GATA1 pathogenic variants disrupt MYH10 silencing during megakaryopoiesis

Paul Saultier, Sandrine Cabantous, Michel Puceat, Franck Peiretti, Timothée Bigot, Noémie Saut, Jean-claude Bordet, Matthias Canault, Johannes Agthoven, Marie Loosveld, et al.

1 GATA1 pathogenic variants disrupt MYH10 silencing during

2 megakaryopoiesis.

3 Paul Saultier,[1,2] Sandrine Cabantous,[1] Michel Puceat,[3] Franck Peiretti,[1] Bigot Timothée,[1] Noémie

4 Saut,[1,4] Jean-Claude Bordet,[5] Matthias Canault,[1] Johannes van Agthoven,[6] Marie Loosveld,[4,7]

5 Dominique Payet-Bornet[7], Delphine Potier[7], Céline Falaise,[2,4] Denis Bernot,[1] Pierre Emmanuel

6 Morange,[1,4] Marie-Christine Alessi,[1,4 *] and Marjorie Poggi[1]

7 [1]Aix Marseille Univ, INSERM, INRAe, C2VN, Marseille, France

8 [2]APHM, La Timone Children's Hospital, Department of pediatric hematology, immunology and

9 oncology, Marseille, France

10 [3]Aix Marseille Univ, INSERM, MMG, Marseille, France

11 [4]APHM, CHU Timone, French Reference Center on Inherited Platelet Disorders, Marseille, France

12 [5]Unité d'Hémostase Biologique, Bron, France

13 [6]Structural Biology Program, Division of Nephrology/Department of Medicine, Massachusetts

14 General Hospital and Harvard Medical School, Charlestown, MA 02129, USA

15 [7]Aix-Marseille Univ, CNRS, INSERM, CIML, Marseille, France.

16

17 *___Corresponding author___ : Marie-Christine Alessi, C2VN, Faculté de Médecine Timone, 27 Boulevard

18 Jean Moulin, 13385 Marseille, France ; E-mail : marie-christine.alessi@univ-amu.fr ; Tel : +33 491

19 324 507 ; Fax : +33 491 254 336.

20

21 **Abstract: 238 words**

22 **Text: 5002 words**

23 **Figures: 5**

24 **Tables: 3**

25 **Supporting Information: 9 figures and 2 tables**

26 **References: 41**

27 **Abstract**

28 **Background**

29 GATA1 is an essential transcription factor for both polyploidization and megakaryocyte (MK)

30 differentiation. The polyploidization defect observed in *GATA1* variant carriers is not well understood.

31 **Objective**

32 To extensively phenotype two pedigrees displaying different variants in the *GATA1* gene and

33 determine if GATA1 controls MYH10 expression levels, a key modulator of MK polyploidization.

34 **Method**

35 146 unrelated *propositi* with constitutional thrombocytopenia were screened on a multigene panel.

36 We described the genotype-phenotype correlation in GATA1 variant carriers and investigated the

37 effect of these novel variants on MYH10 transcription using luciferase constructs.

38 **Results**

39 The clinical profile associated with the p.L268M variant localized in the C terminal Zinc finger was

40 unusual in that the patient displayed bleeding and severe platelet aggregation defects without early-

41 onset thrombocytopenia. p.N206I localized in the N terminal Zinc finger was associated, on the other

42 hand, with severe thrombocytopenia (15G/L) in early life. High *MYH10* levels were evidenced in

43 platelets of GATA1 variant carriers. Analysis of MKs anti-GATA1 ChIP-sequencing data revealed two

44 GATA1 binding sites, located in the 3' untranslated region and in intron 8 of the *MYH10* gene.

45 Luciferase reporter assays showed their respective role in the regulation of MYH10 gene expression.

46 Both GATA1 variants significantly alter intron 8 driven MYH10 transcription.

47 **Conclusion**

48 The discovery of an association between MYH10 and GATA1 is a novel one. Overall, this study

49 suggests that impaired *MYH10* silencing via an intronic regulatory element is the most likely cause of

50 GATA1-related polyploidization defect.

51

52 **Key words**

53 GATA1, MYH10, downstream targets, functional variants, platelets.

54 **Essentials**

55 • MYH10 silencing is an essential step in the megakaryocyte polyploidization.

56 • Two novel GATA1 variants were identified in the N and C-terminal zinc finger domain.

57 • MYH10 expression levels are increased in platelets from both GATA1 variant carriers.

58 • GATA1 alters MYH10 expression via an intronic binding site in the MYH10 gene.

59

60   **Main Text**

61   **Introduction**

62   GATA1 is an essential transcription factor for erythrocyte and megakaryocyte (MK) differentiation

63   [1,2]. It belongs to the GATA family of zinc finger (ZF) transcription factors, which recognize the

64   W(A/T)GATAR(A/G) DNA pattern. The protein is encoded by the *GATA1* gene, located on the X

65   chromosome [3,4]. The GATA1 protein contains two highly conserved ZFs, which are separated by

66   a short linker. The C-terminal ZF binds with high affinity and specificity to (A/T)GATA(A/G) motifs

67   throughout the genome, while the N-terminal ZF can bind protein cofactors, such as FOG1.

68   Several hemizygous loss-of-function *GATA1* variants have been described in patients with varying

69   degrees of macrothrombocytopenia and anemia. Most of these variants affect the N-terminal ZF of

70   GATA1 and impair the GATA1 DNA binding interaction with the GATA1 cofactor FOG1 [5–7].

71   Accordingly, mice with targeted mutations in the *Gata1* promoter display macrothrombocytopenia, in

72   which reduced Gata1 expression [8–10] or conditional knockout of *Gata1* results in a selective loss

73   of Gata1 expression in MKs [11]. This observation underscores the importance of GATA1 in

74   sustaining normal megakaryocytic differentiation and platelet production. Previous studies have

75   shown that *GATA1* is required for both polyploidization and MK differentiation. Bone marrow from

76   *GATA1* variant carriers shows a moderate to marked increase in MK numbers, which are also small

77   and hypolobulated [5,6].*In vivo*, the Gata1^{low} mice display a marked increase in small and hypoploid

78   MKs [9]. GATA1-deficient MKs also exhibit increased proliferation both *in vitro* and *in vivo* with

79   defective nuclear and cytoplasmic maturation [9,10].

80   Some studies have suggested that the polyploidization defect is associated with GATA1 inactivation

81   through deregulation of cyclin D1, STAT1 or CDC6 expression [10,12]. However, we propose an

82   alternative mechanism involving altered silencing of the protein non-muscle heavy chain IIB (MYH10).

83   During polyploidization, expression of MYH10 protein, which localizes to the actomyosin contractile

84   ring, is repressed by the transcription repressor RUNX1, thereby enabling the switch from mitosis to

85   endomitosis [13]. The polyploidization defect observed in *RUNX1* variant carriers is in part associated

86   with the persistence of MYH10 expression [14], which has been detected in circulating platelets [15].

3

87   Similarly, increased levels of MYH10 protein have been detected in platelets from patients carrying

88   *FLI1* variants, thus suggesting cooperation between the two transcription factors, RUNX1 and FLI1,

89   in normal repression of the *MYH10* gene [15–18].

90   Genome-wide analysis of simultaneous GATA1, RUNX1, FLI1 and TAL1 DNA binding in human MKs

91   has revealed that these transcription factors bind to the same binding sites in the genome and

92   cooperate to regulate the expression of target genes during the early stages of megakaryopoiesis

93   [19]. Transcriptional cooperation between RUNX1 and GATA1 in the activation of megakaryocytic

94   promoters has been further established using reporter assays with transient transfection in a variety

95   of mammalian cell lines. These studies have demonstrated the physical interaction between the two

96   transcription factors [20,21]. Therefore, we suggest that deregulation of *MYH10* expression may also

97   be implicated in *GATA1*-related thrombocytopenia.

98   In this study, we report a detailed analysis of two novel GATA1 variants localized in each ZF. Elevated

99   MYH10 protein levels were found in platelets from carriers of each variant. We describe the

100   identification of a *MYH10* intronic region that functions as a GATA1-responsive cis-regulatory

101   element.

102

4

103 **Materials and Methods**

104 Methods concerning high-throughput gene sequencing, structural model of GATA1-DNA interactions,

105 flow cytometry analysis, cell expression and luciferase reporter assays, platelet-rich plasma (PRP)

106 serotonin level, platelet PAI-1 level, mepacrine uptake and release, immunoprecipation and western

107 blot analysis, and epifluorescence microscopy are described in the Online Supplementary Material.

108 _Study limitations:_ Owing to the limited number of patients, statistical comparison between patients

109 and controls was not always possible. Fresh whole blood (required for primary cell culture) was

110 available in only one patient (F1-II1). Consequently, most in vitro experiment was performed using

111 transfected non-hematopoietic cell line. The platelet serotonin levels were available for two patients

112 and the BM smears for one patient. TEM analysis of platelets were performed for only one family.

113 The others experiments were performed for all the patients.

114 **Patient enrollment, research ethics, patient consent, and clinical and laboratory data**

115 From 2014 to 2017, we recruited 146 unrelated _propositi_ with constitutional thrombocytopenia of

116 unknown origin. Genetic analyses were performed at the French Reference Center for Inherited

117 Platelet Disorders at La Timone University Hospital in Marseille, France. All cases were included in

118 the study after obtaining informed written consent in accordance with protocols approved by national

119 institutional review boards and the Declaration of Helsinki principles. Medical and family history data

120 were obtained from medical reports and patient interviews. Bleeding tendency was assessed using

121 the International Society on Hemostasis and Thrombosis-Bleeding Assessment Tool (ISTH-BAT)

122 score. Red blood cell counts, platelet counts, mean corpuscular volume (MCV) and mean platelet

123 volume (MPV) were determined on EDTA-anticoagulated blood samples using the ADVIA 120

124 automated cell counter. Bone marrow smears were obtained for one propositus and stained with May-

125 Grünwald-Giemsa. Fetal hemoglobin levels were assessed using high-performance liquid

126 chromatography in the Biochemistry Department at La Timone University Hospital in Marseille,

127 France.

Platelet aggregation

Platelet aggregation was assessed by measuring light transmission through the stirred platelet-rich plasma (PRP) suspensions for 3 min using an APACT 4004 aggregometer. Platelet aggregation was triggered by adding 1 µM or 2.5 µM adenosine diphosphate (ADP; Helena laboratories), 0.5 mg/ml arachidonic acid (Helena laboratories) and 0.36 µg/ml or 0.72 µg/ml collagen (Bio/Data Corporation). Platelet agglutination was triggered by adding 1.5 mg/ml ristocetin (Sigma-Aldrich).

Evaluation of platelet granules using electron microscopy

Transmission electron microscopy (TEM) analysis of platelets was performed as previously described [18]. Briefly, we assessed platelet surface area ($µm^2$) and diameter (µm), the number of alpha granules per unit area ($µm^{-2}$), and the diameter of alpha granules in a blinded case/control experiment. Dense granules were quantified using whole-mount electron microscopy as previously described [18].

***In vitro* megakaryocyte differentiation and ploidy**

After density gradient separation (Eurobio), circulating CD34+ cells were purified using positive selection with magnetic beads (Miltenyi-Biotec) and cultured in StemSpan Serum-Free Expansion Medium II supplemented with Megakaryocyte Expansion Supplement (Stemcell Technologies). At days 9 and 11 of culture, MKs were analyzed via flow cytometry for CD41 and CD42a markers (anti-CD41-APC and anti-CD42a-PE antibodies, BD Biosciences) as well as for DNA ploidy using Hoechst-33342 (Sigma-Aldrich). The cells were then analyzed using a Navios cytometer (Beckman Coulter).

Statistical analyses

Quantitative variables are expressed as the mean ± standard error. Analyses were performed using GraphPad Prism software. Statistical differences were determined via ANOVA with Dunnett's multiple comparison test. $P<0.05$ was considered statistically significant.

Results

Two novel *GATA1* variants identified via genetic analysis

Next-generation sequencing yielded the identification of two novel *GATA1* variants (NM_002049) in three boys from two unrelated families (Figures 1A-B). For the first family, only one hemizygous variant c.617A>T (p.N206I) causing alteration in the N-terminal ZF was evidenced in the proband (II-1) (Fig 1A). For the second family only one deleterious variant c.802C>A (p. L268M) was shared by the two brothers in the C terminal ZF GATA1 domain. Five other variants worn by the two brothers have been identified. None of them were pathogenic due to benign *in silico* prediction and/or high allelic frequency (>10⁻⁴) in general population databases (Supplemental Table 1). The analysis of the coverage of these genes was also not in favor of deletion or duplication (data not shown). These two missense variants were absent from the GnomAD and ExAC databases, and have a Combined Annotation Dependent Depletion (CADD) score of 24.9 for c.617A>T (Request: Chromosome X, Position 48792341, CADD GRCh38-v1.6) and 23.7 for c.802C>A (Request: Chromosome X, Position 48793229, CADD GRCh38-v1.6). The chart prediction for these two variants based on different prediction tools to identify disease relevant non-synonymous single nucleotide variants is given in supplementary figure (Supplemental Figure 1A and B). The variants occurred in both mothers, while as expected, the two fathers were homozygous for wild type genotypes (Figure 1A).

Structural modeling

Structural analysis provided further insight into the mechanisms by which these two variants may exert deleterious effects. An initial sequence alignment of GATA1 orthologs and paralogs revealed that both variants affect highly conserved sequences and may therefore play a structural role in GATA1 binding to target DNA sequences or protein binding partners (Supplemental Figure 1). As shown in the X-ray structure of GATA1 in complex with double-stranded DNA, the L268 residue is located in the interface between GATA1 and the major groove of the DNA-binding sequence TGATAA (Figure 1D, left), making van der Waals contact with bases 32T, 33T, 9T and 10C (distance cutoff of 4Å). Structural modeling of the p.L268M substitution did not reveal any steric hindrance at the GATA1-

178 DNA interface. However, the loss-of-function effect of p.L268M likely results from distorted

179 hydrophobic packing with the aforementioned bases (shown in the left inset of Figure 1D).

180 Although the N206 residue is not localized within the GATA1-DNA interface (Figure 1D, left), the

181 p.N206I substitution may affect the capacity of GATA1 to bind to FOG1. The nuclear magnetic

182 resonance structures of the GATA1-FOG1 complex revealed that the N206 residue is part of the

183 interface between the two proteins, forming a specific amino-π interaction between the GATA1 side

184 chain amino group and the aromatic ring of FOG1-Y30 (Figure 1D, right). Structural modeling showed

185 that this interaction is disrupted in the N2016I variant. The variant caused destabilization of the

186 GATA1/FOG1 interface by steric hindrance, likely resulting in FOG1-mediated loss of function (shown

187 in the right inset of Figure 1D).

Nuclear localization and FOG1 binding in GATA1 cases

189 To further investigate the subcellular localization of the GATA1 variants, we expressed the wild-type

190 and variant forms of GATA1 in MSR and H9C2 cells. GATA1 variants did not reduced GATA1 protein

191 levels (data not shown). The GATA1 variants entered the nucleus to the same extent as wild-type

192 GATA1 (Supplemental Figure 2A), thereby suggesting that the variants alter GATA1 binding

193 properties to chromatin rather than subcellular localization. The structural prediction was confirmed

194 via GATA1 co-immunoprecipitation using MSR cells overexpressing GATA1 (WT or variants) and

195 FOG1. FOG1 was detected in the GATA1 precipitates of wild-type and p.L268M GATA1-expressing

196 cells, while FOG1 detection was markedly reduced in the p.N206I GATA1 precipitates (Supplemental

197 Figure 2B).

Case descriptions

199 Thrombocytopenia onset kinetics markedly differed between the two families. Thrombocytopenia was

200 present early in life in the p.N206I carrier (F1-II-1) with a severely reduced platelet count (15 G/L) at

201 the age of 18 months that persisted. Bleeding during dental avulsion was prevented by platelet

202 concentrate transfusion prior to the procedure. The patient was referred to our reference center on

203 platelet disorders because of resistance to immunoglobulin, corticosteroids and mycophenolate

204 mofetil treatment, as immune thrombocytopenia was initially suspected. The F1 proband did not suffer

205 from anemia (hemoglobin 118 g/L at seven years of age), and his mean corpuscular volume was

206 rather low (76.8 fL- normal range: 80-95 fL). May-Grunwald-Giemsa staining of bone marrow smears

207 showed mild dyserythropoiesis with erythroblastosis and nuclear karyorrhexis (Figure 2C).

208 By contrast, thrombocytopenia was initially absent in the two F2 brothers carrying the p.L268M

209 variant. At 18 and 9 months of age, F2-II-1 and F2-II-2 exhibited normal platelet counts (261 G/L and

210 233 G/L), respectively despite pronounced mucocutaneous bleedings. The initial hemostasis

211 assessment in the two children showed no defects in coagulation (normal prothrombin and activated

212 partial thromboplastin times, normal levels of the factors VIII, XIII, fibrinogen and von Willebrand

213 factor) or fibrinolysis (normal alpha2 antiplasmin, t-PA and plasma PAI-1 levels and normal von Kaulla

214 test) (data not shown). Bleeding time was prolonged (>15 min) and was not corrected by DDAVP

215 injection. A severe platelet aggregation defect was detected in both brothers with normal platelet

216 counts in PRP samples (338-380G/L). On several occasions (6 times for the older brother and 3 times

217 for the younger brother) and in the absence of medication, collagen, arachidonic acid and TXA2

218 agonist U-46619 failed to induce platelet aggregation. Reduced platelet aggregation was also

219 observed in response to ADP, while ristocetin-induced agglutination was almost normal (Table 1 and

220 Supplemental Figure 3). Both boys were diagnosed with hereditary platelet dysfunction. Diagnosis of

221 Glanzmann thrombasthenia has been evoked but platelet exploration has refuted it. Indeed, platelet

222 aggregation was only slightly reduced upon stimulation with high doses ADP (10 µM) and TRAP-14

223 (50 µM). Exploration of GPIIbIIIa with flow cytometry showed a normal expression level of GPIIbIIIa

224 (Table 2) and a significant increment of PAC-1 fixation after 10 µM ADP stimulation (MFI: basal state=

225 0.31 (normal value 0.21-0.49) stimulated state =6.23 (normal value: 5.4-16.2), positive cell: basal

226 state= 0% (normal value <5%), stimulated state =73% (normal value >70%)). In addition, we did not

227 detect variations, deletion or duplication in the *ITGA2B* and *ITGB3* genes and in genes involved in its

228 activation (*FERMT3, RASGRP2, TLN1*). The altered aggregation profiles was not due to defective

229 expression of GPVI, SYK, PLCγ2 or COX1 in platelets, with no major differences in signal intensity

230 between the two parents and the children. A 30% reduction in TBXAS1 levels was obseverd in both

231 F2 children compared to their parents (Supplemental Figure 4). In patient F2-II-2, dental avulsions

232 and an appendectomy were performed after platelet concentrate transfusions without hemorrhagic

233 complications. At five years of age, the platelet counts for patient F2-II-1 remained in the normal range

234 (194 x 10^9 /L), while his brother's platelet counts started to decrease at three years of age (131 x

235 10^9/L). Thrombocytopenia then gradually developed over time in both brothers, with more severe

236 thrombocytopenia in the younger boy (F2-II-2) (Figure 2A). TEM analysis of platelets derived from

237 patient F2-II-1 showed that, in keeping with previous measurements of mean platelet volume obtained

238 using hematology analyzers, the platelet size distribution was normal (Figure 3B and Supplemental

239 Table).

240 As GATA1 is critical for erythropoiesis, we also aimed to identify any red blood cell defects.

241 Examination of peripheral F2-II-2 blood smears demonstrated abnormally shaped red blood cells with

242 anisocytosis and poikilocytosis (Figure 2B). The F2 probands did not suffer from anemia either, with

243 hemoglobin levels at 111 and 112 g/L at five years of age and 146 and 144 g/L at the date of last

244 follow-up (at 27 and 23 years of age, respectively). Hemoglobin electrophoresis revealed, as already

245 reported in several GATA1 related diseases, an elevated level of fetal hemoglobin. They both

246 presented with elevated fetal hemoglobin levels (25%), albeit without microcytosis (MCV: 87-92 fL

247 between five and eight years of age) and hypochromia (mean corpuscular hemoglobin concentration:

248 330-331 g/L between five and eight years of age). Remarkably, the two F2 brothers displayed a

249 progressive increase in MCV over time, while their platelet counts gradually decreased (Figure 2A);

250 the values at last follow-up were 99 and 103 fL (at 26 and 23 years of age, respectively). Table 3

251 resumes main characteristics of GATA1 variant carriers.

252 **Platelets from GATA1 cases present severe dense granular defect**

253 Electron microscopy revealed a major delta granular deficit in GATA1 variant carriers. Analysis via

254 whole-mount electron microscopy revealed reduced numbers of dense granules, with an average of

255 0.44 ± 1.02 (F1-II-1), 0.56 ± 0.97 (F2-II-1) and 0.13 ± 0.49 (F2-II-2) dense granules per platelet

256 (normal values: 5.46 ± 3.91). Furthermore, 77% (F1-II-1), 65% (F2-II-1) and 90% (F2-II-2) of platelets

257 lacked dense granules (Figure 3A). Mepacrine uptake and release upon TRAP-14 stimulation (40

258 µM) and quantification of mepacrine-labeled dense granules via fluorescent microscopy confirmed

259 the dense granule defect in all variant carriers (Supplemental Figure 5). On several occasions, we

260 observed low platelet serotonin levels for the two F2 brothers (F2-II-1: 0.12 µg/109 platelets in 2015,

261 0.10 µg/109 platelets in 2009 and F2-II-2: 0.16 µg/109 platelets in 2015, normal range: 0.3-1.2 µg/109

262 platelets). Flow cytometry showed reduced expression levels of the dense granule marker CD63 upon

263 TRAP-14 stimulation in all variant carriers (Table 2).

264 Platelet count (Table 1), platelet aggregation (Supplemental Figure 3), flow cytometric evaluation of

265 mepacrine uptake and release (not shown) and platelet serotonin levels (0.39 µg/109 platelets,

266 normal range: 0.3-1.20 µg/109 platelets) were normal in the mother (F2-I-1).

267 **Platelets from GATA1 cases present mild alpha granular defect and structural**

268 **abnormalities**

269 Assessment of peripheral blood smears from F1-II-1 and F2-II-2 did not show gray platelets, although

270 an experienced hematologist identified some hypogranular platelets in both children (Figure 2B).

271 Quantitative morphometric evaluation of F2-II-1 alpha granules indicated a distinct decrease in alpha

272 granule density and consequently a decrease in the alpha granule area: total platelet area ratio (4.1%,

273 reference range 11.0-13.5%). The diameter of the alpha granules was only slightly decreased. The

274 platelets presented many other structural anomalies. Platelet sections showed highly variable

275 granular and mitochondrial content, with large areas of membrane complexes that in some cases

276 occupied the entire section (See Figure 3B legend for complete description). To further assess the

277 alpha granules, we evaluated the levels of PAI-1 antigen in serum. Indeed, 95% of circulating PAI-1

278 is localized in platelets 4. Furthermore, proteomic analysis has found that reduced PAI-1 levels is a

279 feature of gray platelet syndrome5. For the three patients, PAI-1 levels were in the normal range (F1-

280 II-1: 0.35, F2-II-1: 0.76 and F2-II-2: 0.69, normal range 0.33-1.07 ng/106 platelets). Flow cytometry

281 showed moderate decrease in the levels of the alpha granule protein marker CD62P upon TRAP-14

282 stimulation in both the F2 children's whereas CD62P did not change after TRAP-14 stimulation in the

283 F1-II-1 (Table 2).

284

11

GATA1 variants are associated with abnormal megakaryocyte maturation and polyploidization.

Microscopic evaluation of bone marrow aspirate smears from patient F1-II-1 showed normal numbers of MKs with a high proportion of round or hypolobulated polyploid nuclei (Figure 2C) indicating abnormal maturation. We further analyzed MK development using blood-derived CD34$^+$ stem cells. The F1 carriers exhibited a 12- to 48-fold increase in circulating CD34$^+$CD38$^+$CD45^{low}CD19^{low} cells (n = 2) compared with healthy controls (n = 5) (Figure 4A). Analysis of CD41$^+$ and CD42a$^+$ cells at days 9 and 11 revealed a marked reduction in MK differentiation with reduced cell size compared with two controls (Figures 4B-C). The CD41$^+$CD42a$^+$ population was almost absent in the F1-II-1 cells (Figure 4B). Furthermore, the percentage of high-ploidy cells (≥ 4n) in the F1-II-1 sample was reduced at day 9. The relative frequencies of 2n, 4n, 8n and 16n ploidy MKs were 54-58%, 23-27%, 4-8% and 1-2% in the controls, versus 66%, 18%, 0% and 0% in the patient carrying the GATA1 variant, respectively (Figure 4D). Bone marrow smears and MK cultures were not available for the F2 family.

GATA1 cases display increased MYH10 levels in platelets

To further investigate the effect of the *GATA1* variants on MYH10, which is required for MK polyploidization and maturation, MYH10 levels were assessed in platelets derived from patients and healthy subjects. MYH10 protein was not detected in healthy donor platelets via immunoblotting. By contrast, MYH10 expression was detected in platelets from the three probands and one carrier of a *FLI1* variant as previously reported [18]. MYH10 protein was also present in the mother of the F2 probands (F2-I-1), albeit at relatively reduced levels (Figure 4E). The polyploidization defect has been shown associated through deregulation of cyclin D1-Cdk4 or CDC6 expression [10,12]. In MSR cells, GATA1 wild-type or mutants' transfection did not altered their expression levels (Supplemental figure 6).

Evidence of a *MYH10* intronic regulatory region

Our data suggests that GATA1 regulates *MYH10* expression. Therefore, we obtained GATA1, RUNX1, and FLI1 ChIP-sequence data on primary human MKs (deposited by the Göttgens laboratory) from the Gene Expression Omnibus (GEO) database [19]. The peak enrichments on

12

312 chromosome 17, which includes the *MYH10* locus, were extracted and displayed in the UCSC

313 genome browser. This analysis showed two enriched binding regions for the GATA1, RUNX1, and

314 FLI1 transcription factors in the *MYH10* gene (displayed from ENCODE data using the

315 erythroleukemia K562 cell line known to express GATA1) (Figure 5A).

316 The first binding peak (Chr. 17: 8472947-8473157) located in the *MYH10* 3'UTR corresponded to

317 GATA1 and RUNX1 binding sites. Remarkably, sequence analysis did not retrieve the consensus

318 GATA1 binding motif in this region. The second peak (Chr. 17: 8553047-8553739) was located within

319 an intronic region and contained potential binding sites for GATA1, FLI1 and RUNX1. Four potential

320 GATA1 binding sites (referred to as BS-1, BS-2, BS-3 and BS-4) were identified, of which three

321 contain the reverted sequence TATC (BS-2, BS-3 and BS-4) (Supplemental Figure 7). We also

322 uploaded the chromatin interaction analysis with paired-end tag sequencing data for CTCF and POL2

323 performed on the K562 cell line from ENCODE, using the UCSC genome browser. This analysis

324 revealed that the intronic enhancer was indeed not occupied by Pol2 or CTCF (data not shown). The

325 interaction between gene regulatory regions (displayed with data from ENCODE) revealed that the

326 intronic site was likely within a close chromatin region in a DNA repressive loop between two CTCF

327 enriched regions (data not shown).

328 **Transcriptional activity of the *MYH10* intronic element**

329 We compared the 3'UTR and intronic regulatory element-driven expression of *MYH10* using a

330 luciferase reporter assay in two cell lines expressing or not the GATA1 hematopoietic partners. The

331 HEL cell line expresses the GATA1, RUNX1 and FLI1 transcription factors. In HEL cells, both the

332 3'UTR and intronic regions significantly repressed luciferase expression compared with the pGL3

333 empty vector (control) (Figure 5B). Removal of only one of the four GATA1 binding sites in MYH10

334 intron 8 region was sufficient to yield markedly reduced intronic activity.

335 We then look at the effect of GATA1 variant on intron 8 driven MYH10 expression. Since HEL cells

336 express endogenous GATA1 we used a non-hematopoietic model, the MSR epithelial cells which do

337 not express basal levels of GATA1, RUNX1 and FLI1 (Supplemental Figure 7). In this model wild-

338 type GATA1 transfection enhanced luciferase activity via the MYH10 intronic region (Figure 5C left),

339   but not via the 3'UTR region (Figure 5C right). Disruption of all four GATA binding sites in the intronic

340   region yielded a significant reduction in luciferase activity (approximately 40%). Disruption of BS-2

341   (M2) or BS-4 (M4) also led to a significant reduction in luciferase activity, while disruption of BS-1

342   (M1) or BS-3 (M3) did not exhibit any significant effect (Figure 5C left).

343   We then evaluated the variants effects in MSR cells. Both here reported (N206L and L268M GATA1)

344   and already described (V205M and G208S) GATA1 variants were associated with a marked reduction

345   in transcriptional activity compared with wild-type GATA1 (Figure 5D and supplemental figure 8).

346

347   **Discussion**

348   In this study, we assessed two pedigrees displaying a variant in the *GATA1* gene. We report the first

349   description of a variant in the C-terminal ZF of GATA1 that initially manifested as mucocutaneous

350   bleeding with platelet dysfunction but without thrombocytopenia, which appeared later in life. We

351   observed increased MYH10 levels in platelets, thus highlighting the persistence of MYH10 expression

352   during MK differentiation, which may explain the decrease in MK ploidy and reduced platelet

353   formation. Assessing published ChIP-sequence data and using reporter cell and mutagenesis

354   assays, we demonstrated that GATA1 repressed *MYH10* expression via intronic binding sites in the

355   *MYH10* gene.

356   To date, nearly all reported GATA1 variants have been shown to involve the N-terminal ZF (positions

357   205, 208, 216 and 218) and the intron/exon boundaries of exon 2 [22]; with the exception of the

358   terminal p.X414R variant, which cause removal of the terminal codon and adds 42 extra amino acids

359   to the full-length GATA1 protein [23]. Variants in the N-terminal ZF and at position 414 are associated

360   with thrombocytopenia, while variants in exon 2 essentially lead to changes in the red blood cell

361   lineage, leading to severe anemia or Diamond-Blackfan anemia. Currently, variants in the C-terminal

362   ZF have not been described. The p.H289R variant, localized immediately 3' of the C-terminal ZF, has

363   been reported to be not as damaging as variants affecting the N-terminal ZF domain [24].

364   In this study, we describe two novel *GATA1* variants. The patient harboring the variant at position 206

365   suffered from moderate bleeding with severe thrombocytopenia and mild dyserythropoietic features.

14

366   This observation agrees with previous descriptions of N-terminal ZF variants. Accordingly, this

367   variation impairs FOG1 binding. The second variant identified affected the C-terminal ZF. Unlike

368   variations in the N-terminal ZF, the p.L268M variation does not affect the GATA1/FOG1 interaction.

369   The clinical characteristics resemble those observed with the p.R216Q variant localized in the DNA-

370   binding surface of the N-terminal ZF, including macrothrombocytopenia, red blood cell abnormalities

371   and minor beta-thalassemia [25–28]. Presentation was also close to the one described for the H289R

372   variation. As in the F2 patients, the hemizygous males harbouring the H289R substitution exhibit

373   increased MCV but normal hemoglobin levels. Thrombocytopenia was present only in two of the four

374   affected family members [24]. Although both F2 children presented with high level of fetal hemoglobin,

375   the absence of thrombocytopenia did not initially direct towards a GATA1 related disease. Indeed,

376   the disease initially manifested as severe mucocutaneous bleeding associated with reduced platelet

377   aggregation without thrombocytopenia. Platelet exploration carried out on several occasions did not

378   make it possible to link the platelet phenotype to a known thrombopathy. Indeed, GPIIbIIIa was

379   expressed at a normal level and was able to bind PAC1 antibody after activation by 10µM ADP. No

380   sign of immune deficiency was observed in these children excluding a kindlin 3 deficiency [29]. Finally,

381   the results of platelet aggregation were not in favor of a CalDAG-GEFI deficiency as platelet activation

382   induced by a high concentration of arachidonic acid, identical to the one used in this article usually

383   gives normal results [30]. For a long time, these children were considered to be carriers of a storage

384   pool disease because of the early detection of a severe deficit of dense granules. These results are

385   to be connected with those of White *et al* [31] who observed a near absence of dense bodies in the

386   platelets from a patient with the R216Q GATA1 mutation. Although rescue experiments have not be

387   performed and causality has not been shown, our results support the idea that the spectrum of

388   *GATA1*-related diseases may be extended to include X-linked bleeding diathesis without

389   thrombocytopenia, associated with collagen and arachidonic acid aggregation defects. In contrast to

390   studies in Gata1-deficient mice [9,11,27,32] we did not observe reduced protein levels of GPVI, SYK

391   or other key molecules involved in collagen signaling pathways, such as PLCγ2 that could explain

392   the aggregation defect. The decrease in TBXAS1 levels observed in the two F2 brothers may play a

393 role in the reduced response to arachidonic acid. However, the platelets also failed to respond to

394 U46619, thus indicating a more complex mechanism. The granule deficiency observed in the current

395 study may also contribute to the platelet aggregation defects, although a complete lack of arachidonic

396 acid- or collagen-induced platelet aggregation is not a feature of granule deficiency. Overall, further

397 studies are required to understand the cause of platelet dysfunction observed in *GATA1* variant

398 carriers. Thrombocytopenia developed gradually over time associated with an enlargement in red

399 blood cells, indicating progression of the disease in p.L268M carriers. Cell-type-specific gene

400 regulation may change during ontogeny. There are substantial developmental differences between

401 neonates and adults in the process of megakaryopoiesis and in their responses to TPO [33]. Wenqing

402 and colleagues have highlighted fundamental and conserved differences at distinct developmental

403 stages, characterized by simpler promoter-centric regulation of cell-type-specific genes in embryonic

404 cells and increased combinatorial enhancer-driven control in adult cells [34]. Overall, the delayed

405 onset of thrombocytopenia observed in the F2 family suggest developmental stage-specific GATA-1

406 regulation.

407 Assessment of MKs in *GATA1* variant carriers revealed polyploidization and differentiation defects.

408 Assessment of the p.N206I variant revealed a marked decrease in mature CD41^{high}CD42a^{high} MKs,

409 along with an increase in CD41$^-$CD42a$^-$ cells. Although we cannot formally exclude impaired

410 expression of CD41 and CD42a in MKs harboring a GATA1 variant, the reduction in high-ploidy cells

411 confirmed the MK maturation defect.

412 Additionally, the GATA1 N206I carrier displayed an increase in circulating CD34+ progenitor cells, as

413 described for individuals carrying a germline ETV6 variants [27]. Recently, Shin *et al.* [36] have found

414 that Gata1 low megakaryocyte-erythroid progenitors/megakaryocyte precursors cell lines display

415 accelerated replication after approximately 90–100 days of culture. The observed increase in CD34+

416 cells may reflect this high proliferative capacity.

417 The polyploidization defect associated with GATA1 is not well understood. MYH10 has been shown

418 to be required for cytokinesis during meiotic cell division [37] and to contribute to MK polyploidisation

419 [13,38]. RUNX1-mediated silencing of *MYH10* has been found to be required for MK polyploidization

420 and may serve as a biomarker of *RUNX1* variants in patients [15]. FLI1 forms a complex with RUNX1

421 to regulate megakaryopoiesis. Carriers of *FLI1* variants also exhibit elevated MYH10 levels in

422 platelets [15–17]. Interestingly, using a genome-wide analysis of simultaneous transcription factor

423 binding in human MKs and ENCODE ChIP-sequence data, we have shown that FLI1, RUNX1 and

424 GATA1 accumulate at regulatory elements in intron 8 and the 3'UTR of the *MYH10* gene [19].

425 Additionally, these transcription factors cooperate during megakaryopoiesis as components of a large

426 transcriptional complex [21,39], thereby suggesting that GATA1 may regulate MYH10 expression in

427 cooperation with RUNX1 and FLI1. To our knowledge, *MYH10* has not yet been described as a target

428 gene of GATA1. We observed abnormal persistence of MYH10 protein in the platelets of the three

429 *GATA1* variant carriers. The mother of the two F2 brothers also exhibited a slight increase in MYH10

430 levels. In a reporter luciferase assay using the HEL cell line, both regulatory elements in intron 8 and

431 the 3'UTR of *MYH10* were found be functional by inducing reduced transcriptional activity. Disruption

432 of the four intronic binding sites fully reversed the repressive activity, thus indicating that the intronic

433 binding sites are required for *MYH10* repression, despite the presence of the other repressive

434 complex partners (i.e., RUNX1, FLI1). The 3'UTR region of *MYH10* may preferentially bind to a portion

435 of a protein complex that includes GATA1, without binding directly to GATA1. While GATA1 wild-type

436 enhanced luciferase activity via the MYH10 intronic region in MSR cells, the opposite effect was

437 observed in HEL cells. GATA1 is known to have both activating and repressing activity depending on

438 the cellular and promoter/enhancer context [40]. Variations in the expression of GATA1 partners in

439 these two cell models, such as RUNX1 and FLI1, may explain these differences. In MSR cells, the

440 GATA1 responsive activity decreased in presence of each variant, compared with wild-type *GATA1*,

441 indicating that the increase in MYH10 levels observed in patients likely results from reduced

442 interaction between the product of the *GATA1* variants and the *MYH10* intronic region. Intronic

443 GATA1 binding sites have already been shown to be involved in the regulation of 5-aminolevulinate

444 synthase 2 (*ALAS2*) gene expression, which is involved in heme synthesis in erythroid cells [41].

445 In summary, we describe two novel pathogenic variations in the *GATA1* gene characterized by distinct

446 clinical and laboratory findings. Our results suggest that variants affecting the GATA1 C-terminal ZF

17

447 should likely be assessed in patients with normal platelet counts and platelet aggregation defects in

448 response to collagen and arachidonic. Importantly, Both GATA1 variants analyzed in this study led

449 to persistent expression of MYH10 in platelets. We identify a *MYH10* intronic region that functions as

450 a GATA1-responsive cis-regulatory element.

451

18

452 **Acknowledgments**

459

PS, SA, and MaP designed and performed the experiments, analyzed the data and wrote the manuscript. DP, DPB, MiP, FP and MC analyzed the data and revised the manuscript. NS performed DNA sequencing and revised the manuscript. JCB performed TEM analysis and revised the manuscript. JvA performed the structure analysis experiments and revised the manuscript. DB, ML, CF and PEM performed clinical and biological characterization of patients, analyzed the data and revised the manuscript. MP and MCA supervised the study and wrote the manuscript.

468 **Declaration of interest**

469 The authors have declared that no conflict of interest exists.

470

471 **References**

472 1 Crispino JD. GATA1 in normal and malignant hematopoiesis. *Semin Cell Dev Biol* 2005; **16**:
473 137–47.

474 2 Crispino JD, Horwitz MS. GATA factor mutations in hematologic disease. *Blood* 2017; **129**:
475 2103–10.

476 3 Songdej N, Rao AK. Hematopoietic transcription factor mutations: important players in
477 inherited platelet defects. *Blood* 2017; **129**: 2873–81.

478 4 Ko LJ, Engel JD. DNA-binding specificities of the GATA transcription factor family. *Mol Cell Biol*
479 1993; **13**: 4011–22.

480 5 Mehaffey MG, Newton AL, Gandhi MJ, Crossley M, Drachman JG. X-linked thrombocytopenia
481 caused by a novel mutation of GATA-1. *Blood* 2001; **98**: 2681–8.

482 6 Nichols KE, Crispino JD, Poncz M, White JG, Orkin SH, Maris JM, Weiss MJ. Familial
483 dyserythropoietic anaemia and thrombocytopenia due to an inherited mutation in GATA1. *Nat*
484 *Genet* 2000; **24**: 266–70.

485 7 Chang AN, Cantor AB, Fujiwara Y, Lodish MB, Droho S, Crispino JD, Orkin SH. GATA-factor
486 dependence of the multitype zinc-finger protein FOG-1 for its essential role in
487 megakaryopoiesis. *PNAS* National Academy of Sciences; 2002; **99**: 9237–42.

488 8 Shivdasani RA, Fujiwara Y, McDevitt MA, Orkin SH. A lineage-selective knockout establishes
489 the critical role of transcription factor GATA-1 in megakaryocyte growth and platelet
490 development. *EMBO J* 1997; **16**: 3965–73.

491 9 Vyas P, Ault K, Jackson CW, Orkin SH, Shivdasani RA. Consequences of GATA-1 deficiency in
492 megakaryocytes and platelets. *Blood* 1999; **93**: 2867–75.

493 10 Muntean AG, Pang L, Poncz M, Dowdy SF, Blobel GA, Crispino JD. Cyclin D-Cdk4 is regulated by
494 GATA-1 and required for megakaryocyte growth and polyploidization. *Blood* 2007; **109**: 5199–
495 207.

496 11 Meinders M, Hoogenboezem M, Scheenstra MR, De Cuyper IM, Papadopoulos P, Németh T,
497 Mócsai A, van den Berg TK, Kuijpers TW, Gutiérrez L. Repercussion of Megakaryocyte-Specific
498 Gata1 Loss on Megakaryopoiesis and the Hematopoietic Precursor Compartment. *PLoS ONE*
499 2016; **11**: e0154342.

21

500 12 Mazzi S, Lordier L, Debili N, Raslova H, Vainchenker W. Megakaryocyte and polyploidization.
501 *Exp Hematol* 2018; **57**: 1–13.

502 13 Lordier L, Bluteau D, Jalil A, Legrand C, Pan J, Rameau P, Jouni D, Bluteau O, Mercher T, Leon
503 C, Gachet C, Debili N, Vainchenker W, Raslova H, Chang Y. RUNX1-induced silencing of non-
504 muscle myosin heavy chain IIB contributes to megakaryocyte polyploidization. *Nat Commun*
505 2012; **3**: 717.

506 14 Bluteau D, Glembotsky AC, Raimbault A, Balayn N, Gilles L, Rameau P, Nurden P, Alessi MC,
507 Debili N, Vainchenker W, Heller PG, Favier R, Raslova H. Dysmegakaryopoiesis of FPD/AML
508 pedigrees with constitutional RUNX1 mutations is linked to myosin II deregulated expression.
509 *Blood* 2012; **120**: 2708–18.

510 15 Antony-Debré I, Bluteau D, Itzykson R, Baccini V, Renneville A, Boehlen F, Morabito M, Droin
511 N, Deswarte C, Chang Y, Leverger G, Solary E, Vainchenker W, Favier R, Raslova H. MYH10
512 protein expression in platelets as a biomarker of RUNX1 and FLI1 alterations. *Blood* 2012; **120**:
513 2719–22.

514 16 Stockley J, Morgan NV, Bem D, Lowe GC, Lordkipanidzé M, Dawood B, Simpson MA,
515 Macfarlane K, Horner K, Leo VC, Talks K, Motwani J, Wilde JT, Collins PW, Makris M, Watson
516 SP, Daly ME, UK Genotyping and Phenotyping of Platelets Study Group. Enrichment of FLI1
517 and RUNX1 mutations in families with excessive bleeding and platelet dense granule secretion
518 defects. *Blood* 2013; **122**: 4090–3.

519 17 Stevenson WS, Rabbolini DJ, Beutler L, Chen Q, Gabrielli S, Mackay JP, Brighton TA, Ward CM,
520 Morel-Kopp M-C. Paris-Trousseau thrombocytopenia is phenocopied by the autosomal
521 recessive inheritance of a DNA-binding domain mutation in FLI1. *Blood* 2015; **126**: 2027–30.

522 18 Saultier P, Vidal L, Canault M, Bernot D, Falaise C, Pouymayou C, Bordet J-C, Saut N, Rostan A,
523 Baccini V, Peiretti F, Favier M, Lucca P, Deleuze J-F, Olaso R, Boland A, Morange PE, Gachet C,
524 Malergue F, Fauré S, et al. Macrothrombocytopenia and dense granule deficiency associated
525 with FLI1 variants: ultrastructural and pathogenic features. *Haematologica* 2017; **102**: 1006–
526 16.

527 19 Tijssen MR, Cvejic A, Joshi A, Hannah RL, Ferreira R, Forrai A, Bellissimo DC, Oram SH,
528 Smethurst PA, Wilson NK, Wang X, Ottersbach K, Stemple DL, Green AR, Ouwehand WH,
529 Göttgens B. Genome-wide analysis of simultaneous GATA1/2, RUNX1, FLI1, and SCL binding in
530 megakaryocytes identifies hematopoietic regulators. *Dev Cell* 2011; **20**: 597–609.

531 20 Xu G, Kanezaki R, Toki T, Watanabe S, Takahashi Y, Terui K, Kitabayashi I, Ito E. Physical
532 association of the patient-specific GATA1 mutants with RUNX1 in acute megakaryoblastic
533 leukemia accompanying Down syndrome. *Leukemia* 2006; **20**: 1002–8.

534 21 Elagib KE, Racke FK, Mogass M, Khetawat R, Delehanty LL, Goldfarb AN. RUNX1 and GATA-1
535 coexpression and cooperation in megakaryocytic differentiation. *Blood* 2003; **101**: 4333–41.

536   22   Kobayashi E, Shimizu R, Kikuchi Y, Takahashi S, Yamamoto M. Loss of the Gata1 gene IE exon
537   leads to variant transcript expression and the production of a GATA1 protein lacking the N-
538   terminal domain. *J Biol Chem* 2010; **285**: 773–83.

539   23   Singleton BK, Roxby DJ, Stirling JW, Spring FA, Wilson C, Poole J, Anstee DJ. A novel GATA1
540   mutation (Stop414Arg) in a family with the rare X-linked blood group Lu(a-b-) phenotype and
541   mild macrothrombocytic thrombocytopenia. *Br J Haematol* 2013; **161**: 139–42.

542   24   Pereira J, Bento C, Manco L, Gonzalez A, Vagace J, Ribeiro ML. Congenital dyserythropoietic
543   anemia associated to a GATA1 mutation aggravated by pyruvate kinase deficiency. *Ann
544   Hematol* 2016; **95**: 1551–3.

545   25   Yu C, Niakan KK, Matsushita M, Stamatoyannopoulos G, Orkin SH, Raskind WH. X-linked
546   thrombocytopenia with thalassemia from a mutation in the amino finger of GATA-1 affecting
547   DNA binding rather than FOG-1 interaction. *Blood* 2002; **100**: 2040–5.

548   26   Balduini CL, Pecci A, Loffredo G, Izzo P, Noris P, Grosso M, Bergamaschi G, Rosti V, Magrini U,
549   Ceresa IF, Conti V, Poggi V, Savoia A. Effects of the R216Q mutation of GATA-1 on
550   erythropoiesis and megakaryocytopoiesis. *Thromb Haemost* 2004; **91**: 129–40.

551   27   Hughan SC, Senis Y, Best D, Thomas A, Frampton J, Vyas P, Watson SP. Selective impairment of
552   platelet activation to collagen in the absence of GATA1. *Blood* 2005; **105**: 4369–76.

553   28   Åström M, Hahn-Strömberg V, Zetterberg E, Vedin I, Merup M, Palmblad J. X-linked
554   thrombocytopenia with thalassemia displays bone marrow reticulin fibrosis and enhanced
555   angiogenesis: comparisons with primary myelofibrosis. *Am J Hematol* 2015; **90**: E44-48.

556   29   Robert P, Canault M, Farnarier C, Nurden A, Grosdidier C, Barlogis V, Bongrand P, Pierres A,
557   Chambost H, Alessi M-C. A novel leukocyte adhesion deficiency III variant: kindlin-3 deficiency
558   results in integrin- and nonintegrin-related defects in different steps of leukocyte adhesion. *J
559   Immunol* 2011; **186**: 5273–83.

560   30   Canault M, Ghalloussi D, Grosdidier C, Guinier M, Perret C, Chelghoum N, Germain M, Raslova
561   H, Peiretti F, Morange PE, Saut N, Pillois X, Nurden AT, Cambien F, Pierres A, van den Berg TK,
562   Kuijpers TW, Alessi M-C, Tregouet D-A. Human CalDAG-GEFI gene (RASGRP2) mutation affects
563   platelet function and causes severe bleeding. *J Exp Med* 2014; **211**: 1349–62.

564   31   White JG, Thomas A. Platelet structural pathology in a patient with the X-linked GATA-1,
565   R216Q mutation. *Platelets* 2009; **20**: 41–9.

566   32   Ferreira R, Ohneda K, Yamamoto M, Philipsen S. GATA1 function, a paradigm for transcription
567   factors in hematopoiesis. *Mol Cell Biol* 2005; **25**: 1215–27.

568   33   Liu Z-J, Italiano J Jr, Ferrer-Marin F, Gutti R, Bailey M, Poterjoy B, Rimsza L, Sola-Visner M.
569   Developmental differences in megakaryocytopoiesis are associated with up-regulated TPO
570   signaling through mTOR and elevated GATA-1 levels in neonatal megakaryocytes. *Blood* 2011;
571   **117**: 4106–17.

23

572 34 Cai W, Huang J, Zhu Q, Li BE, Seruggia D, Zhou P, Nguyen M, Fujiwara Y, Xie H, Yang Z, Hong D,
573   Ren P, Xu J, Pu WT, Yuan G-C, Orkin SH. Enhancer dependence of cell-type-specific gene
574   expression increases with developmental age. *Proc Natl Acad Sci U S A* 2020; **117**: 21450–8.

575 35 Poggi M, Canault M, Favier M, Turro E, Saultier P, Ghalloussi D, Baccini V, Vidal L, Mezzapesa
576   A, Chelghoum N, Mohand-Oumoussa B, Falaise C, Favier R, Ouwehand WH, Fiore M, Peiretti F,
577   Morange PE, Saut N, Bernot D, Greinacher A, et al. Germline variants in ETV6 underlie reduced
578   platelet formation, platelet dysfunction and increased levels of circulating CD34+ progenitors.
579   *Haematologica* 2017; **102**: 282–94.

580 36 Shin E, Jeong J-G, Chung H, Jung H, Park C, Yoon SR, Kim T-D, Lee SJ, Choi I, Noh J-Y. The
581   Gata1low murine megakaryocyte-erythroid progenitor cells expand robustly and alter
582   differentiation potential. *Biochem Biophys Res Commun* 2020; **528**: 46–53.

583 37 Yang F, Wei Q, Adelstein RS, Wang PJ. Non-muscle myosin IIB is essential for cytokinesis during
584   male meiotic cell divisions. *Dev Biol* 2012; **369**: 356–61.

585 38 Badirou I, Pan J, Legrand C, Wang A, Lordier L, Boukour S, Roy A, Vainchenker W, Chang Y.
586   Carboxyl-terminal-dependent recruitment of nonmuscle myosin II to megakaryocyte
587   contractile ring during polyploidization. *Blood* 2014; **124**: 2564–8.

588 39 Eisbacher M, Holmes ML, Newton A, Hogg PJ, Khachigian LM, Crossley M, Chong BH. Protein-
589   protein interaction between Fli-1 and GATA-1 mediates synergistic expression of
590   megakaryocyte-specific genes through cooperative DNA binding. *Mol Cell Biol* 2003; **23**: 3427–
591   41.

592 40 Welch JJ, Watts JA, Vakoc CR, Yao Y, Wang H, Hardison RC, Blobel GA, Chodosh LA, Weiss MJ.
593   Global regulation of erythroid gene expression by transcription factor GATA-1. *Blood* 2004;
594   **104**: 3136–47.

595 41 Zhang Y, Zhang J, An W, Wan Y, Ma S, Yin J, Li X, Gao J, Yuan W, Guo Y, Engel JD, Shi L, Cheng T,
596   Zhu X. Intron 1 GATA site enhances ALAS2 expression indispensably during erythroid
597   differentiation. *Nucleic Acids Res* 2017; **45**: 657–71.

598 **Figure legend**

599

600 **Figure 1. Identification of two novel GATA1 variants.**
601 A) Pedigrees of the affected families. Squares denote males, and circles represent females. The
602 slashes denote deceased family members. Solid black symbols represent affected family members
603 carrying a hemizygous *GATA1* variant. Circles with a dot represent heterozygous female carriers of
604 the variants. Solid gray symbols represent non-tested family members with a medical history of
605 bleeding events.
606 (B) Sanger sequencing electropherograms of the indicated family members.
607 (C) Schematic diagram of the GATA1 protein domain structure. The functional N-terminal
608 transactivation domain (TD) and both the N- and C-terminal zinc-finger domains (ZF1 and ZF2,
609 respectively) are depicted. The positions of the variations within GATA1 are indicated in red
610 (variations reported in this study) and black (variations reported in previous studies).

24

611 (D) Left: Figure. Structure of Zn fingers of GATA1 bound to palindromic DNA recognition site (PDB
612 code: 3VD6). The diagram shows GATA1 in green cartoon with key residues L268 and N206 in yellow
613 sticks, and the DNA fragment in magenta sticks. Zn2+ ions are in grey spheres with stick
614 representation of coordinating residues. Other GATA1 relevant variants V205M and G208R are
615 mapped in orange sticks. Inset A show 3VD6 superimposed to the same structure harboring a model
616 of the L268M mutation. Right: Lowest energy conformer of NMR structure of Zn fingers of GATA1
617 bound to FOG1 (PDB code: 1Y0J). The diagram shows GATA1 in green cartoon with key residue
618 N206 in yellow stick, and FOG1 in red. Zn2+ ions are in grey spheres with stick representation of
619 coordinating residues. Other GATA1 relevant variants V205M and G208R are mapped in orange
620 sticks. Inset A shows 1Y0J superimposed to the same structure harboring a model of the N206I
621 mutation. In dashes is shown an atomic clash between I206 (GATA1) and Y30 (FOG1) with
622 interatomic distances below 2.5 Å.

623 **Figure 2. Quantitative and morphological modifications of platelets and red blood cells in**
624 **GATA1 variant carriers**
625 (A) Evolution of platelet count and mean corpuscular volume (MCV) of erythrocytes with age in the
626 two F2 probands.
627 (B) **Representative images of peripheral blood smears** stained with May-Grünwald-Giemsa for
628 patients F1-II-1 and F2-II-2 (at 6 and 23 years old, respectively). The peripheral blood smear images
629 revealed hypogranular platelets (left, black arrow, magnification x1000) mild anisocytosis and
630 poikylocytosis (right, magnification x 500).
631 (C) **Representative images of bone marrow aspirate smears** stained with May-Grünwald-Giemsa
632 for patient F1-II-1 (at 17 months old). Assessment of bone marrow aspirates revealed
633 dysmegakaryopoiesis with hypolobulated megakaryocytes (left, black arrow, magnification x 200) and
634 dyserythropoiesis with nuclear karyorrhexis (right, black arrow, magnification x500). Smears were
635 examined for each subject by a well-trained hematologist (M Loosveld).
636
637 **Figure 3. Alpha and dense granule defects associated with GATA1 variants.**
638 (A) Representative micrographs of dense granule distribution in a control and patients F1-II-1 (at 8
639 years old), F2-II-1 (at 27 years old) and F2-II-2 (at 23 years old). Scale bar: 1 µm. The graph shows
640 the cumulative distribution of dense granules per whole platelet deposited on formar film for patient
641 vs. control platelets (healthy subjects, n=54); 100 platelets per whole-platelet mount were examined
642 for each subject by a well-trained engineer blinded to the genotype (JC Bordet, Unité d'Hémostase
643 Biologique, Bron, France).
644 (B) Platelet sections from a healthy subject (control panel) and patient F2-II-1 (all other panels).
645 Platelet section from control shows numerous alpha granules, mitochondria recognizable by the
646 crests, and a large glycogen area. Although the numbers are variable, platelets derived from patient
647 F2-II-1 at 22 years old displayed low alpha granule counts for each section, sometimes lower than
648 mitochondria counts. Overall, the sections display either very few organelles with a relatively
649 amorphous background (i) or normal counts for organelles that show variable granulation (iv). Other
650 notable structures observed included very large areas of membrane complexes (ii, iii) and apoptotic
651 granules (left part of panel iii). We also observed membranous formations that resemble a dense
652 tubular system, although the formations were occasionally arranged in a large ring (i), in the vicinity
653 of the tubulin circle (bottom center panel) or in aggregates in the cytoplasm (v). Scale bar: 0.5 µm.
654 The graphs show the cumulative distribution of platelet area, number of alpha granules/µm^2, major
655 platelet diameter, and alpha granule diameter using classical electron microscopy sections in patient
656 F2-II-1 vs. control platelets (healthy subjects n=10); we examined 100 platelets per section.
657
658 **Figure 4. Increased number of CD34$^+$ cells and abnormal megakaryocyte differentiation in**
659 **GATA1 variant carriers.**
660 **(A)** Flow cytometry analysis of circulating CD34$^+$ cells. Left: Representative dot plots of cells from one
661 healthy control and patient F1-II-1. The plots display the side scatter and CD34 expression analysis
662 in CD19low CD45low cells. The ellipse gates show the CD19low CD45low CD34+ cell population.
663 Right: The percentage of CD34$^+$ cells in the total CD19low CD45low cell population. The control data

25

664 represent five different individuals. The data for patient F1-II-1 was obtained from two independent
665 experiments carried out at 15-month intervals. The data are provided as the median ± interquartile
666 range.
667 **(B-D)** In vitro megakaryocyte differentiation. Circulating CD34$^+$ cells from F1-II-1 patient or controls
668 were isolated and cultured in the presence of TPO and SCF to induce megakaryocytic commitment.
669 **(B)** Megakaryocyte differentiation was monitored via flow cytometry. The density plots display the
670 CD41/CD42a expression profiles of two unrelated controls and F1-II-1 affected patient (at days 9 and
671 11 of differentiation**)**. **(C)** Representative images of cultured cells at day 11 (20X magnification). The
672 graph represents the median ± interquartile range of the cell surface area. Fifty cells were quantified.
673 Differences between the groups were assessed using the Kruskal-Wallis test followed by Dunn's
674 multiple comparison test. **(D)** Ploidy levels were monitored via flow cytometry. The histograms display
675 the frequency distribution of Hoechst 33342 levels in CD34$^+$-derived cells from two unrelated controls
676 and F1-II-1 affected patient at day 9 of differentiation.
677 **(E)** Western blot analysis of MYH10 expression in platelets from the patients with GATA1 variants
678 (F1-II2, F2-II-1 and F2-II-2), one heterozygous carrier (F2-I-1), one unaffected family member (F2-I-
679 2) and four unrelated controls. Platelets from a patient carrying the FLI1 p R337Q variant represents
680 the MYH10 positive control. GAPDH was used as a protein loading control. Quantification of band
681 intensity for MYH10 is shown below the western blot.
682

683 **Figure 5. Identification and functional analysis of the MYH10 regulatory regions.**
684 **(A)** Visualization of *MYH10* regulatory regions with GATA1, FLI1 and RUNX1 binding peaks identified
685 via ChIP-sequencing. Part of the *NDEL* gene is shown upstream of the *MYH10* gene using the UCSC
686 genome browser (chr17:8,435,074-8,669,900 using GRCh38/hg38). The bottom layer shows the
687 position of two GATA1 binding region peaks within the 3'UTR (untranslated transcribed region, blue
688 arrows) and intron 8 (red arrows) of *MYH10* for the transcription factors GATA1, FLI1 and RUNX1.
689 **(B)** Luciferase expression in HEL cells transfected with different pGL3 luciferase reporter vectors
690 including the 3'UTR or intron 8 regions of *MYH10* gene. Mutations on GATA1-binding sites in intron
691 8 were introduced as indicated (M1 or M2 or M3 or M(1,2,3,4)). The pGL3 vector without regulatory
692 regions (empty vector) was used as a control. Each plasmid was assayed in three to seven
693 independent transfection experiments. The dual luciferase assay was performed by sequentially
694 measuring the firefly and renilla luciferase activities of the same sample with the results expressed
695 as the ratio of firefly to renilla activity (Fluc/Rluc). **: p<0.001, ***: p<0.0001 vs. pGL3-intron 8 MYH10
696 wild-type (WT) (one-way ANOVA corrected for multiple comparisons). +++: p<0.0001 vs. pGL3-empty
697 vector (one-way ANOVA corrected for multiple comparisons).
698 **(C)** Luciferase expression in GripTite 293 MSR cells transfected with different pGL3 luciferase
699 reporters and pCDNA3-GATA1 vectors as indicated. Left: pGL3-intron 8 MYH10 wild-type (WT),
700 pGL3-intron 8 MYH10 mutants, empty pCDNA3 and pCDNA3-GATA1 WT. Middle: pGL3-3'UTR
701 MYH10, empty pCDNA3 and pCDNA3-GATA1 WT. Right: empty pGL3, empty pCDNA and pCDNA-
702 GATA1 WT. The pCDNA3 empty vector was used as a control. *: p<0.01, ****: p<0.0001 vs. pGL3-
703 intron8-MYH10 WT in the presence of GATA1 (black bar) (one-way ANOVA corrected for multiple
704 comparisons). ++++: p<0.0001 vs. the empty pGL3 vector in the absence of GATA1 (white bar). Each
705 plasmid was assayed in six separate transfection experiments (one-way ANOVA corrected for
706 multiple comparisons).
707 **(D)** Luciferase expression after transfection in GripTite 293 MSR cells with pGL3-intron 8 MYH10 WT
708 and pCDNA3-GATA1 WT or variant vectors as indicated. *: p<0.01 vs. GATA1 WT (one-way ANOVA
709 corrected for multiple comparisons). Each plasmid was assayed in five separate transfection
710 experiments.
711

712 **Table 1. Platelet aggregation assays**

713 Results correspond to assays performed in 1990 and 1993 (A) and 2015 (B).

714

A 715

Individuals	Age (years)	Blood platelet count (x 10⁹/l)	PRP platelet count (x 10⁹/l)	Platelet aggregation maximal intensity (%)					
				ADP		Collagen		Arachidonic acid	Ristocetine
				1 µM	2.5 µM	0.36 µg/ml	0.72 µg/ml	0.5 mg/ml	1.5 mg/ml
Control *	NA	NA	350	45	68	60	63	69	76
F2-II-1	1.7	233	380	9	25	0	0	19	67
Control *	NA	NA	350	45	68	57	63	69	68
F2-II-2	0.8	225	338	8	23	0	0	11	46

716

717

718

B 719

Individuals	Age (years)	Blood platelet count (x 10⁹/l)	PRP platelet count (x 10⁹/l)	Platelet aggregation maximal intensity (%)					
				ADP			Collagen	Arachidonic acid	Ristocetine
				2.5 µM	5 µM	10 µM	3.3 µg/ml	0.5 mg/ml	1.25 mg/ml
Normal range				83-93	75-90	78-92	79-85	76-91	78-948
F2II-1	26.7	106	159	40	62	56	25	14	73
F2II-2	22.7	58	102	40	62	68	24	11	70

720

721 * Controls correspond to one healthy individual tested at the same time.

722

Table 2. Platelet glycoprotein expression

Platelet glycoprotein MFI (a.u.)		F1-II-1	Normal range (10th-90th range n = 14)	F2-II-1	F2-II-2	Normal range (10th-90th range n = 12)
			Individuals			
αIIbβ3 (CD41)	Basal	24	21-44	24	30	20-29
	TRAP14 50µM	25	30-62	30	34	28-46
GPIbα (CD42b)	Basal	1.7	1.2-.3	2.1	2.4	1.2-2.4
	TRAP14 50µM	1.5	0.9-2.3	1.8	1.8	0.6-1.1
CD63	Basal	1.2	0.9-3.4	1.1	1.3	0.5-1
	TRAP14 50µM	1.6	1.9-6.3	1.7	1.3	1.3-3.6
	Ratio TRAP-14 / basal	1.3	1.7-3.7	1.4	1	1.5-4.8
CD62P (P-selectin)	Basal	0.2	0.09-0.35	0.2	0.4	0.15-0.75
	TRAP14 50µM	0.2	1.9-6.3	1.5	1.7	2.3-5
	Ratio TRAP-14 / basal	1.1	12.4-22.8	7	4.2	5.5-25

28

1 **Table 3. Characteristics of GATA1 variant carriers**

	Family 1 (F1-II-1)	Family 2 (F2-II-1 and F2-II-2)
DNA change	c.802C>A	C617A>T
Protein change	p.N206I	p.L268M
Variant localization	N-terminal zinc finger	C-terminal zinc finger
Clinical manifestations		
Thrombocytopenia	Early in life (18 months) with a severely reduced platelet count (15 G/L).	Initially absent (at 18 and 9 months, 261 and 233 G/L for each brother). Thrombocytopenia gradually developed from 2-5 years of age and then steadily worsened.
Bleeding	Frequent bruising, post vaccinal hematoma.	Pronounced mucocutaneous bleedings.
	Bleeding during dental avulsion prevented by platelet concentrates.	Bleeding during dental avulsion and appendicectomy was prevented by platelet concentrates.
Platelet	Not done	Bleeding time >15min
	No aggregation test performed (because platelet number is < 20G/L).	Aggregation defect (mainly collagen and arachidonic acid) in absence of thrombocytopenia.
	> 90% reduction in dense granule number.	≥ 90% reduction in dense granule number for both children.
	Surface glycoproteins: No basal defect in CD41, CD42b, CD62P, and CD63; reduced expression levels of CD63 and CD62P after TRAP-14 stimulation.	Surface glycoproteins: No basal defect in CD41, CD42b, CD62P, and CD63; reduced expression levels of CD63 and CD62P after TRAP-14 stimulation.
	Serotonin: NA (too low platelet count).	Serotonin low values in both children: 0.12-0.16 µg/10^9 platelets (Normal values: 0.3-1.20).
	PAI-1 antigen: normal value, 0.35 ng/10^6 platelets (Normal values: 0.33-1.07).	PAI-1 antigen: normal values in both child 0.69-0.76 ng/10^6 platelets (Normal values: 0.33-1.07).
Coagulation and fibrinolysis	No defect.	No defect.
Erythrocyte	Absence of anemia.	Absence of anemia.
	MCV : 77-81 fL (Normal values : 80-95).	Progressive increase in MCV over time (from 85-90 to 99 and 103 fL at 26 and 23 years of age) (Normal values: 80-95).
	Discrete anisopoïkilocytosis, few dacryocytes and schizocytes.	Anisopoikilocytosis.
	Fetal hemoglobin not evaluated.	Elevated level of fetal hemoglobin (25%).
Bone marrow	Dysmegakaryopoiesis with hypolobulated megakaryocytes and mild dyserythropoiesis with nuclear karyorrhexis.	Not done.

29

1
2
3
4
5   1
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

30

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
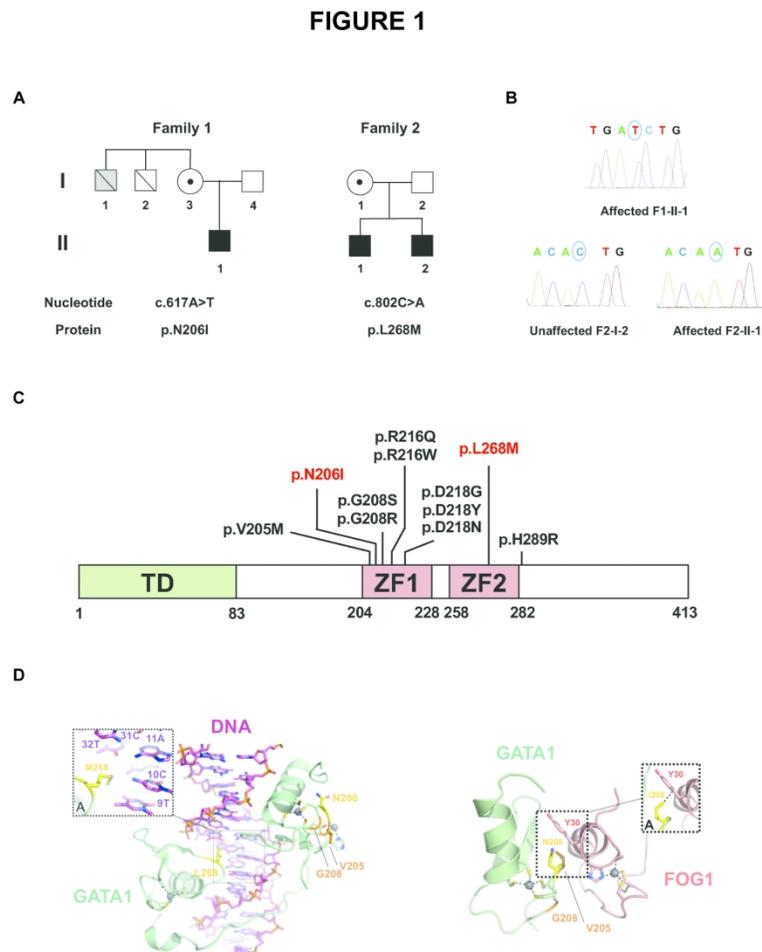43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

FIGURE 1

Figure 1. Identification of two novel GATA1 variants.
A) Pedigrees of the affected families. Squares denote males, and circles represent females. The slashes denote deceased family members. Solid black symbols represent affected family members carrying a hemizygous GATA1 variant. Circles with a dot represent heterozygous female carriers of the variants. Solid gray symbols represent non-tested family members with a medical history of bleeding events.
(B) Sanger sequencing electropherograms of the indicated family members.
(C) Schematic diagram of the GATA1 protein domain structure. The functional N-terminal transactivation domain (TD) and both the N- and C-terminal zinc-finger domains (ZF1 and ZF2, respectively) are depicted. The positions of the variations within GATA1 are indicated in red (variations reported in this study) and black (variations reported in previous studies).

209x296mm (300 x 300 DPI)

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
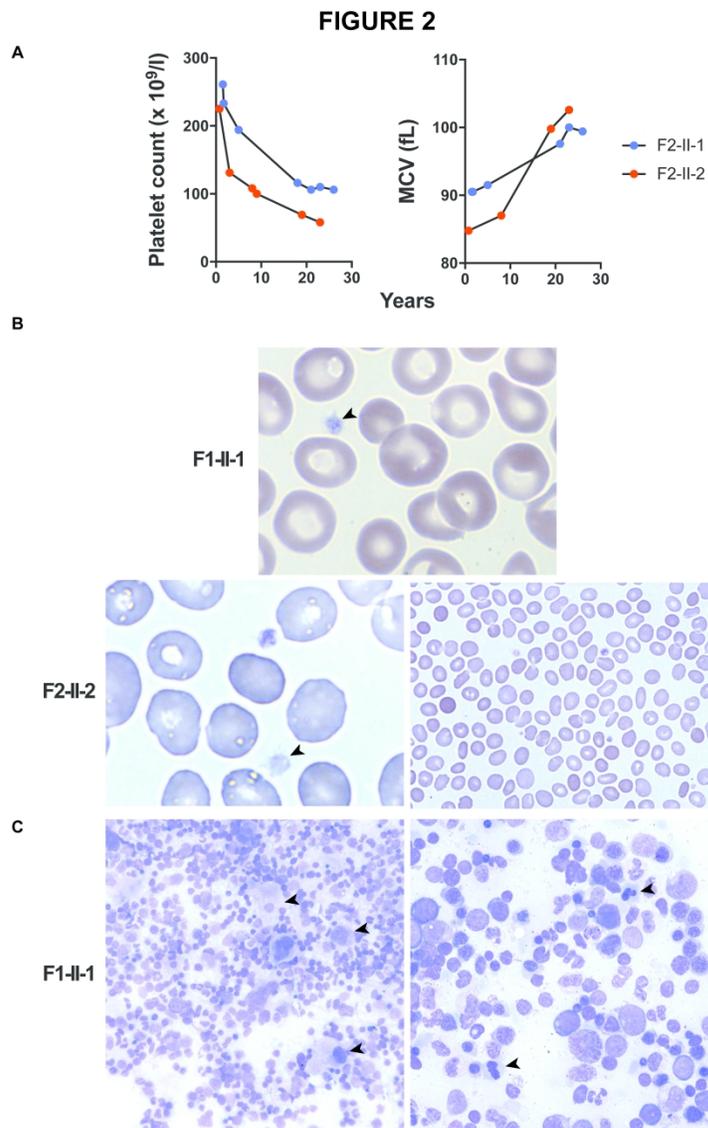52
53
54
55
56
57
58
59
60

FIGURE 2

Figure 2. Quantitative and morphological modifications of platelets and red blood cells in GATA1 variant carriers

(A) Evolution of platelet count and mean corpuscular volume (MCV) of erythrocytes with age in the two F2 probands.(B) Representative images of peripheral blood smears stained with May-Grünwald-Giemsa for patients F1-II-1 and F2-II-2 (at 6 and 23 years old, respectively). The peripheral blood smear images revealed hypogranular platelets (left, black arrow, magnification x1000) mild anisocytosis and poikylocytosis (right, magnification x 500).(C) Representative images of bone marrow aspirate smears stained with May-Grünwald-Giemsa for patient F1-II-1 (at 17 months old). Assessment of bone marrow aspirates revealed dysmegakaryopoiesis with hypolobulated megakaryocytes (left, black arrow, magnification x 200) and dyserythropoiesis with nuclear karyorrhexis (right, black arrow, magnification x500). Smears were examined for each subject by a well-trained hematologist (M Loosveld).

FIGURE 3
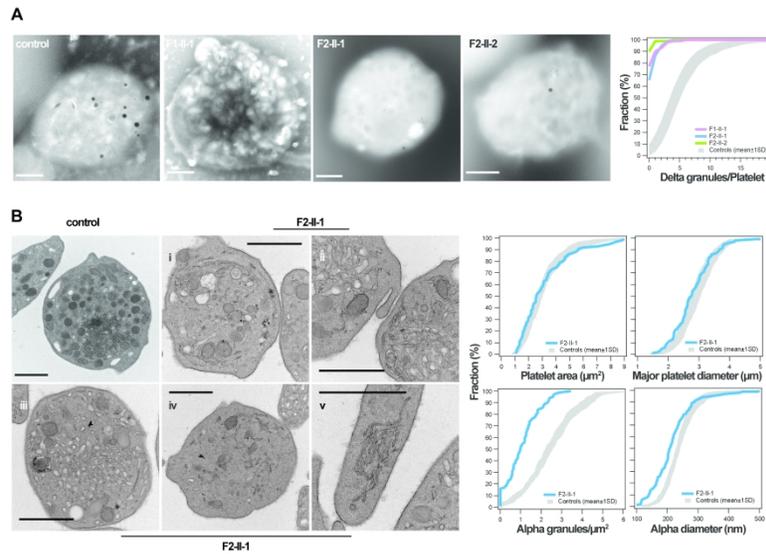
Figure 3. Alpha and dense granule defects associated with GATA1 variants.
(A) Representative micrographs of dense granule distribution in a control and patients F1-II-1 (at 8 years old), F2-II-1 (at 27 years old) and F2-II-2 (at 23 years old). Scale bar: 1 μm. The graph shows the cumulative distribution of dense granules per whole platelet deposited on formar film for patient vs. control platelets (healthy subjects, n=54); 100 platelets per whole-platelet mount were examined for each subject by a well-trained engineer blinded to the genotype (JC Bordet, Unité d'Hémostase Biologique, Bron, France).
(B) Platelet sections from a healthy subject (control panel) and patient F2-II-1 (all other panels). Platelet section from control shows numerous alpha granules, mitochondria recognizable by the crests, and a large glycogen area. Although the numbers are variable, platelets derived from patient F2-II-1 at 22 years old displayed low alpha granule counts for each section, sometimes lower than mitochondria counts. Overall, the sections display either very few organelles with a relatively amorphous background (i) or normal counts for organelles that show variable granulation (iv). Other notable structures observed included very large areas of membrane complexes (ii, iii) and apoptotic granules (left part of panel iii). We also observed membranous formations that resemble a dense tubular system, although the formations were occasionally arranged in a

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

large ring (i), in the vicinity of the tubulin circle (bottom center panel) or in aggregates in the cytoplasm (v). Scale bar: 0.5 µm. The graphs show the cumulative distribution of platelet area, number of alpha granules/µm2, major platelet diameter, and alpha granule diameter using classical electron microscopy sections in patient F2-II-1 vs. control platelets (healthy subjects n=10); we examined 100 platelets per section.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
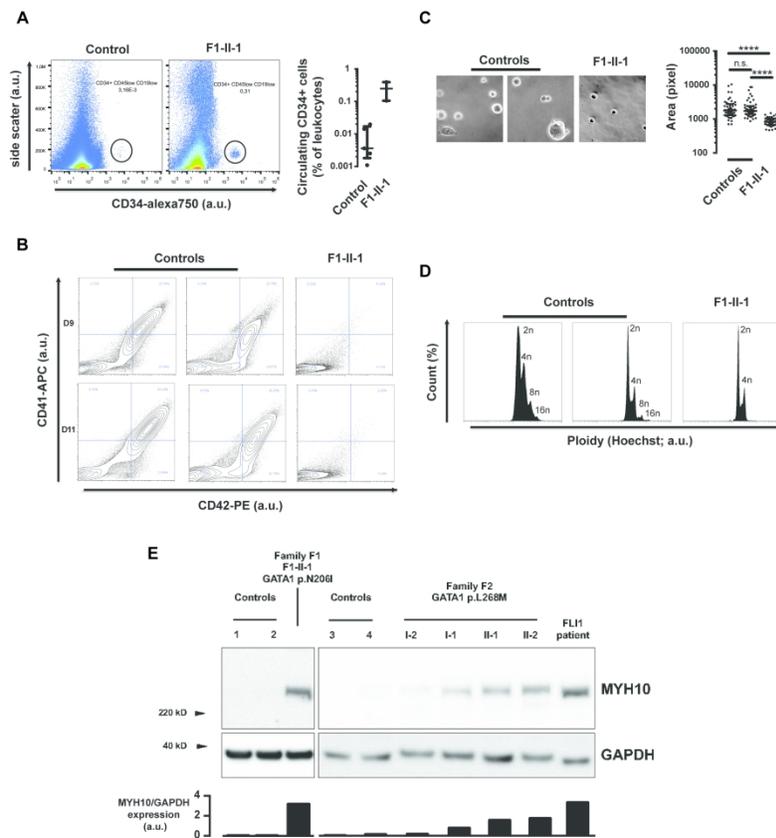51
52
53
54
55
56
57
58
59
60

FIGURE 4

Figure 4. Increased number of CD34+ cells and abnormal megakaryocyte differentiation in GATA1 variant carriers.
(A) Flow cytometry analysis of circulating CD34+ cells. Left: Representative dot plots of cells from one healthy control and patient F1-II-1. The plots display the side scatter and CD34 expression analysis in CD19low CD45low cells. The ellipse gates show the CD19low CD45low CD34+ cell population. Right: The percentage of CD34+ cells in the total CD19low CD45low cell population. The control data represent five different individuals. The data for patient F1-II-1 was obtained from two independent experiments carried out at 15-month intervals. The data are provided as the median ± interquartile range.
(B-D) In vitro megakaryocyte differentiation. Circulating CD34+ cells from F1-II-1 patient or controls were isolated and cultured in the presence of TPO and SCF to induce megakaryocytic commitment. (B) Megakaryocyte differentiation was monitored via flow cytometry. The density plots display the CD41/CD42a expression profiles of two unrelated controls and F1-II-1 affected patient (at days 9 and 11 of differentiation). (C) Representative images of cultured cells at day 11 (20X magnification). The graph represents the median ± interquartile range of the cell surface area. Fifty cells were quantified. Differences

between the groups were assessed using the Kruskal-Wallis test followed by Dunn's multiple comparison test. (D) Ploidy levels were monitored via flow cytometry. The histograms display the frequency distribution of Hoechst 33342 levels in CD34+-derived cells from two unrelated controls and F1-II-1 affected patient at day 9 of differentiation.

(E) Western blot analysis of MYH10 expression in platelets from the patients with GATA1 variants (F1-II2, F2-II-1 and F2-II-2), one heterozygous carrier (F2-I-1), one unaffected family member (F2-I-2) and four unrelated controls. Platelets from a patient carrying the FLI1 p R337Q variant represents the MYH10 positive control. GAPDH was used as a protein loading control. Quantification of band intensity for MYH10 is shown below the western blot.
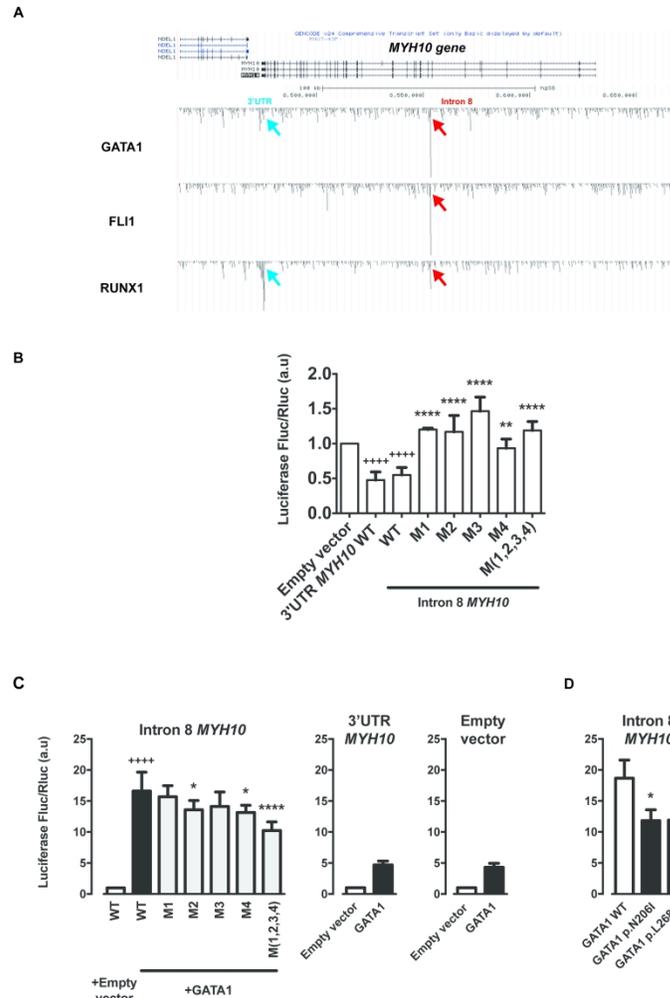
FIGURE 5

Figure 5. Identification and functional analysis of the MYH10 regulatory regions.
(A) Visualization of MYH10 regulatory regions with GATA1, FLI1 and RUNX1 binding peaks identified via ChIP-sequencing. Part of the NDEL gene is shown upstream of the MYH10 gene using the UCSC genome browser (chr17:8,435,074-8,669,900 using GRCh38/hg38). The bottom layer shows the position of two GATA1 binding region peaks within the 3'UTR (untranslated transcribed region, blue arrows) and intron 8 (red arrows) of MYH10 for the transcription factors GATA1, FLI1 and RUNX1.
(B) Luciferase expression in HEL cells transfected with different pGL3 luciferase reporter vectors including the 3'UTR or intron 8 regions of MYH10 gene. Mutations on GATA1-binding sites in intron 8 were introduced as indicated (M1 or M2 or M3 or M(1,2,3,4)). The pGL3 vector without regulatory regions (empty vector) was used as a control. Each plasmid was assayed in three to seven independent transfection experiments. The dual luciferase assay was performed by sequentially measuring the firefly and renilla luciferase activities of the same sample with the results expressed as the ratio of firefly to renilla activity (Fluc/Rluc). **: $p<0.001$, ***: $p<0.0001$ vs. pGL3-intron 8 MYH10 wild-type (WT) (one-way ANOVA corrected for multiple comparisons). +++: $p<0.0001$ vs. pGL3-empty vector (one-way ANOVA corrected for multiple

comparisons).
(C) Luciferase expression in GripTite 293 MSR cells transfected with different pGL3 luciferase reporters and pCDNA3-GATA1 vectors as indicated. Left: pGL3-intron 8 MYH10 wild-type (WT), pGL3-intron 8 MYH10 mutants, empty pCDNA3 and pCDNA3-GATA1 WT. Middle: pGL3-3'UTR MYH10, empty pCDNA3 and pCDNA3-GATA1 WT. Right: empty pGL3, empty pCDNA and pCDNA-GATA1 WT. The pCDNA3 empty vector was used as a control. *: $p < 0.01$, ****: $p < 0.0001$ vs. pGL3-intron8-MYH10 WT in the presence of GATA1 (black bar) (one-way ANOVA corrected for multiple comparisons). ++++: $p < 0.0001$ vs. the empty pGL3 vector in the absence of GATA1 (white bar). Each plasmid was assayed in six separate transfection experiments (one-way ANOVA corrected for multiple comparisons).
(D) Luciferase expression after transfection in GripTite 293 MSR cells with pGL3-intron 8 MYH10 WT and pCDNA3-GATA1 WT or variant vectors as indicated. *: $p < 0.01$ vs. GATA1 WT (one-way ANOVA corrected for multiple comparisons). Each plasmid was assayed in five separate transfection experiments.

1 # Supporting information.

2 **Methods**

3 **DNA sequencing, variant calling, variant confirmation and segregation studies**

4 DNA extraction, next-generation sequencing library preparation, variant calling and annotation were

5 performed as previously described [1]. Briefly, the DNA samples were enriched for the exons of 308

6 genes involved in platelet function and disorders (list available upon request) using HaloPlex

7 technology (Agilent technologies) or SeqCap EZ Choice technology (Roche sequencing). DNA

8 sequencing was then performed on a NextSeq 500 sequencer (Illumina) using 2-channel SBS

9 technology and a Mid-Output kit. Sequence quality was verified using Sequencing Analysis Viewer

10 v1.8.37 software (Illumina). The sequences were then aligned using the Burrows-Wheeler Alignment

11 tool with maximal exact matches (hg19 version of the genome). Variant calling was performed using

12 GATK v3.3.0 and annotated with ANNOVAR [2] and VarAFT software [3]. We applied the following

13 filters for variant analysis: SNPs with a frequency <1% (based on ExAC, 1000G and ESP6500 data),

14 located in coding regions or splice junctions, and designated deleterious by several prediction

15 algorithms (UMD Predictor [4], PROVEAN [5] and Polyphen [6]). We confirmed the variations in

16 probands and pedigree members via Sanger sequencing using a 3500xL Genetic Analyzer.

17 **Structural modeling**

18 Structural modeling of the GATA1 (NP_002040.1) variants was performed using molecular

19 visualization and modeling with Coot and PyMOL software (The PyMOL Molecular Graphics System;

20 Version 1.5.0.4; Schrödinger LLC).

21 **Flow cytometry analysis**

22 Major platelet receptors and granule markers were assessed as previously described [18]. Briefly,

23 PRP samples were incubated either with or without TRAP-14 (Polypeptide group) and antibodies

24 against αIIbβ3 integrin (clone P2, Beckman Coulter), active form of αIIbβ3 (clone PAC-1; BD), GPIbα

25 (clone HIP1, BD Biosciences), CD63 (clone CLB-gran 12; Beckman Coulter) and CD62P (clone CLB-

1

1 Thromb/6, BD Biosciences) for 30 min at room temperature. Scatter signals and fluorescence

2 intensity were analyzed using an FC500 Flow Cytometer (Beckman Coulter). Circulating CD34+ cells

3 were quantified as previously reported [27].

Cell expression and luciferase reporter assays

5 Wild-type *GATA1* cDNA (amplified from pPacGATA1, a gift from Jeffrey W. Taub) [7] was inserted

6 into a pCDNA3.1-cMyc expression vector. The *GATA1* variants were generated using a GENEART

7 Site-directed Mutagenesis System (Life technologies) according to the manufacturer's protocol. The

8 primer sequences used are available upon request. The *MYH10* regulatory regions (intron 8 and the

9 3' untranslated region [UTR]) were cloned into a pGL3 vector (Proteogenix). Polymerase chain

10 reaction mutagenesis was used to substitute the GATA1 binding sites with an AAAA sequence. The

11 plasmids were denoted M1/M2/M3/M4 when the four binding sites were substituted and M1, M2, M3

12 or M4 when a single binding site was substituted. After cloning and transformation, each genotype

13 was confirmed using Sanger DNA sequencing (Beckman Coulter Genomics). To examine cell-type

14 dependence, we used the HEL hematopoietic and GripTite 293 MSR (HEK 293 cell line expressing

15 the human macrophage scavenger captor) cell lines. The cells were transfected with pCDNA3.1-

16 cMyc-GATA1 along with the SV40 drive luciferase plasmid (pGL473-hRLuc) to normalize for

17 transfection efficiency. Polyjet (Tebu-bio) was used for transfection according to the manufacturer's

18 protocol. The cells were harvested and lysed 48 hours after transfection. Firefly and Renilla luciferase

19 activity were then measured consecutively using a Dual-Luciferase Reporter 1000 Assay system

20 (Promega) with an EnSight™ Multimode Microplate Reader (PerkinElmer) according to the

21 manufacturer instructions.

PRP serotonin level

23 PRP was prepared according to standard procedures and then frozen at -70 °C. After thawing and

24 vortexing the sample, the serotonin levels were assessed using high performance liquid

25 chromatography with the following parameters: isocratic flow 0.95 ml/minute, sensor potential 0.5V,

26 sensitivity 10 nA, column temperature 30 ± 1 °C, analysis time 10 minutes, injection volume 20 µl,

27 serotonin retention time 5.2 minutes, eluent retention time 6.8 minutes. The serotonin levels were

2

then expressed according to the PRP platelet count. The plasma serotonin levels were considered negligible as compared with the intraplatelet serotonin content.

Platelet PA1-1 antigen level

Serum was prepared from blood collected in glass tubes and incubated at 37°C for 120 minutes before centrifugation at 2000 g for 10 minutes at 4°C. The supernatant was stored at -70°C until downstream assays were performed. PAI-1 antigen concentrations in plasma and serum samples were assessed using an enzyme-linked immunosorbent assay (Asserachrom, Diagnostica Stago). The serum results are expressed according to the PRP platelet count after subtracting the plasma values.

Mepacrine uptake and release assay

PRP samples were prepared according to standard procedures and diluted to obtain a platelet concentration of 10^7 platelets/mL. The diluted PRP samples were incubated with 1 µM or 2 µM mepacrine (Q3251, Sigma-Aldrich) for 30 min at 37 °C; diluted control PRP samples were incubated without mepacrine. Next, the platelet samples were incubated either with or without 40 µM TRAP-14. Platelet-mepacrine fluorescence was analyzed using a Navios Cytometer (Beckman Coulter). Mepacrine-labeled dense granules were also quantified via fluorescence microscopy. In single-plane views, mepacrine displayed characteristic punctate staining. A total of 100 platelets were examined, and the number of dense granules per platelet was determined.

**Epifluorescence microscopy**

Forty-eight hours after transfection, H9C2 cells (ATCC CRL-1446) were fixed in 4% paraformaldehyde for 20 min at room temperature. The cells were then washed, permeabilized with 0.3% Triton X100 in PBS for 5 min, blocked with 3% BSA PBS, and incubated overnight with rabbit anti-cMyc antibody (Santa Cruz Biotechnology; sc-789). Next, the cells were incubated with anti-rabbit Alexa-546-labeled secondary antibody (Life technologies A11010), and DAPI staining was performed. Finally, after washing the cells, the slides were mounted with Fluoromount and examined using an AXIO Imager M1 microscope (Carl Zeiss, Germany).

Platelet lysates, FOG1/GATA1 immunoprecipitation and western blot analysis

The platelet lysates were prepared as previously described.[8] GripTite 293 MSR cells were co-transfected with the vectors pFN21A-halotag-FOG1 (a gift from Dr. Fujiwara)[9] and pCDNA3-cMyc GATA1. Transfected MSR cells were treated with DTBP reagent (Interchim) for crosslinking and lysed with radio-immunoprecipitation assay (RIPA) buffer (Sigma-Aldrich) supplemented with EDTA-free protease cocktail inhibitor (Roche). The immunoprecipitation was carried out overnight at 4°C by incubating the cell extracts with rabbit anti-cMyc coated with magnetic beads (PureProteome, Millipore). After washing the beads with lysis buffer, the immunoprecipitates were suspended in sample loading buffer (NuPAGE LDS sample buffer, Thermo Fisher Scientific) and boiled for 5 min at 95 ºC. The immunoprecipitated proteins were assessed for halotag-FOG1 via western blot.

The proteins were separated on NuPAGE gels with MES SDS running buffer (Thermo Fisher Scientific) and then transferred onto a polyvinylidene fluoride membrane. The membranes were blocked (5% Blotto buffer) and labeled overnight with the following primary antibodies: rabbit anti-MYH10 (Cell Signaling Technology), anti-GATA1 (Thermo Fisher Scientific), mouse anti-halotag (Promega), anti-GPVI, anti-SYK, anti-PLC2, anti-COX, anti-TBXAS, anti-cMyc (Santa-Cruz Biotechnology) and anti-GAPDH (Millipore). The membranes were then incubated with the appropriate horseradish peroxidase-conjugated secondary antibody (Bio-Rad Laboratories). Staining was performed using ECL detection reagent (Life Technologies), and the chemiluminescent blots were imaged using the charge-coupled device camera system ImageQuant LAS 4000 (GE Healthcare) and quantified using ImageJ software (National Institutes of Health, Bethesda, MD, USA).

4

1
2
3
4 **Supplemental figures**
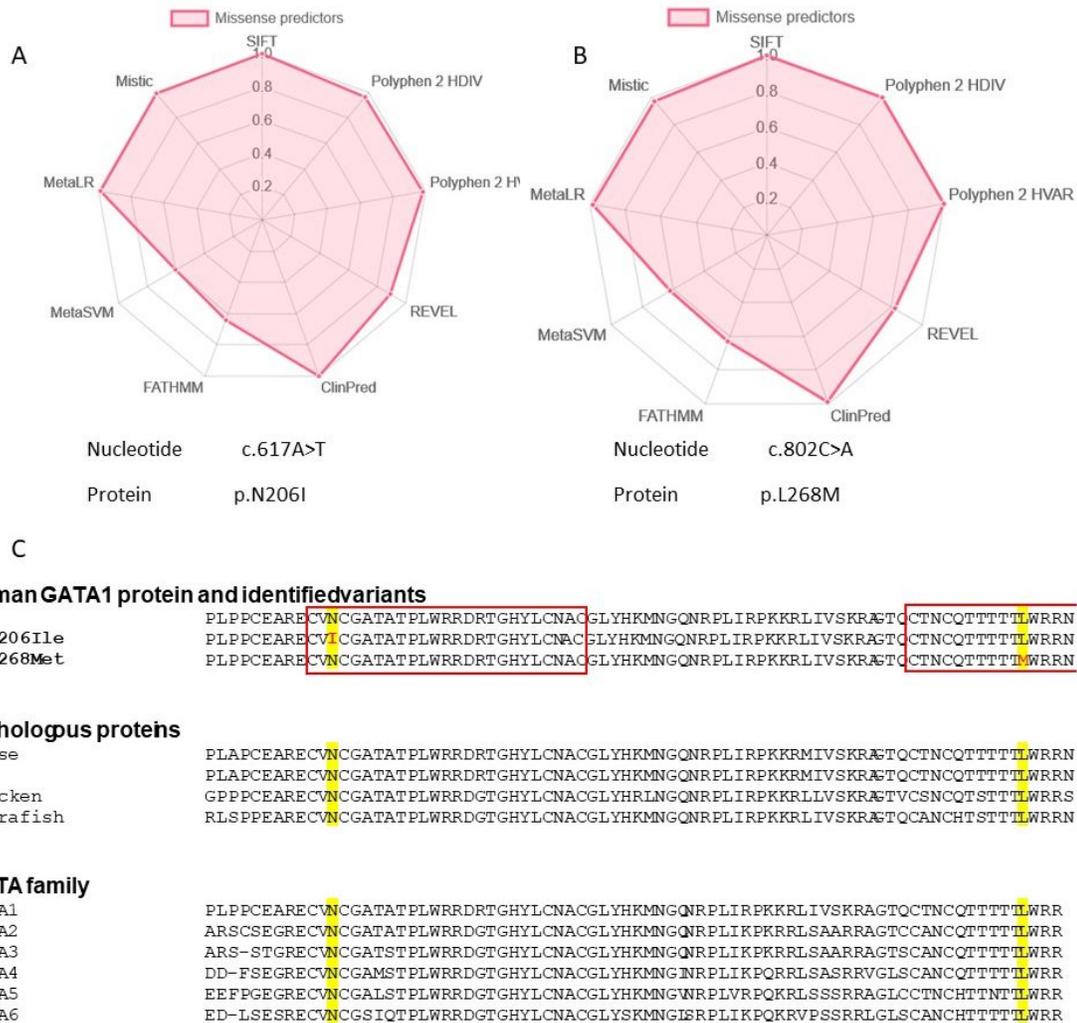
Supplemental Figure 1: Radar view of missense predictors and GATA alignments.
(A-B) Radar view of missense predictors (issued from https://mobidetails.iurc.montp.inserm.fr/MD/).
Values are normalized (0-1) 0 being the less damaging and 1 the most for each predictor algorithms.
(C) The variants reported in this study are indicated in red (top alignment). Alignments of GATA1
orthologous and different members of the GATA family are provided (middle and bottom alignments,
respectively). The boxes delineate both zinc finger domains, ZF1 (position 204-228) and ZF2 (position
258-282). The GATA1 reference variant is NM_002049. WT: wild-type.
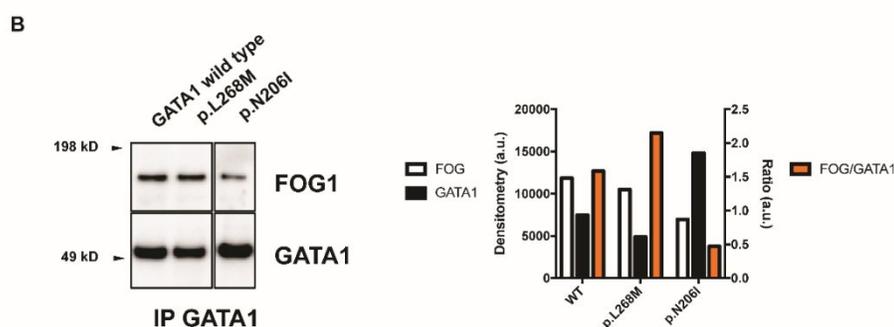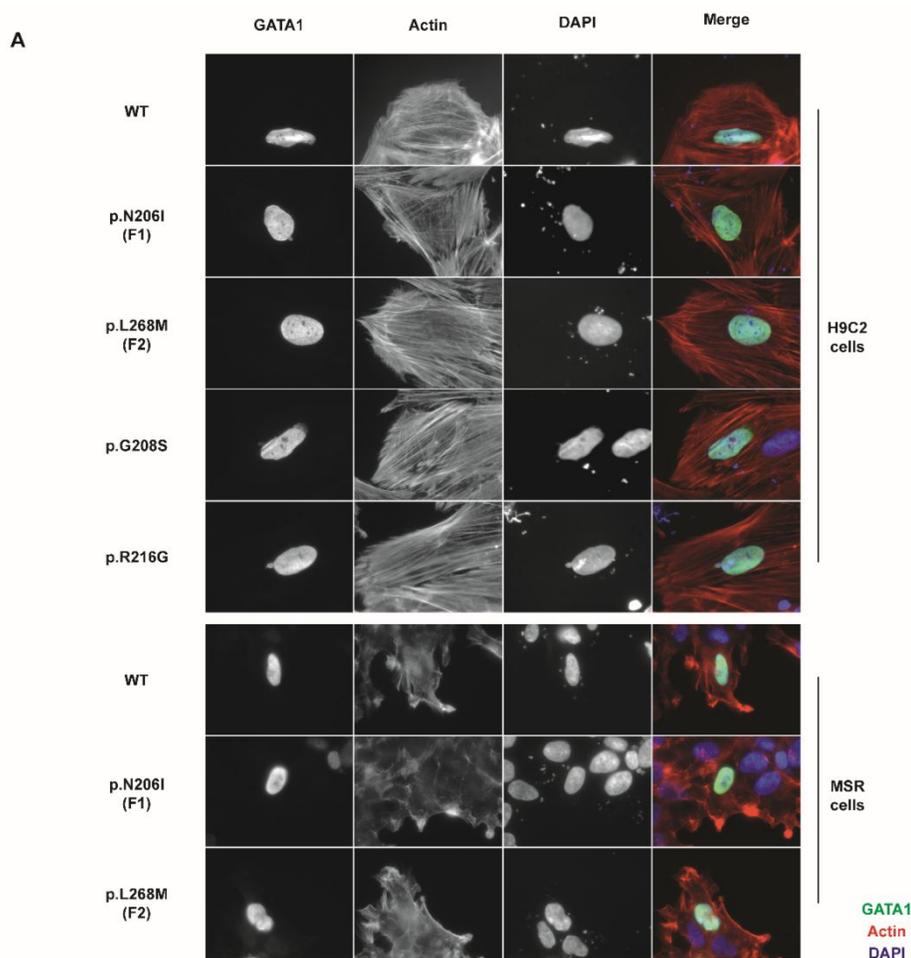
**Supplemental Figure 2: Functional characterization of the GATA1 variants.**
(A) Subcellular localization of the GATA1 variants. Representative immunofluorescence microscopy images of H9C2 (upper panel) and MSR (lower panel) cells transfected with wild-type (WT) GATA1 or variants (pCDNA3-cMyc-GATA1 expression vectors) after cMyc, actin and DAPI staining (100X magnification). The p.N206I and p.L268M variants were compared to the previously reported variants p.G208S [10] and p.R216G [11,12] as well as WT GATA1.
(B) Immunoprecipitation of GATA1-cMyc in GripTite 293 MSR cells co-transfected with pFN21A-HaloTag-FOG1 and pCDNA3-cMyc GATA1 (wild-type or variants) vectors. Immunoprecipitation was performed on cell lysate with rabbit anti-cMyc antibody (Santa-Cruz, sc-789). The immunoprecipitates
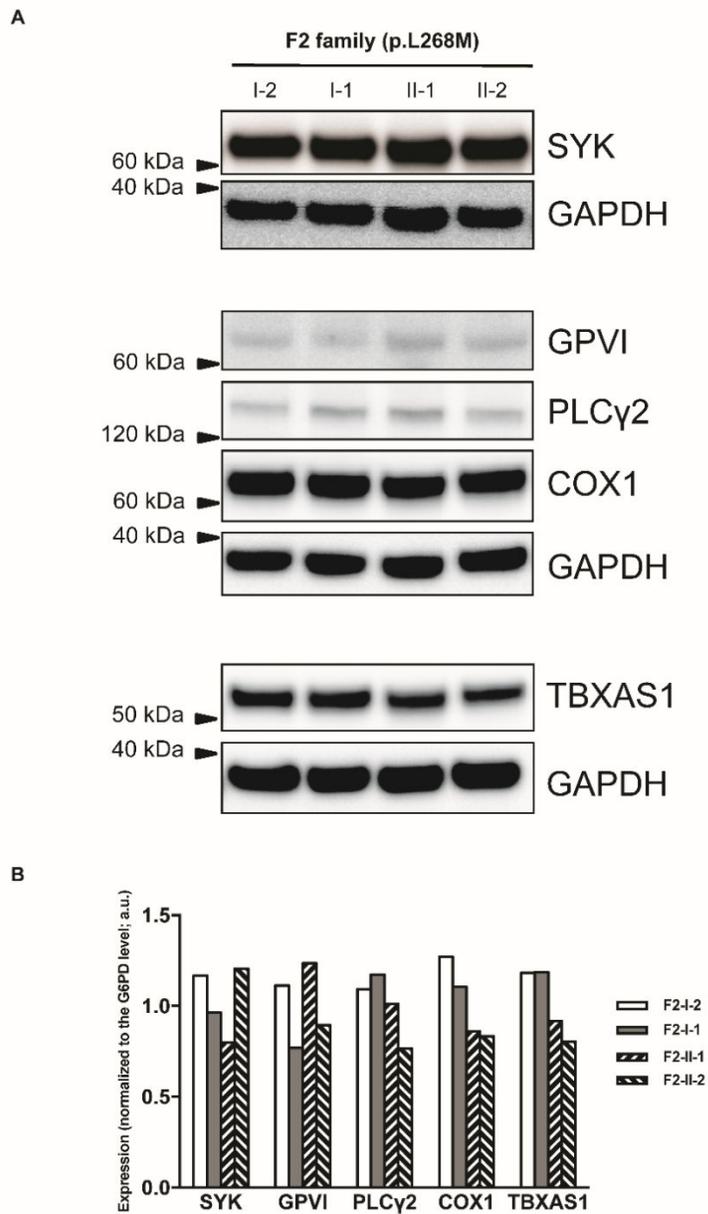
6

1   were analyzed via western blot using murine anti-HaloTag (Promega, G921A) and anti-cMyc (Santa-
2   Cruz, sc-40) antibodies. Input correspond to the GATA1 and FOG1 levels before IP. All samples were
3   run together and the bands were detected simultaneously (the original photograph can be provided
4   on request). Quantification of band intensity for FOG1, GATA1 and ratio GOG1/GATA1 is shown to
5   the right of the western blot.
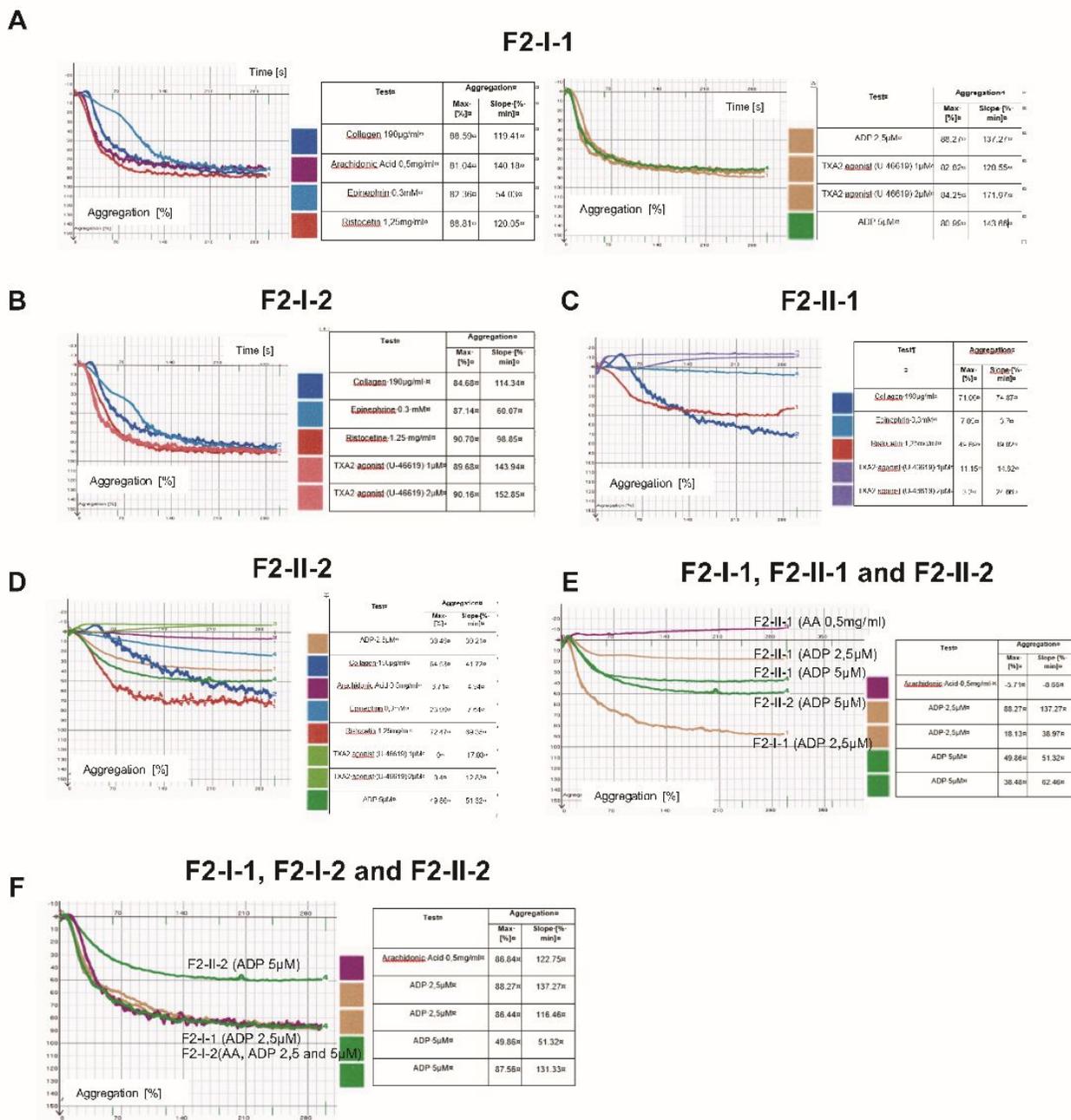
**Supplemental Figure 3: Platelet aggregation tracings for F2 family (pL268M).**
A-D Aggregation tracings for F2-I-1 (mother), F2-I-2 (father), F2-II-1 (firstborn son) and F2-II-2 (second son) family members, on separated graphs. **E**- Aggregation tracings for F2-I-1, F2-II-1 and F2-II-2 on the same graphs. **F**- Aggregation tracings for F2-I-1, F2-I-2 and F2-II-2 on the same graphs. Platelet aggregation was induced by different agonist and different agonist concentrations indicated in the table, on the right to the curves. Percentage max of aggregation and slope were indicated in the table.

8

1
2
3
4   1
5
6
7   2
8
9   3
10
11   4
12
13   5
14
15   6
16
17   7
18   8
19
20   9
21
22   10
23
24   11
25
26   12
27   13
28
29
30   14
31   15
32
33   16
34
35   17
36
37   18
38
39   19
40
41   20
42   21
43
44   22
45   23
46
47   24
48   25
49
50   26
51   27
52
53   28
54
55   29
56
57
58
59
60

A

F2 family (p.L268M)

I-2   I-1   II-1   II-2

60 kDa ▶   SYK
40 kDa ▶

GAPDH

60 kDa ▶   GPVI

PLCγ2
120 kDa ▶

COX1
60 kDa ▶
40 kDa ▶

GAPDH

TBXAS1
50 kDa ▶

40 kDa ▶
GAPDH

B

Expression (normalized to the G6PD level; a.u.)

1.5

1.0

0.5

0.0

SYK   GPVI PLCγ2 COX1 TBXAS1

F2-I-2
F2-I-1
F2-II-1
F2-II-2

Supplemental

Figure 4: Evaluation of proteins involved in the collagen and arachidonic acid signaling pathways

Expression of the proteins involved in the collagen (GPVI, SYK and PLCγ2) and arachidonic acid (COX, TBXAS1) signaling pathways were analyzed via western blot of platelet lysates (30 µg) from the F2 family members: the GATA1 variant carriers (F2-II-1 and F2-II-2), maternal heterozygous carrier (F2-1-1) and unaffected father (F2-I-2). Quantification of band intensity is shown below the western blots.

9

A

B

1
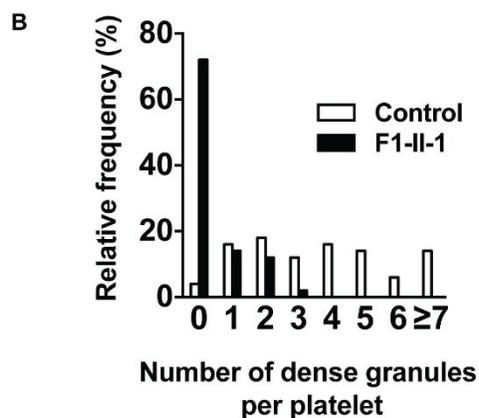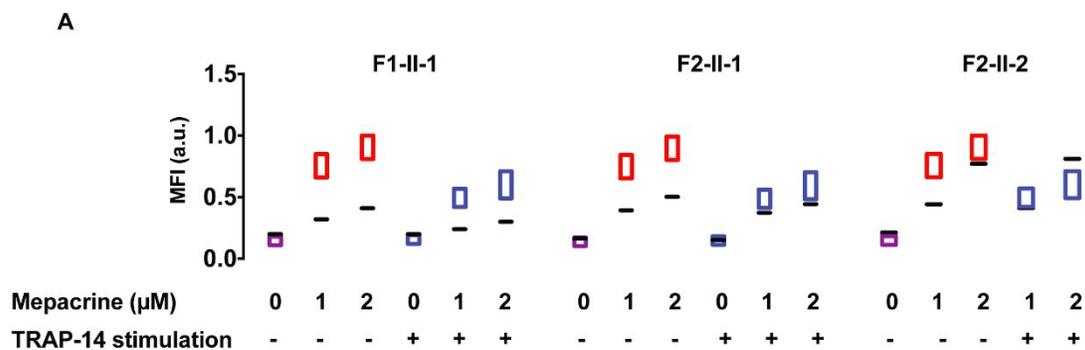
Supplemental Figure 5: Mepacrine test

(A) Flow cytometric mepacrine uptake and release assays of the three variant carriers and unrelated controls. The platelets were incubated with 1 µM or 2 µM mepacrine and stimulated with TRAP-14 (50 µM) to evaluate mepacrine release. Rectangles indicate the reference values, while black horizontal bars indicate the patient values. MFI: mean fluorescence intensity.

(B) Quantification of dense granules in 100 platelets using immunofluorescence microscopy after mepacrine uptake in patient F1-II-1 and a control.
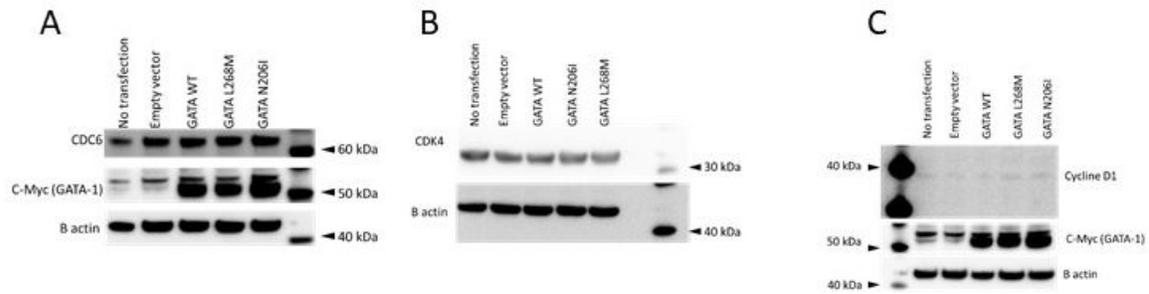
Supplemental Figure 6: Evaluation of proteins involved in GATA-1 mediated polyploidization defect.

Western blot analysis of CDC6 (A), CDK4 (B) and Cyclin-D1 (C) expression in GripTite 293 MSR cell lines transfected with GATA-1 wild-type or variants. GATA-1-cMyc was assessed as a transfection control. Actin expression was assessed as a protein loading control. A and B panels shows the same samples.

11

3'UTR sequence
hg38_dna range=chr17:8472947-8473157

```
CAAAGCCCAGCTGCCATAAACTGTCTTCCTTGTGCTCTGAATGCCCAACGAAGGTTGTATAAACTACC
TTTCCTTAGGTCTGGAGTGTCGGGTCAGTCAAGAAGGGACCCTTTTAGTTATTAAATAGTATCTTTGC
AGAATCAGGGCCTGGAATTGGTTCACACAGATTGTTTCTCCTGGCACGCTGAGTAGTTGGGAAACTGT
GACGTGA
```
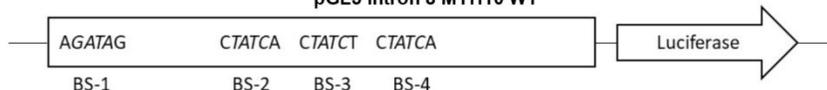
Intronic sequence
hg38_dna range=chr17:8553047-8553739

Wild-type

```
TGGGGCCACCAAAGTAAATTAATTCTGTATGTGTCAGATAGCCTGAATTTTAAGTCAGACTA
TCCCACACATCCTCCACGTCTGGTGAAAGACACACTCATATGAATCATGCTCTGACACACCT
ATATGAAGAAAGCTGAACTGCAGATTCATTCTCCTGTCATTTTCAAGTATTTCCAGATTTTG
CTGGCTGTGAGTCCATTTGGAGGCATAAAATACTGTTCTTAACCAGAAAATCATGGTCTGTC
CCTTGAAGGATAGAAATAATGTATTTCTATCACTGGAGATTATTCTATTAAGTACCCGAAAA
CTGTCATCTATCTTTCTAATTGGCACACTATCATTCAGAATCATTTCTATTTTATTACATCT
ATACTCAAATGTGTTTTACCTGGGTATATATGTGCATACACATATATAAACCTTCTTATGGC
CATTCTATGCAGGTGACATTTCTTAGGACTATTCTATGTACTTGACATTTTATTTTTTTCTC
CAGCTTTATTCTACTTAACACTAATTGAGATTACTTTAACTCTGTAATCTTATAAGGATGTT
TTCAATGATTACATGCTCATCCTTCTAATATCAGAAGTAGATGGTAAAAAGTGGGCTAACAG
CAATAGGGGGAGGGCAACTGCATTCATCTGAATGTCCTATATCACACAAACACTGTGGCACT
CTCCACTTCCC
```

pGL3 intron 8 MYH10 WT

| AGATAG | CTATCA | CTATCT | CTATCA | Luciferase |
| BS-1 | BS-2 | BS-3 | BS-4 | |

Mutated binding-site (BS-M1,2,3,4)

```
TGGGGCCACCAAAGTAAATTAATTCTGTATGTGTCAAAAAGCCTGAATTTTAAGTCAGACTA
TCCCACACATCCTCCACGTCTGGTGAAAGACACACTCATATGAATCATGCTCTGACACACCT
ATATGAAGAAAGCTGAACTGCAGATTCATTCTCCTGTCATTTTCAAGTATTTCCAGATTTTG
CTGGCTGTGAGTCCATTTGGAGGCATAAAATACTGTTCTTAACCAGAAAATCATGGTCTGTC
CCTTGAAGGATAGAAATAATGTATTTCAAAAACTGGAGATTATTCTATTAAGTACCCGAAAA
CTGTCATCAAAATTTCTAATTGGCACACAAAAATTCAGAATCATTTCTATTTTATTACATCT
ATACTCAAATGTGTTTTACCTGGGTATATATGTGCATACACATATATAAACCTTCTTATGGC
CATTCTATGCAGGTGACATTTCTTAGGACTATTCTATGTACTTGACATTTTATTTTTTTCTC
CAGCTTTATTCTACTTAACACTAATTGAGATTACTTTAACTCTGTAATCTTATAAGGATGTT
TTCAATGATTACATGCTCATCCTTCTAATATCAGAAGTAGATGGTAAAAAGTGGGCTAACAG
CAATAGGGGGAGGGCAACTGCATTCATCTGAATGTCCTATATCACACAAACACTGTGGCACT
CTCCACTTCCC
```

pGL3 intron 8 MYH10 with mutated binding-site

| AAAAAG | CAAAAA | CAAAAT | CAAAAA | Luciferase |
| BS-M1 | BS-M2 | BS-M3 | BS-M4 | |
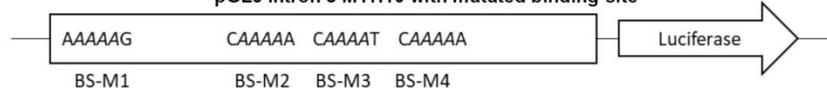
1 **Supplemental Figure 7: DNA sequence of regulatory GATA1 binding regions of the**
2 **MYH10 gene**.
3 The sequences of the 3'UTR (Chr. 17: 8472947-8473157) and intronic region (Chr. 17: 8553047-
4 8553739) of the MYH10 gene. The putative consensus binding sites for GATA1 within intron 8 of the
5 MYH10 gene are displayed in italic and underlined in green. Both fragments were cloned into a
6 luciferase construct, 5' of the luciferase reporter gene, to generate pGL3-3'UTR MYH10 WT and
7 pGL3-Intron 8 MYH10 WT. The binding-site mutant constructs are pGL3-Intron 8 MYH10 M1, pGL3-
8 Intron 8 MYH10 M2, pGL3-Intron 8 MYH10 M3, pGL3-Intron 8 MYH10 M4 and pGL3-Intron 8 MYH10
9 M(1,2,3,4) with the indicated mutations.

12

1
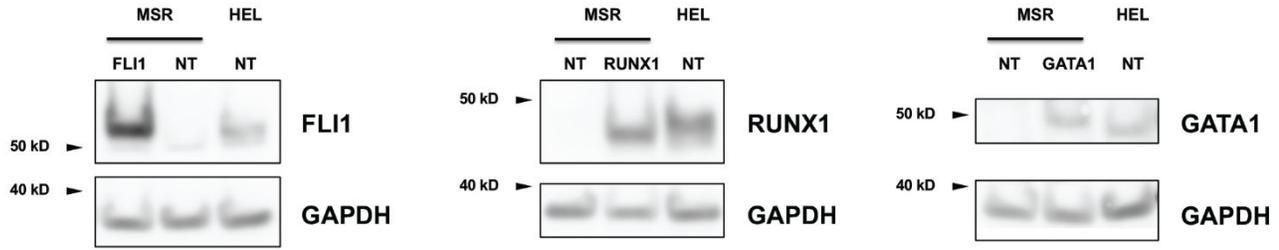
2

3 **Supplemental Figure 8: FLI1, RUNX1 and GATA1 transcription factor cell expression**

4 Western blot analysis of FLI1, RUNX1 and GATA1 expression in native (non-transfected, NT) HEL

5 cells, native GripTite 293 MSR cell lines, and transfected (with RUNX1, FLI1 or GATA1) GripTite 293

6 MSR cell lines. GAPDH expression was assessed as a protein loading control.

7

**MSR cells transfected with
pGL3-intron8 MYH10 and pCDNA3-GATA**

1

2
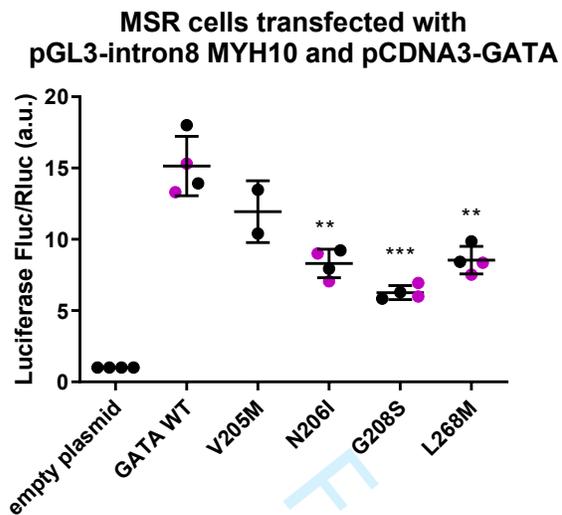
Supplemental Figure 9: Functional analysis of the MYH10 regulatory regions.

Luciferase expression after transfection in GripTite 293 MSR cells with pGL3-intron 8 MYH10 WT
and pCDNA3-GATA1 WT or variant vectors as indicated. ** $p<0.01$ vs. GATA1 WT (one-way ANOVA
corrected for multiple comparisons). Each plasmid was assayed in duplicate in one or two separate
transfection experiments (pink or black symbols).

8

1
2
3
4 1 **Supplemental Table 1 Candidate genes**
5 2
6
7 3
8
9 4

Gene	Variation (NM)	gnomAD exome (v2.0.1 Exomes global MAF)	gnomAD genome (Genomes global MAF)	gnomAD (v3 Genomes global MAF)	dbSNP rsid (*Identifier for NCBI dbSNP*)	Clinvar interpretation (v20210123)	CADD raw: *[-6.41;35.5] The higher the less likely to be observed*	CADD phred *Phred-like scaling of raw score*	MPA score *Raw score [0;10], 10: high impact*	MPA impact
WDR66	c.1915C>T (NM_144668.5)	0.0001	9.686e-05	8.373e-05	rs202238684	No match	6.3	35	10	Nonsense
LYST	c.11268-5del (NM_000081.3)(	0.5087	0.5810	0.5679	No match	Benign / Likely benign	0.079599	1.915	0	Unknown
GCKR	c.217-8C>T (NM_001486.3)	0.0001	No match	2.791e-05	rs755995385	No match	0.26	3.88	0	Unknown
CD226	c.-33delT (NM_001303618.1)	0.0117	0.0089	0.0091	rs3214811	No match	0.017666	1.450	0	Unknown
RNF145	c.1270-9_1270-4del (NM_001199383.1)	0.0015	0.0014	0.0017	rs59997682	No match	0.704591	8.497	0	Unknown

5

32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57 15
58
59
60

1 **Supplemental Table 2 Morphometric parameters from sections of platelet electron**
2 **microscopy.**
3 See the Methods section for the comparison between the patient and the controls

4

	Platelet area (μm^2)	Major platelet diameter (μm)	Minor platelet diameter (μm)	Alpha granule diameter (nm)	Alpha granule per μm^2
Controls (n=10)	2.79±1.60	2.85±1.23	1.25±1.49	235±1.2	2.51±1.29
F2-II-1	2.74±1.73	2.75±1.26	1.27±1.51	208±1.4	1.08±0.81
p-value	0.751	0.104	0.663	<0.001	<0.001

5

16

1
2
3
4 1 **Supplemental references**
5

6 2 1 Saultier P, Vidal L, Canault M, Bernot D, Falaise C, Pouymayou C, Bordet J-C, Saut N, Rostan A,
7 3 Baccini V, Peiretti F, Favier M, Lucca P, Deleuze J-F, Olaso R, Boland A, Morange PE, Gachet C,
8 4 Malergue F, Fauré S, et al. Macrothrombocytopenia and dense granule deficiency associated
9 5 with FLI1 variants: ultrastructural and pathogenic features. *Haematologica* 2017; **102**: 1006–
10 6 16.
11
12
13 7 2 Wang K, Li M, Hakonarson H. ANNOVAR: functional annotation of genetic variants from high-
14 8 throughput sequencing data. *Nucleic Acids Res* 2010; **38**: e164.
15
16 9 3 Desvignes J-P, Bartoli M, Delague V, Krahn M, Miltgen M, Béroud C, Salgado D. VarAFT: a
17 10 variant annotation and filtration system for human next generation sequencing data. *Nucleic
18 11 Acids Res* 2018; **46**: W545–53.
19
20 12 4 Frédéric MY, Lalande M, Boileau C, Hamroun D, Claustres M, Béroud C, Collod-Béroud G.
21 13 UMD-predictor, a new prediction tool for nucleotide substitution pathogenicity -- application
22 14 to four genes: FBN1, FBN2, TGFBR1, and TGFBR2. *Hum Mutat* 2009; **30**: 952–9.
23
24 15 5 Choi Y, Sims GE, Murphy S, Miller JR, Chan AP. Predicting the functional effect of amino acid
25 16 substitutions and indels. *PLoS ONE* 2012; **7**: e46688.
26
27 17 6 Adzhubei IA, Schmidt S, Peshkin L, Ramensky VE, Gerasimova A, Bork P, Kondrashov AS,
28 18 Sunyaev SR. A method and server for predicting damaging missense mutations. *Nat Methods*
29 19 2010; **7**: 248–9.
30
31 20 7 Caldwell JT, Edwards H, Dombkowski AA, Buck SA, Matherly LH, Ge Y, Taub JW.
32 21 Overexpression of GATA1 confers resistance to chemotherapy in acute megakaryocytic
33 22 Leukemia. *PLoS ONE* 2013; **8**: e68601.
34
35 23 8 Poggi M, Canault M, Favier M, Turro E, Saultier P, Ghalloussi D, Baccini V, Vidal L, Mezzapesa
36 24 A, Chelghoum N, Mohand-Oumoussa B, Falaise C, Favier R, Ouwehand WH, Fiore M, Peiretti F,
37 25 Morange PE, Saut N, Bernot D, Greinacher A, et al. Germline variants in ETV6 underlie reduced
38 26 platelet formation, platelet dysfunction and increased levels of circulating CD34+ progenitors.
39 27 *Haematologica* 2017; **102**: 282–94.
40
41 28 9 Fujiwara T, Sasaki K, Saito K, Hatta S, Ichikawa S, Kobayashi M, Okitsu Y, Fukuhara N, Onishi Y,
42 29 Harigae H. Forced FOG1 expression in erythroleukemia cells: Induction of erythroid genes and
43 30 repression of myelo-lymphoid transcription factor PU.1. *Biochem Biophys Res Commun* 2017;
44 31 **485**: 380–7.
45
46 32 10 Mehaffey MG, Newton AL, Gandhi MJ, Crossley M, Drachman JG. X-linked thrombocytopenia
47 33 caused by a novel mutation of GATA-1. *Blood* 2001; **98**: 2681–8.
48
49 34 11 Yu C, Niakan KK, Matsushita M, Stamatoyannopoulos G, Orkin SH, Raskind WH. X-linked
50 35 thrombocytopenia with thalassemia from a mutation in the amino finger of GATA-1 affecting
51 36 DNA binding rather than FOG-1 interaction. *Blood* 2002; **100**: 2040–5.

12   Balduini CL, Pecci A, Loffredo G, Izzo P, Noris P, Grosso M, Bergamaschi G, Rosti V, Magrini U, Ceresa IF, Conti V, Poggi V, Savoia A. Effects of the R216Q mutation of GATA-1 on erythropoiesis and megakaryocytopoiesis. *Thromb Haemost* 2004; **91**: 129–40.

18