Pathogenic and likely pathogenic variants in at least five genes account for ~3% of mild isolated non-syndromic thrombocytopenia

Running title: Panel sequencing in thrombocytopenia

Paul Guéguen[1,2,†], Arnaud Dupuis[3,†], Jean-Yves Py[4], Aurore Desprès[1], Emmanuelle Masson[1,2], Cédric Le Marechal[1,2], David N. Cooper[5], Christian Gachet[3], Jian-Min Chen[1], Claude Férec[1,2,*]

[1]CHU Brest, Service de Génétique Médicale et de Biologie de la Reproduction, Brest, France.

[2]EFS, Univ Brest, Inserm, UMR 1078, GGB, F-29200 Brest, France.

[3]Université de Strasbourg, Institut National de la Santé et de la Recherche Médicale, Etablissement Français du Sang Grand Est, Unité Mixte de Recherche-S 1255, Fédération de Médecine Translationnelle de Strasbourg, Strasbourg, France.

[4]EFS Centre-Pays de la Loire, site d'Orléans, France.

[5]Institute of Medical Genetics, School of Medicine, Cardiff University, Cardiff, United Kingdom

†Authors share co-first authorship.

***Correspondence:**

Prof. Claude Férec, EFS and INSERM UMR1078, 46 rue Félix Le Dantec, 29218 BREST, France. Tél.: +33(02)-98 44 41 38. Fax: 33(02)-98 01 64 74. E-mail: claude.ferec@univ-brest.fr.

SOURCE(S) OF SUPPORT

This work was supported by the Etablissement Français du Sang (EFS), the Association de Transfusion Sanguine et de Biogénétique Gaétan Saleün and the Institut National de la Santé et de la Recherche Médicale (INSERM).

CONFLICT OF INTEREST

The authors declare no conflict of interest.

Abstract

BACKGROUND: Thrombocytopenia has a variety of different etiologies, both acquired and hereditary. Inherited thrombocytopenia may be associated with other symptoms (syndromic forms) or may be strictly isolated. To date, only about half of all the familial forms of thrombocytopenia have been accounted for in terms of well-defined genetic abnormalities. However, data are limited on the nature and frequency of the underlying causative genetic variants in individuals with mild isolated non-syndromic thrombocytopenia.

STUDY DESIGN AND METHODS: Thirteen known or candidate genes for isolated thrombocytopenia were included in a gene panel analysis in which targeted next-generation sequencing was performed on 448 French blood donors with mild isolated non-syndromic thrombocytopenia.

RESULTS: A total of 68 rare variants including missense, splice site, frameshift, nonsense and in-frame variants (all heterozygous) were identified in 11 of the 13 genes screened. 29% (N = 20) of the variants detected were absent from both the French Exome (FrEx) Project and gnomAD exome databases. Using stringent criteria and an unbiased approach, we classified 7 predicted loss-of-function variants (3 in *ITGA2B* and 4 in *TUBB1*) and four missense variants (1 in *GP1BA*, 2 in *ITGB3* and 1 in *ACTN1*) as being pathogenic or likely pathogenic. Altogether, they were found in 13 members (~3%) of our studied cohort.

CONCLUSION: We present the results of gene panel sequencing of known and candidate thrombocytopenia genes in mild isolated non-syndromic thrombocytopenia. Pathogenic and likely pathogenic variants in five known thrombocytopenia genes were identified, accounting for ~3% of individuals with the condition.

Key words: gene panel, mild isolated non-syndromic thrombocytopenia, missense variants, next generation sequencing, pathogenicity prediction, platelet count

INTRODUCTION

Thrombocytopenia is defined in terms of the number of circulating platelets being $<150 \times 10^9$/L of blood. The main consequence of a decrease in the number of blood platelets is to expose an individual to a risk of bleeding, especially following traumatic injury. Thrombocytopenia can result from bone marrow failure, excessive peripheral consumption of platelets or other external factors. It may however also be caused by genetic defects, and two main types of inherited thrombocytopenia have been classically distinguished: syndromic forms that affect multiple organs in combination with thrombocytopenia, and isolated forms. Inherited thrombocytopenia, irrespective of whether it is isolated or syndromic, is a very heterogeneous group of conditions that may or may not be associated with a bleeding tendency, which can range from mild to severe. Typically, platelet counts higher than 50×10^9/L do not lead to clinical problems whereas a count of less than 30×10^9/L leads to spontaneous bruising and purpura.[1]

Over the past 30 years, a series of genes causing isolated and/or syndromic thrombocytopenia have been steadily identified.[2-4] The ability to detect causative genetic variant(s) in subjects with thrombocytopenia is of great diagnostic, and potentially also therapeutic, importance.[4-6] However, to date, most studies have focused on the familial and/or moderate/severe forms of the disease, for which approximately 50% of cases are explicable by well-defined genetic abnormalities. By contrast, data on the frequency of causative genetic variants in the overall population of mild isolated non-syndromic thrombocytopenia subjects are limited[7-10]. In other words, both the fraction of mild isolated non-syndromic thrombocytopenia that is attributable to a genetic cause and the relative contribution of each causative gene, remain unclear. Herein, we have addressed these issues by performing targeted gene panel sequencing of a cohort of French blood donors with mild isolated non-syndromic thrombocytopenia.

MATERIALS AND METHODS

Ethics statement

This study was approved by the Agence Nationale de Sécurité du Médicament et des Produits de Santé (ANSM; registration No. 2014-A00002-45), the French Agency for the Safety of Drugs and Health Products. Informed consent for present and possible follow-up genetic studies was obtained from each participant.

Cases

A total of 448 blood donors (male, 364 (median age = 55, range = 18–72); female, 84 (median age = 44, range = 21–71)) who donated at least twice were recruited from three sites of the French Transfusion Organization (Establishment Français du Sang), Alsace, Bretagne and Centre Atlantique. The donors were selected on the basis of having a platelet count of $<150 \times 10^9$/L measured on at least two occasions between January 1st 2010 and December 31st 2012. On the basis of their platelet counts, they were considered to have mild isolated non-syndromic thrombocytopenia.

Genes included in panel next-generation sequencing

Thirteen known/candidate genes for isolated thrombocytopenia (*ACTN1*, *CYCS*, *FLNA*, *GP1BA*, *GP1BB*, *GP9*, *ITGA2*, *ITGA2B*, *ITGB3*, *MYH9*, *TUBA4A*, *TUBB1*, *VWF*) were included in a panel for targeted next-generation sequencing. The genes, mRNA reference accession numbers, exon numbers, coding sequence lengths, and reasons for inclusion are provided in Table 1.

AmpliSeq™ library preparation and quantification

DNA was extracted from buccal swabs using the Maxwell® 16 Blood DNA Purification Kit (Promega). Primer pairs that targeted all coding sequences, intron/exon junctions and 5'- and 3'-untranslated regions of the aforementioned 13 genes (with the exception of *VWF*, where only the exon 28 mutational hotspot was analyzed) were designed using the Ion AmpliSeq™ Designer online interface. The targeted regions were divided into two panels, each panel being PCR amplified in a 10 μl reaction with the AmpliSeq™ Library Kit 2.0 (ThermoFisher) according to the recommended protocol: activation at 99°C for 2 min followed by 16 cycles of 99°C for 15 sec and 60°C for 4 min. After the two reactions had been combined, the primer end sequences were partially digested with the addition of 2 μl FuPa Reagent, followed by incubations at 50°C for 10 min, 55°C for 10 min, and 60°C for 20 min. Subsequently, a barcode was added to each library using the Ion Xpress™ Barcode Adapters kit (ThermoFisher), followed by incubation firstly at 22°C for 30 min and then at 72°C for 10 min. The libraries were firstly purified using AMPure XP beads (Beckman Coulter) in accordance with the manufacturer's instructions and then quantified by quantitative PCR using the Ion Library Quantitation Kit (ThermoFisher).

Massively parallel sequencing (Proton Ion)

Emulsion PCR, sphere enrichment and chip loading were automated by the Ion Chef instrument in accordance with the supplier's recommendations (Ion PI™ Hi-Q™ PI kit, ThermoFisher). The quantified libraries were mixed and multiplexed at 96 samples per PI chip. Sequencing was carried out on the Ion Proton instrument (ThermoFisher) using the HiQ Ion kit. Sequencing signal processing, sequence generation and base calling, and alignment to the human genome reference sequence hg19 were processed by Torrent Suite Software v5.2.1. Variant annotation

was performed using SeqNext v4.0.1 (JSI) software and the VCF output files from the Torrent Server (variant caller 5.2).

Variant inclusion criteria

Rare variants (defined as having an allele frequency of <1% in accordance with exome data available in the gnomAD database[11]) including (i) small deletions or insertions that affected canonical GT-AG splice sites and/or coding sequence and (ii) single nucleotide substitutions that altered either canonical GT-AG splice sites or resulted in missense or nonsense mutations, were included for analysis. All variants of interest were subjected to validation by Sanger sequencing (corresponding PCR and sequencing primer sequences are available upon request).

Variant classification in terms of pathological relevance

We firstly divided the variants into predicted loss-of-function (pLoF) (in accordance with gnomAD definition[11]) and non-pLoF. The pathogenic relevance of pLoF variants was evaluated in a conservative and unbiased manner, using the corresponding types of variant in public datasets of the French exome (FrEx) project[12] and gnomAD[11] for comparison, and taking into consideration modes of disease inheritance (i.e. recessive versus dominant) with respect to previous relevant findings. Regarding non-pLoF variants, those previously reported to be pathogenic were firstly identified by means of VarSome,[13] followed by verification through cross-checking data deposited in ClinVar[14] and manual review of the corresponding original publications. The remaining variants were subjected to further analysis including pathogenicity prediction.

Pathogenicity predictions for missense variants

We adopted the PP3 rule established by VarSome,[13] whose prediction was based upon a combined score of predictions from 10 *in silico* programs (i.e., DANN, MutationTaster, dbNSFP.FATHMM, MetaSVM, MetaLR, GERP, LRT, MutationAssessor, PROVEAN, and SIFT), and employed the Combined Annotation-Dependent Depletion (CADD) tool[15], for this purpose.

Revisiting the donors carrying pathogenic and likely pathogenic variants

Donors found to carry pathogenic and likely pathogenic variants were revisited by telephone call or email with respect to their personal and familial background of bleeding history. Bleeding scores were determined by means of the International Society on Thrombosis and Haemostasis - Bleeding Assessment Tools (ISTH-BAT).[16]

Statistical analysis

The significance of the differences between variant carrier or allele frequencies in cases and controls was tested by two-sided Fisher's exact test. A P-value < 0.05 was considered to provide evidence of statistical significance.

Variant nomenclature

Variant nomenclature followed Human Genome Variation Society (HGVS) recommendations[17] and verified via VarSome.

RESULTS

Gene panel sequencing in 448 French blood donors with mild isolated non-syndromic thrombocytopenia

We included 13 genes (Table 1) in the panel sequencing of 448 French blood donors with mild isolated non-syndromic thrombocytopenia. The average sequencing depth was >1200× and the average sequencing coverage of the targeted regions was 97%. All variants of interest were subjected to confirmation by Sanger sequencing. A total of 68 rare variants were identified including missense, canonical splice site, frameshift, nonsense and inframe variants (all heterozygous) in 11 of the 13 genes screened; the two genes in which no variants were found were *CYCS* and *GP9*. It should be noted that in cases where two or more variants were found in the same gene, these variants were each counted independently, irrespective of whether or not they were located in *cis* (e.g., *TUBA4A*:c.[541G>A;547G>C]) or in *trans* (e.g., *GP1BA*:c.252delC(;)179A>G). Details of each rare variant, including the affected gene, HGVS nomenclature, rs number if applicable, allele frequencies in the French population (in accordance with FrEx) and all gnomAD populations, are provided in Supp. Table S1. 29% (N = 20) of the 68 variants were absent from both the FrEx and gnomAD exome databases. The 68 variants were found in 64 (14.3%) of the 448 donors.

Adopting a conservative and unbiased strategy to classify the detected variants in terms of pathological relevance

The American College of Medical Genetics has provided detailed guidelines for classifying variants as pathogenic, likely pathogenic, uncertain significance, likely benign and benign[18]. However, unlike those cases with classical familial forms of thrombocytopenia, this classification is often not straightforward in our blood donors with mild isolated non-syndromic

thrombocytopenia. Nonetheless, taking into consideration the established guidelines, we adopted an objective and unbiased strategy to classify the identified variants in terms of their pathological relevance. To this end, we divided the variants into predicted loss-of-function (pLoF) and non-pLoF categories.

Classification of pLoF variants

Nine of the 68 rare variants fell into the pLoF category, including canonical splice site, frameshift and nonsense variants (2 in *GP1BA*, 3 in *ITGA2B*, and 4 in *TUBB1*; Table 2). None were present in the FrEx database. Four of these variants were also absent from the gnomAD exome database. For the other five variants, their population allele frequencies in the gnomAD exome database were extremely low (<0.0001). In addition, a survey of the nine pLoF variants in VarSome indicated that only *TUBB1* c.35delG has previously been reported to be associated with a human disease state, thyroid dysgenesis[19].

To assess objectively the contribution of the pLoF variants to mild isolated non-syndromic thrombocytopenia, a comparison with the corresponding type of variant from a control group is required. In this regard, the FrEx database, which currently contains whole exome sequencing data from 574 French subjects,[12] was the preferred option. A systematic survey of FrEx variants in the 13 genes under study identified only one rare pLoF variant, *ITGA2B* c.1413C>A (p.Tyr471Ter) (Table 2). It was detected once in the 574 French subjects, was absent from the gnomAD exome database, and has not been previously reported in the literature. However, the apparent overrepresentation of aggregated pLoF variants in cases as compared to controls achieved statistical significance only for the *TUBB1* gene (5/448 vs. 0/574, $P = 0.016$) in part

due to small sample sizes. The corresponding *P* values for *GP1BA* (3/448 vs. 0/574) and *ITGA2B* (3/448 vs. 1/574) were 0.08 and 0.32.

We further extracted all rare pLoF variants in each of the three genes from the gnomAD exome dataset in the context of the non-Finnish European population (Supp. Table S2). Erring on the conservative side, we included low-confidence pLoF variants (defined as end-truncating variants or 'quality dubious' by gnomAD) for analysis, although their exclusion would not affect the overall conclusion. In this regard, it is pertinent to point out that none of the pLoF variants found in our cases would have been annotated as end-truncating by reference to variants annotated in their respective genes in gnomAD. As summarized in Supp. Table S3, the aggregated allele frequency of rare *GP1BA* pLoF variants in cases was higher than that in the non-Finnish European population but did not attain statistical significance ($P = 0.10$). Once again, erring on the conservative side, we classified the *GP1BA* pLoF variants found in our cases as being of uncertain significance, although homozygous *GP1BA* pLoF variants cause autosomal recessive Bernard-Soulier syndrome[20] and heterozygous *GP1BA* missense variants cause autosomal dominant thrombocytopenia[21-23].

By contrast, the aggregated allele frequency of rare pLoF variants in *ITGA2B* was significantly higher in cases as compared to the non-Finnish European population ($P = 0.002$; Supp. Table S3). This, taken together with the fact that homozygous *ITGA2B* pLoF variants cause autosomal recessive Glanzmann thrombasthenia[24,25], implied that *ITGA2B* haploinsufficiency may confer susceptibility to mild isolated non-syndromic thrombocytopenia. Here, one may argue that to date, only a single heterozygous variant in the *ITGA2B* gene, p.Arg995Trp, has been reported to cause autosomal dominant thrombocytopenia; and p.Arg995Trp was thought to result in a gain of function[26]. However, it should be pointed out that

the multiple subjects carrying the p.Arg995Trp variant invariably had moderate thrombocytopenia and some of them even had a bleeding tendency[26]; these phenotypes turned out to be more severe than those studied here. It is entirely possible that different types of heterozygous variant in the same gene may be associated with subtle phenotypic differences. Thus, taking a balanced view of these factors, we would tentatively classify the heterozygous *ITGA2B* pLoF variants as being likely pathogenic for mild isolated non-syndromic thrombocytopenia.

The aggregated allele frequency of rare pLoF variants in the *TUBB1* gene was significantly higher in cases as compared to the non-Finnish European population ($P = 0.0003$; Supp. Table S3). Heterozygous loss-of-function variants in the gene, including pLoF variants, have been previously reported to cause autosomal dominant macrothrombocytopenia or inherited platelet disorders of uncertain etiology on the basis of the associated clinical and biological phenotype[8,27,28]. Moreover, *Tubb1*$^{+/-}$ mice had only 75% of the platelet count of wild-type mice[29]. Taken together, heterozygous LoF variants in the *TUBB1* gene could be confidently classified as pathogenic in relation to mild isolated non-syndromic thrombocytopenia.

Classification of non-pLoF variants

The remaining 59 variants (58 missense and one three-base deletion; Supp. Table S1) fell into the category of non-pLoF variants. We employed VarSome to identify those that have been previously described to be pathogenic, resulting in the identification of four missense variants (Table 3). All four missense variants were absent from the FrEx database, and were either absent or extremely infrequent (i.e., allele frequency of <0.0001) in the gnomAD exome database (Table 3). Two of these, *GP1BA* p.Cys225Ser and *ITGB3* p.Arg240Trp, have been classified by

ClinVar as pathogenic whereas the other two, *ITGB3* p.Gly605Ser and *ACTN1* p.Arg46Gln, are not registered. We further evaluated the evidence supporting the pathogenicity of these variants in the original publications. Notably, heterozygous *GP1BA* p.Cys225Ser has recently been identified in a family with macrothrombocytopenia[8] and *ACTN1* p.Arg46Gln has previously been identified in multiple patients with autosomal dominant macrothrombocytopenia[30,31]. These two variants were thus classified as pathogenic in our cases. As for *ITGB3* p.Arg240Trp and p.Gly605Ser, both were pathogenic in the context of autosomal recessive Glanzmann's thrombasthenia[32,33]. By analogy to the earlier discussed pLoF variants in the *ITGA2B* gene, these two variants may, in the heterozygous state, predispose to mild isolated non-syndromic thrombocytopenia and were therefore classified as likely pathogenic.

To provide further evidence supporting the pathogenic role of the four aforementioned rare missense variants, we employed VarSome to identify rare FrEx missense and inframe variants in the 13 genes under study that have been previously described to be pathogenic as we did for variants found in the thrombocytopenia cases. This yielded two hits, namely *MYH9* c.3340T>C (p.Ser1114Pro) and *VWF* c.3797C>T (p.Pro1266Leu). In accordance with VarSome, both were classified by UniProt as ‘disease’ (*MYH9* c.3340T>C for Alport Syndrome with macrothrombocytopenia and *VWF* c.3797C>T for von Willebrand disease type 2). *MYH9* c.3340T>C was found in 2 of the 574 French controls although it was absent from the genomAD exome database; *VWF* c.3797C>T was detected only once in the 574 French controls and has an allele frequency of 0.0008 according to the genomAD exome database. However, these two variants were annotated with “conflicting interpretations of pathogenicity” in ClinVar. We reviewed the publications reporting the two variants and noted strong evidence against pathogenicity. Specifically, a recent study of a large French cohort of MYH9-related disorders in

terms of the mutational spectrum and genotype-phenotype correlations indicated *MYH9* c.3340T>C to be a benign polymorphism;[34] the observation that all six *VWF* c.3797C>T carriers derived from two unrelated families had normal platelet counts, both at baseline and under stress conditions, effectively excluded *VWF* c.3797C>T from being causative for thrombocytopenia.[35]

In summary, whereas four pathogenic or likely pathogenic missense variants were found in five of the 448 blood donors, no such missense variants were identified in the 13 genes under study in the 574 French controls ($P = 0.016$).

Difficulties in classifying the remaining non-pLoF missense variants with respect to possible pathological relevance

Of the remaining 55 non-pLoF variants, 54 are missense variants (the only exception is a 3-base deletion variant, *TUBB1* c.1042_1044delAAC; see Supp. Table S1). We attempted to classify these 54 missense variants in terms of their potential pathogenic relevance, employing a combination of their anticipated rarity in control populations (i.e., absence in the French population and having an allele frequency of <0.0001 in gnomAD populations) and pathogenicity predictions. Whereas FrEx and gnomAD exome data were used to determine the respective population allele frequencies, VarSome[13] and CADD[15] were used to predict pathogenicity. However, no meaningful conclusions could be drawn from this approach once the same analyses were performed on the corresponding FrEx missense variants in the 13 genes. The main issue lay with the uncertainty of interpretation about pathogenicity prediction outcomes. Before we use findings from two genes (*ITGA2B* and *ACTN1*) for illustrative purposes, two clarifications are necessary. With VarSome, we adopted its PP3 rule that classifies a variant as either pathogenic or not. By contrast, CADD provides a ranking rather than a prediction or

default cut-off, with higher scores more likely to be deleterious.[36] In this regard, it is pertinent to mention that in Ensembl, a score of above 30 is annotated as 'likely deleterious' and scores below 30 as 'likely benign'. However, Ensembl strongly recommended that “the actual score is used when assessing a variant and a cut-off appropriate to your requirements is chosen”.

Table 4 summaries the rare *ITGA2B* missense variants that were found in cases and/or French controls. Using the PP3 rule, only three *ITGA2B* missense variants were predicted to be pathogenic, and all three of these were found only in controls. As for the *ACTN1* missense variants, a slightly higher fraction of those found in controls were predicted to be pathogenic as compared to those found in cases (Table 5). In addition, we collated 14 functionally characterized *ACTN1* missense variants from a literature search and also predicted their pathological relevance. A discordance rate of 21% (3/14) was noted between VarSome predictions and *in vitro* functional data (Table 6). With reference to CADD, the only variant that showed no functional defect *in vitro* (i.e., p.Asp666Val) had a score of 28.8 whereas the five variants that showed a functional defect *in vitro* (i.e., p.Gln32Lys, p.Val105Ile , p.Thr737Asn, p.Arg738Trp and p.Glu769Lys) had scores less than 28.8 (Table 6).

Comparison of the carrier frequencies of pathogenic and likely pathogenic variants in cases and controls

Four pLoF variants in *TUBB1* (identified in five cases; Table 2), a missense variant in *GP1BA* (identified in a single case; Table 3) and a missense variant in *ACTN1* (identified in a single case; Table 4) were classified as pathogenic variants. In addition, three pLoF variants in *ITGA2B* (each identified in a single case; Table 2) and two *ITGB3* (identified in three cases; Table 3) were classified as likely pathogenic variants. By contrast, only a single variant deemed to be of

pathogenic relevance, *ITGA2B* p.Tyr471Ter (Table 2), was identified in 574 French controls. Altogether, pathogenic and likely pathogenic variants were significantly enriched in cases compared to controls (2.9% (13/448) vs. 0.17% (1/574); odds ratio = 17.12, $P = 0.00016$).

Findings from revisiting the 13 donors carrying pathogenic and likely pathogenic variants

We revisited the 13 cases carrying pathogenic or likely pathogenic variants, either by telephone or mail, with respect to a personal and familial history of bleeding. Five patients reported a personal bleeding history. Of these five subjects, one had an ISTH-BAT score of 0, three had a score of 1 and the remaining had a score of 3 (Supp. Table S1). Three cases reported a positive family history (Supp. Table S1) but we were unable to obtain DNA samples from the affected family members for genetic analysis.

DISCUSSION

This work represents an attempt to determine the fraction of mild isolated non-syndromic thrombocytopenia to which a genetic etiology may be attributed. To this end, 13 known or candidate genes for isolated thrombocytopenia (at the time of study initiation) were included for panel sequencing 448 French blood donors considered to have mild isolated non-syndromic thrombocytopenia. Of the 13 genes, 12 were included on the basis of either a known or a suspected role in isolated thrombocytopenia based upon the then available literature, the reasons being explained in Table 1. The remaining gene, *TUBA4A*, was included based upon unpublished observations in mice.[37]

A total of 68 rare variants including missense, canonical splice site, frameshift, nonsense and inframe variants were identified in 11 of the 13 genes under study. Adopting a conservative and

unbiased strategy, we classified six variants as pathogenic and another five as likely pathogenic. In particular, having observed a significant enrichment of heterozygous pLoF variants in the *ITGA2B* gene in cases as compared to the non-Finnish European controls, we proposed that haploinsufficiency in the gene may predispose to mild isolated non-syndromic thrombocytopenia and further surmised that different types of heterozygous variant may result in subtle phenotypic differences. Larger studies are however needed to confirm or refute these postulates.

Pathogenic and likely pathogenic variants account for only 16% (11/68) of the rare variants found in our thrombocytopenia cases. It is very likely that some of the remaining 55 rare variants, of which 54 are missense variants, have pathological relevance. A diverse range of *in silico* programs are currently available for predicting the pathogenic relevance of missense variants. We therefore performed this prediction for the identified missense variants and, for the purpose of an unbiased comparison, also the FrEx missense variants in the 13 genes, by means of VarSome and CADD. Although we had expected to identify a larger number of putatively pathogenic missense variants in cases than in controls, this was often not the case. Indeed, we also observed a considerable discordance rate between predictions and *in vitro* functional data for 14 previously reported *ACTN1* missense variants. These observations are consistent with recent reports indicating that currently available prediction tools should be used with caution.[38-40]

There is a final point to make. Our studied subjects were exclusively blood donors who were considered to have mild isolated non-syndromic thrombocytopenia upon detection of a platelet count of $<150 \times 10^9$/L on at least two donations. Some of these donors may however represent hitherto unrecognized cases of familial thrombocytopenia. Indeed, three of the 13 donors carrying pathogenic and likely pathogenic variants had a positive family when these subjects were revisited (Supp. Table S1). Additionally, 5 of the 13 carriers reported a personal bleeding

history but none had an INST-BAT score of >3 (Supp. Table S1). Although the ISTH-BAT scores for all the 13 carriers fell within the normal range (i.e., 0-3),[41] we would recommend that prospective donors found to carry a pathogenic and likely pathogenic variant should be declined for further blood donations on the grounds of safety.

In summary, we have performed a gene panel sequencing of known and candidate thrombocytopenia genes in a cohort of French subjects with mild isolated non-syndromic thrombocytopenia. We identified a total of 11 pathogenic variants in five known thrombocytopenia genes, together accounting for ~3% of the cases under study. In other words, a molecular diagnosis was established in ~3% of these cases. This study also provides a glimpse of the relative contributions of the different causative genes to mild isolated non-syndromic thrombocytopenia, with *TUBB1* lesions being the most prevalent, in the studied cohort. It should however be emphasized that our findings are preliminary in nature due to the growing list of genes implicated in the etiology of inherited thrombocytopenia.[4] Whole exome sequencing or whole genome sequencing of larger cohorts is expected to reveal the full spectrum of genetic determinants of the condition. Additionally, our study has demonstrated the serious limitations of currently available *in silico* tools for predicting the pathogenicity of missense variants.

ACKNOWLEDGMENTS

This work was supported by the Etablissement Français du Sang (EFS), the Association de Transfusion Sanguine et de Biogénétique Gaétan Saleün and the Institut National de la Santé et de la Recherche Médicale (INSERM)

WEB RESOURCES

CADD: https://cadd.gs.washington.edu/ (accessed 23 May 2019).

ClinVar: https://www.ncbi.nlm.nih.gov/clinvar/ (accessed 23 May 2019).

Ensembl: https://www.ensembl.org/info/genome/variation/prediction/protein_function.html (accessed 23 May 2019).

FrEx: http://lysine.univ-brest.fr/FrExAC/ (accessed 23 May 2019).

Ion AmpliSeq™ Designer online interface: http://www.ampliseq.com (accessed 23 May 2019).

ISTH-BAT: https://practical-haemostasis.com/Clinical%20Prediction%20Scores/Formulae%20code%20and%20formulae/Formulae/Bleeding-Risk-Assessment-Score/ISTH_BAT_score.html (accessed 10 June 2020)

gnomAD: https://gnomad.broadinstitute.org/ (accessed 23 May 2019).

VarSome: https://varsome.com/ (accessed 23 May 2019).

REFERENCES

1. Izak M, Bussel JB. Management of thrombocytopenia. F1000Prime Rep 2014;**6**:45.
2. Balduini CL, Savoia A. Genetics of familial forms of thrombocytopenia. Hum Genet 2012;**131**:1821-32.
3. Johnson B, Fletcher SJ, Morgan NV. Inherited thrombocytopenia: novel insights into megakaryocyte maturation, proplatelet formation and platelet lifespan. Platelets 2016;**27**:519-25.
4. Noris P, Pecci A. Hereditary thrombocytopenias: a growing list of disorders. Hematology Am Soc Hematol Educ Program 2017;**2017**:385-99.
5. Balduini CL, Savoia A, Seri M. Inherited thrombocytopenias frequently diagnosed in adults. J Thromb Haemost 2013;**11**:1006-19.
6. Greinacher A, Eekels JJM. Simplifying the diagnosis of inherited platelet disorders? The new tools do not make it any easier. Blood 2019;**133**:2478-83.
7. Simeoni I, Stephens JC, Hu F, et al. A high-throughput sequencing test for diagnosing inherited bleeding, thrombotic, and platelet disorders. Blood 2016;**127**:2791-803.
8. Bastida JM, Lozano ML, Benito R, et al. Introducing high-throughput sequencing into mainstream genetic diagnosis practice in inherited platelet disorders. Haematologica 2018;**103**:148-62.
9. Leinoe E, Zetterberg E, Kinalis S, et al. Application of whole-exome sequencing to direct the specific functional testing and diagnosis of rare inherited bleeding disorders in patients from the Oresund Region, Scandinavia. Br J Haematol 2017;**179**:308-22.
10. Johnson B, Lowe GC, Futterer J, et al. Whole exome sequencing identifies genetic variants in inherited thrombocytopenia with secondary qualitative function defects. Haematologica 2016;**101**:1170-9.
11. Lek M, Karczewski KJ, Minikel EV, et al. Analysis of protein-coding genetic variation in 60,706 humans. Nature 2016;**536**:285-91.
12. Génin E. The French Exome (FREX) Project: a population-based panel of exomes to help filter out common local variants. Genet Epidemiol 2017;**41**:691.
13. Kopanos C, Tsiolkas V, Kouris A, et al. VarSome: The Human Genomic Variant Search Engine. Bioinformatics 2018.
14. Landrum MJ, Lee JM, Benson M, et al. ClinVar: public archive of interpretations of clinically relevant variants. Nucleic Acids Res 2016;**44**:D862-8.
15. Kircher M, Witten DM, Jain P, et al. A general framework for estimating the relative pathogenicity of human genetic variants. Nat Genet 2014;**46**:310-5.
16. Rodeghiero F, Tosetto A, Abshire T, et al. ISTH/SSC bleeding assessment tool: a standardized questionnaire and a proposal for a new bleeding score for inherited bleeding disorders. J Thromb Haemost 2010;**8**:2063-5.
17. den Dunnen JT, Dalgleish R, Maglott DR, et al. HGVS recommendations for the description of sequence variants: 2016 update. Hum Mutat 2016;**37**:564-9.
18. Richards S, Aziz N, Bale S, et al. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. Genet Med 2015;**17**:405-24.
19. Stoupa A, Adam F, Kariyawasam D, et al. *TUBB1* mutations cause thyroid dysgenesis associated with abnormal platelet physiology. EMBO Mol Med 2018;**10**.
20. Ware J, Russell SR, Vicente V, et al. Nonsense mutation in the glycoprotein Ib alpha coding sequence associated with Bernard-Soulier syndrome. Proc Natl Acad Sci U S A 1990;**87**:2026-30.

21. Miller JL, Lyle VA, Cunningham D. Mutation of leucine-57 to phenylalanine in a platelet glycoprotein Ib alpha leucine tandem repeat occurring in patients with an autosomal dominant variant of Bernard-Soulier disease. Blood 1992;**79**:439-46.
22. Kunishima S, Imai T, Hamaguchi M, et al. Novel heterozygous missense mutation in the second leucine rich repeat of GPIbalpha affects GPIb/IX/V expression and results in macrothrombocytopenia in a patient initially misdiagnosed with idiopathic thrombocytopenic purpura. Eur J Haematol 2006;**76**:348-55.
23. Vettore S, Scandellari R, Moro S, et al. Novel point mutation in a leucine-rich repeat of the GPIbalpha chain of the platelet von Willebrand factor receptor, GPIb/IX/V, resulting in an inherited dominant form of Bernard-Soulier syndrome affecting two unrelated families: the N41H variant. Haematologica 2008;**93**:1743-7.
24. Newman PJ, Seligsohn U, Lyman S, et al. The molecular genetic basis of Glanzmann thrombasthenia in the Iraqi-Jewish and Arab populations in Israel. Proc Natl Acad Sci U S A 1991;**88**:3160-4.
25. Burk CD, Newman PJ, Lyman S, et al. A deletion in the gene for glycoprotein IIb associated with Glanzmann's thrombasthenia. J Clin Invest 1991;**87**:270-6.
26. Kunishima S, Kashiwagi H, Otsu M, et al. Heterozygous *ITGA2B* R995W mutation inducing constitutive activation of the alphaIIbbeta3 receptor affects proplatelet formation and causes congenital macrothrombocytopenia. Blood 2011;**117**:5479-84.
27. Kunishima S, Kobayashi R, Itoh TJ, et al. Mutation of the beta1-tubulin gene associated with congenital macrothrombocytopenia affecting microtubule assembly. Blood 2009;**113**:458-61.
28. Kunishima S, Nishimura S, Suzuki H, et al. *TUBB1* mutation disrupting microtubule assembly impairs proplatelet formation and results in congenital macrothrombocytopenia. Eur J Haematol 2014;**92**:276-82.
29. Schwer HD, Lecine P, Tiwari S, et al. A lineage-restricted and divergent beta-tubulin isoform is essential for the biogenesis, structure and function of blood platelets. Curr Biol 2001;**11**:579-86.
30. Kunishima S, Okuno Y, Yoshida K, et al. *ACTN1* mutations cause congenital macrothrombocytopenia. Am J Hum Genet 2013;**92**:431-8.
31. Gueguen P, Rouault K, Chen JM, et al. A missense mutation in the alpha-actinin 1 gene (*ACTN1*) is the cause of autosomal dominant macrothrombocytopenia in a large French family. PLoS One 2013;**8**:e74728.
32. Lanza F, Stierle A, Fournier D, et al. A new variant of Glanzmann's thrombasthenia (Strasbourg I). Platelets with functionally defective glycoprotein IIb-IIIa complexes and a glycoprotein IIIa 214Arg----214Trp mutation. J Clin Invest 1992;**89**:1995-2004.
33. Ambo H, Kamata T, Handa M, et al. Three novel integrin beta3 subunit missense mutations (H280P, C560F, and G579S) in thrombasthenia, including one (H280P) prevalent in Japanese patients. Biochem Biophys Res Commun 1998;**251**:763-8.
34. Saposnik B, Binard S, Fenneteau O, et al. Mutation spectrum and genotype-phenotype correlations in a large French cohort of MYH9-related disorders. Mol Genet Genomic Med 2014;**2**:297-312.
35. Federici AB, Mannucci PM, Castaman G, et al. Clinical and molecular predictors of thrombocytopenia and risk of bleeding in patients with von Willebrand disease type 2B: a cohort study of 67 patients. Blood 2009;**113**:526-34.
36. Rentzsch P, Witten D, Cooper GM, et al. CADD: predicting the deleteriousness of variants throughout the human genome. Nucleic Acids Res 2019;**47**:D886-D94.
37. Strassel C, Magiera MM, Dupuis A, et al. An essential role for alpha4A-tubulin in platelet biogenesis. Life Sci Alliance 2019;**2**.

38. Miosge LA, Field MA, Sontani Y, et al. Comparison of predicted and actual consequences of missense mutations. Proc Natl Acad Sci U S A 2015;**112**:E5189-98.
39. Ernst C, Hahnen E, Engel C, et al. Performance of *in silico* prediction tools for the classification of rare *BRCA1/2* missense variants in clinical diagnostics. BMC Med Genomics 2018;**11**:35.
40. Raraigh KS, Han ST, Davis E, et al. Functional assays are essential for interpretation of missense variants associated with variable expressivity. Am J Hum Genet 2018;**102**:1062-77.
41. Elbatarny M, Mollah S, Grabell J, et al. Normal range of bleeding scores for the ISTH-BAT: adult and pediatric data from the merging project. Haemophilia 2014;**20**:831-5.
42. Morison IM, Cramer Borde EM, Cheesman EJ, et al. A mutation of human cytochrome c enhances the intrinsic apoptotic pathway but causes only thrombocytopenia. Nat Genet 2008;**40**:387-9.
43. Nurden P, Debili N, Coupry I, et al. Thrombocytopenia resulting from mutations in filamin A can be expressed as an isolated syndrome. Blood 2011;**118**:5928-37.
44. Berrou E, Adam F, Lebret M, et al. Heterogeneity of platelet functional alterations in patients with filamin A mutations. Arterioscler Thromb Vasc Biol 2013;**33**:e11-8.
45. Kunishima S, Lopez JA, Kobayashi S, et al. Missense mutations of the glycoprotein (GP) Ib beta gene impairing the GPIb alpha/beta disulfide linkage in a family with giant platelet disorder. Blood 1997;**89**:2404-12.
46. Kurokawa Y, Ishida F, Kamijo T, et al. A missense mutation (Tyr88 to Cys) in the platelet membrane glycoprotein Ibbeta gene affects GPIb/IX complex expression--Bernard-Soulier syndrome in the homozygous form and giant platelets in the heterozygous form. Thromb Haemost 2001;**86**:1249-56.
47. Savoia A, Kunishima S, De Rocco D, et al. Spectrum of the mutations in Bernard-Soulier syndrome. Hum Mutat 2014;**35**:1033-45.
48. Kashiwagi H, Kunishima S, Kiyomizu K, et al. Demonstration of novel gain-of-function mutations of alphaIIbbeta3: association with macrothrombocytopenia and Glanzmann thrombasthenia-like phenotype. Mol Genet Genomic Med 2013;**1**:77-86.
49. Ghevaert C, Salsmann A, Watkins NA, et al. A nonsynonymous SNP in the *ITGB3* gene disrupts the conserved membrane-proximal cytoplasmic salt bridge in the alphaIIbbeta3 integrin and cosegregates dominantly with abnormal proplatelet formation and macrothrombocytopenia. Blood 2008;**111**:3407-14.
50. Gresele P, Falcinelli E, Giannini S, et al. Dominant inheritance of a novel integrin beta3 mutation associated with a hereditary macrothrombocytopenia and platelet dysfunction in two Italian families. Haematologica 2009;**94**:663-9.
51. Althaus K, Greinacher A. MYH9-related platelet disorders. Semin Thromb Hemost 2009;**35**:189-203.
52. Savoia A, Pastore A, De Rocco D, et al. Clinical and genetic aspects of Bernard-Soulier syndrome: searching for genotype/phenotype correlations. Haematologica 2011;**96**:417-23.
53. Sanchez-Guiu I, Anton AI, Padilla J, et al. Functional and molecular characterization of inherited platelet disorders in the Iberian Peninsula: results from a collaborative study. Orphanet J Rare Dis 2014;**9**:213.
54. Bottega R, Marconi C, Faleschini M, et al. ACTN1-related thrombocytopenia: identification of novel families for phenotypic characterization. Blood 2015;**125**:869-72.
55. Westbury SK, Turro E, Greene D, et al. Human phenotype ontology annotation and cluster analysis to unravel genetic defects in 707 cases with unexplained bleeding and platelet disorders. Genome Med 2015;**7**:36.
56. Faleschini M, Melazzini F, Marconi C, et al. *ACTN1* mutations lead to a benign form of platelet macrocytosis not always associated with thrombocytopenia. Br J Haematol 2018;**183**:276-88.

57. Yasutomi M, Kunishima S, Okazaki S, et al. ACTN1 rod domain mutation associated with congenital macrothrombocytopenia. Ann Hematol 2016;**95**:141-4.

Table 1. The 13 genes analyzed in this study

Gene	Chr.	mRNA reference	Number of exons	Coding sequence (bp)	Disease association reported at the time of study initiation
ACTN1	14q24.1	NM_001102.3	21	2679	Autosomal dominant thrombocytopenia[30,31]
CYCS	7p15.3	NM_018947.5	3	318	Autosomal dominant thrombocytopenia[42]
FLNA	Xq28	NM_001110556.1	48	7944	Heterozygous variants reported to cause isolated thrombocytopenia[43,44]
GP1BA	17p13.2	NM_000173.6	2	1959	Heterozygous variants reported to cause autosomal dominant thrombocytopenia[21-23]
GP1BB	22q11.21	NM_000407.4	2	621	Two variants potentially responsible for dominant macrothrombocytopenia[45,46]
GP9	3q21.3	NM_000174.4	3	534	For exploration, since the von Willebrand receptor complex comprises GP1BA, GP1BB, GP5 and GP9; biallelic variants in either *GP1BA*, *GP1BB* or *GP9* cause autosomal recessive Bernard-Soulier syndrome;[47] unlike *GP1BA* and *GP1BB*, no heterozygous variants were known to cause isolated macrothrombocytopenia at the time of study initiation
ITGA2	5q11.2	NM_002203.3	30	3546	For exploration, due to the findings in *ITGA2B* and *ITGB3* genes described below
ITGA2B	17q21.31	NM_000419.4	30	3102	Heterozygous variants reported to cause dominant macrothrombocytopenia[26,48]
ITGB3	17q21.32	NM_000212.2	15	2367	Heterozygous variants reported to cause dominant macrothrombocytopenia[49,50]
MYH9	22q12.3,	NM_002473.5	41	5883	*MYH9* variants may be associated with isolated macrothrombocytopenia[51]
TUBA4A	2q35	NM_006000.2	4	1347	For exploration (based upon then unpublished observations in mice[37])
TUBB1	20q13.32	NM_030773.3	4	1356	Autosomal dominant macrothrombocytopenia[27,28]
VWF	12p13.31	NM_000552.4	Only exon 28 was included	1500 (exon 28)	Chronic thrombocytopenia is typical of platelet-type or type 2B von Willebrand disease

Table 2. Predicted loss-of-function variants in cases and controls

Gene	Carrier code	Nucleotide change	Amino acid change	Allele frequency in French controls[a]	Allele frequency in all gnomAD populations[b]
Variants in the 448 French blood donors with isolated non-syndromic thrombocytopenia					
GP1BA	ABCT0090	c.252delC	p.Lys85SerfsTer27	0	0
	ABCT0253, ABCT0312	c.1601_1602delAT	p.Tyr534CysfsTer82	0	0.00004069
ITGA2B	ABCT0022	c.558C>A	p.Tyr186Ter	0	0.000004699
	ABCT0130	c.1672C>T	p.Gln558Ter	0	0
	ABCT0353	c.2626delA	p.Ser876AlafsTer34	0	0
TUBB1	ABCT0410, ABCT0416	c.35delG	p.Cys12LeufsTer12	0	0.00003252
	ABCT0141	c.58-2A>G	p.?	0	0.000008125
	ABCT0110	c.844C>T	p.Arg282Ter	0	0.00003252
	ABCT0021	c.1080dup	p.Leu361AlafsTer19	0	0
Variants in the 574 French controls					
ITGA2B	Not applicable	c.1413C>A	p.Tyr471Ter	0.000871	0

[a]In accordance with FrEx.
[b]In accordance with the gnomAD exome data.

Table 3. Previously described disease-associated missense variants in French blood donors with isolated non-syndromic thrombocytopenia[a]

Gene	Carrier code	Nucleotide change	Amino acid change	Allele frequency in genomAD populations[b]	Comment on pathogenicity		
					VarSome description	ClinVar classification	Key supporting evidence
GP1BA	ABCT0432	c.673T>A	p.Cys225Ser	0.000004061	"Same amino acid change as a previously established pathogenic variant regardless of nucleotide change"	Pathogenic	Homozygosity in patients with Bernard-Soulier syndrome[8,52,53]; heterozygosity in a family with macrothrombocytopenia[8].
ITGB3	ABCT0253, ABCT0312	c.718C>T	p.Arg240Trp	0.00001231	"UniProt classifies this variant as 'disease' (Glanzmann Thrombasthenia)"	Pathogenic	Homozygosity in a patient with Glanzmann's thrombasthenia[32]
ITGB3	ABCT0242	c.1813G>A	p.Gly605Ser	0.000008122	"Same amino acid change as a previously established pathogenic variant regardless of nucleotide change"	Not registered	Compound heterozygous *ITGB3* variants, c.1813G>A and c.917A>C (p.His306Pro), in a patient with Glanzmann's thrombasthenia; both variants were shown to be of functional significance in transfected CHO cells[33]
ACTN1	ABCT0108	c.137G>A	p.Arg46Gln	0	"UniProt classifies this variant as 'disease' (Bleeding Disorder, Platelet-Type 15)"	Not registered	Reported in multiple patients with autosomal dominant macrothrombocytopenia[30,31]; shown to be of functional significance *in vitro*[30]

[a]All four variants were absent from FrEx.

[b]In accordance with the gnomAD exome data via VarSome.

Table 4. Pathogenicity predictions of rare *ITGA2B* missense variants under study

Nucleotide change	Amino acid change	Carrier code	Allele frequency in French controls[a]	Allele frequency in gnomAD populations[b]	Pathogenic supporting (PP3)[c]	CADD score
Variants found in the 448 French cases but not in the 574 controls						
c.109G>C	p.Val37Leu	ABCT0187	0	0	No	12.24
c.316G>A	p.Glu106Lys	ABCT0266	0	0.00009846	No	23.7
c.655G>A	p.Gly219Ser	ABCT0163	0	0	No	24.5
c.1708G>A	p.Gly570Arg	ABCT0397, ABCT0117	0	0.0002037	No	5.103
c.2033C>T	p.Ala678Val	ABCT0263	0	0.0001477	No	24.9
c.2747G>A	p.Cys916Tyr	ABCT0358	0	0	No	25.3
Variants present in controls but not cases						
c.1125C>A	p.Ser375Arg	Not applicable	0.000871	0	No	15.82
c.1162G>A	p.Gly388Ser	Not applicable	0.000871	0	Yes	30
c.1375G>A	p.Asp459Asn	Not applicable	0.000871	0.00005279	Yes	33
c.1414G>A	p.Gly472Arg	Not applicable	0.000871	0.00003655	Yes	22.9
c.1627C>T	p.Arg543Trp	Not applicable	0.000871	0.0007028	No	23.3
c.1949C>T	p.Thr650Met	Not applicable	0.000871	0.00006457	No	0.401
c.2614C>A	p.Leu872Met	Not applicable	0.000871	0.001398	No	22.1
c.3112G>C	p.Gly1038Arg	Not applicable	0.000871	0	No	19.82
Variants found in both cases and controls						
c.2363C>A	p.Ala788Asp	ABCT0024	0.000871	0.0002884	No	27.1
c.2602G>A	p.Val868Met	ABCT0111, ABCT0035	0.00348	0.002323	No	18.45
c.2965G>A	p.Ala989Thr	ABCT0278, ABCT0367	0.000871	0.0006174	No	14.85

[a]In accordance with FrEx.
[b]In accordance with the gnomAD exome data via VarSome.
[c]In accordance with VarSome's PP3 rule.

Table 5. Pathogenicity predictions of rare *ACTN1* missense variants under study

Nucleotide change	Amino acid change	Carrier code	Allele frequency in French controls[a]	Allele frequency gnomAD populations[b]	Pathogenic supporting (PP3)[c]	CADD score
Variants found in the 448 French cases						
c.179G>A	p.Arg60Gln	ABCT0103	0	0.000008149	No	25.7
c.898C>T	p.Arg300Trp	ABCT0090	0	0	Yes	31
c.950G>A	p.Arg317Gln	ABCT0347	0	0.000004061	Yes	32
c.1427G>A	p.Cys476Tyr	ABCT0279	0	0	No	24.8
c.2239C>G	p.Gln747Glu	ABCT0193	0	0.00006091	No	24.6
Variants present in the 574 French controls						
c.146G>A	p.Gly49Glu	Not applicable	0.000871	0.000008163	No	23.8
c.682G>A	p.Val228Ile	Not applicable	0.000871	0.0001138	Yes	22.2
c.1115A>C	p.Glu372Ala	Not applicable	0.000871	0	Yes	27.7
c.1348C>T	p.Arg450Cys	Not applicable	0.000871	0	Yes	33
c.1760C>T	p.Ala587Val	Not applicable	0.000871	0.00004468	No	23.7
c.2033T>C	p.Ile678Thr	Not applicable	0.000871	0	Yes	25

[a]In accordance with FrEx.
[b]In accordance with the gnomAD exome data via VarSome.
[c]In accordance with VarSome's PP3 rule.

Table 6. Pathogenicity predictions of 14 previously reported rare *ACTN1* missense variants that have been subjected to *in vitro* functional analysis

Nucleotide change	Amino acid change	Pathogenicity confirmed by functional analysis	Original reference(s)	Allele frequency in gnomAD populations[a]	Pathogenic supporting (PP3)[b]	CADD score
c.64G>A	p.Asp22Asn	Yes	Bottega et al. (2015)[54]	0	No	29.8
c.94C>A	p.Gln32Lys	Yes	Kunishima et al. (2013)[30]	0	No	28.2
c.136C>T	p.Arg46Trp	Yes	Bottega et al. (2015)[54]; Westbury et al. (2015)[55]; Faleschini et al. (2018)[56]	0.000004084	Yes	33
c.137G>A	p.Arg46Gln	Yes	Kunishima et al. (2013)[30]; Guéguen et al. (2013)[31]; Westbury et al. (2015)[55]	0	Yes	31
c.313G>A	p.Val105Ile	Yes	Kunishima et al. (2013)[30]; Westbury et al. (2015)[55]; Faleschini et al. (2018)[56]	0	Yes	25.9
c.673G>A	p.Glu225Lys	Yes	Kunishima et al. (2013)[30]; Bottega e al. (2015)[54]; Westbury et al. (2015)[55]; Faleschini et al. (2018)[56]	0.000004062	Yes	31
c.751G>A	p.Gly251Arg	Yes	Bottega e al. (2015)[54]	0	Yes	32
c.1184T>A	p.Leu395Gln	Yes	Yasutomi et al. (2016)[57]	0	Yes	29.1
c.1997A>T	p.Asp666Val	No	Bottega e al. (2015)[54]	0	Yes	28.8
c.2210C>A	p.Thr737Asn	Yes	Bottega et al. (2015)[54]; Westbury et al. (2015)[55]	0	Yes	27
c.2212C>T	p.Arg738Trp	Yes	Kunishima et al. (2013)[30]; Bottega et al. (2015)[54]; Westbury et al. (2015)[55]	0	Yes	28.5
c.2255G>A	p.Arg752Gln	Yes	Kunishima et al. (2013)[30]; Bottega et al. (2015)[54]; Faleschini et al. (2018)[56]	0.00001218	Yes	33
c.2290G>A	p.Gly764Ser	Yes	Bottega e al. (2015)[54]	0.000004089	Yes	31
c.2305G>A	p.Glu769Lys	Yes	Bottega e al. (2015)[54]	0.00003248	No	25.8

[a]In accordance with the gnomAD exome data via VarSome.
[b]In accordance with VarSome's PP3 rule.