



Article scientifique

Article

2022

Accepted version

Open Access

This is an author manuscript post-peer-reviewing (accepted version) of the original publication. The layout of the published version may differ .

---

## A homozygous duplication of the *FGG* exon 8-intron 8 junction causes congenital afibrinogenemia. Lessons learned from the study of a large consanguineous Turkish family

---

Guipponi, Michel; Masclaux, Frédéric; Sloan Bena, Frédérique; Di Sanza, Corinne; Özbek, Namik; Peyvandi, Flora; Menegatti, Marzia; Casini, Alessandro; Malbora, Baris; Neerman Arbez, Marguerite

### How to cite

GUIPPONI, Michel et al. A homozygous duplication of the *FGG* exon 8-intron 8 junction causes congenital afibrinogenemia. Lessons learned from the study of a large consanguineous Turkish family. In: *Haematologica*, 2022, vol. 107, n° 5, p. 1064–1071. doi: 10.3324/haematol.2021.278945

This publication URL: <https://archive-ouverte.unige.ch/unige:154133>

Publication DOI: [10.3324/haematol.2021.278945](https://doi.org/10.3324/haematol.2021.278945)

# **A HOMOZYGOUS DUPLICATION OF THE *FGG* EXON 8-INTRON 8 JUNCTION CAUSES CONGENITAL AFIBRINOGENEMIA**

Lessons learned from the study of a large consanguineous Turkish family

Michel Guipponi<sup>1,2</sup>, Frédéric Masclaux<sup>1</sup>, **Frédérique Sloan-Béna<sup>1</sup>**, Corinne Di Sanza<sup>2</sup>, Namik Özbek<sup>3</sup>, Flora Peyvandi<sup>4,5</sup>, Marzia Menegatti<sup>4</sup>, Alessandro Casini<sup>6</sup>, Baris Malbora<sup>7</sup>, Marguerite Neerman-Arbez<sup>2\*</sup>

<sup>1</sup>Medical Genetics Service, University Hospitals of Geneva, Geneva, Switzerland.

<sup>2</sup>Department of Genetic Medicine and Development, Faculty of Medicine, University of Geneva, Geneva, Switzerland.

<sup>3</sup>Department of Pediatric Hematology, Ankara City Hospital, Ankara, Turkey.

<sup>4</sup>Fondazione IRCCS Ca' Granda Ospedale Maggiore Policlinico, Angelo Bianchi Bonomi Hemophilia and Thrombosis Center and Fondazione Luigi Villa, Milan, Italy.

<sup>5</sup>Università degli Studi di Milano, Department of Pathophysiology and Transplantation, Milan, Italy.

<sup>6</sup>Division of Angiology and Hemostasis, University Hospitals of Geneva, Geneva, Switzerland

<sup>7</sup>Department of Pediatric Hematology and Oncology, Istanbul Yeni Yuzyil University, Istanbul, Turkey.

\*To whom correspondence should be addressed

## **ABSTRACT**

Congenital afibrinogenemia is the most severe congenital fibrinogen disorder, characterised by undetectable fibrinogen in circulation. Causative mutations can be divided into two main classes: null mutations with no protein production at all and missense mutations producing abnormal protein chains that are retained inside the cell. The vast majority of cases are due to single base pair mutations or small insertions **or** deletions in the coding regions or intron-exon junctions of *FGB*, *FGA* and *FGG*. Only a few large rearrangements have been described, all deletions involving *FGA*. Here we report the characterization of a 403 bp duplication of the *FGG* exon 8-intron 8 junction accounting for congenital afibrinogenemia in a large consanguineous family from Turkey. This mutation, which had escaped detection by Sanger sequencing of short PCR amplicons of coding sequences and splice sites, was identified by studying multiple alignments of reads obtained from Whole Exome Sequencing of a heterozygous individual followed by PCR amplification and sequencing of a larger portion of *FGG*. Because the mutation duplicates the donor splice site of intron 8, we predicted that the impact of the mutation would be on *FGG* transcript splicing. Analysis of mRNAs produced by cells transiently transfected with normal or mutant minigene constructs showed that the duplication causes production of several aberrant *FGG* transcripts generating premature truncating codons.

## **Keywords**

**Fibrinogen, mutation, duplication, bleeding, thrombosis**

## INTRODUCTION

The ultimate goal of the coagulation cascade is the controlled conversion by thrombin of fibrinogen into fibrin which forms a polymer to give stability, strength and adhesive surfaces to growing blood clots containing platelets and red blood cells. Human fibrinogen (1, 2) is produced in the liver from three homologous polypeptide chains, B $\beta$ , A $\alpha$  and  $\gamma$  encoded by the fibrinogen gene cluster *FGB*, *FGA* and *FGG*, on human chromosome 4. Two copies of each polypeptide chain assemble to form a hexamer (A $\alpha$ B $\beta$  $\gamma$ )<sub>2</sub> held together by disulphide bonds. Alternative spliced transcripts are produced for *FGA* and *FGG*, these are A $\alpha$ E and  $\gamma$ ' respectively. While A $\alpha$ E chains are present in only 1-2% of circulating fibrinogen,  $\gamma$ ' chains are present in 8-15% of circulating fibrinogen, in heterodimeric or homodimeric form. Normal plasma fibrinogen levels vary between 2 and 4 g L<sup>-1</sup>. Variation in fibrinogen levels is a complex trait, influenced by both the environment and genotype.

Inherited disorders of fibrinogen include Type I disorders (afibrinogenemia and hypofibrinogenemia) which affect the quantity of fibrinogen in circulation and Type II disorders (dysfibrinogenemia and hypodysfibrinogenemia) which affect the quality of circulating fibrinogen (3, 4). Congenital afibrinogenemia is the most severe disorder, characterised by undetectable fibrinogen in circulation (5, 6).

While the first dysfibrinogenemia mutation was identified as early as 1968 (7) the molecular basis of afibrinogenemia was elucidated much later (8). This disorder is characterized by autosomal recessive inheritance and the complete absence of fibrinogen in plasma. In populations where consanguineous marriages are common, the prevalence of afibrinogenemia is increased (9).

We identified the first causative mutation for congenital afibrinogenemia, a large, recurrent deletion in *FGA* in 1999 (8) identified in homozygosity in 4 members of a Swiss family. Since then, the underlying molecular pathophysiology of numerous causative mutations leading to fibrinogen deficiency has been determined by our group and many others (reviewed in (4) (10-12).

Causative mutations can be divided into two main classes: null mutations with no protein production at all and missense mutations producing abnormal protein chains that are retained

inside the cell. The vast majority of cases are due to single base pair mutations or small insertions or deletions in the coding regions or intron-exon junctions of *FGB*, *FGA* and *FGG*. These can easily be identified by PCR amplification followed by Sanger sequencing or by next generation sequencing in particular whole exome sequencing. Only a few large rearrangements have been described. In addition to the recurrent deletion we identified with breakpoints in *FGA* intron 1 and the *FGA–FGB* intergenic region, three other large deletions in the fibrinogen gene cluster have been reported by others, all involving part of the *FGA* gene. These are a deletion of 1.2 kb eliminating the entire *FGA* exon 4 in a Japanese patient (13); a deletion of 15 kb, with breakpoints situated in *FGA* intron 4 and in the *FGA–FGB* intergenic region in a Thai patient (14); and a 4.1-kb deletion encompassing *FGA* exon 1 in an Italian patient (15). All patients were homozygous for the identified deletions except for the Thai patient, for whom complete maternal uniparental disomy was confirmed for the deleted chromosome 4 (14).

Rearrangements of this type cannot be identified by simple PCR analysis of coding regions. Other techniques such as array Comparative Genomic Hybridization (CGH) can be useful in some cases, however the resolution of commercial arrays limits the discovery of rearrangements i.e. deletions, duplications greater than 15kb. Consequently mutations less than 15 kb will escape detection using this technique in most diagnostic settings.

Here we report the characterization of a 403 bp duplication of the *FGG* exon 8-intron 8 junction accounting for congenital afibrinogenemia in a large consanguineous family from Turkey. This mutation, which had escaped detection by Sanger sequencing of short PCR amplicons of coding sequences and splice sites, was identified by studying multiple alignments of reads obtained from Whole Exome Sequencing of a heterozygous individual followed by PCR amplification and sequencing of a larger portion of *FGG*. The mutation duplicates the donor splice site of intron 8 which leads to aberrant splicing of both the major  $\gamma$  transcript and the minor  $\gamma'$  transcript.

## **METHODS**

### **Patient samples**

This study was performed with institutional review board approval and with written informed consent from all patients, in accordance with the Declaration of Helsinki. Platelet-poor-plasma samples were obtained from citrated venous blood and analyzed as described in Supplementary Methods.

### **PCR and Sanger sequencing**

Genomic DNA was extracted from whole blood-EDTA according to standard protocols. PCR amplifications of all *FGB*, *FGA* and *FGG* coding regions and intron-exon junctions were performed as previously described (16). Standard primer sequences and PCR protocols are available on demand. Specific primer sequences for this study are available in Supplementary Methods. Sanger sequencing of purified PCR products was performed by Fasteris AG, Geneva, Switzerland.

### **Array CGH**

The array-CGH analysis was performed using Human Genome CGH Microarray Kit G3 1 M (Agilent Technologies, Palo Alto, USA) with ~2.4KB overall median probe spacing according to protocols provided by the manufacturers. Copy number variant analysis was done using the Agilent Genomic Workbench Software 7.0.4.0. and UCSC Genome Browser Human Genome GRCh37/hg19 .

### **Next Generation Sequencing**

Whole exome sequencing (WES) was performed at the Health 2030 Genome Centre at Campus Biotech, Geneva using IDT Research Exome reagents. Read mapping and variant calling were performed using BWA 0.7.13, Picard 2.9.0, GATK HaplotypeCaller 3.7, aligned to the GRCh37/hg19 reference genome and annotated with Annovar 2017/07/17 and UCSC RefSeq (refGene) downloaded on 2018/08/10.

### **Minigene constructs, transfections**

PCR products including intronic and exonic sequences from *FGG* intron 7 and *FGG* exon 10 were amplified from the genomic DNA of one homozygous affected individual and one normal individual and cloned into the pcDNA3.1 V5His TOPO-TA eukaryotic expression vector (Invitrogen) to obtain mutant and normal minigenes. Presence of the 403 bp duplicated fragment in the mutant clone was confirmed by Sanger sequencing. Transient transfections of HEK-293T cells ( $10^5$  cells/ condition) were performed in 6 well plates using Lipofectamine 2000 (Invitrogen) in OptiMEM (Gibco Invitrogen) and 2 micrograms of normal or mutant construct. Two days post-transfection, cells were lysed in Trizol for RNA extraction using the Turbo DNA free kit (Invitrogen). Reverse transcription and PCR amplification of cDNA for analysis of splicing variants are described in Supplementary Methods.

## RESULTS

The members of the large consanguineous family (Figure 1) all originate from a village in Turkey which keeps a religious faith different to the surrounding villages in the area. While the surrounding villages have a 'Sunni' faith, the village where the patients live has an Alavian-Bektashi faith. Since there are no marriages between these two religious groups, all marriages are between individuals from the same village. Interestingly, the original founders migrated from the Iranian province of Mazandaran, formerly known as Taberistan, during the 9th and 10th centuries. Current inhabitants of Maznandar are also of Alavian-Bektashi faith.

Bleeding severity was assessed according to the score from the EN-RBD study. Patients were divided into clinical bleeding severity categories (asymptomatic and grade 1, 2, or 3 bleeding). Category 1 refers to provoked bleeding episodes, category 2 refers to spontaneous minor bleeding episodes (eg, bruising), and category 3 refers to spontaneous major bleeding episodes (e.g.cerebral bleeds or hemarthrosis) (17). Of the eight patients diagnosed with afibrinogenemia (Figure 1), four were available for genetic analysis. Fibrinogen measurements, both antigenic and coagulable, were performed for 41 additional family members (Table I).

Three patients available for study were male (1283, 1316, and 1317), and one was female (1314). All patients with afibrinogenemia have a severe clinical phenotype (grade 3) while most heterozygous patients with hypofibrinogenemia have a grade 0 (mean 0.5). No patient received fibrinogen on prophylaxis. Two male patients with afibrinogenemia (1283 and 1317), for which no additional clinical information is available, have experienced a thrombotic event. This is not unusual; afibrinogenemia is associated with an increased thrombotic risk even in absence of additional thrombotic risk factors, whether genetic or environmental. In a recent study of 204 afibrinogenemic patients, 37 (18 %) reported a thrombotic event (18).

### Identification of a duplication of the *FGG* exon 8- intron 8 junction

We aimed to identify the causative mutation by first screening three affected patients (1314, 1317 and 1283) and one heterozygous carrier (1288) by PCR amplification of all *FGB*, *FGA* and *FGG* coding regions and intron-exon junctions followed by Sanger sequencing as previously described (16). This approach was unsuccessful, no causative mutation was identified. We then performed array-CGH analysis and identified 23 variants not listed in the database of Genomic Variant



(<http://dgv.tcag.ca/dgv/app/home>). However none of these variants was a candidate for the afibrinogenemia phenotype so this approach was also unsuccessful.

Finally, as part of our ongoing research project determining the causative mutations and genetic modifiers of congenital fibrinogen disorders which uses Whole Exome Sequencing followed by variant calling in a panel of selected genes including the fibrinogen genes, we included one heterozygous carrier, 1292, the mother of affected patient 1283, in the study. A detailed analysis of the reads suggested the presence of an insertion of 36bp in exon 8 of the *FGG* gene NM\_021870.3,exon8,c.959\_960insAATCCACCTGCTGCAAAATATCCAGTAGTTTGGCAT. This abnormality was present in only 14.2% of the reads (14/98) covering this position. The apparent localization of the insertion in *FGG* exon 8 was in contradiction with our previous results demonstrating the absence of a causative mutation in any fibrinogen gene coding sequence. BLAT analysis revealed that the 36bp long sequence was in fact in intron 8 of *FGG* suggesting the presence of a more complex local rearrangement. The visual inspection of the reads aligned to *FGG* exon 8 further supported this hypothesis with the presence of hard clipped and outward-facing read pairs spanning far apart from that expected based on the library insert size (Figure 2). This configuration of the read pairs was highly suggestive of the presence of a heterozygous duplication that could not be fully characterized by our exon-centered analysis.

We therefore amplified by PCR a larger portion of *FGG* from intron 7 to exon 10 which yielded a 4 kb product corresponding to the normal sequence and a 4.4 kb product for affected individuals. Heterozygous individuals showed both bands, as expected (Figure 3A). Sequencing of the larger band revealed the presence of a duplicated sequence of 403 bp (out of 404, one base in a stretch of 4 intronic “A”s is missing) containing the last 169 bases from *FGG* exon 8, the donor splice site GT, and 232 additional bases of *FGG* intron 8 (Figure 3B). Genotypes for all family members are indicated in Table I. The presence of the duplication in homozygosity was associated with absence of fibrinogen in circulation and afibrinogenemia in all four affected individuals, while heterozygosity was associated with decreased fibrinogen levels (Table I).

### **The duplication causes aberrant splicing of both FGG transcripts**

Because the mutation duplicates the donor splice site of intron 8, we predicted that the impact of the mutation would be on FGG transcript splicing. Interestingly, two FGG transcripts are normally produced which differ at the 3’end: the major  $\gamma$  chain mRNA has ten exons while in the

minor  $\gamma'$  chain isoform intron 9 is not spliced out, substituting the four amino acids encoded by exon 10 with twenty  $\gamma'$  COOH-terminal residues (19, 20) (21). The presence of the duplication was thus anticipated to affect splicing of both isoforms. HEK-293T cells were transiently transfected with minigene constructs encompassing intron 7 to exon 10 with and without the duplication (Figure 4A). RNAs produced were reverse transcribed to cDNAs which were used in two different PCR reactions to amplify transcripts containing exon 10, present in the major  $\gamma$  transcript, and transcripts containing the last bases of exon 9, retained in the minor  $\gamma'$  transcript but spliced out in the major  $\gamma$  transcript. The results obtained for transcripts containing exon 10 (Figure 4B) showed one major product, indicated by an asterisk, for the normal minigene which was shown by sequencing to correspond to the correctly spliced mRNA containing exons 8, 9 and 10 encoding the major  $\gamma$  chain. Sequencing of clones of individual PCR products showed that a transcript retaining intron 9 was also produced, thus encoding the minor  $\gamma'$  transcript even though exon 10 is present. For the mutant, the major product obtained, indicated by an asterisk, retained intron 8 with the duplication, resulting in a transcript with a frameshift and a premature truncating codon 13 codons downstream.

Cloning of the PCR products allowed the identification of additional minor aberrant transcripts resulting in frameshifts and premature truncating codons. One included part of the duplicated exonic segment due to utilisation of the normal donor splice site of intron 8 and a cryptic acceptor site situated in the duplicated sequence, followed by normal splicing of intron 9 using the duplicated donor site. Another transcript showed skipping of exon 9, joining together exon 8 and exon 10. The latter product was also identified in a clone of the normal minigene. Finally, a normal correctly spliced transcript was also identified for the mutant minigene, but this is expected to be a rare event based on the intensity of the corresponding band indicated by an arrow (Figure 4B). In order to have a complete picture of the different transcripts produced by alternative splicing of exon 9, we performed a second PCR amplification using a reverse primer localised on the last bases of exon 9 which are retained in the minor  $\gamma'$  transcript but spliced out in the  $\gamma$  transcript. Again, one major product was identified for the normal construct and one for the mutant construct (Figure 4C). Sequencing confirmed production of the normal  $\gamma'$  transcript for cells transfected with the normal minigene while those transfected with the mutant minigene retained intron 8 resulting in a frameshift as described above.

## DISCUSSION

We describe here the identification of a homozygous duplication at the *FGG* exon 8-intron 8 junction accounting for congenital afibrinogenemia in a large consanguineous family from Turkey. In principle identification of the mutation causing the complete deficiency of fibrinogen in the affected individuals, necessarily homozygous given the structure of the family pedigree, should be relatively straightforward for a laboratory equipped for standard genetic screening methods i.e. PCR amplification of the three fibrinogen encoding genes followed by Sanger sequencing. Indeed, we previously studied the molecular epidemiology of causative mutations for congenital fibrinogen disorders (22) with the aim to design a cost and time- effective screening strategy based on the genetic data from 266 unrelated patients genotyped in our laboratory. When we prospectively tested our strategy on 32 consecutive new probands we found that screening of *FGA* exons 2, 4, 5 and *FGG* exon 8 combined with the search for the 11 kb deletion of *FGA* led to the identification of approximately 80% of mutated alleles, including 15 new mutations (22).

In this case the size of the duplicated sequence 403 bp was too large to be detected by standard PCR amplification approaches, which most often involve amplifying single exons with only the immediate intronic sequences, and too small to be detected by array CGH. Identification of the mutation was only possible following a deep investigation of an aberrant sequence picked up by whole exome sequencing analysis but not initially confirmed by Sanger sequencing, an approach which is unlikely to be undertaken by routine diagnostic laboratories. One lesson learned from this study is that amplifying larger overlapping portions of the genes of interest may allow the identification of similar mutations involving large insertions and duplications. In our case, amplifying for example *FGG* exons 7 to 10, around 4400 bp in the normal sequence, and allowing sufficient elongation time to amplify the larger mutant band would have enabled to identify the mutation sooner. Recent technology such as Nanopore DNA sequencing allowing sequencing of DNA fragments up to a few Mbp may be useful to identify such mutations, allowing for example the sequencing of the entire fibrinogen gene cluster with haplotype phasing of variant sites along the sequence.

The duplication is only one of five mutations involving rearrangements of more than 100 bp of the fibrinogen genes, and the first identified in the *FGG* gene. As previously mentioned four deletions of several kilobases have been identified in *FGA* including the recurrent deletion we

identified in a Swiss family (8, 13-15). In addition, an in-frame duplication of 117 bp “Fibrinogen Champagne au Mont d’Or” leads to duplication of 39 amino acids within a repetitive sequence of 13 amino acids in the connector portion of the  $\alpha$ C domain (23) a mutation predicted to cause an extension of the coiled coil. In *FGB*, while fibrinogen New York (24) is described as a deletion of the amino acids encoded by *FGB* exon 2 but the mutation is not characterized at the DNA level and as mentioned by the authors is most probably due to a splice-site mutation leading to exon 2 skipping rather than deletion of exon 2.

It is likely that many more mutations of this sort have remained elusive even for laboratories specialized in mutation identification. We hypothesized that the same duplication was likely to be found in other afibrinogenemic patients and their family members from the same geographical region or ethnic group. We therefore screened 10 unrelated patients from Turkey by PCR amplification and identified one additional afibrinogenemic patient who was homozygous for the same duplication. His parents, his grandmother and his two sisters were all heterozygous (data not shown). Enquiries into the geographical origin of this family revealed that they were originally from Kumbet village, Ortaköy, Aksaray, i.e. the same region as the first family.

While the nature of the mutation and the clear association with the phenotype i.e complete fibrinogen deficiency with a homozygous genotype and partial fibrinogen deficiency with a heterozygous phenotype did not allow any reasonable doubt that we had identified the causative mutation we wished to identify the underlying molecular mechanism. Since the mutation duplicates the donor splice site of intron 8 we predicted that the mutation would impact FGG transcript splicing, of both the major  $\gamma$  transcript and the minor  $\gamma'$  transcript. Analysis of RNAs produced by cells transiently transfected with normal or duplicated minigene constructs showed that in this model system the duplication causes production of several different aberrant transcripts of both isoforms leading to frameshifts and premature truncating codons.

Regarding the clinical significance of our findings, genotyping patients with fibrinogen disorders is now recommended in guidelines from the International Society of Thrombosis and Hemostasis (25). In quantitative fibrinogen disorders, the identification of the causative mutation(s) can help to distinguish between afibrinogenemia and severe hypofibrinogenemia. Providing an accurate diagnosis for these patients is important since in specific clinical settings such as pregnancy or surgery, patient management could be different.

In conclusion, we have identified a large duplication of several hundreds of base pairs at the *FGG* exon 8-intron 8 junction accounting for congenital afibrinogenemia in a large consanguineous Turkish family. The nature and size of the duplication can explain why this mutation was not identified using a standard PCR approach. It is highly likely that other patients with inherited quantitative fibrinogen disorders for whom no causative mutation has been identified harbor similar rearrangements.

**Acknowledgements:**

The authors thank Cédric Howald and Keith Harshman at the Health 2030 Genome Centre at Campus Biotech, Geneva for whole exome sequencing and initial processing of the data and Drs Nermin Keni and Ekrem Unal for providing patient samples from the second Turkish family. This study was funded by a grant from the Swiss National Science Foundation (grant # 31003A\_172864) to MNA.

**Author contributions:** MG, FM, FSB and CDS performed genetic experiments and interpreted the results. NO, FP, MM, AC and BM collected patient samples and clinical information and performed fibrinogen measurements. MNA directed the study and wrote the first draft of the manuscript. All authors contributed to writing and editing the final manuscript.

## References

1. Vilar R, Fish RJ, Casini A, Neerman-Arbez M. Fibrin(ogen) in human disease: both friend and foe. *Haematologica*. 2020;105(2):284-96.
2. Pieters M, Wolberg AS. Fibrinogen and fibrin: An illustrated review. *Res Pract Thromb Haemost*. 2019 Apr;3(2):161-72.
3. Casini A, Brungs T, Lavenu-Bombled C, Vilar R, Neerman-Arbez M, de Moerloose P. Genetics, diagnosis and clinical features of congenital hypodysfibrinogenemia: a systematic literature review and report of a novel mutation. *J Thromb Haemost*. 2017 May;15(5):876-88.
4. Neerman-Arbez M, de Moerloose P, Casini A. Laboratory and Genetic Investigation of Mutations Accounting for Congenital Fibrinogen Disorders. *Semin Thromb Hemost*. 2016 Jun;42(4):356-65.
5. Menegatti M, Peyvandi F. Treatment of rare factor deficiencies other than hemophilia. *Blood*. 2019 Jan 31;133(5):415-24.
6. Casini A, Neerman-Arbez M, de Moerloose P. Heterogeneity of congenital afibrinogenemia, from epidemiology to clinical consequences and management. *Blood Rev*. 2020 Dec 26:100793.
7. Blomback M, Blomback B, Mammen EF, Prasad AS. Fibrinogen Detroit--a molecular defect in the N-terminal disulphide knot of human fibrinogen? *Nature*. 1968 Apr 13;218(5137):134-7.
8. Neerman-Arbez M, Honsberger A, Antonarakis SE, Morris MA. Deletion of the fibrinogen [correction of fibrogen] alpha-chain gene (FGA) causes congenital afibrogenemia. *J Clin Invest*. 1999 Jan;103(2):215-8.
9. Peyvandi F, Mannucci PM. Rare coagulation disorders. *Thromb Haemost*. 1999 Oct;82(4):1207-14.
10. Casini A, de Moerloose P, Neerman-Arbez M. Clinical Features and Management of Congenital Fibrinogen Deficiencies. *Semin Thromb Hemost*. 2016 Jun;42(4):366-74.
11. Casini A, Neerman-Arbez M, Ariens RA, de Moerloose P. Dysfibrinogenemia: from molecular anomalies to clinical manifestations and management. *J Thromb Haemost*. 2015 Jun;13(6):909-19.
12. de Moerloose P, Casini A, Neerman-Arbez M. Congenital fibrinogen disorders: an update. *Semin Thromb Hemost*. 2013 Sep;39(6):585-95.
13. Watanabe K, Shibuya A, Ishii E, Kurihara M, Inoue S, Ono M, et al. Identification of simultaneous mutation of fibrinogen alpha chain and protein C genes in a Japanese kindred. *Br J Haematol*. 2003 Jan;120(1):101-8.
14. Spina S, Duga S, Asselta R, Peyvandi F, Mahasandana C, Malcovati M, et al. Congenital afibrinogenemia caused by uniparental isodisomy of chromosome 4 containing a novel 15-kb deletion involving fibrinogen Aalpha-chain gene. *Eur J Hum Genet*. 2004 Nov;12(11):891-8.
15. Monaldini L, Asselta R, Duga S, Peyvandi F, Karimi M, Malcovati M, et al. Mutational screening of six afibrinogenemic patients: identification and characterization of four novel molecular defects. *Thromb Haemost*. 2007 Apr;97(4):546-51.
16. Neerman-Arbez M, de Moerloose P, Bridel C, Honsberger A, Schonborner A, Rossier C, et al. Mutations in the fibrinogen alpha gene account for the majority of cases of congenital afibrinogenemia. *Blood*. 2000 Jul 1;96(1):149-52.
17. Peyvandil F, Palla R, Menegatti M, Siboni SM, Halimeh S, Faeser B, et al. Coagulation factor activity and clinical bleeding severity in rare bleeding disorders: results from the European Network of Rare Bleeding Disorders. *Journal of Thrombosis and Haemostasis*. 2012;10(4):615-21.
18. Casini A, von Mackensen S, Santoro C, Djambas Khayat C, Belhani M, Ross C, et al. Clinical phenotype, fibrinogen supplementation and health-related quality of life in patients with afibrinogenemia. *Blood*. 2021 Jan 27.
19. Mosesson MW, Siebenlist KR, Meh DA. The structure and biological features of fibrinogen and fibrin. *Ann N Y Acad Sci*. 2001;936:11-30.

20. Standeven KF, Ariens RA, Grant PJ. The molecular physiology and pathology of fibrin structure/function. *Blood Rev.* 2005 Sep;19(5):275-88.
21. de Maat MP, Verschuur M. Fibrinogen heterogeneity: inherited and noninherited. *Curr Opin Hematol.* 2005 Sep;12(5):377-83.
22. Casini A, Blondon M, Tintillier V, Goodyer M, Sezgin ME, Gunes AM, et al. Mutational Epidemiology of Congenital Fibrinogen Disorders. *Thromb Haemost.* 2018 Nov;118(11):1867-74.
23. Hanss MM, Ffrench PO, Mornex JF, Chabuet M, Biot F, De Mazancourt P, et al. Two novel fibrinogen variants found in patients with pulmonary embolism and their families. *J Thromb Haemost.* 2003 Jun;1(6):1251-7.
24. Liu CY, Koehn JA, Morgan FJ. Characterization of fibrinogen New York 1. A dysfunctional fibrinogen with a deletion of B beta(9-72) corresponding exactly to exon 2 of the gene. *J Biol Chem.* 1985 Apr 10;260(7):4390-6.
25. Casini A, Undas A, Palla R, Thachil J, de Moerloose P, Subcommittee on Factor X, et al. Diagnosis and classification of congenital fibrinogen disorders: communication from the SSC of the ISTH. *J Thromb Haemost.* 2018 Sep;16(9):1887-90.

## Figure legends

### Figure 1. Family tree of the large consanguineous family

**Figure 2. Whole exome sequencing read alignments.** A) IGV (Interactive Genome Viewer) screen capture of *FGG* gene showing reads between exon 7 and exon 10 (gene located on reverse strand, transcript NM\_000509.5) aligned to the GRCh37/hg19 reference genome. Reads are colored by pair orientation as defined by standard IGV settings. Green color defines paired-end reads orientation inconsistencies which can delineate tandem duplication with respect to the reference genome. B) Zoom on the reads that span the internal junction of the tandem duplication and only partially align to the reference genome (reads referred to as “hardclipped” at position 155’527’621).

**Figure 3. Identification of a duplicated segment at the *FGG* exon 7-intron 8 junction.** A) PCR amplification of a portion of *FGG* from intron 7 to exon 10 yields a 4 kb product corresponding to the normal sequence and a 4.4 kb product for affected individuals. Heterozygous individuals show both bands. B) Partial sequence of the 4.4 kb band. The acceptor “AG” site at the end of intron 7 and duplicated donor “GT” splice sites at the beginning of intron 8 are highlighted in yellow. Duplicated sequences are shown in blue. The duplicated sequence contains 403 bp out of 404bp of the normal sequence, one base in a stretch of 4 intronic “A”s (shown in red) is missing.

**Figure 4. Minigene constructs and RT-PCR analysis of splicing variants.** A) Minigene constructs for the normal (top) and duplicated (bottom) *FGG* sequence between intron 7 and the 3’UTR of exon 10. HEK-293T cells transiently transfected with these constructs can produce transcripts for both the major  $\gamma$  transcript, containing exon 10, and the minor  $\gamma'$  transcript, which does not splice out the 3’ portion of exon 9. B) cDNAs obtained for transcripts containing exon 10. The major product for the normal minigene, indicated by an asterisk, corresponds to the correctly spliced mRNA containing exons 8, 9 and 10 encoding the major  $\gamma$  chain. For the mutant, the major product, indicated by an asterisk, retains intron 8 with the duplication, resulting in a transcript with a frameshift and a premature truncating codon. A correctly spliced transcript

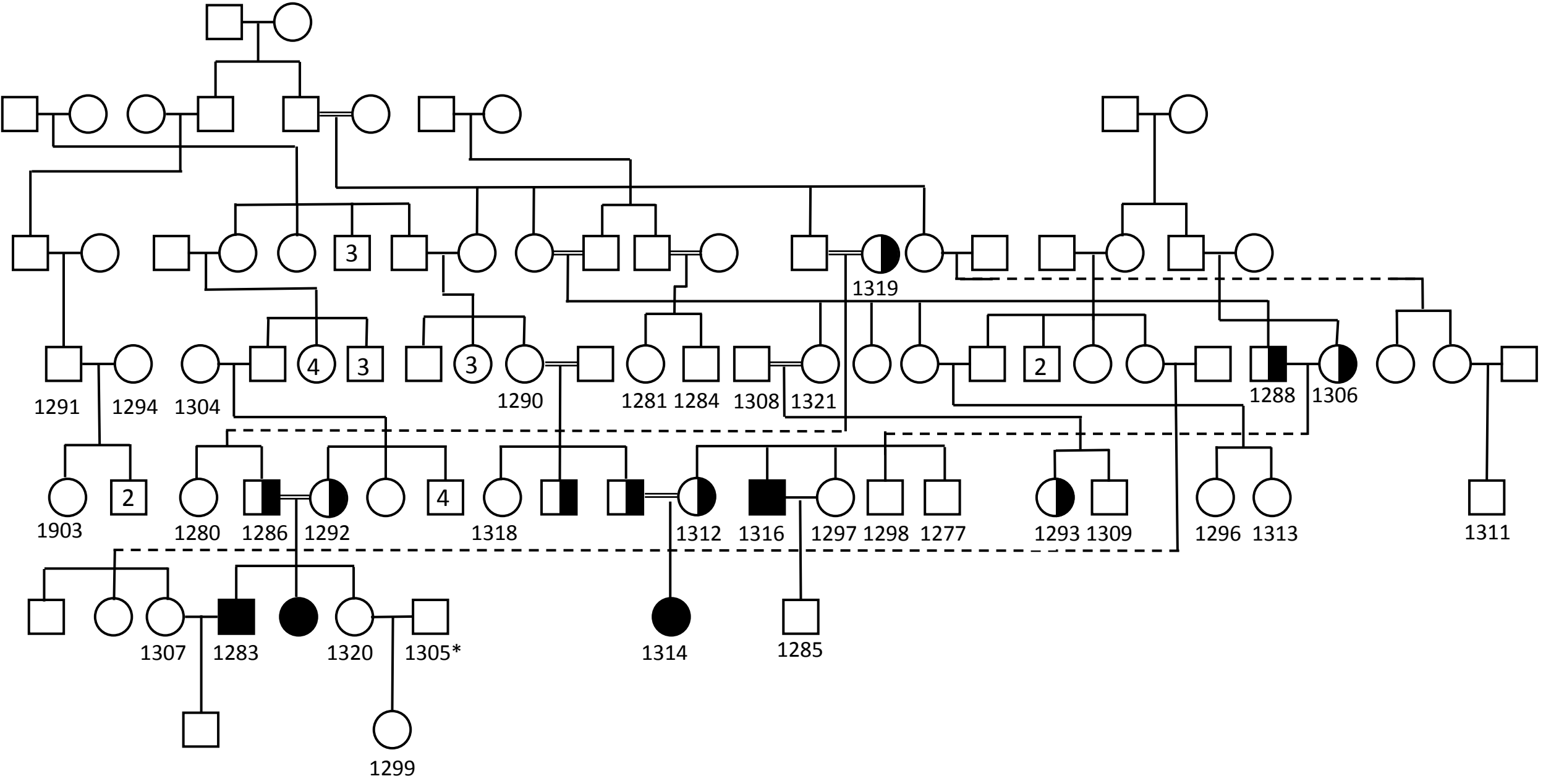


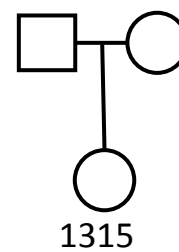
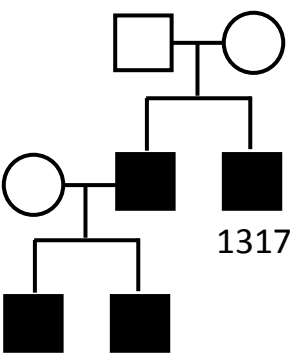
was also identified for the mutant minigene, but this is expected to be a rare event based on the intensity of the corresponding band indicated by an arrow. **C)** cDNAs obtained for transcripts containing the last bases of exon 9 retained in the minor  $\gamma'$  transcript. The major band for the normal minigene, shown with an asterisk, corresponds to the normal  $\gamma'$  transcript while the major transcript produced from the mutant minigene, retained intron 8 resulting in a frameshift as described above. Again, normal splicing is observed (indicated by an arrow) but at low levels.

**Table I Patient symptoms, fibrinogen measurements and genotypes.**

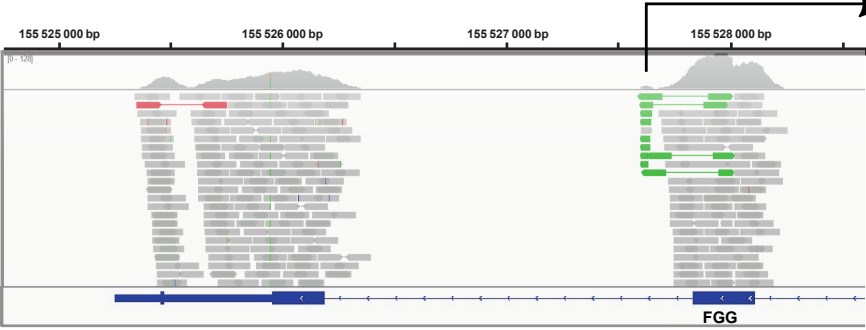
Patient ID	Fibrinogen Clauss (mg/dL)	Fibrinogen Antigen (mg/dL)	Bleeding Score	Thrombosis	Genotype
1277	197	193	0		normal
1278	273	204	0		normal
1279	282	254	0		normal
1280	169	215	0		heterozygous
1281	214	218	0		normal
1282	231	206	0		normal
1283	<20	<2	3	Yes	homozygous
1284	213	209	0		normal
1285	105	150	0		heterozygous
1286	152	177	0		heterozygous
1287	85	126	0		heterozygous
1288	152	170	1		heterozygous
1289	226	260	0		normal
1290	291	322	0		normal
1291	170	161	0		normal
1292	103	122	0		heterozygous
1293	115	110	2		heterozygous
1294	130	145	1		heterozygous
1295	273	320	0		heterozygous
1296	144	142	0		heterozygous
1297	186	175	0		normal
1298	317	334	0		normal
1299	246	267	0		normal
1300	n.d.	n.d.	0		normal
1301	272	260	0		normal
1302	334	342	0		normal
1303	90	86	2		heterozygous
1304	152	157	0		heterozygous
1305	269	267	0		normal
1306	108	132	1		heterozygous
1307	441	448	0		normal
1308	266	248	0		normal
1309	203	192	0		normal
1310	389	450	0		normal
1311	222	216	0		normal
1312	123	161	0		heterozygous
1313	149	187	0		heterozygous
1314	< 20	< 2	3		homzygous

1315	127	157	0		heterozygous
1316	< 20	< 2	3		homzygous
1317	< 20	< 2	3	Yes	homzygous
1318	338	350	0		normal
1319	121	140	1		heterozygous
1320	179	220	1		normal
1321	165	171	0		heterozygous

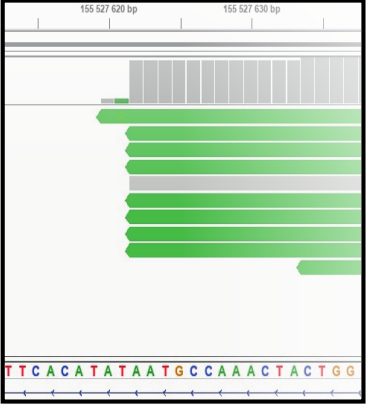




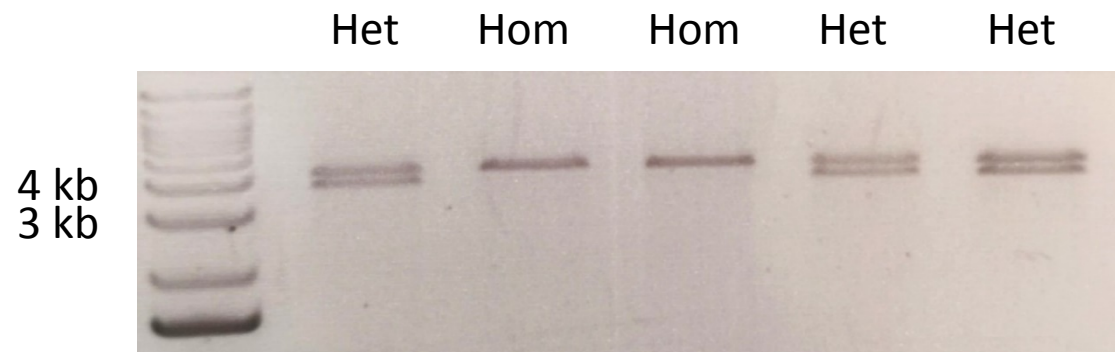
a



b



**A**



**B**

CATGTGATGGTTGTATTTCTTCTTCTCAGTACTGCAGACTATGCcATGTTCAAGGTGGGACCTGAAGCTGACAAGTACC  
GCCTAACATATGCCTACTTCGCTGGTGGGGATGCTGGAGATGCCTTTGATGGCTTTGATTTTGGCGATGATCCTAGTGAC  
AAGTTTTTACATCCCATAATGGCATGCAGTTTCACTACCTGGGACAATGACAATGATAAGTTTGAAGGCAACTGTGCTGA  
ACAGGATGGATCTGGTTGGTGGATGAACAAGTGTACGCTGGCCATCTCAATGGAGTTTATTACCAAGGTATGTTTTCTT  
TTCTTAGATTCCAAGTTAATGTATAGTGTATACTATTTTCATAAAAAATAATAATAGATATGAAGAAATGAAGAATAAT  
TTATAAAGATAGTAGGGATTTTATCATGTTCTTTATTTCAACTAAGTTCTTTGAACTGGAAGTGGATAATACCAAGTTC  
ATGCCTAAAATTAGCCCTTCTAAAGAAATCCACCTGCTGCATAATCCAGTAGTTTGGCATTTTTGGCGATGATCCTAG  
TGACAAGTTTTTACATCCCATAATGGCATGCAGTTTCACTACCTGGGACAATGACAATGATAAGTTTGAAGGCAACTGTG  
CTGAACAGGATGGATCTGGTTGGTGGATGAACAAGTGTACGCTGGCCATCTCAATGGAGTTTATTACCAAGGTATGTTT  
TCCTTTCTTAGATTCCAAGTTAATGTATAGTGTATACTATTTTCATAAAAAATAATAATAGATATGAAGAAATGAAGAA  
TAATTTATAAAGATAGTAGGGATTTTATCATGTTCTTTATTTCAACTAAGTTCTTTGAACTGGAAGTGGATAATACCAA  
GTTTCATGCCTAAAATTAGCCCTTCTAAAGAAATCCACCTGCTGCATAATCCAGTAGTTTGGCATTATATGTGAACTAT  
CACCATCATAGCTGG

**Normal:** The normal gene structure shows three exons (Exon 10, Exon 9, Exon 8) and two gamma genes (gamma and gamma'). The gamma gene is located between Exon 10 and Exon 9, and the gamma' gene is located between Exon 9 and Exon 8. The gamma gene is transcribed from the left (indicated by a right-pointing arrow) and the gamma' gene is transcribed from the right (indicated by a left-pointing arrow).

**Mutant:** The mutant gene structure shows a duplication of Exon 8, indicated by a red box and a bracket labeled "duplicated sequence". The gamma gene is located between Exon 10 and Exon 9, and the gamma' gene is located between Exon 9 and Exon 8. The gamma gene is transcribed from the left (indicated by a right-pointing arrow) and the gamma' gene is transcribed from the right (indicated by a left-pointing arrow).

Normal Mutant

2 kb

1.5 kb

1 kb

\*

\*

gamma

Normal Mutant

2 kb

1.5 kb

1 kb

gamma'