A Sardinian founder mutation in GP1BB that impacts thrombocytopenia.

Authors
Fabio Busonero[1,6,*], Maristella Steri[1,6,*], Valeria Orrù[1], Gabriella Sole[1], Stefania Olla[1], Michele Marongiu[1], Andrea Maschio[1], Carlo Sidore[1], Sandra Lai[1], Antonella Mulas[1], Magdalena Zoledziewska[1], Matteo Floris[1,2], Mauro Pala[1], Paola Forabosco[1], Isadora Asunis[1], Maristella Pitzalis[1], Francesca Deidda[1], Marco Masala[1], Cristian Antonio Caria[1], Susanna Barella[3], Goncalo R. Abecasis[4], David Schlessinger[5], Serena Sanna[1], Edoardo Fiorillo[1], Francesco Cucca[1,2,*].

Affiliations
[1]Istituto di Ricerca Genetica e Biomedica, Consiglio Nazionale delle Ricerche (CNR), 09042 Monserrato (Cagliari), Italy. [2]Dipartimento di Scienze Biomediche, Università degli Studi di Sassari, 07100 Sassari, Italy. [3]Ospedale Pediatrico Microcitemico "Antonio Cao" (A.O.Brotzu), 09100 Cagliari, Italy. [4]Center for Statistical Genetics, University of Michigan, Ann Arbor, 48109 Michigan, USA. [5]Laboratory of Genetics and Genomics, National Institute on Aging, US National Institutes of Health, Baltimore, 21224 Maryland, USA. [6]These authors contributed equally to this work.

Correspondence emails
*Correspondence should be addressed to Francesco Cucca (fcucca@uniss.it), Fabio Busonero (fabio.busonero@irgb.cnr.it) or Maristella Steri (maristella.steri@irgb.cnr.it).

Inherited platelets disorders can be severe, especially after trauma or surgical procedures in some monogenic disorders, as in Bernard-Soulier syndrome (BSS; MIM #231200). BSS is a rare autosomal recessive macro-thrombocytopenia (incidence of about 1 per million). Its hallmark is a defective adhesion of platelets to the sub-endothelium, resulting from quantitative or qualitative defects in the GPIb-IX-V complex, a platelet receptor for von Willebrand Factor (vWF), which is composed of four subunits: GPIbα, GPIbβ, GPIX, and GPV[1]. Laboratory diagnosis is based on prolonged bleeding time, moderate-to-severe thrombocytopenia (platelet count typically ranges from 20 to 100×10^9/L), giant platelets and deficient ristocetin-dependent platelet agglutination (RIPA)[2]. Very little is known about the biochemical and clinical features of heterozygous carriers of mutations causing BSS, and about the impact in general population individuals of variation in genes encoding the GPIb-IX-V complex when present in heterozygosity. In fact, family members with only one mutated allele are generally asymptomatic, with sub-normal platelet count, slightly enlarged platelets, and marginally reduced levels of glycoproteins expression.

Here, to dissect the impact of genetic variability on platelet count, a sequencing-based whole-genome association study was performed in 6,528 volunteers included in the SardiNIA general population cohort[3]. Six signals were identified (Supplementary Table 1, supplementary data for description), including a novel non-synonymous variant (22:19711445:C/T; MAF=0.0045; $P=1.172\times10^{-16}$), mapping in the second exon (c.C79T, p.P27S) of the GP1BB gene (Suppl. Fig. 1). Completely independent of previously reported associations in the same genomic region (Suppl. data), p.P27S is Sardinian-specific, being completely missing in large sequencing datasets such as 1,000 Genomes Project[4], GoNL[5], GnomeAD[6], the Exome Sequencing Project in NHLBI's TOPMed program[7]. No homozygous and 57 carriers for the rare 22:19711445-T allele were found. Platelet count in wild-type homozygous were $242.87\pm117.05\times10^9$/L (mean±1.96*SD), whereas in p.P27S carriers were $174.17\pm91.51\times10^9$/L, corresponding to a reduction of 70.13×10^9/L for each copy of the minor allele (Fig. 1A). With this large effect, the novel founder mutation explains about 1.05% of phenotypic variance for platelet count, representing the largest phenotypic effect among all the independent variants reported so far in the GWAS Catalog (Suppl. Table 2 and 3)[8]. Moreover, in a subset of 2,000 individuals, whose mean platelet volume (MPV) was measured, this variant was associated with notably larger platelets ($P=2.13\times10^{-10}$), consistently with evidence of morphologically enlarged

platelets in BSS patients (Fig. 1B). To assess platelet functionality, a 7-color flow cytometry panel (Suppl. Table 4) was set up in 24 of 57 p.P27S carriers (42.1%) and in an equal number of matched unrelated controls. Monoclonal antibodies directed against the GPIIb-IIIa complex (CD41a and CD61), and the vWF receptor complex (CD42a and CD42b) were used to investigate the basal receptor expression in resting platelets. p.P27S carriers showed increased levels of GPIIb (CD41a, +22.36%, $P=1.61\times10^{-4}$, N=48) and GPIIIa (CD61, +16.20%, $P=6.61\times10^{-4}$, N=48) a typical finding in the presence of enlarged platelets (Suppl. Fig. 2). The expression of GPIX and GPIbα glycoproteins and their correct assembly into the GPIb-IX-V complex are known to be impaired by a defective GPIbβ peptide[9]. Indeed, despite carrying only one mutated allele, p.P27S heterozygous showed appreciably lower basal expression levels of both GPIX (-24.69%, $P=2.66\times10^{-6}$, N=46; Fig. 1C) and GPIbα (-26.51%, $P=3.66\times10^{-8}$, N=48; Fig. 1D), and consequently less of the entire complex, compared to controls. This is far more than the normal expression levels of GPIX and GPIbα in carriers of other missense mutations in GP1BB, as recently reported[10]. Pre-activation and reactivity changes in p.P27S platelets were investigated after exposure to the agonist adenosine diphosphate (ADP). Indeed, activated αIIbβ3 was prominently induced in p.P27S carriers, as shown by the extent of PAC-1 binding to resting and activated platelets (+41.94%, $P=4.84\times10^{-3}$, N=48, Fig. 1E). Notably, no variation in the response of platelets after ADP stimulation, were recently reported in BSS patients and carriers[9]. Remarkably, platelet reactivity turned out to be differentially regulated: no changes were observed in surface exposure of neo P-selectin (CD62P, +35.22%, P=0.138, N=48, Fig. 1F) and neo granulophysin (CD63, +1.86%, P=0.658, N=46, Fig. 1G), markers of granule content release. The unique functional effects of the p.P27S lead us to examine its possible consequences on molecular structure and conformational changes of GPIbβ by molecular modeling analysis based on the X-ray crystal structure[11]. Proline to Serine substitution falls in the Leucine-rich repeat N-terminal (LRRNT) domain of the 206 amino acid long protein encoded by GP1BB (Figg. 2A-B). Proline residues are expected to be disruptive of structure; and indeed, in that highly conserved region and close to cysteine residues involved in the Cys26-Cys32 disulphide bridge, p.P27S could thus modify the stability, and consequently the conformation, of GPIbβ. To test this hypothesis, we first performed in-silico Molecular Dynamic simulations, observing an increased conformational mobility of the amino acid backbone close to p.P27S (Suppl. Fig. 3), suggesting the instability of the GPIbβ glycoprotein in accord with the observed reduction of the expression of GPIX and GPIbα. Strikingly, a greater fluctuation of the amino acids in loop 2 of the p.P27S protein was also recorded, as indicated by Root-Mean-Square-Fluctuation (Fig. 2C).

In summary, all typical findings of macrothrombocytopenias (i.e. BSS) were observed in p.P27S obligate carriers characterized in this study: low levels of large platelets and low expression of GPIX and GPIbα glycoproteins, as shown by flow-cytometry. As one might anticipate, the most severe cases are caused by deletions and nonsense mutations, but some missense mutations are disabling enough to be clinically significant. In one of the reported cases[12], a charge difference is introduced (p.Asn89Asp); in the other[13], as in this case, the Proline residue is replaced (p.Pro27Leu), which is expected to disrupt secondary structure in the protein. That p.P27S influences conformational changes and stability of GPIbβ, in turn affecting GPIb-IX-V complex function, is further clearly supported by the in-silico molecular dynamic analyses. Noteworthy, a critical interaction of GPIbβ with GPIX involves N-terminal residues 15 through 32 of GPIbβ, precisely including Proline 27[14]. According to Hardy-Weinberg expectation, at least 4 p.P27S homozygous individuals, most likely with BSS, are expected in Sardinia, but none has been reported so far: this may suggest that BSS is likely underdiagnosed in Sardinia, consistent with other reports[15]. Thus, clinicians should be aware of the novel p.P27S mutation in the molecular characterization of Sardinian origin patients with a clinical picture of platelet macrocytosis and platelet count $<100\times10^9$/L.

Authorship

F.B., M.Pi. and F.D. collected samples and extracted genomic DNA from blood; F.B., A.M., A.Mu. and M.Z. performed genotyping; F.B., A.M., performed sequencing; V.O. and E.F. designed flow cytometric panels; V.O., E.F. and S.L. performed cytometric analysis; M.S., G.S., C.S., M.F., M.Pa., P.F., Ma.Ma. and S.S. performed statistical analyses; M.S., M.M. and S.S. performed bioinformatic analyses; F.B. and S.S. performed region-specific analysis and selected candidate genes; S.O. performed in-silico analyses; I.A. and C.A.C. performed other functional evaluation; F.C., G.R.A and D.S. provided funds and supervised the work; S.B. provided clinical support; F.B., M.S. and F.C. wrote the paper; V.O., G.S., S.O., A.M., C.S., M.F., D.S., and S.S. revised the paper. All authors read the paper and contributed to its final form.

Declaration of interests
The authors declare no competing interests.

Acknowledgements
We thank all the volunteers who generously participated in this study; we are grateful to mr Mario Lovicu and mr Nazario Olla for the logistic support provided and helpful suggestions. Supported by contracts N01-AG-1-2109 and HHSN271201100005C from the Intramural Research Program of the National Institute on Aging, National Institutes of Health (NIH).

Web Resources
The URLs for data presented herein are as follows
-Online Mendelian Inheritance in Man, https://www.omim.org/
-SardiNIA Project, https://sardinia.irp.nia.nih.gov/;
-1000 Genomes Project data repository, ftp://ftp.1000genomes.ebi.ac.uk/;
-GoNL, Genome of the Netherlands;
-GnomeAD, http://gnomad.broadinstitute.org/;
-Exome Sequencing Project, https://esp.gs.washington.edu/drupal/
-NHLBI TOPMed Program, https://www.nhlbiwgs.org/
-GWAS catalog, https://www.ebi.ac.uk/gwas/;

References
1. Savoia A., Pastore A., De Rocco D., Civaschi E., Di Stazio M., Bottega R., Melazzini F., Bozzi V., Pecci A., Magrin S. et al. (2011). Clinical and genetic aspects of Bernard-Soulier syndrome: searching for genotype/phenotype correlations. Haematologica 96, 417-423.
2. Berndt, M.C. and Andrews, R.K. (2011). Bernard-Soulier syndrome. Haematologica 96, 355-359.
3. Sidore C., Busonero F., Maschio A., Porcu E., Naitza S., Zoledziewska M., Mulas A., Pistis G., Steri M., Danjou F., et al. (2015). Genome sequencing elucidates Sardinian genetic architecture and augments association analyses for lipid and blood inflammatory markers. Nature Genetics 47, 1272-1281.
4. 1000 Genomes Project Consortium, Abecasis G.R., Auton A., Brooks L.D., DePristo M.A., Durbin R.M., Handsaker R.E., Kang H.M., Marth G.T., McVean G.A. et al. (2012). An integrated map of genetic variation from 1,092 human genomes. Nature 491, 56–65.
5. Boomsma D.I., Wijmenga C., Slagboom E.P., Swertz M.A., Karssen L.C., Abdellaoui A., Ye K., Guryev V., Vermaat M., van Dijk F., et al. (2014). The Genome of the Netherlands: design, and project goals. Eur J Hum Genet 22, 221-7.
6. Karczewski K.J., Francioli L.C., Tiao G., Cummings B.B, Alföldi J., Wang Q., Collins R.L., Laricchia K.M., Ganna A., Birnbaum D.P. et al. (2019). Variation across 141,456 human exomes and genomes reveals the spectrum of loss-of-function intolerance across human protein-coding genes. bioRxiv 531210; doi: https://doi.org/10.1101/531210.

7. Sequencing of 53,831 diverse genomes from the NHLBI TOPMed Program. Taliun D., Harris D.N., Kessler M.D., Carlson J., Szpiech Z.A., Torres R., Taliun S.A.G., Corvelo A., Gogarten S.M., Kang H.M., et al. (2019). Preprint from bioRxiv, 06 Mar 2019. DOI: 10.1101/563866. PPR:PPR72371.
8. MacArthur J., Bowler E., Cerezo M., Gil L., Hall P., Hastings E., Junkins H., McMahon A., Milano A., Morales J., et al. (2017). The new NHGRI-EBI Catalog of published genome-wide association studies (GWAS Catalog). Nucleic Acids Res. 45, D896-901.
9. Hadjkacem, B., Elleuch, H., Gargouri, J. & Gargouri, A. (2009). Bernard-Soulier syndrome: novel nonsense mutation in GPIbbeta gene affecting GPIb-IX complex expression. Ann. Hematol. 88, 465-472.
10. Bragadottir G., Birgisdottir E.R., Gudmundsdottir B.R., Hilmarsdottir B., Vidarsson B., Magnusson M.K., Larsen O.H., Sorensen B., Ingerslev J., Onundarson P.T. (2015). Clinical phenotype in heterozygote and biallelic Bernard-Soulier syndrome-a case control study. Am J Hematol. 90, 149-55.
11. McEwan PA, Yang W, Carr KH, et al (2011). Quaternary organization of GPIb-IX complex and insights into Bernard-Soulier syndrome revealed by the structures of GPIbβ and a GPIbβ/GPIX chimera. Blood. 118(19), 5292- 5301.
12. Fiore M, De Thoré C., Ranjatoelina H.R., Baas M.J., Jacquemont M.L., Dreyfus M., Bombled C.L., Li R., Gachet C., Dupuis A., Lanza F. (2020). High prevalence of the natural Asn89Asp mutation in the GP1BB gene associated with Bernard–Soulier syndrome in French patients from the genetic isolate of Reunion Island. British Journal Haematology 189, e67-e71.
13. Bastida J.M., Lozano M.L., Benito R., Janusz K., Palma-Barqueros V., Del Rey M., Hernández-Sánchez J.M., Riesco S., Bermejo N., González-García H., et al., (2018). Introducing high-throughput sequencing into mainstream genetic diagnosis practice in inherited platelet disorders. Haematologica 103(1), 148-162.
14. D. Kenny, P.A. Morateck and R.R. Montgomery (2002). The cysteine knot of platelet glycoprotein Ib (GPIb-beta) is critical for the interaction of GPIb-beta with GPIX. Blood 99(12), 4428-33.
15. Noris P. and C.L. Balduini (2015). Inherited thrombocytopenias in the era of personalized medicine. Haematologica 100, 2.

Figures titles and legends

Figure 1 Effects of chr22:19711445 genotype on platelet-related phenotypes.
(A) Platelet count distribution stratified on 57 heterozygous carriers and 6,471 homozygous wild types. (**B**) Mean platelet volume distribution stratified on 28 heterozygous carriers and 1,972 homozygous wild types. Basal expression levels of the main GPIb-IX-V receptor glycoproteins on resting platelets: (**C**) GPIX on 23 carriers and 23 controls and (**D**) GPIbα on 24 carriers and 24 controls (**E**, **F** and **G**). Expression levels of the most relevant platelet activation-dependent markers (activated αIIbβ3, P-selectin and granulophysin), in basal condition and after stimulation with ADP. Violin plots represent the distribution of the data; the boxplots inside report the median value as a dot, the interquartile range (IQR) as a box, the 1st quartile - 1.5 IQR and the 3rd quartile + 1.5 IQR as whiskers. *P<0.05, **P<0.01, ***P<0.001.

Figure 2 *GPIbβ amino acid sequence with BSS*-causing mutations and molecular modeling analyses. (**A**) Positions of the mutations within the coding regions of platelet glycoprotein (GP)Ibβ according to NCBI Reference Sequence, NP_000398.1. The different domains are indicated with different patterns. Different types of mutation are colour-coded; highlighted in yellow is the Proline27 to Serine27 substitution (P27S) here firstly described. Known mutations were obtained from Savoia A. et al., 2014, Sivapalaratnam S. et al., 2017, Bragadottir G. et al., 2015, Qiao J. et al., 2015, Kunishima S. et al., 2001, Ferrari S. et al., 2018 and Bastida J.M. et al., 2018. (**B**) 3D structure of GPIbβ

sequence, colour-coded according to the schematic representation in (A). (**C**) X-ray structure of protein (26-143 aa code 3RFE) showing the impact of p.P27S on GPIbβ glycoprotein conformation; in particular, the superposition between the first (**teal**) and last (**yellow**) frame of the molecular dynamics for the p.P27S protein (**left**), and the superposition between the first (**pink**) and last (**green**) frame for the WT protein (**right**) are reported.