Received: 8 May 2019 | Revised: 30 July 2019 | Accepted: 4 August 2019

DOI: 10.1002/humu.23890

Human Mutation HGVS HUMAN GENOME VARIATION SOCIETY WILEY

RESEARCH ARTICLE

Identification of splice defects due to noncanonical splice site or deep-intronic variants in *ABCA4*

Zeinab Fadaie[1] | Mubeen Khan[1] | Marta Del Pozo-Valero[1,2] | Stéphanie S. Cornelis[1] | Carmen Ayuso[2] | Frans P. M. Cremers[1] | Susanne Roosing[1] | The ABCA4 study group

[1]Department of Human Genetics, Donders Institute for Brain, Cognition and Behavior, Radboud University Medical Center, Nijmegen, The Netherlands

[2]Department of Genetics, Instituto de Investigación Sanitaria–Fundación Jiménez Díaz University Hospital, Universidad Autónoma de Madrid (IIS-FJD, UAM), Madrid, Spain

Correspondence
Susanne Roosing, PhD, Department of Human Genetics, Donders Institute for Brain, Cognition, and Behavior, Radboud University Medical Center, P. O. Box 9101, 6500 HB Nijmegen, The Netherlands.
Email: Susanne.Roosing@radboudumc.nl

Funding information
Foundation Fighting Blindness, Grant/Award Number: PPA-0517–0717-RAD; Retina UK, Grant/Award Number: UK grant GR591; Fighting Blindness Ireland, Grant/Award Number: FB18CRE

Abstract

Pathogenic variants in the ATP-binding cassette transporter A4 (*ABCA4*) gene cause a continuum of retinal disease phenotypes, including Stargardt disease. Noncanonical splice site (NCSS) and deep-intronic variants constitute a large fraction of disease-causing alleles, defining the functional consequences of which remains a challenge. We aimed to determine the effect on splicing of nine previously reported or unpublished NCSS variants, one near exon splice variant and nine deep-intronic variants in *ABCA4*, using in vitro splice assays in human embryonic kidney 293T cells. Reverse transcription-polymerase chain reaction and Sanger sequence analysis revealed splicing defects for 12 out of 19 variants. Four deep-intronic variants create pseudoexons or elongate the upstream exon. Furthermore, eight NCSS variants cause a partial deletion or skipping of one or more exons in messenger RNAs. Among the 12 variants, nine lead to premature stop codons and predicted truncated *ABCA4* proteins. At least two deep-intronic variants affect splice enhancer and silencer motifs and, therefore, these conserved sequences should be carefully evaluated when predicting the outcome of NCSS and deep-intronic variants.

KEYWORDS

ABCA4, deep-intronic variants, noncanonical splice site variant, splice enhancers, splice silencers, Stargardt disease

1 | INTRODUCTION

Inherited retinal diseases (IRDs) are clinically and genetically heterogeneous conditions (Berger, Kloeckener-Gruissem, & Neidhardt, 2010; Sullivan & Daiger, 1996), which makes it a great challenge for clinicians to come to a genetic diagnosis in affected individuals. However, by defining the genetic cause, the genetic risk of the disease for other family members can be assessed and it provides essential prognostic information for affected family members and possible therapeutic approaches. More knowledge of genetic variability in a gene will also provide better insight and understanding of the disease mechanism (Carss et al., 2017; Ellingford et al., 2015).

The rise of next-generation sequencing technology has drastically changed the opportunities in obtaining a genetic diagnosis in affected individuals (Neveling et al., 2012; Vaz-Drago, Custódio, & Carmo-Fonseca, 2017). Using these techniques, hundreds of thousands of single-nucleotide (nt) variants are detected in each individual.

However, in many instances, the functional significance of variants remains unclear. Several in silico tools predict the putative effect of missense and splice variants. The latter variants can be experimentally assessed for their pathogenicity by performing in vitro midigene splice assays (Sangermano et al., 2018).

Stargardt disease 1 (STGD1; MIM# 600110) is the most common inherited macular disease. It is characterized by bilateral progressive loss of central vision, color vision defect, photophobia, and importantly, delayed dark adaptation and fundus imaging point to accumulation of lipofuscin (Fishman, Farbman, & Alexander, 1991; Stargardt, 1909). STGD1 is an autosomal recessive disease caused by pathogenic variants in the ATP-binding cassette subfamily A member 4 (*ABCA4*) gene (MIM# 601691; NM_000350.2; Allikmets et al., 1997). The protein is comprised of two tandem halves, each of which consists of a nt-binding domain, a cytoplasmic domain, and a transmembrane domain, followed by a large extracellular segment (Bungert, Molday, & Molday, 2001). *ABCA4* is a 128 kb gene containing 50 exons that encodes a polypeptide of 2,273 amino acids and is located in the rod and cone photoreceptor cells, as well as the retinal pigment epithelium (Ahn, Wong, & Molday, 2000; Lenis et al., 2018; Sun, Molday, & Nathans, 1999). One or both copies of this gene were found to be mutated in the majority of patients with STGD1 (Zernant et al., 2014), in 30% of patients with cone-rod dystrophy (Cremers et al., 1998; Maugeri et al., 2000), and in approximately 5% of individuals with pan-retinal dystrophy or a phenotype resembling retinitis pigmentosa (Cremers et al., 1998).

The *ABCA4* gene carries many pathogenic noncanonical splice site (NCSS) variants. A comprehensive study on all *ABCA4* variants published up to 2016 showed a total of 5,962 likely causal variants or alleles in *ABCA4* of which 13.6% (809/5,962) are located in NCSS (Cornelis et al., 2017). These NCSS variants are located at the first and last three nucleotides of an exon as well as the −3 to −14nts from the acceptor site, and +3 to +6 from the donor site. Besides NCSS variants, many pathogenic variants were observed in canonical sequences at the AG-acceptor (−1 and −2) and GT-donor (+1 and +2) nucleotides. The functional consequences of many NCSS variants in *ABCA4* were revealed using in vitro minigene and midigene splice assays in human embryonic kidney 293T (HEK293T) cells (Sangermano et al., 2018; Schulz et al., 2017).

Until recently, a single pathogenic variant was identified in approximately 25% of STGD1 cases worldwide (Zernant et al., 2014, 2017). Approximately 40% of these cases (~10% of all STGD1 cases), many of whom showing late-onset STGD1, were explained by the frequent coding variant c.5603A>T (p.Asn1868Ile) with an allele frequency of 0.07 in most control populations. This variant generally was found in *trans* with severe *ABCA4* variants (Runhart et al., 2018; Zernant et al., 2017). Another 40% of monoallelic STGD1 probands (10% of all STGD1 cases) carried deep-intronic variants in their second alleles (Bauwens et al., 2019; Braun et al., 2013; Sangermano et al., 2019; Zernant et al., 2014). The functional consequences of many of these variants only came to light when performing in vitro splice assays (Bauwens et al., 2019; Sangermano et al., 2019) or by reverse transcription-polymerase chain reaction (RT-PCR) analysis of photoreceptor progenitor cells differentiated from patient-derived induced pluripotent stem cells (Albert et al., 2018; Sangermano et al., 2016).

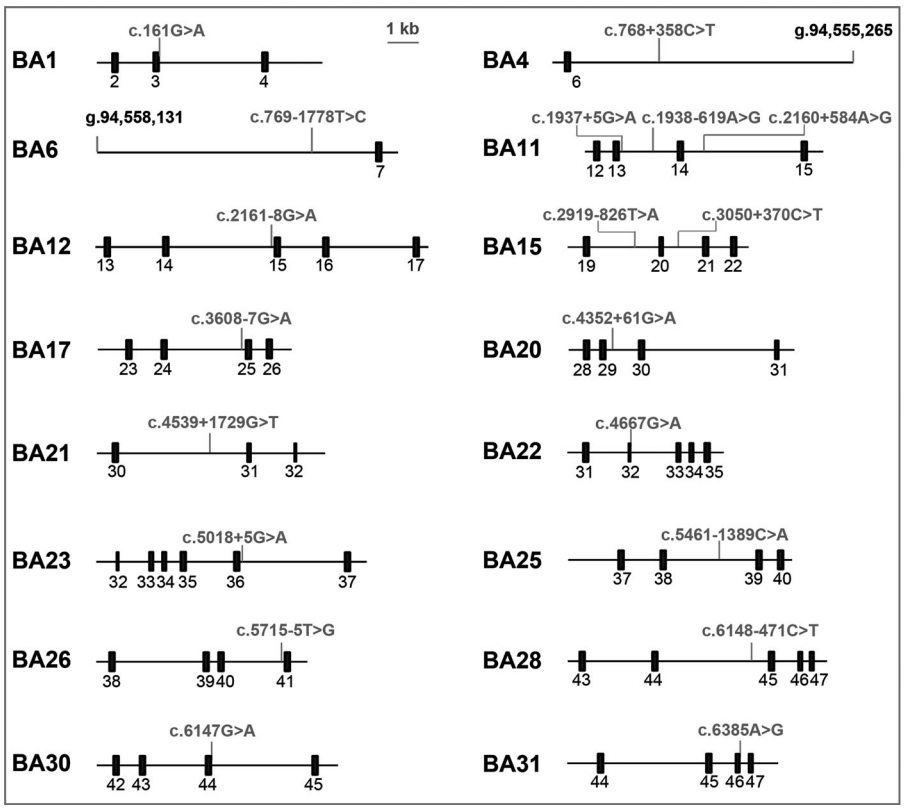

FIGURE 1 Schematic representation of mutant midigene splice constructs of *ABCA4* and corresponding locations of the NCSS and deep-intronic variants tested. The exons are represented as black rectangles. BA depicts the BAC-derived clones used that were previously described (Sangermano et al., 2018). The BA11 and BA15 variants each were introduced separately into the WT constructs. *ABCA4*, ATP-binding cassette subfamily A member 4; BAC, bacterial artificial chromosome; NCSS, noncanonical splice site; WT, wild-type

TABLE 1 In vitro assessed NCSS variants and the observed RNA and predicted protein effects

DNA variant	RNA effect	Protein effect	Splice defect	Variant effect
c.161G>A	r.[161_302del,=][b]	p.[Cys54Serfs*14,Cys54Tyr]	Exon 3 skipping	Moderate[a]
c.1937+5G>A	r.1806_1937del	p.(Tyr603_Ser646del)	132-nt exon 13 deletion	Severe
c.2161−8G>A	r.2161_2382del	p.(His721_Val794)	Exon 15 skipping	Severe
c.3608−7G>A	r.=	p.(=)	None	N.A.
c.4667G>A	r.4635_4667del	p.(Ser1545_Gln1555del)	Exon 32 skipping	Severe
c.5018+5G>A	r.4849_5018del	p.(Val1617Alafs*113)	Exon 35 skipping	Severe
c.5715−5T>G	r.5461_5714delins5715-4_5715-1	p.(Thr1821Serfs*34)	Exon 39/40 skipping	Severe
c.6147G>A	r.6006_6147del	p.(Ser2002Argfs*11)	Exon 44 skipping	Severe
c.6385A>G	r.6340_6386del	p.(Val2114Hisfs*4)	47-nt exon 46 deletion	Severe

Note: The RNA and protein effect annotations show the most abundant product followed by the less abundant RNA product observed.
Abbreviations: N.A., not applicable; NCSS, noncanonical splice site.
[a]The splice defect is classified as moderate. However, the missense variant Cys54Tyr was proposed earlier to be considered a severe variant in Stargardt cases.
[b]The equal sign depicts the presence of >15% WT RNAs.

To expand the knowledge of the consequences of NCSS and deep-intronic variants in *ABCA4*, we studied reported and unpublished NCSS, as well as reported deep-intronic variants in *ABCA4* employing midigene-based splice assays. This study contributes to a deeper understanding of alternative splicing through the activation of cryptic splice sites and the presence of exonic splice enhancers (ESEs) or exonic splice silencers (ESSs) and provides evidence for the pathogenicity of 12 splicing variants, thereby significantly expanding our knowledge of the effects of putative splice defects.

2 | MATERIALS AND METHODS

2.1 | Editorial policies and ethical considerations

The study adhered to the tenets of the Declaration of Helsinki and was approved by the local ethics committees of each participating center. Written informed consent was obtained from patients before inclusion in the study.

2.2 | Clinical studies

The 12 probands carried pathogenic *ABCA4* variants (Table S1) and were diagnosed with macular dystrophies or cone-rod dystrophy (Tables S2 and S3). Medical records of each patient were reviewed for clinical examination including the age of onset, visual acuity, fundus photography, and electroretinogram, where available.

2.3 | Selection of NCSS and deep-intronic *ABCA4* variants

Variants were selected for in vitro analysis when they adhered to the following criteria: (a) NCSS and "near exon aberrant RNA" (NEAR) splice variants were predicted to result in a reduction of at least 2% of the relative strength in at least two of five different splice site prediction algorithms i.e. Splice Site Finder-like (SSFL), MaxEntScan, NNSPLICE, GeneSplicer, and human splicing finder (HSF) (Desmet et al., 2009; Pertea et al., 2001; Reese et al., 1997; Shapiro & Senapathy, 1987; Yeo & Burge, 2004). Moreover, variants were selected when nearby (up to 300 base pairs [bps]) a cryptic splice site was strengthened or created.

Deep-intronic variants were included if a mutant cryptic splice site was predicted by at least two out of five algorithms with ≥75% score in the presence of an already existing other splice site within 300 nts, which together could result in the formation of a pseudoexon (PE) (Sangermano et al., 2019). As it has been shown that ESEs and ESSs have a significant effect on the splicing process in human cells we assessed ESEs through ESEfinder, an in silico prediction tool integrated with Alamut® version 2.10 (Cartegni, Chew, & Krainer, 2002; Fairbrother & Chasin, 2000). ESEfinder determines the presence of five different ESE elements, that is SF2/ASF, SF2/ASF (immunoglobulin M-BRCA1), SC35, SRp40, and SRp55 (Cartegni, Wang, Zhu, Zhang, & Krainer, 2003; Smith et al., 2006). Moreover, suggested pathogenic variants adhered to the selection criteria when differences could be observed in predicted ESEs or ESSs between wild-type (WT) and mutant sequences. ESSs were assessed through algorithms incorporated in HSF as introduced by Wang et al. (2004) and Sironi et al. (2004).

Nineteen variants were selected to be assessed by a midigene splice assays. All in silico splice site prediction scores of the variants investigated in this study are provided in Table S4. Of these, 16 variants were previously reported (Bauwens et al., 2019; Cornelis et al., 2017; Fujinami et al., 2013; Tayebi et al., 2019; Zernant et al., 2014, 2017; Zhang, Arias, Ke, & Chasin, 2009), we also assessed in-house data based on molecular inversion probes-based sequence analysis of 108

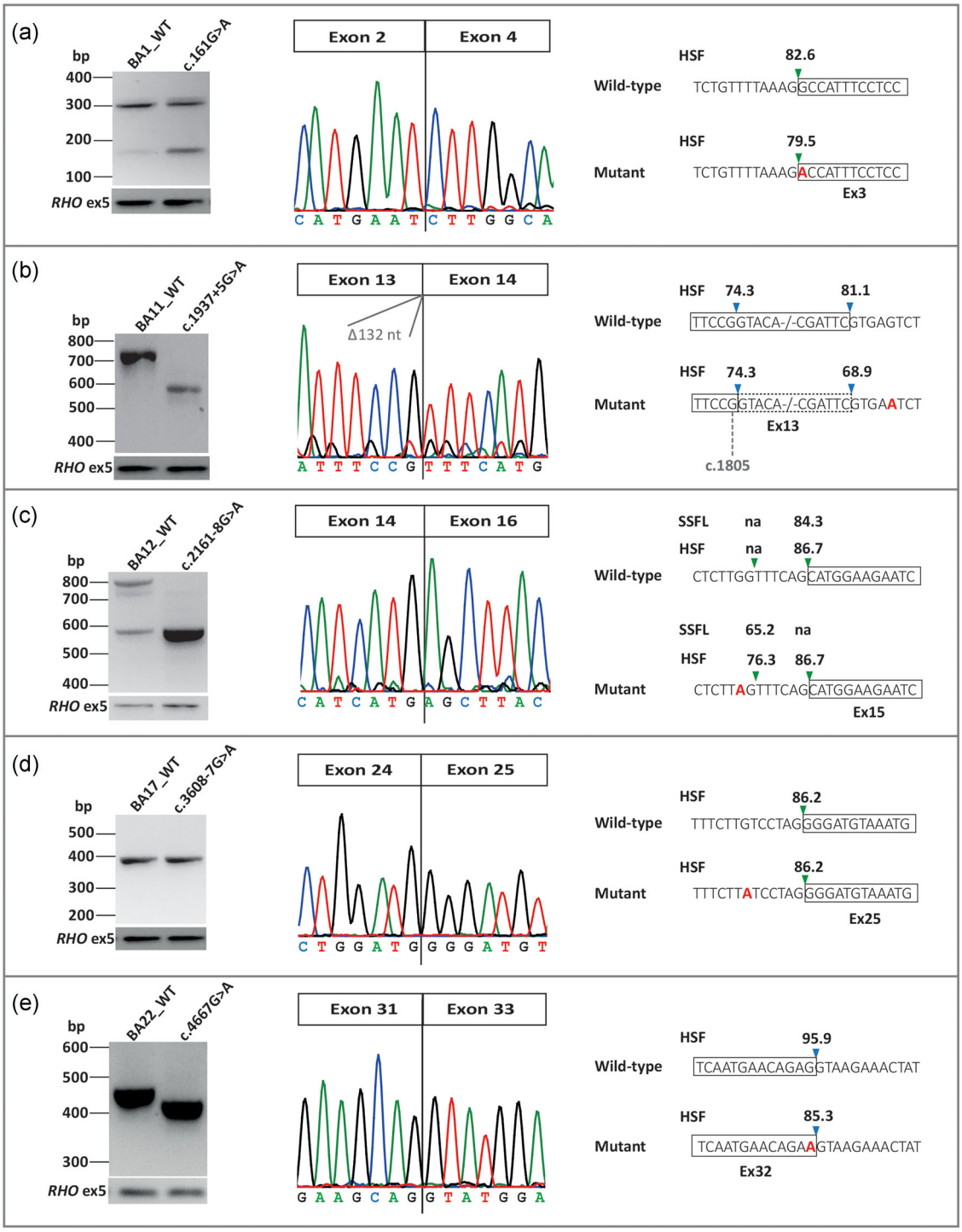

FIGURE 2 Overview of splice defects for nine NCSS variants. Exon 5 *RHO* RT-PCR was used as a control for transfection efficiency. The chromatogram presents the nucleotides identified in the mutant midigene construct. (a) Exon 3 showed weak natural exon skipping in the WT construct, which is significantly increased for the c.161G>A mutant. (b) The recruitment of a cryptic SDS in exon 13 at position c.1806 resulted in a 132-nt deletion. (c) The c.2161−8G>A mutant construct showed full exon 15 skipping. Note that exon 15 also shows natural exon skipping. (d) c.3608−7G>A did not result in a splice defect. (e, f) Variants c.4667G>A and c.5018+5G>A led to complete exon 32 and 35 skipping, respectively. (g) RT-PCR for the c.5715−5T>G mutant construct showed a complex splicing pattern. Exons 39 and 40 are partially skipped in the WT messenger RNA (mRNA); c.5715−5T>G induced exon 39/40 skipping. This variant also created a new splice acceptor site (SAS) upstream of exon 41 which led to the insertion of four nucleotides into the mature mRNA. (h) c.6147G>A caused complete exon 44 skipping. (i) The use of a cryptic SDS in exon 46 caused the 47-nt exon deletion. bp, base pair; HSF, human splicing finder; int, intron; NCSS, noncanonical splice site; nt, nucleotide; RT-PCR, reverse transcription-polymerase chain reaction; SDS, splice donor site; WT, wild-type

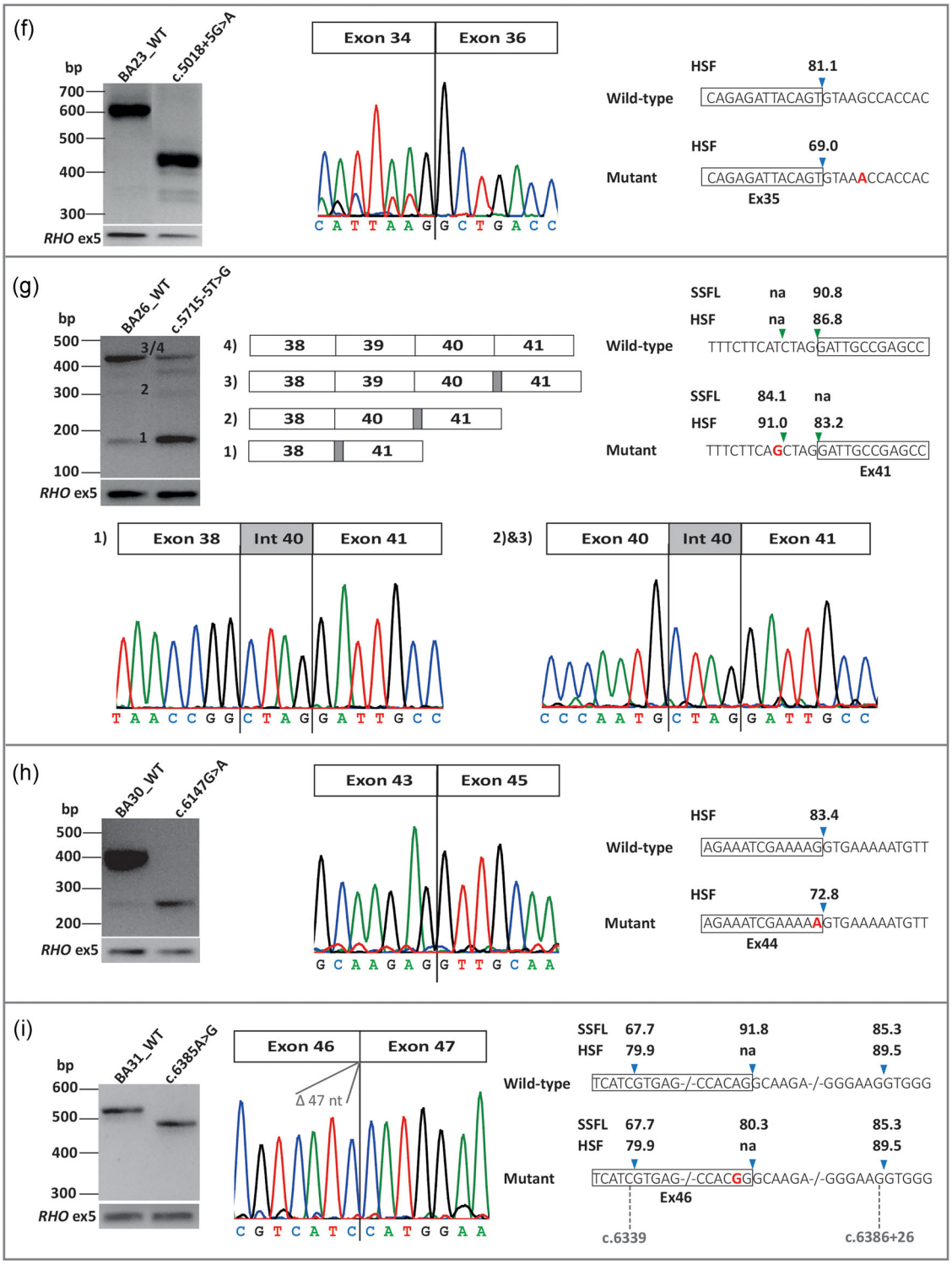

FIGURE 2 Continued

inherited retina-disease associated genes in an approximately 5,000 probands (S. R. and F. P. M. C., unpublished data). The latter analysis led to the inclusion of c.1937+5G>A, c.5715−5T>G, and c.6147G>A, as these fulfilled the criteria described above.

2.4 | Generation of *ABCA4* WT and mutant midigenes

Previously, we generated a library of 31 overlapping WT midigenes (BA1–BA31; Sangermano et al., 2018) (Khan et al., 2019). Through Gateway cloning and subsequent site-directed mutagenesis, mutant

constructs were generated for all 19 variants investigated in this study (Figure 1). Subsequently, WT and mutant constructs were independently transfected in HEK293T cells, assessed through RT-PCR, gel analysis, and followed by Sanger sequencing of the observed fragments. When a multitude of products was observed after gel electrophoresis, they were quantified using Fiji software, as previously described (Sangermano et al., 2018). In addition, for the c.5715−5T>G variant, all observed fragments were cloned via the pGEM-T Easy Vector System I (Promega, Madison, WI) according to the manufacturer's protocol and analyzed by Sanger sequencing. All mutagenesis, exonic primers, and quantification measurements are available in Tables S5–S7.

3 | RESULTS

3.1 | Splicing effect of NCSS variants in *ABCA4*

The nine NCSS variants experimentally assessed with their observed effect on splicing are listed in Table 1. Of all variants, only two were

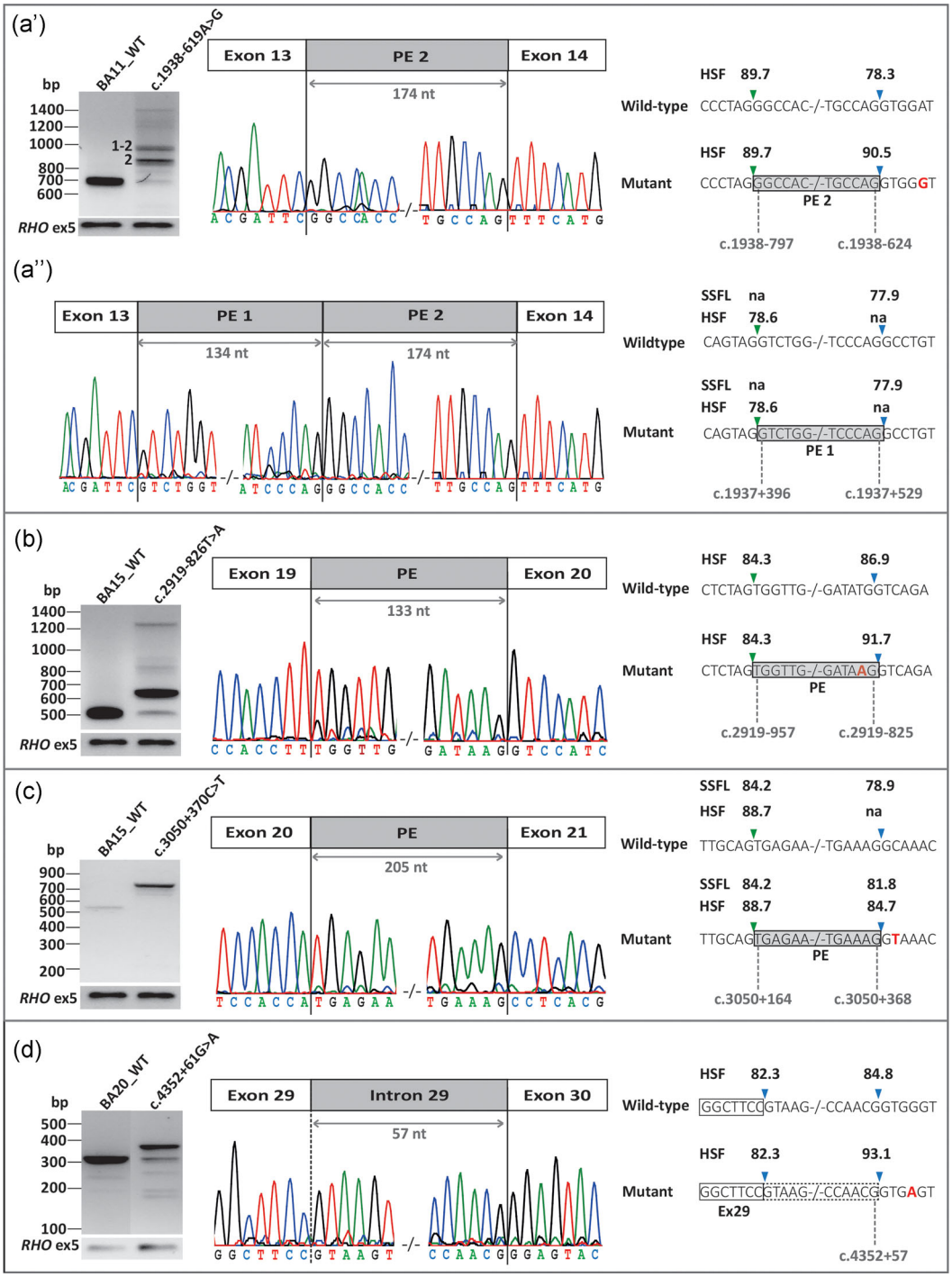

FIGURE 3 Continued.

observed in the "control" population database gnomAD (c.161G>A, allele frequency: 0.00003608; c.3608−7G>A, allele frequency: 0.000004087). Four of the selected variants (c.161G>A, c.4667G>A, c.6147G>A, and c.6385A>G) were in the coding regions of *ABCA4*, whereas the remaining NCSS variants were located in introns. Eight of the nine NCSS variants led to partial or entire exon skipping, while only c.3608−7G>A did not show an effect on splicing. The c.1937+5G>A, c.2161−8G>A, and c.4667G>A variants led to partial in-frame deletions of *ABCA4* exons. The c.6385A>G variant caused an open reading frame disruption by a 47-nt deletion of exon 46 due to the use of a cryptic splice donor site (SDS) at position c.6340 in the exon (Figure 2). The other five NCSS variants caused frameshifts and led to predicted truncated proteins (Table 1).

Variant c.5715−5T>G, located five nucleotides upstream of exon 41, showed a complex defect in splicing compared with the WT product. Among the different fragments, we observed exon 39 skipping as well as exon 39/40 skipping. The most prominent mutant messenger RNA (mRNA) contained a 4-nt insertion of the exon 41 splice acceptor site (SAS) along with exon 39/40 skipping. We did not observe any mRNA with only exon 40 skipping.

3.2 | Splicing effect of deep-intronic variants

The 10 NEAR and deep-intronic variants with their observed effect on splicing from this study are presented in Table 2. None of these variants was observed in gnomAD, indicating that they are very rare. Four showed a splicing defect, while the other six variants did not show any effect on splicing (Figure S1). Of the four variants that did show a splice defect, c.1938−619A>G, c.2919−826T>A, and c.3050+370C>T created a PE in the mature mRNAs, leading to a frameshift and a subsequent predicted truncated protein. The NEAR splicing variant c.4352+61G>A elongated the *ABCA4* mRNA downstream of exon 29 by 57 nts, resulting in a premature stop codon, likely due to an increased cryptic SDS score according to five prediction algorithms, even though there is an inactivation of several ESE protein binding motifs, such as SC35 and SF2/ASF (Figure 3).

Of the six variants that showed no effect on splicing, the c.768+358C>T, c.769−1778T>C, and c.5461−1389C>A variants showed alternative ESEs being recognized in the corresponding mutant and the c.2160+584A>G, c.4539+1729G>T, and c.6148−471C>T variants significantly increased the splicing scores in different algorithms (Table S3). However, none of these six variants showed splicing defects in HEK293T cells in the in vitro assay (Figure S1).

3.3 | Clinical characteristics of STGD1 cases carrying causal NCSS or deep-intronic variants

Clinical data were collected from the corresponding patients carrying the variants assessed in this study. The overview of the *ABCA4* variants observed in trans in these patients as well as phenotypic details are provided in Tables S1 and S2, respectively.

For all variants assessed in this study, we quantified visible fragments from gel electrophoresis analysis to determine the severity of variants. The quantifications observed ≤15% of WT mRNA in four of five variants with visible WT fragments suggesting that they represent severe variants (Table S7). For the analysis of c.161G>A, we noted naturally occurring exon skipping for exon 3 in 14% of the WT construct. Therefore, we normalized the full-length fragment in mutant construct (44%) to the full-length RNA including exon 3 in the WT (86%). The splice defect of c.161G>A was classified as moderately severe due to the resulting 51% (86/44 × 100) remaining product. However, already earlier the missense variant was proposed to be severe in STGD1 cases, and therefore we argue that both the exon skipping and missense mutation are contributing to the severity of this variant. Six other variants did not show any WT fragments besides the mutant fragment and are therefore deemed severe causal variants. Variant c.3608−7G>A did not show a splice defect and is therefore likely not causative until proven otherwise with another experimental setup.

Through this study, we established the splice defects for 12 *ABCA4* variants found in 12 macular dystrophy probands. The c.1938−619A>G variant is located at position +5 of the PE and strengthened a cryptic SDS which likely led to the recognition of ESE SC35 and SRp55 motifs. These ESEs are located in the exon–intron boundaries and are shown to promote the recognition of exons with weak 5′ and 3′ splice sites and are being involved in exon definition by assisting in the recruitment of splicing factors before the removal of the adjacent introns. Therefore, the recognition of these ESE motifs will facilitate the splicing machinery to detect the nearby sequence as an exon (Buvoli, Buvoli, & Leinwand, 2007; Lam & Hertel, 2002; Wu, Zhang, & Zhang, 2005). The individual carrying this

FIGURE 3 Overview of splice defects for three deep-intronic variants and one near exon variant. All WT and mutant midigenes were transfected in HEK293T cells and their splicing effects were identified by RT-PCR. The exon 5 *RHO* RT-PCR was used as a control for transfection efficiency. The chromatograms present the nucleotides identified in the mutant midigene construct. (a) Variant c.1938−619A>G strengthened a cryptic SDS and surprisingly resulted in two independent PE insertions. Fragment 1–2 consists of a fused fragment of two PEs encompassing 134 and 174 nt (PE1–2). PE1 (134 nt) extends from c.1937+396 to c.1937+529 and the boundaries of PE2 (174-nt) are c.1938−797 and c.1938−624. Fragment 2 contains the second PE of 174-nt (PE2) only. (b) Variant c.2919−826T>A strengthened a cryptic SDS and led to a 133-nt PE. (c) Variant c.3050+370C>T created a new canonical SDS which resulted in a 105-nt PE insertion. (d) The NEAR splicing variant c.4352+61G>A enhanced a cryptic SDS's HSF score and elongated exon 29 with 57 nt in 84% percent of the mRNA. bp, base pair; HEK293T, human embryonic kidney 293T; HSF, human splicing factor; mRNA, messenger RNA; NEAR, near exon aberrant RNA; nt, nucleotide; PE, pseudoexon; RT-PCR, reverse transcription-polymerase chain reaction; SDS, splice donor site; SSFL, SpliceSiteFinder-like; WT, wild-type

TABLE 2 In vitro assessed near exon and deep-intronic variants and the observed RNA and predicted protein effects

DNA variant	RNA effect	Protein effect	Splice defect	Variant effect
c.768+358C>T	r.=	p.(=)	None	N.A.
c.769−1778T>C	r.=	p.(=)	None	N.A.
c.1938−619A>G	r.[1937_1938ins [1937+396_1937+529;1938-797_1938-624],1937_1938ins1938-797_1938-624]	p.[Phe647Alafs*22,Phe647-Serfs*22]	174-nt PE and 308-nt PE	Severe
c.2160+584A>G	r.=	p.(=)	None	N.A.
c.2919−826T>A	r.[2918_2919in-s2919−957_2919−825,=][a]	p.[Leu973Phefs*1,=][a]	133-nt PE	Severe
c.3050+370C>T	r.3050_3051ins3050+164_3050+368	p.(Leu1018Glufs*4)	205-nt PE	Severe
c.4352+61G>A	r.[4352_4353ins4352+1_4352+57,=][a]	p.[Glu1452*,=][a]	57-nt exon 29 elongation	Severe
c.4539+172>T	r.=	p.(=)	None	N.A.
c.5461−1389C>A	r.=	p.(=)	None	N.A.
c.6148−471C>T	r.=	p.(=)	None	N.A.

Note: The RNA and protein effect annotations show the most abundant product followed by the less abundant RNA product observed.
Abbreviations: N. A., not applicable; nt, nucleotide; PE, pseudoexon.
[a]The equal sign depicts the presence of >15% WT RNAs.

variant was identified with c.5882G>A (p.Gly1961Glu) as a second variant and was shown to segregate with the disease in the family with a typical STGD1 phenotype (Zernant et al., 2014).

The c.2919−826T>A variant is located at the penultimate nt of the newly recognized PE, increasing the SDS scores at the c.2919−824 positions. The corresponding patient carries the mild c.5882G>A (p.(Gly1961Glu)) variant on the other allele and shows bull's eye maculopathy, which is characteristic for STGD1 cases carrying p.(Gly1961Glu) (Zernant et al., 2014).

4 | DISCUSSION

We found splice defects for 12 of the 19 (63%) assessed NCSS or deep-intronic *ABCA4* variants (Figure S2). Based on the midigene splice assays, 11 could be classified as severe variants and one variant (c.161G>A) is considered to have a moderately severe splice effect. The splicing effects of all tested variants and their RNA and protein annotations were uploaded into the *ABCA4* Leiden Open (source) Variant Database (*ABCA4*-LOVD). In addition, we updated the protein outcome for those variants that were already present in the database (www.lovd.nl/ABCA4).

Among nine NCSS variants, six variants were found to cause skipping of one or two exons. The c.5715−5T>G variant revealed multiple fragments due to the splice defect. All erroneous fragments showed a 5′ elongation of exon 41, however, the variant also leads to skipping of both exon 39 and 40 and of exon 39 only. Whereas most NCSS variants lead to complete exon skipping, the midigene splice assays for both c.1937+5G>A and c.6385A>G variants resulted in partial exon deletions due to the use of cryptic splice sites within exons 13 and 46, respectively. As observed for c.1937+5G>A, an upstream cryptic SDS at c.1806 was utilized which led to an in-frame 132-nt deletion of exon 13, thereby removing amino acid residues 603–646. This affects the first extracellular domain of the *ABCA4* protein, but a residual function of the protein cannot be excluded.

The prediction for the coding variant c.6385A>G was challenging as the canonical SDS of exon 46 contains GC instead of GT, which is recognized only by the SSFL algorithm. The SSFL values were reduced from 91.8% to 80.3% for the mutant variant. While we anticipated a 26-nt exon elongation due to high predicted scores for a cryptic splice site, or intron retention due to the small intron 46 size (73 nt), we observed the use of an upstream cryptic SDS at position c.6340. This resulted in a 47-nt deletion of exon 46, leading to a frameshift resulting in a predicted truncated protein. We hypothesize that the strong cryptic SDS at position c.6386+27 (SSFL score: 85.3) may not be used by the splicing machinery due to the high abundance of silencers preventing the binding of splicing factors (Figure S3). The preferred cryptic SDS within the exon has a relatively low SSFL score of 67.7, but the region has few predicted silencer motifs and therefore likely is preferred over other SDSs. The corresponding patient with a retinitis pigmentosa-like phenotype has the c.5461−10T>C (p.[Thr1821Aspfs*6,Thr1821Valfs*13]) pathogenic variant as the other allele (Braun et al., 2013), which is the most frequent severe variant in *ABCA4* (Sangermano et al., 2016).

Among the 10 NEAR splice and deep-intronic variants evaluated, c.1938−619A>G, c.2919−826T>A, and c.3050+370C>T generated PEs that contain stop codons and thus result in predicted truncated ABCA4 proteins. Vaz-Drago et al. (2017) recently showed that the majority of deep-intronic variants generate a canonical SAS or SDS, while the minority creates or disrupts an ESE or ESS element. In our study, only one (c.3050+370C>T) out of three variants introduced a new splice site leading to a PE. Analysis of the c.1938−619A>G variant revealed a complex effect and introduced two PEs. The variant created a 134-nt PE as well as a second PE of 174 nt which

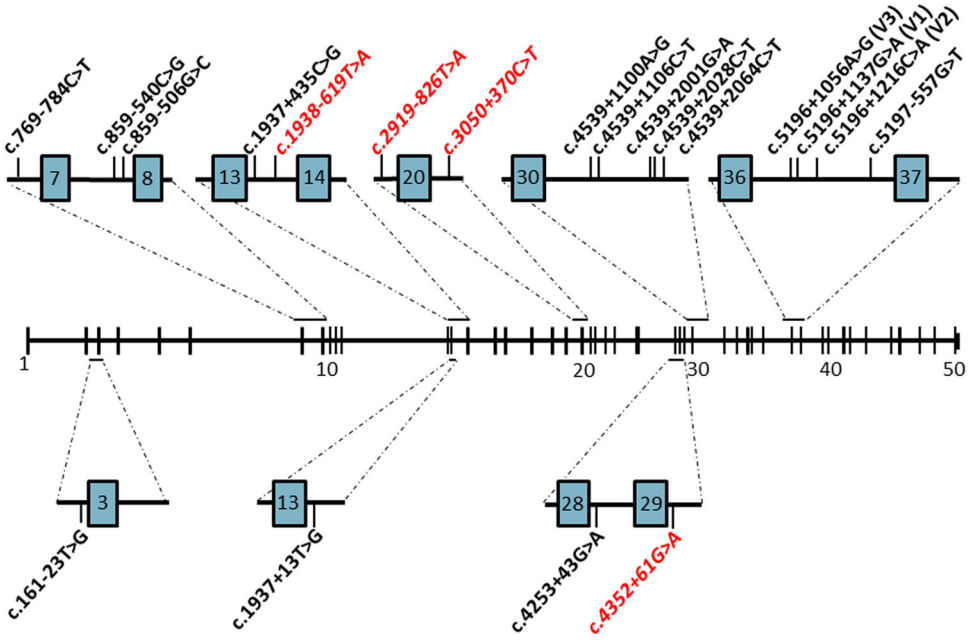

FIGURE 4 The landscape of all currently known pathogenic deep-intronic and NEAR splicing variants in *ABCA4*. The three deep-intronic variants and one NEAR splicing variant deemed as pathogenic in this study are depicted in red/italic. Other variants were described elsewhere (Albert et al., 2018; Bauwens et al., 2019; Braun et al., 2013; Sangermano et al., 2019). *ABCA4*, ATP-binding cassette subfamily A member 4; NEAR, near exon aberrant RNA

resided 491 nt downstream. The 134-nt PE was previously shown to result from the c.1937+435C>G (Sangermano et al., 2019). To our knowledge, the phenomenon of two PEs generated by a single variant has not been described before and the underlying mechanism remains to be elucidated.

The WT c.3050+370C residue is part of a cryptic "GC-type" splice site as predicted by SSFL (SSFL score: 78.9%), but apparently is not employed by the splicing machinery. The c.3050+370C>T variant, however, creates a canonical GT (SSFL score: 81.8%) which was recognized by all five splice site algorithms and led as expected to a PE insertion that contains a premature stop codon after four amino acids p.(Leu1018Glufs*4).

Six deep-intronic variants did not show a splicing effect, while there were strong predictions for cryptic SDSs for the c.2160+584A>G and c.4539−1729G>T variants. An explanation for the absence of PEs may be a paucity of retina-specific splicing motifs and/or an abundance of silencer motifs (Murphy, Cieply, Carstens, Ramamurthy, & Stoilov, 2016). Variants showing no effect on splicing in this study may still be proven to be pathogenic when assessed in induced Pluripotent Stem Cell (iPSC)-derived photoreceptor precursors. For example, pathogenicity was proven when studying *ABCA4* variants c.4539+2001G>A and c.4539+2028C>T in photoreceptor precursor cells derived from patient fibroblasts while no effect on splicing was detected in fibroblasts of the same patients (Albert et al., 2018). Moreover, a retina-specific increase of a 128-nt PE insertion was also observed for the most frequent Leber congenital amaurosis-associated *CEP290* variant, c.2991+1655A>G (den Hollander et al., 2006; Dulla et al., 2018).

As current estimates indicate that NCSS, NEAR splice, and deep-intronic variants represent 15–20% of the causes of recessive human diseases (Carss et al., 2017; Matlin, Clark, & Smith, 2005), we sought to assess pathogenicity of 19 *ABCA4* variants that were previously published, or identified in our cohort. We clearly determined the effect on splicing and, consequently, the highly likely pathogenicity for six NCSS variants as well as three deep-intronic variants, and thereby contribute to a growing list of NCSS variants and the 16 pathogenic deep-intronic variants published previously (Figure 4; Albert et al., 2018; Bauwens et al., 2019; Braun et al., 2013; Sangermano et al., 2019). Moreover, our study revealed a fourth NEAR splicing variant to be pathogenic in addition to the three previously described NEAR splice variants in *ABCA4* (Bauwens et al., 2019; Sangermano et al., 2018). An overview of the currently known pathogenic deep-intronic and NEAR splice variants is presented in Figure 4.

Determining the precise effects of splice site variants will open new opportunities for therapeutic approaches for patients carrying these variants. As previously shown, modulation of *ABCA4* pre-mRNA splicing can be executed through antisense oligonucleotides, which can bind complementarily to mRNA and manipulate the splicing process by skipping PEs (Albert et al., 2018; Bauwens et al., 2019; Sangermano et al., 2019). Moreover, antisense oligonucleotides have proven their effect in in vitro and in vivo studies for the most frequent deep-intronic variant in *CEP290* that causes Leber congenital amaurosis (Collin et al., 2012; Garanto et al., 2016). The latter is currently in Phase II clinical trial (https://clinicaltrials.gov/ct2/show/NCT03140969).

By evaluating the pathogenic effects of putative splicing variants, we gained crucial knowledge for evaluation of yet to be identified NCSS, NEAR or deep-intronic variants. While predictions by splice site algorithms are crucial, we also observed an important, and possibly essential, role of ESE and ESS motifs. Future studies regarding retina-specific splicing motifs and proteins will improve predictions for the effect of novel NCSS and deep-intronic variants in *ABCA4* as well as in other IRD-associated genes.

ACKNOWLEDGMENTS

The work of Z. F. is supported by the Foundation Fighting Blindness USA Project Program Award, grant no. PPA-0517–0717-RAD (to F. P. M. C. and S. R.). The work of M. K. is supported by the Rotterdamse Stichting Blindenbelangen, the Stichting Blindenhulp, the Stichting tot Verbetering van het Lot der Blinden, and the Stichting Blinden-Penning (to F. P. M. C and S. R.). The work of M. K. and S. C. is supported by Retina UK, grant no. GR591 (to F. P. M. C.). The work of S. C. is supported by the Fighting Blindness Ireland (to F. P. M. C. and S. R.). The work of M. D. P. V. is supported by the Conchita Rábago Foundation and the Boehringer Ingelheim Fonds. The work of C. A. is supported by grants PI16/0425 from ISCIII partially supported by the European Regional Development Fund (ERDF), ONCE Foundation, and Ramon Areces Foundation. The work of R. A. is supported, in part, by grants from the National Eye Institute/NIH EY028203, EY019007 (Core Support for Vision Research), Foundation Fighting Blindness USA Project Program Award PPA-1218-0751-COLU, and unrestricted funds from Research to Prevent Blindness (New York, NY) to the Department of Ophthalmology, Columbia University.

DATA ACCESSIBILITY

The authors confirm that the data supporting the finding of this study are available within the article and its supplementary materials.

ABCA4 Study Group

Rando Allikmets, Department of Ophthalmology and Department of Pathology and Cell Biology, Columbia University, New York, New York; Miriam Bauwens, Center for Medical Genetics, Ghent University and Ghent University Hospital, Ghent, Belgium; Mohammad Ghofrani, Cellular and Molecular Research Center, Qom University of Medical Sciences, Qom, Iran; Michael B. Gorin, Department of Ophthalmology, David Geffen School of Medicine, Stein Eye Institute, University of California, Los Angeles, Los Angeles, California; Department of Human Genetics, David Geffen School of Medicine, University of California, Los Angeles, Los Angeles, California; Mohammad Keramatipour, Department of Medical Genetics, Tehran University of Medical Sciences, Tehran, Iran; Francesca Simonelli, Eye Clinic, Multidisciplinary Department of Medical, Surgical and Dental Sciences, University of Campania Luigi Vanvitelli, Naples, Italy; Naeimeh Tayebi, Program in Genetics and Genome Biology, The Hospital for Sick Children, Toronto, Canada; Andrea Vincent, Department of Ophthalmology, New Zealand National Eye Center, Faculty of Medical and Health Sciences, The University of Auckland, Grafton, Auckland, New Zealand; Eye Department, Greenlane Clinical Center, Auckland District Health Board, Auckland, New Zealand; Nicole Weisschuh, Molecular Genetics Laboratory, Center for Ophthalmology, Institute for Ophthalmic Research, University of Tuebingen, Tuebingen, Germany.

ORCID

Zeinab Fadaie http://orcid.org/0000-0003-0440-3846
Mubeen Khan http://orcid.org/0000-0002-7545-5662
Marta Del Pozo-Valero http://orcid.org/0000-0003-3934-0873
Stéphanie S. Cornelis http://orcid.org/0000-0001-7619-927X
Carmen Ayuso http://orcid.org/0000-0002-9242-7065
Frans P.M. Cremers http://orcid.org/0000-002-4954-5592
Susanne Roosing http://orcid.org/0000-0001-9038-0067

REFERENCES

Ahn, J., Wong, J. T., & Molday, R. S. (2000). The effect of lipid environment and retinoids on the ATPase activity of ABCR, the photoreceptor ABC transporter responsible for Stargardt macular dystrophy. *Journal of Biological Chemistry*, *275*(27), 20399–20405.

Albert, S., Garanto, A., Sangermano, R., Khan, M., Bax, N. M., Hoyng, C. B., … Cremers, F. P. M. (2018). Identification and rescue of splice defects caused by two neighboring deep-intronic ABCA4 mutations underlying Stargardt disease. *The American Journal of Human Genetics*, *102*(4), 517–527.

Allikmets, R., Shroyer, N. F., Singh, N., Seddon, J. M., Lewis, R. A., Bernstein, P. S., & Hutchinson, A. (1997). Mutation of the Stargardt disease gene (ABCR) in age-related macular degeneration. *Science*, *277*(5333), 1805–1807.

Bauwens, M., Garanto, A., Sangermano, R., Naessens, S., Weisschuh, N., De Zaeytijd, J., … De Baere, E. (2019). ABCA4-associated disease as a model for missing heritability in autosomal recessive disorders: Novel noncoding splice, cis-regulatory, structural, and recurrent hypomorphic variants. *Genetics in Medicine*, *21*, 1761–1771. https://doi.org/10.1038/s41436-018-0420-y

Berger, W., Kloeckener-Gruissem, B., & Neidhardt, J. (2010). The molecular basis of human retinal and vitreoretinal diseases. *Progress in Retinal and Eye Research*, *29*(5), 335–375.

Braun, T. A., Mullins, R. F., Wagner, A. H., Andorf, J. L., Johnston, R. M., Bakall, B. B., … Stone, E. M. (2013). Non-exomic and synonymous variants in ABCA4 are an important cause of Stargardt disease. *Human Molecular Genetics*, *22*(25), 5136–5145.

Bungert, S., Molday, L. L., & Molday, R. S. (2001). Membrane topology of the ATP binding cassette transporter ABCR and its relationship to ABC1 and related ABCA transporters: Identification ofn-linked glycosylation sites. *Journal of Biological Chemistry*, *276*(26), 23539–23546.

Buvoli, M., Buvoli, A., & Leinwand, L. A. (2007). Interplay between exonic splicing enhancers, mRNA processing, and mRNA surveillance in the dystrophic Mdx mouse. *PLOS One*, *2*(5), e427. https://doi.org/10.1371/journal.pone.0000427

Carss, K. J., Arno, G., Erwood, M., Stephens, J., Sanchis-Juan, A., Hull, S., … Huissoon, A. (2017). Comprehensive rare variant analysis via whole-genome sequencing to determine the molecular pathology of inherited retinal disease. *The American Journal of Human Genetics*, *100*(1), 75–90.

Cartegni, L., Chew, S. L., & Krainer, A. R. (2002). Listening to silence and understanding nonsense: Exonic mutations that affect splicing. *Nature Reviews Genetics*, *3*(4), 285–298.

Cartegni, L., Wang, J., Zhu, Z., Zhang, M. Q., & Krainer, A. R. (2003). ESEfinder: A web resource to identify exonic splicing enhancers. *Nucleic Acids Research*, *31*(13), 3568–3571.

Collin, R. W., Den Hollander, A. I., Van Der Velde-visser, S. D., Bennicelli, J., Bennett, J., & Cremers, F. P. M. (2012). Antisense oligonucleotide (AON)-based therapy for Leber congenital amaurosis caused by a frequent mutation in CEP290. *Molecular Therapy-Nucleic Acids*, *1*, e14.

Cornelis, S. S., Bax, N. M., Zernant, J., Allikmets, R., Fritsche, L. G., den Dunnen, J. T., … Cremers, F. P. M. (2017). In silico functional meta-analysis of 5,962 ABCA4 variants in 3,928 retinal dystrophy cases. *Human Mutation*, *38*(4), 400–408.

Cremers, F. P. M., van de Pol, D. J., van Driel, M., den Hollander, A. I., van Haren, F. J., Knoers, N. V., … Hoyng, C. B. (1998). Autosomal recessive retinitis pigmentosa and cone-rod dystrophy caused by splice site mutations in the Stargardt's disease gene ABCR. *Human Molecular Genetics*, *7*(3), 355–362.

den Hollander, A. I., Koenekoop, R. K., Yzer, S., Lopez, I., Arends, M. L., Voesenek, K. E. J., … Cremers, F. P. M. (2006). Mutations in the CEP290 (NPHP6) gene are a frequent cause of Leber congenital amaurosis. *The American Journal of Human Genetics*, *79*(3), 556–561.

Desmet, F.-O., Hamroun, D., Lalande, M., Collod-Béroud, G., Claustres, M., & Béroud, C. (2009). Human splicing finder: An online bioinformatics tool to predict splicing signals. *Nucleic Acids Research*, *37*(9), e67.

Dulla, K., Aguila, M., Lane, A., Jovanovic, K., Parfitt, D. A., Schulkens, I., … Cheetham, M. E. (2018). Splice-modulating oligonucleotide QR-110 restores CEP290 mRNA and function in human c.2991+1655A>G LCA10 models. *Molecular Therapy - Nucleic Acids*, *12*, 730–740. https://doi.org/10.1016/j.omtn.2018.07.010

Ellingford, J. M., Sergouniotis, P. I., Lennon, R., Bhaskar, S., Williams, S. G., Hillman, K. A., … Black, G. C. M. (2015). Pinpointing clinical diagnosis through whole exome sequencing to direct patient care: A case of Senior-Loken syndrome. *The Lancet*, *385*(9980), 1916.

Fairbrother, W. G., & Chasin, L. A. (2000). Human genomic sequences that inhibit splicing. *Molecular and Cellular Biology*, *20*(18), 6816–6825.

Fairbrother, W. G., Yeo, G. W., Yeh, R., Goldstein, P., Mawson, M., Sharp, P. A., & Burge, C. B. (2004). RESCUE-ESE identifies candidate exonic splicing enhancers in vertebrate exons. *Nucleic Acids Research*, *32*, W187–W190.

Fishman, G. A., Farbman, J. S., & Alexander, K. R. (1991). Delayed rod dark adaptation in patients with Stargardt's disease. *Ophthalmology*, *98*(6), 957–962.

Fujinami, K., Zernant, J., Chana, R. K., Wright, G. A., Tsunoda, K., Ozawa, Y., … Michaelides, M. (2013). ABCA4 gene screening by next-generation sequencing in a British cohort. *Investigative Ophthalmology & Visual Science*, *54*(10), 6662–6674.

Garanto, A., Chung, D. C., Duijkers, L., Corral-Serrano, J. C., Messchaert, M., Xiao, R., … Collin, R. W. (2016). In vitro and in vivo rescue of aberrant splicing in CEP290-associated LCA by antisense oligonucleotide delivery. *Human Molecular Genetics*, *25*(12), 2552–2563.

Khan, M, Cornelis, S. S., Khan, M. I., Elmelik, D., Manders, E., Bakker, S., … Cremers, FPM (2019). Cost-effective molecular inversion probe-based ABCA4 sequencing reveals deep-intronic variants in Stargardt disease. *Human Mutation*, https://doi.org/10.1002/humu.23787

Lam, B. J., & Hertel, K. J. (2002). A general role for splicing enhancers in exon definition. *RNA*, *8*(10), 1233–1241.

Lenis, T. L., Hu, J., Ng, S. Y., Jiang, Z., Sarfare, S., Lloyd, M. B., … Radu, R. A. (2018). Expression of ABCA4 in the retinal pigment epithelium and its implications for Stargardt macular degeneration. *Proceedings of the National Academy of Sciences of the United States of America*, *115*(47), E11120–E11127. https://doi.org/10.1073/pnas.1802519115

Matlin, A. J., Clark, F., & Smith, C. W. J. (2005). Understanding alternative splicing: Towards a cellular code. *Nature Reviews Molecular Cell Biology*, *6*(5), 386–398.

Maugeri, A., Klevering, B. J., Rohrschneider, K., Blankenagel, A., Brunner, H. G., Deutman, A. F., … Cremers, F. P. M. (2000). Mutations in the ABCA4 (ABCR) gene are the major cause of autosomal recessive cone-rod dystrophy. *The American Journal of Human Genetics*, *67*(4), 960–966.

Murphy, D., Cieply, B., Carstens, R., Ramamurthy, V., & Stoilov, P. (2016). The musashi 1 controls the splicing of photoreceptor-specific exons in the vertebrate retina. *PLOS Genetics*, *12*(8), e1006256. https://doi.org/10.1371/journal.pgen.1006256

Neveling, K., Collin, R. W. J., Gilissen, C., van Huet, R. A. C., Visser, L., Kwint, M. P., … Scheffer, H. (2012). Next-generation genetic testing for retinitis pigmentosa. *Human Mutation*, *33*(6), 963–972.

Pertea, M., Lin, X, & Salzberg, SL (2001). GeneSplicer: A new computational method for splice site prediction. *Nucleic Acids Research*, *29*(5), 1185–1190.

Reese, M. G., Eeckman, F. H., Kulp, D., & Haussler, D. (1997). Improved splice site detection in Genie. *Journal of Computational Biology*, *4*(3), 311–323.

Runhart, E. H., Sangermano, R., Cornelis, S. S., Verheij, J. B., Plomp, A. S., Boon, C. J., & Cremers, F. P. M. (2018). The common ABCA4 variant p. Asn1868Ile shows nonpenetrance and variable expression of Stargardt disease when present in trans with severe variants. *Investigative Ophthalmology & Visual Science*, *59*(8), 3220–3231.

Sangermano, R., Bax, N. M., Bauwens, M., Van den Born, L. I., De Baere, E., Garanto, A., … Albert, S. (2016). Photoreceptor progenitor mRNA analysis reveals exon skipping resulting from the ABCA4 c. 5461−10T→ C mutation in Stargardt disease. *Ophthalmology*, *123*(6), 1375–1385.

Sangermano, R., Garanto, A., Khan, M., Runhart, E. H., Bauwens, M., Bax, N. M., … Cremers, F. P. M. (2019). Deep-intronic ABCA4 variants explain missing heritability in Stargardt disease and allow correction of splice defects by antisense oligonucleotides. *Genetics in Medicine*, *21*, 1751–1760. https://doi.org/10.1038/s41436-018-0414-9

Sangermano, R., Khan, M., Cornelis, S. S., Richelle, V., Albert, S., Garanto, A., … Cremers, F. P. M. (2018). ABCA4 midigenes reveal the full splice spectrum of all reported noncanonical splice site variants in Stargardt disease. *Genome Research*, *28*(1), 100–110.

Schulz, H. L., Grassmann, F., Kellner, U., Spital, G., Rüther, K., Jägle, H., & Weber, B. H. (2017). Mutation spectrum of the ABCA4 gene in 335 Stargardt disease patients from a multicenter German cohort—impact of selected deep intronic variants and common SNPs. *Investigative Ophthalmology & Visual Science*, *58*(1), 394–403.

Shapiro, M. B., & Senapathy, P. (1987). RNA splice junctions of different classes of eukaryotes: Sequence statistics and functional implications in gene expression. *Nucleic Acids Research*, *15*(17), 7155–7174.

Sironi, M., Menozzi, G., Riva, L., Cagliani, R., Comi, G. P., Bresolin, N., … Pozzoli, U. (2004). Silencer elements as possible inhibitors of pseudoexon splicing. *Nucleic Acids Research*, *32*(5), 1783–1791.

Smith, P. J., Zhang, C., Wang, J., Chew, S. L., Zhang, M. Q., & Krainer, A. R. (2006). An increased specificity score matrix for the prediction of SF2/ASF-specific exonic splicing enhancers. *Human Molecular Genetics*, *15*(16), 2490–2508.

Stargardt, K. (1909). Über familiäre, Progressive Degeneration in der Maculagegend des Auges. *Albrecht von Græfe's Archiv für Ophthalmologie*, *71*(3), 534–550.

Sullivan, L. S., & Daiger, S. P. (1996). Inherited retinal degeneration: Exceptional genetic and clinical heterogeneity. *Molecular Medicine Today*, *2*(9), 380–386.

Sun, H., Molday, R. S., & Nathans, J. (1999). Retinal stimulates ATP hydrolysis by purified and reconstituted ABCR, the photoreceptor-specific ATP-binding cassette transporter responsible for Stargardt disease. *Journal of Biological Chemistry*, *274*(12), 8269–8281.

Tayebi, N., Akinrinade, O., Khan, M. I., Hejazifar, A., Dehghani, A., Cremers, F. P. M., & Akhlaghi, M. (2019). Targeted next generation sequencing reveals genetic defects underlying inherited retinal disease in Iranian families. *Molecular Vision*, *8*(25), 106–117.

Vaz-Drago, R., Custódio, N., & Carmo-Fonseca, M. (2017). Deep intronic mutations and human disease. *Human Genetics*, *136*(9), 1093–1111.

Wang, Z., Rolish, M. E., Yeo, G., Tung, V., Mawson, M., & Burge, C. B. (2004). Systematic identification and analysis of exonic splicing silencers. *Cell*, *119*(6), 831–845.

Wu, Y., Zhang, Y., & Zhang, J. (2005). Distribution of exonic splicing enhancer elements in human genes. *Genomics*, *86*(3), 329–336. https://doi.org/10.1016/j.ygeno.2005.05.011

Yeo, G., & Burge, C. B. (2004). Maximum entropy modeling of short sequence motifs with applications to RNA splicing signals. *Journal of Computational Biology*, *11*(2-3), 377–394.

Zernant, J., Lee, W., Collison, F. T., Fishman, G. A., Sergeev, Y. V., Schuerch, K., … Allikmets, R. (2017). Frequent hypomorphic alleles account for a significant fraction of ABCA4 disease and distinguish it from age-related macular degeneration. *Journal of Medical Genetics*, *54*(6), 404–412.

Zernant, J., Xie, Y., Ayuso, C., Riveiro-Alvarez, R., Lopez-Martinez, M.-A., Simonelli, F., … Allikmets, R. (2014). Analysis of the ABCA4 genomic locus in Stargardt disease. *Human Molecular Genetics*, *23*(25), 6797–6806.

Zhang, X. H.-F., Arias, M. A., Ke, S., & Chasin, L. A. (2009). Splicing of designer exons reveals unexpected complexity in pre-mRNA splicing. *RNA*, *15*(3), 367–376.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.

How to cite this article: Fadaie Z, Khan M, Del Pozo-Valero M, et al. Identification of splice defects due to noncanonical splice site or deep-intronic variants in *ABCA4*. *Human Mutation*. 2019;40:2365–2376. https://doi.org/10.1002/humu.23890