Whole exome sequencing identifies genetic variants in inherited thrombocytopenia with secondary qualitative function defects

Ben Johnson,1 Gillian C. Lowe,1 Jane Futterer,1 Marie Lordkipanidzé,1 David MacDonald,1 Michael A. Simpson,2 Isabel Sanchez-Guixé,2 Sian Drake,1 Danai Bem,1 Vincenzo Leo,1 Sarah J. Fletcher,1 Ban Dawood,1 José Rivera,1 David Allsup,1 Tina Biss,1 Paula HB Bolton-Maggs,7 Peter Collins,8 Nicola Curry,9 Charlotte Grimley,9 Beki James,15 Mike Makris,1 Jayashree Motwani,12 Sue Pavord,12 Katherine Talks,12 Jecko Thachil,13 Jonathan Wilde,14 Mike Williams,12 Paul Harrison,12 Paul Gissen,12 Stuart Mundell,12 Andrew Mumford,12 Martina E. Daly,1 Steve P. Watson,1 and Neil V. Morgan1 on behalf of the UK GAPP Study Group

1Institute for Cardiovascular Sciences, College of Medical and Dental Sciences, University of Birmingham, UK; 2Division of Genetics and Molecular Medicine, King’s College, London, UK; 3Centro Regional de Hemodonación, Universidad de Murcia, IMIB-Arrixaca, Murcia, Spain; 4Department of Infection, Immunity and Cardiovascular Disease, University of Sheffield Medical School, University of Sheffield, UK; 5Hull Haemophilia Treatment Centre, Hull and East Yorkshire Hospitals NHS Trust, Castle Hill Hospital, Hull, UK; 6Department of Haematology, Royal Victoria Infirmary, Newcastle Upon Tyne, UK; 7Department of Haematology, Manchester Royal Infirmary, Manchester, UK; 8Arthur Bloom Haemophilia Centre, School of Medicine, Cardiff University, UK; 9Oxford Haemophilia & Thrombosis Centre, Churchill Hospital, Oxford, UK; 10Nottingham Haemophilia Centre, Nottingham University Hospital, UK; 11Regional Centre for Paediatric Haematology, Leeds Children’s Hospital, UK; 12Department of Haematology, Birmingham Children’s Hospital, UK; 13Department of Haematology, Oxford University Hospitals NHS Foundation Trust, UK; 14Adult Haemophilia Centre, Queen Elizabeth Hospital, Birmingham, UK; 15School of Immunity and Infection, College of Medical and Dental Sciences, University of Birmingham, UK; 16Medical Research Council, Laboratory for Molecular Cell Biology, University College London, UK; 17School of Physiology, Pharmacology and Neuroscience, University of Bristol, UK; and 18School of Cellular and Molecular Medicine, University of Bristol, UK

ABSTRACT

Inherited thrombocytopenias are a heterogeneous group of disorders characterized by abnormally low platelet counts which can be associated with abnormal bleeding. Next-generation sequencing has previously been employed in these disorders for the confirmation of suspected genetic abnormalities, and more recently in the discovery of novel disease-causing genes. However its full potential has not yet been exploited. Over the past 6 years we have sequenced the exomes from 55 patients, including 37 index cases and 18 additional family members, all of whom were recruited to the UK Genotyping and Phenotyping of Platelets study. All patients had inherited or sustained thrombocytopenia of unknown etiology with platelet counts varying from 11x10^9/L to 186x10^9/L. Of the 51 patients phenotypically tested, 37 (73%), had an additional secondary qualitative platelet defect. Using whole exome sequencing analysis we have identified “pathogenic” or “likely pathogenic” variants in 46% (17/37) of our index patients with thrombocytopenia. In addition, we report variants of uncertain significance in 12 index cases, including novel candidate genetic variants in previously unreported genes in four index cases. These results demonstrate that whole exome sequencing is an efficient method for elucidating potential pathogenic genetic variants in inherited thrombocytopenia. Whole exome sequencing also has the added benefit of discovering potentially pathogenic genetic variants for further study in novel genes not previously implicated in inherited thrombocytopenia.
Introduction

Inherited thrombocytopenias (IT) are a heterogeneous group of disorders characterized by platelet counts of less than 150x10^9/L in whole blood. Platelet counts are considered normal when maintained at levels between 150x10^9/L and 450x10^9/L. This is achieved by homeostatic processes controlling platelet production (thrombopoiesis), platelet senescence and platelet consumption/destruction. Pathogenic mutations can result in a disruption of these balanced processes causing IT. However, the clinical manifestations are often dependent on both a decreased platelet count and a qualitative or acquired platelet defect and can vary dramatically from severe and potentially life-threatening bleeding to no symptoms. This variation is noted among individuals shown to have the same underlying genetic causes of disease, suggesting that bleeding risk and phenotype are complex traits.  

The average incidence of IT is estimated to be approximately 270 cases per 1x10^6 live births. To date there are 27 individual IT disorders with known causative mutations registered within the Online Mendelian Inheritance in Man (OMIM) catalog, although 33 disease-causing genes have been described.

Genetic studies have played a major role in the diagnosis and progressive understanding of IT. The genes implicated in the disease encode proteins that vary widely in function and include transcription factors (ETV6, FLI1, GATA4, GF1B and RUNX1) and proteins involved in cytoskeleton rearrangement and organization (ACTN1, FLNA, GPF1A, GPF1B, GP9, TUBB1 and WAS). However, some protein functions currently remain unknown (SLFN14 and GNE). Although our knowledge of the causes of IT continues to grow, presently a genetic diagnosis is only reported in approximately 50% of individuals.

So far, genetic investigation into IT has focused on candidate gene sequencing and individual cases of whole exome sequencing (WES) when a causative gene is not obvious. With 50% of patients currently undiagnosed, a change in the way we approach genetic diagnosis is necessary. The pathogenicity of variants was determined and called using previously described methods.

Methods

Study approval

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

Platelet counts, morphology and white blood cell counts

Results from patients’ samples were compared to ranges for healthy volunteers for the specific method of morphology used. Platelet counts for light transmission aggregometry and flow cytometry analysis as well as mean platelet volume in platelet-rich plasma were determined using the Sysmex XN-1000 analyzer (n=11). The FLU-F channel was used to measure platelet counts in whole blood and the immature platelet fraction. Mean platelet volume was determined from the platelet-rich plasma. White blood cell counts were obtained using the Sysmex XN-DIFF channel. All samples were tested against a normal range which was established by measuring the counts for 40 healthy individuals using the Sysmex XN-1000.

Platelet preparation and platelet function testing

Platelet function was assessed by light transmission aggregometry, including lumaggregometry, for samples with platelet counts in platelet-rich plasma of >1x10^9/mL (n=15). An in-house flow cytometry assay was developed to assess platelet function in patients with platelet counts in platelet-rich plasma <1x10^9/mL (n=22). Platelets from individuals with borderline platelet counts in platelet-rich plasma, between 1.0 and 1.5x10^9/mL, were assessed using both assays (n=16).

Whole exome sequencing

WES and bioinformatics analysis were performed as described previously. The pathogenicity of variants was determined and called using the consensus guidelines as set out by the American College of Medical Genetics and Genomics and the Association for Molecular Pathology (ACMG guidelines). Segregation was determined by Sanger sequencing of candidate variants in both affected and unaffected family members, when available, and the classification was adapted appropriately for the specific study and small sample size.

Sanger sequencing

To verify candidate mutations and examine their segregation among family members Sanger sequencing was performed using standard methods on an ABI 3730 automated sequencer, as described previously.

Results

Recruitment of patients

To date, 55 patients with a suspected IT or sustained reduced platelet counts have been enrolled from 25 UK Haemophilia Care Centres and investigated as part of the GAPP study. Before enrollment in the study, all patients underwent clinical and genetic work-up to exclude known
platelet disorders (including Bernard-Soulier syndrome and MYH9-related disorders, analyzed initially by blood film), idiopathic thrombocytopenic purpura and other non-platelet disorders including von Willebrand disease and inherited coagulation factor deficiencies. The patients’ bleeding phenotypes are displayed in Table 1. WES was performed on genomic DNA from all patients, including 57 index cases, all of whom met the study’s entry criteria. All patients, excluding F85.I and F35.II, were of white British or mixed British ethnicity. All results following platelet function testing and WES were reported back to the referring hematology consultants to aid in genetic counselling and disease management.

**Platelet counts, morphology and function testing**

Patients were recruited with a platelet count in whole blood, at the time of enrollment, of less than 150x10^9/L. Patients with platelet counts in the range of 150x10^9/L to 200x10^9/L remained enrolled in the study if they showed a similar phenotype to related affected family members and a platelet count below 150x10^9/L had been observed prior to enrollment (patients F4.II, F11.III, F13.I and F30.II). Platelet counts, mean platelet volume and immature platelet fraction are displayed in Table 1. Of the 55 recruited patients, 12 were deemed to have a macrothrombocytopenia and three a microthrombocytopenia (Table 1). White cell counts were within the normal range (3.78x10^9/L - 10.11x10^9/L, n=40) in all patients analyzed (n=11).

Platelet function studies revealed the presence of a secondary qualitative defect in addition to the low platelet count in 37/51 (73%) of the 55 patients whose DNA underwent WES and who were also available for platelet function testing (Table 1). Of the 37 patients with a secondary qualitative defect, 89% (33/37) displayed defects in both alpha and dense granule secretion. Five of these patients with an observed granule secretion defect were also suspected to have an additional Gi defect because of reduced responses to all concentrations of ADP. The remaining four patients without an observable granule secretion defect showed abnormalities in alternative pathways (integrin activation, cyclooxygenase pathway and GPVI surface levels) in addition to low platelet counts (Table 1).

**Whole exome sequencing**

WES was performed on genomic DNA from all 55 patients, including 37 index cases, following platelet function testing. An average fold-coverage of 111 was observed across all DNA samples analyzed by WES with
Whole exome sequencing in inherited thrombocytopenia

an average of 91% of target sequences having >20x coverage. Areas of poor coverage were analyzed manually when occurring in previously IT-associated genes.

WES revealed between 24,000 and 25,000 variants (single nucleotide variants, small scale insertions/deletions, and splice site variants) in the DNA from each patient, with an average of 197 novel variants per exome. On average, per individual, 2401 variants with a mean allele frequency of <0.01 were observed, excluding synonymous variants. By evaluating the specificity of the pipeline in calling small variations, it was found that the sensitivity was over 99% and the false discovery rate was approximately 3%. The percentage of the genes with ≥20x coverage for a panel of 358 platelet-related genes is included within Online Supplementary Table S1.

Copy number variations were detected using ExomeDepth. The analysis revealed an average of 137 copy number variations per exome (n=32). No copy number variations were deemed potential candidates either because of a high allele frequency or a lack of expression or functional role of the gene within the megakaryocyte/platelet lineage.

Variants in known thrombocytopenia-causing genes

WES and downstream analysis identified variants within in 33 known IT-causing genes in 25 index cases (68%). All variants exceeded 30x sequence coverage at the point of variation and have been confirmed by Sanger sequencing. Variants were selected from positive hits to genes within the panel of 358 IT-associated genes (Online Supplementary Table S1). On average, 37 variants per individual (range, 11-52) were noted in genes from the panel of 358 IT-associated genes, of which on average four (range, 0-7) variants were significant per exome analyzed.

In total 28 variants were noted in 14 genes previously known to cause IT (Table 2). Twenty-one index cases possessed a single variant in a gene previously known to cause IT. Four index cases possessed two variants in genes previously known to cause IT. One variant, RUNX1; c.270+1G>T, was noted in two index cases (F13.I and F14.I). Candidate variations were present within ACTN1, the 5′UTR of ANKRD26, CYCS, FLI1, G6B, ITGB3, GPIBA (heterozygous), MYH9, NBEAL2, RUNX4, SLFN14, STIM1, TPM4 and TUBB1. All but six variants were novel and not present within the variant databases previously mentioned. Three variants, ANKRD26; c.-126T>G in F21.I, MYH9; c.3493C>T (rs80338829) in F9.I and RUNX4; c.550G>A in F19.I and F19.II have been previously associated with IT.18,20 The remaining three variants that have been previously observed occurred at frequencies of <0.005 (0.05%) in available databases. One of the databases scrutinized was that of the ExAC consortium (http://exac.broadinstitute.org) which may include data from individuals with low platelet counts who were either undiagnosed or recruited through an unrelated study (Table 2). Seven variants have previously been published as part of two separate publications from the UK-CDP study group.18 Classification of the 28 variants occurring within the known IT-related genes, following the interpretation guidelines set out by Richards et al.,17 revealed four variants to be “pathogenic”, 13 to be “likely pathogenic” and 11 to be of “uncertain significance.” Variants classified as “pathogenic” were either already known to be a genetic cause of IT; ANKRD26; c.-126T>G in F21.I and MYH9; p.Arg1165Cys in F9.I, or were predicted to be loss-of-function variants in genes for which a loss of function is known to cause disease; FLI1; p.Asn331Thr fs*4, in F4.I and F4.II and RUNX1; p.Trp79* in F12.I.

On average, less than one novel variant was expected to be observed in the known IT-causing genes in which variants were observed. The number of variations occurring also exceeded the expected number when extending the analysis to cover variants with a mean allele frequency of <0.01.

Of the 37 index patients, four presented with two candidate variations in known disease-linked genes, which in one case were present in the same gene. These were as follows: F6.I (CFH1: b.676+1G>A and STIM1; c.270+1G>T, p.Trp79* and ITGB3, p.Arg117Trp) and F23.I (TPM4; p.Ala183Val and TUBB1, p.Phe242Leu).

Of the 25 index cases with variants in known disease-causing genes, nine were observed to have variants within the RUNT1-related transcription factor gene; RUNX1. One variant, RUNX1; p.Arg177Gln, observed in F19.I and F19.II has been previously reported as a causative germline mutation of a familial platelet disorder in two individuals from the same pedigree.21 The variations consisted of five missense variants, two splice-site variants and one nonsense variant. One splice-site variation, c.270+1G>T, was present within three affected individuals from two separate families (F13 and F14). All variants, with the exception of a missense substitution (p.D6N), lie within the genetic region encoding the RUNT homology domain (RHD) which mediates DNA binding and heterodimerization with CBFB (Figure 2).22 Platelets from the majority of these patients (10/15) demonstrated a reduction in ATP secretion and, in keeping with previous reports, several of these patients displayed additional clinical features. Variations in RUNX1 are associated with a propensity to myelodysplastic syndrome and acute myeloid leukemia. To date, hematologic malignancies have not been reported in any patients; however, the brother of F16.1 did have a history of acute myeloid leukemia but was unavailable for testing.

Potentially damaging variants in novel candidate genes

After scrutinizing individuals for variants within the panel of 358 platelet-associated genes (Online Supplementary Table S1), individuals without a variant in a previously IT-associated gene were analyzed for variants in novel genes. WES analysis revealed potentially damaging candidate variants in three families with currently unknown genetic etiology (Table 3). All candidate variants are novel (excluding a previously annotated variant in MLK1; p.Val575Met, which occurs at a frequency of 0.0007718 within the ExAC consortium), segregate with the disease status and have been confirmed by Sanger sequencing.

Variants within ANKRD18A, GNE and FRMDP1 in two related individuals from consanguineous relationships

WES analysis of two related patients (F35.I and F35.II) of South Asian ethnicity was approached differently to that of other patients in this study. Both patients displayed a similarly severe clinical phenotype with a significant reduction in circulating platelets (15x10^9/L). Platelet function testing revealed a reduction in P-selectin (CD62P) expression upon stimulation and variable fluorescent fibrinogen binding which was consistent across both affected
individuals. The patients were cousins born from consanguineous relationships within a single consanguineous kindred so the analysis was focused on identification of a shared homozygous variant due to the recessive segregation of disease. Three variants occurring within ANKRD18A; p.Glu799del, GNE; p.Gly447Arg and FRMPD1; p.Ala509Val were present in both affected individuals and within a tightly linked region of homozygosity on chromosome 9p. The variations within ANKRD18A and GNE were novel according to the previously mentioned databases whereas the variant in FRMPD1 has been observed at a frequency of 0.0003708 including 39 times within the South Asian population (rs571037699). There is no ClinVar entry for this variant and all three variants are classified as variants of “uncertain significance”.

**One missense variant in the recently proposed inherited thrombocytopenia-linked gene, MKL1**

One individual was shown to harbor a rare (frequency <0.01) missense variant within the Megakaryoblastic Leukaemia (translocation) 1 gene; MKL1. The variant was the only variant occurring within a gene of hemostatic relevance within 109 significant novel variants. The variant; MKL1; c.1723G>A, p.Val575Met present in patient F37.I has been noted previously at a frequency of 0.0007718 (allele count of 6/7774 in the ExAC consortium). The patient has a mild reduction in platelet count (130x10^9/L) and no secondary qualitative defects in platelet function were observed. The variant is classified as of “unknown significance”.

**Novel missense candidate variants in PADI2 and TTF2**

Three affected individuals and four unaffected related individuals of a large kindred were recruited to the study. Mild thrombocytopenia was observed within the family with platelet counts ranging from 80x10^9/L to 186x10^9/L in the three affected individuals. All three affected individuals presented with a normal platelet size (7.9-8.6 fl) and a mild reduction in secretion was observed in F30.I and F30.III but not in F30.II. All affected individuals shared a similar bleeding phenotype, suffering from spontaneous epistaxis, excessive bruising and prolonged bleeding from minor wounds. WES analysis revealed 14 novel or rare (frequency <0.01) variants shared between the three affected individuals. Sanger sequencing of all 14 variants in four unaffected related individuals narrowed down candi-

### Table 1. Platelet and bleeding phenotypes of 55 patients recruited to the UK-GAPP study.

<table>
<thead>
<tr>
<th>Family</th>
<th>Patient</th>
<th>Platelet count (x10^9/L)</th>
<th>MPV (fl)</th>
<th>IPF (%)</th>
<th>Secondary defect</th>
<th>Bleeding phenotype</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>I</td>
<td>73</td>
<td>9.6</td>
<td>Yes (Fibrinogen)</td>
<td>Cutaneous bruising/bleeding, menorrhagia</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>I</td>
<td>50</td>
<td>8.6</td>
<td>No</td>
<td>Cutaneous bruising</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>I</td>
<td>80</td>
<td>10.3</td>
<td>Yes (Cyclooxygenase)</td>
<td>Cutaneous bruising, menorrhagia</td>
<td></td>
</tr>
<tr>
<td>II</td>
<td>50</td>
<td>12</td>
<td>Yes (Cyclooxygenase)</td>
<td>Cutaneous bruising, epistaxis, purpura</td>
<td></td>
<td></td>
</tr>
<tr>
<td>III</td>
<td>98</td>
<td>10.5</td>
<td>3.2</td>
<td>N/A</td>
<td>Cutaneous bruising, epistaxis, purpura</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>I</td>
<td>142</td>
<td>11.8+</td>
<td>Yes (Secretion and G1)</td>
<td>Oral cavity bleeding, epistaxis, menorrhagia</td>
<td></td>
</tr>
<tr>
<td>II</td>
<td>157</td>
<td>11.4+</td>
<td>Yes (Secretion)</td>
<td>Cutaneous bruising, oral cavity bleeding</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>I</td>
<td>92</td>
<td>8.8</td>
<td>Yes (Secretion)</td>
<td>Cutaneous bruising, epistaxis, bleeding into joints</td>
<td></td>
</tr>
<tr>
<td>II</td>
<td>100</td>
<td>8.6</td>
<td>Yes (Secretion)</td>
<td>Cutaneous bruising, life-threatening bleeding following surgery</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>I</td>
<td>110</td>
<td>8.9</td>
<td>Yes (Secretion)</td>
<td>Cutaneous bruising, excessive bleeding following surgery</td>
<td></td>
</tr>
<tr>
<td>II</td>
<td>100</td>
<td>8.9</td>
<td>Yes (Secretion)</td>
<td>Cutaneous bruising, epistaxis</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>I</td>
<td>50</td>
<td>10.4</td>
<td>Yes (Secretion)</td>
<td>Cutaneous bruising, menorrhagia</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>I</td>
<td>70</td>
<td>10.7+</td>
<td>No</td>
<td>Cutaneous bruising, menorrhagia, post-partum hemorrhage</td>
<td></td>
</tr>
<tr>
<td>II</td>
<td>70</td>
<td>10.1</td>
<td>No</td>
<td>Cutaneous bruising, epistaxis, hematuria, menorrhagia, post-partum hemorrhage</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>I</td>
<td>35</td>
<td>11.4+</td>
<td>No</td>
<td>Epistaxis, cutaneous bruising,</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>I</td>
<td>55</td>
<td>N/A</td>
<td>No</td>
<td>Cutaneous bruising, epistaxis</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>I</td>
<td>62</td>
<td>9.1</td>
<td>Yes (Secretion)</td>
<td>Cutaneous bruising</td>
<td></td>
</tr>
<tr>
<td>II</td>
<td>N/A</td>
<td>9.1</td>
<td>Yes (Secretion)</td>
<td>Cutaneous bruising, epistaxis</td>
<td></td>
<td></td>
</tr>
<tr>
<td>III</td>
<td>146</td>
<td>9.1</td>
<td>Yes (Secretion)</td>
<td>Cutaneous bruising, epistaxis, menorrhagia</td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>I</td>
<td>100</td>
<td>8</td>
<td>Yes (Secretion)</td>
<td>Excessive cutaneous bleeding</td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>I</td>
<td>163</td>
<td>9.1</td>
<td>No</td>
<td>Cutaneous bruising, epistaxis, hematuria</td>
<td></td>
</tr>
<tr>
<td>II</td>
<td>45</td>
<td>11.9+</td>
<td>Yes (GPVI)</td>
<td>Cutaneous bruising, epistaxis</td>
<td></td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>I</td>
<td>139</td>
<td>8</td>
<td>Yes (Secretion)</td>
<td>Epistaxis, hematuria</td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>I</td>
<td>90</td>
<td>7.6-</td>
<td>Yes (Secretion)</td>
<td>Cutaneous bruising, petechiae</td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>I</td>
<td>130</td>
<td>7.1-</td>
<td>Yes (Secretion and G1)</td>
<td>Cutaneous bruising, epistaxis, oral cavity bleeding</td>
<td></td>
</tr>
<tr>
<td>II</td>
<td>70</td>
<td>7.5</td>
<td>Yes (Secretion and G1)</td>
<td>Cutaneous bruising, epistaxis, oral cavity bleeding</td>
<td></td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>I</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>Excessive bruising/bleeding</td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>I</td>
<td>110</td>
<td>8.1</td>
<td>Yes (Secretion)</td>
<td>Cutaneous bruising/bleeding, petechiae, hematuria</td>
<td></td>
</tr>
</tbody>
</table>

continued on the next page
dates to only two missense variants; \textit{PADI2} (p.Lys499Arg) and \textit{TTF2} (p.His1089Asp). Both variants segregate with disease, not being present in the unaffected individuals. Both variants have been observed previously at a low frequency (<0.01) within the EXaC database (Table 3) and are currently classified as being of “uncertain significance”.

**Discussion**

Here we present the first, large-scale application of WES analysis to patients with inherited bleeding diatheses presenting with thrombocytopenia of unknown etiology.

Platelet counts and phenotypic presentations varied considerably among the patients studied, which is consistent with the variability observed in the spectrum of IT. However, the majority of patients (73%) were noted to have a secondary qualitative defect in platelet function which may explain the disproportionate bleeding when compared to the patients’ platelet counts. A lack of consistency was noted in families 13 and 30, which apparently included affected individuals both with and without defects in platelet function. Clinical complications are shared among the affected family members so this most likely represents limitations in the sensitivity of platelet function testing or intra-familial variability.

Overall, when considering pathogenicity WES analysis positively predicted pathogenicity in 46% of index cases (17/37) (results classified as “pathogenic” or “likely pathogenic” in a gene consistent with the patients’ phenotype and zygosity consistent with expected inheritance). Twenty-two percent of the index cases (8/37) had uncertain/possible pathogenicity (results classified as being of “uncertain significance” in known IT-causing genes). The remaining 32% of index cases (12/37) had a negative prediction of pathogenicity (no convincing variants identified in known IT-causing genes). WES is not without its limitations and, as with any genetic analysis, all variants must be functionally confirmed as deleterious to the coded protein. However, our positive variant discovery rate is comparable to or exceeds the rates in previous large-scale WES clinical multicenter studies of Mendelian disorders.\(^{22,23}\)

Focusing our genetic analysis on patients with unknown etiology of disease with minor prior genetic testing has produced a spectrum of variants different from that from previous, large-scale, targeted genetic studies of IT. Patients were recruited to the study with clinically diagnosed bleeding disorders of unknown etiology. One
A caveat about this approach is the possible exclusion of
individuals with known Bernard-Soulier syndrome or
MYH9-related disorders as these two forms of IT are rou-
tinely tested for in many hematology centers in the UK.
However, three index cases with variants in either GP1BA
or MYH9 were noted in our analysis; these patients had
atypical presentations of Bernard-Soulier syndrome or
MYH9-related disorder and were, therefore, potentially
falsely-negatively reported cases. The individuals with
variants within GP1BA and MYH9 showed a slight
increase in mean platelet volume; however, this was not at
the magnitude of giant platelets normally attributed to this
group of disorders and only patient F9.I showed any sec-
condary syndromic symptoms with the individual suffering
from congenital cataracts.

One attribute of excluding patients with known variants
in GP1BA, GP1BB, MYH9 and potentially GP9 was the
discovery of a relatively large percentage of individuals
analyzed (24% of index cases) with variants in RUNX1 as
a primary likely cause of disease. With the exception of
one predicted loss-of-function variant, the variants present
within RUNX1 are currently classified as either “likely
pathogenic” or of “uncertain significance” and need func-
tional confirmation to be considered the cause of disease.
However, the presence of these variants in a large number
of individuals with an often shared secondary functional
defect in secretion does suggest that the prevalence of
RUNX1 variants may be higher than previously thought.
This raises the issue of whether they should be considered
as clinically significant as Bernard-Soulier syndrome and
MYH9-related disorders and be searched for in a primary
genetic screening at the initial diagnosis of IT.

An advantage of using WES is the possibility of finding
candidate variations in novel genes in subjects who do not
possess variants in known IT-causing genes. The determi-
nation of whether these candidate variants are in fact
pathogenic relies on functional confirmation of the dele-
terious effect of the variant. However, WES analysis, espe-
cially with combined segregation analysis by Sanger
sequencing in extensive kindreds, can provide indications
as to which variants may be of scientific and clinical rele-
vance. This strategy has recently been utilized in the dis-
covery of novel candidate variations in SLFN14 initially as
part of the GAPP study.15,24

Table 2. Results of whole exome sequencing analysis of 55 patients with inherited thrombocytopenia showing variants in known thrombocytopenia-caus-
ing genes. 68% of individuals have a predicted genetic etiology in a previously IT-associated gene. When a variant has been previously observed it is
annotated in the prevalence column with the database in which it is included. The ACMG consensus guideline results are also displayed in the final
classification column.22

<table>
<thead>
<tr>
<th>Family</th>
<th>Patient</th>
<th>Gene(s)</th>
<th>Genomic variation</th>
<th>Protein effect</th>
<th>Variation type</th>
<th>Prevalence</th>
<th>Classification</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>ACTN1</td>
<td>c.2647G&gt;C</td>
<td>p.Gly883Arg</td>
<td>Missense</td>
<td>Novel</td>
<td>Likely pathogenic</td>
<td></td>
</tr>
<tr>
<td>II</td>
<td>ANKRD26</td>
<td>c.-126T&gt;G</td>
<td>5'UTR variation</td>
<td>Known</td>
<td>Pathogenic</td>
<td></td>
<td></td>
</tr>
<tr>
<td>III</td>
<td>CYCS</td>
<td>c.155C&gt;T</td>
<td>p.Ala52Val</td>
<td>Missense</td>
<td>Novel</td>
<td>Likely pathogenic</td>
<td></td>
</tr>
<tr>
<td>IV</td>
<td>CYCS</td>
<td>c.155C&gt;T</td>
<td>p.Ala52Val</td>
<td>Missense</td>
<td>Novel</td>
<td>Likely pathogenic</td>
<td></td>
</tr>
<tr>
<td>V</td>
<td>FLI1</td>
<td>c.992_995del</td>
<td>p.Asn331Thr fs*4</td>
<td>Frameshift deletion</td>
<td>Novel</td>
<td>Pathogenic</td>
<td></td>
</tr>
<tr>
<td>VI</td>
<td>FLI1</td>
<td>c.992_995del</td>
<td>p.Asn331Thr fs*4</td>
<td>Frameshift deletion</td>
<td>Novel</td>
<td>Pathogenic</td>
<td></td>
</tr>
<tr>
<td>VII</td>
<td>FLI1</td>
<td>c.1028A&gt;G</td>
<td>p.Tyr340Cys</td>
<td>Missense</td>
<td>Novel</td>
<td>Likely pathogenic</td>
<td></td>
</tr>
<tr>
<td>VIII</td>
<td>FLI1</td>
<td>c.1028A&gt;G</td>
<td>p.Tyr340Cys</td>
<td>Missense</td>
<td>Novel</td>
<td>Likely pathogenic</td>
<td></td>
</tr>
<tr>
<td>IX</td>
<td>GP1BA</td>
<td>c.3493C&gt;T</td>
<td>p.Arg1165Cys</td>
<td>Missense</td>
<td>Novel</td>
<td>Likely pathogenic</td>
<td></td>
</tr>
<tr>
<td>X</td>
<td>GP1BA</td>
<td>c.413G&gt;T</td>
<td>p.Gly138Val</td>
<td>Missense</td>
<td>Novel</td>
<td>Likely pathogenic</td>
<td></td>
</tr>
<tr>
<td>XI</td>
<td>MYH9</td>
<td>c.3493C&gt;T</td>
<td>p.Arg1165Cys</td>
<td>Missense</td>
<td>Novel</td>
<td>Likely pathogenic</td>
<td></td>
</tr>
<tr>
<td>XII</td>
<td>NBEAL2</td>
<td>c.1376delT</td>
<td>p.Leu459Arg fs*13</td>
<td>Frameshift deletion</td>
<td>Novel</td>
<td>Likely pathogenic</td>
<td></td>
</tr>
<tr>
<td>XII</td>
<td>NBEAL2</td>
<td>c.6893A&gt;G</td>
<td>p.Asn2298Ser</td>
<td>Missense</td>
<td>Novel</td>
<td>Likely pathogenic</td>
<td></td>
</tr>
<tr>
<td>XIII</td>
<td>RUNX1</td>
<td>c.16G&gt;A</td>
<td>p.Asp66Asn</td>
<td>Missense</td>
<td>Novel</td>
<td>Likely pathogenic</td>
<td></td>
</tr>
<tr>
<td>XIIII</td>
<td>RUNX1</td>
<td>c.16G&gt;A</td>
<td>p.Asp66Asn</td>
<td>Missense</td>
<td>Novel</td>
<td>Likely pathogenic</td>
<td></td>
</tr>
<tr>
<td>XIII</td>
<td>RUNX1</td>
<td>c.16G&gt;A</td>
<td>p.Asp66Asn</td>
<td>Missense</td>
<td>Novel</td>
<td>Likely pathogenic</td>
<td></td>
</tr>
<tr>
<td>XIV</td>
<td>RUNX1</td>
<td>c.236G&gt;A</td>
<td>p.Try79*</td>
<td>Nonsense</td>
<td>Novel</td>
<td>Pathogenic</td>
<td></td>
</tr>
<tr>
<td>XIVII</td>
<td>RUNX1</td>
<td>c.236G&gt;A</td>
<td>p.Try79*</td>
<td>Nonsense</td>
<td>Novel</td>
<td>Pathogenic</td>
<td></td>
</tr>
<tr>
<td>XIVIII</td>
<td>RUNX1</td>
<td>c.230+1G&gt;T</td>
<td>p.Tyr79*</td>
<td>Nonsense</td>
<td>Novel</td>
<td>Pathogenic</td>
<td></td>
</tr>
<tr>
<td>XV</td>
<td>RUNX1</td>
<td>c.230+1G&gt;T</td>
<td>p.Tyr79*</td>
<td>Nonsense</td>
<td>Novel</td>
<td>Pathogenic</td>
<td></td>
</tr>
<tr>
<td>XVI</td>
<td>RUNX1</td>
<td>c.427+1G&gt;T</td>
<td>p.Gly1429Ser</td>
<td>Missense</td>
<td>Novel</td>
<td>Likely pathogenic</td>
<td></td>
</tr>
<tr>
<td>XVIII</td>
<td>RUNX1</td>
<td>c.427+1G&gt;T</td>
<td>p.Gly1429Ser</td>
<td>Missense</td>
<td>Novel</td>
<td>Likely pathogenic</td>
<td></td>
</tr>
</tbody>
</table>

continued on the next page
Family 35 is an interesting case of two affected related individuals born from consanguineous relationships. The molecular function of ANKRD18A is currently unknown, while FRMPD1 regulates the subcellular localization of activator of G-protein signaling 3 (AGS3). Both genes are expressed weakly in hematopoietic cells. However, GNE, coding for an enzyme in the sialic acid biosynthetic pathway, is expressed in all cells of the hematopoietic lineage. There are currently 88 registered mutations in GNE in the Human Genome Mutation Database (www.hgmd.cf.ac.uk). Mutations are known to be the genetic cause of sialuria (OMIM269921) and hereditary inclusion body myopathy (OMIM600737). Recently, two separate groups have reported patients with compound heterozygous variations in GNE, causing GNE-related myopathy with congenital thrombocytopenia. One interesting phenotypic presentation of the affected individual was an intermittent mild thrombocytopenia with low platelet counts in whole blood of between 50x10^9/L and 150x10^9/L. Here we present a novel variant within GNE, at a highly conserved genetic site. The missense variant observed in F37.I represents the only variant to occur in a gene with previous hematologic implications. One further variant in GNE was observed in addition to a “likely pathogenic” frameshift causing insertion within TUBB1 in patient F25.I. Due to the predicted loss of function of the frameshift causing the TUBB1 variant it is unlikely that the variant with GNE is additive to the phenotype of patient F25.I. However, the variant of uncertain significance in patient F37.I is an interesting candidate to take forward for functional studies.

Whole exome sequencing in inherited thrombocytopenia

| 17 | RUNX1 | c.505A>G | p.Thr168Ala Missense Novel Uncertain significance |
| 18 | RUNX1 | c.512A>T | p.Asp171Val Missense Novel Uncertain significance |
| 19 | RUNX1 | c.503G>A | p.Arg177Gln Missense Known Likely pathogenic |
| 19 | RUNX1 | c.503G>A | p.Arg177Gln Missense Known Likely pathogenic |
| 20 | SLFN14 | c.652G>A | p.Lys218Glu Missense Novel Uncertain significance |
| 21 | SLFN14 | c.657A>T | p.Lys219Asn Missense Novel Uncertain significance |
| 22 | SLFN14 | c.659T>A | p.Val220Asp Missense Novel Likely pathogenic |
| 22 | SLFN14 | c.659T>A | p.Val220Asp Missense Novel Likely pathogenic |
| 22 | SLFN14 | c.659T>A | p.Val220Asp Missense Novel Likely pathogenic |
| 23 | TPM4 | c.548C>T | p.Ala183Val Missense Novel Uncertain significance |
| 24 | TUBB1 | c.721C>T | p.Arg241Trp Missense 0.001071 (ExAC) Novel Likely pathogenic |
| 25 | TUBB1 | c.1000_1001insG | p.Leu334Ala fs*19 Frameshift insertion Novel Likely pathogenic |

MKL1 was initially included in our panel of 358 genes for post-WES analysis due to its role in megakaryocyte maturation elucidated via its binding partner, serum response factor (SRF). Recently, the first case of a homozygous mutation in MKL1 in a patient with severe immunodeficiency and no hematologic malignancies was reported. One interesting phenotypic presentation of the affected individual was an intermittent mild thrombocytopenia with low platelet counts in whole blood of between 50x10^9/L and 150x10^9/L. Here we present a novel variant within MKL1, at a highly conserved genetic site. The missense variant observed in F37.I represents the only variant to occur in a gene with previous hematologic implications. One further variant in MKL1 was observed in addition to a “likely pathogenic” frameshift causing insertion within TUBB1 in patient F25.I. Due to the predicted loss of function of the frameshift causing the TUBB1 variant it is unlikely that the variant with MKL1 is additive to the phenotype of patient F25.I. However, the variant of uncertain significance in patient F37.I is an interesting candidate to take forward for functional studies.

WES and segregation determination using Sanger sequencing revealed candidate variants in PADI2 and TTF2 that segregate with disease in F30.I, F30.II and F30.III. The phenotypic presentations vary between the patients but clinical presentations are consistent, which
may reflect limitations in the sensitivity of platelet function testing. Neither gene has previously been implicated in hematologic abnormalities: mutations in PADI2 have been associated with schizophrenia, breast cancer and rheumatoid arthritis, while mutations in TTF2 have been associated with thyroid dysgenesis.11-14 WES analysis has therefore provided us with the first steps for determining the impact of these two variants of uncertain significance and whether they have the propensity to be disease caus- ing.

In summary, we show that WES can be applied to iden- tify the underlying genetic cause in known IT-causing genes for patients with thrombocytopenia and unclear dis- ease etiology. We show similar positive detection rates when compared to prioritized studies and, with the addition of complementary functional studies, show an improved detection rate when compared to WES analysis of other developmental disorders. We also suggest the applicability of WES in providing preliminary insight into novel genes and their potential mechanism of action through candidate variations of unknown significance.

This approach provides a foundation to enhance our current knowledge on megakaryopoiesis, platelet function and platelet senescence/death through subsequent functional studies.

Acknowledgments

We thank the families for providing samples and our clinical and laboratory colleagues for their help. This work was supported by the British Heart Foundation (RG/PG/13/36/30275; RG/09/007), an ARC Doctoral Training Partnership grant (BJ), a Wellcome Trust Combined Training Programme Fellowship (093994) (GCL), the Healing Foundation (PH) and the Platelet Research Centre Award to Guy’s & St Thomas’ NHS Foundation Trust in partnership with King’s College London and King’s College Hospital NHS Foundation Trust. We thank the Queen Elizabeth Hospital Charity for funding the Sysmex XN-1000.

References

4. Canales ML, Mauer AM. Sex-linked heredi-
Whole exome sequencing in inherited thrombocytopenia


