haematologica
Journal of The Ferrata Storti Foundation

Complications of whole-exome sequencing for causal gene discovery in primary platelet secretion defects

by Marcin M. Gorski, Anna Lecchi, Eti A. Femia, Silvia La Marca, Andrea Cairo, Emanuela Pappalardo, Luca A. Lotta, Andrea Artoni, and Flora Peyvandi

Haematologica 2019 [Epub ahead of print]

Citation: Marcin M. Gorski, Anna Lecchi, Eti A. Femia, Silvia La Marca, Andrea Cairo, Emanuela Pappalardo, Luca A. Lotta, Andrea Artoni, and Flora Peyvandi. Complications of whole-exome sequencing for causal gene discovery in primary platelet secretion defects. Haematologica. 2019; 104:xxx
doi:10.3324/haematol.2018.204990

Complications of whole-exome sequencing for causal gene discovery in primary platelet secretion defects

Marcin M. Gorski,[1,2] Anna Lecchi,[1] Eti A. Femia,[1] Silvia La Marca,[1] Andrea Cairo,[1] Emanuela Pappalardo,[1,2] Luca A. Lotta,[1] Andrea Artoni,[1] and Flora Peyvandi.[1,2]

[1] Fondazione IRCCS Ca' Granda Ospedale Maggiore Policlinico, Angelo Bianchi Bonomi Hemophilia and Thrombosis Center, Milano, Italy

[2] Università degli Studi di Milano, Department of Pathophysiology and Transplantation and Fondazione Luigi Villa, Milano, Italy

Running title: Exome sequencing in platelet secretion defects

Manuscript word count: **3070**

Abstract word count: **228**

Tables: **3**

Figures: **2**

Supplementary file: **1**

Correspondence to:

Flora Peyvandi, MD, PhD

Fondazione IRCCS Ca' Granda Ospedale Maggiore Policlinico

Angelo Bianchi Bonomi Hemophilia and Thrombosis Center

Via Pace 9, 20122 Milano, Italy.

Tel: +39 02 5503 5414. Fax: +39 02 54 100 125.

Email: flora.peyvandi@unimi.it

Keywords: next-generation DNA sequencing, platelet secretion defects, platelet disorders, platelet aggregation.

Abstract

Primary platelet secretion defects are heterogeneous group of functional defects characterized by reduced platelet granule secretion upon stimulation by different agonists. The clinical and laboratory heterogeneity of primary platelet secretion defects warrants a tailored approach. We performed a pilot study in order to develop DNA sequence analysis pipelines for gene discovery and to create a list of candidate causal genes for platelet secretion defects. Whole-exome sequencing analysis of 14 unrelated Italian patients with primary secretion defects and 16 controls was performed on Illumina HiSeq. Variant prioritization was carried out using two filtering approaches: identification of rare, potentially damaging variants in platelet candidate genes or by selecting singletons. To corroborate the results, exome sequencing was applied in a family, in which platelet secretion defects and bleeding diathesis were present.

Platelet candidate gene analysis revealed gene defects in 10/14 patients, which included *ADRA2A*, *ARHGAP1*, *DIAPH1*, *EXOC1*, *FCGR2A*, *ITPR1*, *LTBP1*, *PTPN7*, *PTPN12*, *PRKACG*, *PRKCD*, *RAP1GAP*, *STXBP5L*, and *VWF*. The analysis of singletons identified additional gene defects in *PLG* and *PHACTR2* in two other patients. The familial analysis confirmed a missense variant p.D1144N in the *STXBP5L* gene and p.P83H in the *KCNMB3* gene as potentially causal. In summary, exome sequencing revealed potential causal variants in 12 of 14 patients with primary platelet secretion defects, highlighting the limitations of the genomic approaches for causal gene identification in this heterogeneous clinical and laboratory phenotype.

Introduction

Disorders of platelet function are characterized by highly variable mucocutaneous bleeding manifestations and excessive hemorrhage following surgical procedures or trauma.[1-4] Primary platelet secretion defects (PSD) are the most common platelet functional defects,[5] which displays both clinical and laboratory heterogeneity.[6] From a clinical standpoint, PSD may be associated with mild to severe bleeding tendency.[7] Thus, due to the heterogeneous nature of platelet secretion defects, laboratory testing is limited to specialized laboratories and accurate mechanistic diagnosis remains challenging.

Platelet aggregation and secretion studies with lumi-aggregometry, where dense granule secretion is assessed in parallel with traditional light transmission aggregometry (LTA), provide evidence for platelet dysfunction.[8, 9] PSD is characterized by reduced or absent delta granule secretion upon stimulation by one or more platelet aggregation agonists either at low or high doses.[8, 9] However, lumi-aggregometry, the gold standard technique for platelet function studies, is not always predictive of the molecular mechanisms, rendering mechanistic differentiation of primary PSD difficult.

Multiple inherited alterations of platelet function have been described, including forms with different pattern of inheritance.[2, 4, 10] For conditions where the laboratory phenotype was not discriminating, genotyping using the next-generation DNA sequencing (NGS) could be a comprehensive and cost-effective strategy for the diagnosis of platelet function disorders.[11-13] Indeed, the application of NGS-based approaches, based on the application of whole-exome sequencing (WES) or custom gene panels, proved to be successful for the diagnosis of inherited platelet defects.[11, 13, 14] Leo et al. applied WES to study 329 candidate genes involved in platelet function defects and identified gene variants in patients with defects in Gi signaling and with platelet secretion abnormalities.[15] WES was also successful in identifying causal mutations in the *RASGRP2* gene, which encodes a protein required for signaling and platelet activation[16, 17] or in identifying a causal mutation displaying autosomal dominant inheritance located in the *THBD* gene.[18] However, a standardized pipeline or procedure linking the identified gene defects to the specific sub-phenotype of diverse platelet function disorders is still missing.

Given the positive experience acquired with the use of WES in identifying potentially pathogenic genetic variants in the platelet function defects, the use of NGS-based diagnostics provides a great opportunity in improving causal gene identification and understanding the underlying clinical phenotype. [19-22] For this reason, we decided to apply exome sequencing in a well-characterized

group of patients with primary PSD and clinically relevant bleeding.[5] The aim of our pilot study was to test whether WES will be an adequate diagnostic tool for causal gene discovery in a heterogeneous group of platelet function defects such as primary PSD.

Methods

Study population

Fourteen unrelated patients with diagnosis of primary PSD were enrolled among 360 individuals with suspected platelet function disorder referred to our outpatient clinic at Ospedale Maggiore Policlinico, Milano (IT).

Patients inclusion criteria were: (i) European ancestry; (ii) platelet count >120.000 /µL; (iii) impaired platelet ATP secretion after stimulation with two or more agonists measured by lumiaggregometry; (iv) normal expression of platelet glycoprotein (GP) Ib/IX/V and GPIIb/IIIa to exclude Bernard-Soulier syndrome and Glanzmann thrombasthenia; (v) absence of any other known platelet disorder; and (vi) absence of von Willebrand (VW) disease.

Four family members of one patient (C740) were also included and studied.

All studied subjects abstained from drugs affecting platelet function for two weeks before blood sampling. All platelet function results were compared with our internal normal range.

The study was approved by local Ethical Committee of Ospedale Maggiore Policlinico, Milano (IT) and carried out according to the Declaration of Helsinki. All participants signed informed consent.

Platelet phenotyping

Personal and family history, blood tests including complete blood count, prothrombin time (PT) and activated partial thromboplastin time (aPTT) by standard methods, VW Factor (F) Antigen and VWF Ristocetin cofactor by automated latex enhanced immunoassay (Instrumentation Laboratory, Milano, IT)[23] were collected *(On line Supplementary methods)*. The bleeding severity score (BSS) was calculated for each patient according to Tosetto et al.[24] (v.n.: child<2; men<5; woman<6).

Blood samples were drawn in trisodium citrate for coagulation, VWF measurement, and platelet function studies and in K–EDTA for DNA extraction[25] and blood cell count.

Platelet aggregation and ATP secretion induced by ADP (4 and 20 µM), collagen (2 µg/mL), thrombin receptor activator peptide (TRAP)-14 (10 µM), and thromboxane A2 analogue U46619 (1µM) were measured in platelet-rich plasma (PRP) by lumiaggregometry (Chrono-log 560, Mascia

Brunelli, Milano, IT). [26] PRP was prepared as previously reported.[27] Intraplatelet ADP, ATP, serotonin, and fibrinogen content was measured as previously reported[28, 29] (*On line Supplementary Methods).*

Whole-exome sequencing

Individual exomes were enriched using a SeqCap EZ Human Exome Library kit v2.0 (Roche NimbleGen) and paired-end sequencing was carried out on the HiSeq2000 (Illumina, San Diego, CA) at the Beijing Genomics Institute (www.bgi.com).

The Short Oligonucleotide Analysis Package aligner (soap2.21)[30] was used to align reads to the reference human genome (hg19/GRCh37) and produce individual Binary Alignment Map (BAM) files. The Genome Analysis Tool Kit (GATK) was used for quality recalibration, duplicate read marking, insertions/deletions (indels) realignment, and BAM sorting to produce merged, sample-level variant calling file (VCF) *(On line Supplementary Methods).*

Variant filtering and candidate gene discovery

Variant filtering and candidate gene discovery were performed on the project level, merged VCF file containing 14 unrelated Italian PSD patients and 16 healthy controls by using two different filtering strategies: selection of singletons and filtering for the SNVs reported by Leo et al.[15] (*On line Supplementary Methods).*

Variant pathogenicity was assigned according to the American College of Medical Genetics and Genomics (ACMG) pathogenicity classification.[31] Platelet gene expression was evaluated using the Human Proteome Map (HPM).[32] *(On line Supplementary Methods).*

Results

Clinical characteristics of patients with PSD

Of 360 patients with suspected platelet disorders investigated at our center, 14 unrelated patients (12 females and 2 males; median age 23 years) fulfilled inclusion criteria (Table 1). Patient BSS ranged between 0-15 and 64% of the cases resulted abnormal (Table 1). PT, aPTT, plasma fibrinogen, and VWF levels were within the normal range (data not shown). Platelet count was normal in all PSD patients (median 258 $x10^9$/L, min-max 120-357; v.n. 150-450), except for patient C749 who had a slightly lower platelet count (120 $x10^9$/L).

Platelet functions studies

Platelet aggregation was lower than the normal range in the majority of the patients with all agonists tested (Figure 1A) and rapidly reversible in 60% of the cases when induced by ADP (4µM). Platelet ATP secretion was absent after stimulation by ADP (4µM) in all patients and lower than the normal range in response to the other agonists in the majority of cases (Figure 1B). In particular, platelet secretion was impaired with two stimuli in 4/14 patients, with three in 4/14, and with more than three in 6/14 (Table 1). These findings confirmed the diagnosis of primary PSD in all patients.

The concentrations of total serotonin, ADP and ATP were normal in all patients including the ATP/ADP ratio, which is considered a diagnostic hallmark for δ-storage pool deficiency *(On line Supplementary Table S1)*. Similarly, fibrinogen from platelet α-granules was normal. All together, these data excluded that the secretion defect of these patients was attributable to the presence of α-or δ-storage pool deficiency.

Exome sequencing and candidate gene discovery

NGS data analysis revealed 101,562 variants that passed quality control and were sequenced with an average read depth of 51 over each site. Of those, 96,432 were single nucleotide variants (SNVs) and 5,130 were indels. Singletons defined as private variants occurring exclusively in a single individual were 11,430 (mean = 762) in PSD cases and 23,564 (mean = 1,473) in controls. In addition, we identified 30,973 rare variants with MAF ≤1% and 11,187 of these variants were considered novel, i.e., not listed in dbSNP or any other variant database.

Platelet candidate gene filtering approach

Candidate gene discovery was carried out by two independent filtering approaches: identification of variants in platelet candidate genes and by selecting singletons *(On line Supplementary Figure* S1). In the first approach, we selected from PSD patients all rare, potentially deleterious variants located in the coding regions of 329 candidate platelet genes listed by Leo et al.[15] This prioritizing strategy revealed 37 gene defects, of which six were novel *(On line Supplementary Table S2)*. Since this variant prioritizing strategy yielded multiple SNVs for the following patients C729 (5 SNVs), C732 (4 SNVs), C739 (4 SNVs), C740 (7 SNVs), C831 (4 SNVs), we used the ACMG variant pathogenicity classification,[31] which revealed 14 gene defects classified as variants of uncertain significance (VUS) in eight patients. To provide functional analysis of these genes, we assessed

their expression patterns in platelets using the Human Proteome Map (HPM), which integrates mass spectrometry analysis of different human tissues and cell types as part of the human proteome project.[32] This evaluation identified potential gene defects in seven PSD patients: *EXOC1* (C732), *DIAPH1* (C739), *STXBP5L* and *PRKACG* (C740), *PTPN12* (C749), *VWF* (C831), *PRKCD* (C1075), *PTPN7* and *PRKCD* (C1107).

Singleton filtering approach

Given that the first approach failed to identify gene defects in six patients, we decided to apply another filtering strategy based on the isolation of singletons. To this end, we selected from all 14 patient private variants, which were rare and possibly deleterious and we obtained 2,875 SNVs in 2,162 genes. To prioritize these SNVs for their putative role in platelet secretion defects, we performed functional annotation using the Database for Annotation, Visualization and Integrated Discovery (DAVID).[33] Significantly associated Gene Ontology (GO) annotations were found for gene clusters in the following functional categories: biological process - extracellular matrix organization for 48 genes (*P*=2.1E-07, Bonferroni *P*=9.9E-04); cellular component - basal lamina containing 10 genes (*P*=5.7E-06, Bonferroni *P*=4.4E-03); molecular function - extracellular matrix structural constituent comprising 22 genes (*P*=5.6E-06, Bonferroni *P*=8.3E-3). In addition, the KEGG pathway analysis (www.genome.jp/kegg/pathway.html) revealed once again a cluster of 26 genes with functional annotation associated with extracellular matrix-receptor interactions (*P*=2.9E-06, Bonferroni *P*=7.9E-04). The extracellular matrix functional category can be defined as any material produced by cells and secreted into the surrounding medium, which include collagen, laminin, fibronectin proteins and glycosaminoglycans (http://www.uniprot.org/keywords/?query=Extracellular%20matrix), indicating that our prioritizing method had indeed identified genes potentially affected in platelet secretion defects. Functional overlap between the above-mentioned gene clusters was achieved by enriching for variants present in genes exhibiting GO terms such as platelets and secretion, platelets and granules, platelets and signaling.

In this way, we identified 70 potential gene defects, of which 68 were missense variants. We also found a STOP gain variant in the *PHF14* gene (c.G298T, p.E100X) in patient C749 and a frameshift deletion in the *TBXAS1* gene (c.151_152delGT, p.V51fs) present in patient C831. Importantly, all 37 missense variants identified by filtering for gene defects in platelet candidate genes were also

found in the list of singletons, which together produced a list of 107 candidate gene defects listed in the *(On line Supplementary Table S2)*.

Similar to the previous filtering strategy, the singleton approach revealed an excess of potential gene defects in several patients *(On line Supplementary Table S2)*. To be able to assign causality, further reduction in the number of SNVs was necessary. To this end, we once again used the ACMG variant pathogenicity classification,[31] which resulted in the identification of 22 putative gene defects classified as VUS in 10 patients with primary PSD. However, only 13 of these variants were located in genes expressed in human platelets according to the HPM[32] (Table 2). In summary, this variant prioritization approach provided candidate gene defects for four patients C696, C708, C797 and C847, for whom the previous strategy was ineffective. It is interesting to note that several of these gene defects were missing from the list of Leo et al.,[15] indicating that these genomic loci could potentially become novel candidate genes associated with PSD.

Family analysis of patient C740

Only one notable pedigree, case C740, was investigated. The distribution of PSD phenotype and BSS in his relatives are reported in Figure 2 (father C1300, mother C1301, and two sisters C1302 and 1304). WES was performed in all four individuals and the variant filtering steps were based on MAF≤1%, selecting SNVs with potentially damaging consequences and assuming disease transmission present in affected and absent in unaffected family members *(On line Supplementary Figure S2*). Upon classification according to the ACMG,[31] four SNVs were confirmed in heterozygous state in PSD affected C740 and father C1300, suggesting an autosomal dominant transmission of the disease. Two of those, p.D1144N in the *STXBP5L* gene and p.P83H in the *KCNMB3* gene, classified as VUS (Table 3) may be involved in secretion process, thus being the most probable gene defects responsible for the PSD phenotype in this family.

Discussion

In this pilot study, we performed WES in 14 unrelated Italian patients diagnosed with primary PSD and 16 healthy controls. We selected a group with common phenotype characterized by impaired platelet aggregation and secretion with two or more stimuli as assessed with lumi-aggregometer and a normal platelet content of the granules, confirming the diagnosis of PSD. In our previous study, we demonstrated that PSD is the most abundant diagnosis with a prevalence of almost one fifth of patients with mild bleeding diathesis.[5]

To identify causal genes underlying these defects, we carried out two prioritizing approaches, which were based on the identification of rare, potentially deleterious variants present in 329 platelet candidate genes listed by Leo et al.[15] or by selecting singletons *(On line Supplementary Figure S1)*. These strategies revealed a number of plausible candidate gene defects explaining phenotypical defects of primary PSD. For instance, patient C740 carries a missense variant p.D1144N in the *STXBP5L* gene (Table 2). In a recent report, another missense variant was identified in this gene as potentially causal in platelet secretion abnormalities.[15] Since STXBP5, a paralog of STXBP5L, promotes platelet secretion,[34, 35] perhaps also STXBP5L may play a role in this process. Another interesting candidate is the *KCNMB3* gene that carries the p.P83H missense variant. This gene encodes the Calcium-Activated Potassium Channel Subunit Beta-3 protein involved in a pathway activated in response to elevated platelet cytosolic Ca^{2+}.

For patient C732, a gene defect was found in *EXOC1*, which is another candidate gene that influences platelet granule exocytosis. This gene encodes the Exocyst Complex Component 1 protein that functions as part of the exocyst complex and is required for targeting exocytic vesicles to specific docking sites on the plasma membrane.[36]

We also found a missense variant p.A464P in the *RAP1GAP* gene for patient C831. This variant has been classified as likely benign and for this reason, it was excluded from Table 2. Importantly, the Rap1GAP protein plays a regulatory role in platelet aggregation,[37] suggesting that this missense variant may have a functional role.

As previously reported, PSD can be associated with proteins acting at different levels: signal transduction, platelet activation, degranulation, or exocytosis.[4] Indeed, we found potential gene defects in proteins involved in all of these processes (Table 2). Importantly, several patients in our study had multiple defects in the above-mentioned genes and gene pathways, which may explain the complex and heterogeneous nature of primary PSD. This indicates that an in-depth functional analysis of platelet receptor and signaling pathways will be necessary to discriminate differences in clinical and laboratory phenotypes of affected individuals.

Study limitations

Following a positive experience with the application of WES to identify gene defects underlying inherited platelet function disorders, [19-22] we chose to investigate primary PSD using the same technique, hoping that genomic approach will be effective in identifying causal variants in a heterogeneous clinical and phenotypic such as primary PSD. However, exome sequencing followed

by two independent variant prioritization approaches yielded inconclusive results. The primary reason for this is undoubtedly the heterogeneous clinical and laboratory phenotype of primary PSD, which may have led to the identification of genes not necessarily associated with the disease. For instance, 20 missense variants were detected in the *TTN* gene in 11 PSD patients, of which eight are VUS. However, *TTN* is one of the most frequently mutated genes in the human genome,[38] implying that the variation found in this gene is probably to the size of its coding regions (363 exons).

Another limitation of this study was perhaps the choice of the variant prioritization strategy. We applied a generally accepted filtering method based on the selection of rare (MAF>1%), potentially damaging variants. This approach revealed great abundance of variants for most patients, which required further selection based on the ACMG pathogenic classification of SNVs (Table 2). This revealed 34 putative gene defects classified as VUS in 12 patients with primary PSD, of which 24 were located in genes expressed in human platelets according to the HPM (Table 2). However, it is possible that many potentially causal SNVs, which were classified as likely benign or benign, were excluded due to a lack of supporting evidence or that the gene defects may only manifest at the level of megakaryocyte development or platelet maturation.

In addition, some of the functional defects might have been located in the non-coding parts of the genome such as promoters, intronic sequences or enhancers, which were not covered by exome sequencing. Finally, since the identification of gross chromosomal aberration such as copy number variation (CNV) from the WES data remains a technical challenge, it is likely that these structural variants would not have been detected. Although several bioinformatics methods have been developed for CNV analysis from the WES data, they require uniform coverage and high resolution of the sequencing data across all exons/coding regions as well as specialized bioinformatics pipeline of data analysis validated against the whole-genome data.[39] For this reason, the whole-genome sequencing is the only sure means for identifying the CNVs alongside SNVs and small indels.

In conclusion, we carried out an exome sequencing in 14 patients with primary PSD and 16 healthy controls, followed by two variant prioritization strategies. Our analysis identified potential gene defects in 12 patients, implying that the NGS-based diagnostic strategies for causal gene identification in such heterogeneous clinical and laboratory phenotype as primary PSD may be ineffective. In this cases, a well-defined, common disease phenotyping and properly established pipeline for variant analysis are necessary. The difficulty in assigning causality can be overcome by

genetic screening of affected and unaffected family members, which allows the identification of gene defects that segregate with the clinical phenotype or by functional studies.

The perils of genetic data sharing with patients may involve ethical concerns, lack of confidence in assessing the causality of identified variants, and the implication of some inherited platelet pathologies with other risks.[40] For these reasons, sharing genetic data with patients is still an opened issue that requires further discussion.

Acknowledgements

The study was supported by Bayer Hemophilia Award Program (BHAP) 2011- Special Project Award to Flora Peyvandi and by BHAP 2012 - Early Career Award to Luca A. Lotta.

We thank prof. Pier Mannuccio Mannucci for critical revision of the manuscript.

Conflict of interest statement

Flora Peyvandi has received honoraria for participating as a speaker at satellite symposia and educational meetings organized by Ablynx, Grifols, Novo Nordisk, Roche, Shire and Sobi. She has received consulting fees from Kedrion and she is member of the scientific advisory board of Ablynx. Other authors have no conflict of interest to declare.

References

1. Cattaneo M. Inherited platelet-based bleeding disorders. J Thromb Haemost. 2003;1(7):1628-1636.

2. Hayward CP, Rao AK, Cattaneo M. Congenital platelet disorders: overview of their mechanisms, diagnostic evaluation and treatment. Haemophilia. 2006;12 (Suppl 3):128-136.

3. Hayward CP, Rivard GE. Quebec platelet disorder. Expert Rev Hematol. 2011;4(2):137-141.

4. Nurden A, Nurden P. Advances in our understanding of the molecular basis of disorders of platelet function. J Thromb Haemost. 2011;9 (Suppl 1):76-91.

5. Lotta L, Maino A, Tuana G, et al. Prevalence of disease and relationships between laboratory phenotype and bleeding severity in platelet primary secretion defects. PLoS One. 2013;8(4):e60396.

6. Rao AK, Jalagadugula G, Sun L. Inherited defects in platelet signaling mechanisms. Semin Thromb Hemost. 2004;30(5):525-535.

7. Quiroga T, Goycoolea M, Panes O, et al. High prevalence of bleeders of unknown cause among patients with inherited mucocutaneous bleeding. A prospective study of 280 patients and 299 controls. Haematologica. 2007;92(3):357-365.

8. Cattaneo M. Light transmission aggregometry and ATP release for the diagnostic assessment of platelet function. Sem Thromb Hemost. 2009;35(2):158-167.

9. Pai M, Wang G, Moffat KA, et al. Diagnostic usefulness of a lumi-aggregometer adenosine triphosphate release assay for the assessment of platelet function disorders. Am J Clin Pathol. 2011;136(3):350-358.

10. Cattaneo M. Molecular defects of the platelet P2 receptors. Purinergic Signal. 2011;7(3):333-339.

11. Bastida JM, Lozano ML, Benito R, et al. Introducing high-throughput sequencing into mainstream genetic diagnosis practice in inherited platelet disorders. Haematologica. 2018;103(1):148-162.

12. Peyvandi F, Hayward CP. Genomic approaches to bleeding disorders. Haemophilia. 2016;22 (Suppl 5):42-45.

13. 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.

14. 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;179(2):308-322.

15. Leo VC, Morgan NV, Bem D, 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):643-650.

16. Lozano ML, Cook A, Bastida JM, et al. Novel mutations in RASGRP2, which encodes CalDAG-GEFI, abrogate Rap1 activation, causing platelet dysfunction. Blood. 2016;128(9):1282-1289.

17. Sevivas T, Bastida JM, Paul DS, et al. Identification of two novel mutations in RASGRP2 affecting platelet CalDAG-GEFI expression and function in patients with bleeding diathesis. Platelets. 2018;29(2):192-195.

18. Maclachlan A, Dolan G, Grimley C, Watson SP, Morgan NV, On Behalf Of The Uk Gapp Study G. Whole exome sequencing identifies a mutation in thrombomodulin as the genetic cause of a suspected platelet disorder in a family with normal platelet function. Platelets. 2017;28(6):611-613.

19. Bariana TK, Ouwehand WH, Guerrero JA, Gomez K. Dawning of the age of genomics for platelet granule disorders: improving insight, diagnosis and management. Br J Haematol. 2017;176(5):705-720.

20. Heremans J, Freson K. High-throughput sequencing for diagnosing platelet disorders: lessons learned from exploring the causes of bleeding disorders. Int J Lab Hematol. 2018;40 (Suppl 1):89-96.

21. Lentaigne C, Freson K, Laffan MA, Turro E, Ouwehand W. Inherited platelet disorders: toward DNA-based diagnosis. Blood. 2016;127(23):2814-2823.

22. Westbury SK, Mumford AD. Genomics of platelet disorders. Haemophilia. 2016;22 (Suppl 5):20-24.

23. Lotta LA, Lombardi R, Mariani M, et al. Platelet reactive conformation and multimeric pattern of von Willebrand factor in acquired thrombotic thrombocytopenic purpura during acute disease and remission. J Thromb Haemost. 2011;9(9):1744-1751.

24. Tosetto A, Rodeghiero F, Castaman G, et al. A quantitative analysis of bleeding symptoms in type 1 von Willebrand disease: results from a multicenter European study (MCMDM-1 VWD). J Thromb Haemost. 2006;4(4):766-773.

25. Lotta LA, Wang M, Yu J, et al. Identification of genetic risk variants for deep vein thrombosis by multiplexed next-generation sequencing of 186 hemostatic/pro-inflammatory genes. BMC Med Genomics. 2012;5(7):1-8.

26. Cattaneo M, Cerletti C, Harrison P, et al. Recommendations for the Standardization of Light Transmission Aggregometry: A Consensus of the Working Party from the Platelet Physiology Subcommittee of SSC/ISTH. J Thromb Haemost. 2013;11(6):1183-1186.

27. Femia EA, Pugliano M, Podda G, Cattaneo M. Comparison of different procedures to prepare platelet-rich plasma for studies of platelet aggregation by light transmission aggregometry. Platelets. 2012;23(1):7-10.

28. Cattaneo M, Bettega D, Lombardi R, Lecchi A, Mannucci PM. Sustained correction of the bleeding time in an afibrinogenaemic patient after infusion of fresh frozen plasma. Br J Haematol. 1992;82(2):388-390.

29. Cattaneo M, Lecchi A, Lombardi R, Gachet C, Zighetti ML. Platelets from a patient heterozygous for the defect of P2CYC receptors for ADP have a secretion defect despite normal thromboxane A2 production and normal granule stores: further evidence that some cases of platelet 'primary secretion defect' are heterozygous for a defect of P2CYC receptors. Arterioscler Thromb Vasc Biol. 2000;20(11):101-106.

30. Li R, Yu C, Li Y, et al. SOAP2: an improved ultrafast tool for short read alignment. Bioinformatics. 2009;25(15):1966-1967.

31. Richards S, Aziz N, Bale S, et al. 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. Genet Med. 2015;17(5):405-424.

32. Kim MS, Pinto SM, Getnet D, et al. A draft map of the human proteome. Nature. 2014;509(7502):575-581.

33. Huang DW, Sherman BT, Tan Q, et al. DAVID Bioinformatics Resources: expanded annotation database and novel algorithms to better extract biology from large gene lists. Nucleic Acids Res. 2007;35(Web Server issue):W169-W175.

34. Ye S, Huang Y, Joshi S, et al. Platelet secretion and hemostasis require syntaxin-binding protein STXBP5. J Clin Invest. 2014;124(10):4517-4528.

35. Zhu Q, Yamakuchi M, Ture S, et al. Syntaxin-binding protein STXBP5 inhibits endothelial exocytosis and promotes platelet secretion. J Clin Invest. 2014;124(10):4503-4516.

36. Lipschutz JH, Mostov KE. Exocytosis: the many masters of the exocyst. Curr Biol. 2002;12(6):212-214.

37. Schultess J, Danielewski O, Smolenski AP. Rap1GAP2 is a new GTPase-activating protein of Rap1 expressed in human platelets. Blood. 2005;105(8):3185-3192.

38. Shyr C, Tarailo-Graovac M, Gottlieb M, Lee JJ, van KC, Wasserman WW. FLAGS, frequently mutated genes in public exomes. BMC Med Genomics. 2014;7(64):1-11

39. Eilbeck K, Quinlan A, Yandell M. Settling the score: variant prioritization and Mendelian disease. Nat Rev Genet. 2017;18(10):599-612.

40. Greinacher A, Eekels JJM. Diagnosis of hereditary platelet disorders in the era of next-generation sequencing:"primum non nocere". J Thromb Haemost. 2019;10.1111/jth.14377

41. Kircher M, Witten DM, Jain P, O'Roak BJ, Cooper GM, Shendure J. A general framework for estimating the relative pathogenicity of human genetic variants. Nat Genet. 2014;46(3):310-315.

Table 1. Clinical and biological characteristics of 14 unrelated PSD patients.

ID	Sex	Age	BSS	Parents consang.	First-grade family bleeding history	Platelet secretion stimulus				
						ADP 4 µM	ADP 20 µM	Collagen 2 µg/mL	U46619 1 µM	TRAP 10 µM
C696	F	34	4	no	No	+	-	-	+	-
C708	F	36	5	no	Mother (mild PSD)	+	+	+	+	-
C729	F	3	2	no	No	+	+	-	-	n/p
C732	F	25	15	no	No	+	+	-	+	-
C739	M	5	0	no	Mother	+	+	-	-	n/p
C740	M	19	10	no	Father (mild PSD), Sister (mild bleeding diathesis without PSD)	+	+	+	+	+
C749	F	55	9	no	Mother and sister with thrombocytopenia	+	+	+	-	+
C783	F	63	13	no	Mother	+	+	-	-	-
C797	F	31	5	no	No	+	+	-	+	-
C831	F	20	7	no	Mother	+	-	+	-	+
C847	F	19	7	no	Mother	+	+	-	+	+
C862	F	3	7	no	No	+	+	-	+	n/p
C1075	F	52	15	no	Brother	+	+	+	-	+
C1107	F	60	8	no	No	+	+	-	+	+

BSS – bleeding severity score; **(+)** indicates defective platelet secretion response to the stimulus; **(-)** indicates response within the normal range; **n/p –** data not present.

Table 2. Putative causal variants identified by WES in 12/14 patients with primary PSD according to Leo et al.[15] classification or by selecting singletons (*Online Supplementary Figure S1).*

All variants were heterozygous.

ID	Gene	Nucleotide change	dbSNP	Amino acid change	MAF 1000G	MAF ESP	MAF ExAC	SIFT	Poly phen2	Mutation Taster	CADD C score	Platelet expression (*)	Assess. (**)
C696	COL24A1	c.G4673A	-	p.G1558E	-	-	-	D	D	D	25	-	VUS
C708	TTN	c.G106955A	rs200497615	p.R35652Q	-	0.0007	0.0003	B	B	D	25	+	VUS
	CSRNP1	c.C673T	rs142034027	p.R225W	-	0.0007	0.0001	D	D	D	32	-	VUS
	NRP1	c.G620A	rs148308681	p.R207H	-	0.0001	0.0001	D	D	D	33	-	VUS
C729	TTN	c.C104564A	-	p.S34855Y	-	-	-	D	D	D	20	+	VUS
	ITGA2	c.G305A	rs41392746	p.S102N	-	-	3.01E-05	B	B	B	20	+	VUS
	MYO3A	c.T1525C	rs150793986	p.Y509H	-	0.0003	0.0002	D	D	D	27	-	VUS
	MUC2	c.G6931A	rs200823008	p.V2311I	-	0.0001	0.0008	-	-	-	-	-	VUS
C732	EXOC1	c.G2009A	rs35001804	p.G670E	0.003	0.0086	0.009	D	D	D	32	+	VUS
C739	DIAPH1	c.T3227G	rs143763573	p.F1076C	-	-	0.0001	D	D	D	26	+	VUS
	ITPR3	c.C5720T	-	p.T1907M	-	-	-	D	D	D	33	+	VUS
C740	TTN	c.C72358T	rs372309164	p.L24120F	-	0.0002	0	B	D	D	18	+	VUS
	TTN	c.G1895A	rs150231219	p.G632D	-	0.0002	0	D	B	B	19	+	VUS
	SLC2A7	c.C670T	rs35776221	p.R224C	0.006	0.01	0.008	D	D	D	27	-	VUS
	STXBP5L	c.G3430A	rs139176240	p.D1120N	-	0.0001	0.0001	B	D	D	25	+	VUS
	KCNMB3	c.C248A	rs61734056	p.P83H	-	1.50E-05	0.0001	D	D	D	27	-	VUS
	LCN1	c.G298C	rs117638349	p.G100R	0.006	0.008	0.004	D	D	B	23	-	VUS
	PRKACG	c.C280T	-	p.R94C	-	-	-	D	D	B	23	+	VUS
	MUC2	c.G2594A	-	p.S865N	-	-	-	-	-	-	-	-	VUS
	MUC2	c.A5038G	rs371137719	p.T1680A	0.01	0.0024	0	-	-	-	-	-	VUS
C749	LYST	c.G8806A	rs2753327	p.V2936I	0.001	0.0009	0.0009	B	B	D	22	-	VUS
	TTN	c.G49413T	rs202094100	p.W16471C	-	0.0008	0.0006	D	D	D	24	+	VUS
	PHF14	c.G298T	-	p.E100X	-	-	-	-	-	D	38	-	VUS
	PTPN12	c.C1066T	rs752211731	p.P356S	-	-	0	D	D	D	27	+	VUS

C797	TTN	c.C17T	rs201490999	p.P6L	-	-	-	D	D	D	24	+	VUS
	EGF	c.G3073A	-	p.A1025T	-	-	-	B	B	D	15	+	VUS
C831	TTN	c.T15768A	rs138826545	p.H5256Q	-	0.0002	0.0002	B	B	D	12	+	VUS
	EGF	c.G1723A	rs115396821	p.G575R	0.008	0.0024	0.0027	D	D	D	26	+	VUS
	TBXAS1	c.151_152del	-	p.V51fs	-	-	-	-	-	-	-	+	VUS
	VWF	c.G8171A	-	p.C2724Y	-	-	-	D	D	D	26	+	VUS
C847	TTN	c.C91384T	rs373623340	p.R30462W	-	-	3.01E-05	D	D	D	26	+	VUS
	PHACTR2	c.G1360C	-	p.D454H	-	-	-	D	D	D	26	+	VUS
	NOS3	c.C3385T	rs774447524	p.R1129C	-	-	2.31E-05	D	D	D	34	-	VUS
C1075	PRKCD	c.A1043G	rs33911937	p.N348S	-	0.0015	0.0016	B	B	D	15	+	VUS
C1107	PTPN7	c.G425A	rs115136927	p.R142Q	0.003	0.0072	0.0062	B	D	D	27	+	VUS
	PRKCD	c.G868T	-	p.A290S	-	-	-	B	D	D	25	+	VUS
	MMRN1	c.G3680T	rs147451161	p.R1227L	0.003	0.0031	0.0036	D	D	D	28	+	VUS

dbSNP – Database of Single Nucleotide Polymorphisms v.138. ***MAF –*** minor allele frequency (MAF from European populations is shown). ***1000G –*** the 1000 Genomes Project. ***ExAC –*** the Exome Aggregation Consortium. ***ESP –*** the Exome Sequencing Project. ***SIFT –*** Sorting Intolerant From Tolerant. ***PolyPhen2 –*** Polymorphism Phenotyping v2. ***Mutation Taster***, prediction scores: D **–** Damaging, B **–** Benign.
CADD C score – Combined Annotation Dependent Depletion score.[41] ***VUS –*** variant of uncertain significance.
(*) Platelet gene expression evaluated by the Human Proteome Map (HPM) (http://www.humanproteomemap.org);[32]
(**) Assess. – Assessment of variant pathogenicity assigned according to the American College of Medical Genetics and Genomics pathogenicity classification.[31]

Table 3. Putative causal variants identified by WES in the family of patient C740 (Online *Supplementary Figure S2)*.

Gene	dbSNP	Nucl. change	Amino acid change	C740	C1300	C1301	C1302	C1304	MAF 1000G	MAF ExAC	MAF ESP	SIFT	Poly phen2	Mutation Taster	CADD C score	PLT Exp. (*)	Assess (**)
SLC2A7	rs35776221	c.C670T	p.R224C	het	het	-	-	-	0.006	0.01	0.008	D	D	D	27	-	VUS
STXBP5L	rs139176240	c.G3430A	p.D1144N	het	het	-	-	-	-	0.0004	0.0001	B	D	D	25	+	VUS
KCNMB3	rs61734056	c.C248A	p.P83H	het	het	-	-	-	-	1.50E-05	0.0001	D	D	D	27	-	VUS
LCN1	rs117638349	c.G298C	p.G100R	het	het	-	-	-	0.006	0.008	0.004	D	D	B	23	-	VUS

dbSNP* –** Database of Single Nucleotide Polymorphisms v.138. ***MAF* –** minor allele frequency (MAF from European populations is shown). ***1000G* –** the 1000 Genomes Project. ***ExAC* –** the Exome Aggregation Consortium. ***ESP* –** the Exome Sequencing Project. ***SIFT* –** Sorting Intolerant From Tolerant. ***PolyPhen2* –** Polymorphism Phenotyping v2. ***Mutation Taster, prediction scores: D **–** Damaging, B **–** Benign.
***CADD C score* –** Combined Annotation Dependent Depletion score.[41] ***VUS* –** variant of uncertain significance.
(*) Platelet gene expression evaluated by the Human Proteome Map (HPM) (http://www.humanproteomemap.org);[32]
(**) Assess. – Assessment of variant pathogenicity assigned according to the American College of Medical Genetics and Genomics pathogenicity classification.[31]

Figure legends

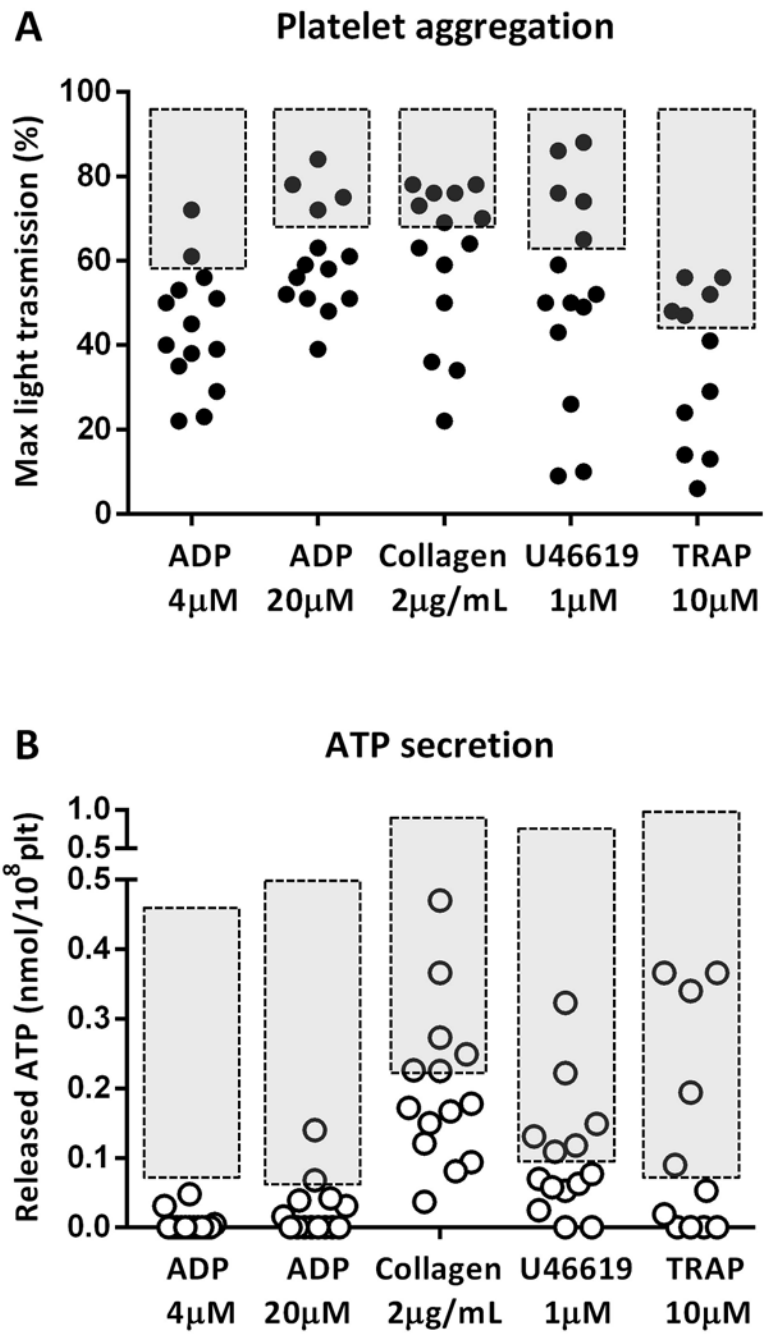
Figure 1. Dot plots of platelet aggregation (A) and secretion (B) of 14 unrelated PSD patients.

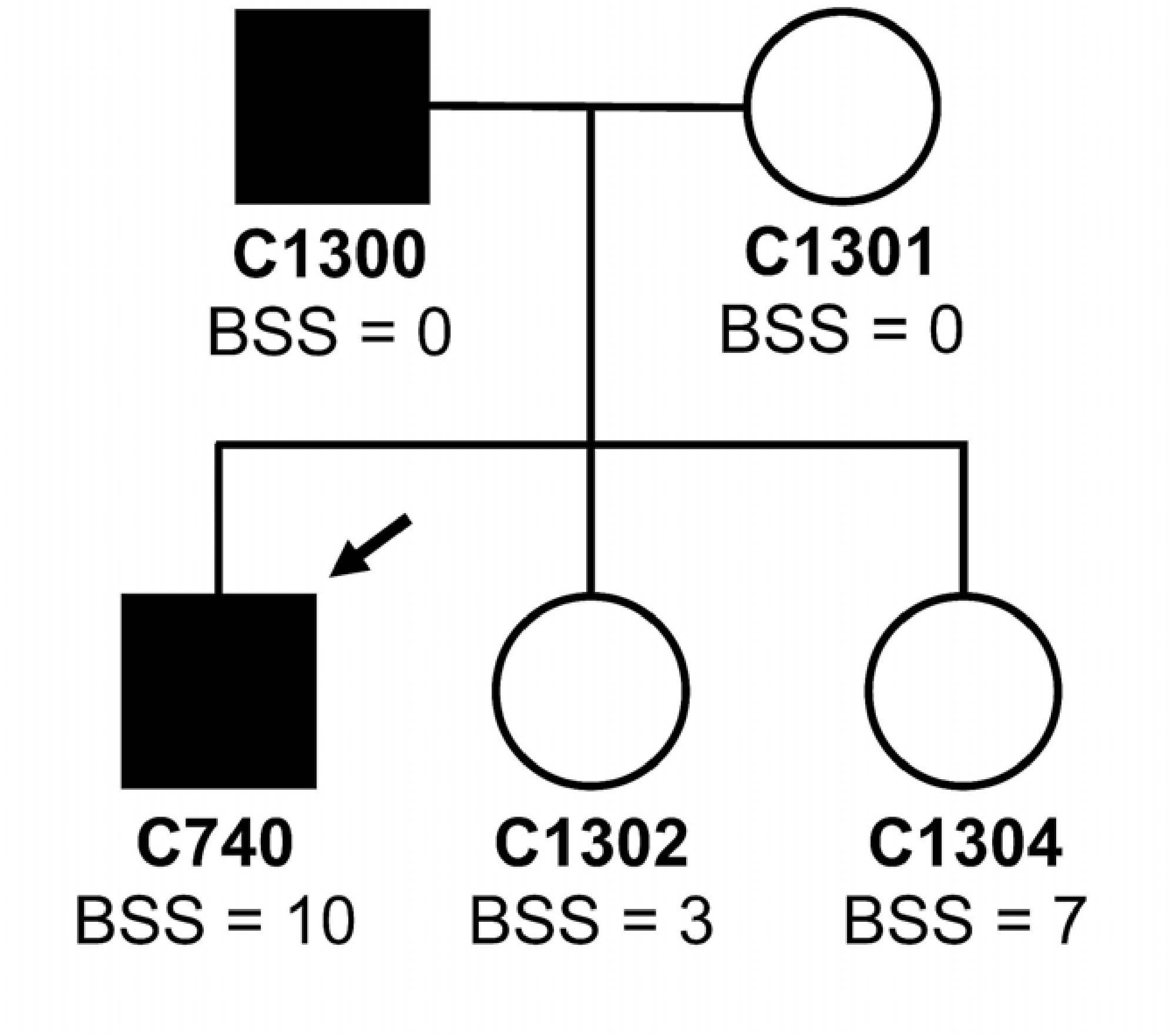
Figure 1 legend. Boxes indicate our internal range of normality (5th -95thpercentiles).

Figure 2. Pedigree of patient C740.

Figure 2 legend. Black and white symbols indicate affected by PSD and unaffected family members, respectively. The arrow indicates the proband C740.

BSS – bleeding severity score


A
Platelet aggregation
Max light trasmission (%)
100
80
60
40
20
0
ADP 4μM
ADP 20μM
Collagen 2μg/mL
U46619 1μM
TRAP 10μM
B
ATP secretion
Released ATP (nmol/10^8 plt)
1.0
0.5
0.5
0.4
0.3
0.2
0.1
0.0
ADP 4μM
ADP 20μM
Collagen 2μg/mL
U46619 1μM
TRAP 10μM


C1300
BSS = 0
C1301
BSS = 0
C740
BSS = 10
C1302
BSS = 3
C1304
BSS = 7

Complications of whole-exome sequencing for causal gene discovery in primary platelet secretion defects

Marcin M. Gorski et al.

Supplementary tables

Supplementary figures

Supplementary methods

Supplementary References

Supplementary Tables

Table S1. ADP, ATP, serotonin, and fibrinogen platelet content.

	PSD patients N=14*	**Internal reference range****
Delta granules		
ADP (nmoles/10^8 platelets)	2.26 (1.19-4.15)	1.30-2.88
ATP (nmoles/10^8 platelets)	5.00 (3.20-9.40)	3.17-7.07
ATP/ADP	2.42 (1.83-2.84)	1.55-3.42
Serotonin (nmoles/10^8 platelets)	0.37 (0.25-0.64)	0.19-0.40
Alpha granules		
Fibrinogen (mg/10^9 platelets)	0.06 (0.04-0.13)	0.03-0.19

*****Median (min-max)

****** (5^{th}-95^{th} percentiles)

Table S2. Single-nucleotide variants (n=107) identified in 14 PSD patients by WES followed by two prioritizing approaches, according to Leo et al.[1] classification or by selecting singletons.

All variants were heterozygous. Variant filtering steps are reported in Figure S1.

ID	Gene	Nucleotide change	dbSNP	Amino acid change	MAF 1000G EUR	MAF ESP EA	ExAC NFE	SIFT	Polyphen2	Mutation Taster	CADD C-score	Platelet expression	Leo et al (JTH, 2015)	ACMG
C696	COL24A1	c.G4673A	.	p.G1558E	.	.	.	D	D	D	25.2	-	-	VUS
C696	LTBP1	c.G3011A	rs141080282	p.R1004Q	0.005	0.0067	0.0059	D	P	D	24.2	+	+	LB
C696	PLEK	c.A322C	rs34515106	p.K108Q	.	0.0007	0.0016	T	P	D	22.4	+	-	LB
C696	MERTK	c.A2305G	rs147899488	p.I769V	.	0.0001	0.0004	T	B	D	16.37	-	-	LB
C696	TUBA3D	c.G331A	rs550660894	p.G111S	.	.	4.5E-05	.	D	D	27.5	-	-	LB
C696	TTN	c.T99179C	rs763888823	p.I33060T	.	.	.	T	B	D	20.7	+	-	LB
C696	TTN	c.G63309T	.	p.M21103I	.	.	.	T	B	D	19.1	+	-	LB
C696	TTN	c.A24973G	rs72648984	p.K8325E	0.008	0.0076	0.0093	T	B	D	13.41	+	-	LB
C696	TTN	c.A15563C	rs72648930	p.Q5188P	0.001	0.0021	0.0015	T	D	D	15.15	+	-	LB
C696	CSRNP1	c.G401A	rs757921966	p.R134H	.	.	1.5E-05	T	B	N	23.5	-	-	LB
C696	MMRN1	c.G1546T	rs141872900	p.V516L	0.007	0.0084	0.0073	T	B	N	0.575	+	-	LB
C696	DGKI	c.G553A	rs779164061	p.V185I	.	.	0.000015	D	P	D	23	+	-	LB
C708	QSOX1	c.G1060A	rs148353050	p.V354M	.	0.0008	0.0003	T	P	N	10.24	+	-	LB
C708	TTN	c.G106955A	rs200497615	p.R35652Q	.	0.0007	0.0003	T	B	D	24.6	+	-	VUS
C708	TTN	c.G97760A	rs55704830	p.R32587H	0.003	0.0038	0.0056	T	D	D	25.5	+	-	LB
C708	SERPINE2	c.G622C	rs375757013	p.V208L	.	0.0002	1.5E-05	T	B	N	0.135	+	+	LB
C708	COL4A4	c.G2630A	rs150979437	p.R877Q	.	0.0033	0.0038	T	B	N	9.424	-	-	LB
C708	ITPR1	c.C5098T	rs540818757	p.P1700S	.	.	.	T	.	D	9.375	+	+	LB

C708	CSRNP1	c.C1389G	.	p.S463R	.	.	.	D	D	N	25.9	-	-	LB
C708	CSRNP1	c.C673T	rs142034027	p.R225W	.	0.0007	0.0001	D	D	D	32	-	-	VUS
C708	MYLK4	c.A1286G	rs35211631	p.Q429R	.	0.0021	0.0011	T	B	N	15.14	-	-	B
C708	PLG	c.T2045A	rs147175166	p.I682N	0.001	0.0008	0.0009	T	D	D	24.8	+	-	LB
C708	NRP1	c.G620A	rs148308681	p.R207H	.	0.0001	0.0001	D	D	D	33	-	-	VUS
C729	FCGR2A	c.A836C	rs146883516	p.D279A	0.002	0.0017	0.002	D	B	N	22.6	+	+	LB
C729	TTN	c.C104564A	.	p.S34855Y	.	.	.	D	D	D	19.45	+	-	VUS
C729	TTN	c.A13364G	rs142304137	p.K4455R	0.002	0.0001	0.0003	D	P	D	8.924	+	-	LB
C729	ITGA2	c.G305A	rs41392746	p.S102N	.	.	3.01E-05	T	B	N	19.85	+	+	LB
C729	ITPR3	c.G2056A	.	p.E686K	.	.	.	T	P	D	23.7	+	-	LB
C729	MYO3A	c.T1525C	rs150793986	p.Y509H	.	0.0003	0.0002	D	D	D	27.4	-	+	VUS
C729	MUC2	c.G6931A	rs200823008	p.V2311I	.	0.0001	0.0008	.	.	.	.	-	+	VUS
C729	ARHGAP1	c.C787T	rs144801476	p.L263F	0.006	0.01	0.0076	D	P	D	23.3	+	+	LB
C732	COL24A1	c.C314T	rs372813075	p.P105L	.	0.0001	1.5E-05	T	P	D	13.79	-	-	LB
C732	LEFTY2	c.A613G	rs770500519	p.T205A	.	.	3.22E-05	D	P	D	23.5	-	-	LB
C732	ITPR1	c.C4236G	rs61757110	p.H1412Q	0.003	0.0015	0.0012	D	D	D	22.9	+	+	LB
C732	EXOC1	c.G2009A	rs35001804	p.G670E	0.003	0.0086	0.0086	D	D	D	32	+	+	VUS
C732	AP3S1	c.A368G	rs199536113	p.N123S	.	.	0.0005	T	B	D	13.77	+	+	LB
C732	BRPF3	c.A3055G	rs145016452	p.S1019G	0.001	0.0031	0.003	T	B	D	17.5	-	-	LB
C732	PLG	c.T1380A	rs116573785	p.S460R	0.001	0.0017	0.0027	T	B	N	7.855	+	-	LB
C732	GNB2	c.A367G	rs771355621	p.I123V	.	.	1.53E-05	T	B	D	11.45	+	+	LB
C739	RAP1GAP	c.A1904G	rs147394161	p.Y635C	0.0099	0.013	0.014	T	P	D	27.8	+	+	B
C739	ABCG5	c.A1567G	rs140899003	p.I523V	.	0.0024	0.002	T	B	N	0.001	-	+	LB
C739	TTN	c.G21202C	.	p.A7068P	.	.	.	T	B	D	20.5	+	-	LB
C739	DGKQ	c.C2596G	rs376714052	p.R866G	.	0.0001	3.37E-05	T	B	D	27.7	-	-	LB

C739	APC	c.A398G	.	p.Y133C	.	.	.	D	D	D	23.9	-	+	LB
C739	DIAPH1	c.T3227G	rs143763573	p.F1076C	.	.	0.0001	D	D	D	26	+	+	VUS
C739	ITPR3	c.C5720T	.	p.T1907M	.	.	.	D	D	D	33	+	-	VUS
C740	TTN	c.C72358T	rs372309164	p.L24120F	.	0.0002	0	T	D	D	17.65	+	-	VUS
C740	TTN	c.G1895A	rs150231219	p.G632D	.	0.0002	0	D	B	N	18.85	+	-	VUS
C740	STXBP5L	c.G3430A	rs139176240	p.D1120N	.	0.0001	0	T	D	D	25.1	+	+	VUS
C740	SLC2A7	c.C670T	rs35776221	p.R224C	0.006	0.01	0.008	D	D	D	27	-	-	VUS
C740	LCN1	c.G298C	rs117638349	p.G100R	0.006	0.008	0.004	D	D	B	23	-	-	VUS
C740	APC	c.C6821T	rs34919187	p.A2274V	.	0.0015	0	T	B	N	16.24	-	+	LB
C740	DNAH11	c.A9935T	rs72657389	p.D3312V	0.008	0.004	0	T	P	D	23.6	-	+	LB
C740	PRKACG	c.C280T	.	p.R94C	.	.	.	D	D	N	22.6	+	+	VUS
C740	ADRA2A	c.G116A	rs539511086	p.R39Q	.	.	0	D	B	N	22.6	+	+	LB
C740	MUC2	c.G2594A	.	p.S865N	.	.	.	.	.	.	.	-	+	VUS
C740	MUC2	c.A5038G	rs371137719	p.T1680A	0.0099	0.0024	0	.	.	.	.	-	+	VUS
C749	F5	c.C3438G	rs6005	p.H1146Q	.	0.0003	6E-05	D	P	N	1.962	+	-	LB
C749	LYST	c.G8806A	rs2753327	p.V2936I	0.001	0.0009	0.0009	T	B	D	22	-	+	VUS
C749	LYST	c.A8224C	rs766760874	p.M2742L	.	.	1.51E-05	T	B	N	16.27	-	+	LB
C749	TTN	c.G49413T	rs202094100	p.W16471C	.	0.0008	0.0006	D	D	D	23.5	+	-	VUS
C749	COL4A3	c.T4421C	rs200302125	p.L1474P	0.003	0.0041	0.0046	D	D	D	23.4	-	-	LB
C749	DGKG	c.T1524G	.	p.F508L	.	.	.	T	B	N	0.172	+	-	LB
C749	PDGFC	c.A113G	rs139145392	p.Q38R	0.008	0.0066	0.007	T	B	D	0.016	+	-	LB
C749	CSF1R	c.T2876C	.	p.I959T	.	.	.	T	B	N	0.001	-	-	LB
C749	PHF14	c.G298T	.	p.E100X	.	.	.	.	.	D	38	-	-	VUS
C749	PTPN12	c.C1066T	rs752211731	p.P356S	.	.	0	D	D	D	27	+	+	VUS
C783	PLAT	c.G1481C	rs61755432	p.G494A	.	0.0007	0.001	T	D	D	23.4	-	-	LB

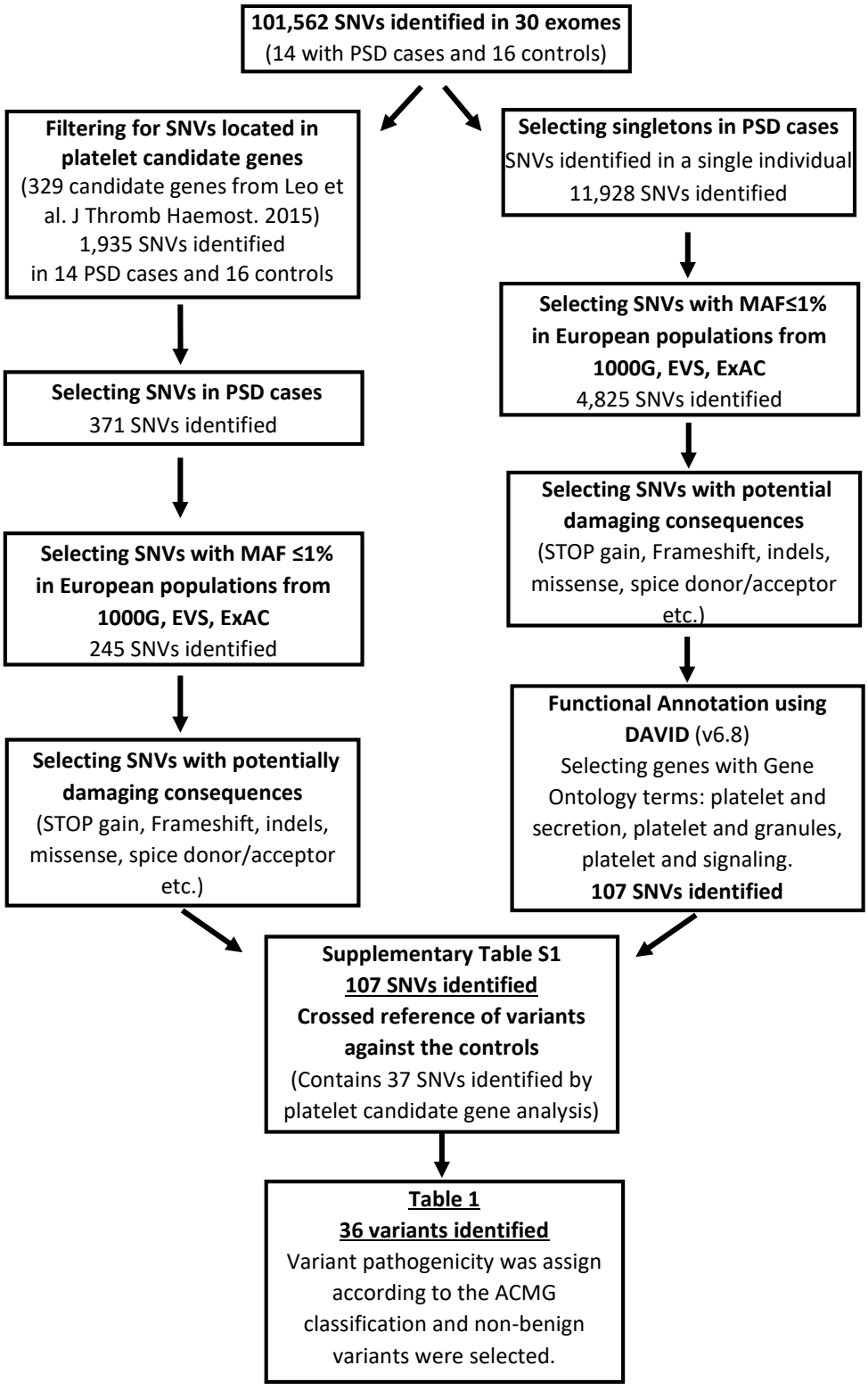
C797	TTN	c.C88394T	rs146181116	p.S29465F	0.007	0.0045	0.0039	T	D	D	22.5	+	-	LB
C797	TTN	c.T62996G	.	p.F20999C	.	.	.	T	P	N	2.844	+	-	LB
C797	TTN	c.C17T	rs201490999	p.P6L	.	.	.	D	D	D	23.8	+	-	VUS
C797	EGF	c.G3073A	.	p.A1025T	.	.	.	T	B	D	15.3	+	-	VUS
C797	PDGFRB	c.G946A	rs41287112	p.V316M	0.003	0.0046	0.0088	T	B	N	14.92	-	-	LB
C831	PRKCZ	c.G1109A	rs147033679	p.R370K	.	.	1.53E-05	T	B	D	7.976	-	-	LB
C831	RAP1GAP	c.G1390C	.	p.A464P	.	.	.	T	B	N	23.1	+	+	LB
C831	WNT3A	c.G527A	rs779729203	p.R176Q	.	.	1.58E-05	T	D	D	29.1	-	-	LB
C831	TTN	c.T15768A	rs138826545	p.H5256Q	.	0.0002	0.0002	T	B	D	11.85	+	-	VUS
C831	FN1	c.A751T	rs55822567	p.N251Y	0.001	0.0017	0.0026	D	P	N	23.7	+	-	LB
C831	FARP2	c.G1552A	rs746757859	p.G518R	.	.	9.36E-05	T	P	N	7.731	-	+	LB
C831	MMRN1	c.A3251G	rs201761344	p.N1084S	.	.	0.0001	D	P	N	22.8	+	-	LB
C831	EGF	c.G1723A	rs115396821	p.G575R	0.008	0.0024	0.0027	D	D	D	26	+	-	VUS
C831	PHF14	c.G2431A	rs61996285	p.V811I	0.003	0.0016	0.0017	T	B	D	19.69	-	-	LB
C831	DNAH11	c.A4282G	rs72657315	p.T1428A	0.002	0.0028	0.0043	D	B	N	22.3	-	+	LB
C831	DGKI	c.C457T	rs61757580	p.L153F	0.0099	0.0073	0.0078	T	P	D	14.82	+	-	B
C831	TBXAS1	c.151_152del	.	p.V51fs	.	.	.	.	.	.	.	+	-	VUS
C831	VWF	c.G8171A	.	p.C2724Y	.	.	.	D	D	D	26	+	+	VUS
C847	CASP9	c.A220G	rs145118493	p.M74V	.	0.0014	0.0013	T	B	N	7.542	+	-	LB
C847	F5	c.G43A	rs9332485	p.G15S	0.001	0.0002	0.0006	D	D	D	29.2	+	-	LB
C847	TTN	c.C91384T	rs373623340	p.R30462W	.	.	3.01E-05	D	D	D	25.7	+	-	VUS
C847	TTC37	c.C3253G	rs202214985	p.Q1085E	.	0.0001	0.0002	T	B	D	10.55	+	+	LB
C847	APC	c.G3949C	rs1801166	p.E1317Q	0.006	0.0093	0.0057	T	B	A	7.737	-	+	LB
C847	F13A1	c.G1861T	rs145180358	p.A621S	.	.	0.0007	T	B	D	23.8	+	-	LB
C847	PHACTR2	c.G1360C	.	p.D454H	.	.	.	D	D	D	25.8	+	-	VUS

C847	NOS3	c.C3385T	rs774447524	p.R1129C	.	.	2.31E-05	D	D	D	34	-	-	VUS
C847	PDGFRL	c.C1046A	rs146087994	p.T349K	.	.	1.5E-05	T	D	D	27.8	-	-	LB
C862	APC	c.C3511T	rs201830995	p.R1171C	0.001	0.0002	0.0003	D	B	N	24.1	-	+	LB
C1075	TTN	c.A53717G	rs727503606	p.K17906R	.	.	0.000015	T	B	N	7.856	+	-	B
C1075	TTN	c.T14477G	.	p.L4826R	.	.	.	D	P	N	1.837	+	-	LB
C1075	PRKCD	c.A1043G	rs33911937	p.N348S	.	0.0015	0.0016	T	B	D	15.06	+	+	VUS
C1075	STX11	c.G799A	rs45574234	p.V267M	0.0089	0.0092	0.0079	D	D	D	24	+	+	LB
C1107	COL11A1	c.G3847T	rs150669855	p.V1283L	0.001	0.0014	0.0013	T	B	N	0.012	-	-	LB
C1107	PTPN7	c.G425A	rs115136927	p.R142Q	0.003	0.0072	0.0062	T	D	D	26.9	+	+	VUS
C1107	TTN	c.T40931C	rs770248490	p.V13644A	.	.	1.57E-05	T	B	N	17.24	+	-	LB
C1107	PRKCD	c.G868T	.	p.A290S	.	.	.	T	D	D	24.5	+	+	VUS
C1107	MMRN1	c.G3680T	rs147451161	p.R1227L	0.003	0.0031	0.0036	D	D	D	27.8	+	-	VUS
C1107	ADCY2	c.C3167T	rs779183904	p.T1056M	.	.	6E-05	T	B	N	18.2	-	-	LB

dbSNP – Database of Single Nucleotide Polymorphisms v.138. ***MAF*** – Minor allele frequency (MAF from European populations are shown). ***1000G*** – the 1000 Genomes Project phase 3 populations. ***ESP*** – the Exome Sequencing Project; ***ExAC*** – the Exome Aggregation Consortium; ***SIFT*** – Sorting Intolerant From Tolerant; ***PolyPhen2*** – Polymorphism Phenotyping v2; ***Mutation Taster***: prediction scores: D – Damaging, B – Benign; ***CADD C score*** – Combined Annotation Dependent Depletion score; **VUS** – variant of uncertain significance; **LB** – likely benign.

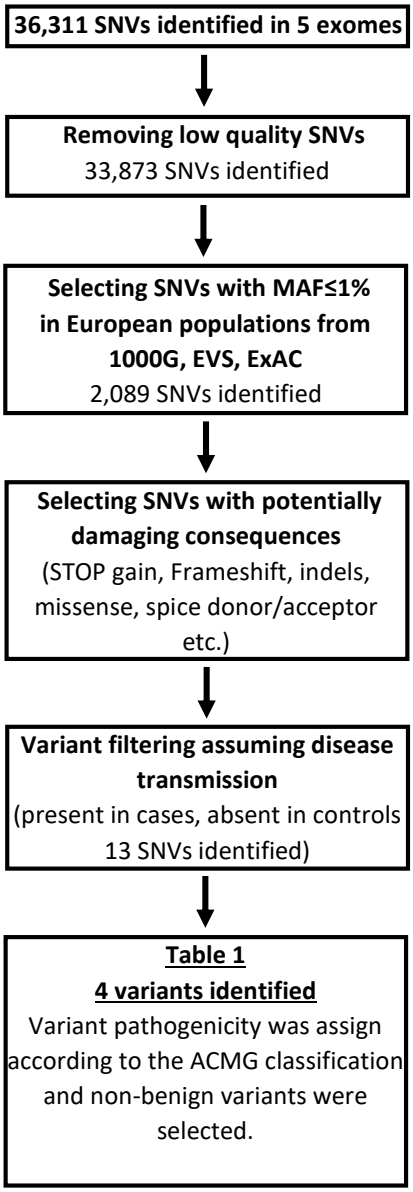
Supplementary Figures

Figure S1. Filtering steps for single nucleotide variants (SNVs) identified by WES in 14 PSD patients and 16 healthy controls.



SNV – Single Nucleotide Variant; ***MAF*** – Minor Allele Frequency; ***1000G*** – the 1000 Genomes Project; ***EVS*** – the Exome Variant Server; ***ExAC*** – the Exome Aggregation Consortium; ***DAVID*** – the Database for Annotation, Visualization and Integrated Discovery; ***ACMG*** – the American College of Medical Genetics and Genomics.

Figure S2. Filtering steps for SNVs identified by WES in four family members of PSD patient C740.



SNV – Single Nucleotide Variant; ***MAF*** – Minor Allele Frequency; ***1000G*** – the 1000 Genomes Project; ***EVS*** – the Exome Variant Server; ***ExAC*** – the Exome Aggregation Consortium; ***DAVID*** – the Database for Annotation, Visualization and Integrated Discovery; ***SIFT*** – Sorting Intolerant From Tolerant;[2] ***PolyPhen2*** – Polymorphism Phenotyping v2.[3] ***Mutation Taster*** (www.mutationtaster.org); ***CADD*** C- score – Combined Annotation Dependent Depletion score.[4]

Supplementary Methods

Materials

Adenosine diphosphate (ADP), adenosine triphosphate (ATP), thromboxane/prostaglandin endoperoxide analogue 9,11-dideoxy-11,9-epoxymethano-prostaglandin F2 (U46619), thrombin receptor activating peptide (TRAP; Ser-Phe-Leu-Leu-Arg-Asn-Pro-Asn-Asp-Lys-Tyr-Glu- Pro-Phe) were from Sigma Aldrich (St. Louis, MO, USA). Horm collagen was from Mascia Brunelli (Milano, IT). Commercial preparations of luciferin/luciferase reagent and protein kinase (Roche Diagnostic, Monza, IT) were used to measure the platelet ATP and ADP contents (ATP Assay Kit, Promega Italia, Milano, IT).

Commercial preparations of luciferin/luciferase (Chrono-lume; Chrono-log Corp, Havertown, PA, USA) were used to measure the platelet ATP released concurrently with platelet aggregation.

Blood sampling

Blood samples were drawn and 3 mL of blood were collected into commercial K-EDTA tubes for complete blood count analysis (ABX Micros 60, Horiba, Milano, IT). Platelet rich plasma (PRP) was prepared from trisodium citrate (129 mM, 1/9 $_{v/v}$) anticoagulated whole blood samples by centrifugation at 200 x g at room temperature for 15 min.[5, 6] Platelet poor plasma (PPP) was obtained by centrifugation at 1400 x g at room temperature for 15 min of samples from which PRP had been removed. Native platelet count of PRP was not modified.[7]

Platelet aggregation and secretion by lumiaggregometry

Platelet aggregation was measured in a lumi-aggregometer (Chrono-log, 560, Mascia Brunelli, Milano, IT) according to International Society on Thrombosis and Haemostasis recommendations.[5] ATP secretion from platelet dense granules was assessed simultaneously with aggregation by using the luciferase/Luciferin reagent (Chrono-lume) added to the PRP Secreted ATP levels were calculated by measuring the maximal amplitude of luminescence during the aggregation. Results were expressed as maximal increase (%) in light transmission for platelet aggregation and in ATP nmoli/10^8 plt for secretion within 3 minutes after platelet stimulation with the agonists: ADP (4 and 20 µM), collagen (2 µg/mL), thrombin receptor activator peptide (TRAP)-14 (10 µM), and thromboxane A2 analogue U46619 (1 µM).

Measurement of adenine nucleotides, serotonin and fibrinogen platelet content

Total platelet ADP and ATP content was measured with a luminometer (LKB 1250, Bio-Orbit Oy, Turku, Finland) by the firefly luciferin/luciferase method.[8] Platelet serotonin (5-HT) content was measured by the o-phthaldialdehyde method.[8] Fibrinogen was measured in washed platelets by a home-made enzyme-linked immunosorbent assay, using a polyclonal anti-fibrinogen antibody as previously reported.[8]

Whole-exome sequencing and variant annotation

Details of DNA extraction and preparation methods have been described elsewhere.[9]

Following variant alignment and calling, variants not meeting the following quality control criteria were removed: variants with more than 3 mismatches, variants-to-read ratio >0.1, variant reads mapping to single strand, total coverage <10 and Qual >30.

Next, variants were annotated onto dbSNPvs138,[10] ClinVar,[11] Sorting Intolerant FromTolerant (SIFT),[2] Polymorphism Phenotyping v2 (Polyphen-2),[3] Mutation Taster,[12] and the Combined Annotation Dependent Depletion (CADD).[4] Minor allele frequencies (MAFs) were obtained from the Exome Variant Server (EVS); (http://evs.gs.washington.edu/EVS/), the 1000 Genomes Project phase 3 populations (1KG)[13] and the Exome Aggregation Consortium (ExAC).[14] In addition, functional annotation of each variant identifying synonymous, non-synonymous, intronic, and spice region variants etc. was performed using the Variant Effect Predictor.[15]

Variant filtering and candidate gene discovery

Exome sequencing of healthy controls was carried out to perform analysis-by-exclusion, which involves prioritizing of rare variants with potential damaging consequences henceforth referred to as deleterious (e.g. missense, STOP gain/loss, insertions/deletions [indels], exon-intron boundaries) that are present exclusively in PSD patients, assuming that if present in controls, by definition, they could not be causal. All variant filtering steps were carried out using VCFtools.[16]

To select singletons, we filtered for private variants in PSD patients, followed by the selection of rare variants with minor allele frequency (MAF) ≤1% in the European populations from the 1KG, EVS and ExAC. Rare variants were further filtered by selecting those with putative functional consequences henceforth referred to as deleterious as described above. Next, functional annotation analysis was carried out using the Database for Annotation, Visualization and Integrated Discovery (DAVID v.6.8; www.david.ncifcrf.gov),[17] which allowed enrichment of genes carrying the Gene Ontology (GO) terms such as platelet secretion and signalling in biological

process, cellular component and molecular function, followed by identification of relevant annotation categories. Statistical significance of annotation terms was based on a DAVID Expression Analysis Systematic Explorer Score, which is based on a Modified Fisher Exact test. Gene clusters were considered significant with a Bonferroni $P<0.05$. GO terms such as platelets and secretion, platelets and granules, and platelets and signaling were used to select potential candidate genes.

The candidate platelet gene analysis was performed exploiting a list of 329 putative genes affected in individuals with platelet function disorders previously described.[1] In this part of analysis, we selected all variants present in the coding regions, 100 base pairs (bp) of 5' and 3' untranslated regions and 10 bp exon-intron boundaries of the 329 candidate genes in PSD cases. Rare variants were selected on the bases of MAF ≤1% followed by selection of putatively deleterious variants as described above.

Variants identified in both filtering strategies were pulled together in one table and cross-referenced against the controls and only SNVs present in PSD patients were selected. Supporting information was gathered using the UniProt Consortium[18] and the ClinVar (www.ncbi.nlm.nih.gov/clinvar/).

Sanger sequencing was performed to confirm NGS results.

Supplementary References

1. Leo VC, Morgan NV, Bem D, 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):643-650.
2. Ng PC, Henikoff S. SIFT: Predicting amino acid changes that affect protein function. Nucleic Acids Res. 2003;31(13):3812-3814.
3. Adzhubei I, Jordan DM, Sunyaev SR. Predicting functional effect of human missense mutations using PolyPhen-2. Curr Protoc Hum Genet. 2013; 76(1):7.20.01-41.
4. Kircher M, Witten DM, Jain P, O'Roak BJ, Cooper GM, Shendure J. A general framework for estimating the relative pathogenicity of human genetic variants. Nature genetics. 2014;46(3):310-315.
5. Cattaneo M, Cerletti C, Harrison P, et al. Recommendations for the Standardization of Light Transmission Aggregometry: A Consensus of the Working Party from the Platelet Physiology Subcommittee of SSC/ISTH. J Thromb Haemost. 2013;11(6):1183-1186.
6. Femia EA, Pugliano M, Podda G, Cattaneo M. Comparison of different procedures to prepare platelet-rich plasma for studies of platelet aggregation by light transmission aggregometry. Platelets. 2012;23(1):7-10.
7. Cattaneo M, Lecchi A, Zighetti M, Lussana F. Platelet aggregation studies: autologous platelet-poor plasma inhibits platelet aggregation when added to platelet-rich plasma to normalize platelet count. Haematologica. 2007;92(5):694-697.
8. Cattaneo M, Lecchi A, Lombardi R, Gachet C, Zighetti ML. Platelets from a patient heterozygous for the defect of P2CYC receptors for ADP have a secretion defect despite normal thromboxane A2 production and normal granule stores: further evidence that some cases of platelet 'primary secretion defect' are heterozygous for a defect of P2CYC receptors. Arterioscler Thromb Vasc Biol. 2000;20(11):101-106.
9. Lotta LA, Wang M, Yu J, et al. Identification of genetic risk variants for deep vein thrombosis by multiplexed next-generation sequencing of 186 hemostatic/pro-inflammatory genes. BMC Med Genomics. 2012;5(7):1-8.
10. Sherry ST, Ward M, Sirotkin K. dbSNP-database for single nucleotide polymorphisms and other classes of minor genetic variation. Genome Res. 1999;9(8):677-679.

11. Landrum MJ, Lee JM, Riley GR, et al. ClinVar: public archive of relationships among sequence variation and human phenotype. Nucleic Acids Res. 2014;42(Database issue):D980-D985.

12. Schwarz JM, Cooper DN, Schuelke M, Seelow D. MutationTaster2: mutation prediction for the deep-sequencing age. Nat Met. 2014;11(4):361-362.

13. Abecasis GR, Auton A, Brooks LD, et al. An integrated map of genetic variation from 1,092 human genomes. Nature. 2012;491(7422):56-65.

14. Lek M, Karczewski KJ, Minikel EV, et al. Analysis of protein-coding genetic variation in 60,706 humans. Nature. 2016;536(7616):285-291.

15. McLaren W, Pritchard B, Rios D, Chen Y, Flicek P, Cunningham F. Deriving the consequences of genomic variants with the Ensembl API and SNP Effect Predictor. Bioinformatics. 2010;26(16):2069-2070.

16. Danecek P, Auton A, Abecasis G, et al. The variant call format and VCFtools. Bioinformatics. 2011;27(15):2156-2158.

17. Huang DW, Sherman BT, Tan Q, et al. The DAVID Gene Functional Classification Tool: a novel biological module-centric algorithm to functionally analyze large gene lists. Genome Biol. 2007;8(9):183.181-183.116.

18. Pundir S, Martin MJ, O'Donovan C. UniProt Protein Knowledgebase. Methods Mol Biol. 2017;1558(2):41-55.