

A comprehensive bioinformatic analysis of 126 patients with an inherited platelet disorder to identify both sequence and copy number genetic variants

Almazni, Ibrahim Abdullah F; Stapley, Rachel; Khan, Abdullah; Morgan, Neil

DOI:
[10.1002/humu.24114](https://doi.org/10.1002/humu.24114)

License:
Creative Commons: Attribution (CC BY)

Document Version
Publisher's PDF, also known as Version of record

Citation for published version (Harvard):
Almazni, IAF, Stapley, R, Khan, A & Morgan, N 2020, 'A comprehensive bioinformatic analysis of 126 patients with an inherited platelet disorder to identify both sequence and copy number genetic variants', *Human Mutation*, vol. 41, no. 11, pp. 1848-1865. <https://doi.org/10.1002/humu.24114>

[Link to publication on Research at Birmingham portal](#)

General rights

Unless a licence is specified above, all rights (including copyright and moral rights) in this document are retained by the authors and/or the copyright holders. The express permission of the copyright holder must be obtained for any use of this material other than for purposes permitted by law.

- Users may freely distribute the URL that is used to identify this publication.
- Users may download and/or print one copy of the publication from the University of Birmingham research portal for the purpose of private study or non-commercial research.
- User may use extracts from the document in line with the concept of 'fair dealing' under the Copyright, Designs and Patents Act 1988 (?)
- Users may not further distribute the material nor use it for the purposes of commercial gain.

Where a licence is displayed above, please note the terms and conditions of the licence govern your use of this document.

When citing, please reference the published version.

Take down policy

While the University of Birmingham exercises care and attention in making items available there are rare occasions when an item has been uploaded in error or has been deemed to be commercially or otherwise sensitive.

If you believe that this is the case for this document, please contact UBIRA@lists.bham.ac.uk providing details and we will remove access to the work immediately and investigate.

Don't settle for just a supplier. Find a custom manufacturing partner.

Your specifications. Your format.
Our scientists waiting to help.



Let's **TALK**
CUSTOM



Selecting a supplier for your biotechnology and biopharma products can be a challenge—especially a supplier who can adapt to your specific needs. Work with a manufacturing partner you can be certain about. With our breadth of technologies and rapid response manufacturing, Promega is here to support your molecular testing needs.



Watch the video or download a PDF:
promega.com/CustomManufacturing



A comprehensive bioinformatic analysis of 126 patients with an inherited platelet disorder to identify both sequence and copy number genetic variants

Ibrahim Almazni | Rachel J. Stapley | Abdullah O. Khan | Neil V. Morgan 

Institute of Cardiovascular Sciences, College of Medical and Dental Sciences, University of Birmingham, Birmingham, UK

Correspondence

Neil V. Morgan, Institute of Cardiovascular Sciences, College of Medical and Dental Sciences, Edgbaston, University of Birmingham, Birmingham B15 2TT, UK.
Email: N.V.Morgan@bham.ac.uk

Funding information

Saudi Arabia Cultural Bureau in London; Wellcome Trust, Grant/Award Number: 218649/Z/19/Z; British Heart Foundation, Grant/Award Numbers: FS/15/18/31317, FS/18/11/33443, PG/13/36/30275, PG/16/103/32650

Abstract

Inherited bleeding disorders (IBDs) comprise an extremely heterogeneous group of diseases that reflect abnormalities of blood vessels, coagulation proteins, and platelets. Previously the UK-GAPP study has used whole-exome sequencing in combination with deep platelet phenotyping to identify pathogenic genetic variants in both known and novel genes in approximately 40% of the patients. To interrogate the remaining “unknown” cohort and improve this detection rate, we employed an IBD-specific gene panel of 119 genes using the Congenica Clinical Interpretation Platform to detect both single-nucleotide variants and copy number variants in 126 patients. In total, 135 different heterozygous variants in genes implicated in bleeding disorders were identified. Of which, 22 were classified pathogenic, 26 likely pathogenic, and the remaining were of uncertain significance. There were marked differences in the number of reported variants in individuals between the four patient groups: platelet count (35), platelet function (43), combined platelet count and function (59), and normal count (17). Additionally, we report three novel copy number variations (CNVs) not previously detected. We show that a combined single-nucleotide variation (SNV)/CNV analysis using the Congenica platform not only improves detection rates for IBDs, suggesting that such an approach can be applied to other genetic disorders where there is a high degree of heterogeneity.

KEYWORDS

CNV, inherited bleeding, platelet disorders, SNV, thrombocytopenia, variant interpretation, whole-exome sequencing

1 | INTRODUCTION

Inherited bleeding disorders (IBDs) are a heterogeneous group of diseases that reflect abnormalities in blood vessels, coagulation proteins, and platelets. They often present after birth or during

childhood, and clinically manifest with variable bleeding tendencies (Blanchette et al., 1991). Although the majority of IBDs are known to be primarily associated with coagulation factor abnormalities such as hemophilia A and B, rarer disorders of platelet count and function are still poorly understood (Sivapalaratnam et al., 2017). Therefore,

Ibrahim Almazni and Rachel J. Stapley contributed equally to this study.

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2020 The Authors. *Human Mutation* published by Wiley Periodicals LLC

to investigate the molecular mechanisms of this group of disorders, it is often best to address the gene(s) already implicated in these bleeding disorders in the first instance, and then specifically to investigate how the genetic variants can disrupt the gene function (Nurden et al., 2012; Peyvandi et al., 2006). An increasing number of new genes and their variants have been discovered, which are implicated in megakaryocyte differentiation and/or platelet production and function (Johnson, Fletcher, et al., 2016).

The UK Genotyping and Phenotyping of Platelets study (UK-GAPP; <https://www.birmingham.ac.uk/research/cardiovascular-sciences/research/platelet-group/platelet-gapp/index.aspx>) has recruited over 1000 patients based on a history of suspected bleeding disorders of unknown cause from over 25 collaborating hemophilia care centers across the United Kingdom. Recruited patients underwent a combination of platelet phenotyping and genotyping to determine the likely causative genes attributable to their specific defects (Jones et al., 2012; Watson et al., 2013). Gross hematological analysis and light transmission aggregometry and/or flow cytometry were used to identify thrombocytopenia (low platelet counts), platelet function, and cell signaling defects. Following this, targeted genetic analysis was employed and revealed variants, both novel and known, to be causative of bleeding in patients.

High-throughput sequencing technologies including whole-exome sequencing (WES) and whole-genome sequencing are valuable tools used to uncover novel variants in platelet-specific genes. Over the past 10 years, such techniques have revealed many causative variants, therefore assisting in providing a clear diagnosis for some patients with severe bleeding disorders (Bastida et al., 2018; Daly et al., 2014; Downes et al., 2019; Leinøe et al., 2017). In addition, targeted next-generation sequencing (NGS) panels can be used to highlight platelet-specific genes that have been previously implicated in bleeding disorders. NGS panels can be employed in a clinical diagnostic setting or used for prescreening, filtering out patients with variants in known genes, and subsequently employing WES for those who may harbor variants in novel genes (Johnson et al., 2018; Simeoni et al., 2016). This approach was applied in the UK-GAPP study where patients with known mutations in hemophilia A and B or coagulation mediated genes, known to cause bleeding were eliminated. However, many of these panels do not search for copy number variations (CNVs), and indeed we, and others have not found definitively causative variants in approximately 40%–50% patients despite a strongly indicative inherited component for their bleeding (Bastida et al., 2018; Johnson et al., 2018; Johnson, Lowe, et al., 2016; Leinøe et al., 2017; Lentaigne et al., 2016). In this study, we address this by applying a newly established, comprehensive genetic analysis software that detects both single-nucleotide variations (SNVs) and CNVs. Congenica (<https://www.congenica.com>) is an automated clinical decision support platform that was used to analyze and rapidly interpret the WES data of 126 patients recruited to the UK-GAPP study. Users are able to prioritize and review genetic variants, as well as assign pathogenicity, after which the software calculates overall pathogenicity based on the American College of Medical Genetics and Genomics (ACMG) guidelines (Richards

et al., 2015). It collates all essential information to make an informed and robust decision for the identification of causal genetic variants.

The Congenica platform is primarily applied for genetic diagnostics and is routinely used in clinical laboratories for variant validation and reporting. For the first time, we show its utility in interrogating a large cohort of patients recruited to the UK-GAPP research study. Using this approach, we perform a robust and comprehensive analysis to find both known and novel genetic variants with plausible association with disease, including rare CNVs not previously detected. Combined with extensive patient phenotypic studies, this provides a potent tool for the dissection of the genetic causes of bleeding in a cohort which, thus far, remains genetically unresolved despite an extensive clinical presentation of familial bleeding.

2 | METHODS

2.1 | Hematological evaluation and platelet phenotyping of patients

To initially classify patients as having a platelet defect and determine their platelet defect subtypes, they underwent an initial hematological workup and extensive platelet function testing workflow. These methods can be seen in detail in the Supporting Information Methods section.

2.2 | WES

WES was performed on the genomic DNA of 117 patients in this study as previously described (Johnson, Lowe, et al., 2016). Briefly following enrichment of coding regions and intron/exon boundaries with the SureSelect human AllExon 50 Mb Kit (Agilent Technologies), captured libraries were sequenced on the Illumina HiSeq 2500 (Illumina) with 100-bp paired-end reads.

2.3 | Processing WES data using Congenica software

First, an Interpretation Request (IR) was completed which included information about the proband and any other family members and related clinical data including HPO terms (abnormal bleeding HP:0001892 and/or thrombocytopenia HP:0001873) for affected individuals. Relevant gene panels (Inherited Bleeding Disorder; High Evidence_Green, Medium Evidence Amber and Low Evidence Red, gene lists) containing 119 genes (Table S1) from Panel app (<https://panelapp.genomicsengland.co.uk/>) were applied in the project and deemed suitable for research purposes. However, of this gene panel only 88 genes from the Genomics England website (R90) are considered as suitable for clinical use at this present time. The WES data (either BAM or FASTQ files) of patients were then transferred to the Congenica SFTP server for processing. The Congenica pipeline could

then be used for sequence alignment and variant calling of SNVs, small insertion/deletion (indels), CNVs (Figure 1), and coverage (Table S3).

The analytical pipeline for the detection of CNVs in genes involved in the IBDs panel was employed using the ExomeDepth coverage approach. The exome read depth of the target patient's sample was compared against the read depth of a reference panel (up to 10 WES samples of each gender) to detect regions with different coverage which could represent a CNV event.

Using the Congenica software, the lower limit that the ExomeDepth calling software uses for CNV calling is ≥ 20 sequence reads. This ensures that ExomeDepth does not consider low quality reads when comparing the reference samples to the target patient.

3 | RESULTS

3.1 | Platelet phenotyping

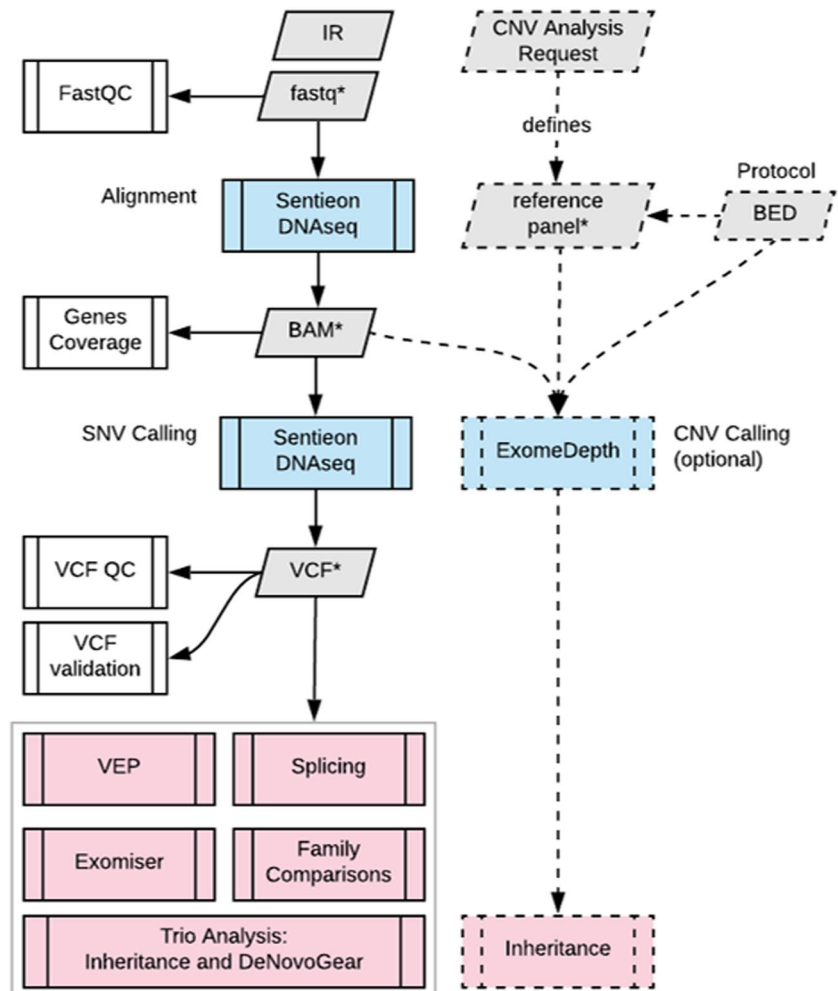
Recruited patients were subjected to an initial hematological analysis and extended deep platelet phenotyping using the previously published workflow (Johnson, Lowe, et al., 2016). Phenotyping outcomes

can be seen in detail in the Supporting Information Results section (Figures S1 and S2; Table S2).

3.2 | Validation of WES analysis with known variants

Validation of the WES analysis in the GAPP study was performed using Congenica software. Five different known genetic variants were identified previously by WES in nine patients with a suspected IBD (Fletcher et al., 2015; Johnson, Lowe, et al., 2016; Table 1). We used two trios (one parent and two affected children) and three unrelated individuals, all with known or likely pathogenic variants in platelet- or megakaryocyte-related genes. This analysis was conducted in a blind manner to assess the reliability and robustness of the software in correctly highlighting all known genetic variants in these patients. Using panels of genes implicated in IBDs, the first trio (Family A) including exomes of patients 1–3 were found to share the same splicing sequence variant in *RUNX1*; c.98-1G>A. The second trio (Family B) including exomes of patients 4–6 all shared a variant c.503G>T p.(Cys168Phe) in *GFI1B*. Patient 7 displayed a homozygous missense mutation c.1246G>A p.(Gly416Arg) in the *GNE* gene. In

FIGURE 1 Congenica pipeline overview for processing of whole-exome sequencing (WES) data. Adapted from <https://www.congenica.com/>. The informatic strategy shown is used to incorporate both single-nucleotide variation (SNV) and copy number variation (CNV) analysis. The raw WES data are inputted as either FASTQ, or BAM files followed by alignment to the reference genome. SNV calling is then performed to generate VCF files and subsequent in silico tools to determine the pathogenicity of variants. Simultaneously, CNV analysis is performed using a predefined reference and sex-matched WES panel and fed into ExomeDepth for CNV calling in the WES samples



| Patient | Gene | Variation | Type |
|--------------|---|-------------------------|-----------------|
| (Family A) 1 | <i>RUNX1</i> (Zhang et al., 2018) | c.98-1G>A | Splice acceptor |
| (Family A) 2 | <i>RUNX1</i> | c.98-1G>A | Splice acceptor |
| (Family A) 3 | <i>RUNX1</i> | c.98-1G>A | Splice acceptor |
| (Family B) 4 | <i>GFI1B</i> (Rabbolini et al., 2017) | c.503G>T p.(Cys168Phe) | Missense |
| (Family B) 5 | <i>GFI1B</i> | c.503G>T p.(Cys168Phe) | Missense |
| (Family B) 6 | <i>GFI1B</i> | c.503G>T p.(Cys168Phe) | Missense |
| 7 | <i>GNE</i> | c.1246G>A p.(Gly416Arg) | Missense |
| 8 | <i>SLFN14</i> (Johnson, Lowe, et al., 2016) | c.659T>A p.(Val220Asp) | Missense |
| 9 | <i>THBD</i> (Rabbolini et al., 2017) | c.1611C>A p.(Cys537Ter) | Stop gain |

Note: NCBI reference sequences: *RUNX1* (NM_001001890.3); *GFI1B* (NM_001371908.1); *GNE* (NM_000157.4); *SLFN14* (NM_001129820.1); *THBD* (NM_000361.2).

patient 8, a missense variant c.659T>A p.(Val220Asp) in *SLFN14* and finally in patient 9, a stop gain mutation c.1611C>A p.(Cys537Ter) located within *THBD* was identified. All known variants found in the patients were successfully verified by Congenica software against our previously analyzed WES data (Table 1).

3.3 | WES analysis to identify new SNVs and CNVs using the Congenica platform

WES data of all 117 patients were analyzed by the Congenica platform based on the ExomeDepth coverage following the phenotyping and platelet function studies workflow. The Congenica pipeline was used for exome sequence alignment and variant calling of SNVs, indels, and CNVs to determine plausible candidate variants from the WES data. An average read depth sequencing coverage of 205 was observed at the site of each variation across all DNA samples analyzed by WES for SNVs (Table S3). The ExomeDepth integrated tool was used to determine CNV based on read coverage (Table 3). First, WES data of the 117 patients were analyzed by filtering using an IBDs gene panel (Table S1). Variants were then filtered within the software based on the exclusion criteria initially stated in the GAPP study. A rare variant cut-off or minor allele frequency (MAF) of <0.01 in each data set was used and synonymous and intron variants ± 5 base pairs away from the exon-intron boundaries were excluded. Non-shared variants between the same affected family members were also eliminated.

Following exclusion of variants based on these criteria, a range of between two and six variants (SNVs, small indels, and splice site) were noted per patient (Table 2). In silico pathogenicity prediction tools that have been integrated into the Congenica software were employed for further analysis (Table 2). A total of 135 variants in genes implicated in bleeding disorders were identified across all the 117 patients and all variants were observed in a heterozygous state (Table 2). In total, 22 variants were classified as pathogenic and 26 were likely pathogenic when considering the ACMG consensus guidelines. The remaining 87 variants were classified as of uncertain

TABLE 1 Nine patients and the five known candidate variants used for validation of the Congenica software

significance. The graphical illustration of this summary is shown in (Figure 2a). There was a marked difference in the number of reported variants between the four classes of variants in patient groups: platelet count (35); platelet function (43); combined platelet count and function (59); and normal count (17) (Figure 2b).

3.4 | Candidate variants identified in patient's cohort

A total of 48/135 (35.5%) variants with MAF of 0, unless otherwise stated, were identified across the 117 patients (Table 2). In total, 14/48 (29.1%) variants have been published previously. The number of variants found to be shared in the same affected family members were 21. Plausible candidate variants were present within the following genes (*RUNX1*, *SLFN14*, *FLI1*, *ETV6*, *HPS3*, *F10*, *P2RY12*, *SMAD4*, *TUBB1*, *GP1BA*, *GBA*, *CYCS*, *VWF*, *THBD*, *LYST*, *ADAMTS13*, *GFI1B*, *ITGA2B*, *NBEAL2*, *MECOM*, and *MYH9*; Table 2). Two rare variants were noted between five related affected family members including *RUNX1*: c.611G>A p.(Arg204Gln) in patients 38.1 and 38.2; *WAS*: c.1456G>A p.(Glu486Lys) in patients 38.3, 38.4, and 38.5; and a stop gain variant within *ADAMTS13*: c.1315G>T p.(Glu439Ter) was shared between two related affected individuals 53.1 and 53.2. Two related affected individuals with macrothrombocytopenia shared a variant within the newly discovered gene (involved in platelet disorders) *MECOM*: c.951G>T p.(Lys317Asn). A novel stop gained variant within *ETV6*: c.1288C>T p.(Arg430Ter) was noted in patient 7, which was subsequently classified as pathogenic.

3.5 | CNVs found in the patient cohort

Overall, the CNV analysis using the integrated ExomeDepth tool revealed an average of four CNVs per exome ($n = 15$; Table 3). There were three rare structural variants covering large regions on chromosomes 11 and 17 and encompassing numerous genes, including

TABLE 2 Variants identified by analysis of whole-exome sequencing of 117 patients with suspected inherited bleeding disorders using the inherited bleeding gene panel

| Patient | Gene(s) | VEP | MAX AF | Genomic variation | Protein effect | PolyPhen | SIFT | Pathogenicity |
|---------|--------------------------------------|--------------|----------|-------------------|---------------------|-------------------|-------------|------------------------|
| 1.1 | NVD | | | | | | | |
| 1.2 | NVD | | | | | | | |
| 1.3 | NVD | | | | | | | |
| 1.4 | NVD | | | | | | | |
| 2.1 | RUNX1 (Stockley et al., 2013) | Splice donor | 0 | c.508+1G>T | p.? | NA | NA | Pathogenic |
| 2.2 | RUNX1 (Stockley et al., 2013) | Splice donor | 0 | c.508+1G>T | p.? | NA | NA | Pathogenic |
| 3.1 | SLFN14 (Johnson, Lowe, et al., 2016) | Missense | 0 | c.659T>A | p.(Val220Asp) | Possibly_damaging | Deleterious | Pathogenic |
| 3.2 | SLFN14 (Johnson, Lowe, et al., 2016) | Missense | 0 | c.659T>A | p.(Val220Asp) | Possibly_damaging | Deleterious | Pathogenic |
| 4.1 | FLI1 (Johnson, Lowe, et al., 2016) | Frameshift | 0 | c.992_995del | p.(Asn331ThrfsTer4) | NA | NA | Pathogenic |
| 4.2 | FLI1 (Johnson, Lowe, et al., 2016) | Frameshift | 0 | c.992_995del | p.(Asn331ThrfsTer4) | NA | NA | Pathogenic |
| 5 | FGA | Missense | 0.00951 | c.1366A>G | p.(Thr456Ala) | Possibly_damaging | Deleterious | Uncertain significance |
| 6.1 | NVD | | | | | | | |
| 6.2 | MPIG6B | Missense | 0 | c.132G>C | p.(Trp44Cys) | Probably_damaging | Deleterious | Uncertain significance |
| | VPS33B | Missense | <0.0001 | c.434T>C | p.(Leu145Ser) | Probably_damaging | Deleterious | Uncertain significance |
| 7 | ETV6 | Stop gained | 0 | c.1288C>T | p.(Arg430Ter) | NA | NA | Pathogenic |
| 8 | VWF | Frameshift | <0.0001 | c.2516del | p.(Gly839GlufsTer4) | NA | NA | Pathogenic |
| | ANKRD26 | Missense | 0.00324 | c.3004G>A | p.(Glu1002Lys) | Possibly_damaging | Deleterious | Uncertain significance |
| 9 | SLC45A2 | Missense | <0.0001 | c.1471G>A | p.(Gly491Arg) | Probably_damaging | Deleterious | Uncertain significance |
| 10.1 | HP53 | Missense | 0 | c.479G>A | p.(Ser160Asn) | Benign | Tolerated | Uncertain significance |
| 10.2 | HP53 | Missense | 0 | c.479G>A | p.(Ser160Asn) | Benign | Tolerated | Uncertain significance |
| 11 | LYST | Missense | 0.0005 | c.8960C>G | p.(Pro2987Arg) | Probably_damaging | Deleterious | Uncertain significance |
| | AP3D1 | Missense | 0.000116 | c.1246G>A | p.(Glu416Lys) | Possibly_damaging | Deleterious | Uncertain significance |
| 12 | COL5A2 | Missense | 0.00264 | c.4067A>G | p.(Asp1356Gly) | Benign | Deleterious | Uncertain significance |
| 13 | F7 | Missense | <0.0001 | c.857C>T | p.(Ala286Val) | Benign | Deleterious | Uncertain significance |
| 14.1 | F10 | Missense | 0 | c.1325G>A | p.(Gly442Asp) | Probably_damaging | Deleterious | Likely pathogenic |
| 14.2 | NBEAL2 | Missense | 0.0066 | c.6866G>A | p.(Arg2289Gln) | Possibly_damaging | Deleterious | Uncertain significance |
| | GBA | Missense | 0.00363 | c.1226A>G | p.(Asn409Ser) | Benign | Deleterious | Uncertain significance |
| | F10 | Missense | 0 | c.1325G>A | p.(Gly442Asp) | Probably_damaging | Deleterious | Likely pathogenic |
| | NBEAL2 | Missense | 0.0066 | c.6866G>A | p.(Arg2289Gln) | Possibly_damaging | Deleterious | Uncertain significance |

TABLE 2 (Continued)

| Patient | Gene(s) | VEP | MAX AF | Genomic variation | Protein effect | PolyPhen | SIFT | Pathogenicity |
|---------|-------------------------------------|--------------|----------|-------------------------|----------------|-------------------|-------------|------------------------|
| | GBA | Missense | 0.00363 | c.1226A>G | p.(Asn409Ser) | Benign | Deleterious | Uncertain significance |
| 15.1 | VWF | Missense | 0.00558 | c.2561G>A | p.(Arg854Gln) | Possibly_damaging | Deleterious | Likely pathogenic |
| 15.2 | NBEA | 5'-UTR | 0 | c.-161C>T | p.? | NA | NA | Uncertain significance |
| | VWF | Missense | 0.00558 | c.2561G>A | p.(Arg854Gln) | Possibly_damaging | Deleterious | Likely pathogenic |
| | NBEA | 5'-UTR | 0 | c.-161C>T | p.? | NA | NA | Uncertain significance |
| 16 | ACVRL1 | Missense | <0.0001 | c.653G>A | p.(Arg218Gln) | Benign | Deleterious | Likely pathogenic |
| | RUNX1 (Stockley et al., 2013) | Stop gained | 0 | c.317G>A | p.(Trp106Ter) | NA | NA | Pathogenic |
| | ITGB3 (Johnson, Lowe, et al., 2016) | Missense | <0.0001 | c.349C>T | p.(Arg117Trp) | Possibly_damaging | Deleterious | Likely pathogenic |
| 17 | RUNX1 (Stockley et al., 2013) | Splice donor | 0 | c.351+1G>T | p.? | NA | NA | Pathogenic |
| | F11 | Stop gained | 0.00127 | c.403G>T | p.(Glu135Ter) | NA | NA | Pathogenic |
| | SERPINC1 | Missense | 0.00276 | c.1246G>T | p.(Ala416Ser) | Probably_damaging | Deleterious | Uncertain significance |
| | F13A1 | Missense | 0.000192 | c.1149G>T | p.(Arg383Ser) | Probably_damaging | Deleterious | Uncertain significance |
| 18 | PTPN11 | Missense | <0.0001 | c.922A>G | p.(Asn308Asp) | Benign | Deleterious | Likely pathogenic |
| 19 | FGB | Missense | 0.00674 | c.794C>T | p.(Pro265Leu) | Probably_damaging | Deleterious | Uncertain significance |
| 20 | GP6 | Missense | 0.00209 | c.172C>T | p.(Arg58Cys) | Possibly_damaging | Deleterious | Uncertain significance |
| | THBD | Missense | 0.00528 | c.1502C>T | p.(Pro501Leu) | Possibly_damaging | Deleterious | Uncertain significance |
| 21 | PLG | Missense | 0.00407 | c.1469G>A | p.(Arg490Gln) | Probably_damaging | Deleterious | Uncertain significance |
| | ARPC1B | Missense | <0.0001 | c.308G>A | p.(Arg103His) | Benign | Deleterious | Uncertain significance |
| 22.1 | P2RY12 (Leo et al., 2015) | Missense | 0.00015 | c.365G>A | p.(Arg122His) | Probably_damaging | Deleterious | Likely pathogenic |
| 22.2 | P2RY12 (Leo et al., 2015) | Missense | 0.00015 | c.365G>A | p.(Arg122His) | Probably_damaging | Deleterious | Likely pathogenic |
| 23 | P2RY12 (Leo et al., 2015) | Missense | 0.0001 | c.772C>A | p.(Pro258Thr) | Probably_damaging | Deleterious | Pathogenic |
| | MCFD2 | Missense | 0.00027 | c.416C>T | p.(Ala139Val) | Benign | Tolerated | Uncertain significance |
| 24 | VPS33B | Missense | 0.00267 | c.1274G>A | p.(Ser425Asn) | Probably_damaging | Deleterious | Uncertain significance |
| | ITGB3 | Missense | 0.00528 | c.197T>G | p.(Leu66Arg) | Probably_damaging | Deleterious | Uncertain significance |
| | LYST | Missense | 0.00264 | c.9017A>G | p.(Lys3006Arg) | Probably_damaging | Tolerated | Uncertain significance |
| 25 | RUNX1 | Missense | 0 | c.403G>A | p.(Gly135Ser) | Probably_damaging | Deleterious | Likely pathogenic |
| | VWF | Missense | 0.00276 | c.7988G>C | p.(Arg2663Pro) | Benign | Tolerated | Uncertain significance |
| 26 | RUNX1 | Missense | 0 | c.593A>T | p.(Asp198Val) | Possibly_damaging | Deleterious | Likely pathogenic |
| | SMAD4 | Splice donor | 0 | c.904+1..904+2in- | p.? | NA | NA | Pathogenic |
| | SMAD4 | Splice donor | 0 | sGCCGTTCACAA | p.? | NA | NA | Uncertain significance |
| | GGCX | Missense | 0 | c.904+3A>G c.2012G>A | p.(Arg671His) | Benign | Deleterious | Uncertain significance |

(Continues)

TABLE 2 (Continued)

| Patient | Gene(s) | VEP | MAX AF | Genomic variation | Protein effect | PolyPhen | SIFT | Pathogenicity |
|---------|--|---|-----------------------------|--|---|---|---|--|
| 27 | TUBB1 TPM4 | Missense Missense | 0.0053 0 | c.13G>A c.440C>T | p.(Val5Ile) p.(Ala147Val) | Probably_damaging Benign | Tolerated Deleterious | Uncertain significance Uncertain significance |
| 28 | PLAT | Missense | 0.00163 | c.928C>T | p.(Arg310Cys) | Possibly_damaging | Deleterious | Uncertain significance |
| 29.1 | TUBB1 (Johnson, Lowe, et al., 2016) | Missense | 0.0008 | c.721C>T | p.(Arg241Trp) | Probably_damaging | Deleterious | Uncertain significance |
| 29.2 | TUBB1 (Johnson, Lowe, et al., 2016) | Missense | 0.0008 | c.721C>T | p.(Arg241Trp) | Probably_damaging | Deleterious | Uncertain significance |
| 30 | GP1BA GP1BA GBA | Frameshift Frameshift Stop gained | 0 0.00235 0 | c.1274_1275del c.1277_1313del c.653G>A | p.(Glu425AlafsTer72) p.(Pro426ArgfsTer34) p.(Trp218Ter) | NA NA NA | NA NA NA | Likely pathogenic Uncertain significance Pathogenic |
| 31 | MPIG6B HPS6 | Splice region Inframe insertion | 0.00022 0 | c.621G>T c.256_264dup | p.? p.(Trp86_Ala88dup) | NA NA | NA NA | Uncertain significance Uncertain significance |
| 32.1 | CYCS (Johnson, Lowe, et al., 2016) | Missense | 0 | c.155C>T | p.(Ala52Val) | Benign | Tolerated | Likely pathogenic |
| 32.2 | GGCX | Missense | 0.0014 | c.1217G>A | p.(Arg406His) | Probably_damaging | Tolerated | Uncertain significance |
| | GGCX | Missense | 0 | c.155C>T | p.(Ala52Val) | Benign | Tolerated | Likely pathogenic |
| | CYCS (Johnson, Lowe, et al., 2016) | Missense | 0.0014 | c.1217G>A | p.(Arg406His) | Probably_damaging | Tolerated | Uncertain significance |
| 33 | COL5A1 TUBB1 (Johnson, Lowe, et al., 2016) | Missense Frameshift | 0 0 | c.1715C>A c.1080dup | p.(Pro572His) p.(Leu361AlafsTer19) | Probably_damaging NA | Deleterious NA | Uncertain significance Likely pathogenic |
| 34 | KLKB1 NBEAL2 | Missense Missense | 0.00132 0.00407 | c.772C>T c.2375G>A | p.(Leu258Phe) p.(Arg792Gln) | Benign Benign | Deleterious Deleterious | Uncertain significance Uncertain significance |
| 35 | COL5A1 TBXAS1 THBD | Missense Missense Missense | 0.00027 0.00162 0.001 | c.145C>T c.1523A>T c.407T>G | p.(His49Tyr) p.(Glu508Val) p.(Leu136Trp) | Probably_damaging Possibly_damaging Probably_damaging | Deleterious NA Deleterious | Uncertain significance Uncertain significance Uncertain significance |
| 36 | RASGRP2 VWF VWF | Missense Missense Missense | 0.0008 0.0024 0.00212 | c.281C>T c.6424C>T c.3365C>T | p.(Pro94Leu) p.(Leu2142Phe) p.(Thr1122Met) | Benign Probably_damaging Possibly_damaging | Tolerated Deleterious Deleterious | Uncertain significance Uncertain significance Uncertain significance |
| | GP1BA (Johnson, Lowe, et al., 2016) | Missense | 0.00417 | c.1761A>C | p.(Gln587His) | Unknown | Deleterious | Uncertain significance |
| 37.1 | GP1BA (Johnson, Lowe, et al., 2016) | Missense | 0 | c.413G>T | p.(Gly138Val) | Probably_damaging | Deleterious | Likely pathogenic |
| 37.2 | GP1BA (Johnson, Lowe, et al., 2016) | Missense | 0 | c.413G>T | p.(Gly138Val) | Probably_damaging | Deleterious | Likely pathogenic |
| 37.3 | GP1BA (Johnson, Lowe, et al., 2016) | Missense | 0 | c.3493C>T | p.(Arg1165Cys) | Probably_damaging | Deleterious | Pathogenic |

TABLE 2 (Continued)

| Patient | Gene(s) | VEP | MAX AF | Genomic variation | Protein effect | PolyPhen | SIFT | Pathogenicity |
|---------|------------------------------------|-----------------|----------|------------------------|--------------------|-------------------|-------------|------------------------|
| | MYH9 (Johnson, Lowe, et al., 2016) | | | | | | | |
| 38.1 | RUNX1 | Missense | 0 | c.611G>A | p.(Arg204Gln) | Possibly_damaging | Deleterious | Likely pathogenic |
| 38.2 | RUNX1 | Missense | 0 | c.611G>A | p.(Arg204Gln) | Possibly_damaging | Deleterious | Likely pathogenic |
| 38.3 | WAS (Johnson et al., 2018) | Missense | 0 | c.1456G>A | p.(Glu486Lys) | Probably_damaging | Deleterious | Uncertain significance |
| 38.4 | WAS (Johnson et al., 2018) | Missense | 0 | c.1456G>A | p.(Glu486Lys) | Probably_damaging | Deleterious | Uncertain significance |
| 38.5 | WAS (Johnson et al., 2018) | Missense | 0 | c.1456G>A | p.(Glu486Lys) | Probably_damaging | Deleterious | Uncertain significance |
| | WAS (Johnson et al., 2018) | | | | | | | |
| 39 | NVD | | | | | | | |
| 40 | NVD | | | | | | | |
| 41 | ACTN1 | Missense | 0 | c.2662G>C | p.(Gly888Arg) | Probably_damaging | Deleterious | Uncertain significance |
| | PLAT | Missense | 0.00139 | c.1481G>C | p.(Gly494Ala) | Probably_damaging | Tolerated | Uncertain significance |
| 42 | ABCG8 | Missense | 0.00157 | c.1924G>A | p.(Ala642Thr) | Benign | Tolerated | Uncertain significance |
| 43 | RUNX1 | Splice acceptor | 0 | c.98-1G>A | p.? | NA | NA | Pathogenic |
| | VWF (Lester et al., 2007) | Missense | 0.001 | c.7390C>T | p.(Arg2464Cys) | Probably_damaging | Deleterious | Pathogenic |
| | GBA | Frameshift | <0.0001 | c.26_27del | p.(Glu9GlyfsTer8) | NA | NA | Likely pathogenic |
| | PROS1 (Alhenc-Gelas et al., 2010) | Missense | 0.000572 | c.284G>A | p.(Gly95Glu) | Possibly_damaging | Deleterious | Likely pathogenic |
| 44.1 | RUNX1 (Stockley et al., 2013) | Splice donor | 0 | c.351+1G>T | p.? | NA | | Pathogenic |
| 44.2 | RUNX1 (Stockley et al., 2013) | Splice donor | 0 | c.351+1G>T | p.? | NA | | Pathogenic |
| 45 | HRG | Missense | 0.00162 | c.1379G>A | p.(Arg460Gln) | Benign | Tolerated | Likely benign |
| 46 | FLNA (Johnson et al., 2018) | Missense | 0.0073 | c.5948C>T | p.(Ser1983Leu) | Probably_damaging | Deleterious | Likely benign |
| 47 | GGCX | Missense | 0.0014 | c.1217G>A | p.(Arg406His) | Probably_damaging | Tolerated | Uncertain significance |
| 48 | RUNX1 (Lamolda et al., 2019) | Missense | 0 | c.586A>G | p.(Thr196Ala) | Possibly_damaging | Deleterious | Likely pathogenic |
| | SLFN14 | Missense | 0 | c.2686T>C | p.(Ser896Pro) | Probably_damaging | Deleterious | Uncertain significance |
| | SLFN14 | Missense | 0 | c.1481A>G | p.(Gln494Arg) | Benign | Deleterious | Uncertain significance |
| | SLFN14 | Missense | 0 | c.859A>G | p.(Lys287Glu) | Probably_damaging | Tolerated | Uncertain significance |
| | SLFN14 | Frameshift | 0 | c.3_4insCTAGTC-GACTATA | p.(Glu2LeufsTer10) | NA | NA | Pathogenic |
| 49 | ABCG5 | Missense | <0.0001 | c.692T>C | p.(Ile231Thr) | Probably_damaging | Deleterious | Uncertain significance |

(Continues)

TABLE 2 (Continued)

| Patient | Gene(s) | VEP | MAX AF | Genomic variation | Protein effect | PolyPhen | SIFT | Pathogenicity |
|---------|------------------------------------|---------------|----------|-------------------|----------------------------|-------------------|-------------|------------------------|
| 50.1 | STXBP2 | Missense | 0.00458 | c.1586G>C | p.(Arg529Pro) | Probably_damaging | Deleterious | Uncertain significance |
| 50.2 | NVD | | | | | | | |
| 51.1 | ADAMTS13 | Splice region | <0.0001 | c.3568+7T>G | p.? | NA | NA | Uncertain significance |
| 51.2 | ADAMTS13 | Splice region | <0.0001 | c.3568+7T>G | p.? | NA | NA | Uncertain significance |
| 51.3 | ADAMTS13 | Splice region | <0.0001 | c.3568+7T>G | p.? | NA | NA | Uncertain significance |
| 52.1 | LYST | Stop gained | 0 | c.4288C>T | p.(Arg1430Ter) | NA | NA | Likely pathogenic |
| 52.2 | THBD (Dargaud et al., 2015) | Stop gained | 0 | c.1611C>A | p.(Cys537Ter) | NA | NA | Pathogenic |
| | | Stop gained | 0 | c.4288C>T | p.(Arg1430Ter) | NA | NA | Likely pathogenic |
| | | Stop gained | 0 | c.1611C>A | p.(Cys537Ter) | NA | NA | Pathogenic |
| 53.1 | ADAMTS13 | Stop gained | 0 | c.1315G>T | p.(Glu439Ter) | NA | NA | Pathogenic |
| 53.2 | ADAMTS13 | Stop gained | 0 | c.1315G>T | p.(Glu439Ter) | NA | NA | Pathogenic |
| 54.1 | NVD | | | | | | | |
| 54.2 | NVD | | | | | | | |
| 55 | F10 | Missense | 0.0007 | c.1222G>A | p.(Asp408Asn) | Benign | Deleterious | Uncertain significance |
| | MPL | Missense | 0.000297 | c.712G>T | p.(Gly238Cys) | Possibly_damaging | Deleterious | Uncertain significance |
| 56 | GF1B (Johnson, Lowe, et al., 2016) | Splice donor | 0 | c.814+1G>A | p.? | NA | NA | Pathogenic |
| 57 | ACVRL1 | Missense | 0.00432 | c.1445C>T | p.(Ala482Val) | Probably_damaging | Deleterious | Uncertain significance |
| | ITGA2B (Johnson et al., 2018) | Stop gained | 0 | c.2176A>T | p.(Lys726Ter) | NA | NA | Pathogenic |
| | THPO | Frameshift | 0.0001 | c.610dup | p.(Glu204- GlyfsTer123) | NA | NA | Uncertain significance |
| | HOXA11 | Missense | <0.0001 | c.248A>G | p.(Tyr83Cys) | Benign | Deleterious | Uncertain significance |
| 58.1 | NBEAL2 | Splice donor | 0 | c.6801+1G>C | p.? | NA | NA | Pathogenic |
| 58.2 | NBEAL2 | Splice donor | 0 | c.6801+1G>C | p.? | NA | NA | Pathogenic |
| 59 | MECOM | Missense | 0.000175 | c.580T>G | p.(Tyr194Asp) | NA | NA | Uncertain significance |
| | F8 | Missense | 0 | c.5441A>T | p.(Asp1814Val) | Benign | Deleterious | Likely pathogenic |
| 60 | MPL | Missense | 0.00472 | c.1063A>G | p.(Lys355Glu) | Benign | Tolerated | Uncertain significance |
| | NBEAL2 | Missense | 0.000184 | c.5866G>A | p.(Val1956Met) | Possibly_damaging | Tolerated | Uncertain significance |
| | PROZ | Missense | 0.00346 | c.647C>T | p.(Thr216Ile) | Probably_damaging | Deleterious | Uncertain significance |
| 61 | COL5A1 | Missense | 0.000124 | c.2146G>A | p.(Glu716Lys) | Possibly_damaging | Deleterious | Uncertain significance |
| 62 | NVD | | | | | | | |
| 63.1 | GF1B (Rabbolini et al., 2017) | Missense | 0.00438 | c.503G>T | p.(Cys168Phe) | Probably_damaging | Deleterious | Likely pathogenic |
| 63.2 | | Missense | 0.00438 | c.503G>T | p.(Cys168Phe) | Probably_damaging | Deleterious | Likely pathogenic |

TABLE 2 (Continued)

| Patient | Gene(s) | VEP | MAX AF | Genomic variation | Protein effect | PolyPhen | SIFT | Pathogenicity |
|---------|-------------------------------------|-------------|----------|-------------------|----------------|-------------------|-------------|------------------------|
| 63.3 | GFI1B (Rabbolini et al., 2017) | Missense | 0.00438 | c.503G>T | p.(Cys168Phe) | Probably_damaging | Deleterious | Likely pathogenic |
| | GFI1B (Rabbolini et al., 2017) | | | | | | | |
| 64 | NVD | | | | | | | |
| 65.1 | MECOM | Missense | 0.0001 | c.951G>T | p.(Lys317Asn) | NA | NA | Uncertain significance |
| 65.2 | SLFN14 | Missense | 0.0014 | c.916G>C | p.(Asp306His) | Benign | Tolerated | Uncertain significance |
| | MECOM | Missense | 0.0001 | c.951G>T | p.(Lys317Asn) | NA | NA | Uncertain significance |
| | SLFN14 | Missense | 0.0014 | c.916G>C | p.(Asp306His) | Benign | Tolerated | Uncertain significance |
| 66.1 | MYH9 (Savoia & Pecci, 2015) | Stop gained | 0.0001 | c.5797C>T | p.(Arg1933Ter) | NA | NA | Pathogenic |
| 66.2 | MYH9 (Savoia & Pecci, 2015) | Stop gained | 0.0001 | c.5797C>T | p.(Arg1933Ter) | NA | NA | Pathogenic |
| 67 | NVD | | | | | | | |
| 68 | RASGRP2 | Missense | 0.000102 | c.1159C>T | p.(Arg387Cys) | Probably_damaging | Deleterious | Uncertain significance |
| 69 | F10 | Missense | 0.000547 | c.1406G>A | p.(Arg469Lys) | Benign | Tolerated | Uncertain significance |
| 70.1 | ETV6 | Stop gained | 0 | c.313C>T | p.(Arg105Ter) | NA | NA | Likely pathogenic |
| 70.2 | NVD | | | | | | | |
| 71 | RUNX1 RUNX1 | Missense | 0 | c.1256T>G | P.(Val419Gly) | Benign | Deleterious | Likely pathogenic |
| | | Missense | 0 | c.1270T>C | p.(Ser424Pro) | Possibly_damaging | Deleterious | Likely pathogenic |
| 72 | FGG | Missense | 0.00792 | c.323C>G | p.(Ala108Gly) | Benign | Tolerated | Likely pathogenic |
| | STXBP2 | Missense | 0.000539 | c.499C>T | p.(Arg167Trp) | Possibly_damaging | Deleterious | Uncertain significance |
| | TUBB1 (Johnson, Lowe, et al., 2016) | Missense | 0.0008 | c.721C>T | p.(Arg241Trp) | Probably_damaging | Deleterious | Uncertain significance |
| 73 | NVD | | | | | | | |
| 74 | RUNX1 | Missense | 0 | c.1265A>C | p.(Glu422Ala) | Benign | Deleterious | Uncertain significance |
| | COL5A1 | Missense | 0.000121 | c.5411C>A | p.(Thr1804Asn) | Benign | Deleterious | Uncertain significance |
| 75 | GFI1B (Rabbolini et al., 2017) | Missense | 0.00438 | c.503G>T | p.(Cys168Phe) | Probably_damaging | Deleterious | Likely pathogenic |
| 76 | THBD | Missense | 0 | c.716C>T | p.(Ala239Val) | Benign | Tolerated | Uncertain significance |
| 77 | THBD | Missense | 0 | c.752G>A | p.(Gly251Asp) | Probably_damaging | Deleterious | Likely pathogenic |
| | COL5A2 | Missense | 0.0001 | c.2786C>T | p.(Ala929Val) | Probably_damaging | Tolerated | Uncertain significance |

(Continues)

TABLE 2 (Continued)

| Patient | Gene(s) | VEP | MAX AF | Genomic variation | Protein effect | PolyPhen | SIFT | Pathogenicity |
|---------|---------|-------------|----------|-------------------|----------------|-------------------|-------------|------------------------|
| 78 | STXBP2 | Missense | 0.000201 | c.911C>T | p.(Thr304Met) | Probably_damaging | Deleterious | Uncertain significance |
| | MCFD2 | Missense | 0 | c.364G>A | p.(Asp122Asn) | Probably_damaging | Deleterious | Likely pathogenic |
| 79 | NBEAL2 | Missense | 0.000128 | c.3184G>A | p.(Val1062Ile) | Possibly_damaging | Deleterious | Uncertain significance |
| | AP3D1 | Missense | 0.0014 | c.1363G>A | p.(Ala455Thr) | Possibly_damaging | Deleterious | Uncertain significance |
| 80 | RUNX1 | Missense | 0 | c.1270T>G | p.(Ser424Ala) | Possibly_damaging | Deleterious | Likely pathogenic |
| | MPL | Missense | 0 | c.305G>A | p.(Arg102His) | Probably_damaging | Deleterious | Uncertain significance |
| 81 | AP3B1 | Missense | 0.000809 | c.2188C>T | p.(Arg730Trp) | Benign | Deleterious | Uncertain significance |
| 82 | TUBB1 | Missense | 0.0001 | c.4C>T | p.(Arg2Cys) | Probably_damaging | Deleterious | Uncertain significance |
| | TUBB1 | Missense | 0.0002 | c.68T>C | p.(Met23Thr) | Benign | Tolerated | Uncertain significance |
| 83 | LPA | Stop gained | 0.001 | c.5081C>G | p.(Ser1694Ter) | NA | NA | Uncertain significance |
| 84.1 | F5 | Missense | 0.00806 | c.5245C>G | p.(Leu1749Val) | Possibly_damaging | Tolerated | Uncertain significance |
| 84.2 | F5 | Missense | 0.00806 | c.5054C>G | p.(Thr1685Ser) | Benign | Tolerated | Uncertain significance |
| NBEAL2 | HP55 | Missense | 0.00276 | c.5021G>A | p.(Arg1674His) | Benign | Deleterious | Uncertain significance |
| | | Missense | 0.00593 | c.345G>A | p.(Met115Ile) | Benign | Tolerated | Uncertain significance |

Note: Variants previously reported in the literature are indicated. NCBI reference sequences: ABCG5 (NM_022436.2); ABCG8 (NM_022437.2); ACTN1 (NM_001102.3); ACVRL1 (NM_001077401.1); ADAMTS13 (NM_139025.4); ANKRD26 (NM_001256053.1); AP3B1 (NM_003664.4); AP3D1 (NM_003938.6); ARPC1B (NM_005720.3); COL5A1 (NM_000093.4); COL5A2 (NM_000393.3); CYCS (NM_018947.5); ETV6 (NM_001987.4); F10 (NM_000504.3); F11 (NM_000128.3); F13A1 (NM_000129.3); F5 (NM_000129.3); F7 (NM_000130.4); F8 (NM_000131.4); F9 (NM_000132.3); FGA (NM_000508.3); FGB (NM_005141.4); FGG (NM_021870.2); FLI1 (NM_002017.4); FLNA (NM_001456.3); GBA (NM_00157.3); GF11B (NM_004188.4); GGCX (NM_000821.5); GP1BA (NM_000173.5); GP6 (NM_001083899.1); HOXA11 (NM_005523.5); HPS3 (NM_032383.3); HPS5 (NM_024747.5); HRG (NM_000412.2); ITGA2B (NM_000419.3); ITGB3 (NM_000212.2); KLKB1 (NM_000892.3); LPA (NM_005577.2); LYST (NM_000081.3); MCFD2 (NM_139279.5); MECOM (NM_004991.3); MP1G6B (NM_138272.2); MPL (NM_005373.2); MYH9 (NM_002473.4); NBEA (NM_015678.4); NBEAL2 (NM_015175.3); P2RY12 (NM_176876.2); PLAT (NM_000930.3); PLG (NM_000301.3); PROS1 (NM_000313.3); PROZ (NM_003891.2); PTPN11 (NM_006949.3); RASGRP2 (NM_153819.1); RUNX1 (NM_001754.4); SERPINC1 (NM_000488.3); SLC45A2 (NM_016180.3); SLFN14 (NM_001129820.1); SMAD4 (NM_005359.5); STXBP2 (NM_006949.3); TBXAS1 (NM_001061.4); THBD (NM_000361.2); THPO (NM_000460.2); TPM4 (NM_003290.2); TUBB1 (NM_030773.3); VPS33B (NM_018668.3); VWF (NM_000552.2); WAS (NM_000377.2).

Abbreviations: NA, not available; NVD, no variant detected; SIFT, sorting intolerance from tolerance; VEP, variant effect predictor.

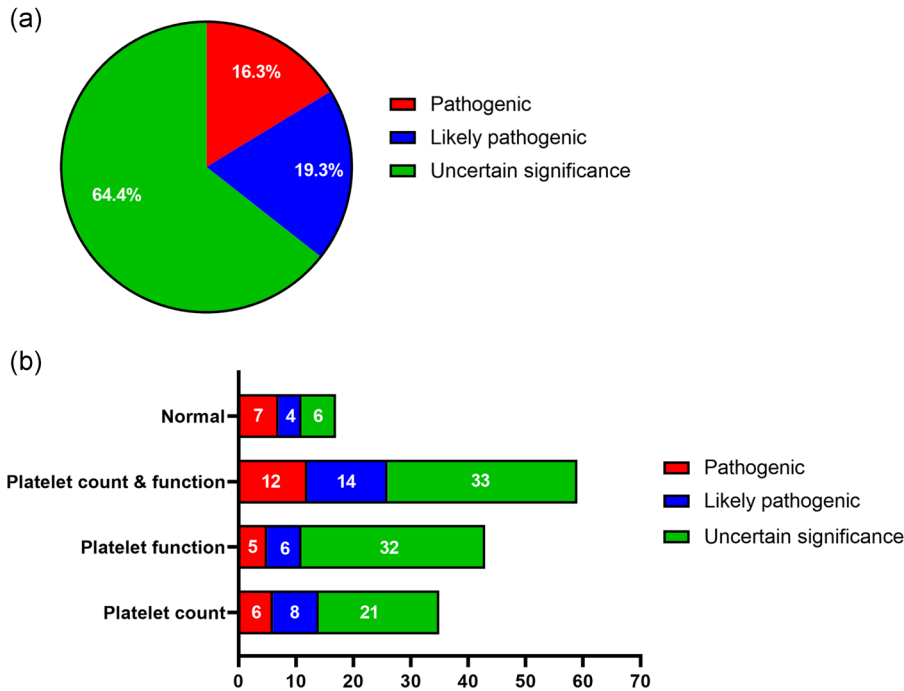


FIGURE 2 Pathogenicity prediction of genetic variants based on the American College of Medical Genetics and Genomics consensus guidelines. (a) Classification of variants based on the pathogenicity prediction analysis. (b) Number of reported variants for the patients in each different class of platelet phenotype observed

candidate genes within the IBD gene panel. First, a rare 604-kbp CNV loss was noted on chromosome 11q24.3 in patient 35 which covered nine genes including *FLI1*. A further rare deletion was found on chromosome 11q24.3 in patient 71 which covered 31 genes including *FLI1* (Figure 3a,b).

Following ExomeDepth alignment with a panel of controls the reads ratio was around 0.5 which indicates heterozygosity, as observed in Table 3. A rare CNV gain was noted in patient 45 within *TBXA2R* on chromosome 19p13.3 and containing four genes in total (Figure 3c). The CNV reads ratio was 2.72 which is indicative of a heterozygous insertion.

3.6 | Oligogenic findings in patient cohort

Within the patient cohort there were several examples of potential oligogenic inheritance involving either two or more gene variations from the IBD gene panel. Of particular interest was patient 16 who demonstrated an apparent pathogenic missense variant in *RUNX1* and a likely pathogenic variant in *ITGB3*, both of which are plausible candidate variants to explain the thrombocytopenia and bleeding history observed. Patient 20 harbored two heterozygous missense variants within *GP6* and *THBD* in which the patient had a platelet function disorder and episodes of bleeding. In patient 30, likely pathogenic and pathogenic variations were found in *GP1BA* and *GBA* respectively. Again this patient had a low platelet count and a history of bleeding.

4 | DISCUSSION

NGS approaches have increasingly been used over the last decade in the molecular diagnosis of IBD. Here, we present a large-scale application of WES analysis by using a robust molecular diagnostics

platform for diagnosis of 117 patients recruited to the UK-GAPP study. The aim was to assess the ability of Congenica software to analyze WES data of the patients for both sequence and structural variants by targeting a known panel of IBD genes. Subsequently, patients with variants in known bleeding disorder genes can be eliminated by a series of filtration steps and WES data targeted for those with undetected variants who may harbor a variant in novel genes. We included and applied a total of 119 genes to our patient cohort for filtering; it is, however, important to note that currently only 88 of these genes are considered clinical-grade genes according to Genomics England. The remainder of the genes may become clinical grade once more variants are identified in patients and deemed pathogenic over time.

Phenotypic presentation and platelet counts varied considerably among our recruited patients, which is consistent with the variability of clinical presentation between patients with suspected IBDs. However, the majority of patients 33/117 (28.2%) were noted with a platelet function defect and 23/117 (19.6%) patients represented thrombocytopenia. Of the 117 recruited patients, 15 (12.8%) were deemed to have a macrothrombocytopenia. Platelet function studies revealed the presence of a combination of platelet defects in addition to thrombocytopenia in 36/117 (30%) of patients. The majority of the patients with platelet defects displayed both secretion and Gi defects (Supplementary figure S2). However, a previous study has shown that some patients with normal lumi-aggregeometry results have platelet spreading defects, indicating the difficulties faced when diagnosing patients with IBDs and the multitude of assays required for platelet phenotype disorders to be diagnosed (Khan et al., 2020).

Overall, a total of 135 variants in genes implicated in bleeding disorders were identified across all 117 patients and all variants were observed in a heterozygous state, implicating dominant inheritance patterns. The study has shown that the majority of

TABLE 3 Copy number variations detected in 15 patients by using ExomeDepth calling approach

| Patient | Gene | Band | Location | Size | Type | Reads expected | Reads observed | Reads ratio | CNV | Bays factor | Patients overlap |
|---------|----------|---------------|------------------------|-----------|------|----------------|----------------|-------------|-------|-------------|------------------|
| 2.1 | ANKR-D26 | 10p12.1 | 10:27280843-27389421 | 10.8 kbp | Loss | 254 | 157 | 0.62 | 1.236 | 8.26 | 1 |
| 3.1 | ANKR-D26 | 10p12.1 | 10:27280843-27389421 | 10.8 kbp | Gain | 271 | 369 | 1.36 | 2.72 | 6.13 | 1 |
| 21 | PIGA | Xp22.31-p21.3 | X:15337573-15353676 | 21.0 Mbp | Loss | 187,776 | 105,800 | 0.56 | 1.126 | 3050 | 16 |
| | GATA1 | Xp21.1-q13.3 | X:48644962-48652716 | 41.5 Mbp | Loss | 458,560 | 256,748 | 0.56 | 1.12 | 6580 | 27 |
| | WAS | Xp21.1-q13.3 | X:48534985-48549818 | 41.5 Mbp | Loss | 458,560 | 256,748 | 0.56 | 1.12 | 6580 | 27 |
| | F9 | Xq25-q27.2 | X:138612917-138645617 | 19.0 Mbp | Loss | 153,271 | 86,478 | 0.56 | 1.128 | 2360 | 24 |
| | F8 | Xq27.3-q28 | X:154064063-154255215 | 8.3 Mbp | Loss | 165,645 | 90,968 | 0.55 | 1.098 | 2410 | 31 |
| | FLNA | Xq27.3-q28 | X:153576892-153603006 | 8.3 Mbp | Loss | 165,645 | 90,968 | 0.55 | 1.098 | 2410 | 31 |
| 26 | SMAD4 | 18q21.2 | 18:48494410-48611415 | 2.0 kbp | Gain | 355 | 495 | 1.39 | 2.78 | 10.8 | 0 |
| 35 | FLI1 | 11q24.3 | 11:128556430-128683162 | 604.6 kbp | Loss | 9862 | 5420 | 0.55 | 1.1 | 422 | 1 |
| 45 | SLFN14 | 17q12 | 17:33875144-33885117 | 5.2 kbp | Loss | 833 | 402 | 0.48 | 0.966 | 6.35 | 4 |
| | TBXA2R | 19p13.3 | 19:3594504-3606838 | 43.5 kbp | Gain | 1467 | 1993 | 1.36 | 2.72 | 12.2 | 0 |
| | GP6 | 19q13.42 | 19:5525073-55549632 | 77.3 kbp | Loss | 1667 | 843 | 0.51 | 1.012 | 39.2 | 0 |
| 46 | FYB1 | 5p13.1 | 5:39105338-39274630 | 145.7 kbp | Gain | 2146 | 2630 | 1.23 | 2.46 | 12.6 | 2 |
| | NBEA | 13q13.3 | 13:35516424-36247159 | 98.7 kbp | Gain | 527 | 742 | 1.41 | 2.82 | 6.86 | 0 |
| | SLFN14 | 17q12 | 17:33875144-33885117 | 9.9 kbp | Loss | 7036 | 3693 | 0.52 | 1.05 | 15.1 | 4 |
| 47 | SLFN14 | 17q12 | 17:33875144-33885117 | 9.9 kbp | Loss | 4964 | 2886 | 0.58 | 1.162 | 6.03 | 4 |
| | GP1BB | 22q11.21 | 22:19710468-19712294 | 37.8 kbp | Loss | 421 | 224 | 0.53 | 1.064 | 8.47 | 0 |
| 48 | SLFN14 | 17q12 | 17:33875144-33885117 | 9.9 kbp | Gain | 4090 | 11,381 | 2.78 | 5.56 | 13.3 | 4 |
| 52.1 | PIGA | Xp22.2-p22.13 | X:15337573-15353676 | 3.6 Mbp | Loss | 49,781 | 32,738 | 0.66 | 1.316 | 275 | 3 |
| | GATA1 | Xp11.3-p11.22 | X:48644962-48652716 | 5.3 Mbp | Loss | 154,142 | 94,557 | 0.61 | 1.226 | 756 | 15 |
| | WAS | Xp11.3-p11.22 | X:48534985-48549818 | 5.3 Mbp | Loss | 154,142 | 94,557 | 0.61 | 1.226 | 756 | 15 |
| | F9 | Xq27.1 | X:138612917-138645617 | 476.4 kbp | Loss | 11,569 | 7853 | 0.68 | 1.358 | 53.5 | 1 |
| | F8 | Xq28 | X:154064063-154255215 | 4.4 Mbp | Loss | 148,620 | 91,530 | 0.62 | 1.232 | 730 | 15 |
| | FLNA | Xq28 | X:154064063-154255215 | 4.4 Mbp | Loss | 148,620 | 91,530 | 0.62 | 1.232 | 730 | 15 |
| 52.2 | PIGA | Xp22.33-p21.3 | X:15337573-15353676 | 23.5 Mbp | Gain | 205,946 | 298,796 | 1.45 | 2.9 | 2890 | 21 |
| | GATA1 | Xp11.3-q13.3 | X:48644962-48652716 | 29.3 Mbp | Gain | 361,570 | 519,986 | 1.44 | 2.88 | 4150 | 25 |
| | WAS | Xp11.3-q13.3 | X:48534985-48549818 | 29.3 Mbp | Gain | 361,570 | 519,986 | 1.44 | 2.88 | 4150 | 25 |
| | F9 | Xq25-q27.3 | X:138612917-138645617 | 14.4 Mbp | Gain | 134,092 | 194,039 | 1.45 | 2.9 | 1690 | 21 |
| | F8 | Xq27.3-q28 | X:154064063-154255215 | 9.9 Mbp | Gain | 153,704 | 221,043 | 1.44 | 2.88 | 1830 | 38 |
| 57 | COL5A1 | 9q34.3 | 9:137533620-137736686 | 190 bp | Loss | 158 | 108 | 0.68 | 1.368 | 5.27 | 0 |

TABLE 3 (Continued)

| Patient | Gene | Band | Location | Size | Type | Reads expected | Reads observed | Reads ratio | CNV | Bays factor | Patients overlap |
|---------|------|---------------|------------------------|-----------|------|----------------|----------------|-------------|-------|-------------|------------------|
| 67 | PIGA | Xp22.2-p22.11 | X:15337573-15353676 | 9.8 Mbp | Gain | 56,608 | 80,502 | 1.42 | 2.84 | 1070 | 10 |
| | F9 | Xq26.3-q27.2 | X:138612917-138645617 | 5.4 Mbp | Gain | 22,574 | 32,057 | 1.42 | 2.84 | 386 | 21 |
| | FLNA | Xq28 | X:153576892-153603006 | 6.0 Mbp | Gain | 58,278 | 86,619 | 1.49 | 2.98 | 1350 | 39 |
| 71 | FLI1 | 11q24.3 | 11:128556430-128683162 | 2.5 Mbp | Loss | 20,510 | 10,805 | 0.53 | 1.054 | 912 | 2 |
| 72 | F9 | Xq26.3-q27.2 | X:138612917-138645617 | 5.4 Mbp | Gain | 23,065 | 32,667 | 1.42 | 2.84 | 469 | 21 |
| | FLNA | Xq28 | X:153576892-153603006 | 260.1 kbp | Gain | 15,952 | 22,830 | 1.43 | 2.86 | 322 | 7 |

Abbreviation: CNV, copy number variation.

plausible candidate variants were associated with IT genes which explain the association of thrombocytopenia with platelet defects in the majority of patients. When considering pathogenicity prediction, 22 patients were classified as pathogenic and 26 patients as likely pathogenic, while 87 patients had uncertain pathogenicity and therefore classified as uncertain significance. A targeted WES analysis was previously carried out on some patients which identified genetic variants in inherited thrombocytopenia with or without secondary qualitative defects (Johnson et al., 2018; Johnson, Lowe, et al., 2016). This study has conclusively identified these genetic variants, which indicates the ability of the Congenica platform to analyze and provide suitable validation of WES data in these patients.

If we compare the performance of the Congenica tool employed here with other bioinformatic platforms we observe the following: 25 variants were identified by the Congenica software as well as other bioinformatic tools and the majority of them were either pathogenic or likely pathogenic. A further 24 variants were classified as pathogenic or likely pathogenic by the Congenica software only. Therefore, this showed that the Congenica software is a more robust tool to analyze WES as it provides a higher variant detection rate compared with other bioinformatic tools. It is also important to note that we did not include variants of uncertain significance here in this evaluation as it is difficult to assign causality but are still plausible pathogenic variants. Congenica software also has the added benefit of detecting CNVs, a process which is notoriously difficult yet valuable in identifying rare causative variants in heterogeneous diseases. Congenica utilizes the integrated ExomeDepth tool to compare a target with reference and here, identified rare CNVs in this population. Congenica software alongside targeted gene panel searching, allows for efficient and accessible detection of variants and with some clinical interpretation will be a valuable tool when analyzing large datasets (Nowakowska, 2017; Valsesia et al., 2012).

Paris-Trousseau syndrome, characterized by a bleeding defect with large α -granules and abnormal megakaryocyte morphology is well documented, which is caused by a dominant inheritance of q23 deletion on chromosome 11 (Stevenson et al., 2015). Patients with this disorder have variable size of chromosomal deletion associated with different components of the syndrome. A hemizygous deletion of *FLI1* was attributed to the platelet defect in two individuals of our cohort. These CNVs noted in patients 35 and 71 cover the deletion region in *FLI1* and are also surrounded by several flanking genes. Both patients presented with thrombocytopenia and a secretion defect which suggest the platelet phenotype and the CNVs in *FLI1* to be associated with their disorders. Thromboxane receptor deficiency is an autosomal recessive or dominant disorder characterized by bleeding symptoms associated with quantitative or qualitative defects within the thromboxane receptor (Mundell & Mumford, 2018). Although we did not find any plausible candidate SNVs in the 119 candidate genes or the thromboxane receptor in patient 45, we did note a rare CNV duplication in the *TBXA2R* gene and deduce that either alone or in combination with variants in *GP6* and *SLFN14*

(a)

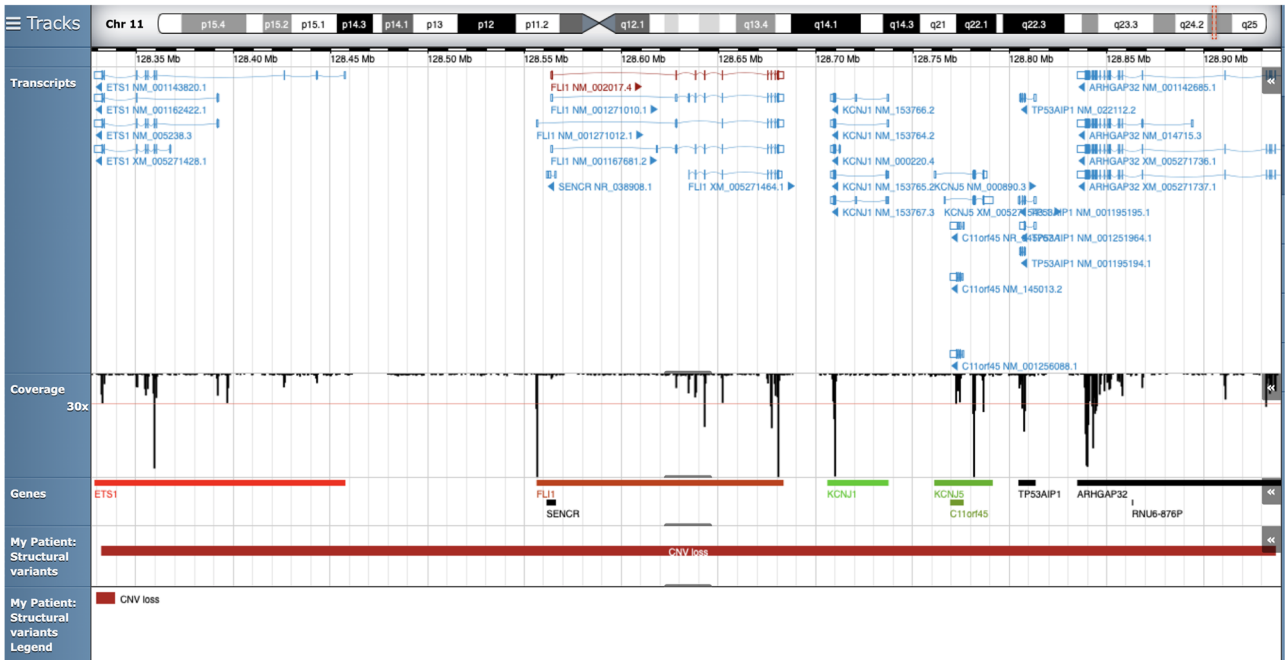


FIGURE 3 Copy number variants found in cohort of GAPP patients. Screenshots from the Congenica software CNV/structural variant tab in patients showing (a) copy number variant (loss) found in patient 35 which includes *FLI1* and eight other genes on chromosome 11q24.3; (b) copy number variant (loss) found in patient 71 which includes *FLI1* and 30 other genes on chromosome 11q24.3; and (c) copy number variant (gain) found including *TBXA2R* in patient 45 and three other genes within chromosome 19p13.3

which were also detected, could be causative of the patient's thrombocytopenia and bleeding. In the future, it would be interesting to investigate these CNVs further to determine the extent of the contiguous deletions or insertions by long-range polymerase chain

reaction and sequencing to determine the breakpoints and mechanisms of the variant, as well as confirming these regions using multiplex ligation-dependent probe amplification, should kits be available for these genomic regions.

(b)

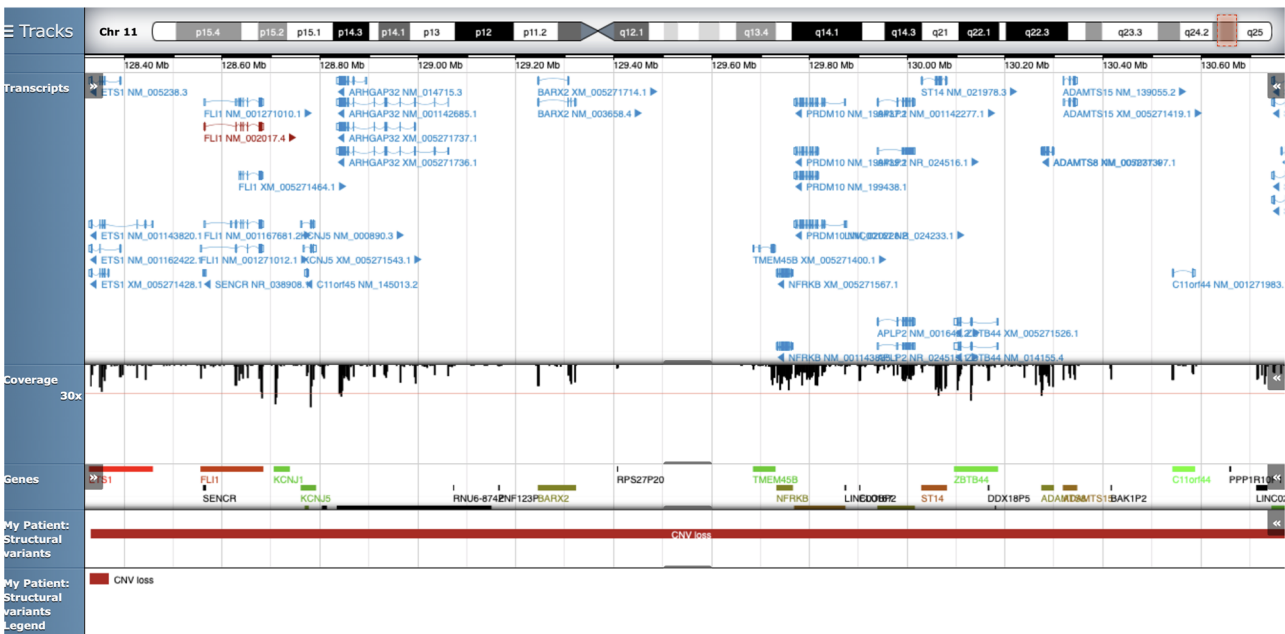
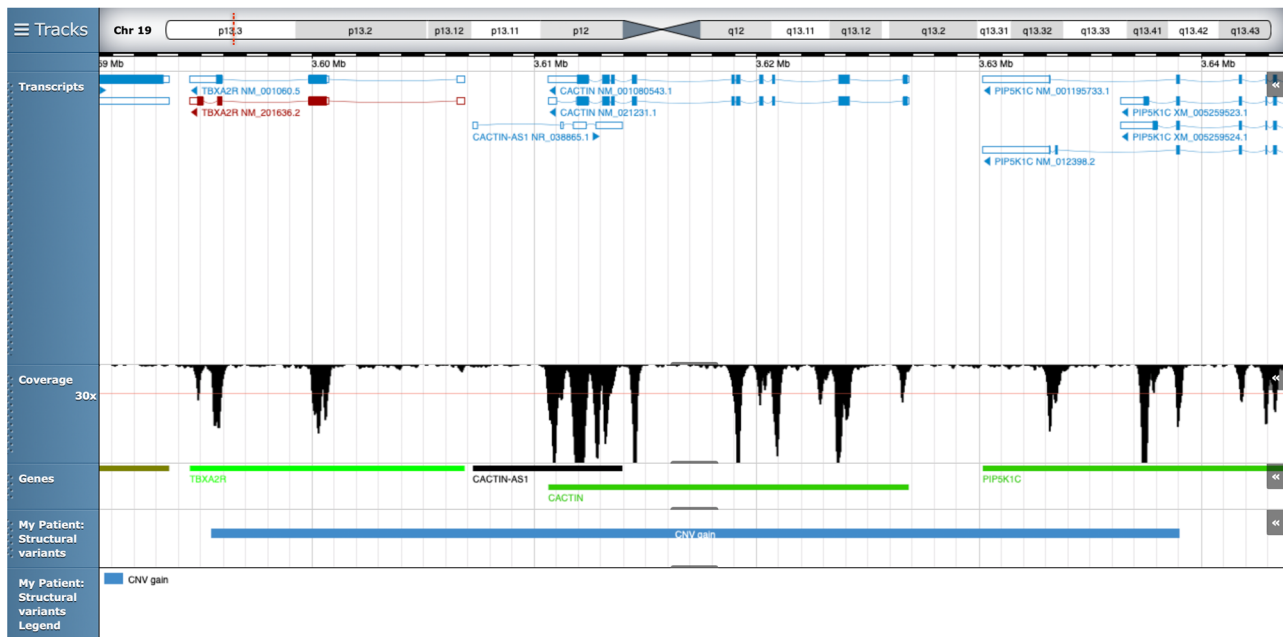


FIGURE 3 Continued

(c)

**FIGURE 3** Continued

In summary, we show validation and a practical approach of a robust diagnostic platform that can be employed for WES analysis. In this study, we use data from a cohort of patients with suspected IBDs; a broad category of diseases, well acknowledged in the hematology field as difficult to classify and associate to single causative genetic abnormalities. This study has shown the ability of the software to detect CNVs with high efficiency with the use of targeted gene panels as a replacement of traditional methods for detecting CNVs.

To conclude, our data reveals use of a highly sensitive and valuable tool which can be used for detecting SNVs and CNVs based on WES data. To our knowledge, this is one of the first studies, although in a research setting, to implement this software for both SNV and CNV analysis. We see this as a leap forward in the ability to classify highly complex disorders with a high degree of heterogeneity within the wider scientific community providing concise and definitive diagnosis for patients.

ACKNOWLEDGMENTS

We thank the Congenica software team, particularly Maria Valencia and Martin Shumanov for help in troubleshooting the software and providing support with the initial sample uploads. The work in the author's laboratories is supported by the British Heart Foundation (PG/13/36/30275; FS/15/18/31317; PG/16/103/32650; FS/18/11/33443; Neil V. Morgan) and the Saudi Arabia Cultural Bureau in London (Ibrahim Almazni). Abdullah O. Khan is a Wellcome-funded Sir Henry Wellcome Fellow (218649/Z/19/Z).

CONFLICT OF INTERESTS

The authors declare that there are no conflict of interests.

AUTHOR CONTRIBUTIONS

Neil V. Morgan designed the research. Ibrahim Almazni performed the experiments and analysis. Rachel J. Stapley performed analysis of the data. All authors contributed to the writing of the manuscript and revised versions.

DATA AVAILABILITY STATEMENT

The variants reported in this manuscript have now been submitted to a public database and can be found at ClinVar (<https://www.ncbi.nlm.nih.gov/clinvar/>).

ORCID

Neil V. Morgan  <http://orcid.org/0000-0001-6433-5692>

REFERENCES

- Alhenc-Gelas, M., Canonico, M., Morange, P. -E., Emmerich, J., & Group, G. G. T. (2010). Protein S inherited qualitative deficiency: Novel mutations and phenotypic influence. *Journal of Thrombosis and Haemostasis*, 8, 2718–2726.
- Bastida, J. M., Lozano, M. L., Benito, R., Janusz, K., Palma-Barqueros, V., Del Rey, M., Hernández-Sánchez, J. M., Riesco, S., Bermejo, N., González-García, H., Rodríguez-Alén, A., Aguilar, C., Sevivas, T., López-Fernández, M. F., Marneth, A. E., van der Reijden, B. A., Morgan, N. V., Watson, S. P., Vicente, V., ... González-Porras, J. R. (2018). Introducing high-throughput sequencing into mainstream genetic diagnosis practice in inherited platelet disorders. *Haematologica*, 103(1), 148–162. <https://doi.org/10.3324/haematol.2017.171132>
- Blanchette, V. S., Sparling, C., & Turner, C. (1991). Inherited bleeding disorders. *Bailliere's Clinical Haematology*, 4(2), 291–332.
- Daly, M. E., Leo, V. C., Lowe, G. C., Watson, S. P., & Morgan, N. V. (2014). What is the role of genetic testing in the investigation of patients with suspected platelet function disorders? *British Journal of Haematology*, 165(2), 193–203.

- Dargaud, Y., Scazecz, J. Y., Wielders, S. J., Trzeciak, C., Hackeng, T. M., Négrier, C., Hemker, H. C., Lindhout, T., & Castoldi, E. (2015). Characterization of an autosomal dominant bleeding disorder caused by a thrombomodulin mutation. *Blood. The Journal of the American Society of Hematology*, *125*, 1497–1501.
- Downes, K., Megy, K., Duarte, D., Vries, M., Gebhart, J., Hofer, S., Shamardina, O., Deevi, S. V. V., Stephens, J., Mapeta, R., Tuna, S., Al Hasso, N., Besser, M. W., Cooper, N., Daugherty, L., Gleadall, N., Greene, D., Haimel, M., Martin, H., ... Freson, K. (2019). Diagnostic high-throughput sequencing of 2396 patients with bleeding, thrombotic, and platelet disorders. *Blood*, *134*(23), 2082–2091. <https://doi.org/10.1182/blood.2018891192>
- Fletcher, S. J., Johnson, B., Lowe, G. C., Bem, D., Drake, S., Lordkipanidzé, M., Guiú, I. S., Dawood, B., Rivera, J., Simpson, M. A., Daly, M. E., Motwani, J., Collins, P. W., Watson, S. P., & Morgan, N. V. (2015). SLFN14 mutations underlie thrombocytopenia with excessive bleeding and platelet secretion defects. *Journal of Clinical Investigation*, *125*(9), 3600–3605. <https://doi.org/10.1172/jci80347>
- Johnson, B., Doak, R., Allsup, D., Astwood, E., Evans, G., Grimley, C., James, B., Myers, B., Stokley, S., Thachil, J., Wilde, J., Williams, M., Makris, M., Lowe, G. C., Wallis, Y., Daly, M. E., Morgan, N. V., & The UK GAPP Study Group. (2018). A comprehensive targeted next-generation sequencing panel for genetic diagnosis of patients with suspected inherited thrombocytopenia. *Research and Practice in Thrombosis and Haemostasis*, *2*(4), 640–652. <https://doi.org/10.1002/rth2.12151>
- Johnson, B., Fletcher, S. J., & Morgan, N. V. (2016). Inherited thrombocytopenia: Novel insights into megakaryocyte maturation, proplatelet formation and platelet lifespan. *Platelets*, *27*(6), 519–525. <https://doi.org/10.3109/09537104.2016.1148806>
- Johnson, B., Lowe, G. C., Futterer, J., Lordkipanidze, M., MacDonald, D., Simpson, M. A., Sanchez-Guiu, I., Drake, S., Bem, D., Leo, V., Fletcher, S. J., Dawood, B., Rivera, J., Allsup, D., Biss, T., Bolton-Maggs, P. H., Collins, P., Curry, N., Grimley, C., ... UK GAPP Study Group. (2016). Whole exome sequencing identifies genetic variants in inherited thrombocytopenia with secondary qualitative function defects. *Haematologica*, *101*(10), 1170–1179. <https://doi.org/10.3324/haematol.2016.146316>
- Jones, M. L., Murden, S. L., Bem, D., Mundell, S. J., Gissen, P., Daly, M. E., Watson, S. P., & Mumford, A. D. (2012). Rapid genetic diagnosis of heritable platelet function disorders with next-generation sequencing: Proof-of-principle with Hermansky-Pudlak syndrome. *Journal of Thrombosis and Haemostasis*, *10*(2), 306–309.
- Khan, A. O., Maclachlan, A., Lowe, G. C., Nicolson, P. L. R., Ghaiithi, R. A., Thomas, S. G., Watson, S. P., Pike, J. A., & Morgan, N. V. (2020). High-throughput platelet spreading analysis: A tool for the diagnosis of platelet-based bleeding disorders. *Haematologica*, *105*(3), e124–e128.
- Lamolda, M., Montes, R., Simón, I., Perales, S., Martínez-Navajas, G., Lopez-Onieva, L., Ríos-Pelegrina, R., Del Moral, R. G., Griñan-Lison, C., & Marchal, J. A. (2019). GENYOi005-A: An induced pluripotent stem cells (iPSCs) line generated from a patient with Familial Platelet Disorder with associated Myeloid Malignancy (FPDMM) carrying a p. Thr196Ala variant. *Stem Cell Research*, *41*, 101603.
- Leinøe, E., Zetterberg, E., Kinalis, S., Østrup, O., Kampmann, P., Norström, E., Andersson, N., Klintman, J., Qvortrup, K., Nielsen, F. C., & Rossing, M. (2017). Application of whole-exome sequencing to direct the specific functional testing and diagnosis of rare inherited bleeding disorders in patients from the Oresund Region, Scandinavia. *British Journal of Haematology*, *179*(2), 308–322. <https://doi.org/10.1111/bjh.14863>
- Lentaigne, C., Freson, K., Laffan, M. A., Turro, E., & Ouwehand, W. H. (2016). Inherited platelet disorders: Toward DNA-based diagnosis. *Blood*, *127*(23), 2814–2823. <https://doi.org/10.1182/blood-2016-03-378588>
- Leo, V., Morgan, N., Bem, D., Jones, M., Lowe, G., Lordkipanidzé, M., Drake, S., Simpson, M., Gissen, P., & Mumford, A. (2015). Use of next-generation sequencing and candidate gene analysis to identify underlying defects in patients with inherited platelet function disorders. *Journal of Thrombosis and Haemostasis*, *13*, 643–650.
- Lester, W. A., Guillaud, A. M., Enayat, M. S., Rose, P., & Hill, F. G. (2007). The R2464C missense mutation in the von Willebrand factor gene causes a novel abnormality of multimer electrophoretic mobility and falls into the subgroup of type 2 von Willebrand disease 'unclassified'. *Thrombosis and Haemostasis*, *97*, 159–160.
- Mundell, S. J., & Mumford, A. (2018). TBXA2R gene variants associated with bleeding. *Platelets*, *29*(7), 739–742.
- Nowakowska, B. (2017). Clinical interpretation of copy number variants in the human genome. *Journal of Applied Genetics*, *58*(4), 449–457. <https://doi.org/10.1007/s13353-017-0407-4>
- Nurden, A., Freson, K., & Seligsohn, U. (2012). Inherited platelet disorders. *Haemophilia*, *18*, 154–160.
- Peyvandi, F., Jayandharan, G., Chandu, M., Srivastava, A., Nakaya, S. M., Johnson, M. J., Thompson, A. R., Goodeve, A., Garagiola, I., Lavoretano, S., Menegatti, M., Palla, R., Spreafico, M., Tagliabue, L., Asselta, R., Duga, S., & Mannucci, P. M. (2006). Genetic diagnosis of haemophilia and other inherited bleeding disorders. *Haemophilia*, *12*, 82–89.
- Rabbolini, D., Morel-Kopp, M. C., Chen, Q., Gabrielli, S., Dunlop, L., Chew, L., Blair, N., Brighton, T., Singh, N., & Ng, A. (2017). Thrombocytopenia and CD 34 expression is decoupled from α -granule deficiency with mutation of the first growth factor-independent 1B zinc finger. *Journal of Thrombosis and Haemostasis*, *15*, 2245–2258.
- Richards, S., Aziz, N., Bale, S., Bick, D., Das, S., Gastier-Foster, J., Grody, W. W., Hegde, M., Lyon, E., Spector, E., Voelkerding, K., & Rehms, H. L. (2015). Standards and guidelines for the interpretation of sequence variants: A joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genetics in Medicine*, *17*(5), 405–424. <https://doi.org/10.1038/gim.2015.30>
- Savoia, A., & Pecci, A. (2015). *MYH9-related disorders*. *GeneReviews® [Internet]*. Seattle: University of Washington.
- Simeoni, I., Stephens, J. C., Hu, F., Deevi, S. V. V., Megy, K., Bariana, T. K., Lentaigne, C., Schulman, S., Sivapalaratnam, S., Vries, M. J. A., Westbury, S. K., Greene, D., Papadia, S., Alessi, M. C., Attwood, A. P., Ballmaier, M., Baynam, G., Bermejo, E., Bertoli, M., ... Turro, E. (2016). A high-throughput sequencing test for diagnosing inherited bleeding, thrombotic, and platelet disorders. *Blood*, *127*(23), 2791–2803. <https://doi.org/10.1182/blood-2015-12-688267>
- Sivapalaratnam, S., Collins, J., & Gomez, K. (2017). Diagnosis of inherited bleeding disorders in the genomic era. *British Journal of Haematology*, *179*(3), 363–376.
- Stevenson, W. S., Rabbolini, D. J., Beutler, L., Chen, Q., Gabrielli, S., Mackay, J. P., Brighton, T. A., Ward, C. M., & Morel-Kopp, M. C. (2015). Paris-Trousseau thrombocytopenia is phenocopied by the autosomal recessive inheritance of a DNA-binding domain mutation in FLI1. *Blood*, *126*(17), 2027–2030.
- Stockley, J., Morgan, N. V., Bem, D., Lowe, G. C., Lordkipanidzé, M., Dawood, B., Simpson, M. A., Macfarlane, K., Horner, K., & Leo, V. C. (2013). Enrichment of FLI1 and RUNX1 mutations in families with excessive bleeding and platelet dense granule secretion defects. *Blood*, *122*, 4090–4093.

- Valsesia, A., Stevenson, B. J., Waterworth, D., Mooser, V., Vollenweider, P., Waeber, G., Jongeneel, C., Beckmann, J. S., Kutalik, Z., & Bergmann, S. (2012). Identification and validation of copy number variants using SNP genotyping arrays from a large clinical cohort. *BMC Genomics*, *13*(1), 241. <https://doi.org/10.1186/1471-2164-13-241>
- Watson, S., Lowe, G., Lordkipanidze, M., Morgan, N., & GAPP Consortium. (2013). Genotyping and phenotyping of platelet function disorders. *Journal of Thrombosis and Haemostasis*, *11*, 351–363.
- Zhang, Y., Wang, F., Chen, X., Zhang, Y., Wang, M., Liu, H., Cao, P., Ma, X., Wang, T., & Zhang, J. (2018). CSF3R mutations are frequently associated with abnormalities of RUNX1, CBFβ, CEBPA, and NPM1 genes in acute myeloid leukemia. *Cancer*, *124*, 3329–3338.

SUPPORTING INFORMATION

Additional Supporting Information may be found online in the supporting information tab for this article.

How to cite this article: Almazni I, Stapley RJ, Khan AO, Morgan NV. A comprehensive bioinformatic analysis of 126 patients with an inherited platelet disorder to identify both sequence and copy number genetic variants. *Human Mutation*. 2020;41:1848–1865. <https://doi.org/10.1002/humu.24114>