High-throughput sequencing for rapid diagnosis of inherited platelet disorders: a case for a European consensus

Alan T. Nurden and Paquita Nurden

Institut de Rythmologie et de Modélisation Cardiaque, Plateforme Technologique d'Innovation Biomédicale, Hôpital Xavier Arnozan, Pessac, France

E-mail: nurdenat@gmail.com

doi:10.3324/haematol.2017.182295

The diagnosis of bleeding disorders caused by inherited defects of platelet function or production (or combinations of both) forms an intrinsic part of the work of every hematological laboratory.[1] The classic procedure for the work-up of inherited platelet disorders (IPDs) is based on an initial examination of the patient by a specialist in the field, followed by a series of often complex biological tests designed to identify molecular pathways likely to be affected in each case. Algorithms have been proposed to aid in this task while human phenotype ontology (HPO) annotation and cluster analysis has been recommended as part of phenotyping.[2,3] However, in many laboratories the tests chosen are often limited by availability, while intensive HPO interrogation is time-consuming and often not performed in the context of an initial hospital consultation. Despite such restrictions, diagnosis advances sufficiently for many patients to justify DNA sequencing of selected genes; a situation usually reserved for patients with classic IPDs and a well-defined phenotype. Frequently the work-up is long and costly, and requires the patient to make multiple visits to his/her local hospital, and possibly to a specialized centre. Too often, the result ends in the lack of a clearly defined diagnosis.

In the current edition of Haematologica, Bastida *et al.*[4] present a high-quality pilot study introducing high-throughput gene sequencing into the mainstream of genetic diagnostic practice for IPDs. Working in the context of a national collaboration, these authors designed a novel platform to identify gene variants in 82 patients, the majority of whom come from the Iberian Peninsula, with IPDs of previously undiagnosed molecular causes. Their procedure consisted of using 1399 probes targeting 1106 regions of 72 genes known to be associated with IPDs and/or encoding proteins significant in platelet physiology. With step-by-step use of bioinformatics tools and informed interpretation of clinical and biological data, they were able to identify likely disease-causing genes in the majority of patients. First, they validated their platform by successfully confirming the mutations previously assigned by Sanger sequencing in an additional group of 10 patients with known IPDs. For 34 patients with a high suspicion of defined IPDs, they identified candidate variants in 30 patients (88.2%) (Figure 1). The success rate for a cohort of 48 patients of uncertain etiology was lower, with 26 variants identified in 16 genes in 26 cases. In 22 cases disease-causing genes were not identified.

Pioneering studies worldwide in the application of next-generation sequencing projects to IPDs came from the BRIDGE-Bleeding and Platelet Disorders (BPD) consortium led by Professor Willem Ouwehand, Cambridge, UK; a project in which we were early members, submitting DNA from 80 well-characterized patients in 2011. BRIDGE initially used whole exome sequencing (WES), and later whole genome sequencing (WGS) to identify gene variants, while also establishing the ThromboGenomics gene platform, which originally contained 63 genes.[5,6] The BRIDGE studies represent the gold standard in the discipline, helped by the sequencing capacity of the Sanger Institute and their extensive back-up bioinformatic analysis. In Europe, in parallel, a Birmingham-based UK group led by Professor Steve Watson looked specifically at UK patients with IPDs in the Genotyping and Phenotyping of Platelets (GAPP) study.[7] The data obtained from these groups and others world-wide[8,9] helped Bastida *et al.*[4] to compose their gene platform. However, one disadvantage of selected gene platforms is that successful identification of disease-causing variants is limited to the genes tested, and this may explain the 22 cases where the causal gene defect was not found in the Spanish study. WES, and particularly WGS, offer the potential to identify new genes, especially when combined with selective approaches such as comparing patient phenotypes with those of knock-out mouse models or gene datasets; approaches that allowed us to identify *TRPM7* (from an existing mouse model) and *DIAPH1* (from a search of genes responsible for deafness) in patients belonging to the French cohort which we submitted to BRIDGE.[10,11] Recent advances in bioinformatic analysis of next-generation sequencing data are enhancing causal gene identification in IPDs, helped by HPO questioning and the study of variant penetrance within large families.[12] It would be interesting to now apply WGS to those cases in the Spanish cohort for whom disease-causing variants were not identified. It should be underlined that finding a novel gene variant is not in itself final proof that it is causing the disease, even with penetration within the family. Cosegregated but non-highlighted genetic variants may also be contributing to phenotype. Thus, structure/function studies on the protein are often necessary, and/or studies on mouse or zebrafish models as was performed for a recently described *SRC* variant;[13] a situation that now applies to several of the gene variants identified by Bastida *et al.*[4]

Improving bioinformatic approaches may also be helpful in identifying gene variants which often display heterozygous expression and only cause bleeding in rare phenotypes when acting in combination. Haplogroups of genetic variants that are individually benign may be widespread, originating in part from now dispersed ethnic and religious minorities. In this regard, we were sur-

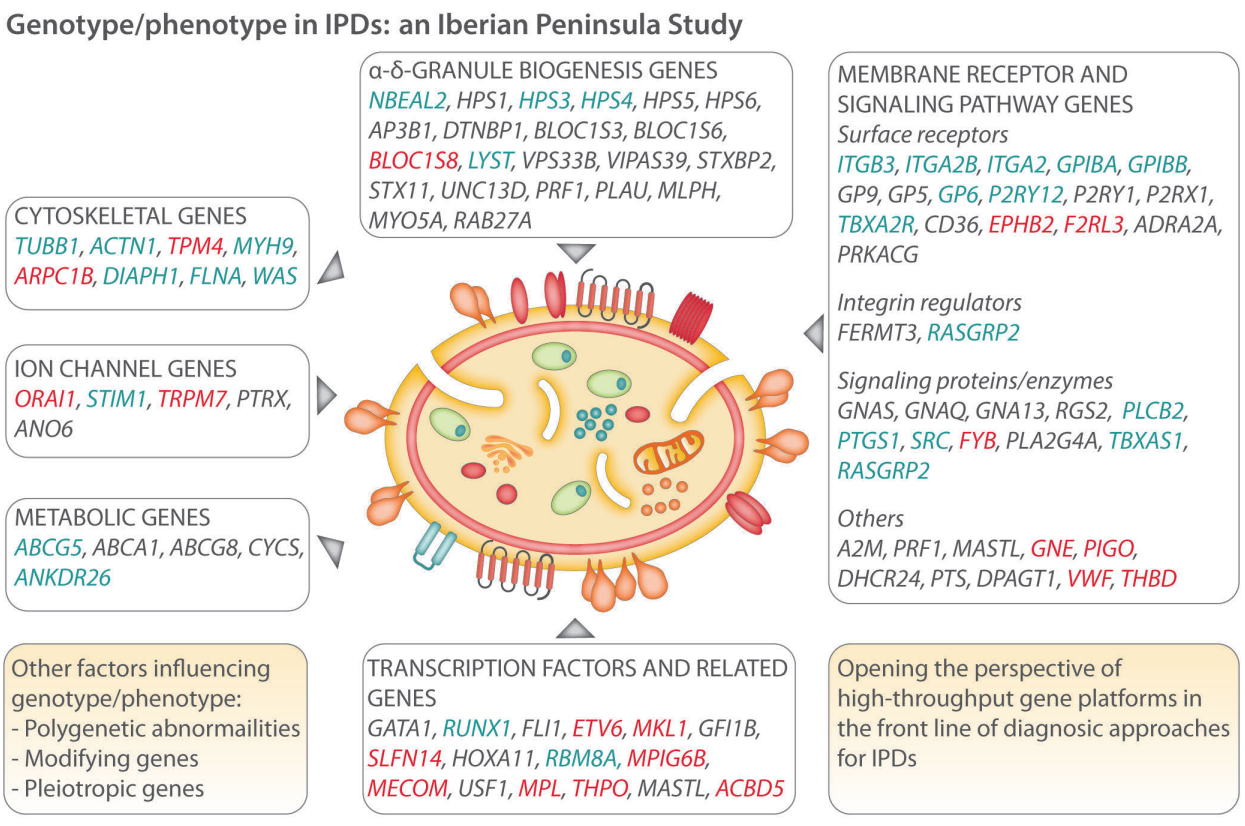

Figure 1. Genotype/phenotype in IPDs: an Iberian Peninsula Study. An illustration showing a platelet (center) surrounded by boxes naming the genes that are arbitrarily grouped into gene families constituting the platform composed by Bastida *et al.*[4] The genes with mutations thought likely or probable to be pathogenic for IPDs within the Iberian Peninsula are highlighted in blue. In red are those genes recently highlighted in the literature to be disease-causing, and which are candidates to be added to the platform.

prised upon discovering that when comparing single nucleotide polymorphisms (SNPs) within selected domains of *ITGA2B* and *ITGB3* in patients with Glanzmann thrombasthenia (GT), combinations of SNPs reoccurred in ostensibly non-related families.[14] Furthermore, studies on GT have clearly shown that while defects in *ITGA2B* and *ITGB3* confirm the disease, other genetic and epidemiological factors define bleeding risk and possibly the frequency of other associated pathologies.[15] This suggests that next-generation sequencing data for IPDs should also be analyzed and gene platforms designed to identify gene variants that may influence bleeding independently of the disease-causing genotype and/or increased susceptibility for associated illnesses – an example being the risk of hematological malignancy with genes such as *RUNX1* or *ETV6* or with defects in immunity, as has been recently performed for inhibitor development in hemophilia A.[16,17] This would be of great benefit with regard to patient care.

So, what is the best way to continue? While highlighting the advances in Europe, the Spanish-based study in the current issue of Haematologica by Bastida *et al.*,[4] or the parallel Scandinavian study[18] show the now obvious need for greater collaboration and standardization. In Europe, many countries already have national or regional sequencing centers specializing in rare diseases, while public databases such as the Exome Aggregation Consortium (ExAC) provide continued analysis of rare gene variants through input from genome-wide association studies and large-scale sequencing, such as the 1000 genomes project. Although easier to manage, high-throughput gene screening platforms need to evolve continuously as disease-causing variants continue to be identified and the whole genome landscape of IPDs emerges.[19,20] The ideal would be to establish a consensus for gene platforms within Europe, permitting local reference centers throughout the European Community to analyze IPD samples. For difficult cases, the next logical step would be to provide back-up links to specialized centers, as highlighted by BRIDGE, with facilities for WGS in order to research not only new disease-causing variants but also the effects of gene enhancers and pleiotropy. Ideally a result should be obtained within days for the bulk of patients, both for the cause of a disease and to predict the risk of bleeding. Thus, next-generation technology can form part of frontline hospital practice, with biological tests performed on the basis of gene testing rather than vice versa. It would also permit uniformity while conserving national identity. In the platelet field this would lead to cost-cutting, increased efficiency and much improved patient care. Advances in Europe would also provide a model for other parts of the world.

References

1. Gresele P, Harrison P, Bury L, et al. Diagnosis of suspected inherited platelet function disorders: results of a worldwide survey. J Thromb Haemost. 2014; 12(9):1562-1569.
2. Noris P, Pecci A, Di Bari F, et al. Application of a diagnostic algorithm for inherited thrombocytopenias to 46 consecutive patients. Haematologica. 2004; 89(10):1219-1225.
3. Westbury SK, Turro E, Greene D, 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; Apr 9;7(1):36. Doi: 10.1186/s13073-015-015-5. eCollection 2015.
4. Bastida JM, Lozano ML, Benito R, et al. Introducing high-throughput sequencing into the mainstream of genetic diagnosis practice in inherited platelet disorders. Haematologica. 2018;103(1):000-000.
5. Lentaigne C, Freson K, Laffan MA, Turro E, Ouwehand WH, on behalf of the BRIDGE-BPD Consortium and the ThromboGenomics Consortium. Inherited platelet disorders: towards DNA-based diagnosis. Blood 2016; 127(23):2815-2823.
6. Simeoni I, Stephens JC, Hu F., et al. A high-throughput sequencing test for diagnosing inherited bleeding, thrombotic, and platelet disorders. Blood 2016; 127(23):2791-2803.
7. 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(10):1170-1179.
8. Kahr WH, Pluthero FG, Elkadri A, et al. Loss of the Arp2/3 complex component ARPC1B causes platelet abnormalities and predisposes to inflammatory disease. Nat Commun 2017; Apr 3;8:14816. doi: 10.1038/ncomms14816.
9. Ferreira CR, Chen D, Abraham SM, et al. Combined alpha-delta platelet storage pool deficiency is associated with mutations in GFI1B. Mol Genet Metab 2017; 120(3):288-294.
10. Stritt S, Nurden P, Favier R, et al. Defects in TRPM7 channel function deregulate thrombopoiesis through altered cellular Mg(2+) homeostasis and cytoskeletal architecture. Nat Commun 2016; Mar 29:711097. Doi: 10. 1038/ncomms 11097.
11. Stritt S, Nurden P, Turro E, et al. A gain-of-function variant in DIAPH1 causes dominant macrothrombocytopenia and hearing loss. Blood 2016; 127(23):2903-2914.
12. Greene D, NIHR BioResource, Richardson S, Turro E. A fast association test for identifying pathogenic variants involved in rare diseases. Am J Hum Genet 2017; 101(1):104-114.
13. Turro E, Greene D, Wijgaerts A, et al. A dominant gain-of-function in universal tyrosine kinase SRC causes thrombocytopenia, myelofibrosis, bleeding, and bone pathologies. Sci Transl Med 2016; 8(328):328ra30. doi: 10.1126/scitranslmed.aad7666. Epub 2016 Mar 2.
14. Pillois X, Nurden AT. Linkage disequilibrium amongst ITGA2B and ITGB3 gene variants in patients with Glanzmann thrombasthenia confirms that most disease-causing mutations are recent. Br J Haematol 2016; 175(4):686-695.
15. Nurden AT. Should studies on Glanzmann thrombasthenia not be telling us more about cardiovascular disease and other major illnesses? Blood Rev 31(5):287-299.
16. Sood R, Kamikubo Y, Liu P. Role of RUNX1 in haematological malignancies. Blood 2017; 129(15):2070-2082.
17. Gorski MM, Blighe K, Lotta LA, et al. Whole-exome sequencing to identify genetic risk variants underlying inhibitor development in severe haemophilia A patients. Blood 2016; 127(23):2924-2933.
18. Leinoe E, Zetterberg E, Kinalis S, et al. 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. Br J Haematol 2017; 178(2):308-322.
19. Astle WJ, Elding H, Jiang T, et al. The allelic landscape of human blood cell trait variation and links to common complex disease. Cell 2016; 167(5):1415-1429.
20. Chami N, Chen MH, Slater AY, et al. Exome genotyping identifies pleiotropic variants associated with red blood cell traits. Am J Hum Genet 2016; 99(1):8-21.