Allergology International 69 (2020) 443–449

Contents lists available at ScienceDirect

Allergology International

journal homepage: http://www.elsevier.com/locate/alit

Original Article

A novel deep intronic *SERPING1* variant as a cause of hereditary angioedema due to C1-inhibitor deficiency

Sofia Vatsiou [a, b], Maria Zamanakou [a], Gedeon Loules [a], Fotis Psarros [c], Faidra Parsopoulou [a, b], Dorottya Csuka [d], Anna Valerieva [e], Maria Staevska [e], Grzegorz Porebski [f], Krystyna Obtulowicz [f], Markus Magerl [g], Marcus Maurer [g], Matthaios Speletas [b], Henriette Farkas [d], Anastasios E. Germenis [a, b, *]

[a] CeMIA SA, Larissa, Greece
[b] Department of Immunology & Histocompatibility, University of Thessaly, School of Health Sciences, Faculty of Medicine, Larissa, Greece
[c] Department of Allergology, Navy Hospital, Athens, Greece
[d] Hungarian Angioedema Reference Center, 3rd Department of Internal Medicine, Semmelweis University, Budapest, Hungary
[e] Clinic of Allergy and Asthma, Medical University of Sofia, Sofia, Bulgaria
[f] Department of Clinical and Environmental Allergology, Jagiellonian University Medical College, Krakow, Poland
[g] Department of Dermatology and Allergy, Charité - Universitätsmedizin Berlin, Berlin, Germany

ARTICLE INFO

Article history:
Received 8 October 2019
Received in revised form
29 November 2019
Accepted 23 December 2019
Available online 17 January 2020

Keywords:
C1-inhibitor deficiency
Hereditary angioedema
Intronic mutations
Next-generation sequencing
SERPING1 gene

Abbreviations:
ACMG, American College of Medical Genetics and Genomics; ASSP, Alternative Splice Site Predictor; AMP, Association for Molecular Pathology; C1-INH, C1-inhibitor; C1-INH-HAE, hereditary angioedema due to C1-INH deficiency; CADD, Combined Annotation Dependent Depletion; GERP, Genomic Evolutionary Rate Profiling; HSF, Human Splicing Finder; ISS, intronic splicing silencer; NGS, next-generation sequencing

ABSTRACT

Background: In about 5% of patients with hereditary angioedema due to C1-inhibitor deficiency (C1-INH-HAE) no mutation in the *SERPING1* gene is detected.
Methods: C1-INH-HAE cases with no mutation in the coding region of *SERPING1* after conventional genotyping were examined for defects in the intronic or untranslated regions of the gene. Using a next-generation sequencing (NGS) platform targeting the entire *SERPING1*, 14 unrelated C1-INH-HAE patients with no detectable mutations in the coding region of the gene were sequenced. Detected variants with a global minor allele frequency lower than the frequency of C1-INH-HAE (0.002%), were submitted to *in silico* analysis using ten different bioinformatics tools. Pedigree analysis and examination of their pathogenic effect on the RNA level were performed for filtered in variants.
Results: In two unrelated patients, the novel mutation c.-22-155G > T was detected in intron 1 of the *SERPING1* gene by the use NGS and confirmed by Sanger sequencing. All bioinformatics tools predicted that the variant causes a deleterious effect on the gene and pedigree analysis showed its co-segregation with the disease. Degradation of the mutated allele was demonstrated by the loss of heterozygosity on the cDNA level. According to the American College of Medical Genetics and Genomics 2015 guidelines the c.-22-155G > T was curated as pathogenic.
Conclusions: For the first time, a deep intronic mutation that was detected by NGS in the *SERPING1* gene, was proven pathogenic for C1-INH-HAE. Therefore, advanced DNA sequencing methods should be performed in cases of C1-INH-HAE where standard approaches fail to uncover the genetic alteration.

Introduction

C1-INH-HAE (OMIM#106100) is a potentially fatal autosomal dominant disease that manifests clinically with episodes of non-pruritic and nonpitting swelling of the deeper layers of the skin or mucosa.[1] The disease is caused by mutations of the *SERPING1* gene, which encodes for C1-INH. *SERPING1* is located on chromosome 11 and consists of 8 exons.[2] Currently, 748 different disease-causing *SERPING1* variants have been published.[3] However, in

* Corresponding author. Department of Immunology & Histocompatibility, Faculty of Medicine, University of Thessaly, 3 Panepistimiou str., Biopolis, 41500, Larissa, Greece.
E-mail address: agermen@med.uth.gr (A.E. Germenis).
Peer review under responsibility of Japanese Society of Allergology.

https://doi.org/10.1016/j.alit.2019.12.009

approximately 5% of C1-INH-HAE patients no causal mutation is identified by standard mutational screening which is ordinarily restricted to the coding exons and exon-intron boundaries of the *SERPING1* gene.[4,5]

Intronic sequences were initially assumed to be largely non-functional and mutations located deep within introns (i.e., more than 100 base pairs away from exon-intron boundaries) were ignored as possible causes of human disease. However, genomic approaches in clinically oriented studies have identified many deep intronic variants with significant association to diseases, and intron functionality is supported by several independent lines of evidence.[6,7] Recently, Vaz-Drago *et al.*[7] have reviewed, between 1983 and 2016, 185 deep intronic mutations across 77 different disease-associated genes.

Hitherto, conventional methods for genotyping of C1-INH-HAE patients did not allow the analysis of *SERPING1* intronic regions. Even when some introns are sequenced along with their upstream and downstream exons, possible detection of deep intronic variants is disregarded. Thus, no deep intronic variant has been reported in association with C1-INH-HAE. To overcome some of the shortcomings of traditional approaches we recently developed a custom next-generation sequencing (NGS) platform that allows analyzing *SERPING1* in its full length.[8] Here, we use this platform for genotyping C1-INH-HAE patients with no detectable mutations in the coding region of *SERPING1*, and we describe our results and their relevance.

Methods

Study population

Fourteen patients (3 Greek, 2 Hungarian, 4 Polish, 3 Bulgarian, 2 German; 5 male, mean age 43 ± 17 years) diagnosed with type I C1-INH-HAE, according to the criteria of the Hereditary Angioedema International Working Group,[9] but without *SERPING1* mutations after conventional molecular analysis, were initially enrolled in the study. Pedigree analysis became feasible in one case. Three healthy first degree relatives of the patient and three first degree relatives with C1-INH-HAE were examined. Informed consent was obtained from all patients and family members investigated, and the study has been approved by the Ethics Committee of the Faculty of Medicine, University of Thessaly.

Genetic analysis

Genomic DNA was extracted from peripheral blood with iPrep PureLink DNA blood kit (Invitrogen, Carlsbad, CA, USA) according to manufacturer's instructions. Conventional genotyping had been performed by sequencing all *SERPING1* translated regions and intron-exon boundaries, long-range PCR and multiplex ligation-dependent probe amplification (MLPA), as previously described.[4]

A custom NGS panel was designed by the Ion AmpliSeq Thermo Fisher Scientific Designer, as previously described,[8] covering the whole *SERPING1* gene (NM_000062.2) including its complete 5′ and 3′ untranslated regions, exonic and intronic regions (Chr11: 57,364,831–57,382,476; GRCh37). Coverage of 99.9% and 79.9% of exonic and intronic regions of *SERPING1* was achieved, respectively. Missing intronic regions were mainly those of introns 3, 4, and 6 (41,4%, 42,2% and 22,5%, respectively). Briefly, amplicon libraries were prepared using the Ion AmpliSeq Library Kit 2.0 (Thermo Fisher Scientific, Waltham, MA, USA) and Ion Xpress™ Barcode Adapter 1–96 Kit (Thermo Fisher Scientific). Pooled, barcoded libraries were clonally amplified using the Ion OneTouch™ system and Ion OT2 HI-Q Template kit (Thermo Fisher Scientific). Ion sphere particles (ISPs) were enriched with the Dynabeads® MyOne™ Streptavidin C1 Beads (Thermo Fisher Scientific) and washed with the Ion OneTouch Wash Solution included in the kit using the Ion One Touch ES system (Thermo Fisher Scientific). NGS was performed on the Personal Genome Machine (PGM) using the Ion PGM Hi-Q sequencing kit (Thermo Fisher Scientific) resulting in a mean depth of coverage 20x. All procedures were performed according to the manufacturer's instructions.

Torrent Suite 5.2 software was used for the analysis of NGS data. Raw data were aligned to the complete *SERPING1* gene of the human reference sequence hg19 (GRCh37), variant calling was performed by the VariantCaller v.5.2 plug-in and coverage analysis by the coverageAnalysis v.5.2.1.2 plug-in. Annotation of variants was performed on Ion Reporter software v.5.2 (Thermo Fisher Scientific). The annotated variants were evaluated and visualized via integrative genome viewer (IGV).

Variants that were present in the reference population data sets [Genome Aggregation Database (gnomAD), Exome Aggregation Consortium[10] release 0.3 (ExAC Browser), Database of Single Nucleotide Polymorphisms[11] build 141 GRCh37.p13 (dbSNP)] at a global minor allele frequency greater than that the frequency C1-INH-HAE (<0.002%) were filtered out from the analysis.

Sanger sequencing was performed for the confirmation of novel variants and family segregation study. The intron 1 of the *SERPING1* gene was amplified by using the primers 5′-CTGCACC-CAAGCTTCCCCGTTCAC-3′ and 5′-CCCCGTCCCCCATCCCACAAG-3'. The PCR was initiated at 94 °C for 2 min, followed by denaturation at 94 °C for 45 s, annealing at 65 °C for 45 s, and extension at 72 °C for 1 min for 30 cycles. The amplified products were sequenced using the BigDye Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems, Foster City, CA, USA) and analyzed in ABI 3730 DNA Analyzer (Applied Biosystems, Foster City, USA). Data were evaluated by the Sequencing Analyses software 5.2 (Applied Biosystems, Foster City, California, USA).[12]

In silico analyses

The putative effect of filtered-in intronic mutations was evaluated by using ten different bioinformatics analysis tools. Sequence segments with and without corresponding changes in their sequences were submitted to five bioinformatics tools examining potential splice effects: NNSPLICE,[13] Netgene2,[14] ASSP,[15] FSPLICE and HSF.[16] All of them provide probability scores for the use of potential donor and acceptor splice sites. Further assessment of the pathogenicity was attempted by the use of 5 additional tools. The CADD provides prediction of the functional impact of the observed variants by integrating different functional aspects and diverse annotations into a single outcome, namely the C-score[17]; the Transcript-inferred Pathogenicity score (Trap score) evaluates a single nucleotide variant's ability to cause disease by damaging the final transcript[18]; DANN is a pathogenicity scoring methodology based on deep neural networks[19]; GERP is a conservation score calculated by quantifying substitution deficits across multiple alignments of orthologues using the genomes of 35 mammals[20]; finally, SpliceAid 2 bioinformatics foresee the splicing pattern alteration and guide the identification of the molecular effect due to the mutations.[21]

Transcriptional analysis

In order to further evaluate the function effect of the variant c.-22-155G > T, whole blood was collected from the proband and his three suffering family members carrying the mutation and being heterozygous for the rs4926 (c.1438G > A) polymorphism. Mononuclear cells were separated by a Ficoll gradient, and total RNA was extracted with phenol-chloroform, ethanol precipitated, and purified by guanidinium thiocyanate dissociation and isopropanol

precipitation.[22,23] First-Strand cDNA was synthesized from 1 μg RNA in 25 μl reactions with 25 μM Primer random p[dN]6 (Roche), 10 mM dNTPs (Invitrogen, Carlsbad, CA, USA) and 200 U/μL M-MLV Reverse Transcriptase (Invitrogen, Carlsbad, CA, USA) employed in 72 °C for 2 min, 42 °C for 75 min and 65 °C for 10 min.

In the produced cDNA, a fragment from exons 6–8 was amplified, sequenced and compared to the gDNA sequence. The forward 5′AACTCAGTTATAAAAGTGCCCATGATGAAT3′[24] and the reverse 5′CCCTTTTGGTGGATAGCG-3′[25] primers were used. The amplification of the cDNA was performed using 10 pmol of each primer and 35 cycles in a Veriti® Thermal Cycler (Applied Biosystems, Foster City, USA), comprising of 1 min at 94 °C, 1 min at 55 °C, and 1 min at 72 °C, followed by a final extension of 10 min at 72 °C. The amplified products were sequenced and analyzed, as described above.

Variant curation

The interpretation of sequence variants was based on the criteria established by the ACMG and the AMP.[26] Data from our previous study[8] in regard with the presence of variants in healthy individuals were used to this aim.

Results

Apart from the known polymorphisms rs28362944 and rs4926, NGS genotyping of our C1-INH-HAE patients confirmed the absence of disease-causing *SERPING1* exonic mutations. However, 35 different intronic mutations were revealed in these patients, 29 of which were filtered out as their global minor allele frequency was greater than

Table 1
Filtered out intronic mutations presenting in the Genome Aggregation Database (gnomAD) with a global minor allele frequency greater than 0.002%.

Locus	Genotype	Coding	dbSNP	Intron	Frequency (%)
chr11:57365895	G/A	c.51 + 101G > A	rs28362945	2	12.56
chr11:57366405	C/CGT	c.51 + 625_51 + 626dupTG	rs3054018	2	38.99
chr11:57366656	C/T	c.52−696C > T	rs1005511	2	39.51
chr11:57367222	C/T	c.52−130C > T	rs1005510	2	38.72
chr11:57369008	C/G	c.551-500C > G	rs28362947	3	22.03
chr11:57369013	A/C	c.551-495A > C	rs28362948	3	22.03
chr11:57369353	A/G	c.551-155A > G	rs2936694	3	36.84
chr11:57369730	G/A	c.685 + 88G > A	rs11229063	4	22.24
chr11:57370742	C/T	c.685 + 1100C > T	rs78364821	4	22.52
chr11:57371911	G/T	c.686-1572G > T	rs28362949	4	22.08
chr11:57371918	G/A	c.686-1565G > A	rs28362950	4	22.42
chr11:57372526	A/G	c.686−957A > G	rs28362951	4	22.19
chr11:57373304	A/G	c.686-179A > G	rs189335964	4	0.1068
chr11:57374280	G/A	c.1029 + 260G > A	rs191053716	6	0.1356
chr11:57374332	T/C	c.1029 + 312T > C	rs11603020	6	22.79
chr11:57374871	C/G	c.1029 + 851C > G	rs11229066	6	22.79
chr11:57374946	G/T	c.1029 + 926G > T	rs11229067	6	22.79
chr11:57375463	G/C	c.1029 + 1443G > C	rs78624400	6	22.82
chr11:57375517	A/G	c.1029 + 1497A > G	rs17661117	6	0.6972
chr11:57376130	T/C	c.1029 + 2110T > C	rs2454659	6	61.20
chr11:57376131	A/G	c.1029 + 2111G > A	rs138770460	6	22.7
chr11:57377215	G/C	c.1030-1975G > C	rs151035150	6	1.1044
chr11:57377676	CT/C	c.1030-1513delT	rs1184255008	6	0.003227
chr11:57377968	A/G	c.1030−1222A > G	rs3824988	6	22.78
chr11:57377992	G/T	c.1030-1198G > T	rs2508443	6	32.92
chr11:57378325	C/T	c.1030-865C > T	rs2511989	6	38.14
chr11:57379170	A/G	c.1030-20A > G	rs2511988	6	61.09
chr11:57381263	T/C	c.1250-538T > C	rs10896631	7	22.76
chr11:57381519	T/C	c.1250-282T > C	rs1557522	7	30.64

Table 2
Bioinformatic analysis of the six filtered-in intronic mutations.

Variant	NNSPLICE[†]	Netgene2[‡]	ASSP[§]	FSPLICE[¶]	HSF[‖]	CADD[#]	Trap score[††]	DANN[‡‡]	GERP RS[§§]
c.-22-155G > T	D:0.96	D:0.79	D:0.46	D:10.16	New intronic cryptic donor site	21.2	0.686	0.9507	1
c.551-156A > G	NI	NI	NI	NI	Creation of an ISE	1.427	0.103	0.3543	NP
c.686−1488_686-1487insT	NI	NI	NI	NI	Creation of an ISE	NP	NP	NP	0.0435
c.686−1335T > A	NI	NI	NI	NI	Alteration of an ISS/Creation of an ISE	0.131	0.078	0.1134	−1.335
c.686−1333A > T	NI	NI	NI	NI	NI	0.969	0.012	0.1669	−1.875
c.1250-154C > G[¶¶]	NI	NI	NI	NI	Creation of an ISE	3.911	0.122	0.5833	0.3333

NI, No impact; NP, Not provided; D, new donor; ISS, Intronic splicing silencer; ISE, Intronic splicing enhancer.
[†] Range 0–1, with minimum score for 5′ and 3′ splice site 0.4 (www.fruitfly.org/seq_tools/splice.html).
[‡] Confidence range 0.5–0.95 (www.cbs.dtu.dk/services/NetGene2/).
[§] Confidence range 0–1, false splice site cutoff for acceptor sites: 2.2, false splice site cutoff for donor sites: 4.5 (Alternative Splice Site Predictor: wangcomputing.com/assp/).
[¶] Acceptor site threshold: 4.175, donor site threshold: 6.099 (www.softberry.com/berry.phtml?topic=fsplice&group=programs&subgroup=gfind).
[‖] www.umd.be/HSF/HSF.shtml.
[#] Range 10–30 (cadd.gs.washington.edu).
[††] Range 0–1 (http://trap-score.org/Search?version=v2).
[‡‡] Range 0–1.
[§§] −12.3 to 6.17.
[¶¶] Observed in 2 patients (trans with the c.-22-155G > T variant) and one healthy member of the studied family.

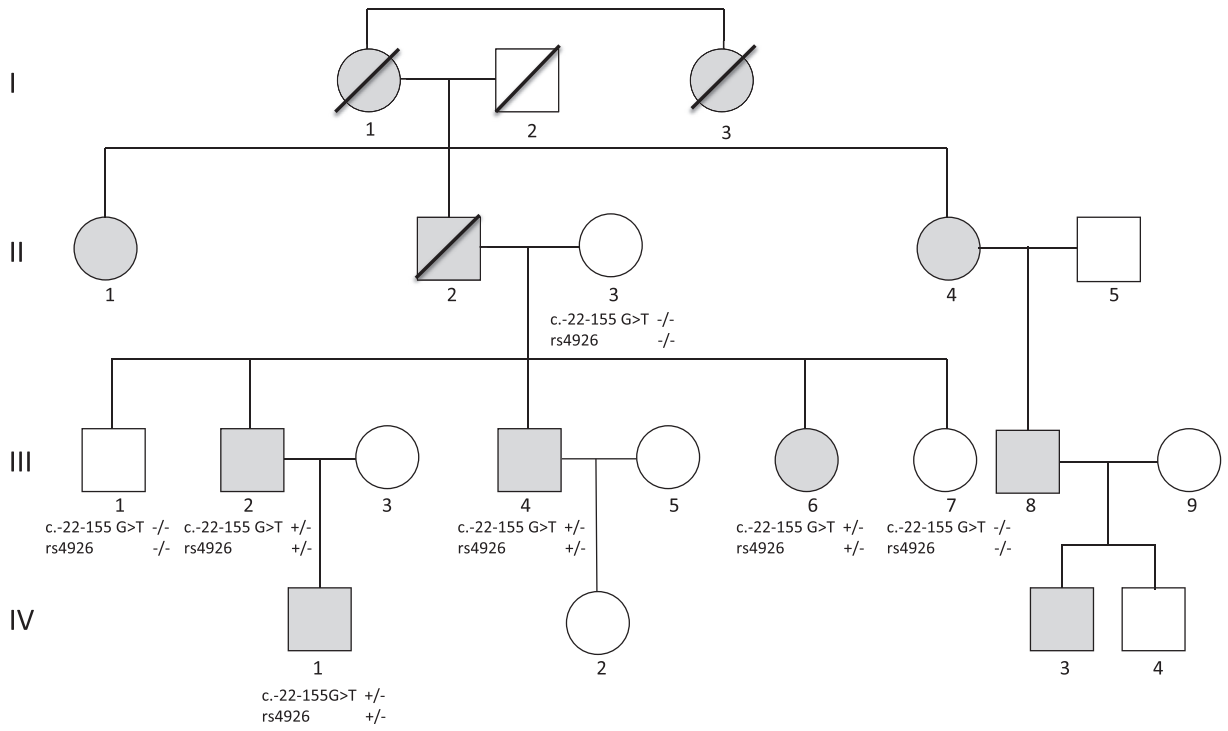

Fig. 1. Pedigree analysis of a C1-INH-HAE family carrying the deep intronic c.-22-155G > T variant. All examined suffering members of the family (III.2, III.4, III.6, IV.1) —carriers of the c.-22-155G > T— were heterozygous for the rs4926 (c.1438G > A) polymorphism, while all examined unaffected members (II.3, III.1, III.7) were homozygous for the wild type of this polymorphism.

0.002% (Table 1). Four out of the remaining 6 variants were unreported (novel), and 2 were reported once in global databases.

Only one of the six filtered-in variants, the novel variant c.-22-155G > T located in intron 1 of the *SERPING1* gene (chr11:57,365,567), was predicted as pathogenic by all used bioinformatics tools (Table 2). Moreover, this variant was also the only one that was detected in two unrelated C1-INH-HAE Greek patients, i.e. with a frequency 14.3% among the cohort of our C1-INH-HAE patients without any mutation in the exonic region of the gene, while it was also undetectable in our healthy control group.[8] Assuming a 5% frequency of C1-INH-HAE patients with no causal mutation identified in the coding exons and exon-intron boundaries of the *SERPING1* gene,[4,5] the estimated frequency of this variant among C1-INH-HAE patients is 0.7%. Thus, the c.-22-155G > T variant was the only one that was further studied.

Pedigree analysis revealed that the c.-22-155G > T variant co-segregated with C1-INH-HAE in all of the 4 analyzed patients belonging to two generations, while it was absent from all of the 3 healthy family members who were also analyzed (Fig. 1). Examined patients were of 8, 40, 44 and 45 (the proband, III.2) years-of-age with an age at disease onset of 8, 7, 24 and 13 years, and a mean frequency of attacks (mainly cutaneous) 2, 5, 14 and 5 per year, respectively. Their antigenic C1-INH concentration at diagnosis was varying from 4% to 40% of the reference. Information about the other suffering members of the family was unobtainable since they are immigrants in Australia. The fifth, unrelated patient examined was a man of 45 years-of-age without family history of angioedema presenting with rare cutaneous attacks since the age of 25 years. No other member of his family was available for analysis.

Further bioinformatic analysis with the use of SpliceAid 2 tool showed that the c.-22-155G > T variant alters the transcriptional motif recognized by hnRNP H1, hnRNP H2, hnRNP F and hnRNP H3 transcriptional factors preventing their binding in the corresponding gene region.

Transcript analysis supported the above indications of the deleterious effect of the c.-22-155G > T variant. By this approach loss of heterozygosity was demonstrated for the exon 8 rs4926 (c.1438G > A) polymorphism located in the same allele with the c.-22-155G > T variant. More specifically, as it is shown in Figure 1, in the gDNA, all

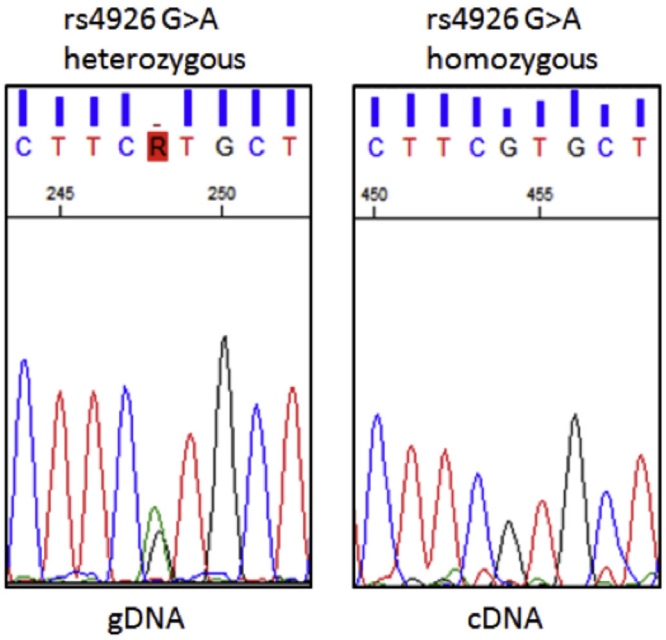

Fig. 2. cDNA and gDNA sequencing of one of the patients (III.6) carrying the c.-22-155 G > T variant and the rs4926 G > A polymorphism. Sequencing of the amplified cDNA demonstrate that while in the gDNA the two alleles (G/A) of the rs4926 are equally represented, in the cDNA only one allele (G) is present.

Table 3
The 49 reported intronic mutations in *SERPING1* gene associated with C1-INH-HAE. The majority of them are located in the donor and acceptor site or a few (up to 5) nucleotides from these regions. Only the c.1029 + 84G > A mutation is located 84 nucleotides from the acceptor site.

Mutation	Intron	Chromosome position	dbSNP	References
c.-22-2A > C	1	chr11:57365720		Aabom *et al.*, 2017[33]
c.-22-2A > G	1	chr11:57365720		Bygum *et al.*, 2011[34]; Xiong *et al.*, 2015[6]
c.-22-1G > A	1	chr11:57365721		Verpy *et al.*, 1996[35]; Gösswein *et al.*, 2008[36]; Xiong *et al.*, 2015[6]; Duponchel *et al.*, 2006[37]; Andrejević *et al.*, 2015[38]
c.51+6T > G	2	chr11:57365800		López-Lera *et al.*, 2011[39]
c.51+5G > A	2	chr11:57365799		Verpy *et al.*, 1996[35]; Duponchel *et al.*, 2006[37]; Xiong *et al.*, 2015[6]
c.51 + 1G > T	2	chr11:57365795		Yamamoto *et al.*, 2012[40]; Cagini *et al.*, 2016[41]
c.51 + 1G > A	2	chr11:57365795	rs1470120365	Gösswein *et al.*, 2008[36]; Pappalardo *et al.*, 2008[4]; Kalmár *et al.*, 2003[42]; Xiong *et al.*, 2015[6]
c.51 + 2T > C	2	chr11:57365796		Cagini *et al.*, 2016[41]
c.51 + 3A > G	2	chr11:57365797		Roche *et al.*, 2005[43]; Pappalardo *et al.*, 2008[4]; Duponchel *et al.*, 2006[37]
c.52−2A > G	2	chr11:57367350		Pappalardo *et al.*, 2008[4]; Xiong *et al.*, 2015[6]
c.52−1 G > A	2	chr11:57367351	rs886041353	Gösswein *et al.*, 2008[36]; Xiong *et al.*, 2015[6]
c.550+1G > A	3	chr11:57367851		Kalmár *et al.*, 2003[42]; Xiong *et al.*, 2015[6]
c.550+2T > C	3	chr11:57367852		Roche *et al.*, 2005;[43]
c.550 + 5G > C	3	chr11:57367855		Roche *et al.*, 2005[43]; Xiong *et al.*, 2015[6]
c.550 + 5G > A	3	chr11:57367855	rs1314284778	Roche *et al.*, 2005[43]; Xiong *et al.*, 2015[6]
c.551-2delA	3	chr11:57369506		Roche *et al.*, 2005;[43]
c.551-1G > A	3	chr11:57369507		Speletas *et al.*, 2015[5]
c.551−2A > G	3	chr11:57369506		Gösswein *et al.*, 2008[36]; Grodecká *et al.*, 2017[44]
c.551−3 C > G	3	chr11:57369505		Gösswein *et al.*, 2008[36]
c.551−5 T > A	3	chr11:57369503		Pappalardo *et al.*, 2008[4]
c.685 + 31G > A	4	chr11: 57369673	rs751335805	Suffritti *et al.*, 2014[45]
c.685 + 1G > T	4	chr11:57369643		Pappalardo *et al.*, 2008[4]; Xiong *et al.*, 2015[6]
c.685 + 1 G > A	4	chr11:57369643	rs113263597	Gösswein *et al.*, 2008[36]; Xiong *et al.*, 2015[6]
c.685 + 2T > A	4	chr11:57369644		Colobran *et al.*, 2014[46]
c.685 + 2T > G	4	chr11:57369644		Pappalardo *et al.*, 2008[4]; Xiong *et al.*, 2015[6]
c.686−12 A > G	4	chr11: 57373471		Pappalardo *et al.*, 2008[4]; Grodecká *et al.*, 2017[44]; Andrejević *et al.*, 2015[38]
c.686-3C > G	4	chr11:57373480		Roche *et al.*, 2005[43]; Xiong *et al.*, 2015[6]
c.889 + 1 G > T	5	chr11:57373687		Gösswein *et al.*, 2008[36]; Xiong *et al.*, 2015[6]
c.889 + 2T > C	5	chr11:57373688		Roche *et al.*, 2005[43]; Xiong *et al.*, 2015[6]
c.889 + 3A > T	5	chr11:57373689		Johnsrud *et al.*, 2015[47]
c.890-14C > G	5	chr11: 57373867		Speletas *et al.*, 2015[5]
c.890-2A > G	5	chr11: 57373879		Bowen *et al.*, 2001[12]; Xiong *et al.*, 2015[6]
c.890-1G > A	5	chr11: 57373880		Sekijima *et al.*, 2004[48]; Xiong *et al.*, 2015[6]
c.1029 + 1 G > T	6	chr11:57374021		Siddique *et al.*, 1991[49]; Xiong *et al.*, 2015[6]
c.1029 + 1 G > A	6	chr11:57374021		Gösswein *et al.*, 2008[36]; Xiong *et al.*, 2015[6]
c.1029 + 84G > A	6	chr11:57374104	rs118132731	Pappalardo *et al.*, 2008[4]
c.1029 + 3_ c.1029 + 6 del4bp	6	chr11:57374023		Gösswein *et al.*, 2008[36]
c.1029 + 4 delA	6	chr11:57374024		Bygum *et al.*, 2011[34]
c.1030−1 G > A	6	chr11:57379189		Speletas *et al.*, 2015[5]
c.1030−1 G > C	6	chr11:57379189		Gösswein *et al.*, 2008[36]; Verpy *et al.*, 1996[35]; Xiong *et al.*, 2015[6]
c.1249 + 1G > C	7	Chr11: 57379410		López-Lera *et al.*, 2011[39]
c.1249 + 1G > A	7	Chr11: 57379410	rs112565881	Gösswein *et al.*, 2008[36]; Xiong *et al.*, 2015[6]
c.1249 + 2T > A	7	Chr11: 57379411		Kawachi *et al.*, 1998[50]; Xiong *et al.*, 2015[6]
c.1249 + 2delT	7	Chr11: 57379411		Roche *et al.*, 2005;[43]
c.1249 + 5G > T	7	Chr11: 57379414		Pappalardo *et al.*, 2008[4]; Grodecká *et al.*, 2017[44]
c.1249 + 5G > A	7	Chr11: 57379414		Colobran *et al.*, 2017[25]; Grodecká *et al.*, 2017[44]
c.1250−13 G > A	7	chr11: 57381788		Gösswein *et al.*, 2008[36]; Kesim *et al.*, 2011[51]
c.1250-1G > A	7	chr11: 57381800		Aabom *et al.*, 2017[33]
c.1250−2 A > G	7	chr11: 57381799		Gösswein *et al.*, 2008[36]

patients of the examined family −carriers of the c.-22-155G > T− were heterozygous for the rs4926 (c.1438G > A) polymorphism, while all unaffected members were homozygous for the wild type. Interestingly, in the cDNA, all patients were found homozygous for the wild type −only one allele (G) is present− indicating that the mRNA from the mutated allele is probably degraded (Fig. 2).

According to the ACMG-AMP 2015 guidelines[26] the pathogenicity potential of the c.-22-155G > T variant is “pathogenic” based on: (a) The strong segregation data. The variant co-segregates with disease in multiple affected family members in a gene definitively known to cause the disease (PP1 criterion used as strong evidence). (b) The absence of the variant from controls in Exome Sequencing Project, 1000 Genomes Project, or Exome Aggregation Consortium (PM2). (c) The observation of the variant in two unrelated probands with the same phenotype, and its absence in controls (PS4). (d) Prediction by multiple bioinformatics tools that the variant causes a deleterious effect on the gene (PP3). (e) The patient's phenotype and family history is highly specific for the disease (PP4). (f) *In vitro* functional studies support a damaging effect on the gene (PS3). Based on the available criteria (BS2, BS4, BP4, BP5) the remaining intronic variants detected in our C1-INH-HAE patients were classified as “benign” or “likely benign”.

Discussion

To identify missing noncoding variants in C1-INH-HAE cases, we sequenced the *SERPING1* locus in 14 patients. In two of them, we detected a novel deep intronic variant (c.-22-155G > T) which was classified as pathogenic, according to ACMG-AMP 2015 guidelines.[26] In recent years, increasing numbers of deep intronic variants located at least 100 bp from the nearest canonical splice site, have been reported and their role in human diseases has been

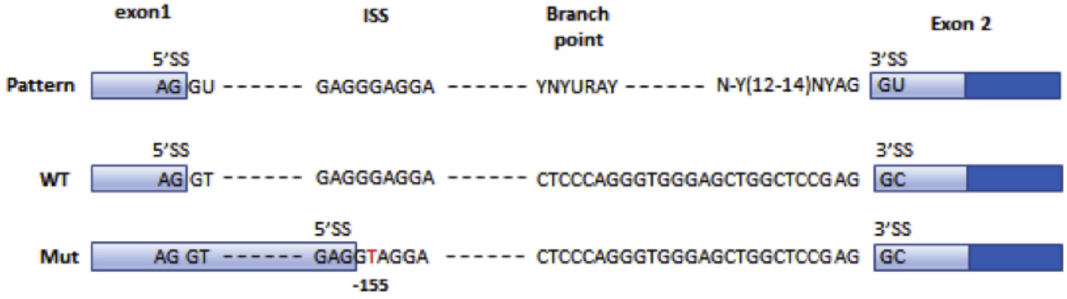

Fig. 3. Graphical presentation of the impact of the c.-22-155G > T variant according to bioinformatics analysis. The variant is predicted to form an alternative donor site leading to a modified mutant RNA longer (>372 bp) than the wild type.

largely demonstrated. To the best of our knowledge, this is the first deep intronic variant of *SERPING1* gene associated with C1-INH-HAE that has been reported in the literature. Heretofore, 49 different intronic mutations of *SERPING1* gene have been associated with C1-INH-HAE, all of them located in the donor and acceptor site or a few nucleotides from these regions[4–6,12,25,33–51] (Table 3).

Deleterious deep intronic variants most commonly lead to pseudo-exon inclusion due to activation of non-canonical splice sites or changes in splicing regulatory elements, or they can disrupt transcription regulatory motifs and non-coding RNA genes.[7] As far as the c.-22-155G > T variant is regarded, the most probable scenario indicated by the findings of the *in silico* and the transcriptional analysis, is that it affects splicing and alters the transcribed RNA. The mutant mRNA is susceptible to degradation by mRNA surveillance pathways. Consequently, only the wild type allele is translated, and this results in C1-INH deficiency.

In more detail, by the use of the NNSPLICE, NetGene2 and FSPLICE, ASSP bioinformatics tools, it has been shown that the c.-22-155G > T variant introduces in the genome a new donor site stronger than the wild type leading to a larger exon 1 in the RNA level (Fig. 3). This was confirmed by HSF indicating, in parallel, that the variant disrupts an ISS recognized by a number of transcriptional factors and, as a result, stops the suppression of the intronic cryptic donor site. Additionally, by analyzing both the wild type and the mutated sequence, SpliceAid 2 concluded that different splicing motifs are destroyed, including the splicing motifs for the transcriptional factors hnRNP H1, hnRNP H2, hnRNP F, hnRNP H3 which act as silencers in the wild type sequence. Transcriptional analysis confirmed that the mutant mRNA is susceptible to degradation. To this aim, an informative exonic SNP, i.e. the polymorphism rs4926 (c.1438G > A) carried by the patients in heterozygous state, was used. On the RNA level, this polymorphism was found in homozygous state, which indicates that the mutated allele is degraded, obviously through mRNA surveillance pathways. Further studies are required in order to define which of the three translation-associated mRNA surveillance pathways (nonsense-mediated mRNA decay, no-go decay, nonstop decay) that target mRNAs for degradation is involved.[7]

In favor of the pathogenicity of the c.-22-155G > T variant is its location in a chromosomal region of intron 1 which, by the use of GERP, a statistically rigorous and biologically transparent framework for constrained element identification,[27] was found highly conserved. Generally, it is considered that the first intron is highly conserved and that its conservation is related to its enrichment with regulatory elements[28] and a specific pattern of chromatin organization.[29] However, the first intron of *SERPING1* is not the longest among all other downstream introns within the gene as it happens in most species.[30]

In conclusion, our study verified the earlier hypothesis that intronic alterations could be the cause of the disease in cases of C1-INH-HAE where standard genotyping approaches cannot uncover any DNA damage, highlighting one more advantage of NGS[8] in the molecular analysis of these patients.[31] Therefore, advanced DNA sequencing methods should be performed in cases of C1-INH-HAE where standard approaches fail to uncover the genetic alteration. Finally, despite that one fifth of *SERPING1* intronic length is escaping the analysis by our NGS panel, the possibility of intronic alterations to be the invisible damage whenever standard genotyping cannot detect the cause of C1-INH-HAE, is minimized. The regulatory mechanisms of gene expression comprise diverse molecular circuits involving multiple dedicated components. Thus mechanisms intervening to the expression of *SERPING1,* other than alterations in non-coding regions, should be considered. As Dirk A. Kleinjan and Veronica van Heyningen have noted[32]: "… the laborious identification of the disease loci and regulatory mechanisms involved in currently 'unsolved' human disorders remains a huge but rewarding task".

Conflict of interest

MZ, GL and FPa are employed by CeMIA SA. AEG is stock owner of CeMIA SA. The rest of the authors have no conflict of interest.

Authors' contributions

SV participated in the study design, carried out the basic experimental analysis and data interpretation and drafted the manuscript. MZ and GL conceived and participated in the design of the study, carried out the data interpretation and performed critical revision of the manuscript. FPa involved in some of the basic experimental analyses. FPs, DC, AV, MS, GP, KO, MMag, MMau, MS and HF were responsible for collection of blood samples and assembly of clinical data. AEG conceived and coordinated the study, participated in the study design, and critically revised of the article for important intellectual content. All authors have read and approved the final version of the manuscript.

References

1. Longhurst H, Cicardi M. Hereditary angio-oedema. *Lancet* 2012;**379**:474–81.
2. Germenis AE, Speletas M. Genetics of hereditary angioedema revisited. *Clin Rev Allergy Immunol* 2016;**51**:170–82.
3. Ponard D, Gaboriaud C, Charignon D, Ghannam A, Wagenaar-Bos IGA, Roem D, et al. *SERPING1* mutation update: mutation spectrum and C1 Inhibitor phenotypes. *Hum Mutat* 2020;**41**:38–57.
4. Pappalardo E, Caccia S, Suffritti C, Tordai A, Zingale LC, Cicardi M. Mutation screening of C1 inhibitor gene in 108 unrelated families with hereditary angioedema: functional and structural correlates. *Mol Immunol* 2008;**45**:3536–44.
5. Speletas M, Szilagyi A, Psarros F, Moldovan D, Magerl M, Kompoti M, et al. Hereditary angioedema: molecular and clinical differences among European populations. *J Allergy Clin Immunol* 2015;**135**:570–3.
6. Xiong HY, Alipanahi B, Lee LJ, Bretschneider H, Merico D, Yuen RK, et al. RNA splicing. The human splicing code reveals new insights into the genetic determinants of disease. *Science* 2015;**347**:1254806.
7. Vaz-Drago R, Custódio N, Carmo-Fonseca M. Deep intronic mutations and human disease. *Hum Genet* 2017;**136**:1093–111.
8. Loules G, Zamanakou M, Parsopoulou F, Vatsiou S, Psarros F, Csuka D, et al. Targeted next-generation sequencing for the molecular diagnosis of hereditary angioedema due to C1-inhibitor deficiency. *Gene* 2018;**667**:76–82.

9. Cicardi M, Aberer W, Banerji A, Bas M, Bernstein JA, Bork K, et al. Classification, diagnosis, and approach to treatment for angioedema: consensus report from the Hereditary Angioedema International Working Group. *Allergy* 2014;**69**:602–16.
10. Lek M, Karczewski KJ, Minikel EV, Samocha KE, Banks E, Fennell T, et al. Analysis of protein-coding genetic variation in 60,706 humans. *Nature* 2016;**536**:285–91.
11. Sherry ST, Ward M-H, Kholodov M, Baker J, Phan L, Smigielski EM, et al. dbSNP: the NCBI database of genetic variation. *Nucleic Acids Res* 2001;**29**:308–11.
12. Bowen B, Hawk JJ, Sibunka S, Hovick S, Weiler JM. A review of the reported defects in the human C1 esterase inhibitor gene producing hereditary angioedema including four new mutations. *Clin Immunol* 2001;**98**:157–63.
13. Reese MG, Eeckman FH, Kulp D, Haussler D. Improved splice site detection in Genie. *J Comput Biol* 1997;**4**:311–23.
14. Brunak S, Engelbrecht J, Knudsen S. Prediction of human mRNA donor and acceptor sites from the DNA sequence. *J Mol Biol* 1991;**220**:49–65.
15. Wang M, Marin A. Characterization and prediction of alternative splice sites. *Gene* 2006;**366**:219–27.
16. Desmet FO, Hamroun D, Lalande M, Collod-Béroud G, Claustres M, Béroud C. Human Splicing Finder: an online bioinformatics tool to predict splicing signals. *Nucleic Acids Res* 2009;**37**:e67.
17. Kircher M, Witten DM, Jain P, O'Roak BJ, Cooper GM, Shendure J. A general framework for estimating the relative pathogenicity of human genetic variants. *Nat Genet* 2014;**46**:310–5.
18. Gelfman S, Wang Q, McSweeney KM, Ren Z, La Carpia F, Halvorsen M, et al. Annotating pathogenic non-coding variants in genic regions. *Nat Commun* 2017;**8**:236.
19. Quang D, Chen Y, Xie X. DANN: a deep learning approach for annotating the pathogenicity of genetic variants. *Bioinformatics* 2015;**31**:761–3.
20. Cooper GM, Stone EA, Asimenos G, NISC Comparative Sequencing Program, Green ED, Batzoglou S, et al. Distribution and intensity of constraint in mammalian genomic sequence. *Genome Res* 2005;**15**:901–13.
21. Piva F, Giulietti M, Burini AB, Principato G. SpliceAid 2: a database of human-splicing factors expression data and RNA target motifs. *Hum Mutat* 2012;**33**:81–5.
22. Chomczynski P, Sacchi N. The single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction: twenty-something years on. *Nat Protoc* 2006;**1**:581–5.
23. Peirson SN, Butler JN. RNA extraction from mammalian tissues. *Methods Mol Biol* 2007;**362**:315–27.
24. Ariga T, Igarashi T, Ramesh N, Parad R, Cicardi M, Davis 3rd AE. Type I C1 inhibitor deficiency with a small messenger RNA resulting from deletion of one exon. *J Clin Invest* 1989;**83**:1888–93.
25. Colobran R, Pujol-Borrell R, Hernández-González M, Guilarte M. A novel splice site mutation in the SERPING1 gene leads to haploinsufficiency by complete degradation of the mutant allele mRNA in a case of familial hereditary angioedema. *J Clin Immunol* 2014;**34**:521–3.
26. Richards S, Aziz N, Bale S, Bick D, Das S, Gastier-Foster J, 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.
27. Davydov EV, Goode DL, Sirota M, Cooper GM, Sidow A, Batzoglou S. Identifying a high fraction of the human genome to be under selective constraint using GERP++. *PLoS Comput Biol* 2010;**6**:e1001025.
28. Gaffney DJ, Keightley PD. Unexpected conserved non-coding DNA blocks in mammals. *Trends Genet* 2004;**20**:332–7.
29. Bieberstein NI, Carrillo Oesterreich F, Straube K, Neugebauer KM. First exon length controls active chromatin signatures and transcription. *Cell Rep* 2012;**2**:62–8.
30. Bradnam KR, Korf I. Longer first introns are a general property of eukaryotic gene structure. *PLoS One* 2008;**3**:e3093.
31. Marcelino-Rodriguez I, Callero A, Mendoza-Alvarez A, Perez-Rodriguez E, Barrios-Recio J, Garcia-Robaina JC, et al. Bradykinin-mediated angioedema: an update of the genetic causes and the impact of genomics. *Front Genet* 2019;**10**:900.
32. Kleinjan DA, van Heyningen V. Long-range control of gene expression: emerging mechanisms and disruption in disease. *Am J Hum Genet* 2005;**76**:8–32.
33. Aabom A, Andersen KE, Fagerberg C, Fisker N, Jakobsen MA, Bygum A. Clinical characteristics and real-life diagnostic approaches in all Danish children with hereditary angioedema. *Orphanet J Rare Dis* 2017;**12**:55.
34. Bygum A, Fagerberg CR, Ponard D, Monnier N, Lunardi J, Drouet C. Mutational spectrum and phenotypes in Danish families with hereditary angioedema because of C1 inhibitor deficiency. *Allergy* 2011;**66**:76–84.
35. Verpy E, Biasotto M, Brai M, Misiano G, Meo T, Tosi M. Exhaustive mutation scanning by fluorescence-assisted mismatch analysis discloses new genotype-phenotype correlations in angiodema. *Am J Hum Genet* 1996;**59**:308–19.
36. Gösswein T, Kocot A, Emmert G, Kreuz W, Martinez-Saguer I, Aygören-Pürsün E, et al. Mutational spectrum of the C1INH (SERPING1) gene in patients with hereditary angioedema. *Cytogenet Genome Res* 2008;**121**:181–8.
37. Duponchel C, Djenouhat K, Frémeaux-Bacchi V, Monnier N, Drouet C, Tosi M. Functional analysis of splicing mutations and of an exon 2 polymorphic variant of SERPING1/C1NH. *Hum Mutat* 2006;**27**:295–6.
38. Andrejević S, Korošec P, Šilar M, Košnik M, Mijanović R, Bonači-Nikolić B, et al. Hereditary angioedema due to C1 inhibitor deficiency in Serbia: two novel mutations and evidence of genotype-phenotype association. *PLoS One* 2015;**10**:e0142174.
39. López-Lera A, Garrido S, Roche O, López-Trascasa M. SERPING1 mutations in 59 families with hereditary angioedema. *Mol Immunol* 2011;**49**:18–27.
40. Yamamoto T, Horiuchi T, Miyahara H, Yoshizawa S, Maehara J, Shono E, et al. Hereditary angioedema in Japan: genetic analysis of 13 unrelated cases. *Am J Med Sci* 2012;**343**:210–4.
41. Cagini N, Veronez CL, Constantino-Silva RN, Buzolin M, Martin RP, Grumach AS, et al. New mutations in SERPING1 gene of Brazilian patients with hereditary angioedema. *Biol Chem* 2016;**397**:337–44.
42. Kalmár L, Bors A, Farkas H, Vas S, Fandl B, Varga L, et al. Mutation screening of the C1 inhibitor gene among Hungarian patients with hereditary angioedema. *Hum Mutat* 2003;**22**:498.
43. Roche O, Blanch A, Duponchel C, Fontán G, Tosi M, López-Trascasa M. Hereditary angioedema: the mutation spectrum of SERPING1/C1NH in a large Spanish cohort. *Hum Mutat* 2005;**26**:135–44.
44. Grodecká L, Hujová P, Kramárek M, Kršjaková T, Kováčová T, Vondrášková K, et al. Systematic analysis of splicing defects in selected primary immunodeficiencies-related genes. *Clin Immunol* 2017;**180**:33–44.
45. Suffritti C, Zanichelli A, Maggioni L, Bonanni E, Cugno M, Cicardi M. High-molecular-weight kininogen cleavage correlates with disease states in the bradykinin-mediated angioedema due to hereditary C1-inhibitor deficiency. *Clin Exp Allergy* 2014;**44**:1503–14.
46. Colobran R, Lois S, de la Cruz X, Pujol-Borrell R, Hernández-González M, Guilarte M. Identification and characterization of a novel splice site mutation in the SERPING1 gene in a family with hereditary angioedema. *Clin Immunol* 2014;**150**:143–8.
47. Johnsrud I, Kulseth MA, Rødningen OK, Landrø L, Helsing P, Waage Nielsen E, et al. A nationwide study of Norwegian patients with hereditary angioedema with C1 inhibitor deficiency identified six novel mutations in SERPING1. *PLoS One* 2015;**10**:e0131637.
48. Sekijima Y, Hashimoto T, Kawachi Y, Koshihara H, Otsuka F, Ikeda S. A novel RNA splice site mutation in the C1 inhibitor gene of a patient with type I hereditary angioedema. *Intern Med* 2004;**43**:253–5.
49. Siddique Z, McPhaden AR, Lappin DF, Whaley K. An RNA splice site mutation in the C1-inhibitor gene causes type I hereditary angio-oedema. *Hum Genet* 1991;**88**:231–2.
50. Kawachi Y, Hibi T, Yamazaki S, Otsuka F. A novel donor splice site mutation in the C1 inhibitor gene of a patient with type I hereditary angioneurotic edema. *J Invest Dermatol* 1998;**110**:837–9.
51. Kesim B, Uyguner ZO, Gelincik A, Mete Gökmen N, Sin AZ, Karakaya G, et al. The Turkish Hereditary Angioedema Pilot Study (TURHAPS): the first Turkish series of hereditary angioedema. *Int Arch Allergy Immunol* 2011;**156**:443–50.