA2</i>	0	4
<i>ITGA2B</i>	3*	6
<i>ITGB1</i>	0	0
<i>ITGB3</i>	0	0
<i>P2RY12</i>	0	0
Total Number	16	60 mutations in 36 individuals
Number of individuals without pathogenic mutations identified: 40		

*Parents heterozygous; child with homozygous mutation giving rise to a Glanzmann thrombasthaenia phenotype.

** These mutations are likely pathogenic. That is, the detected variation is unreported in the literature to date, however, based on the type of variation, it's deleterious effect predicted using bioinformatic tools (see data analysis) and the associated phenotypic data, is of the type to cause the disorder

Table 2. Mutations detected in the candidate genes. Genes affecting the platelet cytoskeleton (top, white shading), the platelet granules (light gray shading) and platelet-related transcription factors (dark gray shading).

Gene	Chromosome	Zygosity	Nucleotide change	Protein alteration	Exon
<i>GP1BA</i> **	17	Heterozygous	c.1432delT	p.Phe478fs	2
<i>MYH9</i>	22	Heterozygous	c.283G>A	p.Ala95Thr	2
		Heterozygous	c.287C>T	p.Ser96Leu	2
		Heterozygous	c.2104C>T	Arg702Cys	17
		Heterozygous	c.4339G>C	p.Asp1447His	31
<i>NBEAL2</i>	3	Compound heterozygous	c.5935C>T	p.Arg1979Trp	37
			c.7103dupA	His2368fs	45
<i>GFI1B</i>	9	Heterozygous	c.503G>T	p.Cys168Phe	4
<i>RUNX1</i> **	21	Heterozygous	c.503–504insAT	p.Ile168fs	3
<i>ITGA2B</i> *	17	Homozygous	c.138–139insT	p.Gly47fs	1

* Parents heterozygous. Child with homozygous mutation giving rise to a Glanzmann thrombasthenia phenotype.

** Mutations are likely pathogenic.

Table 3. Pathogenic genetic variants detected: nucleotide cDNA changes and corresponding protein alterations.

The candidate array was successful in detecting mutations in genes commonly associated with macrothrombocytopenia and included a total of nine *MYH9* mutations (six of which had previously been reported in the literature as pathogenic and three of which are of uncertain significance) (Figure 2) and a compound heterozygous mutation of *NBEAL2* in keeping with Gray platelet syndrome.

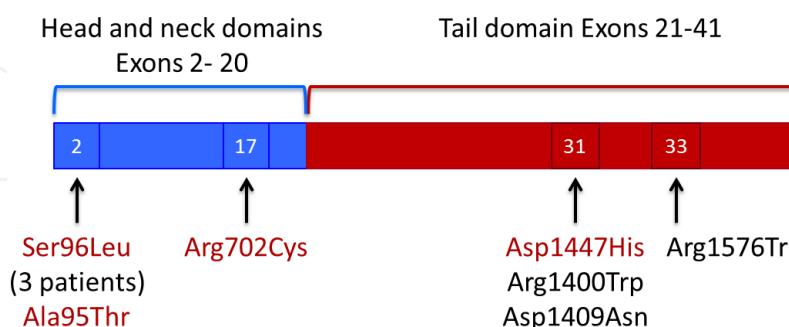


Figure 2. *MYH9* variants detected in the candidate gene panel. Exons 2–20 encode the head and neck domains of NMMHC IIA (Blue block). Exons 21–41 encode the tail domains. Mutations were detected in exons 2, 17, 31 and 33. Six pathogenic mutations (red text) and three variants of uncertain significance (black text) were detected.

A homozygous mutation of *ITGA2B* was also detected and confirmed a suspected Glanzmann thrombasthenia phenotype. Several transcription factor variants were found, including a *FLI1*

mutation of uncertain significance in one patient, three *GATA1* mutations of uncertain significance in three individuals from two families, three pathogenic *GFI1B* mutations in three individuals from two families and two of uncertain significance in two individuals in another two families. *RUNX1* mutations were identified in three individuals from three families; two of these were considered likely pathogenic, whilst the third was shown to represent a false positive result (*RUNX1*, heterozygous, stop/gain, c. 966T>G (p.Tyr322X), exon 6). False positivity was confirmed by Sanger sequencing that showed a wild-type sequence across that region.

Sanger sequencing was also performed in selected samples across regions of low coverage ($Q < 30$) from those genes in which the clinical significance is widely accepted and included, *GP9*, *GP1BA*, *GPIBB*, *FLI1* exon 3, *FLI1* exon 9, *MYH9* exon 20, *MYH9* exon 37 and *GFI1B* exon 5. This confirmatory step detected a novel mutation in *FLI1* [38], not identified by NGS.

4. Discussion

The diagnosis of IPFD and IPNDs using classic phenotypic methods poses a challenge to clinicians and laboratory scientists due to lack of consensus over classification and diagnostic criteria, poor standardization of tests and heterogeneity of traditional diagnostic approaches [6]. This diagnostic conundrum is evident in our cohort where only 11 patients received a suspected diagnosis to a pathway level following multiple previous phenotypic tests. In addition, only 62% of patients received any form of phenotypic test, reflecting the difficulty of accessing these specialized techniques in many centers.

Sanger sequencing is widely regarded as a reliable platform for routine diagnostic genetic testing and small-scale projects. However, effective analysis of numerous disease-associated genes by Sanger sequencing in a diagnostic setting is time-consuming, expensive and not always feasible [18]. A candidate gene array was selected as it has the potential to simultaneously analyze all of the selected coding regions of disease-targeted genes. Moreover, relative to WES and WGS, it provides good gene coverage and representation of exons, is relatively fast and cheap and minimizes the problems with unexpected findings and development of complex downstream bioinformatic pipelines for analysis [39].

We have demonstrated that high-quality sequence data can be generated from a candidate group of platelet genes using the Illumina MiSeq platform. Our candidate gene panel comprised 19 genes associated with IPNDs, predominantly inherited macrothrombocytopenia. Pathogenic mutations were detected in 17.4% of the cohort. The most number of mutations was detected in the *MYH9* gene. *MYH9*-related disorders are the most common forms of inherited thrombocytopenia and are frequently under-recognized or misdiagnosed as immune ITP [40–42]. Immunofluorescence staining of the peripheral blood film demonstrating abnormal clustering of non-muscle myosin heavy chain IIA (NMMHC IIA), seen as Döhle bodies on the blood film is regarded as a suitable diagnostic test [40], but is not available at all centers. A strong genotype–phenotype relationship is recognized in these disorders, with mutations affecting the motor (head and neck) region of NMMHC-IIA causing more severe

thrombocytopenia and a higher risk for nephritis, cataracts and deafness, whilst those mutations affecting the tail region cause less severe thrombocytopenia and extra-hematological manifestations [43, 44]. Genetic confirmation of *MYH9*-related disorders, therefore, has prognostic significance. In our group of patients, three pathogenic mutations in five individuals were detected and were predicted to affect the motor region of NMMHC IIA. Knowledge of these mutations has provided an opportunity to offer advice regarding additional non-hematological surveillance tests such as audiograms, renal function assessments and ophthalmological screening for cataracts [40, 41, 45].

Transcription factors are the key regulators for the development of the hemostatic platelet from blood stem cells. Stem cells differentiate into a bipotent megakaryocyte-erythroid progenitor, then a committed megakaryocyte that undergoes endoreplication prior to extending proplatelet extensions from the cytoplasm into the bone marrow sinusoid forming platelets [46]. This complex differentiation pathway is orchestrated by the activation and repression of groups of genes important for blood cell development via transcription factors [46, 47]. The candidate gene panel contained four genes that encode hemopoietic transcription factors, *FLI1*, *GATA1*, *GFI1B* and *RUNX1*. Definitive diagnosis of platelet disorders caused by mutations in these genes solely by phenotypic testing is not possible. We detected a pathogenic mutation in one of these genes, *GFI1B*, and likely pathogenic mutations, in *RUNX1*. The *RUNX1* gene is responsible for the familial platelet disorder with a predisposition to acute myeloid leukemia (FPD/AML) [48]. The propensity to develop acute leukemia is determined by the action of the variant, with dominant negative and haploinsufficient mutations having different leukemogenic risk. The former has a higher risk (up to 40% in some reports) of progression to AML or myelodysplastic syndrome [49–51]. Other factors include the residual level of activity of wild-type *RUNX1* [52], deregulation induced by dominant negative mutations on hematopoietic stem cell genes such as *NR4A3* [53] as well as effects on p53 genes-dependent genes that induce genomic instability of the granulomonocytic precursors [52]. The median age of onset of progression to myelodysplastic syndrome / acute leukemia is 33 years of age, and therefore, the detection of two, likely pathogenic, *RUNX1* mutations by our candidate gene panel is of obvious importance [49]. Despite their adverse risk, clinical guidelines regarding the best way to counsel, test and manage these patients and their family members are lacking and recommendations are largely based on expert opinion [54]. Initial referral to a specialist team comprising a physician as well as genetic counselor is recommended, as well as, full blood count analysis, bone marrow biopsy (to detect occult malignancy) and full human-leukocyte antigen (HLA) typing of patients and their first-degree relatives (in the event a bone marrow transplant is required in the future). A biannual follow-up schedule thereafter should be established to ensure close hematological surveillance [54]. *GFI1B* is another transcription factor that plays an essential role in hematopoiesis [46, 55]. Two recent publications [22, 23] described mutations in the DNA-binding zinc finger domain of *GFI1B* causing an autosomal dominant bleeding disorder in affected families. Our candidate gene array detected another mutation in a non-DNA-binding zinc finger domain of *GFI1B* (*GFI1B* c.503G>T). Further characterization of this c.503G>T mutation indicates a milder platelet phenotype with less clinical bleeding symptomatology than the DNA-binding mutants [56] (Figure 3). The detection of this non-DNA-binding mutation has afforded us an opportunity to propose a

genotype–phenotype relationship associated with mutations in two different regions of GFI1B. This is important to enable classification, aid diagnosis and inform treatment strategies.

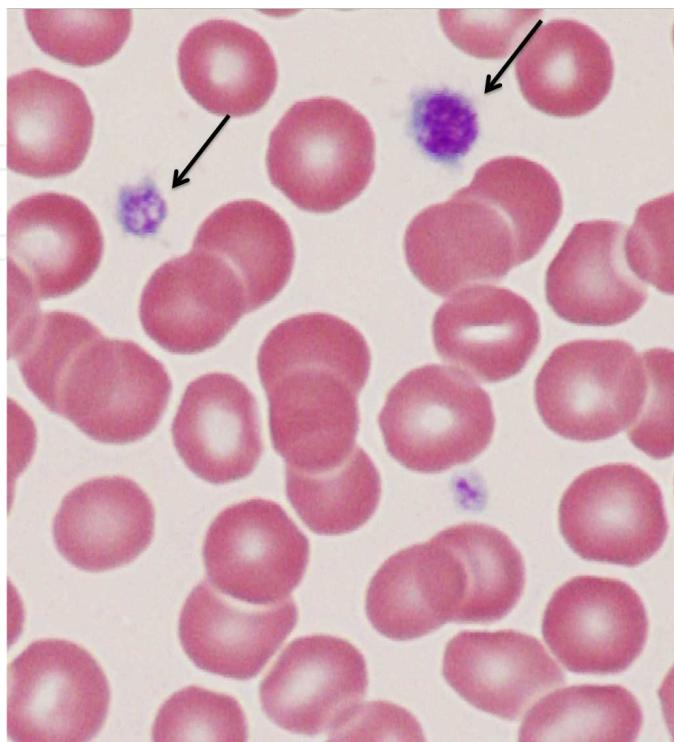


Figure 3. The blood film of an affected individual with the *GFI1B* c.503G>T mutation demonstrating macrothrombocytopenia. Platelets show normal granulation unlike the platelets seen in individuals with the *GFI1B* c.880-881insC mutation (Figure 1D) that have a heterogeneous appearance (some platelets appear hypogranular or gray whilst others have normal granulation).

The yield of pathogenic variants reported above may have been improved by more stringent patient selection criteria. In this study, all patients suspected of an inherited thrombocytopenia by treating hematologists were included regardless of the platelet phenotype. That is, not all patients demonstrated macrothrombocytopenia. In addition, in 16 cases only DNA was received and the platelet phenotype was not known. Noting that 15 of the 19 genes on the candidate panel are known to cause macrothrombocytopenia and that only 5 genes on the panel (*ETS1*, *P2RY12*, *F2R*, *GP6*, *RUNX1*) have an uncertain platelet phenotype or otherwise known to cause functional disorders with normal-sized platelets, the pre-test probability of detecting a pathogenic variant in samples where macrothrombocytopenia was not present was low. Furthermore, this candidate array was performed in a research laboratory and therefore included genes (*ETS1* and *F2R*) where the association with inherited thrombocytopenia is not well delineated. Exclusive inclusion of genes with clear evidence of disease association may further improve the diagnostic yield.

Variants of uncertain significance (VUS) were detected in over a third of the cohort (39.1%). Thirteen samples contained more than one VUS. One sample contained five VUS in five different genes (*GFI1B*, *ITGA2*, *MYH9*, *NBEAL2* and *TUBB1*). In many instances, these variants

were novel. It is likely, as knowledge of the genes causing inherited platelet bleeding disorders increases, this percentage will decrease, the VUS either becoming recognized as pathogenic or definitely non-pathogenic. Our analytical pathway used three bioinformatics tools (SIFT, PolyPhen2, Mutation taster) in variants lacking published literature to assist variant annotation. Bioinformatic tools using sequence and/or structure to predict the effects of amino acid substitutions on protein function have been developed following observations that disease-causing mutations are more likely to occur at positions that show evolutionary conservation and/or common structural features which enable them to be distinguished from neutral substitutions [57–60]. These tools serve to guide future experiments and should not be used solely as a clinical predictor of pathogenicity. Consider the *ACTN1* missense mutation (*ACTN1*, heterozygous, c.580G>A [p.Gly194Arg], exon 6, rs145918825) detected in our candidate gene array. It is predicted to disturb the calponin homology domain (CHD) within the actin-binding domain (ABD) of α -actinin (an important platelet structural protein). All of the mutations described in the literature to date have identified *ACTN1* mutations within the functional domains (ABD and the C-terminal calmodulin-like domain [CaM]) but not within the spacer spectrin repeats [25, 61, 62]. Bioinformatic tools were applied to this variant. It is predicted to be deleterious by SIFT (sequence homology-based tool), whereas PolyPhen-2 (structure/sequence based tool) predicts the amino acid alteration to be benign. This highlights two points. Firstly, it is advisable that predictions are made by integrating the results from several tools as reliance on one tool may lead to incorrect annotation [63], and secondly, that bioinformatic tools provide predictions only. In this case, the functional consequences of the *ACTN1* DNA variant are yet to be described and thus the variant may or may not be significant. Further family studies and additional structural analyses of the protein may clarify the pathogenicity of the variant [35].

Coverage is a crucial metric for establishing accuracy as well as analytical sensitivity and specificity of a NGS testing platform [64]. Coverage requirements depend on the application of the NGS test. In general, sequencing more reads will increase the power of the assay. We determined the necessary coverage level based on recommendations forwarded by the Royal College of Pathologists of Australasia [65] whose guidance is in compliance with National Pathology Accreditation Advisory Council (NPAAC) standards for testing of human nucleic acids [66] and combined this advice with recommendations from published literature and other international bodies such as the ACMG [35]. Our accepted Q score (Q30) was met in 92.3% of all genomic targets and in 97% of exonic targets. The read coverage distribution curve displayed a classic Poisson-like distribution indicating uniformity of coverage, this data accompanied by the high quality of base calls suggested that the NGS platform is able to deliver reliable sequence data. However, there were also areas of lower coverage where the platform did not perform as well, and lacked sensitivity. These regions were identified at genomic targets in *FLI1*, *GP1BA*, *GP1BB*, *GP9*, *ITGB1* and *NBEAL2* and were predicted in the design studio report. Two false negative results were confirmed in regions where coverage was low. The first being the failed detection of *GP1BA* and *GP9* mutations in the second internal control sample and the second was a novel pathogenic mutation in *FLI1* that was confirmed by Sanger sequencing and additional laboratory investigations. To ensure coverage of the respective amplicons over the *GP9* region, parallel Sanger sequencing was performed. Targeted Sanger

sequencing was also performed for *GP1BA* and *GP1BB* in cases in which phenotypic details had been provided by the referring clinician and where confident exclusion of a variant in those genes was necessary. Sanger sequencing performed over these regions did not detect additional mutations. Only a single false positive result was confirmed by Sanger sequencing (*RUNX1*, stop/gain, c.966T>G). This suggested good platform specificity. The question as to whether confirmatory Sanger sequencing need be performed is debated in the literature [39, 67]. Proponents argue that it is required to confirm a diagnosis as well as remove incorrect calls introduced by experimental errors. Whereas, opponents argue, in the setting where the NGS platform performance metrics have been established to be comparable to Sanger sequencing performance measures, a strategy dictated by the degree of coverage per nucleotide be adopted. Suggesting that parallel Sanger sequencing need not be performed as long as the coverage is >30 times per nucleotide at that genomic target, adding that confirmatory testing be performed where coverage is less than 20 times, and be determined by visual inspection with coverage between 20 and 30 times. Authors commented that the laboratory may also simply elect to exclude the target from the report if Sanger sequencing is not performed despite low coverage [39].

An important aspect of the post-analytical process is the timely provision of a genomic test report. In the setting of inherited platelet disorders, a false negative interpretation may lead to a falsely conservative bleeding prophylactic strategy at the time of surgery, in turn, placing the individual at a potentially increased risk of bleeding. A false positive result, on the other hand, may cause undue stress to the individual and their family. A genomic test report was therefore carefully and consistently structured taking into consideration recommendations from professional bodies such as the RCPA [65] and ACMG [68]. The report (Appendix 1) contained a summary of the genes analyzed and reflected the scope and limitation of the assay and indicated the context in which the test was performed. A clear, succinct, interpretative comment was made regarding the detected variant. This indicated whether or not the detected variant was associated with the clinical phenotype and highlighted variants of uncertain significance. The body of the report detailed, in a structured format (see materials and methods), any detected pathogenic or clinically relevant variants and whether these had been previously described. An interpretation on the significance of the detected variant was supported by relevant references where possible, and recommendations regarding additional validation tests and /or genetic counseling and clinical screening were provided. Following the main body of the report, DNA variants that were considered to be non-pathogenic were listed. The report was concluded by a description of the test method and limitations thereof.

In conclusion, our study has demonstrated the potential to successfully diagnose inherited macrothrombocytopenia in cases that remained uncharacterized by traditional phenotypic approaches. Optimization of this format will provide patients an opportunity for a “one stop, one step” testing platform that is cost-effective and not affected by the pre-analytical variables that hinder current testing methods based on functional analysis of platelets. However, the translation of NGS from a powerful research tool into the clinical laboratory will require co-operation from international groups to establish best practice, quality and reporting standards

for these conditions, as well as to generate reliable databases that link platelet phenotypes to genotypes to provide best hemostasis clinician advice.

5. Appendix

Test performed: Candidate gene array of 19 genes (*ACTN1, CD36, F2R, FLI1, ETS1, GATA1, GFI1b, GP1BA, GP1BB, GP6, GP9, ITGA2, ITGA2B, ITGB1, ITGB3, MYH9, NBEAL2, P2RY12, RUNX1, TUBB1*) using the Illumina MiSeq next-generation sequencing platform.

Please Note:

This test has been performed for research purposes only and has not been NATA accredited in our laboratory.

Validation by Sanger sequencing has not been performed on clinically significant or novel detected variants and should be considered by the referring clinician.

Result: A mutation in a gene known or predicted to be associated with decreased platelet counts and/ or function has been identified. A second variant of uncertain significance has also been identified.

DNA variants: Variant 1: *MYH9*, Heterozygous, c.287C>T (p.Ser96Leu), Exon 2, rs121913657, pathogenic.

Variant 2: *NBEAL2*, Heterozygous, c.6178C>T (p.Arg2060Cys), exon37, uncertain significance.

Previously described: Variant 1: Yes (rs121913657)

Variant 2: No.

Interpretation: A heterozygous 287C-T transition in the *MYH9* gene, resulting in a ser96-to-leu (S96L) substitution, has been predicted to disturb the helical region of the protein resulting in *MYH9*- related disorder (Epstein syndrome).

The pathogenicity of variant 2 is uncertain as information regarding this mutation is not available in the reported literature. Note that the classification of variants of uncertain/ unknown significance may change over time if additional information on these conditions becomes available in the reported literature.

References: Arrondel C, et al. Expression of the non-muscle myosin heavy chain IIA in the human kidney and screening for *MYH9* mutations in Epstein and Fechtner syndromes. J Am Soc Nephrol 2002;13: 65–74.

Utsch B, et al. Bladder exstrophy and Epstein type congenital macrothrombocytopenia: evidence for a common cause? (Letter) Am J Med Genet 2006;140A:2251–3.

Kunishima S, et al. Immunofluorescence analysis of neutrophil non-muscle myosin heavy chain-A in *MYH9* disorders: association of subcellular localization with *MYH9* mutations. Lab Invest 2003;83:115–22.

Recommendations: The pathogenicity of detected candidate variants should be validated independently by Sanger sequencing. Where necessary, the functional significance of these variants should be confirmed independently by appropriate biological assays to replicate the phenotype of this patient.

MYH9-related disorders have an autosomal dominant inheritance. Genetic counselling is recommended for this individual and their family. Family screening may be appropriate after appropriate genetic counselling.

DNA variants detected of unlikely clinical significance:

NBEAL2, Heterozygous, c.1531C>G (p.Arg511Gly), Exon 13, rs11720139, likely **non-pathogenic**. *GP6*, Homozygous, c.691G>A (p.Ala231Thr), Exon 6, rs2304167, likely **non-pathogenic**. *MYH9*, Heterozygous, c.4876A>G (p.Ile1626Val), Exon 34, rs2269529, likely **non-pathogenic**.

Test method:

A TruSeq custom amplicon specific for the target regions of 19 genes, *ACTN1*, *CD36*, *F2R*, *FLI1*, *ETS1*, *GATA1*, *GFI1b*, *GP1BA*, *GP1BB*, *GP6*, *GP9*, *ITGA2*, *ITGA2B*, *ITGB1*, *ITGB3*, *MYH9*, *NBEAL2*, *P2RY12*, *RUNX1*, *TUBB1* was designed using Illumina design studio (Illumina, Inc, San Diego, CA, USA). Next-generation sequencing was performed using the MiSeq Illumina sequencer platform (Illumina, Inc.). Obtained sequences were aligned to the reference genome (GRCh37/hg19) using MiSeq reporter software (Illumina, Inc.) and the genomic datasets viewed using the Integrative Genomics viewer (IGV) (www.broadinstitute.org/igv/). Variant calls were generated using ANNOVAR software (<http://www.openbioinformatics.org/annovar>) with an acceptance threshold Q-score of 30, corresponding to a 1:1000 error rate. Sanger sequencing was performed to provide data for bases with insufficient coverage. The University of California, Santa Cruz (UCSC), genome browser (<http://genome.ucsc.edu>) was used for variant analysis and variants were cross-checked against databases including the NHLBI-extended sequencing project (ESP), 1000 genomes project database and the Database of Single-Nucleotide Polymorphisms (dbSNP). Bioinformatic tools (SIFT, PolyPhen-2 and Mutation taster) were used to predict variant effects on protein structure and function in the cases of variants lacking published literature.

Limitations: Overall gene coverage was 97% using this format. Therefore, it is possible that the genomic region where a disease causing mutation exists in the proband was not captured and therefore was not detected.

It is also possible that a particular genetic mutation was not recognised as the underlying cause of the genetic disorder due to incomplete scientific knowledge of the impact of all variants at this point in the literature.

Reported by:

An example of a NGS report.

Author details

David J. Rabbolini^{1,2}, Marie-Christine Morel Kopp^{1,2}, Sara Gabrielli^{1,2}, Qiang Chen^{1,2}, William S. Stevenson^{1,2} and Christopher M. Ward^{1,2*}

*Address all correspondence to: cward@med.usyd.edu.au

1 Northern Blood Research Centre, Kolling Institute of Medical Research, The University of Sydney, Sydney, Australia

2 Department of Haematology and Transfusion Medicine, Royal North Shore Hospital, Sydney, Australia

References

- [1] Brass LF. Thrombin and platelet activation. *Chest* 2003;124(3 Suppl):18S–25S.
- [2] Shattil SJ, Kashiwagi H, Pampori N. Integrin signaling: the platelet paradigm. *Blood* 1998;91(8):2645–57.
- [3] Balduini CL, Cattaneo M, Fabris F, Gresele P, Iolascon A, Pulcinelli FM, et al. Inherited thrombocytopenias: a proposed diagnostic algorithm from the Italian Gruppo di Studio delle Piastrine. *Haematologica* 2003;88(5):582–92.
- [4] Kunishima S, Saito H. Congenital macrothrombocytopenias. *Blood Rev* 2006;20(2):111–21.
- [5] Noris P, Pecci A, Di Bari F, Di Stazio MT, Di Pumo M, Ceresa IF, et al. Application of a diagnostic algorithm for inherited thrombocytopenias to 46 consecutive patients. *Haematologica* 2004;89(10):1219–25.
- [6] Gresele P, Harrison P, Bury L, Falcinelli E, Gachet C, Hayward CP, et al. Diagnosis of suspected inherited platelet function disorders: results of a worldwide survey. *J Thromb Haemost* 2014;12(9):1562–9.
- [7] Gresele P. Diagnosis of inherited platelet function disorders: guidance from the SSC of the ISTH. *J Thromb Haemost* 2015;13(2):314–22.
- [8] Israels SJ, El-Ekiaby M, Quiroga T, Mezzano D. Inherited disorders of platelet function and challenges to diagnosis of mucocutaneous bleeding. *Haemophilia* 2010;16(Suppl 5):152–9.
- [9] Hayward CP, Eikelboom J. Platelet function testing: quality assurance. *Semin Thromb Hemost* 2007;33(3):273–82.

- [10] Hayward CP, Rao AK, Cattaneo M. Congenital platelet disorders: overview of their mechanisms, diagnostic evaluation and treatment. *Haemophilia* 2006;12(Suppl 3):128–36.
- [11] Moffat KA, Ledford-Kraemer MR, Nichols WL, Hayward CP, North American Specialized Coagulation Laboratory A. Variability in clinical laboratory practice in testing for disorders of platelet function: results of two surveys of the North American Specialized Coagulation Laboratory Association. *Thromb Haemost* 2005;93(3):549–53.
- [12] Favaloro EJ, Lippi G, Franchini M. Contemporary platelet function testing. *Clin Chem Lab Med* 2010;48(5):579–98.
- [13] Harrison P, Mackie I, Mumford A, Briggs C, Liesner R, Winter M, et al. Guidelines for the laboratory investigation of heritable disorders of platelet function. *Br J Haematol* 2011;155(1):30–44.
- [14] George JN, Shattil SJ. The clinical importance of acquired abnormalities of platelet function. *N Engl J Med* 1991;324(1):27–39.
- [15] Sanger F, Nicklen S, Coulson AR. DNA sequencing with chain-terminating inhibitors. *Proc Natl Acad Sci USA* 1977;74(12):5463–7.
- [16] Bonetta L. Genome sequencing in the fast lane. *Nat Meth* 2006;3(2):141–7.
- [17] Johnsen JM, Nickerson DA, Reiner AP. Massively parallel sequencing: the new frontier of hematologic genomics. *Blood* 2013;122(19):3268–75.
- [18] Shendure J, Ji H. Next-generation DNA sequencing. *Nat biotechnol* 2008;26(10):1135–45.
- [19] Harismendy O, Ng PC, Strausberg RL, Wang X, Stockwell TB, Beeson KY, et al. Evaluation of next generation sequencing platforms for population targeted sequencing studies. *Genome Biol* 2009;10(3):R32.
- [20] Smith ML, Wheeler KE. Weight-based heparin protocol using antifactor Xa monitoring. *Am J Health-Syst Pharm* 2010;67(5):371–4.
- [21] Bamshad MJ, Ng SB, Bigham AW, Tabor HK, Emond MJ, Nickerson DA, et al. Exome sequencing as a tool for Mendelian disease gene discovery. *Nat Rev Genet* 2011;12(11):745–55.
- [22] Monteferrario D, Bolar NA, Marneth AE, Hebeda KM, Bergevoet SM, Veenstra H, et al. A dominant-negative GFI1B mutation in the gray platelet syndrome. *N Engl J Med* 2014;370(3):245–53.
- [23] Stevenson WS, Morel-Kopp MC, Chen Q, Liang HP, Bromhead CJ, Wright S, et al. GFI1B mutation causes a bleeding disorder with abnormal platelet function. *J Thromb Haemost* 2013;11(11):2039–47.

- [24] Kunishima S, Okuno Y, Yoshida K, Shiraishi Y, Sanada M, Muramatsu H, et al. ACTN1 mutations cause congenital macrothrombocytopenia. *Am J Hum Genet* 2013;92(3):431–8.
- [25] Bottega R, Marconi C, Faleschini M, Baj G, Cagioni C, Pecci A, et al. ACTN1-related thrombocytopenia: identification of novel families for phenotypic characterization. *Blood* 2015;125(5):869–72.
- [26] Thermo Scientific technical bulletin: 260/280 and 260/230 ratios [Internet]. 2008. Available from: <http://www.nanodrop.com/Library/T009-NanoDrop%201000-&-NanoDrop%208000-Nucleic-Acid-Purity-Ratios.pdf> [Accessed: 2015-08-02].
- [27] Illumina-Truseq Custom Amplicon Library Preparation Guide [Internet]. 2013. Available from: http://support.illumina.com/content/dam/illumina-support/documents/documentation/chemistry_documentation/samplepreps_truseq/truseqcstomamplicon/truseq-custom-amplicon-libraryprep-ug-15027983-c.pdf [Accessed: 2015-08-08].
- [28] Illumina- MiSeq System user guide [Internet]. 2014. Available from: https://support.illumina.com/content/dam/illumina-support/documents/documentation/system_documentation/miseq/miseq-system-guide-15027617-o.pdf [Accessed: 2015-08-08].
- [29] Wang K, Li M, Hakonarson H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res* 2010;38(16):e164.
- [30] Robinson JT, Thorvaldsdottir H, Winckler W, Guttman M, Lander ES, Getz G, et al. Integrative genomics viewer. *Nat biotechnol* 2011;29(1):24–6.
- [31] Gunter C. Genomics: a picture worth 1000 genomes. *Nat Rev Genet* 2010;11(12):814.
- [32] Kumar P, Henikoff S, Ng PC. Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. *Nat Protocol* 2009;4(7):1073–81.
- [33] Adzhubei IA, Schmidt S, Peshkin L, Ramensky VE, Gerasimova A, Bork P, et al. A method and server for predicting damaging missense mutations. *Nat Meth* 2010;7(4):248–9.
- [34] Schwarz JM, Cooper DN, Schuelke M, Seelow D. MutationTaster2: mutation prediction for the deep-sequencing age. *Nat Meth* 2014;11(4):361–2.
- [35] Richards CS, Bale S, Bellissimo DB, Das S, Grody WW, Hegde MR, et al. ACMG recommendations for standards for interpretation and reporting of sequence variations: revisions 2007. *Genet Med* 2008;10(4):294–300.
- [36] Quail MA, Smith M, Coupland P, Otto TD, Harris SR, Connor TR, et al. A tale of three next generation sequencing platforms: comparison of Ion Torrent, Pacific Biosciences and Illumina MiSeq sequencers. *BMC Genom* 2012;13:341.

- [37] Ross MG, Russ C, Costello M, Hollinger A, Lennon NJ, Hegarty R, et al. Characterizing and measuring bias in sequence data. *Genome Biol* 2013;14(5):R51.
- [38] Stevenson WS, Rabbolini DJ, Beutler L, Chen Q, Gabrielli S, Mackay JP, et al. Paris-Trousseau thrombocytopenia is phenocopied by the autosomal recessive inheritance of a DNA-binding domain mutation in FLI1. *Blood* [Epub ahead of print]. 2015. Available from: <http://dx.doi.org/10.1182/blood-2015-06-650887>. [Accessed 2015-08-27]
- [39] Sikkema-Raddatz B, Johansson LF, de Boer EN, Almomani R, Boven LG, van den Berg MP, et al. Targeted next-generation sequencing can replace Sanger sequencing in clinical diagnostics. *Hum Mutat* 2013;34(7):1035–42.
- [40] Savoia A, De Rocco D, Panza E, Bozzi V, Scandellari R, Loffredo G, et al. Heavy chain myosin 9-related disease (MYH9 -RD): neutrophil inclusions of myosin-9 as a pathognomonic sign of the disorder. *Thromb Haemost* 2010;103(4):826–32.
- [41] Balduini CL, Pecci A, Savoia A. Recent advances in the understanding and management of MYH9-related inherited thrombocytopenias. *Br J Haematol* 2011;154(2):161–74.
- [42] Althaus K, Greinacher A. MYH9-related platelet disorders. *Semin Thromb Hemost* 2009;35(2):189–203.
- [43] Althaus K, Greinacher A. MYH9-related platelet disorders. *Semin Thromb Hemost* 2009;35(2):189–203.
- [44] Pecci A, Panza E, Pujol-Moix N, Klerys C, Di Bari F, Bozzi V, et al. Position of non-muscle myosin heavy chain IIA (NMMHC-IIA) mutations predicts the natural history of MYH9-related disease. *Hum Mutat* 2008;29(3):409–17.
- [45] Althaus K, Greinacher A. MYH-9 related platelet disorders: strategies for management and diagnosis. *Transfusion medicine and hemotherapy : offizielles Organ der Deutschen Gesellschaft fur Transfusionsmedizin und Immunhamatologie*. 2010;37(5):260–7.
- [46] Doerks T, Copley RR, Schultz J, Ponting CP, Bork P. Systematic identification of novel protein domain families associated with nuclear functions. *Genome Res* 2002;12(1):47–56.
- [47] Tijssen MR, Ghevaert C. Transcription factors in late megakaryopoiesis and related platelet disorders. *J Thromb Haemost* 2013;11(4):593–604.
- [48] Song WJ, Sullivan MG, Legare RD, Hutchings S, Tan X, Kufrin D, et al. Haploinsufficiency of CBFA2 causes familial thrombocytopenia with propensity to develop acute myelogenous leukaemia. *Nat Genet*. 1999;23(2):166–75.
- [49] Liew E, Owen C. Familial myelodysplastic syndromes: a review of the literature. *Haematologica* 2011;96(10):1536–42.

- [50] Beri-Dexheimer M, Latger-Cannard V, Philippe C, Bonnet C, Chambon P, Roth V, et al. Clinical phenotype of germline RUNX1 haploinsufficiency: from point mutations to large genomic deletions. *Eur J Hum Genet* 2008;16(8):1014–8.
- [51] Matheny CJ, Speck ME, Cushing PR, Zhou Y, Corpora T, Regan M, et al. Disease mutations in RUNX1 and RUNX2 create nonfunctional, dominant-negative, or hypomorphic alleles. *Embo J* 2007;26(4):1163–75.
- [52] Antony-Debre I, Manchev VT, Balayn N, Bluteau D, Tomowiak C, Legrand C, et al. Level of RUNX1 activity is critical for leukemic predisposition but not for thrombocytopenia. *Blood* 2015;125(6):930–40.
- [53] Bluteau D, Gilles L, Hilpert M, Antony-Debre I, James C, Debili N, et al. Down-regulation of the RUNX1-target gene NR4A3 contributes to hematopoiesis deregulation in familial platelet disorder/acute myelogenous leukemia. *Blood* 2011;118(24):6310–20.
- [54] Churpek JE, Lorenz R, Nedumgottil S, Onel K, Olopade OI, Sorrell A, et al. Proposal for the clinical detection and management of patients and their family members with familial myelodysplastic syndrome/acute leukemia predisposition syndromes. *Leuk Lymphoma* 2013;54(1):28–35.
- [55] Vassen L, Okayama T, Moroy T. Gfi1b:green fluorescent protein knock-in mice reveal a dynamic expression pattern of Gfi1b during hematopoiesis that is largely complementary to Gfi1. *Blood* 2007;109(6):2356–64.
- [56] Rabbolini DJ, Morel-Kopp MC, Chen Q, Gabrielli S, Dunlop L, Brighton T, et al. Abstracts. *J Thromb Haemost* 2015;13:1–997 (abstract PO421).
- [57] Ng PC, Henikoff S. Predicting the effects of amino acid substitutions on protein function. *Annu Rev Genomics Hum Genet* 2006;7:61–80.
- [58] Miller MP, Kumar S. Understanding human disease mutations through the use of interspecific genetic variation. *Hum Mol Genet* 2001;10(21):2319–28.
- [59] Wang Z, Moult J. SNPs, protein structure, and disease. *Hum Mutat* 2001;17(4):263–70.
- [60] Sunyaev S, Ramensky V, Bork P. Towards a structural basis of human non-synonymous single nucleotide polymorphisms. *Trends Genet* 2000;16(5):198–200.
- [61] Kunishima S, Okuno Y, Yoshida K, Shiraishi Y, Sanada M, Muramatsu H, et al. ACTN1 mutations cause congenital macrothrombocytopenia. *Am J Hum Genet* 2013;92(3):431–8.
- [62] Gueguen P, Rouault K, Chen JM, Raguenes O, Fichou Y, Hardy E, et al. A missense mutation in the alpha-actinin 1 gene (ACTN1) is the cause of autosomal dominant macrothrombocytopenia in a large French family. *PloS one* 2013;8(9):e74728.

- [63] Gilissen C, Hoischen A, Brunner HG, Veltman JA. Disease gene identification strategies for exome sequencing. *Eur J Hum Genet* 2012;20(5):490–7.
- [64] Gargis AS, Kalman L, Berry MW, Bick DP, Dimmock DP, Hambuch T, et al. Assuring the quality of next-generation sequencing in clinical laboratory practice. *Nat biotechnol* 2012;30(11):1033–6.
- [65] The Royal College of Pathologists of Australasia: Massively Parallel Sequencing Implementation Guidelines [Internet]. 2014. Available from: <https://www.rcpa.edu.au/getattachment/7d264a73-938f-45b5-912f-272872661aaa/Massively-Parallel-Sequencing-Implementation.aspx> [Accessed: 2015-08-09]
- [66] National Pathology Accreditation Advisory Council Requirements for medical testing of human nucleic acids. In: Health AGDo (Ed.) 2 edn [Internet]. 2013. Available from: [https://www.health.gov.au/internet/main/publishing.nsf/Content/E688964F88F4FD20CA257BF0001B739D/\\$File/V0.25%20NAD%20Human%20Genetics.pdf](https://www.health.gov.au/internet/main/publishing.nsf/Content/E688964F88F4FD20CA257BF0001B739D/$File/V0.25%20NAD%20Human%20Genetics.pdf) [Accessed: 2015-08-09].
- [67] Zhang W, Cui H, Wong LJ. Application of next generation sequencing to molecular diagnosis of inherited diseases. *Top Curr Chem* 2014;336:19–45.
- [68] Rehm HL, Bale SJ, Bayrak-Toydemir P, Berg JS, Brown KK, Deignan JL, et al. ACMG clinical laboratory standards for next-generation sequencing. *Genet Med* 2013;15(9):733–47.