Dysregulation of oncogenic factors by GFI1B p32: investigation of a novel GFI1B germline mutation

by Michela Faleschini, Nicole Papa, Marie-Christine Morel-Kopp, Caterina Marconi, Tania Giangregorio, Federica Melazzini, Valeria Bozzi, Marco Seri, Patrizia Noris, Alessandro Pecci, Anna Savoia, and Roberta Bottega

Haematologica 2021 [Epub ahead of print]


Publisher’s Disclaimer.
E-publishing ahead of print is increasingly important for the rapid dissemination of science. Haematologica is, therefore, E-publishing PDF files of an early version of manuscripts that have completed a regular peer review and have been accepted for publication. E-publishing of this PDF file has been approved by the authors. After having E-published Ahead of Print, manuscripts will then undergo technical and English editing, typesetting, proof correction and be presented for the authors’ final approval; the final version of the manuscript will then appear in print on a regular issue of the journal. All legal disclaimers that apply to the journal also pertain to this production process.
Dysregulation of oncogenic factors by GFI1B p32: investigation of a novel GFI1B germline mutation

Michela Faleschini¹, Nicole Papa¹, Marie-Christine Morel-Kopp², Caterina Marconi³, Tania Giangregorio¹, Federica Melazzini⁴, Valeria Bozzi⁴, Marco Seri², Patrizia Noris⁴, Alessandro Pecci⁴, Anna Savoia¹,⁵, Roberta Bottega¹

¹Institute for Maternal and Child Health – IRCCS Burlo Garofolo, Trieste, Italy;
²Department of Haematology and Transfusion Medicine, Royal North Shore Hospital and Northern Blood Research Centre, Kolling Institute, University of Sydney, Sydney, Australia;
³Department of Medical and Surgical Sciences, University of Bologna, Bologna, Italy.
⁴Biotechnology Research Laboratories, IRCCS Policlinico San Matteo Foundation, Pavia, Italy;
⁵Department of Medical Sciences, University of Trieste, Trieste, Italy.

Running Heads
Novel GFI1B mutation dysregulates oncogenic factors

Correspondence
Anna Savoia
IRCCS Burlo Garofolo
Department of Medical Sciences, University of Trieste
via Dell'Istria, 65/1
I-34137 Trieste, Italy
Tel +39 040 3785527 - Fax +39 040 3785540
Email: anna.savoia@burlo.trieste.it

Abstract word count: 152
Main text word count: 3015
Number of Tables: 1
Number of Figures: 5
Number of Supplementary files: 1
ABSTRACT

GF11B is a transcription factor essential for the regulation of erythropoiesis and megakaryopoiesis, and pathogenic variants have been associated with thrombocytopenia and bleeding. Analysing thrombocytopenic families by whole exome sequencing, we identified a novel GFI1B variant (c.648+5G>A), which causes exon 9 skipping and overexpression of a shorter p32 isoform. We report the clinical data of our patients and critically review the phenotype observed in individuals with different GFI1B variants leading to the same effect on the p32 expression. Since p32 is increased in acute and chronic leukemia cells, we tested the expression level of genes playing a role in various type of cancers, including hematological tumors and found that they are significantly dysregulated, suggesting a potential role for GFI1B in carcinogenesis regulation. Increasing the number of individuals with GFI1B variants will allow us to better characterize this rare disease and determine whether it is associated with an increased risk of developing malignancies.
INTRODUCTION

Inherited thrombocytopenias represent a group of heterogeneous rare disorders characterized by reduced platelet count that may associate with other defects.\textsuperscript{1} Mutations in at least 40 different genes are responsible for these disorders, including a recently discovered autosomal dominant condition reported as \textit{GFI1B}-associated bleeding disorder or platelet-type bleeding disorder 17 (OMIM 187900), which is caused by pathogenic variants of the \textit{GFI1B} (growth factor-independent 1B) gene.\textsuperscript{2-4}

\textit{GFI1B} is a transcription repressor of the GFI zinc finger family, consisting of three domains: a highly conserved N-terminal repressor domain called Snail/Gfi1, which recruits chromatin regulatory proteins; an intermediary domain of unknown function; and a C-terminal cluster of six zinc-finger domains (ZNF), of which ZNF 3, 4, and 5 bind to DNA, while ZNF 1, 2, and 6 have yet to be characterized.\textsuperscript{5,6} \textit{GFI1B} plays a key role in hematopoiesis through different isoforms. Indeed, in addition to generating a long isoform consisting in 330 amino acids (p37), the \textit{GFI1B} mRNA undergoes an alternative splicing process producing a short isoform of 284 residues (p32) characterized by the skipping of exon 9 (NG_034227) and consequent removal of ZNF 1 and 2. The long p37 form is reported to have a pivotal role in megakaryopoiesis and platelet production, whereas the short p32 form is essential for erythroid lineage.\textsuperscript{7}

\textit{GFI1B} regulates the differentiation of the hematopoietic stem cells through the repression of different promoters, including its self-regulatory regions via an autoregulatory feedback mechanism.\textsuperscript{7-9} Among other targets is the hematopoietic stem cell marker CD34, whose down regulation during hematopoiesis is not quenched when \textit{GFI1B} is mutated. Therefore, one feature of \textit{GFI1B}-associated bleeding disorder is the aberrant expression of CD34 on patients' platelets, which is due to the dysregulation of the respective gene at the transcriptional level.\textsuperscript{10}

Since its first description,\textsuperscript{2} individuals with \textit{GFI1B} mutations have been examined to characterize this novel rare platelet disorder. To our knowledge, these studies have identified at least 15 different mutations in approximately 20 unrelated families, allowing better characterization of the disease, whose features are increased bleeding tendency, thrombocytopenia with enlarged platelets, reduced \(\alpha\)-granule content with abnormal distribution.\textsuperscript{2,3,10-17}

We report a novel heterozygous \textit{GFI1B} variant (c.648+5G>A) in a family with mild thrombocytopenia and no other significant features. The substitution causes skipping of exon 9 and consequent overexpression of the short p32 isoform and dysregulation of \textit{GFI1B} target genes, including CD34 and other genes that are involved in neoplastic transformation, suggesting a potential role for \textit{GFI1B} in carcinogenesis regulation.
METHODS

Family study and clinical features
The propositus (Figure 1), a 65-year-old female with low platelet number, and her family members were studied to determine the molecular basis of thrombocytopenia. Their clinical features, as well as the methods used for blood cell analyses, are described in more detail in Supplementary Information. The institutional review board of the IRCCS “Policlinico San Matteo Foundation” of Pavia approved the study. All subjects provided written informed consent for the study, which was conducted in accordance with the Declaration of Helsinki.

Mutation screening and reverse transcriptase polymerase chain reaction analysis
Mutational screening was performed by whole exome sequencing (WES) in the proband (II-2), as previously reported. Variants were confirmed by Sanger sequencing in her daughters (III-1 and III-2), sisters (II-3 and II-4) and brother (II5) (Table 1). Total RNA was extracted from peripheral blood cells of patients II-2 and III-1, as well as two healthy controls and cDNA amplified as previously reported, using primers enlisted in Table S1 (Supplementary Information).

Bioinformatic analysis
The effect of the splice-site mutation was predicted by in-silico analyses using two dedicated bioinformatic tools: Splice Site Prediction by Neural Network (NNSplice; http://www.fruitfly.org/seq_tools/splice.html) and Human Splicing Finder Version 2.4.1 (http://www.umd.be/HSF/).

Expression vectors and dual-luciferase reporter assay
Wild-type and mutant (c.511_648del due to skipping of exon 9) GFI1B cDNA were cloned into the tagged (myc) expression vector pcDNA3.1 (Invitrogen, Carlsbad, CA, USA). The GFI1B, GFI, MEIS1, and CD34 promoters were cloned into a reporter firefly luciferase vector (pGL4/luc2, Promega). Plasmids were transiently co-transfected in Meg01, as previously described. The luciferase activity was detected using the Dual Luciferase reporter assay system kit (Promega), according to the manufacturer’s instructions. Details of experiments are included in Supplementary Information.

Quantitative Real-Time PCR
Quantitative Real-Time PCR (qPCR) was performed to evaluate the absolute expression of GFI1B isoforms and relative expression of GFI1B major target, using specific primers (Table S2; Supplementary Information). FastStart Universal SYBR Green Master Mix (Roche) and ABI PRISM 7900 detection (Applied Biosystem, Foster City, CA, USA) were used. All experiments were performed three times in triplicate, as previously reported.

Western blot analyses
HEK293T cells were transfected with myc-tagged wild-type and mutant GFI1B cDNA using calcium phosphate method and maintained in DMEM medium (Euroclone, Pero, Italy) with 10% FBS. Protein fractionated cell extracts were analysed by Western blot using primary antibodies anti-myc (9E10; sc-40;
Santa Cruz Biotechnology, Dallas, TX, USA), anti-HSP90 (sc-7947; Santa Cruz Biotechnology), anti-ORC2 (ab68348; Abcam, Cambridge, UK); as previously described.23

Immunofluorescence assay
GF11B transfected in HeLa cells was detected using 9E10 antibody against c-myc (Santa Cruz Biotechnology). Peripheral blood smears were double-labelled with antibodies against CD41 (Santa Cruz Biotechnology) and CD34 (Miltenyi Biotec, Bergisch Gladbach, Germany). General staining and image acquisition have previously been described.24,25 Detailed methods for both immunofluorescence analyses are included in Supplementary Information.
RESULTS

Identification of the novel c.648+5G>A mutation in GFI1B

Whole exome sequencing of proband II-2 revealed a single nucleotide heterozygous substitution (c.648+5G>A) in the GFI1B gene, affecting the splice donor site of intron 9 (Figure 1A). Sanger sequencing confirmed the variant in the proband (II-2), her sister (II-3) and her daughters (III-1 and III-2) but not in her healthy sister (II-5) (Figure 1B). The proband’s brother (II-4) was homozygous for the wild-type allele and the thrombocytopenia was likely due to liver cirrhosis. The c.648+5G>A substitution is reported in GnomAD with a MAF value of 0.000008025 and identified in 2 European (non-Finnish) individuals in heterozygous status. Based on NNSplice software, the c.648+5G>A variant is predicted to drop the score of the splice donor site from 0.89 to 0.24, suggesting potential alternative splicing processes. Therefore, RT-PCR analysis was performed using RNA extracted from affected individuals’ peripheral blood cells (II-2 and III-1) (Figure 1C). In contrast with the healthy control showing the expected product of 246 bp, the patients’ samples amplified an additional band of 108 bp. Sequencing analysis revealed that the lower band corresponded to a transcript characterized by skipping of exon 9, resulting in an in-frame mRNA deletion of 138 bp (p.Val171_Gln216del). The alternative splicing is expected to produce a shorter protein of 32kDa, lacking zinc fingers 1 and 2, corresponding to what was previously described as p32.7,26

Blood cell studies in family members

The major blood parameters of the family members are reported in Table 1. The propositus had moderate thrombocytopenia while the other individuals carrying the heterozygous c.648+5G>A GFI1B mutation had a platelet count only slightly lower or higher than the lower limit of the normal range. Platelet size calculated as MPV by the automated counter was normal in all subjects. The study of MPDs by image analysis of blood smears revealed slightly increased values, indicating mild platelet macrocytosis. The remaining parameters, including hemoglobin concentration, MCV and red blood cell count were within the normal ranges in all investigated subjects. Examination of peripheral blood smears revealed a mild reduction in platelet α-granules and red blood cell anisocytosis in all affected family members. Flow cytometry revealed normal expression of the platelet surface GP complexes Ia-IIa, Iib-IIIa, and Ib-IX-V in individuals II-2 and III-1 (data not shown). Platelet aggregation after stimulation with collagen (4 and 20 µg/ml), ADP (5 and 20 µM,), epinephrine (10 µM), arachidonic acid (1 µM), TRAP (25 µM), and ristocetin (1.5 mg/mL) was normal in individuals II-2 and III-1 carrying the c.648+5G>A variant, and similar to II-4 who is wild-type (data not shown).

Expression level of alternative spliced products

The patient’s cDNA was used to clone the wild-type (330 aa; p37) and the exon 9 skipped (284 aa; p32) forms of GFI1B tagged to myc into an expression vector. Accordingly to what observed during the RT-PCR on patients RNA (Figure 1C), colony screening using primers on exons 8 and10 (Table S1) detected the expected products of 246 bp and of 108 bp, as well as two additional bands of 312 bp and 174 (Figure 2A). Sanger sequencing showed that the additional products retained the last 66 bp (ggatccccgccggggtcatctagcctgacctgcacctgaccccccgggctcatttcctccgcaag) of intron 9 in both p37 and p32 forms due to recognition of a cryptic acceptor splicing site (ggagtgtcctgttccgcagggat) with a score of 0.83.
predicted by NNSpice tool (Figure 2B). Therefore, the two additional products would correspond to GFI1B molecules of expected molecular weight of 39kDa (p39) and 34kDa (p34), respectively for an in-frame insertion of 22 amino acids (p.216_217insGIPAGSSPEPDPAPSPLRQ).

We investigated the expression level of the four forms of GFI1B (Table S2). In control samples (N=2), forms p37 and p32 represents the most expressed products, being approximately 83% and 10%, respectively of the total GFI1B cDNA; the other two, p34 and p39, are overlooked (Figure 2C). In patients, the expression level of p37 and p32 is significantly different than in controls, corresponding to 54% and 32% of the total cDNA, respectively. Of note, p32 is relatively less expressed than p37, suggesting that c.648+5G>A leads to partial skipping of exon 9 or partial degradation of the mRNA. Finally, no significant difference in the expression of p39 and p34 has been seen between patient and control samples.

**Pathogenic role of the c.648+5G>A mutation leading to p32.**

To investigate the effect of c.648+5G>A on GFI1B cellular localization, we performed immunofluorescence and western-blotting analysis in cells transiently transfected with the wild-type or mutant cDNA. Like p37, p32 isoform enters the HeLa cell nucleus (Figure 3A) and both are similarly distributed in the cytoplasmic and nuclear fractions of Hek293 cells (Figure 3B), suggesting that zinc fingers 1 and 2 are dispensable for migration of p32 into the nucleus.

Moreover, we determined the effect on the transcriptional activity, using the luciferase gene under the control of three known GFI1B target gene promoters (MEIS, GFI, and GFI1B itself). The luciferase activity was significantly reduced when the Meg-01 cells were co-transfected with myc-tagged p37, confirming the role of GFI1B as a transcription repressor of those three targets.27 On the contrary, the expression of p32 not only abolished the repression but also increased the transcriptional activity of those three promoters, suggesting that the functional defect is likely due to dominant negative effect of the exogenous p32 on the p37 endogenous transcription factor in Meg-01 cells (Figure 3C). Of note, similar data were obtained using non-tagged p37 and p32 forms generated to test the effect of a different mutation (c.648+1_648+8delGTGGGCAC) of GFI1B.10

**Transcriptional regulation of the CD34 promoter reflects the aberrant expression of CD34 in patients’ platelets**

Previous investigations reported that GFI1B-related thrombocytopenia is associated with aberrant expression of the stem cell antigen CD34 in platelets and megakaryocytes, which could be caused by loss of CD34 downregulation during megakaryocytopenia because of the defective GFI1B function.10 We therefore investigated the transcriptional effect of p32 on the CD34 promoter. Similarly to what observed on MEIS, GFI, and GFI1B promoters, p32 does not repress the CD34 promoter activity (p<0.001) (Figure 4A). Consistent with our *in vitro* findings, immunofluorescence analysis of peripheral blood smears confirmed the aberrant expression of the CD34 antigen in the proband’s platelets (II-2) (Figure 4B).

**p32 dysregulates the transcription of genes involved in oncogenic pathways.**

Since myeloid and lymphoid leukemia cells express higher level of p32 than control cells,26 we hypothesized that p32 could unbalance transcription of genes involved in malignant transformation. For this reason, we investigated the expression level of other GFI1B targets, including proto-oncogenes GFI1 and MYC,
antiapoptotic factors TGFBR3 and BCL2L1, and hematopoietic master regulators, the GATA family members, whose dysregulation is known to occur in oncogenesis. Quantitative PCR performed on patients’ peripheral blood RNA showed that except for GATA2, all the other genes tested (GFI1, BCL2L1, TGFBR3, MYC, GATA1 and GATA3) were significantly overexpressed in patients II-2 and III-1 compared to controls with a fold-change ranging from 1.5 to > 3 (Figure 5), suggesting that patients carrying the c.648+5G>A mutation could have an increased risk for malignancies. In order to confirm the role of p37 and p32 in regulating the oncogenic factors, qPCR was also carried out in HEK cells overexpressing the two GFI1B isoforms. As in patients’ peripheral blood cells, the overexpression of p32 is associated with an increased expression level of the oncogenes, indicating that p32 up-regulates their transcription (Supplementary Figure S1). The only exception is represented by GATA 3, whose expression was significantly decreased in p32-overexpressing cells instead of being increased as in patients’ samples, suggesting that p32 down-regulates this gene at least in the HEK cellular model.
DISCUSSION

GFI1B is a transcriptional repressor that plays a fundamental role in megakaryocyte development and erythropoiesis\(^9\) and pathogenic variants are responsible for the GFI1B-associated bleeding disorder. The novel heterozygous GFI1B splicing variant (c.648+5G>A) identified in our family causes alternative splicing leading to unbalanced expression level of the GFI1B isoforms. Specifically, we detected four alternative splicing events, leading to the three known isoforms, p37, p32 and p39, as well as the novel p34 isoform. The p37 and p32 isoforms are expressed at different level in the individuals carrying the c.648+5G>A substitution, with the p32/p37 ratio 10 times higher in affected than healthy individuals, indicating that this unbalance due to the presence of the c.648+5G>A mutation could be responsible for thrombocytopenia. In fact, considering that p37 has a pivotal role in megakaryopoiesis and platelet production and that p32 is essential for erythroid lineage, distortion of the p32/p37 ratio is likely to dysregulate hemopoiesis. On the contrary, the role of p39 and p34, which represent approximately 3% and 10%, respectively of the total GFI1B mRNA in both specimens of control samples, has not yet been defined.\(^7\)

We further investigated the role of p37 and p32 in terms of cellular localization and ability to control the transcription of targeted genes. In overexpressed conditions, most of the two isoforms are in the nucleus at comparable level, though their transcriptional effect is opposite. Whereas p37 represses the activity of the luciferase that is under the control of GFI1B target promoters, such as MEIS1, GFI1, and GFI1B itself, p32 increased its activity, reversing the transcriptional repression.

The same effect of c.648+5G>A on alternative splicing has previously been reported for c.648+1_648+8delGTGGGCAC, another mutation of GFI1B that was found in members of one family affected with MYH9-related disease.\(^10\) Like in our family, in individuals carrying the GFI1B but not the MYH9 variant, the expression level of p32 was markedly increased, being in at least equal amounts to p37. However, these individuals were not thrombocytopenic (mean value of 228x10^9/L).

Another GFI1B variant, associated with skipping of exon 9 and significant reduction of the p37 expression level, is the synonymous substitution c.576C>T(p.Phe192=) in exon 9, which is reported as rs150813342 with a MAF of 0.0041.\(^28\) Indeed, in individuals heterozygous for c.576C>T the platelet count is reduced in average of 25–30x10^9/L in comparison with controls.

Finally, unbalance between p32 and p37 was identified in association with another mutation (c.551insG/p.Ser185Leufs*3).\(^29\) In patients homozygous for this mutation, whereas the expression level of p37 was significantly decreased, likely due to degradation by nonsense-mediated decay, the expression of p32 was at the same level as in control. Therefore, in these patients the active form of GFI1B is mainly p32. Of note, whereas individuals homozygous for c.551insG had a severe phenotype, those heterozygous were clinically unaffected. Of note, in these patients, as well as in our affected individuals, erythropoiesis was not defective, except for mild anisocytosis, suggesting that red cell production is rather independent of p32/p37 ratio.\(^29\)
Taken together all these observations suggested that the expressivity is extremely variable even when the variants exert the same effect. Indeed, bleeding and platelet count vary significantly, from individuals without defects (c.648+1_648+8delGTGGGCAC) or slight reduction of platelet count (c.576C>T) to those with an intermediate phenotype characterized by mild thrombocytopenia with mild bleeding tendency as in our family (c.648+5G>A) or with severe disease when the c.551insG mutation is homozygous.29 Of note, abnormalities in megakaryocyte maturation were observed in isogenic cell clones homozygous for rs150813342 and in hematopoietic stem cell progenitors in which GFI1B p37 was selectively silenced,28 suggesting that both expressivity and penetrance of the platelet phenotype might depend - at least in the cases enlisted above - on a specific threshold of the p32/p37 ratio, a value that cannot be estimated from the literature or our own data.

One feature that is likely to be pathognomonic of the GFI1B-associated bleeding disorder is the abnormal expression of CD34 on platelet surface.10 Indeed, platelets of affected individual III-2, as well as those mentioned above with altered p32/p37 ratio, express the CD34 antigen. Strong expression has also been reported in association with the Cys168Phe, Arg184Pro, and His181Tyr mutations, all affecting ZNF 1, or H294fsX307, Gln287* and G272fsX2 removing ZNF 5 and 6.3,10–12

Changes in proportion between p37 and p32 significantly increasing p32 levels has been reported in acute and chronic leukemia, which could be interpreted as a triggering event for carcinogenesis.26 This hypothesis is supported by significant up-regulation of GFI1, BCL2L1, TGFBR3, MYC, GATA1, and GATA3 in cells from the patients' peripheral blood. These genes are also differentially expressed in HEK cells overexpressing p32 or p37, confirming that p32 controls their transcription. Considering that these oncogenes are involved in hematopoietic differentiation and are dysregulated in various types of cancer, including hematological malignancies,27 it would be interesting to establish whether the GFI1B-associated bleeding disorder could be part of the nosological entity called “Inherited thrombocytopenia-associated genes with predisposition to neoplasms”.30

Increasing the number of individuals with GFI1B mutations and creating a registry of individuals with GFI1B variants, including a comprehensive medical and family history and regular follow-up, would be of fundamental importance in defining whether there is any increased risk of developing such hematological malignancies. Therefore, any effort should be made to identify patients with this rare form of thrombocytopenia by evaluating the CD34 expression on platelets, a useful assay in defining the pathogenicity of variants that would otherwise be considered of uncertain significance.

ACKNOWLEDGEMENT
This study was supported by IRCCS “Burlo Garofolo” (Ricerca Corrente 01/2018) and AIRC Grant IG-21974. The authors would like to thank patients for their participation in this project and Prof. Gughi D.L. for the constructive discussions.
REFERENCES

<table>
<thead>
<tr>
<th>Subject (Gender£)</th>
<th>GFI1B mutational status</th>
<th>Platelet count (x 10^9/L)</th>
<th>MPV (fL)</th>
<th>MPD (μm)</th>
<th>Hb (g/dL)</th>
<th>MCV (fL)</th>
<th>RBC (x 10^9/μL)</th>
<th>RCDW (%)</th>
<th>Leukocyte count (x 10^9/L)</th>
<th>WHO bleeding score§</th>
</tr>
</thead>
<tbody>
<tr>
<td>II-2 (F)</td>
<td>648+5G&gt;A/+</td>
<td>50</td>
<td>10.7</td>
<td>3.32</td>
<td>13.6</td>
<td>85</td>
<td>4.89</td>
<td>14</td>
<td>6.23</td>
<td>3</td>
</tr>
<tr>
<td>II-3 (F)</td>
<td>648+5G&gt;A/+</td>
<td>149</td>
<td>9.2</td>
<td>2.95</td>
<td>13.4</td>
<td>90.6</td>
<td>4.40</td>
<td>13.7</td>
<td>4.01</td>
<td>3</td>
</tr>
<tr>
<td>III-1 (F)</td>
<td>648+5G&gt;A/+</td>
<td>124</td>
<td>12.6</td>
<td>2.96</td>
<td>12</td>
<td>82.3</td>
<td>4.10</td>
<td>13.4</td>
<td>7.62</td>
<td>0</td>
</tr>
<tr>
<td>III-2 (F)</td>
<td>648+5G&gt;A/+</td>
<td>153</td>
<td>11.7</td>
<td>3.57</td>
<td>13.9</td>
<td>88.3</td>
<td>4.80</td>
<td>15</td>
<td>9.1</td>
<td>0</td>
</tr>
<tr>
<td>II-4 (M)</td>
<td>+/-</td>
<td>94*</td>
<td>10.8</td>
<td>3.09</td>
<td>15.6</td>
<td>98</td>
<td>4.92</td>
<td>13.9</td>
<td>5.7</td>
<td>0</td>
</tr>
<tr>
<td>II-5 (F)</td>
<td>+/-</td>
<td>215</td>
<td>10</td>
<td>3.26</td>
<td>14.6</td>
<td>95.1</td>
<td>4.58</td>
<td>14</td>
<td>5.42</td>
<td>0</td>
</tr>
<tr>
<td>Normal values</td>
<td></td>
<td>150-450</td>
<td>8-13**</td>
<td>2.4-2.7†</td>
<td>11.7-15.5 (F)</td>
<td>13.2-17.3 (M)</td>
<td>82-98</td>
<td>3.8-5.2 (F)</td>
<td>4.4-5.7 (M)</td>
<td>11.6-16</td>
</tr>
</tbody>
</table>

£F, female. M, male

* in this patient thrombocytopenia was likely due to liver cirrhosis

**MPV range given by the automated cell counter

†MPD obtained in 50 healthy volunteers (95% confidence interval)

§World Health Organization (WHO) bleeding scale: 0 - no bleeding tendency; 1 - cutaneous bleeding only (including minimal mucosal bleeding); 2 - mild blood loss (any mucosal bleeding not fulfilling the criteria for grade 1 or 3); 3 - gross blood loss, requiring transfusion; 4 - debilitating blood loss (including retinal or cerebral associated with fatality). See patients and methods for details about bleeding symptoms.
FIGURE LEGENDS

Figure 1 Genetic analyses of the family. A) Sanger sequencing showing exon 9 (e9) and intron 9 (i9) boundary of GFI1B gene. The heterozygous c.648+5G>A substitution is indicated by an arrow. Nucleotide A of the ATG translation initiation start site of the GFI1B cDNA in GenBank sequence NM_004188.6 is indicated as nucleotide +1. B) Family pedigree showing an autosomal dominant pattern of inheritance. Arrow indicates the proband whereas black filled symbols represent affected individuals carrying the mutation. The grey filled symbol indicates a thrombocytopenic brother without the mutation (platelet count in this patient may be influenced by a liver cirrhosis). C) RT-PCR showing different transcription pattern between patients (II-2; III-1) and healthy control (HC). The expected product of 246bp was detected in the HC sample; an additional band of 108 bp corresponding to skipping of exon 9 was found in the affected individuals (M, Molecular weight).

Figure 2. GFI1B splicing isoforms. A) PCR products obtained after amplification of a portion of GFI1B from different plasmids containing subcloned GFI1B cDNAs (individual II-2) using primer aligning on exons 8 and 10. Fragments corresponding to the wild-type (246 bp) and skipped exon 9 (108 bp) forms of GFI1B are shown in lanes 1 and 4, and 2 and 5, respectively. Additional fragments of 312 bp (lane 7) and 174 bp (lanes 3 and 6) corresponding to p39 and p34, respectively, were also detected (M, Molecular weight). B) Schematic structure of the genomic GFI1B gene (gDNA) which includes the coding exons (exons 6-11; NM_004188.6). On bottom, the representation of the full cDNA (p37) and the different isoforms identified. p32 is characterized by skipping of exon 9. p39 and p34 results in p37 and p32, respectively, which also retain the last 66 bp of intron 9 (of which the first and the last are listed in detail) due to recognition of a cryptic acceptor splice site “ag” (in bold). Arrows represent primers used for the amplification. C) Differential expression level of the four GFI1B isoforms between patients (II-2 and III-1) and healthy controls (HC) performed by qPCR using specific primers. Overall the the wild-type isoform (p37) is the most represented in controls while p32 is significantly increased in patients (*p<0.05). Error bars represent the standard deviation of three independent experiments. Statistical analysis was performed using T-test.

Figure 3. Pathogenic role of GFI1B p32. A) Immunofluorescence (IF) and B) Western blot of nuclear (N) and cytoplasmatic (C) fractionated cellular lysates from Hek293 transiently transfected with WT (p37) and mutated (p32) GFI1B expression vector. Results demonstrate that both isoforms are distributed in the nucleus as well as in the cytoplasm with the same proportions. In IF, Propidium Iodide (PI) was used to detect nuclei. HSP90 and ORC2 antibodies were used in WB as loading marker for the cytoplasmic and nuclear fraction, respectively. C) Transcriptional activity of p37 (WT) and the variant p32 GFI1B isoforms measured on GFI, GFI1B and MEIS promoters in Meg01 cells. Results are represented by ratio of the promoter signal (Firefly) and the control promoter (Renilla) and expressed as fold change in luciferase activity. p37 shows the expected repressive activity on targets compared to the empty vector while p32 is an activator (1 to > 1.5) fold in comparison with the p37 for all the three targets (**p<0.001). Error bars represent the standard deviation of three independent experiments. Statistical analysis was performed using T-test.
Figure 4. CD34 expression in c.648+5G>A patients. A) Transcriptional activity of p37 (WT) and the variant p32 GFI1B isoforms measured on the CD34 promoter after transient transfection in Meg01 cells. Results are represented by ratio of the promoter signal (Firefly) and the control promoter (Renilla) and expressed as fold change in luciferase activity. Significative variations occur on CD34 transcriptional levels as a significant p32 impaired repression activity is detected in comparison to p37 active repression (***P<0.001). Error bars represent the standard deviation of three independent experiments. Statistical analysis was performed using T-test. B) Peripheral blood slides of the proband III-2 and healthy volunteers were double-labelled for CD41 (red) and CD34 (green). Platelets were recognized by morphology and CD41 expression. Platelets of the proband III-2 showed a clear aberrant expression of the CD34 antigen compared to healthy control (HC). Scale bars correspond to 5 µm.

Figure 5. Quantitative PCR (qPCR) of GFI1B transcriptional target genes. The GFI, BCL2L1, TGFβR3, MYC, GATA1, and GATA3 mRNA expression levels were significantly increased in patients (II-2 and III-1) compared to controls (two unrelated healthy individuals) (fold change from 1.5 to >3). No significant difference was observed for GATA2. Error bars represent the standard deviation of three independent experiments. Statistical analysis was performed using T-test.
Figure 3

A

PI  MYC  MERGE

p37

<table>
<thead>
<tr>
<th></th>
<th>C</th>
<th>N</th>
<th></th>
<th>C</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>P37</td>
<td>0.29</td>
<td>0.71</td>
<td>P32</td>
<td>0.32</td>
<td>0.68</td>
</tr>
</tbody>
</table>

B

% N/C

<table>
<thead>
<tr>
<th></th>
<th>C</th>
<th>N</th>
<th></th>
<th>C</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>P37</td>
<td></td>
<td></td>
<td>P32</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

C

GFI

GFI1b

MEIS

<table>
<thead>
<tr>
<th></th>
<th>Empty vector</th>
<th>p37</th>
<th>p32</th>
</tr>
</thead>
<tbody>
<tr>
<td>GFI</td>
<td>1</td>
<td>2</td>
<td>1.5</td>
</tr>
<tr>
<td>GFI1b</td>
<td>1</td>
<td>2</td>
<td>1.5</td>
</tr>
<tr>
<td>MEIS</td>
<td>1</td>
<td>1</td>
<td>1.5</td>
</tr>
</tbody>
</table>
Figure 4

A

***

***

CD34

empty vector  p37  p32

B

CD41  CD34

III-2

HC
Figure 5

The bar chart illustrates the fold expression levels of various genes: GFI, BCL2L1, TGFB3, MYC, GATA1, GATA2, and GATA3. The chart compares expression levels between healthy controls (HC) and patients. Significant differences are indicated by *** and **, respectively. The y-axis represents fold expression, and the x-axis lists the genes.
Supplementary Information

Supplementary Methods

Family study and clinical features

The propositus was a 65-year-old female referred for chronic thrombocytopenia discovered at the age of 31 when diagnosed with breast cancer. The degree of thrombocytopenia was mild to moderate, with platelet count always fluctuating between 50 and 100 x10⁹/L. She reported a lifelong history of bleeding tendency, including easy bruising, bleeding from minor wounds, occasional gum bleeding and menorrhagia. She underwent three surgical procedures: quadrantectomy at age 32, bilateral annessiectomy at age 49, and radical mastectomy at age 55 without any bleeding complications. She had three pregnancies, resulting in one early abortion and two vaginal deliveries. The first delivery was complicated with post-partum hemorrhage requiring red blood cells transfusion, whereas the second was carried out without any major bleeding complications.

The proband's oldest daughter (III-1) was found to have mild thrombocytopenia at the age of 14 when in hospital for appendectomy, which was not complicated by excessive bleeding. Thrombocytopenia was persistently mild, ranging from 124 to 135 x10⁹/L. She never presented significant bleeding symptoms and did not report any hemorrhagic complication during two caesarian deliveries. The proband's second daughter (III-2) had a platelet count always around the lower limit of normal (150 - 170 x10⁹/L). She has never reported bleeding symptoms. She had three vaginal deliveries and underwent surgery for carpal tunnel syndrome without any complications.

The proband's sister (II-3) also presented with platelet counts at the lower limit of normal (145-155 x10⁹/L) and no spontaneous bleeding tendency. She had 5 pregnancies, resulting in two miscarriages and three vaginal deliveries without bleeding complications. At age of 55, she received a diagnosis of ovarian cancer that relapsed twice despite surgery and chemotherapy. One of the two surgical procedures for cancer removal was complicated by major hemorrhage requiring red blood cells transfusion, which led us to assign a bleeding score of 3. However, since such a complication can occur also in individuals who are not at risk for hemorrhages due to platelet or coagulation defects, the WHO score of individual II-3 should not regarded as expression of bleeding diathesis.

Of note, during the genetic counselling, the proband (II-2) reported to carry, as well as her sister (II-3), a pathogenic mutation in the BRCA1 gene that was not transmitted to her daughters.

The proband’s brother (II-4), who never had any bleeding symptoms, was incidentally found to have mild thrombocytopenia (120 x10⁹/L) at the age of 55. Platelet count decreased over the last years together with worsening of liver function. A diagnosis of liver cirrhosis was then made. Finally, another proband’s sister (II-5) had no thrombocytopenia or bleeding tendency.

Except for the proband and her sister, the other family members (II-4, II-5, III-1, and III-2) did not develop malignancies except for the proband’s mother (I-2), who, at advance age, was reported to be affected by
myelodysplastic syndrome that was later responsible for her death. She also had chronic thrombocytopenia and suffered from menorrhagia, though she had eight pregnancies resulting in two abortions and six vaginal deliveries, all without bleeding complications. No clinical information was available for I-1, as well as the proband's younger sister and brother, who were not included in this study.

The institutional review board of the IRCCS “Policlinico San Matteo Foundation” of Pavia approved the study. All subjects provided written informed consent for the study, which was conducted in accordance with the Declaration of Helsinki.

**Blood cell studies**

Complete blood cell counts for all subjects were obtained from the same automated cell counter (Cell-Dyn 3700 from Abbott, using the impedance channel of the instrument). Mean Platelet Diameter (MPD) was measured by software-assisted image analysis on May-Grünwald-Giemsa stained blood films, as previously described. Surface expression of platelet glycoproteins (GPs) was investigated by flow cytometry as reported; results were expressed as the percentages of mean fluorescence intensity patient/control run in parallel. Platelet aggregation was measured in platelet rich plasma by the densitometric method of Born, as reported. Platelets were stimulated with collagen (20 and 4 μg/ml) (Mascia Brunelli, Milan, Italy), adenosine diphosphate (ADP, 20 and 5 μM, Sigma-Aldrich, St. Louis, MO, USA), ristocetin (1.5 mg/ml, Sigma-Aldrich), epinephrine 10 μM (Mascia Brunelli), arachidonic acid 1 μM (Sigma Aldrich) and TRAP 25 μM (Sigma-Aldrich). The extent of platelet aggregation was measured 5 minutes after the addition of the stimulating agents. Adenosine triphosphate (ATP) platelet release was measured by lumiaggregometry after stimulation with collagen 20 μg/ml and ADP 20 μM.

**Expression vectors and dual-luciferase reporter assay**

Wild-type (WT) and mutant (c.648+5G>A) GFI1B were amplified from patient's platelet cDNA and cloned into the tagged (myc) expression vector pcDNA3.1 (Invitrogen, Carlsbad, CA, USA). The GFI1B, GFI, MEIS1 and CD34 promoters were generated by PCR from normal human DNA, cloned in the pGEM vector (Promega, Madison, WI, USA) and inserted into a reporter firefly luciferase vector (pGL4/luc2, Promega). Plasmids were transiently co-transfected in Meg01 as previously described. Transfection efficiency was monitored by co-transfection of the Renilla luciferase vector (LucR) acting as internal control and an empty vector was used to clean the background noise due to the possible baseline activation of the promoters. Cell lysates were prepared 48h post-transfection and assayed for luciferase activity using the Dual Luciferase reporter assay system kit (Promega) according to the manufacturer’s instructions. All transfection experiments were performed three times in triplicate (N=9), results were expressed as a ratio of firefly to renilla (LucF/LucR) and normalized to the activity of the empty vector.
**Immunofluorescence assay**

HeLa cells were seeded on chamber slides and transiently transfected with myc-tagged wild-type or mutant GFI1B plasmids. After 16 hours cells were fixed with 4% paraformaldehyde for 20 minutes and then permeabilized with 0.1% Triton X-100. GFI1B was detected using a primary antibody against c-myc (9E10, Santa Cruz Biotechnology) followed by a secondary FITC-labelled anti-mouse antibody (F0479, DakoCytomation – Agilent Technologies, Santa Clara, CA, USA), while nuclei were stained with Propidium Iodide (Sigma Aldrich). Images were obtained with a Nikon C1si confocal microscope, using a 60X Plan Apo objectives. Images were processed for brightness and contrast regulation using ImageJ 1.45 (NIH, Bethesda, USA) as previously described.4

Peripheral blood smears were fixed with methanol/acetone 1:1 at 4°C and double-labelled with a goat antibody against CD41 (Santa Cruz Biotechnology) and the AC136 mouse monoclonal antibody against CD34 (Miltenyi Biotec, Bergisch Gladbach, Germany). The appropriate chicken Alexa Fluor 594- or Fluor 488-conjugated anti-goat or anti-mouse were used as secondary antibodies (Invitrogen). General staining procedures have been already described.5 Blood smear of the proband was stained and analyzed in parallel with blood smears of three healthy volunteers. Images were acquired through an Axioscope 2 Plus microscope (Carl Zeiss, Gottingen, Germany) equipped with a AxioCam MRc5 camera (Carl Zeiss).

**Supplementary Results**

**Figure S1**

![Graph](image)

Figure S1. Quantitative PCR (qPCR) of GFI1B transcriptional target genes in HEK cells. The expression level of GFI, BCL2L1, TGFBR3, MYC, GATA1 and GATA3, the genes up-regulated in patients blood cells was evaluated in HEK cells overexpressing p37 or p32. TGFBR3 was not detected at valuable Ct and therefore was not included in the graph. Error bars represent the standard deviation of three independent experiments. Statistical analysis was performed using T-test.
Supplementary Tables

Table S1

<table>
<thead>
<tr>
<th>Primer</th>
<th>Sequence: 5'→ 3'</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Genomic primers for Sanger sequencing</strong></td>
<td></td>
</tr>
<tr>
<td>4F</td>
<td>TCCCCCTCTCTCACATCCAAC</td>
</tr>
<tr>
<td>4R</td>
<td>ACCCCTCCCAATCTGTGTC</td>
</tr>
<tr>
<td><strong>Cloning primers</strong></td>
<td></td>
</tr>
<tr>
<td>F_kpn_myc</td>
<td>TAAGGTACCATGGAACAAAAACTCAGAAGAGGATCTGATGCCACGTCCTCCTCTG</td>
</tr>
<tr>
<td>R_not</td>
<td>TATGCGGCCGCCTCAGATTGTGCTG</td>
</tr>
<tr>
<td><strong>cDNA primers for RT</strong></td>
<td></td>
</tr>
<tr>
<td>GFI1b_F cDNA</td>
<td>TACCACGTGTGAAGTGCAAC</td>
</tr>
<tr>
<td>GFI1b_R cDNA</td>
<td>CACCTGCTCATCCACTCAGAC</td>
</tr>
</tbody>
</table>

Table S2

<table>
<thead>
<tr>
<th>Gene</th>
<th>GenBank</th>
<th>Primer Forward (5'-3')</th>
<th>Primer Reverse (5'-3')</th>
</tr>
</thead>
<tbody>
<tr>
<td>GFI1b_p37</td>
<td>NM_004188.6</td>
<td>TACCACGTGTGAAGTGCAAC</td>
<td>CTGCGCTCCTGGAAGGTG</td>
</tr>
<tr>
<td>GFI1b_p32</td>
<td>NM_004188.6</td>
<td>CAACCTATGGAACAAAACACTCAGAAGGATCTGATGCCACGTCCTCCTCTG</td>
<td>GAAGGCTGCGCTCCTGAGTTG</td>
</tr>
<tr>
<td>GFI1b_p39</td>
<td>NM_004188.6</td>
<td>TGAAGTGCAACAAAGGGGAAT</td>
<td>GTCTGAGTGATGAGCACAGTG</td>
</tr>
<tr>
<td>GFI1b_p34</td>
<td>NM_004188.6</td>
<td>ACCTGCGCTCCTGGAAGGTG</td>
<td>GTCTGAGTGATGAGCACAGTG</td>
</tr>
<tr>
<td>GFI1</td>
<td>NM_005263</td>
<td>ACCCCTGTCATGACTGTTCCAGAAGG</td>
<td>GAAGGCTGCGCTCCTGAGTTG</td>
</tr>
<tr>
<td>BCL2L1</td>
<td>NM_138578</td>
<td>GCAACTGGAAGGTCTGCTGCTTGT</td>
<td>GGCTGGGATACTCTTGTGG</td>
</tr>
<tr>
<td>TGFB3</td>
<td>NM_003239.2</td>
<td>AAGAATCCATAATTCAGACATGGAC</td>
<td>AAGGTTTCCGCTCAGTTG</td>
</tr>
<tr>
<td>MYC</td>
<td>NM_002467</td>
<td>AATGAAAAAGGCAGAAAGTTATCCAGG</td>
<td>GAAGGCTGCGCTCCTGAGTTG</td>
</tr>
<tr>
<td>GATA1</td>
<td>NM_002049</td>
<td>CCAAGCTTCTGGAAGACTCTC</td>
<td>ATTTCTCTGAAACAGG</td>
</tr>
<tr>
<td>GATA2</td>
<td>NM_032638.4</td>
<td>GTCACGTGAGGAAGAGGAGG</td>
<td>AGTTCCCTGCTGAGAG</td>
</tr>
<tr>
<td>GATA3</td>
<td>NM_002051.2</td>
<td>CTCATTAAGCAGCAGAACAG</td>
<td>CAGAAGCAGAAAG</td>
</tr>
<tr>
<td>TUBB1</td>
<td>NM_030773.3</td>
<td>GAGTATGAATCTCCTCAGA</td>
<td>CACCTGCTGCTTCAGAG</td>
</tr>
</tbody>
</table>

**GF1B: NM_004188.6**

**GFI1b: NM_004188.6**
Supplementary Bibliography


