Westbury, S. K., & Mumford, A. D. (2016). Genomics of platelet disorders. *Haemophilia*, *22*(Suppl S5), 20-24. https://doi.org/10.1111/hae.12964

Genomics of platelet disorders

Sarah K. Westbury[1], Andrew D. Mumford[1,2,3]

[1]. School of Clinical Sciences, University of Bristol, Bristol, UK.

[2]. Bristol Haemophilia Comprehensive Care Centre, Bristol, UK

[3]. West of England Genomic Medicine Centre, Bristol, UK

Key words: platelet number disorders; platelet function disorders; genetic diagnosis; next generation sequencing.

Word count: 2,852; Two tables.

Corresponding author: Dr Sarah Westbury, School of Clinical Sciences, University of Bristol, Level 7 Bristol Royal Infirmary, Bristol, BS2 8HW, UK. +44 117 3423152; a.mumford@bristol.ac.uk

SUMMARY

Genetic diagnosis in families with inherited platelet disorders (IPD) is not performed widely because of the genetic heterogeneity of this group of disorders and because in most cases, it is not possible to select single candidate genes for analysis using clinical and laboratory phenotypes. Next generation sequencing (NGS) technology has revolutionised the scale and cost-effectiveness of genetic testing, and has emerged as a valuable tool for IPD. This review examines the potential utility of NGS as a diagnostic tool to streamline detection of causal variants in known IPD genes and as a vehicle for new gene discovery.

MAIN TEXT

Genetic diagnosis for families with inherited bleeding disorders may enhance clinical care by improving the precision of diagnosis, by enabling better prediction of prognosis and heritability and by informing selection of targeted therapies. Although widely available for coagulation factor disorders such as haemophilia, genetic testing for inherited platelet disorders (IPD) has until recently, been restricted to a small number of specific disorders.

Scope of inherited platelet disorders

One important difference between IPD and coagulation factor disorders is that the IPD comprise a highly heterogeneous group of different disorders of platelet number or function. Recent advances in gene discovery have now defined more than 50 distinct IPD (**tables 1 and 2**). Most of these disorders have been reported in small numbers of families worldwide, suggesting that they are individually rare. However, IPD collectively account for nearly 10% of all bleeding disorders in some national

registries [1], and may be substantially under-reported. Despite the prevalence of IPD, diagnosis at genetic level is achieved in a small minority of families and is usually restricted to centres where there is a clinical or research interest in platelet disorders [2]. For many families with IPD, absence of genetic diagnosis reflects the practical constraint that traditional genetic testing by analysis of single candidate genes cannot be applied to most IPD in which multiple genes are potentially implicated. Causal variants may lies in non-coding regions implicated in the regulation of known IPD candidate genes but which have not been historically examined in clinical diagnostic genetic testing, or may occur in more than one IPD gene giving rise to complex inheritance patterns and practical difficulties in resolving phenotypes. In other IPD families, failure to achieve genetic diagnosis is very likely to reflect that they harbour causal variants in novel genes, not yet associated with platelet disorders.

Inherited thrombocytopenia as a model for improving genetic diagnosis

Amongst the IPD, the potential for genetic diagnosis is best illustrated by the inherited thrombocytopenias. In this group of disorders, molecular testing has historically focussed on the recessive forms of Bernard Soulier syndromes (BSS), caused by variants in *GP1BA, GP1BB or GP9,* which may be associated with a severe bleeding tendency. However, recessive BSS accounts for only approximately 5% of inherited thrombocytopenia cases in large series [3]. More recently, good progress has been made by several large international consortia on unravelling genetic defects associated with inherited thrombocytopenias with less severe bleeding and which are often discovered as an incidental finding during routine haematology testing. With these advances, comprehensive genetic testing has now

been reported to enable molecular diagnosis in up to 50% of patients with the inherited thrombocytopenia phenotype [4]. These analyses have demonstrated that the majority of families with inherited thrombocytopenia for which genetic diagnosis is currently possible, have causal variants in *MYH9*, *ANKRD26*, *ACTN1* and heterozygous forms of BSS. The reminder of inherited thrombocytopenia families have rare disorders described in small numbers of pedigrees worldwide, or have unknown disorders [4]

The inherited thrombocytopenias also illustrate how genetic diagnosis can directly benefit affected families. Although in many cases, the haemostatic defect in inherited thrombocytopenias is small or even absent, this is not necessarily the case for other phenotypic manifestations for which genetic diagnosis can have important prognostic implications. This is best illustrated by inherited thrombocytopenias caused by some variants in *RUNX1* or *ANKR26*, which are associated with an increased risk of myeloid malignancy [5] or by some *ETV6* variants that may be associated with lymphoid malignancy [6]. Similarly, demonstration of some *MYH9* variants in families with *MYH9*-related disorder may be predictive of adverse phenotypes such as hearing loss, cataract formation and rapidly progressive nephropathy that may not be clinically evident in very young affected patients [7], and which may require specific interventions. Detection of causal variants for some inherited thrombocytopenias may also enable specific treatments such as thrombopoietin mimetics to increase platelet counts in settings such as before surgery. These agents have proven efficacy in disorders such as *MYH9*-related disorder [8] and Wiscott Aldrich syndrome [9], but are of uncertain or doubtful benefit in other inherited thrombocytopenias.

Approaches to gene sequencing in inherited platelet disorders

Sanger sequencing remains the gold standard for genetic testing in human disease, but requires amplification of discrete regions of candidate genes by polymerase chain reaction and then sequencing by chain termination techniques. Since this is a time-consuming and low-throughput approach, Sanger sequencing is practically limited to the small number of IPD in which candidate genes can be selected by laboratory phenotype testing (eg. Glanzmann thrombasthenia (*ITGA2B* or *ITGB3*) or because there are additional syndromic feature that are suggestive of a specific IPD (eg. MYH9- related disorder (*MYH9*) and Chediak Higashi syndrome (*LYST*)).

Next generation DNA sequencing (NGS) technology is a significant advance from Sanger sequencing that has revolutionised gene discovery and is now emerging as a clinical diagnostic tool for IPD. This technique enables parallel sequencing of large regions of DNA spanning many genes, is rapid and readily automated and is an increasingly cost effective method of achieving a genetic diagnosis.

One important application of NGS as a diagnostic tool for IPD and other bleeding disorders has been the development of NGS panels that enable analysis of pre-defined groups of genes already implicated in human bleeding and platelet disorders. This has been recently illustrated by the ThromboGenomics NGS platform, which enables rapid and cost-effective sequence analysis of the coding regions and selected non-coding regions of 63 genes relevant to bleeding disorders, including 36 genes previously associated with IPD [10]. In a recent validation study in which 300 DNA samples from patients with IPD and other bleeding disorders, the ThromboGenomics platform showed 100% sensitivity for detecting causal variants in

samples in which genetic diagnosis had already been obtained by Sanger methodology. Genetic diagnosis was also achieved in more than 90% of cases with a phenotypically suspected disorder, but who had not undergone prior genetic analysis [10].

Gene discovery in inherited platelet disorders

In circumstances when analysis of a single candidate gene or gene panels does not yield a plausible genetic explanation for an IPD, NGS may also be an effective means of identifying causal variants in genes not previously associated with IPD. This typically requires extending the target regions for NGS analysis to the exonic regions of all protein coding genes (whole exome sequencing; WES) and has been adopted by several large scale gene discovery projects, including the UK Genotyping and Platelet Phenotyping (GAPP) [11] and the international BRIDGE-Bleeding and Platelet Disorders (BRIDGE-BPD) [12] programmes. These groups and others have achieved notable successes with WES, resulting in the discovery and characterisation of several new IPD (reviewed in [13])

Further advances in NGS technology have also now extended the scope of cost-effective sequencing from exome to whole genome level (whole genome sequencing; WGS). Current WGS techniques enable analysis of more than 98% of all genomic DNA with sufficient confidence to reliably detect sequence variants. This is essential to enable detection of some IPD in which the causal variants occur in regulatory regions outside the coding space, such as thrombocytopenia absent radius syndrome (*RBM8A*) [14] and ANKRD26-related thrombocytopenia (*ANKRD26*) [15]. WGS has recently been used to successfully identify *SRC* as a

new candidate gene for an IPD associated with thrombocytopenia associated with myelofibrosis and skeletal abnormalities [16].

Identification of causal variants for IPD from next generation sequencing data

Although WES and WGS dramatically increase the potential for gene discovery and genetic diagnosis in IPD, there are also some practical barriers for clinical delivery of these approaches. Chief amongst these is that even at exome level, sequence data from individual patients may show greater than 15,000 differences across the estimated 20,000 to 25,000 protein coding genes, compared to reference exome sequences. This may raise significant challenges in resolving variants in a single gene that are causally associated with an IPD from irrelevant bystander variants that are unrelated to the IPD.

One way of overcoming this barrier is to compare candidate sequence variants in IPD families against publicly available databases that record previously identified disease-associated variants. Examples of these resources include ClinVar (http://www.ncbi.nlm.nih.gov/clinvar/), the Exome Variant Server (http://evs.gs.washington.edu/EVS/) and the Human Gene Mutation Database (http://www.hgmd.cf.ac.uk), although it is essential to recognise that some variants listed in public resources may be identified as 'disease-associated', but are in fact bystander variants without functional consequences [17].

For variants not previously associated with IPD, many can be eliminated from further consideration by applying the assumption that variants that are common in the background population of the IPD case are unlikely to be causally associated with a

rare IPD. This filtering step is now more reliable since the publication of catalogues of population genetic variation through initiatives such as the UK10K project [18] and NHLBI exome sequencing project (http://evs.gs.washington.edu/EVS/), which provide increasingly refined estimates of the population frequencies of variants that are specific to ethnicity, and which can be used to assess rarity of candidate variants in IPD families. Variants that are predicted to have a high impact on gene translation (stop-gain, stop-loss, frame shift or splice donor/acceptor site) and missense variants are more commonly disease-associated than synonymous or non-coding region variants. Therefore, further refinement of candidate gene lists may sometimes be achieved by annotation of variants against likely gene transcripts and elimination of those with low predicted impact. More detailed predictions of the likely pathogenicity of missense variants may also be achieved using publically available computational algorithms such as the Combined Annotation Dependent Depletion (CADD) score (http://cadd.gs.washington.edu). Demonstration that a candidate variant segregates with the IPD phenotype within family structures may provide further support of causality.

Using phenotype to resolve causal IPD variants

Detailed analysis of the phenotype of individual IPD families is also a powerful tool to refine further the selection of candidate variants. This approach has been most successful for platelet function disorders in which careful laboratory identification of the defective platelet pathway can be used to narrow down variant lists to genes functionally implicated in specific platelet pathways, as illustrated by the recent discovery of *SLFN14* as a candidate gene for an IPD in which platelet secretion is defective [19].

For most IPD in which there are less distinctive features, there is still potential value in considering phenotype, but in conjunction with statistical genetic techniques to increase the power of resolving candidate variants. This has been best illustrated by the international BRIDGE-BPD consortium (https://bridgestudy.medschl.cam.ac.uk/bpd.shtml), in which a collection of more than 1000 patients with bleeding disorders or IPD has undergone detailed phenotype coding and either WES or WGS analysis. An essential part of this programme has been the systematic description of the clinical and laboratory characteristics of participants using the Human Phenotype Ontology (HPO) vocabulary, which has been updated comprehensively with terms for bleeding and platelet disorders [12]. This enables statistical analysis of the degree of similarity of phenotype terms between unrelated families and thereby, identification of causal variants for IPD by comparing WES or WGS datasets in families with a high degree of similarity. The efficacy of this approach is illustrated by the recent discovery of *DIAPH1* as a new IPD gene in two families sharing HPO terms for macrothrombocytopenia and hearing loss [20].

Ongoing gene discovery analyses are currently in progress for inherited platelet function disorders, which are likely to be more prevalent than platelet number disorders, but have proven more difficult to associate with causal variants in single genes. In part this reflect practical difficulties for many clinical centres in unambiguously defining which platelet pathway is defective using laboratory tests. However, systematic analyses of platelet function disorder patient cohorts by consortia such as the UK GAPP group using standardised light transmission aggregation and ATP secretion assays has successfully delineated specific

subgroups of platelet function disorders, including the major subgroups of Gi-pathway and non-syndromic dense granule secretion defects [21]. Careful analysis of NGS data from these sub-groups is expected to expand the repertoire of new IPD disorders defined at genetic level.

Next generation sequencing as mainstream for genetic diagnosis of IPD

The increasing accessibility of NGS, paralleled by the development of comprehensive population databases and statistical genetics techniques to enable detection of causal variants, is now transforming genetic diagnostic practice. This vision is a core objective for pioneering initiatives such as the UK National Health Service 100,000 Genomes project (www.genomicsengland.co.uk) which aims to perform WGS analysis of 100,000 DNA samples from UK rare disease and cancer patients by the end of 2018 and to introduce WGS into mainstream healthcare. This impetus promises to revolutionise our approach to genetic diagnosis in IPD and heralds the way for genomic medicine in all bleeding disorders.

REFERENCES

1. United Kingdom Haemophilia Doctors Organisation Annual Report 2015: at http://wwwukhcdoorg/annual-reports/ : Acessed 11th March 2016.

2. Gresele P, Harrison P, Bury L, Falcinelli E, Gachet C, Hayward CP, et al. Diagnosis of suspected inherited platelet function disorders: results of a worldwide survey. J Thromb Haemost. 2014;12(9):1562-9.

3. Noris P, Biino G, Pecci A, Civaschi E, Savoia A, Seri M, et al. Platelet diameters in inherited thrombocytopenias: analysis of 376 patients with all known disorders. Blood. 2014;124(6):e4-e10.

4. Balduini CL, Noris P. Innovation in the field of thrombocytopenias: achievements since the beginning of the century and promises for the future. Haematologica. 2016;101(1):2-4.
5. Noris P, Favier R, Alessi MC, Geddis AE, Kunishima S, Heller PG, et al. ANKRD26-related thrombocytopenia and myeloid malignancies. Blood. 2013;122(11):1987-9.
6. Noetzli L, Lo RW, Lee-Sherick AB, Callaghan M, Noris P, Savoia A, et al. Germline mutations in ETV6 are associated with thrombocytopenia, red cell macrocytosis and predisposition to lymphoblastic leukemia. Nat Genet. 2015;47(5):535-8.
7. Pecci A, Klersy C, Gresele P, Lee KJ, De Rocco D, Bozzi V, et al. MYH9-related disease: a novel prognostic model to predict the clinical evolution of the disease based on genotype-phenotype correlations. Hum Mutat. 2014;35(2):236-47.
8. Pecci A, Gresele P, Klersy C, Savoia A, Noris P, Fierro T, et al. Eltrombopag for the treatment of the inherited thrombocytopenia deriving from MYH9 mutations. Blood. 2010;116(26):5832-7.
9. Gerrits AJ, Leven EA, Frelinger AL, 3rd, Brigstocke SL, Berny-Lang MA, Mitchell WB, et al. Effects of eltrombopag on platelet count and platelet activation in Wiskott-Aldrich syndrome/X-linked thrombocytopenia. Blood. 2015;126(11):1367-78.
10. Simeoni I, Stephens JC, Hu F, Deevi S, Megy K, Bariana TK et al. A comprehensive high-throughput sequencing test for the diagnosis of inherited bleeding, thrmbotic and platelt disorders. Blood 2016 (currently under review).
11. Leo VC, Morgan NV, Bem D, Jones ML, Lowe GC, Lordkipanidze M, et al. Use of next-generation sequencing and candidate gene analysis to identify underlying

defects in patients with inherited platelet function disorders. J Thromb Haemost. 2015;13(4):643-50.
12. Westbury SK, Turro E, Greene D, Lentaigne C, Kelly AM, Bariana TK, et al. Human phenotype ontology annotation and cluster analysis to unravel genetic defects in 707 cases with unexplained bleeding and platelet disorders. Genome Med. 2015;7(1):36.
13. Lentaigne C, Freson K, Laffan M, Turro E, Ouwehand W. Progress in genetic diagnosis of inherited platelt disorders. Blood. 2016. (currently under review)
14. Albers CA, Paul DS, Schulze H, Freson K, Stephens JC, Smethurst PA, et al. Compound inheritance of a low-frequency regulatory SNP and a rare null mutation in exon-junction complex subunit RBM8A causes TAR syndrome. Nat Genet. 2012;44(4):435-9, S1-2.
15. Noris P, Perrotta S, Seri M, Pecci A, Gnan C, Loffredo G, et al. Mutations in ANKRD26 are responsible for a frequent form of inherited thrombocytopenia: analysis of 78 patients from 21 families. Blood. 2011;117(24):6673-80.
16. Turro E, Greene D, Wijgaerts A, Thys C, Lentaigne C, Bariana TK, et al. A dominant gain-of-function mutation in universal tyrosine kinase SRC causes thrombocytopenia, myelofibrosis, bleeding, and bone pathologies. Sci Transl Med. 2016;8(328):328ra30.
17. Xue Y, Chen Y, Ayub Q, Huang N, Ball EV, Mort M, et al. Deleterious- and disease-allele prevalence in healthy individuals: insights from current predictions, mutation databases, and population-scale resequencing. Am J Hum Genet. 2012;91(6):1022-32.

18. Kaye J, Hurles M, Griffin H, Grewal J, Bobrow M, Timpson N, et al. Managing clinically significant findings in research: the UK10K example. Eur J Hum Genet. 2014;22(9):1100-4.
19. Fletcher SJ, Johnson B, Lowe GC, Bem D, Drake S, Lordkipanidze M, et al. SLFN14 mutations underlie thrombocytopenia with excessive bleeding and platelet secretion defects. J Clin Invest. 2015;125(9):3600-5.
20. Stritt S, Nurden P, Turro E, Greene D, Jansen SB, Westbury SK, et al. A gain-of-function variant in DIAPH1 causes dominant macrothrombocytopenia and hearing loss. Blood. 2016. doi: 10.1182/blood-2015-10-675629.
21. Dawood BB, Lowe GC, Lordkipanidze M, Bem D, Daly ME, Makris M, et al. Evaluation of participants with suspected heritable platelet function disorders including recommendation and validation of a streamlined agonist panel. Blood. 2012;120(25):5041-9.

Platelet Function Disorder*	**Genes**	**Non-hematological manifestations**
Arthrogryposis, renal dysfunction and cholestasis 1 and 2 (#208085, #613404)	*VPS33B; VIPAS39*	Arthrogryposis, renal dysfunction, cholestasis
Bernard-Soulier syndrome (#231200)	*GP1BA; GP1BB; GP9*	None
Chediak-Higashi syndrome (#214500)	*LYST*	Partial albinism, impaired phagocytosis
Deficiency of phospholipase A2, group IVA (no #)	*PLA2G4A*	None
Dense granule abnormalities (no #)	*NBEA*	None
Familial haemophagocytic lymphohistiocytosis type 5 (#613101)	*STXBP2*	None
Ghosal hematodiaphyseal dysplasia (#231095)	*TBXAS1*	Skeletal abnormalities
Glanzmann thrombasthenia (#273800)	*ITGA2B; ITGB3*	None
Hermansky-Pudlak syndromes 1-9 (#203300, #608233, #614072, #614073, #614074, #614075, #614076, #614077, #614171)	*HPS1; AP3B1; HPS3; HPS4; HPS5; HPS6; DTNBP1; BLOC1S3; BLOC1S6*	Oculocutaneous albinisim, pulmonary fibrosis, colitis (dependent on subtype)
Leucocyte integrin adhesion deficiency, type III (#612840)	*FERMT3*	None
PAR4 defect (no #)	*F2RL3*	None
Platelet-type bleeding disorder 8 (#609821)	*P2RY12*	None
Platelet-type bleeding disorder 11 (#614201)	*GP6*	None
Platelet-type bleeding disorder 13 (#614009)	*TBXA2R*	None
Platelet-type bleeding disorder 18 (#615888)	*RASGRP2*	None
Quebec platelet disorder (#601709)	*PLAU*	None
Scott syndrome (#262890)	*ANO6*	None

Platelet Number Disorders*	Genes	Non-hematological manifestations
Amegakaryocytic thrombocytopenia with radio-ulnar synostosis 1 and 2 (#605432, #616738)	*HOXA11, MECOM*	Skeletal abnormalities
Autosomal dominant thrombocytopenia 2 (#188000)	*ANKRD26*	None
Autosomal dominant thrombocytopenia 4 (#612004)	*CYCS*	None
Congenital amegakaryocytic thrombocytopenia (#604498)	*MPL*	None
Dominant macrothrombocytopenia and hearing loss (no #)	*DIAPH1*	Sensorineural hearing impairment
Familial platelet disorder with associated myeloid malignancy (#601399)	*RUNX1*	None
Gray platelet syndrome (#139090)	*NBEAL2*	None
Inherited thrombocytopenia with early onset myelofibrosis (no #)	*SRC*	Skeletal abnormalities, dysmorphic facies
Inherited thrombocytopenia with excessive bleeding (no #)	*SLFN14*	None
Macrothrombocytopenia (no # for isolated platelet phenotype)	*FLNA*	Periventricular nodular heterotopia, otopalatodigital syndrome
May-Hegglin anomaly and other *MYH9*-related disorders (#155100, #153640, #153650, #605249)	*MYH9*	Cataract, renal dysfunction, sensorineural hearing impairment
Paris-Trousseau type thrombocytopenia and Jacobson syndrome (#188025, #147791)	*FLI1*	Multiple features including developmental delay, skeletal abnormalities, congenital cardiac defects
Platelet-type bleeding disorder 15 (#615193)	*ACTN1*	None
Platelet-type bleeding disorder 17 (#187900)	*GFI1B*	None
Pseudo-von Willebrand disease (#177820)	*GP1BA*	None
Stormorken syndrome (#185070)	*STIM1*	Myopathy, mioisis, asplenia, ichthyosis
Thrombocytopenia 5 (#616216)	*ETV6*	Solid organ malignancy
Thrombocytopenia absent radius syndrome (#274000)	*RBM8A*	Skeletal abnormalities, congenital cardiac defects
TUBB1-related autosomal dominant macrothrombocytopenia (#613112)	*TUBB1*	None
Wiskott-Aldrich syndrome (#301000)	*WAS*	Immune deficiency, eczema

Table 2: Example disorders of platelet number defined at gene level. * Disorder names are according to Online Mendelian Inheritance In Man nomenclature (http://www.ncbi.nlm.nih.gov/omim) with phenotype reference numbers where appropriate.