UAM

Universidad Autónoma
de Madrid

UNIVERSIDAD AUTÓNOMA DE MADRID

PROGRAMA DE DOCTORADO EN BIOCIENCIAS MOLECULARES

ESTUDIO CLÍNICO Y MOLECULAR DE LAS DISTROFIAS HEREDITARIAS DE RETINA ASOCIADAS A *ABCA4* Y *PROM1*

TESIS DOCTORAL

Marta del Pozo Valero

Madrid, 2020

UAM

Universidad Autónoma
de Madrid

UNIVERSIDAD AUTÓNOMA DE MADRID

Facultad de Medicina

Departamento de Bioquímica

ESTUDIO CLÍNICO Y MOLECULAR DE LAS DISTROFIAS HEREDITARIAS DE RETINA ASOCIADAS A *ABCA4* Y *PROM1*

Memoria presentada por **Marta del Pozo Valero**, Graduada en Biotecnología, para optar al título de Doctor por la Universidad Autónoma de Madrid

Directora de Tesis:

Dra. Carmen Ayuso García

Tutor de Tesis:

Dr. Leandro Sastre Garzón

iiS FJD | INSTITUTO DE INVESTIGACIÓN SANITARIA FUNDACIÓN JIMÉNEZ DÍAZ

Realizado en el Instituto de Investigación Sanitaria Fundación Jiménez Díaz (IIS-FJD)

Dña. CARMEN AYUSO GARCÍA, jefa del Departamento de Genética del Hospital Universitario Fundación Jiménez Díaz y Directora Científica del Instituto de Investigación Sanitaria Fundación Jiménez Díaz,

CERTIFICA:

Que la presente memoria titulada: "*Estudio clínico y molecular de las distrofias hereditarias de retina asociadas a ABCA4 y PROM1*", que presenta Dña. Marta del Pozo Valero para obtener el grado de Doctor, ha sido realizada bajo mi dirección. Por ello, la autorizo para su presentación ante un Tribunal Calificador.

Madrid, 26 de agosto de 2020

Fdo. Dra. Carmen Ayuso García

Directora de la Tesis Doctoral

Fdo. Dr. Leandro Sastre Garzón

Tutor de la Tesis Doctoral

If you think research is expensive,

try disease!

Mary Lasker

#SinCienciaNoHayFuturo

Agradecimientos

Llegó la parte que dicen que es más difícil, pero la que más ganas tenía de escribir.

En primer lugar, me gustaría dar las gracias a mi directora de Tesis, Carmen. Primero, por responderme al correo en el que le mandaba mi CV, y después, por permitirme formar parte de este grupo y valorar mi trabajo. Gracias por la cercanía y las palabras de motivación.

También me gustaría agradecer a la Fundación Conchita Rábago, y en especial a Marta, su ayuda todos estos años.

Gracias a todos mis compañeros de Genética, que han hecho que tuviera ganas de venir cada día al laboratorio. A quienes ya se han ido, Patri, Clara, Lilián, Valentina y a María, la otra Conchita, a quien espero que de ahora en adelante le vaya muy muy bien como Doctora. También Ana A. y Esther, que se fueron hace poquito, y Jennifer, a la que mi sitio también echa de menos. A Inma, que me ha acompañado desde el principio y con la que puedo discutir cosas serias y no tan serias. A Marta Cortón, por ayudarme y enseñarme cuando lo he necesitado, Olga y Loli, por estar ahí desde el principio y seguirme el rollo. A Fiona, Rosa, Almu, Marta Rodríguez, Ruth y María José por poder contar con ellas. A Inés, la otra pucelana, por ayudarme siempre que lo necesito. A Camilo, Jesús y Miguel, por hacerme reír aún diciendo solo "Delpo" o vacilándome (con cariño). A Aurora y Sonia, por su simpatía al entrar por la puerta. Al resto de mis compis de investigación, Cris, Irene, Andrea, Lorena, Rosario, Berta, por hacer más amenos los días de trabajo. También agradecer a los bioinfos Pablo, Perce, Ionut, Raquel y Lorena, y a Alejandra T. y Alejandra D. (Jr. or Sr.), que ayudan a que el ambiente de trabajo, y de no trabajo, sea muy (o más) llevadero. Y gracias al resto de compañeros del Servicio de Genética, Chony, Ana B., Carol, Isabel, Saoud, Elvira, Lidia, Fernando, Laura, Rocío, Laura, Fermina. Gracias también a las oftalmólogas de la FJD, Blanca, Belén y Ester, porque siempre se puede contar con su ayuda.

I want to thank my colleagues of Nijmegen, Frans, for allowing me to learn more about this field and be part of his group, Mubeen, for teaching me everything she knew, and Susanne, Stephanie, Suzanne, Janine, Zeinab, Ellen, Femke, Marco and Jan for making me feel nice during my stay.

También me gustaría recordar a otras personas que han formado parte de esta etapa, a Fahimeh, por enseñarme a tener paciencia, a Pilar, Elisabeth, Amelia y Julia, a los editores y revisores de las revistas científicas, por enseñarme cómo poner en valor mi trabajo, a la gente de gestión de la Escuela de Doctorado, mi tutor Leandro, coordinadores y secretarios del programa de Doctorado, por su ayuda y responder siempre a mis insistentes preguntas, la mayoría en esta difícil etapa final, a los revisores externos de esta tesis, Carlo y Sandro por su interés, y a los miembros del tribunal por aceptar la invitación. Y a los pacientes, a los que espero que sigamos ayudando con muchos más proyectos de investigación.

Agradecimientos

Por supuesto, gracias a toda mi familia y amigos. En especial a Bea, que ha tenido que aprender nociones básicas de genética y oftalmología para diseñar esta portada.

Gracias a mi abuela, por transmitirme su fuerza y por su envidiable personalidad.

Gracias a mi hermana, Raquel, por estar en lo importante, y en lo nada importante.

Gracias a mi padre, por todo lo que ha hecho y sigue haciendo.

Y ya para finalizar, le dedico esta tesis a la persona que más se la merece, mi madre.

Resumen, *Abstract*

Las distrofias hereditarias de retina (DHR) son un conjunto de enfermedades que provocan la degeneración de ambos tipos de fotorreceptores, conos y bastones. Las distrofias maculares (DM) están caracterizadas por la pérdida de visión central y atrofia de la mácula, que es la región central de la retina. La enfermedad de Stargardt es la maculopatía más frecuente, y está causada por mutaciones bialélicas en el gen *ABCA4* (STGD1).

En este trabajo se estudian familias con DM, principalmente diagnosticadas de enfermedad de Stargardt. El gen *ABCA4* explica la causa molecular en el 73% de los pacientes diagnosticados con enfermedad de Stargardt, y en el 25% de los pacientes con otras maculopatías. La mayoría de las mutaciones identificadas se encuentran en regiones codificantes, sin embargo, en este trabajo se describen y estudian funcionalmente nuevas variantes intrónicas que alteran el *splicing*, así como una nueva deleción de origen español. El espectro mutacional del gen *ABCA4* en la cohorte española demuestra que la mayoría de las variantes son de tipo *missense*, seguidas de las *frameshift* y *stop*, y en un pequeño porcentaje encontramos variantes *deep intronic* y "grandes" deleciones.

Además, se describen correlaciones genotipo-fenotipo en 506 familias con mutaciones en el gen *ABCA4*. La mayoría de ellos son pacientes con STGD1, pero en nuestra cohorte también encontramos pacientes con distrofia de conos y bastones (DCB). La edad de inicio de los síntomas de estos últimos es más temprana, y presentan en su mayoría variantes truncantes, aunque también hay tres variantes *missense* que se asocian a este fenotipo más severo. Por otra parte, la variante más frecuente en la población española, c.3386G>T; p.(Arg1129Leu), se asocia a un fenotipo de STGD1, y los homocigotos presentan una forma más leve de la enfermedad, e incluso, respeto foveal, por lo que no presentan síntomas a la edad del diagnóstico.

En algunos pacientes con diagnóstico de enfermedad de Stargardt se identifican variantes en otros genes (4%). Uno de ellos es *PROM1*, del que se amplía su espectro genético y fenotípico con una cohorte de 22 pacientes (~1%). Los pacientes presentan retinosis pigmentaria, DM, o, en su mayoría, DCB, y todos muestran atrofia macular. El patrón de herencia en estas familias puede ser autosómico dominante o recesivo, en función de las variantes que se identifican en *PROM1*.

Los resultados de este trabajo ponen de manifiesto la importancia del diagnóstico genético para el pronóstico de la enfermedad de los pacientes, y su posible inclusión en ensayos clínicos. Además, la descripción genética y fenotípica de grandes cohortes de pacientes posibilita el desarrollo de más ensayos clínicos para estos pacientes.

Inherited retinal dystrophies (IRD) are a group of diseases that affect both type of photoreceptors, cones and rods. Macular Dystrophies (MD) are characterized by central visual loss and atrophy of the macula, the central region of the retina. Stargardt disease is the most frequent maculopathy, and it is caused by biallelic mutations in the *ABCA4* gene (STGD1).

In this work, families diagnosed with MD are studied, mainly those suspected with Stargardt disease. *ABCA4* is the genetic cause of the disease of 73% of Stargardt diagnosed patients, and 25% of patients diagnosed with other maculopathies. Most of the identified mutations are found in coding regions of *ABCA4*, however, in this work we also describe and study by functional assays new intronic variants affecting splicing, as well as one novel deletion with a probably Spanish origin. The mutational spectrum of the *ABCA4* gene in our Spanish cohort show that most variants are missense, followed by frameshift and stop variants, while deep intronic variants and "gross" deletions represent a small percentage of them.

In addition, we describe genotype-phenotype correlations for 506 families with *ABCA4* mutations. Most of them are STGD1 patients, but in our cohort, we also find patients suffering from cone-rod dystrophy (CRD). Age of onset of symptoms in the latter group is earlier, and they mainly carry truncating variants, although three missense variants are associated to this severe phenotype. On the other hand, the most prevalent variant in the Spanish cohort, c.3386G>T; p.(Arg1129Leu), is associated to STGD1 phenotype, and homozygotes for this variant develop a milder phenotype, and may even present foveal sparing, so that these patients have not symptoms at age of diagnosis.

Some patients with Stargardt diagnosis carry variants in genes other than *ABCA4* (4%). One of them is *PROM1*, of which, in this work, we expand its genetic and phenotypic spectrum, with a cohort of 22 families (~1%). Patients with mutations in this gene present retinitis pigmentosa, MD, or, most of them, CRD, all showing macular atrophy. The inherited pattern of these families can be autosomal dominant or recessive, depending on the variants identified in *PROM1*.

Results of this work reveal the importance of the genetic diagnosis for the prognosis of the patients' disease, and for their possible inclusion in clinical trials that are currently being carried out. Furthermore, the genetic and phenotypic description of large cohorts of patients enables the development of more clinical trials for these patients.

Índice

Índice de tablas y figuras

Índice de Tablas

Índice de Figuras

Clave de abreviaturas

ACHM	Acromatopsia
AD	Autosómico dominante
ADN/DNA	Ácido desoxirribonucleico
AR	Autosómico recesivo
ARN/RNA	Ácido ribonucleico
ARNm/ mRNA	ARN mensajero
AO	Age of onset (Edad de inicio)
AON	Oligonucleótidos antisentido
AV	Agudeza visual
AVVs	Virus adenoasociados
BCM	Blue cone monocromatism (Monocromatismo de conos azules)
CV	Campo visual
CN	Ceguera nocturna
CNV	Copy number variation (Variación en el número de copias)
CRD	Cone-rod dystrophy
CSNB	Congenital Stationary Night Blindness (Ceguera nocturna estacionaria congénita)
DC	Distrofia de conos
DCB	Distrofia de conos-bastones
DHR	Distrofias hereditarias de retina
DM	Distrofia macular
DMAE	Degeneración macular asociada a la edad
ECD	Dominio extracellular
EPR	Epitelio pigmentario de la retina
ERG	Electrorretinograma
ESEs	Exonic *splicing* enhancers (Activadores del splicing exónico)
ESSs	Exonic *splicing* silencers (Silenciadores del splicing exónico)
FFM	Fundus flavimaculatus
GMP	Guanosín monofosfato
GMPc	Guanosín monofosfato cíclico
GDP	Guanosín difosfato
GTP	Guanosín Trifosfato
hESC	Células embrionarias humanas

Clave de abreviaturas

HGMD	Human Gene Mutation Database
IPSc	Induced pluripotent stem cells (Células madre pluripotentes inducidas)
ISEs	Intronic splicing enhancers (Activadores del splicing intrónico)
ISSs	Intronic splicing silencers (Silenciadores del splicing intrónico)
LCA	Leber congenital amaurosis (Amaurosis congénita de Leber)
LOVD	Leiden Open Variation Database
MAF	Minor allele frequency (Frecuencia del alelo minoritario)
Mb	Megabases
MD	Macular dystrophy
MLPA	Multiplex ligation-dependent probe amplification (Amplificación de sondas dependiente de ligandos múltiples)
NBD	Dominio de unión a nucleótidos
NCSS	Sitios no canónicos de *splicing*
NGS	Next generation sequencing
STGD/STGD1	Enfermedad de Stargardt/ Enfermedad de Stargardt autosómica recesiva
OCT	Tomografía de coherencia óptica
OMIM	Online Mendelian Inheritance in Man
PDE	Fosfodiesterasa
PE	Fosfatidiletanolamina
Kb	Kilobases
pb	Pares de bases
RP	Retinosis pigmentaria
smMIPs	Single molecule molecular inversion probes
SNP	Single nucleotide polymorphisms (Polimorfismos de nucleótido simple)
TMD	Dominio transmembrana
VUS	Variant of uncertain significance (Variante de significado clínico incierto)
WES	Whole exome sequencing (Secuenciación del exoma completo)
WGS	Whole genome sequencing (Secuenciación del genoma completo)
XL	Ligado al cromosoma X

Introducción

1. La retina y el ciclo visual

1.1. Estructura de la retina

La retina humana es un tejido sensible a la luz que recubre el interior del ojo. La retina deriva del ectodermo interno del tubo neural y su formación comienza entre los 20 y 23 días de vida embrionaria, mediante la evaginación del tubo neural formada por dos capas que reciben el nombre de copa óptica. La capa interna de la copa origina la neurorretina, formada por varias capas de células nerviosas, mientras que la capa externa origina el epitelio pigmentario de la retina (EPR). Las bases de la organización de la retina fueron sentadas por Santiago Ramón y Cajal en 1891 (Ramon y Cajal, 1891). De acuerdo a esta descripción, la retina se organiza en 10 capas paralelas:

- El EPR constituido por células epiteliales cuya función es aportar los nutrientes y reciclar los discos de los segmentos externos de los fotorreceptores.
- Capa de los fotorreceptores: formada por los segmentos externos de los fotorreceptores.
- Capa limitante externa: formada por las uniones de los fotorreceptores y las células de Müller.
- Capa granular externa: capa nuclear formada por los cuerpos de los fotorreceptores.
- Capa plexiforme externa: capa de conexiones sinápticas entre los fotorreceptores y las células bipolares.
- Capa granular interna: capa nuclear formada por los cuerpos de las células bipolares, horizontales, amacrinas y las células gliales de Müller.
- Capa plexiforme interna: capa de conexiones sinápticas entre células bipolares, horizontales y amacrinas.
- Capa de células ganglionares: capa nuclear que contiene los cuerpos de las células ganglionares.
- Capa de fibras del nervio óptico: formada por los axones de las células ganglionares que forman el nervio óptico, encargado de transmitir las señales nerviosas de este circuito celular al cerebro.
- Capa limitante interna: separa la retina del humor vítreo.

Esta estructura se detalla en la Figura I1.
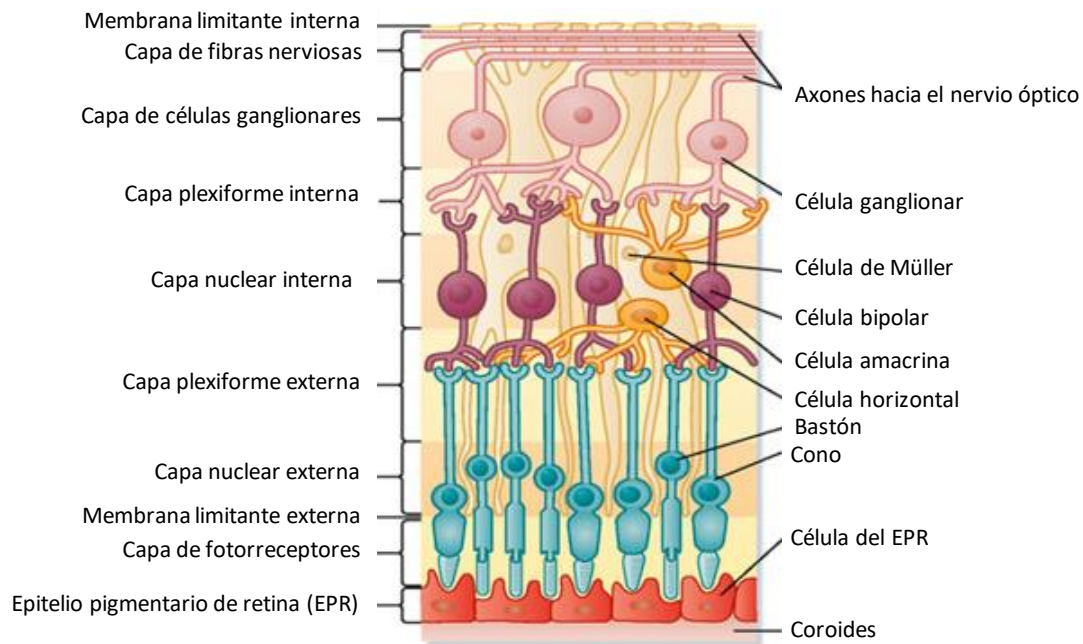
Figura I1. *Capas y células de la retina. Imagen modificada de www.inof.es.*

1.2. Fotorreceptores

Los fotorreceptores de la retina son las células receptoras de los estímulos luminosos que posteriormente se transformarán en imágenes en el cerebro. Existen dos tipos de fotorreceptores, los bastones y los conos (Figura I2). Ambos tipos celulares presentan una estructura formada por el segmento externo, el cilio de conexión y el segmento interno, encargados de la fototransducción, el transporte entre ambos segmentos, y el mantenimiento de la homeostasis, respectivamente. Además del cuerpo celular, en el que se localiza el núcleo de los fotorreceptores, presentan una expansión interna con una fibra conductora y el terminal sináptico para establecer conexiones entre ellos y con las células nerviosas de la capa nuclear externa (Urtubia Vicario, 1997). Las principales diferencias morfológicas entre conos y bastones se encuentran en el segmento externo. En los bastones, los segmentos externos presentan una estructura alargada y cilíndrica, y en los conos son cortos, cónicos y afilados. Además, los discos apilados que forman este segmento externo, y en los que se localizan los fotopigmentos (las opsinas), se encuentran empaquetados y rodeados por la membrana plasmática en los bastones (Gilliam *et al.*, 2012) y en los conos mantienen su continuidad con la membrana celular. También existen diferencias en el número y la distribución de conos y bastones en la retina. Los conos representan aproximadamente el 5% del total de fotorreceptores, y se localizan principalmente en la parte central de la retina, concretamente en la fóvea, y son los responsables de la visión diurna y la agudeza visual. En los conos, existen 3 tipos de opsinas sensibles para diferentes longitudes de onda de la luz (rojo, azul y verde). Por otra parte, los bastones están ampliamente distribuidos por

la periferia de la retina y ausentes en la fóvea, y son los encargados de la visión nocturna. La rodopsina es la única opsina que presentan, responsable de la visión en blanco y negro.

El proceso de renovación de los discos tiene lugar en el cilio de conexión y en el segmento interno de los fotorreceptores. En el cilio de conexión se forman los nuevos discos mediante plegamientos de la membrana celular y en el segmento interno se sintetizan las nuevas proteínas, incluidas las opsinas, que se incorporan a los pliegues membranosos durante la formación de los discos. Los discos antiguos localizados en el extremo distal de los segmentos externos se desprenden y son fagocitados por las células del EPR. La renovación completa del segmento externo tiene una duración de 10-15 días (Lee *et al.*, 2001). El material no fagocitado por el EPR (por ejemplo, los lípidos) queda retenido en estas células y puede acumularse formando gránulos de lipofuscina.
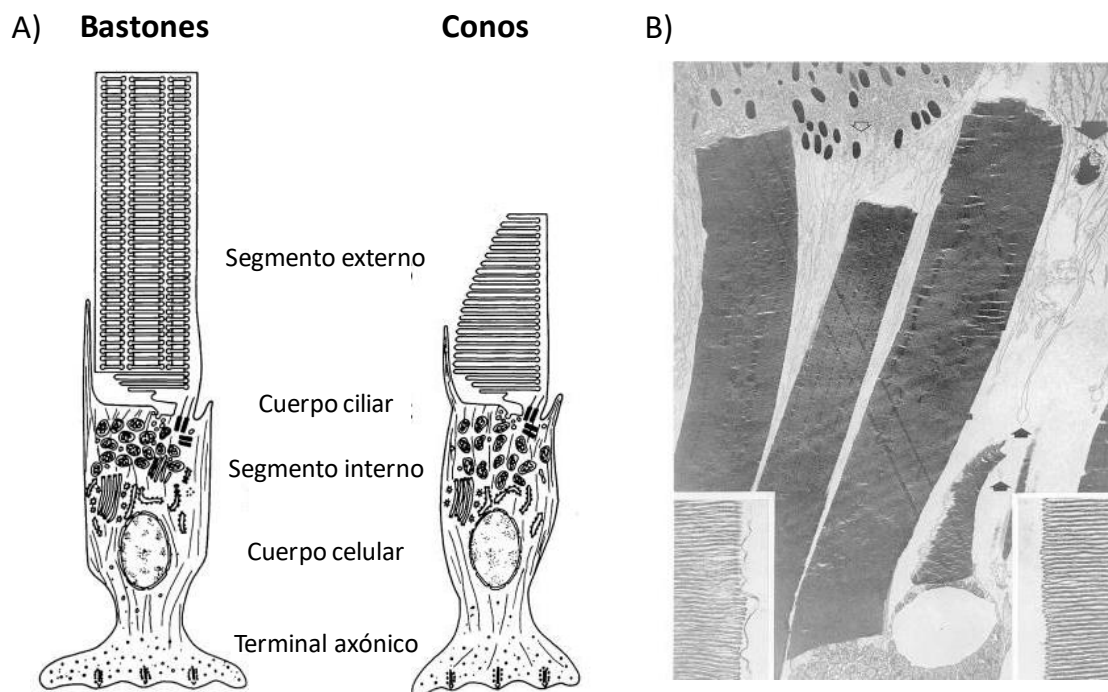
Figura 12. *Estructura de los fotorreceptores.* *A) Esquema B) Vista microscópica. Imagen modificada de Goldberg et al., 2016.*

1.3. Fototransducción y ciclo visual

La fototransducción tiene lugar cuando un fotón atraviesa las diferentes capas de la retina hasta alcanzar los fotorreceptores y generar una respuesta nerviosa que se transmite a la corteza cerebral (Figura I3A). La rodopsina es la molécula sensible a la luz presente en los bastones, y está formada por la opsina y un cromóforo, el 11-cis-retinal, que es un derivado de la vitamina A. La vitamina A ingerida en la dieta es el all-trans-retinol, que se absorbe y pasa a la sangre, y desde allí entra al EPR, donde se convierte a 11-cis-retinal.

En presencia de luz, la rodopsina absorbe un fotón, y como consecuencia se produce la isomerización del 11-cis-retinal a all-trans-retinal. La rodopsina fotoactivada interacciona con la transducina, que es una proteína G compuesta por 3 subunidades (α, β, y γ), que se encuentra inactiva cuando está unida a GDP (guadenosina difosfato). La rodopsina fotoactivada provoca el cambio de GDP a GTP (guadenosina trifosfato) sobre la subunidad α de la transducina, lo que induce su activación y la subunidad α se libera del resto del complejo proteico (Tα-GTP). La disociación de la transducina en sus subunidades permite que Tα-GTP transporte la señal de activación a la fosfodiesterasa (PDE), mediante la eliminación de su subunidad inhibitoria. La PDE activada hidroliza el GMCc a GMP. Como consecuencia de la disminución de GMPc, se cierran los canales de Na^+/Ca^{2+} y se acumulan los iones Na^+ en el exterior de la membrana plasmática. Ese proceso provoca la hiperpolarización del fotorreceptor. Esta hiperpolarización da lugar al cese de liberación del neurotransmisor glutamato por parte del fotorreceptor, cesa la inhibición sobre las células bipolares y, de esta forma, se transmite el impulso nervioso a las células ganglionares y éstas lo transmiten al núcleo geniculado lateral, en el cerebro, que constituye el centro visual del tálamo.

En la fase oscura se produce la cascada de inactivación para devolver a los fotorreceptores al estado de reposo. La disminución de los niveles de Ca^{2+} produce la estimulación de la enzima guanylyl ciclasa para estabilizar la cantidad de GMPc a partir de GTP. La rodopsina es fosforilada por la rodopsina quinasa y se une a la arrestina para su inactivación, y la hidrólisis del GTP unido a Tα-GTP junto al complejo multiproteico RGS9 inactiva a la PDE.

A) **Cascada de fototransducción**
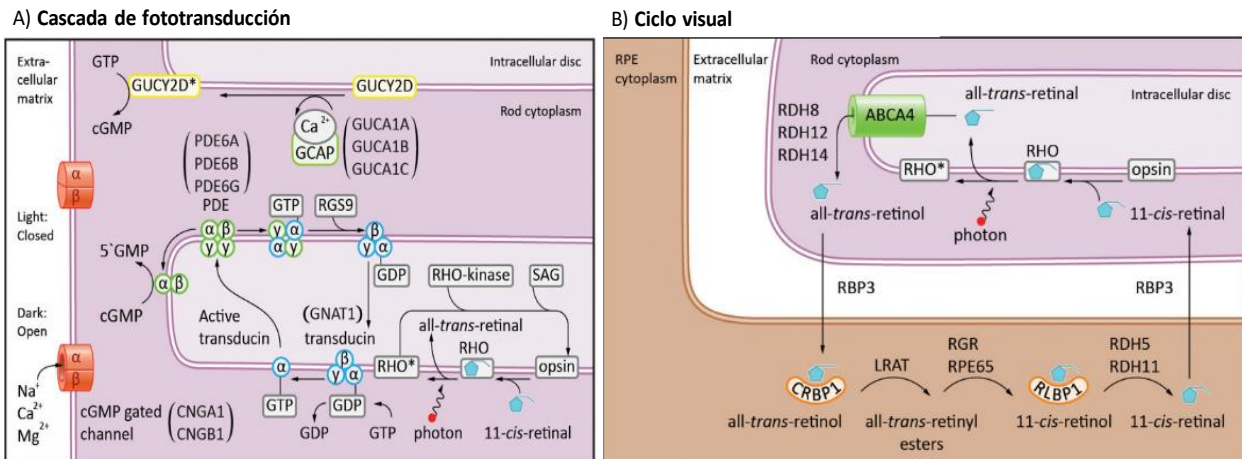
B) **Ciclo visual**

Figura I3. *Representación esquemática de la fototransducción y el ciclo visual. A) Proceso de fototransducción. B) Ciclo visual canónico. Imagen adaptada de* Verbakel *et al., 2018.*

El ciclo visual es un proceso complejo que consiste en regenerar el 11-cis-retinal a partir del all-trans-retinal que se genera en la cascada de fototransducción, y ocurre simultáneamente a ésta. El all-trans-retinal es liberado por la rodopsina en la membrana de los discos de los segmentos externos, donde reacciona con la fosfatidiletanolamina (PE) y forma el all-trans-N-ret-PE (N-RPE). El all-trans-retinal es transportado desde la superficie del lumen hacia la superficie citoplasmática de la membrana de los discos de los fotorreceptores por la actividad flipasa del transportador ABCA4 y se reduce a all-trans-retinol por la enzima retinol deshidrogenasa (RDH8, RDH12 y RDH14). El all-trans-retinol se transporta hasta el EPR a través de su unión a la proteína IRBP. En el citoplasma de las células EPR se une a la proteína CRBP1 y se re-isomeriza mediante las enzimas lecitin-retinol acil transferasa (LRAT) y la isomerohidrolasa retinoide (RPE65) hasta 11-cis-retinol, que es oxidado hasta 11-cis-retinal (RDH5 y RDH11) y transportado de nuevo al citoplasma de los fotorreceptores para que se pueda unir a la opsina y regenerar la rodopsina (Figura I3B).

Existe otro ciclo visual "no canónico" que ocurre adicionalmente en los conos y en las células de Müller, en el que la regeneración del 11-cis-retinal es mucho más rápida para abastecer a los conos en condiciones de luz (Mata *et al.*, 2002).

2. Aspectos generales de las distrofias hereditarias de la retina

Las distrofias hereditarias de retina (DHR) son un conjunto de enfermedades genéticas caracterizadas por la degeneración de los fotorreceptores, generalmente progresiva, que provoca la pérdida de visión en los pacientes (Ayuso and Millan, 2010). La prevalencia de estas enfermedades es de 1:3000 personas, siendo consideradas enfermedades raras.

Los patrones de herencia en las DHR incluyen el autosómico recesivo (AR), el autosómico dominante (AD) y el ligado al cromosoma X (XL) en la mayoría de los casos, sin embargo,

también se ha descrito herencia mitocondrial y formas digénicas asociadas a estas enfermedades (Parmeggiani *et al.*, 2011).

En cuanto a su clasificación clínica, las DHR se dividen en formas periféricas y formas centrales en función de los fotorreceptores primariamente afectados, bastones o conos, respectivamente. Los síntomas asociados a la degeneración de los bastones son la ceguera nocturna y la pérdida de visión periférica, mientras que la degeneración de los conos provoca la disminución de la agudeza visual y del campo visual central, debido a la localización específica de ambos tipos celulares en la retina. Entre las formas periféricas se encuentra la retinosis pigmentaria (RP), que representa la forma más frecuente de DHR con una prevalencia de 1:4000 personas (Ayuso and Millan, 2010). Las formas centrales agrupan las que afectan principalmente a la mácula, conocidas como distrofias maculares (DM). Esta clasificación de las DHR se utilizaba principalmente antes de la era de la genética molecular, y actualmente, debido a las características solapantes de muchas DHR, en particular las diagnosticadas a edades tardías, estos criterios de clasificación han cambiado.

La base genética de las DHR es también muy heterogénea. En el ciclo visual intervienen muchas rutas biológicas, y las mutaciones en las proteínas implicadas en este proceso pueden provocar el desarrollo de estas enfermedades. Actualmente, se han descrito 271 genes asociados a DHR y patologías asociadas (RetNet, fecha de acceso: julio 2020). Algunos subtipos de DHR, como la RP o la amaurosis congénita de Leber (LCA) presentan una elevada heterogeneidad genética (Figura I4), ya que existen varios genes que dan lugar a estas enfermedades, mientras que otras DHR son genéticamente homogéneas, y están causadas por un único gen, como la retinosquisis (gen *RS1*), la coroideremia (gen *CHM*), y la enfermedad de Stargardt (gen *ABCA4*). Además, variaciones en los mismos genes pueden dar lugar a distintos fenotipos de DHR, como es el caso del gen *ABCA4*, y, en algunos genes se encuentran variantes que pueden dar lugar a formas sindrómicas o no sindrómicas de DHR, como es el caso del gen *CEP290*. Por último, varios genes asociados a RP han sido descritos en familias con diferentes patrones de herencia, como los genes *BEST1*, *PROM1*, *RHO* y *RP1*, cuyas variantes pueden dar lugar a enfermedad, ya hayan sido heredadas de forma autosómica dominante o de forma autosómica recesiva.
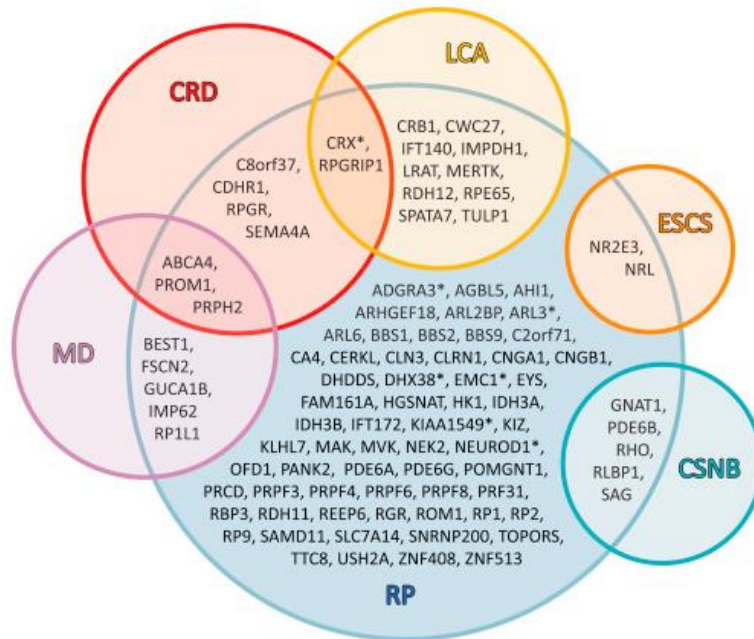
Figura I4. *Heterogeneidad genética en las distrofias hereditarias de retina. MD: distrofia macular; CRD: distrofia de conos y bastones; LCA: amaurosis congénita de Leber; RP: retinosis pigmentaria; ESCS: síndrome del incremento de conos S; CSNB: ceguera nocturna estacionaria. Imagen tomada de* Verbakel *et al.,* 2018*.*

2.1. Distrofias de conos y bastones

Subgrupos de DHR son las distrofias de conos (DC) o distrofias de conos-bastones (DCB), caracterizadas por la degeneración primaria de los conos seguida a menudo de la degeneración del otro tipo de fotorreceptores, los bastones. Su incidencia estimada es de 1:20000-100000 individuos (Hamel, 2007; Roosing *et al.*, 2014). Estas DHR están caracterizadas por la pérdida de la agudeza y el campo visual, alteraciones en la visión de los colores y un grado variable de nistagmus y fotofobia. La respuesta de los conos es ausente o está severamente afectada en los electrorretinogramas (ERG). Además, los pacientes con DCB desarrollan posteriormente alteraciones en los bastones que llevan a que los pacientes presenten ceguera nocturna en el proceso de su enfermedad.

Las enfermedades con afectación de los conos se suelen dividir en estacionarias (Aboshiha *et al.*, 2016) o progresivas (Michaelides *et al.*, 2006). Las formas estacionarias se conocen como síndromes de disfunción de conos, y son congénitas o de comienzo precoz, mientras que las progresivas presentan una edad de inicio más tardía y normalmente los bastones también están implicados. Dentro de las formas estacionarias se incluyen la acromatopsia y el monocromatismo de conos azules, y las formas progresivas son las distrofias de conos y conos-bastones.

2.2. Enfermedad de Stargardt

La enfermedad de Stargardt (STGD) se conoce como la distrofia macular hereditaria juvenil más común, con una prevalencia de 1:10000 personas (Blacharski, 1988). Recientemente, un estudio basado en datos genéticos poblacionales, la sitúa como la DHR más frecuente, con una prevalencia de 1:6578 individuos (Hanany, Rivolta and Sharon, 2020). Esta enfermedad tiene un modelo autosómico recesivo de herencia y está causada por variantes en un único gen, *ABCA4* (STGD1; MIM 248200) (Allikmets *et al.*, 1997).

Los pacientes con STGD1 comúnmente presentan pérdida de la agudeza visual, iniciándose los síntomas en las primeras dos décadas de vida. Oftalmológicamente, en el fondo de ojo se observa una atrofia macular característica y lesiones puntiformes blanco-amarillentas a nivel del EPR conocidas como "flecks", diseminadas en el polo posterior y media periferia (Anderson *et al.*, 1995). En la autofluorescencia de fondo se pueden encontrar lesiones hiperfluorescentes en la zona de atrofia e hipofluorescentes en las áreas que presenten manchas o un aspecto de maculopatía en ojo de buey (Rotenstreich, Fishman and Anderson, 2003; Gomes *et al.*, 2009; Chen *et al.*, 2010). En la tomografía de coherencia óptica (OCT) se observa una pérdida de la arquitectura normal de la retina que comienza en el centro de la mácula (Strauss *et al.*, 2016). Los hallazgos del electrorretinograma pueden ser normales o anormales dependiendo del grado de extensión de las lesiones, y se pueden clasificar en 3 subtipos en función de si presentan respuestas normales de conos y bastones (grupo 1), con afectación de conos (grupo 2) o con afectación también de bastones (grupo 3) (Lois *et al.*, 2001).

El curso de la enfermedad es muy variable, y en la actualidad no existen tratamientos aprobados para esta enfermedad.

3. Aspectos genéticos de las distrofias hereditarias de la retina

Hasta finales de 2019, se habían reportado mutaciones en 32 genes causantes de DC o DCB (Gill *et al.*, 2019). Seis genes causan predominantemente DC y 22 genes DCB, sin embargo, también existe solapamiento entre los fenotipos asociados a estos genes. Las proteínas codificadas por estos genes están implicadas en diversas funciones de los fotorreceptores, incluyendo la fototransducción, la morfogénesis del segmento externo, el transporte intraflagelar y la liberación de neurotransmisores (Tabla 1). En cuanto a la enfermedad de Stargardt, asociada principalmente al gen *ABCA4*, se han descrito fenotipos similares asociados a otros genes: formas dominantes severas para *ELOVL4* y formas benignas de fundus flavimaculatus para *PLA2G5* (Tabla 2).

DHR Herencia	Gen	Proteína	Función	Otras DHR

Tabla 1. Resumen de los genes asociados a enfermedad de Stargardt y Distrofias de conos y bastones.
Modificada de Gill et al., 2019

DHR Herencia	Gen	Proteína	Función	Otras DHR
Enfermedad de Stargardt				
AR	*ABCA4*	*ATP-binding cassette subfamily A member 4R*	Ciclo retinoides	DC, DCB, MD, RP
Distrofia de conos progresiva (DC)				
AR	*CACNA2D4*	*Voltage-dependent calcium channel alpha-2/delta-4*	Liberación del neurotransmisor	
	CNGB3	*Cyclic nucleotide-gated channel beta-3*	Fototransducción	ACHM
	PDE6C	*Cone-specific phosphodiesterase alpha subunit*	Fototransducción	ACHM
	PDE6H	*Cone-specific phosphodiesterase gamma subunit*	Fototransducción	ACHM
XL	*OPN1LW*	*Long-wave-sensitive opsin 1*	Fototransducción	BCM
	OPN1MW	*Medium-wave-sensitive opsin 1*	Fototransducción	BCM
Distrofia de conos y bastones progresiva (DCB)				
AD	*AIPL1*	*Aryl-hydrocarbon-interacting protein-like 1*	Desarrollo tisular	LCA, RP
	GUCY2D	*Guanylate cyclase 2D*	Recuperación del fotorreceptor	LC
	PITPNM3	*Membrane-associated phosphatidylinositol transfer protein 3*	Señalización tirosina-kinasa	
	PRPH2	*Peripherin 2*	Morfogénesis del segmento externo	LCA, MD, RP
	RIMS1	*Protein regulating synaptic membrane exocytosis 1*	Liberación del neurotransmisor	RP
	UNC119	*Human homologue of C.elegans UNC119 protein*	Liberación del neurotransmisor	
AR	*ADAM9*	*A disintegrin and metalloproteinase domain 9*	Unión segmento externo-EPR	
	C21ORF2	*Chromosome 21 open reading frame 2*	Ciliogénesis	
	C8ORF37	*Chromosome eight open reading frame 37*	Desconocida	RP
	CDHR1	*Cadherin-related family member 1*	Morfogénesis del segmento externo	RP
	CEP78	*Centrosomal protein 78-kD*	Ciliogénesis	
	CERKL	*Ceramide kinase-like*	Supervivencia de fotorreceptores	
	KCNV2	*Potassium voltage-gated channel subfamily V 2*	Desconocida	
	POC1B	*Proteome of the centriole 1B*	Transporte intraflagelar	
	RAB28	*Ras-associated protein 28*	Transporte intraflagelar	
	RPGRIP1	*Retinitis pigmentosa GTPase regulator protein 1*	Tráfico intracelular	LCA
	SEMA4A	*Semaphorin 4A*	Desarrollo tisular	RP
	TTLL5	*Tubulin tyrosine ligase-like family member 5*	Señalización del receptor esteroide	
AR/AD	*CACNA1F*	*Voltage-dependent calcium channel alpha-1F*	Liberación del neurotransmisor	
	PROM1	*Prominin 1*	Morfogénesis del segmento externo	MD, RP
	RAX2	*Retina and anterior neural fold homeobox 2*	Desarrollo tisular	
	CRX	*Cone-rod homeobox-containing gene*	Desarrollo tisular	
Distrofia de conos y distrofia de conos y bastones (DC/DCB)				
AD	*GUCA1A*	*Guanylate cyclase activator 1A*	Recuperación del fotorreceptor	
AR	*CNGA3*	*Cyclic nucleotide-gated channel alpha-3*	Fototransducción	ACHM
XL	*RPGR*	*Retinitis pigmentosa GTPase regulator*	Transporte intraflagelar	MD, RP

Abreviaturas: DHR: distrofia hereditaria de retina; AD: autosómico dominante; AR: autosómico recesivo; XL: ligado al cromosoma X; ACHM: acromatopsia; LCA: Amaurosis congénita de Leber; BCM: monocromatismo de conos azules; RP: retinosis pigmentaria; MD: distrofia macular; EPR: epitelio pigmentario de retina.

DHR Herencia	Gen	Proteína	Función	Otras DHR
Tabla 2. Otros genes asociados a Stargardt-like y Fundus flavimaculatus.				
Fenotipos Stargardt-like (STGD-like)				
AD	*ELOVL4*	*Elongation of very long chain fatty acids 4*	Desconocida	
	PRPH2	*Peripherin 2*	Morfogénesis del segmento externo	LCA, MD, RP
AD/AR	*BEST1*	*Bestrophin*	Ciclo retinoides	Best
	PROM1	*Prominin 1*	Morfogénesis del segmento externo	MD, RP
Fundus flavimaculatus/ *Flecks* benignos				
AR	*PLA2G5*	*Phospholipase A2 group V*	Desconocida	
Abreviaturas: DHR: distrofia hereditaria de retina; AD: autosómico dominante; AR: autosómico recesivo; LCA: Amaurosis congénita de Leber; RP: retinosis pigmentaria; MD: distrofia macular. En negrita los genes presentes en Tabla 1.				

3.1. El gen *PROM1*

El gen *PROM1* se localiza en el cromosoma 4, comprende 108 Kb del genoma y tiene 27 exones. Codifica la proteína Prominin 1 o PROM1 (también llamada PROML1, AC133 y CD133), altamente conservada en el reino animal, y conocida por utilizarse como un antígeno de superficie celular en el aislamiento de células madre y células progenitoras en varios tejidos y órganos, principalmente en los sistemas neural y hematopoyético (Fargeas *et al.*, 2006). Además de expresarse en esos tejidos, también los fotorreceptores, las células de la glía y las células epiteliales expresan PROM1 (Weigmann *et al.*, 1997; Corbeil *et al.*, 2000; Florek *et al.*, 2005). Sin embargo, la pérdida de la funcionalidad de Prominin 1 debido a mutaciones en el gen *PROM1* solamente se ha observado en el ciclo visual, por lo que la mayoría de los estudios sobre su función se han centrado en el contexto de la retina.

La proteína PROM1 es una glicoproteína localizada en las protrusiones de la membrana plasmática, con una estructura consistente en un dominio N terminal, 5 dominios transmembrana separados por 2 largos dominios extracelulares glicosilados y 2 dominios intracelulares, y un dominio intracelular C terminal (Figura I5A). Se han reportado siete diferentes isoformas de la proteína PROM1 en tejidos humanos, aunque los exones afectados por el *splicing* alternativo solo afectan a los dominios N y C terminal (Fargeas, Huttner and Corbeil, 2007).

En la retina, PROM1 se localiza en la base de la membrana de los segmentos externos de los fotorreceptores, y está implicado en la morfogénesis de los discos presentes en los segmentos externos de los fotorreceptores conos y bastones (Yang *et al.*, 2008) (Figura I5B). El primer estudio en el que se relaciona el gen *PROM1* con una enfermedad es en una familia con enfermedad de Stargardt autosómica dominante (Kniazeva *et al.*, 1999). En el año 2000 se describe una mutación *frameshift* en homocigosis en el gen *PROM1* en una familia con degeneración retiniana, y se observa que la proteína mutada no alcanza la superficie celular (Maw

et al., 2000). En 2008 Yang y colaboradores identifican una mutación en heterocigosis en *PROM1* en la familia con herencia dominante publicada por Zkniazeva et al. y relacionan las mutaciones en este gen con una morfogénesis de los discos de los fotorreceptores defectuosa (Yang *et al.*, 2008). Por esta razón, las mutaciones en el gen *PROM1* se asocian a una forma dominante de enfermedad de Stargardt (STGD4; MIM 603786). Sin embargo, hasta el desarrollo de esta Tesis Doctoral (2018) se habían reportado alrededor de 70 variantes en el gen *PROM1* (*Human Gene Mutation Database, HGMD)* asociadas a otros tipos de DHR, incluyendo RP autosómica recesiva con degeneración macular, DCB autosómica recesiva y dominante, y maculopatía en ojo de buey (Maw *et al.*, 2000; Michaelides *et al.*, 2003; Yang *et al.*, 2008; X. Zhang *et al.*, 2014) .
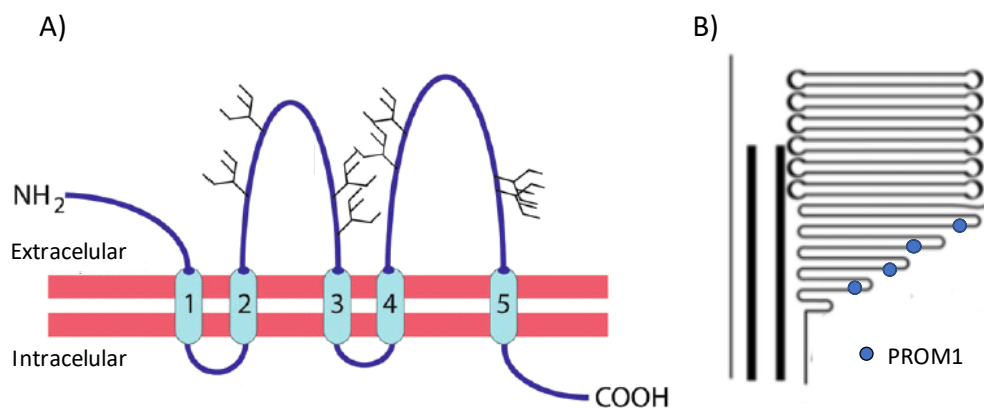
Figura I5. *Estructura y localización de PROM1. A) Estructura. B) Localización de PROM1 en los bordes de los discos externos nacientes. Su función se relaciona con el mantenimiento de la curvatura de la membrana de los discos nacientes (Goldberg et al., 2016). Imagen tomada de* Corbeil et al, 2013 *(A).*

3.2. Gen *ABCA4*: función y modelo de correlación genotipo-fenotipo

La primera vez que el gen *ABCA4* se asoció a una DHR concreta fue con la enfermedad de Stargardt (Allikmets *et al.*, 1997). Actualmente, se considera como el gen más mutado en DHR (Hanany, Rivolta and Sharon, 2020). Este gen se localiza en el cromosoma 1, comprende 128 Kb del genoma y tiene 50 exones. Codifica la proteína ABCA4, que forma parte de la superfamilia de transportadores ABC (ATP-binding cassette), encargadas de transportar compuestos a través de las membranas, desde la membrana intradiscal al citoplasma de los fotorreceptores celulares. ABCA4 consiste en un polipéptido de 2273 aminoácidos que contiene 2 dominios transmembrana (TMD) para la unión del sustrato, dos dominios citoplasmáticos que contienen los dominios de unión a nucleótidos (NBD), que proporcionan energía para el transporte de los sustratos mediante la unión e hidrólisis de ATP, y dos dominios extracelulares que presentan residuos glicosilados (ECD) (Figura I6A).
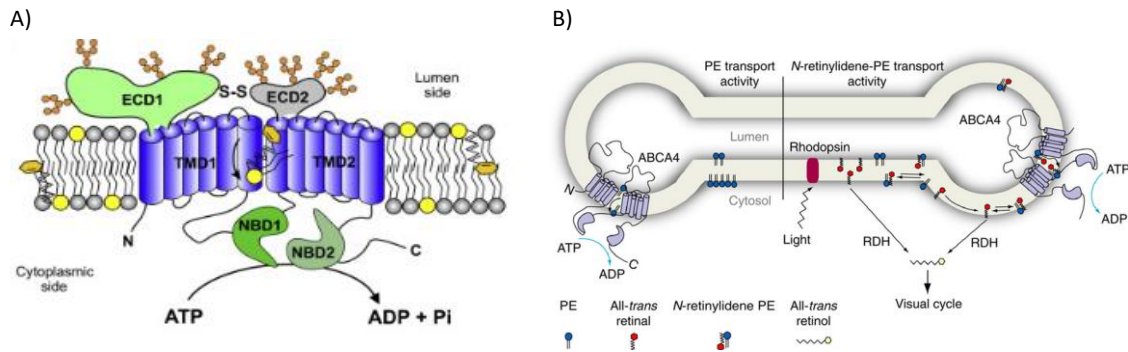
Figura I6. *Estructura y función de ABCA4. A) Estructura. B) Función. Imágenes tomadas de* Molday, 2015 *y* Quazi et al., 2012.

Esta proteína se localiza en los bordes de la membrana de los discos de los segmentos externos de los fotorreceptores (Papermaster *et al.*, 1978; Papermaster, Reilly and Schneider, 1982; Molday, Rabin and Molday, 2000), y actúa como una flipasa de compuestos retinoides (Tsybovsky, Molday and Palczewski, 2010; Molday, 2015). De hecho, es la única proteína de la familia ABC que actúa como importadora (Quazi, Lenevich and Molday, 2012) (Figura I6B):

- Por un lado, cuando se produce la fotoexcitación, el compuesto all-trans retinal reacciona con la fosfatidil etanolamina (PE), formando el all-trans-N-ret-PE (N-RPE), que es transportado por ABCA4 en la membrana de los discos, desde el lado del lumen hacia el del citoplasma, donde se hidroliza a PE y all-trans-retinal, y éste último se reduce a all-trans-retinol como parte del ciclo visual.

- Por otro lado, ABCA4 también interviene en la eliminación del exceso de 11-cis retinal que no es necesario para la regeneración de las opsinas. Este compuesto reacciona con la PE y será transportado por ABCA4 hacia el lado citoplasmático de la membrana de los discos externos, donde se isomeriza a N-RPE, que vuelve al ciclo visual.

El gen *ABCA4* se expresa mayoritariamente en los fotorreceptores, aunque también se ha descrito una expresión limitada en queratinocitos (Braun *et al.*, 2013), fibroblastos (Aukrust *et al.*, 2017) y EPR (Lenis *et al.*, 2018). Las mutaciones bialélicas en *ABCA4* producen una pérdida de función de la proteína ABCA4 o una mislocalización de ésta por alteraciones en el plegamiento. Esto provoca el fallo en el transporte del N-RPE en la membrana de los discos hacia la superficie citoplasmática para su disociación y posterior entrada del all-trans-retinal al ciclo visual. El compuesto N-RPE se acumula en los discos externos y al reaccionar con una segunda molécula de N-RPE, forma el compuesto A2PE, que es un derivado bisretinoide que al ser hidrolizado hasta A2E durante la fagocitosis de los segmentos externos de los fotorreceptores por las células del EPR, se acumula en estas células formando depósitos de lipofuscina. Estos compuestos bisretinoides no pueden ser digeridos por el EPR provocando la muerte de estas células, y por consiguiente, la muerte de los fotorreceptores (Molday, 2007).

Mutaciones en ambas copias del gen *ABCA4* se han asociado a un amplio espectro de enfermedades. La enfermedad de Stargardt es la más frecuente pero también se han descrito mutaciones en este gen en pacientes con DCB, RP, fundus flavimaculatus y coriorretinopatías (Allikmets *et al.*, 1997; Cremers *et al.*, 1998; Martínez-Mir *et al.*, 1998; Klevering *et al.*, 2005; Bertelsen *et al.*, 2014; Tanaka *et al.*, 2018).

Desde 1998 existe un modelo de correlación genotipo-fenotipo que relaciona el tipo de mutaciones en el gen *ABCA4* y la severidad del fenotipo, que está relacionada a su vez con la función residual de la proteína ABCA4 (van Driel *et al.*, 1998; Maugeri *et al.*, 2000). Actualmente se considera que una combinación de variantes leves con baja penetrancia (*mild*) y severas (*severe*) daría lugar a STGD1 de inicio tardío (*late-onset* STGD1), variantes leves y severas o dos variantes moderadas darían un fenotipo de STGD1, mientras que una combinación de variantes severas y moderadas o dos variantes severas darían lugar a DCB (Sangermano, 2018) o STGD1 de inicio temprano (Figura I7). El fenotipo de STGD1 de inicio tardío es un fenotipo leve con un mejor pronóstico y asociado a respeto foveal, que hace que los pacientes presenten una buena agudeza visual a edades tardías (Rotenstreich, Fishman and Anderson, 2003; Fujinami *et al.*, 2011; Westeneng-van Haaften *et al.*, 2012; Fujinami *et al.*, 2013).

Fenotipo esperado	Normal	Normal	STGD1 inicio tardío	STGD1	DCB	DCB severa
Combinación variantes en función proteína	normal/normal	severa/normal mod./normal leve/normal	severa/leve	severa/leve mod./mod.	severa/mod.	severa/severa
Actividad ABCA4						
Casos STGD1			~20%	~60%	~15%	~5%
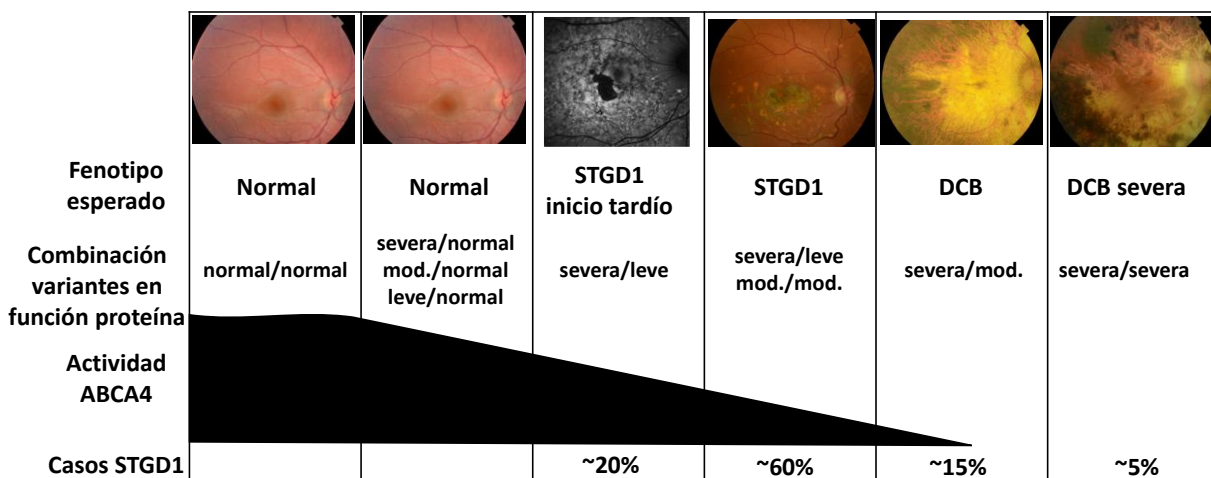
Figura I7. *Modelo de correlación genotipo-fenotipo para ABCA4. Las variantes se clasifican en leves, moderadas (mod.) o severas en función de la cantidad producida de tránscrito normal o wild type, y, por tanto, de la cantidad de proteína* ABCA4 *funcional. Elaboración propia basada en Sangermano, 2018.*

3.3. Heredabilidad perdida en la enfermedad de Stargardt

Se estima que en el 25-30% de los pacientes con diagnóstico de Stargardt, en los que se secuencian las regiones codificantes del gen *ABCA4,* no se identifican los dos alelos del gen mutados, sino que en la mayoría se identifica un único alelo patogénico (Zernant *et al.*, 2011). El elevado número de pacientes STGD monoalélicos para *ABCA4* se ha descrito en varios estudios en diferentes poblaciones (Zernant *et al.*, 2011; Chacón-Camacho *et al.*, 2013; Riveiro-Alvarez *et al.*, 2013; Xin *et al.*, 2015; Zaneveld *et al.*, 2015). Algunas de las razones por las que se podría explicar son:

1) Otros genes causantes de un fenotipo STGD-*like*

Se han descrito mutaciones en los genes *PRPH2*, *ELOVL4* y *PROM1* en casos con fenotipos semejantes a STGD1 (Stone *et al.*, 1994; Kniazeva *et al.*, 1999; Boon *et al.*, 2007), que se conocen como fenotipos STGD-*like* ya que no están causados por el gen *ABCA4*. Además, estudios mediante técnicas de nueva generación (NGS) en cohortes de pacientes no resueltos han identificado mutaciones en *BEST1, CDH23, CRB1* y *USH2A* (Zaneveld *et al.*, 2015). Hay que tener en cuenta que debido a la gran variabilidad clínica que presenta el gen *ABCA4*, el diagnóstico correcto de los pacientes con enfermedad de Stargardt es complicado, y en ocasiones se identifican genes asociados a otras DHR, como sería el caso de *USH2A*.

2) Variantes con penetrancia reducida

Actualmente, se han descrito más de 1200 variantes en el gen *ABCA4* asociadas a DHR (HGMD y *Leiden Open Variation Database*, LOVD, Fecha de acceso: julio 2020). Las variantes con una frecuencia alélica superior al 1% (MAF<0.01) se suelen descartar como la causa de la enfermedad en los análisis de priorización, sin embargo, puede haber casos en los que la frecuencia de algunas variantes sea más elevada en pacientes que en individuos sanos, aun siendo alta en los últimos, y sugiera una causalidad de su enfermedad. En el caso de *ABCA4*, hay ciertas variantes con una frecuencia >5% en la población general. Entre ellas se incluye la variante c.5603A>T (p.Asn1868Ile), cuya frecuencia alélica en la población general es ~6%, pero está presente en los pacientes STGD1 3 o 4 veces más frecuentemente de lo que se esperaría (Zernant *et al.*, 2017). Esta variante era considerada una variante benigna, sin embargo, estudios recientes la consideran una variante patogénica con penetrancia reducida (Zernant *et al.*, 2017; Cremers *et al.*, 2018; Runhart *et al.*, 2018). Se calcula que esta variante causaría la enfermedad en una reducida proporción de pacientes (~5%) portadores de esta variante y una variante severa en *ABCA4*. Otras de las variantes consideradas leves o "*mild*" en este gen son la c.3113C>T; p.(Ala1038Val) y la c.5882G>A; p.(Gly1961Glu) ya que los pacientes homocigotos para estas variantes aparecen en un porcentaje mucho menor de lo que se esperaría según su frecuencia en la población general (Cornelis *et al.*, 2017). Todavía no se conoce el mecanismo por el que estas variantes podrían

actuar como variantes de baja penetrancia, se propone que elementos *cis* y *trans* reguladores podrían actuar como modificadores genéticos.

3) Variantes en regiones no codificantes

Los pacientes STGD en los que no se identifican variantes en la región codificante del gen *ABCA4* podrían tener variantes fuera de estas regiones, que en total representan el 94% del tamaño del gen y no se incluyen normalmente para su estudio en las técnicas convencionales y NGS. El estudio del papel de variantes que se encuentran en la región intrónica de *ABCA4* está llevándose a cabo desde hace pocos años, siendo Braun y colaboradores en el año 2013 los primeros autores que identificaron e hicieron estudios funcionales en variantes no codificantes en *ABCA4* para demostrar su implicación en STGD1 (Braun *et al.*, 2013). En este estudio, se realizó el haplotipo de 28 pacientes monoalélicos para *ABCA4* y se secuenció ARN procedente de retina humana para identificar secuencias intrónicas que pudieran tener un papel en el *splicing* del gen *ABCA4*. De este modo, identificaron exones alternativos minoritarios en el 5% de los tránscritos de *ABCA4*, correspondientes a una fracción menor del 1% del ARNm total. El análisis de estas regiones en las que se identificaron exones alternativos en los casos STGD no resueltos permitió la identificación de 5 variantes *deep intronic* (c.5196+1137G>A, c.5196+1216C>A, c.5196+1056A>G, c.4539+2001G>A y c.4539+2028C>T) localizadas en sitios alternativos de *splicing*, y que, además, se encontraban en *trans* con el alelo previamente identificado en los pacientes. El estudio de ARN de queratinocitos derivados de pacientes portadores de 3 de estas variantes demostró la incorporación de pseudoexones. Estudios posteriores han demostrado la existencia de nuevas variantes *deep intronic* en pacientes STGD1 previamente no resueltos, enriquecidas en estos pacientes respecto de controles sanos, y en algunos casos, testadas funcionalmente antes y durante el desarrollo de esta Tesis Doctoral (Zernant *et al.*, 2014; Bax *et al.*, 2015; Bauwens *et al.*, 2019; Fadaie *et al.*, 2019; Khan *et al.*, 2019; Sangermano *et al.*, 2019; Khan *et al.*, 2020).

3.4. Estudios funcionales para la identificación de defectos en el *splicing* del gen *ABCA4*

3.4.1. Mecanismo de *splicing*

La expresión génica conlleva etapas altamente reguladas que parten de la transcripción del ADN, que genera un ARNm que, cuando es procesado, sirve como molde para ser traducido a proteína. El *splicing* es el proceso bioquímico en el que se eliminan los intrones y se empalman los exones del tránscrito de ARNm prematuro. Este proceso está altamente regulado ya que se reconocen secuencias consenso características para la identificación de exones e intrones, y es llevado a cabo por el spliceosoma, un complejo macromolecular compuesto por 5 partículas ribonucleoproteicas (U1, U2, U4, U5 y U6 snRNPs – uridine-rich small nuclear ribonucleiprotein) en el centro catalítico, y más de 200 proteínas adicionales.

El correcto proceso de *splicing* requiere tanto elementos reguladores en *cis*, localizados en la secuencia del ARN prematuro, como factores en *trans*, proteínas capaces de unirse a estos elementos (Singh and Cooper, 2012). El reconocimiento de los intrones para su eliminación se produce mediante unas secuencias altamente conservadas evolutivamente localizadas en los extremos 5' y 3' de los exones. El sitio de *splicing* 5' se conoce como sitio donador, y está formado por los dos primeros nucleótidos del intrón (GU). El sitio 3' de *splicing* es el sitio aceptor y está formado por la secuencia *branch point*, el tracto de polipirimidinas y los dos nucleótidos del sitio 3' (AG) (Figura I8). Tras el reconocimiento de estos sitios por la maquinaria del spliceosoma se forman varios complejos que finalmente dan lugar al tránscrito maduro de ARNm.
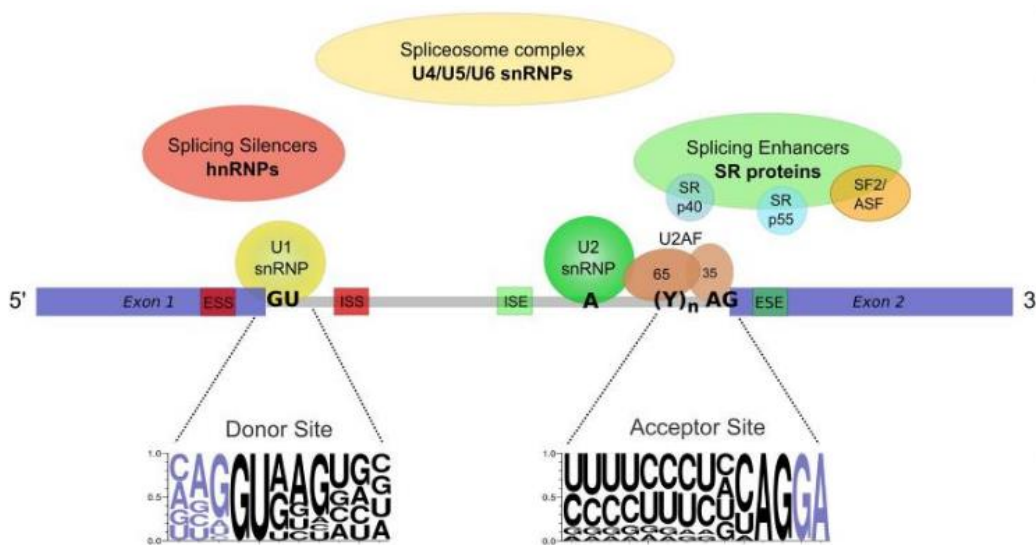
Figura I8. *Elementos implicados en la regulación del proceso de splicing*. *Imagen de* Dubois, et al., 2014.

El *splicing* alternativo consiste en que a partir de una molécula de ARNm prematuro se generan diferentes ARNm maduros, llamadas isoformas, que darán lugar a proteínas diferentes codificadas por el mismo gen. Se estima que más del 90% de los genes humanos son regulados por *splicing* alternativo (Pan *et al.*, 2008; Wang *et al.*, 2008) y alrededor de un tercio de las mutaciones asociadas a enfermedades hereditarias han sido relacionadas con la alteración del proceso de *splicing* (Lim *et al.*, 2011; Sterne-Weiler *et al.*, 2011). Los exones que finalmente formarán parte del ARNm maduro en el proceso de *splicing* alternativo están definidos por la interacción de los elementos en *cis* y los factores en *trans*. Los elementos en *cis* incluyen los ESEs (*exonic splicing enhancers*) y los ISEs (*intronic splicing enhancers*) que se unen a proteínas en *trans* que favorecen el *splicing* (proteínas SR), y los ESSs (*exonic splicing silencers*) e ISSs (*intronic splicing silencers*) que se unen a proteínas que inhiben el *splicing* (hnRNPs). Las interacciones entre estos elementos resultan en potenciar o inhibir la maquinaria del spliceosoma en los sitios de *splicing* débiles, respectivamente (Wang and Burge, 2008; Wang *et al.*, 2008).

Por último, el *splicing* aberrante sucede como consecuencia de variaciones en las secuencias que alteran el correcto procesamiento del ARNm prematuro, y es una causa frecuente de enfermedades hereditarias. La eliminación de exones parcial o total (*skipping*) y la incorporación de secuencias intrónicas en el ARNm suelen ser causadas por mutaciones en los sitios reguladores de *splicing* (Havens, Duelli and Hastings, 2013; Scotti and Swanson, 2016).

3.4.2. Estudios *in vitro* en el gen *ABCA4*

Estudios funcionales en variantes del gen *ABCA4* han sido reportados desde poco después de su identificación como gen asociado a STGD1. El estudio de variantes *missense* y variantes localizadas en sitios no canónicos de *splicing* (NCSS) (del nucleótido +3 a +10 y -14 a -3 en los intrones, y en los primeros o últimos 2 nucleótidos de los exones) utilizando ARNm de pacientes, o minigenes en los que se clonan fragmentos de ADN con las mutaciones de los pacientes, han permitido demostrar patrones de *splicing* aberrante (Maugeri *et al.*, 1999; Rivera *et al.*, 2000). Debido a que *ABCA4* se expresa en muy bajo nivel en tejidos somáticos y no es posible obtener tejido de la retina de los pacientes, también se han utilizado células precursoras de fotorreceptores inducidas a partir de fibroblastos de pacientes portadores de las variantes de estudio (Albert. *et al.*, 2015; Sangermano *et al.*, 2016). Sin embargo, esta estrategia supone una gran inversión en tiempo y económica, por lo que no es posible utilizarla para estudiar funcionalmente un gran número de variantes, como sería el caso del gen *ABCA4*. Como se mencionaba anteriormente, los estudios que se realizan para estudiar el *splicing* consisten en introducir fragmentos de ADN con las variantes de estudio en vectores, clonar estas construcciones en vectores de expresión y posteriormente transfectar células (normalmente células HEK293T) para estudiar el patrón de *splicing* aberrante en el ARNm. La limitación principal de estos estudios es el tamaño del vector, que condiciona la longitud del fragmento de ADN a clonar, y, por lo tanto, el número de exones adyacentes a la variante a testar. Esta limitación para el estudio del *splicing* es muy importante ya que hay muchos elementos en el contexto genómico que influencian el reconocimiento de los sitios de *splicing*. Por esta razón, Sangermano y colaboradores (Sangermano *et al.*, 2018) diseñaron un protocolo para preparar "midigenes" largos, en los que se clonan fragmentos de ADN genómico del gen *ABCA4* de hasta 12 Kb, que posibilitan el estudio de variantes en sitios no canónicos de *splicing* y *deep intronic* con un adecuado contexto genómico (Sangermano *et al.*, 2018). Gracias a esta estrategia se ha demostrado el efecto patogénico de variantes no canónicas de *splicing*, así como de variantes *deep intronic* nuevas y previamente descritas, también durante el desarrollo de esta Tesis Doctoral (Sangermano *et al.*, 2018; Fadaie *et al.*, 2019; Khan *et al.*, 2019, 2020).

4. Aproximaciones terapéuticas en DHR: enfermedad de Stargardt.

Principalmente, el manejo clínico de estas enfermedades consiste en aliviar los síntomas. Únicamente existe un tratamiento aprobado para tratar una DHR causada por mutaciones en el gen *RPE65*, Luxturna™ (https://luxturna.com). La caracterización molecular es imprescindible para el preciso diagnóstico de estas enfermedades, el asesoramiento genético y pronóstico de la enfermedad, así como la posible participación en los ensayos clínicos que se están desarrollando. Entre los ensayos clínicos en desarrollo en DHR, existen varios que involucran a la enfermedad de Stargardt y el gen *ABCA4*, debido a que es una enfermedad monogénica, muy frecuente, y se ha estudiado ampliamente la fisiopatología de las células que sufren degeneración (fotorreceptores y células del EPR). Por ello, y siendo esta enfermedad una de las principales DHR estudiadas en esta Tesis Doctoral, las aproximaciones terapéuticas que se describen aquí están enfocadas en el tratamiento de la enfermedad de Stargardt y no en las DHR en general.

En primer lugar, los pacientes con STGD1 pueden utilizar lentes de baja visión y otras tecnologías de asistencia para optimizar su visión, además de aconsejarles llevar una dieta baja en vitamina A y reducir su exposición a la luz UV para ralentizar la progresión de esta enfermedad.

Existen **terapias farmacológicas** para estos pacientes, principalmente dirigidas a actuar en el ciclo visual para evitar el acúmulo de los derivados bisretinoides que provocan los acúmulos de lipofuscina en el EPR. Fármacos como el soraprazan, emixuxtat, ALK-001, STG-001, fenretinide y A1120 son moduladores del ciclo visual. ALK-001 es una forma deuterada de la vitamina A que hace que disminuya su tasa de dimerización y no se producen elementos tóxicos, a diferencia de la forma natural (Issa *et al.*, 2015) (NTC02402660). STG-001, fenretinide y A1120 actúan como competidores de la proteína de unión RBP haciendo que se reduzca la cantidad de retinol en el ojo (Rahman *et al.*, 2019). El fármaco emixustat es un inhibidor de la enzima RPE65 que participa en la regeneración del 11-cis-retinal y su acción es reducir los compuestos tóxicos y la tasa metabólica de los fotorreceptores (NTC03772665 y NTC03033108).

La **terapia génica** también se ha desarrollado en enfermedades del ojo, ya que es un órgano anatómicamente accesible, pequeño y cerrado, a la vez que se considera inmunoprivilegiado por la existencia de barreras mecánicas que crean un microambiente ocular en el que se inhiben las respuestas inflamatorias (Kumaran *et al.*, 2018). Además, con una pequeña cantidad de vector es suficiente, y se reducen los efectos secundarios tóxicos relacionados con el vector utilizado (Garoon and Stout, 2016).

En la **terapia génica de sustitución**, el objetivo en los pacientes STGD1 es introducir un gen *ABCA4* funcional en la retina que incremente los niveles de la proteína normal en los fotorreceptores. Estudios preclínicos de terapia de aumento génico (*gene augmentation*) en ratones *abca4*-/- han demostrado mejoría en su fenotipo (Allocca *et al.*, 2008; Han *et al.*, 2012)

animando el desarrollo de estudios también en humanos. Los vectores basados en virus adenoasociados (AVVs) son los más utilizados en terapia génica, y en DHR ha sido demostrada su seguridad y efectividad (Colella and Auricchio, 2012). La limitación de estos vectores es su capacidad máxima de 4.7 Kb, que imposibilita introducir genes grandes como *ABCA4*. Sin embargo, se han desarrollado estrategias en las que el gen se divide en 2 y se empaqueta en 2 diferentes AVVs que se ensamblan después mediante recombinación homóloga y/o el mecanismo de *trans-splicing* (Trapani *et al.*, 2014). Los vectores lentivirales también se han desarrollado para su uso en genes grandes, y en el caso del gen *ABCA4*, el vector StarGen™ dio resultados positivos en estudios con animales, y actualmente hay ensayos clínicos en fases tempranas con pacientes STGD1 (NCT01367444 y NCT01736592).

Otra aproximación de la terapia génica es el uso de **oligonucleótidos antisentido** (AON) que consiste en pequeñas moléculas (15-30 nucleótidos) complementarias a una secuencia de ARNm prematuro (diana) que codifica para una proteína. Esta unión va a interferir en el proceso de *splicing*, por lo que se suelen diseñar dirigidos a mutaciones que alteran este mecanismo. Actualmente, los ensayos clínicos en DHR utilizando AON se están desarrollando para el gen *CEP290* (NCT3140969). En el caso de STGD1, al identificarse cada vez más mutaciones *deep intronic* en el gen *ABCA4*, se han diseñado AON para mutaciones específicas (Albert. *et al.*, 2017; Garanto *et al.*, 2019), sin embargo estos estudios se encuentran en fase preclínica.

La **edición del ADN genómico** ha adquirido un enorme protagonismo en los últimos años. Actualmente, hay dos enfoques principales, un enfoque *in vivo*, en el que las mutaciones que causan la enfermedad se corrigen directamente en las células de la retina mediante la administración de AAVs con el material de edición génica necesario, y el enfoque *ex vivo*, en el que se corrige la mutación en células reprogramadas derivadas del paciente (iPSC), que después se diferencian en células de retina y se trasplantan en la retina del paciente (Sanjurjo-Soriano and Kalatzis, 2018). Actualmente, la edición génica se realiza utilizando el sistema CRISPR/Cas9 (*Clustered Regularly Interspaced Short Palindromic Repeats*), tanto en estudios *in vivo* (Ruan *et al.*, 2017; Yu *et al.*, 2017), como *in vitro* (Burnight *et al.*, 2017; Fuster-García *et al.*, 2017), aunque también hay estudios previos utilizando las nucleasas *zinc finger*, ZFN (Greenwald, Cashman and Kumar-Singh, 2010; Overlack *et al.*, 2012), y *transcription activator-like effector*, TALENs (Low *et al.*, 2013). El primer ensayo que utiliza CRISPR/Cas9 para eliminar *in vivo* una mutación en un paciente se ha llevado a cabo el pasado mes de marzo, en un paciente con LCA y mutaciones en el gen *CEP290* ('First CRISPR therapy dosed', 2020; Hampton, 2020), lo que supone que la edición génica puede ser una aproximación prometedora para que se desarrollen ensayos clínicos en genes asociados a DHR.

Finalmente, otra de las aproximaciones terapéuticas utilizadas en DHR es la **terapia celular (con células madre)**, que, en la enfermedad de Stargardt, consiste en regenerar las células del EPR para evitar la pérdida de visión de los pacientes. Las células madre derivadas de células embrionarias humanas (hESC) son pluripotentes y se pueden diferenciar a muchos tipos celulares (Odorico, Kaufman and Thomson, 2001). La utilización de estas células para remplazar el EPR, que contiene células que se encargan de transportar nutrientes, eliminar los productos bisrretinoides y reciclar el compuesto 11-cis-retinal ha mostrado resultados positivos en estudios preclínicos con ratones (Lu *et al.*, 2009). Por esta razón, se desarrollaron dos estudios clínicos prospectivos para establecer la seguridad y tolerancia del trasplante subretiniano de estas células diferenciadas a EPR en pacientes STGD1 (NCT01345006 y NCT01469832), y sus resultados mostraron que no surgieron efectos adversos derivados de la terapia de células madre, pero si en la cirugía e inmunosupresión de algunos pacientes (Schwartz *et al.*, 2016). Además, la visión de 3 pacientes mejoró y en el 72% de los pacientes se habían incrementado los niveles de pigmento subretiniano en las zonas atróficas del EPR, consistente con la eficiencia del trasplante de estas células (Schwartz *et al.*, 2016). Aun así, se necesitan más investigaciones ya que el tamaño de muestra era pequeño y no había grupo control. Otro estudio clínico actualmente en desarrollo está enfocado a evaluar este tratamiento a largo plazo (NCT02445612). Estudios futuros utilizando esta estrategia serán útiles para evaluar su potencial en el tratamiento de enfermedades con degeneración retiniana.

Hipótesis y objetivos

Hipótesis

La enfermedad de Stargardt es la maculopatía hereditaria más frecuente y se considera una distrofia de retina genéticamente homogénea, ya que está causada por mutaciones en el gen *ABCA4*. Sin embargo, no todos los pacientes diagnosticados de Stargardt son caracterizados con mutaciones bialélicas en este gen, en algunos de ellos se identifica una variante y en otros ninguna. Profundizar en el estudio del gen *ABCA4*, y de otros genes también asociados a formas similares de esta patología puede ayudar a encontrar la causa genética fuera de las regiones comúnmente estudiadas, así como expandir el espectro fenotípico de los pacientes con estas patologías.

Además, la caracterización molecular de los pacientes con distrofias de retina está permitiendo su inclusión en ensayos clínicos.

Objetivos

Objetivo general

El objetivo principal de esta Tesis Doctoral ha sido identificar la causa molecular de pacientes con distrofias maculares, principalmente pacientes diagnosticados de enfermedad de Stargardt, y describir su fenotipo asociado con el fin de evaluar el curso de su enfermedad.

Objetivos específicos

1. Caracterización genética de pacientes con enfermedad de Stargardt mediante secuenciación del gen *ABCA4*, incluyendo regiones no codificantes, y estudio de variantes en el número de copias (CNVs).
2. Identificación y caracterización funcional de nuevas variantes intrónicas en el gen *ABCA4* en pacientes Stargardt no caracterizados que potencialmente provocan alteraciones en el correcto mecanismo de *splicing* de este gen.
3. Descripción genética y fenotípica de los pacientes con mutaciones en el gen *ABCA4* estableciendo nuevas correlaciones genotipo-fenotipo.
4. Estudio molecular y clínico de pacientes con distrofia macular y mutaciones en el gen *PROM1*, previamente asociado a un fenotipo Stargardt-*like.*

Resultados

Caracterización de pacientes no resueltos y estudio funcional de variantes

Artículo 1: Molecular characterization of 945 families with autosomal recessive maculopathies

Del Pozo-Valero M, Ayuso C, FJD Genetics Department.

Resumen:

En este trabajo se describe la cohorte de familias con distrofias maculares autosómicas recesivas, incluyendo pacientes con sospecha de enfermedad de Stargardt, distrofias de conos y bastones, y otras distrofias maculares.

Estos pacientes se han dividido en dos grupos, según su diagnóstico de sospecha, en: a) pacientes con diagnóstico de enfermedad de Stargardt, y b) pacientes con diagnóstico de otras maculopatías.

El objetivo de este trabajo era evaluar la implementación de nuevas técnicas de secuenciación masiva en el laboratorio para el estudio de estos pacientes.

Al principio de este estudio, en diciembre de 2016, la cohorte estaba formada por 802 pacientes, y al final de este estudio (diciembre de 2019), por 945 pacientes. El 59% de los pacientes (292/493) con diagnóstico de enfermedad de Stargardt estaban caracterizados con el gen *ABCA4*, llegando hasta el 73% (407/553) al final del periodo del estudio, tras la implementación de otras técnicas que permitían la secuenciación de la región codificante completa, junto con zonas intrónicas en otros casos. En el caso de los pacientes con otras maculopatías, al principio del estudio se había caracterizado el 38% de los pacientes (117/309), y al final, este porcentaje fue del 59% (195/392), observando que el número de casos caracterizados con el gen *ABCA4* y otros genes es similar (25% y 24%, respectivamente).

Respecto a los casos monoalélicos para el gen *ABCA4*, éstos representaban el 13% (63/493) y el 4% (12/309) de los grupos a) y b), respectivamente. Se encontró un segundo alelo mutado en este gen en el 73% (46/63) de los casos con sospecha de Stargardt, mientras que solo el 41% (5/12) de los casos monoalélicos con otras maculopatías presentaban un segundo alelo mutado en *ABCA4*.

Estos resultados ponen de relieve la importancia del diagnóstico genético en casos en los que la información clínica es escasa, a su vez que la reevaluación clínica de los pacientes en los que no se identifican variantes en los genes conocidos.

Contribución de la autora:

La autora ha recopilado toda la información sobre el diagnóstico molecular llevado a cabo en las 945 familias con distrofia macular, en los dos momentos del estudio. También ha participado en el análisis estadístico y ha elaborado el manuscrito que se presenta a continuación.

Molecular characterization of 945 families with autosomal recessive maculopathies

ABSTRACT

We describe a cohort of families with autosomal recessive maculopathies, including Stargardt disease, cone-rod dystrophies and other macular dystrophies studied in two periods of time, before and after the implementation of NGS technologies. These patients were divided into two groups, according to their diagnosis, as STGD and Other (non STGD). At the beginning of the study (December 2016), the cohort was comprised of 802 patients, and at the end (December 2019) of 945 patients. 59% (292/493) of STGD patients were characterized with *ABCA4*, leading to 73% (407/553) at the end of the study, because of the implementations of technologies where the *ABCA4* gene and non-coding regions were sequenced. In the group of Other maculopathies, at the beginning of this study, 38% (117/309) of patients were characterized, and at the end this percentage was 59% (195/392), and the number of solved cases with the *ABCA4* gene or other genes was similar (25% and 24%, respectively). Monoallelic *ABCA4* cases at the beginning of the study represent the 13% (63/493) and 4% (12/309) of STGD and non STGD groups. A second pathogenic variant was found in 73% (46/63) of the former and 41% (5/12) of the later groups. With these results, it is established the importance of genetic diagnosis in cases with por clinical information, as well as the clinical reevaluation in patients in which no variants in known genes are found.

INTRODUCTION

The inherited macular dystrophies (MD) comprise a heterogeneous group of disorders characterized by bilateral central visual loss and atrophy of the macula and underlying retinal pigment epithelium (RPE). The hallmark of these diseases is a loss of visual acuity, and they affect people regardless the age. The genetic spectrum of MD is very heterogeneous. All patterns of inheritance have been reported associated with these diseases. The different forms of macular degeneration encompass a wide range of clinical, psychophysical and histological findings (Michaelides, 2003). Different types of MD include Stargardt disease (STGD), Best disease, adult vitelliform macular disease, pattern dystrophy, X linked juvenile retinoschisis, and age-related macular degeneration, among others.

Other group of inherited retinal dystrophies are cone and cone-rod dystrophies (CCRD), commonly related with MD because they also involve progressive degeneration, dysfunction, and vision loss of the central retina due to photoreceptor degeneration (Birtel *et al.*, 2018). Full field electroretinography (ffERG) findings are needed to distinguish among diseases: normal in MD, only reduced cone responses in CD and both cone and rod responses reduced in CRD.

Stargardt disease is the most frequent maculopathy and it is considered as a monogenic disease caused by biallelic mutations in the *ABCA4* gene (STGD1). The *ABCA4* gene has been also associated to other retinal phenotypes, regarding the severity of the variants found in this gene (F. P. M. M. Cremers *et al.*, 1998; Martínez-Mir *et al.*, 1998; Bertelsen *et al.*, 2014; Tanaka *et al.*, 2018; Verbakel *et al.*, 2018). Currently, we know that several non-canonical splice site (NCSS) and deep intronic variants lead to defects in splicing (Sangermano *et al.*, 2018, 2019; Bauwens

et al., 2019; Fadaie *et al.*, 2019; Khan *et al.*, 2019, 2020) which remark the importance of the screening of *ABCA4* coding and non-coding regions in these patients.

Now, with the Next Generation Sequencing technologies (NGS) it is possible to study a high number of genes at the same time, and design gene panels including non-coding regions. However, in the past, the genetic diagnosis of patients with inherited disorders was performed using genotyping microarrays or Sanger sequencing of known mutations or genes.

Here, we report a summary of the results achieved after the implementation of NGS to study 945 autosomal recessive MD and CCRD (hereinafter, arMD) families between 2016 and 2020 at the Genetic Department of the University Hospital Fundacion Jimenez Diaz, in Madrid.

MATERIAL AND METHODS

Patients

A cohort of 945 unrelated families with arMD were recruited at the Fundación Jiménez Díaz University Hospital (FJD, Madrid, Spain) up to December 2019. This study was performed in accordance with the tenets of the Helsinki Declaration, it was approved by the Research Ethics Committee of the Fundación Jimenez Diaz University Hospital. DNA samples were collected from the FJD biobank and all participants signed an Informed consent.

Diagnosis and molecular characterization

The 945 probands were classified into two categories:

- Suspected Stargardt diagnosis (STGD cases): Patients referred with Stargardt disease and an autosomal recessive pattern of inheritance.
- Other maculopathies (non STGD cases): Patients referred with macular dystrophies and an autosomal recessive pattern of inheritance.

The 945 probands were studied by one or more genetic technologies: genotyping microarrays, gene panels, clinical exome, smMIPs and WES (Aguirre-Lamban *et al.*, 2009, 2011; Riveiro-Alvarez *et al.*, 2009, 2013; Corton *et al.*, 2013; Zernant *et al.*, 2014; Martin-Merida *et al.*, 2018, 2019; Del Pozo-Valero *et al.*, 2019, 2020; Khan *et al.*, 2020). Variants were prioritized as we described before (Del Pozo-Valero *et al.*, 2019).

Statistical analysis

For statistical data analysis between categories of patients, and the two points of time of this study (December 2016 and December 2019) we used R 3.6.3 Version. A Fisher exact test (2x2 contingency table) was used. P-values adjusted by FDR <0.05 were considered significant.

RESULTS

General overview

At the beginning of this study (December 2016) a total of 802 arMD cases were recruited and studied in our Department. Of them, 493 were suspected of Stargardt disease, and 309 presented other maculopathies.

Of 493 patients with STGD diagnosis, 292 (59%) presented biallelic mutations in the *ABCA4* gene, and only one patient was characterized with mutations in other gene, *PRPH2*. The remaining patients were uncharacterized, 63 of them carrying one pathogenic allele in the *ABCA4* gene.

Regarding non STGD cases, 117 of 309 were characterized by 2016, and *ABCA4* was found in 65 of them. The rest of cases carried mutations in other genes (unpublished data). Of 192 uncharacterized patients, 12 were monoallelic for *ABCA4*.

With the implementation of other NGS technologies to study unsolved and new cases, a total of 945 arMD cases had been studied by January 2020, including the previous 802 cases and additional 143 patients.

By the end of this study (January 2020), 407 (74%) and 20 (4%) patients with STGD diagnosis had been characterized with *ABCA4* , or other genes (*BBS1, BEST1, CERKL, CNGB3, CRB1, EYS, OPA1, PLA2G5, PROM1, PRPH2, WFS1*), respectively. Sixteen of the unsolved cases carried one mutated *ABCA4* allele.

About a total of 392 non STGD cases by December 2019, we found *ABCA4* mutations in 100 cases (26%), and mutations in other genes in 95 cases (24%) (unpublished data). Approximately, 48% of cases were unsolved, remaining monoallelic for *ABCA4* a total of 10 patients.

All these results are summarized in Figure 1.
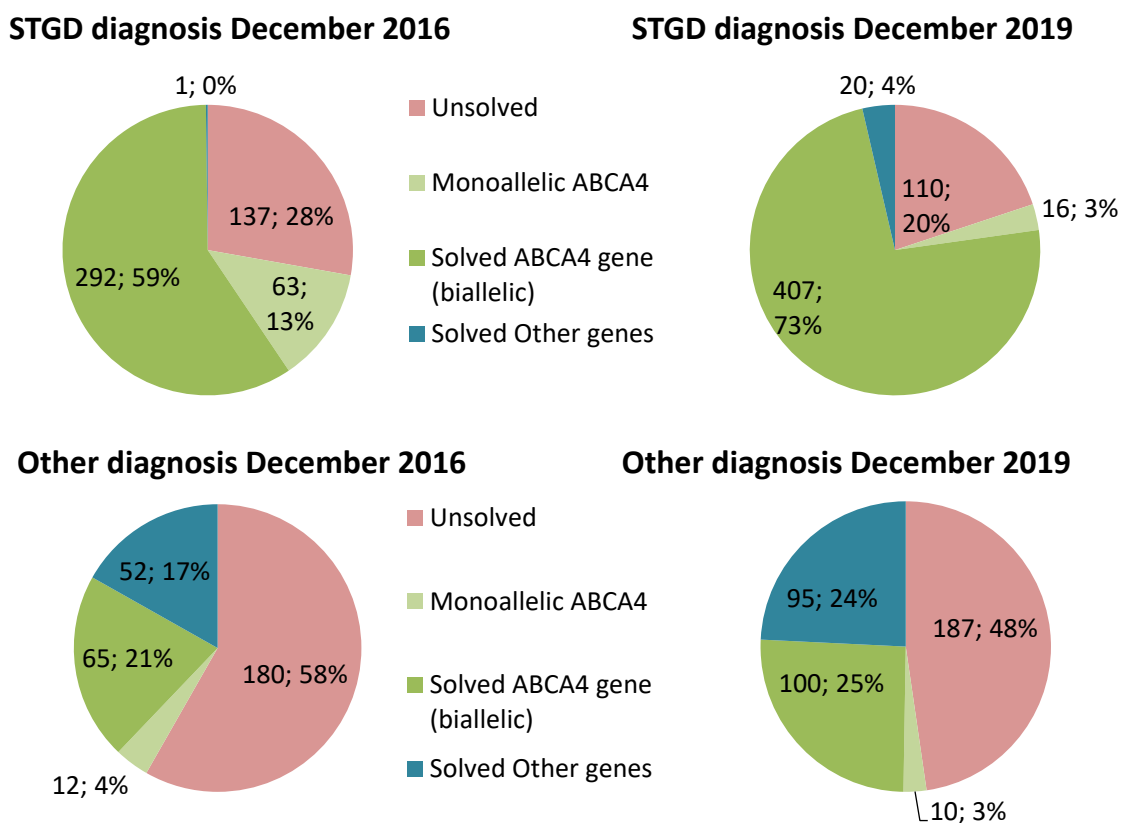
Figure 1. Overview of studied cases with diagnosis of STGD and Other maculopathies.

Comparisons between the genetic diagnosis of STGD patients at the beginning and the end of this study showed statistically significant differences. This was also true in the case of patients with a clinical suspicion of other maculopathies solved with other genes and unsolved (Table 1).

Regarding the genetic diagnosis of both group of patients at the end of this study (December 2019), we also observed statistically significant differences between STGD and non STGD phenotypes excepting monoallelic cases (Table 1).

Monoallelic *ABCA4* cases

At the beginning of this study, 75 patients were monoallelic for *ABCA4*, 63 of them with STGD diagnosis. In these cases, NGS including gene panels with intronic regions (Del Pozo-Valero *et al.*, 2020) or sequencing of the entire *ABCA4* locus by smMIPs (Khan *et al.*, 2020) were performed when possible.

In 46 STGD (73%) and 5 non STGD (41%) diagnosed patients a second pathogenic variant in the *ABCA4* was identified. Of them, coding variants were found in 25 cases, deep intronic variants in 19 cases, non-canonical splice site (NCSS) variants in 3 cases, and a deletion in 4 cases. Mutations in other genes were found in 3 cases (*BEST1, CERKL* and *WFS1*).

By December 2019, 26 cases remained with one pathogenic variant in *ABCA4*, in which the complete gene had been screened in 12 of them.

Table 1. Comparisons between the genetic status and clinical diagnosis of patients at the beginning and end of the study.

STGD diagnosis	Solved **other genes**	Solved ***ABCA4***	**Monoallelic** *ABCA4*	**Unsolved**	Total
December 2016	1 (0%)	292 (59%)	63 (13%)	137 (28%)	493
December 2019	20 (4%)	407 (73%)	16 (3%)	110 (20%)	553
P-value	*p<0.05*	*p<0.05*	*p<0.05*	*p<0.05*	
Other diagnosis	Solved **other genes**	Solved ***ABCA4***	**Monoallelic** *ABCA4*	**Unsolved**	Total
December 2016	52 (17%)	65 (21%)	12 (4%)	180 (58%)	309
December 2019	95 (24%)	100 (25%)	10 (3%)	187 (48%)	392
P-value	*p<0.05*	NS	NS	*p<0.05*	
December 2019	Solved **other genes**	Solved ***ABCA4***	**Monoallelic** *ABCA4*	**Unsolved**	Total
STGD	20 (4%)	407 (73%)	16 (3%)	110 (20%)	553
Other diagnosis	95 (24%)	100 (25%)	10 (3%)	187 (48%)	392
P-value	*p<0.05*	*p<0.05*	NS	*p<0.05*	

DISCUSSION

In this work, we report the genetic diagnosis of 945 arMD families studied over 28 years at the Genetics Department of HUFJD. First approaches to study these families consisted in using genotyping microarrays or Sanger sequencing of specific genes or mutations. Since *ABCA4* was associated to the most frequent maculopathy, STGD1 (Allikmets, Singh, *et al.*, 1997), many of the families presented here were characterized before the implementation of NGS.

Our patients were divided in two groups by their suspected diagnosis: STGD and non STGD patients. In fact, we found *ABCA4* biallelic variants in 74% of patients diagnosed with STGD, and in 25% of patients diagnosed with other maculopathies. A clinical reevaluation in these cases revealed that *ABCA4* patients presented with STGD1 or cone-rod dystrophy (Del Pozo-Valero *et al.*, 2020). Wolock et al (Wolock *et al.*, 2019) identified other genes in patients with a compatible *ABCA4*-disease. The same could occur in our non-characterized patients, therefore, other ophthalmological examinations are needed to establish a correct diagnosis. Some patients of the group of STGD diagnosed patients presented mutations in genes that cause Stargardt-like disease due to similar clinical features. These genes were *PRPH2*, *PROM1*, and *PLA2G5*, the last one associated to benign fleck retina (Sergouniotis *et al.*, 2011).

It has been reported that 10-15% of cases remained monoallelic for *ABCA4* after the study of coding regions (Zernant *et al.*, 2018). Deep intronic and NCSS *ABCA4* variants lead to splice defects and explain the missed molecular cause in almost 50% of cases (Khan *et al.*, 2020). In this study, 68% of monoallelic cases were solved after the study of coding and/or non-coding regions of *ABCA4*. Most of them (49%) carried a second coding variant not detected previously by the technology used or considered benign by its high frequency (Del Pozo-Valero *et al.*, 2020). Deep intronic variants were found in 37% of the solved cases.

With this brief report, we provide a general overview of the characterization of families with arMD diagnosis, most of them diagnosed with STGD. *ABCA4* is the most mutated gene in our cohort, explaining 53% of arMD cases. We remark the importance of genetic diagnosis in cases in which clinical information is not available. Additional genetic studies involving a higher number or genes and/or non-coding regions, together with a clinical reevaluation, will be useful to complete the molecular characterization of unsolved patients.

REFERENCES

Aguirre-Lamban, J. et al. (2009) 'Molecular analysis of the ABCA4 gene for reliable detection of allelic variations in Spanish patients: Identification of 21 novel variants', British Journal of Ophthalmology. doi: 10.1136/bjo.2008.145193.

Aguirre-Lamban, J. et al. (2011) 'Further associations between mutations and polymorphisms in the ABCA4 gene: Clinical implication of allelic variants and their role as protector/risk factors', Investigative Ophthalmology and Visual Science. doi: 10.1167/iovs.10-5743.

Allikmets, R. et al. (1997) 'A photoreceptor cell-specific ATP-binding transporter gene (ABCR) is mutated in recessive Stargardt macular dystrophy', Nature Genetics, 15(3), pp. 236–246. doi: 10.1038/ng0397-236.

Bauwens, M. et al. (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. doi: 10.1038/s41436-018-0420-y.

Bertelsen, M. et al. (2014) 'Generalized choriocapillaris dystrophy, a distinct phenotype in the spectrum of ABCA4-associated retinopathies', Investigative Ophthalmology and Visual Science. doi: 10.1167/iovs.13-13391.

Birtel, J. et al. (2018) 'Clinical and genetic characteristics of 251 consecutive patients with macular and cone/cone-rod dystrophy', Scientific Reports. doi: 10.1038/s41598-018-22096-0.

Corton, M. et al. (2013) 'Exome sequencing of index patients with retinal dystrophies as a tool for molecular diagnosis.', PloS one. doi: 10.1371/journal.pone.0065574.

Cremers, F. P. M. M. et al. (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), pp. 355–362. doi: 10.1093/hmg/7.3.355.

Fadaie, Z. et al. (2019) ' Identification of splice defects due to noncanonical splice site or deep-intronic variants in ABCA4 ', Human Mutation. doi: 10.1002/humu.23890.

Khan, M. et al. (2019) 'Cost-effective molecular inversion probe-based ABCA4 sequencing reveals deep-intronic variants in Stargardt disease', Human Mutation. doi: 10.1002/humu.23787.

Khan, M. et al. (2020) 'Resolving the dark matter of ABCA4 for 1054 Stargardt disease probands through integrated genomics and transcriptomics', Genetics in Medicine. doi: 10.1038/s41436-020-0787-4.

Martin-Merida, I. et al. (2018) 'Toward the mutational landscape of autosomal dominant retinitis pigmentosa: A comprehensive analysis of 258 Spanish families', Investigative Ophthalmology and Visual Science, 59(6), pp. 2345–2354. doi: 10.1167/iovs.18-23854.

Martin-Merida, I. et al. (2019) 'Genomic Landscape of Sporadic Retinitis Pigmentosa', Ophthalmology. doi: 10.1016/j.ophtha.2019.03.018.

Martínez-Mir, A. et al. (1998) 'Retinitis pigmentosa caused by a homozygous mutation in the Stargardt disease gene ABCR', Nature Genetics, 18(1), pp. 11–12. doi: 10.1038/ng0198-11.

Michaelides, M. (2003) 'The genetics of inherited macular dystrophies', Journal of Medical Genetics, 40(9), pp. 641–650. doi: 10.1136/jmg.40.9.641.

Del Pozo-Valero, M. et al. (2019) 'Expanded phenotypic spectrum of retinopathies associated with autosomal recessive and dominant mutations in PROM1', American Journal of Ophthalmology. doi: 10.1016/j.ajo.2019.05.014.

Del Pozo-Valero, M. et al. (2020) 'Genotype-phenotype correlations in a Spanish cohort of 506 families with bi-allelic ABCA4 pathogenic variants', American Journal of Ophthalmology. doi: 10.1016/j.ajo.2020.06.027.

Riveiro-Alvarez, R. et al. (2009) 'Frequency of ABCA4 mutations in 278 Spanish controls: An insight into the prevalence of autosomal recessive Stargardt disease', British Journal of Ophthalmology. doi: 10.1136/bjo.2008.148155.

Riveiro-Alvarez, R. et al. (2013) 'Outcome of ABCA4 disease-associated alleles in autosomal recessive retinal dystrophies: Retrospective analysis in 420 Spanish families', Ophthalmology, 120(11), pp. 2332–2337. doi: 10.1016/j.ophtha.2013.04.002.

Sangermano, R. et al. (2018) 'ABCA4 midigenes reveal the full splice spectrum of all reported noncanonical splice site variants in Stargardt disease', Genome Research, 28(1), pp. 100–110. doi: 10.1101/gr.226621.117.

Sangermano, R. et al. (2019) 'Deep-intronic ABCA4 variants explain missing heritability in Stargardt disease and allow correction of splice defects by antisense oligonucleotides', Genetics in Medicine. doi: 10.1038/s41436-018-0414-9.

Sergouniotis, P. I. et al. (2011) 'Biallelic mutations in PLA2G5, encoding group v phospholipase A 2, cause benign fleck retina', American Journal of Human Genetics. doi: 10.1016/j.ajhg.2011.11.004.

Tanaka, K. et al. (2018) 'The Rapid-Onset Chorioretinopathy Phenotype of ABCA4 Disease', Ophthalmology. doi: 10.1016/j.ophtha.2017.07.019.

Verbakel, S. K. et al. (2018) 'Non-syndromic retinitis pigmentosa', Progress in Retinal and Eye Research. doi: 10.1016/j.preteyeres.2018.03.005.

Wolock, C. J. et al. (2019) 'A case–control collapsing analysis identifies retinal dystrophy genes associated with ophthalmic disease in patients with no pathogenic ABCA4 variants', Genetics in Medicine. doi: 10.1038/s41436-019-0495-0.

Zernant, J. et al. (2014) 'Analysis of the ABCA4 genomic locus in Stargardt disease', Human molecular genetics, 23(25), pp. 6797–6806. doi: 10.1093/hmg/ddu396.

Zernant, J. et al. (2018) 'Extremely hypomorphic and severe deep intronic variants in the ABCA4 locus result in varying Stargardt disease phenotypes', Cold Spring Harbor molecular case studies. doi: 10.1101/mcs.a002733.

Resultados

Artículo 2: Identification of splice defects due to noncanonical splice site or deep-intronic variants in *ABCA4*

Fadaie Z, Khan M, Del Pozo-Valero M, et al.

Publicado en *Human Mutation*, 2019.

Resumen:

En este trabajo se estudiaron funcionalmente variantes del gen *ABCA4*, situadas en los intrones y en los sitios no canónicos de *splicing* (NCSS), con el objetivo de evaluar su posible patogenicidad al alterar potencialmente el *splicing*.

Se seleccionaron las variantes previamente publicadas en otros trabajos y variantes nuevas identificadas en 12 pacientes (no españoles) con enfermedad de Stargardt. Los criterios de selección fueron: las variantes que reducen un 2% la potencia predicha y en las que al menos 2 de los programas de predicción muestren la creación de un sitio críptico de *splicing* con una diferencia superior al 75% de su puntuación si este sitio existía previamente. De esta forma, 19 variantes se estudiaron funcionalmente utilizando midigenes que contienen varios exones del gen *ABCA4*, en los que se introduce la variante a estudiar mediante mutagénesis dirigida. Utilizando el sistema de clonaje Gateway, se transfectaron células HEK293T. Finalmente, el ARN extraído de estas células se sometió a retrotranscripción para estudiar si se ha producido alteración en el patrón de *splicing*, bien produciéndose la pérdida parcial o total de exones, o insertándose parte de los intrones en el transcrito de ARNm. La cuantificación del ARNm aberrante y el ARNm silvestre o "*wild type*" permitió también clasificar estas variantes en cuanto a su severidad.

De las 19 variantes estudiadas, 9 estaban situadas en sitios de *splicing* no canónicos (NCSS), y únicamente una de ellas no mostraba el patrón de *splicing* alterado. Las 10 variantes restantes eran *deep intronic*, y en 4 de ellas se identificaron defectos del *splicing*, la mayoría provocaban la inserción de pseudoexones.

En conclusión, se enfatiza la importancia de realizar estos estudios para determinar la patogenicidad de estos tipos de variantes intrónicas. Gracias a este estudio, 19 variantes se pudieron clasificar como patogénicas (12) o no patogénicas (7), lo que beneficia el diagnóstico en los casos en los que se identifiquen.

Contribución de la autora:

En este trabajo, realizado durante la estancia predoctoral en el Department of Human Genetics Radboud UMC Nijmegen, la autora seleccionó variantes candidatas previamente publicadas para estudiar su efecto en base a que cumplían los criterios establecidos. Concretamente, realizó el estudio funcional completo de 2 de las 19 variantes que se estudian en este trabajo, y las cuales muestran defectos en el patrón de *splicing*. Además, participó en la revisión crítica del manuscrito y en su aprobación final.

Human Mutation HGVS HUMAN GENOME VARIATION SOCIETY WILEY

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

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).
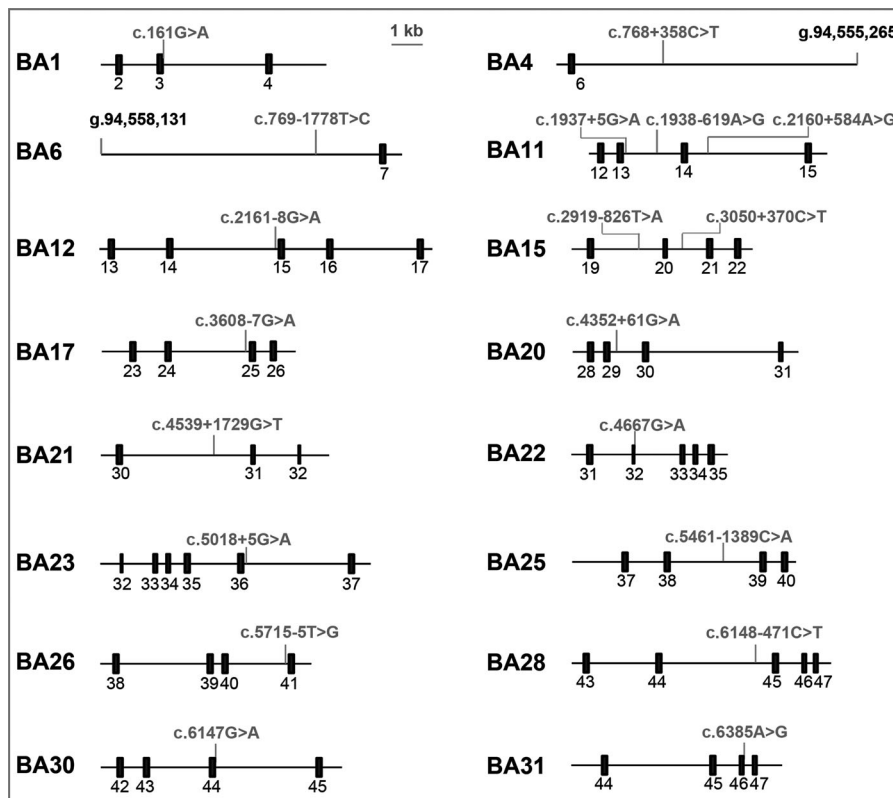
FIGURE 1 Schematic representation of mutant midigene splice constructs of ABCA4 and corresponding locations of the NCSS and deep-intronic variants tested. The exons are represented as black rectangles. BA depicts the BAC-derived clones used that were previously described (Sangermano et al., 2018). The 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,=]	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,=][b]	p.[His721_Val794del,=][b]	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_5714;5460_5715-del_ins (5715−4_5715−1),=][b]	p.[Thr1821Serfs*34,=][b]	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
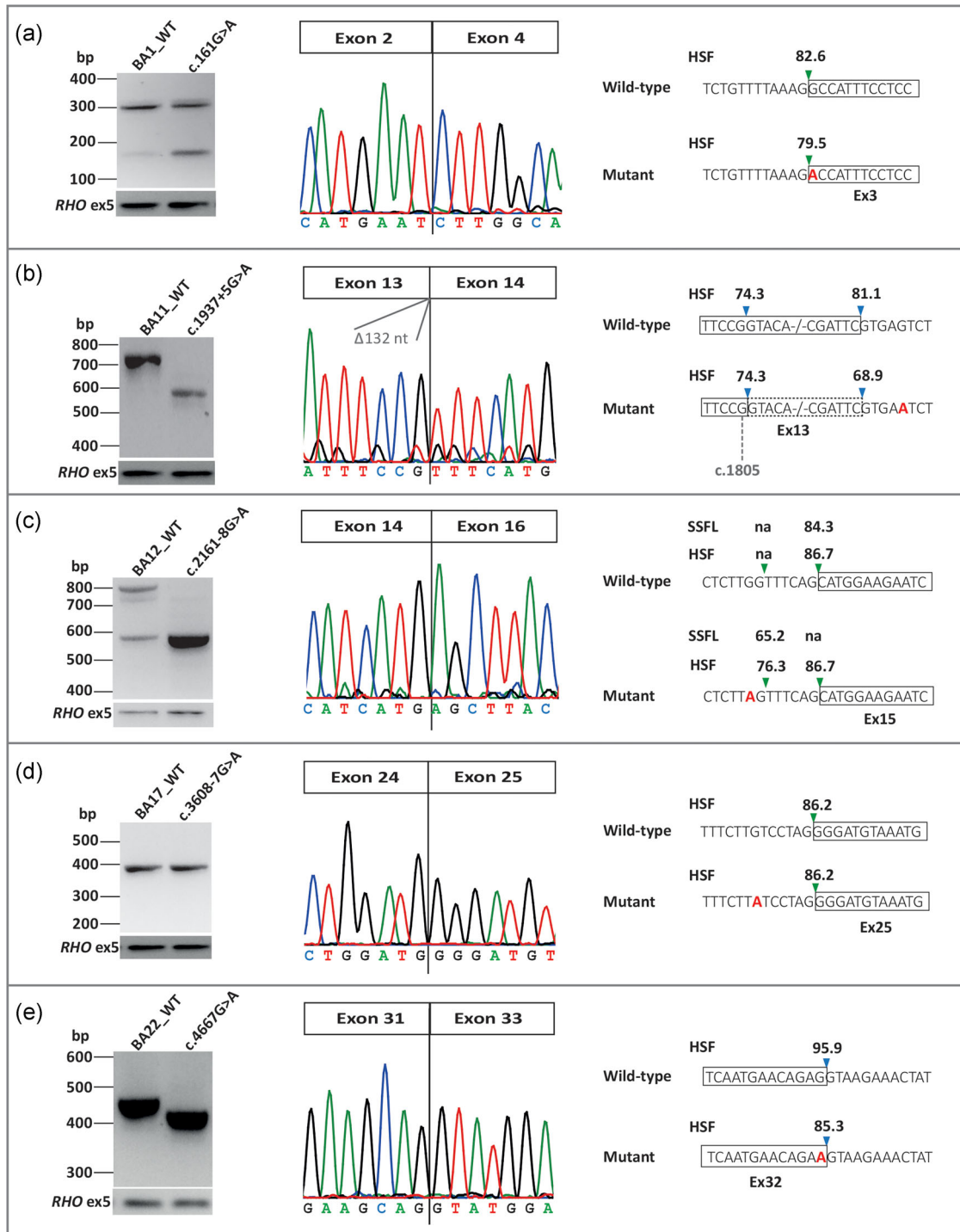
FIGURE 2 Overview of splice defects for nine NCSS variants. Exon 5 *RHO* RT-PCR was used as a control for transfection efficiency. The chromatogram presents the nucleotides identified in the mutant midigene construct. (a) Exon 3 showed weak natural exon skipping in the WT construct, which is significantly 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
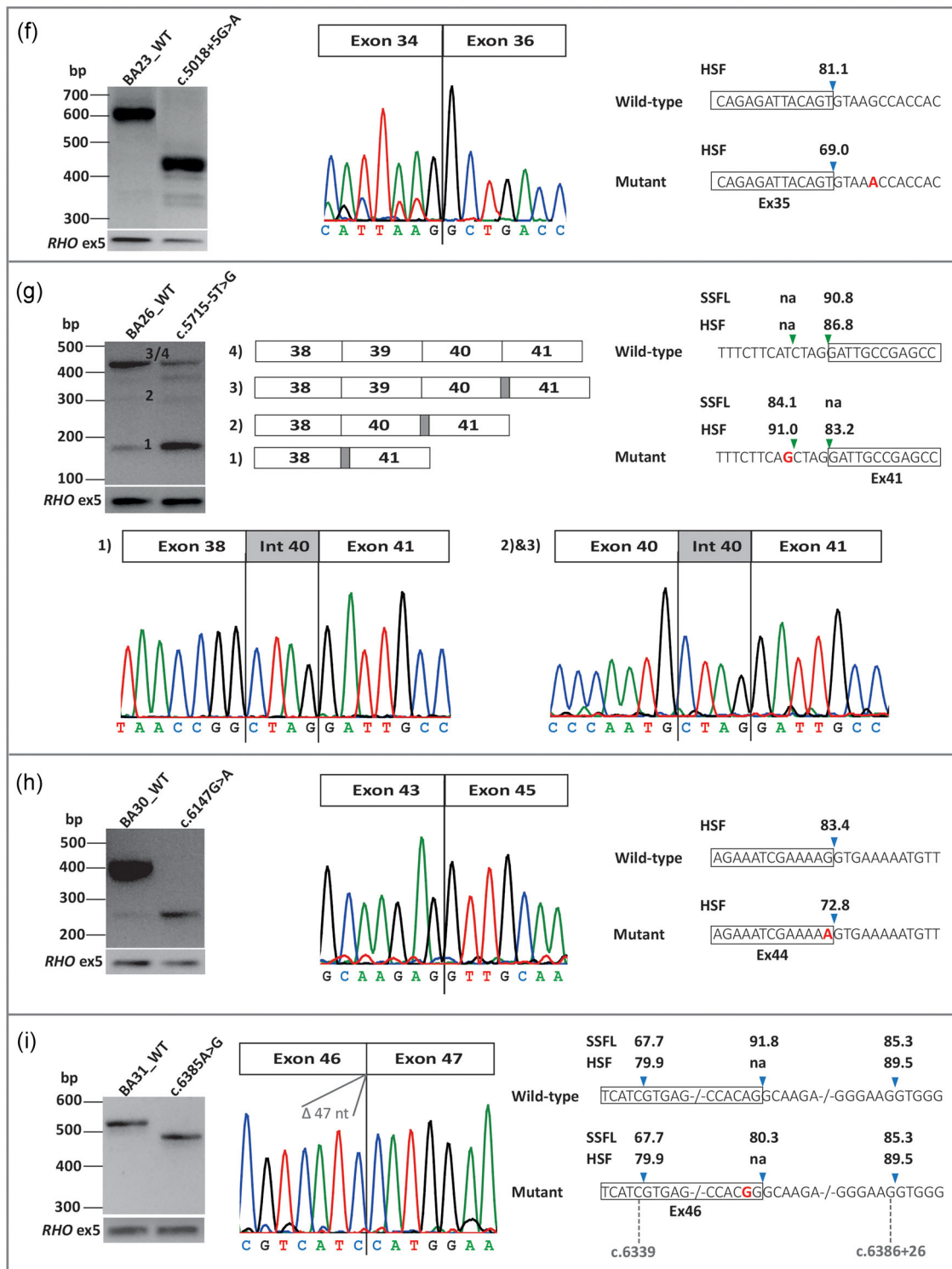
FIGURE 2 Continued

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
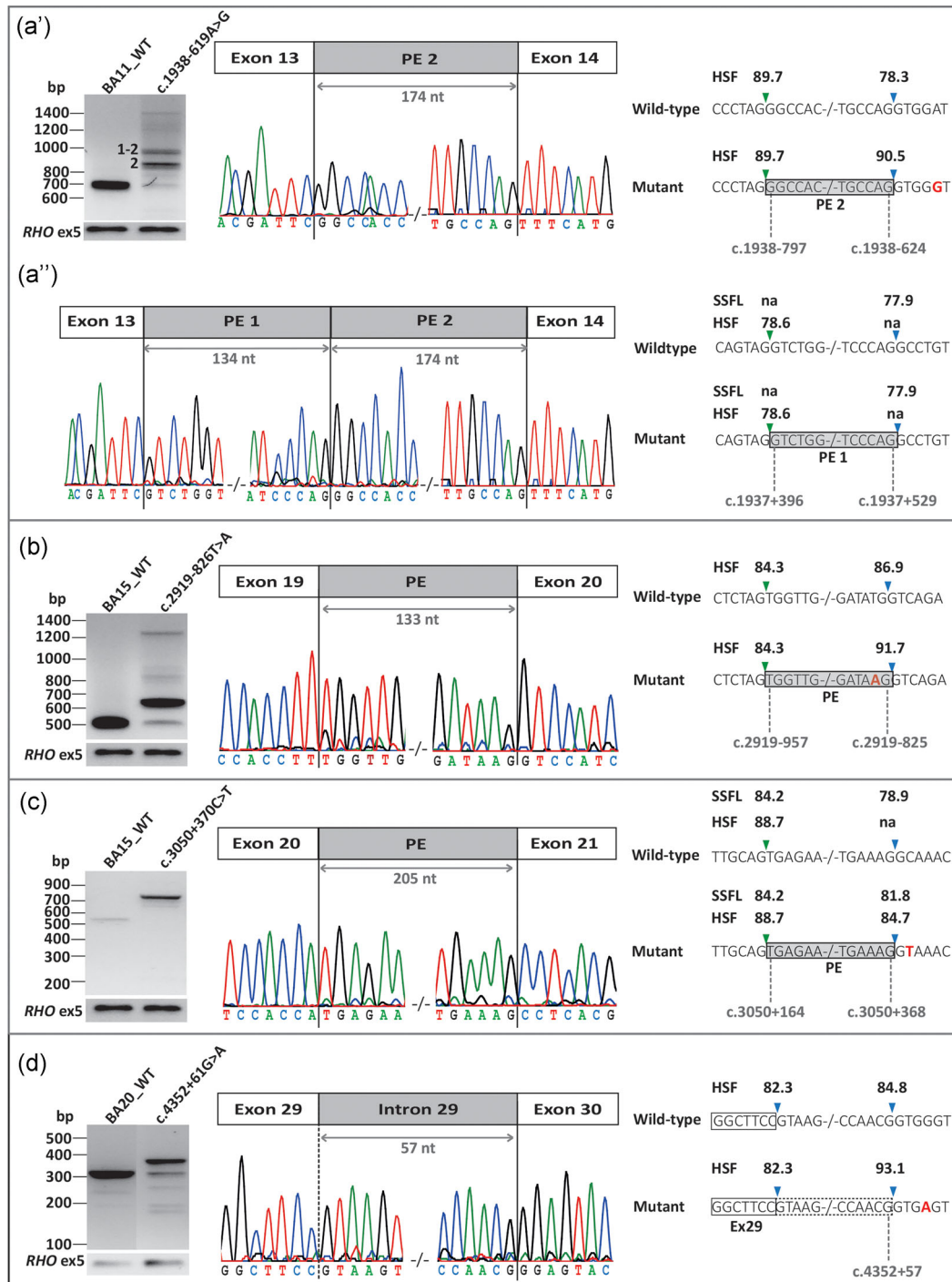
FIGURE 3 Continued.

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 (1938−797_1938−624, 1937+396_1937+529; 1938−797_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_2919ins2919−957_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
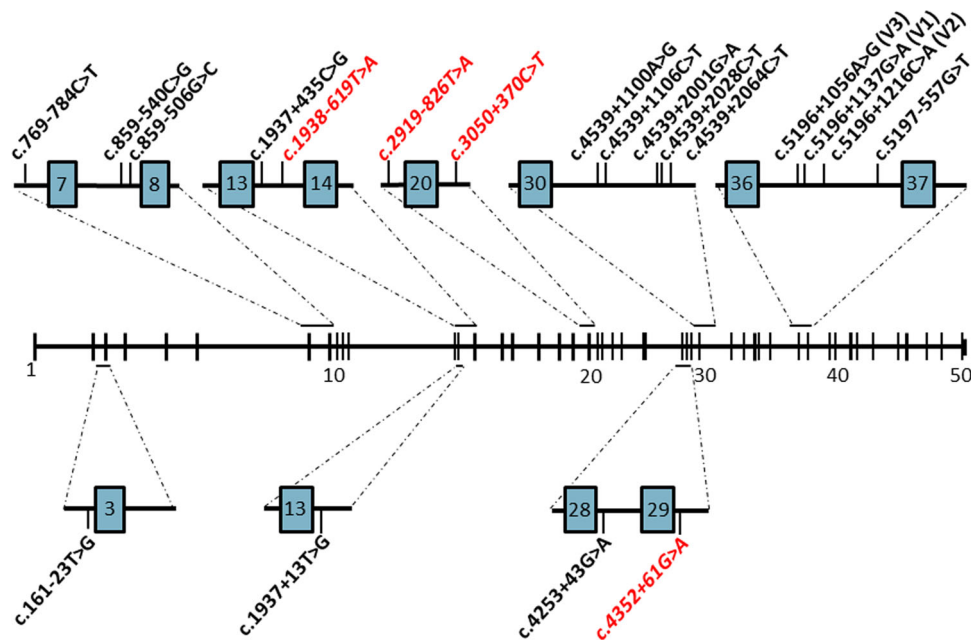
FIGURE 4 The landscape of all currently known pathogenic deep-intronic and NEAR splicing variants in *ABCA4*. The three deep-intronic variants and one NEAR splicing variant deemed as pathogenic in this study are depicted in red/italic. Other variants were described elsewhere (Albert et al., 2018; Bauwens 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.

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, Green-lane 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 of n-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;1–12. https://doi.org/10.1002/humu.23890

Supplemental files

Supplemental Table S1: Pathogenicity proven variants from this study with their identified variants in *trans* in the probands.

Proband ID	Variant assessed in this study		Variant described/identified in *trans*			Source
	Allele 1 (DNA)	Allele 1 (Protein)	Allele 2 (DNA)	Allele 2 (RNA)	Allele 2 (Protein)	
A-1, A-2	c.161G>A	p.(Cys54Serfs*14,Cys54Tyr]	c.6079C>T	r.6079C>T	p.(Leu2027Phe)	LOVD[#], this study
B-1, B-2	c.1937+5G>A	p.(Tyr603_Ser646del)	c.1937+5G>A	r.1806_1937del	p.(Tyr603_Ser646del)	This study
C	c.1938-619A>G	p.Phe647Alafs*22)	c.[3758C>T;5882G>A]	r.[3758C>T;5882G>A]	p.[Thr1253Met;Gly1961Glu]	(Zernant et al., 2014)
D	c.2161-8G>A	P.(His721_Val794del)	c.5882G>A	r.5882G>A	p.(Gly1961Glu)	(Zernant et al., 2014)
E	c.2919-826T>A	p.[Leu973Phefs*1,=]	c.5882G>A	r.5882G>A	p.(Gly1961Glu)	(Zernant et al., 2014)
F	c.3050+370C>T	p.Leu1018Glufs*4)	c.5087G>A	r.5087G>A	p.(Ser1696Asn)	(Zernant et al., 2014)
G	c.4352+61G>A	p.[Glu1452*,=]	c.634C>T	r.634C>T	p.(Arg212Cys)	(Zernant et al., 2014)
H	c.4667G>A	p.(Ser1545_Gln1555del)	c.4667G>A	r.4635_4667del	p.(Ser1545_Gln1555del)	(Tanaka et al., 2018; Zernant et al., 2017)
I	c.5018+5G>A	p.(Val1617Alafs*113)	c.[1622T>C;3113C>T(;)5603A>T]	r.[1622T>C;3113C>T(;)5603A>T]	p.[Leu541Pro,Ala1038Val(;)Asn1868Ile]	(Zernant et al., 2017)
J-1, J-2	c.5715-5T>G	p.[(Thr1821Serfs*34),=]	unknown	unknown	unknown	This study
K	c.6147G>A	p.(Ser2002Argfs*11)	c.3113C>T	r.3113C>T	p.(Ala1038Val)	This study
L	c.6385A>G	p.(Val2114Hisfs*4)	c.5461-10T>C	r.5461_5714del	p.(Thr1821Aspfs*6)	LOVD[$], this study

[#]Variant described by Zhang et al. (Zhang et al., 1999); these siblings have not been previously described. [$] Variant described elsewhere (Fujinami et al., 2013), this patient has been described elsewhere (Tayebi et al., 2019). LOVD, Leiden Open (source) Variant Database.

Supplemental Table S2: Ophthalmologic characteristics of probands carrying pathogenic splice variants

Proband ID	Clinical features	Age at last exam	(subjective) Age of onset	BCVA RE	BCVA LE	Fundus Photos	Electroretinogram	Source
A-1, A-2	CRD	53	12, 10	CF	CF	A large atrophic lesion at the macula, surrounding the optic nerve, scattered flecks in the fundus but extensive hyperautofluorescent flecks out to the mid periphery	n.a.	LOVD*
B-1	CRD	23	5	CF	CF	Macular pigmented scar, yellowish subretinal lesions (salt and pepper)	Severely diminished responses at both scotopic and photopic conditions	This study
B-2	CRD	16	6	CF	LP	Macular degeneration, severe optic atrophy	Absent responses at both scotopic and photopic conditions	This study
C	STGD1	33	18	20/150	20/150	Mild atrophy confined to the central macula (BEM, little to no flecks)	Normal ffERG responses	(Zernant et al., 2014)
D	STGD1	37	28	20/25	20/32	Macular dystrophy surrounded by rare yellow-white flecks at the level of the retinal pigment epithelium	Normal ffERG responses	(Zernant et al., 2014)
E	STGD1	52	12	40/400	40/400	Well-delineated chorioretinal atrophy lesion in the macula; lesion-centric flecks and reticular flecks just beyond the vascular arcades and nasal to the optic disc (BEM)	Normal ffERG responses	(Zernant et al., 2014)
F	STGD1	37	31	20/80	20/80	Macular atrophy with yellow-white flecks at the level of the retinal pigment epithelium at the posterior pole reaching the vascular arcades	Normal ffERG responses	(Zernant et al., 2014)
G	STGD1	49	37	20/160	20/200	Classic hyper autofluorescent deposits in the posterior pole in a "reticulated" pattern, including nasal to the disc with peripapillary sparing. Central atrophic changes (mottled) in both eyes with initially a patch of discrete geographic atrophy that was parafoveal (RE) and multiple small patches of atrophy including foveal (LE)	ffERG: normal rod responses, cone responses were reduced (~40%) with normal implicit times	(Zernant et al., 2014)
H	Rapid-Onset Chorio-retinopathy	41	8	HM	HM	Diffuse, widespread chorioretinal atrophy across the posterior pole; paramacular bone-spicule pigment deposition	Non-recordable scotopic and photopic responses	(Tanaka et al., 2018; Zernant et al., 2017)
I	STGD1	15	n.a	20/25	20/25	Mild granular atrophy confined to the central macula	n.a.	(Zernant et al., 2017)
J-1	CRD	21	10	20/400	CF	Extensive macular atrophy extending outside arcades and nasal to disc with corresponding hypo FAF centrally, mottled peripheral retina with patchy hyper and hypo FAF throughout retina	n.a.	This study
J-2	CRD	35	13	3/60	CF	Extensive macular atrophy extending outside arcades and nasal to disc with corresponding hypo FAF centrally, mottled peripheral retina with patchy hyper and hypo FAF throughout retina	n.a.	This study
K	Classic STGD1	41	<23	20/800	20/800	Fundus flavimaculatus	ffERG: photopic normal with borderline latencies mfERG centrally with reduced amplitudes and normal latencies, peripherally normal	This study
L	RP	23	6	LP	LP	Bone spicule deposits predominant in the peripheral retina, arteriolar narrowing and waxy palor of the optic nerve head	n.a.	LOVD[$]

*Variant described by Zhang et al. (Zhang et al., 1999), current siblings have not been previously described.[$] Variant described by Fujinami et al. (Fujinami et al., 2013), current patient has been described by Tayebi et al. (Tayebi et al., 2019). BCVA, Best corrected visual acuity; CRD, cone-rod dystrophy; CF, counting fingers; FAF, fundus autofluorescence; ffERG, full-field electroretinogram; HM, hand motion; LE, left eye; LOVD, Leiden Open Variant Database; LP, light perception; n.a., not analyzed; RE, right eye; RP, retinitis pigmentosa, STGD1, Stargardt disease, BEM, Bull's eye maculopathy.

Supplemental Table S3: variants with no splice defect from this study with their identified variants in *trans* in the probands.

Proband ID	Variant assessed in this study		Variant described/identified in *trans*			Source
	Allele 1 (DNA)	Allele 1 (Protein)	Allele 2 (DNA)	Allele 2 (RNA)	Allele 2 (Protein)	
M	c.768+358C>T	p.(=)	c.5603A>T	r.5603A>T	p.(Asn1868Ile)	(Zernant et al., 2014)
N*	c.768+358C>T	p.(=)	c.[2160+584A>G;5603A>T;5882G>A]	r.[=;5603A>T;5882G>A]	p.[=,Asn1868Ile,Gly1961Glu]	(Zernant et al., 2014)
	c.[2160+584A>G;5603A>T;5882G>A]	p.[=,Asn1868Ile,Gly1961Glu]	c.768+358C>T	r.=	p.(=)	(Zernant et al., 2014)
O	c.769-1778T>C	p.(=)	c.[4457C>T;4539+2064C>T]	r.[4457C>T;4539_4540ins4539+1891_4540-2162]	p.[Pro1486Leu,Arg1514Leufs*36]	(Zernant et al., 2014)
P*	c.769-1778T>C	p.(=)	c.[4253+43G>A;4539+1729G>T]	r.[4129_4253del;=]	p.[Ile1377Hisfs*3,=]	(Zernant et al., 2014)
	c.[4253+43G>A;4539+1729G>T]	p.[Ile1377Hisfs*3,=]	c.769-1778T>C	r.=	p.(=)	(Zernant et al., 2014)
Q	c.1015T>G;c.3608-7G>A	p.(=)	c.5603A>T	r.5603A>T	p.(Asn1868Ile)	(Bauwens et al., 2019)
R	c.[4539+2064C>T;5461-1389C>A]	p.[Arg1514Leufs*36,=]	unkown	unknown	unknown	(Zernant et al., 2014)
S	c.[5603A>T;6148-471C>T]	p.[Asn1868Ile,=]	c.[5461-10T>C;5603A>T]	r.[5461_5714del,5461_5584del;5603A>T]	p.[Thr1821Aspfs*6,Thr1821Valfs*13,Asn1868Ile]	(Zernant et al., 2014)

*Deep-intronic variants on both alleles were assessed in this study.

Supplemental Table S4: Wildtype and mutant *in silico* prediction scores for all variants assessed in this study.

DNA variant	Splicing effect	c. position splice site	Canonical or non-canonical splice site	SSFL			Max EntScan			NN SPLICE			Gene Splicer			HSF		
				WT	Mut	Delta %	WT	Mut	Delta %	WT	Mut	Delta %	WT	Mut	Delta %	WT	Mut	Delta %
c.161G>A	Exon 3 skipping	c.161	Canonical	88.2	84.3	3.9	8.1	6.5	13.3	0.9	0.8	10	12.0	11.2	3.3	82.6	79.5	3.1
c.768+358C>T	None	c.768+344	Non-canonical	85.9	85.9	0	9.5	9.5	0	1.0	1.0	0	1.2	0.8	1.6	92.0	92.0	0
c.769-1778T>C	None	c.769-1795	Non-canonical	83.2	83.2	0	6.6	6.6	0	1.0	1.0	0	5.1	5.4	-1.4	85.1	85.1	0
c.1937+5G>A	132-nt exon 13 deletion	c.1937	Canonical	69.7	57.5	12.2	7.9	-	65.8	0.8	0	80	4.2	-	17.5	81.1	68.9	12.2
		c.1805	Non-canonical	68.1	68.1	0	5.7	5.7	0	0.8	0.8	0	1.8	1.8	0	74.3	74.3	0
c.1938-619A>G	174-nt pseudoexon	c.1938-624	Non-canonical	73.4	85.6	-12.2	3.6	8.6	-41.6	0.2	0.9	-70	-	3.4	-14.1	78.3	90.5	-12.2
c.2160+584A>G	None	c.2160+583	Non-canonical	-	83.4	-83.4	-	8.4	-70	-	1.0	-100	-	0.7	-2.9	-	91.6	-91.6
c.2161-8G>A	Exon 15 skipping	c.2161	Canonical	84.3	-	84.3	5.4	-	33.7	0.2	0.1	10	3.5	-	16.6	86.7	86.7	0
		c.2161-6	Non-canonical	-	65.2	-65.2	-	4.0	-25	-	0.2	-20	-	-	0	-	76.3	-76.3
c.2919-826T>A	133-nt pseudoexon	c.2919-825	Non-canonical	76	84.5	-8.5	3.6	6.8	-26.6	0.1	0.8	-70	-	2.8	-11.6	86.9	91.7	-4.8
c.3050+370C>T	205-nt pseudoexon	c.3050+368	Non-canonical	78.9	81.8	-2.9	-	8.1	-67.5	-	0.7	-70	-	5.1	-21.2	-	84.7	-84.7
c.3608-7G>A	None	c.3608	Canonical	89.4	89.4	0	7.9	7.2	4.3	1.0	1.0	0	7.2	7.2	0	86.2	86.3	-0.1
c.4352+61G>A	57-nt exon 29 elongation	c.4352+57	Non-canonical	76.3	86.4	-10.1	8.3	11.5	-26.6	0.7	1.0	-30	5.4	9.9	-18.7	84.8	93.1	-8.3
c.4539+1729G>T	None	c.4539+1727	Non-canonical	-	84.6	-84.6	-	9.8	-81.6	-	0.9	-90	-	-	0	-	89.3	-89.3
c.4667G>A	Exon 32 skipping	c.4667	Canonical	91.6	79.5	12.1	10.1	6.4	30.8	1.0	1.0	0	-	-	0	95.9	85.3	10.6
c.5018+5G>A	Exon 35 skipping	c.5018	Canonical	72.3	60.2	12.1	5.3	-	44.1	0.9	0	90	1.1	-	4.5	81.1	69	12.1
c.5461-1389C>A	None	c.5461-1389	Non-canonical	-	-	0	-	-	0	-	0	0	-	-	0	-	71.2	-71.2
c.5715-5T>G	Exon 39/40 skipping	c.5715	Canonical	90.8	-	90.8	10.1	3.6	40.6	1.0	0.9	10	9.7	2.9	32.3	86.8	83.2	3.6
		c.5715-4	Non-canonical	-	84.1	-84.1	-	6.3	-39.3	-	0.9	-90	-	6.6	-31.4	-	91	-91
c.6147G>A	Exon 44 skipping	c.6147	Canonical	77.9	65.7	12.2	4.5	-	37.5	0.6	0	60	5.1	-	21.2	83.4	72.8	10.6
c.6148-471C>T	None	c.6148-458	Non-canonical	81.8	85	-3.2	7.7	8.5	-5	0.7	0.9	-20	2.4	3.8	-6.6	87.6	87.6	0
c.6385A>G	47-nt exon 46 deletion	c.6385	Canonical	91.8	80.3	11.5	-	-	0	-	-	0	-	-	0	-	-	0
		c.6339	Non-canonical	67.7	67.7	0	4.4	4.4	0	0.6	0.6	0	5.4	5.5	-0.4	79.9	79.9	0

SSFL, SpliceSiteFinder-Like (0-100); MaxEnScan (0-12 for SDSs_0-16 for SASs); NNSPLICE (0-1); GeneSplicer (0-24 for SDSs_0-21 for SASs); HSF, Human Splicing Finder (0-100). "-" indicates an unrecognized splice site.

Supplemental Table S5: The mutagenesis primer sequences with GC content and melting temperature tm(°C).

Name	Sequence	Tm (°C)	GC%
c.161G>A_F	cgccttgttggggaaatggtctttaaaacagaaaatgagg	80.3	42.5
c.161G>A_R	cctcattttctgttttaaagaccatttccccaacaaggcg	80.3	42.5
c.768+358C>T_F	actgggaccagaaaagattcaatatcttaccatgatcacc	76.5	40.0
c.768+358C>T_R	ggtgatcatggtaagatattgaatcttttctggtcccagt	76.5	40.0
c.769-1778T>C_F	gtgaaatgatatgttgtctgtggtttgcttcaaaaaaacttcagga	79.3	34.8
c.769-1778T>C_R	tcctgaagttttttgaagcaaaccacagacaacatatcatttcac	79.3	34.8
c.1937+5G>A_F	cgcgaacttcagattcacgaatcgtccacgaagc	82.7	52.9
c.1937+5G>A_R	gcttcgtggacgattcgtgaatctgaagttcgcg	82.7	52.9
c.1938-619A>G_F	ggagaccctggtatacccacctggcaaacag	78.1	58.1
c.1938-619A>G_R	ctgtttgccaggtgggtataccagggtctcc	78.1	58.1
c.2160+584A>G_F	ggctataaaaactccttaccactgaggaaggatggac	75.0	45.9
c.2160+584A>G_R	gtccatccttcctcagtggtaaggagtttttatagcc	75.0	45.9
c.2161-8G>A_F	ttcttccatgctgaaactaagaggccatgcgtcag	80.1	48.6
c.2161-8G>A_R	ctgacgcatggcctcttagtttcagcatggaagaa	80.1	48.6
c.2919-826T>A_F	agctctctgaccttatcagttagcatggagagaaag	73.4	44.4
c.2919-826T>A_R	ctttctctccatgctaactgataaggtcagagagct	73.4	44.4
c.3050+370C>T_F	ccccctgaaaagtttacctttcaggccgtgg	79.2	54.8
c.3050+370C>T_R	ccacggcctgaaaggtaaacttttcaggggg	79.2	54.8
c.3608-7G>A_F	catcccctaggataagaaaaaagactgatgccagc	76.0	45.7
c.3608-7G>A_R	gctggcatcagtcttttttcttatcctaggggatg	76.0	45.7
c.4352+61G>A_F	cctgccatcttgaactcaccgttgggtcctc	80.8	58.1
c.4352+61G>A_R	gaggacccaacggtgagttcaagatggcagg	80.8	58.1
c.4539+1729G>T_F	cccattatttgctcctacctgctgccttagcca	78.5	51.5
c.4539+1729G>T_R	tggctaaggcagcaggtaggagcaaataatggg	78.5	51.5
c.4667G>A_F	attctgataaaaatagtttcttacttctgttcattgacccagaatttgc	75.7	30.6
c.4667G>A_R	gcaaattctgggtcaatgaacagaagtaagaaactattttttatcagaat	75.8	30.6
c.5018+5G>A_F	gggctgtggtggtttacactgtaatctctgagagct	77.1	50.0
c.5018+5G>A_R	agctctcagagattacagtgtaaaccaccacagccc	77.2	50.0
c.5461-1389C>A_F	acactgtggcttatatctttgcaaacttcgtgggc	77.7	45.7
c.5461-1389C>A_R	gcccacgaagtttgcaaagatataagccacagtgt	77.7	45.7
c.5715-5T>G_F	ggctcggcaatcctagctgaagaaaaggggtca	80.8	54.5
c.5715-5T>G_R	tgaccccttttcttcagctaggattgccgagcc	80.8	54.5
c.6147G>A_F	acaacaaaacatttttcactttttcgatttcttctgctggtacacc	78.9	34.8
c.6147G>A_R	ggtgtaccagcagaagaaatcgaaaaagtgaaaaatgttttgttgt	78.9	34.8
c.6148-471C>T_F	aggcactggagaacaatatgaacacagccattgct	78.6	45.7
c.6148-471C>T_R	agcaatggctgtgttcatattgttctccagtgcct	78.6	45.7
c.6385A>G_F	tgggaatctcttgcccgtgggatgtgaggac	82.5	58.1
c.6385A>G_R	gtcctcacatcccacgggcaagagattccca	82.5	58.1

7

Supplemental Table S6: The *ABCA4* (NM_000350.2) primer sequences used for RT-PCR and corresponding wild-type product size.

Forward primer	Sequence	Reverse primer	Sequence	Product size (bp)
Exon 2	attcgctttgtggtggaact	Exon 4	tctctggtgcattcatgagg	290
Exon 12	ccaacgtgccctctctctac	Exon 15	ccttggagaagaaggtgctg	669
Exon 13 (1)	gcctatctgcaggacatggt	Exon 17	caaggaagtggggttccata	789
Exon 13 (2)	ggatcacaaggagccaggt	Exon 14	cgcaactccttctccaagac	190
Exon 13 (2)	ggatcacaaggagccaggt	PE2$	tggcaaacaggatagatttga	258
PE1$	ataatgcagccccttcccta	Exon 14	cgcaactccttctccaagac	220
Exon 19	tctttgaacgtgagcatcca	Exon 22	aatcaccaccttggcatctc	504
Exon 23	tgcgcaagatgaaaaacatc	Exon 26	cctccgtgaccttcagaaaa	358
Exon 29	tgcagacgtcctcctgaata	Exon 31	atgttcctgtccgtcaggtc	301
Exon 30	aaacatcacccagctgttcc	Exon 32	tcattgacccagaatttgctc	259
Exon 31	acctgacggacaggaacatc	Exon 35	ctgagagctgctccttggtc	437
Exon 32	gcaaattctgggtcaatgaac	Exon 37	cccagcactcacggaataat	579
Exon 37	tgaattattccgtgagtgctg	Exon 40	attgggagaggaagaagtgg	517
Exon 38	gcggtcattcccatgatgta	Exon 41	aacaatgggctccttagtgg	430
Exon 39	tgaggaagctgctcattgtc	Exon 44	tgttctcgtcctgtgagcag	611
Exon 42	atttatccaggcacctccag	Exon 45	gagagtttccgcttgttgc	405
Exon 43	aagatgctcactggggacac	Exon 47	gaatggtgcccatacatcg	518
Exon 44	gggctactgtcctcagtttg	Exon 47	gaatggtgcccatacatcg	423
RHO# exon 3	cggaggtcaacaacgagtct	*RHO* exon 5*	aggtgtaggggatgggagac	514
RHO exon 5	atctgctgcggcaagaac	*RHO* exon 5*	aggtgtaggggatgggagac	140

#*Rhodopsin* (NM_000539.3); * *RHO* 3'UTR; $ Pseudoexon 1 and 2 generated by c.1938-619A>G variant

Supplemental Table S7: The percentage of wild-type and mutant bands in RT-PCRs for variants resulting in multiple mRNA molecules.

	Fragment size (nt)	Raw intensity	Correction factor	Corrected intensity	Percentage	Severity
c.161G>A_ p.[Cys54Serfs*14,Cys54Tyr]						Moderate
WT_c.161G>A_upper band	289	7129.8	1	7129.8	86%	
WT_c.161G>A_lower band	147	549.8	0.50	1169.5	14%	
Mut_c.161G>A_upper band	289	6524.2	1	6524.2	44%	
Mut_c.161G>A_lower band	147	4216.4	0.50	8289.5	56%	
c.1938-619A>G_ p.(Phe647Alafs*22)						Severe
WT_c.1938-619A>G_single band	668	9666.9	1	9666.9	100%	
Mut_c.1938-619A>G_upper band	970	3189.1	1.45	2196.2	30%	
Mut_c.1938-619A>G_middle band	842	5324.2	1.26	4223.9	58%	
Mut_c.1938-619A>G_lower band	668	899.9	1	899.9	12%	
c.2919-826T>A_ p.[Leu973Phefs*1,=]						Severe
WT_c.2919-826T>A_single band	503	8992.0	1	8992.0	100%	
Mut_c.2919-826T>A_upper band	1250	2208.2	2.48	888.6	5%	
Mut_c.2919-826T>A_middle band 1	900	178.4	1.78	99.7	1%	
Mut_c.2919-826T>A_middle band 2	830	958.6	1.65	580.9	3%	
Mut_c.2919-826T>A_middle band 3	636	17681.4	1.26	13983.9	74%	
Mut_c.2919-826T>A_lower band	503	3466.6	1	3466.6	18%	
c.4352+61G>A_p.[Glu1452*,=]						Severe
WT_c.4352+61G>A_single band	301	11468.1	1	11468.1	100%	
Mut_c.4352+61G>A_upper band	358	3412.1	1.18	2868.9	84%	
Mut_c.4352+61G>A_lower band	301	550.2	1	550.2	16%	
c.5715-5T>G_ p.[Thr1821Serfs*34,=]						Severe
WT_c.5715-5T>G_upper band	429	9556.4	1	9556.4	39%	
WT_c.5715-5T>G_middle band 1	390	213.8	0.90	235.2	1%	
WT_c.5715-5T>G_middle band 2	309	678.8	0.72	942.4	4%	
WT_c.5715-5T>G_lower band	179	5812.5	0.41	13930.5	56%	
Mut_c.5715-5T>G_upper band	429	1702.4	1	1705.4	4%	
Mut_c.5715-5T>G_middle band 1	390	1653.5	0.90	1818.8	4%	
Mut_c.5715-5T>G_middle band 2	309	1850.3	0.72	2568.9	5%	
Mut_c.5715-5T>G_lower band	179	17195.0	0.41	41210.4	87%	

Measurements were generated using Fiji 3 software. Data points were generated by experiments performed in duplicate and the mean is provided. The values were normalized for the size of different fragments present in RT-PCR product mixture using correction factors as previously described (Sangermano et al., 2018).
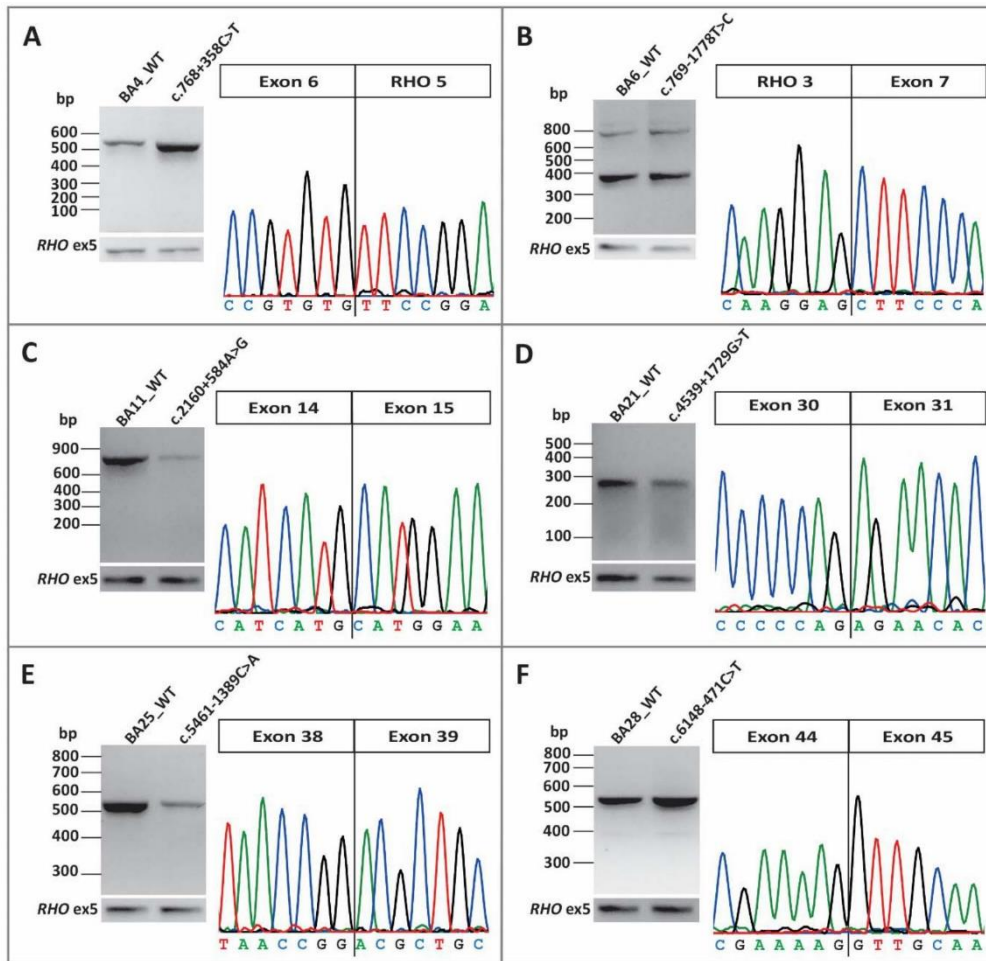
Supplemental Figure 1: Six deep-intronic variants which showed no effect on splicing in HEK293T cells. A-F) Variants assessed which only show a fragment at the size of wild-type. The chromatogram show the nucleotides identified in the mutant midigene construct. Bp, base pairs; WT, wild-type.

11
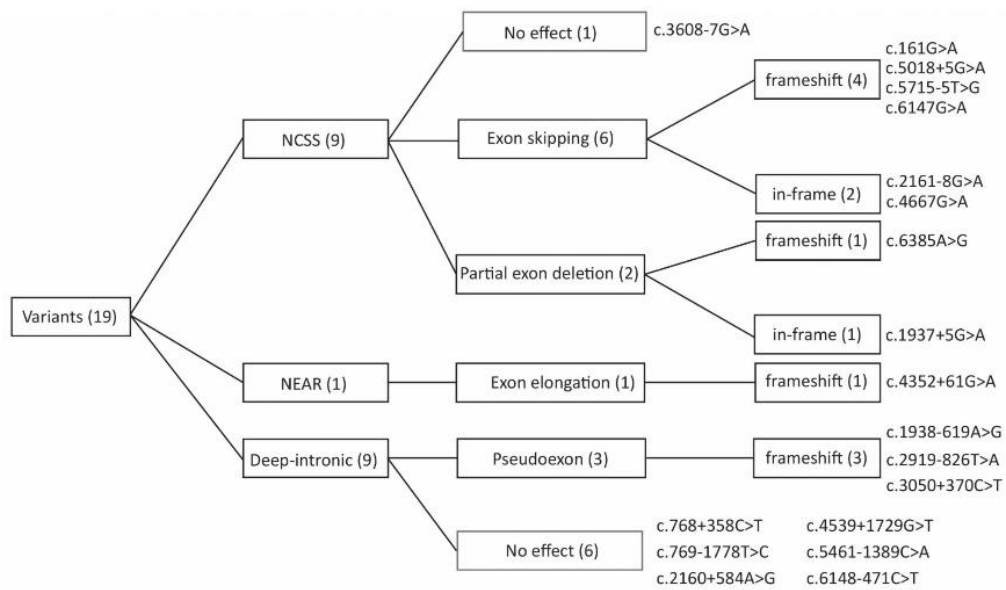
Supplemental Figure 2: The effect of the studied variants categorized by effect. The effect in the mRNA and corresponding reading frame of variants are shown. The numbers in the brackets are the number of variants in each subgroup.

12

92
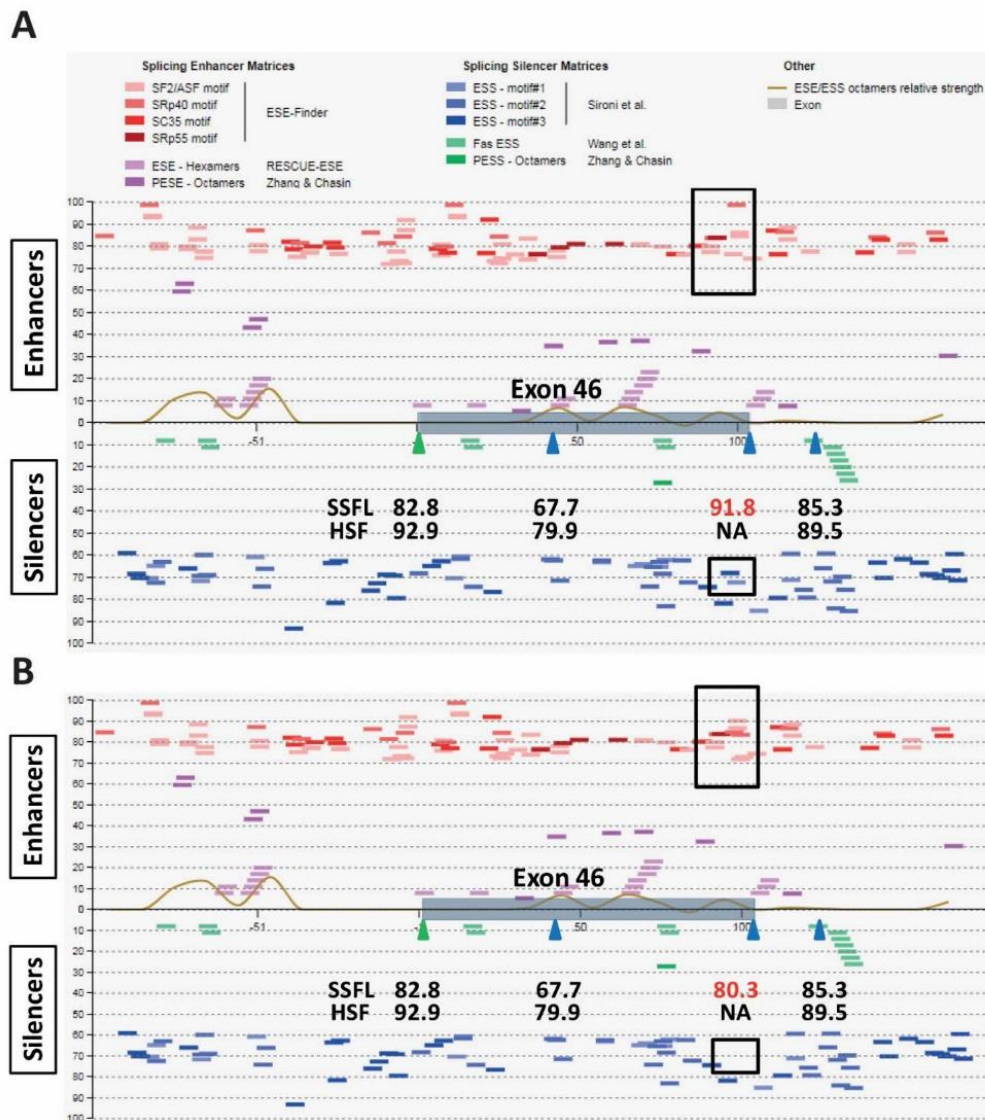
Supplemental Figure 3: Prediction of enhancer and silencer motifs for c.6385A>G. A) Enhancer (top) and silencer (bottom) motifs located in exon 46 and 100-nt up- and downstream regions in the wild-type RNA. **B)** Enhancer and silencer motifs in exon 46 in mutant RNA. Different algorithms identify various motifs which are depicted in different colors. The red and pink bars represent ESEs, whereas blue and green bars colors depict ESSs. The SSFL and HSF scores predictions score for SAS (green triangle) and SDS (blue triangle) sites are provided. The black boxes indicate differences between ESEs and ESSs due to the c.6385A>G variant.

References

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. *Genet Med*. doi:10.1038/s41436-018-0420-y

Fujinami, K., Zernant, J., Chana, R. K., Wright, G. A., Tsunoda, K., Ozawa, Y., . . . Allikmets, R. (2013). ABCA4 gene screening by next-generation sequencing in a British cohort. *Investigative ophthalmology & visual science, 54*(10), 6662-6674.

Sangermano, R., Khan, M., Cornelis, S. S., Richelle, V., Albert, S., Garanto, A., . . . van den Born, L. I. (2018). ABCA4 midigenes reveal the full splice spectrum of all reported noncanonical splice site variants in Stargardt disease. *Genome research, 28*(1), 100-110.

Tanaka, K., Lee, W., Zernant, J., Schuerch, K., Ciccone, L., Tsang, S. H., . . . Allikmets, R. (2018). The Rapid-Onset Chorioretinopathy Phenotype of ABCA4 Disease. *Ophthalmology, 125*(1), 89-99. doi:10.1016/j.ophtha.2017.07.019

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

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., . . . Bertelsen, M. (2014). Analysis of the ABCA4 genomic locus in Stargardt disease. *Human molecular genetics, 23*(25), 6797-6806.

Zhang, K., Garibaldi, D. C., Kniazeva, M., Albini, T., Chiang, M. F., Kerrigan, M., . . . Allikmets, R. (1999). A novel mutation in the ABCR gene in four patients with autosomal recessive Stargardt disease. *American journal of ophthalmology, 128*(6), 720-724.

14

Artículo 3: Resolving the dark matter of *ABCA4* for 1,054 Stargardt disease probands through integrated genomics and transcriptomics

Khan M, Cornelis SS, Del Pozo-Valero M, et al.

Publicado en *Genetics in Medicine*, 2020.

Resumen:

Este trabajo describe el espectro del gen *ABCA4* en los casos previamente no caracterizados y que se reanalizan utilizando otra tecnología.

En este trabajo se estudió el locus completo del gen *ABCA4* junto a las regiones codificantes del gen *PRPH2* en 1054 pacientes con diagnóstico de enfermedad de Stargardt o con características clínicas similares, procedentes de diferentes países de todo el mundo. En concreto, se incluyeron 80 pacientes españoles de nuestra cohorte que habían sido previamente estudiados mediante otras técnicas.

La tecnología que se utilizó en este trabajo para analizar el gen *ABCA4* fue *single molecule molecular inversion probes* (smMIPs). En total, se diseñaron 3866 sondas smMIPs, que cubrían el 97.4% del gen *ABCA4* con una cobertura media de 377 lecturas.

Se identificaron variantes bialélicas en 448 pacientes (42.5% de casos caracterizados). En cuanto a las variantes *deep intronic*, estaban presentes en 117 alelos (13%). También se realizaron estudios funcionales con midigenes en 58 variantes nuevas, 13 de las cuales mostraron un patrón alterado y se consideraron causales. Además, se identificaron variantes estructurales (11 deleciones distintas), y un paciente presentaba isodisomía uniparental. Tras estudiar el gen completo se identifica la causa molecular ("perdida") en casi el 50% de ellos.

Con respecto a nuestra cohorte, de los 80 pacientes incluidos en este estudio, se caracterizaron con el gen *ABCA4* el 70% de los pacientes en los que se había identificado previamente un alelo patogénico en este gen (34/48) y únicamente el 6% (2/32) de los que no presentaban ninguna mutación en los estudios previos.

En conclusión, esta tecnología permite secuenciar la región genómica completa de los genes ya descritos de una forma coste-eficiente, sin la posibilidad de encontrar hallazgos incidentales como el estudio del genoma completo WGS. El estudio funcional de las variantes *deep intronic* encontradas revela la necesidad de realizar estos ensayos para caracterizar completamente a los pacientes. Con este estudio, se redefine el número actual de variantes *deep intronic* asociadas a esta enfermedad, al igual que el número de variantes estructurales identificadas en el gen *ABCA4*.

Contribución de la autora:

En este trabajo, la autora seleccionó a los 80 pacientes de nuestra cohorte que cumplían los criterios para participar en este estudio, y preparó las muestras para su envío. Durante la estancia predoctoral en el Department of Human Genetics Radboud UMC Nijmegen, la autora analizó 58 de ellos, además del proyecto piloto de 15 pacientes. También seleccionó las variantes candidatas de *splicing* para realizar estudios funcionales, diseñando los *primers* para realizar la mutagénesis, y realizando los clonajes, cultivo celular, RT-PCR, extracción del ARNm y la secuenciación para establecer la patogenicidad de las variantes candidatas en los pacientes españoles. Además, durante su reincorporación al laboratorio, realizó los estudios de segregación en los casos resueltos. Finalmente, participó en la revisión crítica del manuscrito, y aprobó la versión final.

Resolving the dark matter of *ABCA4* for 1054 Stargardt disease probands through integrated genomics and transcriptomics

A full list of authors and affiliations appears at the end of the paper.

Purpose: Missing heritability in human diseases represents a major challenge, and this is particularly true for *ABCA4*-associated Stargardt disease (STGD1). We aimed to elucidate the genomic and transcriptomic variation in 1054 unsolved STGD and STGD-like probands.

Methods: Sequencing of the complete 128-kb *ABCA4* gene was performed using single-molecule molecular inversion probes (smMIPs), based on a semiautomated and cost-effective method. Structural variants (SVs) were identified using relative read coverage analyses and putative splice defects were studied using in vitro assays.

Results: In 448 biallelic probands 14 known and 13 novel deep-intronic variants were found, resulting in pseudoexon (PE) insertions or exon elongations in 105 alleles. Intriguingly, intron 13 variants c.1938-621G>A and c.1938-514G>A resulted in dual PE insertions consisting of the same upstream, but different downstream PEs. The intron 44 variant c.6148-84A>T resulted in two PE insertions and flanking exon deletions. Eleven distinct large deletions were found, two of which contained small inverted segments. Uniparental isodisomy of chromosome 1 was identified in one proband.

Conclusion: Deep sequencing of *ABCA4* and midigene-based splice assays allowed the identification of SVs and causal deep-intronic variants in 25% of biallelic STGD1 cases, which represents a model study that can be applied to other inherited diseases.

INTRODUCTION

High-throughput genome sequencing has made a huge impact in biology and is considered the most powerful genetic test to elucidate inherited human diseases.[1] It allows the unbiased detection of a wide spectrum of genetic variants including coding and noncoding single-nucleotide variants (SNVs), as well as structural variants (SVs). However, sequencing and data storage costs as well as the possibility of secondary genetic findings hamper the use of genome sequencing.

Based on the advantages and limitations mentioned above, genome sequencing is not the best method to perform sequence analysis of one or a few genes that are associated with a clinically distinct condition. This is illustrated by autosomal recessive Stargardt disease (STGD1), which is caused by variants in the *ABCA4* gene. STGD1 is the most frequently inherited macular dystrophy with an estimated prevalence of 1/10,000.[2] Thus far, 1180 unique *ABCA4* variants have been reported in 8777 alleles of 6684 cases (www.lovd.nl/ABCA4).[3] A large proportion of the variants affect noncanonical splice site (NCSS) sequences, with variable effects on messenger RNA (mRNA)

processing,[4–6] and several deep-intronic (DI) variants have been identified.[5,7–13] Most of these DI variants strengthen cryptic splice sites resulting in the insertion of pseudoexons (PEs) in the mature *ABCA4* mRNA. SVs seem to be rare in *ABCA4*,[7,10,12,14] although systematic copy-number variant (CNV) analyses have not been performed in most STGD1 cases.

Due to the relatively large size of the *ABCA4* gene (50 exons; 128,313 bp), variant screening initially was restricted to the scanning of the exons and flanking splice sites with poor sensitivity, leaving 50–70% of STGD1 probands genetically unsolved.[14–17] Recently, sequence analysis of the entire 128-kb gene was performed using next-generation sequencing platforms using Raindance microdroplet polymerase chain reaction (PCR) target enrichment or Illumina TruSeq Custom Amplicon target enrichment,[10] HaloPlex-based sequence enrichment,[7,9] or genome sequencing.[1,9]

Identification of two pathogenic alleles is important to confirm the clinical diagnosis because several promising clinical trials are underway based on RNA modulation with antisense oligonucleotides,[7,9,18] drug based therapies,[19] as well as gene augmentation[20] and stem cell therapies.[21] STGD1

cases will only be eligible for these therapies if both causal alleles are known. In addition, recent studies identified alleles carrying a coding variant in *cis* with a DI variant, and only these combinations represented fully penetrant alleles,[7,9] pointing toward the importance of analyzing noncoding regions in the STGD1 cases.

Recently, we reported on the use of 483 single-molecule molecular inversion probes (smMIPs) to sequence the 50 exons and 12 intronic regions carrying 11 pathogenic DI variants of 412 genetically unsolved STGD1 cases.[5] In this study, we aimed to design a semiautomated, high-throughput, cost-effective, and comprehensive sequence analysis of the entire *ABCA4* gene, which could serve as a model study to investigate human inherited diseases due to variants in one or a few genes. Using 3866 smMIPs we sequenced 1054 genetically unsolved STGD or STGD-like probands and 138 biallelic controls carrying known *ABCA4* variants. Novel NCSS and DI variants were tested in vitro for splice defects. Additionally, a very high and reproducible read coverage allowed us to perform CNV analysis.

MATERIALS AND METHODS

Samples
Twenty-one international and four national centers ascertained 1054 genetically unsolved probands in whom STGD was part of the differential diagnosis as determined by the local ophthalmologists specializing in inherited retinal diseases. Since *ABCA4* disease is known for its clinical heterogeneity, a spectrum of (overlapping) *ABCA4*-associated phenotypes were part of this study, as well as a STGD1 phenocopy: central areolar choroidal dystrophy (CACD). The clinical findings specific to a certain clinical diagnosis and the main phenotypic characteristics used in the differential diagnosis are described in Table S1. Also, 19 cases with a clinical diagnosis of macular dystrophy without further specification were included.

Among 1054 cases 833 probands were previously screened by employing different screening methods, i.e., exome sequencing, targeted gene panel sequencing including all *ABCA4* coding regions, and Sanger sequencing of all coding *ABCA4* exons. Details are provided in Table S2.

We discerned two patient groups. The first patient group consisted of 993 genetically unsolved probands who carried one ($n = 345$) or no ($n = 648$) *ABCA4* allele. For two subjects, DNA was not available and both parents of the probands were studied, assuming autosomal recessive inheritance. The second patient group consisted of 61 "partially solved" probands, carrying the c.5603A>T (p.Asn1868Ile) variant in *trans* with other alleles. This last group was also investigated as it was suspected that there could be unidentified DI variants in *cis* with c.5603A>T, as the penetrance of c.5603A>T, when in *trans* with a severe *ABCA4* variant, was ~5% in the population.[22,23]

This study was approved by the Medical Ethical Committee 2010-359 (Protocol nr. 2009-32; NL nr. 34152.078.10) and the Commissie Mensgebonden Onderzoek Arnhem-Nijmegen (Dossier no. 2015-1543; dossier code sRP4h). All samples were collected according to the tenets of the Declaration of Helsinki and written informed consent was obtained for all patients participating in the study.

smMIPs design and *ABCA4* sequence analysis
Detailed information on the smMIPs-based *ABCA4* sequencing, selection of candidate splice variants, and inclusion criteria is provided in the Supplementary Materials and Methods.

Midigene-based splice assay
The effect of nine NCSS variants and 58 DI variants was assessed by midigene-based splicing assays employing 23 wild-type (WT) BA clones previously described[4] and the newly designed BA32, BA33, BA34, and BA35. WT and mutant constructs were transfected in HEK293T cells and the extracted total RNA was subjected to reverse transcription (RT)-PCR as described previously.[4] Details are provided in Supplementary Materials and Methods.

Identification of CNVs and assessment of the underlying mechanism
An Excel-based script was employed to detect CNVs using smMIP read number. Microhomology at the breakpoints was assessed using ClustalW, breakpoint regions were analyzed for non-B motifs by tool (nBMST and QGRS Mapper) (for details see Supplementary Materials and Methods).

Semiquantification of RT-PCR products
To quantify the ratios between correct and aberrant RT-PCR products, densitometric analysis was performed using ImageJ software.

Uniparental disomy detection
To test the presence of uniparental disomy (UPD), haplotype analysis was performed in one STGD1 case (DNA14-33085) using exome sequencing data.

RESULTS

smMIPs performance and *ABCA4* sequencing
A pilot sequencing study was conducted using 15 STGD1 samples and five DNA samples of control individuals, revealing all 34 previously identified variants (Table S3). The average number of reads for the 20 DNA samples ranged from 10 to 152,500 per smMIP, with an overall average coverage of 933× for each smMIP.

In total 1192 DNA samples were analyzed for variants in *ABCA4* using six NextSeq500 runs. The average number of reads of the 3866 smMIPs was 377×. As most nucleotide positions are targeted with two smMIPs, the effective average coverage was ~700×. To determine the coverage of *ABCA4* in more detail, we calculated the average coverage of each nucleotide position for runs 1 to 5 combined (Table S4). To visualize the results, nucleotide positions that were not covered or poorly covered (≤10 reads), moderately covered
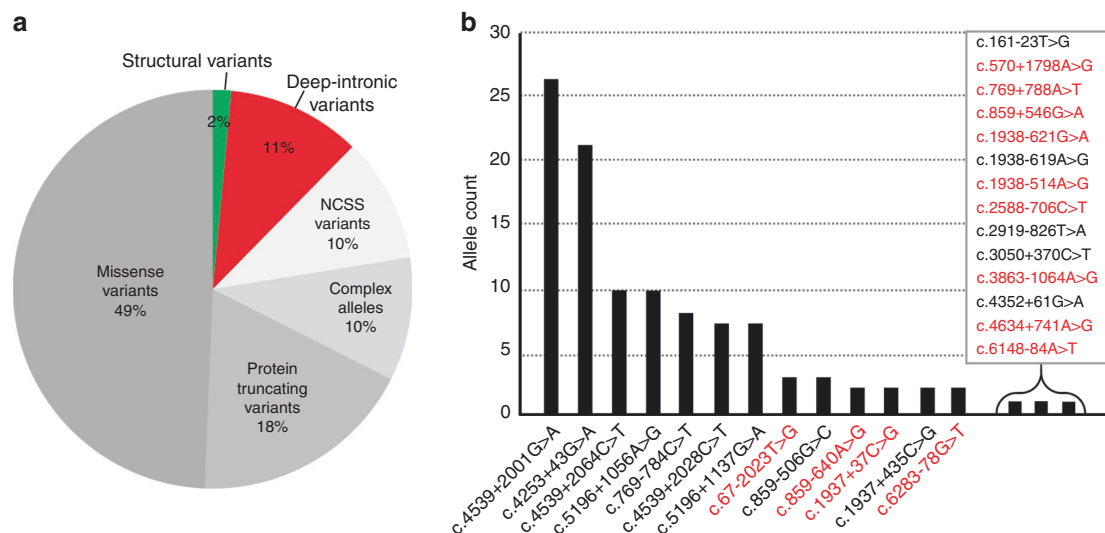
Fig. 1 Distribution of different types of alleles and deep-intronic variants in *ABCA4*. a The contribution of each type of variant or allele in biallelic and monoallelic cases except those carrying c.5603A>T is represented. Protein truncating variants comprise nonsense, frameshift, and canonical splice site variants. The 10% complex alleles only consist of combinations of missense variants, the most frequent of which were c.[1622T>C;3113C>T] (*n* = 30; 27% of all complex alleles) and c.[4469G>A;5603A>T] (*n* = 27; 25% of all complex alleles). They do not include the complex alleles that contain noncanonical splice site (NCSS) variants, deep-intronic variants, or protein truncating variants, when present in *cis* with other variants. If these would have been included, 16% of the alleles would consist of complex alleles. **b** Deep-intronic variant allele count in this study. Novel deep-intronic variants are highlighted in red. One hundred seventeen causal deep-intronic variants were identified. The deep-intronic variants c.4539+2001G>A (*n* = 26) and c.4253+43G>A (*n* = 21) were found most frequently. Most of the novel deep-intronic variants were found in single STGD1 probands.

(11–49 reads), or well covered (≥50 reads) are depicted in Fig. S1. From the 128,366 nt of *ABCA4*, 1980 nt (1.5%) were not or poorly covered, 1410 nt (1.1%) were moderately covered, and 124,976 nt (97.4%) were well covered. Although *ABCA4* introns carry several repetitive elements (Fig. S1), they only had a small effect on smMIPs design. Several larger repeats are present in up- and downstream regions of *ABCA4*, which resulted in the absence—or poor performance—of smMIPs. Sequencing of 1192 samples yielded a total of 7756 unique *ABCA4* variants that are listed in Table S5.

Sensitivity and specificity of the smMIPs-based sequencing
To assess the sensitivity of the new smMIPs sequencing platform, we tested 123 previously genotyped samples[5,9] in three series (runs 2, 3, and 6) (Table S6) as well as 15 control DNA samples carrying 13 different SVs spread throughout the *ABCA4* gene (run 6) (Table S7). All previously known SNVs (*n* = 300) and 13 SVs could be identified, yielding a sensitivity of 100%. Six additional variants were found due to low coverage in the previous studies, and three variants had not been annotated correctly previously.

***ABCA4* gene sequencing and identification of variants**
ABCA4 sequencing was performed for 1054 genetically unsolved STGD and STGD-like patients. This revealed 323 unique (likely) pathogenic SNVs and 11 SVs in 1144 alleles. Sixty-four of 323 SNVs (26%) and all 11 SVs were novel (Table S8). Detailed in silico analysis of novel SNVs is provided in Table S9. Thirteen percent of these alleles were represented by DI variants and SVs and another 10%

accounted for NCSS variants (Fig. **1a**). All variants and the respective cases were uploaded into the *ABCA4* variant and STGD1 cases database LOVD at www.lovd.nl/ABCA4.

Two (likely) pathogenic variants were found in 323 probands, three probands carried p.Asn1868Ile in a homozygous manner, and one (likely) causal variant in *trans* with p.Asn1868Ile was found in 125 probands. Only one (likely) causal variant was identified in 174 probands. Additionally, in 65 probands, the p.Asn1868Ile variant was the only identified variant (Table S10). No (likely) causal variants were found in 364 cases.

Among the SNVs, the most common causal alleles were c.5603A>T (*n* = 134), c.5882G>A (*n* = 84), c.[5461-10T>C;5603A>T] (*n* = 44), c.[1622T>C;3113C>T] (*n* = 30), c.[4469G>A;5603A>T] (*n* = 27), c.4539+2001G>A (*n* = 26), c.6079C>T (*n* = 23), and c.4253+43G>A (*n* = 21) (Table S8). To visualize the relative frequency of causal STGD1-causing alleles, we excluded 65 heterozygous c.5603A>T alleles that were found as the only *ABCA4* allele in these cases, as they were most likely present because of its high allele frequency (0.06) in the general population (Fig. S2).[23,24]

Splice defects due to noncanonical splice site variants
The effect on splicing of nine NCSS variants was tested in nine wild-type splice constructs previously described[4] (Fig. S3). All of the nine tested novel NCSS variants showed a splice defect when tested in HEK293T cells. Severity was assigned according to the percentage of remaining WT mRNA, as described previously.[4] Five NCSS variants were deemed severe as they showed ≤30 of WT mRNA, three were

considered to have a moderate effect with WT RNA present between >30 and ≤70% correct RNA and only one was mild as it showed >70% of WT RNA (Table S11, Fig. S4).

Deep-intronic variants identification and functional characterization

Based on the defined selection criteria, 58 DI variants were selected for splice assays. To test their effects, 27 WT midigene splice constructs were employed, 23 of which were described previously,[4] and four of which were new (Fig. S3). Thirteen of 58 tested DI variants showed a splice defect upon RT-PCR and Sanger validation (Figs. 2 and 3). For the variants that did not show any splice defect, RT-PCR results are shown in Fig. S5.

Six of the novel DI variants, i.e., variants c.570+1798A>G, c.769-788A>T, c.859-640A>G, c.1938-514A>G, c.2588-706C>T, and c.4634+741A>G, resulted in out-of-frame PE inclusions in the RNA and were deemed severe (Figs. 2 and 3). Variants c.67-2023T>G and c.859-546G>A were classified to have a moderate effect as 33% and 36% of the WT RNA products were present, respectively. As predicted due to the presence of a downstream cryptic splice donor site (SDS), variant c.1937+37C>G led to an elongation of exon 13 by 36 nucleotides, which resulted in the introduction of a premature stop codon (p.Phe647*). Moreover, two intron 13 variants, c.1938-621G>A and c.1938-514A>G, showed a complex splice pattern that led to the generation of two mutant transcripts each (Fig. 3a–c). Each of these products contained a shared PE of 134 nt (PE1) as well as variant-specific PEs, denoted PE2 (174 nt) or PE3 (109 nt) for c.1938-621G>A and c.1938-514A>G, respectively (Fig. S6). For variant c.1938-621G>A only 7% of the total complementary DNA (cDNA) product showed PE inclusion whereas for c.1938-514A>G, 87% of the cDNA products were mutant. To investigate the nature of the PE1 insertions, we studied the exon 12–17 segment of the mRNA obtained from photoreceptor progenitor cells (PPCs) derived from a control person. As depicted in Fig. S7, transcripts containing PE1 or PE1 and PE2 were identified when PPCs were grown under nonsense-mediated decay–suppressing conditions. The sum quantity of these two products was 2.9% of total mRNA suggesting that there are small amounts of PE insertions involving PE1 in the healthy retina.

Intriguingly, DI variant c.6148-84A>T showed four RNA splice products, namely a normal spliced RNA, the skipping of exon 45, the insertion of a 221-nt PE (pe1a) coupled with the deletion of exon 44, and finally, the insertion of a 173-nt PE (pe1b) that consist of the same SDS as pe1a but a different splice acceptor site (SAS) (Fig. 3d–f). Finally, variant c.3863-1064A>G showed a complex splice pattern compared with the WT and variant c.6283-78G>T led to the insertion of a 203-nt PE in intron 45 (Fig. S6). However, the exact boundaries of the presumed PE for variant c.3863-1064A>G could not yet be determined due to technical difficulties.

Overall, 13 novel DI variants were found in 18 alleles. Next to the novel variants, 14 previously reported pathogenic DI variants[7–9,13,18] were found in a total of 99 alleles, details of which are shown in Fig. 1b and Table S8.

Identification of novel structural variants in STGD1 cases

Among 1054 STGD and STGD-like patients analyzed, we identified 11 unique novel heterozygous SVs, all exon-spanning deletions, in 16 patients. The corresponding deletions encompass between 1 and 33 exons, ranging from 411 bp to 55.7 kb (Fig. 4, Tables S12–S17). All deletions were found in a heterozygous state in single cases, except the smallest (c.699_768+341del), which encompassed 70 bp of exon 6 and 341 bp of intron 6, and was found in six unrelated patients of Spanish origin. Deletion breakpoints were determined employing genomic PCR and Sanger sequencing for 9 of the 11 deletions. Two deletion junctions (deletions 7 and 11) could not be amplified as the 3' breakpoints were located downstream of the gene beyond the regions targeted by smMIPs. Surprisingly, Sanger sequencing revealed two complex rearrangements as deletions 5 and 6 carried inverted fragments of 279 and 224 bp respectively, residing between large deletions. These small inversions could not be identified with the CNV detection tool.

Microhomologies, repetitive elements, and non-B DNA conformation at deletion breakpoints

The breakpoints of the deletions were subjected to bioinformatic analysis to find elements underlying their formation. The presence of microhomology, repetitive elements, and non-B DNA conformations was investigated except for deletions 7 and 11 as exact boundaries could not be determined by Sanger sequencing. All other studied SVs presented microhomology at the breakpoint junctions, ranging in size from 1 to 6 bp (Fig. S8), four of which presented short insertions (Table S18). In 8 of 11 (72.7%) of the deletion breakpoints, a known repetitive element was observed, including seven non–long terminal repeats (non-LTR) retrotransposons, among which there were one short interspersed nuclear element (SINE) and four long interspersed nuclear elements (LINEs), three DNA transposons from the *hAT* superfamily, and two retrotransposons from the LTR superfamily. However, no breakpoint was part of a known element belonging to the same class and no *Alu* sequence was observed at the breakpoint junctions. Finally, the most prevalent non-B conformations observed among our breakpoints are Oligo(G)n tracts as 21 of these repeats were found in seven SVs (Tables S18, S19). Inverted repeats were observed in five breakpoint regions. No direct repeats or mirror repeats have been detected, therefore excluding triplex and slipped hairpin structure formation, respectively.

Uniparental isodisomy of chromosome 1

In STGD1 proband DNA14-33085, a causal homozygous DI variant, c.859-506G>C (p.[Phe287Thrfs*32,=]), was identified. Segregation analysis revealed this variant to be present in his unaffected father, but not in his unaffected mother. To test the possibility that the mother carried a deletion

ARTICLE
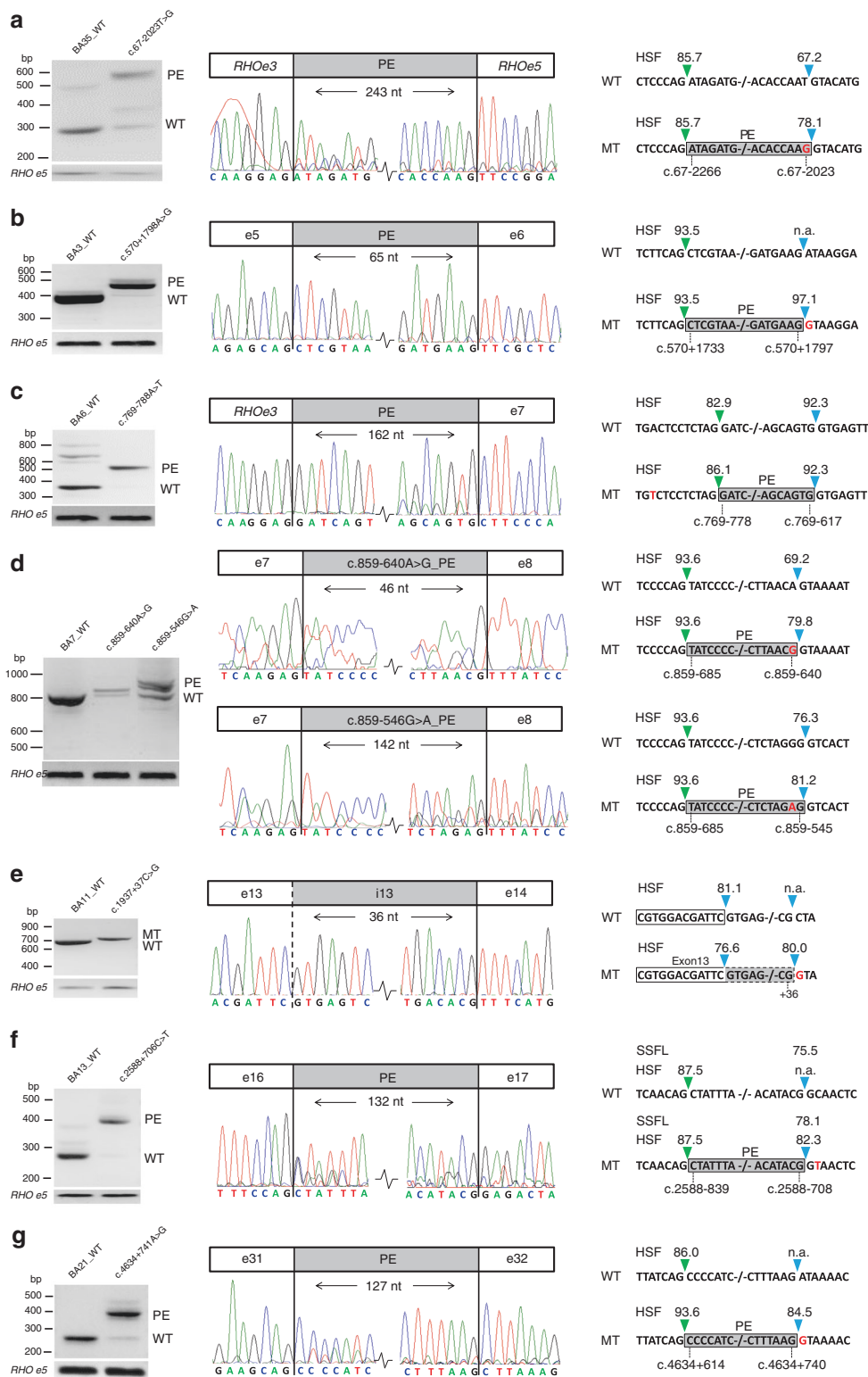
Fig. 2 Novel splice defects due to deep-intronic *ABCA4* variants. Wild-type (WT) and mutant (MT) midigenes were transfected in HEK293T cells and the extracted RNA was subjected to reverse-transcription polymerase chain reaction (RT-PCR). Left panels show the *ABCA4*-specific RT-PCR products with Rhodopsin exon 5 (*RHO* e5) RT-PCR as a transfection efficiency control. In the middle panels, Sanger sequencing results of the RT-PCR products are given. At the right side, pseudoexons (PEs) and an exon elongation are depicted with splice site strength predictions for WT and MT sequences, with green rectangles representing the splice acceptor sites and blue rectangles representing the splice donor sites. Red highlighted nucleotides represent the variants. Except for c.1937+37C>G (**e**), which resulted in a 36-nt exon 13 elongation, all deep-intronic variants lead to PEs (**a–d**, **f**, **g**). The intron 7 variants in (**d**) result in partially overlapping PEs that share the same splice acceptor site at position c.859-685. *HSF* Human Splicing Finder, *na* not applicable, *PE* pseudoexon, *SSFL* splice site finder–like.
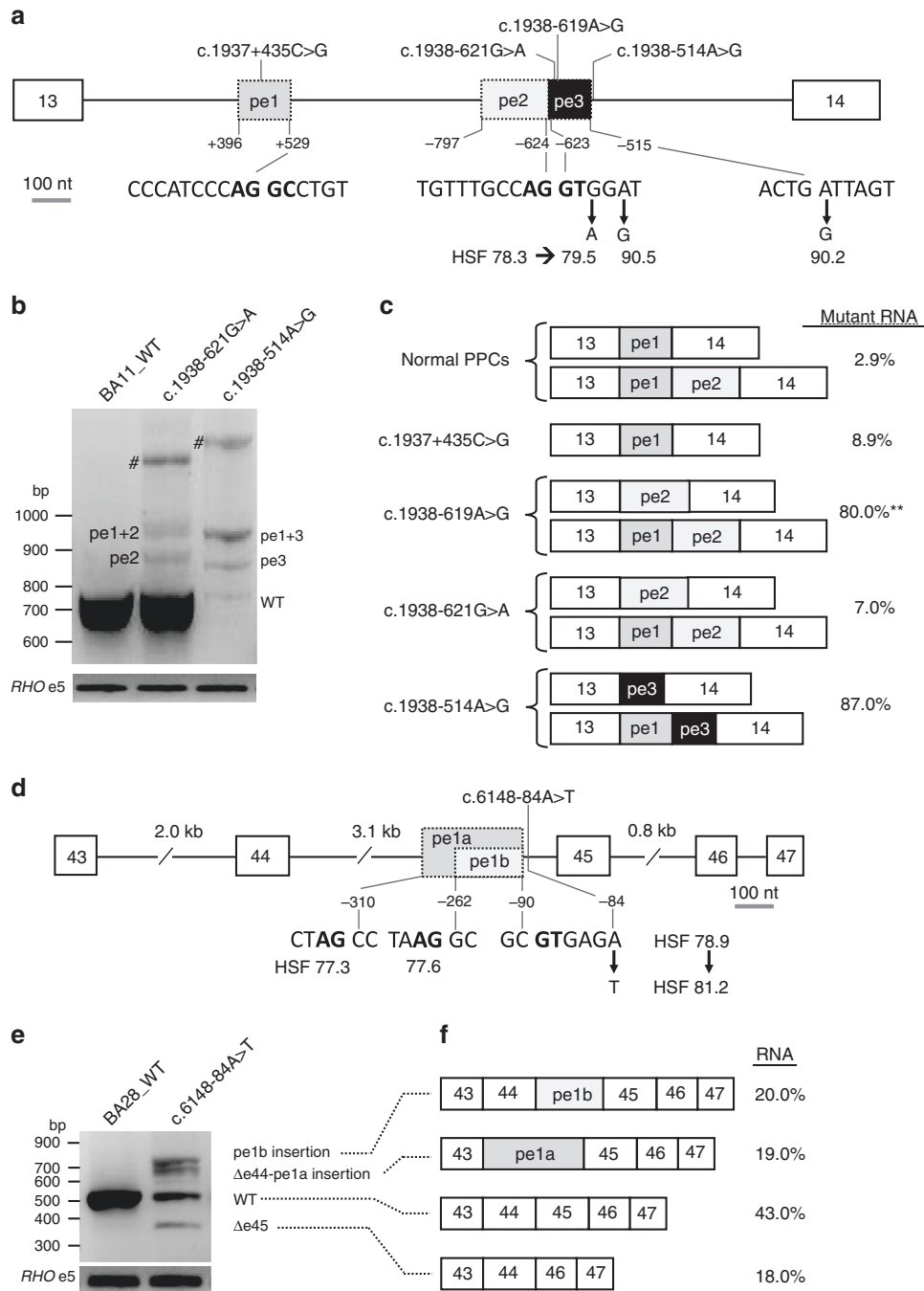
Fig. 3 Splice defects due to variants in *ABCA4* introns 13 and 44. a Genomic structure of intron 13 containing three pseudoexons (PEs) due to four deep-intronic variants. PE2 and PE3 share a splice donor site (SDS) (for PE2) and splice acceptor site (for PE3). Variants c.1938-621G>A and c.1938-619A>G strengthen the same cryptic SDS of PE2 slightly or strongly, respectively, as based on the Human Splicing Finder (HSF). Variant c.1938-514A>G creates a new strong SDS of PE3. The canonical and putative canonical splice sequences are given in bold lettering. The first and last positions of the PEs are provided. **b** Agarose gel analysis of reverse-transcription polymerase chain reaction (RT-PCR) products for intron 13 variants upon HEK293T cell splice assays. PE2 and PE3 were observed as single insertions, but also in combination with PE1. #Heteroduplex fragments of the lower bands. **c** Schematic representation of all mutant transcripts identified upon RT-PCR in HEK293T cell splice assays and of PE1 and PE1/PE2 observed as naturally occurring PEs when analyzing photoreceptor progenitor cells (PPCs) derived from a healthy individual. Interestingly, PE1 was previously shown to be induced by variant c.1937+435C>G (*Sangermano et al.)[9] and also can be part of mutant transcripts, together with PE2 or PE3. This is surprising as it is located far upstream of the other causal variants. **Reported by Fadaie et al.[13] **d** Variant c.6148-84A>T strengthens a SDS and results in PE1a or PE1b by employing upstream or downstream splice acceptor sites, respectively. These splice acceptor sites are comparable in predicted strength based on HSF. The canonical splice sequences are given in bold. **e** Agarose gel analysis of RT-PCR products due to c.6148-84A>T. **f** The largest fragment shows a 173-nt PE insertion between exons 44 and 45. The second largest band contains a 221-nt PE insertion (PE1a) and skipping of exon 44. The third-largest fragment represents the WT messenger RNA (mRNA) and the smallest fragment misses exon 45. The relative amounts of the products are listed at the right side.
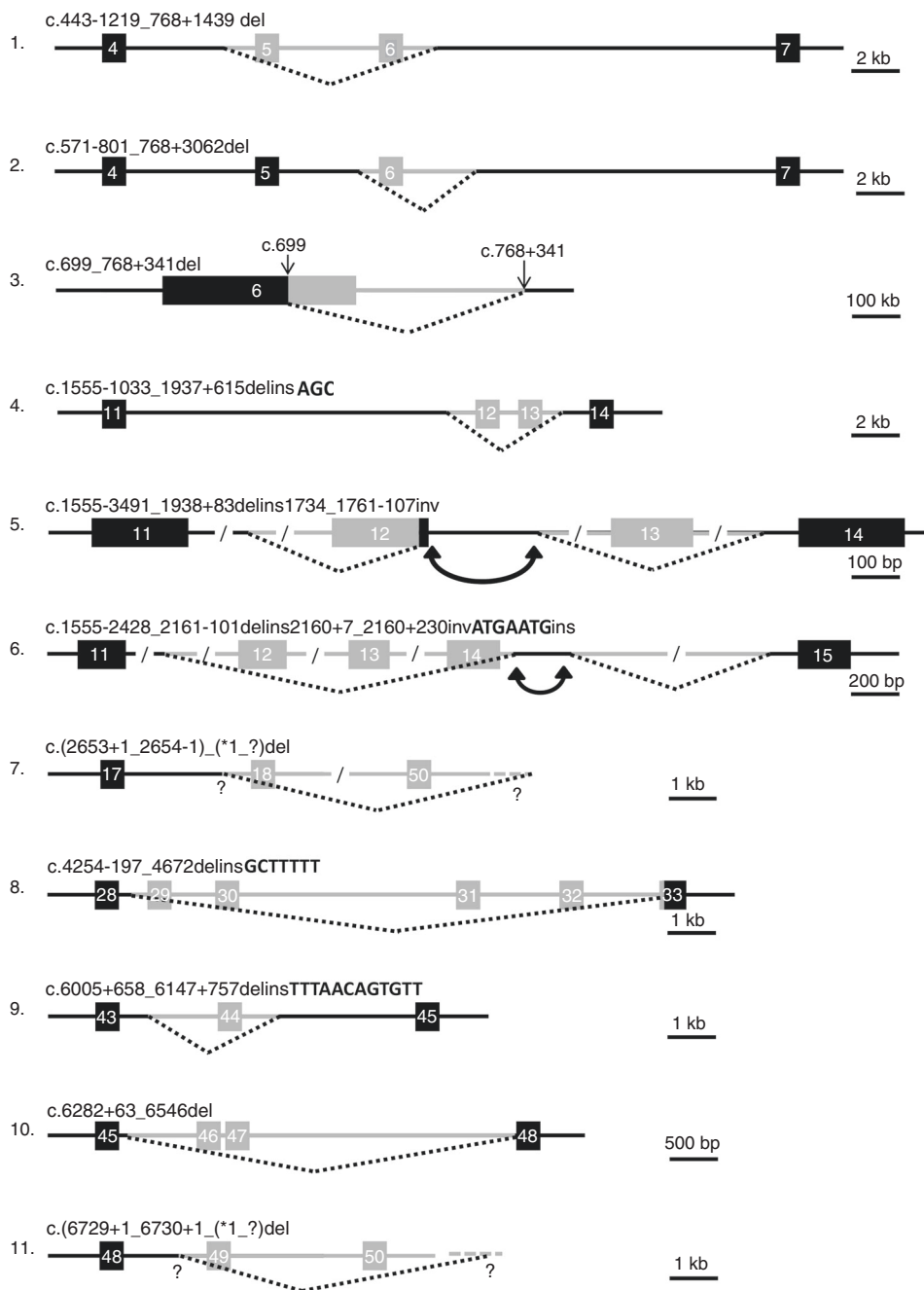
Fig. 4 Novel heterozygous structural variants in *ABCA4*. Schematic representation of the 11 structural variants identified. Exons are represented as boxes, black when they are not deleted and gray when they are deleted. Introns are represented as continuous lines, whereas stippled lines depict the deleted regions. Question marks denote that the exact location of breakpoints were unknown. Inverted double arrows represent inverted sequences.

spanning this variant, we performed CNV analysis in the proband's *ABCA4* gene. No deletion was identified (Table S17, column AU) and no heterozygous SNPs were observed in or near *ABCA4* in the proband's DNA. To test whether the chromosome 1 of the father carrying the c.859-506G>C *ABCA4* variant was passed on to the proband as two copies (UPD), exome sequencing was conducted for the proband's DNA. As shown in Fig. S9, chromosome 1 of the proband carries only homozygous SNPs, strongly suggesting the occurrence of UPD.

DISCUSSION

Employing 3866 smMIPs, 97.4% of the 128-kb *ABCA4* gene could be sequenced robustly in 1054 genetically unsolved probands with a STGD or a STGD-like phenotype. In this way, 448 (42.5%) of the probands could be genetically solved. We not only identified nine novel NCSS variants and 13 novel DI variants, but also 11 novel heterozygous SVs. The large setup of this study allowed us to provide a "landscape" overview of the different variant types underlying STGD1. As depicted in Fig. **1a**, we can appreciate that DI variants

constitute a significant cause of STGD1, i.e., 11.7% of the alleles in biallelic cases, identified in 22.5% of biallelic probands. Deletions constitute 1.8% of alleles and were found in 3.5% of biallelic cases. Seven probands carried two DI variants or one DI variant and one SV. Taken together, "dark matter" alleles were found in 113/448 (25.2%) biallelic STGD1 probands. Together, these results strongly argue for a complete sequence analysis of the *ABCA4* gene to fully appreciate its mutational landscape.

Complex splice defects due to intron 13 and 44 variants

Interestingly, the two intron 13 DI variants, i.e., c.1938-621G>A and c.1938-514G>A, were in close vicinity of two previously described variants, c.1937+435C>G[9] and c.1938-619A>G.[10,13] As shown in Fig. 3a–c the PE resulting from c.1937-514A>G (PE3) is located adjacent to PE2 as they share a dual SAS/SDS (Fig. 3a–c). The involvement of PE1, located 491, 493, and 775 nt upstream of variants c.1937-621G>A, c.1937-619A>G, and c.1937-514A>G, respectively, is very surprising. Control PPCs also show a small percentage (2.9%) of mRNAs containing PE1 or PE1–PE2. Interestingly, the SDS of PE1 also can be employed as a SAS, which, in theory could render this intronic SAS/SDS a target for recursive splicing.[25] Together, these findings suggest that there is a "natural sensitivity" for PE1 to be recognized as a PE even if the splice defect is located far downstream. Intron 44 variant c.6148-84A>T interestingly resulted in three abnormal splice products involving different PE insertions with or without flanking exon 44 or 45 deletions. Follow-up studies employing patient-derived retinal-like cells are required to validate these complex splicing patterns.

In Table S20, we listed all published 353 DI variant alleles.[5–13,18,26,27] The three most frequent are c.4253+43G>A (*n* = 100), c.4539+2001G>A (*n* = 64), and c.5196+1137G>A (*n* = 47). For some DI variants, the splice defects in HEK293T cells or patient-derived PPCs are very small (c.769-784C>T, c.1937+435G>C, c.1937-621G>A)[9,28] (this study) or smaller than expected (c.4539+2001G>A, c.4539+2028C>T). We hypothesize that retina-specific splice factors play roles that are largely missing (HEK293T cells) or underrepresented (PPCs) compared with the normal retina.

Current state of knowledge on structural variants in *ABCA4*

In this study, 11 unique SVs with sizes ranging from 411 bp to 55.7 kb were readily identified employing an easy-to-use visual detection tool taking advantage of the high number of reads obtained from smMIPs-based sequencing. Although this tool needs further automation to increase its performance for the detection of smaller deletions or duplications, it demonstrated its efficiency for deletions as small as 411 bp. To our knowledge, 47 different SVs have been identified in STGD1 patients (Table S21), 25 of which have been published elsewhere. Forty SVs are deletions, ranging in size from 23 bp to complete deletion of the *ABCA4* gene. There are six duplications, ranging from 24 bp to 26 kb, two indels, and one small insertion of 24 bp. As shown in Fig. S10, these SVs are spread over the entire gene. All SVs are rare, except for a 23-bp deletion affecting the splicing of exons 28 and 29 in 15 Israeli probands, as well as deletions spanning exons 20–22 and exon 6, both found in 6 probands, in Belgium/Germany/Netherlands and from Iberic origin, respectively, suggesting founder effects.

This genomic instability could be explained by the local genomic architecture (the presence of microhomology, repetitive elements, sequences forming non-B DNA conformations, and sequence motifs), leading to genomic rearrangements by impairing the replication process. For example, a microhomology of 1–4 bp may facilitate nonhomologous end joining (NHEJ)[29] and longer microhomologies of between 5 and 25 bp may favor microhomology-mediated end joining (MMEJ).[30] The assessment of the local architecture of deletions identified in this study lead us to rule out the non-allelic homologous recombination (NAHR) hypothesis (as no *Alu* sequence or L1 at any breakpoint was observed) and to propose the NHEJ or replication slippage models as the main implicated mechanisms (Table S18, Fig. S8). Indeed, the presence of microhomologies <5 bp in most of the junctions, and of scars characterized by insertion of several random nucleotides, could be a signature for NHEJ. Alternatively, several examples of an impaired replication fork have been noted that supports the replicative-based repair model. Indeed, despite the absence of repetitive elements of the same class at both sides of the breakpoints, their presence may initiate the formation of secondary structures, as repetitive elements could be more difficult to replicate, leading to an increased chance of replication fork stalling or collapsing.[31] Finally, Oligo(G)n tracts displayed a significant overrepresentation in the breakpoint regions. Such structures can induce tetraplex formation[32] and could also trigger rearrangement.

Uniparental isodisomy chromosome 1

UPD was found in one STGD1 case in this study, which represents the third STGD1 case showing UPD thus far reported.[33,34] UPD is a rare event, with an estimated occurrence of 1 in 5000 or even fewer individuals.[35] UPD was also described in six other inherited retinal dystrophy patients in which chromosomes 1, 2, and 6 were implicated.[36–41] We cannot exclude that there are additional UPD cases in our cohort as segregation analysis was not performed for all homozygous cases. Our finding stresses the importance of segregation analysis in the parents' DNAs as the recurrence risk for future offspring is very low in UPD families.

Missing heritability

In 174/1054 (16.5%) of probands, we identified only one (likely) causal allele. In view of the high carrier frequency of *ABCA4* variants in the general population, estimated to be ~5%,[3,14] about one-third of these monoallelic cases may be explained in this way. This may even be higher as we intentionally recruited monoallelic STGD and STGD-like probands for this study. Some causal variants may have

ARTICLE

escaped our attention. First, we have not focused on variants affecting transcription regulation. Thus far, there is limited evidence for *ABCA4* variants affecting transcription,[7] but the reported putative regulatory variants were not found in this study. As in silico tools (Alamut algorithms, SpliceAI)[42] may not predict retina-specific splice defects, we may have missed some causal variants. Also, smMIPs-based sequencing may miss heterozygous deletions smaller than ~400 bp and will not detect insertions or inversions larger than ~40 bp. In addition, more refined functional tests of coding and noncoding *ABCA4* variants are needed to understand the full genetic landscape of STGD1.

The major advantages of smMIPs-based *ABCA4* sequencing compared with genome sequencing are that it (1) is at least an order of magnitude cheaper than genome sequencing, (2) results in much smaller data storage, and (3) requires no separate informed consent regarding secondary findings. Disadvantages of smMIPs are that (1) it is restricted to one or a few genes if including introns, (2) it is more cost-effective when large series are analyzed, (3) the analysis is suitable for the detection of CNVs but not for inversions and insertions, and (4) the sequencing procedure and variant calling require a specialized setup.

In our study a significant fraction of probands carried one (likely) causal variant or c.5603A>T as a single allele (239; 22.7%) or no causal variant (364; 34.5%). A more comprehensive smMIPs-based screening platform for these STGD-like cases would likely require the sequence analysis of an additional ~80 genes associated with inherited central vision defects.

As shown in this study, smMIPs-based analysis of the complete sequence(s) of one or a few genes implicated in clinically well-defined human diseases may allow the (re) analysis of hundreds to thousands of samples, in particular by targeting cohorts in developing countries in which low-cost analysis is crucial. A similar approach can be applied to all other frequent monogenic disorders to find missing variants in noncoding regions to provide a genetic diagnosis.

In conclusion, comprehensive sequence analysis of *ABCA4* in 1054 unsolved STGD and STGD-like probands, splice assays in HEK293T cells, and SV analysis resulted in the identification of "dark matter" variants in 25% of biallelic STGD1 probands. Novel complex types of splice defects were identified for intron 13 and 44 variants. Together with published causal DI variants and SVs, a detailed genomic and transcriptomic landscape of *ABCA4*-associated STGD1 was thereby established.

SUPPLEMENTARY INFORMATION
The online version of this article (https://doi.org/10.1038/s41436-020-0787-4) contains supplementary material, which is available to authorized users.

ACKNOWLEDGEMENTS

DISCLOSURE

The authors declare no conflicts of interest.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

REFERENCES

1. Carss KJ, Arno G, Erwood M, et al. Comprehensive rare variant analysis via whole-genome sequencing to determine the molecular pathology of inherited retinal disease. Am J Hum Genet. 2017;100:75–90.
2. Blacharski PA, Newsome DA. Bilateral macular holes after Nd:YAG laser posterior capsulotomy. Am J Ophthalmol. 1988;105:417–418.
3. Cornelis SS, Bax NM, Zernant J, et al. In silico functional meta-analysis of 5,962 ABCA4 variants in 3,928 retinal dystrophy cases. Hum Mutat. 2017;38:400–408.
4. Sangermano R, Khan M, Cornelis SS, et al. ABCA4 midigenes reveal the full splice spectrum of all reported noncanonical splice site variants in Stargardt disease. Genome Res. 2018;28:100–110.
5. Khan M, Cornelis SS, Khan MI, et al. Cost-effective molecular inversion probe-based ABCA4 sequencing reveals deep-intronic variants in Stargardt disease. Hum Mutat. 2019;40:1749–1759.
6. Schulz HL, Grassmann F, Kellner U, et al. 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. Invest Ophthalmol Vis Sci. 2017;58:394–403.
7. Bauwens M, Garanto A, Sangermano R, et al. ABCA4-associated disease as a model for missing heritability in autosomal recessive disorders: novel noncoding splice, cis-regulatory, structural, and recurrent hypomorphic variants. Genet Med. 2019;21:1761–1771.
8. Braun TA, Mullins RF, Wagner AH, et al. Non-exomic and synonymous variants in ABCA4 are an important cause of Stargardt disease. Hum Mol Genet. 2013;22:5136–5145.
9. Sangermano R, Garanto A, Khan M, et al. Deep-intronic ABCA4 variants explain missing heritability in Stargardt disease and allow correction of splice defects by antisense oligonucleotides. Genet Med. 2019;21:1751–1760.
10. Zernant J, Xie YA, Ayuso C, et al. Analysis of the ABCA4 genomic locus in Stargardt disease. Hum Mol Genet. 2014;23:6797–6806.
11. Bauwens M, De Zaeytijd J, Weisschuh N, et al. An augmented ABCA4 screen targeting noncoding regions reveals a deep intronic founder variant in Belgian Stargardt patients. Hum Mutat. 2015;36:39–42.
12. Bax NM, Sangermano R, Roosing S, et al. Heterozygous deep-intronic variants and deletions in ABCA4 in persons with retinal dystrophies and one exonic ABCA4 variant. Hum Mutat. 2015;36:43–47.
13. 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. Hum Mutat. 2019;40:2365–2376.
14. Maugeri A, van Driel MA, van de Pol DJR, et al. The 2588G -> C mutation in the ABCR gene is a mild frequent founder mutation in the western European population and allows the classification of ABCR mutations in patients with Stargardt disease. Am J Hum Genet. 1999;64:1024–1035.
15. Rivera A, White K, Stohr H, et al. A comprehensive survey of sequence variation in the ABCA4 (ABCR) gene in Stargardt disease and age-related macular degeneration. Am J Hum Genet. 2000;67:800–813.
16. Jaakson K, Zernant J, Kulm M, et al. Genotyping microarray (gene chip) for the ABCR (ABCA4) gene. Hum Mutat. 2003;22:395–403.
17. Maia-Lopes S, Aguirre-Lamban J, Castelo-Branco M, Riveiro-Alvarez R, Ayuso C, Silva ED. ABCA4 mutations in Portuguese Stargardt patients: identification of new mutations and their phenotypic analysis. Mol Vis. 2009;15:584–591.
18. Albert S, Garanto A, Sangermano R, et al. Identification and rescue of splice defects caused by two neighboring deep-intronic ABCA4 mutations underlying Stargardt disease. Am J Hum Genet. 2018;102:517–527.
19. Charbel Issa P, Barnard AR, Herrmann P, Washington I, MacLaren RE. Rescue of the Stargardt phenotype in Abca4 knockout mice through inhibition of vitamin A dimerization. Proc Natl Acad Sci U S A. 2015;112:8415–8420.
20. Allocca M, Doria M, Petrillo M, et al. Serotype-dependent packaging of large genes in adeno-associated viral vectors results in effective gene delivery in mice. J Clin Invest. 2008;118:1955–1964.
21. Lu B, Malcuit C, Wang S, et al. Long-term safety and function of RPE from human embryonic stem cells in preclinical models of macular degeneration. Stem Cells. 2009;27:2126–2135.
22. Cremers FPM, Cornelis SS, Runhart EH, Astuti GDN. Author response: penetrance of the ABCA4 p.Asn1868Ile allele in Stargardt disease. Invest Ophthalmol Vis Sci. 2018;59:5566–5568.
23. Runhart EH, Sangermano R, Cornelis SS, et al. The common ABCA4 variant p.Asn1868Ile shows nonpenetrance and variable expression of Stargardt disease when present in trans with severe variants. Invest Ophthalmol Vis Sci. 2018;59:3220–3231.
24. Zernant J, Lee W, Collison FT, et al. Frequent hypomorphic alleles account for a significant fraction of ABCA4 disease and distinguish it from age-related macular degeneration. J Med Genet. 2017;54:404–412.
25. Hafez M, Hausner G. Convergent evolution of twintron-like configurations: one is never enough. RNA Biol. 2015;12:1275–1288.
26. Zernant J, Lee W, Nagasaki T, et al. Extremely hypomorphic and severe deep intronic variants in the ABCA4 locus result in varying Stargardt disease phenotypes. Cold Spring Harb Mol Case Stud. 2018;4:a002733.
27. Nassisi M, Mohand-Said S, Andrieu C, et al. Prevalence of ABCA4 deep-intronic variants and related phenotype in an unsolved "one-hit" cohort with Stargardt disease. Int J Mol Sci. 2019;20:5053.
28. Runhart EH, Valkenburg D, Cornelis SS, et al. Late-onset Stargardt disease due to mild, deep-intronic ABCA4 alleles. Invest Ophthalmol Vis Sci. 2019;60:4249–4256.
29. Lieber MR. The mechanism of double-strand DNA break repair by the nonhomologous DNA end-joining pathway. Annu Rev Biochem. 2010; 79:181–211.
30. McVey M, Lee SE. MMEJ repair of double-strand breaks (director's cut): deleted sequences and alternative endings. Trends Genet. 2008;24: 529–538.
31. Vissers LE, Bhatt SS, Janssen IM, et al. Rare pathogenic microdeletions and tandem duplications are microhomology-mediated and stimulated by local genomic architecture. Hum Mol Genet. 2009;18:3579–3593.
32. Bacolla A, Wells RD. Non-B DNA conformations, genomic rearrangements, and human disease. J Biol Chem. 2004;279:47411–47414.
33. Fingert JH, Eliason DA, Phillips NC, Lotery AJ, Sheffield VC, Stone EM. Case of Stargardt disease caused by uniparental isodisomy. Arch Ophthalmol. 2006;124:744–745.
34. Riveiro-Alvarez R, Valverde D, Lorda-Sanchez I, et al. Partial paternal uniparental disomy (UPD) of chromosome 1 in a patient with Stargardt disease. Mol Vis. 2007;13:96–101.
35. Liehr T. Cytogenetic contribution to uniparental disomy (UPD). Mol Cytogenet. 2010;3:8.
36. Rivolta C, Berson EL, Dryja TP. Paternal uniparental heterodisomy with partial isodisomy of chromosome 1 in a patient with retinitis pigmentosa without hearing loss and a missense mutation in the Usher syndrome type II gene USH2A. Arch Ophthalmol. 2002;120:1566–1571.

37. Roosing S, van den Born LI, Hoyng CB, et al. Maternal uniparental isodisomy of chromosome 6 reveals a TULP1 mutation as a novel cause of cone dysfunction. Ophthalmology. 2013;120: 1239–1246.

38. Thompson DA, Gyurus P, Fleischer LL, et al. Genetics and phenotypes of RPE65 mutations in inherited retinal degeneration. Invest Ophthalmol Vis Sci. 2000;41:4293–4299.

39. Thompson DA, McHenry CL, Li Y, et al. Retinal dystrophy due to paternal isodisomy for chromosome 1 or chromosome 2, with homoallelism for mutations in RPE65 or MERTK, respectively. Am J Hum Genet. 2002; 70:224–229.

40. Wiszniewski W, Lewis RA, Lupski JR. Achromatopsia: the CNGB3 p. T383fsX mutation results from a founder effect and is responsible for the visual phenotype in the original report of uniparental disomy 14. Hum Genet. 2007;121:433–439.

41. Souzeau E, Thompson JA, McLaren TL, et al. Maternal uniparental isodisomy of chromosome 6 unmasks a novel variant in TULP1 in a patient with early onset retinal dystrophy. Mol Vis. 2018;24:478–484.

42. Jaganathan K, Kyriazopoulou Panagiotopoulou S, McRae JF, et al. Predicting splicing from primary sequence with deep learning. Cell. 2019;176:535–48 e524.

Mubeen Khan, MSc[1,2], Stéphanie S. Cornelis, MSc[1,2], Marta Del Pozo-Valero, MSc[1,3], Laura Whelan, BSc[4], Esmee H. Runhart, MD[2,5], Ketan Mishra, MSc[1,2], Femke Bults, BSc[1], Yahya AlSwaiti, MD[6], Alaa AlTalbishi, MD, PhD[6], Elfride De Baere, MD, PhD[7], Sandro Banfi, MD[8], Eyal Banin, MD, PhD[9], Miriam Bauwens, PhD[7], Tamar Ben-Yosef, PhD[10], Camiel J. F. Boon, MD, PhD[11,12], L. Ingeborgh van den Born, MD, PhD[13,14], Sabine Defoort, MD[15], Aurore Devos, MSc[16], Adrian Dockery, PhD[4], Lubica Dudakova, PhD[17], Ana Fakin, MD, PhD[18], G. Jane Farrar, PhD[4], Juliana Maria Ferraz Sallum, MD, PhD[19,20], Kaoru Fujinami, MD, PhD[21,22,23,24], Christian Gilissen, PhD[1,25], Damjan Glavač, PhD[26], Michael B. Gorin, MD, PhD[27,28], Jacquie Greenberg, PhD[29], Takaaki Hayashi, MD, PhD[30], Ymkje M. Hettinga, MD[31], Alexander Hoischen, PhD[1], Carel B. Hoyng, MD, PhD[2,5], Karsten Hufendiek, MD[32], Herbert Jägle, MD[33], Smaragda Kamakari, PhD[34], Marianthi Karali, PhD[8], Ulrich Kellner, MD[35,36], Caroline C. W. Klaver, MD, PhD[5,37,38], Bohdan Kousal, MD[17,39], Tina M. Lamey, PhD[40,41], Ian M. MacDonald, MD, CM[42], Anna Matynia, PhD[27,28], Terri L. McLaren, BSc[40,41], Marcela D. Mena, PhD[43], Isabelle Meunier, MD, PhD[44], Rianne Miller, BSc[1], Hadas Newman, MD[45,46], Buhle Ntozini, BSc[29], Monika Oldak, MD, PhD[47], Marc Pieterse, MSc[1], Osvaldo L. Podhajcer, PhD[43], Bernard Puech, MD[15], Raj Ramesar, MBA, PhD[29], Klaus Rüther, MD, PhD[48], Manar Salameh, BSc[6], Mariana Vallim Salles, MD[19,20], Dror Sharon, PhD[9], Francesca Simonelli, MD, PhD[49], Georg Spital, MD[50], Marloes Steehouwer, BSc[1], Jacek P. Szaflik, MD, PhD[51], Jennifer A. Thompson, PhD[41], Caroline Thuillier, MSc[52], Anna M. Tracewska, PhD[53], Martine van Zweeden, BSc[1], Andrea L. Vincent, MBChB, MD[54,55], Xavier Zanlonghi, MD, PhD[56], Petra Liskova, MD, PhD[17,39], Heidi Stöhr, PhD[57], John N. De Roach, PhD[40,41], Carmen Ayuso, MD, PhD[3], Lisa Roberts, PhD[29], Bernhard H. F. Weber, PhD[57], Claire-Marie Dhaenens, MD, PhD[1,16] and Frans P. M. Cremers, PhD[1,2]

[1]Department of Human Genetics, Radboud University Medical Center, Nijmegen, The Netherlands; [2]Donders Institute for Brain, Cognition and Behaviour, Radboud University Medical Center, Nijmegen, The Netherlands; [3]Department of Genetics, IIS-Fundación Jiménez Díaz, CIBERER, Madrid, Spain; [4]The School of Genetics & Microbiology, Trinity College Dublin, Dublin, Ireland; [5]Department of Ophthalmology, Radboud University Medical Center, Nijmegen, The Netherlands; [6]St John of Jerusalem Eye Hospital Group, East Jerusalem, Palestine; [7]Center for Medical Genetics Ghent, Ghent University and Ghent University Hospital, Ghent, Belgium; [8]Department of Precision Medicine, University of Campania Luigi Vanvitelli, Naples and Telethon Institute of Genetics and Medicine (TIGEM), Pozzuoli, Italy; [9]Department of Ophthalmology, Hadassah Medical Center, Faculty of Medicine, The Hebrew University of Jerusalem, Jerusalem, Israel; [10]Ruth and Bruce Rappaport Faculty of Medicine, Technion-Israel Institute of Technology, Haifa, Israel; [11]Department of Ophthalmology, Leiden University Medical Center, Leiden, The Netherlands; [12]Department of Ophthalmology, Amsterdam University Medical Centers, Amsterdam, The Netherlands; [13]The Rotterdam Eye Hospital, Rotterdam, The Netherlands; [14]The Rotterdam Ophthalmic Institute, Rotterdam, The Netherlands; [15]Service d'exploration de la vision et neuro-ophtalmologie, Centre Hospitalier Universitaire de Lille, Lille, France; [16]Univ. Lille, Inserm, CHU Lille, U1172 - LilNCog - Lille Neuroscience & Cognition, F-59000 Lille, France; [17]Research Unit for Rare Diseases, Department of Paediatrics and Adolescent Medicine, First Faculty of Medicine, Charles University and General University Hospital in Prague, Prague, Czech Republic; [18]Eye Hospital, University Medical Centre Ljubljana, Ljubljana, Slovenia; [19]Department of Ophthalmology and Visual Sciences, Universidade Federal de São Paulo, São Paulo, SP, Brazil; [20]Instituto de Genética Ocular, São Paulo, SP, Brazil; [21]UCL Institute of Ophthalmology, London, UK; [22]Laboratory of Visual Physiology, Division of Vision Research, National Institute of Sensory Organs, National Hospital Organization Tokyo Medical Center, Tokyo, Japan; [23]Graduate School of Health Management, Keio University, Tokyo, Japan; [24]Moorfields Eye Hospital, London, UK; [25]Radboud Institute for Molecular Life Sciences, Radboud University Medical Center, Nijmegen, The Netherlands; [26]Department of Molecular Genetics, Institute of Pathology, University of

Ljubljana, Ljubljana, Slovenia; [27]Department of Ophthalmology, David Geffen School of Medicine, Stein Eye Institute, University of California–Los Angeles, Los Angeles, CA, USA; [28]Department of Human Genetics, David Geffen School of Medicine, University of California–Los Angeles, Los Angeles, CA, USA; [29]University of Cape Town/MRC Genomic and Precision Medicine Research Unit, Division of Human Genetics, Department of Pathology, Institute of Infectious Disease and Molecular Medicine (IDM), Faculty of Health Sciences, University of Cape Town, Cape Town, South Africa; [30]Department of Ophthalmology, The Jikei University School of Medicine, Tokyo, Japan; [31]Bartiméus Diagnostic Center for Complex Visual Disorders, Zeist, The Netherlands; [32]University Eye Hospital Hannover Medical School, Hannover, Germany; [33]Department of Ophthalmology, University Hospital, University Regensburg, Regensburg, Germany; [34]Ophthalmic Genetics Unit, OMMA Ophthalmological Institute of Athens, Athens, Greece; [35]Rare Retinal Disease Center, AugenZentrum Siegburg, MVZ ADTC Siegburg GmbH, Siegburg, Germany; [36]RetinaScience, Bonn, Germany; [37]Department of Ophthalmology, Erasmus Medical Centre, Rotterdam, The Netherlands; [38]Department of Epidemiology, Erasmus Medical Centre, Rotterdam, The Netherlands; [39]Department of Ophthalmology, First Faculty of Medicine, Charles University and General University Hospital in Prague, Prague, Czech Republic; [40]Centre for Ophthalmology and Visual Science, The University of Western Australia, Nedlands, WA, Australia; [41]Australian Inherited Retinal Disease Registry and DNA Bank, Department of Medical Technology and Physics, Sir Charles Gairdner Hospital, Nedlands, WA, Australia; [42]Departments of Ophthalmology and Medical Genetics, University of Alberta, Edmonton, AB, Canada; [43]Laboratory of Molecular and Cellular Therapy, Fundacion Instituto Leloir-CONICET, Buenos Aires, Argentina; [44]Institut des Neurosciences de Montpellier, INSERM, Université de Montpellier, Montpellier, France; [45]Department of Ophthalmology, Tel Aviv Sourasky Medical Center, Tel Aviv, Israel; [46]Sackler Faculty of Medicine, Tel Aviv University, Tel Aviv, Israel; [47]Department of Histology and Embryology, Medical University of Warsaw, Warsaw, Poland; [48]Augenarztpraxis, Dorotheenstraße, Berlin, Germany; [49]Eye Clinic, Multidisciplinary Department of Medical, Surgical and Dental Sciences, University of Campania Luigi Vanvitelli, Naples, Italy; [50]Department of Ophthalmology, St. Franziskus-Hospital, Münster, Germany; [51]Department of Ophthalmology, Medical University of Warsaw SPKSO Ophthalmic University Hospital, Warsaw, Poland; [52]CHU Lille, Institut de Génétique Médicale, Lille, France; [53]DNA Analysis Unit, ŁUKASIEWICZ Research Network—PORT Polish Center for Technology Development, Wroclaw, Poland; [54]Department of Ophthalmology, New Zealand National Eye Centre, Faculty of Medical and Health Sciences, The University of Auckland, Grafton, Auckland, New Zealand; [55]Eye Department, Greenlane Clinical Centre, Auckland District Health Board, Auckland, New Zealand; [56]Centre de Compétence Maladie Rare, Clinique Jules Verne, Nantes, France; [57]Institute of Human Genetics, University of Regensburg, Regensburg, Germany.

Descripción genética y fenotípica de la cohorte de pacientes *ABCA4*

Artículo 4: Genotype-phenotype correlations in a Spanish cohort of 506 families with bi-allelic *ABCA4* mutations

Del Pozo-Valero M, Riveiro-Alvarez R, Blanco-Kelly F, et al.

Resumen:

En este trabajo se recogió toda la información genética y clínica disponible de 506 familias caracterizadas con el gen *ABCA4* en nuestra cohorte con el objetivo de describir el espectro mutacional y fenotípico del gen *ABCA4* en la cohorte más grande publicada hasta la fecha.

Los pacientes se habían caracterizado con el gen *ABCA4* utilizando diferentes tecnologías dependiendo del momento del estudio, y se clasificaron según sus hallazgos oftalmológicos en pacientes con enfermedad de Stargardt o distrofia de conos y bastones. A su vez, las variantes identificadas se clasificaron en *missense* o truncantes para hacer las correlaciones genotipo-fenotipo.

En cuanto a los hallazgos genéticos, se identificaron 228 variantes, de las cuales 50 no se habían descrito previamente. La mayoría de las variantes identificadas eran variantes *missense*, seguidas de las variantes *frameshift* y *stop*. Por su parte, las variantes *deep intronic* y CNVs solo representaban el 3% de las variantes identificadas. Sin embargo, se identificó una nueva deleción de parte del exón e intrón 6 del gen *ABCA4* en 6 familias de nuestra cohorte, y el estudio de sus haplotipos reveló un origen común de esta mutación en todas las familias.

En cuanto al fenotipo, se demostró, con datos de edad de inicio de los síntomas de 372 pacientes, que la mayoría de los pacientes con STGD1 presentaban variantes *missense*, mientras que los pacientes con DCB presentaban en su mayoría variantes truncantes, sugiriendo la posible disfunción de bastones en los casos diagnosticados con STGD1 con variantes truncantes. Sin embargo, también se asociaron a DCB 3 variantes específicas *missense*, c.1804C>T; p.(Arg602Trp), c.3056C>T; p.(Thr1019Met) y c.6320G>C; p.(Arg2107Pro) y solo una variante a STGD1, c.3386G>T; p.(Arg1129Leu). Además, se observó que la distribución de la edad de inicio en estos dos grupos de pacientes era estadísticamente significativa, presentando una uniformidad en la aparición más temprana de los síntomas en los pacientes con DCB. También se realizaron estas correlaciones con las variantes más prevalentes en nuestra cohorte, la c.3386G>T; p.(Arg1129Leu) y c.5882G>A; p.(Gly1961Glu), observando una correlación entre los homocigotos y una evolución más leve de la enfermedad, ya que en el caso de la segunda variante no identificamos pacientes homocigotos en nuestra cohorte. También se identificaron

pacientes con respeto foveal, en los que la variante c.3386G>T; p.(Arg1129Leu) estaba sobrerrepresentada.

Con este trabajo se reveló el espectro mutacional del gen *ABCA4* en la cohorte más grande reportada hasta la fecha, actualizando el modelo de genotipo-fenotipo descrito previamente, el cual contaba con un número limitado de pacientes. Además, se pone de manifiesto que las variantes truncantes y ciertas *missense* podrían provocar una disfunción de los bastones, por lo que el diagnóstico genético es fundamental para poder establecer el pronóstico de la enfermedad de estos pacientes.

Contribución de la autora:

La autora participó en el diseño del estudio, el análisis y los estudios de segregación de algunas familias caracterizadas en este periodo, así como los estudios de haplotipos de las familias con la nueva deleción. Recopiló todos los datos genéticos y clínicos de los pacientes incluidos en este estudio, curando la anotación de variantes y clasificándoles según el tipo de variante y su fenotipo clínico. Participó en el análisis de los datos estadísticos de las correlaciones genotipo-fenotipo, evaluando las conclusiones más importantes de este trabajo. Finalmente, escribió y revisó críticamente las sucesivas versiones de este manuscrito y participó en su proceso de publicación.

Genotype-phenotype correlations in a Spanish cohort of 506 families with bi-allelic *ABCA4* pathogenic variants

Marta Del Pozo-Valero[1,2], MSc; Rosa Riveiro-Alvarez[1,2], PhD; Fiona Blanco-Kelly[1,2], MD, PhD; Jana Aguirre-Lamban[1], PhD; Inmaculada Martin-Merida[1,2], PhD; Ionut-Florin Iancu[1,2], MSc; Saoud Swafiri[1,2], MD; Isabel Lorda-Sanchez[1], MD, PhD; Elvira Rodriguez-Pinilla[1], MD, PhD; Mª Jose Trujillo-Tiebas[1,2], PhD; Belen Jimenez-Rolando[3], MD; Ester Carreño[3], MD, PhD; Ignacio Mahillo-Fernandez[4], PhD; Carlo Rivolta[5,6,7], PhD; Marta Corton[1,2], PhD; Almudena Avila-Fernandez[1,2], PhD; Blanca García Sandoval[2,3], MD, PhD; Carmen Ayuso[1,2], MD, PhD.

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

2. Center for Biomedical Network Research on Rare Diseases (CIBERER), ISCIII, Madrid, Spain.

3. Department of Ophthalmology, Instituto de Investigación Sanitaria–Fundación Jiménez Díaz University Hospital, Universidad Autónoma de Madrid (IIS-FJD, UAM), Madrid, Spain.

4. Department of Epidemiology, Instituto de Investigación Sanitaria–Fundación Jiménez Díaz University Hospital, Universidad Autónoma de Madrid (IIS-FJD, UAM), Madrid, Spain.

5. Institute of Molecular and Clinical Ophthalmology Basel (IOB), Basel, Switzerland

6. Department of Ophthalmology, University Hospital Basel, Switzerland.

7. Department of Genetics and Genome Biology, University of Leicester, Leicester, United Kingdom.

*Correspondence to: Carmen Ayuso: cayuso@fjd.es. Servicio de Genética, Instituto de Investigación Sanitaria-Fundación Jiménez Díaz, Universidad Autónoma de Madrid, Av. Reyes Católicos, 2. Madrid, 28040 Spain. Tel: +34 609612728; Fax: +34 915504849

Supplemental Material available at AJO.com

Short title: Genotype-phenotype correlations in 506 *ABCA4* families

Competing interests

The authors declare that they have no competing interests

ABSTRACT

Purpose: To define genotype-phenotype correlations in the largest cohort study worldwide of patients: 434 with Stargardt disease (STGD1) and 72 with cone-rod dystrophy (CRD), all carrying biallelic ABCA4 variants.

Design: Cohort study.

Methods: We characterized 506 patients with ABCA4 variants using conventional genetic tools and Next Generation Sequencing technologies. Medical history and ophthalmological data were obtained from 372 patients. Genotype-phenotype correlation studies were carried out for the following variables: variant type, age at onset of symptoms (AO), and clinical phenotype.

Results: A total of 228 different pathogenic variants were identified in 506 ABCA4 patients, 50 of which were novel. Genotype-phenotype correlations showed that most of the patients with biallelic truncating variants presented CRD and these cases had a significantly earlier AO than STGD1 patients. Three missense variants are associated with CRD for first time (c.1804C>T;p.(Arg602Trp), c.3056C>T;p.(Thr1019Met), c.6320G>C;p.(Arg2107Pro)). Analysis of the most prevalent ABCA4 variant in Spain, c.3386G>T;p.(Arg1129Leu), revealed that is correlated to STGD1, a later AO and foveal sparing.

Conclusions: Our study, conducted in the largest ABCA4-associated disease cohort reported to date, updates the genotype-phenotype model established for ABCA4 variants, and broadens the mutational spectrum of the gene. According to our observations, ABCA4 patients presenting with two truncating variants may first present features of STGD1 but eventually develop rod dysfunction, and specific missense variants may be associated with a different phenotype, underscoring the importance of an accurate genetic diagnosis. Also, it is a prerequisite for enrollment in clinical trials, and to date, no other treatment has been approved for STGD1.

INTRODUCTION

Causative variants in the *ABCA4* gene (photoreceptor-specific ATP-binding cassette transporter 4; MIM: 601691) are associated with several inherited retinal dystrophies (IRDs). Biallelic *ABCA4* variants are mostly found in patients with Stargardt disease (STGD1)(Allikmets, Singh, *et al.*, 1997) but have also been described in cone-rod dystrophy (CRD) and retinitis pigmentosa (RP) patients (Frans P M Cremers *et al.*, 1998; Martínez-Mir *et al.*, 1998).

ABCA4 comprises 50 exons and encodes the multidomain transmembrane protein ABCA4, located at the rim of disc membranes in the outer segments of both cone and rod photoreceptors of the human retina (Molday, Rabin and Molday, 2000). The role of ABCA4 in the visual cycle is to transport or flip N-retinylidene-phosphatidylethanolamine (PE) from the lumen to the cytoplasmic side of the disc membrane (Molday, Rabin and Molday, 2000). Mutant ABCA4 proteins usually induce the accumulation in disc membranes of all-trans retinal and N-retinylidene-PE, which react to produce fluorophore A2E precursors, leading to photoreceptor degeneration (Molday, 2007).

STGD1 (#248200) is the most common juvenile macular dystrophy, with an estimated prevalence of 1:10000 and a carrier frequency of approximately 2% (Blacharski and PA, 1988). However, previous studies suggested a higher prevalence (6%) of carriers in Spain(Riveiro-Alvarez *et al.*, 2013). STGD1 is characterized by a disease onset usually within the first two

decades of life-but also early and late-onset cases exist- affecting central vision. Ophthalmoscopic examinations reveal atrophy of the retinal pigment epithelium, and presence of yellow flecks around the macula and midperiphery (Anderson *et al.*, 1995). In contrast, CRD is defined as a progressive loss of cone function followed by rod-function loss, resulting in further impairment of peripheral vision and night blindness. Ophthalmoscopic examinations in CRD patients showed perifoveal atrophy of the outer retina and bull's eye maculopathy (Michaelides *et al.*, 2006; Hamel, 2007).

To explain the differences in the clinical IRD subtypes induced by *ABCA4* variants, a genotype-phenotype model was proposed, based on the functional consequences of the combination of *ABCA4* variants(van Driel *et al.*, 1998; Maugeri *et al.*, 2000). Persons carrying two severe variants are expected to present severe forms of CRD, which could be misdiagnosed as RP due to the progression of the disease, which closely resembles CRD (Riveiro-Alvarez *et al.*, 2013).

To date, more than 1200 *ABCA4* variants have been reported in the Human Gene Mutation Database (HGMD). Most STGD1 patients seem to carry biallelic coding *ABCA4* variants, whereas unsolved cases carrying no *ABCA4* mutated alleles or one such allele can be explained by the presence of deep intronic variants (Braun *et al.*, 2013; Zernant *et al.*, 2014; Bauwens *et al.*, 2019; Khan *et al.*, 2019; Sangermano *et al.*, 2019) or the low penetrant c.5603A>T; p.(Asn1868Ile) variant (Zernant *et al.*, 2017; Runhart *et al.*, 2018).

In this study, we report findings from the largest cohort of *ABCA4* patients described to date, consisting of 506 families with biallelic variants. In addition to precisely assess the prevalence of *ABCA4* variants in this Spanish cohort, we describe new genotype-phenotype correlations for *ABCA4* causal variants and STGD1 or CRD phenotypes.

MATERIAL AND METHODS

Subjects and samples

Five hundred six Spanish families with a clinical diagnosis of STGD1 or CRD were recruited at the Fundación Jiménez Díaz University Hospital (FJD, Madrid, Spain). A solved genotype with biallelic *ABCA4* variants was used for the inclusion criteria. This study was performed in accordance with the tenets of the Helsinki Declaration and subsequent reviews, and the procedure for patient enrolment was approved by the Research Ethics Committee of the Fundación Jimenez Diaz University Hospital. DNA samples were collected from the FJD biobank. Informed consent was obtained from all subjects.

Molecular screening

Index cases from 506 unrelated families that had undergone molecular characterization over the past 29 years. A total of 299 index cases were characterized using conventional genetic tools described before (Valverde *et al.*, 2006; Riveiro-Alvarez *et al.*, 2013) and 207 index cases were studied by different NGS strategies, including targeted gene panels, clinical exome, and/or whole-exome sequencing (WES), as previously described (Martin-Merida *et al.*, 2018; Del Pozo-Valero *et al.*, 2019). Depending on the screening technique used at the time of diagnosis, subjects with one identified *ABCA4* allele underwent either Sanger sequencing of known deep intronic variants or multiplex ligation probe amplification (MLPA) using *ABCA4* probes (Probemis P-151 and P-152) (MRC-Holland, Amsterdam) or copy number variation (CNV) analysis of NGS

data or a combination of these. Additionally, to complete the genotype data for 34 cases, the entire *ABCA4* gene was sequenced using smMIPs-based technology(Khan *et al.*, 2020).

The pathogenicity of *ABCA4* variants was established according to their allele frequency appearing in gnomAD (http://gnomad.broadinstitute.org/); *in silico* prediction tools were used to classify new splice and missense variants, including SIFT(Sim *et al.*, 2012), PolyPhen(Adzhubei *et al.*, 2010), CADD(Kircher *et al.*, 2014) and M-CAP(Jagadeesh *et al.*, 2016). Additionally, we conducted cosegregation studies in family members when other relatives were available for study. For variant classification, we followed the guidelines of the American College of Medical Genetics and Genomics (ACMG)(Richards *et al.*, 2015) and the recent study by Cornelis et al. (Cornelis *et al.*, 2017). Stop, frameshift, and splice variants were considered as truncating variants due to their presumable effect on the protein, including unreported non-canonical splice site variants (Khan *et al.*, 2020). Complex alleles are defined when 2 *ABCA4* variants were present on the same allele. Complex alleles carrying a truncating variant were considered to be truncating alleles.

Five microsatellite markers (D1S2804, D1S2868, D1S236, D1S2664, and D1S2793) and 3 SNPs (rs769211, rs1801555, and rs4148058) flanking 3.74 Mb around *ABCA4* were studied in 6 families with the variant c.699_768+341del.

Clinical assessment

A comprehensive review of the ophthalmological data available in the clinical examination notes of the 506 *ABCA4* patients was carried out to record the following data: age at onset of visual acuity (VA) loss, visual field (VF) constriction, and night blindness (NB); best-corrected visual acuity (BCVA) measurements, in decimal scale; full field electroretinography (ffERG) responses; and fundus appearance. In some cases, spectral domain optical coherence tomography (SD-OCT) and fundus autofluorescence (FAF) images were examined. The age of onset (AO) of the disease was defined as the patient age at visual acuity loss or at initial diagnosis.

Diagnosis of STGD1 or CRD were based on the following criteria:

Diagnosis of STGD1 was determined according to initial symptoms of visual acuity loss; fundus images showing orange-yellow flecks in the retina, a beaten-bronze appearance; and normal or cone altered ffERG results.

Diagnosis of CRD was based on initial symptoms of loss of central vision and/or night blindness; fundus images showing atrophic macular degeneration and peripheral alterations including pigment epithelial thinning, pigment deposits, or both; and a decrease in cone-rod ffERG responses.

When not available clinical information, the diagnosis referred by their ophthalmologists was used.

Genotype-phenotype correlations

To perform genotype-phenotype correlations, all 1012 alleles from the 506 families were classified into the following 12 categories:

- Categories A, B, and C: Patients carrying missense-missense, missense-truncating and truncating-truncating variants, respectively;

- Categories A2, B2, and C2: Patients carrying missense-missense, missense-truncating, and truncating-truncating variants excluding the c.3386G>T variant, respectively;

- Categories D, E, and F: Patients carrying the c.3386G>T variant in homozygosis, in combination with a different missense variant, and in combination with a truncating variant, respectively; and

- Categories G, H, I: Patients carrying the c.5882G>A variant in homozygosis, in combination with a different missense variant, and in combination with a truncating variant, respectively.

Categories were compared based on AO and clinical diagnosis of STGD1 or CRD. Due to the non-normal distribution of the data, Wilcoxon rank sum test was used to perform comparisons between groups. Medians and interquartile ranges (IQR) were represented. For missense variants, 95% confidence intervals for percentages were calculated using the binomial exact method. Odds ratios (OR) and their respective 95% confidence intervals (CI), were calculated by median unbiased estimation. Statistical analyses and graphical representation were done using R version 3.6.0.

RESULTS

Mutational spectrum of *ABCA4* variants

A total of 228 different variants in the *ABCA4* gene were found in 1012 alleles from our Spanish cohort of 506 index patients (Supplementary Tables S1 and S2). Their classification by type of variant is shown in Figure 1A.

Thirty-three different variants were part of 21 complex assortments and accounted for 5% (48/1012) of all alleles. Ten variants were only found in complex alleles and not as single alleles (Supplementary Table S2 and Figure 1A and 1B). The following were the 3 most frequent complex variant combinations: the previously reported c.[3322C>T;6320G>A] and c.[4926C>G;5044_5058del] (Cornelis *et al.*, 2017), as well as one novel variant, c.[3386G>T;6718A>G], representing 12.5% of the total complex alleles in our Spanish cohort. In 7 families with 3 *ABCA4* variants identified, the correct phase could not be established since no samples from relatives were available, therefore these 7 complex alleles could be in other combination in these patients (Table S3).

The most frequent variants are shown in Table 1, with the missense c.3386G>T; p.(Arg1129Leu) being present in 33.6% of the patients with an allelic frequency of 18.8% (190/1012). This variant was found in 183 single and 7 complex alleles.

In this genetic screening, 50 variants were as yet unpublished, representing 22% of the total number of different variants and 7.2% of all patient alleles (73/1012) (Figure 1C and Supplementary Table S4). Three (c.6071A>G, c.2481del and c.2483C>T) were present as 2 complex allele assortments, since the last 2 variants were observed in *cis*. All novel variants were

segregating with the disease or predicted as pathogenic by at least 3 of the 4 programs used, and their population frequency was absent or <0.002.

CNVs were found in 7 families, representing 0.7% of the total number of alleles. Family MD-0401 carried a deletion of intron 11. Besides, a novel 411bp deletion [c.699_768+341del; p.(Gln234Phefs*5)] covering 70 bp of exon 6 and 341 bp of intron 6 was identified in 6 unrelated Spanish families (MD-0162, MD-0039, RP-2668, MD-0166, MD-0460 and RP-2531). This deletion was found in a heterozygous state in 5 families and in homozygosis in one family. Segregation studies confirmed the presence of the novel deletion in combination with a second unshared variant in *trans* in 4 families. Haplotype analysis of 8 markers in *ABCA4* revealed a common minimal and maximal shared region of 1.58 Mb (chr1:94360107-95946135) and 3.29Mb (chr1:93335742-96628133) in all the families, respectively (Supplementary Figure S1). In addition, 3 families shared the same haplotype for all the markers used (MD-0039, MD-0162, and RP-2531).

Deep intronic variants were found in 28 patients, with 2.8% allele frequency (Supplementary Table S5). The most prevalent was c.4539+2064C>T, which was present in 14 patients (one homozygote), representing 1.5% of all alleles.

The screening of the complete *ABCA4* gene in 7 patients with c.6148G>C; p.(Val2050Leu) a variant that was previously classified as pathogenic but now recognized as benign, allowed us to identify additional pathogenic variants in all of them. In addition, 9 patients with the low-penetrant variant c.5603A>T; p.(Asn1868Ile) also underwent this screening. In this case, further pathogenic intronic variants in *cis* were found in only 2 of them (MD-1075 and MD-1279) (Supplementary Table S1 and S3).

Homozygous variants were carried by 81 patients (Figure 1D). Cosegregation and existence of consanguinity or endogamy allowed us to confirm their homozygous state in 56 (69%) of cases. Sixteen of the remaining patients in whom cosegregation analysis was not performed carried variants found to be prevalent among the Spanish population shown in Table 1, thus explaining homozygosity. In homozygotes for c.3386G>T, CNV studies including MLPA or NGS were performed to discard gross deletions.

Clinical characteristics of *ABCA4* patients

Diagnosis of STGD1 was established for 434 patients; the remaining 72 patients presented CRD. Clinical information of 372 patients including AO, age at diagnosis, BCVA, and ffERG results is summarized in Supplementary Table S1.

The median AO (IQR) of 66 CRD and 306 STGD1 patients was 10 (6) and 16 (15) years, respectively (Supplementary Table S6). CRD patients presented an onset of disease during the first and early second decade of life, while the disease onset in STGD1 patients was in the second and third decades, revealing a statistically significant difference in distribution according to this variable (Figure 2A).

Some patients presented with a good BCVA at age at diagnosis, not showing symptoms of loss of visual acuity. A well-preserved foveal structure together with a very good BCVA was described in 8 STGD1 patients from families MD-0853, MD-0991, MD-0959, MD-1106, MD-1110, MD-1146,

MD-1356, and MD-1381, ranging in age from 19 to 72 years. SD-OCT and FAF images of 4 of them are shown in Figure 3. FAF images revealed macular atrophy sparing the fovea in patients MD-0959, MD-1146 and MD-1381 while MD-1356 revealed a hyperautofluorescent halo surrounding areas of non definitive dark hypoautofluorescence in the macula.

<u>Genotype-phenotype correlations</u>

To determine whether the combination of *ABCA4* variants in our cohort, regardless of diagnosis, reflected the established genotype-phenotype model, the AO of 372 index patients was compared between genotype categories A, B, and C. The median AO (IQR) was 17 (15), 14 (14), and 9 (3.5) years, respectively. There were statistically significant differences between patients with biallelic truncating variants (category C) and those with both biallelic missense (category A) and missense-truncating (category B) variants (Supplementary Figure S2 and Table S7). Patients carrying two missense (category A) and missense-truncating variants (category B) also showed statistically significant differences.

Analysis of patients with the c.3386G>T variant revealed that the median AO (IQR) in categories D, E, and F was 21.5 (18.5), 17 (8.5) and 14 (10) years, respectively (Table 2). There were statistically significant differences between c.3386G>T homozygotes (category D) and compound heterozygotes carrying a truncating variant (category F), and between patients carrying the c.3386G>T in combination with a missense variant (category E) and patients carrying the c.3386G>T in combination with a truncating variant (category F). Comparisons between the patients carrying the c.3386G>T variant and non-3386G>T patients showed statistically significant differences when all patients were taken into account, the median AO (IQR) among these patients was 16.5 (10.8) and 13 (15) years, respectively (Table 2). There were no statistically significant differences when comparing categories A2-C2 with D-F (Supplementary Table S8 and S9). The same analysis was carried out for the c.5882G>A variant, excluding category G since there were no homozygous patients in our cohort. In this case, median AO (IQR) for categories H and I were 17 (13.8) and 20 (15) years, respectively, and no statistically significant differences were found (Table 2). The comparison between patients carrying c.3386G>T and c.5882G>A did not reveal statistically significant differences (Supplementary Table S10).

Genotype-phenotype correlation regarding the clinical CRD and the STGD1 phenotypes evidenced statistically significant differences between patients from the two classes carrying biallellic missense variants (12.5 (8.3) and 17 (16) years, respectively) and a missense with a truncating variant (10 (5) and 15 (15.5) years, respectively) (Supplementary Table S6). CRD and STGD1 patients carrying biallelic truncating variants had a similar AO (9 (5.5) and 9 (3) years) (Figure 2B and Table S6). Remarkably, most of the CRD patients (41%, 27/66) belonged to the latter group, thus contrasting with STGD1 patients (6.5%, 20/306).

The number of alleles carrying different missense variants was compared between CRD and STGD1 patients. Four variants showed statistically significant differences (Supplementary Table S11). Variants c.1804C>T; p.(Arg602Trp) (OR = 5.31; 95%CI = 2.27-11.7), c.3056C>T; p.(Thr1019Met) (OR = 7.58; 95%CI = 2.12-25.1), and c.6320G>C; p.(Arg2107Pro) (OR = 10.5; 95%CI = 1.08-102) were over-represented in CRD patients, while the c.3386G>T variant was the only variant over-represented in STGD1 patients (OR = 0.37; 95%CI = 0.14-0.80). In addition,

c.3386G>T variant was also over-represented in the foveal sparing cohort: it was carried by MD-0959 and MD-1356 in homozygosis, and MD-1106 carried it in heterozygosis. Missense variants previously described as severe variants did not show statistically significant differences (Table S11).

DISCUSSION

We report the largest cohort of patients with *ABCA4* variants ever analyzed to date, consisting of 434 STGD1 and 72 CRD patients, providing an accurate analysis of the genomic and phenotypic landscape of different combinations of variants in this gene. A total of 228 different DNA changes were identified, most of which were missense changes (56%).

Novel variants accounted for 22% of all variants and 7.2% of *ABCA4* patient alleles, and other studies based on large cohorts of STGD1 patients have found similar rates of novel variants (Schulz *et al.*, 2017; Fujinami *et al.*, 2019; Khan *et al.*, 2019). Using comprehensive targeted NGS-based screening, we were able to observe the highly diverse allelic and mutational spectrum of the *ABCA4* gene.

By screening CNVs and/or deep intronic variants we were able to solve 8 and 25 families, respectively. CNVs in the *ABCA4* gene do not usually account for a representative proportion of variants (Yatsenko *et al.*, 2003; Zernant *et al.*, 2014); the same holds for our cohort as well, for which they represent less than 1% of all alleles. Interestingly, a novel 411 bp deletion partially encompassing the sixth exon and intron of *ABCA4* was found in 6 families, in whom a common region of 1.58 Mb was found, thus suggesting a possible founder mutation in the Spanish population. Sequencing of *ABCA4* introns enabled us to explain some of the missing heritability, thanks to the identification of several deep intronic variants that affect the correct splicing of primary *ABCA4* transcripts, as previously reported (Braun *et al.*, 2013; Zernant *et al.*, 2014; Bauwens *et al.*, 2019; Khan *et al.*, 2019; Sangermano *et al.*, 2019). In our cohort, 2.6% of all alleles were found to be deep intronic variants, a rate that closely matches the 2-2.4% reported by Schulz et al (Schulz *et al.*, 2017) and Fujinami et al (Fujinami *et al.*, 2019) in large cohorts of more than 300 STGD1 cases. However, in Khan et al (Khan *et al.*, 2019), these variants represented 15% of the missing alleles, most likely due to the fact these patients had been previously screened for coding variants and because all studied probands were analyzed for deep-intronic variants.

A recent *in silico* meta-analysis provided a pathogenic classification for all reported *ABCA4* variants based on their frequency in controls and in IRD patients (Cornelis *et al.*, 2017). Based on these findings, the complete gene was also sequenced in a parallel study in 7 cases carrying c.6148G>C, a variant previously classified as pathogenic. The variant was found in combination with another pathogenic *ABCA4* variant in *cis* in all cases. One of these cases is family RP-0674, previously reported in Corton et al (Corton *et al.*, 2013). The new variant identified was a coding variant filtered out on the WES analysis, owing to extremely low coverage. According to these data and findings from recent studies (González-del Pozo *et al.*, 2018), c.6148G>C should therefore be considered a likely benign variant. On the other hand, the frequent variant c.5603A>T, recently considered a low-penetrant variant (Zernant *et al.*, 2017; Runhart *et al.*, 2018), was identified in 9 cases allowing us to consider them solved, and only 2 carried an

additional intronic variant in *cis*. Further analysis of negative results together with review and reclassification of variants is needed to solve these cases.

Genotype-phenotype correlations for the AO of disease in patients were assessed following stratification by *ABCA4*-variant categories, regardless of phenotype, and by clinical STGD1 or CRD phenotypes. Our results demonstrated that patients with biallelic truncating variants have a statistically significant earlier AO than other combinations of variants and most of them presented a CRD phenotype. Combinations of missense variants with another missense or truncating variant were over-represented in STGD1 patients. Our data suggest that STGD1 patients carrying 2 truncating variants could evolve to be CRD and as a result further ophthalmological examinations including ffERG should be considered. The proposed genotype-phenotype correlation model suggests that the phenotype can be predicted by the *ABCA4* variant type, depending on the residual function of the ABCA4 protein (Maugeri *et al.*, 1999). We believe that our findings provide further insights into the accuracy of this model based on AO data of 372 patients, a sample size that confers greater statistical weight. It is also true that the classification of truncating variants included splice variants that produce partial truncations, and there are also missense variants that cause severe functional effects (N. Zhang *et al.*, 2014; Tanna *et al.*, 2017; Molday *et al.*, 2018). None of these missense variants (c.[1622T>C;3113C>T];p.[Leu541Pro;Ala1038Val], c.2894A>G; p.(Asn965Ser) and c.4918C>T; p.(Arg1640Trp)) were related with a CRD phenotype in our cohort. However, variants c.1804C>T; p.(Arg602Trp), c.3056C>T; p.(Thr1019Met), and c.6320G>C; p.(Arg2107Pro) were associated with a CRD phenotype, while c.3386G>T was correlated with STGD1 patients. To our knowledge, this is the first time that these specific *ABCA4* missense variants are clinically associated to a different phenotype, which could be used to evaluate the prognosis of patients diagnosed at early ages with mild clinical manifestations.

We also performed genotype-phenotype correlations for the most prevalent Spanish variant, c.3386G>T (Valverde *et al.*, 2006), as well as the common c.5882G>A variant. Homozygous patients for c.3386G>T presented later AO and represent only 11% of all patients with this variant, as seen also in homozygous cases for the c.5882G>A variant, though these cases were absent from our cohort. It has been reported that c.5882G>A in a homozygous state typically causes a milder phenotype than when it is present in combination with other variants (Tanna *et al.*, 2017). Our data could suggest that homozygous patients for c.5882G>A could have a very mild phenotype, even without manifestation of visual symptoms. The Spanish variant c.3386G>T should be considered mild, although we previously proposed that it could have a moderately severe effect (Valverde *et al.*, 2006).

Severe CRD phenotypes could be diagnosed as RP(Riveiro-Alvarez *et al.*, 2013). In this work, all CRD patients presented with rod disfunction, due to ERG findings or symptoms associated with rod degeneration. Six cases in which these data were not available were referred with CRD diagnosis. At the other end of the severity spectrum, 8 patients with clinical features of STGD1but without clinical symptoms at the age of diagnosis had good visual acuity and well-preserved foveal structure. Later onset or preserved visual acuity has been described in STGD1 patients (Armstrong *et al.*, 1998; Yatsenko *et al.*, 2001) associated with a milder phenotype and foveal sparing (Rotenstreich, Fishman and Anderson, 2003; Fujinami *et al.*, 2011; Westeneng-van Haaften *et al.*, 2012; Fujinami, Sergouniotis, *et al.*, 2013). A previous study reported that the

c.6089G>A; p.(Arg2030Gln) change, which we did not identify in our patients, was over-represented in cases with foveal-sparing, compared to typical STGD1 cases (Fujinami, Sergouniotis, *et al.*, 2013). In our cohort, 2 patients were homozygous for the Spanish c.3386G>T variant, a finding which supports the mild effect of this variant and the possibility of an underdiagnosis of additional homozygotes, due to the lack of visual disabling symptoms. However, also one CRD patient carried this variant in homozygosis. Further studies sequencing the entire *ABCA4* gene or regulatory regions would be needed to determine if additional variants in *cis* could be modifying the penetrance of these variants in homozygotes.

In summary, this study supports the role played by genetic diagnosis in predicting the progression of the disease, and the difficulty of obtaining a correct clinical diagnosis when non-typical STGD1 features are present or electrophysiology data are not available. Certain combinations of variants in homozygosis state may not always be associated with a diagnosed clinical phenotype. Alternatively, onset of symptoms may occur later in life, as in the case of foveal-sparing patients. Given the wealth of gene-based therapy initiatives under way involving patients with *ABCA4* causative variants, a precise identification of the genetic makeup of STGD1 or CRD cases, including the presence of missing alleles, is a crucial step toward enrolling these patients in clinical trials.

REFERENCES

1. Allikmets R, Singh N, Sun H, et al. A photoreceptor cell-specific ATP-binding transporter gene (ABCR) is mutated in recessive Stargardt macular dystrophy. Nat Genet. 1997;15(3):236-246.

2.   Cremers FPM, Van De Pol DJR, Van Driel M, et al. Autosomal recessive retinitis pigmentosa and cone-rod dystrophy caused by splice site mutations in the Stargardt's disease gene ABCR. Hum Mol Genet. 1998;7(3):355-362.

3.   Martínez-Mir A, Paloma E, Allikmets R, et al. Retinitis pigmentosa caused by a homozygous mutation in the Stargardt disease gene ABCR. Nat Genet. 1998;18(1):11-12.

4.   Molday LL, Rabin AR, Molday RS. ABCR expression in foveal cone photoreceptors and its role in Stargardt macular dystrophy. Nat Genet. 2000.

5.   Molday RS. ATP-binding cassette transporter ABCA4: molecular properties and role in vision and macular degeneration. J Bioenerg Biomembr. 2007;39(5-6):507-517.

6.   Blacharski, PA. Fundus flavimaculatus. Retin Dystrophies Degener. 1988:135-159.

7.   Riveiro-Alvarez R, Lopez-Martinez MA, Zernant J, et al. Outcome of ABCA4 disease-associated alleles in autosomal recessive retinal dystrophies: Retrospective analysis in 420 Spanish families. Ophthalmology. 2013;120(11):2332-2337.

8.   Anderson KL, Baird L, Lewis RA, et al. A YAC contig encompassing the recessive Stargardt disease gene (STGD) on chromosome 1p. Am J Hum Genet. 1995;57(6):1351-1363.

9.   Hamel CP. Cone rod dystrophies. Orphanet J Rare Dis. 2007.

10.   Michaelides M, Hardcastle AJ, Hunt DM, Moore AT. Progressive Cone and Cone-Rod Dystrophies: Phenotypes and Underlying Molecular Genetic Basis. Surv Ophthalmol. 2006.

11.   van Driel MA, Maugeri A, Klevering BJ, Hoyng CB, Cremers FP. ABCR unites what ophthalmologists divide(s). Ophthalmic Genet. 1998;19(3):117-122.

12.   Maugeri A, Klevering BJ, Rohrschneider K, et al. Report Mutations in the ABCA4 (ABCR) Gene Are the Major Cause of Autosomal Recessive Cone-Rod Dystrophy. Am J Hum Genet. 2000;67(4):960-966.

13.   Braun TA, Mullins RF, Wagner AH, et al. Non-exomic and synonymous variants in ABCA4 are an important cause of Stargardt disease. Hum Mol Genet. 2013;22(25):5136-5145.

14.   Sangermano R, Garanto A, Khan M, et al. Deep-intronic ABCA4 variants explain missing heritability in Stargardt disease and allow correction of splice defects by antisense oligonucleotides. Genet Med. 2019.

15.   Bauwens M, Garanto A, Sangermano R, et al. ABCA4-associated disease as a model for missing heritability in autosomal recessive disorders: novel noncoding splice, cis-regulatory, structural, and recurrent hypomorphic variants. Genet Med. 2019.

16.   Zernant J, Xie YA ngel., Ayuso C, et al. Analysis of the ABCA4 genomic locus in Stargardt disease. Hum Mol Genet. 2014;23(25):6797-6806.

17.   Khan M, Cornelis SS, Khan MI, et al. Cost-effective molecular inversion probe-based ABCA4 sequencing reveals deep-intronic variants in Stargardt disease. Hum Mutat. 2019.

18.   Runhart EH, Sangermano R, Cornelis SS, et al. The common ABCA4 variant p.Asn1868ile shows nonpenetrance and variable expression of stargardt disease when present in trans with severe variants. Investig Ophthalmol Vis Sci. 2018;59(8):3220-3231.

19.   Zernant J, Lee W, Collison FT, et al. Frequent hypomorphic alleles account for a significant fraction of ABCA4 disease and distinguish it from age-related macular degeneration. J Med Genet. 2017;54(6):404-412.

20. Valverde D, Riveiro-Alvarez R, Bernal S, et al. Microarray-based mutation analysis of the ABCA4 gene in Spanish patients with Stargardt disease: evidence of a prevalent mutated allele. Mol Vis. 2006.

21. Martin-Merida I, Aguilera-Garcia D, Fernandez-San Jose P, et al. Toward the mutational landscape of autosomal dominant retinitis pigmentosa: A comprehensive analysis of 258 Spanish families. Investig Ophthalmol Vis Sci. 2018;59(6):2345-2354.

22. Del Pozo-Valero M, Martin-Merida I, Jimenez-Rolando B, et al. Expanded phenotypic spectrum of retinopathies associated with autosomal recessive and dominant mutations in PROM1. Am J Ophthalmol. 2019.

23. Khan M, Cornelis SS, Pozo-Valero M Del, et al. Resolving the dark matter of ABCA4 for 1054 Stargardt disease probands through integrated genomics and transcriptomics. Genet Med. 2020.

24. Sim N-L, Kumar P, Hu J, Henikoff S, Schneider G, Ng PC. SIFT web server: predicting effects of amino acid substitutions on proteins. Nucleic Acids Res. 2012;40(W1):W452-W457.

25. Adzhubei IA, Schmidt S, Peshkin L, et al. A method and server for predicting damaging missense mutations. Nat Methods. 2010;7(4):248-249.

26. 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(3):310-315.

27. Jagadeesh KA, Wenger AM, Berger MJ, et al. M-CAP eliminates a majority of variants of uncertain significance in clinical exomes at high sensitivity. Nat Genet. 2016.

28. Richards S, Aziz N, Bale S, et al. Standards and guidelines for the interpretation of sequence variants: A joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. Genet Med. 2015;17(5):405-424.

29. Cornelis SS, Bax NM, Zernant J, et al. In Silico Functional Meta-Analysis of 5,962 ABCA4 Variants in 3,928 Retinal Dystrophy Cases. Hum Mutat. 2017;38(4):400-408.

30. Schulz HL, Grassmann F, Kellner U, et al. 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. Investig Ophthalmol Vis Sci. 2017.

31. Fujinami K, Strauss RW, Chiang J, et al. Detailed genetic characteristics of an international large cohort of patients with Stargardt disease: ProgStar study report 8. Br J Ophthalmol. 2019.

32. Yatsenko AN, Shroyer NF, Lewis RA, Lupski JR. An ABCA4 genomic deletion in patients with Stargardt disease. Hum Mutat. 2003.

33. Corton M, Nishiguchi KM, Avila-Fernández A, et al. Exome sequencing of index patients with retinal dystrophies as a tool for molecular diagnosis. PLoS One. 2013.

34. González-del Pozo M, Martín-Sánchez M, Bravo-Gil N, et al. Searching the second hit in patients with inherited retinal dystrophies and monoallelic variants in ABCA4, USH2A and CEP290 by whole-gene targeted sequencing. Sci Rep. 2018.

35. Maugeri A, van Driel MA, van de Pol DJ, et al. The 2588G-->C mutation in the ABCR gene is a mild frequent founder mutation in the Western European population and allows the

classification of ABCR mutations in patients with Stargardt disease. Am J Hum Genet. 1999;64(4):1024-1035.

36.   Tanna P, Strauss RW, Fujinami K, Michaelides M. Stargardt disease: Clinical features, molecular genetics, animal models and therapeutic options. Br J Ophthalmol. 2017.

37.   Molday LL, Wahl D, Sarunic M V., Molday RS. Localization and functional characterization of the p.Asn965Ser (N965S) ABCA4 variant in mice reveal pathogenic mechanisms underlying Stargardt macular degeneration. Hum Mol Genet. 2018.

38.   Zhang N, Tsybovsky Y, Kolesnikov A V., et al. Protein misfolding and the pathogenesis of ABCA4-associated retinal degenerations. Hum Mol Genet. 2014.

39.   Yatsenko AN, Shroyer NF, Lewis RA, Lupski JR. Late-onset Stargardt disease is associated with missense mutations that map outside known functional regions of ABCR (ABCA4). Hum Genet. 2001.

40.   Armstrong JD, Meyer D, Xu S, Elfervig JL. Long-term follow-up of Stargardt's disease and fundus flavimaculatus. Ophthalmology. 1998.

41.   Fujinami K, Sergouniotis PI, Davidson AE, et al. Clinical and molecular analysis of stargardt disease with preserved foveal structure and function. Am J Ophthalmol. 2013.

42.   Westeneng-van Haaften SC, Boon CJF, Cremers FPM, Hoefsloot LH, den Hollander AI, Hoyng CB. Clinical and genetic characteristics of late-onset Stargardt's disease. Ophthalmology. 2012;119(6):1199-1210.

43.   Fujinami K, Akahori M, Fukui M, Tsunoda K, Iwata T, Miyake Y. Stargardt disease with preserved central vision: identification of a putative novel mutation in ATP-binding cassette transporter gene. Acta Ophthalmol. 2011;89(3):e297-8.

44.   Rotenstreich Y, Fishman GA, Anderson RJ. Visual acuity loss and clinical observations in a large series of patients with Stargardt disease. Ophthalmology. 2003;110(6):1151-1158.

FIGURES AND TABLES

Figure 1. Mutational spectrum of *ABCA4* in 506 Spanish families

A) Number of different variants — Total: 228 variants

Number of alleles — Total: 1012 alleles

Legend: Missense; Frameshift, stop; Complex; Canonical splice site; Non canonical splice site; Deep intronic; In-frame indels; CNV

B) Single and complex alleles: Single alleles 964; 95% — Complex alleles 48; 5%

C) Reported and novel variants: Previously reported variants 178; 78% — Novel variants 50; 22%

D) Compound heterozygotes and homozygotes patients: Homozygotes 81; 16% — Compound heterozygotes 425; 84%

Figure 1. Mutational spectrum of *ABCA4* in 506 Spanish patients. A) Percentage distribution of 228 different variants identified in 1012 patient alleles by variant type. B) Percentage distribution of single and complex alleles. C) Percentage distribution of novel and reported variants. D) Representation of persons carrying homozygous or compound heterozygous *ABCA4* variants. Complex variants include variants present in complex alleles and not in single alleles.
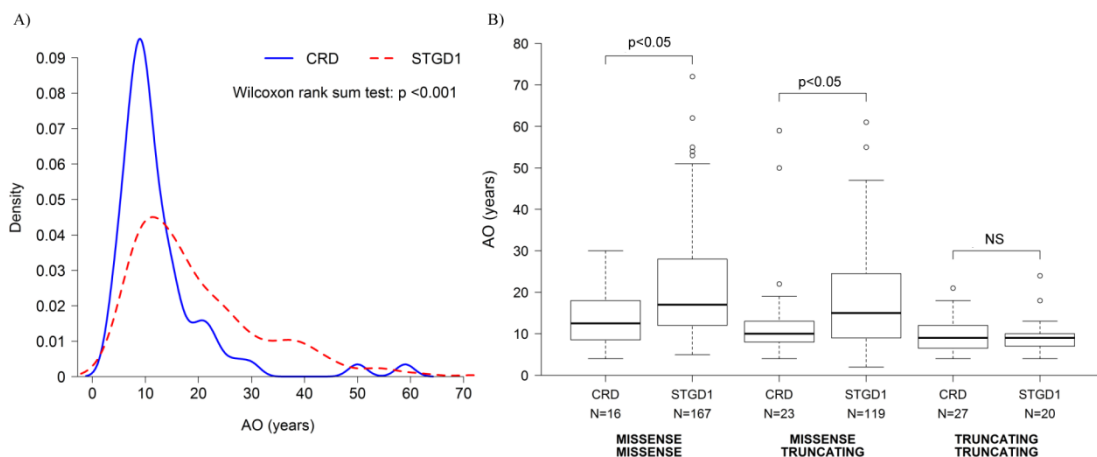
Figure 2. Distribution of age of onset and genotype-phenotype correlation in CRD vs. STGD1 patients. A) CRD patients presented a statistically significant earlier AO than STGD1 patients. B) AO of CRD and STGD1 carrying at least one missense variant showed statistically significant differences. Patients with biallelic truncating variants presented similar AO. Missense variants c.1804C>T; p.(Arg602Trp), c.3056C>T; p.(Thr1019Met), and c.6320G>C; p.(Arg2107Pro) were over-represented in CRD patients, and c.3386G>T; p.(Arg1129Leu) in STGD1 patients. Abbreviations: AO, age of onset, CRD, cone-rod dystrophy; STGD1, Stargardt disease; N, number of cases.
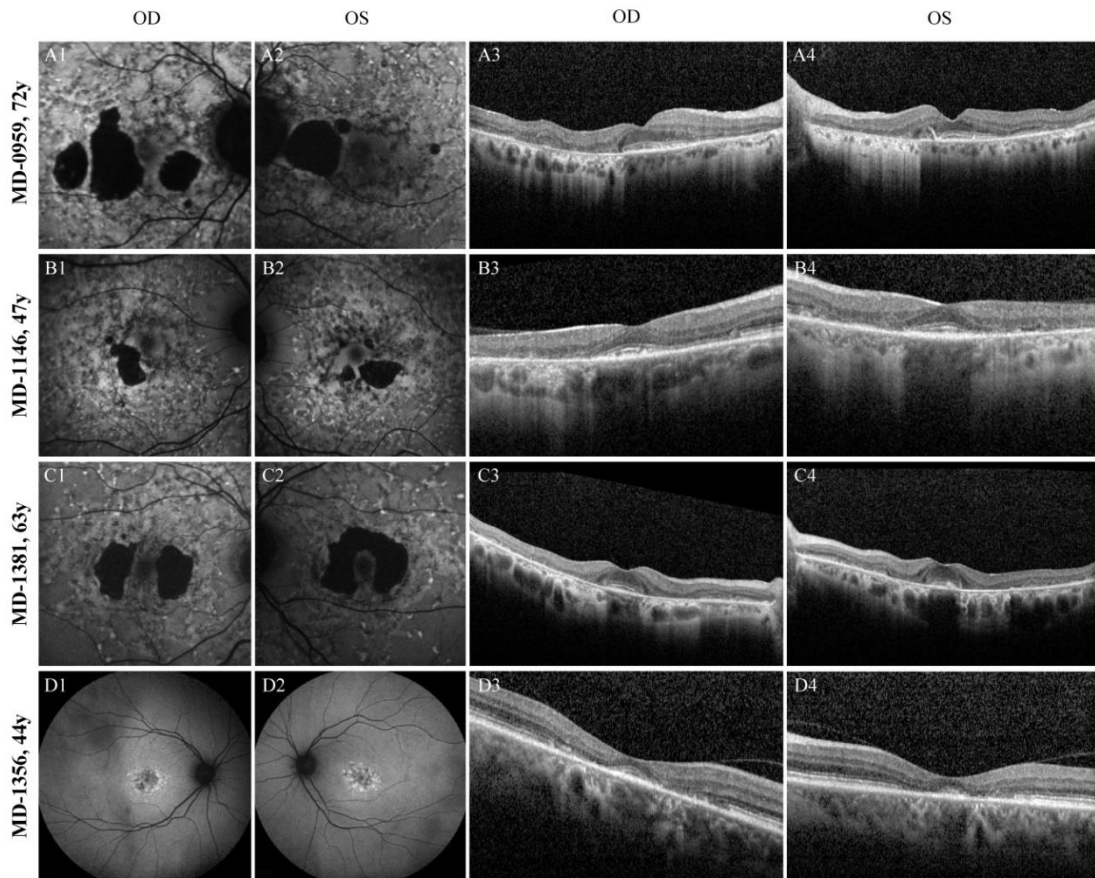
Figure 3. FAF and SD-OCT images of patients presenting foveal sparing. FAF images A1-A2 to C1-C2 and D1-D2 at 35º and 55º center in the macula, respectively, showing areas of definitive dark autofluorescence sparing in the foveal area in patients MD-0959, MD-1146 and MD-1381. Images obtained of patient MD-1356 images show non-definitive dark autofluorescence with scattered hyperautofluorescent lesions in the perifoveal area. SD-OCT images A3-A4 to D3-D4 evidence disruption of the ellipsoid zone and external layers in the perifoveal area with subfoveal preservation in all patients. Abbreviations: FAF, fundus autofluorescence; SD-OCT, spectral domain optical coherence tomography.

Table 1. Most prevalent ABCA4 variants found in 506 Spanish families.

	ABCA4 variants			
Single variants				
Exon	**Nucleotide**	**Protein**	**Number of families**	**Number of alleles**
23	c.3386G>T	p.(Arg1129Leu)	170	190
42	c.5882G>A	p.(Gly1961Glu)	64	64
22	c.3210_3211dup	p.(Ser1071Cysfs*14)	30	34
13	c.1804C>T	p.(Arg602Trp)	26	30
41	c.5819T>C	p.(Leu1940Pro)	26	29
30	c.4457C>T	p.(Pro1486Leu)	23	26
19	c.2888del	p.(Gly963Alafs*14)	23	25
45	c.6179T>G	p.(Leu2060Arg)	21	24
Complex variants				
22;46	c.[3322C>T;6320G>A]	p.[Arg1108Cys;Arg2107His]	12	13
23;48	c.[3386G>T;6718A>G]	p.[Arg1129Leu;Thr2240Ala]	5	6
35;36	c.[4926C>G;5044_5058del]	p.[Ser1642Arg;Val1681_Cys1685del]	4	6

Table 2. Genotype-phenotype correlation for prevalent ABCA4 variants c.3386G>T;p.(Arg1129Leu) and c.5882G>A;p.(Gly1961Glu). Abbreviations: AO, age of onset; IQR, interquartile range; N, number of cases; NS, non significant.# Three patients were excluded of this category because they carried another variant in cis together c.3386G>T variant.

Genotype-phenotype correlation for c.3386G>T variant			
Category	**D: c.3386G>T-c.3386G>T**	**E: c.3386G>T-MISSENSE**	**p**
Median AO (IQR) (N)	21.5 (18.5) years (N=12)	17.0 (8.50) years (N=68)	NS
Category	**D: c.3386G>T-c.3386G>T**	**F: c.3386G>T-TRUNCATING**	**p**
Median AO (IQR) (N)	21.5 (18.5) years (N=12)	14.0 (10.0) years (N=43)	<0.05
Category	**E: c.3386G>T-MISSENSE**	**F: c.3386G>T-TRUNCATING**	**p**
Median AO (IQR) (N)	17.0 (8.50) years (N=68)	14.0 (10.0) years (N=43)	<0.05
Category	**c.3386G>T patients**	**Non- c.3386G>T patients**	**p**
Median AO (IQR) (N)	16.5 (10.8) years (N=126)	13.0 (15.0) years (N=246)	<0.05
Category	**c.3386G>T patients**	**All patients**	**p**
Median AO (IQR) (N)	16.5 (10.8) years (N=126)	15.0 (15.0) years (N=372)	NS
Category	**Non- c.3386G>T patients**	**All patients**	**p**
Median AO (IQR) (N)	13.0 (15.0) years (N=246)	15.0 (15.0) years (N=372)	NS
Genotype-phenotype correlation for c.5882G>A variant			
Category	**H: c.5882G>A-MISSENSE**	**I: c.5882G>A-TRUNCATING**	**p**
Median AO (IQR) (N)	17.0 (13.8) years (N=28)	20.0 (15.0) years (N=15)	NS

Supplementary Material

Del Pozo-Valero_ Genotype-phenotype correlations in a Spanish cohort of 506 families with bi-allelic *ABCA4* pathogenic variants

Title and description of data:

Figure S1. Pedigrees of the six families with the novel c.699_768+341del; p.(Gln234Phefs*5) variant in the *ABCA4* gene.

Figure S2. Genotype-phenotype correlation for 372 patients carrying biallelic missense, a combination of a missense and truncating, or biallelic truncating variants in the *ABCA4* gene.
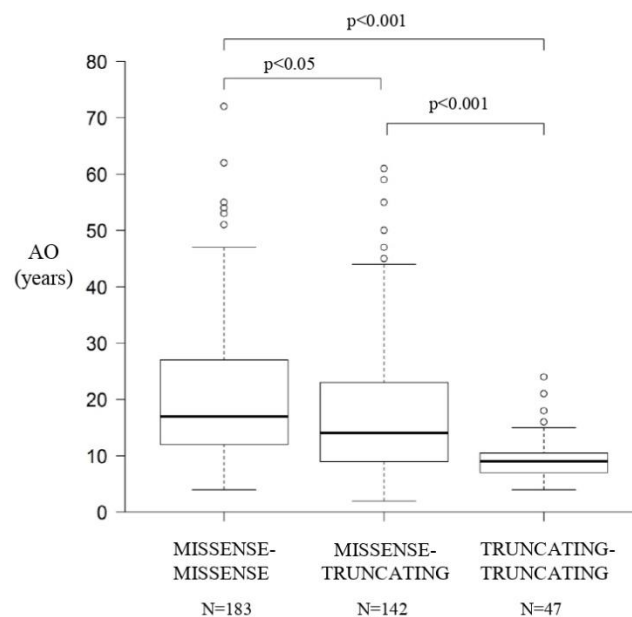
Table S1. Genetic and clinical information of 506 Spanish families with *ABCA4* mutations.

Table S2. Total of *ABCA4* variants identified in 506 Spanish families.

Table S3. Variants identified in 21 complex alleles.

Table S4. *In-silico* predictions of novel *ABCA4* variants identified in 506 Spanish families.

Table S5. Deep intronic *ABCA4* variants identified in 506 Spanish families.

Table S6. Genotype-phenotype correlation of cone-rod dystrophy (CRD) and Stargardt disease (STGD1) patients regarding type of *ABCA4* variant.

Table S7. Genotype-phenotype correlation regarding type of *ABCA4* variant.

Table S8. Genotype-phenotype correlation regarding type of *ABCA4* variant excepting c.3386G>T;p.(Arg1129Leu).

Table S9. Genotype-phenotype correlation between c.3386G>T Patients and Non-c.3386G>T Patients.

Table S10. Comparison between patients carrying the c.3386G>T variant vs patients carrying the c.5882G>A variant in *ABCA*.

Table S11. Missense variants associated with CRD or STGD1 phenotypes.

1

Table S1. Genetic and clinical information of 506 Spanish families with *ABCA4* mutations. Abbreviations: STGD1, Stargardt disease; CRD, cone-rod dystrophy; IVS, intron; VA, visual acuity; VF, visual field; NB, night blindness; y, years; ERG, electroretinogram; BCVA, best corrected visual acuity; OD, right eye; OS, left eye; CF, counting fingers; HM, hand movement

Family ID	Phenotype	ABCA4 variants						Segregation	Symptoms onset (age in years)			Age at ophthalmological examination	ERG	BCVA (OD/OS)	Reference
		Allele1_Exon	Allele1_cDNA	Allele 1_Protein	Allele2_Exon	Allele2_cDNA	Allele2_Protein		VA loss	VF loss	NB				
MD-0012	STGD1	25;42	c.[3758C>T;5882G>A]	p.[Thr1253Met; Gly1961Glu]	27	c.3943C>T	p.(Gln1315*)	Yes	38	38	38	-	-	-	Riveiro-Alvarez et al.,2013
MD-0014	STGD1	22;46	c.[3322C>T;6320G>A]	p.[Arg1108Cys;Arg2107His]	22;46	c.[3322C>T;6320G>A]	p.[Arg1108Cys;Arg2107His]	Yes	5	20	-	12y	Normal	0.1/0.1	Riveiro-Alvarez et al.,2013
MD-0015	STGD1	17	c.[2588G>C;3163C>T]	p.[Gly863Ala,Gly863del;Arg1055Trp]	19	c.2888del	p.(Gly963Alafs*14)	Yes	11	11	-	12y	Normal	0.1/0.05	Riveiro-Alvarez et al.,2013
MD-0017	STGD1	22;46	c.[3322C>T;6320G>A]	p.[Arg1108Cys;Arg2107His]	46	c.6320G>C	p.(Arg2107Pro)	Yes	8	No	-	15y	Normal	0.15/0.1	Riveiro-Alvarez et al.,2013
MD-0022	STGD1	19	c.2888del	p.(Gly963Alafs*14)	45	c.6179T>G	p.(Leu2060Arg)	Yes	12	No	-	-	-	-	Riveiro-Alvarez et al.,2013
MD-0031	STGD1	23	c.3386G>T	p.(Arg1129Leu)	35	c.4926C>G	p.(Ser1642Arg)	Yes	41	No	-	41y	Normal	-	This study
MD-0033	STGD1	3	c.223T>G	p.(Cys75Gly)	27	c.3881_3885del	p.(Arg1294Lysfs*126)	Yes	13	13	25	28y	-	CF/CF	This study
MD-0038	STGD1	13	c.1804C>T	p.(Arg602Trp)	33	c.4739del	p.(Leu1580*)	Yes	6	No	-	28y	-	0.05/0.05	Riveiro-Alvarez et al.,2013
MD-0039	STGD1	6	c.699_768+341del	p.(Gln234Phefs*5)	IVS40	c.5714+5G>A	p.[=,Glu1863Leufs*33]	Yes	10	11	-	22y	-	0.2/0.3	This study
MD-0040	CRD	21	c.3056C>T	p.(Thr1019Met)	27	c.3943C>T	p.(Gln1315*)	Yes	9	No	-	21y	Cone-rod pattern	CF/CF	Riveiro-Alvarez et al.,2013
MD-0047	STGD1	23	c.3386G>T	p.(Arg1129Leu)	23	c.3386G>T	p.(Arg1129Leu)	Yes	21	19	-	22y	Cone-pattern	0.1/0.1	Riveiro-Alvarez et al.,2013
MD-0051	STGD1	22	c.3292C>T	p.(Arg1098Cys)	35	c.4855T>C	p.(Phe1619Leu)	-	18	38	-	48y	Normal	0.12/0.16	This study
MD-0057	STGD1	23	c.3386G>T	p.(Arg1129Leu)	42	c.5882G>A	p.(Gly1961Glu)	Yes	15	15	-	19y	Normal	-	Riveiro-Alvarez et al.,2013
MD-0060	STGD1	13	c.1804C>T	p.(Arg602Trp)	42	c.5882G>A	p.(Gly1961Glu)	Yes	17	17	-	22y	Normal	0.7/0.6	Riveiro-Alvarez et al.,2013
MD-0061	STGD1	23	c.3386G>T	p.(Arg1129Leu)	23	c.3386G>T	p.(Arg1129Leu)	Yes	33	No	-	35y	-	0.5/0.6	Riveiro-Alvarez et al.,2013
MD-0062	CRD	23	c.3386G>T	p.(Arg1129Leu)	43	c.5929G>A	p.(Gly1977Ser)	Yes	27	45	45	50y	Cone-rod pattern	CF/CF	Riveiro-Alvarez et al.,2013
MD-0064	CRD	6	c.634C>T	p.(Arg212Cys)	43	c.5929G>A	p.(Gly1977Ser)	Yes	15	15	-	40y	-	0.1/0.05	Riveiro-Alvarez et al.,2013
MD-0065	STGD1	42	c.5882G>A	p.(Gly1961Glu)	IVS7	c.859-506G>C	p.[Phe287Thrfs*32,=]	Yes	12	No	-	21y	-	0.3/0.2	This study
MD-0066	STGD1	23	c.3386G>T	p.(Arg1129Leu)	41	c.5819T>C	p.(Leu1940Pro)	-	17	28	-	-	-	-	Riveiro-Alvarez et al.,2013
MD-0072	CRD	13	c.1804C>T	p.(Arg602Trp)	22	c.3287C>T	p.(Ser1096Leu)	Yes	12	No	-	27y	Cone-rod pattern	0.07/0.07	Riveiro-Alvarez et al.,2013
MD-0076	STGD1	6	c.768G>T	p.(Val256Val)	23	c.3386G>T	p.(Arg1129Leu)	-	11	No	-	24y	-	0.1/0.1	Riveiro-Alvarez et al.,2013
MD-0078	STGD1	23	c.3386G>T	p.(Arg1129Leu)	48	c.6559C>T	p.(Gln2187*)	Yes	12	No	-	38y	-	0.1/0.16	Riveiro-Alvarez et al.,2013
MD-0079	CRD	19	c.2888del	p.(Gly963Alafs*14)	19	c.2888del	p.(Gly963Alafs*14)	Yes	8	25	18	10y	Cone-rod pattern	<0.05/<0.05	Riveiro-Alvarez et al.,2013
MD-0081	STGD1	29	c.4297G>A	p.(Val1433Ile)	42	c.5882G>A	p.(Gly1961Glu)	Yes	28	No	-	-	-	-	Riveiro-Alvarez et al.,2013
MD-0082	STGD1	23	c.3364G>A	p.(Glu1122Lys)	23	c.3386G>T	p.(Arg1129Leu)	Yes	15	15	15	15y	Cone-pattern	-	Riveiro-Alvarez et al.,2013
MD-0084	STGD1	47	c.6410G>A	p.(Cys2137Tyr)	47	c.6410G>A	p.(Cys2137Tyr)	Yes	7	7	-	23y	-	<0.05/<0.05	Riveiro-Alvarez et al.,2013
MD-0086	CRD	19	c.2888del	p.(Gly963Alafs*14)	19	c.2888del	p.(Gly963Alafs*14)	-	9	40	18	49y	Scotopic and photopic extinguise	CF/CF	Riveiro-Alvarez et al.,2013
MD-0088	STGD1	28	c.4139C>T	p.(Pro1380Leu)	IVS40	c.5714+5G>A	p.[=,Glu1863Leufs*33]	-	16	No	-	25y-30y	-	0.25/0.25-0.1/0.1	Riveiro-Alvarez et al.,2013
MD-0090	STGD1	IVS22	c.3329-2A>T	p.(?)	43	c.5929G>A	p.(Gly1977Ser)	Yes	8	No	-	22y	-	CF/CF	Riveiro-Alvarez et al.,2013
MD-0094	STGD1	14	c.1964T>G	p.(Phe655Cys)	16	c.2481del	p.(Thr829Argfs*14)	-	6	43	No	-	-	-	This study
MD-0096	STGD1	23	c.3386G>T	p.(Arg1129Leu)	28	c.[4222T>C;4918C>T]	p.[Trp1408Arg;Arg1640Trp]	Yes	20	No	-	37y	-	0.15/0.1	Riveiro-Alvarez et al.,2013
MD-0110	STGD1	22;46	c.[3322C>T;6320G>A]	p.[Arg1108Cys;Arg2107His]	42	c.5882G>A	p.(Gly1961Glu)	Yes	8	No	-	19y	-	0.16/0.2	Riveiro-Alvarez et al.,2013
MD-0111	STGD1	23	c.3386G>T	p.(Arg1129Leu)	23	c.3386G>T	p.(Arg1129Leu)	Yes	-	-	-	35y	Normal	0.1/0.8	Riveiro-Alvarez et al.,2013
MD-0116	STGD1	22;46	c.[3322C>T;6320G>A]	p.[Arg1108Cys;Arg2107His]	30	c.4469G>A	p.(Cys1490Tyr)	Yes	12	No	-	13y	Normal	0.2/0.2	Riveiro-Alvarez et al.,2013
MD-0119	STGD1	23	c.3386G>T	p.(Arg1129Leu)	23	c.3386G>T	p.(Arg1129Leu)	Yes	19	No	-	21y	-	0.4/0.1	Riveiro-Alvarez et al.,2013
MD-0125	STGD1	22	c.3210_3211dup	p.(Ser1071Cysfs*14)	40	c.5630_5644dup	p.(Lys1877_Ala1881dup)	Yes	9	9	-	14y	-	0.06/0.06	Riveiro-Alvarez et al.,2013
MD-0126	STGD1	43	c.5929G>A	p.(Gly1977Ser)	43	c.5929G>A	p.(Gly1977Ser)	Yes	9	28	12	-	-	-	Riveiro-Alvarez et al.,2013
MD-0128	STGD1	23	c.3386G>T	p.(Arg1129Leu)	23	c.3386G>T	p.(Arg1129Leu)	Yes	12	No	-	24y	-	0.1/0.1	Riveiro-Alvarez et al.,2013
MD-0133	STGD1	1	c.32T>C	p.(Leu11Pro)	19	c.2888del	p.(Gly963Alafs*14)	Yes	8	No	-	-	-	-	Riveiro-Alvarez et al.,2013
MD-0135	STGD1	8	c.1029dup	p.(Asn344*)	42	c.5882G>A	p.(Gly1961Glu)	-	-	25	-	25y	Normal	0.4/0.6	Riveiro-Alvarez et al.,2013
MD-0137	STGD1	22;46	c.[3322C>T;6320G>A]	p.[Arg1108Cys;Arg2107His]	IVS40	c.5714+5G>A	p.[=,Glu1863Leufs*33]	Yes	13	No	-	17y	Cone-pattern	0.3/0.3	Riveiro-Alvarez et al.,2013
MD-0138	STGD1	17	c.2588G>C	p.(Gly863Ala)	30	c.4537dup	p.(Gln1513Profs*42)	Yes	24	24	24	42y	-	CF/CF	Riveiro-Alvarez et al.,2013
MD-0139	STGD1	22;46	c.[3322C>T;6320G>A]	p.[Arg1108Cys;Arg2107His]	23	c.3386G>T	p.(Arg1129Leu)	Yes	20	No	-	40y	-	0.1/0.1	Riveiro-Alvarez et al.,2013
MD-0146	STGD1	21	c.3056C>T	p.(Thr1019Met)	44	c.6140T>A	p.(Ile2047Asn)	Yes	13	13	-	22y	-	0.1/0.1	Riveiro-Alvarez et al.,2013
MD-0153	STGD1	22	c.3210_3211dup	p.(Ser1071Cysfs*14)	42	c.5881G>A	p.(Gly1961Arg)	Yes	24	No	-	35y	Normal	0.1/0.05	Riveiro-Alvarez et al.,2013
MD-0155	STGD1	13	c.1804C>T	p.(Arg602Trp)	23	c.3386G>T	p.(Arg1129Leu)	Yes	17	No	-	19y	Normal	0.3/0.3	Riveiro-Alvarez et al.,2013
MD-0158	STGD1	22	c.3210_3211dup	p.(Ser1071Cysfs*14)	30	c.4537del	p.(Gln1513Argfs*13)	-	10	10	10	10y	Cone-pattern	0.2/0.2	Riveiro-Alvarez et al.,2013
MD-0162	STGD1	IVS6-Ex6	c.699_768+341del	p.(Gln234Phefs*5)	23	c.3386G>T	p.(Arg1129Leu)	-	20	-	-	-	-	-	This study
MD-0163	STGD1	30	c.4457C>T	p.(Pro1486Leu)	36	c.5172G>A	p.(Trp1724Cys)	Yes	32	No	32	-	-	-	Riveiro-Alvarez et al.,2013
MD-0164	STGD1	6	c.700C>T	p.(Gln234*)	23	c.3386G>T	p.(Arg1129Leu)	-	-	-	-	-	-	-	Riveiro-Alvarez et al.,2013
MD-0166	STGD1	IVS6-Ex6	c.699_768+341del	p.(Gln234Phefs*5)	46	c.6320G>A	p.(Arg2107His)	Yes	39	No	-	-	-	-	This study
MD-0167	STGD1	22	c.3210_3211dup	p.(Ser1071Cysfs*14)	22	c.3281C>G	p.(Pro1094Arg)	-	23	No	-	27y	-	<0.05/<0.05	Riveiro-Alvarez et al.,2013
MD-0168	STGD1	42	c.5882G>A	p.(Gly1961Glu)	45	c.6179T>G	p.(Leu2060Arg)	Yes	-	-	-	16y	-	0.16/0.1	Riveiro-Alvarez et al.,2013
MD-0170	CRD	9	c.1222C>T	p.(Arg408*)	30	c.4457C>T	p.(Pro1486Leu)	Yes	14	14	14	15y	Cone-rod pattern	0.5/0.5	Riveiro-Alvarez et al.,2013
MD-0173	STGD1	IVS30	c.4539+2064C>T	p.[=,Arg1514Leufs*36]	IVS30	c.4539+2064C>T	p.[=,Arg1514Leufs*36]	Yes	7	7	No	29y	-	CF/CF	This study
MD-0174	CRD	25	c.4918C>T	p.(Arg1640Trp)	IVS44	c.6147+2T>A	p.(?)	Yes	4	24	-	27y	Cone-rod pattern	0.01/0.025	Riveiro-Alvarez et al.,2013
MD-0176	CRD	19	c.2888del	p.(Gly963Alafs*14)	45	c.6179T>G	p.(Leu2060Arg)	Yes	10	No	-	-	-	-	Riveiro-Alvarez et al.,2013
MD-0178	STGD1	22	c.3210_3211dup	p.(Ser1071Cysfs*14)	46	c.6320G>C	p.(Arg2107Pro)	-	9	9	-	-	-	-	Riveiro-Alvarez et al.,2013
MD-0181	STGD1	22	c.3323G>A	p.(Arg1108His)	IVS38	c.5460+5G>A	p.(Trp1772Argfs*9)	Yes	16	No	-	29y	-	0.1/0.09	Riveiro-Alvarez et al.,2013
MD-0183	STGD1	43	c.5929G>A	p.(Gly1977Ser)	44	c.6079C>T	p.(Leu2027Phe)	Yes	55	50	50	56y	Cone-pattern	0.5/0.8	Riveiro-Alvarez et al.,2013
MD-0187	STGD1	28	c.4139C>T	p.(Pro1380Leu)	42	c.5882G>A	p.(Gly1961Glu)	-	39	42	-	55y	Normal	0.1/0.05	Riveiro-Alvarez et al.,2013
MD-0190	CRD	22	c.3210_3211dup	p.(Ser1071Cysfs*14)	38	c.5395A>G	p.(Asn1799Asp)	Yes	8	20	20	14y	-	CF/CF	This study
MD-0191	STGD1	13	c.1804C>T	p.(Arg602Trp)	23	c.3386G>T	p.(Arg1129Leu)	Yes	18	No	-	14y	-	0.1/0.1	Riveiro-Alvarez et al.,2013
MD-0194	STGD1	19	c.2791G>A	p.(Val931Met)	22	c.3210_3211dup	p.(Ser1071Cysfs*14)	Yes	-	-	-	-	-	-	Riveiro-Alvarez et al.,2013
MD-0196	STGD1	23	c.3386G>T	p.(Arg1129Leu)	40	c.5644A>G	p.(Met1882Val)	Segregation	27	No	-	40y	-	0.05/0.05	Riveiro-Alvarez et al.,2013
MD-0198	STGD1	19	c.2888del	p.(Gly963Alafs*14)	23	c.3386G>T	p.(Arg1129Leu)	Yes	19	19	-	26y	-	0.1/0.1	Riveiro-Alvarez et al.,2013
MD-0200	STGD1	14	c.2041C>T	p.(Arg681*)	23	c.3386G>T	p.(Arg1129Leu)	Yes	23	28	24	34y	-	0.1/0.1	Riveiro-Alvarez et al.,2013
MD-0203	STGD1	35	c.4918C>T	p.(Arg1640Trp)	42	c.5882G>A	p.(Gly1961Glu)	Yes	13	No	-	18y	-	0.2/0.1	Riveiro-Alvarez et al.,2013
MD-0206	CRD	35;36	c.[4926C>G;5044_5058del]	p.[Ser1642Arg;Val1682_Val1686del]	35;36	c.[4926C>G;5044_5058del]	p.[Ser1642Arg;Val1682_Val1686del]	Yes	18	18	16	65y	-	PL/PL	Riveiro-Alvarez et al.,2013

Family ID	Phenotype	ABCA4 variants						Segregation	Symptoms onset (age in years)			Age at ophthalmological examination	ERG	BCVA (OD/OS)	Reference
		Allele1_Exon	Allele1_cDNA	Allele 1_Protein	Allele2_Exon	Allele2_cDNA	Allele2_Protein		VA loss	VF loss	NB				
MD-0207	STGD1	30	c.4537dup	p.(Gln1513Profs*42)	42	c.5882G>A	p.(Gly1961Glu)	Yes	12	No	-	23y	Normal	0.09/0.15	Riveiro-Alvarez et al.,2013
MD-0213	CRD	13	c.1804C>T	p.(Arg602Trp)	13	c.1804C>T	p.(Arg602Trp)	Yes	15	19	-	40y	Cone-rod pattern	CF/CF	Riveiro-Alvarez et al.,2013
MD-0215	STGD1	36	c.5044_5058del	p.(Val1682_Val1686del)	47	c.6449G>A	p.(Cys2150Tyr)	-	-	-	-	-	-	-	Riveiro-Alvarez et al.,2013
MD-0216	STGD1	31	c.4577C>T	p.(Thr1526Met)	IVS40	c.5714+5G>A	p.[=,Glu1863Leufs*33]	Yes	38	35	36	43y	-	0.5/CF	Riveiro-Alvarez et al.,2013
MD-0218	STGD1	19	c.2894A>G	p.(Asn965Ser)	19	c.2894A>G	p.(Asn965Ser)	Yes	20	20	-	52y	-	CF/CF	Riveiro-Alvarez et al.,2013
MD-0225	STGD1	42	c.5882G>A	p.(Gly1961Glu)	48	c.6559C>T	p.(Gln2187*)	Yes	25	No	-	30y	Normal	0.1/0.7	Riveiro-Alvarez et al.,2013
MD-0227	STGD1	23	c.3386G>T	p.(Arg1129Leu)	41	c.5819T>C	p.(Leu1940Pro)	-	13	No	-	13y	-	CF/0.1	Riveiro-Alvarez et al.,2013
MD-0238	STGD1	23	c.3386G>T	p.(Arg1129Leu)	41	c.5819T>C	p.(Leu1940Pro)	Yes	15	No	-	15y	-	0.16/0.16	Riveiro-Alvarez et al.,2013
MD-0240	STGD1	15	c.2285C>A	p.(Ala762Glu)	42	c.5882G>A	p.(Gly1961Glu)	Yes	-	-	-	14y	-	0.2/0.16	Riveiro-Alvarez et al.,2013
MD-0242	STGD1	12	c.1715G>C	p.(Arg572Pro)	37	c.5242G>A	p.(Gly1748Arg)	Yes	18	No	-	24y	Normal	0.1/0.1	This study
MD-0244	STGD1	14	c.1957C>T	p.(Arg653Cys)	23	c.3386G>T	p.(Arg1129Leu)	Yes	16	16	16	-	-	-	This study
MD-0245	STGD1	IVS38	c.4253+5G>A	p.(Ile1377Hisfs*3)	33	c.4672G>A	p.(Gly1558Arg)	Yes	6	No	-	10y	Normal	0.1/0.1	This study
MD-0247	STGD1	23	c.3386G>T	p.(Arg1129Leu)	47	c.6410G>A	p.(Cys2137Tyr)	Yes	12	12	-	22y	Cone-pattern	0.05/0.1	This study
MD-0249	STGD1	36	c.5044_5058del	p.(Val1682_Val1686del)	IVS40	c.5714+5G>A	p.[=,Glu1863Leufs*33]	Yes	-	-	-	14y	-	0.2/0.2	Riveiro-Alvarez et al.,2013
MD-0252	STGD1	23	c.3386G>T	p.(Arg1129Leu)	23;48	c.[3386G>T;6718A>G]	p.[Arg1129Leu;Thr2240Ala]	Yes	40	No	-	-	-	-	Riveiro-Alvarez et al.,2013
MD-0254	STGD1	5	c.454C>T	p.(Arg152*)	23	c.3386G>T	p.(Arg1129Leu)	Yes	13	No	-	17y	-	0.1/0.05	Riveiro-Alvarez et al.,2013
MD-0260	STGD1	13	c.1804C>T	p.(Arg602Trp)	IVS40	c.5714+5G>A	p.[=,Glu1863Leufs*33]	Yes	-	-	-	-	-	-	Riveiro-Alvarez et al.,2013
MD-0262	STGD1	13	c.1819G>A	p.(Gly607Arg)	23	c.3386G>T	p.(Arg1129Leu)	-	19	19	19	19y	-	0.16/0.16	Riveiro-Alvarez et al.,2013
MD-0264	STGD1	23	c.3386G>T	p.(Arg1129Leu)	47	c.6449G>A	p.(Cys2150Tyr)	Yes	-	-	-	-	-	-	Riveiro-Alvarez et al.,2013
MD-0266	STGD1	42	c.5882G>A	p.(Gly1961Glu)	45	c.6179T>G	p.(Leu2060Arg)	Yes	24	No	-	29y	-	0.2/0.2	Riveiro-Alvarez et al.,2013
MD-0267	STGD1	19	c.2791G>A	p.(Val931Met)	IVS40	c.5714+5G>A	p.[=,Glu1863Leufs*33]	Yes	-	-	-	-	-	-	Riveiro-Alvarez et al.,2013
MD-0270	STGD1	6	c.634C>T	p.(Arg212Cys)	23	c.3386G>T	p.[=,Arg1129Leu]	Yes	-	-	-	-	-	-	Riveiro-Alvarez et al.,2013
MD-0277	STGD1	19	c.2888del	p.(Gly963Alafs*14)	23	c.3386G>T	p.(Arg1129Leu)	Yes	7	No	-	44y	-	CF/0.1	This study
MD-0279	STGD1	15	c.2300T>A	p.(Val767Asp)	43	c.5929G>A	p.(Gly1977Ser)	Yes	14	No	-	31y	-	CF/CF	This study
MD-0280	STGD1	12	c.1648G>A	p.(Gly550Arg)	23	c.3386G>T	p.(Arg1129Leu)	Yes	13	No	-	30y	-	0.1/<0.05	This study
MD-0281	STGD1	23	c.3386G>T	p.(Arg1129Leu)	36	c.5044_5058del	p.(Val1682_Val1686del)	Yes	9	No	-	24y	Normal	0.1/0.1	Riveiro-Alvarez et al.,2013
MD-0283	STGD1	19	c.2888del	p.(Gly963Alafs*14)	23	c.3386G>T	p.(Arg1129Leu)	-	-	-	-	30y	-	0.1/0.1	Riveiro-Alvarez et al.,2013
MD-0284	STGD1	42	c.5882G>A	p.(Gly1961Glu)	45	c.6179T>G	p.(Leu2060Arg)	-	15	No	-	30y	-	0.1/0.1	Riveiro-Alvarez et al.,2013
MD-0286	CRD	21	c.3056C>T	p.(Thr1019Met)	IVS36	c.5196+1056A>G	p.(Met1733Valfs*2)	Yes	9	14	-	14y	Scotopic and photopic extinguise	0.1/0.1	This study
MD-0287	STGD1	21	c.3056C>T	p.(Thr1019Met)	23	c.3386G>T	p.(Arg1129Leu)	Yes	9	No	-	18y	-	0.1/0.1	Riveiro-Alvarez et al.,2013
MD-0288	STGD1	8	c.982G>T	p.(Glu328*)	22	c.3322C>T	p.(Arg1108Cys)	Yes	6	6	-	-	-	-	Riveiro-Alvarez et al.,2013
MD-0290	CRD	13	c.1804C>T	p.(Arg602Trp)	35	c.4919G>A	p.(Arg1640Gln)	-	9	9	-	28y	-	<0.05/<0.05	This study
MD-0291	STGD1	23	c.3386G>T	p.(Arg1129Leu)	45	c.6179T>G	p.(Leu2060Arg)	Yes	10	No	-	16y	Normal	0.125/0.1	Riveiro-Alvarez et al.,2013
MD-0293	STGD1	30	c.4457C>T	p.(Pro1486Leu)	38	c.5395A>G	p.(Asn1799Asp)	-	53	No	-	59y	-	<0.05/<0.05	This study
MD-0295	STGD1	43	c.5917del	p.(Val1973*)	43	c.5917del	p.(Val1973*)	Yes	7	No	-	8y	Normal	0.3/0.2	Riveiro-Alvarez et al.,2013
MD-0298	CRD	8	c.982G>T	p.(Glu328*)	43	c.5929G>A	p.(Gly1977Ser)	Yes	4	4	-	12y	Cone-rod pattern	0.05/<0.05	Riveiro-Alvarez et al.,2013
MD-0299	STGD1	1	c.52C>T	p.(Arg18Trp)	45	c.6179T>G	p.(Leu2060Arg)	Yes	13	No	-	33y	-	0.1/0.1	Riveiro-Alvarez et al.,2013
MD-0300	STGD1	23	c.3386G>T	p.(Arg1129Leu)	39	c.5549T>C	p.(Leu1850Pro)	-	-	-	-	-	-	-	Riveiro-Alvarez et al.,2013
MD-0301	STGD1	13	c.1792G>A	p.(Val598Met)	43	c.5914G>A	p.(Gly1972Glu)	-	25	No	25	41y	-	0.1/0.1	This study
MD-0302	STGD1	12	c.[1622T>C;3113C>T]	p.[Leu541Pro;Ala1038Val]	42	c.5882G>A	p.(Gly1961Glu)	Yes	17	-	-	17y	Normal	0.5/0.4	Riveiro-Alvarez et al.,2013
MD-0305	STGD1	21	c.3056C>T	p.(Thr1019Met)	22	c.3323G>A	p.(Arg1108His)	Yes	17	17	-	34y	-	CF/0.1	Riveiro-Alvarez et al.,2013
MD-0307	STGD1	28	c.4200C>A	p.(Tyr1400*)	IVS35	c.5018+2T>C	p.(?)	Yes	9	No	-	31y	Cone-pattern	-	Riveiro-Alvarez et al.,2013
MD-0308	STGD1	12	c.1592A>G	p.(Glu531Gly)	IVS28	c.4253+4C>T	p.(Ile1377Hisfs*3)	Yes	21	27	-	26y	Normal	0.5/0.5	Riveiro-Alvarez et al.,2013
MD-0317	STGD1	23	c.3386G>T	p.(Arg1129Leu)	23	c.3386G>T	p.(Arg1129Leu)	-	-	-	-	-	-	-	Riveiro-Alvarez et al.,2013
MD-0318	STGD1	27	c.4069G>A	p.(Ala1357Thr)	IVS28	c.4253+4C>T	p.(Ile1377Hisfs*3)	Yes	15	15	-	27y	Cone-pattern	0.07/0.07	Riveiro-Alvarez et al.,2013
MD-0323	STGD1	27	c.3899G>A	p.(Arg1300Gln)	41	c.5819T>C	p.(Leu1940Pro)	-	-	-	-	-	-	-	This study
MD-0324	STGD1	1	c.3G>A	p.(Met1Ile)	23	c.3386G>T	p.(Arg1129Leu)	Yes	15	15	-	17y	Cone-pattern	0.2/0.2	Riveiro-Alvarez et al.,2013
MD-0326	STGD1	30	c.4457C>T	p.(Pro1486Leu)	43	c.5929G>A	p.(Gly1977Ser)	Yes	31	No	22	31y	Normal	0.5/0.5	Riveiro-Alvarez et al.,2013
MD-0329	STGD1	41	c.5819T>C	p.(Leu1940Pro)	41	c.5819T>C	p.(Leu1940Pro)	Yes	6	6	8	10y	-	0.1/0.16	Riveiro-Alvarez et al.,2013
MD-0331	STGD1	21	c.3056C>T	p.(Thr1019Met)	23	c.3386G>T	p.(Arg1129Leu)	Yes	15	No	-	37y	Normal	0.1/0.7	Riveiro-Alvarez et al.,2013
MD-0334	STGD1	23	c.3386G>T	p.(Arg1129Leu)	IVS38	c.5461-10T>C	p.(Thr1821Aspfs*6)	Yes	14	No	-	14y	Cone-pattern	0.2/0.2	Riveiro-Alvarez et al.,2013
MD-0335	STGD1	27	c.3943C>T	p.(Gln1315*)	IVS36	c.5196+1137G>A	p.[Met1733Glufs*78,=]	-	-	-	-	-	-	-	This study
MD-0336	CRD	44	c.6088C>T	p.(Arg2030*)	44	c.6088C>T	p.(Arg2030*)	Yes	16	22	-	47y	-	PL/HM	This study
MD-0340	STGD1	6	c.671del	p.(Thr224Argfs*17)	12	c.1633A>T	p.(Asn545Tyr)	Yes	12	No	-	-	-	-	This study
MD-0341	STGD1	30	c.4519G>A	p.(Gly1507Arg)	38	c.5377A>G	p.(Val1793Met)	-	-	-	-	-	-	-	This study
MD-0342	STGD1	9	c.1222C>T	p.(Arg408*)	23	c.3386G>T	p.(Arg1129Leu)	-	15	No	-	-	-	-	Riveiro-Alvarez et al.,2013
MD-0345	STGD1	13	c.1804C>T	p.(Arg602Trp)	23	c.3386G>T	p.(Arg1129Leu)	-	22	No	-	-	-	-	Riveiro-Alvarez et al.,2013
MD-0349	STGD1	20	c.2966T>C	p.(Val989Ala)	27	c.3988G>T	p.(Glu1330*)	-	11	No	-	18y	Normal	0.1/0.1	Riveiro-Alvarez et al.,2013
MD-0353	STGD1	15	c.2285C>A	p.(Ala762Glu)	45	c.6179T>G	p.(Leu2060Arg)	-	-	-	-	32y	-	<0.05/<0.05	This study
MD-0354	STGD1	8	c.1025_1038del	p.(Asp342Glyfs*6)	23	c.3386G>T	p.(Arg1129Leu)	Yes	12	No	32	-	-	-	Riveiro-Alvarez et al.,2013
MD-0359	STGD1	23	c.3386G>T	p.(Arg1129Leu)	38	c.5395A>G	p.(Asn1799Asp)	Yes	-	-	-	-	-	-	Riveiro-Alvarez et al.,2013
MD-0360	STGD1	IVS28	c.4253+4C>T	p.(Ile1377Hisfs*3)	IVS28	c.4253+4C>T	p.(Ile1377Hisfs*3)	Yes	9	No	-	14y	-	<0.05/<0.05	Riveiro-Alvarez et al.,2013
MD-0364	CRD	22	c.3210_3211dup	p.(Ser1071Cysfs*14)	28	c.4139C>T	p.(Pro1380Leu)	Yes	11	No	-	11y	Cone-rod pattern	0.1/0.1	Riveiro-Alvarez et al.,2013
MD-0370	STGD1	21	c.3056C>T	p.(Thr1019Met)	23	c.3386G>T	p.(Arg1129Leu)	Yes	10	No	14	14y	-	0.4/0.4	Riveiro-Alvarez et al.,2013
MD-0373	STGD1	16	c.2401G>A	p.(Ala801Thr)	23	c.3364G>A	p.(Glu1122Lys)	-	-	-	-	-	-	-	This study
MD-0388	STGD1	45	c.6230G>A	p.(Arg2077Gln)	47	c.6449G>A	p.(Cys2150Tyr)	-	54	No	-	56y	Normal	0.3/0.2	Riveiro-Alvarez et al.,2013
MD-0390	STGD1	19	c.2888del	p.(Gly963Alafs*14)	23	c.3386G>T	p.(Arg1129Leu)	-	2	No	-	29y	Normal	0.1/0.2	Riveiro-Alvarez et al.,2013
MD-0392	CRD	46	c.6320G>C	p.(Arg2107Pro)	46	c.6320G>C	p.(Arg2107Pro)	Yes	-	-	-	-	-	-	Riveiro-Alvarez et al.,2013
MD-0394	STGD1	5	c.454C>T	p.(Arg152*)	14	c.2023G>A	p.(Val675Ile)	-	26	No	-	44y	-	0.1/0.1	This study
MD-0395	STGD1	8	c.950del	p.(Gly317Alafs*57)	23	c.3386G>T	p.(Arg1129Leu)	-	11	No	-	40y	-	0.05/0.05	This study
MD-0400	STGD1	38	c.5451G>T	p.(Glu1817Asp)	38	c.5451G>T	p.(Glu1817Asp)	-	55	No	-	76y	Normal	HM/HM	This study
MD-0401	STGD1	19	c.2791G>A	p.(Val931Met)	11Del	c.1357_1554del	p.(Asp453_Glu518del)	-	12	-	-	-	-	-	This study

Table S1. Genetic and clinical information of 506 Spanish families with *ABCA4* mutations

Family ID	Phenotype	ABCA4 variants							Symptoms onset (age in years)			Age at ophthalmological examination	ERG	BCVA (OD/OS)	Reference
		Allele1_Exon	Allele1_cDNA	Allele 1_Protein	Allele2_Exon	Allele2_cDNA	Allele2_Protein	Segregation	VA loss	VF loss	NB				
MD-0402	STGD1	12;21	c.[1622T>C;3113C>T]	p.[Leu541Pro; Ala1038Val]	23	c.3386G>T	p.(Arg1129Leu)	Yes	-	-	-	-	-	-	This study
MD-0408	CRD	41	c.5819T>C	p.(Leu1940Pro)	47	c.6449G>A	p.(Cys2150Tyr)	-	10	No	12	29y	Cone-rod pattern	0.1/HM	Riveiro-Alvarez et al.,2013
MD-0410	STGD1	23	c.3386G>T	p.(Arg1129Leu)	23	c.3386G>T	p.(Arg1129Leu)	Yes	22	-	-	53y	-	CF/0.1	Riveiro-Alvarez et al.,2013
MD-0412	STGD1	23	c.3386G>T	p.(Arg1129Leu)	28	c.4139C>T	p.(Pro1380Leu)	Yes	18	No	-	-	-	-	Riveiro-Alvarez et al.,2013
MD-0414	STGD1	21	c.3113C>T	p.(Ala1038Val)	35	c.4918C>T	p.(Arg1640Trp)	Yes	-	-	-	-	-	-	This study
MD-0416	STGD1	41	c.5819T>C	p.(Leu1940Pro)	42	c.5882G>A	p.(Gly1961Glu)	-	10	No	-	-	-	-	Riveiro-Alvarez et al.,2013
MD-0420	STGD1	22	c.3210_3211dup	p.(Ser1071Cysfs*14)	48	c.6563T>C	p.(Phe2188Ser)	Yes	11	No	-	-	-	-	Riveiro-Alvarez et al.,2013
MD-0423	STGD1	IVS28	c.4253+4C>T	p.(Ile1377Hisfs*3)	IVS30	c.4539+2064C>T	p.[=,Arg1514Leufs*36]	-	11	No	-	13y	Cone-pattern	0.1/0.2	This study
MD-0427	STGD1	13	c.1832T>C	p.(Leu611Pro)	23	c.3386G>T	p.(Arg1129Leu)	Yes	9	No	-	12y	-	0.1/0.1	Riveiro-Alvarez et al.,2013
MD-0428	STGD1	23	c.3386G>T	p.(Arg1129Leu)	43	c.5929G>A	p.(Gly1977Ser)	-	24	No	-	24y	Normal	-	Riveiro-Alvarez et al.,2013
MD-0431	STGD1	22	c.3210_3211dup	p.(Ser1071Cysfs*14)	30	c.4457C>T	p.(Pro1486Leu)	-	20	No	-	23y	Normal	0.2/0.3	Riveiro-Alvarez et al.,2013
MD-0432	CRD	5;22	c.[560G>A;3210_3211dup]	p.[Arg187His;Ser1071Cysfs*14]	14	c.2041C>T	p.(Arg681*)	-	6	No	-	46y	Scotopic and photopic extinguise	PL/PL	Riveiro-Alvarez et al.,2013
MD-0433	STGD1	23	c.3386G>T	p.(Arg1129Leu)	30	c.4457C>T	p.(Pro1486Leu)	-	36	No	36	65y	Normal	0.4/0.5	Riveiro-Alvarez et al.,2013
MD-0435	STGD1	12	c.1609C>T	p.(Arg537Cys)	23	c.3386G>T	p.(Arg1129Leu)	-	-	-	-	-	-	-	Riveiro-Alvarez et al.,2013
MD-0437	STGD1	6	c.634C>T	p.(Arg212Cys)	23	c.3386G>T	p.(Arg1129Leu)	Yes	-	-	-	-	-	-	Riveiro-Alvarez et al.,2013
MD-0439	STGD1	8	c.871C>G	p.(Pro291Ala)	23	c.3364G>A	p.(Glu1122Lys)	-	-	-	-	-	-	-	Riveiro-Alvarez et al.,2013
MD-0445	STGD1	19	c.2888del	p.(Gly963Alafs*14)	23	c.3386G>T	p.(Arg1129Leu)	Yes	-	-	-	-	-	-	Riveiro-Alvarez et al.,2013
MD-0450	STGD1	8	c.1025_1038del	p.(Asp342Glyfs*6)	42	c.5882G>A	p.(Gly1961Glu)	Yes	37	No	37	37y	Normal	0.3/0.25	Riveiro-Alvarez et al.,2013
MD-0451	STGD1	IVS15	c.2382+5G>C	p.[=,His721_Val794del]	23	c.3386G>T	p.(Arg1129Leu)	Yes	16	No	-	29y	-	0.1/0.1	Riveiro-Alvarez et al.,2013
MD-0452	STGD1	23	c.3386G>T	p.(Arg1129Leu)	45	c.6229C>T	p.(Arg2077Trp)	Yes	18	26	-	47y	-	0.05/0.05	Riveiro-Alvarez et al.,2013
MD-0456	STGD1	13;29	c.[1819G>A;4283C>T]	p.[Gly607Arg;Thr1428Met]	41	c.5761G>A	p.(Val1921Met)	-	14	-	-	-	-	-	This study
MD-0460	STGD1	IVS36	c.5196+1137G>A	p.[Met1733Glufs*78,=]	Ex6-IVS6	c.699_768+341del	p.(Gln234Phefs*5)	-	-	-	-	-	-	-	This study
MD-0463	STGD1	8	c.1035T>G	p.(Tyr345*)	22	c.3210_3211dup	p.(Ser1071Cysfs*14)	Yes	8	No	-	-	-	-	Riveiro-Alvarez et al.,2013
MD-0464	STGD1	17	c.2588G>C	p.[Gly863Ala,Gly863del]	IVS44	c.6147+2T>A	p.(?)	Yes	33	No	-	39y	-	0.05/0.1	Riveiro-Alvarez et al.,2013
MD-0465	STGD1	23	c.3386G>T	p.(Arg1129Leu)	IVS40	c.5714+5G>A	p.[=,Glu1863Leufs*33]	-	25	No	-	26y	Normal	CF/CF	Riveiro-Alvarez et al.,2013
MD-0466	STGD1	22	c.3210_3211dup	p.(Ser1071Cysfs*14)	23	c.3386G>T	p.(Arg1129Leu)	Yes	8	No	-	13y	-	0.3/0.3	Riveiro-Alvarez et al.,2013
MD-0467	CRD	12	c.1622T>C	p.(Leu541Pro)	43	c.5917del	p.(Val1973*)	-	7	7	7	10y	Cone-pattern	0.05/0.05	Riveiro-Alvarez et al.,2013
MD-0474	STGD1	12	c.1622T>C	p.(Leu541Pro)	28	c.4234C>T	p.(Gln1412*)	Yes	-	-	-	-	-	-	Riveiro-Alvarez et al.,2013
MD-0479	STGD1	12;12	c.1751_1753delinsAT	p.(Ile584Asnfs*65)	45	c.6179T>G	p.(Leu2060Arg)	Yes	-	-	-	-	-	-	Riveiro-Alvarez et al.,2013
MD-0481	STGD1	23	c.3386G>T	p.(Arg1129Leu)	39;42	c.[5512C>G;5882G>A]	p.[His1838Asp;Gly1961Glu]	-	-	-	-	-	-	-	Riveiro-Alvarez et al.,2013
MD-0482	STGD1	14	c.2057T>C	p.(Leu686Ser)	27	c.4069G>A	p.(Ala1357Thr)	-	-	-	-	-	-	-	Riveiro-Alvarez et al.,2013
MD-0486	STGD1	17	c.2588G>C	p.[Gly863Ala,Gly863del]	45	c.6179T>G	p.(Leu2060Arg)	-	17	No	-	31y	Cone-pattern	0.1/0.1	Riveiro-Alvarez et al.,2013
MD-0493	STGD1	30	c.4457C>T	p.(Pro1486Leu)	IVS30	c.4539+2064C>T	p.[=,Arg1514Leufs*36]	Yes	44	No	-	46y	Cone-pattern	0.05/0.05	This study
MD-0494	STGD1	22;46	c.[3322C>T;6320G>A]	p.[Arg1108Cys;Arg2107His]	42	c.5882G>A	p.(Gly1961Glu)	Yes	-	-	-	-	-	-	Riveiro-Alvarez et al.,2013
MD-0497	STGD1	1	c.52C>T	p.(Arg18Trp)	23	c.3386G>T	p.(Arg1129Leu)	-	23	No	-	27y	Normal	0.1/0.1	Riveiro-Alvarez et al.,2013
MD-0498	STGD1	14	c.2041C>T	p.(Arg681*)	23	c.3386G>T	p.(Arg1129Leu)	-	14	No	-	34y	-	0.1/0.1	Riveiro-Alvarez et al.,2013
MD-0502	STGD1	5	c.560G>A	p.(Arg178His)	27	c.3871C>T	p.(Gln1291*)	Yes	24	No	-	36y	-	0.3/0.3	This study
MD-0506	STGD1	22	c.3292C>T	p.(Arg1098Cys)	35	c.4919G>A	p.(Arg1640Gln)	-	12	No	-	33y	-	0.05/0.05	Riveiro-Alvarez et al.,2013
MD-0509	STGD1	17	c.2588G>C	p.[Gly863Ala,Gly863del]	44	c.6118C>T	p.(Arg2040*)	-	23	No	-	28y	Cone-pattern	-	Riveiro-Alvarez et al.,2013
MD-0512	STGD1	30	c.4537dup	p.(Gln1513Profs*42)	IVS40	c.5714+5G>A	p.[=,Glu1863Leufs*33]	-	18	18	-	19y	-	0.2/0.2	This study
MD-0514	CRD	22	c.3323G>A	p.(Arg1108His)	36	c.5044_5058del	p.(Val1682_Val1686del)	Yes	12	No	-	25y	Cone-rod pattern	0.16/0.16	Riveiro-Alvarez et al.,2013
MD-0516	STGD1	36	c.5044_5058del	p.(Val1682_Val1686del)	36	c.5044_5058del	p.(Val1682_Val1686del)	Yes	-	-	-	-	-	-	Riveiro-Alvarez et al.,2013
MD-0518	STGD1	6	c.742_768+29del	p.(Val248_Val256del)	23	c.3386G>T	p.(Arg1129Leu)	-	19	19	-	-	-	-	Riveiro-Alvarez et al.,2013
MD-0519	STGD1	23	c.3386G>T	p.(Arg1129Leu)	43	c.5929G>A	p.(Gly1977Ser)	Yes	15	15	-	-	-	-	Riveiro-Alvarez et al.,2013
MD-0521	STGD1	22	c.3210_3211dup	p.(Ser1071Cysfs*14)	23	c.3386G>T	p.(Arg1129Leu)	-	15	No	-	23y	Cone-pattern	0.1/0.16	This study
MD-0523	STGD1	41	c.5819T>C	p.(Leu1940Pro)	42	c.5882G>A	p.(Gly1961Glu)	-	40	49	40	-	-	-	Riveiro-Alvarez et al.,2013
MD-0524	STGD1	28;35	c.[4222T>C;4918C>T]	p.[Trp1408Arg;Arg1640Trp]	42	c.5882G>A	p.(Gly1961Glu)	Yes	30	No	-	32y	Normal	0.1/0.1	Riveiro-Alvarez et al.,2013
MD-0528	STGD1	41	c.5819T>C	p.(Leu1940Pro)	44	c.6089G>A	p.(Arg2030Gln)	-	-	-	-	-	-	-	Riveiro-Alvarez et al.,2013
MD-0529	STGD1	39	c.5531G>A	p.(Gly1844Asp)	42	c.5882G>A	p.(Gly1961Glu)	Yes	17	No	-	19y	Normal	0.4/0.5	Riveiro-Alvarez et al.,2013
MD-0530	STGD1	39	c.5512C>G	p.(His1838Asp)	42	c.5882G>A	p.(Gly1961Glu)	-	-	-	-	-	-	-	Riveiro-Alvarez et al.,2013
MD-0534	STGD1	IVS28	c.4253+5G>A	p.(Ile1377Hisfs*3)	48	c.6718A>G	p.(Thr2240Ala)	-	-	-	-	-	-	-	Riveiro-Alvarez et al.,2013
MD-0535	STGD1	5	c.457A>T	p.(Ile153Leu)	35	c.4918C>T	p.(Arg1640Trp)	-	-	-	-	-	-	-	Riveiro-Alvarez et al.,2013
MD-0536	STGD1	21	c.3113C>T	p.(Ala1038Val)	IVS28	c.4253+5G>A	p.(Ile1377Hisfs*3)	-	55	-	-	-	-	-	This study
MD-0537	STGD1	8	c.1022A>T	p.(Glu341Val)	35	c.4918C>T	p.(Arg1640Trp)	-	-	-	-	-	-	-	This study
MD-0539	STGD1	13	c.1804C>T	p.(Arg602Trp)	30	c.4457C>T	p.(Pro1486Leu)	-	36	-	-	36y	Normal	0.4/0.7	Riveiro-Alvarez et al.,2013
MD-0544	STGD1	22	c.3292C>T	p.(Arg1098Cys)	IVS33	c.4773+1G>T	p.(?)	-	28	-	-	-	-	-	Riveiro-Alvarez et al.,2013
MD-0545	STGD1	23	c.3386G>T	p.(Arg1129Leu)	46	c.6329G>A	p.(Trp2110*)	-	-	-	-	41y	-	0.15/0.15	Riveiro-Alvarez et al.,2013
MD-0547	STGD1	23	c.3386G>T	p.(Arg1129Leu)	39;42	c.[5512C>G;5882G>A]	p.[His1838Asp;Gly1961Glu]	Yes	7	7	7	-	-	-	Riveiro-Alvarez et al.,2013
MD-0548	STGD1	6	c.634C>T	p.(Arg212Cys)	5;46	c.6310C>T	p.(Gln2104*)	Yes	9	-	-	24y	-	0.1/0.1	This study
MD-0553	STGD1	6	c.735T>G	p.(Tyr245*)	44	c.6089G>A	p.(Arg2030Gln)	-	44	44	-	-	-	-	Riveiro-Alvarez et al.,2013
MD-0554	STGD1	22	c.3299T>C	p.(Ile1100Thr)	22	c.3299T>C	p.(Ile1100Thr)	Consanguinity	51	53	58	64y-67y	-	0.9/0.2-0.1/HM	This study
MD-0555	STGD1	1;44	c.[1A>G;6089G>A]	p.[Met1Val;Arg2030Gln]	23	c.3386G>T	p.(Arg1129Leu)	Yes	22	25	-	35y	Normal	0.05/0.01	This study
MD-0556	STGD1	19	c.2888del	p.(Gly963Alafs*14)	23	c.3386G>T	p.(Arg1129Leu)	-	25	25	-	50y	-	0.05/0.02	This study
MD-0558	STGD1	23	c.3386G>T	p.(Arg1129Leu)	23	c.3386G>T	p.(Arg1129Leu)	-	10	No	-	-	-	-	This study
MD-0559	STGD1	35	c.4919G>A	p.(Arg1640Gln)	41	c.5819T>C	p.(Leu1940Pro)	Yes	6	6	-	7y	-	0.2/0.2	Riveiro-Alvarez et al.,2013
MD-0560	STGD1	23	c.3386G>T	p.(Arg1129Leu)	47	c.6410G>A	p.(Cys2137Tyr)	Yes	16	20	-	53y	Normal	0.1/0.1	Riveiro-Alvarez et al.,2013
MD-0563	STGD1	23	c.3364G>A	p.(Glu1122Lys)	23	c.3386G>T	p.(Arg1129Leu)	Yes	16	16	-	-	-	-	This study
MD-0565	STGD1	23	c.3380G>A	p.(Gly1127Glu)	41	c.5819T>C	p.(Leu1940Pro)	-	8	Yes	8	-	-	-	This study
MD-0568	STGD1	IVS47	c.6480-1G>T	p.(?)	40	c.5603A>T	p.(Asn1868Ile)	-	20	No	-	31y	Normal	-	This study
MD-0571	STGD1	23	c.3409A>G	p.(Arg1137Gly)	IVS38	c.5461-10T>C	p.(Thr1821Aspfs*6)	Yes	29	29	29	33y	-	0.1/0.07	This study
MD-0572	STGD1	23	c.3386G>T	p.(Arg1129Leu)	40	c.5655del	p.(Val1887Trpfs*6)	-	9	18	-	18y	-	0.1/0.1	This study

Table S1. Genetic and clinical information of 506 Spanish families with *ABCA4* mutations

Family ID	Phenotype	ABCA4 variants							Symptoms onset (age in years)			Age at ophthalmological examination	ERG	BCVA (OD/OS)	Reference
		Allele1_Exon	Allele1_cDNA	Allele 1_Protein	Allele2_Exon	Allele2_cDNA	Allele2_Protein	Segregation	VA loss	VF loss	NB				
MD-0577	STGD1	23	c.3386G>T	p.(Arg1129Leu)	43	c.5929G>A	p.(Gly1977Ser)	Yes	10	10	-	17y	Normal	0.1/0.1	Riveiro-Alvarez et al.,2013
MD-0578	STGD1	13	c.1804C>T	p.(Arg602Trp)	23	c.3386G>T	p.(Arg1129Leu)	Yes	14	18	19	18y	-	0.14/0.14	This study
MD-0580	STGD1	26	c.3832G>T	p.(Glu1278*)	42	c.5882G>A	p.(Gly1961Glu)	-	-	-	-	-	-	-	This study
MD-0581	STGD1	IVS32	c.4668-1G>A	p.(?)	42	c.5882G>A	p.(Gly1961Glu)	Yes	8	No	-	-	-	-	Riveiro-Alvarez et al.,2013
MD-0582	STGD1	34	c.4793C>A	p.(Ala1598Asp)	41	c.5819T>C	p.(Leu1940Pro)	-	-	-	-	-	-	-	Riveiro-Alvarez et al.,2013
MD-0583	STGD1	23	c.3386G>T	p.(Arg1129Leu)	42	c.5858del	p.(Pro1953Glnfs*21)	-	-	-	-	-	-	-	This study
MD-0584	CRD	6	c.634C>T	p.(Arg212Cys)	21	c.3056C>T	p.(Thr1019Met)	Yes	13	-	-	36y	Cone-rod pattern	0.1/0.1	This study
MD-0585	STGD1	13	c.1766G>A	p.(Trp589*)	23	c.3386G>T	p.(Arg1129Leu)	-	19	19	-	20y	Normal	0.2/0.3	Riveiro-Alvarez et al.,2013
MD-0588	STGD1	22	c.3323G>A	p.(Arg1108His)	IVS26	c.3862+1G>A	p.(?)	-	17	No	-	18y	Normal	0.16/0.6	Riveiro-Alvarez et al.,2013
MD-0589	STGD1	23	c.3386G>T	p.(Arg1129Leu)	35;36	c.[4926C>G;5044_5058del]	p.[Ser1642Arg;Val1681_Cys1685del]	-	-	-	-	-	-	-	This study
MD-0590	STGD1	27	c.4003_4004del	p.(Pro1335Argfs*86)	42	c.5882G>A	p.(Gly1961Glu)	-	14	15	-	14y	Normal	0.5/0.4	Riveiro-Alvarez et al.,2013
MD-0595	STGD1	31	c.4576A>G	p.(Thr1526Ala)	IVS44	c.6147+2T>A	p.(?)	-	-	-	-	20y	Normal	-	Riveiro-Alvarez et al.,2013
MD-0597	STGD1	12	c.1714C>T	p.(Arg572*)	42	c.5882G>A	p.(Gly1961Glu)	Yes	-	-	-	-	-	-	Riveiro-Alvarez et al.,2013
MD-0599	STGD1	23	c.3386G>T	p.(Arg1129Leu)	IVS30	c.4539+2064C>T	p.[=,Arg1514Leufs*36]	-	8	No	-	30y	Cone-pattern	0.1/0.1	This study
MD-0600	STGD1	23	c.3364G>A	p.(Glu1122Lys)	42	c.5882G>A	p.(Gly1961Glu)	Yes	11	No	-	-	-	-	Riveiro-Alvarez et al.,2013
MD-0602	STGD1	30	c.4457C>T	p.(Pro1486Leu)	46	c.6320G>A	p.(Arg2107His)	-	8	No	-	30y	-	CF/0.1	This study
MD-0604	CRD	4	c.393del	p.(Leu132Cysfs*22)	19	c.2888del	p.(Gly963Alafs*14)	Yes	10	-	-	10y	Cone-rod pattern	0.2/0.2	This study
MD-0605	STGD1	IVS38	c.5461-10T>C	p.(Thr1821Aspfs*6)	IVS38	c.5461-10T>C	p.(Thr1821Aspfs*6)	Endogamy	4	No	-	9y	-	0.05/0.05	Riveiro-Alvarez et al.,2013
MD-0607	STGD1	23	c.3386G>T	p.(Arg1129Leu)	43	c.5929G>A	p.(Gly1977Ser)	-	21	21	-	28y	Cone-pattern	0.1/0.1	Riveiro-Alvarez et al.,2013
MD-0611	STGD1	3	c.184C>G	p.(Pro62Ala)	13	c.1933G>A	p.(Asp645Asn)	-	30	35	45	44y	-	<0.1/.0.1	This study
MD-0617	STGD1	IVS38	c.5461-10T>C	p.(Thr1821Aspfs*6)	46	c.6320G>A	p.(Arg2107His)	Yes	9	-	-	10y	-	0.05/0.01	This study
MD-0619	STGD1	15	c.2300T>A	p.(Val767Asp)	42	c.5882G>A	p.(Gly1961Glu)	Yes	15	No	-	32y	Normal	0.1/0.1	This study
MD-0621	STGD1	23	c.3386G>T	p.(Arg1129Leu)	IVS19	c.2919-826T>A	p.(Leu973Phefs*1)	-	42	No	-	44y	-	0.4/0.5	This study
MD-0622	STGD1	30	c.4457C>T	p.(Pro1486Leu)	IVS38	c.5461-10T>C	p.(Thr1821Aspfs*6)	-	17	-	-	-	-	-	This study
MD-0625	STGD1	42	c.5882G>A	p.(Gly1961Glu)	45	c.6179T>G	p.(Leu2060Arg)	-	12	12	-	-	-	-	This study
MD-0626	CRD	23	c.3386G>T	p.(Arg1129Leu)	44	c.6088C>T	p.(Arg2030*)	-	8	No	-	36y	Cone-rod pattern	0.05/0.05	This study
MD-0631	STGD1	8	c.982G>T	p.(Glu328*)	30	c.4457C>T	p.(Pro1486Leu)	Yes	-	-	-	-	-	-	This study
MD-0632	STGD1	6	c.634C>T	p.(Arg212Cys)	23	c.3386G>T	p.(Arg1129Leu)	Yes	-	-	-	-	-	-	This study
MD-0635	STGD1	23	c.3386G>T	p.(Arg1129Leu)	35	c.4918C>T	p.(Arg1640Trp)	-	-	-	-	-	-	-	This study
MD-0640	STGD1	45	c.6179T>G	p.(Leu2060Arg)	45	c.6179T>G	p.(Leu2060Arg)	-	-	-	-	58y	-	0.05/0.05	This study
MD-0653	STGD1	14	c.2057T>C	p.(Leu686Ser)	40	c.5603A>T	p.(Asn1868Ile)	-	35	37	-	38y	-	0.1/0.1	This study
MD-0658	STGD1	45	c.6179T>G	p.(Leu2060Arg)	45	c.6179T>G	p.(Leu2060Arg)	-	9	-	-	35y	-	CF/CF	This study
MD-0663	STGD1	23	c.3386G>T	p.(Arg1129Leu)	27	c.3898C>T	p.(Arg1300*)	-	15	15	-	16y	-	0.7/0.2	This study
MD-0675	STGD1	23	c.3386G>T	p.(Arg1129Leu)	23;48	c.[3386G>T;6718A>G]	p.[Arg1129Leu;Thr2240Ala]	-	25	-	-	-	-	-	This study
MD-0678	STGD1	44	c.6089G>A	p.(Arg2030Gln)	44	c.6147_6147+7del	p.(Val2050Leufs*11)	-	11	15	15	28y	Normal	-	This study
MD-0683	STGD1	41	c.5819T>C	p.(Leu1940Pro)	47	c.6449G>A	p.(Cys2150Tyr)	Yes	-	-	-	-	-	-	This study
MD-0691	STGD1	8	c.871C>G	p.(Pro291Ala)	13	c.1804C>T	p.(Arg602Trp)	-	17	No	-	29y	-	0.01/0.01	This study
MD-0692	STGD1	12;12	c.1751_1753delinsAT	p.(Ile584Asnfs*65)	42	c.5882G>A	p.(Gly1961Glu)	-	9	No	-	-	-	-	This study
MD-0694	STGD1	3	c.287del	p.(Asn96Thrfs*19)	23;30	c.[3386G>T;4537dup]	p.[Arg1129Leu;Gln1513Profs*42]	Yes	6	6	-	7y	Cone-pattern	0.1/0.1	This study
MD-0696	STGD1	30	c.4457C>T	p.(Pro1486Leu)	43	c.5914G>A	p.(Gly1972Arg)	-	-	-	-	-	-	-	This study
MD-0698	CRD	13	c.1755del	p.(Lys585Lysfs*63)	30	c.4457C>T	p.(Pro1486Leu)	Yes	19	26	26	39y	Cone-rod pattern	0.1/0.1	This study
MD-0701	STGD1	42	c.5882G>A	p.(Gly1961Glu)	IVS46	c.6387-1G>A	p.(?)	Yes	14	14	-	15y	Cone-pattern	0.15/0.15	This study
MD-0704	CRD	16	c.2481del	p.(Thr829Argfs*14)	46	c.6179T>G	p.(Leu2060Arg)	-	-	-	-	-	-	-	This study
MD-0705	STGD1	19	c.2888del	p.(Gly963Alafs*14)	23	c.3386G>T	p.(Arg1129Leu)	-	-	-	-	-	-	-	This study
MD-0714	STGD1	23	c.3386G>T	p.(Arg1129Leu)	23	c.3386G>T	p.(Arg1129Leu)	-	6	30	-	31y	-	0.01/0.1	This study
MD-0716	STGD1	IVS40	c.5714+5G>A	p.(=,Glu1863Leufs*33)	42;45	c.[5843C>T;6179T>G]	p.[Pro1948Leu;Leu2060Arg]	-	13	13	-	21y	-	0.05/0.05	This study
MD-0717	STGD1	14	c.2041C>T	p.(Arg681*)	48	c.6718A>G	p.(Thr2240Ala)	-	40	45	45	-	-	-	This study
MD-0720	STGD1	23	c.3386G>T	p.(Arg1129Leu)	23	c.3386G>T	p.(Arg1129Leu)	Yes	31	29	-	35y	Normal	0.1/0.1	This study
MD-0723	STGD1	23	c.3386G>T	p.(Arg1129Leu)	47	c.6410G>A	p.(Cys2137Tyr)	-	13	13	Yes	15y	Cone-pattern	0.2/0.2	This study
MD-0724	STGD1	42	c.5882G>A	p.(Gly1961Glu)	45	c.6221G>A	p.(Gly2074Asp)	Yes	23	-	-	30y	Normal	0.2/0.1	This study
MD-0731	STGD1	13	c.1819G>C	p.(Gly607Arg)	42	c.5882G>A	p.(Gly1961Glu)	-	-	-	-	-	-	-	This study
MD-0736	STGD1	6	c.634C>T	p.(Arg212Cys)	28	c.4139C>T	p.(Pro1380Leu)	Yes	32	32	-	-	-	-	This study
MD-0740	STGD1	47	c.6410G>A	p.(Cys2137Tyr)	IVS28	c.4253+43G>A	p.[=, Ile1377Hisfs*3]	-	61	No	-	62y	-	0.8/0.16	This study
MD-0741	STGD1	23	c.3386G>T	p.(Arg1129Leu)	IVS38	c.5461-10T>C	p.(Thr1821Aspfs*6)	-	-	-	-	-	-	-	This study
MD-0746	STGD1	14	c.2041C>T	p.(Arg681*)	14	c.2041C>T	p.(Arg681*)	-	10	12	-	15y	Normal	0.05/0.05	This study
MD-0747	CRD	13	c.1804C>T	p.(Arg602Trp)	35	c.4849del	p.(Val1617Cysfs*45)	Yes	7	7	-	25y	-	CF/0.1	This study
MD-0748	STGD1	9	c.1222C>T	p.(Arg408*)	42	c.5882G>A	p.(Gly1961Glu)	-	14	No	-	26y	Normal	-	This study
MD-0750	STGD1	12	c.1751_1753delinsAT	p.(Ile584Asnfs*65)	42	c.5882G>A	p.(Gly1961Glu)	-	36	-	-	-	-	-	This study
MD-0751	STGD1	19	c.2888del	p.(Gly963Alafs*14)	23	c.3386G>T	p.(Arg1129Leu)	-	22	No	-	53y	-	PL/PL	This study
MD-0754	CRD	41	c.5819T>C	p.(Leu1940Pro)	42	c.5882G>A	p.(Gly1961Glu)	-	11	-	-	13y	Cone-rod pattern	0.3/0.2	This study
MD-0759	STGD1	14	c.2041C>T	p.(Arg681*)	42	c.5882G>A	p.(Gly1961Glu)	-	27	27	-	45y	-	0/0.004	This study
MD-0763	STGD1	12	c.1667T>G	p.(Met556Arg)	23	c.3386G>T	p.(Arg1129Leu)	Yes	12	13	-	14y	Normal	0.7/0.6	This study
MD-0765	STGD1	41	c.5819T>C	p.(Leu1940Pro)	41	c.5819T>C	p.(Leu1940Pro)	-	-	-	-	-	-	-	This study
MD-0768	STGD1	20;38	c.[2971G>C;3899G>A]	p.[Gly991Arg;Arg1300Gln]	38	c.5317_5318insA	p.(Ala1773Aspfs*14)	-	-	-	-	-	-	-	This study
MD-0769	STGD1	14	c.2023G>A	p.(Val675Ile)	16	c.2488G>T	p.(Glu830*)	-	41	-	-	-	-	-	This study
MD-0770	CRD	19	c.2888del	p.(Gly963Alafs*14)	30	c.4537dup	p.(Gln1513Profs*42)	-	10	10	14	39y	-	HM/HM	This study
MD-0771	STGD1	15	c.2285C>A	p.(Ala762Glu)	23	c.3386G>T	p.(Arg1129Leu)	-	-	-	-	17y	Normal	0.1/0.1	This study
MD-0778	STGD1	34	c.4793C>A	p.(Ala1598Asp)	34	c.4793C>A	p.(Ala1598Asp)	Endogamy	26	26	35	41y	Cone-pattern	0.05/0.1	This study
MD-0779	STGD1	21	c.3056C>T	p.(Thr1019Met)	IVS28	c.4253+43G>A	p.[=,Ile1377Hisfs*3]	-	25	No	25	48y	Cone-pattern	CF/CF	This study
MD-0780	STGD1	8	c.982G>T	p.(Glu328*)	39	c.5549T>C	p.(Leu1850Pro)	Yes	-	-	-	-	-	-	This study
MD-0781	STGD1	13	c.1804C>T	p.(Arg602Trp)	13	c.1804C>T	p.(Arg602Trp)	Consanguinity	11	11	-	28y	-	0.05/CF	This study
MD-0783	STGD1	IVS30;42	c.[4540-8T>C;5882G>A]	p.[Gln1513insProGln;Gly1961Glu]	8	c.950del	p.(Gly317Alafs*57)	Yes	14	17	No	17y	Cone-pattern	0.4/0.3	This study

Table S1. Genetic and clinical information of 506 Spanish families with *ABCA4* mutations

Family ID	Phenotype	Allele1_Exon	Allele1_cDNA	Allele 1_Protein	Allele2_Exon	Allele2_cDNA	Allele2_Protein	Segregation	VA loss	VF loss	NB	Age at ophthalmological examination	ERG	BCVA (OD/OS)	Reference
MD-0787	STGD1	43	c.5917del	p.(Val1973*)	43	c.5917del	p.(Val1973*)	-	-	-	-	-	-	-	This study
MD-0790	STGD1	12	c.1715G>C	p.(Arg572Pro)	35	c.4918C>T	p.(Arg1640Trp)	-	-	-	-	-	-	-	This study
MD-0791	STGD1	22;46	c.[3322C>T;6320G>A]	p.[Arg1108Cys;Arg2107His]	23	c.3386G>T	p.(Arg1129Leu)	-	26	26	-	45y	-	0.1/0.1	This study
MD-0794	STGD1	23	c.3386G>T	p.(Arg1129Leu)	27	c.3988A>G	p.(Glu1330*)	-	6	-	-	20y	Normal	0.2/0.2	This study
MD-0797	STGD1	19	c.2888del	p.(Gly963Alafs*14)	23	c.3386G>T	p.(Arg1129Leu)	-	8	No	8	25y	-	0.1/0.1	This study
MD-0799	STGD1	43	c.5929G>A	p.(Gly1977Ser)	IVS28	c.4253+43G>A	p.[=, Ile1377Hisfs*3]	-	47	47	-	-	-	-	This study
MD-0807	STGD1	6;16	c.[634C>T;2546T>C]	p.[Arg212Cys;Val849Ala]	42	c.5882G>A	p.(Gly1961Glu)	-	12	-	-	-	-	-	This study
MD-0819	STGD1	23	c.3386G>T	p.(Arg1129Leu)	43	c.5929G>A	p.(Gly1977Ser)	-	-	-	-	-	-	-	This study
MD-0821	STGD1	IVS28	c.4253+4C>T	p.(Ile1377Hisfs*3)	42	c.5882G>A	p.(Gly1961Glu)	Yes	20	No	-	-	-	-	This study
MD-0822	STGD1	23	c.3386G>T	p.(Arg1129Leu)	41	c.5819T>C	p.(Leu1940Pro)	-	-	-	-	-	-	-	This study
MD-0826	STGD1	42	c.5882G>A	p.(Gly1961Glu)	IVS30	c.4539+2064C>T	p.[=,Arg1514Leufs*36]	-	29	29	-	35y	-	0.1/0.1	This study
MD-0829	STGD1	14	c.1957C>T	p.(Arg653Cys)	23	c.3386G>T	p.(Arg1129Leu)	-	16	20	-	19y	-	0.1/0.1	This study
MD-0832	STGD1	23	c.3386G>T	p.(Arg1129Leu)	47	c.6415C>T	p.(Arg2139Trp)	-	35	No	-	42y	Normal	0.16/0.16	This study
MD-0834	STGD1	13	c.1804C>T	p.(Arg602Trp)	23	c.3386G>T	p.(Arg1129Leu)	Yes	22	No	-	32y	Normal	-	This study
MD-0846	STGD1	19	c.2888del	p.(Gly693Alafs*14)	48	c.6718A>G	p.(Thr2240Ala)	Yes	26	No	-	28y-44y	Normal	0.2/0.2-0.3/0.5	This study
MD-0847	STGD1	22;46	c.[3322C>T;6320G>A]	p.[Arg1108Cys;Arg2107His]	23	c.3386G>T	p.(Arg1129Leu)	Yes	22	22	-	24y-28y	Normal	0.6/0.6-0.1/0.1	This study
MD-0852	STGD1	27	c.3874C>T	p.(Gln1292*)	21	c.3113C>T	p.(Ala1038Val)	-	-	-	-	43y	-	0.4/0.5	This study
MD-0853	STGD1	30	c.4537dup	p.(Gln1513Profs*42)	40	c.5603A>T	p.(Asn1868Ile)	-	-	-	-	57y	Normal	1/1	This study
MD-0856	STGD1	42	c.5882G>A	p.(Gly1961Glu)	43	c.5917del	p.(Val1973*)	-	29	29	-	32y	-	CF/0.3	This study
MD-0859	STGD1	6	c.611C>A	p.(Ala204Asp)	23	c.3386G>T	p.(Arg1129Leu)	Yes	5	14	49	40y	-	0.05/0.05	This study
MD-0863	STGD1	IVS30	c.4539+2064C>T	p.[=,Arg1514Leufs*36]	42	c.5882G>A	p.(Gly1961Glu)	Yes	-	-	-	37y	-	0.05/0.05	This study
MD-0864	STGD1	13	c.1868A>G	p.(Gln623Arg)	22	c.3259A>G	p.(Glu1087Lys)	-	-	-	-	-	-	-	This study
MD-0869	STGD1	13	c.1832T>C	p.(Leu611Pro)	23	c.3386G>T	p.(Arg1129Leu)	Yes	9	9	-	21y	Cone-pattern	0.05/0.05	This study
MD-0870	STGD1	3	c.223T>G	p.(Cys75Gly)	23	c.3386G>T	p.(Arg1129Leu)	-	-	-	-	-	-	-	This study
MD-0877	STGD1	13	c.1804C>T	p.(Arg602Trp)	42	c.5882G>A	p.(Gly1961Glu)	Yes	17	No	-	-	-	-	This study
MD-0878	CRD	8	c.982G>T	p.(Glu328*)	8	c.982G>T	p.(Glu328*)	Yes	-	Yes	-	25y	-	CF/0.05	This study
MD-0881	STGD1	23	c.3386G>T	p.(Arg1129Leu)	44	c.6089G>A	p.(Arg2030Gln)	Yes	30	No	30	40y	-	0.05/0.1	This study
MD-0882	STGD1	IVS38	c.5461-10T>C	p.(Thr1821Aspfs*6)	IVS39	c.5714+5G>A	p.[=,Glu1863Leufs*33]	-	13	No	-	-	-	-	This study
MD-0885	STGD1	13	c.1804C>T	p.(Arg602Trp)	46	c.6316C>T	p.(Arg2106Cys)	Yes	38	No	-	-	-	-	This study
MD-0886	STGD1	35;36	c.[4926C>G;5044_5058del]	p.[Ser1642Arg;Val1682_Val1686del]	46	c.6380C>T	p.(Ser2127Phe)	Yes	8	No	-	-	-	-	This study
MD-0894	CRD	22	c.3210_3211dup	p.(Ser1071Cysfs*14)	22	c.3210_3211dup	p.(Ser1071Cysfs*14)	-	5	No	-	20y	-	CF/CF	This study
MD-0895	STGD1	23	c.3364G>A	p.(Glu1122Lys)	IVS28	c.4253+43G>A	p.[=,Ile1377Hisfs*3]	-	8	50	-	70y	-	<0.1/<0.1	This study
MD-0906	STGD1	13	c.1804C>T	p.(Arg602Trp)	23	c.3364G>A	p.(Glu1122Lys)	Yes	27	-	-	34y	Normal	0.2/0.2	This study
MD-0908	STGD1	8	c.1025_1038del	p.(Asp342Glyfs*6)	23	c.3386G>T	p.(Arg1129Leu)	Yes	27	-	-	19y	Cone-pattern	0.2/0.2	This study
MD-0911	STGD1	13	c.1879G>T	p.(Glu627*)	41	c.5819T>C	p.(Leu1940Pro)	Yes	8	No	-	9y	-	0.05/0.05	This study
MD-0918	CRD	44	c.6088C>T	p.(Arg2030*)	45	c.6216T>A	p.(Ser2072Arg)	-	10	10	-	32y	Cone-rod pattern	-	This study
MD-0925	STGD1	23	c.3386G>T	p.(Arg1129Leu)	43	c.5929G>A	p.(Gly1977Ser)	Yes	36	36	36	-	-	-	This study
MD-0929	CRD	36	c.5044_5058del	p.(Val1682_Val1686del)	36	c.5044_5058del	p.(Val1682_Val1686del)	-	6	8	-	26y	-	0.05/0.05	This study
MD-0932	STGD1	28;35	c.[4222T>C;4918C>T]	p.[Trp1408Arg;Arg1640Trp]	42	c.5882G>A	p.(Gly1961Glu)	-	12	No	-	-	-	-	This study
MD-0939	STGD1	13	c.1804C>T	p.(Arg602Trp)	IVS30	c.4539+2064C>T	p.[=,Arg1514Leufs*36]	-	-	-	-	20y	-	CF/CF	This study
MD-0939	STGD1	IVS36	c.5196+1137G>A	p.[Met1733Glufs*78,=]	IVS38	c.5461-10T>C	p.(Thr1821Aspfs*6)	-	4	21	-	25y	-	0.1/0.1	This study
MD-0941	STGD1	19	c.2894A>G	p.(Asn965Ser)	45	c.6272T>A	p.(Leu2091Gln)	Yes	16	11	22	37y	-	0.05/0.1	This study
MD-0943	STGD1	23	c.3386G>T	p.(Arg1129Leu)	40	c.5644A>G	p.(Met1882Val)	-	22	22	-	26y	Normal	0.4/0.4	This study
MD-0945	STGD1	23	c.3386G>T	p.(Arg1129Leu)	IVS29	c.4353-1G>A	p.(?)	-	5	5	-	37y	-	0.08/0.05	This study
MD-0946	STGD1	23	c.3386G>T	p.(Arg1129Leu)	29;48	c.[3386G>T;6718A>G]	p.[Arg1129Leu;Thr2240Ala]	-	33	45	-	53y	-	0.1/0.1	This study
MD-0952	STGD1	23	c.3386G>T	p.(Arg1129Leu)	42	c.5882G>A	p.(Gly1961Glu)	-	26	26	26	-	-	-	This study
MD-0953	STGD1	22	c.3210_3211dup	p.(Ser1071Cysfs*14)	23	c.3386G>T	p.(Arg1129Leu)	Yes	-	-	-	-	-	-	This study
MD-0954	STGD1	21	c.3113C>T	p.(Ala1038Val)	IVS21	c.3050+5G>A	p.(Leu973_His1017delinsPhe)	Yes	-	-	-	-	-	-	This study
MD-0955	STGD1	19	c.2888del	p.(Gly693Alafs*14)	23	c.3386G>T	p.(Arg1129Leu)	-	10	No	-	41y	Normal	CF/CF	This study
MD-0956	STGD1	17/IVS38	c.[2588G>C;5461-10T>C]	p.[Gly863Ala,Gly863del;Thr1821Aspfs*6)]	40	c.5603A>T	p.(Asn1868Ile)	-	18	No	-	19y	Normal	0.16/0.16	This study
MD-0959	STGD1	23	c.3386G>T	p.(Arg1129Leu)	23	c.3386G>T	p.(Arg1129Leu)	Yes	72	No	No	72	-	1/1	This study
MD-0960	STGD1	23	c.3383A>G	p.(Asp1128Gly)	23	c.3383A>G	p.(Asp1128Gly)	Yes	11	11	-	14y	Normal	0.2/0.2	This study
MD-0962	STGD1	29	c.4254C>A	p.(Ser1418Arg)	29	c.4254C>A	p.(Ser1418Arg)	Yes	-	-	-	10y	Normal	CF/0.3	This study
MD-0963	STGD1	14	c.2023G>A	p.(Val675Ile)	45	c.6179T>G	p.(Leu2060Arg)	-	-	-	-	-	-	-	This study
MD-0964	STGD1	15	c.2285C>A	p.(Ala762Glu)	34	c.4793C>A	p.(Ala1598Asp)	-	15	No	-	-	-	-	This study
MD-0966	STGD1	39	c.5549T>C	p.(Leu1850Pro)	42	c.5882G>A	p.(Gly1961Glu)	Yes	-	-	-	-	-	-	This study
MD-0967	STGD1	6	c.634C>T	p.(Arg212Cys)	IVS34	c.4848+3A>G	p.[Gly1592_Lys1616del,=]	Yes	9	No	-	12y	Normal	0.1/0.1	This study
MD-0968	STGD1	23	c.3386G>T	p.(Arg1129Leu)	27	c.3988G>T	p.(Glu1330*)	Yes	19	19	-	23y-32y	-	0.1/0.1	This study
MD-0969	STGD1	30	c.4457C>T	p.(Pro1486Leu)	IVS36	c.5196+1137G>A	p.[Met1733Glufs*78,=]	-	-	-	-	51y	-	0.2/0.2	This study
MD-0979	STGD1	23	c.3386G>T	p.(Arg1129Leu)	27	c.3943C>T	p.(Gln1315*)	Yes	-	-	-	11y	Cone-pattern	0.3/0.3	This study
MD-0980	STGD1	45	c.6179T>G	p.(Leu2060Arg)	48	c.6718A>G	p.(Thr2240Ala)	-	-	-	-	34y	-	0.16/0.1	This study
MD-0981	STGD1	13	c.1804C>T	p.(Arg602Trp)	IVS30	c.4539+2064C>T	p.[=,Arg1514Leufs*36]	-	15	15	-	16y	-	0.05/0.1	This study
MD-0984	STGD1	23	c.3386G>T	p.(Arg1129Leu)	41	c.5819T>C	p.(Leu1940Pro)	-	24	24	24	-	-	-	This study
MD-0985	STGD1	13	c.1766G>C	p.(Trp589Ser)	42	c.5882G>A	p.(Gly1961Glu)	-	-	-	-	-	-	-	This study
MD-0989	STGD1	41	c.5819T>C	p.(Leu1940Pro)	41	c.5819T>C	p.(Leu1940Pro)	-	7	7	-	15y	-	CF/CF	This study
MD-0991	STGD1	30	c.4457C>T	p.(Pro1486Leu)	30	c.4457C>T	p.(Pro1486Leu)	Yes	25	-	-	33y	Cone-rod pattern	1.2/1.2	This study
MD-0993	STGD1	13	c.1804C>T	p.(Arg602Trp)	23	c.3386G>T	p.(Arg1129Leu)	Yes	14	No	-	15y	-	0.05/0.05	This study
MD-0995	STGD1	11	c.1364T>A	p.(Leu455Gln)	14	c.2099G>A	p.(Trp700*)	-	-	-	-	-	-	-	This study
MD-0997	STGD1	23	c.3383A>G	p.(Asp1128Gly)	23	c.3386G>T	p.(Arg1129Leu)	Yes	-	-	-	-	-	-	This study
MD-1000	STGD1	16	c.[2483C>T;2481del]	p.[Pro828Leu;Thr829Argfs*14]	40	c.5603A>T	p.(Asn1868Ile)	-	-	-	-	-	-	-	This study
MD-1001	STGD1	44	c.6089G>A	p.(Arg2030Gln)	47	c.6410G>A	p.(Cys2137Tyr)	-	-	-	-	-	-	-	This study
MD-1002	STGD1	42	c.5882G>A	p.(Gly1961Glu)	IVS30	c.4539+2064C>T	p.[=,Arg1514Leufs*36]	Yes	-	-	-	-	-	-	This study

Table S1. Genetic and clinical information of 506 Spanish families with *ABCA4* mutations

Family ID	Phenotype	*ABCA4* variants							Symptoms onset (age in years)			Age at ophthalmological examination	ERG	BCVA (OD/OS)	Reference
		Allele1_Exon	Allele1_cDNA	Allele 1_Protein	Allele2_Exon	Allele2_cDNA	Allele2_Protein	Segregation	VA loss	VF loss	NB				
MD-1007	STGD1	23	c.3383A>G	p.(Asp1128Gly)	39	c.5549T>C	p.(Leu1850Pro)	-	15	No	-	51y	-	CF/0.05	This study
MD-1010	STGD1	12	c.1622T>C	p.(Leu541Pro)	17	c.2588G>C	p.(Gly863Ala,Gly863del)	Yes	-	-	-	-	-	-	This study
MD-1015	STGD1	38	c.5384T>C	p.(Leu1795Ser)	38	c.5384T>C	p.(Leu1795Ser)	Yes	39	No	-	54y	-	CF/CF	This study
MD-1022	STGD1	23	c.3386G>T	p.(Arg1129Leu)	IVS30	c.4539+2064C>T	p.[=,Arg1514Leufs*36]	-	18	No	-	30y	-	0.1/0.1	This study
MD-1025	STGD1	35	c.4918G>T	p.(Arg1640Trp)	45	c.6179T>G	p.(Leu2060Arg)	-	8	8	-	9y	-	0.05/0.1	This study
MD-1034	STGD1	23	c.3383A>G	p.(Asp1128Gly)	23	c.3383A>G	p.(Asp1128Gly)	Yes	38	No	-	40y	Normal	0.6/0.2	This study
MD-1036	STGD1	22;44	c.[3322C>T;6071A>G]	p.[Arg1108His;Asp2024Gly]	30	c.4457C>T	p.(Pro1486Leu)	Yes	12	No	13	18y	Normal	0.2/0.16	This study
MD-1041	CRD	22	c.3210_3211dup	p.(Ser1071Cysfs*14)	36	c.5196+1G>A	p.(?)	Yes	10	12	-	13y	Cone-rod pattern	0.05/0.1	This study
MD-1042	STGD1	23	c.3383A>G	p.(Asp1128Gly)	IVS41	c.5836-1G>C	p.(?)	-	-	-	-	-	-	-	This study
MD-1045	STGD1	13	c.1819G>C	p.(Gly607Arg)	42	c.5882G>A	p.(Gly1961Glu)	-	17	No	-	26y	-	0.1/0.1	This study
MD-1046	STGD1	23	c.3386G>T	p.(Arg1129Leu)	IVS33	c.4773+1del	p.(?)	Yes	17	No	-	-	-	-	This study
MD-1047	STGD1	22	c.3210_3211dup	p.(Ser1071Cysfs*14)	22	c.3210_3211dup	p.(Ser1071Cysfs*14)	Yes	9	9	-	10y	-	0.1/0.1	This study
MD-1048	STGD1	22	c.3210_3211dup	p.(Ser1071Cysfs*14)	23	c.3386G>T	p.(Arg1129Leu)	Yes	10	10	-	12y	Normal	0.1/0.12	This study
MD-1049	STGD1	4	c.428C>T	p.(Pro143Leu)	IVS40	c.5714+5G>A	p.[=,Glu1863Leufs*33]	Yes	-	-	-	-	-	-	This study
MD-1050	STGD1	23	c.3386G>T	p.(Arg1129Leu)	23	c.3386G>T	p.(Arg1129Leu)	Yes	35	-	-	55y	-	0.1/0.05	This study
MD-1054	STGD1	IVS40	c.5714+5G>A	p.(=,Glu1863Leufs*33)	IVS30	c.4539+2064C>T	p.[=,Arg1514Leufs*36]	-	24	-	-	-	-	-	This study
MD-1059	STGD1	15	c.2285C>A	p.(Ala762Glu)	23	c.3386G>T	p.(Arg1129Leu)	-	14	14	14	15y	Cone-pattern	0.05/0.6	This study
MD-1063	STGD1	35;36	c.[4926C>G;5044_5058del]	p.[Ser1642Arg; Val1681_Cys1685del]	35;36	c.[4926C>G;5044_5058del]	p.([Ser1642Arg; Val1681_Cys1685del]	-	7	No	-	8y	-	0.125/0.04	This study
MD-1064	STGD1	23	c.3386G>T	p.(Arg1129Leu)	IVS30	c.4539+2064C>T	p.[=, Arg1514Leufs*36]#	-	-	-	-	-	-	-	This study
MD-1067	STGD1	45	c.6229C>T	p.(Arg2077Trp)	45	c.6229C>T	p.(Arg2077Trp)	-	9	9	-	10y	-	0.16/0.4	This study
MD-1070	STGD1	13	c.1927G>A	p.(Ala643Met)	44	c.6089G>A	p.(Arg2030Gln)	-	23	23	-	48y	-	0.2/HM	This study
MD-1071	STGD1	23	c.3386G>T	p.(Arg1129Leu)	23	c.3386G>T	p.(Arg1129Leu)	Yes	-	-	-	-	-	-	This study
MD-1072	STGD1	23	c.3386G>T	p.(Arg1129Leu)	43	c.5929G>A	p.(Gly1977Ser)	Yes	24	No	18	26y	Normal	0.15/0.8	This study
MD-1075	STGD1	9	c.1222C>T	p.(Arg408*)	IVS28;40	c.[4253+43G>A;5603A>T]	p.[=,Ile1377Hisfs*3];(.)(Asn1868Ile)	-	-	-	-	-	-	-	This study
MD-1077	STGD1	8	c.1025_1038del	p.(Asp342Glyfs*6)	22	c.3292C>T	p.(Arg1098Cys)	Yes	-	-	-	-	-	-	This study
MD-1086	STGD1	30	c.4457C>T	p.(Pro1486Leu)	IVS30	c.4539+2064C>T	p.[=,Arg1514Leufs*36]	Yes	-	-	-	-	-	-	This study
MD-1087	STGD1	IVS34	c.4848+3A>G	p.[Gly1592_Lys1616del,=]	42	c.5882G>A	p.(Gly1961Glu)	-	-	-	-	-	-	-	This study
MD-1100	STGD1	3	c.286A>G	p.(Asn96Asp)	23	c.3386G>T	p.(Arg1129Leu)	-	-	-	-	48y	Normal	-	This study
MD-1101	STGD1	17	c.2588G>T	p.(Gly863Ala)	IVS30	c.4539+2064C>T	p.[=,Arg1514Leufs*36]	-	44	46	-	48y	Normal	0.8/0.6	This study
MD-1102	STGD1	23;48	c.[3386G>T;6718A>G]	p.[Arg1129Leu;Thr2240Ala]	IVS28	c.4253+5G>A	p.(Ile1377Hisfs*3)	Yes	5	9	9	14y	Cone-pattern	0.16/0.16	This study
MD-1106	STGD1	23	c.3386G>T	p.(Arg1129Leu)	40	c.5644A>G	p.(Met1882Val)	Yes	9	No	-	19y	Normal	1/1	This study
MD-1107	STGD1	22	c.3210_3211dup	p.(Ser1071Cysfs*14)	22	c.3277G>A	p.(Asp1093Asn)	Yes	-	-	-	-	-	-	This study
MD-1110	STGD1	29	c.4322G>A	p.(Gly1441Asp)	48	c.6718A>G	p.(Thr2240Ala)	-	47	47	47	49y	-	1/1	This study
MD-1114	STGD1	23	c.3386G>T	p.(Arg1129Leu)	30	c.4469G>A	p.(Cys1490Tyr)	-	22	No	-	25y	Normal	0.3/0.4	This study
MD-1116	STGD1	6	c.634C>T	p.(Arg212Cys)	30	c.4457C>T	p.(Pro1486Leu)	-	10	No	45	53y	Cone-pattern	0.05/0.05	This study
MD-1119	STGD1	23	c.3386G>T	p.(Arg1129Leu)	41	c.5819T>C	p.(Leu1940Pro)	Yes	-	-	-	-	-	-	This study
MD-1120	STGD1	19	c.2888del	p.(Gly963Alafs*14)	40	c.5603A>T	p.(Asn1868Ile)	-	-	-	-	50y	-	0.3/0.2	This study
MD-1121	STGD1	ISV40	c.5714+1G>A	p.(?)	48	c.6718A>G	p.(Thr2240Ala)	-	-	-	-	-	-	-	This study
MD-1123	STGD1	6	c.634C>T	p.(Arg212Cys)	23	c.3386G>T	p.(Arg1129Leu)	-	19	No	-	27y	Normal	0.2/0.2	This study
MD-1130	STGD1	39	c.5549T>C	p.(Leu1850Pro)	39	c.5549T>C	p.(Leu1850Pro)	Yes	9	-	-	-	-	-	This study
MD-1137	STGD1	23	c.3386G>T	p.(Arg1129Leu)	41	c.5819T>C	p.(Leu1940Pro)	-	13	18	-	50y	-	0.05/0.05	This study
MD-1138	STGD1	4	c.378G>A	p.(Trp126*)	6	c.634C>T	p.(Arg212Cys)	Yes	9	No	-	10y	Normal	0.2/0.16	This study
MD-1139	STGD1	23	c.3386G>T	p.(Arg1129Leu)	48	c.6559C>T	p.(Gln2187*)	-	-	-	-	-	-	-	This study
MD-1146	STGD1	14	c.2023G>A	p.(Val675Ile)	47	c.6410G>A	p.(Cys2137Tyr)	Yes	46	-	-	46y	Normal	1/1	This study
MD-1147	STGD1	23	c.3386G>T	p.(Arg1129Leu)	28	c.4139C>T	p.(Pro1380Leu)	-	30	No	-	31y	Normal	0.05/0.1	This study
MD-1156	CRD	3	c.287del	p.(Asn96Thrfs*19)	12	c.1667T>C	p.(Met556Arg)	Yes	9	9	-	9y	Cone-rod pattern	0.05/0.05	This study
MD-1157	CRD	8	c.982G>T	p.(Glu328*)	42	c.5882G>A	p.(Gly1961Glu)	Yes	22	No	-	22y	-	0.5/0.6	This study
MD-1164	STGD1	30	c.4457C>T	p.(Pro1486Leu)	30	c.4457C>T	p.(Pro1486Leu)	-	42	No	-	46y	-	0.2/0.3	This study
MD-1184	STGD1	23	c.3386G>T	p.(Arg1129Leu)	ivs28	c.4253+4C>T	p.(Ile1377Hisfs*3)	-	20	No	-	-	-	-	This study
MD-1193	STGD1	22	c.3292C>T	p.(Arg1098Cys)	40	c.5630_5644dup	p.(Lys1877_Ala1881dup)	-	-	-	-	-	-	-	This study
MD-1202	STGD1	13	c.1804C>T	p.(Arg602Trp)	42	c.5882G>A	p.(Gly1961Glu)	Yes	20	20	-	70y	-	0.05/0.05	This study
MD-1205	STGD1	41	c.5819T>C	p.(Leu1940Pro)	42	c.5882G>A	p.(Gly1961Glu)	-	30	30	-	52y	-	0.014/0.014	This study
MD-1207	STGD1	23	c.3386G>T	p.(Arg1129Leu)	30	c.4436G>A	p.(Trp1479*)	-	14	18	-	30y	Normal	0.1/0.1	This study
MD-1208	STGD1	23	c.3386G>T	p.(Arg1129Leu)	47	c.6437G>T	p.(Gly2146Val)	-	12	No	-	44y	-	0.1/0.05	This study
MD-1211	STGD1	6	c.634C>T	p.(Arg212Cys)	23	c.3386G>T	p.(Arg1129Leu)	-	15	No	-	20y	-	0.1/0.05	This study
MD-1212	STGD1	23	c.3386G>T	p.(Arg1129Leu)	38	c.5318C>T	p.(Ala1773Val)	Yes	16	No	-	37y	-	0.05/0.1	This study
MD-1215	STGD1	13	c.1819G>C	p.(Gly607Arg)	40	c.5603A>T	p.(Asn1868Ile)	-	-	-	-	-	-	-	This study
MD-1216	STGD1	13	c.1766G>A	p.(Trp589*)	23	c.3386G>T	p.(Arg1129Leu)	-	25	25	-	62y	-	CF/CF	This study
MD-1219	STGD1	8	c.874A>C	p.(Ser292Arg)	22	c.3251T>C	p.(Ille1084Thr)	-	35	40	-	55y	-	0.3/0.1	This study
MD-1220	STGD1	42	c.5882G>A	p.(Gly1961Glu)	47	c.6449G>A	p.(Cys2150Tyr)	-	43	43	-	56y	-	0.1/0.1	This study
MD-1223	STGD1	21	c.3056C>T	p.(Thr1019Met)	35	c.4849del	p.(Val1617Cysfs*45)	Yes	10	No	-	13y	-	0.1/0.1	This study
MD-1228	STGD1	23	c.3386G>T	p.(Arg1129Leu)	39	c.5549T>C	p.(Leu1850Pro)	-	17	20	-	19y	-	0.1/0.1	This study
MD-1230	STGD1	22	c.3210_3211dup	p.(Ser1071Cysfs*14)	38	c.5318C>T	p.(Ala1773Val)	-	7	7	-	8y	-	0.1/0.1	This study
MD-1233	STGD1	22;46	c.[3322C>T;6320G>A]	p.[Arg1108Cys;Arg2107His]	42	c.5882G>A	p.(Gly1961Glu)	-	39	44	-	58y	-	0.1/0.05	This study
MD-1235	STGD1	27	c.3988G>T	p.(Glu1330*)	IVS38	c.5461-1G>T	p.(?)	-	8	-	-	-	-	-	This study
MD-1236	STGD1	23	c.3386G>T	p.(Arg1129Leu)	48	c.6718A>G	p.(Thr2240Ala)	-	28	-	-	-	-	-	This study
MD-1238	STGD1	22	c.3277G>A	p.(Asp1093Asn)	23	c.3386G>T	p.(Arg1129Leu)	-	-	-	-	60y	-	1/40 AO	This study
MD-1239	STGD1	22	c.3259G>A	p.(Glu108Lys)	36	c.5044_5058del	p.(Val1682_Val1686del)	-	-	-	-	-	-	-	This study
MD-1240	STGD1	1	c.32T>C	p.(Leu11Pro)	30	c.4457C>T	p.(Pro1486Leu)	-	-	-	-	-	-	-	This study
MD-1243	STGD1	23	c.3386G>T	p.(Arg1129Leu)	27	c.3943C>T	p.(Gln1315*)	-	-	-	-	-	-	-	This study
MD-1246	STGD1	35	c.4918G>T	p.(Arg1640Trp)	46	c.6329G>A	p.(Trp2110*)	-	-	-	-	-	-	-	This study
MD-1249	STGD1	8	c.871C>G	p.(Pro291Ala)	41	c.5819T>C	p.(Leu1940Pro)	-	-	-	-	-	-	-	This study

Family ID	Phenotype	ABCA4 variants							Symptoms onset (age in years)			Age at ophthalmological examination	ERG	BCVA (OD/OS)	Reference
		Allele1_Exon	Allele1_cDNA	Allele 1_Protein	Allele2_Exon	Allele2_cDNA	Allele2_Protein	Segregation	VA loss	VF loss	NB				
MD-1251	STGD1	22;46	c.[3322C>T;6320G>A]	p.[Arg1108Cys;Arg2107His]	IVS38	c.5461-10T>C	p.(Thr1821Aspfs*6)	-	10	-	-	18y	-	0.04/0.05	This study
MD-1255	STGD1	35	c.4918C>T	p.(Arg1640Trp)	42	c.5882G>A	p.(Gly1961Glu)	-	-	-	-	-	-	-	This study
MD-1256	STGD1	IVS14	c.2161-8G>A	p.(His721_Val794del)	17	c.2613G>A	p.(Trp871*)	-	-	-	-	-	-	-	This study
MD-1257	STGD1	3	c.214G>A	p.(Gly72Arg)	43	c.5929G>A	p.(Gly1977Ser)	-	7	7	-	25y	-	0.02/0.02	This study
MD-1258	STGD1	22	c.3259G>A	p.(Glu1087Lys)	22	c.3259G>A	p.(Glu1087Lys)	-	-	-	-	-	-	-	This study
MD-1264	STGD1	23	c.3386G>T	p.(Arg1129Leu)	27	c.4070C>T	p.(Ala1357Val)	-	-	-	-	-	-	-	This study
MD-1271	STGD1	22;46	c.[3322C>T;6320G>A]	p.[Arg1108Cys;Arg2107His]	23	c.3386G>T	p.(Arg1129Leu)	Yes	21	No	-	24y	-	0.25/0.2	This study
MD-1272	STGD1	IVS6	c.768+2T>G	p.(?)	43	c.5899T>G	p.(Cys1967Gly)	Yes	-	-	-	-	-	-	This study
MD-1273	STGD1	23;48	c.[3386G>T;6718A>G]	p.[Arg1129Leu;Thr2240Ala]	23;48	c.[3386G>T;6718A>G]	p.[Arg1129Leu;Thr2240Ala]	Yes	-	-	-	-	-	-	This study
MD-1274	STGD1	23	c.3386G>T	p.(Arg1129Leu)	Intrón 28	c.4253+4C>T	p.(Ile1377Hisfs*3)	-	14	14	-	15y	-	0.1/0.1	This study
MD-1279	STGD1	23	c.3386G>T	p.(Arg1129Leu)	IVS38;40	c.[5461-10T>C;5603A>T]	p.[Thr1821Aspfs*6,Thr1821Valfs*13;Asn1868Ile]	-	-	-	-	-	-	-	This study
MD-1298	STGD1	35	c.4919G>A	p.(Arg1640Gln)	Intron 40	c.5714+5G>A	p.[=,Glu1863Leufs*33]	-	-	-	-	13y	-	0.4/0.16	This study
MD-1300	STGD1	23	c.3386G>T	p.(Arg1129Leu)	Intron 38	c.5461-10T>C	p.(Thr1821Aspfs*6)	-	-	-	-	-	-	-	This study
MD-1302	STGD1	30	c.4457C>T	p.(Pro1486Leu)	30	c.4457C>T	p.(Pro1486Leu)	-	-	-	-	-	-	-	This study
MD-1339	STGD1	22	c.3210_3211dup	p.(Ser1071Cysfs*14)	23	c.3386G>T	p.(Arg1129Leu)	-	38	37	No	65y	Cone-pattern	0.05/0.05	This study
MD-1343	STGD1	38	c.5318C>T	p.(Ala1773Val)	38	c.5318C>T	p.(Ala1773Val)	Consanguinity	11	30	No	56y	-	CF/CF	This study
MD-1348	STGD1	22	c.3210_3211dup	p.(Ser1071Cysfs*14)	23	c.3386G>T	p.(Arg1129Leu)	-	-	-	-	20y	Cone-pattern	0.2/0.2	This study
MD-1349	STGD1	13	c.1804C>T	p.(Arg602Trp)	42	c.5882G>A	p.(Gly1961Glu)	-	26	26	No	29y	Normal	0.15/0.15	This study
MD-1350	STGD1	17	c.2588G>C	p.[Gly863Ala,Gly863del]	19	c.2888del	p.(Gly963Alafs*14)	-	-	-	-	-	-	-	This study
MD-1353	STGD1	12	c.1609C>T	p.(Arg537Cys)	43	c.5981G>A	p.(Gly1994Glu)	-	20	-	-	-	-	-	This study
MD-1356	STGD1	23	c.3386G>T	p.(Arg1129Leu)	23	c.3386G>T	p.(Arg1129Leu)	-	40	No	No	43y	Normal	0.9/0.9	This study
MD-1359	STGD1	12;21	c.[1622T>C;3113C>T]	p.[Leu541Pro;Ala1038Val]	42	c.5882G>A	p.(Gly1961Glu)	-	-	-	-	-	-	-	This study
MD-1375	STGD1	23	c.3386G>T	p.(Arg1129Leu)	23	c.3386G>T	p.(Arg1129Leu)	-	-	-	-	-	-	-	This study
MD-1376	STGD1	23	c.3386G>T	p.(Arg1129Leu)	38	c.5318C>T	p.(Ala1773Val)	-	-	-	-	-	-	-	This study
MD-1377	STGD1	13	c.1853G>T	p.(Gly618Val)	39	c.5531_5557dup	p.(Gly1844_Gln1852dup)	-	-	-	-	-	-	-	This study
MD-1381	STGD1	IVS28	c.4253+43G>A	p.[=, Ile1377Hisfs*3]	35	c.4919G>A	p.(Arg1640Gln)	-	45	No	No	63y	Normal	0.7/1	This study
RP-0033	CRD	13	c.1848del	p.(Glu616Aspfs*33)	13	c.1848del	p.(Glu616Aspfs*33)	Yes	15	9	9	-	-	-	This study
RP-0193	STGD1	45	c.6179T>G	p.(Leu2060Arg)	45	c.6179T>G	p.(Leu2060Arg)	Endogamy	30	11	30	-	-	-	Martin-Merida et al., 2019
RP-0266	CRD	IVS28	c.4253+5G>A	p.(Ile1377Hisfs*3)	46	c.6179T>G	p.(Leu2060Arg)	-	12	No	8	37y	Cone-rod pattern	CF/CF	Riveiro-Alvarez et al.2013
RP-0267	STGD1	36	c.5044_5058del	p.(Val1682_Val1686del)	36	c.5044_5058del	p.(Val1682_Val1686del)	Yes	6	No	-	30y	-	CF/<0.05	Riveiro-Alvarez et al.,2013
RP-0280	STGD1	38	c.5413A>G	p.(Asn1805Asp)	38	c.5413A>G	p.(Asn1805Asp)	Yes	-	-	-	26y	-	0.1/0.1	This study
RP-0298	CRD	8	c.950del	p.(Gly317Alafs*57)	33	c.4720G>T	p.(Glu1574*)	yes	10	10	8	36y	-	HM/HM	This study
RP-0532	CRD	23	c.3386G>T	p.(Arg1129Leu)	38	c.5318C>T	p.(Ala1773Val)	Yes	24	no	no	27y	-	0.2/0.2	This study
RP-0674	CRD	3	c.287del	p.(Asn96Thrfs*19)	30	c.4537dup	p.(Gln1513Profs*42)	Yes	9	9	4	-	-	-	This study
RP-0714	CRD	IVS28	c.4253+4C>T	p.(Ile1377Hisfs*3)	IVS28	c.4253+4C>T	p.(Ile1377Hisfs*3)	yes	10	10	30	-	-	-	Riveiro-Alvarez et al.,2013
RP-0741	CRD	43	c.5917del	p.(Val1973*)	43	c.5917del	p.(Val1973*)	Endogamy	8	8	8	27y	-	CF/CF	Riveiro-Alvarez et al.,2013
RP-0998	CRD	14	c.2041C>T	p.(Arg681*)	22	c.3210_3211dup	p.(Ser1071Cysfs*14)	Yes	9	32	No	35y	-	PL/0.07	Riveiro-Alvarez et al.,2013
RP-1102	CRD	15	c.2285C>A	p.(Ala762Glu)	15	c.2285C>A	p.(Ala762Glu)	Yes	8	-	8	27y	-	PL/PL	This study
RP-1112	CRD	1	c.1A>G	p.(Met1Val)	22	c.3210_3211dup	p.(Ser1071Cysfs*14)	-	7	7	7	9y	-	<0.1/<0.1	Martin-Merida et al., 2019
RP-1126	CRD	IVS26	c.3862+1G>A	p.(?)	46	c.6329G>A	p.(Trp2110*)	-	7	7	7	36y	-	CF/CF	Riveiro-Alvarez et al.,2013
RP-1367	CRD	IVS26	c.3862+1G>A	p.(?)	IVS26	c.3862+1G>A	p.(?)	Endogamy	13	NO	13	32y	-	0.05/0.05	This study
RP-1455	CRD	19	c.2888del	p.(Gly963Alafs*14)	48	c.6688del	p.(Leu2230Serfs*17)	-	4	-	17	-	-	-	This study
RP-1481	STGD1	21	c.3113C>T	p.(Ala1038Val)	23	c.3386G>T	p.(Arg1129Leu)	-	36	33	42	51y	-	0.12/0.12	Martin-Merida et al., 2019
RP-1539	CRD	22	c.3210_3211dup	p.(Ser1071Cysfs*14)	30	c.4417C>A	p.(Leu1437Met)	-	59	70	59	83y	-	0.2/0.4	Martin-Merida et al., 2019
RP-1543	CRD	28	c.4234C>T	p.(Gln1412*)	43	c.5917del	p.(Val1973*)	-	8	8	23	29y	-	HM/LP	Martin-Merida et al., 2019
RP-1578	CRD	22	c.3287C>T	p.(Ser1096Leu)	34	c.4793C>A	p.(Ala1598Asp)	-	9	No	9	12y	-	0.15/0.2	This study
RP-1603	CRD	16	c.2568C>A	p.(Tyr856*)	16	c.2568C>A	p.(Tyr856*)	-	6	8	10	6y	-	HM/HM	This study
RP-1669	CRD	22	c.3210_3211dup	p.(Ser1071Cysfs*14)	22	c.3210_3211dup	p.(Ser1071Cysfs*14)	Consanguinity	21	-	38	-	-	-	This study
RP-1680	CRD	13	c.1804C>T	p.(Arg602Trp)	36	c.5044_5058del	p.(Val1682_Val1686del)	-	10	44	11	12y	-	0.05/0.025	Riveiro-Alvarez et al.2013
RP-1715	CRD	13	c.1804C>T	p.(Arg602Trp)	13	c.1804C>T	p.(Arg602Trp)	Yes	-	-	-	-	-	-	Riveiro-Alvarez et al.2013
RP-1742	CRD	20	c.2980A>G	p.(Ile994Val)	38	c.5383T>G	p.(Leu1795Val)	-	15	8	15	-	-	-	Martin-Merida et al., 2019
RP-1769	STGD1	5	c.560G>A	p.(Arg187His)	22	c.3210_3211dup	p.(Ser1071Cysfs*14)	-	8	No	8	12y	-	0.16/0.25	Martin-Merida et al., 2019
RP-1819	CRD	22	c.3210_3211dup	p.(Ser1071Cysfs*14)	22	c.3210_3211dup	p.(Ser1071Cysfs*14)	Endogamy	14	-	14	58y	-	NPL/PL	This study
RP-1844	CRD	30	c.4537dup	p.(Gln1513Profs*42)	30	c.4537dup	p.(Gln1513Profs*42)	-	-	-	-	-	-	-	This study
RP-2028	STGD1	15	c.2300T>A	p.(Val767Asp)	IVS41	c.5836-1G>C	p.(?)	-	9	No	No	11y	-	0.1/0.05	Martin-Merida et al., 2019
RP-2104	CRD	13	c.1879C>T	p.(Glu627*)	41	c.5819T>C	p.(Leu1940Pro)	Yes	3	-	5	29y	-	PL/PL	Martin-Merida et al., 2019
RP-2133	CRD	6	c.634C>T	p.(Arg212Cys)	36	c.5044_5058del	p.(Val1682_Val1686del)	-	12	-	12	26y	-	0.5/0.05	This study
RP-2143	CRD	9	c.1222C>T	p.(Arg408*)	IVS38	c.5461-1G>T	p.(?)	-	6	-	6	17y	-	CF/CF	Martin-Merida et al., 2019
RP-2198	CRD	23	c.3386G>T	p.(Arg1129Leu)	23	c.3386G>T	p.(Arg1129Leu)	Endogamy	16	Yes	16	42y	-	<0.1/0.4	Martin-Merida et al., 2019
RP-2210	CRD	21	c.3056C>T	p.(Thr1019Met)	21	c.3056C>T	p.(Thr1019Met)	Consanguinity	4	-	4	65y	-	<0.1/<0.1	Martin-Merida et al., 2019
RP-2252	CRD	19	c.2878G>A	p.(Ala960Thr)	42	c.5882G>A	p.(Gly1961Glu)	-	30	65	70	89y	Scotopic and photopic extinguise	LP/0.05	This study
RP-2343	CRD	6	c.613T>G	p.(Cys205Gly)	41	c.5819T>C	p.(Leu1940Pro)	-	8	-	9	8y	-	0.05/0.3	Martin-Merida et al., 2019
RP-2386	CRD	13	c.1804C>T	p.(Arg602Trp)	13	c.1804C>T	p.(Arg602Trp)	-	8	No	-	-	-	-	Martin-Merida et al., 2019
RP-2419	STGD1	11	c.1364T>A	p.(Leu455Gln)	34	c.4793C>A	p.(Ala1598Asp)	Yes	34	34	28	-	-	-	Martin-Merida et al., 2019
RP-2492	STGD1	22	c.3210_3211dup	p.(Ser1071Cysfs*14)	33 VUS	c.4672G>A	p.(Gly1558Arg)	-	8	-	12	11y	-	0.1/0.1	Martin-Merida et al., 2019
RP-2500	STGD1	6	c.671del	p.(Thr224Argfs*17)	42	c.5882G>A	p.(Gly1961Glu)	-	-	No	No	41y	-	0.1/0.1	This study
RP-2520	STGD1	IVS28	c.4253+5G>A	p.(Ile1377Hisfs*3)	IVS28	c.4253+5G>A	p.(Ile1377Hisfs*3)	Yes	13	-	13	-	-	-	This study
RP-2531	STGD1	Ex6-IVS6	c.699_768+341del	p.(Gln234Phefs*5)	Ex6-IVS6	c.699_768+341del	p.(Gln234Phefs*5)	Endogamy	10	10	No	21y	-	0.02/0.02	This study
RP-2668	CRD	IVS6-Ex6	c.699_768+341del	p.(Gln234Phefs*5)	30	c.4457C>T	p.(Pro1486Leu)	-	22	22	22	33y	Cone-rod pattern	0.4/0.6	This study
RP-2680	CRD	19	c.2888del	p.(Gly963Alafs*14)	30	c.4457C>T	p.(Pro1486Leu)	Yes	50	50	50	-	-	-	This study
RP-2711	STGD1	IVS20	c.3050+5G>A	p.(Leu973_His1017delinsPhe)#	23	c.3386G>T	p.(Arg1129Leu)	-	12	No	No	-	-	-	Martin-Merida et al., 2019
RP-2867	STGD1	23	c.3386G>T	p.(Arg1129Leu)	IVS28	c.4253+5G>A	p.(Ile1377Hisfs*3)	-	-	-	-	-	-	-	This study

Table S1. Genetic and clinical information of 506 Spanish families with *ABCA4* mutations

| Family ID | Phenotype | *ABCA4* variants | | | | | | | Symptoms onset (age in years) | | | Age at ophthalmological examination | ERG | BCVA (OD/OS) | Reference |
		Allele1_Exon	Allele1_cDNA	Allele 1_Protein	Allele2_Exon	Allele2_cDNA	Allele2_Protein	Segregation	VA loss	VF loss	NB				
RP-2948	CRD	23	c.3386G>T	p.(Arg1129Leu)	41	c.5819T>C	p.(Leu1940Pro)	-	20	20	20	-	-	-	This study
RP-2956	CRD	28	c.4253+5G>A	p.(Ile1377Hisfs*3)	28	c.4253+5G>A	p.(Ile1377Hisfs*3)	Consanguinity	11	20	16	-	-	-	This study
RP-3080	STGD1	45	c.6179T>G	p.(Leu2060Arg)	48	c.6647C>T	p.(Ala2216Val)	-	62	-	-	-	-	-	This study
RP-3093	CRD	33	c.4773+1G>T	p.(?)	33	c.4773+1G>T	p.(?)	-	-	-	-	-	-	-	This study
RP-3118	STGD1	23	c.3386G>T	p.(Arg1129Leu)	25	c.4918C>T	p.(Arg1640Trp)	-	15	15	No	-	-	-	This study
RP-3136	STGD1	23	c.3420C>G	p.(Cys1140Trp)	23	c.3386G>T	p.(Arg1129Leu)	-	-	-	-	-	-	-	This study

Table S2. Total of *ABCA4* variants identified in 506 Spanish families. Abbreviations: IVS, intron; CNV, copy number variant. #Protein effect based on functional assays reported by Braun (2013) Sangermano (2018) Sangermano (2019) Bauwens (2019) Fadaie (2019) and Khan (2019). $Variants only found in combination with other variants in cis.

ABCA4 variants (NM_000350)						
Exon/Intron	Nucleotide	Protein	Reference	Number of individual alleles	Number of complex alleles	Type of variant
1	c.1A>G	p.(Met1Val)	Briggs (2001) Invest Ophthalmol Vis Sci 42, 2229	1	1	missense
1	c.32T>C	p.(Leu11Pro)	Rozet (1998) Eur J Hum Genet 6, 291	2		missense
1	c.3G>A	p.(Met1Ile)	Riveiro-Alvarez (2013) Ophthalmology 120, 2332	1		missense
1	c.52C>T	p.(Arg18Trp)	Gerber (1998) Genomics 48, 139	2		missense
3	c.184C>G	p.(Pro62Ala)	Novel	1		missense
3	c.214G>A	p.(Gly72Arg)	Rivera (2000) Am J Hum Genet 67, 800	1		missense
3	c.223T>G	p.(Cys75Gly)	Lewis (1999) Am J Hum Genet 64, 422	2		missense
3	c.286A>G	p.(Asn96Asp)	Papaioannou (2000) Invest Ophthalmol Vis Sci 41, 16	1		missense
3	c.287del	p.(Asn96Thrfs*19)	Corton (2013) PLoS One 8, e65574	3		frameshift
4	c.378G>A	p.(Trp126*)	Novel	1		stop gained
4	c.393del	p.(Leu132Cysfs*22)	Novel	1		frameshift
4	c.428C>T	p.(Pro143Leu)	Jaakson (2003) Hum Mutat 22, 395	1		missense
5	c.454C>T	p.(Arg152*)	Souied (1999) Invest Ophthalmol Vis Sci 40, 2740	2		stop gained
5	c.457A>T	p.(Ile153Leu)	Riveiro-Alvarez (2013) Ophthalmology 120, 2332	1		missense
5	c.560G>A	p.(Arg187His)	Aguirre-Lamban (2009) Hum Genet 126 330	2	1	missense
6	c.611C>A	p.(Ala204Asp)	Novel	1		missense
6	c.613T>G	p.(Cys205Gly)	Novel	1		missense
6	c.634C>T	p.(Arg212Cys)	Gerber (1998) Genomics 48, 139	13	1	missense
6	c.671del	p.(Thr224Argfs*17)	Stenirri (2007) Eur J Ophthalmol 17, 749	2		frameshift
6	c.699_768+341del	p.(Gln234Phefs*5)	Novel	7		frameshift CNV
6	c.700C>T	p.(Gln234*)	Aguirre-Lamban (2007) Hum Genet 121, 648	1		stop gained
6	c.735T>G	p.(Tyr245*)	Stenirri (2004) Clin Chem 50, 1336	1		stop gained
6	c.742_768+29del	p.(Val248_Val256del)	Riveiro-Alvarez (2013) Ophthalmology 120, 2332	1		inframe deletion
6	c.768G>T	p.(Val256Val)	Maugeri (1999) Am J Hum Genet 64, 1024	1		missense
IVS6	c.768+2T>G	p.(?)	Novel	1		splice donor
IVS7	c.859-506G>C	p.[Phe287Thrfs*32,=]#	Sangermano (2019) Genet Med	1		deep intronic
8	c.871C>G	p.(Pro291Ala)	Riveiro-Alvarez (2013) Ophthalmology 120, 2332	3		missense
8	c.874A>C	p.(Ser292Arg)	Novel	1		missense
8	c.950del	p.(Gly317Alafs*57)	Corton (2013) PLoS One 8, e65574	3		frameshift
8	c.982G>T	p.(Glu328*)	Pasadhika (2009) Am J Ophthalmol 148, 260	7		stop gained
8	c.1022A>T	p.(Glu341Val)	Novel	1		missense
8	c.1025_1038del	p.(Asp342Glyfs*6)	Yatsenko (2001) Hum Genet 108, 346	4		frameshift
8	c.1029dup	p.(Asn344*)	Aguirre (2007) Hum Genet 122, 548	1		stop gained
8	c.1035T>G	p.(Tyr345*)	Zaneveld (2015) Genet Med 17, 262	1		stop gained
9	c.1222C>T	p.(Arg408*)	Birch (2001) Exp Eye Res 73, 877	5		stop gained
11	c.1357_1554del	p.(Asp453_Glu518del)	Novel	1		inframe deletion CNV
11	c.1364T>A	p.(Leu455Gln)	Salles (2018) Mol Vis 24, 546	2		missense
12	c.1592A>G	p.(Glu531Gly)	Aguirre-Lamban (2010) Hum Genet 127 119	1		missense
12	c.1609C>T	p.(Arg537Cys)	Jaakson (2003) Hum Mutat 22, 395	2		missense
12	c.1622T>C	p.(Leu541Pro)	Rozet (1998) Eur J Hum Genet 6, 291	3	3	missense
12	c.1633A>T	p.(Asn545Tyr)	Novel	1		missense
12	c.1648G>A	p.(Gly550Arg)	Shroyer (2001) Hum Mol Genet 10, 2671	1		missense
12	c.1667T>G	p.(Met556Arg)	Khan (2020) Genet Med	2		missense
12	c.1714C>T	p.(Arg572*)	Stenirri (2008) Clin Chem Lab Med 46, 1250	1		stop gained
12	c.1715G>C	p.(Arg572Pro)	Lewis (1999) Am J Hum Genet 64, 422	2		missense
12	c.1751_1753delinsAT	p.(Ile584Asnfs*65)	Novel	3		frameshift
13	c.1755del	p.(Lys585Lysfs*63)	Novel	1		frameshift
13	c.1766G>A	p.(Trp589*)	Riveiro-Alvarez (2013) Ophthalmology 120, 2332	2		stop gained
13	c.1766G>C	p.(Trp589Ser)	Khan (2020) Genet Med	1		missense
13	c.1792G>A	p.(Val598Met)	Birtel (2018) Sci Rep 8,	1		missense
13	c.1804C>T	p.(Arg602Trp)	Lewis (1999) Am J Hum Genet 64, 422	30		missense
13	c.1819G>A	p.(Gly607Arg)	Rivera (2000) Am J Hum Genet 67, 800	1	1	missense
13	c.1819G>C	p.(Gly607Arg)	Bravo-Gil (2016) Sci Rep 6, 23910	3		missense
13	c.1832T>C	p.(Leu611Pro)	Riveiro-Alvarez (2013) Ophthalmology 120, 2332	2		missense
13	c.1848del	p.(Glu616Aspfs*33)	Martinez-Mir (1998) Nat Genet 18, 11	2		frameshift
13	c.1853G>T	p.(Gly618Val)	Novel	1		missense
13	c.1868A>G	p.(Gln623Arg)	Zernant (2011) Invest Ophthalmol Vis Sci 52, 8479	1		missense
13	c.1879G>T	p.(Glu627*)	Novel	2		stop gained
13	c.1927G>A	p.(Ala643Met)	Briggs (2001) Invest Ophthalmol Vis Sci 42, 2229	1		missense
13	c.1933G>A	p.(Asp645Asn)	Lewis (1999) Am J Hum Genet 64, 422	1		missense
14	c.1957C>T	p.(Arg653Cys)	Rivera (2000) Am J Hum Genet 67, 800	2		missense
14	c.1964T>G	p.(Phe655Cys)	Downs (2007) Arch Ophthalmol 125, 252	1		missense
14	c.2023G>A	p.(Val675Ile)	Fujinami (2013) Am J Ophthalmol 155, 1075	4		missense
14	c.2041C>T	p.(Arg681*)	Maugeri (1999) Am J Hum Genet 64, 1024	8		stop gained
14	c.2057T>C	p.(Leu686Ser)	Paloma (2001) Hum Mutat 17, 504	2		missense
14	c.2099G>A	p.(Trp700*)	Fumagalli (2001) Hum Genet 109, 326	1		stop gained
IVS14	c.2161-8G>A	p.(His721_Val794del)#	Khan (2020) Genet Med	1		non canonical splice site
15	c.2285C>A	p.(Ala762Glu)	Aguirre-Lamban (2008) Hum Genet 123 546	7		missense
15	c.2300T>A	p.(Val767Asp)	Simonelli (2000) Invest Ophthalmol Vis Sci 41, 892	3		missense
IVS15	c.2382+5G>C	p.[=,His721_Val794del]#	Khan (2020) Genet Med	1		non canonical splice site
16	c.2401G>A	p.(Ala801Thr)	Downs (2007) Arch Ophthalmol 125: 252	1		missense
16	c.2481del	p.(Thr829Argfs*14)	Novel	2	1	frameshift
16	c.2483C>T $	p.(Pro828Leu)	Novel		1	missense
16	c.2488G>T	p.(Glu830*)	Novel	1		stop gained
16	c.2546T>C $	p.(Val849Ala)	Webster (2001) Invest Ophthalmol Vis Sci 42, 1179		1	missense
16	c.2568C>A	p.(Tyr856*)	Fujinami (2013) Invest Ophthalmol Vis Sci 54, 6662	2		stop gained
17	c.2588G>C	p.[Gly863Ala,Gly863del]	Gerth (2002) Graefes Arch Clin Exp Ophthalmol 240, 628	6	2	missense
17	c.2588G>T	p.(Gly863Ala)	Novel	1		missense
17	c.2613G>A	p.(Trp871*)	Jespersgaard (2019) Sci Rep 9	1		stop gained
19	c.2791G>A	p.(Val931Met)	Allikmets (1997) Nat Genet 15, 236	3		missense

Exon/Intron	Nucleotide	Protein	Reference	Number of individual alleles	Number of complex alleles	Type of variant
			***ABCA4* variants (NM_000350)**			
19	c.2878G>A	p.(Ala960Thr)	Novel	1		missense
19	c.2888del	p.(Gly963Alafs*14)	Paloma (2001) Hum Mutat 17, 504	25		frameshift
19	c.2894A>G	p.(Asn965Ser)	Allikmets (1997) Nat Genet 15, 236	3		missense
IVS19	c.2919-826T>A	p.(Leu973Phefs*1)#	Fadaie (2019) Hum Mutat	1		deep intronic
20	c.2966T>C	p.(Val989Ala)	Briggs (2001) Invest Ophthalmol Vis Sci 42, 2229	1		missense
20	c.2971G>C $	p.(Gly991Arg)	Jaakson (2003) Hum Mutat 22, 395		1	missense
20	c.2980A>G	p.(Ile994Val)	Novel	1		missense
IVS20	c.3050+5G>A	p.(Leu973_His1017delinsPhe)#	Rivera (2000) Am J Hum Genet 67, 800	2		non canonical splice site
21	c.3056C>T	p.(Thr1019Met)	Rozet (1998) Eur J Hum Genet 6, 291	12		missense
21	c.3163C>T $	p.(Arg1055Trp)	Paloma (2001) Hum Mutat 17, 504		1	missense
21	c.3113C>T	p.(Ala1038Val)	Allikmets (1997) Nat Genet 15, 236	5	3	missense
22	c.3210_3211dup	p.(Ser1071Cysfs*14)	Allikmets (1997) Nat Genet 15, 236	33	1	frameshift
22	c.3251T>C	p.(Ille1084Thr)	Novel	1		missense
22	c.3259G>A	p.(Glu1087Lys)	Allikmets (1997) Nat Genet 15, 236	4		missense
22	c.3277G>A	p.(Asp1093Asn)	Novel	2		missense
22	c.3281C>G	p.(Pro1094Arg)	Riveiro-Alvarez (2013) Ophthalmology 120, 2332	1		missense
22	c.3287C>T	p.(Ser1096Leu)	Riveiro-Alvarez (2013) Ophthalmology 120, 2332	2		missense
22	c.3292C>T	p.(Arg1098Cys)	Rivera (2000) Am J Hum Genet 67, 800	5		missense
22	c.3299T>C	p.(Ile1100Thr)	Novel	2		missense
22	c.3322C>T	p.(Arg1108Cys)	Rozet (1998) Eur J Hum Genet 6: 291	1	14	missense
22	c.3323G>A	p.(Arg1108His)	Webster (2001) Invest Ophthalmol Vis Sci 42, 1179	4		missense
IVS22	c.3329-2A>T	p.(?)	Riveiro-Alvarez (2006) Hum Genet 118 784	1		splice acceptor
23	c.3364G>A	p.(Glu1122Lys)	Lewis (1999) Am J Hum Genet 64, 422	7		missense
23	c.3380G>A	p.(Gly1127Glu)	Duno (2012) Ophthalmic Genet 33, 225	1		missense
23	c.3383A>G	p.(Asp1128Gly)	Novel	7		missense
23	c.3386G>T	p.(Arg1129Leu)	Allikmets (1997) Science 277, 1805	183	7	missense
23	c.3409A>G	p.(Arg1137Gly)	Nassisi (2018) Int J Mol Sci 19,	1		missense
23	c.3420C>G	p.(Cys1140Trp)	Zhang (2014) PLoS One 9, 95528	1		missense
25	c.3758C>T $	p.(Thr1253Met)	Paloma (2001) Hum Mutat 17, 504		1	missense
25	c.4918C>T	p.(Arg1640Trp)	Rozet (1998) Eur J Hum Genet 6, 291	11	3	missense
26	c.3832G>T	p.(Glu1278*)	Novel	1		stop gained
IVS26	c.3862+1G>A	p.(?)	Briggs (2001) Invest Ophthalmol Vis Sci 42, 2229	4		splice donor
27	c.3871C>T	p.(Gln1291*)	Zaneveld (2015) Genet Med 17, 262	1		stop gained
27	c.3874C>T	p.(Gln1292*)	Ernest (2009) Mol Vis 15, 2841	1		stop gained
27	c.3881_3885del	p.(Arg1294Lysfs*126)	Novel	1		frameshift
27	c.3898C>T	p.(Arg1300*)	Rivera (2000) Am J Hum Genet 67, 800	1		stop gained
27	c.3899G>A	p.(Arg1300Gln)	Briggs (2001) Invest Ophthalmol Vis Sci 42, 2229	1	1	missense
27	c.3943C>T	p.(Gln1315*)	Riveiro-Alvarez (2006) Hum Genet 118 777	5		stop gained
27	c.3988G>T	p.(Glu1330*)	Riveiro-Alvarez (2013) Ophthalmology 120, 2332	4		stop gained
27	c.4003_4004del	p.(Pro1335Argfs*86)	Riveiro-Alvarez (2013) Ophthalmology 120,2332	1		frameshift
27	c.4069G>A	p.(Ala1357Thr)	Duno (2012) Ophthalmic Genet 33, 225	2		missense
27	c.4070C>T	p.(Ala1357Val)	Nõupuu (2016) Graefes Arch Clin Exp Ophthalmol 254, 865	1		missense
28	c.4139C>T	p.(Pro1380Leu)	Lewis (1999) Am J Hum Genet 64, 422	6		missense
28	c.4200C>A	p.(Tyr1400*)	Maugeri (1999) Am J Hum Genet 64, 1024	1		stop gained
28	c.4222T>C $	p.(Trp1408Arg)	Lewis (1999) Am J Hum Genet 64, 422		3	missense
28	c.4234C>T	p.(Gln1412*)	Maugeri (1999) Am J Hum Genet 64, 1024	2		stop gained
IVS28	c.4253+43G>A	p.[=,Ile1377Hisfs*3]#	Zernant (2018) Cold Spring Harb Mol Case Stud	5	1	deep intronic
IVS28	c.4253+4C>T	p.(Ile1377Hisfs*3)#	Ozgul (2004) Hum Mutat 23, 523	10		non canonical splice site
29	c.4254C>A	p.(Ser1418Arg)	Novel	2		missense
29	c.4283C>T $	p.(Thr1428Met)	Allikmets (1997) Science 277, 1805		1	missense
29	c.4297G>A	p.(Val1433Ile)	Lewis (1999) Am J Hum Genet 64, 422	1		missense
29	c.4322G>A	p.(Gly1441Asp)	Novel	1		missense
IVS29	c.4353-1G>A	p.(?)	Novel	1		splice acceptor
30	c.4417C>A	p.(Leu1437Met)	Stenirri (2008) Clin Chem Lab Med 46, 1250	1		missense
30	c.4436G>A	p.(Trp1479*)	Jaakson (2003) Hum Mutat 22, 395	1		stop gained
30	c.4457C>T	p.(Pro1486Leu)	Lewis (1999) Am J Hum Genet 64, 422	26		missense
30	c.4469G>A	p.(Cys1490Tyr)	Lewis (1999) Am J Hum Genet 64, 422	2		missense
30	c.4519G>A	p.(Gly1507Arg)	Fujinami (2013) Am J Ophthalmol 155, 1075	1		missense
30	c.4537del	p.(Gln1513Argfs*13)	Aguirre (2008) Hum Genet 123 544	1		frameshift
30	c.4537dup	p.(Gln1513Profs*42)	Briggs (2001) Invest Ophthalmol Vis Sci 42, 2229	8		frameshift
IVS30	c.4539+2064C>T	p.[=,Arg1514Leufs*36]#	Zernant (2014) Hum Mol Genet 23, 6797	15		deep intronic
IVS30	c.4540-8T>C $	p.(Gln1513insProGln)#	Khan (2020) Genet Med		1	non canonical splice site
31	c.4576A>G	p.(Thr1526Ala)	Novel	1		missense
31	c.4577C>T	p.(Thr1526Met)	Lewis (1999) Am J Hum Genet 64, 422	1		missense
IVS32	c.4668-1G>A	p.(?)	Riveiro-Alvarez (2013) Ophthalmology 120, 2332	1		splice acceptor
33	c.4672G>A	p.(Gly1558Arg)	Novel	2		missense
33	c.4720G>T	p.(Glu1574*)	Maia-Lopes (2009) Mol Vis 15, 584	1		stop gained
33	c.4739del	p.(Leu1580*)	Riveiro-Alvarez (2006) Hum Genet 119 671	1		stop gained
33	c.4773+1G>T	p.(?)	Pang (2002) Hum Mutat 19, 189	3		splice donor
IVS33	c.4773+1del	p.(?)	Novel	1		splice donor
34	c.4793C>A	p.(Ala1598Asp)	Maugeri (2000) Am J Hum Genet 67, 960	6		missense
IVS34	c.4848+3A>G	p.[Gly1592_Lys1616del,=]#	Khan (2020) Genet Med	2		non canonical splice site
35	c.4849del	p.(Val1617Cysfs*45)	Novel	2		frameshift
35	c.4855T>C	p.(Phe1619Leu)	Aguirre-Lamban (2008) Hum Genet 124 314	1		missense
35	c.4919G>A	p.(Arg1640Gln)	Simonelli (2000) Invest Ophthalmol Vis Sci 41, 892	5		missense
35	c.4926C>A	p.(Ser1642Arg)	Birch (2001) Exp Eye Res 73, 877	1	6	missense
IVS35	c.5018+2T>C	p.(?)	Cideciyan (2009) Hum Mol Genet 18, 931	1		splice donor
36	c.5044_5058del	p.(Val1682_Val1686del)	Allikmets (1997) Nat Genet 15, 236	13	6	inframe deletion
36	c.5172G>T	p.(Trp1724Cys)	Stenirri (2008) Clin Chem Lab Med 46, 1250	1		missense
36	c.5196+1G>A	p.(?)	Stone (2017) Ophthalmology 124, 1314	1		splice donor
IVS36	c.5196+1056A>G	p.(Met1733Valfs*2)#	Braun (2013) Hum Mol Genet 22, 5136	1		deep intronic

Table S2. Total of *ABCA4* variants identified in 506 Spanish families

			ABCA4 variants (NM_000350)			
Exon/Intron	Nucleotide	Protein	Reference	Number of individual alleles	Number of complex alleles	Type of variant
IVS36	c.5196+1137G>A	p.[Met1733Glufs*78,=]#	Braun (2013) Hum Mol Genet 22, 5136	4		deep intronic
37	c.5242G>A	p.(Gly1748Arg)	Paloma (2001) Hum Mutat 17, 504	1		missense
38	c.5317_5318insA	p.(Ala1773Aspfs*14)	Paloma (2001) Hum Mutat 17, 504	1		frameshift
38	c.5318C>T	p.(Ala1773Val)	Webster (2001) Invest Ophthalmol Vis Sci 42, 1179	6		missense
38	c.5377G>A	p.(Val1793Met)	Stone (2017) Ophthalmology 124, 1314	1		missense
38	c.5383T>G	p.(Leu1795Val)	Novel	1		missense
38	c.5384T>C	p.(Leu1795Ser)	Nassisi (2018) Int J Mol Sci 19	2		missense
38	c.5395A>G	p.(Asn1799Asp)	Paloma (2001) Hum Mutat 17, 504	3		missense
38	c.5413A>G	p.(Asn1805Asp)	Paloma (2001) Hum Mutat 17, 504	2		missense
38	c.5451G>T	p.(Glu1817Asp)	Webster (2001) Invest Ophthalmol Vis Sci 42, 1179	2		missense
IVS38	c.4253+5G>A	p.(Ile1377Hisfs*3)#	Rivera (2000) Am J Hum Genet 67: 800	10		non canonical splice site
IVS38	c.5460+5G>A	p.(Trp1772Argfs*9)#	Aguirre-Lamban (2008) Hum Genet 123 547	1		non canonical splice site
IVS38	c.5461-10T>C	p.(Thr1821Aspfs*6)#	Klevering (2005) Graefes Arch Clin Exp Ophthalmol 243, 90	11	2	non canonical splice site
IVS38	c.5461-1G>T	p.(?)	Riera (2017) Sci Rep 7, 42078	2		splice acceptor
39	c.5512C>G	p.(His1838Asp)	Riveiro-Alvarez (2013) Ophthalmology 120, 2332	1	2	missense
39	c.5531_5557dup	p.(Gly1844_Gln1852dup)	Novel	1		inframe insertion
39	c.5531G>A	p.(Gly1844Asp)	Riveiro-Alvarez (2013) Ophthalmology 120, 2332	1		missense
39	c.5549T>C	p.(Leu1850Pro)	Aguirre-Lamban (2010) Hum Genet 127 119	7		missense
40	c.5603A>T	p.(Asn1868Ile)	Webster (2001) Invest Ophthalmol Vis Sci 42: 1179	7	2	missense
40	c.5630_5644dup	p.(Lys1877_Ala1881dup)	Aguirre-Lamban (2009) Br J Ophthalmol 93, 614	2		inframe insertion
40	c.5644A>G	p.(Met1882Val)	Fukui (2002) Invest Ophthalmol Vis Sci 43, 2819	3		missense
40	c.5655del	p.(Val1887Trpfs*6)	Novel	1		frameshift
ISV40	c.5714+1G>A	p.(?)	Novel	1		splice donor
IVS40	c.5714+5G>A	p.[=,Glu1863Leufs*33]#	Cremers (1998) Hum Mol Genet 7, 355	14		non canonical splice site
41	c.5761G>A	p.(Val1921Met)	Jaakson (2003) Hum Mutat 22, 395	1		missense
41	c.5819T>C	p.(Leu1940Pro)	Paloma (2001) Hum Mutat 17, 504	29		missense
IVS41	c.5836-1G>C	p.(?)	Novel	2		splice acceptor
42	c.5843C>T $	p.(Pro1948Leu)	Novel		1	missense
42	c.5858del	p.(Pro1953Glnfs*21)	Novel	1		frameshift
42	c.5881G>A	p.(Gly1961Arg)	Riveiro-Alvarez (2006) Hum Genet 118 774	1		missense
42	c.5882G>A	p.(Gly1961Glu)	Allikmets (1997) Science 277, 1805	60	4	missense
43	c.5899T>G	p.(Cys1967Gly)	Novel	1		missense
43	c.5914G>A	p.(Gly1972Arg)	Jaakson (2003) Hum Mutat 22, 395	2		missense
43	c.5917del	p.(Val1973*)	Rivera (2000) Am J Hum Genet 67, 800	9		stop gained
43	c.5929G>A	p.(Gly1977Ser)	Rozet (1998) Eur J Hum Genet 6, 291	18		missense
43	c.5981G>A	p.(Gly1994Glu)	Novel	1		missense
44	c.6071A>G $	p.(Asp2024Gly)	Novel		1	missense
44	c.6079C>T	p.(Leu2027Phe)	Allikmets (1997) Nat Genet 15, 236	1		missense
44	c.6088C>T	p.(Arg2030*)	Lewis (1999) Am J Hum Genet 64, 422	4		stop gained
44	c.6089G>A	p.(Arg2030Gln)	Lewis (1999) Am J Hum Genet 64, 422	6	1	missense
44	c.6118C>T	p.(Arg2040*)	Baum (2003) Ophthalmologica 217, 111	1		stop gained
44	c.6140T>A	p.(Ile2047Asn)	Aguirre-Lamban (2008) Hum Genet 123 544	1		missense
44	c.6147_6147+7del	p.(Val2050Leufs*11)	Novel	1		frameshift
IVS44	c.6147+2T>A	p.(?)	Valverde (2007) Invest Ophthalmol Vis Sci 48, 985	3		splice donor
45	c.6179T>G	p.(Leu2060Arg)	Paloma (2001) Hum Mutat 17, 504	23	1	missense
45	c.6216T>A	p.(Ser2072Arg)	Novel	1		missense
45	c.6221G>A	p.(Gly2074Asp)	Consugar (2015) Genet Med 17, 253	1		missense
45	c.6229C>T	p.(Arg2077Trp)	Allikmets (1997) Nat Genet 15, 236	3		missense
45	c.6230G>A	p.(Arg2077Gln)	Riveiro-Alvarez (2013) Ophthalmology 120, 2332	1		missense
45	c.6272T>A	p.(Leu2091Gln)	Novel	1		missense
46	c.6316C>T	p.(Arg2106Cys)	Allikmets (1997) Nat Genet 15, 236	1		missense
46	c.6320G>A	p.(Arg2107His)	Rozet (1998) Eur J Hum Genet 6: 291	3	13	missense
46	c.6320G>C	p.(Arg2107Pro)	Riveiro-Alvarez (2009) Br J Ophthalmol 93, 1359	4		missense
46	c.6329G>A	p.(Trp2110*)	Maugeri (2000) Am J Hum Genet 67, 960	3		stop gained
46	c.6380C>T	p.(Ser2127Phe)	Khan (2020) Genet Med	1		missense
46	c.6310C>T	p.(Gln2104*)	Novel	1		stop gained
IVS46	c.6387-1G>A	p.(?)	Novel	1		splice acceptor
47	c.6410G>A	p.(Cys2137Tyr)	Aguirre-Lamban (2008) Hum Genet 123 547	8		missense
47	c.6415C>T	p.(Arg2139Trp)	Lewis (1999) Am J Hum Genet 64, 422	1		missense
47	c.6437G>T	p.(Gly2146Val)	Novel	1		missense
47	c.6449G>A	p.(Cys2150Tyr)	Fishman (1999) Arch Ophthalmol 117, 504	6		missense
IVS47	c.6480-1G>T	p.(?)	Jiang (2016) Invest Ophthalmol Vis Sci 57, 145	1		splice acceptor
48	c.6559C>T	p.(Gln2187*)	Riveiro-Alvarez (2006) Hum Genet 118 774	3		stop gained
48	c.6563T>C	p.(Phe2188Ser)	Fukui (2002) Invest Ophthalmol Vis Sci 43, 2819	1		missense
48	c.6647C>T	p.(Ala2216Val)	Schulz (2017) Invest Ophthalmol Vis Sci 58, 394	1		missense
48	c.6688del	p.(Leu2230Serfs*17)	Martin-Merida (2018) Invest Ophthalmol Vis Sci 59, 2345	1		frameshift
48	c.6718A>G	p.(Thr2240Ala)	Downs (2007) Arch Ophthalmol 125: 252	7	6	missense

Table S3. Variants identified in 21 complex alleles.
#In these cases cosegregation of variants was not possible

ABCA4 complex alleles (NM_000350)		
Nucleotide	**Protein**	**Number of alleles**
c.[3322C>T;6320G>A]	p.[Arg1108Cys;Arg2107His]	13
c.[3386G>T;6718A>G]	p.[Arg1129Leu;Thr2240Ala]	6
c.[4926C>G;5044_5058del]	p.[Ser1642Arg; Val1681_Cys1685del]	6
c.[1622T>C;3113C>T]	p.[Leu541Pro; Ala1038Val]	3
c.[4222T>C;4918C>T]	p.[Trp1408Arg;Arg1640Trp]	3
c.[5512C>G;5882G>A]	p.[His1838Asp;Gly1961Glu]	2
c.[560G>A;3210_3211dup]#	p.[Arg187His;Ser1071Cysfs*14]	1
c.[2588G>C;3163C>T]	p.[Gly863Ala;Arg1055Trp]	1
c.[1A>G;6089G>A]	p.[Met1Val;Arg2030Gln]	1
c.[3758C>T;5882G>A]	p.[Thr1253Met; Gly1961Glu]	1
c.[3322C>T;6071A>G]	p.[Arg1108His;Asp2024Gly]	1
c.[2971G>C;3899G>A]#	p.[Gly991Arg;Arg1300Gln]	1
c.[2483C>T;2481del]	p.[Pro828Leu;Thr829Argfs*14]	1
c.[2588G>C;5461-10T>C]	p.[Gly863Ala, Gly863del;Thr1821Aspfs*6)]	1
c.[1819G>A;4283C>T]#	p.[Gly607Arg;Thr1428Met]	1
c.[634C>T;2546T>C]#	p.[Arg212Cys;Val849Ala]	1
c.[4540-8T>C;5882G>A]	p.[Gln1513insProGln;Gly1961Glu]	1
c.[3386G>T;4537dup]	p.[Arg1129Leu;Gln1513Profs*42]	1
c.[5461-10T>C;5603A>T]#	p.[Thr1821Aspfs*6,Thr1821Valfs*13;Asn1868Ile]	1
c.[4253+43G>A;5603A>T]#	p.[=,Ile1377Hisfs*3](;)(Asn1868Ile)	1
c.[5843C>T;6179T>G]#	p.[Pro1948Leu;Leu2060Arg]	1

Table S4. In-silico predictions of novel *ABCA4* variants identified in 506 Spanish families

ABCA4 variants (NM_000350)
Exon/Intron	**Nucleotide**	**Protein**	**Number of alleles**	**Type of variant**	**MAF (gnomAD)**	**SIFT**	**PolyPhen**	**CADD**	**M-CAP**
3	c.184C>G	p.(Pro62Ala)	1	missense	-	deleterious	possibly damaging	25.7	damaging
4	c.378G>A	p.(Trp126*)	1	stop gained	-	-	-	36	-
4	c.393delC	p.(Leu132Cysfs*22)	1	frameshift	-	-	-	-	-
6	c.611C>A	p.(Ala204Asp)	1	missense	-	tolerated	benign	25.8	damaging
6	c.613T>G	p.(Cys205Gly)	1	missense	-	deleterious	probably damaging	25.7	damaging
6	c.699_768+341del	p.(Gln234Phefs*5)	7	frameshift	-	-	-	-	-
IVS6	c.768+2T>G	p.(?)	1	splice_donor	-	-	-	23.5	-
8	c.1022A>T	p.(Glu341Val)	1	missense	-	deleterious	possibly damaging	31	damaging
8	c.874A>C	p.(Ser292Arg)	1	missense	-	deleterious	possibly damaging	24.5	damaging
11	c.1357_1554del	p.(Asp453_Glu518del)	1	inframe_deletion	-	-	-	-	-
12	c.1633A>T	p.(Asn545Tyr)	1	missense	-	deleterious	benign	26.6	damaging
12	c.1751_1753delinsAT	p.(Ile584Asnfs*65)	3	frameshift	-	-	-	35	-
12	c.1755del	p.(Lys585Lysfs*63)	1	frameshift	-	-	-	-	-
13	c.1853G>T	p.(Gly618Val)	1	missense	-	deleterious	benign	33	damaging
13	c.1879G>T	p.(Glu627*)	2	stop gained	-	-	-	35	-
16	c.2481del	p.(Thr829Argfs*14)	3	frameshift	-	-	-	-	-
16	c.2483C>T	p.(Pro828Leu)	1	missense	-	deleterious	benign	24.2	damaging
16	c.2488G>T	p.(Glu830*)	1	stop gained	-	-	-	42	-
17	c.2588G>T	p.(Gly863Ala)	1	missense	-	deleterious	probably damaging	31	-
19	c.2878G>A	p.(Ala960Thr)	1	missense	1.60E-03	deleterious	possibly damaging	23.4	damaging
20	c.2980A>G	p.(Ile994Val)	1	missense	1.19E-05	tolerated	possibly damaging	23.6	damaging
22	c.3251T>C	p.(Ille1084Thr)	1	missense	-	deleterious	benign	26.7	damaging
22	c.3277G>A	p.(Asp1093Asn)	2	missense	-	deleterious	probably damaging	26.7	damaging
22	c.3299T>C	p.(Ile1100Thr)	2	missense	-	deleterious	possibly damaging	26.1	damaging
23	c.3383A>G	p.(Asp1128Gly)	7	missense	-	deleterious	possibly damaging	27.6	damaging
26	c.3832G>T	p.(Glu1278*)	1	stop gained	-	-	-	50	-
27	c.3881_3885del	p.(Arg1294Lysfs*126)	1	frameshift	-	-	-	-	-
29	c.4254C>A	p.(Ser1418Arg)	2	missense	-	deleterious	probably damaging	26.8	-
29	c.4322G>A	p.(Gly1441Asp)	1	missense	-	deleterious	probably damaging	33	damaging
IVS29	c.4353-1G>A	p.(?)	1	splice_acceptor	-	-	-	22.8	-
31	c.4576A>G	p.(Thr1526Ala)	1	missense	-	deleterious	benign	23.1	damaging
33	c.4672G>A	p.(Gly1558Arg)	2	missense	-	deleterious	probably damaging	34	damaging
IVS33	c.4773+1del	p.(?)	1	splice_donor	-	-	-	-	-
35	c.4849del	p.(Val1617Cysfs*45)	2	frameshift	-	-	-	-	-
38	c.5317_5318insA	p.(Ala1773Aspfs*14)	1	frameshift	-	-	-	-	-
38	c.5383T>G	p.(Leu1795Val)	1	missense	7.95E-06	deleterious	probably damaging	24.2	damaging
39	c.5531_5557dup	p.(Gly1844_Gln1852dup)	1	stop gained,inframe_insertion	-	-	-	-	-
40	c.5655del	p.(Val1887Trpfs*6)	1	frameshift	-	-	-	-	-
IVS40	c.5714+1G>A	p.(?)	1	splice_donor	3.98E-06	-	-	26.5	-
IVS41	c.5836-1G>C	p.(?)	2	splice_acceptor	-	-	-	26.2	-
42	c.5858del	p.(Pro1953Glnfs*21)	1	frameshift	-	-	-	-	-
43	c.5899T>G	p.(Cys1967Gly)	1	missense	-	deleterious	probably damaging	26	-
43	c.5981G>A	p.(Gly1994Glu)	1	missense	-	deleterious	probably damaging	34	damaging
44	c.6071A>G	p.(Asp2024Gly)	1	missense	-	tolerated	probably damaging	25	damaging
44	c.6147_6147+7del	p.(Val2050Leufs*11)	1	frameshift	-	-	-	-	-
45	c.6216T>A	p.(Ser2072Arg)	1	missense	-	deleterious	probably damaging	25.5	damaging
45	c.6272T>A	p.(Leu2091Gln)	1	missense	-	deleterious		32	damaging
46	c.6310C>T	p.(Gln2104*)	1	stop gained	-	-	-	48	-
IVS46	c.6387-1G>A	p.(?)	1	splice_acceptor	-	-	-	27.1	-
47	c.6437G>T	p.(Gly2146Val)	1	missense	-	deleterious	probably damaging	33	damaging

Table S5. Deep intronic *ABCA4* variants identified in 506 Spanish families

ABCA4 deep intronic variants (NM_000350)				
Intron	Nucleotide	Protein	Number of individual alleles	Number of complex alleles
30	c.4539+2064C>T	p.[=,Arg1514Leufs*36]	15	
36	c.5196+1137G>A	p.[Met1733Glufs*78, =]	4	
28	c.4253+43G>A	p.[=,Ile1377Hisfs*3]	5	1
19	c.2919-826T>A	p.(Leu973Phefs*1)	1	
36	c.5196+1056A>G	p.(Met1733Valfs*2)	1	
7	c.859-506G>C	p.[Phe287Thrfs*32,=]	1	

Table S6. Genotype-phenotype correlation of cone-rod dystrophy (CRD) and Stargardt disease (STGD1) patients regarding type of *ABCA4* variant

Phenotype	All patients	MISSENSE-MISSENSE	MISSENSE-TRUNCATING	TRUNCATING-TRUNCATING
	Median AO (IQR) (N)	Median AO (IQR) (N)	Median AO (IQR) (N)	Median AO (IQR) (N)
CRD	10.0 (6.0) years (N=66)	12.5 (8.3) years (N=16)	10.0 (5.0) years (N=23)	9.0 (5.5) years (N=27)
STGD1	16.0 (15.0) years (N=306)	17.0 (16.0) years (N=167)	15.0 (15.5) years (N=119)	9.0 (3.0) years (N=20)
p	<0.001	<0.05	<0.05	NS

Abbreviations: AO, age of onset; IQR, interquartile range; N, number of cases; NS, non significant

Table S7. Genotype-phenotype correlation regarding type of *ABCA4* variant

Category	A: MISSENSE-MISSENSE	B: MISSENSE-TRUNCATING	p
Median AO (IQR) (N)	17.0 (15.0) years (N=183)	14.0 (14.0) years (N=142)	<0.05
Category	A: MISSENSE-MISSENSE	C: TRUNCATING-TRUNCATING	p
Median AO (IQR) (N)	17.0 (15.0) years (N=183)	9.0 (3.5) years (N=47)	<0.001
Category	B: MISSENSE-TRUNCATING	C: TRUNCATING-TRUNCATING	p
Median AO (IQR) (N)	14.0 (14.0) years (N=142)	9.0 (3.5) years (N=47)	<0.001

Abbreviations: AO, age of onset; IQR, interquartile range; N, number of cases.

Table S8. Genotype-phenotype correlation regarding type of *ABCA4* variant excepting c.3386G>T;p.(Arg1129Leu)

Category	A2: MISSENSE-MISSENSE	B2: MISSENSE-TRUNCATING	p
Median AO (IQR) (N)	16.5 (19.0) years (N=100)	13.5 (16.0) years (N=100)	NS
Category	A2: MISSENSE-MISSENSE	C2: TRUNCATING-TRUNCATING	p
Median AO (IQR) (N)	16.5 (19.0) years (N=100)	9.0 (3.8) years (N=46)	<0.001
Category	B2: MISSENSE-TRUNCATING	C2: TRUNCATING-TRUNCATING	p
Median AO (IQR) (N)	13.5 (16.0) years (N=100)	9.0 (3.8) years (N=46)	<0.001

Abbreviations: AO, age of onset; IQR, interquartile range; N, number of cases; NS, non significant.

Table S9. Genotype-phenotype correlation between c.3386G>T Patients and Non-c.3386G>T Patients.

Category	A2: MISSENSE-MISSENSE	D: c.3386G>T-c.3386G>T	p
Median AO (IQR) (N)	16.5 (19.0) (N=100)	21.5 (18.5) years (N=12)	NS
Category	A2: MISSENSE-MISSENSE	E: c.3386G>T-MISSENSE	p
Median AO (IQR) (N)	16.5 (19.0) (N=100)	17.0 (8.5) years (N=68)	NS
Category	B2: MISSENSE-TRUNCATING	F: c.3386G>T-TRUNCATING	p
Median AO (IQR) (N)	13.5 (16.0) (N=100)	14.0 (10.0) years (N=43)	NS

Abbreviations: AO, age of onset; IQR, interquartile range; N, number of cases; NS, non significant.

Table S10. Comparison between patients carrying the c.3386G>T variant vs. patients carrying the c.5882G>A variant in *ABCA4*

Variant	Median AO (IQR) (N)		
	Homozygous	Variant-MISSENSE	Variant-TRUNCATING
c.3386G>T	21.5 (18.5) years (N=12)	17.0 (8.5) years (N=68)	14.0 (10.0) years (N=43)
c.5882G>A	-	17.0 (13.8) years (N=28)	20.0 (15.0) years (N=15)
p		NS	NS

Abbreviations: AO, age of onset; IQR, interquartile range; N, number of cases; NS, non significant

Table S11. Missense variants associated with CRD or STGD1 phenotypes.

	ABCA4 variant	Alleles STGD1	Alleles CRD	Proportion STGD1	Proportion CRD	OR (95% CI)
Associated with a clinical phenotype (This study)	c.1804C>T; p.(Arg602Trp)	20	10	3.15 (1.93, 4.82)	16.7 (8.29, 28.5)	5.31 (2.27, 11.7)*
	c.3056C>T; p.(Thr1019Met)	7	5	1.10 (0.44, 2.26)	8.33 (2.76, 18.4)	7.58 (2.12, 25.1)*
	c.3386G>T; p.(Arg1129Leu)	177	6	27.9 (24.4, 31.5)	10.0 (3.76, 20.5)	0.37 (0.14, 0.80)#
	c.6320G>C; p.(Arg2107Pro)	2	2	0.31 (0.04, 1.13)	3.33 (0.41, 11.5)	10.5 (1.08, 102)*
Previously reported severe missense variants	c.1622T>C; p.(Leu541Pro)	2	1	0.31 (0.04, 1.13)	1.67 (0.04, 8.94)	5.60 (0.18, 70.1)
	c.[1622T>C;3113C>T]; p.[Leu541Pro;Ala1038Val]	3	0	0.47 (0.10, 1.37)	0.00 (0.00, 5.96)	
	c.2894A>G; p.(Asn965Ser)	3	0	0.47 (0.10, 1.37)	0.00 (0.00, 5.96)	
	c.3113C>T; p.(Ala1038Val)	5	0	0.79 (0.26, 1.83)	0.00 (0.00, 5.96)	
	c.4918C>T; p.(Arg1640Trp)	10	1	1.57 (0.76, 2.88)	1.67 (0.04, 8.94)	1.19 (0.05, 6.47)

Abbreviations: STGD1: Stargardt disease; CRD: Cone-rod dystrophy; OR: odd ratio; CI: confidence interval. *Variant associated with CRD, #Variant associated with STGD1.

Análisis genético y fenotípico de pacientes con variantes en *PROM1*

Artículo 5: Expanded phenotypic spectrum of retinopathies associated with autosomal recessive and dominant mutations in *PROM1*

Del Pozo-Valero M, Martin-Merida I, Jimenez-Rolando B, et al.

Publicado en *American Journal of Ophthalmology*, 2019.

Resumen:

Este trabajo recoge las familias de nuestra cohorte en las que se identificaron variantes raras en el gen *PROM1*.

El objetivo de este estudio fue realizar un análisis detallado de las variantes y fenotipos de los pacientes en los que se habían identificado mutaciones en este gen, y así establecer su prevalencia en nuestra cohorte, ya que hasta la fecha había muy pocas publicaciones relacionadas con este gen.

Se recogió una cohorte de 2216 pacientes en los que el gen *PROM1* había sido cribado. De los pacientes con variantes en el gen *PROM1*, se analizó su patrón de herencia y su clínica oftalmológica para establecer correlaciones genotipo-fenotipo.

En total, 32 familias presentaron variantes en el gen *PROM1*, de las cuales 4 y 18 se consideraron caracterizadas con una herencia autosómica dominante y recesiva, respectivamente. En las 10 familias restantes se identificaron variantes monoalélicas en las que haría falta una segunda variante en este gen para explicar su patología, o variantes de significado clínico incierto, que, debido a estudios de segregación, el patrón de herencia, o su frecuencia, se descartan como causa de la enfermedad, a falta de más estudios en este gen. En las familias con herencia dominante, se identificó una nueva variante de *splicing*, c.303+1G>A;p.(?) en varias familias, siendo una de las variantes más prevalentes en nuestra cohorte junto a la c.1354dup; p.(Tyr452Leufs*13) y c.1984-1G>T; p.(?), en las que se realizaron estudios de haplotipos y revelaron un origen común, siendo las familias portadoras de la variante c.1984-1G>T; p.(?) las que presentaron una región mínima compartida más grande (457 Kb).

En cuanto al fenotipo, la mayoría de los pacientes presentaban distrofia de conos y bastones, aunque también había casos con retinosis pigmentaria y distrofia macular. La característica común a todos los pacientes es que presentaban atrofia macular. En este trabajo también se describe una nueva asociación fenotípica, la variante c.303+1G>A; p.(?) estaría asociada a una maculopatía leve de inicio tardío. Además, en un paciente con RP, se identificó la variante situada en el mismo sitio donador del intrón 3, c.303+2T>C; p.(?), en homocigosis, por lo que se realizaron estudios

funcionales para intentar determinar si la expresión o el patrón de *splicing* era diferente en ambas variantes. Sin embargo, no se pudo dilucidar este hecho ya que el exón 3 de *PROM1* no está presente en las isoformas de los tejidos estudiados, al contrario que en la retina, donde la mayoría de las isoformas presentan este exón, como había sido demostrado en publicaciones anteriores.

En conclusión, este estudio recoge la cohorte más grande de pacientes con variantes en el gen *PROM1*, estableciendo una prevalencia de este gen del 1%, aportando nueva información genética y clínica, y estableciendo una nueva asociación fenotípica.

Contribución de la autora:

En este trabajo la autora estuvo implicada en el diseño del estudio y recopiló todas las variantes anotadas del gen *PROM1* en nuestra cohorte, así como los trabajos sobre este gen publicados hasta ese momento. Realizó los estudios de haplotipos y la segregación de las variantes en algunas de las familias, estableciendo su patogenicidad. En el caso de las variantes del intrón 3, recogió las muestras y realizó los estudios de ARN de expresión del gen *PROM1,* y la posterior secuenciación y análisis de los tránscritos en pacientes y controles. También recopiló los datos clínicos para establecer las correlaciones genotipo-fenotipo en el gen *PROM1*. Finalmente, escribió el manuscrito y revisó críticamente las versiones sucesivas de éste, y participó en el proceso de la publicación del artículo.

Expanded Phenotypic Spectrum of Retinopathies Associated with Autosomal Recessive and Dominant Mutations in *PROM1*

MARTA DEL POZO-VALERO, INMACULADA MARTIN-MERIDA, BELEN JIMENEZ-ROLANDO, ANA ARTECHE, ALMUDENA AVILA-FERNANDEZ, FIONA BLANCO-KELLY, ROSA RIVEIRO-ALVAREZ, CAROLINE VAN CAUWENBERGH, ELFRIDE DE BAERE, CARLO RIVOLTA, BLANCA GARCIA-SANDOVAL, MARTA CORTON, AND CARMEN AYUSO

• PURPOSE: To describe the genetic and phenotypic characteristics of a cohort of patients with *PROM1* variants.

• DESIGN: Case-case study.

• METHODS: We screened a cohort of 2216 families with inherited retinal dystrophies using classical molecular techniques and next-generation sequencing approaches. The clinical histories of 25 patients were reviewed to determine age of onset of symptoms and the results of ophthalmoscopy, best-corrected visual acuity, full-field electroretinography, and visual field studies. Fundus autofluorescence and spectral-domain optical coherence tomography were further assessed in 7 patients.

• RESULTS: *PROM1* variants were identified in 32 families. Disease-causing variants were found in 18 autosomal recessive and 4 autosomal dominant families. Monoallelic pathogenic variants or variants of unknown significance were identified in the remaining 10 families. Comprehensive phenotyping of 25 patients from 22 families carrying likely disease-causing variants revealed clinical heterogeneity associated with the *PROM1* gene. Most of these patients presented cone-rod dystrophy and some exhibited macular dystrophy or retinitis pigmentosa, while all presented with macular damage. Phenotypic association of a dominant splicing variant with late-onset mild maculopathy was established. This variant is one of the 3 likely founder variants identified in our Spanish cohort.

• CONCLUSIONS: We report the largest cohort of patients with *PROM1* variants, describing in detail the phenotype in 25 of them. Interestingly, within the variability of phenotypes related to this gene, macular involvement is a common feature in all patients. (Am J Ophthalmol 2019;207:204–214. © 2019 Elsevier Inc. All rights reserved.)

INHERITED RETINAL DYSTROPHY (IRD) REFERS TO A group of progressive and degenerative diseases that affect the photoreceptor cells and lead to visual impairment.[1] Its prevalence is estimated at 1 in 3500 to 1 in 5000 people worldwide. IRDs are subclassified as retinitis pigmentosa (RP), which is the most common type, cone-rod dystrophy (CRD), and cone dystrophy (CD). Patterns of inheritance include autosomal dominant (ad), autosomal recessive (ar), and X-linked (XL) transmission, though mitochondrial and digenic forms have also been described.[2] The first noticeable symptoms are commonly night blindness and a progressive loss of the peripheral visual field among patients with RP[3] because of the dysfunction or loss of rod photoreceptors and loss of visual acuity (VA) and photophobia in patients with CD/CRDs[4] caused by cone degeneration. To date, 307 loci have been found to cause IRD, with 271 genes identified (https://sph.uth.edu/retnet), one of which is *PROM1* (RefSeq NM_006017; OMIM 604365).

Located on chromosome 4p15.32, *PROM1* comprises 27 exons and encodes prominin-1, a 5-transmembrane domain glycoprotein originally identified as a hematopoietic stem cell antigen, CD133/AC133.[5,6] Prominin-1 is expressed in several stem and progenitor cells originating from various sources, including the neural and hematopoietic systems.[7,8] Photoreceptor, glial, and epithelial cells from developing and adult organs express prominin-1.[9–11] In the retina, PROM1 is located at the base of photoreceptor outer segments, where it is involved in disk membrane morphogenesis.[12] Seven alternative splice variants have been reported in human tissue,[13] 2 of which are the most highly expressed isoforms in the retina.[14] Disease-causing *PROM1* variants have been associated with arRP with macular degeneration, arCRD, adCRD,

and macular phenotypes, such as Stargardt-like and bull's eye maculopathies.[15,16] We used state-of-the-art genotyping techniques to study the involvement of PROM1 as the potential cause of a variety of IRDs in a large cohort of 2216 families, mostly from Spain. We provide genetic data from a total of 32 families carrying novel and previously described PROM1 variants. We also describe the phenotype associated with likely disease-causing PROM1 variants by further detailed ophthalmic evaluation in 25 patients. Overall, these data are important for the development of future therapeutic trials, including gene-specific therapies for PROM1-related phenotypes.

SUBJECTS AND METHODS

• **SUBJECTS AND SAMPLES:** All 2216 patients were clinically diagnosed with nonsyndromic IRD and recruited at Fundación Jiménez Díaz University Hospital (FJD; Madrid, Spain). This study was designed in compliance with the tenets of Declaration of Helsinki, and patient enrollment was approved by the FJD ethics committee. DNA samples were collected from the FJD biobank. Total RNA of 4 patients and 3 control subjects was obtained from saliva, whole blood, blood-isolated mononuclear cells, or fibroblasts using NucleoSpin RNA kits (NucleoSpin; Macherey-Nagel, Duren, Germany) or TRI reagent (Thermo Fisher Scientific, Waltham, Massachusetts, USA), following the manufacturer's instructions. Before RNA extraction, saliva samples were collected using the Oragene RNA kit (DNA Genotek; Ottawa, Ontario, Canada).

• **MOLECULAR SCREENING:** Molecular screening of the IRD patients making up this cohort has been performed over the past 27 years using different genetic tools. First, commercial genotyping microarrays (ASPER Ophthalmics, Tartu, Estonia) and/or sequencing of the most prevalent genes, depending on the patient's phenotype, were applied to 1762 index patients. Following this, 957 remained uncharacterized and were subsequently studied through different next-generation sequencing (NGS) strategies, including targeted gene panels,[17] clinical exome, and/or whole-exome sequencing, as previously described.[18] An additional 454 new cases were studied by NGS as the first approach.

The pathogenicity of unreported variants was established according to their allele frequency in the Exome Aggregation Consortium and gnomAD (http://gnomad. broadinstitute.org/), in silico prediction tools for new splice and missense variants, including Combined Annotation Dependent Depletion[19] and those included in the commercial Alamut software (v2.9.0; Interactive Biosoftware, Rouen, France), and cosegregation studies in the family when other relatives were available. The guidelines of the American College of Medical Genetics and Genomics were used for variant classification.[20]

In addition, we performed high-resolution copy number variant analysis of PROM1 using the customized arrEYE Agilent CGH-8x60K microarray[21] (Agilent Technologies, Inc., Santa Clara, California, USA) and Sanger sequencing to discard the intronic variants c.2077-521A>G[22] and the c.2281-20_2281-11del[23] in PROM1 in 6 patients carrying monoallelic variants in PROM1 (RP-0365, RP-1740, MD-0289, RP-1243, RP-2831, and MD-0875) as well as the affected daughter of MD-0873, who presented with another phenotype (Leber congenital amaurosis and congenital heart disease).

• **HAPLOTYPE ANALYSIS:** Three microsatellite markers (D4S1602, D4S2960, and D4S1567) and 7 single nucleotide polymorphisms (rs73125627, rs3213710, rs2072313, rs6449209, rs3815344, rs2286455, and rs2078622) flanking 2 Mb around PROM1 were studied in 14 families with the variants c.303+1G>A, c.1354dup; p.Tyr452Leufs*13, and c.1984-1G>T.

• **EXPRESSION ANALYSIS OF THE PROM1 GENE:** Total RNA obtained from the whole blood, blood mononuclear cells, and saliva of 4 individuals carrying the pathogenic splicing variants c.303+1G>A and c.303+2T>C, as well 3 wild-type controls, was transcribed by using the SuperScript IV cDNA synthesis kit (Thermo Fisher Scientific) in a final volume of 20 µL. Fibroblast RNA from 1 control subject was also used specifically to evaluate PROM1 expression in this tissue. Reverse transcription polymerase chain reaction assays were performed using specific primers for exons 2 and 5.

• **CLINICAL ASSESSMENT:** IRDs were classified as CD/ CRD or RP phenotypes based on clinical examination together with presence and onset of either VA loss and photophobia or visual field constriction and night blindness as the initial symptoms, respectively. A macular dystrophy (MD) phenotype was defined in the patients with macular lesions and normal full-field electroretinogram responses. A comprehensive ophthalmologic examination including a detailed medical history, measurements of best-corrected visual acuity (BCVA), and ophthalmoscopy was available for 25 patients carrying PROM1 variants. In addition, RP patients and patients carrying the c.303+1G>T variant were also tested using spectral-domain optical coherence tomography (OCT) and fundus autofluorescence (FAF). Fluorescein angiography (FA) was also performed in 1 patient from family MD-0934.

RESULTS

• **MOLECULAR FINDINGS FROM PROM1 SCREENING:** PROM1 was studied using conventional methods and/or NGS (Supplemental Figure 1). We analyzed a cohort of 2216 patients comprised primarily of Spanish individuals,

TABLE 1. Likely Causative *PROM1* Variants in Our Cohort of Patients with Inherited Retinal Dystrophy

Family	Phenotype	Allele 1				Allele 2				Segregation
		Exon/Intron	Nucleotide	Protein	Reference	Exon/Intron	Nucleotide	Protein	Reference	
Causative variants in ad families										
MD-0235	MD	Intron 3	c.303+1G>A	Splicing	Bryant and associates[27]	—	—	—	—	Yes
MD-0934	CRD/MD	Intron 3	c.303+1G>A	Splicing	Bryant and associates[27]	—	—	—	—	Yes
MD-1074	CRD	Intron 3	c.303+1G>A	Splicing	Bryant and associates[27]	—	—	—	—	NA
RP-2604	RP	Exon 10	c.1117C>T	p.Arg373Cys	Maw and associates[30]	—	—	—	—	NA
Causative variants in ar families										
RP-0878	RP	Intron 3	c.303+2T>C	Splicing	This study	Intron 3	c.303+2T>C	Splicing	This study	Yes
MD-0100	CRD	Exon 5	c.622del	p.Thr208Leufs*23	Zhao and associates[34]	Intron 5	c.630+1G>A	Splicing	This study	Yes
MD-0649	CRD	Intron 5	c.630+1G>A	Splicing	This study	Intron 5	c.630+1G>A	Splicing	This study	NA
RP-0586	CRD	Exon 8	c.869del	p.Ser290fs*2	Permanyer and associates[14]	Exon 11	c.1177_1178del	p.Ile393Argfs*21	Carss and associates[35]	NA
MD-0383	CRD	Exon 12	c.1354dup	p.Tyr452Leufs*13	Pras and associates[29]	Exon 12	c.1354dup	p.Tyr452Leufs*13	Pras and associates[29]	NA
MD-0803	CRD	Exon 12	c.1354dup	p.Tyr452Leufs*13	Pras and associates[29]	Exon 12	c.1354dup	p.Tyr452Leufs*13	Pras and associates[29]	NA
RP-1899	CRD	Exon 12	c.1354dup	p.Tyr452Leufs*13	Pras and associates[29]	Exon 12	c.1354dup	p.Tyr452Leufs*13	Pras and associates[29]	Yes
RP-0855	CRD	Exon 12	c.1354dup	p.Tyr452Leufs*13	Pras and associates[29]	Exon 12	c.1354dup	p.Tyr452Leufs*13	Pras and associates[29]	Yes
MD-0682	CRD	Exon 12	c.1354dup	p.Tyr452Leufs*13	Pras and associates[29]	Exon 12	c.1354dup	p.Tyr452Leufs*13	Pras and associates[29]	NA
RP-2752	CRD	Exon 12	c.1354dup	p.Tyr452Leufs*13	Pras and associates[29]	Exon 12	c.1354dup	p.Tyr452Leufs*13	Pras and associates[29]	NA
RP-2924	CRD	Exon 12	c.1354dup	p.Tyr452Leufs*13	Pras and associates[29]	Exon 12	c.1354dup	p.Tyr452Leufs*13	Pras and associates[29]	NA
RP-1110	CRD	Exon 12	c.1354dup	p.Tyr452Leufs*13	Pras and associates[29]	Exon 15	c.1709_1710insAA	p.Tyr570*	This study	Yes
MD-0838	CRD	Exon 12	c.1354dup	p.Tyr452Leufs*13	Pras and associates[29]	Intron 19	c.2130+2del	Splicing	Patel and associates[32]	Yes
MD-0654	CRD	Exon 12	c.1414del	p.Arg472Glufs*18	This study	Exon 12	c.1414del	p.Arg472Glufs*18	This study	NA
MD-0123	CRD	Intron 17	c.1984-1G>T	Splicing	De Castro-Miró and associates[24]	Intron 17	c.1984-1G>T	Splicing	De Castro-Miró and associates[24]	NA
MD-0762	CRD	Intron 17	c.1984-1G>T	Splicing	De Castro-Miró and associates[24]	Intron 17	c.1984-1G>T	Splicing	De Castro-Miró and associates[24]	Yes
MD-0873	CRD	Intron 17	c.1984-1G>T	Splicing	De Castro-Miró and associates[24]	Intron 17	c.1984-1G>T	Splicing	De Castro-Miró and associates[24]	NA
Likely causative variant										
RP-1852	CRD	Exon 12	c.1435G>A	p. Gly479Arg	This study	Exon 12	c.1435G>A	p. Gly479Arg	This study	NA

CRD = cone-rod dystrophy; MD = macular dystrophy; NA = not available; RP = retinitis pigmentosa.

TABLE 2. Other *PROM1* Variants Identified in Our Cohort of Patients with Inherited Retinal Dystrophy

Family	Inheritance	Phenotype	PROM1 Variant in Heterozygosis						Causative Variants in Other Genes
			Exon/Intron	Nucleotide	Protein	Reference	Classification	Comments	
Families with a noncausative variant in *PROM1*									
MD-0875	AR (sporadic)	CRD	Exon 4	c.314A>G	p.Tyr105Cys	This study	VUS	No cosegregation with the disease	—
MD-0001	AD	MD	Exon 4	c.437G>A	p. Arg146Gln	This study	VUS	No cosegregation with the disease	—
RP-0365	AR	RP	Exon 5	c.604C>G	p.Arg202Gly	Patel and associates[32]	VUS	Variant reclassified to VUS supported by ClinVar, population frequency, and this study	—
RP-1740	AR (sporadic)	RP	Exon 5	c.604C>G	p.Arg202Gly	Patel and associates[32]	VUS	Variant reclassified to VUS supported by ClinVar, population frequency, and this study	RP1
MD-0059	AD	CRD	Exon 13	c.1472T>C	p.Phe491Ser	This study	VUS	No cosegregation with the disease	—
RP-2645	AR	RP	Exon 23	c.2408T>C	p.Val803Ala	This study	VUS	No cosegregation with the disease	COL11A1
RP-2680	AR	CRD	Intron 23	c.2490-2A>G	Splicing	This study	Likely pathogenic	No cosegregation with the disease	ABCA4
Families incompletely characterized with a pathogenic recessive variant in *PROM1*									
MD-0289	AR (sporadic)	CRD	Exon 12	c.1354dup	p.Tyr452Leufs*13	Pras and associates[29]	Pathogenic	Lack of a second allele	—
RP-1243	AR (sporadic)	RP	Exon 12	c.1354dup	p.Tyr452Leufs*13	Pras and associates[29]	Pathogenic	Lack of a second allele	—
RP-2831	AR (sporadic)	RP	Exon 12	c.1354dup	p.Tyr452Leufs*13	Pras and associates[29]	Pathogenic	Lack of a second allele	—

AD = autosomal dominant; AR = autosomal recessive; CRD = cone-rod dystrophy; MD = macular dystrophy; RP = retinitis pigmentosa; VUS = variant of unknown significance.

149

TABLE 3. Clinical Characteristics of Retinitis Pigmentosa and Cone-Rod Dystrophy/Macular Dystrophy Patients with Causative Variants in *PROM1*

Characteristics	RP	CRD/MD	c.1354dup	c.1984-1G>T	c.303+1G>A
Age at diagnosis, n	2	21	9	2	5
Mean ± SD (y)	29.5 ± 4.95	23.76 ± 17.64	15.77 ± 8.21	22 ± 5.66	45.6 ± 24.42
Range (y)	26-33	4-65	3-28	18-26	4-65
First symptom[a]					
Nyctalopia, n (%)	2/2 (100)*	12/17 (71)	6/8 (75)	1/1 (100)	2/4 (50)
Mean ± SD	24 ± 8.48	26.83 ± 15.19	24.8 ± 17.96	25	46 ± 11.31
Visual field loss, n (%)	2/2 (100)	15/20 (75)	6/9 (67)	3/3 (100)	1/5 (20)
Mean ± SD	28 ± 2.83	17 ± 10.33	15 ± 5.59	12 ± 6.56	50
Visual acuity loss, n (%)	2/2 (100)	20/20 (100)*	9/9 (100)*	3/3 (100)*	4/4 (100)*
Mean ± SD	32	18.21 ± 16.78	12.13 ± 6.56	10.67 ± 8.08	46.75 ± 11.70
Funduscopic examination					
Typical RP, n (%)	2/2 (100)	7/21 (33)	5/9 (55)	—	—
Macular RPE changes, n (%)	2/2 (100)	20/20 (100)	8/8 (100)	3/3 (100)	5/5 (100)
Full-field electroretinography, n (%)					
Scotopic and photopic extinguished	2/2 (100)	1/13 (8)	—	—	—
Rod-cone pattern	—	—	—	—	—
Cone-rod pattern	—	3/13 (23)	3/5 (60)	—	—
Cone pattern	—	—	—	—	—
Scotopic and photopic reduced, pattern not further specified	—	6/13 (46)	2/5 (40)	2/2 (100)	2/5 (20)
Normal	—	3/13 (23)	—	—	3/5 (60)

Patients with the c.1354dup, c.1984-1G>T, and c.303+1G>A variants are shown separately.
CRD = cone-rod dystrophy; MD = macular dystrophy; RP = retinitis pigmentosa; RPE = retinal pigment epithelium; SD = standard deviation.
[a]First symptom reported.

all of whom presented with nonsyndromic IRD, including RP, CD/CRD, and/or MD. Tables 1 and 2 summarize the variants identified in *PROM1* from a total of 32 families (Supplemental Figure 2).

In total, we identified 19 *PROM1* variants, 10 of which are novel. The pathogenicity of these variants was established in accordance with the American College of Medical Genetics and Genomics classification[20] (Supplemental Table 1). The location of the likely pathogenic variants in the coding sequence of *PROM1* is shown in Supplemental Figure 3.

Thirteen different variants were considered causative/likely causative; these appeared in 18 arRP or arCRD families, 3 adCRD/MD families, and 1 adRP family (Table 1). One was a novel missense homozygous variant (c.1435G>A; p.Gly479Arg) in family RP-1852.

The 6 remaining *PROM1* variants appeared in heterozygosis in 7 families (Table 2). In 4 families we identified 4 novel missense variants of unknown significance (VUSs), including 4 heterozygous missense variants that were also present in their apparently healthy relatives (families MD-0875, MD-0001, MD-0059, and RP-2645). The previously reported variant c.604C>G; p.Arg202Gly was identified in families RP-0365 and RP-1740, and family RP-2680 carried the novel splicing variant c.2490-2A>G. In addition, the known recessive variant

c.1354dup was present in 3 monoallelic families (MD-0289, RP-1243, and RP-2831; Table 2).

No copy number variants or intronic variants could be identified in sporadic or recessive cases carrying monoallelic *PROM1* variants. We also assessed the presence of additional pathogenic variants in other IRD genes by means of targeted NGS, enabling us to identify 4 likely pathogenic variants in 3 cases (Supplemental Table 2). In proband RP-1740, who presented with RP, we also identified a likely pathogenic recessive variant in heterozygosis in *RP1* in addition to the p.Arg202Gly variant in *PROM1*. In family RP-2680, presenting with arCRD, biallelic variants were found in *ABCA4* in addition to the splicing variant c.2490-2A>G in *PROM1*. Segregation analysis showed recessive segregation of *ABCA4* alleles in the 2 affected siblings, together with a lack of segregation for the *PROM1* variant. Remarkably, the index case of family RP-2645 carried a novel de novo VUS in *COL11A1* along with the c.2408T>C; p.Val803Ala *PROM1* VUS. Upon clinical reevaluation, we considered the variant in *COL11A1* to be likely causative in this patient.

The most frequent pathogenic variant in our cohort was the c.1354dup variant, which leads to a frameshift p.Tyr452Leufs*13; this variant was discovered in 12 unrelated families, 1 of them from Ecuador (MD-0682). Nine of these variants fully explained the IRD phenotype because it
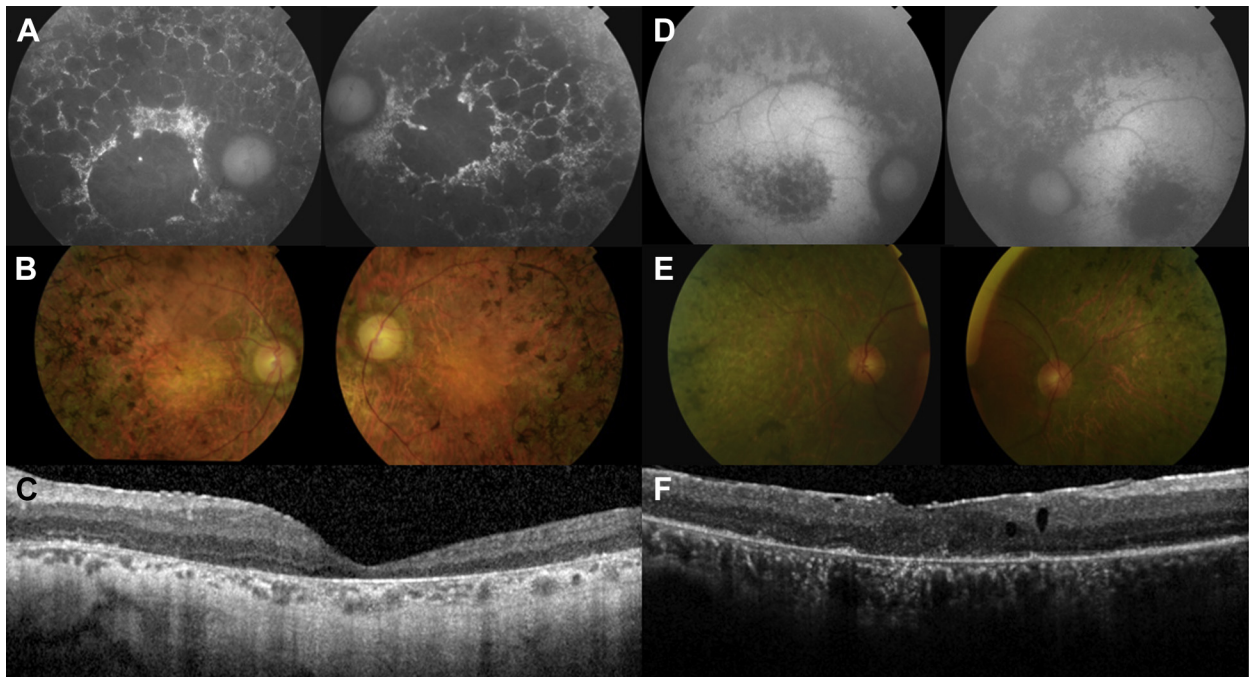
FIGURE 1. Imaging findings in patients with retinitis pigmentosa (RP) carrying *PROM1* variants. (A–C) Proband of RP-0878, carrying the c.303+2T>C variant in homozygosis, 46 years of age. (D–F) Proband of RP-2604, carrying the c.1117C>T; p.Arg373Cys variant in heterozygosis, 56 years of age. Fundus autofluorescence images show extensive hypoautofluorescent lesions located in the mid-periphery and the macular area. Fundus photographs reveal bone spicule pigmentation in the periphery, retinal vessel attenuation, and severe retinal atrophy. Spectral-domain optical coherence tomography showed extensive loss of the ellipsoid band in both cases, as well as a severe foveal and choroidal atrophy in patient RP-0878, and an epiretinal membrane with intraretinal cystic spaces in patient RP-2604. The absence of the ellipsoid and retinal pigment epithelium layers in the spectral-domain optical coherence tomography images is consistent with the fundus autofluorescence findings.

was found in homozygosis (7 families) or compound heterozygosis (2 families). In family RP-1110 we identified this variant segregating with a novel nonsense variant in the proband, and in homozygosis in an affected fourth-degree relative.

In 11 families we identified 6 splicing variants, 3 of which were novel. The variant c.303+1G>A, affecting the donor splice site of intron 3, was found in 5 patients from 3 families (MD-0235, MD-0934, and MD-1074), and behaved in a dominant manner. In family MD-1074 we could not establish the inheritance mode, as the heterozygous proband was conceived by in vitro fertilization with egg donation. No additional likely pathogenic variants were identified by whole-exome sequencing in family MD-0235. The variants c.303+2T>C and c.630+1G>A, affecting the donor sites in introns 3 and 5, respectively, showed a recessive pattern of inheritance in 3 families (RP-0878, MD-0100, and MD-0649), and c.2490-2A>G, located at the acceptor site in intron 23, was identified in heterozygosis in the *ABCA4*-mutated patient mentioned above (RP-2680). Two previously reported splice variants were found in 4 unrelated families: the c.1984-1G>T in homozygosis in 3 families (MD-0123,

MD-0762, and MD-0873) and the c.2130+2del variant in compound heterozygosis in family MD-0654.

RNA expression analyses were performed to study the splicing effect of the c.303+1G>A and c.303+2T>C variants. *PROM1* was expressed in whole blood, blood mononuclear cells, saliva, and fibroblasts. However, this study demonstrated that exon 3 was skipped in the isoforms expressed in these tissues (Supplemental Figure 4).

Considering all the variants identified in this study, the prevalence of *PROM1* variants in our IRD cohort was 1.4%, explaining the molecular defect in 22 families (∼1%).

• IDENTIFICATION OF FOUNDER PATHOGENIC VARIANTS IN *PROM1*: Haplotype analysis was carried out in a total of 14 unrelated Spanish families to shed light on the possibility of founder effects for the 3 most prevalent pathogenic variants in our *PROM1* cohort. Haplotyping in 9 families (10 patients) carrying the common variant c.1354dup, in 2 families with the dominant allele c.303+1G>A, and in 3 families with the c.1984-1G>T variant revealed shared minimal genomic regions of 99 Kb, 17 Kb, and 457 Kb, respectively (Supplemental Figure 2).

print & web 4C/FPO
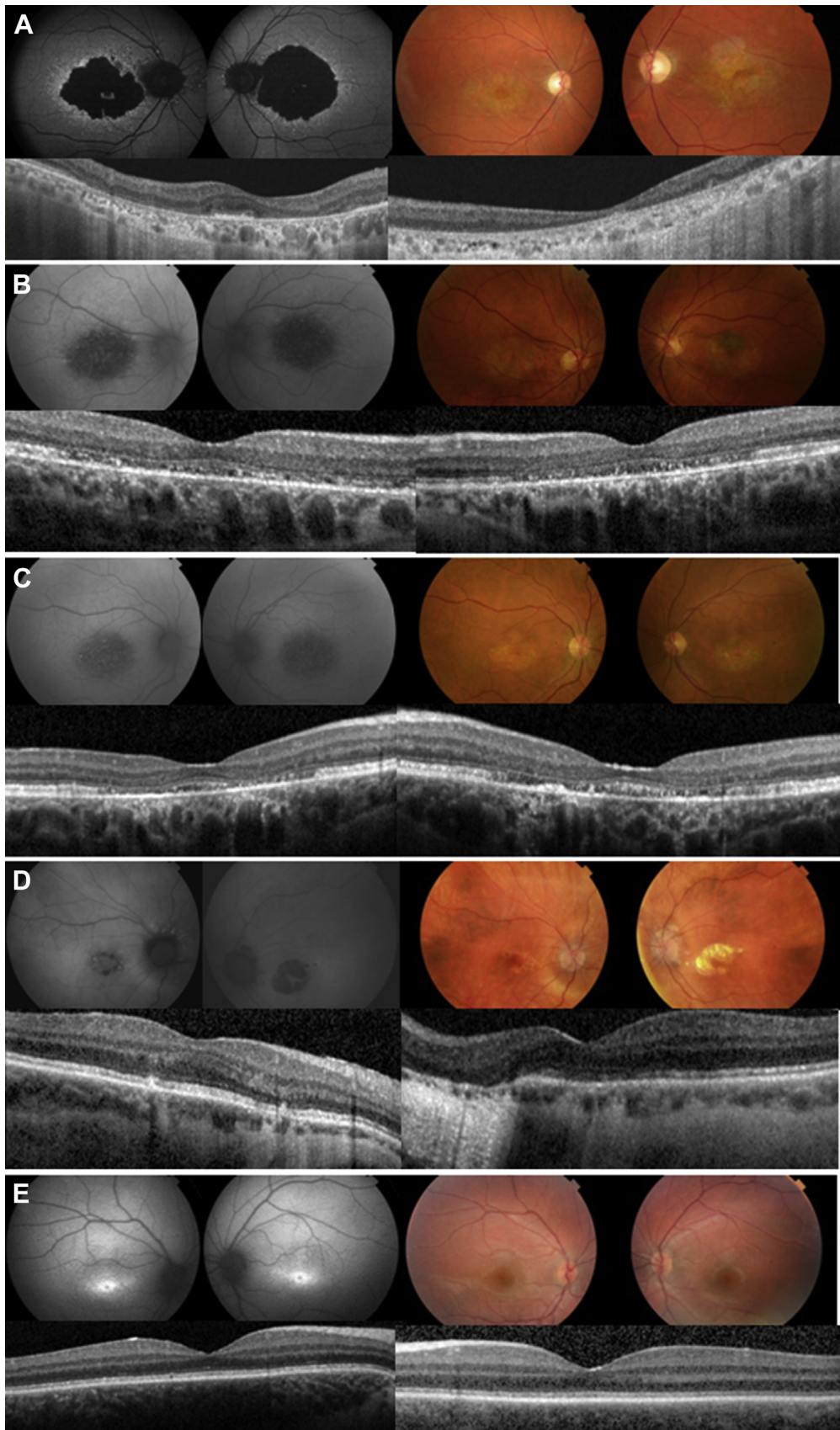
• PHENOTYPE OF PATIENTS CARRYING LIKELY DISEASE-CAUSING VARIANTS IN PROM1: The PROM1-associated phenotype and the clinical data available for 25 patients from 22 unrelated families were reviewed in detail and appear in Table 3 and Supplemental Table 3. According to the clinical assessment, most patients harboring PROM1 variants in our cohort had CRD, and only 3 and 2 patients were diagnosed with MD and RP, respectively. The CRD phenotype of the RP-1852 patient who carried a novel missense variant in homozygosis in PROM1 is also described in Supplemental Table 3. The first symptoms of the RP patients (RP-0878 and RP-2604) were night blindness (mean 24 ± 8.48 y) and loss of visual field (mean 28 ± 2.83 y). At the age of examination, these RP patients presented with low BCVA and extinguished rod and cone responses in the full-field electroretinogram (Table 3). Fundus images, FAF, and spectral-domain OCT revealed typical bone spicule-like pigmentation in the periphery as well as macular atrophy (Figure 1).

Considering the 20 patients with CRD/MD for whom information was available, VA loss was uniformly the earliest presenting symptom (mean 18.21 ± 16.78 y) in all of them. The primary symptoms were similar between patients carrying the recessive variants c.1354dup and c.1984-1G>T (Table 3). Phenotypically, high variability was seen for nearly all ophthalmologic parameters analyzed (BCVA, ophthalmoscopy, visual field, and full-field electroretinogram), even within patients showing the same genotype and belonging to the same family. Four patients with the c.303+1G>A splicing variant showed a distinct phenotype characterized by a later onset (Table 3 and Figure 2). They presented VA loss in their fifth decade of life (mean 46.75 ± 11.70 y) and their BCVA was well preserved until that age. At the age of diagnosis, the only symptom presented by the youngest patient with the c.303+1G>A variant (MD-1074) was an alteration in color perception. FAF and spectral-domain OCT images of all patients showed changes of varying severity in the macular region and the retinal layers (Figure 2). In the

case of the child of MD-1074, the ophthalmologic findings were normal except for macular hyperfluorescence in the FAF. Patient MD-0934 III:1 has recently developed a choroidal neovascular membrane in the right eye (Supplemental Figure 5).

DISCUSSION

IN THIS STUDY WE PRESENT THE LARGEST COHORT OF IRD patients carrying variants in PROM1, describing genetic data from 32 IRD families and phenotypic data from 25 patients carrying likely disease-causing variants. We found 19 rare PROM1 variants in these families, 10 of which are novel. The prevalence of PROM1 variants that seems to reliably explain the IRD phenotype is about 1%. The most prevalent pathogenic variant found in 12 unrelated families was c.1354dup. Haplotyping revealed a common 99-Kb genomic region linked to the variant in the probands analyzed, including 1 Latin American patient likely of Spanish ancestry. These findings indicate a possible founder effect for this pathogenic variant in our cohort of Spanish origin.

Around 12 pathogenic splicing PROM1 variants have been reported. We report a remarkably high proportion of splicing site alleles, explaining the phenotype in 10 families. Found in a total of 6 families, 2 splicing variants (c.303+1G>A and c.1984-1G>T) are frequent alleles in our cohort, likely because of founder effects as revealed by haplotype analysis. The longest shared region comprises the c.1984-1G>A splicing variant, which has been described in Spanish and French populations,[24–26] thereby suggesting a founder effect in this geographic area. In addition, the variants c.303+1G>A and c.303+2T>C showed different behavior in our families as concerns the underlying inheritance pattern, acting as dominant or recessive alleles, respectively. Although the heterozygous variant c.303+1G>A has been recently reported in a patient with Stargardt disease,[27] no family data were available. Both variants affected the

---

◀

FIGURE 2. Imaging findings in patients with the c.303 + 1G > A variant in heterozygosis. (A) Fundus photographs obtained at 52 years of age, fundus autofluorescence (FAF), and spectral-domain optical coherence tomography (OCT) images obtained at 67 years of age from patient MD-0235. (B) FAF, fundus photographs, and spectral-domain OCT images obtained of patient MD-0934 III:8 at 51 years of age. (C) FAF, fundus photographs, and spectral-domain OCT images obtained of patient MD-0934 III:6 (sister of MD-0934 III:8) at 53 years of age. (D) FAF, fundus photographs, and spectral-domain OCT images obtained of patient MD-0934 III:1 (cousin of MD-0934 III:8) at 62 years of age. (E) FAF, fundus photographs, and spectral-domain OCT images obtained of patient MD-1074 at 6 years of age. The FAF highlights lesions of variable severity limited to the macular area from a marked plaque of atrophy in patient A, speckled hypoautofluorescent in patients B and C, and hyperautofluorescent in patient E. Patient D showed a preserved macula in the OD and a plaque of atrophy surrounding the foveal area in the OS. Fundus photographs reveal macular atrophy in patients A–D sparing the fovea in the OD of patient A and the OS of patient D. The retinal fundus photograph of patient E has a normal appearance. Spectral-domain OCT images demonstrate perifoveal thinning of the outer retina in patient A and a marked loss of photoreceptor inner segment/outer segment band along with central choroid thickening in patients B and C. In patient C the fovea was preserved, as evidenced by the "flying saucer" sign. Foveal sparing is observed in both eyes of patient A and in the OS of patient D. Patient E shows a normal appearance in the spectral-domain OCT images.

canonical donor site of exon 3, which is an alternative exon included in the 2 main retinal PROM1 isoforms.[14] Skipping of exon 3 seems to occur naturally in extraocular tissue, which has prevented us from studying the potential splicing effect of these variants. It is therefore not possible to determine the molecular mechanism of the dominant splicing variant c.303+1G>A given the unavailability of retinal tissue from the patients. Moreover, because exon 3 is presented in 97% of the isoforms in the retina[14] we suggest that exon 3 could play an important role in photoreceptor function and/or maintenance that should be further studied.

Variable phenotypes and different inheritance patterns are associated with pathogenic variants in PROM1. It was previously considered that heterozygous missense variants are associated with ad Stargardt-like and bull's eye maculopathies,[28] while nonsense and frameshift variants are associated with arRP with macular degeneration.[14,29,30] However, our findings do not seem to support these observations. In our cohort, most of the identified PROM1 variants were biallelic and seem to be associated with arCRD. These CRD patients showed both variable intra- and interfamilial age at onset, making it difficult to establish a genotype–phenotype correlation in PROM1 regarding specific variants. The RP-associated variants were less frequent, with only 2 cases. However, regardless of the initial diagnosis of primary cone or rod loss, all patients with PROM1 disease-causing variants developed macular involvement, which therefore should be considered a characteristic phenotypic finding of PROM1.

Remarkably, patients carrying the splicing variant c.303+1G>A showed particular phenotypic features not seen in other PROM1-associated cases in our cohort, including a later onset in the fifth decade with well-preserved VA until 50 years of age. However, this variant was also found in a younger patient with color disturbances. As clinical data were not available in the other patients at such an early age, we could not establish whether the disease in this child would progress slowly or if there could be a modifying factor making the patient's phenotype more severe.

We identified a novel missense variant in homozygosis in family RP-1852. Though unable to perform segregation studies, we consider this variant a candidate to explain the CRD phenotype of the patient. Moreover, homozygous missense variants in PROM1 have been described before in a patient with CRD.[31]

Monoallelic PROM1 variants, including 5 novel variants predicted to be pathogenic by in silico programs and 2 variants reported previously, were also found in ad and ar families. We were able to rule out PROM1 as a causal factor in 2 families with likely disease-causing variants in other IRD genes. The variant p.Arg202Gly reported as a dominant disease-causing variant[32] was identified in 2 ar families. The inconsistency of its inheritance pattern and its high frequency in population databases (Exome Aggregation Consortium; 0.015%) did not lend support to dominant-type pathogenicity[33]; therefore, it should be considered a VUS with unlikely causality in our patients. Three VUSs in PROM1 were segregating in healthy relatives of 3 families, excluding them as possibly disease-causing dominant alleles unless incomplete penetrance was present or symptoms were not yet developed. Further in-depth studies of the intronic regions are needed to elucidate whether PROM1 is the causal gene in the rest of ar cases carrying monoallelic PROM1 variants.

In summary, we describe the clinical and genetic findings of the largest cohort of IRD patients with variants in PROM1 reported to date. NGS allows us to evaluate the contribution of PROM1 to IRD. This approach is crucial for accurate genetic counseling, disease prognosis, and for the development of future gene therapies. Disease-causing variants explain 1% of our cohort, supporting the idea that this gene is a rare cause of IRD. This study also highlights the great heterogeneity of PROM1-associated phenotypes regardless of genotype. All patients harboring disease-causing variants in this gene present characteristic macular involvement, which seems to support the pathogenicity of PROM1 variants.

REFERENCES

1. Hartong DT, Berson EL, Dryja TP. Retinitis pigmentosa. *Lancet* 2006;368(9549):1795–1809.
2. Ferrari S, Di Iorio E, Barbaro V, et al. Retinitis pigmentosa: Genes and disease mechanisms. *Curr Genomics* 2011;12(4): 238–249.
3. Ayuso C, Millan JM. Retinitis pigmentosa and allied conditions today: A paradigm of translational research. *Genome Med* 2010;2(5):34.
4. Thiadens AA, Phan TM, Zekveld-Vroon RC, et al. Clinical course, genetic etiology, and visual outcome in cone and cone-rod dystrophy. *Ophthalmology* 2012;119(4): 819–826.
5. Miraglia S, Godfrey W, Yin AH, et al. A novel five-transmembrane hematopoietic stem cell antigen: isolation, characterization, and molecular cloning. *Blood* 1997;90(12): 5013–5021.
6. Yin AH1, Miraglia S, Zanjani ED, et al. AC133, a novel marker for human hematopoietic stem and progenitor cells. *Blood* 1997;90(12):5002–5012.
7. Fargeas C. Prominin-1 (CD133): from progenitor cells to human diseases. *Future Lipidol* 2006;1(2):213–225.
8. Bauer N, Fonseca AV, Florek M, et al. New insights into the cell biology of hematopoietic progenitors by studying prominin-1 (CD133). *Cells Tissues Organs* 2008;188(1-2): 127–138.
9. Weigmann A, Corbeil D, Hellwig A, Huttner WB. Prominin, a novel microvilli-specific polytopic membrane protein of the apical surface of epithelial cells, is targeted to plasmalemmal protrusions of non-epithelial cells. *Proc Natl Acad Sci U S A* 1997;94(23):12425–12430.
10. Corbeil D, Röper K, Hellwig A, et al. The human AC133 hematopoietic stem cell antigen is also expressed in epithelial cells and targeted to plasma membrane protrusions. *J Biol Chem* 2000;275(8):5512–5520.
11. Florek M, Haase M, Marzesco AM, et al. Prominin-1/CD133, a neural and hematopoietic stem cell marker, is expressed in adult human differentiated cells and certain types of kidney cancer. *Cell Tissue Res* 2005;319(1):15–26.
12. Yang Z, Chen Y, Lillo C, et al. Mutant prominin 1 found in patients with macular degeneration disrupts photoreceptor disk morphogenesis in mice. *J Clin Invest* 2008;118(8): 2908–2916.
13. Fargeas CA, Huttner WB, Corbeil D. Nomenclature of prominin-1 (CD133) splice variants - An update. *Tissue Antigens* 2007;69(6):602–606.
14. Permanyer J, Navarro R, Friedman J, et al. Autosomal recessive retinitis pigmentosa with early macular affectation caused by premature truncation in PROM1. *Invest Ophthalmol Vis Sci* 2010;51(5):2656–2663.
15. Michaelides M, Johnson S, Poulson A, et al. An autosomal dominant bull's-eye macular dystrophy (MCDR2) that maps to the short arm of chromosome 4. *Invest Ophthalmol Vis Sci* 2003;44(4):1657–1662.
16. Kniazeva MF, Chiang MF, Cutting GR, et al. Clinical and genetic studies of an autosomal dominant cone-rod dystrophy with features of Stargardt disease. *Ophthalmic Genet* 1999; 20(2):71–81.
17. Perez-Carro R, Corton M, Sánchez-Navarro I, et al. Panel-based NGS reveals novel pathogenic mutations in autosomal recessive retinitis pigmentosa. *Sci Rep* 2016;6:19531.
18. Martin-Merida I, Aguilera-Garcia D, Fernandez-San Jose P, et al. Toward the mutational landscape of autosomal dominant retinitis pigmentosa: A comprehensive analysis of 258 Spanish families. *Invest Ophthalmol Vis Sci* 2018;59(6): 2345–2354.
19. Kircher M, Witten DM, Jain P, et al. A general framework for estimating the relative pathogenicity of human genetic variants. *Nat Genet* 2014;46(3):310–315.
20. Richards S, Aziz N, Bale S, et al. Standards and guidelines for the interpretation of sequence variants: A joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet Med* 2015;17(5):405–424.
21. Van Cauwenbergh C, Van Schil K, Cannoodt R, et al. ArrEYE: A customized platform for high-resolution copy number analysis of coding and noncoding regions of known and candidate retinal dystrophy genes and retinal noncoding RNAs. *Genet Med* 2017;19(4): 457–466.
22. Mayer AK, Rohrschneider K, Strom TM, et al. Homozygosity mapping and whole-genome sequencing reveals a deep intronic PROM1 mutation causing cone-rod dystrophy by pseudoexon activation. *Eur J Hum Genet* 2016;24(3): 459–462.
23. Eidinger O, Leibu R, Newman H, et al. An intronic deletion in the PROM1 gene leads to autosomal recessive cone-rod dystrophy. *Mol Vis* 2015;21:1295–1306.
24. De Castro-Miró M, Pomares E, Lorés-Motta L, et al. Combined genetic and high-throughput strategies for molecular diagnosis of inherited retinal dystrophies. *PLoS One* 2014; 9(2):e88410.
25. Boulanger-Scemama E, El Shamieh S, Démontant V, et al. Next-generation sequencing applied to a large French cone and cone-rod dystrophy cohort: Mutation spectrum and new genotype-phenotype correlation. *Orphanet J Rare Dis* 2015;10:85.
26. Bravo-Gil N, González-Del Pozo M, Martín-Sánchez M, et al. Unravelling the genetic basis of simplex retinitis pigmentosa cases. *Sci Rep* 2017;7:41937.
27. Bryant L, Lozynska O, Maguire AM, et al. Prescreening whole exome sequencing results from patients with retinal degeneration for variants in genes associated with retinal degeneration. *Clin Ophthalmol* 2018;12:49–63.
28. Michaelides M, Gaillard MC, Escher P, et al. The PROM1 mutation p.R373C causes an autosomal dominant bull's eye maculopathy associated with rod, rod-cone, and macular dystrophy. *Investig Ophthalmol Vis Sci* 2010;51(9): 4771–4780.
29. Pras E, Abu A, Rotenstreich Y, et al. Cone-rod dystrophy and a frameshift mutation in the PROM1 gene. *Mol Vis* 2009;15: 1709–1716.

30. Maw MA, Corbeil D, Koch J, et al. A frameshift mutation in prominin (mouse)-like 1 causes human retinal degeneration. *Hum Mol Genet* 2000;9(1):27–34.

31. Tiwari A, Bahr A, Bähr L, et al. Next generation sequencing based identification of disease-associated mutations in Swiss patients with retinal dystrophies. *Sci Rep* 2016;6:28755.

32. Patel N, Aldahmesh MA, Alkuraya H, et al. Expanding the clinical, allelic, and locus heterogeneity of retinal dystrophies. *Genet Med* 2016;18(6):554–562.

33. Sharon D, Kimchi A, Rivolta C. OR2W3 sequence variants are unlikely to cause inherited retinal diseases. *Ophthalmic Genet* 2016;37(4):366–368.

34. Zhao L, Wang F, Wang H, et al. Next-generation sequencing-based molecular diagnosis of 82 retinitis pigmentosa probands from Northern Ireland. *Hum Genet* 2015;134(2):217–230.

35. Carss K, Arno G, Erwood M, et al. Comprehensive rare variant analysis via whole-genome sequencing to determine the molecular pathology of inherited retinal disease. *Am J Hum Genet* 2017;100(1):75–90.

Supplementary data

Supplementary Figure 1. Molecular strategy followed for the screening of *PROM1* in our cohort of patients with inherited retinal dystrophy. [a] The daughter of MD-0873 presented a different phenotype and was also screened to discard a second allele in the *PROM1* gene. Abbreviations: N, number of families; NGS, next-generation sequencing; CNV, copy number variation; ar, autosomal reccesive; ad, autosomal dominant; VUS, variant of unknown significance.
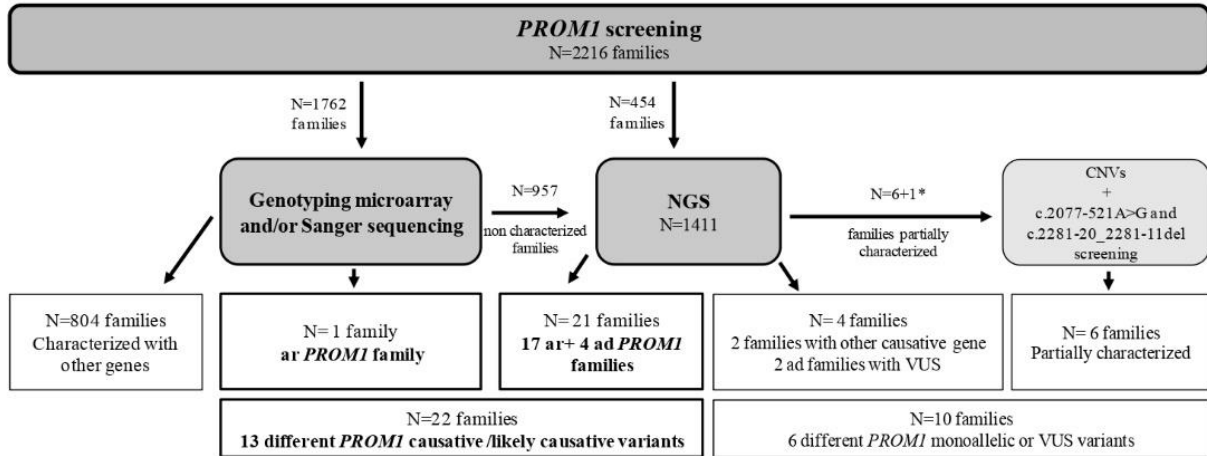
Supplementary Figure 2. Pedigrees of the 31 families with variants in the *PROM1* gene. Families with the c.303+1G>A, c.1354dup and the c.1984-1G>T variant showed a shared region of 17 Kb, 99 Kb, and 467 Kb, respectively. Family MD-1074 is not shown. Abbreviations: ad, autosomal dominant; ar, autosomal recessive; CRD, cone-rod dystrophy; LCA, Leber congenital amaurosis; VUS, variant of unknown significance; m, mutation allele; wt, wild type allele.

ad families with causative variants
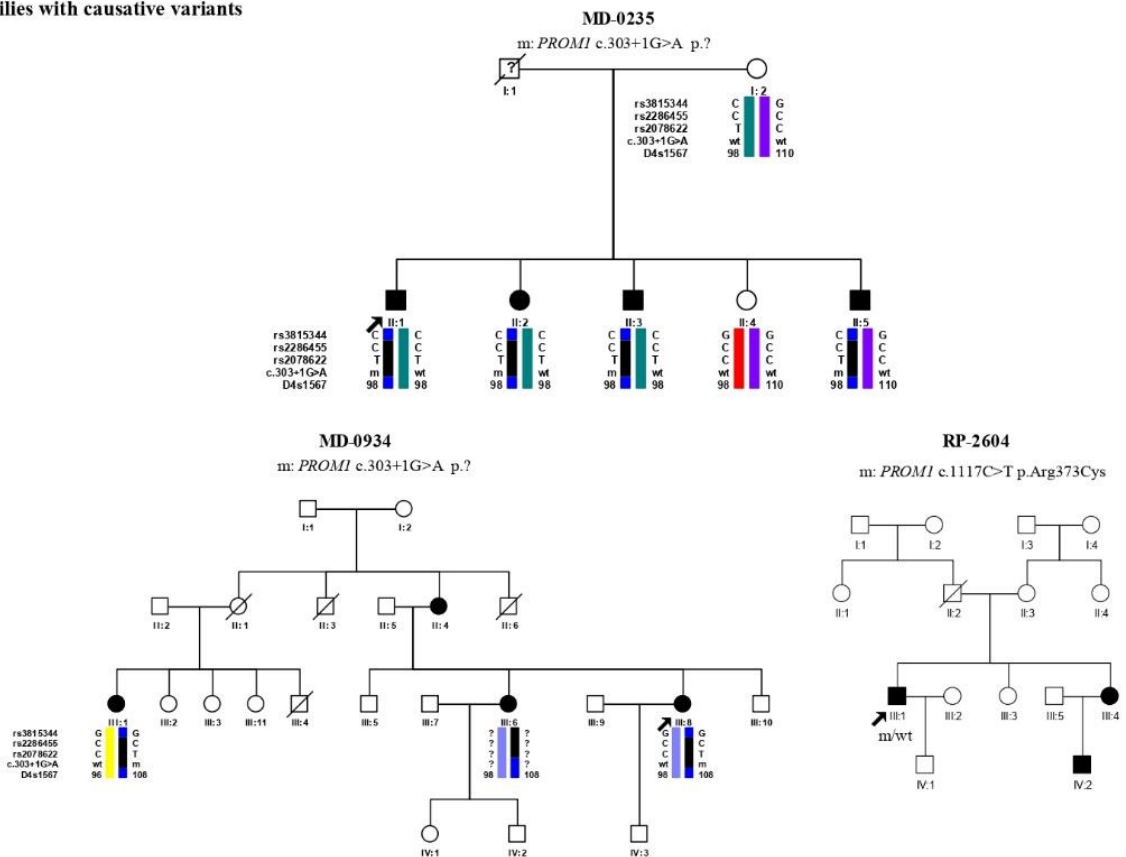
Supplementary Figure 2 (continued). ar families with the c.1354dup variant
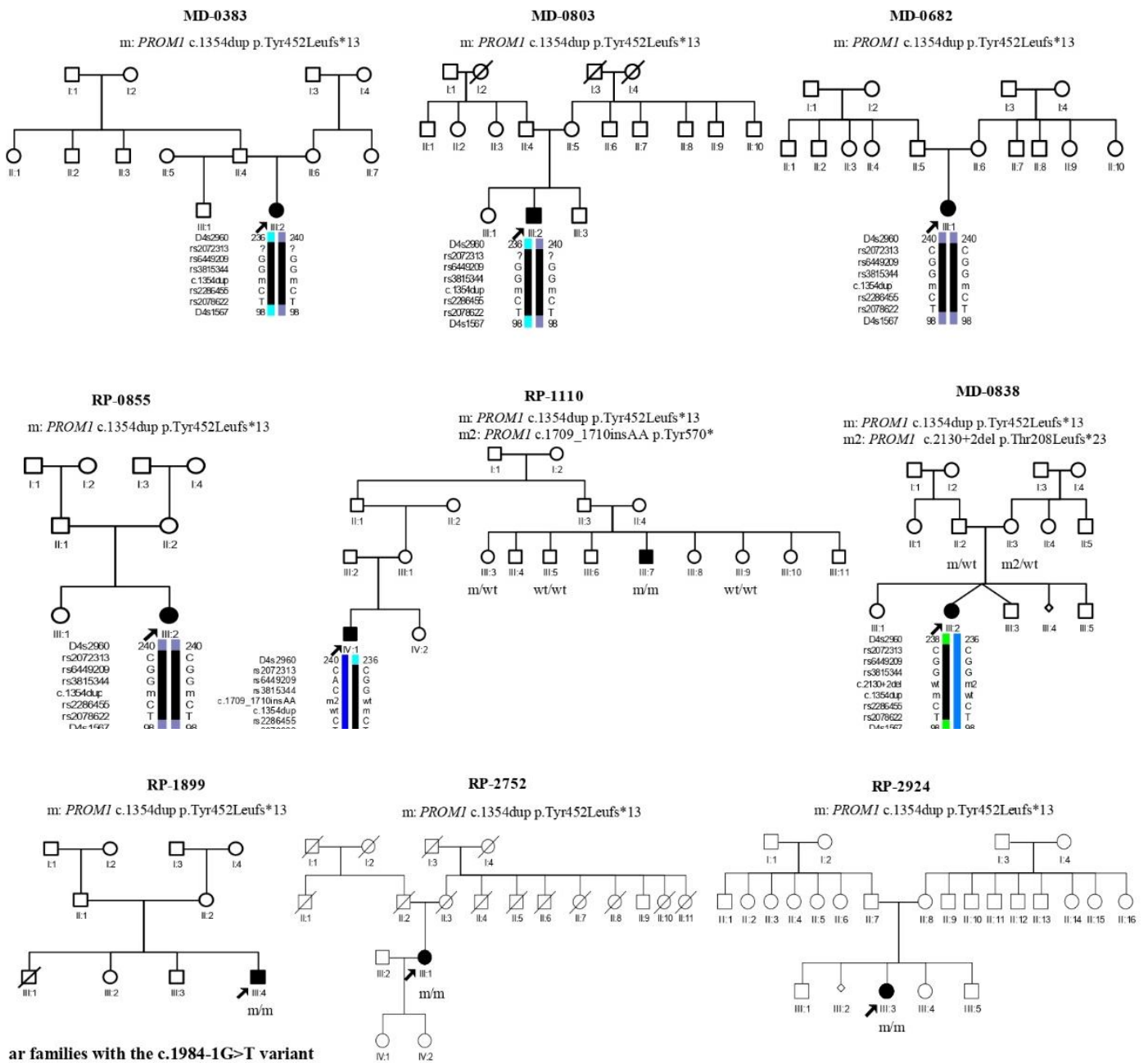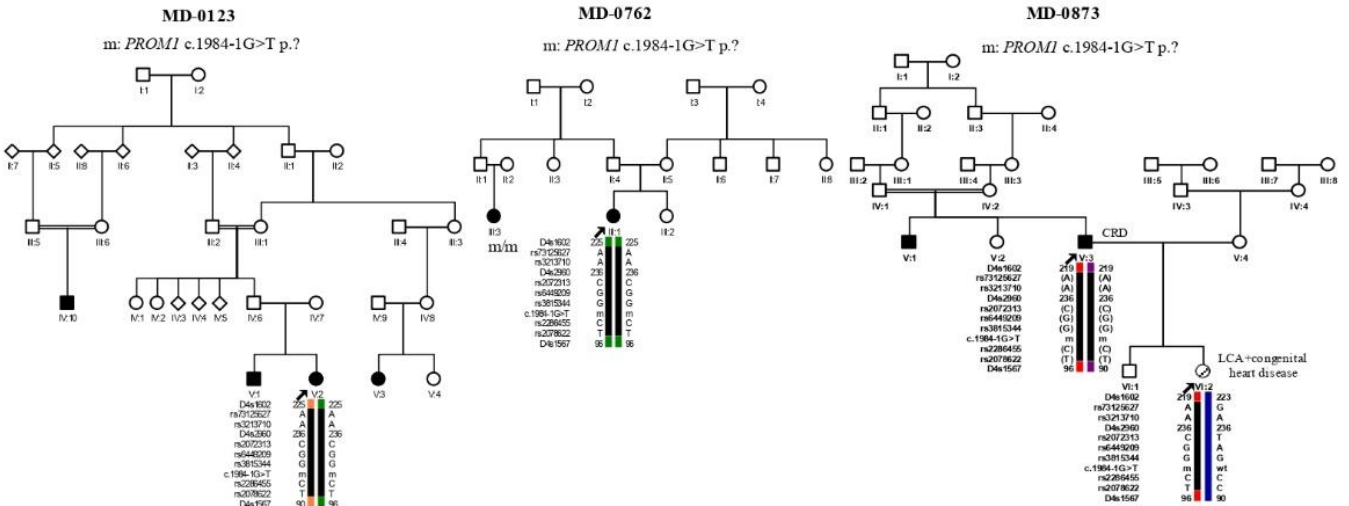
MD-0383

m: *PROM1* c.1354dup p.Tyr452Leufs*13

MD-0803

m: *PROM1* c.1354dup p.Tyr452Leufs*13

MD-0682

m: *PROM1* c.1354dup p.Tyr452Leufs*13

RP-0855

m: *PROM1* c.1354dup p.Tyr452Leufs*13

RP-1110

m: *PROM1* c.1354dup p.Tyr452Leufs*13
m2: *PROM1* c.1709_1710insAA p.Tyr570*

MD-0838

m: *PROM1* c.1354dup p.Tyr452Leufs*13
m2: *PROM1* c.2130+2del p.Thr208Leufs*23

RP-1899

m: *PROM1* c.1354dup p.Tyr452Leufs*13

RP-2752

m: *PROM1* c.1354dup p.Tyr452Leufs*13

RP-2924

m: *PROM1* c.1354dup p.Tyr452Leufs*13

ar families with the c.1984-1G>T variant

MD-0123

m: *PROM1* c.1984-1G>T p.?

MD-0762

m: *PROM1* c.1984-1G>T p.?

MD-0873

m: *PROM1* c.1984-1G>T p.?

158

Supplementary Figure 2 (continued). ar families with causative and likely causative variants
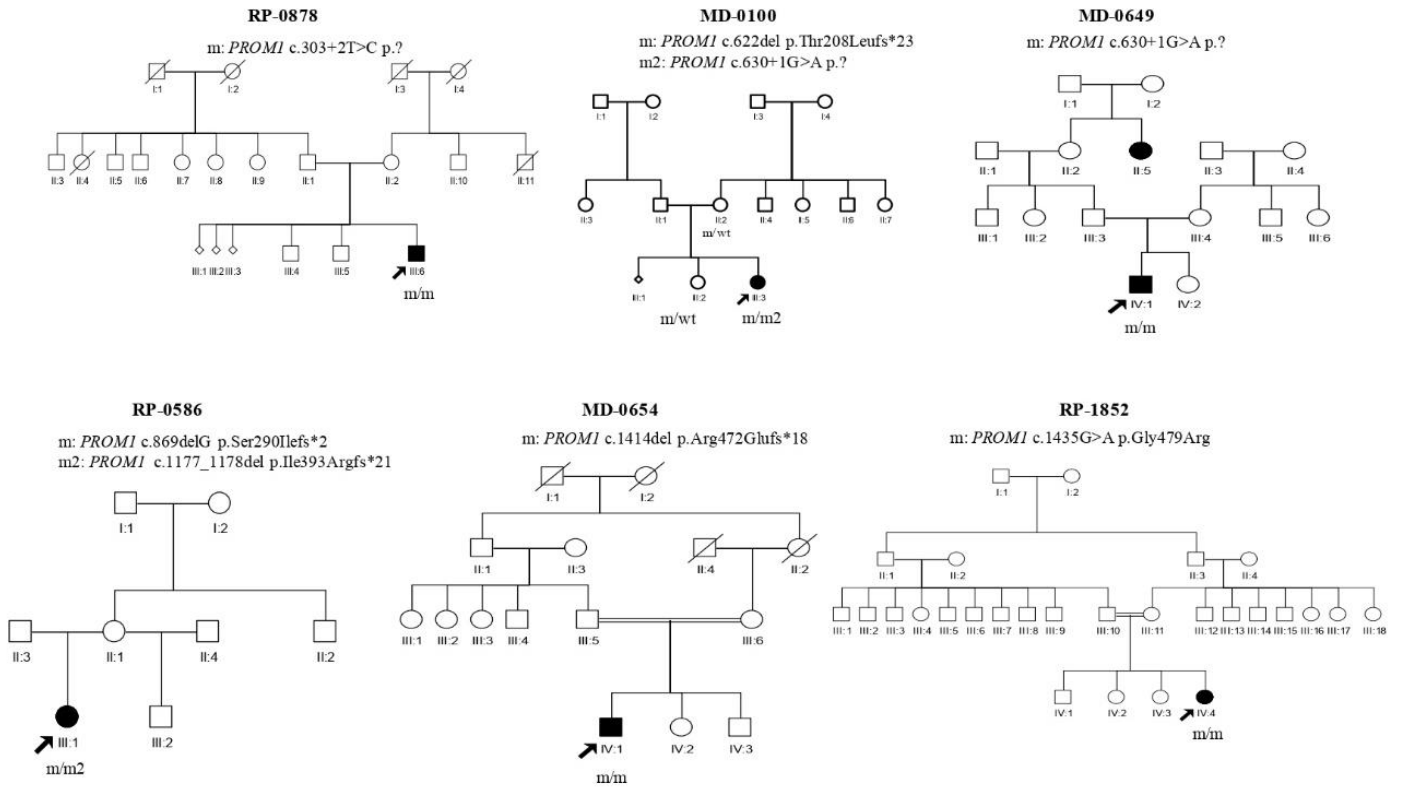
Supplementary Figure 2 (continued). Families with non causative variants
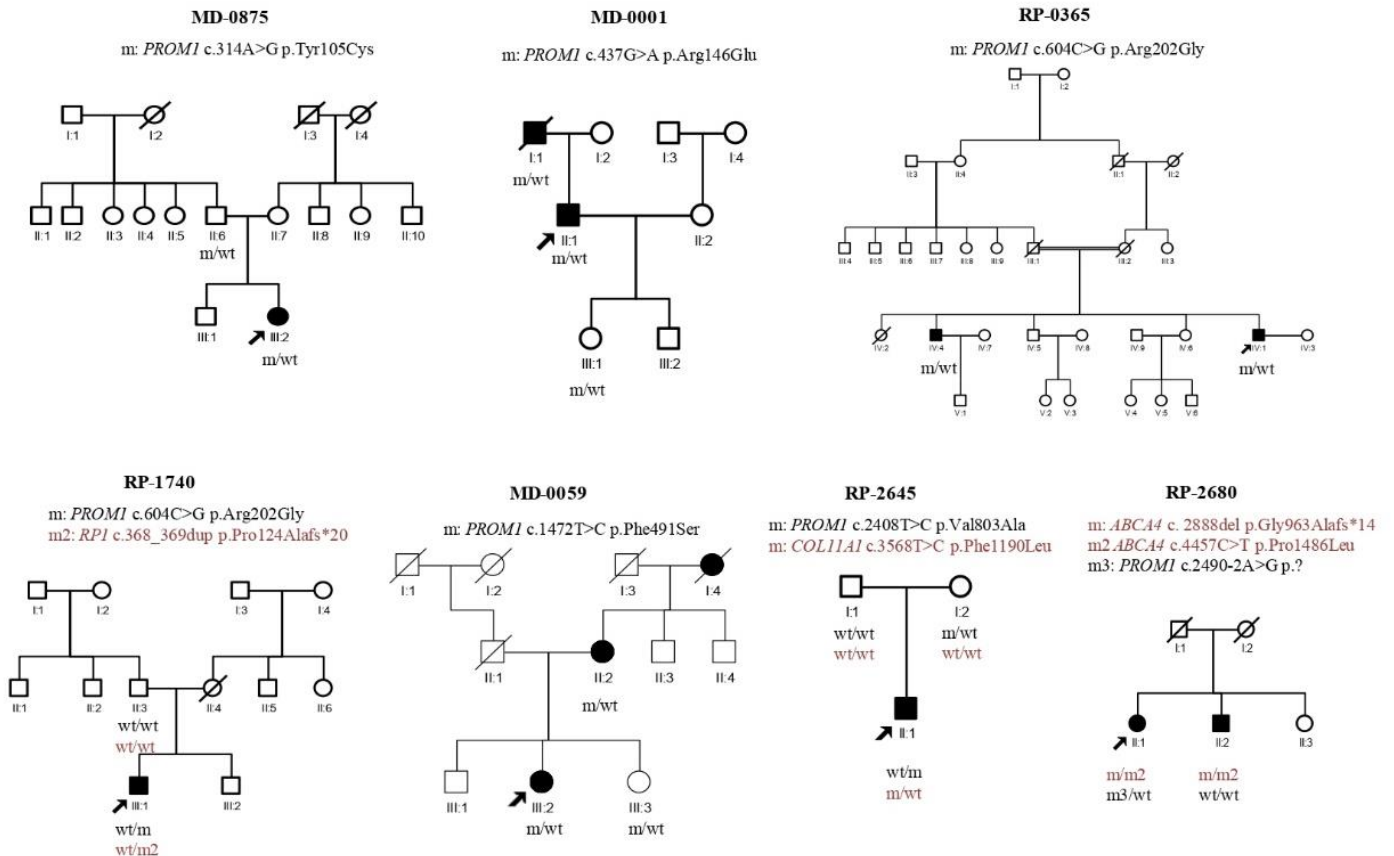
Supplementary Figure 2 (continued). Families incompletely characterized with a recessive variant

MD-0289

m: *PROM1* c.1354dup p.Tyr452Leufs*13

RP-1243

m: *PROM1* c.1354dup p.Tyr452Leufs*13

RP-2831

m: *PROM1* c.1354dup p.Tyr452Leufs*13
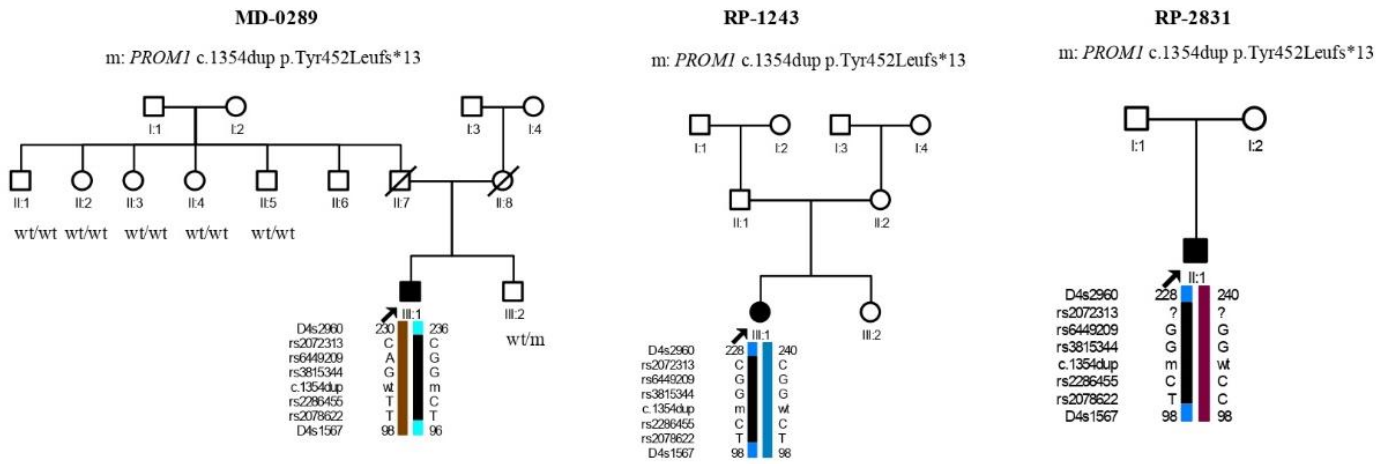
Supplementary Figure 3. Structure of the *PROM1* gene and location of the likely pathogenic/pathogenic variants found in this study. Above the figure: recessive mutations present in patients from this study. Below the figure: dominant mutations present in patients from this study. Bold: novel mutations identified in this study. Purple: coding sequence of protein PROM1 transmembrane domains. [a]Variant of unknown significance likely causative. Abbreviations: ar, autosomal recessive; ad, autosomal dominant; cds, coding sequence.
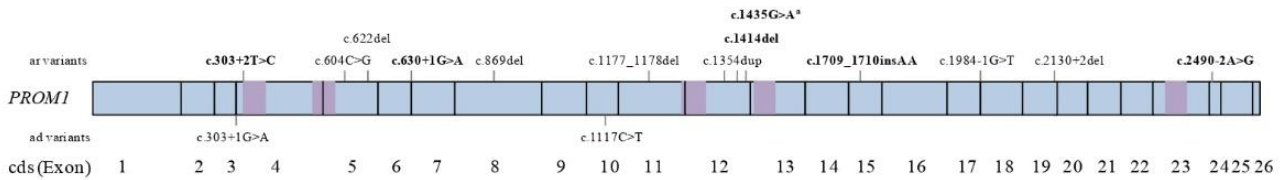
Supplementary Figure 4. *PROM1* expression in available tissues obtained from controls and patients with *PROM1* variants c.303+1G>A and c.303+2T>C.
A: *PROM1* is expressed in whole blood, blood mononuclear cells, saliva, and fibroblasts. B: Sequence analysis of DNA obtained from a control, a patient with the variant c.303+1G>A in heterozygosis and a patient with the variant c.303+2T>C in homozygosis. C: The sequence analysis of cDNA obtained from the tissues presented in Figure A in controls, patients with the variants c.303+1G>A or c.303+2T>C in *PROM1* revealed that the exon 3 is not present in the isoforms expressed in these tissues.
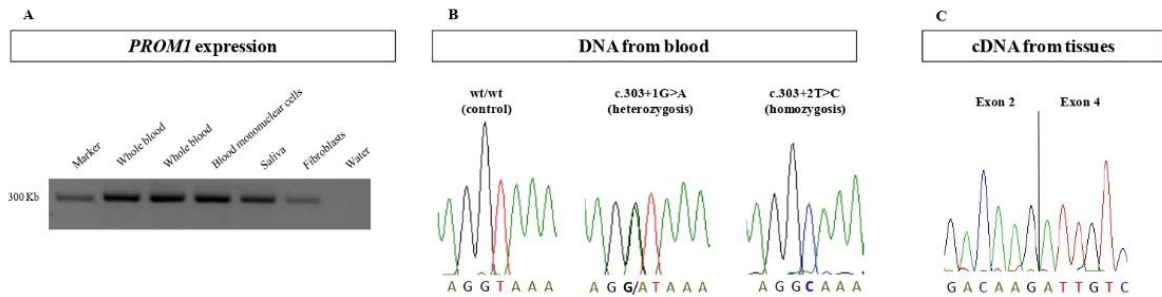
Supplementary Figure 5. Fundus angiography (FA) and spectral domain optical coherence tomography (SD-OCT) images showing the choroidal neovascular membrane developed in the right eye of the patient III:1 of family MD-0934, at 63 y.
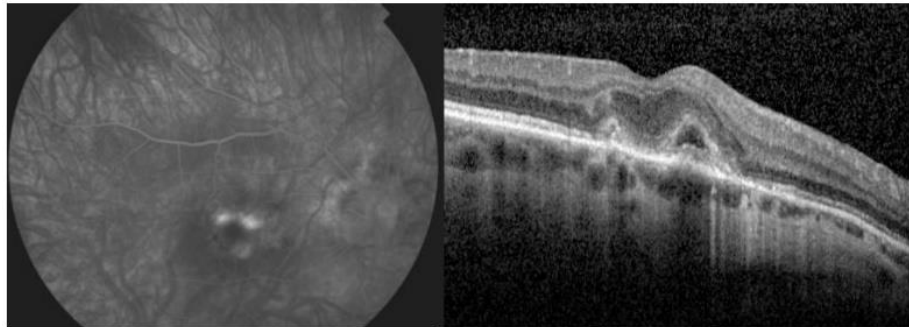
Supplementary Table 1. *In-sílico* predictions of the novel *PROM1* variants identified in this study. ExAC, Exome Agregation Consortium; MAF, minor allele frequency; VUS, variant of unknown significance.

Missense, frameshift, and stop variants								
Family_ID	cDNA change	Protein change	ExAC (MAF)	SIFT [0–1]	Polyphen [0–1]	Mutation Taster	CADD phred [0–40]	Classification in this study
MD-0875	c.314A>G	p.Tyr105Cys	0.000022	Deletereous (score 0)	Probably damaging (1)	Disease causing	25.1	VUS
MD-0001	c.437G>A	p.Arg146Gln	0.000077	Tolerated (score 0.19)	Probably damaging (1)	Disease causing	35	VUS
MD-0654	c.1414del	p.Arg472Glufs*18	-	-	-	-	23.9	Likely pathogenic
RP-1852	c.1435G>A	p.Gly479Arg	0.0000092	Deletereous (score 0)	Probably damaging (1)	Disease causing	33	Likely pathogenic
MD-0059	c.1472T>C	p.Phe491Ser	-	Deletereous (score 0.02)	Probably damaging (1)	Disease causing	34	VUS
RP-1110	c.1709_1710insAA	p.Tyr570*	-	-	-	-	28.2	Likely pathogenic
RP-2645	c.2408T>C	p.Val803Ala	-	Deletereous (score 0.03)	Probably damaging (0.975)	Polymorphism	23.3	VUS

Splicing variants								
Family_ID	cDNA change	Predicted Effect	ExAC (MAF)	SSF [0–100]	MaxEnt [0-12]	NNSPLICE [0-1]	GeneSplicer [0-15]	HSF [0-100]
RP-0878	c.303+2T>C	Broken donor site	-	82.5→79.3	9.1 →0	0.8 →0	3.5 →0	84.9 →0
MD-0100 MD-0649	c.630+1G>A	Broken donor site	-	79.5 →0	7.3 →0	1 →0	4.7 →0	83.7 →0
RP-2680	c.2490-2A>G	Broken acceptor site	-	85.8→0	10 →0	1 →0	4.5 →0	78.3 →0

Supplementary Table 2. Variants found in other genes in our cohort of patients with *PROM1* variants.

Family	Gene	Transcript	Allele 1				Allele 2			
			Exon/Intron	Nucleotide	Protein	Reference	Exon/Intron	Nucleotide	Protein	Reference
RP-1740	*RP1*	NM_006269.1	2	c.368_369dup	p.Pro124Alafs*20	(1)				
RP-2680	*ABCA4*	NM_000350	19	c.2888del	p.Gly963Alafs*14	(2)	30	c.4457C>T	p.Pro1436Leu	(3)
RP-2645	*COL11A1*	NM_001854.3	46	c.3568T>C	p.Phe1190Leu	Novel				

References:

1. Neveling K, Collin RWJ, Gilissen C, et al. Next-generation genetic testing for retinitis pigmentosa. *Hum Mutat*. 2012;33(6):963-972.

2. Paloma E, Martínez-Mir A, Vilageliu L, Gonzàlez-Duarte R, Balcells S. Spectrum of ABCA4 (ABCR) gene mutations in Spanish patients with autosomal recessive macular dystrophies. *Hum Mutat*. 2001;17(6):504-510.

3. Lewis RA, Shroyer NF, Singh N, et al. Genotype/Phenotype analysis of a photoreceptor-specific ATP-binding cassette transporter gene, ABCR, in Stargardt disease. *Am J Hum Genet*. 1999;64(2):422-434.

Supplementary Table 3. Clinical Data of 25 patients from 22 families. Abbreviations: Hom, homozygous; Het, heterozygous; CF, counting fingers; NB, night blindness; VF, visual field; VA, visual acuity; BCVA, best corrected visual acuity; OD, right eye; OS, left eye; LP, light perception; NA, not available; y, years; EOG, electrooculogram; ERG, electroretinogram; ISe, inner segment ellipsoid; Typical RP fundus: pale optic disc, narrowed retina vessels and pigmentary changes (bone-spicules); RPE, retinal pigment epithelium; HM, Hand motion.

Family ID	Patient ID	*PROM1* mutation	Symptoms onset (age in years)				BCVA OD/OS (decimal)	Fundus Examination	VF	ERG	Other findings
			Diagnosis	NB	VF loss	VA loss					
RP patients											
RP-0878	III:6	c.303+2T>C (Hom)	26	18	26	32	LP/0.2 (41y)	Typical RP and atrophic maculopathy (41y)	10° (41y)	Rod and cone responses estinguished (41y)	Posterior subcapsular cataract.
RP-2604	III:1	c.1117C>T (Het)	33	30	30	32	0.1/0.05 (57y)	Typical RP, ISe and RPE atrophy. Epiretinal membrane (57y)	Tubular and patchy central scotoma (57y)	Rod and cone responses estinguished (57y)	Cataract surgery (56y). Cystoid macular edema (OS)
CRD and MD patients with the c.303+1G>A variant											
MD-0235	II:1	c.303+1G>A (Het)	65	No (56y)	50	40	0.7/0.2 (56y)	Atrophic maculopathy (56y)	Central scotoma (56y)	Normal (56y)	Normal EOG (50y) Dischromatopsia. Foveal sparing
MD-0934	III:8	c.303+1G>A (Het)	45	38	No (49y)	34	0.8/0.7 (51y)	Atrophic maculopathy (51y)	NA	Normal (48y)	Dischromatopsia
MD-0934	III:6	c.303+1G>A (Het)	54	54	No (54y)	54	1/1 (52y)	Atrophic maculopathy (54y)	NA	Normal (51y)	
MD-0934	III:1	c.303+1G>A (Het)	60	No (61y)	No (61y)	59	1/0.05 (62y)	RPE changes (OD) and atrophic maculopathy (OS) (62y)	Central scotoma (OS) (62y)	Rod and cone responses slighty reduced (62y)	Cataract. Foveal sparing
MD-1074	NA	c.303+1G>A (Het)	4	NA (6y)	No (6y)	NA (6y)	1/1 (6y)	Normal (6y)	NA	Rod and cone responses slighty reduced (6y)	Dischromatopsia
CRD patients with causative variants											
MD-0100	III:3	c.622del (Het) c.630+1G>A (Het)	14	18	14	4	0.05/0.05 (23y)	Typical RP and bull's eye maculopathy (20y)	Tubular 5° (20y)	NA	Dischromatopsia. Nystagmus. Photophobia
MD-0649	IV:1	c.630+1G>A (Hom)	16	16	16	7	NA	Bull's eye maculopathy (16y)	NA	NA	Dischromatopsia. Photophobia
RP-0586	III:1	c.869delG (Het) c. 1177_1178del (Het)	12	15	13	6	0.4/0.5 (18y)	Typical RP and bull's eye maculopathy (18y)	Absolute central and relative peripheral scotoma (18y)	Rod and cone responses estinguished (18y)	Dischromatopsia
MD-0383	III:2	c.1354dup (Hom)	3	Yes	No	Yes	0.5/0.5 (20y)	Bull's eye maculopathy and RPE atrophy (20y)	30° (20y)	Rod response reduced, cone response estingused (20y)	
MD-0803	III:2	c.1354dup (Hom)	9	10	No (9y)	8	0.4/0.3 (13y)	Bull's eye maculopathy (13y)	Central scotoma (13y)	Rod and cone responses reduced (13y)	Strabismus
RP-1899	III:4	c.1354dup (Hom)	28	No (39y)	20	23	NA	NA	NA	NA	
RP-0855	III:2	c.1354dup (Hom)	13	No (42y)	7	7	CF/CF (23y)	Typical RP and atrophic maculopathy (23y)	Tubular (23y)	Rod and cone responses severely reduced (23y)	
MD-0682	III:1	c.1354dup (Hom)	20	25	20	15	0.2/0.25 (20y)	Atrophic maculopathy (23y)	Generalized absolute field constriction (24y)	Rod response reduced and cone response estinguished (23y)	
RP-2752	III:1	c.1354dup (Hom)	25	NA	14	7	<0.1/<0.1 (48y)	Typical RP (48y)	NA	NA	
RP-2924	III:3	c.1354dup (Hom)	18	54	10	8	0.1/0.1 (39y)	Typical RP and atrophic maculopathy (60y)	Relative central and absolute peripheral scotoma (39y)	NA	Dischromatopsia. Nystagmus (OD)
RP-1110	III:7	c.1354dup (Hom)	NA	NA	NA	NA	LP/LP (39y)	Myopic chorioretinal degeneration, Typical RP and macular atrophy (39y)	NA	NA	Cataract. Myopia magna. Nystagmus
RP-1110	IV:1	c.1354dup (Het) c.1709_1710insAA (Het)	18	25	19	19	LP/LP (52y)	RP inversa (52y)	Generalized absolute field constriction (52y)	NA	Cataract. Nystagmus
MD-0838	III:2	c.1354dup (Het) c.2130+2del (Het)	8	10	No (9y)	4	0.6/0.5 (11y)	ISe and RPE atrophy in the macula (11y)	Central scotoma (11y)	Rod response reduced and cone response estinguished (11y)	Pachychoroid
MD-0654	IV:1	c.1414del (Hom)	21	32	13	13	HM/CF (35y)	Typical RP (limited pigmentation) and atrophic maculopathy (35y)	<10° (35y)	NA	Nystagmus
MD-0123	V:2	c.1984-1G>T (Hom)	26	NA	11	6	0.5/0.3 (26y)	RPE atrophy and atrophic maculopathy (26y)	Central scotoma (26y)	Rod and cone responses reduced (26y)	
MD-0762	III:1	c.1984-1G>T (Hom)	22	25	19	20	0.05/0.05(25y)	RPE atrophy and yellow-white lesions in the macula (25y)	NA	NA	Dyschromatopsia
MD-0873	V:3	c.1984-1G>T (Hom)	18	NA	6	6	CF/CF (32y)	Narrowed vessels and atrophic maculopathy (32y)	Central scotoma (32y)	Rod and cone responses reduced (32y)	Dyschromatopsia. Cataract (OS). Fuchs heterochromic iridocyclitis (OS)
CRD patient with likely causative variants											
RP-1852	IV:4	c.1435G>A (Hom)	40	No (59y)	40	23	CF/CF (57y)	Typical RP and atrophic maculopathy (57y)	30° (57y)	Rod and cone responses estinguished (57y)	

Discusión general

A pesar de la gran cantidad de estudios publicados sobre el gen *ABCA4*, aún se desconocen muchos aspectos sobre este gen. Por un lado, muchos de los pacientes diagnosticados de enfermedad de Stargardt no conocen su causa molecular, y, además, se han asociado diferentes distrofias de retina a este mismo gen. Por esta razón, en este trabajo se intenta caracterizar a estos pacientes mediante la búsqueda de nuevas variantes, y se profundiza en el fenotipo asociado a *ABCA4* y *PROM1*, otro gen relacionado con distrofias de retina.

I. Caracterización genética de pacientes con sospecha de enfermedad de Stargardt

Estudios en pacientes no resueltos

Como se ha descrito anteriormente, el espectro genético de las DHR, y de las DM es muy heterogéneo (Michaelides, Hunt and Moore, 2003). Además, gracias a las nuevas técnicas de NGS se incrementa el diagnóstico genético de estas enfermedades (Eisenberger *et al.*, 2013; Bernardis *et al.*, 2016). Previamente al desarrollo de esta Tesis Doctoral, el número de pacientes que había sido estudiado con sospecha de enfermedad de Stargardt u otras maculopatías de herencia recesiva era 802, y se habían identificado variantes en el gen *ABCA4* en 357 (44%) pacientes, de los cuales 292 (81%) tenían sospecha de STGD1. Además, 75 pacientes presentaban un único alelo mutado en el gen *ABCA4*.

A finales del año 2019 había un total de 945 familias con distrofia macular y herencia recesiva, de las cuales 553 se remitieron con sospecha de Stargardt, y 392 con sospecha de otras maculopatías. Con el desarrollo de esta Tesis Doctoral, la implementación de la secuenciación masiva en el laboratorio, y el estudio de regiones no codificantes, se habían caracterizado el 77% (427/945) y el 49% (195/392) de las familias de cada sospecha clínica, respectivamente (**Artículo 1**). Además, en 46 (73%) y en 5 (41%) de los casos monoalélicos diagnosticados de Stargardt u otra maculopatía, respectivamente, se identificó una segunda variante en el gen *ABCA4*. De las familias con sospecha de Stargardt, 407 (73%) presentaban variantes bialélicas en el gen *ABCA4*, mientras que 20 (4%) tenían variantes en otros genes. De los casos remitidos con sospecha de otras maculopatías, en 100 casos (26%) se identificaron variantes en *ABCA4*, que fueron reclasificados a STGD1 o DCB en el **Artículo 4** (Del Pozo-Valero *et al.*, 2020).

En algunos pacientes con diagnóstico de Stargardt no se identifica su causa molecular (Zernant *et al.*, 2011), lo que se conoce como "heredabilidad perdida". La mayoría de los casos resueltos durante este trabajo, y principalmente los casos monoalélicos, presentaban variantes en regiones codificantes del gen *ABCA4*. La causa por la que no habían sido previamente identificadas reside en la utilización de diferentes tecnologías a lo largo de los 29 años en los que se han estudiado estos pacientes en la Fundación Jiménez Díaz. Inicialmente se utilizaba un microarray de genotipado de mutaciones concretas en el gen *ABCA4* (Valverde *et al.*, 2006). Posteriormente, se utilizaron otras técnicas que permitían identificar más variantes en este gen (Valverde *et al.*, 2007;

Aguirre-Lamban *et al.*, 2009, 2010; Riveiro-Alvarez *et al.*, 2013). Actualmente, las tecnologías de secuenciación masiva permiten la secuenciación de la región codificante y las regiones intrónicas flanqueantes de *ABCA4*.

Además, otras variantes resultan difíciles de identificar. Para detectar las variaciones en el número de copias utilizando NGS se necesitan herramientas bioinformáticas posteriores a la secuenciación (Yamamoto *et al.*, 2016; Gross *et al.*, 2019), y, además, en el gen *ABCA4* son eventos raros (Zernant *et al.*, 2014). En 4 de las familias monoalélicas previamente no resueltas se identificó una deleción de 411 pb que incluye parte del exón e intrón 6 del gen *ABCA4*. Esta deleción también fue identificada en 2 familias adicionales de la cohorte total española (**Artículo 4**).

Aparte de estas limitaciones técnicas, también existen variantes codificantes que no se consideraban patogénicas en el momento del análisis y que actualmente se han reclasificado o se consideran hipomórficas. La variante c.5603A>G; p.(Asn1868Ile), previamente considerada una variante benigna por su elevada frecuencia, tiene penetrancia reducida y sería patogénica cuando se encuentra en *trans* con una variante severa, o con algún elemento en *cis* (Zernant *et al.*, 2017; Runhart *et al.*, 2018). En este trabajo, 9 pacientes presentan esta variante, y solo en 2 se ha identificado otra variante en *cis* en *ABCA4*. Estos pacientes se consideran caracterizados, siendo necesario realizar estudios futuros para determinar qué otros elementos en *cis* pueden explicar la penetrancia de esta variante.

Por otra parte, la región no codificante de *ABCA4* sería la región candidata para el estudio de los pacientes con sospecha de STGD no resueltos. Debido al estudio de estas regiones y a la realización de estudios funcionales para establecer la patogenicidad de las variantes encontradas, se caracterizaron 19 de los pacientes monoalélicos con el gen *ABCA4* como causal (**Artículo 1**)*.* Las variantes estudiadas funcionalmente durante esta Tesis Doctoral incluyen variantes situadas en regiones no canónicas de *splicing* (en regiones codificantes y en intrones, NCSS), y variantes *deep intronic* que habían sido previamente publicadas, o se habían identificado en pacientes no resueltos por primera vez. En el **Artículo 2** (Fadaie *et al.*, 2019) se describen defectos en el patrón de *splicing* en 12 de las 19 variantes estudiadas (63%). Las variantes NCSS producían la pérdida de 1 o 2 exones en 6 casos, y la pérdida de parte del exón en 2 casos, mientras que las variantes *deep intronic* y cercanas al exón (NEAR) provocan la inserción de pseudoexones en 3 casos, y la elongación del exón en el caso de la variante c.4539+61G>A. En el **Artículo 3** (Khan *et al.*, 2020), en el que se estudiaron 1054 pacientes no resueltos con sospecha de Stargardt (80 de ellos corresponden a los procedentes de nuestra cohorte española), se analizaron 9 variantes NCSS y 58 *deep intronic*, de las cuales las 9 NCSS y 13 de las *deep intronic* mostraban defectos en el patrón de *splicing*. Con estos trabajos, actualmente se conocen 35 variantes *deep intronic* y 60

NCSS en las que se ha demostrado un defecto en el patrón de *splicing* del gen *ABCA4*, y, por tanto, se consideran variantes patogénicas (Anexo I).

Por otra parte, el estudio del gen *ABCA4* completo en el **Artículo 3** (Khan *et al.*, 2020) permitió identificar la causa molecular de la enfermedad en el 55% (44/80) de los pacientes españoles que se incluyeron. El diagnóstico *a priori* de estos pacientes era enfermedad de Stargardt, sin embargo, no disponemos de historias clínicas y oftalmológicas de todos los pacientes ya que el HUFJD es un centro de referencia en distrofias de retina y muchos pacientes proceden de otros centros. Es posible que estos pacientes tengan otro tipo de DHR con algunos aspectos similares a STGD1, y por ello hayan sido erróneamente diagnosticados de STGD1, ya que es la maculopatía más frecuente. Varios genes que originan características clínicas similares y otros, aún no descubiertos, podrían explicar su fenotipo. De hecho, 7 de los 44 pacientes no resueltos están incluidos en los 20 casos remitidos con sospecha de Stargardt (4% del total), en los que se identificaron mutaciones en otros genes al final de este estudio (**Artículo 1**). De ellos, algunos genes pueden dar lugar a una clínica similar (*PLA2G5, PRPH2* y *PROM1*) (Kniazeva *et al.*, 1999; Boon *et al.*, 2007; Sergouniotis *et al.*, 2011) , pero otros se alejan bastante de este fenotipo (*BBS1, OPA1* y *WFS1*) (Dhalla, Desai and Zuckerbrod, 2006; Estrada-Cuzcano *et al.*, 2012; Nasca *et al.*, 2017). Por ello, en los casos negativos, sería necesaria una reevaluación clínica del fenotipo para identificar la causa molecular de su enfermedad. Un estudio reciente que involucra a pacientes con enfermedades compatibles con *ABCA4* (*ABCA4-disease*), sin variantes en este gen, también identifica otros genes como probablemente causales (*PRPH2, PROM1, CDHR1, CERKL, CRX* y *RPE65)* (Wolock *et al.*, 2019).

En los casos en los que el fenotipo es compatible con Stargardt y continúan no resueltos a pesar de haberles cribado el gen completo hay que destacar que se han estudiado las variantes no codificantes en las que en al menos dos de los predictores de *splicing* se mostraban alteraciones, como se ha descrito previamente (Sangermano *et al.*, 2018).

En nuestra serie se estudiaron 12 variantes (**Artículo 3**), algunas de las cuales no cumplían todos los criterios anteriormente descritos, cuyos resultados fueron no concluyentes (Anexo II), probablemente debido al tipo celular en el que se realizó el estudio. Estos estudios se realizaron en células HEK293T, derivadas de la línea celular de células embrionarias de hígado HEK293, por ser fácilmente transfectables. Sin embargo, en el contexto de la retina no se puede saber si sucedería lo mismo. De hecho, el uso de células retina-*like,* en lugar de células no retinianas, ha permitido ver diferencias en ensayos de *splicing* (Rivera *et al.*, 2000).

La utilización de la estrategia de midigenes permite disponer de construcciones con un contexto genómico adecuado (intrones completos y exones adyacentes), y ha demostrado el defecto de *splicing* de variantes de *ABCA4* que no se observaban utilizando minigenes más pequeños

(Sangermano *et al.*, 2018). Otro método utilizando células precursoras de fotorreceptores derivadas de pacientes (Sangermano *et al.*, 2016), ha permitido estudiar el papel de otras variantes intrónicas en el gen *ABCA4,* cuyo efecto no se apreciaba claramente al utilizar los fibroblastos del paciente sin transformar (Albert *et al.*, 2018), o las células HEK293T (Sangermano *et al.*, 2019). Esto sugiere que sería apropiado utilizar células de retina para realizar estos estudios. Los fotorreceptores expresan un programa específico de *splicing* e isoformas específicas (Murphy *et al.*, 2016), y podrían existir motivos de *splicing* específicos de los fotorreceptores que están ausentes en las células HEK293T. Por este motivo, el reestudio de variantes con resultados no concluyentes en otro tipo celular, de retina o incluso en células derivadas del paciente, podría demostrar defectos en el *splicing* que no se observaron con las células HEK293T, y sería recomendable sobre todo en los pacientes con una sospecha clara de STGD y monoalélicos en *ABCA4*.

Otra de las posibilidades sería que variantes en las que aparentemente no se predice alteración del *splicing* utilizando programas de predicción, y que no se ensayan funcionalmente, estuvieran provocando defectos en este mecanismo. Aparte de los elementos *cis* situados en la secuencia del ARNm prematuro, los factores *trans* que se unen a él para promover o inhibir el *splicing* también son importantes. Para ayudar a priorizar estas variantes conociendo los sitios de unión al ARN de las proteínas implicadas en el *splicing*, se creó el programa SpliceAid database con secuencias de ARN diana de estas proteínas obtenidas experimentalmente (Piva *et al.*, 2009, 2012).

Finalmente, la secuencia promotora de *ABCA4* no ha sido muy estudiada hasta la fecha (Zernant *et al.*, 2014). Se han identificado algunas variantes en regiones reguladoras de este gen que podrían afectar a la expresión de *ABCA4*, como la c.768+3223C>T y c.2919-383C>T (Bauwens *et al.*, 2019), aunque se necesitan más estudios para dilucidar su implicación en la enfermedad (Cremers *et al.*, 2020).

Por todas estas razones, no es posible asegurar que la causa genética en estos pacientes no esté relacionada con el gen *ABCA4*.

En otros estudios publicados de pacientes con STGD se reportan tasas de caracterización con variantes bialélicas del 47 al 76% de los casos analizados, mientras que se reporta caracterización incompleta (con una variante) del 10 al 46% de los casos (Fujinami *et al.*, 2013; Utz *et al.*, 2014; Jiang *et al.*, 2016; Schulz *et al.*, 2017; Nassisi *et al.*, 2018). Los estudios centrados en encontrar la segunda causa molecular en los pacientes STGD monoalélicos o no resueltos para *ABCA4* revelan tasas de caracterización variables: el 48% al secuenciar toda la región codificante y las regiones intrónicas adyacentes del gen (Zernant *et al.*, 2011); del 20% (Nassisi *et al.*, 2019) al 38% (Khan *et al.*, 2019) al incluir regiones intrónicas además de la región codificante; y el 42.5%

(**Artículo 3**) al incluir el 97% del locus de *ABCA4*. Estos resultados, similares a los realizados en nuestra cohorte española, apoyan las ideas reflejadas anteriormente.

Reclasificación de la patogenicidad de las variantes

El porcentaje de variantes nuevas que se identifican en el gen *ABCA4* en estudios con pacientes es similar al 22% de nuestro trabajo (Schulz *et al.*, 2017; Fujinami *et al.*, 2019; Khan *et al.*, 2019). Por esta razón, establecer la patogenicidad de estas nuevas variantes es importante a la hora de considerarlas la causa de la enfermedad, especialmente si tenemos en cuenta los ensayos clínicos que hay actualmente en los que se necesitan tener identificadas variantes bialélicas en este gen.

En un metaanálisis de 2017 (Cornelis *et al.*, 2017) se estableció la patogenicidad de 913 variantes en *ABCA4* recopiladas hasta el año 2016, teniendo en cuenta su frecuencia en la población general y en los pacientes publicados. En ese estudio, se observó que 30 variantes consideradas causales tenían una frecuencia poblacional superior a la frecuencia en la cohorte de pacientes, por lo que se reclasificaron como benignas. Una de ellas es la variante c.6148G>C; p.(Val2050Leu), que se encontraba en 7 de nuestros pacientes considerados como caracterizados. Por esta razón, en estos pacientes también se estudió el locus completo de *ABCA4* y en 6 de ellos se identificó una variante patogénica adicional. En el único caso en el que no se identificó la otra variante en *ABCA4* (familia RP-1875) se tendrán que realizar estudios adicionales para identificar el defecto molecular.

En el resto de los pacientes caracterizados, se hizo una revisión exhaustiva de las variantes que habían sido identificadas, y en los portadores de variantes nuevas, asimismo se realizaron estudios de segregación y/o una revisión clínica que apoyara el diagnóstico genético, de modo que estos datos se tuvieron en cuenta para considerarlas causales.

Espectro mutacional del gen *ABCA4* en España

En el **Artículo 4** (Del Pozo-Valero *et al.*, 2020) se describe la mayor cohorte de pacientes caracterizados con variantes en el gen *ABCA4* hasta la fecha (julio 2020), que corresponde a 506 pacientes con ascendencia española estudiados en el Servicio de Genética de la FJD. Este estudio contiene los casos caracterizados durante los 29 años desde el inicio del reclutamiento de esta cohorte. De ellos, también hay casos caracterizados durante el desarrollo de esta Tesis Doctoral, incluyendo los casos caracterizados en el **Artículo 3** (Khan *et al.*, 2020).

Desde la primera publicación de la asociación del gen *ABCA4* con la enfermedad de Stargardt en 1997 (Allikmets *et al.*, 1997) se han descrito más de 1200 variantes patogénicas en este gen (Cremers *et al.*, 2020), y se considera el gen más frecuentemente mutado en distrofias de retina. Debido a su gran tamaño existen muchas variantes, algunas de ellas difíciles de clasificar. En el presente trabajo, se identificaron en total 228 variantes diferentes en 1012 alelos mutados. El 56% de las variantes son *missense*, que corresponden al 65% de los alelos mutados, seguidas de las

variantes que generan codones de parada prematura *frameshift* y *stop*, que representan el 24% del total de variantes y el 17% de los alelos mutados. En otros estudios de cohortes en los que se analiza el gen *ABCA4* el porcentaje de variantes *missense* y *frameshift/stop* varía del 55 al 59% y del 15 al 30%, respectivamente (Jiang *et al.*, 2016; Schulz *et al.*, 2017; Nassisi *et al.*, 2018).

La mayoría de los estudios de grandes cohortes con pacientes STGD1 están orientados a su caracterización genética y se dan las tasas en función del porcentaje de los casos que se han resuelto. En este trabajo el enfoque es distinto ya que se utilizaron únicamente los casos resueltos para tener una visión general de las mutaciones frecuentes. En dos estudios de cohortes grandes de pacientes (Khan *et al.*, 2019, 2020) enfocados a encontrar la causa en los pacientes no resueltos, se encontraron mayoritariamente variantes no codificantes, reflejándose porcentajes del 15% y 12% de alelos portadores de variantes *deep intronic*, respectivamente. Sin embargo, estos casos no reflejan la frecuencia real de casos resueltos, ya que en la mayoría de ellos las regiones codificantes ya habían sido previamente analizadas, por lo que las variantes adicionales que podrían explicar el fenotipo están en las regiones no codificantes. Vistas en el contexto general, como el de nuestro estudio, solo el 1% de las variantes son *deep intronic*, y representan el 3% de los alelos mutados, por lo que podemos deducir que no son eventos frecuentes en población española y que lo común es encontrar variantes codificantes en estos pacientes.

De forma similar, las CNVs no son frecuentes en el gen *ABCA4* tampoco en nuestra cohorte, ya que sólo se identifican deleciones en 7 familias, y en 6 de ellas es la misma deleción de 411 pb, no descrita anteriormente e identificada utilizando diferentes tecnologías diagnósticas (MLPA, exoma clínico, panel de genes y smMIPs). Por esta razón, se realizaron estudios de haplotipos con el fin de determinar si estaban relacionadas ancestralmente y esta variante tenía un efecto fundador. En todas las familias se compartía una región de 1.58 Mb, y 3 de ellas compartían todos los marcadores (3.29 Mb). Además, en el **Artículo 3** (Khan *et al.*, 2020) se encontró la misma deleción en pacientes procedentes de Argentina, por lo que es una variante fundadora de población hispánica.

II. Análisis de las correlaciones genotipo-fenotipo en pacientes *ABCA4*

Descripción fenotípica de nuestra cohorte

El gen *ABCA4* ha sido asociado a varios fenotipos, incluyendo distrofias coriorretinianas (Bertelsen *et al.*, 2014; Tanaka *et al.*, 2018). Los fenotipos más comunes, aparte de la enfermedad de Stargardt son las distrofias de conos y las distrofias de conos y bastones, y también hay artículos en los que se reportan pacientes con RP que presentan variantes en este gen (Cremers *et al.*, 1998; Martínez-Mir *et al.*, 1998; Verbakel *et al.*, 2018). Respecto a este último fenotipo, actualmente se descarta que *ABCA4* sea causante de RP (Riveiro-Alvarez *et al.*, 2013). Así, en los casos descritos, se observa una maculopatía que progresa rápida y severamente, y que en edades avanzadas presenta aspectos típicos de RP (pigmentos o espículas) (Cremers *et al.*, 2020). Sin embargo, el curso de la enfermedad es el opuesto a la RP ya que la afectación primaria es central, y después avanza hacia la periferia, por lo que estos casos se consideran como DCB severa. En el **Artículo 4** (Del Pozo-Valero *et al.*, 2020) de esta Tesis Doctoral profundizamos en el fenotipo asociado de la mayor cohorte de pacientes que se ha descrito hasta la fecha.

Los pacientes fueron clasificados en 2 fenotipos, STGD1 y DCB. De los 506 pacientes, pudimos obtener datos oftalmológicos de 372, mientras que los 134 restantes se clasificaron según el diagnóstico de sospecha. El criterio principal para incluir a los pacientes en el grupo de DCB fue que presentaran síntomas de ceguera nocturna y/o un electrorretinograma con afectación de conos y de bastones. Los pacientes con afectación únicamente de conos fueron clasificados como STGD1. Los pacientes con un fenotipo STGD1 eran 434 mientras que 72 presentaban DCB, de los cuales teníamos información clínica y oftalmológica de 306 y 66, respectivamente.

Algunos pacientes presentaban una agudeza visual cercana a la unidad a la edad del diagnóstico, incluso en edades avanzadas (50 años). Se conoce como respeto foveal y se ha descrito previamente en casos con mutaciones en *ABCA4* (Rotenstreich, Fishman and Anderson, 2003; Fujinami *et al.*, 2011; Westeneng-van Haaften *et al.*, 2012; Fujinami *et al.*, 2013). Esta característica consiste en que los pacientes preservan una zona (fóvea) en la que la agudeza visual es óptima mientras que el resto de la mácula está atrófica. Debido a esto, los pacientes no presentan síntomas y algunos se diagnostican incidentalmente en una consulta en la que les observan el fondo de ojo. Solo se ha reportado un estudio en el que se realiza una correlación entre las variantes de *ABCA4* en estos pacientes con respeto foveal, y únicamente se observó que la variante c.6089G>A; p.(Arg2030Gln) podría estar asociada con esta característica clínica (Fujinami *et al.*, 2013), la cual no se ha identificado en ninguno de nuestros 506 pacientes. En cambio, la variante sobrerrepresentada en nuestros pacientes con este fenotipo es la variante española c.3386G>T; p.(Arg1129Leu), que se comentará en el siguiente apartado.

Correlaciones genotipo-fenotipo

El modelo que existe de correlación genotipo-fenotipo para el gen *ABCA4* fue propuesto en el año 1999, tras las primeras publicaciones de pacientes con este gen mutado (Cremers *et al.*, 1998; van Driel *et al.*, 1998; Maugeri *et al.*, 1999), que se actualiza con estudios en los que clasifican la severidad de ciertas variantes, bien según el fenotipo, o según la cantidad de ARN no aberrante determinada mediante estudios funcionales (Sangermano *et al.*, 2016, 2018; Zernant *et al.*, 2017). Por esta razón, nosotros realizamos un estudio de correlación genotipo-fenotipo en los pacientes de los que disponíamos de datos clínicos teniendo en cuenta diversos criterios (**Artículo 4**). En primer lugar, se dividieron los alelos mutados en función del tipo de variante, clasificándolos en *missense* o truncantes. Además, se realizó la misma correlación atendiendo al fenotipo.

Agrupando los pacientes según el tipo de variante sin tener en cuenta el fenotipo, observamos que la edad de inicio de los síntomas (mediana (rango intercuartil)) va disminuyendo según aumenta la severidad del tipo de variante, con resultados estadísticamente significativos: 17 (15), 14 (14) y 9 (3.5) años para las combinaciones de genotipos *missense-missense*, *missense*-truncante y truncante-truncante, respectivamente. Estos resultados apoyan que la severidad del fenotipo, atendiendo a un temprano desarrollo de los síntomas, está asociada a la presencia de variantes truncantes en los pacientes.

Por otra parte, clasificando los pacientes por fenotipos en STGD1 y DCB, observamos que la distribución de la edad de inicio es diferente y estadísticamente significativa en ambos grupos: mientras que los pacientes DCB presentan una edad de inicio uniforme en la primera década de la vida, en los pacientes STGD1 esta distribución es más heterogénea y se observa que la mayoría presentan su debut entre la primera y la tercera década de la vida.

Analizando los genotipos según estos dos fenotipos, hemos observado que el 93% de los pacientes STGD1 presentan al menos una variante *missense* (286/306), mientras que el 41% de los pacientes DCB presentan dos variantes truncantes (27/66). Este resultado sugiere que los pacientes diagnosticados como STGD1 que presentan dos variantes truncantes probablemente vayan a evolucionar a DCB, por lo que identificar variantes truncantes en *ABCA4* en estos pacientes puede ayudar a refinar su pronóstico. En cuanto a las variantes *missense*, hemos observado que 3 de ellas se asocian a un fenotipo de DCB: c.1804C>T; p.(Arg602Trp) (OR = 5.31; 95%CI = 2.27-11.7), c.3056C>T; p.(Thr1019Met) (OR = 7.58; 95%CI = 2.12-25.1), y c.6320G>C; p.(Arg2107Pro) (OR = 10.5; 95%CI = 1.08-102). Y únicamente la variante *missense* c.3386G>T; p.(Arg1129Leu) se asocia a STGD1 (OR = 0.37; 95%CI = 0.14-0.80). Al igual que en el caso de las mutaciones truncantes, identificar estas variantes *missense* puede dar un valor pronóstico a otros pacientes que se diagnostican temprano y en los que la sintomatología es leve. Para evaluar el efecto que estas variantes podrían tener en la proteína ABCA4, ya que se esperaría que al estar asociadas a

174

un fenotipo más severo provoquen un cambio estructural que impida su funcionalidad, se necesitaría tener un modelo estructural de ABCA4. Sin embargo, la caracterización estructural de esta proteína es difícil debido a su complejidad, disponibilidad, y naturaleza transmembrana. Esto limita su estudio a modelos animales, estudios *in vitro*, y modelos estructurales obtenidos a partir de proteínas de otras especies y estudios computacionales (Tsybovsky *et al.*, 2013; Tsybovsky and Palczewski, 2014; Molday, 2015). Los cambios asociados a DCB se sitúan en los dominios ECD1 (Arg602), NBD1 (Thr1019) y NBD2 (Arg2107), siendo los dos últimos los dominios más conservados de estas proteínas (Tsybovsky, Molday and Palczewski, 2010). Los aminoácidos que cambian (Arg→Trp, Thr→Met y Arg→Pro) si difieren bastante en cuanto a su tamaño y polaridad, lo que podría implicar un cambio estructural que interfiera en la correcta funcionalidad de la proteína. Sin embargo, la variante asociada a STGD1 (p.Arg1129Leu), también se encuentra en el NBD1 y el cambio es un aminoácido polar por uno no polar, pero podría no implicar un cambio estructural importante.

Otro análisis que se llevó a cabo fue el estudio de correlación genotipo-fenotipo en las <u>variantes más prevalentes</u> de nuestra cohorte. Así, se estudió la variante c.3386G>T; p.(Arg1129Leu), presente en el 18% de los alelos mutados, y la c.5882G>A; p.(Gly1961Glu), que representa el 6% de los mismos.

Se realizaron comparaciones entre los pacientes con la variante <u>c.3386G>T; p.(Arg1129Leu)</u> respecto a todos los pacientes caracterizados con variantes diferentes a esta para observar si había diferencias en cuanto a la severidad de esta variante con respecto a otras variantes *missense*. Se demostró que, *a priori*, ser portador de la variante c.3386G>T; p.(Arg1129Leu) no estaría relacionado con presentar un fenotipo más leve en cuanto a la edad de inicio se refiere. Sin embargo, los homocigotos para esta variante presentaban una edad de inicio superior a los otros grupos, 21.5 (18.5) años, además del hecho de que solamente representaban al 11% de los pacientes portadores de esta variante (18/170), siendo tan frecuente en población española.

En el grupo de los 8 pacientes que presentan respeto foveal, la variante c.3386G>T; p.(Arg1129Leu) estaba en homocigosis en 2 de ellos, y en heterocigosis compuesta en uno. Ninguna otra variante se encontraba tan representada en este grupo. Estos hallazgos fenotípicos, además de la edad de inicio tardía en los pacientes homocigotos, sugieren que la severidad de esta variante es leve-moderada. Además, fue la única variante *missense* que se asociaba a STGD1 (OR = 0.37; 95%CI = 0.14-0.80). Sin embargo, en nuestra cohorte, también había 6 pacientes con DCB que presentaban esta variante, la mayoría en combinación con otra variante *missense*, e incluso había un paciente homocigoto. Estos pacientes fueron clasificados como DCB en base al diagnóstico referido por sus clínicos y al presentar ceguera nocturna, sin embargo, no tenemos datos de ERG para confirmarlo, por lo que recomendamos la reevaluación clínica.

En el caso de la segunda variante más prevalente, la c.5882G>A; p.(Gly1961Glu), el dato a resaltar es que no había pacientes homocigotos para esta variante en nuestra cohorte. Esta variante es la más frecuente a nivel mundial, y se ha descrito que los homocigotos presentan un fenotipo más leve (Tanna *et al.*, 2017).

Nuestra hipótesis para explicar la baja frecuencia y ausencia de homocigotos con las variantes c.3386G>T; p.(Arg1129Leu) y c.5882G>A; p.(Gly1961Glu), respectivamente, podría recaer en que estos individuos presenten un fenotipo leve y que no hayan desarrollado síntomas, o que se traten de variantes de reducida penetrancia que necesiten otro elemento en *cis* para desarrollar el fenotipo cuando están en homocigosis, mientras que si están en combinación con otras variantes actúan como penetrantes.

Las limitaciones de este estudio principalmente son clínicas, ya que en la mayoría de los pacientes no disponemos de todas las pruebas oftalmológicas para hacer un diagnóstico preciso, por lo que no podemos asegurar que nuestra clasificación sea totalmente fidedigna. Además, el espectro fenotípico del gen *ABCA4* es tan amplio, que restringir su clasificación a enfermedades concretas, aun siendo necesario en aras de facilitar su análisis, supone una limitación respecto a la riqueza de espectro fenotípico que se asocia a las variantes en este gen. En cuanto a la clasificación de las variantes en *missense* y truncantes, también sabemos que no es exacta, ya que las variantes clasificadas como truncantes incluyen diferentes tipos, y, por ejemplo, los estudios del ARN en variantes de *splicing* demuestran que algunas variantes intrónicas provocan tránscritos truncantes, pero también normales, al igual que puede pasar con las variantes situadas en los sitios canónicos que no se estudian funcionalmente. Además, se han descrito variantes *missense* severas, en función de estudios funcionales (Zhang *et al.*, 2014; Tanna *et al.*, 2017; Molday *et al.*, 2018) que en nuestro estudio no se relacionan con un fenotipo más grave. Con todo, creemos que la aproximación llevada a cabo en este trabajo es la apropiada para abordar y establecer correlaciones en más de 500 pacientes con mutaciones en el gen *ABCA4*.

III. Hallazgos en otros genes asociados a fenotipos Stargardt-*like*

Uno de los objetivos de esta Tesis Doctoral fue describir los hallazgos moleculares y clínicos de otros genes que también se han asociado a la enfermedad de Stargardt. Principalmente, estos genes son *PRPH2*, *ELOVL4*, y *PROM1*. El gen *PRPH2* es uno de los genes más prevalentes en nuestra cohorte de DHR asociados a RP (Martin-Merida *et al.*, 2018), mientras que en nuestra cohorte de pacientes nunca se han identificado variantes patogénicas en el gen *ELOVL4*.

Las mutaciones en el gen *PROM1* son también raras y poco prevalentes. Al inicio de esta Tesis Doctoral se habían reportado alrededor de 70 variantes diferentes. Sin embargo, la mayoría de las publicaciones recientes sobre variantes en este gen no disponían de datos familiares, genéticos o fenotípicos para evaluar correctamente su patogenicidad. Por esta razón, este gen fue seleccionado para expandir su espectro genético y fenotípico (**Artículo 5**).

Se seleccionaron 2216 familias en las que se había cribado el gen *PROM1* mediante diferentes tecnologías a lo largo de los años. Sólo se identificaron variantes raras en *PROM1* en 32 familias, lo que revela que es un gen causal muy poco frecuente. En total, fueron identificadas 19 variantes, 10 de las cuales se describían por primera vez en este trabajo. Las familias consideradas como caracterizadas con variantes en el gen *PROM1* fueron 22. Esto supone que la prevalencia de casos caracterizados con el gen *PROM1* en nuestra cohorte es ~1%. En 18 familias se identificaron variantes con herencia autosómica recesiva y en 4 con herencia dominante. Además, en 3 familias se identificó una variante recesiva en heterocigosis, por lo que no estaban totalmente caracterizadas, y otras 7 familias eran portadoras de variantes de significado clínico incierto (estudios no concluyentes).

El estudio fenotípico se realizó con 25 pacientes de las 22 familias caracterizadas. Se clasificaron según sus hallazgos clínicos en DCB, RP y DM. En primer lugar, no había diferencias entre fenotipos clasificando las familias por su herencia (dominante o recesiva) así como por el tipo de variante (*missense* o truncantes), como se había descrito previamente (Liu *et al.*, 2016).

Respecto al fenotipo asociado, cabe mencionar que *PROM1* se había asociado a la enfermedad de Stargardt con herencia dominante, y en concreto, la variante c.1117C>T; p.(Arg373Cys) (Yang *et al.*, 2008). Nuestro único paciente con esta variante y herencia dominante presentaba un fenotipo de RP.

En cuanto al resto de nuestros pacientes con herencia dominante, 5 pacientes de 3 familias eran portadores de una nueva variante de *splicing* prevalente en nuestra cohorte, c.303+1G>A; p.(?), y su fenotipo era de DM, con 2 de ellos presentando además una leve alteración de conos y bastones en el ERG. La media de la edad de inicio en estos pacientes era 46.75 ± 11.7 años, y algunos presentaban valores de agudeza visual de la unidad hasta edades superiores a los 50 años.

Por esta razón, el fenotipo que se reportó como asociado a esta variante fue el de una maculopatía leve de inicio tardío. No obstante, un paciente con esta variante era un niño cuyo único síntoma era la alteración de los colores en el momento del diagnóstico, y cuyas pruebas oftalmológicas fueron normales exceptuando la autofluorescencia, en la que se observó hiperfluorescencia en una pequeña área de la mácula y el ERG, que mostraba una leve alteración de conos y bastones. Al no tener pruebas oftalmológicas del resto de pacientes con esta variante a edades tan tempranas, no podemos asegurar si este niño va a evolucionar con una progresión lenta, o puede tratarse de un caso en el que haya factores modificadores que agraven su fenotipo.

Por último, comparamos la gravedad y evolución de los síntomas de los pacientes, según su clasificación clínica. Las categorías de clasificación fenotípica que utilizamos para agrupar a todos los pacientes fueron RP y DCB o DM. De esta forma, pudimos observar que el primer síntoma de los 2 pacientes con RP era la ceguera nocturna, y que ambos presentaban atrofia macular, mientras que en el caso de los pacientes con DCB, el primer síntoma era la pérdida de agudeza visual, y también presentaban atrofia macular en mayor o menor medida. Por esta razón, consideramos que la característica fenotípica distintiva del gen *PROM1* en pacientes con DHR es la presencia de atrofia macular, independientemente del fenotipo que presenten. Este rasgo, que apoya la patogenicidad específica de variantes en *PROM1*, podría eventualmente ayudar a reclasificar las VUS en este gen, apoyando su clasificación hacia una variante benigna o patogénica.

El espectro de fenotipos de DHR asociados al gen *PROM1*: RP, DC y DCB (con ambos tipos de fotorreceptores afectados), se ve apoyado por la distribución de la proteína PROM1 en la retina. Así, *PROM1* se expresa en conos y bastones, sin embargo, su localización difiere. En los bastones, la proteína PROM1 se localiza en la membrana de los discos nacientes, en la base de los segmentos externos. Por su parte, en los conos, parece que se distribuye por todas las evaginaciones de la membrana plasmática de los segmentos externos, las lamelas, que al estar abiertas, las proteínas están más expuestas al espacio extracelular, interfiriendo con otras proteínas y haciendo que se degeneren primariamente los conos en la mayoría de los casos. Aunque la patogénesis de *PROM1* en el ciclo visual se ha atribuido a una función estructural, regulando la morfogénesis de los discos de los segmentos externos, también se ha descrito que esta proteína es importante en el correcto posicionamiento de las opsinas en los segmentos externos, a través del cilio conector (Zacchigna *et al.*, 2009), así como en la regulación de la autofagia de las células del EPR (Bhattacharya *et al.*, 2017).

En las variantes más prevalentes de nuestra cohorte se realizaron análisis de haplotipos, revelando regiones compartidas de 17Kb, 99Kb y 457Kb entre las familias portadoras de las variantes c.303+1G>A; p.(?) (3 familias), c.1354dup; p.(Tyr452Leufs*13) (12 familias), y c.1984-1G>T;

p.(?) (3 familias), respectivamente. Esto revela un posible efecto fundador de estas variantes en nuestra población. En el caso de la variante c.1984-1G>T, únicamente se ha reportado en casos españoles y franceses (De Castro-Miró *et al.*, 2014; Boulanger-Scemama *et al.*, 2015; Bravo-Gil *et al.*, 2017), mientras que la c.1354dup; p.(Tyr452Leufs*13) ha sido descrita también en otras poblaciones, tunecinos y alemanes entre otros (Habibi *et al.*, 2016; Birtel *et al.*, 2018).

Uno de los hallazgos novedosos de este estudio fue el gran número de pacientes con variantes de *splicing* en *PROM1*. Hasta ese momento, se habían reportado alrededor de 12 variantes de este tipo, sin embargo, en nuestra cohorte, este tipo de variantes explican el fenotipo de 10 familias, lo que supone el 45% de las familias caracterizadas. Además, 3 de ellas eran nuevas, y 2 estaban localizadas en el sitio donador del intrón 3, en el primer y segundo nucleótido (c.303+1G>A y c.303+2T>C). El otro hallazgo a resaltar es el hecho de que la primera variante, localizada en el sitio +1 del donador canónico, fue identificada en 3 familias con una herencia aparentemente dominante (en una de ellas el probandus fue concebido por donación de gametos y no se pudo dilucidar el modo de herencia), mientras que la segunda, situada en el sitio +2, fue identificada en una familia recesiva, en la que el probandus la portaba en homocigosis. Se realizaron estudios funcionales en los pacientes con estas variantes para determinar si había diferencias en el patrón de *splicing* que pudieran explicar este hallazgo. Sin embargo, lo que pudimos comprobar fue que, aunque *PROM1* se expresaba en sangre, saliva y fibroblastos, ninguna de las isoformas de estos tejidos presentaba el exón 3 del gen *PROM1*, por lo que no había diferencias al utilizar los controles y los pacientes con estas variantes. Se han reportado varias isoformas asociadas a este gen, siendo 2 de ellas las isoformas mayormente expresadas en la retina (Permanyer *et al.*, 2010), las cuales presentan el exón 3 (97%). Nuestro estudio revela que el exón 3 no se expresa en tejidos extraoculares, lo que indica que puede ser un exón importante en los fotorreceptores de la retina. Nuestra hipótesis, también teniendo en cuenta que el fenotipo de los pacientes con la c.303+1G>A es tardío y leve, y el paciente con la c.303+2T>C presenta RP con atrofia macular, es que la causa del distinto efecto fenotípico de estas variantes no sería debido a que ambas presenten un patrón diferente de *splicing*, ya que las predicciones *in sílico* y el sitio en el que se localizan predicen la eliminación del exón 3 en ambos casos. Lo que puede ocurrir es que no se mantiene el equilibrio de isoformas normales (con el exón 3) y aberrantes (sin el exón 3) en la retina, lo que provocaría la patogénesis de PROM1 (Figura D1). En el caso de los pacientes heterocigotos, el alelo no mutado daría lugar al tránscrito mayoritario en la retina, mientras que el mutado daría el tránscrito sin el exón 3. Aun así, seguiría existiendo la proteína normal en la retina, y, por lo tanto, el fenotipo asociado es más leve. En el caso del paciente homocigoto, todos los tránscritos serían aberrantes, por lo que no habría proteína PROM1 funcional y el fenotipo es más severo. Los padres del paciente homocigoto eran portadores de la variante c.303+2T>C en heterocigosis, lo que según nuestra hipótesis tendría un desbalance de isoformas normales y mutadas. No tenemos

datos clínicos para confirmarlo ni descartarlo, aunque el paciente refiere que sus padres fueron diagnosticados de degeneración macular asociada a la edad.
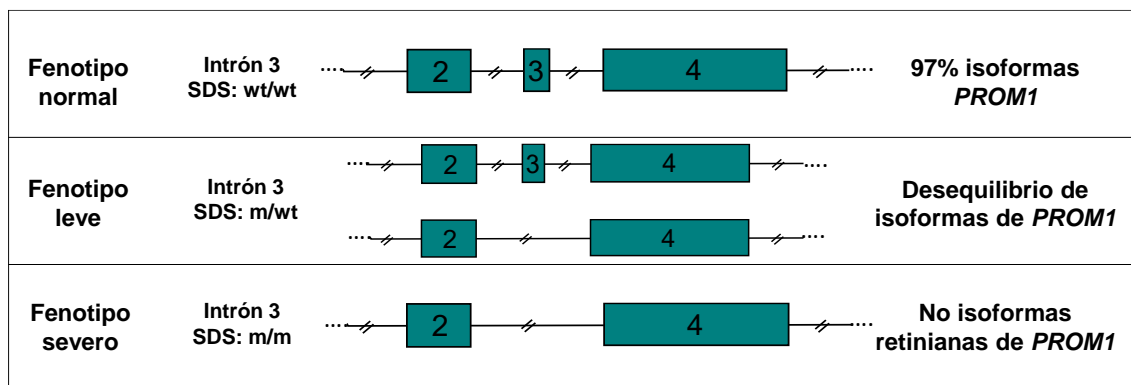

Figura D1. *Hipótesis sobre la patogenicidad asociada a las mutaciones en el sitio donador de splicing del intrón 3 del gen PROM1. Abreviaturas: SDS: sitio donador de splicing; wt/m: alelo silvestre/mutante.*

En un estudio muy reciente en el que se reporta la segunda cohorte más grande de pacientes con variantes en *PROM1* después de nuestro trabajo, los autores llegan a algunas conclusiones similares y otras diferentes. Cehajic-Kapetanovic y colaboradores (Cehajic-Kapetanovic *et al.*, 2019) establecen correlaciones entre las familias con herencia recesiva y dominante. En el primer caso, los pacientes presentan una edad de inicio más temprana y un fenotipo más severo que involucra a la periferia de la retina, mientras que los pacientes con herencia dominante, portadores de la variante c.1117C>T; p.(Arg373Cys), presentan una edad tardía y un fenotipo que involucra principalmente a los conos, confinado a la mácula. Como se ha dicho anteriormente, en nuestro trabajo no se puede llegar a estas conclusiones, principalmente porque hay un paciente con RP y herencia portador de esta variante con herencia dominante. Además, establecemos una nueva asociación fenotípica con una variante nueva en familias dominantes y un fenotipo leve de inicio tardío. Sin embargo, al igual que nosotros, estos autores destacan la falta de datos sobre los patrones de herencia en las publicaciones anteriores con variantes en el gen *PROM1*, también demuestran que todos los pacientes presentan afectación macular, y, por último, como en nuestro estudio, observan que el fenotipo mayoritario asociado al gen *PROM1* es una DCB.

180

IV. Algoritmo diagnóstico para el estudio de pacientes con sospecha de enfermedad de Stargardt

Basándonos en los resultados de esta Tesis Doctoral, proponemos el algoritmo diagnóstico que se detalla en el Anexo III para el estudio molecular de los pacientes remitidos con sospecha de enfermedad de Stargardt. Actualmente, los pacientes se estudian mediante exoma clínico, utilizándose paneles virtuales para filtrar por los genes de interés. En este paso, proponemos utilizar el panel de genes de distrofias de retina no sindrómica, que incluye los genes actualmente descritos asociados a estas enfermedades. Puesto que no siempre se dispone de información oftalmológica precisa de los pacientes, éste sería el mejor abordaje para caracterizar a los pacientes que presentan otro tipo de distrofias de retina, solapantes o no con la enfermedad de Stargardt. Si después de este estudio sólo se identificase una variante patogénica en el gen *ABCA4*, proponemos secuenciar las variantes intrónicas más prevalentes (en nuestro caso, c.4253+43G>A, c.4539+2064C>T y c.5196+1137G>A). En un futuro, estas regiones se incluirán en el exoma clínico por lo que no será necesario este segundo paso. Si finalmente no se identifican variantes intrónicas o el paciente no se ha caracterizado mediante exoma clínico, se necesitaría ampliar el número de genes y regiones estudiadas, bien mediante el análisis del gen *ABCA4* completo, o mediante exoma (WES) o genoma completo (WGS), en un contexto de investigación.

Conclusiones, *Conclusions*

1. El gen *ABCA4* es responsable de la patología del 73% de los casos remitidos con sospecha de enfermedad de Stargardt y del 25% de los remitidos con otras maculopatías. Además, en el 73% de los pacientes monoalélicos diagnosticados de enfermedad de Stargardt, se encuentra una segunda mutación en este gen.

2. El cribado del locus completo del gen *ABCA4*, así como el estudio funcional de variantes intrónicas, es necesario para la caracterización de la mayoría de los pacientes monoalélicos en los que se había secuenciado la región codificante. Actualmente, 35 variantes *deep intronic* y 60 variantes de *splicing* no canónicas alteran el mecanismo de *splicing*.

3. En el 4% de los casos con sospecha clínica de enfermedad de Stargardt se encuentran mutaciones en otros genes, y en el 20% no se identifica su causa molecular. Por lo tanto, estos casos necesitan una reevaluación clínica. Aunque no se ha podido establecer la implicación de *ABCA4* en los casos no resueltos, no se puede descartar que variantes intrónicas no estudiadas o situadas en regiones reguladoras estén jugando un papel en la patología de estos pacientes.

4. La distribución de la edad de inicio en los grupos de pacientes con STGD1 y DCB muestra una aparición más temprana y uniforme de los síntomas en los pacientes con DCB en la primera década de la vida, mientras que en los pacientes STGD1 aparecen a lo largo de la primera y tercera década de la vida.

5. El espectro mutacional del gen *ABCA4* en la cohorte más grande reportada, correspondiente a este trabajo, de 506 familias con mutaciones bialélicas en este gen, demuestra que las variantes *missense*, seguidas de las *frameshift* y *stop* son las más prevalentes. Las variantes *deep intronic* y CNVs suponen únicamente el 3% y 1%, respectivamente, de los alelos mutados en este gen.

6. La variante de *ABCA4* más prevalente en población española, p.(Arg1129Leu), se asocia a un fenotipo de STGD1 y aparece sobrerrepresentada en los pacientes con respeto foveal. Los pacientes homocigotos presentan un fenotipo más leve.

7. El diagnóstico genético puede ayudar al pronóstico de la evolución de la enfermedad en los pacientes con variantes en *ABCA4*. Las variantes truncantes, así como las tres variantes *missense* p.(Arg602Trp), p.(Thr1019Met) y p.(Arg2107Pro), se asocian a un fenotipo de DCB, mientras que ser portador del resto de variantes *missense* se asocia a presentar un fenotipo de STGD1, como se ha demostrado en el caso de la variante p.(Arg1129Leu).

8. Las mutaciones en el gen *PROM1* explican la causa del 1% de los casos de nuestra cohorte de DHR, que corresponde a 22 familias. Se asocia principalmente a un fenotipo de DCB, describiéndose también casos de RP y maculopatías leves. El rasgo común de todos estos pacientes es la presencia de atrofia macular.

9. En las familias con mutaciones en el gen *PROM1* se identifican los patrones de herencia tanto autosómica dominante como recesiva, dependiendo de la variante identificada. Dos nuevas variantes de *splicing* situadas en el intrón 3 se asocian a un fenotipo distinto dependiendo de su genotipo. La posible explicación que se propone es que la falta de la mayoría de las isoformas canónicas en la retina se relaciona con un fenotipo más grave. Por tanto, el exón 3 del gen *PROM1* debe desarrollar un papel importante en los fotorreceptores.

10. El algoritmo de diagnóstico molecular para los pacientes remitidos con sospecha de enfermedad de Stargardt debería incluir en el primer análisis el estudio de los genes asociados a distrofias de retina no sindrómica, especialmente aquellos con características clínicas similares (*PROM1* y *PRPH2*), para caracterizar posibles fenotipos solapantes. Además, se recomienda el estudio del gen *ABCA4* completo en el caso de identificar un único alelo mutante.

.

1. *ABCA4* is the gene responsible for the pathogenesis of 73% of cases diagnosed with Stargardt disease and 25% of cases referred with other maculopathies, most of them suffering cone-rod dystrophy. Moreover, in 73% of monoallelic *ABCA4* patients with Stargardt diagnosis a second mutated allele in this gene was found.

2. Complete screening of *ABCA4* locus, followed by functional splice assays for new intronic variants, are needed in order to characterize most of the monoallelic patients in which coding regions were previously sequenced. To date, 35 deep intronic variants and 60 non canonical splice site variants affecting splicing have been described.

3. Four percent of the patients clinically diagnosed with Stargardt disease were carriers of mutations in other genes different from *ABCA4*; meanwhile 20% of these cases were unsolved. A clinical reevaluation of these patients is therefore needed. Although it has not been possible to establish the involvement of *ABCA4* in those unsolved cases, it cannot be discarded that other intronic variants not studied or located in regulatory regions could be playing a role in the pathology of these patients.

4. Distribution of age of onset in STGD1 and CRD patients with *ABCA4* mutations show that CRD patients present an earlier and uniform age of onset during the first decade of life, while STGD1 patients develop the first symptoms along the first and third decade of life.

5. The mutational spectrum of the *ABCA4* gene in the largest cohort reported in this work, 506 biallelic *ABCA4* families, reveals that missense followed by frameshift and stop mutations are the most prevalent variants found. Deep intronic variants and CNVs only represent, respectively, 3% and 1% of the mutated alleles in this gene.

6. The most prevalent *ABCA4* variant in the Spanish population, p.(Arg1129Leu), is associated with STGD1 and is over-represented in foveal sparing patients. Homozygous patients present with a mild phenotype.

7. The genetic diagnosis can help to establish a prognosis of the disease, since truncating and three missense variants, p.(Arg602Trp), p.(Thr1019Met) and p.(Arg2107Pro), are associated to CRD, while being a carrier of the remaining missense variants is related to presenting a STGD1 phenotype, as demonstrated in the case of p.(Arg1129Leu) variant.

8. Mutations in *PROM1* explain 1% of the cases of our cohort, what corresponds to 22 families. It is mainly associated with a CRD phenotype, although there are also cases with RP and mild maculopathies. Macular atrophy is the common sign in all *PROM1* patients.

9. Families with *PROM1* mutations have an autosomal dominant or recessive pattern of inheritance depending on the variant identified. Two novel splicing variants affecting donor site of intron 3 are associated with a different phenotype according to their genotype. As a possible explanation we propose that the lack of most of the canonical

isoforms found in the retina induces a more severe phenotype. Therefore, exon 3 of *PROM1* should play an important role in photoreceptors.

10. The molecular diagnostic algorithm for patients suspected of Stargardt disease should include a first step to screen all genes associated with non-syndromic IRDs, especially those genes causing similar clinical features (*PROM1* and *PRPH2*), in order to identify other overlapping phenotypes. Besides, the screening of the entire *ABCA4* gene in those cases where an *ABCA4* mutated allele has been identified is recommended.

Bibliografía

Aboshiha, J. *et al.* (2016) 'The cone dysfunction syndromes', *British Journal of Ophthalmology*. doi: 10.1136/bjophthalmol-2014-306505.

Adzhubei, I. A. *et al.* (2010) 'A method and server for predicting damaging missense mutations', *Nature Methods*, pp. 248–249. doi: 10.1038/nmeth0410-248.

Aguirre-Lamban, J. *et al.* (2009) 'Molecular analysis of the ABCA4 gene for reliable detection of allelic variations in Spanish patients: Identification of 21 novel variants', *British Journal of Ophthalmology*. doi: 10.1136/bjo.2008.145193.

Aguirre-Lamban, J. *et al.* (2010) 'Comparison of high-resolution melting analysis with denaturing high-performance liquid chromatography for mutation scanning in the ABCA4 gene', *Investigative Ophthalmology and Visual Science*. doi: 10.1167/iovs.09-4518.

Aguirre-Lamban, J. *et al.* (2011) 'Further associations between mutations and polymorphisms in the ABCA4 gene: Clinical implication of allelic variants and their role as protector/risk factors', *Investigative Ophthalmology and Visual Science*. doi: 10.1167/iovs.10-5743.

Albert., S. *et al.* (2015) 'Towards the identification of deepintronic ABCA4 mutations in Stargardt patients by using induced pluripotent stem cell-derived photoreceptor progenitor cells', *Investigative Ophthalmology and Visual Science*.

Albert., S. *et al.* (2017) 'Identification, RNA splice defect assessment and AON correction of non-coding variants of ABCA4 in Stargardt disease', *Investigative Ophthalmology and Visual Science*.

Albert, S. *et al.* (2018) 'Identification and Rescue of Splice Defects Caused by Two Neighboring Deep-Intronic ABCA4 Mutations Underlying Stargardt Disease', *American Journal of Human Genetics*. doi: 10.1016/j.ajhg.2018.02.008.

Allikmets, R. *et al.* (1997) 'A photoreceptor cell-specific ATP-binding transporter gene (ABCR) is mutated in recessive Stargardt macular dystrophy', *Nature Genetics*, 15(3), pp. 236–246. doi: 10.1038/ng0397-236.

Allikmets, R. *et al.* (1997) 'Mutation of the Stargardt disease gene (ABCR) in age-related macular degeneration', *Science*. doi: 10.1126/science.277.5333.1805.

Allocca, M. *et al.* (2008) 'Serotype-dependent packaging of large genes in adeno-associated viral vectors results in effective gene delivery in mice', *Journal of Clinical Investigation*. doi: 10.1172/JCI34316.

Anderson, K. L. *et al.* (1995) 'A YAC contig encompassing the recessive Stargardt disease gene (STGD) on chromosome 1p.', *American journal of human genetics*, 57(6), pp. 1351–63. Available at: http://www.ncbi.nlm.nih.gov/pubmed/8533764 (Accessed: 24 July 2019).

Armstrong, J. D. *et al.* (1998) 'Long-term follow-up of Stargardt's disease and fundus flavimaculatus', *Ophthalmology*. doi: 10.1016/S0161-6420(98)93026-3.

Aukrust, I. *et al.* (2017) 'The intronic ABCA4 c.5461-10T>C variant, frequently seen in patients with Stargardt disease, causes splice defects and reduced ABCA4 protein level', *Acta Ophthalmologica*. doi: 10.1111/aos.13273.

Ayuso, C. and Millan, J. M. (2010) 'Retinitis pigmentosa and allied conditions today: A paradigm of translational research', *Genome Medicine*, p. 34. doi: 10.1186/gm155.

Bauwens, M. *et al.* (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*. doi: 10.1038/s41436-018-0420-y.

Bibliografía

Bax, N. M. *et al.* (2015) 'Heterozygous deep-intronic variants and deletions in ABCA4 in persons with retinal dystrophies and one exonic ABCA4 variant', *Human Mutation*, 36(1), pp. 43–47. doi: 10.1002/humu.22717.

Bernardis, I. *et al.* (2016) 'Unravelling the Complexity of Inherited Retinal Dystrophies Molecular Testing: Added Value of Targeted Next-Generation Sequencing', *BioMed Research International*. Hindawi Limited, 2016. doi: 10.1155/2016/6341870.

Bertelsen, M. *et al.* (2014) 'Generalized choriocapillaris dystrophy, a distinct phenotype in the spectrum of ABCA4-associated retinopathies', *Investigative Ophthalmology and Visual Science*. doi: 10.1167/iovs.13-13391.

Bhattacharya, S. *et al.* (2017) 'Prominin-1 is a novel regulator of autophagy in the human retinal pigment epithelium', *Investigative Ophthalmology and Visual Science*. doi: 10.1167/iovs.16-21162.

Birtel, J. *et al.* (2018) 'Clinical and genetic characteristics of 251 consecutive patients with macular and cone/cone-rod dystrophy', *Scientific Reports*. doi: 10.1038/s41598-018-22096-0.

Blacharski PA (1988) 'Fundus flavimaculatus', *Retinal Dystrophies and Degenerations*. Raven Press, pp. 135–159.

Boon, C. J. F. *et al.* (2007) 'Mutations in the peripherin/RDS gene are an important cause of multifocal pattern dystrophy simulating STGD1/fundus flavimaculatus', *British Journal of Ophthalmology*. doi: 10.1136/bjo.2007.115659.

Boulanger-Scemama, E. *et al.* (2015) 'Next-generation sequencing applied to a large French cone and cone-rod dystrophy cohort: Mutation spectrum and new genotype-phenotype correlation', *Orphanet Journal of Rare Diseases*, 10(1). doi: 10.1186/s13023-015-0300-3.

Braun, T. A. *et al.* (2013) 'Non-exomic and synonymous variants in ABCA4 are an important cause of Stargardt disease', *Human Molecular Genetics*, 22(25), pp. 5136–5145. doi: 10.1093/hmg/ddt367.

Bravo-Gil, N. *et al.* (2017) 'Unravelling the genetic basis of simplex Retinitis Pigmentosa cases', *Scientific Reports*, 7, p. 41937. doi: 10.1038/srep41937.

Burnight, E. R. *et al.* (2017) 'Using CRISPR-Cas9 to Generate Gene-Corrected Autologous iPSCs for the Treatment of Inherited Retinal Degeneration', *Molecular Therapy*. doi: 10.1016/j.ymthe.2017.05.015.

De Castro-Miró, M. *et al.* (2014) 'Combined Genetic and high-throughput strategies for molecular diagnosis of inherited retinal dystrophies', *PLoS ONE*, 9(2), p. e88410. doi: 10.1371/journal.pone.0088410.

Cehajic-Kapetanovic, J. *et al.* (2019) 'Clinical and Molecular Characterization of PROM1-Related Retinal Degeneration', *JAMA network open*. doi: 10.1001/jamanetworkopen.2019.5752.

Chacón-Camacho, O. F. *et al.* (2013) 'ABCA4 mutational spectrum in Mexican patients with Stargardt disease: Identification of 12 novel mutations and evidence of a founder effect for the common p.A1773V mutation', *Experimental Eye Research*. doi: 10.1016/j.exer.2013.02.006.

Chen, B. *et al.* (2010) 'Analysis of Autofluorescent retinal images and measurement of atrophic lesion growth in Stargardt disease', *Experimental Eye Research*. doi: 10.1016/j.exer.2010.03.021.

Colella, P. and Auricchio, A. (2012) 'Gene therapy of inherited retinopathies: A long and successful road from viral vectors to patients', *Human Gene Therapy*. doi: 10.1089/hum.2012.123.

Corbeil, D. *et al.* (2000) 'The human AC133 hematopoietic stem cell antigen is also expressed in epithelial cells and targeted to plasma membrane protrusions.', *The Journal of biological chemistry*, 275(8), pp. 5512–20. Available at: http://www.ncbi.nlm.nih.gov/pubmed/10681530.

Corbeil, D. (2013) *Prominin-1 (CD133): New Insights on Stem & Cancer Stem Cell Biology*. 1st edn. Edited by D. Corbeil. Springer-Verlag New York. doi: 10.1007/978-1-4614-5894-4.

Cornelis, S. S. *et al.* (2017) 'In Silico Functional Meta-Analysis of 5,962 ABCA4 Variants in 3,928 Retinal Dystrophy Cases', *Human Mutation*, 38(4), pp. 400–408. doi: 10.1002/humu.23165.

Corton, M. *et al.* (2013) 'Exome sequencing of index patients with retinal dystrophies as a tool for molecular diagnosis.', *PloS one*. doi: 10.1371/journal.pone.0065574.

Cremers, F P M *et al.* (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), pp. 355–362. doi: 10.1093/hmg/7.3.355.

Cremers, F. P.M. *et al.* (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), pp. 355–362. doi: 10.1093/hmg/7.3.355.

Cremers, F. P. M. *et al.* (2018) 'Author response: Penetrance of the ABCA4 p.Asn1868Ile allele in stargardt disease', *Investigative Ophthalmology and Visual Science*. doi: 10.1167/iovs.18-25944.

Cremers, F. P. M. *et al.* (2020) 'Clinical spectrum, genetic complexity and therapeutic approaches for retinal disease caused by ABCA4 mutations.', *Progress in retinal and eye research*. England, p. 100861. doi: 10.1016/j.preteyeres.2020.100861.

Cremers, F. P. M. M. *et al.* (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), pp. 355–362. doi: 10.1093/hmg/7.3.355.

Dhalla, M. S., Desai, U. R. and Zuckerbrod, D. S. (2006) 'Pigmentary maculopathy in a patient with Wolfram syndrome', *Canadian Journal of Ophthalmology*. doi: 10.1016/S0008-4182(06)80064-5.

van Driel, M. A. *et al.* (1998) 'ABCR unites what ophthalmologists divide(s)', *Ophthalmic genetics*, 19(3), pp. 117–22. Available at: http://www.ncbi.nlm.nih.gov/pubmed/9810566 (Accessed: 24 July 2019).

Dubois, J., Terrier, O. and Rosa-Calatrava, M. (2014) 'Influenza viruses and mRNA splicing: Doing more with less', *mBio*. doi: 10.1128/mBio.00070-14.

Eisenberger, T. *et al.* (2013) 'Increasing the yield in targeted next-generation sequencing by implicating CNV analysis, non-coding exons and the overall variant load: The example of retinal dystrophies', *PLoS ONE*, 8(11). doi: 10.1371/journal.pone.0078496.

Estrada-Cuzcano, A. *et al.* (2012) 'BBS1 mutations in a wide spectrum of phenotypes ranging from nonsyndromic retinitis pigmentosa to bardet-biedl syndrome', *Archives of Ophthalmology*. Arch Ophthalmol, 130(11), pp. 1425–1432. doi: 10.1001/archophthalmol.2012.2434.

Fadaie, Z. *et al.* (2019) ' Identification of splice defects due to noncanonical splice site or deep-intronic variants in ABCA4 ', *Human Mutation*. doi: 10.1002/humu.23890.

Fargeas, C. A. *et al.* (2006) 'Prominin-1 (CD133): from progenitor cells to human diseases', *Future Lipidology*, 1(2), pp. 213–225. doi: 10.2217/17460875.1.2.213.

Fargeas, C. A., Huttner, W. B. and Corbeil, D. (2007) 'Nomenclature of prominin-1 (CD133) splice

variants - An update', *Tissue Antigens*, 69(6), pp. 602–606. doi: 10.1111/j.1399-0039.2007.00825.x.

'First CRISPR therapy dosed' (2020) *Nature biotechnology*. NLM (Medline), p. 382. doi: 10.1038/s41587-020-0493-4.

Florek, M. *et al.* (2005) 'Prominin-1/CD133, a neural and hematopoietic stem cell marker, is expressed in adult human differentiated cells and certain types of kidney cancer', *Cell and Tissue Research*, 319(1), pp. 15–26. doi: 10.1007/s00441-004-1018-z.

Fujinami, K. *et al.* (2011) 'Stargardt disease with preserved central vision: identification of a putative novel mutation in ATP-binding cassette transporter gene.', *Acta ophthalmologica*, 89(3), pp. e297-8. doi: 10.1111/j.1755-3768.2009.01848.x.

Fujinami, K. *et al.* (2013) 'ABCA4 gene screening by next-generation sequencing in a British cohort', *Investigative Ophthalmology and Visual Science*, 54(10), pp. 6662–6674. doi: 10.1167/iovs.13-12570.

Fujinami, K. *et al.* (2013) 'Clinical and molecular analysis of stargardt disease with preserved foveal structure and function', *American Journal of Ophthalmology*. doi: 10.1016/j.ajo.2013.05.003.

Fujinami, K. *et al.* (2019) 'Detailed genetic characteristics of an international large cohort of patients with Stargardt disease: ProgStar study report 8', *British Journal of Ophthalmology*. doi: 10.1136/bjophthalmol-2018-312064.

Fuster-García, C. *et al.* (2017) 'USH2A Gene Editing Using the CRISPR System', *Molecular Therapy - Nucleic Acids*. doi: 10.1016/j.omtn.2017.08.003.

Garanto, A. *et al.* (2019) 'Antisense oligonucleotide screening to optimize the rescue of the splicing defect caused by the recurrent deep-intronic ABCA4 variant c.4539+2001G>A in stargardt disease', *Genes*. doi: 10.3390/genes10060452.

Garoon, R. B. and Stout, J. T. (2016) 'Update on ocular gene therapy and advances in treatment of inherited retinal diseases and exudative macular degeneration', *Current Opinion in Ophthalmology*. doi: 10.1097/ICU.0000000000000256.

Gill, J. S. *et al.* (2019) 'Progressive cone and cone-rod dystrophies: Clinical features, molecular genetics and prospects for therapy', *British Journal of Ophthalmology*. doi: 10.1136/bjophthalmol-2018-313278.

Gilliam, J. C. *et al.* (2012) 'Three-dimensional architecture of the rod sensory cilium and its disruption in retinal neurodegeneration', *Cell*. doi: 10.1016/j.cell.2012.10.038.

Gomes, N. L. *et al.* (2009) 'A comparison of fundus autofluorescence and retinal structure in patients with stargardt disease', *Investigative Ophthalmology and Visual Science*. doi: 10.1167/iovs.08-2657.

González-del Pozo, M. *et al.* (2018) 'Searching the second hit in patients with inherited retinal dystrophies and monoallelic variants in ABCA4, USH2A and CEP290 by whole-gene targeted sequencing', *Scientific Reports*. doi: 10.1038/s41598-018-31511-5.

Greenwald, D. L., Cashman, S. M. and Kumar-Singh, R. (2010) 'Engineered zinc finger nuclease-mediated homologous recombination of the human rhodopsin gene', *Investigative Ophthalmology and Visual Science*. doi: 10.1167/iovs.10-5781.

Gross, A. M. *et al.* (2019) 'Copy-number variants in clinical genome sequencing: deployment and interpretation for rare and undiagnosed disease', *Genetics in Medicine*, 21(5), pp. 1121–1130.

doi: 10.1038/s41436-018-0295-y.

Habibi, I. *et al.* (2016) 'Identifying mutations in Tunisian families with retinal dystrophy', *Scientific Reports*. doi: 10.1038/srep37455.

Hamel, C. P. (2007) 'Cone rod dystrophies', *Orphanet Journal of Rare Diseases*. doi: 10.1186/1750-1172-2-7.

Hampton, T. (2020) 'With First CRISPR Trials, Gene Editing Moves Toward the Clinic', *JAMA - Journal of the American Medical Association*. doi: 10.1001/jama.2020.3438.

Han, Z. *et al.* (2012) 'DNA nanoparticle-mediated ABCA4 delivery rescues Stargardt dystrophy in mice', *Journal of Clinical Investigation*. doi: 10.1172/JCI64833.

Hanany, M., Rivolta, C. and Sharon, D. (2020) 'Worldwide carrier frequency and genetic prevalence of autosomal recessive inherited retinal diseases', *Proceedings of the National Academy of Sciences of the United States of America*. doi: 10.1073/pnas.1913179117.

Havens, M. A., Duelli, D. M. and Hastings, M. L. (2013) 'Targeting RNA splicing for disease therapy', *Wiley Interdisciplinary Reviews: RNA*. doi: 10.1002/wrna.1158.

Issa, P. C. *et al.* (2015) 'Rescue of the Stargardt phenotype in Abca4 knockout mice through inhibition of vitamin A dimerization', *Proceedings of the National Academy of Sciences of the United States of America*. doi: 10.1073/pnas.1506960112.

Jagadeesh, K. A. *et al.* (2016) 'M-CAP eliminates a majority of variants of uncertain significance in clinical exomes at high sensitivity', *Nature Genetics*. doi: 10.1038/ng.3703.

Jiang, F. *et al.* (2016) 'Screening of ABCA4 gene in a chinese cohort with stargardt disease or cone-rod dystrophy with a report on 85 novel mutations', *Investigative Ophthalmology and Visual Science*. doi: 10.1167/iovs.15-18190.

Khan, M. *et al.* (2019) 'Cost-effective molecular inversion probe-based *ABCA4* sequencing reveals deep-intronic variants in Stargardt disease', *Human Mutation*. doi: 10.1002/humu.23787.

Khan, M. *et al.* (2020) 'Resolving the dark matter of ABCA4 for 1054 Stargardt disease probands through integrated genomics and transcriptomics', *Genetics in Medicine*. doi: 10.1038/s41436-020-0787-4.

Kircher, M. *et al.* (2014) 'A general framework for estimating the relative pathogenicity of human genetic variants', *Nature Genetics*, 46(3), pp. 310–315. doi: 10.1038/ng.2892.

Klevering, B. J. *et al.* (2005) 'The spectrum of retinal phenotypes caused by mutations in the ABCA4 gene', *Graefe's Archive for Clinical and Experimental Ophthalmology*.

Kniazeva, M *et al.* (1999) 'A new locus for autosomal dominant stargardt-like disease maps to chromosome 4.', *American journal of human genetics*, 64(5), pp. 1394–9. doi: 10.1086/302377.

Kniazeva, M. *et al.* (1999) 'Clinical and genetic studies of an autosomal dominant cone-rod dystrophy with features of Stargardt disease', *Ophthalmic Genetics*, 20(2), pp. 71–81. doi: 10.1076/opge.20.2.71.2287.

Kumaran, N. *et al.* (2018) 'Retinal gene therapy', *British Medical Bulletin*. doi: 10.1093/bmb/ldy005.

Lee, J. *et al.* (2001) 'The metabolism of fatty acids in human Bietti crystalline dystrophy', *Investigative Ophthalmology and Visual Science*.

Lenis, T. L. *et al.* (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*. doi: 10.1073/pnas.1802519115.

Lim, K. H. *et al.* (2011) 'Using positional distribution to identify splicing elements and predict pre-mRNA processing defects in human genes', *Proceedings of the National Academy of Sciences of the United States of America*. doi: 10.1073/pnas.1101135108.

Liu, S. *et al.* (2016) 'Whole-exome sequencing identifies a novel homozygous frameshift mutation in the PROM1 gene as a causative mutation in two patients with sporadic retinitis pigmentosa', *International Journal of Molecular Medicine*, 37(6), pp. 1528–1534. doi: 10.3892/ijmm.2016.2551.

Lois, N. *et al.* (2001) 'Phenotypic subtypes of Stargardt macular: Dystrophy-fundus flavimaculatus', *Archives of Ophthalmology*. doi: 10.1001/archopht.119.3.359.

Low, B. E. *et al.* (2013) 'Correction of the Crb1rd8 allele and retinal phenotype in C57BL/6N mice via TALEN-mediated homology-directed repair', *Investigative Ophthalmology and Visual Science*. doi: 10.1167/iovs.13-13278.

Lu, B. *et al.* (2009) 'Long-term safety and function of RPE from human embryonic stem cells in preclinical models of macular degeneration', *Stem Cells*. doi: 10.1002/stem.149.

Martin-Merida, I. *et al.* (2018) 'Toward the mutational landscape of autosomal dominant retinitis pigmentosa: A comprehensive analysis of 258 Spanish families', *Investigative Ophthalmology and Visual Science*, 59(6), pp. 2345–2354. doi: 10.1167/iovs.18-23854.

Martin-Merida, I. *et al.* (2019) 'Genomic Landscape of Sporadic Retinitis Pigmentosa', *Ophthalmology*. doi: 10.1016/j.ophtha.2019.03.018.

Martínez-Mir, A. *et al.* (1998) 'Retinitis pigmentosa caused by a homozygous mutation in the Stargardt disease gene ABCR', *Nature Genetics*, 18(1), pp. 11–12. doi: 10.1038/ng0198-11.

Mata, N. L. *et al.* (2002) 'Isomerization and oxidation of vitamin A in cone-dominant retinas: A novel pathway for visual-pigment regeneration in daylight', *Neuron*. doi: 10.1016/S0896-6273(02)00912-1.

Maugeri, A. *et al.* (1999) 'The 2588G→C mutation in the ABCR gene is a mild frequent founder mutation in the western European population and allows the classification of ABCR mutations in patients with Stargardt disease', *American Journal of Human Genetics*, 64(4), pp. 1024–35. doi: 10.1086/302323.

Maugeri, A. *et al.* (2000) 'Report Mutations in the ABCA4 (ABCR) Gene Are the Major Cause of Autosomal Recessive Cone-Rod Dystrophy', *Am. J. Hum. Genet*, 67(4), pp. 960–966. doi: 10.1086/303079.

Maw, M. A. *et al.* (2000) 'A frameshift mutation in prominin (mouse)-like 1 causes human retinal degeneration.', *Human molecular genetics*, 9(1), pp. 27–34. doi: Doi 10.1093/Hmg/9.1.27.

Michaelides, M. *et al.* (2003) 'An autosomal dominant bull's-eye macular dystrophy (MCDR2) that maps to the short arm of chromosome 4', *Investigative Ophthalmology and Visual Science*, 44(4), pp. 1657–1662. doi: 10.1167/iovs.02-0941.

Michaelides, M. (2003) 'The genetics of inherited macular dystrophies', *Journal of Medical Genetics*, 40(9), pp. 641–650. doi: 10.1136/jmg.40.9.641.

Michaelides, M. *et al.* (2006) 'Progressive Cone and Cone-Rod Dystrophies: Phenotypes and Underlying Molecular Genetic Basis', *Survey of Ophthalmology*. doi:

10.1016/j.survophthal.2006.02.007.

Michaelides, M., Hunt, D. M. and Moore, A. T. (2003) 'The genetics of inherited macular dystrophies', *Journal of Medical Genetics*. doi: 10.1136/jmg.40.9.641.

Molday, L. L. *et al.* (2018) 'Localization and functional characterization of the p.Asn965Ser (N965S) ABCA4 variant in mice reveal pathogenic mechanisms underlying Stargardt macular degeneration', *Human Molecular Genetics*. doi: 10.1093/hmg/ddx400.

Molday, L. L., Rabin, A. R. and Molday, R. S. (2000) 'ABCR expression in foveal cone photoreceptors and its role in Stargardt macular dystrophy', *Nature Genetics*. doi: 10.1038/77004.

Molday, R. S. (2007) 'ATP-binding cassette transporter ABCA4: molecular properties and role in vision and macular degeneration.', *Journal of bioenergetics and biomembranes*, 39(5–6), pp. 507–17. doi: 10.1007/s10863-007-9118-6.

Molday, R. S. (2015) 'Insights into the Molecular Properties of ABCA4 and Its Role in the Visual Cycle and Stargardt Disease', in *Progress in Molecular Biology and Translational Science*. doi: 10.1016/bs.pmbts.2015.06.008.

Murphy, D. *et al.* (2016) 'The Musashi 1 Controls the Splicing of Photoreceptor-Specific Exons in the Vertebrate Retina', *PLoS Genetics*. doi: 10.1371/journal.pgen.1006256.

Nasca, A. *et al.* (2017) 'Not only dominant, not only optic atrophy: expanding the clinical spectrum associated with OPA1 mutations', *Orphanet Journal of Rare Diseases*. BioMed Central Ltd., 12(1), pp. 1–10. doi: 10.1186/s13023-017-0641-1.

Nassisi, M. *et al.* (2018) 'Expanding the mutation spectrum in ABCA4: Sixty novel disease causing variants and their associated phenotype in a large french stargardt cohort', *International Journal of Molecular Sciences*. MDPI AG, 19(8). doi: 10.3390/ijms19082196.

Nassisi, M. *et al.* (2019) 'Prevalence of ABCA4 Deep-Intronic Variants and Related Phenotype in An Unsolved "One-Hit" Cohort with Stargardt Disease'. MDPI AG, 20(20). doi: 10.3390/ijms20205053.

Odorico, J. S., Kaufman, D. S. and Thomson, J. A. (2001) 'Multilineage Differentiation from Human Embryonic Stem Cell Lines', *Stem Cells*. doi: 10.1634/stemcells.19-3-193.

Overlack, N. *et al.* (2012) 'Gene repair of an Usher syndrome causing mutation by zinc-finger nuclease mediated homologous recombination', *Investigative Ophthalmology and Visual Science*. doi: 10.1167/iovs.12-9812.

Pan, Q. *et al.* (2008) 'Deep surveying of alternative splicing complexity in the human transcriptome by high-throughput sequencing', *Nature Genetics*. doi: 10.1038/ng.259.

Papermaster, D. S. *et al.* (1978) 'Immunocytochemical localization of a large intrinsic membrane protein to the incisures and margins of frog rod outer segment disks', *Journal of Cell Biology*. doi: 10.1083/jcb.78.2.415.

Papermaster, D. S., Reilly, P. and Schneider, B. G. (1982) 'Cone lamellae and red and green rod outer segment disks contain a large intrinsic membrane protein on their margins: An ultrastructural immunocytochemical study of frog retinas', *Vision Research*. doi: 10.1016/0042-6989(82)90204-8.

Parmeggiani, F. *et al.* (2011) 'Retinitis Pigmentosa: Genes and Disease Mechanisms', *Current Genomics*, 12(4), pp. 238–249. doi: 10.2174/138920211795860107.

Bibliografía

Permanyer, J. *et al.* (2010) 'Autosomal recessive retinitis pigmentosa with early macular affectation caused by premature truncation in PROM1', *Investigative Ophthalmology and Visual Science*, 51(5), pp. 2656–2663. doi: 10.1167/iovs.09-4857.

Piva, F. *et al.* (2009) 'SpliceAid: A database of experimental RNA target motifs bound by splicing proteins in humans', *Bioinformatics*. doi: 10.1093/bioinformatics/btp124.

Piva, F. *et al.* (2012) 'SpliceAid 2: A database of human splicing factors expression data and RNA target motifs', *Human Mutation*. doi: 10.1002/humu.21609.

Del Pozo-Valero, M. *et al.* (2019) 'Expanded phenotypic spectrum of retinopathies associated with autosomal recessive and dominant mutations in PROM1', *American Journal of Ophthalmology*. doi: 10.1016/j.ajo.2019.05.014.

Del Pozo-Valero, M. *et al.* (2020) 'Genotype-phenotype correlations in a Spanish cohort of 506 families with bi-allelic ABCA4 pathogenic variants', *American Journal of Ophthalmology*. doi: 10.1016/j.ajo.2020.06.027.

Quazi, F., Lenevich, S. and Molday, R. S. (2012) 'ABCA4 is an N-retinylidene-phosphatidylethanolamine and phosphatidylethanolamine importer', *Nature Communications*. doi: 10.1038/ncomms1927.

Rahman, N. *et al.* (2019) 'Macular dystrophies: Clinical and imaging features, molecular genetics and therapeutic options', *British Journal of Ophthalmology*. doi: 10.1136/bjophthalmol-2019-315086.

Ramon y Cajal, S. (1891) 'Notas preventivas sobre la retina y el gran simpático.', in *Barcelona: Casa Provincial de la Caridad*.

Richards, S. *et al.* (2015) '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', *Genetics in Medicine*, 17(5), pp. 405–424. doi: 10.1038/gim.2015.30.

Riveiro-Alvarez, R. *et al.* (2009) 'Frequency of ABCA4 mutations in 278 Spanish controls: An insight into the prevalence of autosomal recessive Stargardt disease', *British Journal of Ophthalmology*. doi: 10.1136/bjo.2008.148155.

Riveiro-Alvarez, R. *et al.* (2013) 'Outcome of ABCA4 disease-associated alleles in autosomal recessive retinal dystrophies: Retrospective analysis in 420 Spanish families', *Ophthalmology*, 120(11), pp. 2332–2337. doi: 10.1016/j.ophtha.2013.04.002.

Rivera, A. *et al.* (2000) 'A comprehensive survey of sequence variation in the ABCA4 (ABCR) gene in Stargardt disease and age-related macular degeneration', *American Journal of Human Genetics*, 67(4), pp. 800–813. doi: 10.1086/303090.

Roosing, S. *et al.* (2014) 'Causes and consequences of inherited cone disorders', *Progress in Retinal and Eye Research*. doi: 10.1016/j.preteyeres.2014.05.001.

Rotenstreich, Y., Fishman, G. A. and Anderson, R. J. (2003) 'Visual acuity loss and clinical observations in a large series of patients with Stargardt disease.', *Ophthalmology*, 110(6), pp. 1151–8. doi: 10.1016/S0161-6420(03)00333-6.

Ruan, G. X. *et al.* (2017) 'CRISPR/Cas9-Mediated Genome Editing as a Therapeutic Approach for Leber Congenital Amaurosis 10', *Molecular Therapy*. doi: 10.1016/j.ymthe.2016.12.006.

Runhart, E. H. *et al.* (2018) 'The common ABCA4 variant p.Asn1868ile shows nonpenetrance and variable expression of stargardt disease when present in trans with severe variants',

198

Investigative Ophthalmology and Visual Science, 59(8), pp. 3220–3231. doi: 10.1167/iovs.18-23881.

Sangermano, R. *et al.* (2016) 'Photoreceptor Progenitor mRNA Analysis Reveals Exon Skipping Resulting from the ABCA4 c.5461-10T→C Mutation in Stargardt Disease', *Ophthalmology*, 123(6), pp. 1375–1385. doi: 10.1016/j.ophtha.2016.01.053.

Sangermano, R. *et al.* (2018) 'ABCA4 midigenes reveal the full splice spectrum of all reported noncanonical splice site variants in Stargardt disease', *Genome Research*, 28(1), pp. 100–110. doi: 10.1101/gr.226621.117.

Sangermano, R. (2018) *Unraveling the missing heritability in ABCA4-associated Stargardt disease*. Radboud University Nijmegen.

Sangermano, R. *et al.* (2019) 'Deep-intronic ABCA4 variants explain missing heritability in Stargardt disease and allow correction of splice defects by antisense oligonucleotides', *Genetics in Medicine*. doi: 10.1038/s41436-018-0414-9.

Sanjurjo-Soriano, C. and Kalatzis, V. (2018) 'Guiding Lights in Genome Editing for Inherited Retinal Disorders: Implications for Gene and Cell Therapy', *Neural plasticity*. doi: 10.1155/2018/5056279.

Schulz, H. L. *et al.* (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 and Visual Science*. doi: 10.1167/iovs.16-19936.

Schwartz, S. D. *et al.* (2016) 'Subretinal transplantation of embryonic stem cell–derived retinal pigment epithelium for the treatment of macular degeneration: An assessment at 4 years', *Investigative Ophthalmology and Visual Science*. doi: 10.1167/iovs.15-18681.

Scotti, M. M. and Swanson, M. S. (2016) 'RNA mis-splicing in disease', *Nature Reviews Genetics*. doi: 10.1038/nrg.2015.3.

Sergouniotis, P. I. *et al.* (2011) 'Biallelic mutations in PLA2G5, encoding group v phospholipase A 2, cause benign fleck retina', *American Journal of Human Genetics*. doi: 10.1016/j.ajhg.2011.11.004.

Sim, N.-L. *et al.* (2012) 'SIFT web server: predicting effects of amino acid substitutions on proteins', *Nucleic Acids Research*, 40(W1), pp. W452–W457. doi: 10.1093/nar/gks539.

Singh, R. K. and Cooper, T. A. (2012) 'Pre-mRNA splicing in disease and therapeutics', *Trends in Molecular Medicine*. doi: 10.1016/j.molmed.2012.06.006.

Sterne-Weiler, T. *et al.* (2011) 'Loss of exon identity is a common mechanism of human inherited disease', *Genome Research*. doi: 10.1101/gr.118638.110.

Stone, E. M. *et al.* (1994) 'Clinical Features of a Stargardt-Like Dominant Progressive Macular Dystrophy With Genetic Linkage to Chromosome 6q', *Archives of Ophthalmology*. doi: 10.1001/archopht.1994.01090180063036.

Strauss, R. W. *et al.* (2016) 'Assessment of estimated retinal atrophy progression in Stargardt macular dystrophy using spectral-domain optical coherence tomography', *British Journal of Ophthalmology*. doi: 10.1136/bjophthalmol-2015-307035.

Tanaka, K. *et al.* (2018) 'The Rapid-Onset Chorioretinopathy Phenotype of ABCA4 Disease', *Ophthalmology*. doi: 10.1016/j.ophtha.2017.07.019.

Tanna, P. *et al.* (2017) 'Stargardt disease: Clinical features, molecular genetics, animal models

and therapeutic options', *British Journal of Ophthalmology*. doi: 10.1136/bjophthalmol-2016-308823.

Trapani, I. *et al.* (2014) 'Effective delivery of large genes to the retina by dual AAV vectors', *EMBO Molecular Medicine*. doi: 10.1002/emmm.201302948.

Tsybovsky, Y. *et al.* (2013) 'Molecular organization and ATP-induced conformational changes of ABCA4, the photoreceptor-specific ABC Transporter', *Structure*. doi: 10.1016/j.str.2013.03.001.

Tsybovsky, Y., Molday, R. S. and Palczewski, K. (2010) 'The ATP-binding cassette transporter ABCA4: Structural and functional properties and role in retinal disease', *Advances in Experimental Medicine and Biology*. doi: 10.1007/978-1-4419-5635-4_8.

Tsybovsky, Y. and Palczewski, K. (2014) 'Expression, purification and structural properties of ABC transporter ABCA4 and its individual domains', *Protein Expression and Purification*. doi: 10.1016/j.pep.2014.02.010.

Urtubia Vicario, C. (1997) *Neurobiología de la visión*. Edited by Universitat Politècnica de Catalunya.

Utz, V. M. *et al.* (2014) 'Predictors of visual acuity and genotype-phenotype correlates in a cohort of patients with Stargardt disease', *British Journal of Ophthalmology*. BMJ Publishing Group, 98(4), pp. 513–518. doi: 10.1136/bjophthalmol-2013-304270.

Valverde, D. *et al.* (2006) 'Microarray-based mutation analysis of the ABCA4 gene in Spanish patients with Stargardt disease: evidence of a prevalent mutated allele.', *Molecular vision*.

Valverde, D. *et al.* (2007) 'Spectrum of the ABCA4 gene mutations implicated in severe retinopathies in Spanish patients.', *Investigative ophthalmology & visual science*, 48(3), pp. 985–90. doi: 10.1167/iovs.06-0307.

Verbakel, S. K. *et al.* (2018) 'Non-syndromic retinitis pigmentosa', *Progress in Retinal and Eye Research*. doi: 10.1016/j.preteyeres.2018.03.005.

Wang, E. T. *et al.* (2008) 'Alternative isoform regulation in human tissue transcriptomes', *Nature*. doi: 10.1038/nature07509.

Wang, Z. and Burge, C. B. (2008) 'Splicing regulation: From a parts list of regulatory elements to an integrated splicing code', *RNA*. doi: 10.1261/rna.876308.

Weigmann, A. *et al.* (1997) 'Prominin, a novel microvilli-specific polytopic membrane protein of the apical surface of epithelial cells, is targeted to plasmalemmal protrusions of non-epithelial cells', *Proceedings of the National Academy of Sciences*, 94(23), pp. 12425–12430. doi: 10.1073/pnas.94.23.12425.

Westeneng-van Haaften, S. C. *et al.* (2012) 'Clinical and genetic characteristics of late-onset Stargardt's disease.', *Ophthalmology*, 119(6), pp. 1199–210. doi: 10.1016/j.ophtha.2012.01.005.

Wolock, C. J. *et al.* (2019) 'A case–control collapsing analysis identifies retinal dystrophy genes associated with ophthalmic disease in patients with no pathogenic ABCA4 variants', *Genetics in Medicine*. doi: 10.1038/s41436-019-0495-0.

Xin, W. *et al.* (2015) 'Identification of genetic defects in 33 probands with Stargardt disease by WES-based bioinformatics gene panel analysis', *PLoS ONE*. doi: 10.1371/journal.pone.0132635.

Yamamoto, T. *et al.* (2016) 'Challenges in detecting genomic copy number aberrations using next-generation sequencing data and the eXome Hidden Markov Model: A clinical exome-first diagnostic approach', *Human Genome Variation*. Nature Publishing Group, 3. doi:

10.1038/hgv.2016.25.

Yang, Z. *et al.* (2008) 'Mutant prominin 1 found in patients with macular degeneration disrupts photoreceptor disk morphogenesis in mice', *Journal of Clinical Investigation*, 118(8), pp. 2908–2916. doi: 10.1172/JCI35891.

Yatsenko, A. N. *et al.* (2001) 'Late-onset Stargardt disease is associated with missense mutations that map outside known functional regions of ABCR (ABCA4)', *Human Genetics*. doi: 10.1007/s004390100493.

Yatsenko, A. N. *et al.* (2003) 'An ABCA4 genomic deletion in patients with Stargardt disease', *Human Mutation*. doi: 10.1002/humu.10219.

Yu, W. *et al.* (2017) 'Nrl knockdown by AAV-delivered CRISPR/Cas9 prevents retinal degeneration in mice', *Nature Communications*. doi: 10.1038/ncomms14716.

Zacchigna, S. *et al.* (2009) 'Loss of the cholesterol-binding protein Prominin-1/CD133 causes disk dysmorphogenesis and photoreceptor degeneration', *Journal of Neuroscience*. doi: 10.1523/JNEUROSCI.2034-08.2009.

Zaneveld, J. *et al.* (2015) 'Comprehensive analysis of patients with Stargardt macular dystrophy reveals new genotype-phenotype correlations and unexpected diagnostic revisions', *Genetics in Medicine*. doi: 10.1038/gim.2014.174.

Zernant, J. *et al.* (2011) 'Analysis of the ABCA4 gene by next-generation sequencing', *Investigative Ophthalmology and Visual Science*. doi: 10.1167/iovs.11-8182.

Zernant, J. *et al.* (2014) 'Analysis of the ABCA4 genomic locus in Stargardt disease', *Human molecular genetics*, 23(25), pp. 6797–6806. doi: 10.1093/hmg/ddu396.

Zernant, J. *et al.* (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), pp. 404–412. doi: 10.1136/jmedgenet-2017-104540.

Zernant, J. *et al.* (2018) 'Extremely hypomorphic and severe deep intronic variants in the ABCA4 locus result in varying Stargardt disease phenotypes', *Cold Spring Harbor molecular case studies*. doi: 10.1101/mcs.a002733.

Zhang, N. *et al.* (2014) 'Protein misfolding and the pathogenesis of ABCA4-associated retinal degenerations', *Human Molecular Genetics*. doi: 10.1093/hmg/ddv073.

Zhang, X. *et al.* (2014) 'Molecular diagnosis of putative Stargardt disease by capture next generation sequencing', *PLoS ONE*, 9(4), p. e95528. doi: 10.1371/journal.pone.0095528.

Anexos

**Anexo I: Variantes que alteran el *splicing* en *ABCA4***

Variantes de sitios de *splicing* no canónicos (NCSS) en *ABCA4* que alteran el *splicing*					
Exón Intrón	**Nucleótido**	**Referencia**	**Exón Intrón**	**Nucleótido**	**Referencia**
IVS1	c.160+5G>C	Sangermano et al, 2018	30	c.4538A>G	Sangermano et al, 2018
	c.161G>A	Fadaie et al, 2019		c.4538A>C	Sangermano et al, 2018
	c.161G>T	Sangermano et al, 2018	IVS30	c.4540-8T>A	Khan et al, 2020
IVS2	c.302+4A>C	Sangermano et al, 2018	32	c.4667G>A	Fadaie et al, 2019
	c.303-3C>G	Sangermano et al, 2018		c.4667G>C	Sangermano et al, 2018
6	c.768G>T	Sangermano et al, 2018	33	c.4773G>C	Sangermano et al, 2018
IVS7	c.859-9T>C	Sangermano et al, 2018	IVS33	c.4773+3A>G	Sangermano et al, 2018
IVS8	c.1100-6T>A	Sangermano et al, 2018		c.4773+5G>A	Sangermano et al, 2018
IVS11	c.1554+3A>T	Khan et al, 2020	IVS34	c.4848+3A>G	Khan et al, 2020
IVS13	c.1937+5G>A	Fadaie et al, 2019	IVS35	c.5018+5G>A	Fadaie et al, 2019
IVS14	c.2161-8G>A	Fadaie et al, 2019	IVS36	c.5196+3_5196+6del	Sangermano et al, 2018
IVS15	c.2382+5G>C	Sangermano et al, 2018	IVS37	c.5312+3A>T	Sangermano et al, 2018
17	c.2588G>C	Sangermano et al, 2018		c.5313-3C>G	Sangermano et al, 2018
IVS17	c.2654-8T>G	Khan et al, 2020	IVS38	c.5460+5G>A	Sangermano et al, 2018
IVS19	c.2919-10T>C	Sangermano et al, 2018		c.5461-10T>C	Sangermano et al, 2018
	c.2919-6C>A	Sangermano et al, 2018		c.5461-8T>G	Sangermano et al, 2018
IVS20	c.3050+5G>A	Sangermano et al, 2018		c.5461-6T>G	Khan et al, 2020
IVS23	c.3522+5del	Sangermano et al, 2018	39	c.5584G>C	Sangermano et al, 2018
IVS24	c.3607G>A	Sangermano et al, 2018	IVS39	c.5584+5G>A	Sangermano et al, 2018
	c.3607+3A>T	Sangermano et al, 2018		c.5584+6T>C	Sangermano et al, 2018
25	c.3812A>G	Sangermano et al, 2018	IVS40	c.5714+5G>A	Sangermano et al, 2018
	c.3813G>C	Sangermano et al, 2018		c.5715-5T>G	Fadaie et al, 2019
26	c.3862G>A	Khan et al, 2020	IVS41	c.5836-3C>A	Sangermano et al, 2018
IVS26	c.3862+3A>G	Sangermano et al, 2018	IVS42	c.5898+5G>A	Khan et al, 2020
27	c.4128G>A	Sangermano et al, 2018		c.5898+5del	Sangermano et al, 2018
IVS27	c.4129-3C>T	Khan et al, 2020	44	c.6147G>A	Fadaie et al, 2019
IVS28	c.4253+4C>T	Sangermano et al, 2018	46	c.6385A>G	Fadaie et al, 2019
	c.4253+5G>A	Sangermano et al, 2018	IVS46	c.6386+3A>G	Khan et al, 2020
	c.4253+5G>T	Sangermano et al, 2018	47	c.6478A>G	Sangermano et al, 2018
			IVS47	c.6479+4A>G	Sangermano et al, 2018
			IVS48	c.6729+5_6729+19del	Sangermano et al, 2018

Variantes *deep intronic* en *ABCA4* que alteran el *splicing*					
Intron	**Nucleótido**	**Referencia**	**Intron**	**Nucleótido**	**Referencia**
IVS1	c.67-2023T>G	Khan et al, 2020	IVS21	c.3191-11T>A	Bauwens et al, 2019
IVS2	c.161-23T>G	Bauwens et al, 2019	IVS26	c.3863-1064A>G	Khan et al, 2020
IVS5	c.570+1798A>G	Khan et al, 2020	IVS28	c.4253+43G>A	Sangermano et al, 2019
IVS6	c.769-784C>T	Sangermano et al, 2018	IVS29	c.4352+61G>A	Fadaie et al, 2019
	c.769-788A>T	Khan et al, 2020	IVS30	c.4539+1100A>G	Sangermano et al, 2019
IVS7	c.859-640A>G	Khan et al, 2020		c.4539+1106C>T	Sangermano et al, 2019
	c.859-546G>A	Khan et al, 2020		c.4539+2001G>A	Braun et al, 2013
	c.859-540C>G	Bauwens et al, 2019		c.4539+2028C>T	Braun et al, 2013
	c.859-506G>C	Sangermano et al, 2019		c.4539+2064C>T	Bauwens et al, 2019
IVS13	c.1937+13T>G	Sangermano et al, 2018		c.4539+2065C>G	Khan et al, 2019
	c.1937+37C>G	Khan et al, 2020	IVS31	c.4634+741A>G	Khan et al, 2020
	c.1937+435C>G	Sangermano et al, 2019	IVS36	c.5196+1056A>G	Braun et al, 2013
	c.1938-621G>A	Khan et al, 2020		c.5196+1137G>A	Braun et al, 2013
	c.1938-619T>A	Fadaie et al, 2019		c.5196+1216C>A	Braun et al, 2013
	c.1938-514A>G	Khan et al, 2020		c.5197-557G>T	Bauwens et al, 2019
IVS16	c.2588-706C>T	Khan et al, 2020	IVS44	c.6148-84A>T	Khan et al, 2020
IVS19	c.2919-826T>A	Fadaie et al, 2019	IVS45	c.6283-78G>T	Khan et al, 2020
IVS20	c.3050+370C>T	Fadaie et al, 2019			

Anexo II: Variantes de *ABCA4* estudiadas funcionalmente en las células HEK293T que no alteran el *splicing*

Variante	Efecto predicho	Frecuencia (gnomAD)	SSF [0-100]	MaxEnt [0-12]	NNSPLICE [0-1]	GeneSplicer [0-24]	HSF [0-100]
c.160+367G>A	Creación SAS	0.001913 (ALL) 0.003243 (NFE)	0→77	0→0	0→0.1	0→0	0→85.9
c.768+5203C>T	Creación SDS	Ausente	57.6→62.7	0→2.4	0→0.1	0→0	0→73.1
c.769-7025G>A	Creación SDS	0.008362 (ALL) 0.01332 (NFE)	67→77	0.3→3.1	0→0.3	0→0	79.8→88.1
c.1554+3771C>T	Creación SDS	0.0002230 (ALL) 0.0003889 (NFE)	51.5→64.9	0→1.9	0→0.1	0→0	0→79.2
c.1937+376G>A	Creación SAS	Ausente	0→0	0→2.8	0→0.5	0→0	0→79.3
c.1937+630T>G	Creación SDS	Ausente	65.8→71.3	0→2.2	0→0	0→0	78.9→82.8
c.1938-447G>A	Creación SAS	0.001115 (ALL)	0→69.6	0→1.4	0→0	0→0	0→77.4
c.1938-436T>A	Creación SAS	Ausente	0→67.8	0→4.7	0→0	0→0	0→80.1
c.2654-193A>C	Creación SAS	0.001274 (ALL)	62→71.9	0→0	0→0	0→0	71.7→81.1
c.2918+34C>T	Creación SDS	0.0004501 (All) 0.00003889 (NFE)	66.7→71.7	0→0	0→0	0→0	0→81.9
c.3328+305A>T	Creación SAS	0.001911 (ALL) 0.001362 (NFE)	61.2→73.4	2.3→3.6	0→0	0→2.2	81→84.3
c.6147+1155G>C	Creación SAS	0.00003185 (ALL)	0→63.2	0→3.6	0→0	0→0	72.9→74.8

Anexo III: Algoritmo diagnóstico propuesto para el estudio de pacientes con sospecha de enfermedad de Stargardt
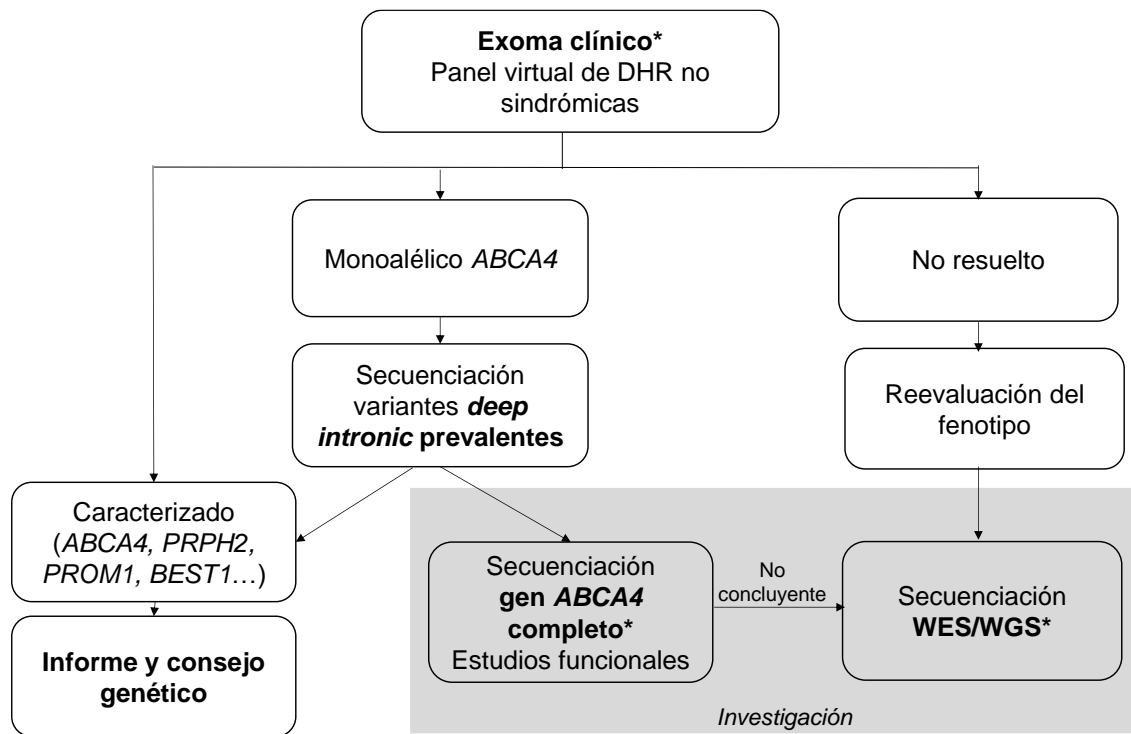
*El análisis de CNVs está incluido en los algoritmos diagnósticos utilizados.

Anexo IV: Listado de publicaciones derivadas de esta Tesis Doctoral.

Martin-Merida I, Avila-Fernandez A, <u>Del Pozo-Valero M</u>, et al. Genomic Landscape of Sporadic Retinitis Pigmentosa. *Ophthalmology. 2019*.

<u>Del Pozo-Valero M</u>, Martin-Merida I, Jimenez-Rolando B, et al. Expanded phenotypic spectrum of retinopathies associated with autosomal recessive and dominant mutations in *PROM1*. *Am J Ophthalmol. 2019*.

Fadaie Z, Khan M, <u>Del Pozo-Valero M</u>, et al. Identification of splice defects due to noncanonical splice site or deep-intronic variants in *ABCA4*. *Hum Mutat. 2019*.

Beigi F*, <u>Del Pozo-Valero M</u>*, Martin-Merida I, et al. Posterior column ataxia with retinitis pigmentosa (PCARP) in an Iranian patient associated with the *FLVCR1* gene. *Ophthalmic Genet. 2020.*

Khan M, Cornelis SS, <u>Del Pozo-Valero M</u>, et al. Resolving the dark matter of *ABCA4* for 1054 Stargardt disease probands through integrated genomics and transcriptomics. *Genet Med. 2020.*

<u>Del Pozo-Valero M</u>, Riveiro-Alvarez R, Blanco-Kelly F, et al. Genotype-phenotype correlations in a Spanish cohort of 506 families with bi-allelic *ABCA4* pathogenic variants. *Am J Ophthalmol. 2020.*

Runhart EH, Khan M, Cornelis SS, et al. Association of Sex With Frequent and Mild *ABCA4* Alleles in Stargardt Disease. *JAMA Ophthalmol. 2020*.