UNIVERSITY OF BIRMINGHAM
University of Birmingham Research at Birmingham

Novel gene variants in patients with platelet-based bleeding using combined exome sequencing and RNAseq murine expression data

Khan, Abdullah; Stapley, Rachel; Pike, Jeremy; Wijesinghe, Susanne; Reyat, Jasmeet; Almazni, Ibrahim Abdullah F; Machlus, Kellie; Morgan, Neil

DOI:
10.1111/jth.15119

License:
Creative Commons: Attribution (CC BY)

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

Citation for published version (Harvard):
Khan, A, Stapley, R, Pike, J, Wijesinghe, S, Reyat, J, Almazni, IAF, Machlus, K & Morgan, N 2020, 'Novel gene variants in patients with platelet-based bleeding using combined exome sequencing and RNAseq murine expression data: Novel genes in patients with platelet-based bleeding', *Journal of Thrombosis and Haemostasis*. https://doi.org/10.1111/jth.15119

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.

Download date: 28. Feb. 2021

Received: 5 June 2020 | Accepted: 28 September 2020

DOI: 10.1111/jth.15119

BRIEF REPORT

Novel gene variants in patients with platelet-based bleeding using combined exome sequencing and RNAseq murine expression data

Abdullah O. Khan[1] | Rachel J. Stapley[1] | Jeremy A. Pike[1] | Susanne.N. Wijesinghe[2] | Jasmeet S. Reyat[1] | Ibrahim Almazni[1] | Kellie R. Machlus[3] | Neil V. Morgan[1] | On behalf of the UK GAPP Study Group

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

[2]Institute of Inflammation and Ageing, College of Medical and Dental Sciences, University of Birmingham, Birmingham, UK

[3]Hematology Division, Department of Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, MA, USA

Correspondence
Abdullah O. Khan and Neil V. Morgan, Institute of Cardiovascular Sciences, College of Medical and Dental Sciences, Edgbaston, University of Birmingham, Birmingham, B15 2TT, UK.
Emails: a.khan.4@bham.ac.uk; N.V.Morgan@bham.ac.uk

Funding information
British Heart Foundation, Grant/Award Number: FS/15/18/31317, FS/18/11/33443, PG/13/36/30275 and PG/16/103/32650; Wellcome Trust, Grant/Award Number: 218649/Z/19/Z; Saudi Arabia Cultural Bureau in London; National Institute of Diabetes and Digestive and Kidney Diseases, Grant/Award Number: K01DK111515

Abstract

Background: The UK Genotyping and Phenotyping of Platelets study has recruited and analyzed 129 patients with suspected heritable bleeding. Previously, 55 individuals had a definitive genetic diagnosis based on whole exome sequencing (WES) and platelet morphological and functional testing. A significant challenge in this field is defining filtering criteria to identify the most likely candidate mutations for diagnosis and further study.

Objective: Identify candidate gene mutations for the remaining 74 patients with platelet-based bleeding with unknown genetic cause, forming the basis of future re-recruitment and further functional testing and assessment.

Methods: Using python-based data frame indexing, we first identify and filter all novel and rare variants using a panel of 116 genes known to cause bleeding across the full cohort of WES data. This identified new variants not previously reported in this cohort. We then index the remaining patients, with rare or novel variants in known bleeding genes against a murine RNA sequencing dataset that models pro-platelet-forming megakaryocytes.

Results: Filtering against known genes identified candidate variants in 59 individuals, including novel variants in several known genes. In the remaining cohort of "unknown" patients, indexing against differentially expressed genes revealed candidate gene variants in several novel unreported genes, focusing on 14 patients with a severe clinical presentation.

Conclusions: We identified candidate mutations in a cohort of patients with no previous genetic diagnosis. This work involves innovative coupling of RNA sequencing and

Manuscript handled by: Matthew T. Rondina.

Final decision: Matthew T. Rondina , 28 September 2020.

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. Journal of Thrombosis and Haemostasis published by Wiley Periodicals LLC on behalf of International Society on Thrombosis and Haemostasis

WES to identify candidate variants forming the basis of future study in a significant number of undiagnosed patients.

KEYWORDS
bleeding, platelets, megakaryocytes, genetics, RNAseq

1 | INTRODUCTION

The UK Genotyping and Phenotyping of Platelets (GAPP)[1] study previously reported a cohort of 55 patients with whole exome sequencing (WES) recruited from hemophilia centers nationwide with a significant clinical history of bleeding.[2] In this group, candidate variants and a genetic diagnosis of platelet-based bleeding was provided in 40 patients, yielding a detection rate of 72.72% candidate variants in this initial cohort using the "manual" filtering method of Johnson et al[2] (Figure 1A).

More recently, the GAPP study has expanded patients with available WES data to include 129 individuals with evidence of clinical bleeding. Of these, 55 individuals have had a definitive genetic diagnosis to identify novel disease-causing genes (eg, *SLFN14*)[3] or new pathogenic variants in known genes (eg, *GP1BA/B*, *WAS*) yielding an overall detection rate of 42.63% (Figure 1B).[4,5] Through an extensive panel of platelet function testing[6] and morphological assessment,[7] these genetic findings have been correlated to the physiological and clinical effects of these rare genetic variants. However, the genetic causes of bleeding in a significant number of GAPP-recruited patients remain unknown, despite an established clinical and family history. Here, we assess these individuals through a series of iterative bioinformatic filtering approaches to identify novel and "known" candidate disease-causing variants.

Essentials

- Identifying genetic variants in platelet disorders is challenging due to its heterogenous nature.
- We combine WES, RNAseq, and python-based bioinformatics to identify novel gene variants.
- We find novel candidates in patient data by cross-referencing against a murine RNAseq model of thrombopoiesis.
- This innovative combined bioinformatic approach provides novel data for future research in the field.

2 | METHODS

2.1 | Patient recruitment and testing

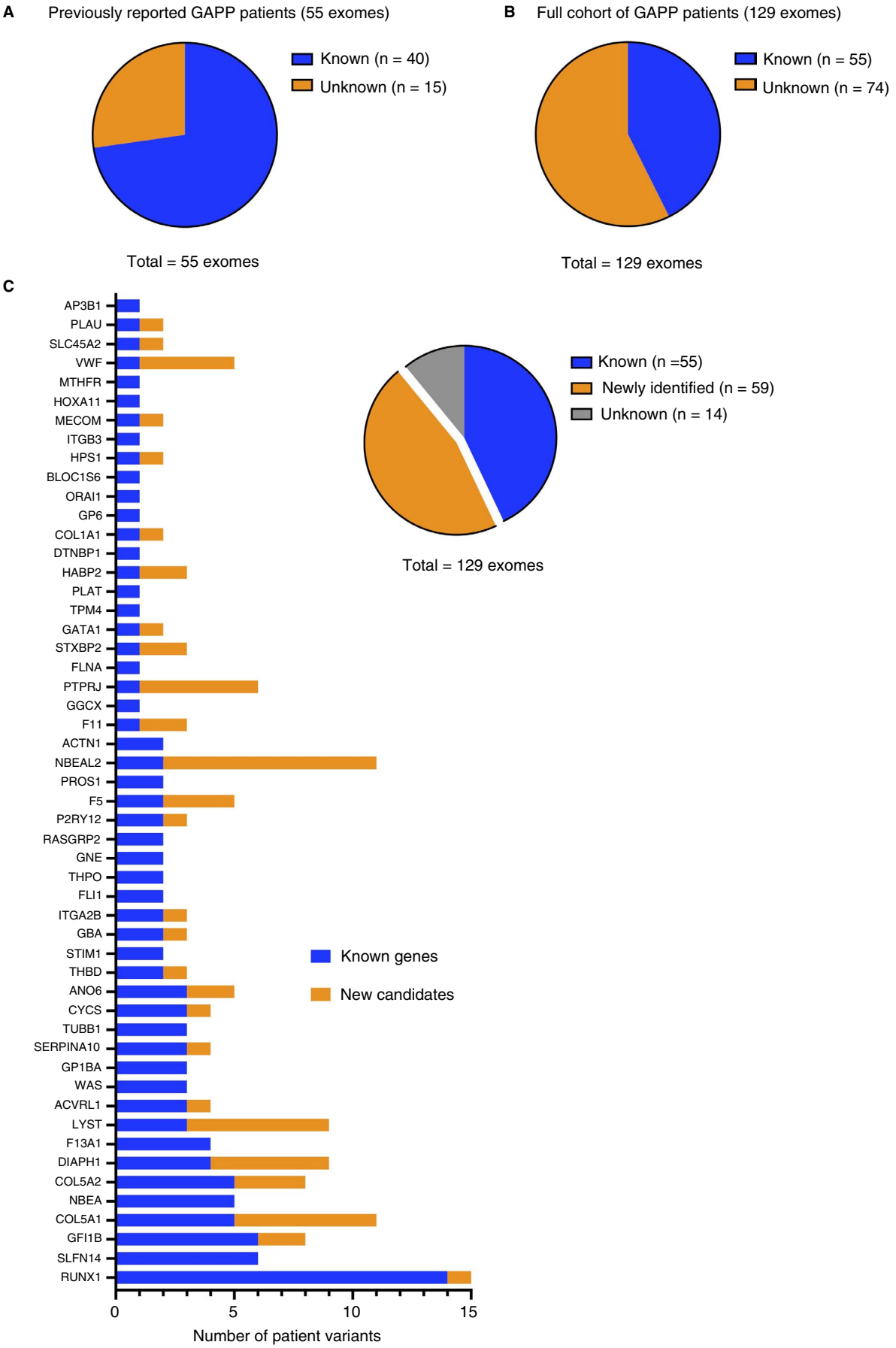
Patients were consented and recruited to the GAPP study from multiple collaborating hemophilia centers across the United Kingdom and Ireland, as previously described,[1] and approved by the UK National Research Ethics Service by the Research Ethics Committee of West Midlands (06/MRE07/36). Peripheral blood was collected from patients and platelet phenotyping using lumi-aggregometry or flow cytometry was performed on platelet-rich plasma, as previously described.[6,7]

Our study cohort consisted of 129 patients with a strong history of bleeding and suspected of having a platelet function disorder of unknown cause as previously described.[1,2] The ISTH Bleeding Assessment Tool (BAT) showed a mean score of 9.825 overall (range 2-23). Platelet counts for the patient group ranged from 43 to 428 $\times$ 10^9/L and the average platelet count was 232 $\times$ 10/L. The mean platelet volume in patients tested ranged between 8.3 fL and 15.1 fL.

2.2 | RNA Sequencing

RNA sequencing (RNAseq) data was provided by KRM and generated as reported in Machlus et al.[8] Briefly, two independent isolations of mouse megakaryocytes (MK; at the round MK, proplatelet, and preplatelet/releasate stages) were performed (C57b1/6, 1 male and 1 female pool, 4-8 mice). Sequencing and analysis were performed as described by Rowley et al,[9] using the Useq analysis package (applying DESeq's negative binomial test).[10,11] For analysis, a P value $< .05$ was used and a false discovery rate of 5%, to yield a total of 7094 (3235 up-regulated and 3859 down-regulated) and 8845 (4125 up-regulated and 4720 down-regulated) differentially expressed genes in round versus proplatelet forming MKs and proplatelet MK versus preplatelet releasate respectively.

FIGURE 1 Distribution of known and unknown GAPP patients and the identification of new variants in bleeding genes. (A) In our previous report of the GAPP cohort in Johnson et al, a 55 exomes were reported with a total of 40 patients identified as having a clear genetic candidate for their inherited platelet-based bleeding. (B) Because the cohort expanded to include 129 exomes, only 55 patients have an established genetic cause of bleeding. To date, the cohort with unresolved causes remains at 74 patients. (C) By cross-indexing rare and novel variants in GAPP exomes with known bleeding related genes, we identified new candidate genes in a further 59 patients, reducing the size of the unknown cohort to 14


A
Previously reported GAPP patients (55 exomes)
Known (n = 40)
Unknown (n = 15)
Total = 55 exomes
B
Full cohort of GAPP patients (129 exomes)
Known (n = 55)
Unknown (n = 74)
Total = 129 exomes
C
Known (n =55)
Newly identified (n = 59)
Unknown (n = 14)
Total = 129 exomes
AP3B1
PLAU
SLC45A2
VWF
MTHFR
HOXA11
MECOM
ITGB3
HPS1
BLOC1S6
ORAI1
GP6
COL1A1
DTNBP1
HABP2
PLAT
TPM4
GATA1
STXBP2
FLNA
PTPRJ
GGCX
F11
ACTN1
NBEAL2
PROS1
F5
P2RY12
RASGRP2
GNE
THPO
FLI1
ITGA2B
GBA
STIM1
THBD
ANO6
CYCS
TUBB1
SERPINA10
GP1BA
WAS
ACVRL1
LYST
F13A1
DIAPH1
COL5A2
NBEA
COL5A1
GFI1B
SLFN14
RUNX1
Known genes
New candidates
0
5
10
15
Number of patient variants

Principal component analysis plots of the samples are included as supplementary Figure 3.

2.3 | Analysis approaches

WES analysis was performed in patient genomic DNA, as previously reported.[2] To improve filtering of candidate genetic variants, further bioinformatic analysis was performed using python in the Pycharm IDE. Exomes were first uploaded as pandas data frames, before filtration to identify rare (as in previous work this is classified as below a frequency of 0.0001 in exome variant server and the 1000 genomes) and novel nonsynonymous variants as described by Johnson et al. These were cross-referenced against a known list of genes that have been previously implicated in inherited causes of bleeding. Relevant gene panels used were Inherited Bleeding Disorder, high evidence (green), medium evidence (amber), and low evidence (red) gene lists, containing 116 genes (supplementary Table S1) derived from the Panel app (https://panelapp.genomicsengland.co.uk/) and 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 suitable for clinical use at this time.

The resulting variants were collated and graphed using PRISM GraphPad 8. DESeq2 output data was collated as pandas data frames depending on the samples analyzed, and visualized using matplotlib and seaborn. Unsupervised hierarchical clustering of RNAseq data was performed to stratify groups of variants according to their level of expression (up-regulated/down-regulated) over the course of thrombopoiesis. The annotated scripts are available in GitHub for public use (https://github.com/aokhan/gapp-python-processing) and further details are available on request.

3 | RESULTS AND DISCUSSION

Of the 129 patients recruited in this study and assessed by the GAPP study, 55 have been reported in previous work and are now classified as having known causes of bleeding (Figure 1B). The unknown cohort comprises 74 patients with extensive clinical bleeding histories, platelet function testing, and WES analysis (Figure 1C). Therefore, to identify new candidate gene variants, we developed a python-based indexing platform that first filtered the 129 patient exomes for rare and novel nonsynonymous and indel variants, before cross-referencing the resulting data frame with a list of 116 genes already established as clinically causative of bleeding (detailed in Methods and supplementary Figure 1). We reasoned that an automated approach that uses an up-to-date list of clinically relevant platelet-based bleeding genes would better identify candidate mutations in the full GAPP cohort. When we compared the detection rate using our initial methodology in the cohort of 55 exomes/patients reported by Jonson et al to our python workflow, we identified the original 40 candidate genes, as well as an additional 8 candidates through cross-indexing with the list of clinical bleeding genes. We observe an increase from 72.3% (40 of 55 exomes) to 87.3% (48 of 55).

Expanding this analysis to the full cohort of 129 exomes, rare and novel genetic variants in a further 59 individuals, previously part of the "unexplained platelet disorder with bleeding" cohort, were identified (Figure 1C). These include new and unreported variants in established genes (*NBEAL2*, *TUBB1*, *RUNX1*, *DIAPH1*, *VWF*),[12] as well as candidates in lesser-known genes more recently associated with bleeding (*PTPRJ*, *MECOM*, *TPM4*).[13-15]

Our automated indexing of rare and novel variants has effectively highlighted a number of new candidates in known gene categories (which we have labeled "newly identified"), but there still remained a cohort of 14 patients with no rare variants in any known disease-causing genes. Given that this remaining cohort of patients is a challenging one to dissect, we developed a filtering approach based on an expression data resource generated from maturing and proplatelet forming murine MKs published by Machlus et al.[8]

RNAseq was performed on populations of murine mature round MKs, proplatelet-forming MKs, and released proplatelets to generate a palette of gene expression data representative of thrombopoiesis using an established in vitro model (Figure 2A).[8] The data were analyzed and processed by DESeq2 and filtered to omit low read counts before plotting the resulting 13,062 genes that were significantly differentially expressed across the maturation of MKs into the process of platelet production (Figure 2B, 2C). These genes were plotted as log2 fold change in expression in the round MK versus proplatelet-forming MK population (y-axis, Figure 2C) and the proplatelet-forming MK versus released proplatelets (x-axis, Figure 2C). Therefore, these data represent an array of genes, which are up- and down-regulated during and throughout thrombopoiesis.

We next filtered this expression data set against the array of known disease-causing genes to validate this approach and found that the 116 candidates in this list are evident as a population of genes that are differentially expressed in this sample, with base mean read sizes visualized by point size and color (Figure 2D). Of these, 25 genes were downregulated during thrombopoiesis, whereas the remaining 91 were upregulated. An annotated crop of this plot in Figure 2D shows established genes well represented by the data (*GATA1*, *THPO*, *TUBB1*, *MYH9*, *RUNX1*). We then cross-referenced our rare GAPP-study patient variants with genes represented in these plots to generate a list of 7034 candidate patient variants that were present in this thrombopoiesis data. We then performed an unsupervised hierarchical clustering of these genes to stratify groups of variants according to their level of expression (up-regulated/down-regulated) over the course of thrombopoiesis. Finally, we restricted this analysis to our remaining cohort of 14 patients with unidentified genetic causes of bleeding (supplementary Figure 2), and upon filtration and clustering, identified 360 novel variants in genes differentially upregulated during thrombopoiesis. These genes have not been previously associated with bleeding in the literature and therefore represent an opportunity to identify

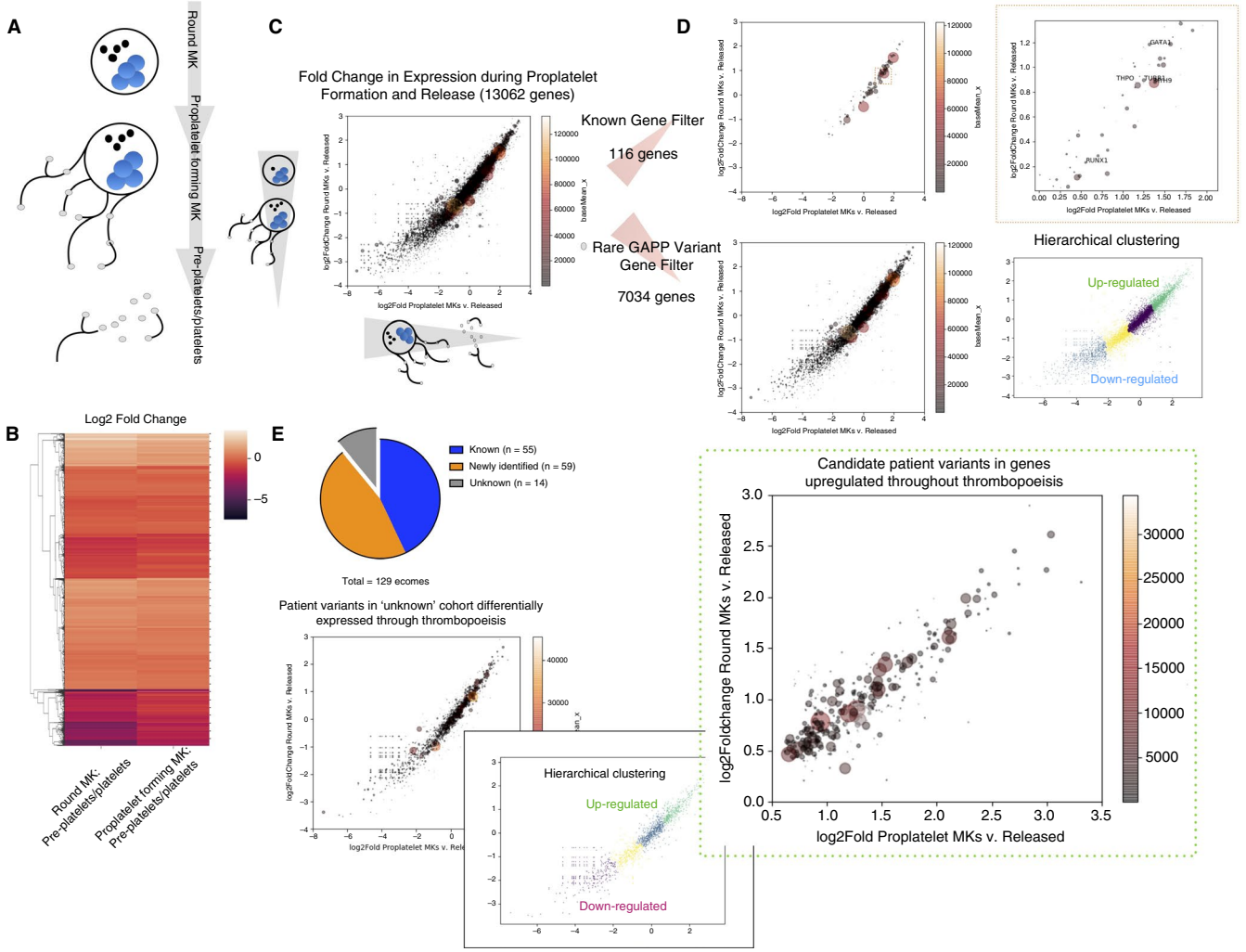


FIGURE 2 Cross-indexing RNAseq data of maturing and proplatelet producing MKs yields novel candidates in an unknown cohort of patients. (A) An RNAseq dataset derived from mature, round MKs, proplatelet-forming MKs, and pre-/platelet releasate was obtained (n = 2). (B) A heatmap of differential expression and (C) a plot of fold change across round MK: proplatelet MKs and proplatelet MKs: released proplatelets showed that this data set effectively models gene expression during thrombopoiesis. (D) Filtering these data against known disease-causing genes and GAPP variants identified up-regulated genes and candidate variants, which include a number of key known genes. These data can be further clustered to segment populations of differentially up-regulated and down-regulated genes. (E) Cross-referencing rare and novel nonsynonymous variants from the GAPP cohort of 14 unknown patients allowed for the segmentation of candidate novel variants which were significantly up-regulated

novel players in thrombopoiesis that may be the cause of bleeding in this patient cohort.

Among these novel candidates, we found a novel variant in *DNAI1*[16] in a patient with a documented platelet secretion defect (Table 1). Recent work from our group investigating the mechanisms by which the polymodification of *TUBB1* mediates platelet production reported *DNAI1* as a motor protein, which is carefully spatially regulated by the posttranslational modification of *TUBB1*.[17] Mutations in *DNAI1* have previously been documented to cause primary ciliary dyskinesia in a very small number of cases with no evidence of a platelet disorder documented.[18] Similarly, in a patient with a severe thrombocytopenia (platelet count of 60×10^9/L) and severe bleeding (BAT score 13) we found a frameshift insertion in *MTMR11* (encoding human myotubularin 11), a protein that remains unstudied but is abundant in MKs in human bone marrow sections.[19] This variant has not been reported in a patient or population previously (gnomAD accessed August 2020) and with a significant predicted pathogenicity score of 37 (combined annotation dependent depletion). In another unrelated individual with a severe macrothrombocytopenia (platelet count of 43×10^9/L, size of 15 fL), a stop/gain mutation in an alpha-granule protein, *MEGF11*,[20] was found which was again absent in population databases (gnomAD accessed August 2020) (Table 1, supplementary Figure 2).

In summary, one of the largest challenges in the study of inherited bleeding lies in interrogating a significant cohort of patients who, despite extensive study, have yet to yield a conclusive genetic cause of disease. This is not only true of the GAPP cohort, but of other studies focusing on such rare heritable disease.[12,21] Therefore, we have created an iterative filtration process to highlight new candidate variants in a population of patients that, to date, had no identified

TABLE 1 Summary of novel candidates in GAPP patients with unexplained platelet related bleeding

Sample ID	Platelet Defect	Platelet Count ($\times 10^9$/L)	Platelet Size (fL)[4]	BAT Score	Candidate Gene	Mutation	Variant	Consequence	Frequency (Gnomad)
1010	Secretion	N/A	N/A	8	*DNAI1*	Missense	c.738G > C	p.E246D	Novel
					TRIM17	Missense	c.1289C > T	p.S430F	Novel
2521	Thrombocytopenia	53	11.7	5	*COL7A1*	Frameshift	c.7786delG	p.G2596fs	Novel
					MAP7	Missense	c.1822G > C	p.E608Q	Novel
1859	Thrombocytopenia	9.3	9.5	10	*RAB36*	Splicing	c.426-4A > G	p.?	Novel
					MYH7B	Splicing	c.4316-2T > C	p.?	Novel
1255	Thrombocytopenia	60	8.6	13	*MTMR11*	Frameshift ins.	c.8_9insG	p.W3fs	Novel
1248	Thrombocytopenia	43	14	5	*MEGF11*	Stop/gain	c.438C > A	p.C146X	Novel

Note: Individuals with established platelet defects (secretion and thrombocytopenia) and significant BAT scores are reported with newly uncovered candidate mutations in genes expressed in megakaryocytes, but previously unstudied. Frequencies reported according to the latest data on gnomAD (https://gnomad.broadinstitute.org/)

cause of bleeding. This approach is open source, automated, and easily adaptable to apply filtration to any list of genes of interest.

We use a murine MK model of thrombopoiesis as a model RNAseq data set against which we filter our unknown cohort of patient genomic data. To our knowledge, this is the first such attempt to interrogate a clinical cohort of patients with bleeding using a combined WES and RNAseq bioinformatics approach. Although murine models of thrombopoiesis are well established and extensively used in the study of platelet-based bleeding because of the high degree of homology between the species, restricting our analysis to this data set is a likely limitation of the study. Future work will focus on applying this approach in parallel to human models of thrombopoiesis, including CD34- and iPSC-derived MKs.

By applying our methods, we generated a large pool of candidates that will become the focus of future studies as we attempt to dissect a well-studied but poorly understood cohort of patients. Further study is needed to definitively prove the involvement of candidate genes and as such, future work will focus on the re-recruitment of patients for mechanistic investigation.

ACKNOWLEDGMENTS

A.O.K. acknowledges Arthur Ruski for his help in troubleshooting python.

The work in the authors' laboratories is supported by the British Heart Foundation (PG/13/36/30275; FS/15/18/31317; PG/16/103/32650; FS/18/11/33443) (N.V.M.) and the Saudi Arabia Cultural Bureau in London (I.A.). A.O.K. is a Wellcome-funded Sir Henry Wellcome Fellow (218649/Z/19/Z). K.R.M. is supported by the National Institute of Diabetes and Digestive and Kidney Diseases (K01DK111515) at the National Institutes of Health, and is an American Society of Hematology Scholar.

CONFLICT OF INTEREST

The authors report no conflicts of interest.

AUTHOR CONTRIBUTIONS

A.O.K., R.S., and N.V.M. designed the research. A.O.K. performed the analysis and generated the relevant scripts. K.R.M. performed experiments and provided RNAseq data. All authors revised the manuscript and curated the data.

ORCID

Abdullah O. Khan https://orcid.org/0000-0003-0825-3179
Kellie R. Machlus https://orcid.org/0000-0002-2155-1050
Neil V. Morgan https://orcid.org/0000-0001-6433-5692

TWITTER

Neil V. Morgan @neilvmorgan

REFERENCES

1. Watson SP, Lowe GC, Lordkipanidze M, Morgan NV. Genotyping and phenotyping of platelet function disorders. *J Thromb Haemost*. 2013;11(suppl 1):351-363.

2. Johnson B, Lowe GC, Futterer J, et al. Whole exome sequencing identifies genetic variants in inherited thrombocytopenia with secondary qualitative function defects. *Haematologica*. 2016;101:1170-1179.
3. Fletcher SJ, Johnson B, Lowe GC, et al. SLFN14 mutations underlie thrombocytopenia with excessive bleeding and platelet secretion defects. *J Clin Invest*. 2015;125:3600-3605.
4. Othman M, Emsley J. Gene of the issue: GP1BA gene mutations associated with bleeding. *Platelets*. 2017;28:832-836.
5. Albert MH, Notarangelo LD, Ochs HD. Clinical spectrum, pathophysiology and treatment of the Wiskott-Aldrich syndrome. *Curr Opin Hematol*. 2011;18(1):42-48.
6. Dawood BB, Lowe GC, Lordkipanidze M, et al. Evaluation of participants with suspected heritable platelet function disorders including recommendation and validation of a streamlined agonist panel. *Blood*. 2012;120:5041-5049.
7. Khan AO, Maclachlan A, Lowe GC, et al. High-throughput platelet spreading analysis: a tool for the diagnosis of platelet-based bleeding disorders. *Haematologica*. 2020;105(3):e124-e128.
8. Machlus KR, Wu SK, Stumpo DJ, et al. Synthesis and dephosphorylation of MARCKS in the late stages of megakaryocyte maturation drive proplatelet formation. *Blood*. 2016;127:1468-1480.
9. Rowley JW, Oler AJ, Tolley ND, et al. Genome-wide RNA-seq analysis of human and mouse platelet transcriptomes. *Blood*. 2011;118(14):e101-e111.
10. Nix DA, Courdy SJ, Boucher KM. Empirical methods for controlling false positives and estimating confidence in ChIP-Seq peaks. *BMC Bioinform*. 2008;9:523.
11. Anders S, Huber W. Differential expression analysis for sequence count data. *Genome Biol*. 2010;11:R106.
12. Downes K, Megy K, Duarte D, et al. Diagnostic high-throughput sequencing of 2396 patients with bleeding, thrombotic, and platelet disorders. *Blood*. 2019;134:2082-2091.
13. Pleines I, Woods J, Chappaz S, et al. Mutations in tropomyosin 4 underlie a rare form of human macrothrombocytopenia. *J Clin Invest*. 2017;127:814-829.
14. Germeshausen M, Ancliff P, Estrada J, et al. MECOM-associated syndrome: a heterogeneous inherited bone marrow failure syndrome with amegakaryocytic thrombocytopenia. *Blood Adv*. 2018;2:586-596.
15. Marconi C, Di Buduo CA, LeVine K, et al. Loss-of-function mutations in PTPRJ cause a new form of inherited thrombocytopenia. *Blood*. 2019;133:1346-1357.
16. Andjelkovic M, Minic P, Vreca M, et al. Genomic profiling supports the diagnosis of primary ciliary dyskinesia and reveals novel candidate genes and genetic variants. *PLoS One*. 2018;13:e0205422.
17. Khan AO, Slater A, Maclachlan A, et al. Post-translational polymodification of β1 tubulin regulates motor protein localisation in platelet production and function. *bioRxiv*. 2020.
18. Failly M, Saitta A, Muñoz A, et al. DNAI1 mutations explain only 2% of primary ciliary dykinesia. *Respiration*. 2008;76:198-204.
19. Kabir NN, Rönnstrand L, Kazi JU. Deregulation of protein phosphatase expression in acute myeloid leukemia. *Med Oncol*. 2013;30:517.
20. Krivtsov AV, Rozov FN, Zinovyeva MV, et al. Jedi—a novel transmembrane protein expressed in early hematopoietic cells. *J Cell Biochem*. 2007;101:767-784.
21. Bastida JM, Lozano ML, Benito R, et al. Introducing high-throughput sequencing into mainstream genetic diagnosis practice in inherited platelet disorders. *Haematologica*. 2018;103:148.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.

How to cite this article: Khan AO, Stapley RJ, Pike JA, et al; the UK GAPP Study Group. Novel gene variants in patients with platelet-based bleeding using combined exome sequencing and RNAseq murine expression data. *J Thromb Haemost*. 2020;00:1–7. https://doi.org/10.1111/jth.15119