



**HAL**  
open science

## Novel manifestations of immune dysregulation and granule defects in gray platelet syndrome

Matthew C Sims, Louisa Mayer, Janine H Collins, Tadbir K Bariana, Karyn Megy, Cecile Lavenu-Bombled, Denis Seyres, Laxmikanth Kollipara, Frances S Burden, Daniel Greene, et al.

### ► To cite this version:

Matthew C Sims, Louisa Mayer, Janine H Collins, Tadbir K Bariana, Karyn Megy, et al.. Novel manifestations of immune dysregulation and granule defects in gray platelet syndrome. *Blood*, 2020, 136 (17), pp.1956-1967. 10.1182/blood.2019004776 . hal-03162108

**HAL Id: hal-03162108**

**<https://amu.hal.science/hal-03162108>**

Submitted on 9 Mar 2021

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

## PLATELETS AND THROMBOPOIESIS

# Novel manifestations of immune dysregulation and granule defects in gray platelet syndrome

Matthew C. Sims,<sup>1-3,\*</sup> Louisa Mayer,<sup>1,2,\*</sup> Janine H. Collins,<sup>1,2,4,\*</sup> Tadbir K. Bariana,<sup>1,4,5,\*</sup> Karyn Megy,<sup>1,2,6</sup> Cecile Lavenu-Bombled,<sup>7</sup> Denis Seyres,<sup>1,2,6</sup> Laxmikanth Kollipara,<sup>8</sup> Frances S. Burden,<sup>1,2,6</sup> Daniel Greene,<sup>1,6,9</sup> Dave Lee,<sup>10</sup> Antonio Rodriguez-Romera,<sup>1,2</sup> Marie-Christine Alessi,<sup>11</sup> William J. Astle,<sup>2,9</sup> Wadie F. Bahou,<sup>12</sup> Loredana Bury,<sup>13</sup> Elizabeth Chalmers,<sup>14</sup> Rachael Da Silva,<sup>10</sup> Erica De Candia,<sup>15,16</sup> Sri V. V. Deevi,<sup>1,6</sup> Samantha Farrow,<sup>1,2,6</sup> Keith Gomez,<sup>5</sup> Luigi Grassi,<sup>1,2,6</sup> Andreas Greinacher,<sup>17</sup> Paolo Gresele,<sup>13</sup> Dan Hart,<sup>18</sup> Marie-Françoise Hurtaud,<sup>7</sup> Anne M. Kelly,<sup>1</sup> Ron Kerr,<sup>19</sup> Sandra Le Quellec,<sup>20</sup> Thierry Leblanc,<sup>7</sup> Eva B. Leinøe,<sup>21</sup> Rutendo Mapeta,<sup>1,6</sup> Harriet McKinney,<sup>1,2,6</sup> Alan D. Michelson,<sup>22</sup> Sara Morais,<sup>23,24</sup> Diane Nugent,<sup>25</sup> Sofia Papadia,<sup>1,2,6</sup> Soo J. Park,<sup>26</sup> John Pasi,<sup>18</sup> Gian Marco Podda,<sup>27</sup> Man-Chiu Poon,<sup>28</sup> Rachel Reed,<sup>10</sup> Mallika Sekhar,<sup>29</sup> Hanna Shalev,<sup>30</sup> Suthesh Sivapalaratnam,<sup>1,4</sup> Orna Steinberg-Shemer,<sup>31,32</sup> Jonathan C. Stephens,<sup>1,6</sup> Robert C. Tait,<sup>33</sup> Ernest Turro,<sup>1,2,6,9</sup> John K. M. Wu,<sup>34</sup> Barbara Zieger,<sup>35</sup> NIHR BioResource, Taco W. Kuijpers,<sup>36,37</sup> Anthony D. Whetton,<sup>10</sup> Albert Sickmann,<sup>8,38,39</sup> Kathleen Freson,<sup>40</sup> Kate Downes,<sup>1,6</sup> Wendy N. Erber,<sup>41,42</sup> Mattia Frontini,<sup>1,2,43</sup> Paquita Nurden,<sup>44</sup> Willem H. Ouwehand,<sup>1,2,6,45</sup> Remi Favier,<sup>7,46,†</sup> and Jose A. Guerrero<sup>1,2,†</sup>

<sup>1</sup>Department of Haematology, University of Cambridge, and <sup>2</sup>National Health Service Blood and Transplant, Cambridge Biomedical Campus, Cambridge, United Kingdom; <sup>3</sup>Oxford Haemophilia and Thrombosis Centre, Oxford University Hospitals NHS Foundation Trust, NIHR Oxford Biomedical Research Centre, Oxford, United Kingdom; <sup>4</sup>Department of Haematology, Barts Health NHS Trust, London, United Kingdom; <sup>5</sup>Royal Free London NHS Foundation Trust, London, United Kingdom; <sup>6</sup>NIHR BioResource, Cambridge University Hospitals, Cambridge Biomedical Campus, Cambridge, United Kingdom; <sup>7</sup>Assistance Publique-Hôpitaux de Paris, Centre de Référence des Pathologies Plaquettaires, Hôpitaux Armand Trousseau, Bicêtre, Robert Debré, Paris, France; <sup>8</sup>Leibniz-Institut für Analytische Wissenschaften-ISAS-e. V., Dortmund, Germany; <sup>9</sup>Medical Research Council Biostatistics Unit, Forvie Site, Cambridge Biomedical Campus, Cambridge, United Kingdom; <sup>10</sup>Stoller Biomarker Discovery Centre, Division of Cancer Sciences, School of Medical Sciences, Faculty of Biology, Medicine and Health, University of Manchester, Manchester Academic Health Science Centre, Manchester, United Kingdom; <sup>11</sup>Centre for CardioVascular and Nutrition Research, INSERM 1263, INRAE 1260, Marseille, France; <sup>12</sup>Department of Medicine, Stony Brook University, Stony Brook, NY; <sup>13</sup>Department of Medicine, Section of Internal and Cardiovascular Medicine, University of Perugia, Perugia, Italy; <sup>14</sup>Royal Hospital for Sick Children, Glasgow, United Kingdom; <sup>15</sup>Institute of Internal Medicine and Geriatrics, Catholic University School of Medicine, Rome, Italy; <sup>16</sup>Fondazione Policlinico Universitario Agostino Gemelli Istituto di Ricovero e Cura a Carattere Scientifico, Rome, Italy; <sup>17</sup>Institut für Immunologie und Transfusionsmedizin, Universitätsmedizin Greifswald, Greifswald, Germany; <sup>18</sup>The Royal London Hospital Haemophilia Centre, Barts and The London School of Medicine and Dentistry, Queen Mary University of London, London, United Kingdom; <sup>19</sup>Department of Haematology, Ninewells Hospital and Medical School, Dundee, United Kingdom; <sup>20</sup>Service d'Hématologie Biologique, Hospices Civils de Lyon, Lyon, France; <sup>21</sup>Department of Haematology, Rigshospitalet, Copenhagen, Denmark; <sup>22</sup>Center for Platelet Research Studies, Dana-Farber/Boston Children's Cancer and Blood Disorders Center, Harvard Medical School, Boston, MA; <sup>23</sup>Serviço de Hematologia Clínica, Hospital de Santo António, Centro Hospitalar Universitário do Porto, Porto, Portugal; <sup>24</sup>Unidade Multidisciplinar de Investigação Biomédica, Instituto de Ciências Biomédicas, Universidade do Porto, Porto, Portugal; <sup>25</sup>Center for Inherited Bleeding Disorders, Children's Hospital of Orange County, Orange, CA; <sup>26</sup>Division of Hematology and Oncology, Moores Cancer Center, University of California, San Diego, La Jolla, CA; <sup>27</sup>Unità di Medicina 2, ASST Santi Paolo e Carlo, Dipartimento di Scienze della Salute, Università degli Studi di Milano, Milan, Italy; <sup>28</sup>University of Calgary Cumming School of Medicine and Southern Alberta Rare Blood and Bleeding Disorders Comprehensive Care Program, Calgary, AB, Canada; <sup>29</sup>Department of Haematology, Royal Free London NHS Trust, London, United Kingdom; <sup>30</sup>Department of Pediatric Hematology/Oncology, Soroka Medical Center, Faculty of Medicine, Ben-Gurion University, Beer Sheva, Israel; <sup>31</sup>Department of Hematology-Oncology, Schneider Children's Medical Center of Israel, Petach Tikva, Israel; <sup>32</sup>Sackler School of Medicine, Tel Aviv University, Tel Aviv, Israel; <sup>33</sup>Department of Haematology, Royal Infirmary, Glasgow, United Kingdom; <sup>34</sup>Division of Hematology-Oncology, University of British Columbia and BC Children's Hospital, Vancouver, BC, Canada; <sup>35</sup>Division of Pediatric Hematology and Oncology, Department of Pediatrics and Adolescent Medicine, Medical Center-Faculty of Medicine, University of Freiburg, Freiburg, Germany; <sup>36</sup>Department of Pediatric Immunology, Rheumatology and Infectious Diseases, Emma Children's Hospital, Amsterdam University Medical Center, Amsterdam, The Netherlands; <sup>37</sup>Sanquin Research Institute, Department of Blood Cell Research, University of Amsterdam, Amsterdam, The Netherlands; <sup>38</sup>Department of Chemistry, College of Physical Sciences, University of Aberdeen, Aberdeen, United Kingdom; <sup>39</sup>Medizinische Fakultät, Medizinisches Proteom Center, Ruhr-Universität Bochum, Bochum, Germany; <sup>40</sup>Department of Cardiovascular Sciences, Center for Molecular and Vascular Biology, University of Leuven, Leuven, Belgium; <sup>41</sup>Faculty of Health and Medical Sciences, The University of Western Australia, Crawley, Australia; <sup>42</sup>PathWest Laboratory Medicine, The University of Western Australia, Nedlands, Australia; <sup>43</sup>British Heart Foundation, Cambridge Centre for Research Excellence, University of Cambridge, Cambridge Biomedical Campus, Cambridge, United Kingdom; <sup>44</sup>Institut Hospitalo-Universitaire L'Institut de Rythmologie et Modélisation Cardiaque, Plateforme Technologique d'Innovation Biomédicale, Hôpital Xavier Arnozan, Pessac, France; <sup>45</sup>Wellcome Sanger Institute, Wellcome Genome Campus, Hinxton, Cambridge, United Kingdom; and <sup>46</sup>INSERM Unité Mixte de Recherche 1170, Gustave Roussy Cancer Campus, Université Paris-Saclay, Villejuif, France

**Gray platelet syndrome (GPS) is a rare recessive disorder caused by biallelic variants in *NBEAL2* and characterized by bleeding symptoms, the absence of platelet  $\alpha$ -granules, splenomegaly, and bone marrow (BM) fibrosis. Due to the rarity of GPS, it has been difficult to fully understand the pathogenic processes that lead to these clinical sequelae. To discern the spectrum of pathologic features, we performed a detailed clinical genotypic and phenotypic study of 47 patients with GPS and identified 32 new etiologic variants in *NBEAL2*. The GPS patient cohort exhibited known phenotypes, including macrothrombocytopenia, BM fibrosis, megakaryocyte emperipolesis of neutrophils, splenomegaly, and elevated serum vitamin B<sub>12</sub> levels. Novel clinical phenotypes were also observed, including reduced leukocyte counts and increased presence of autoimmune disease and positive autoantibodies. There were widespread differences in the transcriptome and proteome of GPS platelets, neutrophils, monocytes, and CD4 lymphocytes. Proteins less abundant in these cells were enriched for constituents of granules, supporting a role for *Nbeal2* in the function of these organelles across a wide range of blood cells. Proteomic analysis of GPS plasma**

**showed increased levels of proteins associated with inflammation and immune response. One-quarter of plasma proteins increased in GPS are known to be synthesized outside of hematopoietic cells, predominantly in the liver. In summary, our data show that, in addition to the well-described platelet defects in GPS, there are immune defects. The abnormal immune cells may be the drivers of systemic abnormalities such as autoimmune disease. (*Blood*. 2020;136(17):1956-1967)**

## Introduction

The discovery that biallelic variants in *NBEAL2* cause gray platelet syndrome (GPS) has been a key advance in hematology<sup>1-3</sup> and has improved our understanding of platelet and granule biology.<sup>4,5</sup> The *Nbeal2*<sup>-/-</sup> mouse phenocopies the features of GPS, including bleeding, thrombocytopenia, absence of platelet  $\alpha$ ( $\alpha$ )-granules, splenomegaly, and bone marrow (BM) fibrosis.<sup>6-8</sup> In addition to the known defects of the megakaryocyte-platelet axis, we and others have reported the susceptibility of *Nbeal2*<sup>-/-</sup> mice to bacterial and viral infection, which is implicated in increased organ damage, lower survival rates, and longer time to recovery upon infection.<sup>9,10</sup> These observations have been attributed to defects of secretory granules in natural killer cells and neutrophils. It has also been shown that mast cells of *Nbeal2*<sup>-/-</sup> mice are deficient in storage vesicles<sup>11</sup> and that monocytes have reduced granularity.<sup>9</sup> Together with our observation that *Nbeal2*<sup>-/-</sup> megakaryocytes (MKs) have a proinflammatory profile,<sup>7</sup> these findings highlight that *Nbeal2* is important for normal granule function in MKs, platelets, and a variety of myeloid and lymphoid cells in mice. However, it has been unclear whether the immune cell defects seen in the murine model are relevant to the pathophysiology of GPS. Small series of patients with GPS have reported recurrent infection<sup>12-14</sup> and atypical presentations of autoimmune lymphoproliferative syndrome.<sup>15,16</sup> Neutrophil ultrastructure has been evaluated in several patients with GPS, with conflicting results: some reports have shown that neutrophil granules are preserved,<sup>17,18</sup> whereas others have noted their absence.<sup>12,13</sup>

Given the paucity of published studies on patients with GPS with confirmed etiologic variants in *NBEAL2*, we established an international collaboration to systematically evaluate clinical and laboratory phenotypes in a large collection of patients. We focused on features related to immunity and followed this up with a detailed molecular characterization of plasma, platelets, and 3 different leukocyte populations in a subset of the patients

using both protein mass spectrometry (MS) and RNA sequencing (RNA-seq).

## Methods

This section contains a short description of the methods we have used. The supplemental Methods (available on the *Blood* Web site) provide further details.

### Enrollment

All study participants provided written informed consent. The majority of participants enrolled into UK or French studies, all of which were approved by research ethics committees (UK: REC 13/EE/0325, REC 10/H0304/65, REC 10/H0304/66; France: INSERM RBM-014). The remaining participants consented using ethics procedures approved at the recruiting center.

### Sequencing and variant interpretation

DNA was extracted and sequenced by using the ThromboGenomics next-generation sequencing platform, long-range PCR, or whole-genome sequencing.<sup>1,19-21</sup> Pathogenicity classification of variants was determined in multidisciplinary meetings according to international guidelines<sup>22-25</sup> and using computational tools.<sup>26-31</sup> Only patients with etiologic variants in *NBEAL2*, defined as pathogenic (PV), likely pathogenic (LPV), and uncertain significance variants with at least 3 supporting or 1 moderate and 1 supporting piece of evidence (supplemental Table 1), were included.

### Phenotyping of patients

Clinical information and the results of laboratory investigations, including the complete blood counts (CBCs), of the patients were obtained by the enrolling physicians. This information was used to record Human Phenotype Ontology (HPO) terms for each participant.<sup>32</sup>

## BM biopsies

BM biopsies were performed on 23 patients with GPS. BM fibrosis was defined by a reticulin grade  $\geq 2$ .<sup>33</sup> Quantification of emperipolesis, the presence of intact neutrophils in the cytoplasm of MKs, was performed by manual counting on 3 samples.

## Enhanced cellular phenotyping and plasma proteomics

Five patients with GPS and 5 healthy control participants attended a central center for a more complete evaluation under standardized conditions. A CBC was performed on a Sysmex XN-1000 analyzer. Additional measured blood cell parameters, not normally included in CBC reports, including forward scatter (FSC) and side scatter (SSC), were analyzed. In addition, platelets, neutrophils, monocytes, and CD4 lymphocytes were isolated from citrate anticoagulated blood.<sup>34</sup> For each cell type, RNA was extracted and sequenced, and the lysate was analyzed by using protein MS. Plasma was obtained from EDTA anticoagulated blood of 11 patients with GPS and 13 healthy control participants and analyzed by using Sequential Window Acquisition of All Theoretical Mass Spectra–Mass Spectrometry. All control subjects were selected from a group of volunteers without comorbidities and were matched on age and sex as closely as possible to the patients with GPS (supplemental Tables 2 and 3).

## Statistical analysis

The Fisher's exact test was used for analyses unless stated otherwise.

## Results

### A wide spectrum of *NBEAL2* variants cause GPS

Forty-seven patients from 38 genetically independent pedigrees were enrolled at 21 hospitals in 11 countries. We identified a total of 70 etiologic variants in *NBEAL2*, 32 of which were novel (Figure 1A; supplemental Table 1). Forty-three percent of variants were inherited in homozygosity and 57% in compound heterozygosity. Fifty-six variants were unique, 64% of which were classified as PV or LPV (Figure 1B). Of those unique variants, 38 were absent from the genotyping results of 162 100 individuals in the Genome Aggregation Database and the 100 000 Genomes Project; the minor allele frequencies of the remaining 18 were  $< 0.00004$ .<sup>35,36</sup>

Missense variants accounted for 49% of all etiologic variants identified, with the remaining ones categorized (in descending order of frequency) as frameshift, stop gained, splice, and short insertion/deletion (indel) variants. We mapped the position of the missense variants onto the 5 known domains of *Nbeal2* and determined the evolutionary conservation score of the affected residues<sup>37</sup> (Figure 1C). Missense variants were enriched ( $P = 4.91 \times 10^{-5}$ ) in the highly conserved BEACH domain, consistent with a suggested key role of this domain in the function of *Nbeal2*.<sup>5,38</sup>

We then mapped the 13 BEACH missense variants to the crystal structure of another BEACH domain-containing protein (BDGP), *Nbea*.<sup>39</sup> Five variants occurred in the hydrophobic core of BEACH, a portion of which interacts with neighboring PH, a 129-residue domain with a neutral conservation score.<sup>37</sup> However, we identified a novel missense variant, R1979W, in a portion of PH

that is highly conserved; the position of this variant corresponds to Nbea R2208, shown to be functionally relevant for the PH–BEACH interface.<sup>40</sup> To our knowledge, this is the first etiologic variant to be described in PH.

### Bleeding and platelet phenotypes in GPS

Patients recruited into the study were between the ages of 6 and 70 years (median, 35 years). The median age at presentation was 11.5 years but this varied between 2 months and 67 years. All patients were thrombocytopenic, and  $\alpha$ -granule deficiency was noted in the platelets of all cases assessed by electron or light microscopy (Figure 2A; supplemental Figure 2.1; supplemental Table 2). Consistent with previous reports,<sup>17,40,41</sup> a wide spectrum of bleeding symptoms was reported, ranging from subcutaneous to intracranial hemorrhage. Five patients were notable for their lack of a bleeding diathesis.

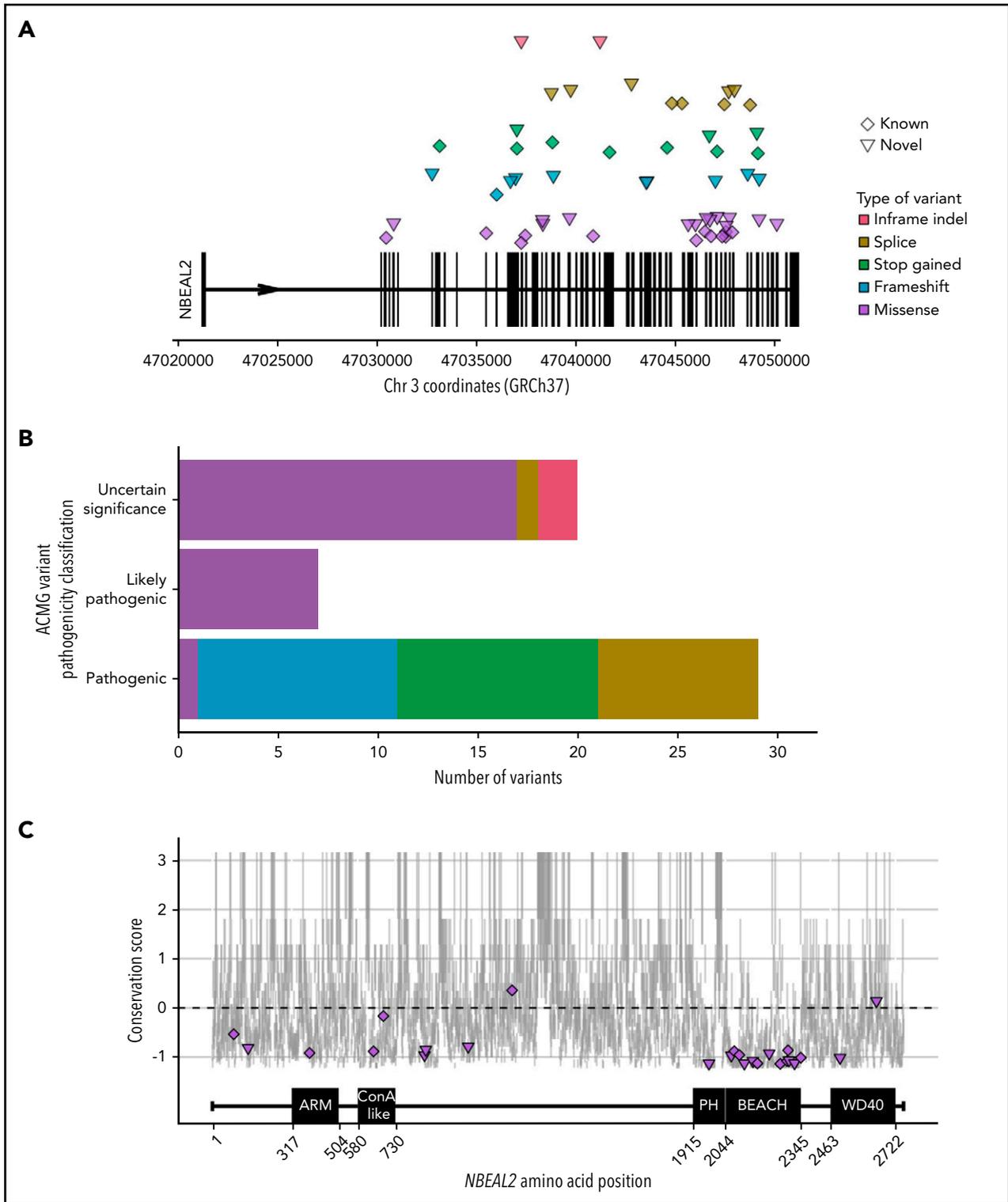
### Raised vitamin B<sub>12</sub> levels and stable BM fibrosis

Raised serum vitamin B<sub>12</sub> levels (B12) were recorded for 31 of 34 patients, expanding the previous observations made in a smaller GPS collection.<sup>17</sup> B12 was increased by at least 1.5 times the upper limit of the local reference range in two-thirds of patients (supplemental Figure 2.1; supplemental Table 2).

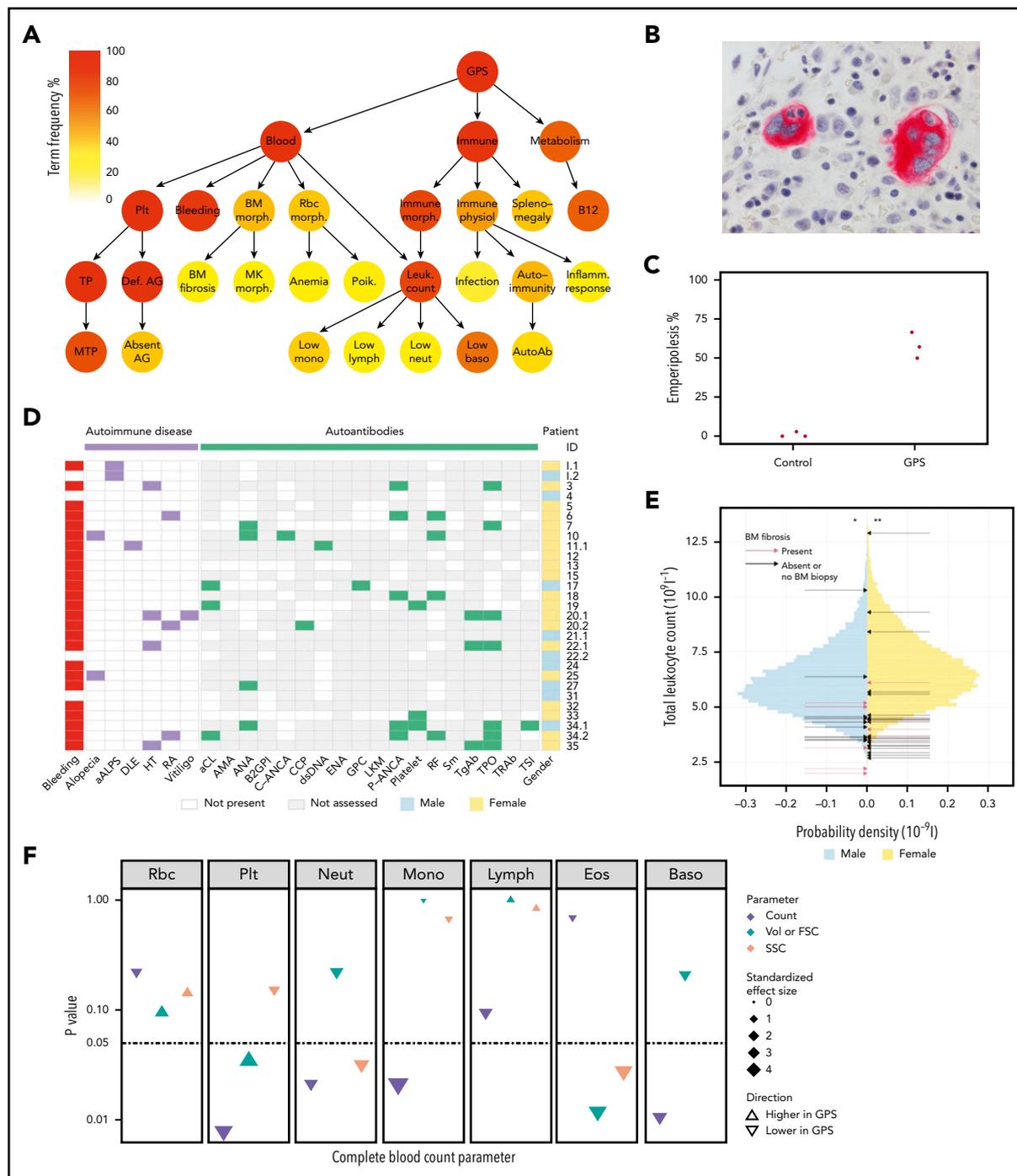
BM biopsies were performed in 23 patients, and 13 (57%) patients were diagnosed with BM fibrosis. Morphologic examination of 5 centrally reviewed trephine samples showed preservation of trilineage hematopoiesis (supplemental Figure 2.2). BM fibrosis was diagnosed at a median age of 28.5 years (range, 10–52 years), with a median of 16 years since this diagnosis to the present time. Patient 20.3 received an allogeneic stem cell transplant at age 32 years,<sup>42</sup> but the rest of the patients have not required treatment of BM fibrosis. Neutrophil emperipolesis was observed in a mean of 58% MKs in 3 GPS BM trephines compared with only 1% in 3 trephines from control subjects (Figure 2B–C).

### Patients with GPS have abnormalities of the immune system, including autoimmune disease affecting multiple organ systems

Coding of clinical and laboratory information by using HPO terms identified phenotypes associated with “Abnormality of the immune system” in 91% of patients (Figure 2A; supplemental Table 2). “Splenomegaly” (determined clinically or by ultrasound) was observed in 40% of patients and did not significantly overlap with BM fibrosis ( $\chi^2$  test,  $P = .26$ ) (supplemental Figure 2.2). “Abnormal immune system morphology” (81%) was largely due to cytopenia of at least one leukocyte type (77%). “Abnormal immune system physiology” (51%) included “Recurrent infections” (17%), most commonly respiratory tract infections and otitis media, and “Autoimmunity” (43%), which consisted of autoimmune disease and positive autoantibodies. Twelve patients (26%) were diagnosed with an autoimmune disease with a wide spectrum of organ systems affected. These included the endocrine (Hashimoto's thyroiditis), skeletal (rheumatoid arthritis), integumentary (alopecia, discoid lupus erythematosus, and vitiligo), and immune (atypical autoimmune lymphoproliferative syndrome) systems (Figure 2D). Twenty-nine patients had autoantibody tests performed and, of these, 17 (59%) had at least 1 positive test result. The 4 most frequent positive test results (in descending order of frequency) were



**Figure 1. NBEAL2 variants identified in patients with GPS participating in this study.** (A) Position of the 56 unique variants relative to NBEAL2 on chromosome 3 (Chr 3) using Genome Reference Consortium human genome build 37 (GRCh37). Vertical black bars represent exons. The color and shape used to represent variants in the legend apply throughout the figure. Inframe indel refers to variants which are inframe small insertions and/or deletions  $\leq 50$  base pairs. (B) The frequency of unique variants classified by pathogenicity, according to the American College of Medical Genetics and Genomics (ACMG) guidelines.<sup>22</sup> (C) The location of missense variants compared with the amino acid sequence of Nbeal2 and known functional domains (on the x-axis). The y-axis position of the variant is the estimated evolutionary conservation score of the affected residue calculated by using ConSurf.<sup>37</sup> A positive score corresponds to lower conservation, and a negative score represents higher conservation. Vertical gray lines represent the range of conservation scores of the 50% confidence interval: the bottom and top of the bar are the 25th and 75th percentile, respectively, of the inferred evolutionary rate distribution.



**Figure 2. Novel clinical phenotypes.** (A) Summary HPO tree, showing the 3 most frequent HPO organ systems represented in the 47 patients with GPS: Blood ("Abnormality of blood and blood-forming tissues"), Immune ("Abnormality of the immune system"), and Metabolism ("Abnormality of metabolism/homeostasis"). HPO terms affecting  $\geq 8$  patients are shown except terms associated with "Bleeding," which are displayed in supplemental Figure 2.1. HPO term labels: Plt, "Abnormal thrombocyte morphology"; Bleeding, "Abnormal bleeding"; BM morph., "Abnormality of bone marrow cell morphology"; Rbc morph., "Abnormal erythrocyte morphology"; Immune morph., "Abnormal immune system morphology"; Immune physiol., "Abnormality of immune system physiology"; Spleno-megaly, "Spleno-megaly"; B12, "Abnormal vitamin B<sub>12</sub> level"; TP, "Thrombocytopenia"; Def. AG, "Abnormal number of alpha granules"; BM fibrosis, "Myelofibrosis"; MK morph., "Abnormal megakaryocyte morphology"; Anemia, "Anemia"; Poik., "Poikilocytosis"; Leuk. count, "Abnormal leukocyte count"; Infection, "Recurrent infections"; Autoimmunity, "Autoimmunity"; Inflamm. response, "Increased in-inflammatory response"; MTP, "Macrothrombocytopenia"; Absent AG, "Absence of alpha granules"; Low mono, "Monocytopenia"; Low lymph, "Lymphopenia"; Low neut, "Neutropenia"; Low baso, "Decreased basophil count"; AutoAb, "Autoimmune antibody positivity." (B) BM trephine image with CD61 stain (1000 $\times$ ) of patient ID 20.3, which shows neutrophil emperipolesis by 2 MKs. (C) Dot plot showing percentage of MKs with neutrophil emperipolesis in 3 patients with GPS (patient IDs 18, 33, and 34.1) (supplemental Figure 2.2 presents representative BM trephine images) and 3 control subjects. (D) Representation of autoimmune disease, results of autoantibody tests, and presence of bleeding symptoms in 29 patients with GPS (labeled according to patient ID) in whom autoantibody tests were performed. Autoantibodies tested in at least 3 patients are included (listed per order in graphic): aALPS, atypical autoimmune lymphoproliferative syndrome; DLE, discoid lupus erythematosus; HT, Hashimoto's thyroiditis; RA, rheumatoid arthritis; and autoantibodies against: aCL, cardiolipin; AMA, mitochondria; ANA, nuclear; B2GPI, beta2-glycoprotein I; C-ANCA, neutrophil cytoplasmic; CCP, cyclic citrullinated peptide; dsDNA, double-stranded DNA; ENA, extractable nuclear antigen; GPC, gastric parietal cell; LKM, liver-kidney microsome; P-ANCA, neutrophil perinuclear; RF, rheumatoid factor; Sm, spliceosomal; TgAb, thyroglobulin; TPO, thyroperoxidase; TRAb, thyroid-stimulating hormone receptor; TSI, thyroid stimulating

anti-thyropoxidase, perinuclear anti-neutrophil cytoplasmic antibodies, rheumatoid factor, and antinuclear antibodies.

### Altered blood cell parameters of multiple cell lineages in patients with GPS

We compared the complete and differential leukocyte counts of all 47 patients with GPS against those of 45 032 healthy blood donors from the INTERVAL (Efficiency and Safety of Varying the Frequency of Whole Blood Donation) study.<sup>43</sup> The median total leukocyte count of both male and female patients with GPS was significantly lower than for the INTERVAL participants (one-sample Wilcoxon signed-rank test,  $P = 3 \times 10^{-3}$  and  $4 \times 10^{-4}$ , respectively) (Figure 2E). All median differential leukocyte counts (neutrophil, monocyte, lymphocyte, eosinophil, and basophil counts), as well as platelet counts and hemoglobin concentrations, were significantly lower than the corresponding INTERVAL values (supplemental Figures 2.1 and 2.3; supplemental Table 2). Although we found no association between total leukocyte, granulocyte, or monocyte counts and BM fibrosis (logistic regression,  $P > .05$ ), platelet and lymphocyte counts were inversely associated with the latter (logistic regression,  $P < .05$ ). Furthermore, there was no association between total or any of the differential leukocyte counts and splenomegaly (logistic regression,  $P > .05$ ).

In addition, 5 patients with GPS and 5 control subjects participated in a single-center follow-up study in which CBC measurements were conducted that included additional parameters (eg, FSC, SSC). This analysis confirmed that the neutrophil, monocyte, and basophil counts were significantly reduced in patients with GPS (Mann-Whitney  $U$  test,  $P < .05$ ) (Figure 2F; supplemental Table 2). Moreover, the FSC and SSC measurements exhibited a significant reduction in GPS eosinophil size and granularity, the latter also holding true for neutrophils (Mann-Whitney  $U$  test,  $P < .05$ ).

### Genotype–phenotype analysis

We compared subgroups of the patients according to the type, position, and pathogenicity classification of their etiologic variant(s). The goal was to determine whether there were phenotypic differences as reported for other inherited disorders caused by pathogenic variants in BDCPs.<sup>44</sup>

First, patients were grouped into those with biallelic loss-of-function variants (frameshift, stop gained, and splice) and those with biallelic variants causing a change in the amino acid sequence but without truncation (hereafter called “protein-altering variants” [ie, missense and in-frame indels]), excluding patients with compound heterozygosity for one loss-of-function variant and one protein-altering variant. There were 23 patients in the loss-of-function group and 18 patients in the protein-altering group. There were no significant differences in age (current and at presentation), spleen size, or CBC parameters between the groups and no significant overrepresentation of BM fibrosis,

recurrent infection, or autoimmune diseases in one group compared with the other (Welch two-sample  $t$ -test and Wilcoxon rank sum test,  $P > .05$ ) (supplemental Table 2).

We then performed two additional comparisons repeating the aforementioned analysis using the same variables and statistical tests: patients with biallelic BEACH missense variants ( $n = 7$ ) vs those with missense variants outside of BEACH ( $n = 5$ ); and patients with biallelic variants of uncertain significance ( $n = 10$ ) vs those with biallelic LPVs or PVs ( $n = 28$ ). Although the statistical power was limited because of small group sizes, no significant genotype–phenotype associations were observed.

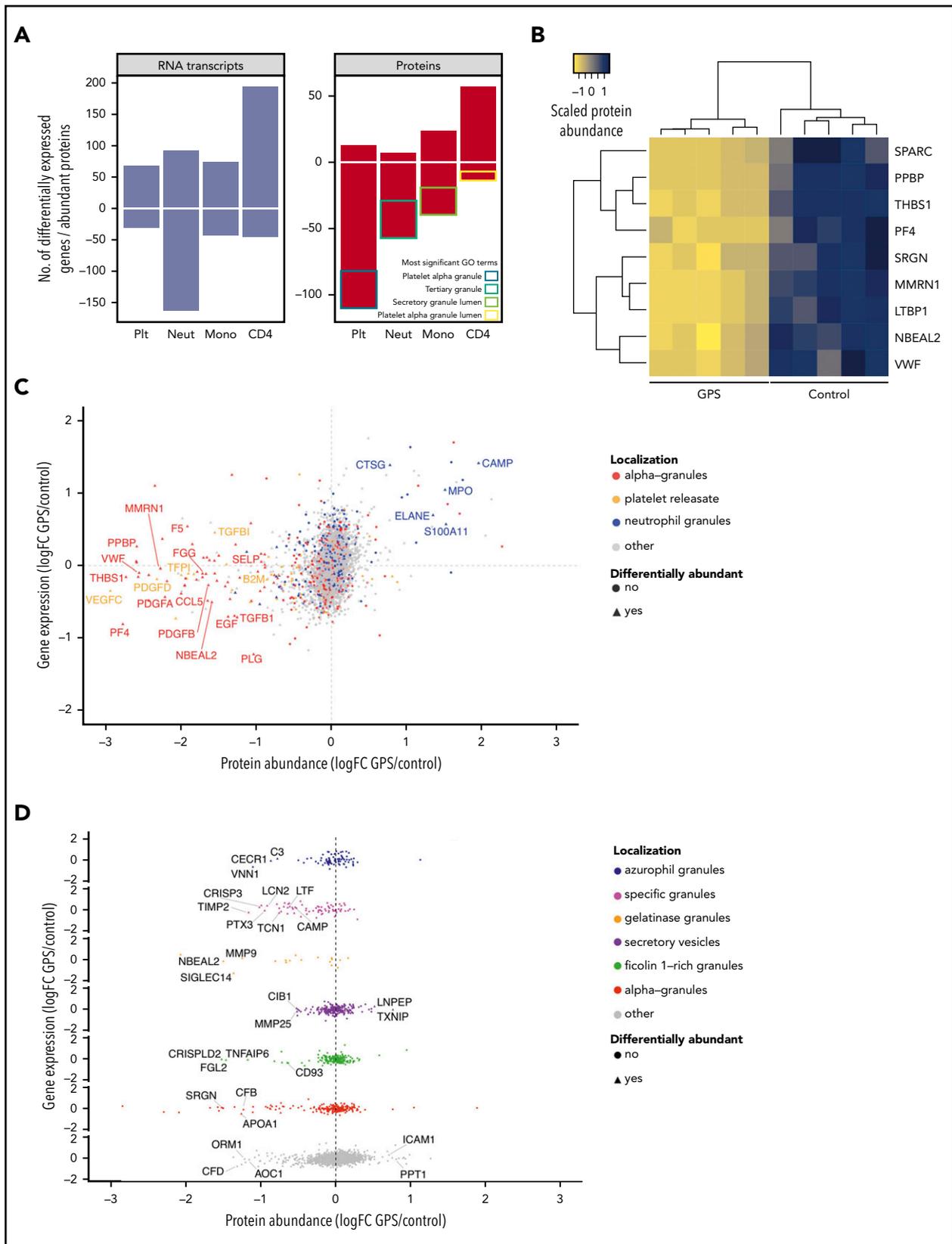
### Platelets and leukocytes have altered transcriptome and proteome profiles in GPS

Prompted by the propensity to autoimmune diseases and the altered CBC parameters, we set out to obtain a more comprehensive understanding of the differences in cellular phenotypes of the same 5 patients with GPS who had additional CBC parameters measured. Platelets, neutrophils, monocytes, and CD4 lymphocytes were analyzed by using RNA-seq and protein MS, and the results were compared with those obtained for 5 control subjects. Principal component analysis of these data clearly delineated GPS and control samples, except for platelet RNA-seq (supplemental Figure 3.1). Across each of the 4 blood cell types, the number of differentially expressed genes ranged from 95 to 255 and differentially abundant proteins from 63 to 123 (Figure 3A; supplemental Table 3). Of the 123, 65, and 63 differentially abundant proteins in platelets, neutrophils, and monocytes, 89%, 86%, and 62% were reduced in patients with GPS, respectively. These reduced proteins were enriched in gene ontology (GO) terms pertaining to cell granules and their lumens. Nine proteins, including Nbeal2, were significantly reduced in patients with GPS in at least 3 of the blood cell types, and all are known to localize to blood cell granules (Figure 3B).

### GPS platelets are diminished in $\alpha$ -granule cargo but contain neutrophil granule proteins

Of the proteins differentially less abundant in GPS platelets, there was a significant overrepresentation of proteins known to be present in  $\alpha$ -granules and/or the platelet releasate (73 of 110,  $P < 2.2 \times 10^{-16}$ ) (Figure 3C; the supplemental Methods provide classification of protein localization). None of these proteins was differentially expressed in the RNA-seq analysis, suggesting that a loss of function of Nbeal2 does not affect the transcriptional output of  $\alpha$ -granule-associated genes in platelets. We then inspected the 13 proteins differentially more abundant in GPS platelets. Five of these, including elastase (Elane) and myeloperoxidase (Mpo), are known to localize to neutrophil granules<sup>45</sup> (supplemental Figure 3.2). Furthermore, when analyzing the platelet proteomics data using a less stringent cutoff ( $\log_2$  fold change  $>0.5$ ) than that used to determine differential protein abundance ( $\log_2$  fold change  $\geq 2$ ), we identified 54 proteins more abundant in GPS platelets at this threshold. Fourteen of

**Figure 2 (continued)** immunoglobulin. (E) Histogram showing the total leukocyte count of 45 032 blood donors in the INTERVAL study<sup>43</sup> stratified according to sex, upon which the results of the patients with GPS are represented by arrows. The median total leukocyte count of both male and female patients with GPS was significantly ( $P = 3 \times 10^{-3}$  and  $4 \times 10^{-4}$ , respectively) lower than INTERVAL participants using a one-sample Wilcoxon signed-rank test, as represented by \* and \*\*. (F) CBC results for 5 patients with GPS vs 5 control subjects. The data point for each cell type and CBC parameter shows the absolute standardized effect size and directionality. On the y-axis ( $\log_{10}$  scale), the  $P$  value (Mann-Whitney  $U$  test) is shown; the horizontal dotted line represents  $P < .05$ . Volume (Vol) or forward scatter (FSC) refers to the following measurements: mean cell volume (MCV) for red blood cells (Rbc), platelet mean frequent volume (P-MFV) for platelets (Plt); FSC for neutrophils (Neut), monocytes (Mono), lymphocytes (Lymph), eosinophils (Eos), and basophils (Baso). SSC was available for all parameters except basophils and reflects the complexity of cellular contents, including granularity.



**Figure 3. The transcriptome and proteome of platelets, neutrophils, monocytes, and CD4 lymphocytes.** For all panels, log<sub>2</sub> fold change (log<sub>2</sub> fold change [logFC]) for gene expression (by RNA-seq) and protein abundance (by MS) was determined by log<sub>2</sub>-transforming the ratio of the means of individual normalized values of the 5 patients with GPS and 5 control subjects. Negative values represent genes/proteins that are reduced in GPS and vice versa. The threshold for differential abundance/expression was determined by each experimental method and the corresponding analysis used (supplemental Methods). (A) Bar chart showing differentially expressed genes by RNA-seq (left) and differentially abundant proteins by MS (right). The y-axis represents the absolute number of genes or proteins determined by each method, and negative numbers represent significant reduction in GPS cells and vice versa. The most significant GO cellular component terms for the differentially less abundant proteins of each of the cell types is

these proteins are normally localized in neutrophil granules, which indicated a significant overrepresentation of neutrophil granule proteins in GPS platelets ( $P = 6.6 \times 10^{-4}$ ). Importantly, none of these 14 proteins is known to localize to platelet granules,<sup>46-48</sup> only 5 of these 14 have been detected previously in the platelet proteome,<sup>49</sup> and only 3 of the 14 were differentially expressed genes in the GPS platelet transcriptome.

### GPS neutrophils and monocytes are deficient in granule proteins

To explore the GPS neutrophil proteome, we annotated the differentially abundant proteins according to their presence or absence in granules. For known granule proteins, these were further annotated according to granule subtype (supplemental Methods). Of the 56 differentially abundant proteins that were reduced in GPS neutrophils (Figure 3A), 36 (65%) could be assigned to different granule subtypes (Figure 3D; supplemental Figure 3.3; supplemental Table 3) and there was a significant overrepresentation of proteins that normally localize to the gelatinase and specific granules ( $P < 1.3 \times 10^{-8}$  and  $< 2.2 \times 10^{-16}$ , respectively). Given the reduced counts of monocytes in our patients with GPS, we hypothesized that a lack of functional Nbeal2 could also influence the granularity of these cells. Because there is no authoritative publication on the monocyte granule proteome, we annotated the differentially abundant proteins as granule proteins if they overlapped with proteins detected in platelet releasate,  $\alpha$ -granules, or neutrophil granules. Of the 39 proteins reduced in GPS monocytes, 29 had a granular localization (supplemental Figure 3.4). Similar to GPS platelets and neutrophils, the genes encoding these 29 proteins were not differentially expressed in our RNA-seq analysis.

### GPS CD4 lymphocytes upregulate markers of immune responses

Given the associations between CD4 lymphocytes and autoimmune disorders,<sup>50</sup> we also included these cells for evaluation by using RNA-seq and protein MS. In contrast to the other 3 cell types analyzed, the majority of differentially abundant proteins in GPS CD4 lymphocytes (57 of 71) were increased (Figure 3A; supplemental Table 3), and there was a correlation with gene expression (supplemental Figure 3.4). Among these proteins/genes differentially more abundant and more highly expressed, there was an overrepresentation of proteins involved in immunomodulatory functions; for example, Bruton tyrosine kinase (Btk) and the  $\alpha$ -chain of the receptor for the Fc domain of IgE (Fc $\epsilon$ 1a).

### The GPS plasma proteome has a proinflammatory and hepatic signature

We postulated that the granular abnormalities in the blood cells evaluated might lead to systemic changes in the blood circulation. To this end, we analyzed the plasma of 11 GPS patients and 13 control subjects using Sequential Window Acquisition of

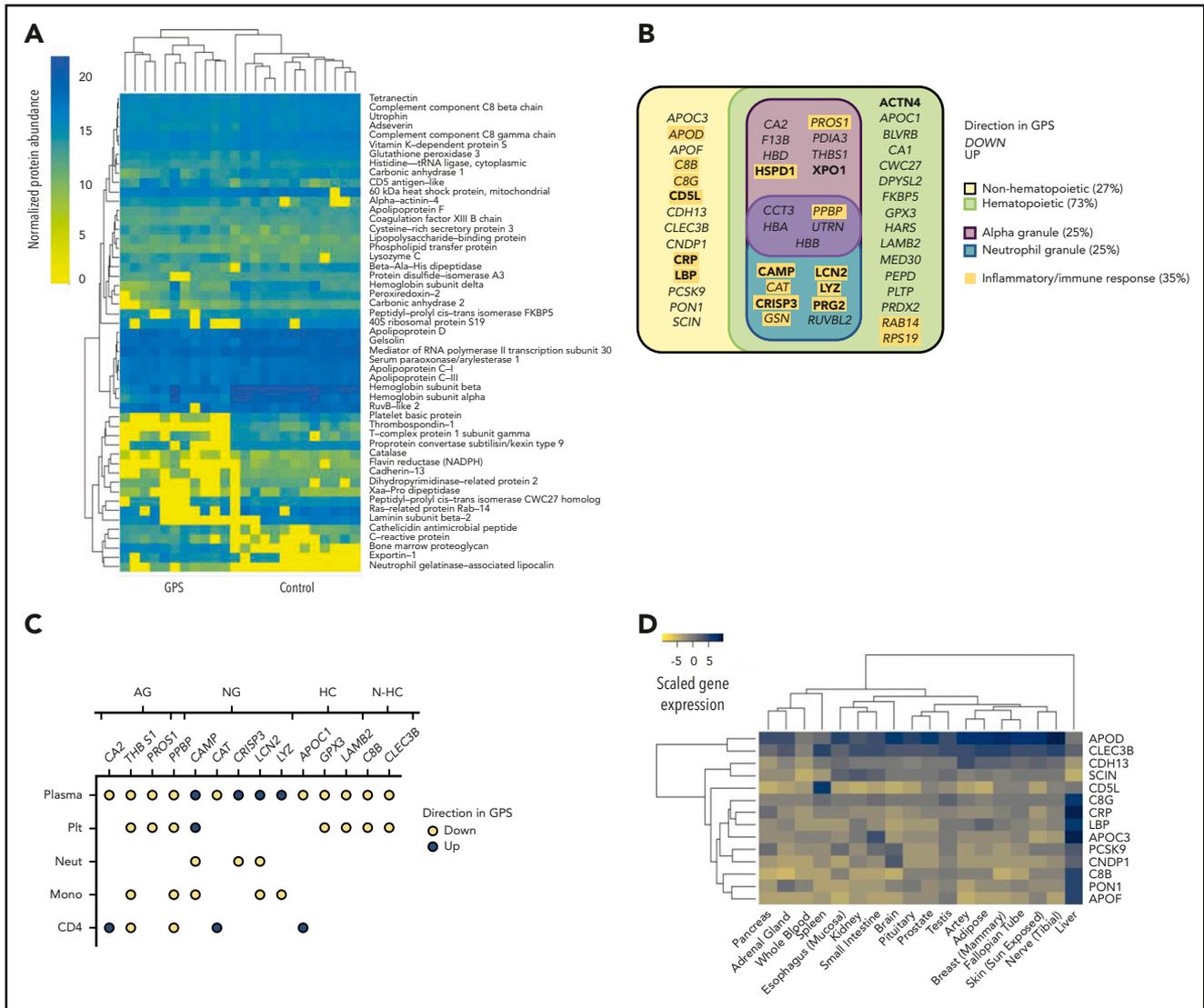
All Theoretical Mass Spectra-Mass Spectrometry. First, an unsupervised random forest analysis of the normalized concentration of plasma proteins segregated patient and control samples and determined 51 discriminatory proteins (Figure 4A; supplemental Table 3). The analysis showed that 11 and 40 discriminatory proteins had higher and lower concentrations, respectively, in patients. GO enrichment analysis highlighted the presence of terms associated with inflammation and immune response, particularly in the top-ranked GO terms for the 11 proteins more abundant in GPS plasma (Figure 4B). We then overlapped the discriminatory plasma proteins with the proteome of the 4 cell types analyzed. Fourteen of the discriminatory plasma proteins were differentially abundant in the proteome of at least 1 cell type, but the pattern of directionality differed according to protein (Figure 4C). Interestingly, proteins that were less abundant in the GPS platelet proteome were also less abundant in GPS plasma, including proteins localized to  $\alpha$ -granules. In contrast, all of the 4 overlapping proteins with higher levels in plasma are known to be present in neutrophil granules. Of these, cathelicidin antimicrobial peptide (Camp), cysteine-rich secretory protein 3 (Crisp3), and neutrophil gelatinase-associated lipocalin (Lcn2) were all differentially less abundant in GPS neutrophils, and all localize to the neutrophil-specific granules.<sup>45</sup>

Cross-referencing the discriminatory plasma proteins with the BLUEPRINT consortium gene expression data<sup>51</sup> revealed that 14 of the discriminatory plasma proteins are not known to be expressed in hematopoietic cells (Figure 4B; supplemental Figure 4). We then analyzed the gene expression of these 14 proteins in the Genotype-Tissue Expression database,<sup>52,53</sup> which showed that 9 of these nonhematopoietic proteins are predominantly synthesized by liver-residing cells. These include known acute-phase reactants such as C-reactive protein and lipopolysaccharide-binding protein (Figure 4D), both of which had increased levels in the plasma of patients with GPS.

## Discussion

Through our international collaboration, we brought together the largest collection, to date, of patients with GPS. We report on 56 unique etiologic NBEAL2 variants. Thirty-two variants were newly identified, including 1 in the PH domain. By comprehensively re-assessing established phenotypes of GPS in 47 patients, we confirm that there is marked heterogeneity of bleeding symptoms, a near-universal increase in B12, and BM neutrophil emperipolesis.<sup>17,40,54</sup> Our data further suggest that splenomegaly and BM fibrosis each occur in one-half of patients but without significant overlap. Importantly, in the majority of patients, BM fibrosis seems to be nonprogressive and is likely to be a reactive phenomenon that does not generally require intervention. We show that the median count for all leukocyte types is reduced in GPS and reveal that neutrophil granularity, as

**Figure 3 (continued)** illustrated in the "Proteins" panel. (B) For each individual, the protein abundance was calculated for each of the 9 proteins that are differentially less abundant in at least 3 types of cells and averaged across all of the cells. The results were scaled from -1 to 1 and plotted as a dendrogram with heatmap, such that positive (dark blue) and negative (yellow) values represent higher and lower protein abundance, respectively. Hierarchical clustering was applied by using the complete linkage method, and dissimilarity between rows and columns was based on the Euclidean distance. Each column represents an individual GPS patient or control subject. (C) The comparison of gene expression and protein abundance of platelets (GPS vs control) highlights the high proportion of "alpha-granules" and "releasate" proteins (in red and orange, respectively), which are significantly diminished in GPS platelets at the protein level ( $P < 2.2 \times 10^{-19}$ ), and also shows that proteins normally resident in neutrophil granules (blue) are increased in GPS platelets. (D) The comparison of gene expression and protein abundance of neutrophils (GPS vs control) separated vertically by subcellular localization shows a significantly lower abundance of specific and gelatinase granule proteins in GPS neutrophils ( $P < 1.3 \times 10^{-8}$  and  $< 2.2 \times 10^{-16}$ , respectively).



**Figure 4. The GPS plasma proteome.** (A) A heatmap of normalized protein abundance of the 51 discriminatory proteins detected in plasma from 11 patients with GPS and 13 control subjects, which showed segregation by unsupervised clustering. Each column represents an individual GPS patient or control subjects. (B) Classification of the 51 discriminatory proteins of the plasma proteome by directionality in GPS plasma, cellular origin (hematopoietic or nonhematopoietic), and subcellular localization (platelet  $\alpha$  or neutrophil granules). Proteins involved in inflammatory or immune responses are highlighted in yellow. (C) A dot plot representing the 14 discriminatory plasma proteins that are also differentially abundant proteins in at least one of the cell types evaluated by using Sequential Window Acquisition of All Theoretical Mass Spectra-Mass Spectrometry (platelets [Plt], neutrophils [Neut], monocytes [Mono], and CD4 lymphocytes [CD4]). Each data point represents a protein and is color-coded according to its directionality in the respective proteome. Absence of a protein in an explicit cell-specific proteome denotes a protein that was determined not to be differentially abundant. Proteins are annotated by granule localization (AG = platelet  $\alpha$  granule, NG = neutrophil granule) or, for proteins not localized in granules, gene expression in hematopoietic cells (HC) or non-hematopoietic cells (N-HC) in the BLUEPRINT consortium atlas.<sup>51</sup> (D) A heatmap showing the gene expression of the 14 discriminatory plasma proteins of nonhematopoietic origin in the Genotype-Tissue Expression Project.<sup>52,53</sup> Scaled gene expression represents log<sub>2</sub> transformation of transcripts per million base pairs mapped.

assessed by SSC, is reduced. We also show, for the first time, that eosinophil SSC is diminished, suggesting a potential role for *Nbeal2* in the granule biology of these cells.

Immune dysregulation in patients was common, with more than one-half having detectable autoantibodies and one-quarter clinically manifesting autoimmune disease. Comparison with HPO terms in a control collection of 428 patients with suspected inherited thrombocytopenia<sup>21</sup> revealed autoimmunity only in 3%, suggesting an overrepresentation of this phenotype in GPS. We observed a wide spectrum of immune pathologies, including newly described associations with rheumatoid arthritis, alopecia, and skin-related autoimmune disorders. In

addition, we extend a previous single case report of Hashimoto's thyroiditis<sup>12</sup> in GPS to 4 patients. Almost 20% of the patients had recurrent infections corroborating our observations in *Nbeal2*<sup>-/-</sup> mice.<sup>10</sup> The wide spectra of autoimmune diseases and immunodeficiency in the context of granule defects are also the consequence of Mendelian disorders of other BDCP genes, such as *LYST*,<sup>38</sup> and highlight the disruptive effect of etiologic variants in *NBEAL2* on a variety of immune cells. This was shown in the RNA-seq and protein MS analysis of GPS platelets, neutrophils, monocytes, and CD4 lymphocytes, in which there were widespread differences in the transcriptome and proteome of all cells evaluated. This finding is in stark contrast to the isolated platelet defect seen in our recent

description of *IKZF5*-related thrombocytopenia, a disorder also accompanied by reduced  $\alpha$ -granularity.<sup>34</sup>

GPS platelets were markedly diminished in proteins known to localize to  $\alpha$ -granules, but unexpectedly, we observed that proteins which were differentially more abundant in GPS platelets were enriched with neutrophil granule constituents. Given the lack of a concordant increase in expression of the genes corresponding to these proteins in our RNA-seq analysis, 2 potential mechanisms to explain this are endocytosis by circulating platelets or emperipolesis by MKs. Although we cannot discard the former, we reason that emperipolesis is likely to have consequences in the platelet progeny produced by MKs that have engulfed neutrophils. The transfer of membrane between MKs and neutrophils during emperipolesis has been reported,<sup>55</sup> and a recent functional genomic study reported an association of *Mpo*, known to localize to neutrophil-specific granules, with mean platelet volume highlighting the cross-talk between neutrophils and platelets.<sup>56</sup> Hence, the release of neutrophil granule proteins inside MKs during emperipolesis would provide a plausible explanation for the higher level of these proteins in GPS platelets.

Our proteomic data set strongly supports the reduction of specific granule proteins in GPS neutrophils, an observation that has previously been contested,<sup>12,13,18</sup> and shows for the first time that gelatinase granule constituents are also depleted. Our results also show a significant depletion of granule proteins from the monocytes of patients with GPS, supporting findings in *Nbeal2*<sup>-/-</sup> mice.<sup>9</sup>

The majority of CD4 lymphocytes do not contain secretory granules, which may provide an explanation for the underrepresentation of granule proteins in the differential transcriptome and proteome of these cells in patients with GPS. However, GPS CD4 lymphocytes showed the strongest correlation between the transcriptome and the proteome with a significant upregulation of multiple transcripts and their corresponding proteins. This upregulation is enriched in GO terms related to immune responses. For instance, *BTK*, whose protein is critical for B-cell development, was both more highly expressed and its protein differentially more abundant in GPS CD4 lymphocytes compared with those in control subjects. The role of *Btk* in the regulation of T cells may be relevant to the autoimmune disorders seen in our patients with GPS and warrants exploration in further research studies.<sup>57,58</sup>

Although we observed a depletion of granule proteins in GPS neutrophils and monocytes, some granule proteins such as Camp, which has a function in innate immunity, as well as other proteins related to the acute phase response were increased in the patient plasma. Furthermore, more than one-quarter of the discriminatory plasma proteins are not known to be synthesized by blood cells, the majority of which are produced in liver-residing cells. Possible mechanisms to explain this observation in patients with GPS include the systemic activation of the liver by proinflammatory myeloid granule proteins and/or the abnormal infiltration of atypical blood cells into the liver.

In conclusion, our study identifies novel features of GPS, including reduced leukocyte counts, autoimmune diatheses, hypogranularity of myeloid cells, and a proinflammatory plasma

proteome. By evaluating GPS patient phenotypes at both the clinical and cellular level, our study provides further evidence for the role of *Nbeal2* in the granule and cell biology of platelets and immune cells and highlights the prevalence of immune dysregulation not previously associated with this rare bleeding disorder.

## Acknowledgments

The authors thank Rachel Linger (National Institute for Health Research BioResource [NIHR-BR]) for assistance with recruitment of patients and healthy donors, Amy Fray and Carly Kempster (Department of Haematology, National Health Service Blood and Transplant [NHSBT]) for phlebotomy and sample collection/processing, Joana Batista (Department of Haematology, NHSBT) for assistance with patient visits to Cambridge and sample processing, Daniel Duarte (NIHR-BR) for DNA sample preparation of ThromboGenomics samples, Lisa Skeates and Luis Campos (Hematopathology and Oncology Diagnostic Service, Cambridge University Hospitals NHS Trust) for preparation of BM trephine specimens, Pedro Martin-Cabrera for providing control BM trephine specimens (Hematopathology and Oncology Diagnostic Service, Cambridge University Hospitals NHS Trust), Daniel Cutler (University College London) for review of electron microscopy images, James Thaventhiran (University of Cambridge) for evaluation of CD4 lymphocyte data, Parsa Akbari (University of Cambridge) for information provided about FSC and SSC measurements in INTERVAL participants, Paul Sims for editorial review of the manuscript, and Jo Westmoreland for production of the visual abstract. The authors also thank NIHR-BR volunteers for their participation and gratefully acknowledge NIHR-BR centers, NHS Trusts, and staff for their contribution. They also thank the NIHR and NHSBT.

M.C.S and J.H.C. are supported by Medical Research Council Clinical Research Training Fellowships (MR/R002363/1 and MR/P02002X/1). L.M. is supported by a studentship grant from the Rosetrees Trust. T.K.B. was supported by the NHSBT and the British Society for Haematology. D.S. was in part funded by an Isaac Newton Trust/Wellcome Trust Institutional Strategic Support Fund fellowship to M.F. L.K. and A.S. acknowledge support by the Ministerium für Kultur und Wissenschaft des Landes Nordrhein-Westfalen, the Regierende Bürgermeister von Berlin-inkl. Wissenschaft und Forschung, and the Bundesministerium für Bildung und Forschung and A.S. is supported by the Deutsche Forschungsgemeinschaft (DFG, SFB 1116 and TR 240). L.B. was supported by a fellowship from Fondazione Umberto Veronesi. P.G. was supported by a Telethon Foundation Grant (GGP15063). K.F. is supported by the Research Council of the KULeuven (C14/19/096) and by unrestricted grants of the Swedish Orphan Biovitrum AB, Bayer, and CSL Behring. M.F. is supported by the British Heart Foundation (FS/18/53/33863). This work was performed by using resources provided by the Cambridge Service for Data Driven Discovery operated by the University of Cambridge Research Computing Service, provided by Dell EMC and Intel using Tier-2 funding from the Engineering and Physical Sciences Research Council (EP/P020259/1), and DiRAC funding from the Science and Technology Facilities Council.

The views expressed are those of the authors and not necessarily those of the National Health Service, the NIHR, or the Department of Health and Social Care in England. This research was made possible through access to the data and findings generated by the 100 000 Genomes Project. The 100 000 Genomes Project is managed by Genomics England Limited (a wholly owned company of the Department of Health and Social Care). The 100 000 Genomes Project is funded by the NIHR and NHS England. The Wellcome Trust, Cancer Research UK, and the Medical Research Council have also funded research infrastructure. The 100 000 Genomes Project uses data provided by patients and collected by the National Health Service as part of their care and support.

## Authorship

Contribution: M.C.S., L.M., J.H.C., W.H.O., and J.A.G. wrote the manuscript; M.C.S., L.M., J.H.C., T.K.B., K.M., C.L.-B., M.-C.A., W.F.B., L.B.,

E.C., E.D.C., K.G., A.G., P.G., D.H., M.-F.H., A.M.K., R.K., S.L.Q., T.L., E.B.L., A.D.M., S.M., D.N., S.P., S.J.P., J.P., G.M.P., M.-C.P., M.S., H.S., S.S., O.S.-S., R.C.T., J.K.M.W., B.Z., K.F., P.N., and R.F. recruited participants into the study and collected phenotype data; M.C.S., L.M., J.H.C., T.K.B., L.K., F.S.B., S.F., R.D.S., R.M., H.M., R.R., J.C.S., W.N.E., and J.A.G. handled and processed samples and performed experiments; M.C.S., L.M., J.H.C., K.M., D.S., L.K., F.S.B., D.G., D.L., A.R.-R., W.J.A., S.V.V.D., L.G., E.T., T.W.K., K.D., W.N.E., and J.A.G. analyzed data; E.T., A.D.W., A.S., K.D., W.N.E., M.F., W.H.O., and R.F. provided oversight of recruitment processes and analysis pipelines; and W.H.O. and J.A.G. supervised the project.

Conflict-of-interest disclosure: The authors declare no competing financial interests.

A complete list of the members of the NIHR BioResource is provided in supplemental Table 4.

ORCID profiles: M.C.S., 0000-0003-4503-0265; L.M., 0000-0003-4905-0669; J.H.C., 0000-0002-8716-3261; K.M., 0000-0002-2826-3879; D.S., 0000-0002-2066-6980; L.K., 0000-0002-2673-0488; W.J.A., 0000-0001-8866-6672; E.D.C., 0000-0003-0942-2819; S.V.V.D., 0000-0002-0405-4335; K.G., 0000-0002-8934-0700; A.G., 0000-0001-8343-7336; S.L.Q., 0000-0002-6203-3946; E.B.L., 0000-0003-0739-7293; A.D.M., 0000-0002-1746-4819; S.M., 0000-0003-4266-4457; D.N., 0000-0002-1109-0879; S.P., 0000-0002-9222-3812; J.P., 0000-0003-3394-2099; G.M.P., 0000-0002-1791-8905; M.S., 0000-0002-5783-1855; R.C.T., 0000-0003-4842-6783; E.T., 0000-0002-1820-6563; J.K.M.W., 0000-0002-6002-4878; B.Z., 0000-0002-4954-7029; T.W.K., 0000-0002-7421-3370; K.F., 0000-0002-4381-2442; K.D., 0000-0003-0366-1579; M.F., 0000-0001-8074-6299; P.N., 0000-0001-5152-1381; W.H.O., 0000-0002-7744-1790.

## REFERENCES

- Albers CA, Cvejic A, Favier R, et al. Exome sequencing identifies NBEAL2 as the causative gene for gray platelet syndrome. *Nat Genet.* 2011;43(8):735-737.
- Gunay-Aygun M, Falik-Zaccari TC, Vilboux T, et al. NBEAL2 is mutated in gray platelet syndrome and is required for biogenesis of platelet  $\alpha$ -granules. *Nat Genet.* 2011;43(8):732-734.
- Kahr WH, Hinckley J, Li L, et al. Mutations in NBEAL2, encoding a BEACH protein, cause gray platelet syndrome. *Nat Genet.* 2011; 43(8):738-740.
- Lo RW, Li L, Leung R, Pluthero FG, Kahr WH. NBEAL2 (Neurobeachin-Like 2) is required for retention of cargo proteins by  $\alpha$ -granules during their production by megakaryocytes. *Arterioscler Thromb Vasc Biol.* 2018;38(10): 2435-2447.
- Mayer L, Jaszal M, Pardo M, et al. Nbeal2 interacts with Dock7, Sec16a, and Vac14. *Blood.* 2018;131(9):1000-1011.
- Deppermann C, Cherpokova D, Nurden P, et al. Gray platelet syndrome and defective thrombo-inflammation in Nbeal2-deficient mice. *J Clin Invest.* 2013;123(8):3331-3342.
- Guerrero JA, Bennett C, van der Weyden L, et al. Gray platelet syndrome: proinflammatory megakaryocytes and  $\alpha$ -granule loss cause myelofibrosis and confer metastasis resistance in mice. *Blood.* 2014;124(24): 3624-3635.
- Kahr WH, Lo RW, Li L, et al. Abnormal megakaryocyte development and platelet function in Nbeal2(-/-) mice. *Blood.* 2013; 122(19):3349-3358.

- Claushuis TAM, de Stoppelaar SF, de Vos AF, et al. Nbeal2 deficiency increases organ damage but does not affect host defense during gram-negative pneumonia-derived sepsis. *Arterioscler Thromb Vasc Biol.* 2018; 38(8):1772-1784.
- Sowerby JM, Thomas DC, Clare S, et al. NBEAL2 is required for neutrophil and NK cell function and pathogen defense. *J Clin Invest.* 2017;127(9):3521-3526.
- Drube S, Grimlowski R, Deppermann C, et al. The neurobeachin-like 2 protein regulates mast cell homeostasis. *J Immunol.* 2017; 199(8):2948-2957.
- Chedani H, Dupuy E, Massé JM, Cramer EM. Neutrophil secretory defect in the gray platelet syndrome: a new case. *Platelets.* 2006;17(1):14-19.
- Drouin A, Favier R, Massé JM, et al. Newly recognized cellular abnormalities in the gray platelet syndrome. *Blood.* 2001;98(5): 1382-1391.
- Kahr WH, Dror Y. Gray platelet syndrome: macrothrombocytopenia with deficient  $\alpha$ -granules. *Blood.* 2012;120(13):2543.
- Rensing-Ehl A, Pannicke U, Zimmermann SY, et al. Gray platelet syndrome can mimic autoimmune lymphoproliferative syndrome. *Blood.* 2015;126(16):1967-1969.
- Steinberg-Shemer O, Tamary H. Gray platelet syndrome mimicking atypical autoimmune lymphoproliferative syndrome: the key is in the blood smear. *Blood.* 2018;131(24):2737.
- Gunay-Aygun M, Zivony-Elboum Y, Gumruk F, et al. Gray platelet syndrome: natural history of a large patient cohort and locus assignment

to chromosome 3p. *Blood.* 2010;116(23): 4990-5001.

- White JG, Brunning RD. Neutrophils in the gray platelet syndrome. *Platelets.* 2004;15(5): 333-340.
- Downes K, Megy K, Duarte D, et al; NIHR BioResource. Diagnostic high-throughput sequencing of 2396 patients with bleeding, thrombotic, and platelet disorders. *Blood.* 2019;134(23):2082-2091.
- 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.
- Turro E, Astle WJ, Megy K, et al; NIHR BioResource for the 100,000 Genomes Project. Whole-genome sequencing of patients with rare diseases in a national health system. *Nature.* 2020;583(7814):96-102.
- Richards S, Aziz N, Bale S, et al; ACMG Laboratory Quality Assurance Committee. 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.
- Kleinberger J, Maloney KA, Pollin TI, Jeng LJ. An openly available online tool for implementing the ACMG/AMP standards and guidelines for the interpretation of sequence variants. *Genet Med.* 2016;18(11):1165.
- Jarvik GP, Browning BL. Consideration of cosegregation in the pathogenicity classification of genomic variants. *Am J Hum Genet.* 2016;98(6):1077-1081.

Correspondence: Janine H. Collins, Department of Haematology, University of Cambridge, NHS Blood and Transplant Building, Long Road, Cambridge CB2 0PT, United Kingdom; e-mail: jc2018@medschl.cam.ac.uk; and Jose A. Guerrero, Department of Haematology, University of Cambridge, NHS Blood and Transplant Building, Long Road, Cambridge CB2 0PT, United Kingdom; e-mail: jg652@medschl.cam.ac.uk.

## Footnotes

Submitted 10 February 2020; accepted 2 July 2020; prepublished online on *Blood* First Edition 21 July 2020. DOI 10.1182/blood.2019004776.

\*M.C.S., L.M., J.H.C., and T.K.B. contributed equally to this work.

†R.F. and J.A.G. contributed equally to this work.

All newly identified NBEAL2 variants have been deposited in ClinVar (NIHR\_Bioresource\_Rare\_Diseases\_GPS). RNA-seq data have been deposited in the European Genome-phenome Archive (EGAD00001005950). Cell type-specific and plasma proteomics data have been deposited in the Proteomics Identifications Database, PRIDE (PXD016366 and PXD017227, respectively).

25. Association for Clinical Genomic Science. Best Practice Guidelines for Variant Classification in Rare Disease. <https://www.acgs.uk.com/media/11631/uk-practice-guidelines-for-variant-classification-v4-01-2020.pdf>. Accessed 10 May 2020.
26. Adzhubei IA, Schmidt S, Peshkin L, et al. A method and server for predicting damaging missense mutations. *Nat Methods*. 2010;7(4):248-249.
27. Ng PC, Henikoff S. SIFT: predicting amino acid changes that affect protein function. *Nucleic Acids Res*. 2003;31(13):3812-3814.
28. Ioannidis NM, Rothstein JH, Pejaver V, et al. REVEL: an ensemble method for predicting the pathogenicity of rare missense variants. *Am J Hum Genet*. 2016;99(4):877-885.
29. 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.
30. Rentzsch P, Witten D, Cooper GM, Shendure J, Kircher M. CADD: predicting the deleteriousness of variants throughout the human genome. *Nucleic Acids Res*. 2019;47(D1):D886-D894.
31. Davydov EV, Goode DL, Sirota M, Cooper GM, Sidow A, Batzoglou S. Identifying a high fraction of the human genome to be under selective constraint using GERP++. *PLoS Comput Biol*. 2010;6(12):e1001025.
32. Westbury SK, Turro E, Greene D, et al; BRIDGE-BPD Consortium. 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.
33. Arber DA, Orazi A, Hasserjian R, et al. The 2016 revision to the World Health Organization classification of myeloid neoplasms and acute leukemia. *Blood*. 2016;127(20):2391-2405.
34. Lentaingne C, Greene D, Sivapalaratnam S, et al; NIHR BioResource. Germline mutations in the transcription factor IKZF5 cause thrombocytopenia. *Blood*. 2019;134(23):2070-2081.
35. Karczewski KJ, Francioli LC, Tiao G, et al; Genome Aggregation Database Consortium. The mutational constraint spectrum quantified from variation in 141,456 humans. *Nature*. 2020;581(7809):434-443.
36. Caulfield M, Davies J, Denny M, et al. The National Genomics Research and Healthcare Knowledge base. <https://doi.org/10.6084/m9.figshare.4530893.v5>. Accessed 12 May 2020.
37. Ashkenazy H, Abadi S, Martz E, et al. ConSurf 2016: an improved methodology to estimate and visualize evolutionary conservation in macromolecules. *Nucleic Acids Res*. 2016;44(W1):W344-W350.
38. Cullinane AR, Schäffer AA, Huizing M. The BEACH is hot: a LYST of emerging roles for BEACH-domain containing proteins in human disease. *Traffic*. 2013;14(7):749-766.
39. Jogl G, Shen Y, Gebauer D, et al. Crystal structure of the BEACH domain reveals an unusual fold and extensive association with a novel PH domain. *EMBO J*. 2002;21(18):4785-4795.
40. Nurden AT, Nurden P. The gray platelet syndrome: clinical spectrum of the disease. *Blood Rev*. 2007;21(1):21-36.
41. Pluthero FG, Di Paola J, Carcao MD, Kahr WHA. NBEAL2 mutations and bleeding in patients with gray platelet syndrome. *Platelets*. 2018;29(6):632-635.
42. Favier R, Roussel X, Audia S, et al. Correction of severe myelofibrosis, impaired platelet functions and abnormalities in a patient with gray platelet syndrome successfully treated by stem cell transplantation. *Platelets*. 2020;31(4):536-540.
43. Di Angelantonio E, Thompson SG, Kaptoge S, et al; INTERVAL Trial Group. Efficiency and safety of varying the frequency of whole blood donation (INTERVAL): a randomised trial of 45 000 donors. *Lancet*. 2017;390(10110):2360-2371.
44. Karim MA, Suzuki K, Fukai K, et al. Apparent genotype-phenotype correlation in childhood, adolescent, and adult Chediak-Higashi syndrome. *Am J Med Genet*. 2002;108(1):16-22.
45. Rørvig S, Østergaard O, Heegaard NH, Borregaard N. Proteome profiling of human neutrophil granule subsets, secretory vesicles, and cell membrane: correlation with transcriptome profiling of neutrophil precursors. *J Leukoc Biol*. 2013;94(4):711-721.
46. Maynard DM, Heijnen HF, Gahl WA, Gunay-Aygun M. The  $\alpha$ -granule proteome: novel proteins in normal and ghost granules in gray platelet syndrome. *J Thromb Haemost*. 2010;8(8):1786-1796.
47. Maynard DM, Heijnen HF, Horne MK, White JG, Gahl WA. Proteomic analysis of platelet  $\alpha$ -granules using mass spectrometry. *J Thromb Haemost*. 2007;5(9):1945-1955.
48. Zufferey A, Schwartz D, Nolli S, Reny JL, Sanchez JC, Fontana P. Characterization of the platelet granule proteome: evidence of the presence of MHC1 in  $\alpha$ -granules. *J Proteomics*. 2014;101:130-140.
49. Senzel L, Natenko DV, Bahou WF. The platelet proteome. *Curr Opin Hematol*. 2009;16(5):329-333.
50. Takeuchi A, Saito T. CD4 CTL, a cytotoxic subset of CD4<sup>+</sup> T cells, their differentiation and function. *Front Immunol*. 2017;8:194.
51. Grassi L, Izuogu OG, Jorge NAN, et al. Cell type specific novel lncRNAs and circRNAs in the BLUEPRINT haematopoietic transcriptomes atlas [published online ahead of print 23 July 2020]. *Haematologica*. doi: 10.3324/haematol.2019.238147.
52. Lonsdale J, Thomas J, Salvatore M, et al; GTEx Consortium. The genotype-tissue expression (GTEx) project. *Nat Genet*. 2013;45(6):580-585.
53. Battle A, Brown CD, Engelhardt BE, et al; eQTL manuscript working group. Genetic effects on gene expression across human tissues [published correction appears in *Nature*. 2018;553(7689):530]. *Nature*. 2017;550(7675):204-213.
54. Larocca LM, Heller PG, Podda G, et al. Megakaryocytic emperipolesis and platelet function abnormalities in five patients with gray platelet syndrome. *Platelets*. 2015;26(8):751-757.
55. Cunin P, Bouslama R, Machlus KR, et al. Megakaryocyte emperipolesis mediates membrane transfer from intracytoplasmic neutrophils to platelets. *eLife*. 2019;8:e44031.
56. Lee DH, Yao C, Bhan A, et al. Integrative genomic analysis and functional studies reveal GP5, GRN, MPO and MCAM as causal protein biomarkers for platelet traits. *bioRxiv*. 2019; doi:10.1101/854216.
57. Rip J, Van Der Ploeg EK, Hendriks RW, Corneth OBJ. The role of Bruton’s tyrosine kinase in immune cell signaling and systemic autoimmunity. *Crit Rev Immunol*. 2018;38(1):17-62.
58. Xia S, Liu X, Cao X, Xu S. T-cell expression of Bruton’s tyrosine kinase promotes auto-reactive T-cell activation and exacerbates aplastic anemia [published online ahead of print 20 August 2019]. *Cell Mol Immunol*. doi: 10.1038/s41423-019-0270-9.