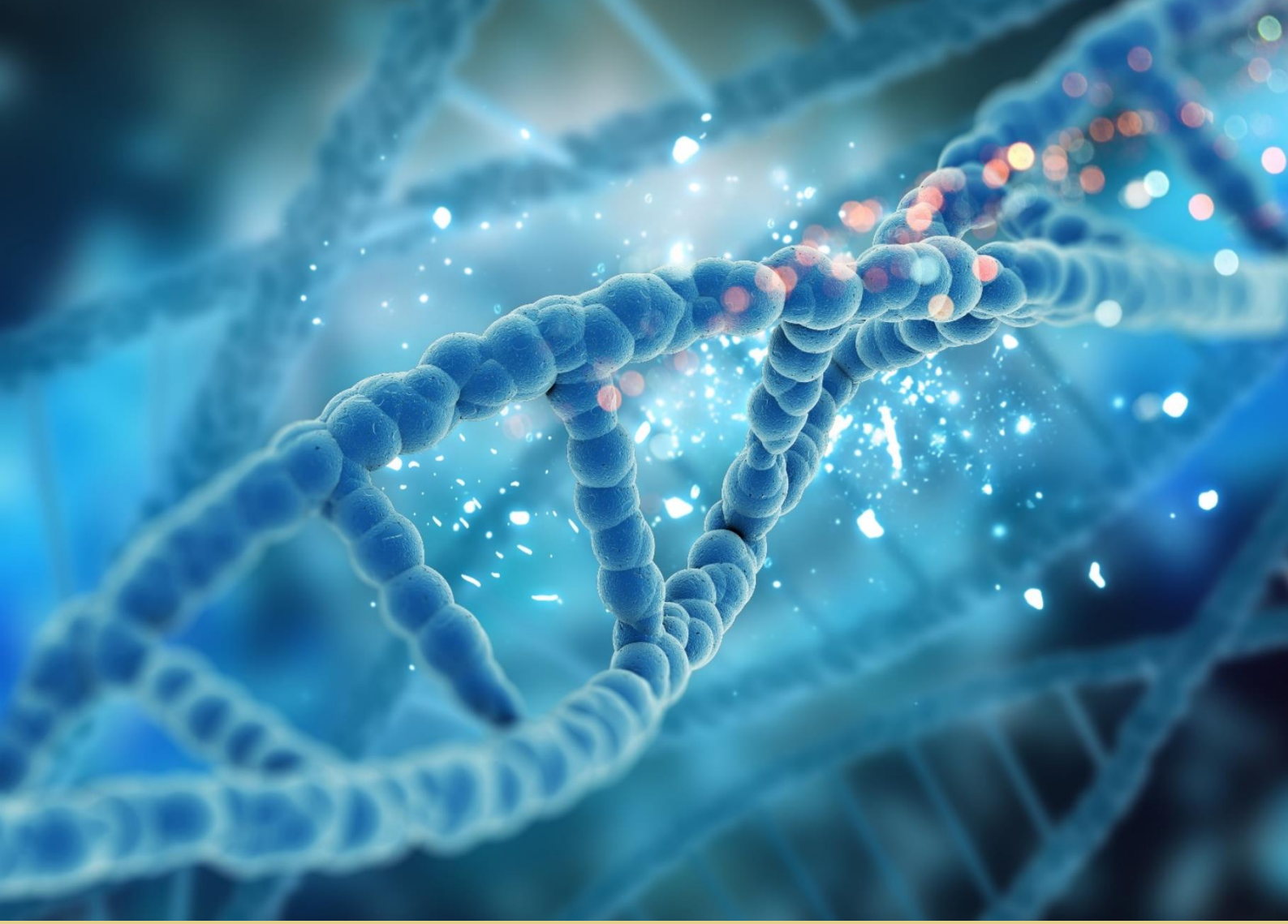


Susana Margarida Martins Carmona

Identification of candidate genes associated with primary congenital glaucoma in patients negative for *CYP1B1* mutations.

Tese de doutoramento do Programa de Doutoramento em Ciências da Saúde, Ciências Biomédicas,
orientada pelo Professor Doutor Eduardo Silva e pela Doutora Conceição Egas,
apresentada à Faculdade de Medicina da Universidade de Coimbra

Janeiro de 2017

· U C ·

Universidade de Coimbra

Identification of candidate genes associated with primary congenital glaucoma in patients negative for *CYP1B1* mutations.

Susana Margarida Martins Carmona

Tese apresentada à Faculdade de Medicina da Universidade de Coimbra para prestação de provas de doutoramento em Ciências da Saúde, ramo de Ciências Biomédicas.

Thesis submitted to the Faculty of Medicine of the University of Coimbra for the attribution of the Doctor degree in Health Sciences, in the field of Biomedical Sciences.

January 2017

UNIVERSIDADE DE COIMBRA

Identification of candidate genes associated with primary congenital glaucoma in patients negative for *CYP1B1* mutations.

The research work presented in this thesis was performed at Biocant - Technology Transfer Association, Portugal, UC-Biotech/Center for Neuroscience and Cell Biology of the University of Coimbra, Portugal and the Feinberg Cardiovascular Research Institute and Division of Nephrology/Hypertension of the Northwestern University, Feinberg School of Medicine, Chicago, USA, under the supervision of Professor Eduardo Silva and Doctor Conceição Egas.

O trabalho experimental apresentado nesta tese foi elaborado no Biocant - Associação de Transferência de Tecnologia, Portugal, UC-Biotech/ Centro de Neurociências e Biologia Celular da Universidade de Coimbra, Portugal e no Instituto de Investigação Cardiovascular Feinberg e Departamento de Nefrologia/Hypertensão da Universidade de Northwestern, Escola de Medicina Feinberg, Chicago, Estados Unidos da América, sob a supervisão do Professor Doutor Eduardo Silva e da Doutora Conceição Egas.

Susana Carmona was the recipient of an FCT Ph.D. fellowship SFRH/BD/90445/2012.

Susana Carmona foi suportada por uma bolsa de doutoramento da FCT, referência SFRH/BD/90445/2012.

Front cover:

DNA molecule image designed by Kjpargeter - Freepik.com

"*Rather than believe that Watson and Crick made the DNA structure, I would rather stress that the structure made Watson and Crick.*"

Francis Crick, What Mad Pursuit (1988)

Acknowledgements / Agradecimentos

Ao Professor Doutor Eduardo Silva, à Doutora Conceição Egas e à Doutora Maria José Simões o meu mais sincero agradecimento pela confiança que depositaram em mim desde o primeiro minuto, dando-me a oportunidade de fazer convosco o meu doutoramento. Ao longo dos últimos quatro anos sempre mostraram disponibilidade, apoiando-me em todos os momentos e depositando em mim a confiança necessária para levar a cabo este projecto. Obrigada pelos ensinamentos científicos que me transmitiram e por me ajudarem a escolher o caminho a seguir em cada passo, sem os quais este trabalho não se teria concretizado. Obrigada pelas palavras de encorajamento que tanto me ajudaram ao longo desta importante etapa.

Ao Professor Doutor Carlos Faro agradeço a oportunidade de poder desenvolver este trabalho no Biocant – Associação de Transferência de Tecnologia. To Professor Susan Quaggin, I would like to acknowledge the possibility of doing part of the research work in the Feinberg Cardiovascular Research Institute and Division of Nephrology/Hypertension of the Northwestern University, Feinberg School of Medicine.

À Fundação para a Ciência e Tecnologia agradeço o seu financiamento sem o qual esta tese não teria sido possível.

À Ana Luisa Carvalho agradeço todo o esforço prestado para conseguir aumentar o número de doentes deste estudo, bem como a disponibilidade total demonstrada sempre que surgia qualquer questão.

To Doctor Tomokasu Souma, I would like to acknowledge his contribution to the design and teaching of the experiments required for *TEK* validation, as well as for his kindness, friendship and encouraging words in the final step of this thesis. It was a pleasure to work with you.

To Doctor Martin Reijns, I would like to acknowledge all the provided information regarding *RNASEH2C* and his help in the interpretation of the results.

To Professor Alan Prescot, I would like to acknowledge the sending of the Human Trabecular Meshwork Cell Line.

Ao Hugo Foufe e Felipe Santos agradeço toda a ajuda na realização da parte Bioinformática desta tese.

A todos os meus colegas e amigos do Biocant, em especial aos da Unidade de Sequenciação, não esquecendo todos os que por lá passaram durante a minha estadia, agradeço a amizade, boa disposição, ajuda e encorajamento que tornaram todo este percurso a melhor experiência possível. Agradeço especialmente aos "residentes" ainda não mencionados, à Cristina Barroso,

pela ajuda laboratorial sempre que solicitada e ao Diogo Pinho, não só pela ajuda laboratorial como por tudo o que estivesse relacionado com o mundo de um doutoramento.

A todos os meus amigos, em especial à Raquel Belo, à Lia Felipe, à Ângela Moreira, à Andreia Neves e ao Tiago Sousa agradeço pela partilha de bons momentos que me fizeram esquecer as preocupações relacionadas com esta tese e pela ajuda prestada em situações relacionadas com este doutoramento. À Teresa Carminho e restante grupo de amigas do Serviço de Genética do Hospital Pediátrico agradeço todo o entusiasmo que sempre sentiram e me transmitiram. Agradeço especialmente ao Abel Sousa, que conheci durante o desenrolar deste trabalho e que se tornou no meu mais recente e melhor amigo. Obrigada por todos os bons momentos que passamos juntos, por ouvires os meus "dramas" cientifícos e pessoais e por, mesmo longe, ainda continuares a fazer parte da minha vida.

Por fim, um agradecimento especial aos meus familiares que sempre me incentivaram durante todo este percurso. Aos meus pais por todo o apoio e incentivo e à Janete, o novo elemento da família, por todos os conselhos, em especial os relacionados com a estadia em Chicago. E claro ao meu irmão Vitor que tanto me apoiou em todos os bons e maus momentos, desde a hora em que me informou da minha aceitação para Bolsa de Doutoramento até ao dia da entrega desta tese. Obrigada pelas imensas ajudas científicas que me deste, irmão, e por teres suportado os meus imensos dramas. Sem a tua ajuda este trabalho teria sido mais difícil. Um agradecimento final à minha cadelinha Moquinhas, que durante os últimos momentos deste doutoramento me fez lembrar que há mais coisas para além do trabalho.

Esta tese é-vos dedicada.

Remissive Index

List of Abbreviations

a.a.: Amino acid

AGS: Aicardi-Goutières syndrome

Ang1: Angiopoietin-1

Ang2: Angiopoietin-2

ARS: Axenfeld-Rieger syndrome

BSA: Bovine serum albumin

ExAC: Exome Aggregation Consortium

BWA: Burrows-Wheeler Aligner

CNV: Copy number variation

ddG: free protein energy

DMEM: Dulbecco's modified Eagle's medium

DMSO: Dimethyl sulfoxide

ESE: Exonic splicing enhancer

ESP: Exome Sequencing Project

ESS: Exonic splicing silencer

FHD: Forkhead domain

GATK: Genome Analysis Tool Kit

HPO: Human phenotype ontology

HRP: Horseradish peroxidase

HTM: Human trabecular cell

IGF1: Insulin-like growth factor I

IGV: Integrative Genomics Viewer

IOP: Intraocular pressure

ISE: Intronic splicing enhancer

ISS: Intronic splicing silencer

JCT: Juxtacanalicular tissue

JOAG: Juvenile-onset primary open angle glaucoma

LoF: Loss-of-function

MAF: Minor allele frequency

mRNA: Messenger RNA

MRP4: Multidrug, resistance-associated protein-4

NGS: Next-generation sequencing

PACG: Primary angle closure glaucoma

PCG: Primary congenital glaucoma

POAG: Primary open angle glaucoma

PVDF: Polyvinylidene fluoride

SC: Schlemm's canal

SDS-PAGE: Sodium dodecyl sulfate – polyacrylamide gel electrophoresis

SNP: Single nucleotide polymorphism

TBS: Tris-buffered saline

TM: Trabecular meshwork

UTR: Untranslated region

WES: Whole exome sequencing

1000 G: 1000 Genome Project

Abstract

Primary congenital glaucoma (PCG) is a optic neuropathy characterized by high intra-ocular pressure (IOP) due to isolated trabeculodysgenesis during the neonatal or early infantile period. Classified as a rare disease in non-consanguineal populations, PCG assumes particular relevance in consanguineal populations, being one of the leading causes of blindness in children.

The first genetic studies identified loci GLC3A, GLC3B, and GLC3C as associated with PCG and *CYP1B1* as a causative gene. Posterior studies confirmed involvement of *CYP1B1*. In fact, mutations in this gene are the principal genetic cause of PCG,. However, a significant number of patients remain without identification of the genetic cause. Several studies have tried to discover new mutated genes in these patients, but only a few cases associated with *LTBP2*, *MYOC*, *FOXC1*, and *TEK* have been reported.

In this work, a Portuguese group of 37 PCG patients was studied for the genetic cause behind PCG development. The search for *CYP1B1* mutations revealed a high number of patients with *CYP1B1* mutations, consistent with the recessive inheritance model. Nine previously reported *CYP1B1* mutations (p.179RfsX18, p.R355HfsX69, p.T404SfsX30, p.D449MfsX8, p.L378Q, p.E387K, p.P437L, p.R444X, and R468_S476dup) were found in sixty-five percent of the patients. No mutations were identified in the non-coding regions of this gene.

The negative *CYP1B1* patients were then sequenced for the whole exome to identify new PCG causal genes. Another negative *CYP1B1* patient with congenital glaucoma and other symptoms was also sequenced. Rare and predicted pathogenic exonic and splice site variants, with the exclusion of synonymous variants, were selected.

Analysis of the PCG candidate genes revealed the novel *TEK* p.V188G heterozygous variant in patient 4. Experimental validation proved its pathogenicity involving a reduction of protein expression due to formation of aggregates that are degraded by the ubiquitin-proteasome pathway, and a reduction of autophosphorylation. Patient 4 also had the *MYOC* p.A445V variant that in addition to the pathogenic *TEK* variant may be determinant for PCG phenotype.

The second search approach was based on the recessive inheritance model accepted for PCG. The homozygous *RNASEH2C* c.G468T variant was found in patient 31. This variant was predicted to disrupt exon 3 donor splice site. *RNASEH2C* is a causative gene of Aicardi-Goutières syndrome, a rare early-onset encephalopathy that usually results in severe intellectual and physical handicap. Congenital glaucoma is one of the associated phenotypes. Since mild forms of the disease have been described and are compatible with symptoms

presented in patient 31, it was hypothesized that *RNASEH2C* c.G468T variant might be causative of the phenotype of this patient. However, the experimental study of the *RNASEH2C* altered transcript did not reveal significant differences comparatively to the wild-type. The *BCOR* p.S1298F and the *TNC* p.R103H and p.V295M variants were identified in patient 33 and patient 41, respectively. *BCOR* is related with ocular syndromes, and *TNC* encodes a component of the trabecular meshwork, the structure involved in PCG pathogenesis.

The absence of strong candidates for all samples conducted to the analysis based on the dominant inheritance model. A possible digenic model involving *S1PR3* and *MMP1* or *NEDD4* was found for patient 1. The novel p.R596T in *TRPM3*, a gene proposed to have an impact in IOP, was identified in patient 25. Patient 45 had the novel *ABCC4* p.G529E variant. This gene is expressed in HTM cells and participates in IOP regulation.

Secondary forms of congenital glaucoma and other ocular conditions resembling some of the PCG symptoms may lead to the incorrect diagnosis of PCG. In these situations, genetic information can be useful to reach the accurate diagnosis. Genes related to these specfic conditions were selected and evaluated. An additional patient with ocular and non-ocular features, suggestive of a syndrome associated with secondary congenital glaucoma, was also studied. In fact, the presence of the novel *de novo FOXC1* p.Y64X in this patient confirmed the secondary cause of congenital glaucoma related to Axenfeld-Rieger syndrome. In the remaining patients, only p.Q309H in *ELP4* could be associated with phenotype of patient 25.

The absence of a commonly mutated gene in the majority of the samples suggests the heterogeneity of PCG and the existence of other genetic mechanisms involved in PCG pathogenesis. Still, this study contributed not only to the confirmation of *TEK* in PCG pathogenesis as also to the identification of possible candidate genes that, in the future, may be confirmed as contributing to PCG development. The identification of new genes will lead a deeper comprehension of the molecular mechanisms behind PCG pathogenesis, increasing the possibility of development of new effective PCG therapeutic approaches.

Resumo

O glaucoma congénito primário (GCP) é uma neuropatia ótica que se caracteriza por um aumento da pressão intraocular (PIO) devido a uma trabeculodisgénese durante o período neonatal ou início do período infantil. Apesar de ser considerada uma doença rara em populações não consanguíneas, o GCP assume especial relevância em populações consanguíneas, sendo uma das principais causas de cegueira infantil.

Os primeiros estudos genéticos desta doença identificaram os loci GLC3A, GLC3B, e GLC3C e o gene causal *CYP1B1*. Mutações no gene *CYP1B1* revelaram ser a principal causa genética de GCP,. Contudo, um número significativo de doentes permanece sem causa genética conhecida. Vários estudos têm sido desenvolvidos com o intuito de descobrir novos genes mutados nestes doentes, mas apenas foram descritos alguns casos associados aos genes *LTBP2*, *MYOC*, *FOXC1* e *TEK*.

No decorrer deste trabalho, 37 doentes com GCP foram estudados para caracterização genética da doença. A pesquisa de mutações no gene *CYP1B1* identificou um elevado número de doentes com mutações. Nove mutações já descritas (p.179RfsX18, p.R355HfsX69, p.T404SfsX30, p.D449MfsX8, p.L378Q, p.E387K, p.P437L, p.R444X, e R468_S476dup) foram encontradas em 65% dos doentes, seguindo o modelo de doença autossómica recessiva. Não foram detetadas mutações nas regiões não codificantes.

De forma a identificar as causas genéticas do GCP nestes doentes sem mutações no gene *CYP1B1*, procedeu-se à sequenciação dos seus exomas. Um doente negativo para mutações no *CYP1B1* e com outros sintomas para além de GCP foi igualmente estudado. Numa fase inicial da análise, as alterações exónicas e de *splicing*, raras e previstas como patogénicas, foram selecionadas e as sinónimas excluídas.

Uma primeira abordagem baseada na pesquisa de alterações nos genes candidatos identificou no doente 4 a alteração nova p.V188G no gene *TEK*, em heterozigotia. Estudos experimentais confirmaram a patogenicidade desta alteração através da diminuição da expressão da proteina, formação de agregados, posteriormente degradados pela via ubiquitina-proteasoma, e uma redução da auto-fosforilação. Este doente revelou ainda a alteração p.A445V no gene *MYOC*, que na presença da alteração do gene *TEK* poderá ter impacto para o desenvolvimento do GCP neste doente.

A segunda abordagem teve por base o modelo recessivo aceite para o CGP. Seguindo esta estratégia, no doente 31 foi identificada a alteração c.G468T no gene *RNASEH2C,* prevista como levando à perda do dador de splicing do exão 3. Este gene está associado ao desenvolvimento do síndrome de Aicardi-Goutières, uma encefalopatia severa o associada por

vezes a glaucoma congénito. Formas mais ligeiras da doença já foram descritas, sendo essas compatíveis com os sintomas do doente 31. Assim sendo, considerou-se a hipótese desta alteração estar associada ao fenótipo do doente 31. Contudo, a validação experimental não revelou diferenças entre os transcriptos com e sem a alteração. Outros dois genes identificados com alterações foram o gene *BCOR* com a alteração p.S1298F no doente 33 e o gene *TNC* com as alterações p.R103H e p.V295M no doente 41. *BCOR* está associado a fenótipos oculares, enquanto *TNC* codifica um dos componentes da malha trabecular, a principal estrutura ocular envolvida no GCP.

A ausência de um forte gene candidato para todas as amostras conduziu a uma análise baseada no modelo dominante. No doente 1 foi identificado um possível modelo digénico envolvendo os genes *S1PR3* e *MMP1* ou *NEDD4*. A alteração p.R596T no gene *TRPM3* foi identificada no doente 25 e o doente 45 apresentou a alteração p.G529E no gene *ABCC4*. Ambos os genes foram descritos como envolvidos na regulação da PIO.

O diagnóstico de glaucoma secundário ou de doenças com características semelhantes aos do CGP é por vezes difícil, podendo levar a um diagnóstico não assertivo de CGP. Nestas situações, os dados genéticos podem auxiliar o diagnóstico clínico. Assim sendo, genes associados a estas doenças foram avaliados. Um doente com glaucoma congénito e outros sintomas indicativos de um sindrome associado a glaucoma secundário foi também estudado. Esta suspeita foi confirmada com a identificação da nova alteração *de novo* p.Y64X no gene *FOXC1*, indicativa do síndrome de Axenfeld-Riger. Relativamente aos restantes doentes, a alteração p.Q309H no gene *ELP4* poderá ter impacto no fenótipo do doente 25.

A ausência de um gene mutado comum à maioria dos doentes sugere a heterogeneidade do GCP, bem como o envolvimento de outros mecanismos genéticos. Ainda assim, este estudo permitiu a confirmação do envolvimento do gene *TEK* no GCP, bem como a identificação de possíveis genes candidatos, que, no futuro, poderão revelar-se como tendo um papel importante no desenvolvimento do GCP. A identificação de novos genes permitirá uma compreensão mais extensa dos mecanismos moleculares inerentes ao GCP, o que aumentará a possibilidade de desenvolvimento de novas terapias.

CHAPTER 1

Introduction

1.1. Human eye

One of the most precious senses is vision, and the eye is the organ that allows us to have this amazing experience. This important sensory organ is composed not only by the eyeball, but also of protective structures as the eyelids, the eyebrow, the lacrimal apparatus, the conjunctiva, and the orbit. The eye movement apparatus is formed by the extrinsic ocular muscles and Tenon's capsule (Schlote et al., 2006; Velayutham, 2009).

The eyeball has three layers, the sclera and the cornea (located anteriorly to the sclera) form the outer layer, the vascular choroid the middle layer, and the neurosensory retina the inner layer (Velayutham, 2009). There are two chambers in the eyeball filled with aqueous humor, the anterior and the posterior chambers. The cornea (anteriorly) and the iris (posteriorly) delimit the anterior chamber, which also includes the trabecular meshwork (TM). The posterior chamber is the space between the posterior face of the iris and the anterior part of the lens and the ciliary body. The vitreous is located posteriorly to the lens that is a transparent, biconvex structure that helps to refract light. The vitreous is a transparent and colorless gelatinous mass and supports the ciliary body and the retina. The retina is a multilayer tissue located on the inner surface of the eye and has two main components: the neurosensory retina (composed by the photoreceptors rods and cones, among other cells) and the retinal pigmented epithelium. The optic nerve is a central nerve, with axons originated from the retinal ganglion cells, and courses posteriorly into the brain (Figure 1.1) (Velayutham, 2009; Cummings, 2013).

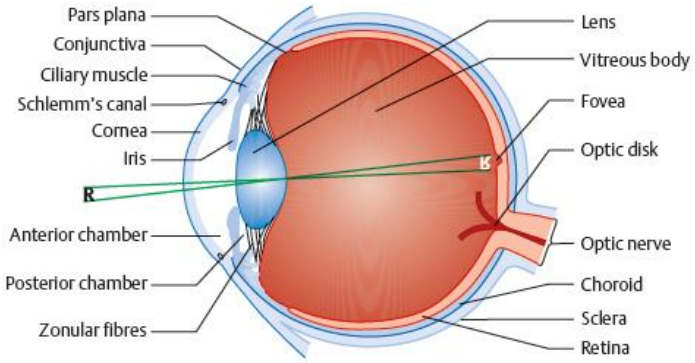


Figure 1.1 – Structure of the eyeball and vision process. The light crosses the eye from the cornea to the retina, creating an inverted image (R). The image information is then transferred to the brain through the optic nerve. Image from Schlote *et al.*, 2006.

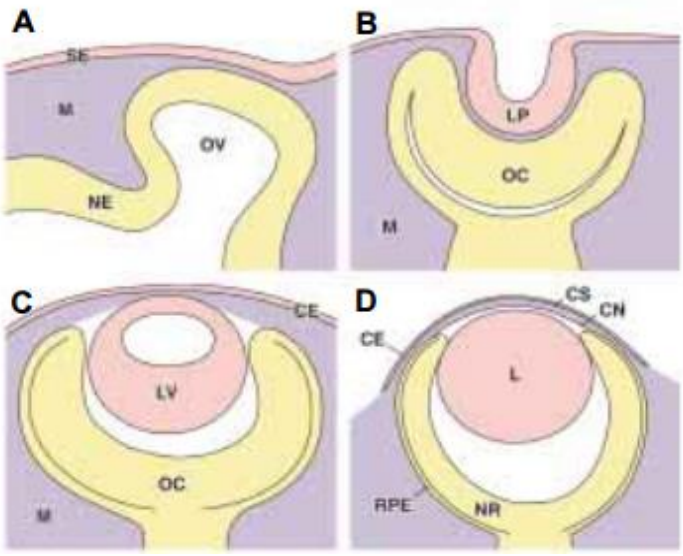
In the vision process, the light rays pass through the cornea, the center of the iris (pupil), through the lens and are gathered and focused onto the retina. The photoreceptors convert

the light stimuli into cellular signals releasing neurotransmitters, and these signals are sent through the optic nerve to higher brain areas (Figure 1.1) (Schlote et al., 2006).

1.1.1.Embryology of the eye and structures of the anterior chamber

The first molecular evidence of eye development is a single structure called the eye field that forms within the neural field on the dorsum of the early embryo. Then eye field is bisected in the midline to lead to the formation of the two separate eyes. All these processes are regulated by the expression of transcription factors (the eye field transcription factors)(Hoyt and Taylor, 2013).

Formation of the optic vesicle that evaginates from the forebrain towards the surface ectoderm starts by the beginning of the 4th week (Figure 1.2A). When the optic vesicle reaches the surface ectoderm, the lens placode is formed from the invagination of the surface ectoderm (Figure 1.2B), leading to the lens vesicle. This vesicle is later released from the surface ectoderm. As the lens vesicle is developing, the optic vesicle invaginates to form the double-walled optic cup. After the lens vesicle formation, it sits in the optic cup (Figure 1.2C). Simultaneously, the mesenchyme that covers the optic vesicle differentiates into the thin inner choroid and the fibrous outer sclera (Larsen, 1993; Gould et al., 2004).



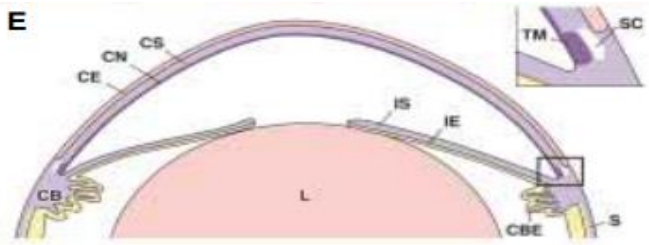


Figure 1.2 – Development of the eye with incidence in the anterior segment. (A) Evagination of neural ectoderm to form the emerging optic vesicle (OV), which moves until it reaches the surface ectoderm. (B) Invagination of the lens placode and formation of the optic cup. (C) Detachment of the lens vesicle from the surface ectoderm, which becomes the epithelial layer of the future cornea. (D) Formation of the corneal stroma and corneal endothelium from the mesenchyme that has migrated. (E) In mature eye iris epithelia, ciliary body epithelia, iris stroma, stroma and muscle of the ciliary body, TM and SC are formed from the periocular mesenchyme. SE: surface ectoderm; M: mesenchyme; OV: optic vesicle; NE: neural ectoderm; LP: lens placode; OC: optic cup; LV: lens vesicle; CE: corneal epithelium; CS: corneal stroma; CN: corneal endothelium; L: lens; RPE: retinal pigmented epithelium; NR: neural retina; CB: ciliary body; CBE: ciliary body epithelia; IS: iris stroma; IE: iris epithelia; S: sclera; TM: trabecular meshwork; SC: Schlemm's canal. Image from Gould *et al*., 2004.

Mesenchymal progenitor cells, cells of the anterior region of the optic cup and cells of the overlying surface ectoderm give rise to the cornea, iris, and drainage structures of the iridocorneal angle. The mesenchyme surrounding the optic cup migrates to the region between the lens and the surface ectoderm and then it splits into two layers creating the anterior chamber. The external wall is continuous with the sclera and the internal wall with the choroid. After lens establishment, the formation of the cornea is stimulated. The ectodermal cells change to form the corneal epithelium and initiate the formation of the corneal stroma. The mesenchymal cells lead to the formation of the corneal endothelium and further development of the corneal stroma with the synthesis of collagen in a lamellar arrangement (Figure 1.2D). The anterior rim of the optic cup expands anteriorly and centrally and differentiates into the iris and ciliary epithelium. The iris and ciliary body stroma have thus a mesenchymal origin. The formation of the iris creates a new space, the posterior chamber, between the iris and the lens (Figure 1.2E) (Larsen, 1993; Gould et al., 2004; Sowden, 2007; Moore and Persaud, 2008).

The development of the trabecular meshwork starts around the 12th week with the appearance of densely packed cell mass forming the primitive iridocorneal angle. The ciliary body is inserted anteriorly (Figure 1.3A). Later, immature trabecular beams with few intertrabecular spaces begin to appear (Figure 1.3B). Schlemm's canal (SC) develops from the blood vasculature adjacent to the trabecular angle at the 5th month of gestation (Hamanaka et al., 1994; Kizhatil et al., 2014). The maturing meshwork is lined by uveal trabecular endothelial cells. Intercellular gaps perforate this lining and increase in size and density. The TM increases

in length by increasing the intertrabecular space, while maintaining the number of beams (Figure 1.3C).

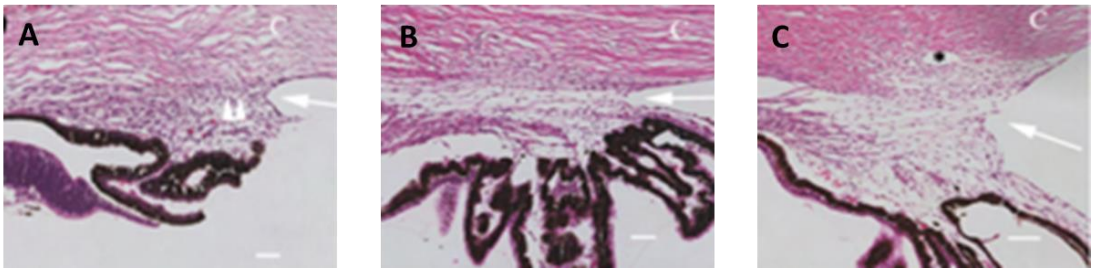


Figure 1.3 – Histological development of the anterior chamber angle. (A) Formation of the primitive anterior chamber angle (arrow) by densely packed cells in the 15th gestation week. (B) Gradual appearance of immature trabecular beams with sparse inter-trabecular spaces and lined by spindle-shaped TM cells (arrow) at the 21st gestation week. Anterior insertion of the ciliary body and incomplete cleaved angle. (C) Well developed TM (arrow) and structurally mature SC (*), at the 27th gestation week. Image adapted from Meghpara *et al.*, 2008.

During TM development, the ciliary body migrates posteriorly, exposing the TM. At birth, the TM is anterior to the iris root and exposed to the aqueous humor (Figure 1.3 and Figure 1.4) (Anderson, 1981; Mizokami et al., 1994; Chan et al., 2007; Sowden, 2007; Meghpara et al., 2008).

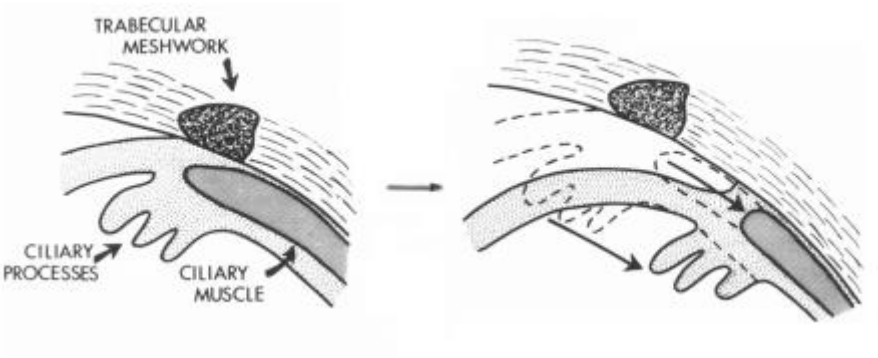


Figure 1.4 – Anterior exposure of the trabecular meshwork to the iris root. In the course of TM development, the ciliary body moves posteriorly, leading TM exposure to the aqueous humor. Image adapted from Anderson, 1981.

1.1.2.Eye structures relevant to glaucoma

1.1.2.1. Ciliary body

The ciliary body is a component of the uveal tract together with the iris and choroid. The ciliary body has a right triangle shape and has two parts: the *pars plana* and the *pars plicata*. The *pars plana* is the posterior part of the ciliary body, has a flatter inner surface and connects with the choroid. The *pars plicata* is located in the most anterior portion of the ciliary body and has several radiating ridges, called ciliary processes (Figure 1.4). Two layers form the epithelium of the ciliary processes: a non-pigmented layer that contacts with the aqueous humor in the posterior chamber, and a pigmented layer in contact with the ciliary process stroma. The

aqueous humor is produced by the ciliary processes. The ciliary body also avoids the intercellular diffusion of large biomolecules from the blood into the aqueous humor. The ciliary muscle layer is composed of bundles of nonstriated fibers. The ciliary body connects to the lens through zonular fibers. When the ciliary muscle contracts, the zonular fibers relax, and the lens moves forward, reaching its accommodative state. Both sympathetic and parasympathetic nerves supply the ciliary body (Velayutham, 2009; Goel et al., 2010; Janssen et al., 2012).

1.1.2.2. Iris

The iris is the most anterior structure of the uveal tract, located between the cornea and the lens. It is a contractile diaphragm with a central aperture: the pupil, which regulates the entry of light into the eye. The iris is composed of the stroma and the bilaminar posterior pigment epithelium. Two groups of smooth muscle are part of the iris: the sphincter pupillae muscle that surrounds the pupil leading to the contraction of the pupil; and the dilatator pupillae muscle, radial group, which contraction originates pupil dilation. At the opposite margin of the pupil, the iris is broadly connected with the ciliary body (Schlote et al., 2006; Velayutham, 2009).

1.1.2.3. Trabecular meshwork and Schlemm's canal

The TM is a triangular-shaped tissue located in the angle of the anterior chamber where the cornea, sclera, and base of the iris converge. The TM bridges the sulcus between the end of Descemet's membrane (Schwalbe's line) and the junction of the ciliary body and iris and the scleral spur (Figure 1.5) (Tamm, 2009).

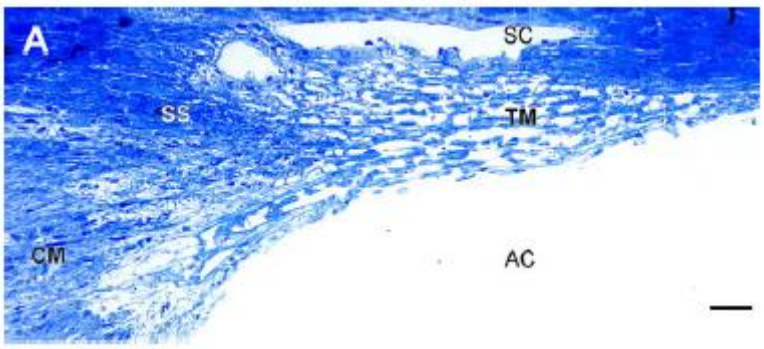


Figure 1.5 – Location of TM. The TM is located in the angle of the anterior chamber, between the cornea and the sclera. The Schwalbe's line is not represented in the figure. AC: anterior chamber; CM: ciliary muscle; TM: trabecular meshwork; SS: scleral spur; SC: Schlemm's canal. Image from Tamm, 2009.

Connective tissue beams form the TM, with a core of collagenous, elastic fibers and microfibril sheath-derived material. The beams are covered by flat cells, the human trabecular cells (HTMs) (Acott and Kelley, 2008; Tamm, 2009). Layers of several beams attached to one another lead to a porous filter-like structure through which the aqueous humor crosses to exit the anterior chamber. Therefore, this structure is an important regulator of the aqueous humor outflow, controlling the intraocular pressure (IOP) (Lutjen-Drecoll, 1999; Goel et al., 2010).

The TM can be divided into three regions that differ in structure: uveal meshwork, corneoscleral meshwork and the juxtacanalicular tissue (JCT) or cribriform region (Figure 1.6).

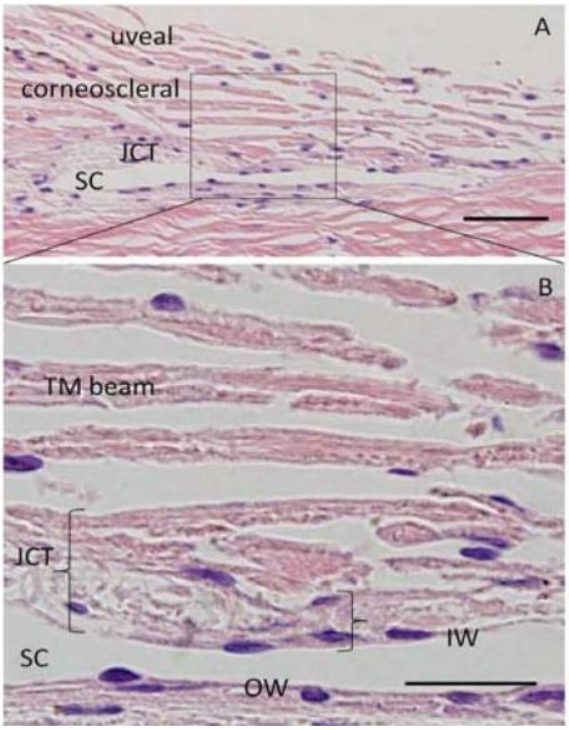


Figure 1.6 – Structure of the TM. The TM has three different regions: the uveal meshwork, the corneoscleral meshwork, and the JCT. Contiguous to the TM is the SC. (B) is an amplification of (A). JCT: juxtacanalicular tissue; SC: Schlemm's canal; IW: inner wall of the Schlemm's canal; OW: outer wall of the Schlemm's canal. Image from Keller and Acott, 2013.

The uveal meshwork is the region closest to the anterior chamber and expanded between the cornea and ciliary body and iris root. The uveal meshwork has irregular openings with a length between 25 and 75 µm (Lutjen-Drecoll, 1999; Tamm, 2009; Goel et al., 2010). The corneoscleral meshwork originates from the scleral spur and extends between the cornea and scleral spur. It is the most extensive portion of the TM, with 8 to 15 trabecular layers that are thicker than those of the uveal meshwork (Lutjen-Drecoll, 1999; Tamm, 2009; Goel et al., 2010). These two regions are composed of fenestrated beams, with large inter-trabecular spaces between them. In the corneoscleral meshwork closer to the JCT region, the trabecular

beams become more flattened (Lutjen-Drecoll, 1999; Acott and Kelley, 2008; Keller and Acott, 2013). The JCT is the smallest region of the TM, with a thickness of 2 to 20 µm. It is not organized in layers of beams like the other two parts of TM but is formed by a loosely arranged extracellular matrix with 2 to 5 layers of scattered cells embedded in an irregular network. The cells contact with each other via extended processes using gap and *adherens* junctions and also attach to extracellular matrix fibrils and endothelial cells of the inner wall of the SC. Aqueous humor crosses the spaces between the cells and the extracellular matrix fibers, which contain a ground substance consisting of various proteoglycans and hyaluronic acid (Figure 1.6) (Tamm, 2009; Keller and Acott, 2013).

Next to the JCT region is the SC. The SC is an endothelium-lined circular canal that is connected to the episcleral and conjunctival veins through the external collector channels, the intrascleral venous plexus, the deep scleral plexus and the aqueous veins (Goel et al., 2010; Dautriche et al., 2015). Its main function is to deliver aqueous humor into the collecting channels, after passage through the TM. The inner and outer walls of the SC are different regarding the morphology and function of the endothelial cells (Figure 1.6 and Figure 1.7) (Dautriche et al., 2015).

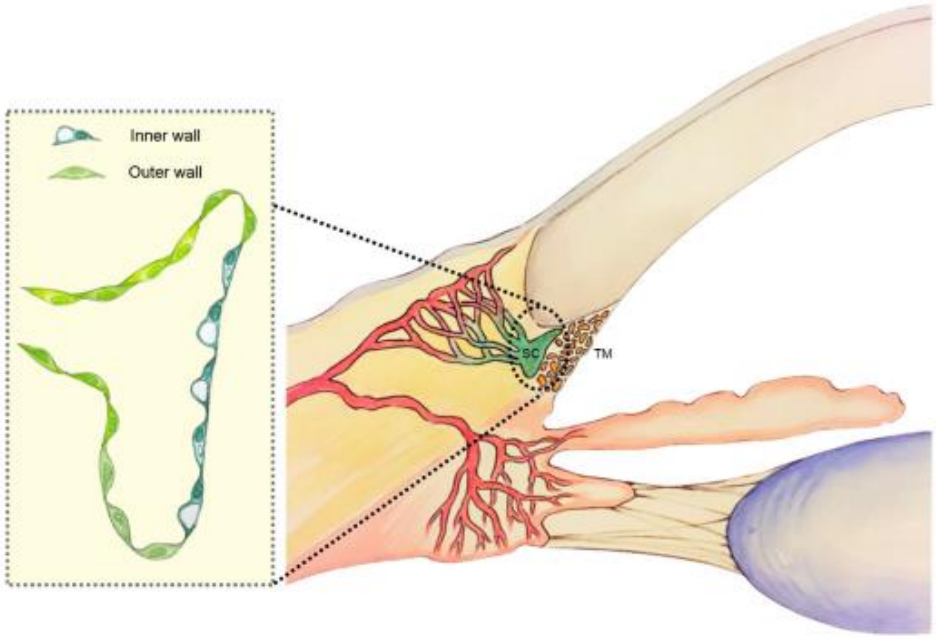


Figure 1.7 – SC and the collection system. The SC continues in the collection channels that communicate with the blood vasculature. Cells of the inner wall of the SC are morphological and functionally different from the cells of the outer wall. TM: trabecular meshwork; SC: Schlemm's canal. Image from Dautriche *et al.*, 2015.

The SC inner wall cells rest on a discontinuous basal lamina, and some of them are in contact with the open spaces of the JCT and aqueous humor. SC cells are connected through tight junctions and some gap junctions (Acott and Kelley, 2008; Tamm, 2009; Dautriche et al., 2015).

Inner wall cells have pores and giant vacuoles (Figure 1.8) due to the basal-apical pressure associated with the aqueous humor flow (Tamm, 2009; Dautriche et al., 2015).

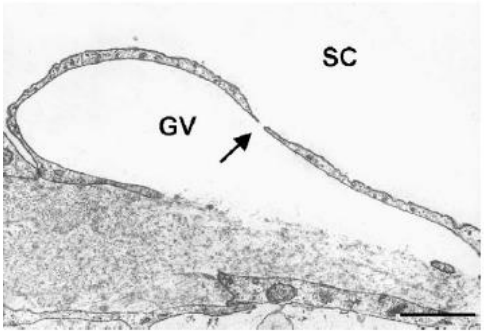


Figure 1.8 – Giant vacuole. Inner wall of SC with a giant vacuole that has an intracellular pore (arrow). GV: giant vacuole; SC: Schlemm's canal. Image from Tamm, 2009.

The nature of the SC cells is controversial with doubts about its lymphatic or vascular origin. A recent study showed that SC develops from blood vessels through canalogenesis, a process different from angiogenesis and vasculogenesis. The authors demonstrated that SC endothelial cells have a hybrid molecular phenotype, expressing proteins that are characteristic of blood and lymphatic vessels (Kizhatil et al., 2014).

1.1.2.4. Optic nerve head

Optic nerve head has approximately 1.2 million retinal ganglion cells axons and is composed of supportive tissue, including glia, blood vessels, and connective tissue. The axons make a 90-degree turn at the optic disc and continue as a series of fascicles or bundles, separated from one another by helical columns of glial cells (astrocytes) and vascular connective tissue septa, forming the optic nerve (Remington, 2005; Albert et al., 2008).

The optic disc has a central depression called cup and between the cup and the edge of the optic disc is the rim, a band of orange-pink neural tissue. The optic disc has a slightly oval shape, with the greatest dimension in the vertical axis. The cup is usually nearly circular, but some individuals may also have a slightly oval shape. The cup-to-disc ratio is the ratio of the vertical diameter of the cup to the vertical diameter of the disc (Figure 1.9) (Albert et al., 2008).

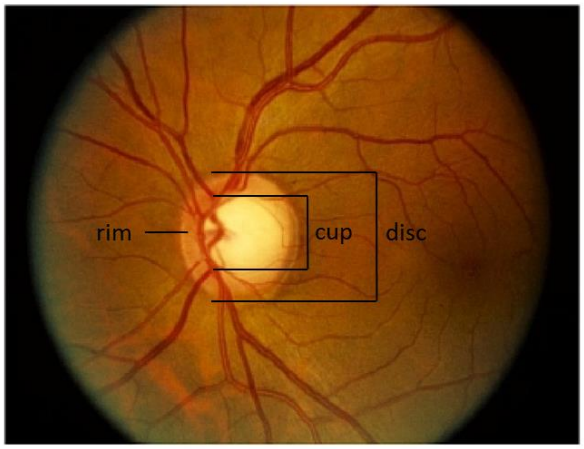


Figure 1.9 – Optic disc. The disc is formed by a central depression (cup) and rim. Image from Almazroa, 2015.

1.1.3.Physiology of the aqueous humor outflow

The aqueous humor is a clear fluid that fills the anterior and posterior chambers of the eye. It is necessary for the optical transparency, structural integrity, nutrition of regions of the eye that do not have blood vessels (cornea, lens, or TM), removal of excretory products from metabolism, and transport of neurotransmitters. In pathological conditions, aqueous humor allows inflammatory cells and mediators to circulate in the eye.

Aqueous humor derives from a plasma filtrate and is produced by the ciliary epithelium of the ciliary processes. Its formation includes three mechanisms: diffusion, ultrafiltration, and active secretion. The last mechanism is the major contributor to aqueous humor formation. Organic and inorganic ions, carbohydrates, glutathione, urea, amino acids and proteins, oxygen, carbon dioxide, and water are the principal components of aqueous humor (Goel et al., 2010).

After production by the ciliary body, aqueous humor flows around the lens and crosses the pupil into the anterior chamber. It reaches the anterior chamber angle, and exits the eye by two different routes: the conventional pathway and the uveoscleral pathway, also called as unconventional pathway (Figure 1.10) (Goel et al., 2010).

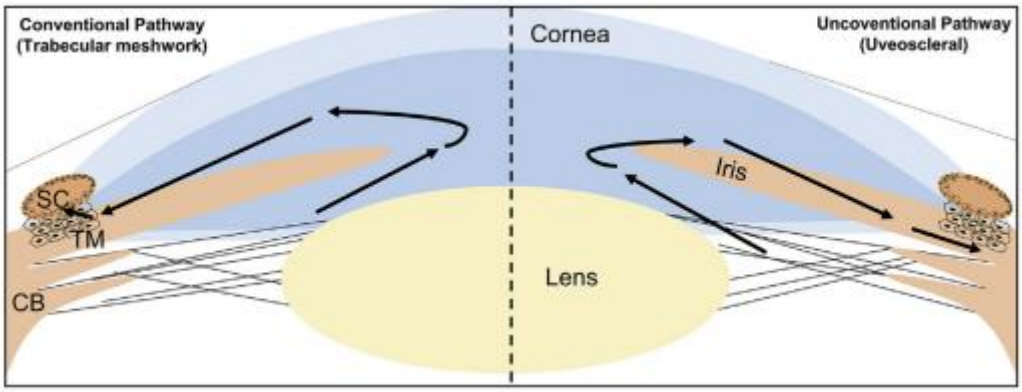


Figure 1.10 – Routes of aqueous humor flow. Aqueous humor is secreted by the ciliary body in the posterior chamber, then passes between the iris and the lens, and goes through the pupil into the anterior chamber. It can exit the eye by two pathways: the conventional pathway, involving the TM (left image), and the unconventional pathway (right image). Image from Carreon *et al.*, 2016.

The conventional outflow is responsible for approximately 85% of the aqueous outflow. In this route, aqueous humor passes through the channels between the uveal and corneoscleral TM beams, crosses the JCT and enters in the SC through micron-sized pores. After exiting the SC, the aqueous humor enters into the collector channels and, subsequently, mixes with blood in the episcleral veins (Figure 1.11) (Goel et al., 2010; Swaminathan et al., 2014).

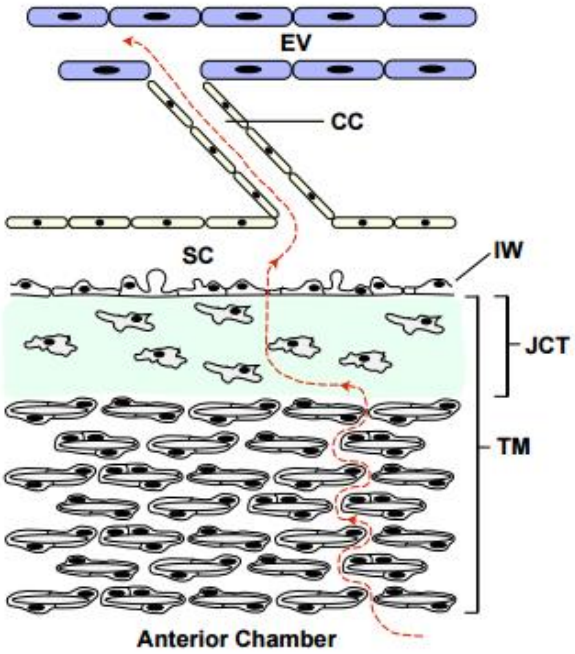


Figure 1.11 – Conventional outflow pathway. Aqueous humor crosses the three regions of the TM, enters in the SC, goes to the collector channel, and finally reaches the episcleral vein. TM: trabecular meshwork; JCT: juxtacanalicular tissue; IW: inner wall of the Schlemm's canal; SC: Schlemm's canal; CC: collector channel; EV: episcleral vein. Image from Swaminathan et al., 2014.

The major regions that confer resistance to aqueous outflow are the JCT and the inner wall of the SC. Glycosaminoglycans, which are a main component of the extracellular matrix of TM, are thought to be partly responsible for increasing the resistance to outflow, as well as other extracellular matrix proteins. Alterations in the cells of TM and the inner wall of SC and the limited presence of pores in the SC that lead to a non-uniform flow throughout the TM and SC also contribute to resistance outflow. Contraction of the ciliary muscle results in a dispersion of the TM and SC dilation, with a decrease in the outflow resistance. The collector channels and distal outflow regions are also thought to contribute to this resistance (Acott and Kelley, 2008; Goel et al., 2010; Swaminathan et al., 2014; Carreon et al., 2016).

Contrarily to the conventional pathway, in the unconventional route, the aqueous humor does not cross the TM to exit the anterior chamber but passes through the uvea, the ciliary body and muscle, and into the choroid and sclera. The unconventional outflow, independently of the structures involved, must always pass through the interstitial spaces of the ciliary muscle. This pathway decreases with age (Goel et al., 2010; Johnson et al., 2016).

1.2. Glaucoma

Glaucoma is a group of optic neuropathies characterized by the progressive degeneration and death of retinal ganglion cells and their axons, leading to visual field loss. Consequently, the optic nerve head has a typical appearance of structural damage with neuroretinal rim thinning, excavation (cupping), and sectoral retinal nerve fiber layer defects (Chang and Goldberg, 2012). Elevated IOP is considered one of the key factors in glaucoma development (Quigley et al., 1994), despite the existence of normal tension glaucoma.

Glaucoma is a leading cause of irreversible blindness worldwide, and the number of glaucoma patients is increasing. It is estimated that the number of glaucoma patients will rise to 79.6 million by the year 2020 (Quigley and Broman, 2006).

1.2.1. Glaucoma classification

This condition is clinically and genetically heterogeneous and can be classified according to the etiology (primary and secondary), the anatomy of anterior chamber (open angle and closed angle) and the age of onset (congenital, juvenile and adult-onset). In general, the three most important categories of glaucoma are: primary open-angle glaucoma (POAG; OMIM 137760), primary angle closure glaucoma (PACG; no OMIM entry), and primary congenital glaucoma (PCG; OMIM 60975) (Figure 1.12) (Vasiliou and Gonzalez, 2008).

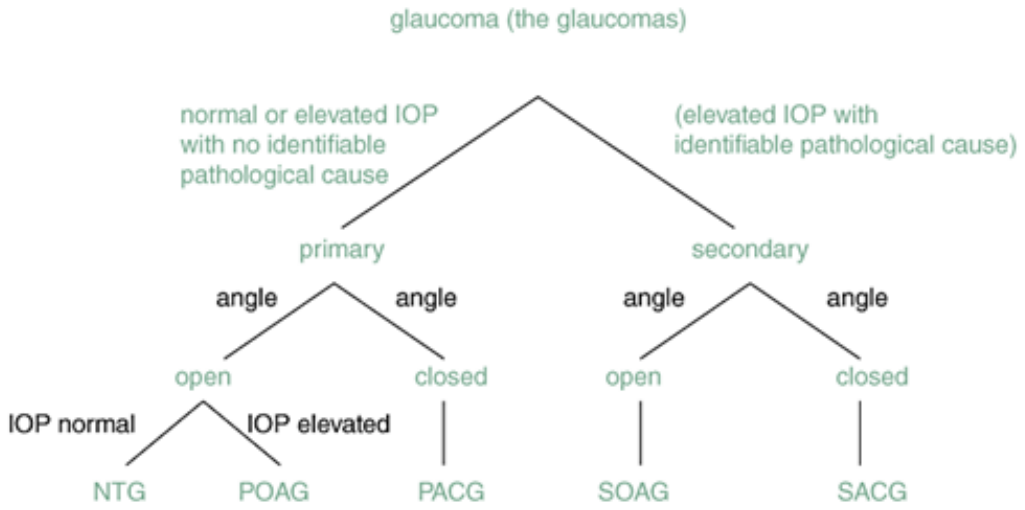


Figure 12 – Glaucoma classification. This classification is according to the etiology (primary vs. secondary) and the anatomy of the anterior chamber (open vs. closed). IOP: intraocular pressure, NTG: normal tension glaucoma, POAG: primary open-angle glaucoma, PACG: primary angle closure glaucoma, SOAG: secondary open-angle glaucoma, SACG: secondary angle closure glaucoma. Image adapted from Casson *et al.*, 2012.

1.2.1.1. Primary open angle glaucoma

The most common type of glaucoma is POAG (Quigley and Broman, 2006), a slow and progressive form of glaucoma. It can be divided according to the age of onset into juvenile-onset POAG (JOAG), with an age of onset from 3 to 35 years old, and adult-onset POAG, usually after the 50th decade of life (Wiggs et al., 1996; Vasiliou and Gonzalez, 2008). POAG can also be classified based on IOP into normal tension glaucoma and increased IOP POAG (Figure 1.12) (Abu-Amero et al., 2015). The susceptibility for POAG depends on risk factors, such as elevated IOP, age, gender, ethnicity, central corneal thickness, and myopia (Abu-Amero et al., 2015). Genetic factors also contribute to the development of POAG, especially for the JOAP subtype, considered an autosomal dominant Mendelian trait with high penetrance (Wiggs et al., 1996). Until date, at least 29 genetic loci have been identified by linkage studies (Kumar et al., 2016). However, only three genes have been described as causative genes: *MYOC* (Stone et al., 1997), *OPTN* (Rezaie et al., 2002) and *WDR36* (Monemi et al., 2005) associated with an autosomal dominant inheritance pattern.

1.2.1.2. Primary angle closure glaucoma

PACG is characterized by the physical obstruction of the TM, usually by the iris. Iris bowing due to inhibition of the aqueous flow, anterior lens movement and compression of the iris between the TM and another anatomical structure can cause the obstruction (Wright et al., 2016). In some cases, the sudden obstruction may lead to a rapid increase in the IOP,

originating an emergency situation, requiring urgent evaluation and treatment (Mantravadi and Vadhar, 2015). Several risk factors have a role in the development of this form of glaucoma, including increasing age, female gender, shallow anterior chamber, short axial length of the eye in hyperopia, small corneal diameter, steep corneal curvature, shallow limbal chamber depth, and thick, anteriorly positioned lenses (Wright et al., 2016). The majority of PACG cases are sporadic. Three genomic loci were identified as susceptibility loci for PACG, rs11024102 on *PLEKHA7*, rs3753841 on *COL11A1* and rs1015213 on chromosome 8q, but the pathologic mechanisms behind these genes are still unknown (Vithana et al., 2012).

1.2.1.3. Primary congenital glaucoma

PCG is a rare but severe form of glaucoma characterized by high IOP during the neonatal or early infantile period being a leading cause of blindness in children (Yang et al., 2009). This form of glaucoma belongs to the subgroup of primary childhood glaucoma and is classified according to the age of onset into neonatal or newborn onset (0-1 month), infantile onset (>1-24 months), and juvenile onset or late-recognized (>2 years) (Kaur et al., 2014).

The typical PCG symptoms include photophobia, blepharospasm, epiphora, buphthalmos, corneal edema and opacification, and breaks in Descemet's membrane (Haab's striae) (deLuise and Anderson, 1983).

1.2.1.3.1. Epidemiology

The disease occurs in all ethnic groups, but its prevalence varies worldwide from 1:5,000 to 1:22,000 with most populations close to 1:10,000 in the Western countries. In the Middle East, this disease is more frequent (1:2500), but the population with the highest incidence is the Roms (Gypsies) in Slovakia (1:1250) (Plasilova et al., 1999). PCG is slightly more common in males and in most cases affects the two eyes (bilateral disease). Higher incident rates of PCG are commonly associated with family consanguinity (Tamcelik et al., 2014).

1.2.1.3.2. Pathophysiology

PCG is characterized by IOP increase, as a consequence of an abnormal development of the anterior chamber angle structures. The exact cause and nature of this abnormal development are still uncertain, and several hypotheses have been considered. Initially, the presence of an imperforate membrane covering the angle, also known as Barkan's membrane, was postulated by Barkan as the cause of the aqueous outflow blockage. However, further studies have not confirmed the presence of this membrane (Anderson, 1981; Rojas et al., 2006). Anderson noted a failure in the posterior migration of the iris and the ciliary body, leading to an

incomplete uncover and therefore exposure to the anterior chamber. The failure in the posterior migration was associated with the premature or excessive formation of unyielding collagenous beams (Anderson, 1981). The abnormal anterior insertion of the longitudinal and circular fibers of the ciliary muscle into the TM, leading to the obstruction of the aqueous humor outflow has also been proposed (Maumenee, 1958). Abnormally anterior insertion of the iris was found in patients with PCG (Anderson, 1981; Rojas et al., 2006; Perry et al., 2012). TM beams were also noted to be tauter and thickened (Maul et al., 1980; Anderson, 1981; Rojas et al., 2006) and JTC appears less porous and thicker (Maul et al., 1980; Anderson, 1981). Rarely, the SC may be absent probably associated with severe maldevelopment (Tawara and Inomata, 1981; Perry et al., 2012).

1.2.1.3.3. Genetics

Genetic causes of PCG have been studied over the years, and it is still a developing field. Linkage analysis identified three loci for this disease: GLC3A (2p2-p21) (Sarfarazi et al., 1995), GLC3B (1p36.2-p36.1) (Akarsu et al., 1996) and GLC3C (14q23) (Stoilov and Marfarazi, 2002). However, the only gene associated with the disease in these loci is *CYP1B1*, located in GLC3A.

CYP1B1 has three exons, with the first and most of the third exon corresponding to the 5' and 3' untranslated regions, respectively (Stoilov et al., 1997). This gene encodes a membrane-bound monomeric monooxygenase expressed in the eye, namely in the trabecular meshwork (Stoilov et al., 1997). CYP1B1 participates in the metabolism of steroids, arachidonic acid, retinoic acid, and melatonin, among other endogenous compounds (Nakano et al., 2014). Mutations in *CYP1B1* are the predominant cause of autosomal recessive inherited PCG (Stoilov et al., 1997; Yang et al., 2009). Cases of incomplete penetrance and variable expressivity of *CYP1B1* mutations have also been described (Campos-Mollo et al., 2009; Suri et al., 2009). The prevalence of PCG patients with *CYP1B1* mutations varies according to the population being more common in consanguineous populations (Fuse et al., 2010).

Despite being recognized that all PCG *CYP1B1* mutations reduce the activity or the abundance of the enzyme, leading to an abnormal collagen distribution and more fragmented microfibrils associated with fibrillin-1 (Zhao et al., 2013; Safari et al., 2016), the exact mechanism behind *CYP1B1* pathogenesis is still unknown. CYP1B1 is thought to be involved in the metabolism of a molecule that participates in eye development or the metabolism of a toxin that when in excess disrupts these developing tissues (Faiq et al., 2015). Recent studies identified periostin (an extracellular matrix protein) and retinol as possible molecules involved in *CYP1B1*-mediated PCG pathogenesis (Zhao et al., 2013; Banerjee et al., 2016).

Other genes have been linked to PCG. *LTBP*2, the latent transforming growth factor beta binding protein 2, caused autosomal recessive disease in Pakistani and Gypsy patients (Ali et al., 2009; Narooie-Nejad et al., 2009; Micheal et al., 2016). This gene encodes an extracellular matrix protein that associates with fibrillin-1 containing microfibrils. The identification of the *LTBP2* as a causative gene for PCG highlighted the role of the extracellular matrix in pathways that lead to glaucoma (Suri et al., 2015). *LTBP2* lies very close to GLC3C on chromosome 14q24 but is not strictly within the locus (Suri et al., 2015).

A few PCG patients had *MYOC* mutations, either independently or in association with *CYP1B1* (Kaur et al., 2005). *MYOC,* the gene encoding for myocilin, is expressed in the TM and the protein is found in the extracellular environment. Mutations in this gene lead to the intracellular sequestration of abnormal wild-type/mutant complexes instead of its secretion to the extracellular environment (Gobeil et al., 2004), disrupting MYOC effect on cell-matrix adhesion and cellular migration (Wentz-Hunter et al., 2004).

*FOXC*1 mutations alone or combined with heterozygous *CYP1B1* mutations have also been described as responsible for PCG (Chakrabarti et al., 2009; Medina-Trillo et al., 2015). FOXC1 is a transcription factor present in the periocular mesenchyme cells that lead to ocular drainage structures such as the iris, cornea, and trabecular meshwork (Chakrabarti et al., 2009). The mechanism behind *FOXC1*-mediated PCG is still not clearly understood. This transcription factor is involved in response to oxidative stress in TM its dysregulation results in TM cell death (Berry et al., 2008; Ito et al., 2014).

Very recently, *TEK* was proposed as a causative gene for PCG. Heterozygous mutations were identified in PCG patients associated with an autosomal dominant model with variable expressivity (Souma et al., 2016). Tie2, the receptor tyrosine kinase encoded by *TEK*, is a regulator of vascular development and stability and is expressed in the SC endothelium. (Jeansson et al., 2011; Kizhatil et al., 2014). The found mutations reduced Tie2 expression, phosphorylation or cellular localization (Souma et al., 2016).

1.2.1.3.4. Clinical diagnosis

An early diagnosis of PCG is essential to prevent the development of the disease. A detailed medical and familiar histories and clinical examination are required.

1.2.1.3.4.1. Clinical examination

Typically, PCG patients present epiphora, blepharospasm, and photophobia. As a result of high IOP, these patients may have buphthalmos, which is the enlargement of the eyeball. Corneal changes are also frequent and may include the enlargement of the corneal diameter

(megalocornea). A horizontal corneal diameter >12 mm is highly suggestive of glaucoma. Pateints may also present Haab's striae (Figure 1.13B), which are breaks in one of the layers of the inner cornea, the Descemet's membrane. These breaks occur because of corneal stretching caused by the increased IOP. Corneal edema is also frequent. The cornea appears cloudy, limiting the view of the iris, because of the high IOP. It initiates as a superficial edema (epithelial edema) and can evolve into a more severe deeper edema (stromal edema). The edema will cause epiphora and photophobia. In severe cases, the sclera also expands and becomes thinner, leading to an increased visibility of the underlying choroid and a "blue sclera" appearance (Figure 1.13) (deLuise and Anderson, 1983; Mandal and Chakrabarti, 2011; Ganesh et al., 2013).

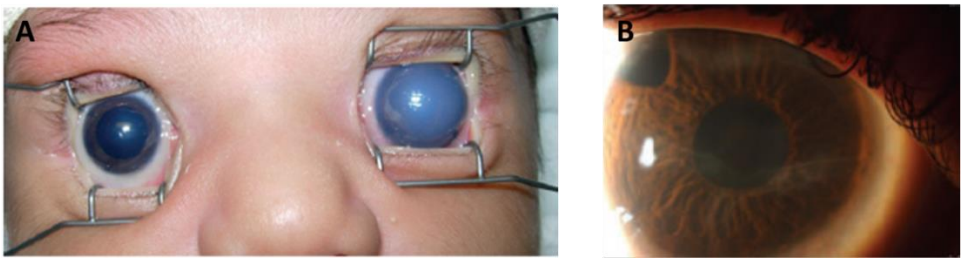


Figure 1.13 – Patients with primary congenital glaucoma. (A) Asymmetric primary congenital glaucoma with megalocornea in the right eye and opacificied megalocornea and buphthalmos in the left eye. Image from Alodhayb *et al.*, 2013. (B) Haab's striae. Image from (Khan, 2011).

1.2.1.3.4.2. Intraocular pressure measurement

The procedure to determine the intraocular pressure, also known as tonometry, must be done in a calm and cooperative patient. Otherwise, the IOP values will be falsely higher. Because this is difficult for young children, usually it requires examination under anesthesia or sedation. However, all anesthetics alter the IOP of the patients (deLuise and Anderson, 1983; Yu Chan et al., 2015).

IOP can be measured by two methods: contact and non-contact. IOP in infants is lower than in adults. In normal infants IOP is between 11 and 14 mmHg. It is a suspicious sign of glaucoma when IOP is greater than 20 mmHg in a calm patient, or there is an asymmetry greater than 5 mmHg between the two eyes (Ganesh et al., 2013).

1.2.1.3.4.3. Gonioscopy

This technique allows the examination of the anterior chamber angle. The healthy infant angle is different from the adult angle with a less distinct Schwalbe's line, a less pigmented TM, and a translucent uveal meshwork, which leads to a less clear junction between the scleral spur and ciliary body. The angle of PCG patients has an anterior insertion of the iris directly into the

trabecular meshwork. This insertion is commonly flat. The angle is usually avascular, but in these patients anomalous iris vessels may also be seen as loops branching from the major arterial circle, and peripheral iris may be covered by thin fluffy tissue (Figure 1.14) (Mandal and Chakrabarti, 2011; Yu Chan et al., 2015).

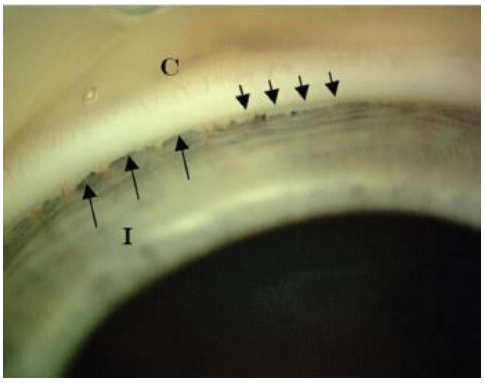


Figure 1.14 – Gonioscopy of a PCG patient. Short arrows indicate anterior TM. Long arrows indicate anteriorly positioned iris. C: cornea; I: iris. Image from Ganesh *et al*., 2013.

1.2.1.3.4.4. Fundoscopy

The optic nerve head must be observed to identify the presence of IOP-induced changes. In PCG patients the cup-to-disc ratio is increased, with a concentric enlargement of the cup (Figure 1.15) (Gupta et al., 2016).

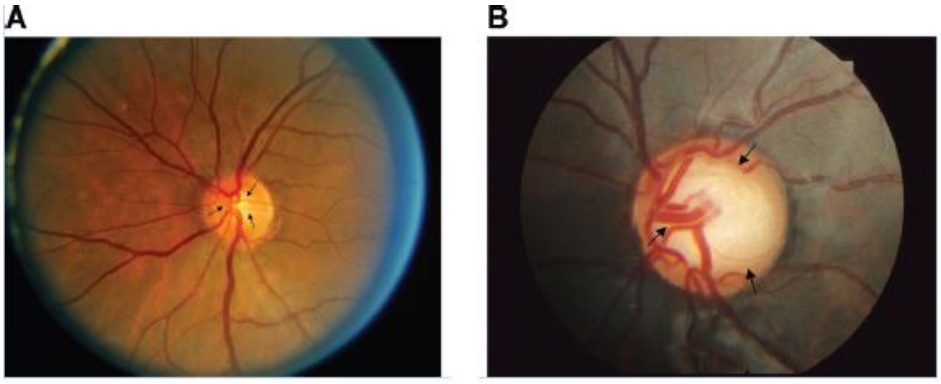


Figure 1.15 – Optic nerve head by fundoscopy. (A) Normal optic nerve head of a child without glaucoma. (B) Concentric increased cup-to-disc ratio of a child with glaucoma. Image from Ganesh *et al*., 2013.

The lesion of the optic nerve is faster in children comparatively to the adults (Quigley, 1977) and appears to be caused by a nerve fiber layer axonal damage, rather than changes in the optic nerve compliance (Ely et al., 2014). Cups in excess of 0.5 – 0.6 are concerning (Ganesh et al., 2013).

If the IOP is promptly reduced by treatment, the optic nerve head cupping can be reversible. Multiple theories have been proposed to explain this phenomenon, such as glial tissue

proliferation, forward movement of the lamina cribrosa or shrinkage of a stretched scleral canal (Quigley, 1977; Mochizuki et al., 2011; Ely et al., 2014).

1.2.1.3.5. Differential diagnosis

PCG, a subgroup of childhood glaucomas, is characterized by the presence of isolated trabeculodysgenesis not associated with other developmental ocular anomalies or ocular disease that can raise the IOP. The group of childhood glaucomas also includes secondary childhood glaucoma, which is developed due to damage of the aqueous outflow system as a result of congenital or acquired ocular diseases or systemic disorders (Table 1.1) (Marchini et al., 2014).

Table 1.1 – Secondary glaucomas. Each anomaly or disease can be considered in one of the four different categories of secondary glaucoma (Kaur *et al*., 2014; Marchini *et al*., 2014).

Classification	Anomaly or disease
Glaucoma associated with a non-acquired ocular anomaly	Axenfeld-Rieger anomaly/syndrome
	Peters anomaly/ syndrome
	Aniridia
	Ectropion uveae
	Persistent fetal vasculature
	Sclerocornea
Glaucoma associated with a non-acquired systemic disease or syndrome	Sturge-Weber syndrome
	Neurofibromatosis
	Marfan syndrome
	Homocystinuria
	Weill-Marchesani syndrome
	Congenital rubella syndrome
	Lowe syndrome
	Rubinstein-Taybi syndrome
Glaucoma associated with acquired conditions	Uveitis
	Trauma
	Steroid induced
	Tumors
	Retinopathy of prematurity
Glaucoma following congenital cataract surgery	

Primary congenital glaucoma has the typical symptoms of epiphora, photophobia, and blepharospasm. Buphthalmos and Haab striae can be present, but the other types of the

neonate and infant glaucomas also manifest the same symptoms (Marchini et al., 2014) and some of the PCG symptoms may be present in other pediatric ophthalmic conditions (Table 1.2).

Table 1.2 – Differential diagnosis of features commonly found in childhood glaucoma. These conditions must be considered during PCG diagnosis (Khan, 2011; (Hoyt and Taylor, 2013))

Signs and symptoms	Conditions
Increased IOP	Anesthesia
	Low level of patient cooperation
Buphthalmos	Blue sclera
	Lens subluxation
	Retinal detachment
	Phthisis
Descemet membrane tears	Birth injury
	Posterior polymorphous corneal dystrophy
Corneal edema or opacification	Congenital hereditary endothelial dystrophy
	Posterior polymorphous corneal dystrophy
	Congenital stromal corneal dystrophy
	Posterior amorphous corneal dystrophy
	Mucopolysaccharidoses and other storage diseases
	Peters anomaly
	Sclerocornea
Epiphora and “red eye”	Nasolacrimal duct obstruction
	Conjunctivitis
	Corneal epithelial defect/abrasion
	Ocular inflammation
Optic nerve cupping	Physiological cupping
	Papillorenal syndrome
	Optic nerve hypoplasia

An assertive diagnosis is essential for an appropriate and early treatment. In some occasions, because children examination can be difficult, pediatric glaucoma may be at risk of being misdiagnosed or sub treated with the possibility of irreversible damage, such as optic nerve damage (Kaur et al., 2014). Glaucoma should not be made considering the elevated IOP only, and the exclusion of glaucoma must be rigorous. (Marchini et al., 2014). However, the

identification of a non-glaucomatous disorder does not eliminate the possibility of glaucoma or further glaucoma development.

1.2.1.3.6. Treatment

The main goal of the PCG treatment is to reduce and maintain lower IOP values to prevent further ocular complication, such as vision loss. Primary congenital glaucoma treatment follows several approaches, however almost all PCG patients undergo surgical procedures to eliminate the resistance to aqueous humor outflow.

1.2.1.3.6.1. Medical therapy

Primary congenital glaucoma treatment uses several groups of drugs. Beta-blockers, carbonic anhydrase inhibitors, prostaglandin analogs, alpha-adrenergic agonists, and parasympathomimetics (miotics) are the most used drugs. These drugs help to reduce IOP usually by decreasing the rate of aqueous humor production and/or increasing uveoscleral outflow (Samant et al., 2016). The administration can be local or systemic.

However, medical therapy in PCG patients has a limited action in IOP reduction. The efficacy of this type of treatments is not high, and it may have severe side effects, caused by the drug and the preservatives. The administration of eye drops is also considered to be difficult in these patients. Therefore, the pharmacological treatment has a supportive role, and it is used to reduce the cornea opacity before the surgery, to prevent optic nerve damage while the patient is waiting for surgery, and postoperatively to control IOP (Yu Chan et al., 2015; Samant et al., 2016).

1.2.1.3.6.2. Goniotomy

Goniotomy is a surgical procedure in which the anomalous uveal TM is cut to reduce the resistance to aqueous humor outflow, thereby restoring the access of aqueous humor to SC. Because a goniotomy knife must enter and cross the anterior chamber until the angle (Figure 1.16), this procedure cannot be done in eyes with cloudy corneas. In mild or moderate PCG patients, this technique is usually the first surgical treatment performed. Sometimes more than one goniotomy must be done (Lister, 1965; Alodhayb et al., 2013; Yu Chan et al., 2015).

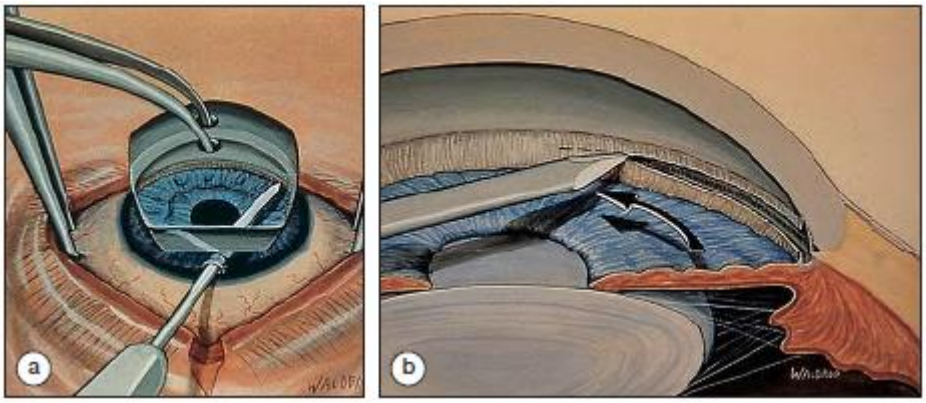


Figure 1.16 – Goniotomy. (A) After anesthesia, the eye is fixed, and the goniotomy knife enters into the anterior chamber through the cornea at the opposite site of the angle to be operated. It goes parallel to the iris until it reaches the angle, which is observed using a surgical goniolens. (B) The tip of the knife cuts the TM. Image from Alberts *et al*., 2008.

1.2.1.3.6.3. Trabeculotomy

During this technique, a trabeculotome is inserted into the SC through a superficial scleral flap. After SC cannulation, the trabeculotome tears through the TM into the anterior chamber, creating a direct communication between the anterior chamber and SC (Figure 1.17). Unlike goniotomy, trabeculotomy can be performed in eyes with cloudy corneas and is technically easier. It is often performed in more severe cases of PCG and can be done simultaneously with trabeculectomy (Burian, 1960; Albert et al., 2008; Mandal and Chakrabarti, 2011; Yu Chan et al., 2015).

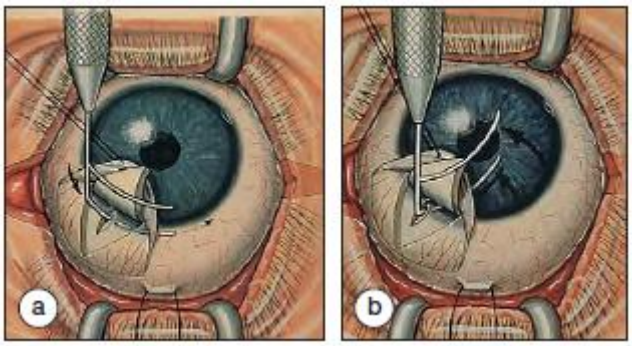


Figure 1.17 – Trabeculotomy. (A) A triangular or rectangular scleral flap is created at the corneo-scleral junction after separation of the conjunctiva from the sclera; the SC is identified, and the internal arm of the trabeculotome is inserted through the cut into the SC. (B) The trabeculotome rotates into the anterior chamber (black arrows) breaking the TM. Image from Alberts et al., 2008.

1.2.1.3.6.4. Trabeculectomy

Trabeculectomy is also an *ab externo* procedure involving a scleral flap at the corneo-scleral junction. However, in this surgery, a section of the TM and SC under the scleral flap is removed, and part of the iris can also be removed (Figure 1.18). The scleral flap is repositioned

and loosely sutured and the conjunctiva replaced. The aqueous humor flows through the sclera into a zone under the conjunctiva called bleb and absorbed by capillaries and lymphatics within the conjunctiva or evaporates across the conjunctiva (Thyer and Wilson, 1972; Wilkins et al., 2001).

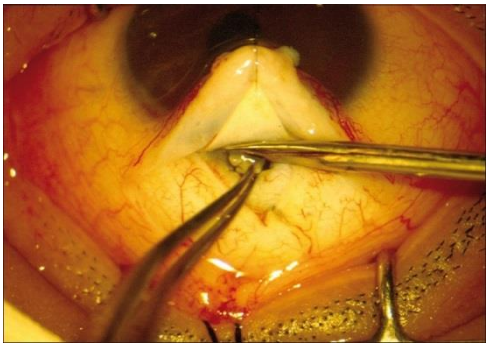


Figure 1.18 – Trabeculectomy. After creation of a scleral flap at the corneo-scleral junction a section of the TM and SC is removed. Image from Mandal, 2011.

The most common cause of trabeculectomy failure is the scarring in the operation site preventing the aqueous humor from exiting the eye. Antimetabolites including mitomycin-C can be used to prevent scarring and increase the success of the trabeculectomy (Wilkins et al., 2001).

1.2.1.3.6.5. Combined trabeculotomy-trabeculectomy

In some patients, these two techniques can be performed simultaneously, with trabeculectomy done after trabeculotomy. Some studies have found a higher success rate with the combined strategy (Elder, 1994). However, other studies have shown no differences when comparing trabeculotomy with combined trabeculotomy-trabeculectomy (Khalil and Abdelhakim, 2016).

1.2.1.3.6.6. Glaucoma draining implants

In the case of failure of the anteriorly described surgeries in refractory cases of PCG, draining implants is a valid alternative. Glaucoma drainage implants may be open-tube devices (non-restrictive) or valved devices (flow-restrictive). The last ones are intended to decrease the incidence of complications associated with hypotony in the postoperative period. Besides the hypotony, tube-cornea touch, tube obstruction, tube or plate exposition, endophthalmitis and choroidal or retinal detachment may also occur (Mandal and Chakrabarti, 2011; Alodhayb et al., 2013).

1.2.1.3.6.7. Cyclodestructive procedures

Contrary to the other procedures, cyclodestructive procedures reduce the formation of aqueous outflow by lesion of the ciliary processes. As the results are often unpredictable and complications are frequent, this type of treatment is reserved for refractory patients or with a very poor prognosis. The most common procedures are cyclophotocoagulation (transscleral Nd:YAG, transscleral diode and endoscopic diode) and cyclocryotherapy (Mandal and Chakrabarti, 2011; Yu Chan et al., 2015).

1.3. Objectives and thesis organization

As a congenital disease primary congenital glaucoma presupposes the presence of gene mutations in PCG patients, and the discovery of the GLC3A, GLC3B, and GLC3C loci proved this assumption. A very significant mark in PCG genetics was the identification of *CYP1B1* as a causative gene for this disorder. Since then several mutations in this gene have been identified in PCG patients from different populations worldwide.

However, and despite the recent discovery of additional disease candidate genes, a significant number of patients remain without identification of causal mutations. Although in this situation the diagnosis of the patient is established based on clinical features and the adequate treatment initiated, it is difficult for the clinician to confirm the clinical diagnosis and provide genetic counseling to the family. Moreover, without knowledge of the mutated genes associated with the disease, it is challenging to build adequate research studies to elucidate molecular mechanisms of primary congenital glaucoma and contribute to the identification of new drug targets and therapeutic approaches. The goals of this study were to discover associations between new mutated genes and primary congenital glaucoma and contribute to the understanding of PCG pathogenetics. The following specific objectives were established to address these goals:

1. Characterize a Portuguese primary congenital glaucoma population and select patients without mutations in *CYP1B1*;
2. Obtain a variant catalog for the PCG patients without *CYP1B1* mutations;
3. Discover putative pathogenic variants in new genes;
4. Validate the pathogenicity of variants and possible association with disease.

In Chapter 1 a brief introduction to the human eye development and organization is presented. This chapter also includes information on the main eye structures related to

glaucoma that may play a role in the disease mechanism. Glaucoma types are described as well as genes already associated, clinical features and current treatments.

Chapter 2 describes the study of a Portuguese population of primary congenital glaucoma patients regarding *CYP1B1* mutations in the enhancer, promoter, 5' and 3' untranslated regions (UTRs), intronic and exonic regions, by Sanger sequencing. Haplotypes of *CYP1B1* were also determined and studied the genotype-phenotype correlation. Patients without *CYP1B1* mutations were the subject of study in the remaining chapters. This chapter addresses objective 1. Part of the results included in this chapter were published in the Ophthalmic Genetics journal.

Chapter 3 describes the strategies and main results of the search for new genes associated with PCG. The exomes of 14 patients were sequenced, and the identified variants in coding regions and splice sites prioritized according to the prevalence of PCG and the recessive, dominant and *de novo* models of inheritance. The most interesting rare variants predicted as pathogenic by functional impact tools were analyzed in detail. This chapter addresses objectives 2 and 3.

Chapter 4 reports the experimental validation of the two most interesting candidate variants identified in the previous chapter. Selection of variants took into account the prevalence in databases and functional consequence predicted. This chapter addresses objective 4.

In Chapter 5, the exomes of patients under study were additionally analyzed for a group of genes already associated with other forms of glaucoma, and that may have an impact on PCG development. The presence of glaucoma and mutations in these genes may indicate other conditions or disorders with secondary glaucoma or other types of glaucoma. This chapter contributed for objective 3.

Chapter 6 summarizes main results, discusses emerging conclusions and their contribution to the state-of-the-art knowledge on primary congenital glaucoma molecular mechanisms. Possible future research lines are suggested.

CHAPTER 2

Genetic study of *CYP1B1* of a primary congenital glaucoma Portuguese population

Part of this chapter has been published in Simões MJ, Carmona S, Roberts R, Wainwright G, Faro C, Silva E, Egas C. CYP1B1 mutational screening in a Portuguese cohort of primary congenital glaucoma patients. Ophthalmic Genet. 2016. 7:1-3. DOI: 10.1080/13816810.2016.1188121

2.1. Introduction

Primary congenital glaucoma is a rare childhood ocular disease characterized by an increased IOP in the presence of typical symptoms, such as epiphora, blepharospasm, and photophobia. Buphthalmos and megalocornea are also found in the majority of the patients (deLuise and Anderson, 1983). These symptoms are usually revealed during the first year of life, but their appearance may occur until the third year of life (Yang et al., 2009). The trabeculodysgenesis evident in PCG patients leads to a reduced aqueous outflow and, consequently, to a rise in IOP and further symptoms (Yang et al., 2009). However, the molecular mechanisms behind the pathogenesis of this disease are not entirely understood. Genetics studies have been performed to identify the genetic cause associated with PCG and to clarify the mechanisms that lead to the anterior chamber angle malformation. The first studies identified three loci for this disease: GLC3A (2p2-p21) (Sarfarazi et al., 1995), GLC3B (1p36.2-p36.1) (Akarsu et al., 1996) and GLC3C (14q23) (Stoilov and Marfarazi, 2002) through linkage analysis. So far the only gene found mutated in these loci in PCG patients was *CYP1B1*, located in GLC3A (Stoilov et al., 1997).

CYP1B1 encodes a 543 amino acid protein of the P450 cytochrome family (Tang et al., 1996). It is a membrane-bound monomeric monooxygenase and participates in the metabolism of steroids, arachidonic acid, retinoic acid, and melatonin, among other endogenous compounds (Nakano et al., 2014). CYP1B1 is detected in the eye, namely in the trabecular meshwork (Stoilov et al., 1997).

The prevalence of PCG patients with *CYP1B1* mutations is variable according to the population: 20-30% of PCG patients in ethnically mixed populations have *CYP1B1* mutations, whereas its prevalence increases to 85-100% in consanguineous populations (Fuse et al., 2010). Up to date, more than 150 *CYP1B1* mutations have been associated with PCG (Sheikh et al., 2014). *CYP1B1* mutations linked to an autosomal recessive inheritance model are the most common cause of PCG (Stoilov et al., 1997; Yang et al., 2009). Nevertheless, rare cases of incomplete penetrance and variable expressivity of *CYP1B1* mutations have also been reported (Campos-Mollo et al., 2009; Suri et al., 2009).

All mutations reduce the activity or the abundance of the CYP1B1 enzyme (Chavarria-Soley et al., 2008). This may result in the absence of an unknown metabolite indispensable for TM development, in the prolonged exposure of the developing tissues to a toxin that is detrimental to proper eye development, or in a mechanism where CYP1B1 reduction may influence the expression of important genes relevant to TM and anterior chamber formation

(Faiq et al., 2015). Zhao and collaborators reported increased IOP in Cyp1b1-deficient (Cyp1b1(-/-)) mice, simultaneously with an abnormal collagen distribution, increased oxidative stress and decreased levels of periostin (a secreted extracellular matrix protein essential for collagen fibrillogenesis) in TM tissues. These results were also found in human glaucomatous TM tissues (Zhao et al., 2013). *CYP1B1* PCG mutations have been demonstrated to affect the extracellular matrix, with less abundant and more fragmented microfibrils associated with fibrillin-1 (Safari et al., 2016), or the retinol metabolism, leading to a total absence or higher levels of retinoic acid. Optimal levels of retinol and related compounds (retinoids) are essential for the normal eye developmentt (Banerjee et al., 2016).

More recently, other genes have been linked to PCG: *LTBP2*, *MYOC*, *FOXC1*, and *TEK*. However, these genes were identified only in a few PCG families.

The purpose of this study was to determine the prevalence and spectrum of *CYP1B1* mutations causative of PCG and to evaluate the correlation between genotype and clinical PCG features in a PCG Portuguese population.

2.2. Material and Methods

2.2.1.Primary congenital glaucoma patients

A group of 37 unrelated patients, clinically diagnosed for PCG, was selected at the Ophthalmology or Medical Genetics Services of the Coimbra Hospital University Center, Coimbra, Portugal. All participants, or parents in the case of children, gave their informed consent for inclusion in this study. The research followed the tenets of the Declaration of Helsinki and was approved by the Ethical Committee of the Faculty of Medicine, University of Coimbra, Coimbra, Portugal.

The patient population had 18 males (48.6%) and 19 females (51.4%). All patients except patient 41 had Portuguese ancestors and lived in the central region of the country. Cases of congenital glaucoma associated with other systemic abnormalities were excluded. All patients exhibited typical PCG symptoms, such as megalocornea, tearing, and photophobia. The clinical characterization of the patients is summarized in Table S2.1, in the annex. Primary congenital glaucoma diagnosis occurred during the first 12 months of age for all patients except for patient 45, with the majority of the patients presenting both eyes affected (patients 4, 9, 33, and 42 had only the right eye affected). Most of the patients exhibited an aggressive form of PCG and had undergone multiple surgeries for IOP control. Three probands (patients 2, 26 and 29) had a positive familial history of PCG without parental consanguinity. Patient 9 also had a positive familial history of PCG, but with parental consanguinity. The remaining 34 patients were simplex cases.

2.2.2.*CYP1B1* amplification

Blood samples were collected from all participants and the DNA extracted from peripheral blood leukocytes using the DNeasy Blood & Tissue Kit (QIAGEN, Hilden, Germany), according to manufacturer's instructions.

Three oligonucleotide primer pairs were designed with the Oligo Explorer software, version 1.2 (T. Kuulasma, University of Kuopio, Finland) to amplify and sequence the *CYP1B1* coding regions (exon 2 and part of the exon 3). DNA samples were amplified by PCR. Amplification reactions contained 0.2 µM of each primer described in Table 2.1, 100 ng of genomic DNA, 1 U of Platinum® Taq DNA Polymerase High Fidelity (Invitrogen, Carlsbad, CA, USA), 0.2 mM dNTPs mixture, 2 mM $MgSO_4$, and 1X High Fidelity PCR Buffer, in 50 µl reactions. Amplification conditions consisted of an initial denaturation of 2 minutes at 94°C followed by 30 cycles of denaturation at 94°C for 30 seconds, annealing according to the temperatures described in Table 2.1 for 45 seconds, and extension at 68°C for 1 minute, with a final step of extension during 10 minutes at 68°C.

Table 2.1 – Oligonucleotide primers used for amplification and sequencing of the *CYP1B1* coding regions and exon-intron boundaries. Primer sequence, amplicon size and main PCR conditions are listed.

ID	Exon	Primer sequence (5' – 3')	Amplicon length (bp)	Annealing temperature (°C)	PCR additives
E2_1F	2	GGCCATTTCTCCAGAGAGTC	768	63	DMSO 6%
E2_1R		CGAAACACACGGCACTCATG			
E2_2F	2	CATGATGCGCAACTTCTTCA	650	58	DMSO 6%
E2_2R		TACTCCGCCTTTTTCAGAGG			
E3_F	3	AGTGAGAAATTAGGAAGCTGTT	784	60	
E3_R		TTTCAGCTTGCCTCTTGCTTC			

Family members available for the *CYP1B1* positive PCG patients (2, 6, 10, 26, 28, 29, 38, 39, 40 and 44) were also screened for the mutations found in the index case to confirm inheritance.

Patients negative for *CYP1B1* mutations in the coding regions were further sequenced for mutation detection in the enhancer, core promoter and 5'UTR (sequenced in the same fragment), introns and 3'UTR (the remaining part of the exon 3) with a similar procedure as described for the coding region of the gene. The oligonucleotides and amplification conditions are listed in Table 2.2. DNA samples were amplified in 50 µL reactions with the following reagents: 100 ng of genomic DNA, 1X Phusion GC Buffer, 0.2 mM of each dNTP, 0.5 µM of each

primer, and 0.02 U/µL of Phusion DNA Polymerase (Invitrogen, Carlsbad, CA). A touchdown-PCR was performed for all specific fragment amplification: one denaturing step at 98°C for 3 minutes followed by another denaturing step at 99°C for 1 minute; a touchdown phase with a decrease of 1°C in the annealing temperature per cycle: 14 cycles of denaturation at 98°C for 10 seconds, initial annealing temperature 14°C superior to the normal annealing temperature and decreasing 1°C per cycle during 30 seconds, and extension at 72°C for 30 seconds; another generic amplification phase: 21 cycles of denaturation at 98°C for 10 seconds, annealing temperature for 30 seconds and extension at 72°C for 30 seconds; and a final extension step at 72°C for 5 minutes. Annealing temperatures are indicated in Table 2.2. Electrophoresis of all PCR products was undertaken on a 1% agarose gel stained with ethidium bromide.

Table 2.2 – Oligonucleotide primers used for amplification and sequencing of the *CYP1B1* enhancer, promoter, 5'-UTR, introns and 3'-UTR. Primer sequence, amplicon size, and main PCR conditions are listed.

ID	Primer sequence (5' – 3 ')	Amplicon length (bp)	Annealing temperature (°C)	PCR additives
enhancer_F	GGTTGTCCCTGGTGAACCTA	813	63	Betaine 1.5 M
enhancer_R	CCCTGAGATTTCCCGCGTA			
5'UTR_F	CATGAAAGCCTGCTGGTAG	744	58	DMSO 6%
5'UTR_R	GTTTCCACCTCGCTGTAAC			
I1_F	CACTGACGACGCCAAGAGAC	821	61	DMSO 6%
I1_R	CGGATCTGGAAAACGTCGCC			
I2_1_F	TACCGGCCACTATCACTGAC	726	60	DMSO 6%
I2_1_R	CTTGTCTTCTACCCCGGCAT			
I2_2_F	CAGAAGGCGGCCACAACT	737	60	DMSO 6%
I2_2_R	ACCTATGGAACCTATTCCGGG			
I2_3_F	GTCAACCAGAGATCAGGGAAAAGG	971	59	DMSO 6%
I2_3_R	CCCGCCTCATTTTAAACCCAGT			
I2_4_F	CTGAAAATACATTTCTAGGCATGG	983	58	DMSO 6%
I2_4_R	AAAGGAAGGCCAGGACATAG			
3'UTR_1F	GCTCCTTGATAGTGCTGTCC	560	58	
3'UTR_1R	CTTATCACCTTCAATCACACAC			
3'UTR_2F	CCAAACTACTGAATGAAGCAG	683	56	
3'UTR_2R	CTTGGTTGCGTTAGTTGTAC			
3'UTR_3F	CTAACGCAACCAAGTGTGC	592	60	
3'UTR_3R	GCTTTCCCAGCTACCAAAG			

ID	Primer sequence (5' – 3 ')	Amplicon length (bp)	Annealing temperature (°C)	PCR additives
3'UTR_4F	GACAGTGGAGATGAGGTCAG	591	60	
3'UTR_4R	CCTGCTTTGTGTAGTTGACTC			
3'UTR_5F	GATGTTTGGCGTAGCAGAG	713	56	
3'UTR_5R	GCTTTGACATACAAATGAAGC			
3'UTR_6F	CCCTCATTGTGTTTCTACCG	816	56	
3'UTR_6R	CCTCCGTGTTGGTTCTTAG			

2.2.3.Sanger sequencing

After purification with AgencourtT® AMPure DNA magnetic beads (Beckman Coulter, Brea, California), the PCR products were directly sequenced in both directions by Sanger sequencing using the primers described in Table 2.1 and Table 2.2. Forward and reverse sequencing electropherograms were aligned, the sequence compared to the reference *CYP1B1* gene sequence (ENSG00000138061) using BioEdit version 7.0.1 package (http://www.mbio.ncsu.edu/bioedit/bioedit.html) and mutations identified. Mutations were annotated according to the ENST00000610745 transcript, from assembly GRCh37.

2.2.4.Haplotype analysis

CYP1B1 haplotypes were determined according to the genotype of five previously described intragenic polymorphisms: R48G (rs10012), A119S (rs1056827), V432L (rs1056836), D449D (rs1056837), and N453S (rs1800440) (Bejjani et al., 2000; Stoilov et al., 2002). Haplotype frequencies were estimated using HaploView 4.2 software (Barrett et al., 2005).

2.2.5.Frequency of the *CYP1B1* mutations in the control population

The presence of *CYP1B1* mutations was evaluated in an ethnically matched control population of 55 Portuguese individuals. These controls were not diagnosed with any form of glaucoma and belonged to an internal database that contains genome and exome sequencing data.

2.2.6.Statistical analysis

The chi-square test was used to compare the frequency of *CYP1B1* positive individuals among the 37 index cases of PCG and the 55 controls.

The comparisons between the *CYP1B1* positive and *CYP1B1* negative PCG patients for the unilateral/bilateral eye disease and family history were performed using the Fisher's exact test and for the gender using Qui-square test. The Mann-Whitney U test was applied for the

number of surgeries, IOP at diagnosis, post-surgery IOP, corneal diameter, and cup-to-disc ratio comparisons. All tests were performed using SPSS Statistics software, version 20 (Armonk, NY: IBM Corp.). Differences between the haplotype frequencies among patients with *CYP1B1* mutations, without mutations and the control group were evaluated through chi-square tests in the HaploView 4.2 software.

2.3. Results

2.3.1. *CYP1B1* mutation screening in PCG patients

Sequencing of the *CYP1B1* coding regions and intron/exon boundaries in the 37 PCG patients identified nine different disease-causing mutations in 24 patients (64.9%). Mutations corresponded to 4 frameshifts (p.A179RfsX18, p.R355HfsX69, p.T404SfsX30, and p.D449MfsX8), 3 missenses (p.L378Q, p.E387K, and p.P437L), 1 nonsense (p.R444X), and 1 duplication (p.R468_S476dup) (Figure 2.1).

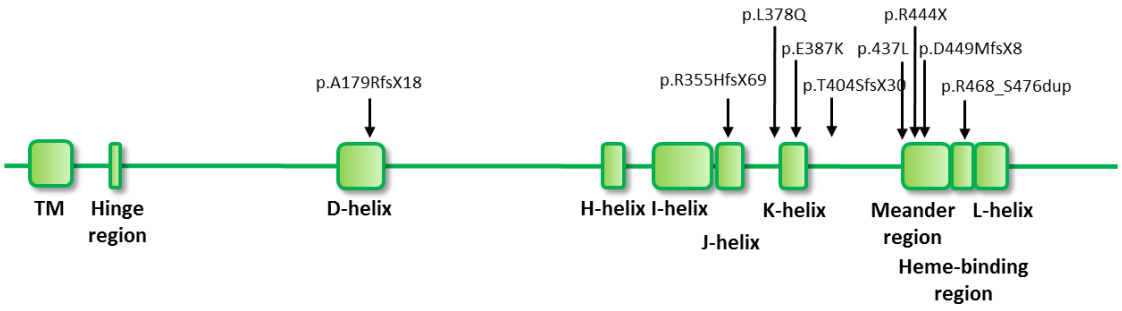


Figure 2.1 – CYP1B1 protein structure showing the location of the identified mutations. The majority of the mutations are located between helix J and L, disturbing the heme-binding region. The left side of the protein scheme represents the N-terminal.

Ten patients carried homozygous mutations (27.0%) and 14 compound heterozygous mutations (37.8%). There were no patients with only one heterozygous *CYP1B1* mutation (Table S2.2). Three of the four patients with a familial history of PCG were positive for mutations in *CYP1B1*. These 3 cases corresponded to the non-consanguineous families (Table S2.1). Two of these familial cases shared the same homozygous mutation, R468_S476dup.

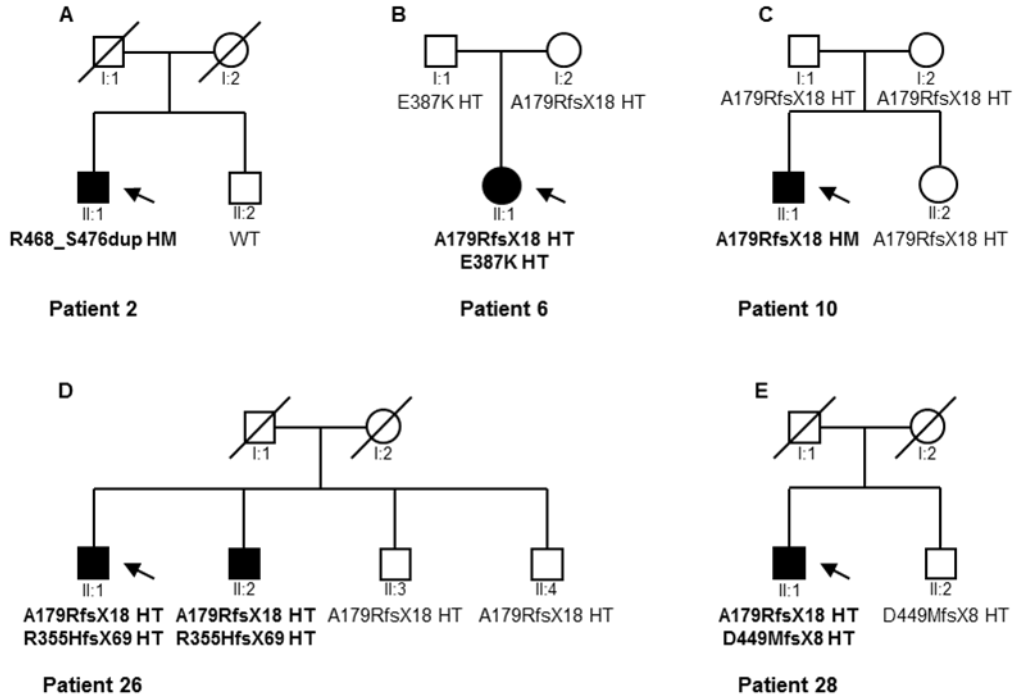
The most frequent mutation identified in this population was the p.A179RfsX18 frameshift mutation (15/74, 20.3%), being the only mutation found in exon 2. All other mutations were located at the 5' end of exon 3. In terms of allele frequencies, the p.A179RfsX18 mutation was followed by p.E387K (11/74, 14.9%), p.R355HfsX69 (8/74, 10.8%), p.T404SfsX30 (6/74, 8.1%) and p.R468_S476dup (4/74, 5.4%) mutations. The p.D449MfsX8, p.L378Q, p.P437L and p.R444X mutations were the least frequent of the studied patients (1/74, 1.4%).

None of the mutations found in our PCG patients was identified in the 55 individuals of the control population. Two control individuals carried one missense variant each, p.E229K or p.A443G, in heterozygosity. The association of the homozygous or compound heterozygous *CYP1B1* mutations with PCG was extremely significant ($p = 3.719 \times 10^{-12}$).

Sequencing of the enhancer, core promoter, 5'UTR, introns, and 3'UTR in the patients negative for *CYP1B1* mutations, did not detect relevant variants, suggesting the involvement of other genetic factors in disease development among this group of patients.

2.3.2.*CYP1B1* mutation analysis of PCG families

Relatives of patients 2, 6, 10, 26, 28, 29, 38, 39, 40, and 44 were sequenced for the *CYP1B1* mutations identified in the index case. Segregation of *CYP1B1* mutations was consistent with an autosomal recessive inheritance with 100% of penetrance. Parents of individuals with homozygous mutations had the same mutation in the heterozygous state. When the proband was a compound heterozygous, each parent showed only one of the mutations. Proband unaffected siblings were negative or heterozygous for *CYP1B1* mutations. The two affected siblings studied (brother of patient 26 and sister of patient 29) had the same genotype of the probands (Figure 2.2).



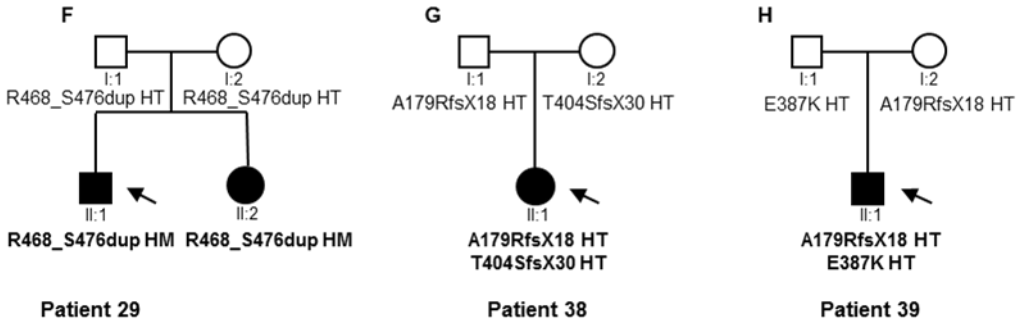


Figure 2.2 – Pedigrees of the available PCG families from patients with *CYP1B1* mutations. *CYP1B1* mutations found in each individual are indicated. HM: homozygous, HT: heterozygous and WT: wild-type. Filled symbols represent affected individuals. Circles specify women and squares men. Arrows indicate the index patients. Oblique lines indicate deceased individuals.

2.3.3.Haplotype analysis

Analysis of the five intragenic polymorphisms in *CYP1B1* identified four different haplotypes among patients and controls. The frequency distributions were particularly different between the group of patients with *CYP1B1* mutations and both the groups without mutations and control. CGGTA was the predominant haplotype observed, with a significantly higher frequency among the group with mutations compared with the groups without mutations and control (0.708 vs. 0.269, p=0.0003 and 0.708 vs. 0.363, p=6.5219 x 10^{-5}, respectively). However, the haplotypes CGCCG and CGCCA were more prevalent in the patients negative for *CYP1B1* mutations and in the control group, when comparing with the mutated group (Table 2.3). All patients with *CYP1B1* homozygous mutations were also homozygous for all the 5 single nucleotide polymorphisms (SNPs), suggesting the association with a particular haplotype (Table S2.2).

Table 2.3 – Estimated *CYP1B1* haplotype frequencies among PCG patients with and without *CYP1B1* mutations and control population. Haplotype prediction and statistical analysis were done using Haploview software.

Haplotype	**Patients with *CYP1B1* mutations** (n=24)	**Patients without *CYP1B1* mutations** (n=9)	**p-value**	**Patients with *CYP1B1* mutations** (n=24)	**Controls** (n=55)	**p-value**	**Patients without *CYP1B1* mutations** (n=10)	**Controls** (n=55)	**p-value**
CGGTA	0.708	0.269	0.0003	0.708	0.363	6.5219E-5	0.269	0.363	0.3666
GTCCA	0.271	0.269	0.9882	0.271	0.354	0.3029	0.269	0.354	0.4126
CGCCG	0.021	0.192	0.0099	0.021	0.200	0.0033	0.192	0.200	0.9295
CGCCA	0.000	0.269	0.0002	0.000	0.074	0.0559	0.269	0.074	0.0043

2.3.4.Genotype-phenotype correlation

Clinical features of PCG were compared between patients with and without *CYP1B1* mutations. The number of patients with the bilateral disease was significantly higher in the group of patients with *CYP1B1* mutations (100.0 % vs. 69.2 %, p=0.011). The mean IOP at the diagnosis was higher for the group with mutations (34.0 vs. 27.2 mm Hg, p=0.001). However, the mean IOP after surgery did not show differences (16.6 vs. 17.2 mm Hg, p=0.579). The number of surgeries was higher for the *CYP1B1* positive group (4.5 vs. 2.6, p=0.025). Sex and family history were not different between the two patient groups (p=0.418 and p=1.000, respectively). The corneal diameter and cup-to-disc ratio were also similar in the groups with and without *CYP1B1* mutations (13.2 vs. 13.4 mm, p=0.807; 0.8 vs. 0.7, p=0.124, respectively).

2.4. Discussion

According to this study, mutations in *CYP1B1* gene are the major cause of PCG in the studied Portuguese population. Most of the PCG patients had *CYP1B1* mutations (64.9%) in the homozygous or the compound heterozygous state. Given the absence of parental consanguinity, this value is much higher than expected. The prevalence of *CYP1B1* mutations in non-consanguineal PCG patients ranges from 22.2% in Japanese (Fuse et al., 2010), 33.3% in Indonesians (Sitorus et al., 2003), 36.0% in Spanish (Milla et al., 2013), 44% among Indians (Chakrabarti et al., 2010), to 50.0% among Brazilians (Stoilov et al., 2002). In a recent study of 21 Portuguese PCG patients, only six (28.6%) had mutations in *CYP1B1* (Cardoso et al., 2015), a much lower percentage compared to the one obtained in our study. Such high difference was not expected but may be explained by the different geographic origin of the patients. This population represents the central region of the country, which is more stable and less subject to immigration flows. Cardoso and collaborators studied a Portuguese population from Lisbon, the capital city of Portugal, which, by internal migration events in the last two centuries, has had an increase in population coming from the different regions of the country, being, therefore, more heterogeneous.

In this study nine different disease-causing mutations were identified: p.A179RfsX18 (Belmouden et al., 2002), p.R355HfsX69 (Stoilov et al., 1997), p.T404SfsX30 (Stoilov et al., 1998), p.D449MfsX8 (Stoilov et al., 1998), p.R468_S476dup (Stoilov et al., 1998), p.L378Q (Della Paolera et al., 2010), p.E387K (Stoilov et al., 1998), p.P437L (Stoilov et al., 1998), and p.R444X (Colomb et al., 2003). All were previously reported, confirming the high heterogeneity of *CYP1B1* mutations among PCG patients and the absence of inbreeding of the population. The p.A179RfsX18 mutation, which leads to the creation of a premature stop codon with the loss of 63% of the peptide, was the most frequently mutated allele (20.3%). Similarly, this

mutation was also the most common in another Portuguese patient study (25.0%) (Cardoso et al., 2015), as well as in the Brazilian (20.2%) (Della Paolera et al., 2010) and Moroccan (25.0%) patients (Belmouden et al., 2002). These observations are indeed interesting and could reflect in some way the Portuguese expansion overseas that began in 1415 with the conquest of Ceuta in Morocco and extended along the coasts of Africa, India, and Brazil. The other frameshift mutations p.R355HfsX69, p.T404SfsX30, and p.D449MfsX8 originate premature stop codons and, therefore, truncated proteins with the loss of 189, 140, and 95 amino acids, respectively (Stoilov et al., 1998). The nonsense p.R444X mutation leads to the loss of 99 amino acids. These truncating mutations eliminate the carboxyl terminus and at least, the heme-binding region of the CYP1B1 protein that is essential for proper function (Figure 2.1). Therefore, these mutations are expected to originate functional null alleles (Stoilov et al., 1998). The non-frameshift p.R468_S476dup mutation leads to the insertion of 9 amino acids in the L-helix (Stoilov et al., 1998), which may disrupt the protein folding. This mutation was identified in the homozygous state in a Spanish patient and was also considered to be a null allele (Lopez-Garrido et al., 2013). The p.E387K and p.P437L non-synonymous mutations occur in highly conserved positions among different species and different cytochromes P450 (Della Paolera et al., 2010). p.E387K is located in the helix K region and P437L in the meander region that precedes the heme-binding region. The functional study of p.E387K and p.P437L revealed that these mutations also create null alleles (Lopez-Garrido et al., 2013; Medina-Trillo et al., 2016). p.L378Q immediately precedes the helix K region, in an evolutionary-preserved position (Della Paolera et al., 2010). It is expected that these non-synonymous mutations at the end of the protein influence the heme-binding ability and proper folding, also creating functional null alleles (Figure 2.1) (Stoilov et al., 1998). All the mutations found in this study segregated with the disease phenotype and were consistent with the autosomal recessive model with complete penetrance (Suri et al., 2009).

No mutations were detected in the enhancer, promoter, 5'UTR, introns, and 3' UTR, which could also contribute to the production of a defective protein or with a lower protein expression. Until date, no intronic mutations have been identified in PCG patients, as well mutations in the *CYP1B1* enhancer. Two studies detected two different variants in the *CYP1B1* promoter. Chakrabarti and colleagues found a strong functional association of the rs2567206 polymorphism in PCG cases harboring homozygous *CYP1B1* mutations, associated with a more severe phenotype. *In vitro* studies revealed that this variant significantly reduced the promoter activity of *CYP1B1* in TM cells, suggesting its involvement in PCG pathogenesis (Chakrabarti et al., 2010). In a more recent study, Lopez-Garrido and collaborators found a PCG patient with a novel heterozygous c.-337G>T mutation, inherited from the mother. This patient also

presented the p.A179fsX18 mutation in the heterozygous state transmitted by the father. His grandmother diagnosed with POAG was a carrier of the c.-337G>T mutation. The authors considered that this mutation could lead to lower or absent CYP1B1 levels through the impairment of *CYP1B1* transcription suggesting a functional analysis of this mutation to prove its pathogenic effect (Lopez-Garrido et al., 2013). Despite the absence of mutations in our patients, other mechanisms related to transcription, translation or post-translation regulation might be impaired and so contributing to a lower or even absence of CYP1B1 levels.

None of the mutations found in the PCG patients were identified in the control group. In addition to the well-known polymorphisms, two heterozygous missense mutations, p.E229K and p.A443G, were found in two controls. The p.A443G has been considered a polymorphism, classified as benign according to SIFT (Ng and Henikoff, 2003) and PolyPhen-2 (Adzhubei et al., 2013) pathogenicity predictors and reported with an allele frequency of 2.6% and 7.0% in Spanish and Ethiopian healthy individuals, respectively (Aklillu et al., 2002; Aklillu et al., 2005). The p.E229K was initially described as a PCG-causing mutation, but further studies associated this mutation with the creation of a hypomorphic allele with incomplete penetrance (Campos-Mollo et al., 2009). It is particularly frequent in the NHLBI Exome Sequencing Project (ESP) white population (41 of the 4296 individuals) (Wiggs et al., 2014), suggesting that it may not be a mutation, but a risk allele for POAG as previously described by Chavarria-Soley and colleagues (Chavarria-Soley et al., 2008). As reported by Wiggs and colleagues, a higher number of missense mutations rather than nonsense and frameshifts are expected in the population without glaucoma (Wiggs et al., 2014). These authors suggested that missense mutations could be less severe not only in the PCG but also in the other later onset types of glaucoma, which is consistent with the variable expressivity identified for some of the missense mutations. In fact, our patients also have more protein truncating mutations than missense mutations (Table S2.2) and all samples with missense mutations are in association with a truncating mutation, except for p.E387K. This total or almost complete absence of CYP1B1 function may be related with the complete penetrance observed in this study.

Haplotype analysis showed that CGGTA was the most frequent haplotype associated with *CYP1B1* mutations. This haplotype had a significantly higher frequency among the group with mutations comparing with the patients without mutations and the controls. Other studies had also demonstrated the association of several *CYP1B1* mutations with this haplotype background, even in very different populations (Chakrabarti et al., 2006). p.A179RfsX18 (Belmouden et al., 2002; Stoilov et al., 2002), p.R355HfsX69 (Stoilov et al., 2002), p.T404SfsX30 (Stoilov et al., 2002), p.P437L (Stoilov et al., 2002), and p.R468_S476dup (Chavarria-Soley et al., 2006) mutations were previously linked with this haplotype. All of our patients with these

mutations also have the CGGTA haplotype. In Morocco, the p.A179RfsX18 allele has been associated with a founder effect to the haplotype CGGTA (Belmouden et al., 2002). The same haplotype was also associated with this mutation in Brazilian patients (Stoilov et al., 2002). It is an ancient haplotype with a high number of ancient PCG-related *CYP1B1* mutations, probably due to founder effects (Chakrabarti et al., 2006). Despite the high frequency of this haplotype in *CYP1B1* positive patients, more recent mutations occurred in other haplotype backgrounds (Chakrabarti et al., 2006). The p.E387K associated with the GTCCA haplotype (Plasilova et al., 1999), and the p.D449MfsX8 with the CGCCG haplotype (Stoilov et al., 2002). In this study, the *CYP1B1* mutations and the expected haplotype also matched the associations already reported.

Several studies have addressed possible correlations between the genotype and PCG clinical features, as well as response to treatment. In this study, statistically significant differences in the number of bilateral eye disease cases, in the mean value of IOP at diagnosis and the number of surgeries, which were higher in the group with *CYP1B1* mutations, were found. A recent study in Moroccan patients also demonstrated a significant increase in these three parameters in *CYP1B1* mutation carriers (Berraho et al., 2015). In the recent study on Portuguese patients, a statistically significant high number of bilateral eye disease cases in the group of patients with *CYP1B1* mutations was found, as well as a lower age at diagnosis. Similarly to our study, there were no differences in postoperative IOP and corneal diameter. However, contrary to our study, the number of surgical interventions was not statistically significantly different (Cardoso et al., 2015). Della Paolera and collaborators found a higher rate of bilateral glaucoma and number of surgeries in the presence of *CYP1B1* gene mutations, but no differences for the corneal diameter and initial IOP variables (Della Paolera et al., 2010). Abu-Amero and collaborators also demonstrated that PCG patients with *CYP1B1* mutations did tend to have a lower operation success rate (Abu-Amero et al., 2011). However, de Melo and collaborators used a logistic regression in a large cohort of Brazilian and Indian patients and did not find genotype-phenotype correlations (de Melo et al., 2015). According to our and the majority of studies, *CYP1B1* mutations associated with more aggressive glaucoma phenotypes.

In conclusion, our study identified nine different *CYP1B1* mutations in 64.9% of the PCG patients, an unexpectedly high prevalence which may be explained by the geographic origin of our patients. The p.A179RfsX18 frameshift mutation was the predominant altered allele, which is in agreement with the recent study on Portuguese PCG patients (Cardoso et al., 2015). This mutation was linked to the ancient CGGTA haplotype background, suggesting a founder effect. Additionally, patients with *CYP1B1* mutations had more aggressive clinical phenotypes with a higher rate of bilateral eye disease and mean IOP at diagnosis. These observations strengthen

the importance of *CYP1B1* mutational screening in PCG patients and at-risk relatives when disease-causing alleles of an affected family member are identified. Molecular diagnostic methods should be implemented in routine clinical practice to assist the medical community in the management of this devastating condition since early diagnosis and the most suitable surgical intervention result in a better prognosis.

CHAPTER 3

Whole exome sequencing analysis of primary congenital glaucoma patients without *CYP1B1* mutations

3.1. Introduction

Primary congenital glaucoma is a rare infantile ocular disease that may have a huge impact on patients' life. In fact, PCG is considered a leading cause of vision loss in children. If not diagnosed and treated promptly it can lead to irreversible blindness. (Yang et al., 2009). The genetic causes of the disease have been studied over the years to understand the pathogenic mechanisms behind PCG development. As previously described, *CYP1B1* was the first gene associated with PCG (Stoilov et al., 1997). The number of PCG patients with *CYP1B1* mutations depends on the population: in consanguineous populations almost all patients have *CYP1B1* mutations. On the contrary, the prevalence in non-consanguineous patients ranges from 22% in Japanese (Fuse et al., 2010) to 50% among Brazilians (Stoilov et al., 2002). These values suggest that in patients without *CYP1B1* mutations other genes may be involved in PCG pathogenesis.

Primary congenital glaucoma has been considered an autosomal recessive disease. Almost all patients with *CYP1B1* mutations have homozygous or two compound heterozygous mutations (Yang et al., 2009). However, a few PCG patients have been reported with only one heterozygous mutation in *CYP1B1*. These patients may have undetected variants, for example in regulatory sequences such as the promoter or 5' and 3' UTRs. Another possible explanation for PCG associated with *CYP1B1* heterozygous mutations is a digenic disease model, with a second mutated gene (Campos-Mollo et al., 2009; Weisschuh et al., 2009; Li et al., 2011). A few cases have been reported for this disease model. Two heterozygous mutations in different genes were found in one PCG patient, one in *CYP1B1* and another in *MYOC* (Kaur et al., 2005) Additionally, two PCG patients had heterozygous mutations in *CYP1B1* and *FOXC1* (Chakrabarti et al., 2009). Furthermore, cases of unaffected individuals with the same genotype of PCG patients were described, suggesting incomplete penetrance (Bejjani et al., 2000; Hollander et al., 2006; Campos-Mollo et al., 2009; Lopez-Garrido et al., 2013; Chen et al., 2015). The presence of a rare dominant locus may be responsible for this incomplete penetrance (Bejjani et al., 2000). Variable expressivity was also reported for *CYP1B1* mutations. Individuals with associated PCG mutated alleles were diagnosed with JOAG and POAG (Bayat et al., 2008; Suri et al., 2009).

Although *CYP1B1* is the most prevalent gene associated with the PCG Mendelian autosomal recessive model, homozygous mutations were also identified in *LTBP2* in Pakistani and Gypsy patients (Ali et al., 2009; Narooie-Nejad et al., 2009; Micheal et al., 2016). Recent genetic findings in *FOXC1* and *TEK* suggest a new PCG disease model, the autosomal dominant model (Chakrabarti et al., 2009; Medina-Trillo et al., 2015; Souma et al., 2016). The study of Chakrabarti and collaborators revealed for the first time five PCG patients with heterozygous

FOXC1 mutations. The study of Medina-Trillo and collaborators on this gene, in 2015, confirmed the association of *FOXC1* mutations with PCG. These authors reported two additional patients with heterozygous mutations. Very recently, in June of 2016, Souma and collaborators reported 10 PCG patients with heterozygous mutations in *TEK*. With this study the authors included *TEK* in the list of PCG candidate genes, reinforcing the possibility of an autosomal dominant model with incomplete penetrance and variable expressivity that warrants future investigation.

The PCG patients negative for *CYP1B1* mutations identified in Chapter 2 constituted an excellent population for genetic studies to discover new genes and mechanisms associated with disease development. At the beginning of this study, besides *CYP1B1*, only three genes in a few cases had been reported as associated with PCG (*LTBP2* in nine families, *MYOC* in four patients, and *FOXC1* in five). For that reason the probability of finding causative mutations in these three candidate genes was low. Whole exome sequencing (WES) was the method selected to analyze the group of patients with no genetic cause known yet to guarantee a complete study. This technique gave the opportunity to study the coding regions of the candidate genes and to discover new mutated genes, reducing the time and cost of the study. Firstly, known candidate genes were evaluated for rare and pathogenic variants. Secondly, since the autosomal recessive pattern of inheritance is the most common model for PCG, the recessive model was considered. After, alternative models of inheritance were taken into account, considering the dominant pattern and the presence of de novo mutations when possible.

3.2. Material and Methods

3.2.1. Patients description

The 13 PCG patients without *CYP1B1* mutations from the initially studied population (Chapter 2) were selected for WES. The number of male patients was superior to female patients (8 males (61.5%) vs. 5 females (38.5%)). Four patients (30.8%) were diagnosed with unilateral glaucoma. Patient 9 was the only individual with a positive familial history of PCG and parental consanguinity (Table S3.1).

Another patient with congenital glaucoma and without *CYP1B1* mutations was included in this study: patient 31 (Table S3.1). Patient 31 was diagnosed with unilateral PCG in the right eye. This patient suffered from neonatal sepsis with petechiae, although without positive bacterial cultures. He had delayed psychomotor development in the first year of life, with poor dynamic postural control and lack of expressive speech. An interauricular communication was also detected. When the boy was two years old, he continued to show delayed development, with

a dynamic postural control characteristic of a 10-month-old child and other developmental areas similar to a 15-month-old child. This delayed development observed for this young boy suggested that congenital glaucoma could be associated with a not-yet-identified condition and not a true PCG case. The parents of this child gave their informed consent for inclusion in this study.

WES validation involved sequencing of a PCG patient with compound heterozygous mutations in *CYP1B1*. Patient 19 was thus sequenced as a positive control.

Parents of patients 1 and 31 were also sequenced, allowing a trio mutation discovery approach. Samples from both parents of the other patients were not available for trio sequencing at the time exome sequencing took place.

3.2.2.Whole exome sequencing

Genomic DNA of patient 31 was extracted from peripheral blood using DNeasy Blood & Tissue Kit (QIAGEN, Hilden, Germany), as described in Chapter 2. Saliva samples provided by the parents of patients 1 and 31 were collected with Oragene DNA (OG-500) collection system (DNA Genotek, Ottawa, Canada) and DNA was extracted according to the manufacturer's instructions. DNA integrity and purity were checked by agarose gel electrophoresis and O.D. ratio measurement at 260/280 and 260/230. DNA was quantified by a fluorometric method using Qubit® dsDNA HS (High Sensitivity) Assay Kit (Thermo Fisher Scientific, Eugene, USA) in the Qubit® Fluorometer (Thermo Fisher Scientific, Carlsbad, USA).

All samples, except 31, 32 and 33, were amplified using Ion AmpliSeq™ RDY Exome Kit (Thermo Fisher Scientific, Carlsbad, CA) and sequenced on the Ion Proton™ System (Thermo Fisher Scientific, Carlsbad, USA). Briefly, whole exonic regions of the genome were amplified from 75 ng of DNA using 12 primer pools. After partial digestion of primer sequences, barcode adaptors were ligated to the amplicons. Purified libraries were quantified by qPCR. The following steps were library amplification by emulsion PCR using Ion PI™ Hi-Q™ OT2 200 Kit and sequencing in the Ion PI chip v3 with Ion PI™ Hi-Q™ Sequencing 200 Kit on the Ion Proton™ Sequencer (Thermo Fisher Scientific, Carlsbad, USA). All the procedures were done according to the manufacturer's instructions. After the run, the reads, the sequences of nucleotides obtained in the sequencing process, were quality trimmed and aligned to the hg19/GRCh37 reference genome using the TMAP software, version 4.0.6 (Thermo Fisher Scientific, Carlsbad, USA). Torrent Variant Caller, version 4.4.2.1 (TVC, Thermo Fisher Scientific, Carlsbad, USA) was used for variant calling (SNPs and Indels), with default parameters for Ion AmpliSeq™ RDY Exome Kit.

WES of samples 31, 32 and 33 was performed by Oxford Gene Technology company (Oxford, UK) using SureSelect Human All Exon V4+UTR capture kit for library preparation and Illumina HiSeq 2000 sequencer. Briefly, exonic blunt-ended DNA fragments with indexing-specific adaptors were captured with probes. The amplified library was sequenced on the Illumina HiSeq 2000 platform using TruSeq v3 chemistry. Reads were mapped to their location in the human genome (hg19/GRCh37) using the Burrows-Wheeler Aligner (BWA) package, version 0.6.1 (Li and Durbin, 2010). Duplicate reads were marked with Picard version 1.62. A local realignment of the mapped reads around potential insertion/deletion (indel) sites was carried out with the Genome Analysis Tool Kit (GATK), version 1.6 (McKenna et al., 2010). This software also recalibrated base quality (Phred scale) scores and called SNPs and Indels.

Exome variants were annotated with the Annovar software using the web server wANNOVAR. ANNOVAR is an annotation tool that uses information documented in several databases to annotate detected genetic variants. Besides transcript and protein changes, this tool provides information regarding the presence of the variant in the dbSNP, allele frequency, functional impact, and conservation among others (Yang and Wang, 2015). Annotation was performed using the RefSeq set of transcripts.

3.2.3. Variants prioritization

A set of filters was applied to SNPs and Indels of all exomes to identify the disease-causing variants and genes. Exonic and splice-site (3 exonic and 8 intronic splice site nucleotides, according to the Ensembl database) variants were selected, and synonymous variants were excluded. Variants were further selected for minor allele frequency (MAF) inferior to 1%, at least in one of the databases, according to the information in the European population of the 1000 Genome Project (1000 G)(Auton et al., 2015), in the European Americans of the NHLBI ESP (http://evs.gs.washington.edu/EVS/), in the non-Finnish European of the Exome Aggregation Consortium (ExAC) (Lek et al., 2016) and in an in-house database with 55 exomes and genomes of Portuguese individuals without congenital disorders. The alignments of all selected variants were manually verified using Integrative Genomics Viewer (IGV) software to exclude evident false positives due to sequencing errors.

The discovery of disease-causing variants and genes involved different approaches: variants in genes already associated with PCG, new genes compatible with the recessive inheritance model, new genes following the dominant inheritance model and genes with *de novo* variants.

For the strategy based on genes associated with PCG all rare variants of *LTBP2*, *MYOC*, *FOXC1*, and *TEK* were selected.

Genes with homozygous variants and with two or more heterozygous variants were chosen to study the recessive inheritance model. In patients 1 and 31, where the trio was sequenced, variant segregation was analyzed. Only homozygous variants present in the heterozygous state in parents continued for evaluation. In cases of compound heterozygosity, each variant must have to be inherited from a different progenitor. For the remaining samples, a strategy based on linkage was adopted. The public frequencies of heterozygous variants of a gene were compared. When MAF values were identical, a strong possibility of linkage was considered, indicating that the variants were inherited from the same parent. For this reason, these variants were excluded. Due to the high number of variants identified in each patient, variants were further evaluated to select those to be analyzed in detail according to the following criteria. Considering PCG is a rare disease in all world populations variants were selected if additionally to the European populations, MAF was also inferior to 1% in other populations. Variants were also chosen if there were no individuals with homozygous variants in the databases. Finally, variants were classified for functional impact using nine different software and those considered pathogenic (6 out of 9, section 3.2.4) were selected.

Analysis of the dominant inheritance model assumed that only very rare and pathogenic variants could be responsible for PCG. As such, only variants absent in all populations in the public databases and predicted as pathogenic by the software were considered. As the number of variants was high, variants to analyze in detail were selected according to gene function, known expression in TM, pathogenicity predicted by all software (section 3.2.4), or variant type: nonsense, stop loss, frameshift, and non-frameshift in non-repeated protein regions. In trios, variants were chosen after segregation analysis.

De novo variants were analyzed in patients 1 and 31, due to the availability of the parent's exomes. Variants were selected if heterozygous in the patient and wild type in parents or homozygous in the patient and heterozygous in only one of the parents. Additionally, all variants not in accordance with the genotype of parents were inspected. Alignments of the identified variants were manually verified to validate correct zygosity calling.

Exomes of patients were also analyzed in Exomiser (Smedley et al., 2015) to help in the discovery of new PCG associated genes. This program can be used to prioritize genes and variants obtained from next-generation sequencing (NGS) studies for novel disease-gene discovery or differential diagnostics of Mendelian diseases. Cross-species phenotype comparisons, clinical relevance, and protein interaction networks are the base of this application. Other filters like variant frequency, predicted pathogenicity and pedigree analysis can also be applied. Exomiser was run for the autosomal recessive, autosomal dominant and X-recessive models. The hiPHIVE method was used with a MAF threshold of 1% and with the

following HPO (Human Phenotype Ontology): primary congenital glaucoma (HP: 0008007), congenital glaucoma (HP: 0001087), buphthalmos (HP: 0000557), megalocornea (HP: 0000485), corneal opacity (HP: 0007957), blepharospasm (HP: 0000643), photophobia (HP: 0000613), and increased intraocular pressure (HP: 0007906).

3.2.4. Gene characterization and *in silico* analysis

Information on gene function, tissue expression, association with diseases, protein domains and loss-of-function (LoF) variant intolerance were obtained to have a complete characterization of the selected genes. This information was collected from several public databases: GeneCards, GeneDistiller, OMIM, HGMD, ClinVar, Uniprot, ExAC, and from literature.

Evaluation of the functional impact of the variants was based on the scores provided by SIFT (Ng and Henikoff, 2003), PolyPhen-2 (Adzhubei et al., 2013), MutationTaster (Schwarz et al., 2014), LRT (Chun and Fay, 2009), MutationAssessor (Reva et al., 2011), FATHMM (Shihab et al., 2013), RadialSVM (Dong et al., 2015), LR (Dong et al., 2015), and CADD (Kircher et al., 2014)and on the inter-species conservation scores provided by GERP++ (Davydov et al., 2010) and PhyloP (Siepel et al., 2006). A variant was classified as pathogenic when the majority of functional impact software predicted it as pathogenic (6 out of 9). Splicing variants were initially analyzed with Human Splicing Finder 3.0 (HSF) (Desmet et al., 2009). After, a detailed analysis was performed for a better classification of the predicted modifying splice site variants. For this Splice-Site Analyzer tool (http://ibis.tau.ac.il/ssat/SpliceSiteFrame.htm), NNSplice (Reese et al., 1997), SplicePort (Dogan et al., 2007), SplicePredictor (Brendel et al., 2004), MaxEntScan (http://genes.mit.edu/burgelab/maxent/Xmaxentscan_scoreseq.html), and Hbond (https://www2.hhu.de/rna/html/hbond_score.php) were used. Differences in free protein energies (ddG) between mutant and wild-type protein were calculated using I-Mutant 2.0 (Capriotti et al., 2005). Protein structural variations were visualized with FoldX plugin for YASARA (Schymkowitz et al., 2005) when protein structures were available in the Protein Data Bank. Sequences were aligned across species using Clustal Omega (Sievers et al., 2011), a multiple sequence alignment program. The probability of LoF intolerance of a gene was collected from the ExAC database.

3.2.5.Sanger sequencing

Validation of the identified variants in *TEK*, *MYOC*, *RNASEH2C* and *CNNM3* in patients or parents was performed by Sanger sequencing. *TEK* promoter and 3'UTR and regions without or lower coverage in *FOXC1* were also Sanger sequenced. DNA samples were amplified by

touchdown PCR for all fragments in 50 µL reactions using the following conditions: 100 ng of cDNA, 1X Phusion GC Buffer, 0.2 mM of each dNTP, 0.2 µM of each primer (Table 3.1), 6% dimethyl sulfoxide (DMSO) or 1.5 M of betaine, and 0.02 U/µL of Phusion DNA Polymerase (Invitrogen, Carlsbad, CA). The PCR program consisted in one denaturing step at 98°C for 3 minutes followed by another denaturing step at 99°C for 1 minute; a touchdown phase with a decrease of 1°C in the annealing temperature per cycle: 14 cycles of denaturation at 98°C for 10 seconds, specific annealing temperature plus 14°C with a decrease of 1°C per cycle for 25 seconds and extension at 72°C for 25 seconds; another generic amplification phase: 21 cycles of denaturation at 98°C for 10 seconds, specific annealing temperature for 25 seconds and extension at 72°C for 25 seconds; and a final extension step at 72°C for 5 minutes. The amplicons were sequenced in both directions and electropherograms aligned to reference sequences of *TEK* (ENSG00000120156), *MYOC* (ENSG00000034971), *FOXC1* (ENSG00000054598), *RNASEH2C* (ENSG00000172922), and *CNNM3* (ENSG00000168763), respectively, using the ContigExpress software.

Table 3.1 – Oligonucleotide primers used for amplification and sequencing of *TEK* variant, promoter and 3'UTR, *MYOC* variant, *FOXC1* low coverage regions, *RNASEH2C* variant and *CNNM3* variant. Primer sequence, amplicon size and main PCR conditions are listed. Prom: promoter, bp: base pair.

ID	Primer sequence (5'- 3')	Amplicon length (bp)	Annealing temperature (°C)	PCR additives
TEK_F	CTTGACCATGTCAGGGAAAGC	379	56	DMSO 6%
TEK_R	ATACAACCTCCAACCAATGCCT			
TEK_prom_1_F	GAACCTGGGCTTTTTCACACC	798	56	DMSO 6%
TEK_prom_1_R	GGGCTACTGGGATCTCTGAC			
TEK_prom_2_F	AAAAGCCAAAACGGCTTCCC	860	56	DMSO 6%
TEK_prom_2_R	TAACACAGGGGCCAATCACC			
TEK_3'UTR_F	GACCTACGTGAATACCACGCT	1167	56	DMSO 6%
TEK_3'UTR_R	GGAGACCTTTTCAATCTGATGAGC			
MYOC_F	TACCACGGACAGTTCCCGTA	612	56	DMSO 6%
MYOC_R	TGTCTACGCCCTCAGACTACA			
FOXC1_F	AGGACAGGCTGCACCTCAA	983	61	Betaine 5M
FOXC1_R	TGACTCGAACATCTCCCGCA			
RNASEH2C_F	CCAGGACTCGAAGTGTCGTT	604	55	DMSO 6%
RNASEH2C_R	GGAGTCGCCTCTACTGTTGG			
CNNM3_F	CAGGTGCTGCGCGAGAG	487	58	DMSO 6%

ID	Primer sequence (5'- 3')	Amplicon length (bp)	Annealing temperature (°C)	PCR additives
CNNM3_R	CGTCCAGCATGAAGCAGTCT			

3.3. Results

3.3.1.Exome sequencing quality control and metrics

All samples passed the quality control for WES (Table 3.2). Most of the sequenced reads mapped to the reference human genome (99.11%). Mean coverage represents the average number of times that nucleotides are read, and values superior to 40-fold for samples sequenced by HiSeq 2000 and superior to 130-fold for samples sequenced by Ion Proton™ were obtained, leading to a high confidence in putative variants. All samples except sample 31 had more than 80% of the target bases with 20X coverage or higher. For samples sequenced with Ion Proton™ the uniformity of coverage was superior to 85%, the value predefined by Ion Torrent for high-quality exome sequenced with AmpliSeq™ libraries. Uniformity of coverage is the percentage of bases in target regions with a read depth at least 20% of the average sample coverage. Values for each sample are presented in Table S3.2 in the annex.

Table 3.2 – WES quality control parameters for samples sequenced with Ion Proton™ and HiSeq 2000. The number of reads obtained and the percentage of reads mapping the reference human genome hg19/ GRCh37 were similar. Exomes sequenced with Ion Proton had a higher coverage. SD: standard deviation, NA: not available.

Parameters	Ion Proton™ (Average±SD)	HiSeq 2000 (Average±SD)
Total number of reads	46,679,328.56±5,292,629.41	46,311,860.67±9,191,900.88
% of reads aligned	99.11±0.46	99.11±0.02
Mean target coverage	140.89±22.35	59.74±12.06
% of target bases ≥ 20X coverage	93.95±1.16	83.80±5.87
Uniformity of coverage	90.79±1.05	NA

During the filtration steps, the number of remaining variants was similar for all samples (Table 3.3). Two exceptions were verified. Patient 41 had a high number of variants (62,070 variants). This patient had nearly 10,000 more initial variants than the other individuals, and consequently, the number of variants after the filtration steps was also superior to others. Oppositely, patient 32 had an inferior number of variants after prioritization. The values for each sample are presented in Table S3.3 in the annex.

Table 3.3 – Initial number of variants and after each filter application during the prioritization step. SD: standard deviation.

Filters	Average±SD
Initial number of variants in the VCF	52,095.14±4,366.90
Selection of exonic and splice site variants	22,755.71±1,784.77
Exclusion of synonymous variants	12,576.21±1,122.53
Exclusion of variants present in the 1000G, ESP, ExAC and in-house database with MAF>1% (EUR)	732.79±96.03

Analysis of the control sample with known *CYP1B1* mutations (sample 19) revealed the presence of two heterozygous variants for this gene: A179RfsX18 and T404SfsX30 (Figure 3.1). As expected, the two mutations previously found by Sanger sequencing were now identified using the candidate gene approach based on the recessive model of inheritance. This sample validated the implemented WES methodology, supporting its use in the discovery of new genes associated with PCG.

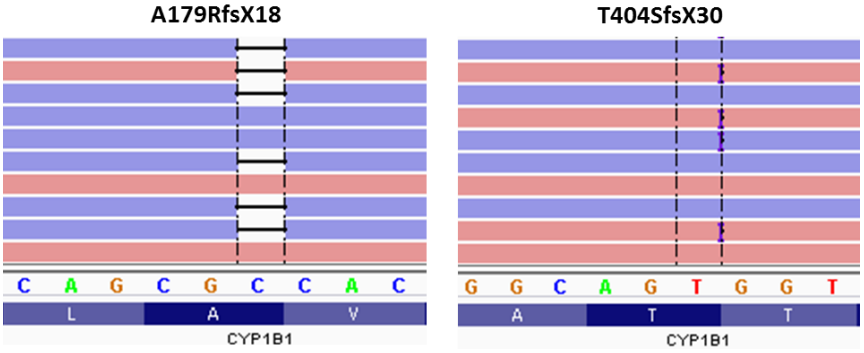


Figure 3.1 – **Mutations found by WES in patient 19.** A179RfsX18 and T404SfsX30 in the heterozygous state were previously identified by Sanger sequencing.

3.3.2.Variants in the PCG candidate genes

LTBP2, *MYOC*, *FOXC1* and *TEK*, already associated with PCG, were analyzed to detect causal variants. Three heterozygous variants were found in *TEK* (c.563T>G and c.1172C>T) and *MYOC* (c.1334C>T) in samples 4 and 36.

TEK c.563T>G variant detected in patient 4 is a novel variant considering the public and the in-house exome variant databases and a new Spanish database (Dopazo et al., 2016). The variant results from the change of valine to glycine at amino acid 188 (p.V188G). Sanger sequencing confirmed the variant (Figure 3.2A). *In silico* predictions classified the variant as pathogenic, except MutationAssessor and FATHMM that considered the variant as low impact and tolerated, respectively (Table S3.19B). Valine 188 is a conserved amino acid in *M. musculus* e *B.*

taurus with GERP++ and PhyloP scores also high (Figure 3.2B and Table S3.19B). The p.V188G variant is in the protein ligand binding Ig2 domain (Figure 3.2C). FoldX calculated an average ddG of 4.93, suggesting the presence of an unstable protein.

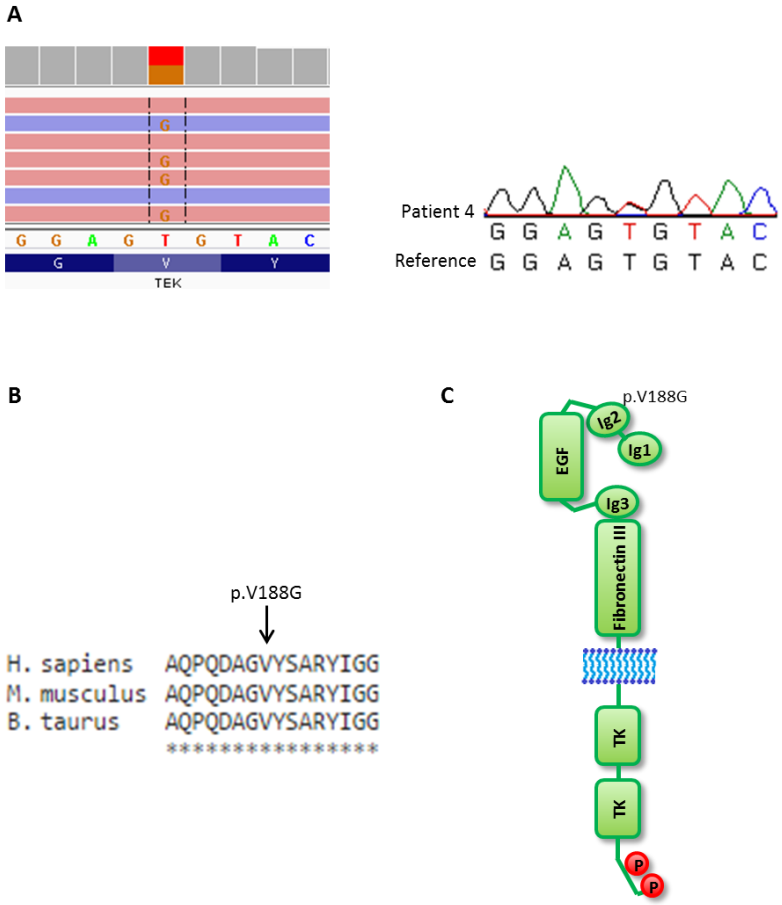


Figure 3.2 – Characterization of the *TEK* p.V188G variant. (A) Thymine changes to a guanine at position 9:27169562. Sanger sequencing confirmed the heterozygous variant in patient 4. (B) The variant occurs in a conserved amino acid across mammals. The alignments were performed with Clustal Omega. (C) The variant locates in the extracellular Ig2 domain.

To guarantee the complete study of *TEK* in patient 4, the promoter and the 5' and 3' UTRs were Sanger sequenced. This supplementary sequencing revealed 4 variants: 9:g.27108800G>A (rs630023); 9:g. 27108975C>T (rs657867); c.*314C>G (rs7037246); and 4:c.*556G>A (rs857). These variants are common in the control populations. In the European population of 1000 G rs630023 has a MAF(G) of 0.214, similar to rs657867 with a MAF(C) of 0.213, and rs7037246 has a MAF(G) of 0.191, almost the same value of rs857 with a MAF(A) of 0.190.

Moreover, patient 4 also presented the c.1334C>T variant in *MYOC* (Figure 3.3A). This p.A445V variant was already reported as rs140967767. It is extremely rare for the European population (MAF(A) of 0.001 for both 1000 G and ESP projects and 0.0002547 for the ExAC population). Despite being associated with POAG (HGMD code CM981350), pigmentary glaucoma and

normal tension glaucoma, this variant was classified as non-pathogenic for all but one (FATHMM) predictors (Table S3.19B). p.A445V is located in the Olfactomedin-like domain (Figure 3.3B), where the majority of the disease-causing variants are reported. However, the amino acid is not highly conserved across species (Figure 3.3C), in accordance with the positive but not very high GERP++ and PhyloP scores (Table S3.19B).

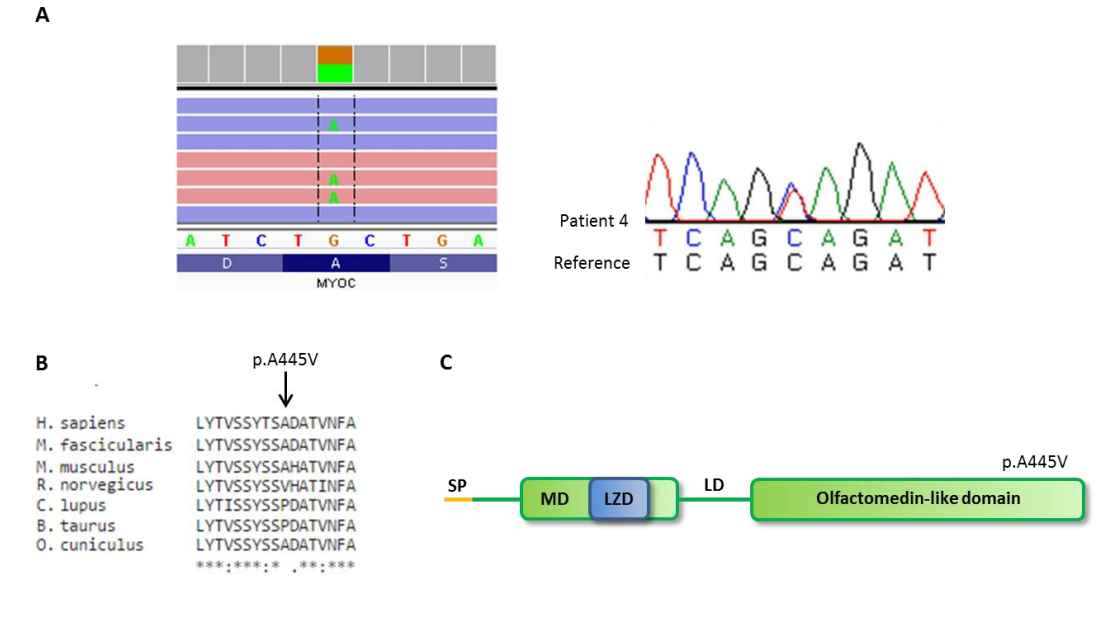


Figure 3.3 – Characterization of the *MYOC* p.A445V variant. (A) A cytosine changes to a thymine at position 1:171605246. Sanger sequencing validated the result obtained by WES. (B) The amino acid 445 is not highly conserved across species. (C) The variant is located in the Olfactomedin-like domain. The majority of the disease-causing variations cluster in this domain. SP: signal peptide, MD: myosin-like domain, LZD: leucine zipper domain, LD: linker domain. Protein N-terminal is located on the left side of the scheme.

The second *TEK* variant was found in sample 36: c.1172C>T, p.T391I, rs34032300, located in the Ig3 domain. Despite the rareness of this variant in the in European populations (MAF(T) of 0 for 1000 G, 0.0003488 for ESP and 0.0001501 for ExAC), it is common in other populations, especially in African populations: MAF(T) of 0.048, 0.044 and 0.04871 for 1000 G, ESP and ExAC. According to these databases, twelve African individuals have this mutation in the homozygous state, suggesting that the variant may not be pathogenic. From the *in silico* analysis, only three software predicted p.T391I as pathogenic, and the variant is not in a highly conserved region between species. These data suggest p.T391I may not be a pathogenic mutation associated with PCG development.

No variants in the candidate genes were found for all the other samples. *FOXC1* had two regions without or with lower values of coverage that were studied posteriorly through Sanger sequencing. These regions did not reveal probably pathogenic variants.

3.3.3.Mutated genes following the recessive model of inheritance

Following the recessive model of inheritance, a different set of genes was identified for each sample. No gene was common to all samples, though some samples shared altered genes, in a total of 12 (Table 3.4). However, none of these common genes had variants with MAF <1% in all populations and were predicted as pathogenic.

Table 3.4 – Genes shared by PCG samples. Patients ID represents the samples that share the gene with variants.

Gene	Patients ID
FCGBP	41, 45
FLNB	41, 45
FRMPD4	41, 43
HYDIN	25, 41
KAT8	18, 36, 42, 43, 45
OBSCN	5, 18
OTOF	42, 43
PDE4DIP	41, 43
PLEC	4, 5
RP1L1	5, 33
TTN	4, 5, 25, 36, 41, 45
XPNPEP2	36, 43

After selection of genes with the rarest variants (without homozygous control individuals described and with MAF<1% in all populations) and predicted as pathogenic by the majority of software or with previous association with glaucoma 16 genes remained. For patients 1 and 31, sequenced in trios, each allele must have to be inherited from each parent. These genes are described in Table 3.5. For samples 1, 4, 32, 36, 42, and 43 no candidate genes following these parameters were identified. Description and characterization of genes and variants according to the autosomal recessive disease model can be found in Tables S3.4 to S3.17.

Additionally, exomes were analyzed in Exomiser for the autosomal recessive model. No gene with a high value (near 1) of Exomiser Score or Phenotype Score was identified, and only *ABCA4* and *COL6A2* were selected (Table 3.6). For the recessive X-linked model Exomiser identified *ARSE*, *BCOR*, and *FLNA* but with low score values (Table 3.6). The potential of this software in gene prioritization for the recessive inheritance pattern was validated by analyzing the *CYP1B1*-positive sample 19 with the same parameters. *CYP1B1* was the first gene on the

top of the list with an Exomiser score of 0.944, a phenotype score of 0.795 and a variant score of 0.853.

Table 3.5 – Candidate genes following the recessive inheritance pattern. Indication of two variants for the same gene represents compound heterozygosity, while a single variant indicates homozygosity. Gene description summary was based on information of GeneCards, GeneDistiller 2014, Orphanet database and Exomiser. LoF: loss of function. *denotes genes also selected by Exomiser.

Sample ID	Gene	Variants	Gene Description
5	*FRMD4A*	p.A813_G819del	Encodes the FERM domain-containing protein 4A. This protein regulates epithelial cell polarity through the connection of ARF6 with the Par protein complex, which regulates the remodeling of adherens junctions and linear actin cable formation during epithelial cell polarization. Gene with LoF intolerance. Polymorphisms associated with Alzheimer's disease. The non-frame deletion leads to the loss of 6 a.a. (GlyGlyAlaGlyGlyAla) in a 3X GlyGlyAla repetition. May not be pathogenic.
5	*STON2*	p.V680LfsX5	Encodes stonin-2, an adaptor membrane protein involved in synaptic vesicle recycling. May facilitate clathrin-coated vesicle uncoating. Gene with LoF intolerance. No known disease association.
9	*ABCA4**	p.G863A	Encodes retinal-specific ATP-binding cassette transporter, an ABC transporter of N-retinylidene-PE. Exclusively expressed in retina photoreceptor cells. Gene linked to a spectrum of phenotypes ranging from retinitis pigmentosa (glaucoma, decreased corneal thickness, and photophobia) to Stargardt disease 1, cone-rod dystrophy, and severe early-onset retinal dystrophy. p.G863A is a mild allele, requiring another null allele to lead to disease.
18	*RYR1*	c.957+5_957+29 delGTGGGGTTTGTG GCGCCCTCCCTCA p.D590N	Encodes ryanodine receptor 1, a calcium release channel in the sarcoplasmic reticulum. Connects the sarcoplasmic reticulum and transverse tubule. Required for normal embryonic development of muscle fibers and skeletal muscle. Associated with susceptibility to malignant hyperthermia, central core disease, and minicore myopathy with external ophthalmoplegia.
25	*GRIK4*	p.H860QfsX188	Encodes glutamate receptor ionotropic, kainate 4. Gene with LoF intolerance. No known disease association.
31	*COL6A2**	p.G171R p.E879K	Encodes collagen alpha-2(VI) chain, one of the three alpha chains of type VI collagen that is a beaded filament collagen found in most connective tissues. Acts as a cell-binding protein. Interacts in mice with COL4A2 that leads to buphthalmos, abnormal lens morphology and microphthalmia. Associated with Bethlem Myopathy 1 and Ullrich Congenital Muscular Dystrophy 1.
31	*MID2*	p.R653H	Located on chromosome X. Encodes probable E3 ubiquitin-protein ligase MID2 that may play a role in microtubule

Sample ID	Gene	Variants	Gene Description
			stabilization involving neural tube closure and other epithelial remodeling processes during development. Gene with LoF intolerance. Associated with X-Linked Mental Retardation, and Non-Syndromic X-Linked Intellectual Disability.
31	*RNASEH2C*	c.G468T	Encodes ribonuclease H2 subunit C, a non catalytic subunit of RNase H2 that specifically degrades the RNA of RNA:DNA hybrids. Participates in DNA replication, possibly by mediating the removal of lagging-strand Okazaki fragment RNA primers during DNA replication. Mediates the excision of single ribonucleotides from DNA:RNA duplexes. Gene with LoF intolerance. Associated with Aicardi-Goutières syndrome, which may cause congenital glaucoma.
33	*ARSE**	p.G172R	Gene located on chromosome X. Encodes arylsulfatase E, a member of the sulfatase family. Interacts with ARSB that is associated with mucopolysaccharidosis type VI (glaucoma, opacification of the corneal stroma and visual impairment) and in mouse leads to abnormal cornea morphology. May be essential for the correct composition of cartilage and bone matrix during development. Has no activity toward steroid sulfates. Associated with Chondrodysplasia Punctata 1, X-linked recessive (cataract).
33	*BCOR**	p.S1298F	Gene located on chromosome X. Encodes BCL-6 corepressor that acts as a transcriptional corepressor when recruited to promoter regions by sequence-specific DNA-binding proteins such as BCL6. Interacts with PRPF31 that is associated with retinitis pigmentosa. Leads in the mouse to retinal degeneration and in fish to abnormal(ly) decreased size eye. Gene with LoF intolerance. Associated with Lenz microphthalmia syndrome (glaucoma, microcornea and visual impairment) and oculo-facio-cardio-dental syndrome with a described case of congenital glaucoma.
33	*SLC38A5*	c.1212C>T	Gene located on chromosome X. Encodes sodium-coupled neutral amino acid transporter 5 that functions as a sodium-dependent amino acid transporter which countertransport protons. Mediates the saturable, pH-sensitive, and electrogenic cotransport of several neutral amino acids including glycine, asparagine, alanine, serine, glutamine and histidine with sodium. Alternative splicing results in multiple transcript variants. Gene with LoF intolerance. No known disease association.
41	*FLNA**	p.R2333H	Gene located on chromosome X. Encodes filamin-A, an actin-binding protein that crosslinks actin filaments and links actin filaments to membrane glycoproteins. Involved in remodeling the cytoskeleton to effect changes in cell shape and migration. Interacts with integrins,

Sample ID	Gene	Variants	Gene Description
			transmembrane receptor complexes, and second messengers. Interacts with *DCN* that is associated with congenital stromal corneal dystrophy (glaucoma, increased corneal thickness, progressive visual loss and congenital corneal dystrophy).Gene with LoF intolerance. Associated with several syndromes, including terminal osseous dysplasia (iris coloboma), periventricular nodular heterotopias, otopalatodigital syndromes, frontometaphyseal dysplasia, Melnick-Needles syndrome, and X-linked congenital idiopathic intestinal pseudoobstruction.
41	*NEUROD4*	p.H282Y	Encodes neurogenic differentiation factor 4 that probably acts as a transcriptional activator. Mediates neuronal differentiation. May be required for the regulation of amacrine cell fate specification in the retina. No known disease association.
41	*TNC*	p.R103H p.V295M	Encodes tenascin, an extracellular matrix protein implicated in guidance of migrating neurons and axons during development, in synaptic plasticity and neuronal regeneration. With a spatially and temporally restricted tissue distribution. Associated with autosomal dominant non-syndromic sensorineural deafness type DFNA.
45	*ATN1*	p.Q488_492Qdel	Encodes atrophin-1, a transcriptional corepressor. Recruits NR2E1 to repress transcription. Promotes vascular smooth cell migration and orientation. Corepressor of MTG8 transcriptional repression. Gene with LoF intolerance. Associated with Dentatorubral-pallidoluysian atrophy.
45	*MICALL1*	p.R320W p.R710P	Encodes MICAL-like protein 1, a probable lipid-binding protein with higher affinity for the phosphatidic acid that is a lipid enriched in recycling endosome membranes. On endosome membranes, may act as a downstream effector of Rab proteins recruiting cytosolic proteins to regulate membrane tubulation. May be involved in a late step of receptor-mediated endocytosis regulating, for instance, endocytosed-EGF receptor trafficking. Alternatively, may regulate slow endocytic recycling of endocytosed proteins back to the plasma membrane. May indirectly play a role in neurite outgrowth. No known disease association.

Table 3.6 – Exomiser information of autosomal and X-linked recessive models genes.

Gene	Exomiser score	Phenotype score	Variant score
ABCA4	0.475	0.654	0.696
COL6A2	0.384	0.504	0.825

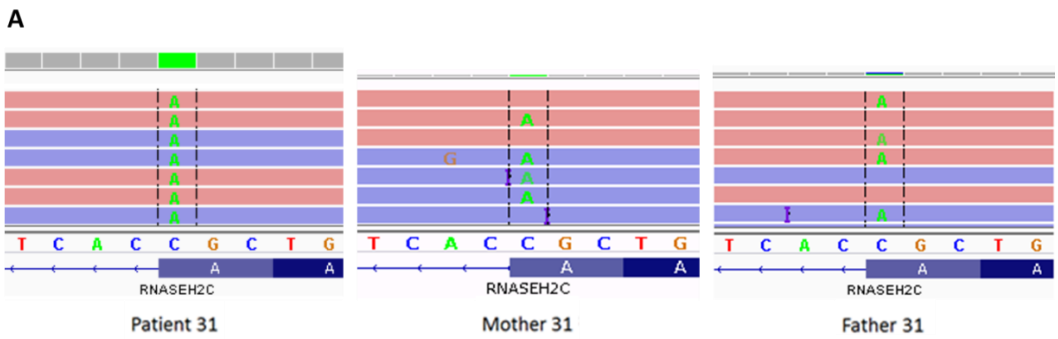
Gene	Exomiser score	Phenotype score	Variant score
ARSE	0.762	0.507	1.000
BCOR	0.714	0.612	0.854
FLNA	0.446	0.502	0.855

Patient 31 presented one of the most interesting genes of the list, *RNASEH2C*, identified with a homozygous synonymous and splicing variant: c.G468T, p.A156A, rs61736590. This rare variant has a MAF(A) of 0.003 and 0.004 for the European populations of 1000 G and ExAC, respectively. Several splicing effect predictor software based on different methods indicated this variant has a strong possibility to lead to the loss of the donor splice site at the end of exon 3 (Table 3.7).

Table 3.7 – Scores obtained with different splicing effect software for the wild-type and the c.G468T *RNASEH2C* alleles. A higher score indicates a higher degree of similarity to the consensus sequence or a higher probability or confidence of a site being a true splice site.

Software	Wild-type score	Mutated Score
Human Splicing Finder	92.33	81.46
Splice-Site Analyzer tool	82.81	70.30
NNSplice	0.99	0.68
SplicePort	1.31	0.50
SplicePredictor	0.97	0.00
MaxEntScan	10.49	7.39
Hbond	17.30	13.80

Parents WES showed that both carried the heterozygous variant (Figure 3.4A). Sanger sequencing validated the presence of this variant in patient 31 and parents (Figure 3.4B).



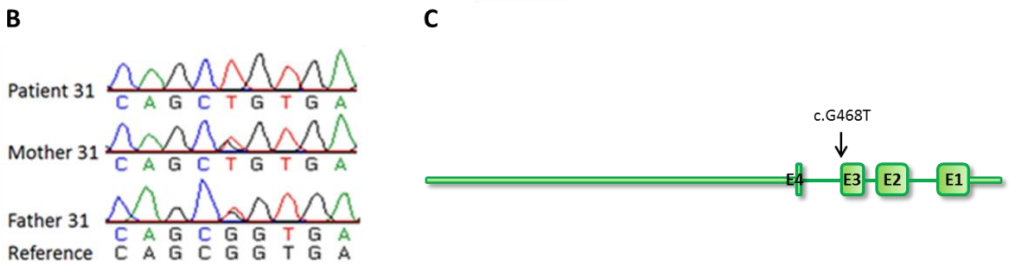


Figure 3.4 – Characterization of the *RNASEH2C* c.G468T variant. (A) A guanine changes to a thymine at position 11:65487516. Patient 1 is homozygous for the variant, while both parents are heterozygous. (B) Sanger sequencing confirmed the results obtained through WES. (C) The variant occurs in the last nucleotide of exon 3. E1: exon 1, E2: exon 2, E3: exon 3, E4: exon 4.

In addition to the *RNASEH2C* variant, two heterozygous variants were found in *COL6A2*, a component of the TM (Acott and Kelley, 2008). Both variants were previously described: c.G511A, p.G171R (rs200710788) and c.G2635A, p.E879K (rs145838734), with a very low frequency in European populations. p.G171R has a MAF(A) of 0.002 in 1000G, 0.0008 in ESP and 0.0013 in ExAC while p.E879K has a MAF(A) of 0.0012 in ESP and 0.0008 in ExAC. p.G171R and p.E879K were considered pathogenic for all but one software, FATHMM and SIFT, respectively (Table S3.10B). Glycine 171 is conserved among species with GERP++ and PhyloP scores also high (4.34 and 3.49), but the glutamate 879 is only present in one isoform of *COL6A2* and is not conserved. The two variants were predicted to result in a slight decrease of protein stability, more relevant for p.G171R (average ddG of -1.14 and -0.59 for p.G171R and p.E879K). Both variants are in Von Willebrand factor A domains (Figure 3.5).

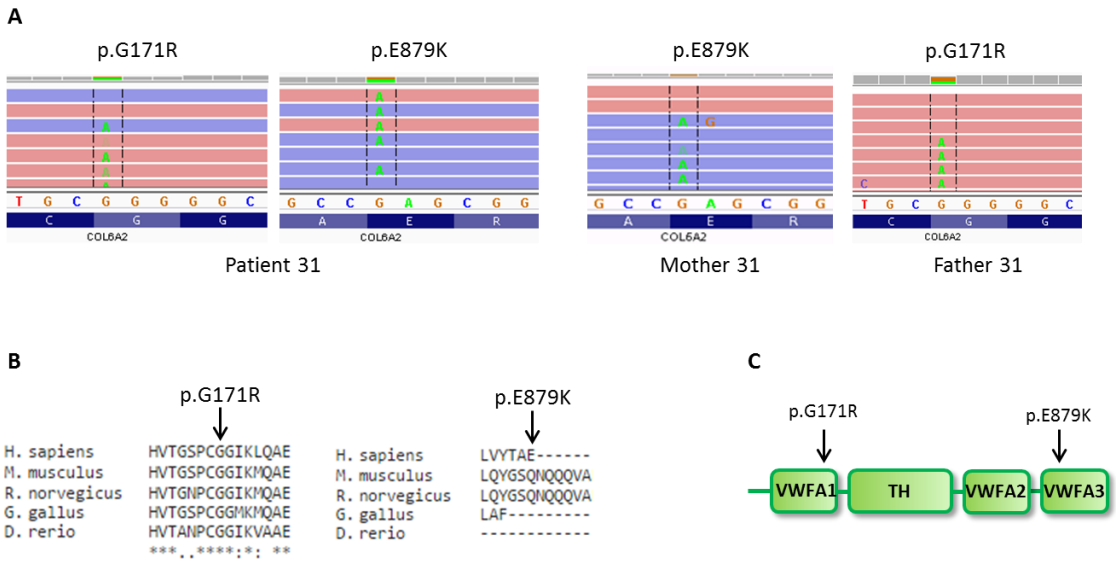


Figure 3.5 – Characterization of the variants p.G171R and p.E879K in *COL6A2*. (A) A guanine changes to an adenine at position 21:47532288 for p.G171R variant and the same nucleotide change occurs at position 21:47549283 for p.E879K. p.G171R was inherited from the father while p.E879K from the mother. (B) The a.a. 171 is highly conserved across species, but the a.a. 879 is not. (C) Variants are located at different Von Willebrand factor A domains.

VWFA1: Von Willebrand factor A domain 1, TH: Triple-helical region, VWFA2: Von Willebrand factor A domain 2, VWFA3: Von Willebrand factor A domain 3. The protein N-terminal is located on the left side of the protein scheme.

Patient 33 had a homozygous variant (c.C3893T, p.S1298F) in *BCOR*, a gene already causative of glaucoma and congenital glaucoma but in association with other symptoms. *BCOR* is located on chromosome X. p.S1298F variant was only reported in the heterozygous state in two African females in ExAC (MAF(A) of 0.0008) and was predicted pathogenic by SIFT, PolyPhen-2, MutationTaster and CADD. Serine 1298 is conserved in mammals with high scores of GERP++ and PhyloP software (Table S3.12B), but a low impact on protein stability is expected (average ddG of -0.19). The variant is located after the AF9 binding domain (Figure 3.6).

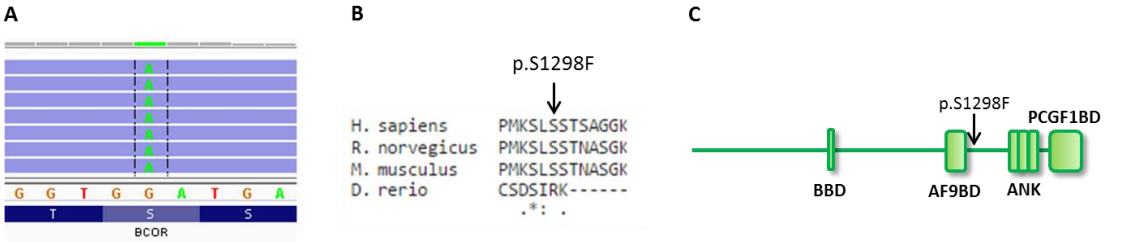


Figure 3.6 – Characterization of the *BCOR* p.S1298F variant. (A) A guanine changes to an adenine at position X:39922279. (B) The amino acid 1298 is conserved across the mammals. (C) p.S1298F occurs after the AF9 binding domain. BBD: BCL6 binding domain, AF9BD: AF9 binding domain, ANK: Ankyrin repeats, PCGF1BD: PCGF1 binding domain. The protein N-terminal is located on the left side of the protein scheme.

Patient 41 had variants in two interesting genes: *FLNA*, previously associated with congenital glaucoma among other characteristics for a specific mutation (Kondoh et al., 2007), and *TNC*, expressed in the TM (Keller et al., 2008). *FLNA* is located in chromosome X and patient 41 had the c.G6998A, p.R2333H variant in this gene. The variant was not reported in the European population, but it was reported in 5 (MAF(T) of 0.0007) and 4 (MAF(T) of 0.0011) African individuals of ExAC and ESP, respectively, under rs377274096. This variant was predicted as pathogenic for all software except SIFT, FATHMM and CADD (Table S3.14B). It locates in a conserved region across mammals of the filamin repeat 22 (Figure 3.7). Decreased protein stability may occur as predicted by the average ddG of -1.57.

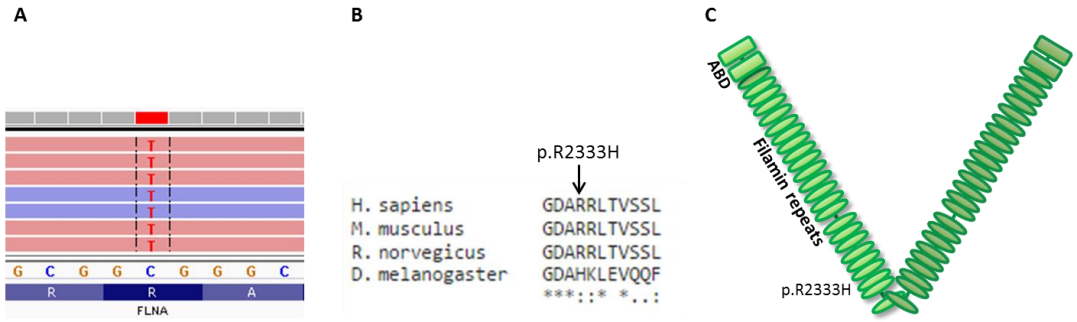


Figure 3.7 – **Characterization of the *FLNA* p.R2333H variant.** (A) A cytosine changes to an thymine at position X:153579974. (B) p.R2333H occurs in a conserved residue across the mammals. (C) The variant localizes in filamin repeat 22, near the self-association site of filamin repeat 24. ABD: actin binding domain. The protein N-terminal is located on the top of the protein scheme.

In *TNC* two heterozygous variants c.G308A, p.R103H (rs150265065) and c.G883A, p.V295M were found. Both variants were not detected in the European population, but they were identified in the African population of ExAC and ESP (p.R103H with a MAF(A) of 0.0009 and 0.0005 in ExAC and ESP, respectively, and p.V295M with a MAF(A) of 0.0005 in ExAC). Both variants are located at the beginning of the protein, conserved across species in accordance with GERP++ and PhyloP scores (Figure 3.8 and Table S3.14B).

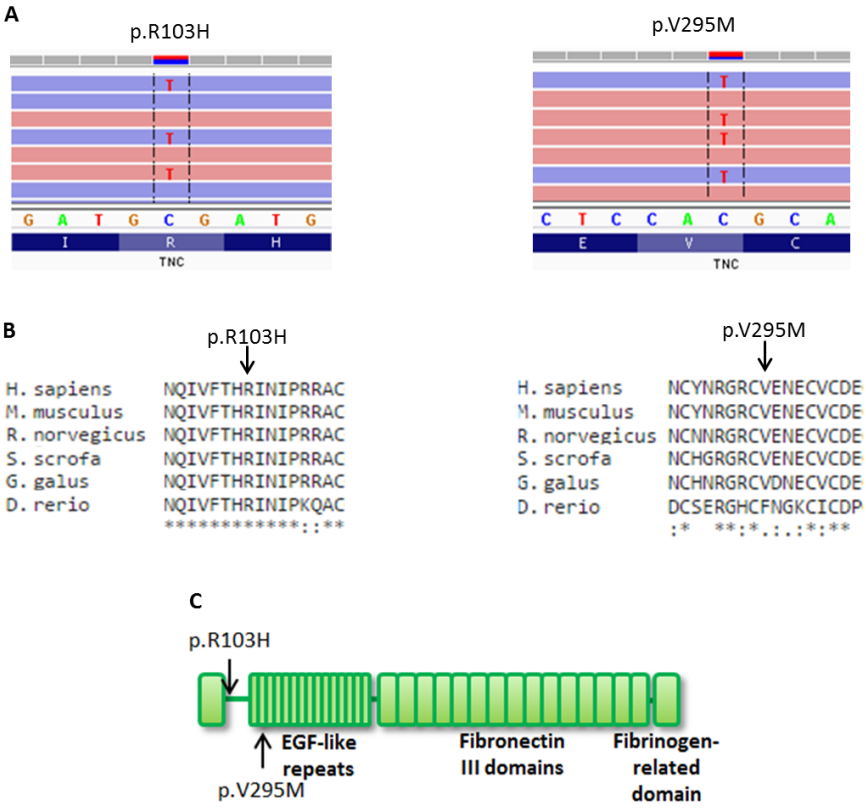


Figure 3.8 – **Characterization of the *TNC* p.R103H and p.V295M variants.** (A) A guanine changes to an adenine at position 9:117852990 for p.R103H variant and the same nucleotide change occurs at position 9:117849127 for p.V295M. (B) Both amino acids are highly conserved across species. (C) p.R103H localizes before the EGF-like

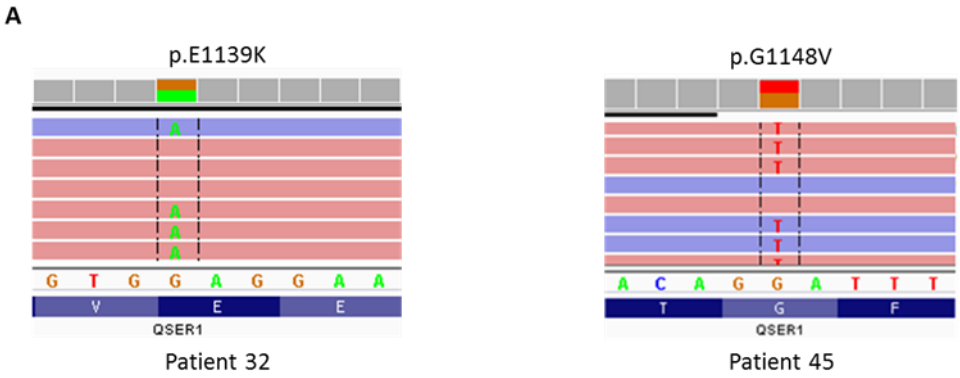
repeats, while p.V295M occurs in the EGF-like 2. EGF: epidermal growth factor. The protein N-terminal is located on the left side of the protein scheme.

Till date, the remaining genes selected as possible PCG candidates and presented in Table 3.4 have not been described as expressed in the TM neither associations with cases of glaucoma are known. However, possible involvement in PCG development should not be discarded.

3.3.4.Genes with variants following the dominant model of inheritance

Recent evidence of a PCG dominant inheritance model led to the analysis of the heterozygous variants found in each patient. Taking into account that PCG is a rare disease in the Portuguese population, variants chosen were those predicted as pathogenic by the majority of the software and not described in public and in-house databases. Each sample revealed its list of genes and variants (Tables S3.18 to S3.31). Only two samples shared two genes: *QSER1* and *ACAN*.

Samples 32 and 45 had two pathogenic *QSER1* c.G3415A (p.E1139K) and c.G3443T (p.G1148V) variants, respectively. p.E1139K was classified as pathogenic except by FATHMM, RadialSVM and LR predictors (Table S3.25B), while for p.G1148V only FATHMM predicted this variant as tolerated (Table S3.31B). The two variants are very close in the protein, in a conserved region, but no domains are known in this region yet (Figure 3.9C). Analysis with I-Mutant2.0 demonstrated a reduction in the stability, especially for p.G1148V (average ddG of -0.89 and -1.45 for p.E1139K and p.G1148V, respectively). This gene has a probability of LoF intolerance of 1, which means that this gene has a high probability of a LoF variant leading to disease, having only 3 LoF variants in ExAC, a much lower value than the predicted 40.3 variants indicated in ExAC.



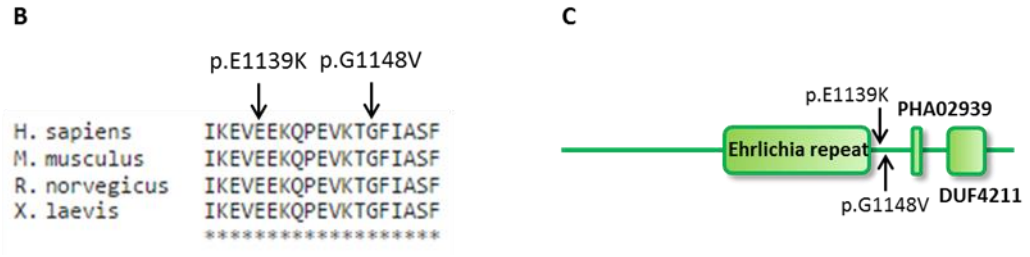


Figure 3.9 – Characterization of the *QSER1* p.E1139K and p.G1148V variants. (A) A guanine changes to an adenine at position 11:32956606 for p.E1139K variant and a guanine changes to a thymine at position 11:32956634 for p.G1148V. (B) Both amino acids are highly conserved across species and (C) Variants are located after the Ehrlichia repeat. The protein N-terminal is located on the left side of the protein scheme.

Patients 33 and 45 harbored *ACAN* variants: c.G1639C (p.G547R) and c.C8A (p.T3N), respectively. Both variants were pathogenic for all software except FATHMM, RadialSVM and LR (Table S3.26B and Table S3.31B) and are located in conserved regions among species (Figure 3.10). I-mutant2.0 software indicated that probably only p.G547R will lead to decreased protein stability (average ddG of -1.35 for p.G547R and -0.53 for p.T3N). ExAC indicated fewer LoF variants (42.2 predicted vs. 2 obtained), being *ACAN* also a gene with a LoF intolerance probability of 1.

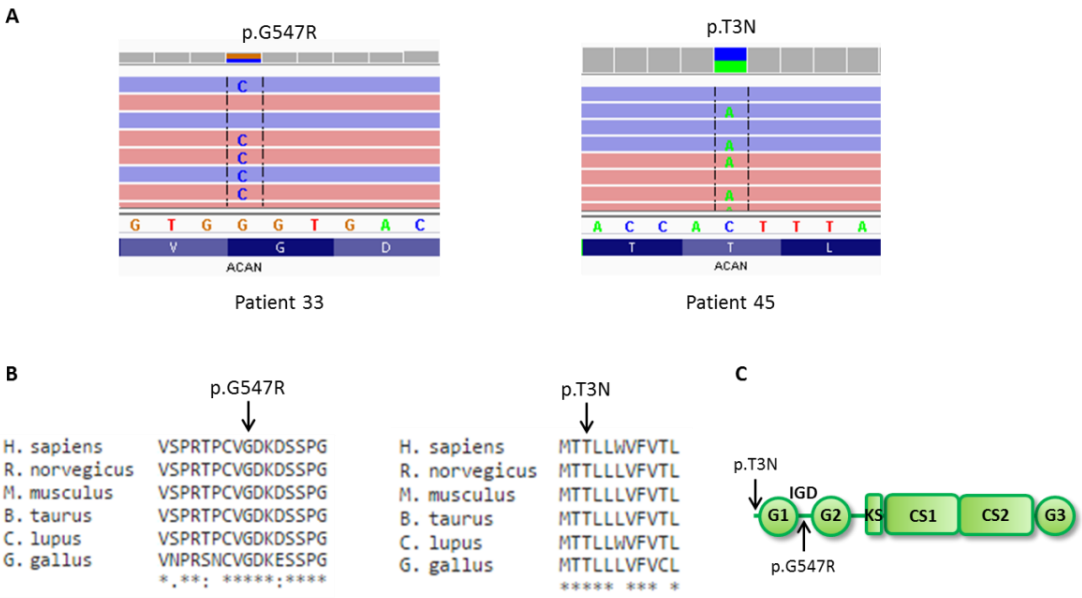


Figure 3.10 – Characterization of the *ACAN* p.G547R and p.T3N variants. (A) A guanine changes to a cytosine at position 15:89391176 for p.G547R variant and a cytosine changes to a adenine at position 15:89379445 for p.T3N. (B) Both amino acids are highly conserved across species. (C) p.T3N is located before the G1 domain, while p.G547R is located in the IGD domain. G1: globular domain 1, IGD: interglobular domain, G2: globular domain 2, KS: keratan sulfate attachment domain, CS1: chondroitin sulfate attachment domain 1, CS2: chondroitin sulfate attachment domain 2, G3: globular domain. The protein N-terminal is located on the left side of the protein scheme.

Independent analysis of samples identified a different set of genes per sample as previously reported. As can be verified in the annex (Tables S3.18 to S3.31) the number of genes for each

sample was high, leading to the selection of only a subset of genes for detailed evaluation. This selection took into account the gene function, known expression in TM, functional impact with variants predicted as probably pathogenic by all software, or specific types of variants such as nonsense or loss of stop, frameshift and non-frameshift variants in a non-repeated protein region.

Following these criteria, 13 genes were chosen from the initial lists of 25 genes (Tables S3.18) for patient 1. All variants were in the heterozygous state and inherited from the parents. Table 3.8 shows the main characteristics of these genes and variants. This patient had three nonsense variants: p.Y146X, p.Q335X, and p.Q1800X in *S1PR3*, *SLC13A2*, and *URB1*, respectively, leading to the loss of about half of the protein for *S1PR3* and *SLC13A2* and the loss of the last 1/5 of the *URB1* encoded protein. No specific domains are known for these three proteins. Two frameshift variants, p.R795KfsX6 and p.Y31X, in *LARS* and *RRH* also led to a premature protein. Similar to *S1PR3*, *NOTCH1* is involved in the angiogenesis, and the p.A533T variant found in this gene was located in the calcium-binding EGF-like domain 14.

Two genes encoding matrix metalloproteinases were identified with possible pathogenic variants: *MMP1* with two heterozygous variants (p.F343L and c.C1029G+4A>G) inherited from the father and located in the hemopexin 2 region of the protein and *MMP19* with the c.520+2T>C splicing variant predicted to affect the donor splice site at the end of exon 4. *NEDD4*, which also contributes to the extracellular matrix, had the missense p.F858V. This variant is located in the WW 3 domain, which participates in the interaction with other molecules. The variant p.A254_V255del in *HECA* was also identified in patient 1 as well as three missense variants predicted as pathogenic by all software in *CLDN9*, *HKDC1*, and *SAMD11* (Table S3.18B). However, no clear association with mechanisms associated with glaucoma are known Table 3.7 lists the genes identified in patient 1, and detailed information is available in Tables S3.18 in the annex.

Table 3.8 – Patient 1 best candidate genes and respective variants following the dominant inheritance model. All variants are in the heterozygous state. Gene description was based on information from GeneCards and ExAC database, Exomiser and literature. LoF: loss of function. * denotes variants predicted as pathogenic by all functional impact predictors. NA: not applicable.

Gene	Variants	a.a. position	Average ddG	Gene and variant description
*CLDN9**	p.G37R	37/217	-1.10	Encodes claudin-9 that plays a major role in tight junction-specific obliteration of the intercellular space, through calcium-independent cell adhesion activity.
HECA	p.A254_ V255del	254/ 543	NA	Encodes headcase protein homolog, a cytoplasmic protein that may be part of the regulatory mechanism

Gene	Variants	a.a. position	Average ddG	Gene and variant description
				in the development of epithelial tube networks such as the circulatory system and lungs. Gene with LoF intolerance.
*HKDC1**	p.M296T	296/917	-0.80	Encodes putative hexokinase HKDC1, whose function is unknown. Located in the hexokinase large subdomain 1.
LARS	p.R795KfsX6	795/ 1,176	NA	Encodes leucine-tRNA ligase, cytoplasmic, a cytosolic leucine-tRNA synthetase that catalyzes the specific attachment of an amino acid to its cognate tRNA.
MMP1	p.F343L c.C1029G +4A>G	343/469	-1.79	Encodes interstitial collagenase, a matrix metalloproteinase that cleaves collagens of types I, II, and III at one site in the helical domain. Also cleaves collagens of types VII and X. Expressed in HTM cell cultures. Both variants located in hemopexin 2 region.
MMP19	c.520+2T>C	173/508	NA	Encodes matrix metalloproteinase-19, a matrix metalloproteinase that degrades various components of the extracellular matrix, such as aggrecan and cartilage oligomeric matrix protein, during development, hemostasis and pathological conditions. Expressed in HTM cell cultures.
*NEDD4**	p.F858V	858/ 1,319	-2.37	Encodes E3 ubiquitin-protein ligase NEDD4, an E3 ubiquitin-protein ligase which accepts ubiquitin from an E2 ubiquitin-conjugating enzyme in the form of a thioester and then directly transfers the ubiquitin to targeted substrates. NEDD4 upregulates expressions of fibronectin and type 1 collagen and contributed to the excessive accumulation of extracellular matrix. Variant located in the WW domain 3 that interacts with other molecules. Exomiser score: 0.755, in humans interact with *NTRK1* that leads to opacification of the corneal stroma and corneal ulceration and in the mouse to blepharoptosis and corneal opacity.
NOTCH1	p.A533T	533/ 2,555	-0.72	Encodes neurogenic locus notch homolog protein 1, a transmembrane receptor that affects the implementation of differentiation, proliferation, and apoptotic programs. Involved in angiogenesis; negatively regulates endothelial cell proliferation and migration and angiogenic sprouting. Expressed in HTM cell cultures. Variant located in the located in the calcium-binding EGF-like domain 14. Exomiser score: 0.761, causative of Adams-Oliver syndrome, which is associated with cataracts. Interacts with *ADAM17* that in mouse leads to increased incidence of corneal inflammation, microphthalmia, corneal opacity, and decreased cornea thickness. Gene with LoF intolerance.

Gene	Variants	a.a. position	Average ddG	Gene and variant description
RRH	p.Y31X	31/337	NA	Encodes visual pigment-like receptor peropsin, a membrane-bound G protein-coupled receptor that may play a role in retinal pigment epithelium physiology either by detecting light directly or by monitoring the concentration of retinoids or other photoreceptor-derived compounds.
S1PR3	p.Y146X	146/ 378	NA	Encodes sphingosine 1-phosphate receptor 3, a receptor for sphingosine 1-phosphate that likely contributes to the regulation of angiogenesis and vascular endothelial cell function. Expressed by the inner wall of SC.
*SAMD11**	p.G572E	572/681	-0.47	Encodes sterile alpha motif domain-containing protein 11 that may play a role in photoreceptor development. Variant located in the SAM domain.
SLC13A2	p.Q335X	335/ 592	NA	Encodes solute carrier family 13 member 2, a cotransport of sodium ions and dicarboxylates such as succinate and citrate.
URB1	p.Q1800X	1,800/ 2,271	NA	Encodes nucleolar pre-ribosomal-associated protein 1, whose function is unknown.

Patient 4, in addition to the p.V188G variant in *TEK,* had 24 novel and possible pathogenic variants (Tables S3.19). Four of these were nonsense variants: p.S43X, p.Q19X, p.Y236X, and p.W64X in *CMYA5*, *DEFB116*, *KLHDC8A*, and *MYCL*, respectively. The variants of *CMYA5*, *DEFB116*, and *MYCL* lead to the deletion of the majority of the a.a. sequences, originating the loss of the protein functional domains. The nonsense p.Y236X in *KLHDC8A* occurs later in the Kelch domain-containing protein 8A, leading to the loss of Kelch repeat 7 only. The frameshift variant p.Q147HfsX18 in *CYTH3* was also found in this patient, leading to the loss of part of the SEC7 domain and the PH domain, which allows the binding to the inositol head group of phosphatidylinositol 3,4,5-trisphosphate. The variant probably leads to the loss of protein function. All predictor software classified the missenses p.N2156S and p.G94E of *ACACB* and *C1QTNF2* as pathogenic (Table S3.19B). p.N2156S is located in the CoA carboxyltransferase C-terminal domain and p.G94E in the collagen-like domain of the complement C1q tumor necrosis factor-related protein 2. Until date, none of these genes are known to be expressed in the TM and their functions do not reveal a strong association with the molecular mechanisms already known for PCG development (Table 3.9), except the previously reported *TEK*.

Table 3.9 – Patient 4 best candidate genes and respective variants following the dominant inheritance model. All variants are in the heterozygous state. Gene description was based on information from GeneCards and ExAC database, Exomiser and literature. LoF: loss of function. * denotes variants predicted as pathogenic by all functional impact predictors. NA: not applicable.

Gene	Variants	a.a. position	Average ddG	Gene and variant description
*ACACB**	p.N2156S	2,156/ 2,458	0.06	Encodes Acetyl-CoA carboxylase 2, an acetyl-CoA carboxylase beta that catalyzes the ATP-dependent carboxylation of acetyl-CoA to malonyl-CoA. Carries out three functions: biotin carboxyl carrier protein, biotin carboxylase and carboxyltransferase. Involved in inhibition of fatty acid and glucose oxidation and enhancement of fat storage. May play a role in the regulation of mitochondrial fatty acid oxidation through malonyl-CoA-dependent inhibition of carnitine palmitoyltransferase 1.
*C1QTNF2**	p.G94E	94/285	0.41	Encodes Complement C1q tumor necrosis factor-related protein 2, whose function is unknown.
CMYA5	p.S43X	43/4,069	NA	Encodes cardiomyopathy-associated protein 5 that may serve as an anchoring protein that mediates the subcellular compartmentation of PKA via binding to PRKAR2A. May function as a repressor of calcineurin-mediated transcriptional activity. May attenuate calcineurin ability to induce slow-fiber gene program in muscle and may negatively modulate skeletal muscle regeneration.
CYTH3	p.Q147HfsX18	147/400	NA	Encodes cytohesin-3, a member of the pleckstrin homology, Sec7 and coiled-coil domains family involved in the control of Golgi structure and function, and it may have a physiological role in regulating ADP-ARF6 functions, in addition to acting on ARF1. Gene with LoF intolerance.
DEFB116	p.Q19X	19/102	NA	Encodes beta-defensin 116, a peptide with antibacterial activity.
KLHDC8A	p.Y236X	236/350	NA	Encodes Kelch domain-containing protein 8A, a protein that when upregulated may provide an alternative pathway for tumors to maintain aggressiveness in the absence of epidermal growth factor receptor dependence.
MYCL	p.W64X	64/394	NA	Encodes protein L-Myc, a protein with function unknown. Gene with LoF intolerance.
TEK	p.V188G	188/ 1,124	-2.99	Encodes angiopoietin-1 receptor, a tyrosine-protein kinase receptor that mediates a signaling pathway that functions in embryonic vascular development. A known candidate gene for PCG. Gene with LoF intolerance.

Patient 5 harbored 30 genes with novel and predicted pathogenic variants (Tables S3.20). Two of these genes had nonsense mutations: p.Q336X in *CRB2* and p.R27X in *TRPT1*. p.Q336X leads

to the loss of BRCT domain and the region of interaction with dimethylated histone H4, while p.R27X causes the loss of almost all the protein sequence (Table 3.10). The non-frameshift insertion p.D151delinsGSDDD found in *PBXIP1* leads to the insertion of 4 a.a. in a conserved region among species at the beginning of the protein. All predictors classified missense variants in *ABCA12*, *ALPI*, *GUCY1B3*, *KATNAL1*, *KCNC1*, and *MSH3* as pathogenic (Table S3.20B). All these variants except p.Y415C in *KATNAL1* decreased the protein stability slightly, as predicted by I-Mutant2.0 software (Table 3.10). The missense variant, p.L1487F in *CPAMD8*, is not located in a specific domain of the protein and appeasr to have slight impact on protein stability (Table 3.10). However, this gene has been recently associated with a recessive form of anterior segment dysgenesis (Cheong et al., 2016). In addition, *FRAS1*, which encodes an extracellular matrix protein, had the splicing variant c.1817A>G that probably disturbs the donor splice site after exon 16 (Table 3.10).

Table 3.10 – Patient 5 best candidate genes and respective variants following the dominant inheritance model. All variants are in the heterozygous state. Gene description was based on information from GeneCards and ExAC database, Exomiser and literature. LoF: loss of function. * denotes variants predicted as pathogenic by all functional impact predictors. NA: not applicable.

Gene	Variants	a.a. position	Average ddG	Gene and variant description
*ABCA12**	p.P1864S	1,864/ 2,595	-2.14	Encodes ATP-binding cassette sub-family A member 12, a probable transporter involved in lipid homeostasis.
*ALPI**	p.G452R	452/528	-1.16	Encodes intestinal-type alkaline phosphatase, a digestive brush-border enzyme that is a component of the gut mucosal defense system and is thought to function in the detoxification of lipopolysaccharide, and in the prevention of bacterial translocation in the gut.
CPAMD8	p.L1487F	1,487/ 1,932	-1.58	Encodes C3 and PZP-like alpha-2-macroglobulin domain-containing protein 8, a member of complement component-3/alpha-2-macroglobulin family that is involved in innate immunity and damage control. Associated with anterior segment dysgenesis.
CRB2	p.Q336X	336/778	NA	Encodes protein crumbs homolog 2, a component of the crumbs cell polarity complex that may play a role in polarized cells morphogenesis.
FRAS1	c.1817A>G	606/ 4,008	NA	Encodes extracellular matrix protein FRAS1 that appears to function in the regulation of epidermal-basement membrane adhesion and organogenesis during development. Exomiser: causative gene of Fraser syndrome associated with corneal opacity, visual impairment e abnormality of the palpebral fissures. Is associated with microphthalmia and cryptophthalmos in mouse.

Gene	Variants	a.a. position	Average ddG	Gene and variant description
*GUCY1B3**	p.D226Y	226/619	-0.69	Encodes guanylate cyclase soluble subunit beta-1, a subunit of the soluble guanylate cyclase that mediates responses to nitric oxide by catalyzing the biosynthesis of the signaling molecule cGMP. Exomiser: in human and mouse interact with *PDE6G*, associated with Retinitis Pigmentosa.
*KATNAL1**	p.Y415C	415/490	0.98	Encodes katanin p60 ATPase-containing subunit A-like 1 that regulates microtubule dynamics in Sertoli cells, a process essential for spermiogenesis and male fertility. Gene with LoF intolerance.
*KCNC1**	p.E580G	580/585	-1.52	Encodes potassium voltage-gated channel subfamily C member 1 that mediates the voltage-dependent potassium ion permeability of excitable membranes. Gene with LoF intolerance.
*MSH3**	p.E755Q	755/ 1,137	-2,.8	Encodes DNA mismatch repair protein Msh3, a component of the post-replicative DNA mismatch repair system.
PBXIP1	p.D151delins GSDDD	151/ 731	NA	Encodes pre-B-cell leukemia transcription factor-interacting protein 1 that interacts with the PBX1 homeodomain protein, inhibiting its transcriptional activation potential by preventing its binding to DNA. Can interact with estrogen receptors alpha and beta and promote the proliferation of breast cancer, brain tumors, and lung cancer.
TRPT1	p.R27X	27/253	NA	Encodes tRNA 2'-phosphotransferase 1, a phosphotransferase that catalyzes the last step of tRNA splicing, the transfer of the splice junction 2-phosphate from ligated tRNA to NAD to produce ADP-ribose 1-2 cyclic phosphate.

The identification of genes with novel and pathogenic variants strategy applied to sample 9 identified 21 genes. Tables S3.21 shows the detailed information for each gene and respective variants. Nonsense variants were found in *EXOC1* and *RNF14*, p.L118X and p.Y242X, respectively. The variant in *EXOC1* is located at the beginning of the protein, while the variant in *RNF14* occurs in the middle of E3 ubiquitin-protein ligase RNF14 leading to the loss of the region responsible for the interaction with the androgen receptor. Moreover, two frameshift variants found in *ADAMTS20* and *DCDC2* also result in truncated proteins. In *C6orf226* the p.X102R variant leads to the loss of codon stop with the gain of 56 a.a. in this patient. The missense p.L108W in *HSD3B2* was the only missense variant predicted as pathogenic by all predictors despite the low average ddG predicted the value of -0.42. According to Exomiser this gene interacts with *HMX1* that is associated with ocular phenotypes such as increased IOP, microcornea and sclerocornea. p.R751G variant was also predicted to have an impact in

HSPG2, the gene of perlecan (Table S3.21B). The variant is located in laminin EGF-like 1 repeat of the domain III and probably causes a slight decrease in protein stability (Table 3.11), as indicated by predictors. Perlecan is a proteoglycan that binds to and cross-links many extracellular matrix components and is expressed in TM.

Table 3.11 – Patient 9 best candidate genes and respective variants following the dominant inheritance model. All variants are in the heterozygous state. Gene description was based on information from GeneCards and ExAC database, Exomiser and literature. LoF: loss of function. * denotes variants predicted as pathogenic by all functional impact predictors. NA: not applicable.

Gene	Variants	a.a. position	Average ddG	Gene and variant description
ADAMTS20	p.L773QfsX23	773/ 1,906	NA	Encodes A disintegrin and metalloproteinase with thrombospondin motifs 20, a metallopeptidase that may play a role in the tissue remodeling process occurring in both normal and pathological conditions. May have a protease-independent function in the transport from the endoplasmic reticulum to the Golgi apparatus of secretory cargos.
C6orf226	p.X102R	102/102	NA	Encodes uncharacterized protein C6orf226, whose function is unknown.
DCDC2	p.G315EfsX32	315/476	NA	Encodes doublecortin domain-containing protein 2 that plays a role in the inhibition of canonical Wnt signaling pathway. May be involved in neuronal migration during development of the cerebral neocortex. Participating in the control of ciliogenesis and ciliary length.
EXOC1	p.L118X	118/894	NA	Encodes exocyst complex component 1, a component of the exocyst complex involved in the docking of exocytic vesicles with fusion sites on the plasma membrane.
*HSD3B2**	p.L108W	108/372	-0.42	Encodes 3 beta-hydroxysteroid dehydrogenase/delta 5-->4-isomerase type 2, a bifunctional enzyme that catalyzes the oxidative conversion of delta(5)-ene-3-beta-hydroxy steroid, and the oxidative conversion of ketosteroids. It plays a crucial role in the biosynthesis of all classes of hormonal steroids. Exomiser score: 0.754, in human interacts with *HMX1*, a causative gene of oculoauricular syndrome, associated with increased intraocular pressure, microcornea and sclerocornea. Interacts with *HMX1* leading to microphthalmia in the mouse.
HSPG2	p.R751G	751/ 4,391	-1.21	Encodes basement membrane-specific heparan sulfate proteoglycan core protein, a proteoglycan that binds to and cross-links many extracellular matrix components and cell-surface molecules. It has been shown that this protein interacts with laminin, prolargin, collagen type IV, FGFBP1, FBLN2, FGF7 and

Gene	Variants	a.a. position	Average ddG	Gene and variant description
				transthyretin, and others, and it plays essential roles in multiple biological activities. Expressed in TM. Exomiser score: 0.186, the causative gene of Schwartz-Jampel syndrome, associated with microcornea. Leads to microphthalmia in the mouse.
RNF14	p.Y242X	242/474	NA	Encodes E3 ubiquitin-protein ligase RNF14 that might act as an E3 ubiquitin-protein ligase that accepts ubiquitin from specific E2 ubiquitin-conjugating enzymes and then transfers it to substrates, which could be nuclear proteins. Could play a role as a coactivator for androgen- and, to a lesser extent, progesterone-dependent transcription.

Patient 18 had 16 genes with novel and probably pathogenic variants (Tables S3.22). However, only three variants lead to the creation of truncated proteins: p.Q1691X, p.K1619NfsX2 and p.S210CfsX15 in *CEP295*, *POLQ*, and *SYNJ2*, respectively. These variants cause the loss of important domains in the proteins, especially p.S210CfsX15 leading to the loss of SAC and RRM domains and the catalytic region. *KRTAP5-1* had a delection causing the loss of 6 a.a. in the beginning of the protein. *TNXB*, the gene that encodes tenascin-X, harbored a missense variant (Table 3.12). Tenascin-X is an extracellular matrix glycoprotein from the same family of tenascin-C, which had two heterozygous variants in patient 41. According to most predictors p.S3825L is pathogenic (Table S3.22B), is located in the fibronectin type-III domain 28, and according to the I-Mutant2.0 software prevision contributes to increased protein stability (Table 3.12).

Table 3.12 – Patient 18 best candidate genes and respective variants following the dominant inheritance model. All variants are in the heterozygous state. Gene description was based on information from GeneCards and ExAC database, Exomiser and literature. LoF: loss of function. NA: not applicable.

Gene	Variants	a.a. position	Average ddG	Gene and variant description
CEP295	p.Q1691X	1,691/ 2,601	NA	Encodes centrosomal protein of 295 kDa that mediates centriole-to-centrosome conversion at late mitosis, but is dispensable for cartwheel removal or centriole disengagement.
KRTAP5-1	p.V44_C49del	44/278	NA	Encodes Keratin-associated protein 5-1, a component of interfilamentous matrix that are essential for the formation of a rigid and resistant hair shaft through their extensive disulfide bond cross-linking with abundant cysteine residues of hair keratins. The matrix proteins include the high-sulfur and high-glycine-tyrosine keratins.

Gene	Variants	a.a. position	Average ddG	Gene and variant description
POLQ	p.K1619NfsX2	1,619/ 2,590	NA	Encodes DNA polymerase theta, a DNA polymerase that promotes microhomology-mediated end-joining, an alternative non-homologous end-joining machinery triggered in response to double-strand breaks in DNA.
SYNJ2	p.S210CfsX15	210/ 1,496	NA	Encodes synaptojanin-2, an inositol 5-phosphatase that may be involved in distinct membrane trafficking and signal transduction pathways. May mediate the inhibitory effect of Rac1 on endocytosis. Exomiser score: 0.656, interacts with *OCRL*, a causative gene of Lowe syndrome, associated with buphthalmos, opacification of the corneal stroma, and glaucoma.
TNXB	p.S3825L	3,825/ 4,242	1.02	Encodes tenascin-X, an extracellular matrix glycoprotein that has anti-adhesive effects, as opposed to fibronectin which is adhesive. Functions in matrix maturation during wound healing. Expressed in rat TM. Gene with LoF intolerance. Exomiser score: 0.753, the causative gene of Ehlers-Danlos syndrome. Interacts with *DCN*, a causative gene of congenital stromal corneal dystrophy, associated with glaucoma, increased corneal thickness, progressive visual loss and congenital corneal dystrophy.

After the selection of novel and predicted pathogenic variants patient 25 had 22 altered genes (Tables S3.23). In this set of genes, four frameshift variants were found in *C3orf30*, *GBP6*, *GRIK4*, and *SPDL1* (Table 3.13), with the p.L142RfsX4 in *GBP6* leading to the loss of the majority of the protein, including part of the GTPase domain. All predictors indicated two missense variants as pathogenic: p.D848L and p.P267A in *ATP13A4* and *GLDC*, both encoding proteins with different functions (Table S3.23B). From the list of 22 genes, the missense p.G1362C located in laminin EGF-like domain 11 of *LAMA3* and the missense p.R596T in *TRPM3* were analyzed in more detail due to its presence in TM and association with glaucoma, respectively. *LAMA3* encodes the subunit α3 of the laminin, a constituent of the basement membrane of the cells, including TM cells (Acott and Kelley, 2008). Despite its predicted pathogenicity (only FATHMM software predicted it as tolerated (Table S3.23B), the variant does not influence the protein stability. Meanwhile, the *TRPM3* p.R596T variant is located before the transmembrane domains (Grimm et al., 2003) and appears to decrease protein stability slightly (Table 3.13). This gene was recently associated with inherited cataract and glaucoma (Bennett et al., 2014).

Table 3.13 – Patient 25 best candidate genes and respective variants following the dominant inheritance model. All variants are in the heterozygous state, except *GRIK4* p.H860QfsX188. Gene description was based on information

from GeneCards and ExAC database, Exomiser and literature. LoF: loss of function. * denotes variants predicted as pathogenic by all functional impact predictors. NA: not applicable.

Gene	Variants	a.a. position	Average ddG	Gene and variant description
*ATP13A4**	p.D848L	848/ 1,196	-0.95	Encodes probable cation-transporting ATPase 13A4, a transmembrane cation-transporting P-type ATPase.
C3orf30	p.Q391fsX11	391/536	NA	Encodes the uncharacterized protein C3orf30 that may play an important role in spermatogenesis.
GBP6	p.L142RfsX4	142/633	NA	Encodes guanylate-binding protein 6 that hydrolyzes GTP to both GDP and GMP.
*GLDC**	p.P267A	267/ 1,020	-2.40	Encodes glycine dehydrogenase (decarboxylating), mitochondrial, a member of glycine cleavage system that catalyzes the degradation of glycine. Binds to the alpha-amino group of glycine through its pyridoxal phosphate cofactor.
GRIK4	p.H860QfsX188	860/946	NA	Encodes glutamate receptor ionotropic, kainate 4. Gene with LoF intolerance.
LAMA3	p.G1362C	1,362/ 3,333	0.02	Encodes laminin subunit alpha-3 responsive to several epithelial-mesenchymal regulators including keratinocyte growth factor, epidermal growth factor and insulin-like growth factor. Expressed in the TM.
SPDL1	p.M524CfsX3	524/605	NA	Encodes protein spindly, required for the localization of dynein and dynactin to the mitotic kinetochore. Also required for correct spindle orientation. Does not appear to be necessary for the removal of spindle assembly checkpoint proteins from the kinetochore upon bipolar spindle attachment.
TRPM3	p.R596T	596/ 1,732	-1.28	Encodes transient receptor potential cation channel subfamily M member 3, a calcium channel mediating constitutive calcium ion entry. Associated with inherited cataract and glaucoma.

A total of 29 variants present in 28 genes remained after the application of the previously described parameters to patient 31 (Tables S3.24). Two nonsense and two frameshift variants in *MTMR2*, *PRPH*, *TRIOBP*, and *WRAP73* originated truncated proteins, leading to the loss of some functional domains (Table 3.14). The proteins encoded by these genes have non-related functions. The deletion of nine nucleotides in *PTAFR*, at the end of the coding sequence, does not affect any known domain. All software predicted the two missense variants in *MFGE8* and *ZNF74* as pathogenic, but only p.F182L in *MFGE8* had more probability to reduce protein stability (average ddG of -2.06). To date, none of these genes is known to be expressed in TM. However, *ECM2* was expressed in HTM cell cultures and is a component of the extracellular matrix. Patient 31 also had a missense variant that, although not predicted to reduce protein stability very strongly, is pathogenic (except for FATHMM) (Table S3.24B).

Table 3.14 – Patient 31 best candidate genes and respective variants following the dominant inheritance model. All variants are in the heterozygous state. Gene description was based on information from GeneCards and ExAC database, Exomiser and literature. LoF: loss of function. * denotes variants predicted as pathogenic by all functional impact predictors. NA: not applicable.

Gene	Variants	a.a. position	Average ddG	Gene and variant description
ECM2	p.N377H	377/699	-0.69	Encodes extracellular matrix protein 2, a matrix glycoprotein that promotes matrix assembly and cell adhesiveness. Expressed in HTM cell culture.
*MFGE8**	p.F182L	182/387	-2.06	Preproprotein that is proteolytically processed to form multiple protein products. The major encoded protein product, lactadherin, is a membrane glycoprotein that promotes phagocytosis of apoptotic cells. This protein has also been implicated in wound healing, autoimmune disease, and cancer. Lactadherin can be further processed to form a smaller cleavage product, medin, which comprises the major protein component of aortic medial amyloid.
MTMR2	p.R364X	364/643	NA	Encodes myotubularin-related protein 2, a phosphoinositide lipid phosphatase that possesses phosphatase activity towards phosphatidylinositol-3-phosphate and phosphatidylinositol-3,5-bisphosphate.
PRPH	p.E150X	150/346	NA	Encodes peripherin-2, a type III intermediate filament cytoskeletal protein found in neurons of the peripheral nervous system.
PTAFR	p.T324_T326 del	324/342	NA	Encodes platelet-activating factor receptor, a chemotactic phospholipid mediator that possesses potent inflammatory, smooth-muscle contractile and hypotensive activity. Seems to mediate its action via a G-protein that activates a phosphatidylinositol-calcium second messenger system.
TRIOBP	p.N890KfsX12	890/ 2,365	NA	Encodes TRIO and F-actin-binding protein that may regulate actin cytoskeletal organization, cell spreading and cell contraction by directly binding and stabilizing filamentous F-actin. The localized formation of TARA and TRIO complexes coordinates the amount of F-actin present in stress fibers. May also serve as a linker protein to recruit proteins required for F-actin formation and turnover.
WRAP73	p.A187PfsX67	187/460	NA	Encodes WD repeat-containing protein WRAP73, a member of the WD repeat protein family that regulates the spindle anchoring at the mitotic centrosome. Involved in ciliary vesicle formation at the mother centriole and required for the docking of vesicles to the basal body during ciliogenesis.
*ZNF74**	p.C390R	390/644	-0.49	Encodes zinc finger protein 74 that may play a role in RNA metabolism.

Similar to the recessive model, patient 32 obtained an inferior number of novel and predicted pathogenic variants (Tables S3.25), with variants detected in 11 genes. Two variants were considered more relevant: p.G318EfsX36 and p.S204X in *CYP27B1* and *OTX1*, respectively. The two variants lead to the loss of almost half of the proteins (Table 3.15). *OTX1* participates in the development of sense organs, being expressed in the ciliary body and anterior retina (Larsen et al., 2009).

Table 3.15 – Patient 32 best candidate genes and respective variants following the dominant inheritance model. All variants are in the heterozygous state. Gene description was based on information from GeneCards and ExAC database, Exomiser and literature. LoF: loss of function. NA: not applicable.

Gene	Variants	a.a. position	Average ddG	Gene and variant description
CYP27B1	p.G318EfsX36	318/508	NA	Encodes 25-hydroxyvitamin D-1 alpha hydroxylase, mitochondrial, a monooxygenase localized to the inner mitochondrial membrane where it hydroxylates 25-hydroxyvitamin D3 at the 1alpha position. This enzyme regulates the level of biologically active vitamin D and plays an important role in calcium homeostasis.
OTX1	p.S204X	204/354	NA	Encodes homeobox protein OTX1, a transcription factor that probably plays a role in the development of the brain and the sense organs. Gene with LoF intolerance.

Patient 33 revealed 19 genes with novel and possible pathogenic variants (Tables S3.26). Three of these variants were frameshift variants in *DCDC5*, *GGA1*, and *WNK4*. The variant in *GGA1* occurred at the end of the protein, but the variants in *DCDC5* and *WNK4* were at the beginning of the coding region leading to the loss of essential domains (Table 3.16). Deletion and duplication correspondent to 2 a.a. occurred in *DPEP2* and *RREB1*, respectively. *GGA1* and *RREB1* are genes reported with no LoF variant in ExAC database. All functional impact predictors classified the missense variants p.G172R and p.G661D in *ARSE* and *MUSK* as pathogenic (Table S3.26B). p.G661D is located in the protein kinase domain and may cause a slight decrease protein stability (average ddG of -1.15).

Table 3.16 – Patient 33 best candidate genes and respective variants following the dominant inheritance model. All variants are in the heterozygous state, except *ARSE* p.G172R. Gene description was based on information from GeneCards and ExAC database, Exomiser and literature. LoF: loss of function. * denotes variants predicted as pathogenic by all functional impact predictors. NA: not applicable.

Gene	Variants	a.a. position	Average ddG	Gene and variant description
*ARSE**	p.G172R	172/589	0.07	Encodes arylsulfatase E, a sulfatase that may be essential for the correct composition of cartilage and bone matrix during development. Has no activity

				toward steroid sulfates.
DCDC5	p.L20FfsX31	20/648	NA	Encodes doublecortin domain-containing protein 5, a member of the doublecortin family that may bind tubulin and enhance microtubule polymerization.
DPEP2	p.V462_S463 del	462/486	NA	Encodes dipeptidase 2, a membrane-bound dipeptidase that hydrolyzes leukotriene D4 into leukotriene E4.
GGA1	p.K428RfsX88	428/639	NA	Encodes Golgi-associated, gamma adaptin ear containing, ARF binding protein 1 that plays a role in protein sorting and trafficking between the trans-Golgi network and endosomes. Mediates the ARF-dependent recruitment of clathrin to the trans-Golgi network and binds ubiquitinated proteins and membrane cargo molecules with a cytosolic acidic cluster-dileucine motif. Gene with LoF intolerance.
*MUSK**	p.G661D	661/869	-1.15	Encodes muscle, skeletal receptor tyrosine-protein kinase that may play a role in clustering of the acetylcholine receptor in the postsynaptic neuromuscular junction.
RREB1	p.E1346_ G1377dup	1,346/ 1,687	NA	Encodes Ras-responsive element-binding protein 1, a transcription factor that binds to RAS-responsive elements of gene promoters. May be involved in Ras/Raf-mediated cell differentiation by enhancing calcitonin expression. Represses the angiotensinogen gene. Negatively regulates the transcriptional activity of AR. Potentiates the transcriptional activity of NEUROD1. Gene with LoF intolerance.
WNK4	p.E107RfsX38	107/ 1,243	NA	Encodes serine/threonine-protein kinase WNK4 that plays an important role in the regulation of electrolyte homeostasis, cell signaling, survival and proliferation.

Performing the same variant selection, patient 36 harbored a set of 18 variants and corresponding genes (Tables S3.27). Five novel variants lead to a premature stop codon in *KIAA1324*, *LRBA*, *MUC12*, *SLC41A3*, and *UTP3* causing the loss of functional domains (Table 3.17). In addition, all predictor software classified the missense variants p.G960R, p.D1020V, and p.Y293D in *CHD9*, *EP400*, and *SLC25A38* as pathogenic (Table S3.27B). None of these genes has data regarding TM expression.

Table 3.17 – Patient 36 best candidate genes and respective variants following the dominant inheritance pattern. All variants are in the heterozygous state. Gene description was based on information from GeneCards and ExAC database, Exomiser and literature. LoF: loss of function. * denotes variants predicted as pathogenic by all functional impact predictors. NA: not applicable.

Gene	Variants	a.a. position	Average ddG	Gene and variant description
*CHD9**	c.G2878A;	960/	0.03	Encodes chromodomain-helicase-DNA-binding

Gene	Variants	a.a. position	Average ddG	Gene and variant description
	p.G960R	2,897		protein 9, a transcriptional coactivator for PPARA and possibly other nuclear receptors. Proposed to be a ATP-dependent chromatin remodeling protein. Has DNA-dependent ATPase activity and binds to A/T-rich DNA. Associates with A/T-rich regulatory regions in promoters of genes that participate in the differentiation of progenitors during osteogenesis. Gene with LoF intolerance. Exomiser score: 0.755, interacts with *YAP1*, which is associated with posterior embryotoxon, glaucoma, and opacification of the corneal stroma.
*EP400**	p.D1020V	1,020/ 3,159	-1.41	Encodes E1A-binding protein p400, a component of the NuA4 histone acetyltransferase complex which is involved in transcriptional activation of select genes principally by acetylation of nucleosomal histones H4 and H2A. Gene with LoF intolerance.
KIAA1324	p.G163AfsX15	163/ 1,013	NA	Encodes UPF0577 protein KIAA1324, a transmembrane protein that may protect cells from cell death by inducing cytosolic vacuolization and upregulating the autophagy pathway.
LRBA	p.T2003DfsX4	2,003/ 2,863	NA	Encodes lipopolysaccharide-responsive and beige-like anchor protein that may be involved in coupling signal transduction and vesicle trafficking to enable polarized secretion and/or membrane deposition of immune effector molecules. Gene with LoF intolerance.
MUC12	p.T698CfsX82	698/ 5,335	NA	Encodes mucin-12, a mucin involved in epithelial cell protection, adhesion modulation, and signaling. May be involved in epithelial cell growth regulation. Stimulated by both cytokine TNF-alpha and TGF-beta in the intestinal epithelium.
*SLC25A38**	p.Y293D	293/304	-0.63	Encodes solute carrier family 25 member 38, a mitochondrial carrier required during erythropoiesis. Probably involved in the biosynthesis of heme, possibly by facilitating 5-aminolevulinate production. May act by importing glycine into mitochondria or by exchanging glycine for ALA across the mitochondrial inner membrane. Gene with LoF intolerance.
SLC41A3	p.Q399X	399/507	NA	Encodes solute carrier family 41 member 3, whose function is unknown.
UTP3	p.R8LfsX30	8/479	NA	Encodes the small subunit processome component Homolog that is essential for gene silencing: has a role in the structure of silenced chromatin. Plays a role in the developing brain.

Patient 41 revealed a superior number of novel and predicted as pathogenic variants. More detailed information about the 30 variants is described in Tables S3.28. This patient had two frameshift variants: p.I1772FfsX15 and p.E589DfsX17 in *HYDIN* and *ZNF493*. Three novel missense variants predicted to be pathogenic by all software in *KMT2C*, *NSMCE1*, and *TMC1* (Table S3.28B). However, only p.R365K may decrease protein stability (average ddG of -2.13) (Table 3.18). No information is available about the expression of the thirty genes in the TM.

Table 3.18 – Patient 41 best candidate genes and respective variants following the dominant inheritance model. All variants are in the heterozygous state. Gene description was based on information from GeneCards and ExAC database, Exomiser and literature. LoF: loss of function. * denotes variants predicted as pathogenic by all functional impact predictors. NA: not applicable.

Gene	Variants	a.a. position	Average ddG	Gene and variant description
HYDIN	p.I1772FfsX15	1,772/ 5,121	NA	Encodes hydrocephalus-inducing protein homolog that is required for ciliary motility.
*KMT2C**	p.P3635L	3,635/ 4,911	-0.52	Encodes histone-lysine N-methyltransferase 2C that methylates Lys-4 of histone H3. H3 Lys-4 methylation represents a specific tag for epigenetic transcriptional activation. A central component of the MLL2/3 complex, a coactivator complex of nuclear receptors, involved in transcriptional coactivation. KMT2C/MLL3 may be a catalytic subunit of this complex. Gene with LoF intolerance. Exomiser score: 0.757, interacts with *PLOD1* a causative of Ehlers-Danlos syndrome, which leads to microcornea, glaucoma, and epicanthus.
*NSMCE1**	p.C231R	231/266	-0.66	Encodes non-structural maintenance of chromosomes element 1 homolog, a component of the SMC5-SMC6 complex involved in DNA double-strand breaks by homologous recombination. The complex is required for telomere maintenance. Is involved in positive regulation of response to DNA damage stimulus.
*TMC1**	p.R365K	365/760	-2.13	Encodes transmembrane channel-like protein 1, a probable ion channel required for the normal function of cochlear hair cells.
ZNF493	p.E589DfsX17	589/646	NA	Encodes zinc finger protein 493 that may be involved in transcriptional regulation.

Patient 42 had 19 genes with novel and possible pathogenic variants. These genes and variants are detailed in Tables S3.29. Three nonsense variants in *AKAP13*, *BBX*, and *SLC17A5* originate truncated proteins with a significant loss of a.a., especially for *AKAP13* and *BBX* (Table 3.19). *CCDC105* presented a frameshift variant at the end of the protein. *CCDC141*, *DCTPP1*, and *PRM3* had one a.a. deletion each. The encoded proteins have no described protein domains at the deletion position yet. All functional impact predictors classified the missense p.L735P in

VPS51 as pathogenic (Table S3.29B). This variant is located at the end of the protein. This patient also presented p.D282N in *SEC31A* that is located in the SEC13 interaction domain, with possible disruption of the domain, and had an elevated CADD score of 36.

Table 3.19 – Patient 42 best candidate genes and respective variants following the dominant inheritance model. All variants are in the heterozygous state. Gene description was based on information from GeneCards and ExAC database, Exomiser and literature. LoF: loss of function. * denotes variants predicted as pathogenic by all functional impact predictors. NA: not applicable.

Gene	Variants	a.a. position	Average ddG	Gene and variant description
AKAP13	p.R1483X	1,483/ 2,813	NA	Encodes A-kinase anchor protein 13 that anchors PKA and acts as an adapter protein to selectively couple G alpha-13 and Rho. Activates estrogen receptor beta by a p38 MAPK-dependent pathway. Stimulates exchange activity on Rho proteins *in vitro*, but not on CDC42, Ras or Rac and may bind calcium ions. Gene with LoF intolerance.
BBX	p.K344X	344/941	NA	Encodes HMG box transcription factor BBX, a component of one form of the SWI/SNF chromatin remodeling complex. Acts both as a coactivator and as a corepressor. May play a role in chromatin remodeling. Gene with LoF intolerance.
CCDC105	p.E475GfsX6	475/499	NA	Encodes coiled-coil domain-containing protein 105, whose function is unknown.
CCDC141	p.E546del	546/ 1,450	NA	Encodes coiled-coil domain-containing protein 141 that plays a critical role in radial migration and centrosomal function.
DCTPP1	p.L58del	58/170	NA	Encodes dCTP pyrophosphatase 1, a pyrophosphatase that converts dCTP to dCMP and inorganic pyrophosphate. Displays weak activity against dTTP and dATP, but none against dGTP. This protein may be responsible for eliminating excess dCTP after DNA synthesis and may prevent overmethylation of CpG islands. Gene with LoF intolerance.
PRM3	p.S45del	45/103	NA	Encodes protamine-3 that substitutes for histones in the chromatin of sperm during the haploid phase of spermatogenesis. Compacts sperm DNA into a highly condensed, stable and inactive complex.
SEC31A	p.D282N	282/ 1,220	-0.41	Encodes protein transport protein Sec31A, a component of the coat protein complex II that promotes the formation of transport vesicles from the endoplasmic reticulum. The coat has two main functions, the physical deformation of the endoplasmic reticulum membrane into vesicles and the selection of cargo molecules.
SLC17A5	p.W485X	485/495	NA	Encodes sialin, a membrane transporter that exports glucuronic acid and free sialic acid out of the lysosome after it is cleaved from sialoglycoconjugates

Gene	Variants	a.a. position	Average ddG	Gene and variant description
				undergoing degradation, this is required for normal CNS myelination. Mediates aspartate and glutamate membrane potential-dependent uptake into synaptic vesicles and synaptic-like microvesicles. Also functions as an electrogenic 2NO(3)(-)/H(+) cotransporter in the plasma membrane of salivary gland acinar cells, mediating the physiological nitrate efflux, 25% of the circulating nitrate ions are typically removed and secreted in saliva.
*VPS51**	p.L735P	735/782	-0.93	Encodes vacuolar protein sorting-associated protein 51 homolog, a component of the GARP complex that participates in the retrograde transport of acid hydrolase receptors, likely by promoting tethering and SNARE-dependent fusion of endosome-derived carriers to the trans-Golgi network. Acts as component of the EARP complex that is involved in endocytic recycling.

Continuing with the same strategy for patient 43 eighteen genes were selected (Tables S3.30). p.R317X p.W194X and p.Y345X in *ANO4, FRG2B* and *KCTD20* originated premature stop codons (Table 3.20). *ANO4* transcript was previously found in HTM cell cultures (Banerjee et al., 2012). *C1orf95* presented an deletion leading to the loss of eight a.a. at the beginning of the protein encoded, while *RASD1* had a duplication leading to the loss of two a.a. at the end of the encoded protein (Table 3.20). Two missense variants in *GID4* and *SLC4A3* were pathogenic according to all the predictors (Table S3.30B). The missense variant in *GID4* may cause a slight decrease in protein stability (average ddG of -1.33). The missense variant in *COL14A1*, p.D1080G, is located in the VWFA2 domain. While not pathogenic according to SIFT, it encodes the collagen alpha-1(XIV) chain, a known constituent of the TM extracellular matrix (Acott and Kelley, 2008).

Table 3.20 – Patient 43 best candidate genes and respective variants following the dominant inheritance model. All variants are in the heterozygous state. Gene description was based on information from GeneCards and ExAC database, Exomiser and literature. LoF: loss of function. * denotes variants predicted as pathogenic by all functional impact predictors. NA: not applicable.

Gene	Variants	a.a. position	Average ddG	Gene and variant description
ANO4	p.R317X	317/955	NA	Encodes anoctamin-4, a transmembrane protein that has calcium-dependent phospholipid scramblase activity; scrambles phosphatidylserine, phosphatidylcholine and galactosylceramide. Does not exhibit calcium-activated chloride channel activity. Expressed in TM cell culture. Gene with LoF

Gene	Variants	a.a. position	Average ddG	Gene and variant description
				intolerance.
C1orf95	p.A11_A18del	11/141	NA	Encodes protein stum homolog whose function is unknown.
COL14A1	p.D1080G	1,080/ 1,796	-0.90	Encodes collagen alpha-1(XIV) chain that plays an adhesive role by integrating collagen bundles. It is probably associated with the surface of interstitial collagen fibrils via COL1. The COL2 domain may then serve as a rigid arm which sticks out from the fibril and protrudes the large N-terminal globular domain into the extracellular space, where it might interact with other matrix molecules or cell surface receptors. Expressed in TM.
FRG2B	p.W194X	194/278	NA	Encodes protein FRG2-like-1 which function is unknown.
*GID4**	p.G256R	256/300	-1.33	Encodes glucose-induced degradation protein 4 homolog, a subunit of the Mediator complex that is a coactivator required for activation of RNA polymerase II transcription by DNA bound transcription factors. Gene with LoF intolerance.
KCTD20	p.Y345X	345/419	NA	Encodes BTB/POZ domain-containing protein KCTD20 that promotes the phosphorylation of AKT family members.
RASD1	p.S232_G233 dup	233/281	NA	Encodes dexamethasone-induced Ras-related protein 1, an activator of G-protein signaling and acts as a direct nucleotide exchange factor for Gi-Go proteins. This protein interacts with the neuronal nitric oxide adaptor protein CAPON, and a nuclear adaptor protein FE65, which interacts with the Alzheimer's disease amyloid precursor protein. This gene may play a role in dexamethasone-induced alterations in cell morphology, growth and cell-extracellular matrix interactions.
*SLC4A3**	p.V1078M	1,078/ 1,232	-0.84	Encodes anion exchange protein 3, a plasma membrane anion exchange protein of wide distribution. Mediates at least a part of the Cl(-) /HCO3(-) exchange in cardiac myocytes. Gene with LoF intolerance.

Lastly, patient 45 had 22 genes with novel and possible pathogenic variants. These genes and variants are detailed in Tables S3.31. Patient 45 had two nonsense variants in *LMBR1L* and *LTBP4* that lead to the loss of important protein domains and three frameshift variants (Table 3.21). The frameshift in *CNTNAP4* occurs at the beginning of the protein causing the loss of all important domains. The other two frameshifts in *PTPN7* and *ZKSCAN8* occur near the end of the protein (Table 3.21). Patient 45 also presented a deletion of 9 a.a. in *PKD1.* This deletion

occurs at the end of the protein, in a region with no known domains. p.G529E present in *ABCC4* is pathogenic according to all predictors (Table S3.31B) and is located in ABC transporter domain 1. Expression of this gene was previously reported in TM tissue (Pattabiraman et al., 2013).

Table 3.21 – Patient 45 best candidate genes and respective variants following the dominant inheritance model. All variants are in the heterozygous state. Gene description was based on information from GeneCards and ExAC database, Exomiser and literature. LoF: loss of function. * denotes variants predicted as pathogenic by all functional impact predictors. NA: not applicable.

Gene	Variants	a.a. position	Average ddG	Gene and variant description
*ABCC4**	p.G529E	529/ 1,325	-0.11	Encodes multidrug resistance-associated protein 4, a transporter that may be an organic anion pump relevant to cellular detoxification. Expressed in TM.
CNTNAP4	p.E24SfsX13	24/ 1,308	NA	Encodes contactin-associated protein-like 4, a presynaptic protein involved in both dopaminergic synaptic transmission and GABAergic system, thereby participating in the structural maturation of inhibitory interneuron synapses. Involved in the dopaminergic synaptic transmission by attenuating dopamine release through a presynaptic mechanism. Also participates in the GABAergic system.
LMBR1L	p.E49X	49/489	NA	Encodes protein LMBR1L, a probable LCN1 receptor that may mediate LCN1 endocytosis.
LTBP4	p.C898X	898/ 1,624	NA	Encodes latent-transforming growth factor beta-binding protein 4, an extracellular matrix protein that may be involved in the assembly, secretion and targeting of TGFB1 to sites at which it is stored and/or activated. May play critical roles in controlling and directing the activity of TGFB1. May have a structural role in the extracellular matrix. Gene with LoF intolerance.
PKD1	p.C3370_ S3378del	3,370/ 4,303	NA	Encodes polycystin-1, a transmembrane glycoprotein that is involved in renal tubulogenesis and involved in fluid-flow mechanosensation by the primary cilium in renal epithelium. May be an ion-channel regulator. Involved in adhesive protein-protein and protein-carbohydrate interactions. Gene with LoF intolerance.
PTPN7	p.I326MfsX2	326/399	NA	Encodes tyrosine-protein phosphatase non-receptor type 7, a tyrosine phosphatase that acts preferentially on tyrosine-phosphorylated MAPK1. Plays a role in the regulation of T and B-lymphocyte development and signal transduction.
ZKSCAN8	p.I478NfsX17	478/578	NA	Encodes zinc finger protein with KRAB and SCAN domains 8 that may be involved in transcriptional regulation.

3.3.5.Genes with *de novo* variants

Trio exome sequencing in families 1 and 31 allowed for the identification of *de novo* pathogenic variants. However, neither of the patients harbored these type of mutations. Variant search also included variants not in accordance with the genotype of parents, without success.

Exomes of patient 31 and his parents were sequenced with different technologies. As such there were a few variants in the patient without parents' genotype information. These variants were benign, according to MAF and pathogenicity prevision (Table S3.32), except the missense heterozygous p.R189C variant in *CNNM3.* This novel variant is pathogenic according to most functional impact predictors except LRT, which classified as unknown, and CADD with a score of 14.38. However, variant study in parents through Sanger sequencing identified the same variant in the mother, also in the heterozygous state (Figure 3.11).

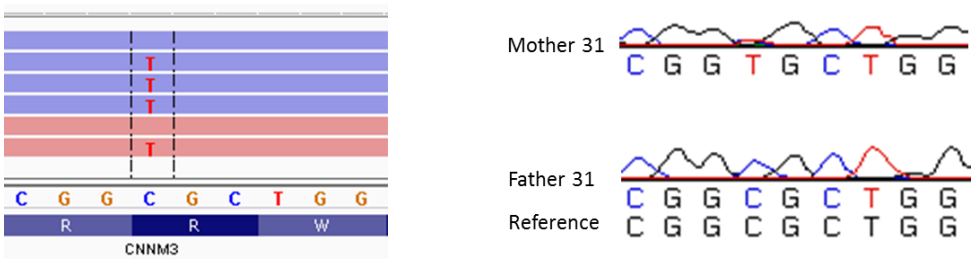


Figure 3.11 – Identification of *CNNM3* p.R189C. The variant was present in the heterozygous state in patient 31 and inherited from the mother, which is also heterozygous fot the variant.

3.4. Discussion

The advent of NGS technology contributed to a rapid evolution in several areas, including in the clinical genetic field. This technology proved to be able to identify new mutated genes in rare and complex diseases. Due to its success in samples without diagnosis by conventional genetic tests, NGS has begun to be implemented in clinical diagnosis, revolutionizing medical practice (Seaby et al., 2016). WES is now cost-effective to analyze nearly all variants present in the coding region of the genome. Variants in the coding region constitute 85% of all Mendelian disease mutations (Gilissen et al., 2012). In addition to the reduced cost, WES also allows a better coverage and interpretability of the variants than genome sequencing. The diagnosis rate using WES ranges from 21 to 40% for the rare disease of unknown etiology (Sawyer et al., 2016; Seaby et al., 2016).

The challenge of WES consists in the identification of the disease-causing variants among all variants distinguishing the rare and true pathogenic variants from rare and specific population polymorphisms. In this study, an average of 52,095 initial variants and 22,700 coding variants

were found in the sequenced samples. These values are similar to other studies (Bamshad et al., 2011; Frebourg, 2014), including in a study performed with 267 Spanish individuals where the authors found near 19,000 single nucleotide variants (Dopazo et al., 2016). Patient 41 revealed a higher initial number of variants (62,070 variants). Patient 41 and his parents were born in Portugal, however both parents have African and Asiatic descent. A study compared the number of coding variants present in African Americans, and European Americans and almost 3,700 variants more were found in African Americans (around 24,000 coding variants in African Americans vs. near 20,300 variants in European Americans) (Bamshad et al., 2011). Furthermore, each population has different local polymorphisms and rare variants (Dopazo et al., 2016). The African descent and private variants from European, African and Asiatic populations justify the high number of variants found in patient 41. No corresponding studies are available for the Asiatic population. Oppositely, patient 32 had the lowest number of variants after prioritization. Consanguinity could explain this low number, but no information about consanguinity is reported for this patient. Nevertheless, both parents may belong to a small group sharing private community variants. More information about this patient and her parents could help explaining this low number of variants. The number of coding variants for samples sequenced with Illumina technology was inferior to the other samples. This difference may be associated with a lower error rate for Indels (Boland et al., 2013).

3.4.1.Variants in the PCG candidate genes

Whole exome sequencing has been successfully applied in the identification of the genetic cause behind several ophthalmological disorders, such as age-related macular degeneration, retinitis pigmentosa, cataracts or myopia, among others (Nishiguchi et al., 2013; Reis et al., 2013; Yu et al., 2014; Li et al., 2016). Till date, in glaucoma, only a few new plausible candidate genes have been identified by WES. *EFEMP1* was found mutated in a family with POAG (Mackay et al., 2015) and *TRPM3* and *COL1A1* had mutations in patients with congenital glaucoma but associated with cataract and retinal detachment (Bennett et al., 2014; Mauri et al., 2016). The only WES-discovered gene associated with PCG is *TEK*, which encodes the Tie2 protein, a receptor tyrosine kinase.

Interestingly, in this study two different heterozygous *TEK* variants were identified, one in patient 4 and another in patient 36. Angiopoietin-1 (Ang1) and angiopoietin-2 (Ang2) bind to the Tie2 receptor. While Ang1 always acts as an agonist, Ang2 may act as an agonist or as an antagonist depending on the context. The mediated-ligand activation of Tie2 leads to the activation of different signaling pathways. Tie2 is constitutively expressed in vascular and

lymphatic endothelial cells and is involved with Ang1 in developmental angiogenesis and vascular development, stabilization and remodeling (Augustin et al., 2009; Fukuhara et al., 2010). In the eye, Tie2 is present in the SC endothelial cells (Kizhatil et al., 2014).

A study of Thomson and collaborators demonstrated the important role of Ang1, Ang2 and Tie2 in SC formation in mice. Animals with deletion of both ligands revealed an absence of SC and lymphatic capillaries of the corneal limbus and a disturbed patterning of vascular capillaries in the corneal limbus at 16.5 weeks. Knockout mice for the ligand or the receptor developed a phenotype similar to glaucoma patients including increased intraocular pressure, buphthalmos, corneal enlargement and increased anterior chamber depth. These results demonstrate the participation of Tie2 receptor in the molecular basis of glaucoma (Thomson et al., 2014). Very recently Souma and collaborators confirmed the involvement of *TEK* in the pathogenesis of glaucoma, more commonly in PCG. They found rare *TEK* variants in ten families that revealed to be loss-of-function mutations through multiple cellular mechanisms. The authors also proved that *TEK* haploinsufficiency was sufficient to develop hypoplastic SC and TM and elevated IOP. These results were consistent with an autosomal dominant pattern with variable expressivity for *TEK* mutations (Souma et al., 2016).

The *TEK* p.V188G found in patient 4 has been predicted as pathogenic. In fact, this variant is located in the Ig2 domain in the extracellular portion of Tie2 and, interestingly, all the previously *TEK* mutations identified in PCG patients are also located in the extracellular domains. In contrast, mutations related with venous malformations are found in the intracellular domains. Ig2 domain is extremely important for ligand binding (Barton et al., 2006; Yu et al., 2013) and also for stable protein expression (Souma et al., 2016). Because the V188 residue does not directly interact with Ang1, the level of possible disruption of the ligand binding is difficult to predict without experimental studies. Protein structure information indicates that this variant leads to structural changes that decrease stability, which is usually verified in disease-causing mutations. This may indirectly affect ligand binding or even Tie2 dimerization required for activation of the signal transduction pathway. Since ligand binding may not be totally abolished, a quantitative assay could help understand the real impact of the p.V188G variant. Another possibility may be a disruption on Tie2 translocation to the cell junctions after Ang1 induction (Fukuhara et al., 2008). The absence of extra *TEK* mutations in patient 4 is in agreement with the proposed dominant inheritance model.

Curiously, in patient 4, an *MYOC* variant was also found. This variant was previously associated with POAG (Alward et al., 1998; Fingert et al., 1999; Faucher et al., 2002; Lopez-Martinez et al., 2007; Souzeau et al., 2013), pigmentary glaucoma (Faucher et al., 2002), normal tension glaucoma (Weisschuh et al., 2005) and in one individual with ocular hypertension and family

history of POAG (Vincent et al., 2002). *MYOC* encodes myocilin, a highly expressed protein in the eye structures, especially in the TM. After synthesis, this glycoprotein is secreted to the extracellular environment (Yue, 2011). This secretion appears to be disrupted in the presence of mutations (Gobeil et al., 2004). The p.A445V variant is located as other glaucoma mutations in the olfactomedin domain. Although the variant has never been found in the control populations of the previous studies, it is described in public databases. I*n silico* analysis and experimental studies suggests that this variant could be classified as a polymorphism. No differences in the charge, size, or polarity of the mutated protein were observed (Alward et al., 1998) and melting temperature, solubility and secretion are similar to the wild-type protein (Burns et al., 2011). However, increased propensity for aggregation was observed suggesting a possible pathogenic effect (Burns et al., 2011).

The autosomal dominant pattern with variable expressivity described by Souma and collaborators is a possible PCG model in family 4. As proposed by the authors, other mutated gene(s) or modulator agents may be present in PCG patients, being the cause of variable expressivity. In patient 4, the p.V188G *TEK* variant might be the main pathogenic mutation, while p.A445V *MYOC* would act as a modulator agent or as a second mutated gene. Each variant would be inherited from a different parent. As such the segregation study of p.V188G and p.A445V will be essential to confirm the inheritance pattern in family 4. Cases of heterozygous *CYP1B1* and *MYOC* mutations in the same patient have already been reported (Kaur et al., 2005). Two different studies demonstrated that overexpression of myocilin leads to an increase of Ang2, a ligand of Tie2 (Borras et al., 2006), while mutated MYOC leads to a downregulation of Ang2 (Kennedy et al., 2012). According to these facts, mutated myocilin may modify the Tie2 function in the TM through a reduction of Ang2 levels, contributing to changes in the TM and SC and consequently to glaucoma pathogenesis. The future experimental validation of the p.V188G *TEK* functional impact will help understand how this variant can contribute to PCG, as well as the study of the possible interaction between the Tie2 and myocilin.

The *TEK* p.T391I variant present in patient 36 has different characteristics from the p.V188G of patient 4. Similarly to p.V188G and the others PCG-associated *TEK* mutations, it is located in the extracellular part of Tie2, but in the Ig3 domain. However, despite being a very rare variant in the European population, it is relatively common in the African population with homozygous individuals. This elevated MAF and the benign classification obtained by almost all the functional impact prediction software suggest that this variant could be considered as a polymorphism. On the other hand, in the study of Souma and collaborators, a loss-of-function mutation found in an African individual had been previously described in non-African

populations. Other mutated gene(s) or modulator agents may be absent in these non-PCG individuals. Extrapolating to p.T391I, the variant may not be a LoF mutation, but a hypomorphic mutation that only leads to disease in the presence of another gene or modulator, which should be very rare in African individuals.

Another interesting fact is the higher incidence of unilateral PCG observed in *TEK* mutations patients. Half of the PCG samples were reported as having unilateral glaucoma (Souma et al., 2016). Patient 4 also developed PCG only in the right eye, which is in accordance with the previous study, while patient 36 was diagnosed with bilateral glaucoma.

Possible carriers of the variant should also be submitted to ophthalmological examination, because of the possibility of development of a later form of glaucoma or the presence of mild symptoms not identified yet, as described for one family of the study of Souma and collaborators (Souma et al., 2016). To date, relatives' samples and their ophthalmological data were not available.

3.4.2.Mutated genes following the recessive model of inheritance

Until recently, PCG has been classified as an autosomal recessive disease according to the information provided by pedigrees. As such, good candidate genes following this inheritance pattern were expected, as well as genes common to several patients since all of them share the same disease. However, no gene with rare and predicted pathogenic variants was common in these patients. These results may be explained by a different inheritance model as described for *FOXC1* and *TEK* (dominant) or for *CYP1B1* and *MYOC* (polygenic), a different genetic cause for each sample, the presence of mutations in other genomic regions such as introns or promoters, or regions without a good coverage by WES methodology.

In the 16 genes identified in the patients, one of the most interesting was *RNASEH2C*. Patient 31 had a homozygous splicing variant, with both parents having the same variant in the heterozygous state. The c.G468T *in silico* functional impact analysis indicated a probable defect in the donor splice site at the end of exon 3, which may lead to a non-functional protein. *RNASEH2C* encodes the auxiliary subunit C of RNase H2 complex. This complex has another auxiliary subunit – the subunit B, and the catalytic subunit A. The C-terminal of the RNASEH2A subunit interacts with the C-terminal of the RNASEH2C to form a stable and enzymatically active heterotrimer. The main function of the RNase H2 complex is to degrade the RNA strand of the RNA:DNA hybrids. It can also break the 5-phosphodiester bond of one or more ribonucleotides inserted in a DNA duplex. These functions are essential during DNA replication in the removal of lagging-strand Okazaki fragment RNA primers and to maintain genomic

integrity (Reijns et al., 2011). It is well-known that mutations in the subunits of RNase H2 complex can lead to the Aicardi-Goutières syndrome (AGS). AGS is a rare early-onset encephalopathy characterized by calcification of the basal ganglia, white matter changes and progressive brain atrophy in the most severe cases. This disease usually results in severe intellectual and physical handicap. Cerebral anomalies and increased cerebrospinal fluid and serum interferon α resemble a congenital infection. Also frequent are skin lesions, usually chilblains, cardiomyopathy and features of autoimmunity. Thyroiditis and less frequently lupus-like disease can also occur, and most important glaucoma can also be present (Crow et al., 2015; Livingston and Crow, 2016). In fact in a large study conducted by Crow and collaborators glaucoma was the second most common associated phenotype (in 6.3% of the AGS patients) and all but one of 23 patients had congenital glaucoma. The only exception was a patient with bilateral glaucoma diagnosed at the age of 6 years (Crow et al., 2015). Nevertheless, the mechanism behind glaucoma development in AGS patients is still unknown.

Patient 31 was initially studied as a unilateral PCG male patient. Access to additional clinical information raised the possibility of congenital glaucoma associated with another disease. In fact, this patient suffered a neonatal sepsis with petechiae without positive cultures and revealed a developmental delay during the first year of life. However, he had no skin lesions. The patient was not tested for the identification of changes characteristic of ARS: cerebral alterations and an increase of the interferon α and its regulators.

AGS caused by *RNASEH2C* mutations is usually associated with early and severe cases (Crow et al., 2015). However, there is clinical variability as observed in a family with a *RNASEH2C* mutation. Two sisters with the homozygous p.R69W mutation showed a different phenotype. The younger child had a severe AGS phenotype, including an onset in infancy and a profound developmental delay, while the older sister presented an entirely normal intellect with a normal head circumference. In fact, this child was only diagnosed with AGS after the diagnosis of her younger sister because of the presence of chilblains and a mild hemiplegia (Vogt et al., 2013). AGS clinical variability may explain the phenotype observed in patient 31, probably related to an incomplete effect of the variant on the donor splice site, leading to the presence of normal and mutated transcripts. Experimental validation of the impact of the variant on the *RNASEH2C* transcript is reported in Chapter 4.

Furthermore, patient 31 also harbored two heterozygous variants in *COL6A2* (p.G171R and p.E879K). *COL6A2* encodes one of the three alpha chains of type VI collagen, a collagen expressed in the TM beams, JCT region and SC. Tetramers of type VI collagen join to form microfibrils that in the TM are in the heteromorphic aggregates and interband longitudinal sheets form. Type VI collagen binds to hyaluronan and heparin through the N-terminal

domains, which are also responsible for cell association, and the C-terminal domains are involved in self-assembly (Acott and Kelley, 2008). The variants found in patient 31 may also be related with PCG. However, this gene has been associated with myopathies such as Bethlem myopathy (Foley et al., 2009), myosclerosis myopathy (Merlini et al., 2008) and Ullrich congenital muscular dystrophy (Zhang et al., 2014), with symptoms not described for patient 31.

Patient 31 also presented a predicted pathogenic variant in *MID2*, inherited from the mother. This chromosome X gene encodes a protein which belongs to a subclass of E3 ubiquitin ligases and appears to have a role in microtubule stabilization (Suzuki et al., 2010). Mutations in the COS domain have been associated with X-linked intellectual disability (Geetha et al., 2014). The variant identified in patient 31 localizes in the SPRY domain. Information about the genotype of male relatives and their clinical status would be fundamental to evaluate the impact of this variant, which does not appear to be related to PCG pathogenesis.

Another interesting gene was *BCOR,* found in patient 33 with the predicted pathogenic p.S1298F variant. This gene on chromosome X encodes a transcriptional corepressor that interacts with BCL6, AF9, members of the Polycomb group repression complex and histone deacetylases class I and II (Huynh et al., 2000; Srinivasan et al., 2003; Gearhart et al., 2006). p.S1298F variant is located after the AF9 binding site near some of the already described mutations, in a relatively conserved region among species. *BCOR* is associated with Lenz microphthalmia syndrome and oculo-facio-cardio-dental syndrome, both rare X-linked dominant disorders. Ocular (anophthalmia, microphthalmia, coloboma and congenital cataracts), dental, cardiac and skeletal anomalies and mental retardation characterize these diseases (Hilton et al., 2009). Reported cases include glaucoma, and there is one case of congenital glaucoma. However, all these patients also presented non-ocular symptoms (Hilton et al., 2009; Zhu et al., 2015). Patient 33 is a young boy diagnosed with unilateral glaucoma with megalocornea and anterior segment dysgenesis. Studies have demonstrated that disruption of *BCOR* leads to suppression of *PIXT2*, a gene associated with anterior segment dysgenesis (Hilton et al., 2007; Tanaka et al., 2014). Furthermore, two African females are heterozygous for this mutation in ExAC. These females may not have symptoms because this allele may be hypomorphic only leading to disease in the homozygous state. Information about other non-ocular symptoms was not available. This patient can have an undiagnosed case of Lenz microphthalmia syndrome, he can be a mild case associated only with PCG, or the variant cannot be in fact pathogenic. Clinical information of his relatives and functional studies are necessary to further evaluate the impact of this variant. This patient also presented rare and predicted pathogenic variants in *ARSE* and *SLC38A5* on chromosome X. Despite *ARSE* being

associated with X-linked recessive chondrodysplasia punctate, a disease that can lead to cataracts development, there is no described association between this disease and congenital glaucoma. *SLC38A5* had no evident relationship with congenital glaucoma.

Patient 41 presented a rare variant in *FLNA*, the gene for filamin A, located on chromosome X. This protein is an actin-binding that crosslinks actin filaments and links actin filaments to membrane glycoproteins. Mutations in this gene have been found, among other disorders, in otopalatodigital syndrome type II, a severe syndrome characterized by hearing impairment, cleft palate, deformities of extremities, characteristic facies such as frontal bossing, marked micrognathia, hypertelorism, and downslant palpebral fissures. Moreover, Kondoh and collaborators reported a 12-year-old boy diagnosed with otopalatodigital syndrome type II with various anomalies at birth, including bilateral glaucoma (Kondoh et al., 2007). However, ExAC reports p.R2333H in 4 male and 1 female African individuals and ESP reports in 3 male and 1 female African individuals and patient 41 does not show other otopalatodigital syndrome type II symptoms, suggesting that this variant is not pathogenic nor is associated with incomplete penetrance. Clinical information and the genetic study of his relatives would contribute to understanding the impact of this variant.

In addition, this patient also had variants in two other genes: *TNC* and *NEUROD4*. *TNC* encodes tenascin-C, a matricellular protein. This protein participates in cellular adhesion and migration, and its expression is enhanced in response to mechanical stress (Chatterjee et al., 2014). Tenascin-C is expressed in the JCT of the TM, predominantly in the basement membrane underlying the inner wall of SC where it co-localizes with fibronectin (Ueda et al., 2002; Keller et al., 2008; Keller et al., 2013). Studies demonstrated that mechanical stretch of TM increases the expression of tenascin-C and alternative transcripts (Vittal et al., 2005; Keller et al., 2007). Increased levels of tenascin-C were detected in the Rho GTPase-mediated resistance to aqueous humor drainage (Pattabiraman and Rao, 2010), although this protein does not appear to contribute directly to outflow resistance, as its suppression did not significantly affect outflow rate with decreased levels of IOP in knockout mice (Keller et al., 2013). However, the protein appears to be necessary for TM maintenance with or without an IOP increase (Keller et al., 2013). Although the outflow rate has not been found affected in knockout mice, this gene may have some effect on the adhesion and migration of cells during the formation of the human TM, which could contribute to the development of glaucoma in patient 41. The two heterozygous variants identified in patient 41 locate at the beginning of the protein, p.R103H variant before the first EGF-like repeat and p.V295M in the EGF-like 2 repeat. Based on these sites, both variants may not affect the ligation through the fibronectin 3 repeats to other proteins, but may influence the formation of trimers via coiled-coil interactions and the

formation of a hexabrachion, the typical assembly of tenascin-C (Tucker and Chiquet-Ehrismann, 2015).

Patient 41 also revealed the rare and predicted pathogenic variant p.H282Y in *NEUROD4*. Despite the pathogenic potential of this homozygous variant, this protein does not have a described relation with glaucoma. In mouse, this gene was expressed throughout the developing nervous system and after birth it becomes highly expressed in the neural retina (Tsuda et al., 1998). Functional studies and the study of patient 41 relatives are essential to evaluate the impact of TNC and *NEUROD4* in PCG development.

The recessive model approach led to the identification of other pathogenic variants in other samples. The *ABCA4* p.G863A variant was the only variant already associated with a disease: the Stargardt disease, the most prevalent inherited macular dystrophy. However, p.G863A is a mild allele that requires the presence of another pathogenic mutation to lead to disease. To date, no patient was identified with p.G863A in the homozygous state (Maugeri et al., 1999; Maugeri et al., 2002). Because of the advanced state of glaucoma in patient 9 with cloudy cornea, no retina data is available. Meanwhile, although patients with Stargardt disease most commonly start to manifest the symptoms in childhood, others may only begin in early adulthood and least frequently in later adulthood (Tanna et al., 2016). As such this patient may not manifest all symptoms yet. Associations between Stargardt disease and congenital glaucoma have not been reported so far indicating *ABCA4* may not be related to glaucoma. Patient 9 had no other homozygous predicted pathogenic and rare mutation despite the consanguinity history. PCG in this patient may be linked to other types of variants, such as copy number variations (CNVs) or to mutations located in non-coding regions.

Remaining samples had variants in genes that at the time are difficult to associate with the mechanisms involved in PCG pathogenesis.

3.4.3.Mutated genes following the dominant model of inheritance

Although PCG has been classified as an autosomal recessive disease, recent evidence shows that PCG may also be an autosomal dominant disorder at least for some patients (Chakrabarti et al., 2009; Medina-Trillo et al., 2015; Souma et al., 2016). Considering this new information the dominant inheritance pattern was also evaluated in this study. The search for heterozygous variants for each sample retrieved a high number of candidates, with inherent difficulties in the study of all putative causal mutations in detail. Assuming a complete penetrance of the variants, control Individuals with a pathogenic variant for PCG would manifest the disease symptoms. As such, variant selection leaned on those novel and

predicted as pathogenic. However, the probability of incomplete penetrance or expressivity can not be completely ruled out.

Selected filtering parameters identified two genes, *QSER1* and *ACAN,* with pathogenic and novel variants shared among patients, each gene common in two samples. The fact that only two patients shared the same mutated gene is not uncommon. Actually, in the studies of Chakrabarti and collaborators and Souma and collaborators only a few individuals had mutations in *FOXC1* and *TEK* from large populations of PCG patients.

QSER1 encodes the glutamine and serine-rich protein 1, with function still unknown. No data is available regarding its expression in the TM. Further studies are necessary to confirm the importance of this gene to PCG development as well as the real functional impact of the variants. Segregation study in these families is also essential to evaluate if the two variants are *de novo* or inherited and to possibly identify relatives with some symptoms not detected yet.

ACAN encodes the extracellular matrix proteoglycan aggrecan. This molecule is a major component of extracellular matrix of cartilaginous tissues, where it occurs in the form of proteoglycan aggregates. These aggregates consist of several molecules of aggrecan and other link proteins attached to a central hyaluronan filament (Roughley and Mort, 2014). In addition to cartilaginous tissues, aggrecan was also found in the sclera, an eye constituent, with a higher concentration in the posterior sclera (Rada et al., 1997; Rada et al., 2000). Proteoglycans such as perlecan, syndecan, decorin, versican, among others are important constituents of TM (Acott and Kelley, 2008). In TM tissues of goniodysgenetic glaucoma patients, proteoglycans were more frequent however had the same profile and distribution comparatively to normal TM tissue (Tawara and Inomata, 1994). Also, some proteoglycans genes were upregulated in POAG patients' TM (Diskin et al., 2006). However, no transcripts of aggrecan were detected in this tissue neither the protein (Wirtz et al., 1997). This study was performed in TM cells in organ explant and cell culture. The authors refer the inability to detect aggrecan transcripts should not be considered as a strong evidence of the absence of this glycoprotein. A low expression or alternative isoforms of aggrecan can be present in the TM or its expression can occur only in a specific period, such as eye development. Altered aggrecan expression, location or function during TM development due to mutations may contribute to an increase outflow resistance with consequently elevated IOP. The two variants, although not located in important protein domains, occur in regions conserved across species and may lead to reduced protein stability, especially p.G547R. Curiously, *ACAN* is the second gene found with predicted pathogenic variants in the PCG patients that is a component of extracellular matrix. The other gene found was *TNC*, a matricellular protein, in patient 41. Variants in these two genes appear to indicate that the components of extracellular matrix

impact TM development and glaucoma pathogenesis and further studies should be performed to validate the involvement of the extracellular matrix in PCG. Similarly to *QSER1*, only expression and functional studies, as well as segregation studies, can prove the role of *ACAN* in congenital glaucoma development in patients 33 and 45.

Similarly to the recessive model, each sample possessed a different set of genes with possible pathogenic and novel variants. One interesting gene in patient 1 was *S1PR3*. This gene encodes the sphingosine 1-phosphate receptor 3, found in the inner wall of SC (Stamer et al., 2009). Sphingosine 1-phosphate, the membrane phospholipid metabolites involved in cardiac development, cell movement, vascular permeability, among other functions, bind to this receptor (Stamer et al., 2009). These metabolites are present in the aqueous humor. Activation of sphingosine 1-phosphate receptors in porcine eyes decreased the outflow facility (Liliom et al., 1998) with an increase in the density of giant vacuoles in the analog porcine structure of Schlemm's canal (Mettu et al., 2004) probably due to an increased strength of cell-cell junctions between the endothelial cells. In human eyes, Stamer and collaborators described the presence of sphingosine 1-phosphate receptor 3 in the inner wall of SC, but its activation did not culminate in adherens junction proteins changes (Stamer et al., 2009). Other mechanisms may be responsible for this sphingosine 1-phosphate receptor 3 decreased outflow facility. In patient 1 this receptor may be constitutively activated due to the loss of some regulatory domain caused by the nonsense variant. Meanwhile, a more recent study demonstrated that only sphingosine 1-phosphate receptor 2, and not sphingosine 1-phosphate receptor 3, is involved in aqueous outflow regulation (Sumida and Stamer, 2011). This variant in patient 1 is inherited from the carrier mother. Experimental studies on this nonsense variant and the clinical evaluation of the mother would help clarify the participation of this variant in PCG. Curiously, another gene involved in the regulation of endothelial cell proliferation (Phng and Gerhardt, 2009) with a possible pathogenic missense variant present in patient 1 is *NOTCH1*. This gene was also found expressed in HTM cell cultures (Morgan et al., 2014). The variant was also inherited from the mother, which in this specific case does not support the digenic model of variants in two different genes involved in the endothelial cell function.

Patient 1 also had three genes with novel and possible pathogenic variants associated with the extracellular matrix: *MMP1*, *MMP19*, and *NEDD4*, all variants inherited from the father. HTM cell cultures express matrix metalloproteinases. MMP-1 can degrade collagen I, II and III, being responsible for the homeostatic extracellular matrix turnover in the TM (De Groef et al., 2013). Low levels of active matrix metalloproteinases were found in POAG and pseudoexfoliation syndrome/glaucoma patients, especially for MMP-2 (Schlotzer-Schrehardt et al., 2003). Recently, a new glucocorticoid-inducible *MMP1* gene therapy vector reversed and protected

against increased IOP caused by corticosteroid in the sheep model (Gerometta et al., 2010; Spiga and Borras, 2010). NEDD4 has not been described in the TM yet, but it is known to induce the accumulation of collagen type I and fibronectin (Chung et al., 2011), which, if expressed in the TM, may lead to reduced aqueous outflow facility. One of the two mentioned variants inherited from the mother, especially the variant in *S1PR3*, in the presence of the variants in *MMP1*, *MMP19* or *NEDD4* inherited from the father can contribute to the disease in patient 1 in a digenic inheritance model. Experimental studies are imperative to better understand the role of these genes in PCG development.

Patient 5 had two novel heterozygous pathogenic in *FRAS1* and *GUCY1B3*. *FRAS1* is causative of the Fraser syndrome that has as main characteristics cryptophthalmos (absence of palpebral fissures), syndactyly (two or more digits are fused together) and abnormal genitalia. This gene is an extracellular matrix protein and a component of the sublamina densa region of embryonic epithelial basement membranes (Pavlakis et al., 2011). *GUCY1B3*, despite having a variant predicted as pathogenic by all functional impact software, does not have a clear association with the known mechanisms of glaucoma, being difficult to prove its association without experimental studies. Patient 5 also harbored a variant in *CPAMD8.* Recently, mutations in *CPAMD8* following the recessive model were associated with a previously unclassified form of anterior segment dysgenesis including bilateral iris hypoplasia, ectopia lentis, corectopia, ectropion uveae, and cataracts in three families (Cheong et al., 2016). Cheong and collaborators showed expression of *CPAMD8* in the developing human lens, retina, iris, and cornea. Despite the heterozygous genotype of the variant found in patient 5, this gene should not be discarded as a possible candidate for patient 5 phenotype. A detailed study of the expression of this gene in the TM during eye development should help validate the possible impact of this gene in PCG development. The available ophthalmological data of patient 5 does not report other anterior segment anomalies.

A novel variant was found in *HSPG2*, the perlecan gene, in patient 9. Perlecan participates in permeability and cell adhesion. The p.R751G variant is located in the domain III of the protein, involved in efflux of perlecan into the extracellular space and in cell surface binding (Gubbiotti et al., 2016). Perlecan is a known component of the basement membranes of the TM beams and SC inner wall cells (Wirtz et al., 1997) and was up-regulated in the presence of elevated IOP (Vittitow and Borras, 2004). Since perlecan has an important role in tissue integrity, a disruption of this protein may lead to a defective TM with consequent increased IOP.

Patient 18 has a novel variant in the gene that encodes tenascin-X (*TNXB*). This gene is also a matricellular protein as the tenascin-C previously described for patient 41 in the recessive model. *TNXB* is usually associated with Ehlers-Danlos syndrome, a disorder of the connective

tissues affecting skin and joints, among other systems (Demirdas et al., 2016). Tenascin-X was found expressed in the TM of rat eyes (Tuori et al., 1999). However, in a study with *TNX* knockout mice no differences in IOP were found comparatively to their wild-type counterparts (Keller et al., 2013).

Patient 25 had a novel variant in *TRPM3*. *TRPM3* was recently associated with a dominant form of cataracts and glaucoma in a large family. From twenty-five individuals diagnosed with congenital, infantile or developmental cataracts (the diagnosis age ranging from birth to 35 years), fifteen (60%) were also diagnosed with glaucoma (ranged from birth into the fifth decade). The majority of the first relatives developed cataracts. However, one patient was initially diagnosed with glaucoma and another with the two conditions simultaneously (congenital). Some individuals also reported other ocular abnormalities, such as persistent pupillary membrane and retinal detachment (Bennett et al., 2014). Interestingly, the identification of the *TRPM3* variant in patient 25 conducted to a revision of the clinical ophthalmological data, and after the early diagnosis of glaucoma a cataract was discovered in the left eye, as well as discreet macular hypoplasia and anisometropic amblyopia in the same eye. Mouse *Trpm3* has been found expressed in several eye structures including retina, ciliary body and in lens epithelial cells (Karali et al., 2007; Brown et al., 2015). The molecular mechanism behind ocular conditions development is not known yet. Bennett and collaborators proposed that *Trpm3* expression in the ciliary body may have an impact in IOP, since contractions of the ciliary muscle are known to modulate aqueous humor outflow. In the lens and retina, the regulation of Ca^{2+} may be essential for the transparency and light/dark transitions, respectively (Bennett et al., 2014). The authors also suggested a close functional synergy between *TRPM3* and *PAX6* during eye development due to the ocular phenotypic overlap caused by these two genes (Bennett et al., 2014). Considering the ocular phenotypic similarity between patient 25 and the affected and mutated elements of the previously reported family, *TRPM3* should be considered as a possible candidate and functional and segregation studies must be performed to prove the role of this gene in the ocular conditions present in this patient. Differences in ocular phenotypes between patients may be associated with the variable expressivity factor.

Patient 31 had a missense variant in the TM-expressed *ECM2*. This gene encodes a component of the extracellular matrix and was found up-regulated in the presence of glaucoma *MYOC* mutants (Kennedy et al., 2012), suggesting the involvement of *ECM2* in glaucoma development. So, p.N377H should not be excluded as a potential causative variant of patient 31 phenotype. A frameshift *OTX1* variant was also found in this patient, probably leading to the loss of protein function. This gene participates in eye development, being expressed in

early in the ciliary body and anterior retina (Larsen et al., 2009). Knockout mice for *Otx1* have iris reduction and lack ciliary process (Acampora et al., 1996). Because ciliary body participates in the aqueous outflow, a pathogenic variant in this gene may be associated with glaucoma. This patient had another frameshift variant in *CYP27B1*. This gene contributes to the regulation of vitamin D, which may be essential to the ocular barrier function and ocular immune privilege (Alsalem et al., 2014). However, there are no described associations between the vitamin D metabolism and aqueous outflow.

In patient 33 two genes *GGA1* and *RREB1* with a frameshift and a deletion of two a.a., respectively, did not have LoF variants in the ExAC database. The protein encoded by *GGA1* participates in the sorting and trafficking between the trans-Golgi network and endosomes (Hirst et al., 2000), while Ras-responsive element-binding protein 1 encoded by *RREB1* is a transcription factor that binds to ras-responsive element. This protein promotes the expression of calcitonin (Thiagalingam et al., 1996), represses the transcription of the androgen receptor (Mukhopadhyay et al., 2007) and may participate in the cell to cell adhesion and regulation of cell movement (Melani et al., 2008). Till date, no information suggests the involvement of these two genes in the mechanisms behind PCG development.

Patient 43 had one novel missense variant in the gene that encodes collagen type XIV alpha 1 chain. This molecule belongs to the group of fibril-associated collagen with interrupted triple helices and the variant is located in the VWFA2 domain. Collagen type XIV has been found in tissues subjeced to mechanical stress, such as skin, tendon, cornea, and articular cartilage (Ansorge et al., 2009). This collagen was also identified in TM and upregulated in TM cell cultures of POAG patients (Liton et al., 2006). As such, variants in this gene may be related to PCG development.

The missense p.G529E in patient 45 may disturb the ABC transporter domain 1 of the protein encoded by *ABCC4*, the multidrug, resistance-associated protein-4 (MRP4). Pattabiraman and collaborators documented the expression of MRP4 in HTM cell cultures and distribution in human aqueous humor outflow pathway. Inhibition of MRP4 induced TM cells relaxation and increased intracellular levels of cAMP and cGMP. Topical application of MRP4 inhibitor led to a significant IOP reduction in rabbits (Pattabiraman et al., 2013). The absence of a functional transporter in patient 45 may have contributed to PCG development. Functional studies and study of other PCG individuals are essential to prove the role of *ABCC4* in glaucoma.

3.4.4.Genes with *de novo* variants

One of the advantages of parent-offspring trios analysis in a genetic study of a patient is the possibility to find *de novo* mutations. A healthy individual is expected to have 44 to 82 *de novo* mutations with 1 to 2 located in the coding region of the genome (Acuna-Hidalgo et al., 2016). However, these mutations, a consequence of mutagenic events during gametogenesis or postzygotically, are also a major cause of severe early-onset genetic disorders. Recently, studies reveal that WES-based diagnosis of sporadic cases can be explained between 60 and 75% by de novo mutations (Acuna-Hidalgo et al., 2016). However, patients 1 and 31 did not have *de novo* mutations. To date, only *de novo* mutations found in two *FOXC1*-mutated PCG cases associated with the dominant inheritance pattern (Chakrabarti et al., 2009), but no *de novo* mutations were identified in *TEK*-positive PCG individuals (Souma et al., 2016). In the recessive model caused by *CYP1B1* mutations, cases of *de novo* mutations were already suggested. Patients revealed to have homozygous mutations while only one parent was a carrier of the mutation, after exclusion of nonpaternity (Stoilov et al., 1998; Panicker et al., 2004).

In this study, WES exposed several genes with possible pathogenic variants that may be associated with PCG development. This technique proved to be efficient in providing data to the identification of the genetic cause behind PCG pathogenesis. Some patients had good candidates genes. The best was *TEK* with the p.V188G variant in patient 4. The splicing variant c.G468T in *RNASEH2C* found in patient 31 also has a strong probability to be responsible for the phenotype of this patient. The functional study of these two variants and respective genes is presented in Chapter 4. Patient 31 also had two heterozygous variants in *COL6A2*. In patient 1 the *S1PR3* variant inherited from the mother together with *MMP1* variant or *NEDD4* variant both inherited from the father may be behind a digenic model of PCG. Other variants such as p.R751G in *HSPG2* in patient 9, p.R596T in *TRPM3* in patient 25, p.S1298F in *BCOR* in patient 33, p.R2333H in *FLNA*, p.R103H and p.V295M in *TNC*, both in patient 41, and p.G529E in *ABCC4* in patient 45 are also interesting and further studies should be performed to understand their impact on PCG development.

Contrarily to the expected, only two genes with rare and probably pathogenic variants were found in the dominant model of inheritance. However, *ACAN* was reported as not being expressed in the TM and the function of *QSER1* is unkown. This absence of a good candidate gene strongly suggests the genetic heterogeneity of PCG or a different origin of the variants, such as intronic variants or CNVs.

CHAPTER 4

Experimental validation of the pathogenicity of *TEK* and *RNASEH2C* variants

4.1. Introduction

Whole exome sequencing analysis of the PCG population under study identified *in silico* predicted pathogenic variants. However, several tiers of information are required to consider a variant as deleterious and probably causal to the disease. The main challenge is thus to distinguish true pathogenic variants from polymorphisms.

The large-scale population studies through whole genome and exome sequencing reveal polymorphisms exclusive to each population and an unexpectedly high number of variants with putative deleterious impact. Nearly one-third of all variants are private (not previously reported in public databases). For example, the recent study of the exomes of 267 healthy Spanish individuals identified 37% new variants and predicted an average of 353 variants per individual as pathogenic by *in silico* methods (Dopazo et al., 2016). These data show that prioritization by frequency and pathogenicity prediction through *in silico* methods are not enough to guarantee the pathogenicity of a variant in the disease context. In fact, some variants previously classified as mutations have been recently reclassified as polymorphisms.

According to the Standards and Guidelines of the American College of Medical Genetics and Genomics, a variant should be classified as pathogenic based on genetic, *in silico* and experimental data (Richards et al., 2015).

Genetic evidence should demonstrate a significant statistical difference of rare or *de novo* predicted pathogenic variants in cases compared to controls, segregation with the disease within affected families and a frequency consistent with the disease prevalence and expected inheritance model.

In silico data regarding evolutionary conservation and predicted effect on protein function or structure are also important for variant classification. The location of the variant within a functional protein domain may cause activity disruption. For example, the variant can change the enzyme active site inactivating the protein or can modify the protein-binding region preventing ligand binding. Changes in protein stability may also be determinant.

Experimental evidence has an important role in the variant classification process. Expression or localization of the mutated protein can change comparatively to the wild-type protein and function can also be altered. These changes may be observed either in patient cells or in a well-validated *in vitro* model system. The introduction of the altered gene into a cell line or animal model should result in a phenotype consistent with the disease. The wild-type phenotype should be rescued after addition of wild-type gene product or knockdown of the variant allele. Nevertheless, caution is necessary when performing *in vitro* assays or working with model organisms because in some cases expression of the mutated protein may result in a phenotype different from the one observed in the human disease and thus not informative.

Gene expression should be tested when possible in tissues relevant to the disease and proteins tested for interactions with other proteins already associated with the disease (Sunyaev, 2012; MacArthur et al., 2014; Richards et al., 2015).

Whole exome sequencing analysis of patient 4 identified the novel predicted pathogenic p.V188G *TEK* variant. As previously reported in Chapter 3 this variant is located in Ig2 Tie2 domain, the domain responsible for Ang1 binding with consequent Tie2 phosphorylation. It was demonstrated that variants located in the extracellular region of Tie2 might be pathogenic through different mechanisms. These variants may reduce Tie2 expression, lead to the formation of Tie2 intracellular aggregates, reduce Tie2 phosphorylation or inhibit Tie2 translocation (Souma et al., 2016).

Patient 31 harbored the c.G468T homozygous variant in *RNASEH2C* predicted to affect the splice site. Variants affecting splicing can be located in exons or introns. The splicing mechanism depends on the nucleotide sequences (*cis*-acting elements) and the different proteins that bind to them (*trans*-acting elements). The first include the 5' splice site (junction between an exon and an intron, called the donor site), the 3' splice site (junction between an intron and an exon, called the acceptor site), the branch point sequence (located 21–34 nucleotides upstream of the 3' end of the intron (Gao et al., 2008)), and the exonic splicing enhancers (ESEs) and silencers (ESSs), and the intronic splicing enhancers (ISEs) and silencers (ISSs) that recruit *trans*-acting elements, which are splicing regulatory proteins. Mutations that affect the splicing mechanism may have several consequences in the messenger RNA (mRNA) sequence. Exon skipping, intron inclusion, cryptic splicing, leaky splicing, or pseudo-exons introduction into the processed mRNA (less frequently) are some examples of these consequences. Different molecular phenotypes are possible depending on the type and severity of the mutation (Caminsky et al., 2014). Despite the use of *in silico* tools to predict the splicing effect of a variant, the most reliable and straightforward method to detect splicing defects is the analysis of the patient's mRNA.

In this chapter, the damaging effect of p.V188G *TEK* and c.G468T *RNASEH2C* variants was studied to correlate the variants to the patients' phenotype. Impact on expression, solubility, and auto-phosphorylation were analyzed for Tie2 p.V188G *TEK* variant. Transcript size and sequence analysis were performed to study the splicing effect of the c.G468T *RNASEH2C* variant at the end of exon 3.

4.2. Material and methods

4.2.1.p.V188G *TEK* variant experimental studies

4.2.1.1. Cell culture

HEK293 cells were cultured in Dulbecco's modified Eagle's medium (DMEM) high glucose (Corning, Manassas, USA) with 10% fetal bovine serum (Sigma-Aldrich, USA) and 1% penicillin-streptomycin solution (HyClone, Utah, USA) in a 37ºC incubator with a 5% CO_2 atmosphere.

4.2.1.2. Transfection

Plasmids containing the wild-type form of *TEK*, the p.V188G variant, and an artificial kinase-dead mutant due to p.D982A substitution, disrupting DFG motif of kinase domain (KD plasmid) all tagged with Flag were created by Dr. Tomokasu Souma from the Feinberg Cardiovascular Research Institute and Division of Nephrology/Hypertension of the Northwestern University Feinberg School of Medicine, as previously described (Souma et al., 2016). pcDNA3 wild-type vector and site-directed mutagenesis with 5'-GCTGGAGGGTACTCGGCCAGGTATATAG-3' and 5'-CGAGTACCCTCCAGCATCCTGGGGCTGA-3' primers were used to in the construction of the V188G plasmid. HEK293 cells at a confluence of 95% were transfected with 2.5 µg of the respective plasmid diluted in 250 µL of optiMEM (Thermo Fisher Scientific, Grand Island, USA) using 5 µL of lipofectamine 2000 (Thermo Fisher Scientific, Carlsbad, USA) diluted in 250 µL of optiMEM, as represented in Figure 4.1. Cells were incubated for 2 days at 37˚C with a 5% CO_2 atmosphere. One microliter of MG132, a proteasome inhibitor, was added to two wells of the wild-type and p.V188G plasmids in a second plate followed by incubation for 4 or 24 hours (Figure 4.1). DMSO was added as a negative control with an incubation period of 24 hours.

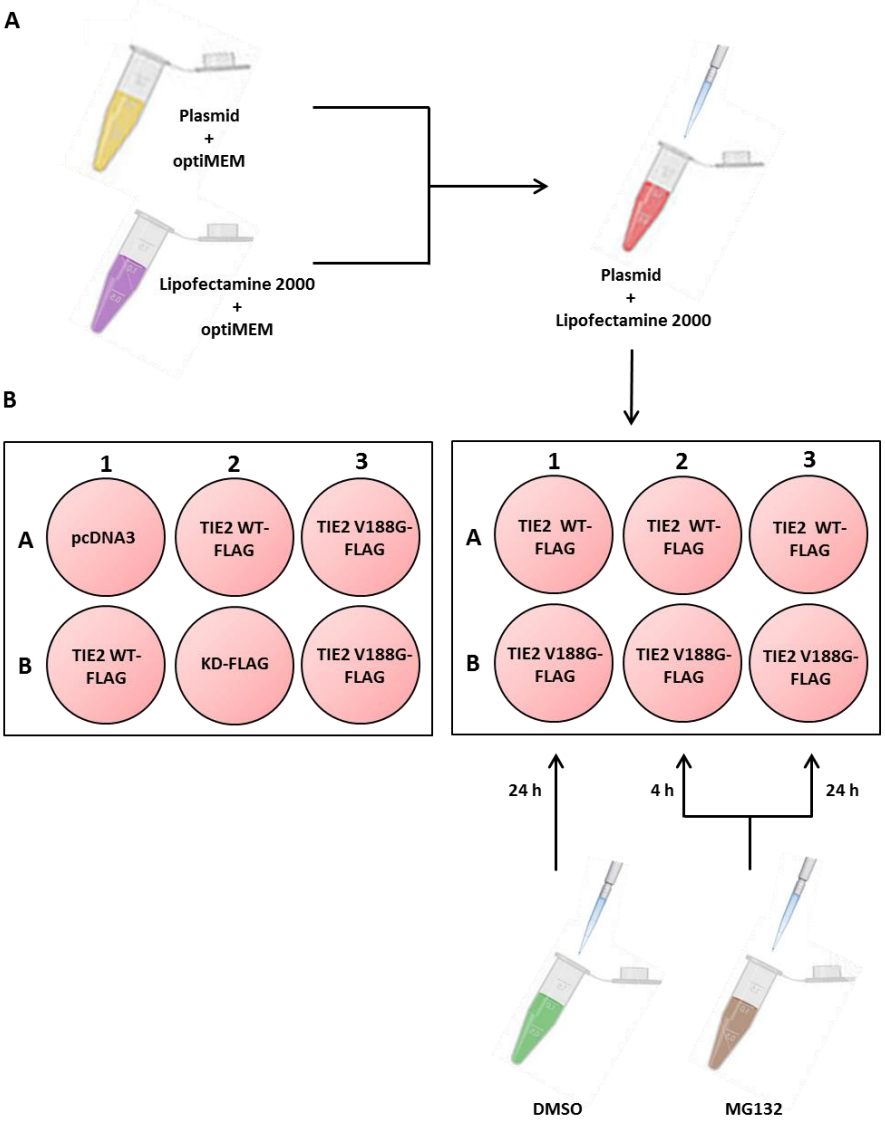


Figure 4.1 – Experimental transfection of HEK293 cells for *TEK* variant validation. (A) Each plasmid was diluted in optiMEM. After 20 minutes incubation at room temperature lipofectamine diluted in optiMEM was added. (B) Twelve wells with HEK293 cells at a confluence near 95% were transfected with the plasmids. In the second plate DMSO was added to the first well of WT and V188G and incubated for 24 hours, the second wells were incubated during 4 hours with MG132 and the last wells during 24 h with MG132.

4.2.1.3. Protein extraction

After DMEM medium removal, cells were lysed with NP-40 buffer containing 150 mM of NaCl, 0.5% of NP-40 (Sigma-Aldrich, USA), 50 mM of Tris-Cl (pH 7.5), 1 mM of DTT, 5 mM of EDTA, 5% of glycerol, SIGMAFAST™ protease inhibitor (Sigma-Aldrich, USA) and phosphatase inhibitor cocktail 3 (Sigma-Aldrich, USA) both at a final concentration of 1X, and the specific phosphatases inhibitors phenylmethylsulfonyl fluoride (Sigma-Aldrich, Germany), sodium fluoride (Sigma-Aldrich, USA) and sodium orthovanadate (Sigma-Aldrich, USA) at the final concentration of 1 mM, 1 mM, and 2 mM, respectively. After 15 minutes incubation with shaking at 4˚C cell lysates were collected and incubated for 30 minutes with rotation at 4˚C.

The lysates were centrifuged at 11,000 x g for 15 minutes at 4˚C. The supernatants were transferred into new tubes. Lysates pellets were saved for expression evaluation (pcDNA3, wild-type Tie2 or variant plasmids from the first plate) or proteasomal degradation (wild-type Tie-2 and variant plasmids from the second plate).

One hundred and thirty microliters of 4X Laemmli Sample Buffer (BIO-RAD, USA) with 10% of 2-mercaptoethanol (FisherBiotech, Fair Lawn, USA) were added to the supernatants and the samples boiled for 5 minutes. The pellets of these wells were washed three times with Tris-buffered saline (TBS) with 0.1% of Triton X-100 (Sigma-Aldrich, USA) and the protease and phosphatase inhibitors previously reported. Two hundred microliters of 2X Laemmli Sample Buffer (BIO-RAD, USA) with 10% of 2-mercaptoethanol was added, followed by 4 series of sonication during 5 seconds and a final heating for 5 minutes at 100˚C.

For auto-phosphorylation evaluation, the supernatant of the cells transfected with wild-type, KD or p.V188G plasmids, were mixed with 15 µL of ANTI-FLAG® M2 Affinity Gel (Sigma-Aldrich. USA) to select the FLAG-proteins from the remaining components of the supernatant. After incubation for 2 hours, the samples were centrifuged at 8,000 rpm for 1 minute at 4˚C and the pellets washed 3 times with TBS with 0.1% of Triton X-100 and protease and phosphatase inhibitors. Forty microliters of 2X Laemmli Sample Buffer (BIO-RAD, USA) with 10% of 2-mercaptoethanol was added to the pellets. All protein lysates were stored at -80ºC.

4.2.1.4. Western Blot

Tie2 expression, solubility, auto-phosphorylation and proteasomal degradation were evaluated by Western blot. For expression evaluation, 20 µL of the denaturated supernatant soluble fraction and insoluble pellets were resolved by electrophoresis on 4-15% sodium dodecyl sulfate – polyacrylamide gel electrophoresis (SDS-PAGE) using Mini-PROTEAN® TGX™ precast gels (BIO-RAD, USA) and proteins transferred onto polyvinylidene fluoride (PVDF) membranes (BIO-RAD, USA). Membranes were cut to separate Tie2 from the α-tubulin loading control and blocked by incubation in 5% non-fat dry milk (Santa Cruz Biotechnology, Dallas, USA), 0.1% Tween 20 in TBS and incubated overnight at 4°C with the anti-Flag M2 antibody (1:5,000; Sigma-Aldrich, USA) for Tie2 and with anti-α-tubulin antibody (1:2,000, DM1A clone, Santa Cruz Biotechnology, USA). After the primary antibody incubation, the membranes were incubated with the horseradish peroxidase (HRP)-linked secondary goat anti-mouse antibody (1:5,000, Thermo Fisher Scientific, USA).

For total auto-phosphorylation, 12 µL of the immunoprecipitated and denaturated samples were separated and immunoblotted as described above, but using 5% bovine serum albumin, fraction V (BSA) (AkronBiotech, USA), 0.1% Tween 20 in TBS and anti-Flag M2 antibody

(1:10,000; Sigma-Aldrich, USA) for Tie2 or anti-pY antibody (1:10,000, 4G10 Platinum, EMD Millipore, Billerica, USA) for phosphotyrosine.

Bands were visualized with enhanced chemiluminescence substrate (ECL) using Clarity™ Western ECL Blotting Substrate (BIO-RAD, USA) and chemiluminescence imaging using Azure c600 (Azure Biosystems).

4.2.2.c.G468T *RNASEH2C* variant experimental studies

4.2.2.1. Cell culture

Adenovirus-transformed HTM cells were gently provided by Dr. Alan Prescott from the Division of Cell Signalling and Immunology of the University of Dundee, Scotland, UK. Cells were cultured in DMEM low glucose, GlutaMAX™, pyruvate supplemented (Thermo Fisher Scientific, Paisley, UK) with 10% fetal bovine serum (Thermo Fisher Scientific, Carlsbad, USA), 1% penicillin/streptomycin (Thermo Fisher Scientific, Carlsbad, USA) and 1% Amphotericin B (Thermo Fisher Scientific, Carlsbad, USA) in a 37ºC incubator with a 5% CO_2 atmosphere.

4.2.2.2. RNA extraction

Total RNA was extracted from the HTM cells. Briefly, the cells were counted in a Neubauer chamber using trypan blue, after trypsinization. TRIzol® Reagent (Thermo Fisher Scientific, Carlsbad, USA), 450 µL, was added to the cell pellet obtained after centrifugation at 300 × g for 5 minutes and the RNA extracted with the Direct-zol™ RNA MiniPrep kit (Zymo Research, USA). Treatment with DNase was performed before adding the Wash buffers. RNA was eluted in 25 µL of nuclease-free water.

Total RNA of patient 31, his mother and a wild-type individual was extracted from peripheral blood samples. After addition of an equal volume of nuclease-free water to blood (150 µL), 900 µL of TRIzol® Reagent (Thermo Fisher Scientific, Carlsbad, USA) were added. The mixture was incubated for 5 minutes at room temperature, and 240 µL of chloroform (Sigma-Aldrich, USA) added. The mixture was homogenized and centrifuged at 12,000 × g for 15 minutes at 4°C. The aqueous phase was placed into a new tube and mixed with the same volume of ethanol (95-100%). Direct-zol™ RNA MiniPrep kit (Zymo Research, USA) was used to purify the RNA as previously described. RNA was extracted twice for each sample and the eluates combined and concentrated using Speedvac (Eppendorf, Hamburg, Germany).

Total RNA concentration was determined by fluorimetry using Qubit® RNA HS Assay Kit (Life Technology, Eugene, USA). Purity was verified by 260/280 and 260/230 O.D. ratio measurements in Nanodrop (NanoDrop Technologies, USA). RNA integrity was evaluated

through observation of the two bands corresponding to the 28S and 18S ribosomal RNA after a 1.2% agarose gel run.

4.2.2.3. Reverse transcription and *RNASEH2C* transcript amplification

Single-stranded cDNA from total RNA was synthesized using the RNA-to-cDNA Kit (Thermo Fisher Scientific, Foster City, USA). After addition of around 1 µg of RNA to 2X RT Buffer Mix (the buffer includes dNTPs, random octamers, and oligo dT-16) and 20X Enzyme mix (MuLV and RNase inhibitor protein) in a final volume of 20 µL, the cDNA was synthesized during 60 minutes at 37˚C followed by 5 minutes at 95˚C.

The impact of the c.G468T variant in *RNASEH2C* transcript sequence was studied by amplification of the cDNA of the *RNASEH2C* transcript from patient 31, his heterozygous mother and a wild-type control individual. Eight oligonucleotide primers were designed with PrimerBlast to amplify seven fragments of the *RNASEH2C* transcripts (Table 4.1). Transcript fragments were amplified by touchdown PCR in 50 µL reactions using the following reagents: 100 ng of cDNA, 1X Phusion GC Buffer, 0.2 mM of each dNTP, 0.2 µM of each primer, 6% DMSO, and 0.02 U/µL of Phusion DNA Polymerase (Invitrogen, Carlsbad, CA). The touchdown PCR program comprised one denaturing step at 98°C for 3 minutes followed by another denaturing step at 99°C for 1 minute; a touchdown phase with a decrease of 1°C in the annealing temperature per cycle: 14 cycles of denaturation at 98°C for 10 seconds, initial annealing temperature at 70°C-1°C per cycle for 30 seconds and extension at 72°C for 30 seconds; another generic amplification phase: 21 cycles of denaturation at 98°C for 10 seconds, annealing temperature at 56°C for 30 seconds and extension at 72°C for 30 seconds; and a final extension step at 72°C for 5 minutes. Amplicons were visualized in a 2% agarose gel electrophoresis stained with ethidium bromide.

Table 4.1 – Oligonucleotide primers used for amplification of the *RNASEH2C* transcript fragments. Primer sequences are listed. All amplifications used the 3'UTR_R primer as the reverse primer.

Primer ID	Primer sequence (5' – 3')
5'UTR_F	AGGGGGTAAGGGGACTACAC
5'UTR_2_F	GCCAGCTTCAGTGTCAGCTC
5'UTR_3_F	TGGGGATCGTGGTCCGTC
E2_F	TCGTGGGATACGTGATGGTG
I2_F	GAGCTGTAGTCCCGATGCTG
E3_F	ATTGGAGCCACTGCCAACTT

Primer ID	Primer sequence (5' – 3')
I3_F	AGACGAGGAGTTCCGTCTTG
3'UTR_R	CTGTGAAGGGATCGCAGCTT

An intronic region of *CYP1B1* was amplified using the I2_3_F and I2_3_R primers to guarantee the efficiency of DNase treatment. Amplification reaction and primers used were the same as described in section 2.2.2.

4.2.2.4. Sanger sequencing

Fragments of patient 31 and wild-type individual amplified with forward primers 5'UTR_2_F and E3_F and the reverse primer 3'UTR_R were purified with AgencourtT® AMPure XP (Beckman Coulter, Brea, USA) and the amplicons sequenced in both directions. Electropherograms were aligned to the ENST00000308418 transcript, from assembly GRCh37, using ContigExpress software.

4.3. Results

4.3.1.p.V188G *TEK* variant experimental studies

Functional impact of p.V188G *TEK* variant was evaluated in cell-based assays to determine its effect on Tie2 protein expression and phosphorylation.

4.3.1.1. Protein expression and solubility

Expression of wild-type and mutant Tie2 in HEK293 cells was evaluated by Western blot. The p.V188G variant resulted in distinctly reduced protein levels when compared to the wild-type both in soluble and insoluble fractions (Figure 4.2). Detection of similar quantities of the α-tubulin loading control for all fractions confirmed identical amounts of the loaded protein.

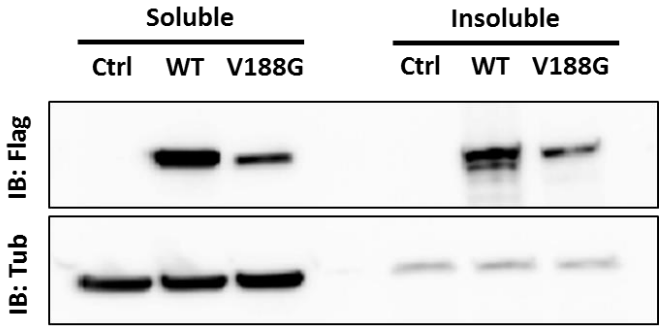


Figure 4.2 – Expression of the Tie2 protein. Western blotting of the WT and the missense p.V188G variant led to a reduction of the protein in the soluble and the insoluble fractions. Tubulin was used as a loading control. Ctrl: control (pcDNA3), WT: wild-type, V188G: Tie2 with p.V188G, IB: immunoblot, Flag: anti-Flag M2 antibody, Tub: anti-α-tubulin antibody.

4.3.1.2. Tie2 auto-phosphorylation

The WT protein auto-phosphorylates when overexpressed. As such, Tie2 wild-type and mutant proteins were overexpressed in HEK293 cells to evaluate the impact of the p.V188G in Tie2 phosphorylation. Selection of Tie2-Flag by immunoprecipitation before Western blot guaranteed the absence of other phosphorylated proteins. The Tie2 protein with no functional kinase domain (Tie2 KD), the domain responsible for the phosphorylation capacity, was used as the negative control. Western blot revealed the absence of phosphorylated tyrosines for p.V188G and the negative control Tie2 KD (Figure 4.3).

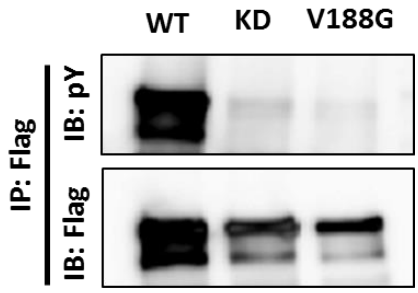


Figure 4.3 – Tie2 auto-phosphorylation. Absence of phosphorylation on both mutant forms of Tie2 was observed by western blot after immunoprecipitation. Despite the reduction of mutated Tie2-Flag comparatively to the wild-type protein, phosphorylated tyrosines were present even in a much lower quantity in the mutant forms than in the wild-type. WT: wild-type, KD: Tie2 with no functional kinase domain, V188G: Tie2 with p.V188G, IP: immunoprecipitation, IB: immunoblot, pY: anti-pY antibody, Flag: anti-Flag M2 antibody.

4.3.1.3. Degradation through the proteasomal pathway

Protein proteasome-mediated degradation is one of the main pathways involved in protein degradation (Hochstrasser, 1992). The study of this important pathway was performed to understand whether proteasomal degradation underlies the reduced levels of protein. Protein expression analysis showed an increased expression of the Tie2-V188G mutant form after addition of the proteasomal inhibitor MG132 (Figure 4.4). This increase was especially observed in the insoluble fraction of cellular lysates, suggesting p.V188G leads to the formation of aggregates. These results indicate that Tie2 level is regulated by proteasome-mediated degradation and structural defects lead to protein degradation through this mechanism.

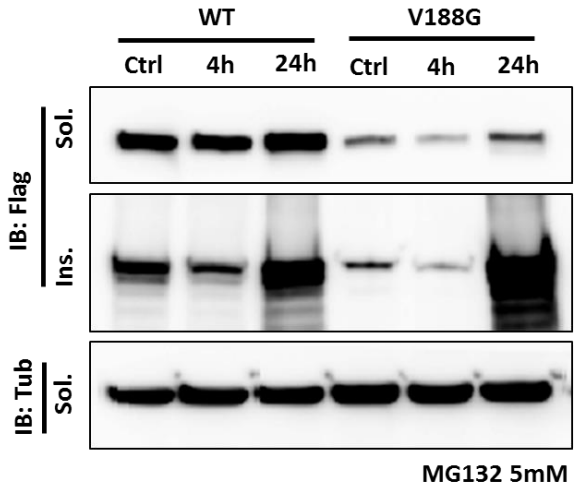


Figure 4.4 – Tie2 expression after treatment with proteasomal inhibitor MG132. Twenty-four-hour Incubation with MG132 increased protein expression. The most significant increase was detected in the insoluble fraction of cellular lysates. Ctrl: control (pcDNA3), WT: wild-type, V188G: Tie2 with p.V188G, IB: immunoblot, Flag: anti-Flag M2 antibody, Tub: anti-α-tubulin antibody, Sol: soluble fraction, Ins: insoluble fraction.

4.3.2.c.G468T *RNASEH2C* variant experimental studies

Functional impact of c.G468T *RNASEH2C* variant was evaluated through the analysis of the transcripts present in blood.

4.3.2.1. RNA extraction and quality control

All extracted RNA samples had high integrity with ribosomal 28S and 18S bands well defined, except a faint smear in the wild-type control sample (Figure 4.5) and no observable DNA contamination with amplification of *CYP1B1* intronic region (Figure 4.6).

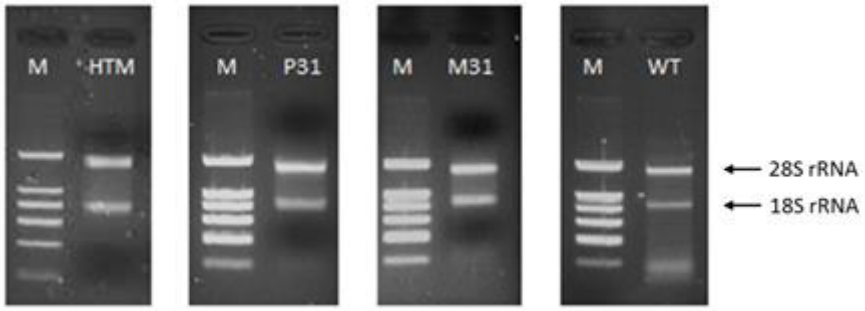


Figure 4.5 – Integrity of RNA samples. The ribosomal 28S and 18S bands are visible in all samples. No smear is detected, except a faint smear in the wild-type control sample. M: marker ladder 1.8 Kb (band size from gel bottom to top: 200, 400, 600, 800, 1000 and 1800 bp), HTM: human trabecular meshwork cell line, P31: patient 31, M31: mother of patient 31, WT: wild-type control individual.

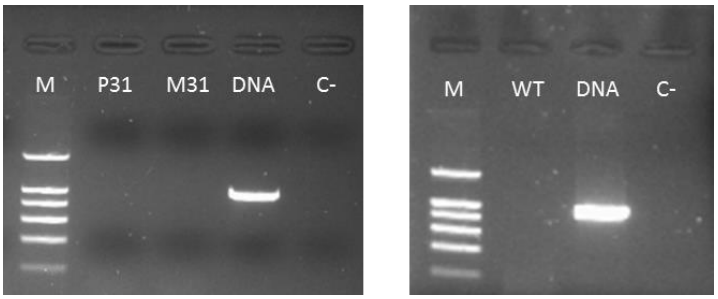


Figure 4.6 – Absence of DNA contamination from the total RNA samples. The *CYP1B1* intronic region of 726 bp only amplified in the DNA control sample. M: marker ladder 1.8 Kb (band size from gel bottom to top: 200, 400, 600, 800, 1000 and 1800 bp), P31: patient 31, M31: mother of patient 31, DNA: a genomic DNA sample, C-: negative control, WT: wild-type control individual.

4.3.2.2. Analysis of the *RNASEH2C* transcript size and sequence

The effect of the *RNASEH2C* variant should ideally be tested on the TM of the patient. However, this tissue is usually not available for total RNA extraction. Since expression of *RNASEH2C* is ubiquitous (Crow et al., 2006), gene expression was evaluated on an HTM cell line to prove gene expression on TM, and the blood samples of the patient, his mother, and a wild-type control to validate the effect of the variant in the transcripts.

The first amplifications were performed using forward primers 5'UTR_F, E3_F and I3_F (Figure 4.7, lanes 5', E3 and I3). Primer 5'UTR_F should provide information on the effect of the variant in the complete *RNASEH2C* transcript, primer E3_F to specifically study the effect of the variant in the loss of the donor splice site at exon 3 as predicted *in silico*, and primer I3_F to guarantee the absence of intron 3 in wild-type individuals. Considering the only transcript described in RefSeq, the NM_032193 transcript, the expected size of the amplicons for the wild-type would be a single band of 678 bp for the amplification using forward primer 5'UTR_F, a band with 178 bp with primer E3_F and no amplification with I3_F.

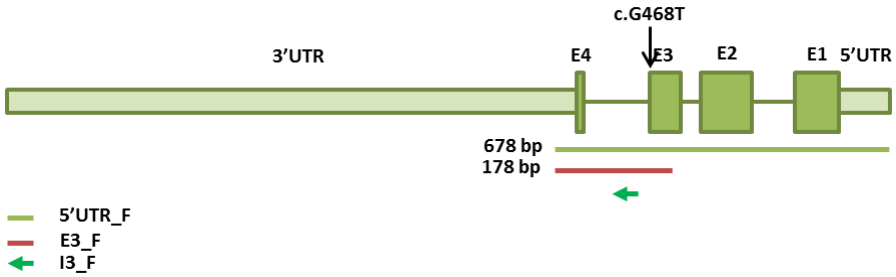


Figure 4.7 – Predicted amplicons expected for 5'UTR_F, E3_F and I3_F forward primers for the NM_032193 transcript. The scheme represents the RefSeq *RNAEH2C* transcript, with light green rectangles representing the UTRs, green rectangles indicating the exons, E1 –E4, and the thin gree line the introns. The colored lines bellow the scheme indicate the position of expected amplicons for the wild-type, with the corresponding size in bp. A small green arrow signals the location of the primer within the intron. Amplification with the intronic primer should not produce any bands.

RNASEH2C transcripts were amplified from HTM cells confirming gene expression. However, the amplification profile was unexpected. Amplification using E3_F led to one band with the expected size (178 bp) suggesting excision of intron 3, two bands for the 5'UTR_F primer, one with the expected size of 678 bp and another with higher size than expected from intron 3 inclusion, and finally an unexpected band for the amplification with I3_F, consistent with inclusion of intron 3 (296 bp) (Figure 4.8).

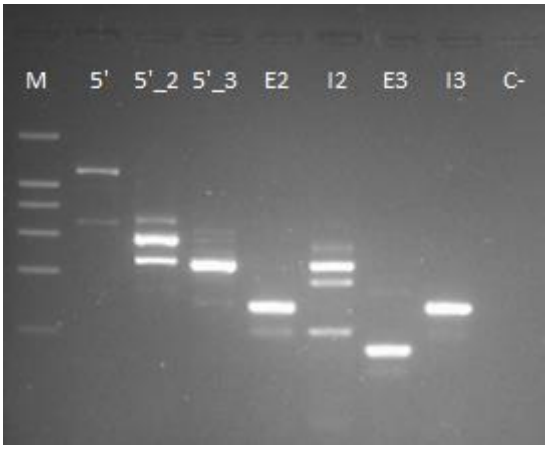


Figure 4.8 – Amplification profile of *RNASEH2C* transcripts in the human trabecular meshwork cell line. M: marker ladder 1.8 Kb (band size from gel bottom to top: 200, 400, 600, 800, 1000 and 1800 bp); 5': 5'UTR_F primer; 5'_2: 5'UTR_2_F primer; 5'_3: 5'UTR_3_F primer; E2: E2_F primer I2: I2_F primer, E3: E3_F primer and I3: I3_F primer; C-: negative control.

Blood samples also expressed *RNASEH2C.* Transcript profile was similar to that obtained for HTM when using E3_F and I3_F primers. However, there was no amplification with the 5'UTR_F primer (Figure 4.9, lines 5', E3 and I3).

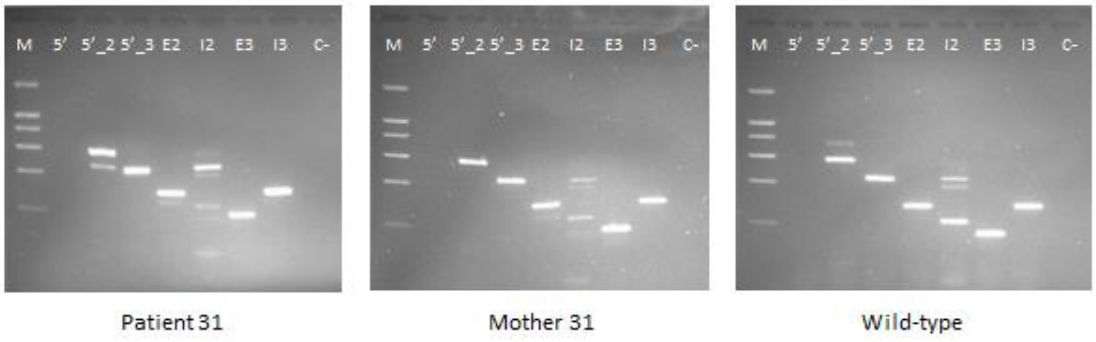


Figure 4.9 – Amplification profile of *RNASEH2C* transcripts in the blood of patient 31, his mother and a wild-type control individual. The profile was similar to that of HTM cells. Amplicon size was similar between the homozygous mutated patient, his heterozygous mother and the control individual. M: marker ladder 1.8 Kb (band size from gel bottom to top: 200, 400, 600, 800, 1000 and 1800 bp); 5': 5'UTR_F primer; 5'_2: 5'UTR_2_F primer; 5'_3: 5'UTR_3_F primer; E2: E2_F primer I2: I2_F primer, E3: E3_F primer and I3: I3_F primer; C-: negative control.

These results implied the presence of additional transcripts in the HTM cell line and blood. Searches in the Ensembl database revealed the RefSeq transcript corresponding to

ENST00000308418 and two additional coding transcripts, ENST00000528220 and ENST00000527610, the latter with a different organization from the other two and retaining intron 3. No information was found for the occurrence of *RNASEH2C* pseudogenes by search in GeneCards and Ensembl. New primers were designed to test the presence of these transcripts. Primer 5'UTR_2_F should amplify the ENST00000308418 (RefSeq) and ENST00000528220 with different amplicon lengths (584 bp and 753 bp, respectively). Primer 5'UTR_3_F specifically aimed ENST00000528220 transcript amplification (458 bp). All three transcripts can amplify with primer E2-F, with a 300 bp amplicon for two of the transcripts and a longer amplicon (535 bp) for transcript ENST00000527610. Primer E3-F was designed to confirm the ENST00000527610 transcript expression (413 bp) by size difference comparison to the other transcripts (178 bp). Finally, primer I2_F should confirm the absence of intron 2 in all transcripts (Figure 4.10).

All primer pairs produced amplicons with a similar profile for HTM cells and blood, except for the RefSeq transcript not observed in blood. Unexpectedly the intronic primer I2_F also generated amplicons. According to Ensembl, intron 2 is not present in any *RNASEH2C* transcript described. In blood samples, the use of primer 5'UTR_2_F led to the formation of one amplicon of the expected size (without the intron) and two additional bands except for the mother of patient 31. Patient 31 presented a smaller size amplicon, the control individual showed a longer amplicon, which may correspond to the amplification of the ENST00000528220 transcript according to the expected length. The corresponding primer amplification in the HTM cells produced the three amplicons (Figure 4.8 and Figure 4.9). Primer 5'UTR_3_F amplified a single band with the expected size for transcript ENST00000528220. The use of primers E2_F and E3_F led to the amplification of fragments of the expected sizes for the wild-type amplicons of ENST00000308418 (RefSeq) and ENST00000528220. Transcript ENST00000527610 was not amplified with the expected size, 535 and 296 bp, respectively.

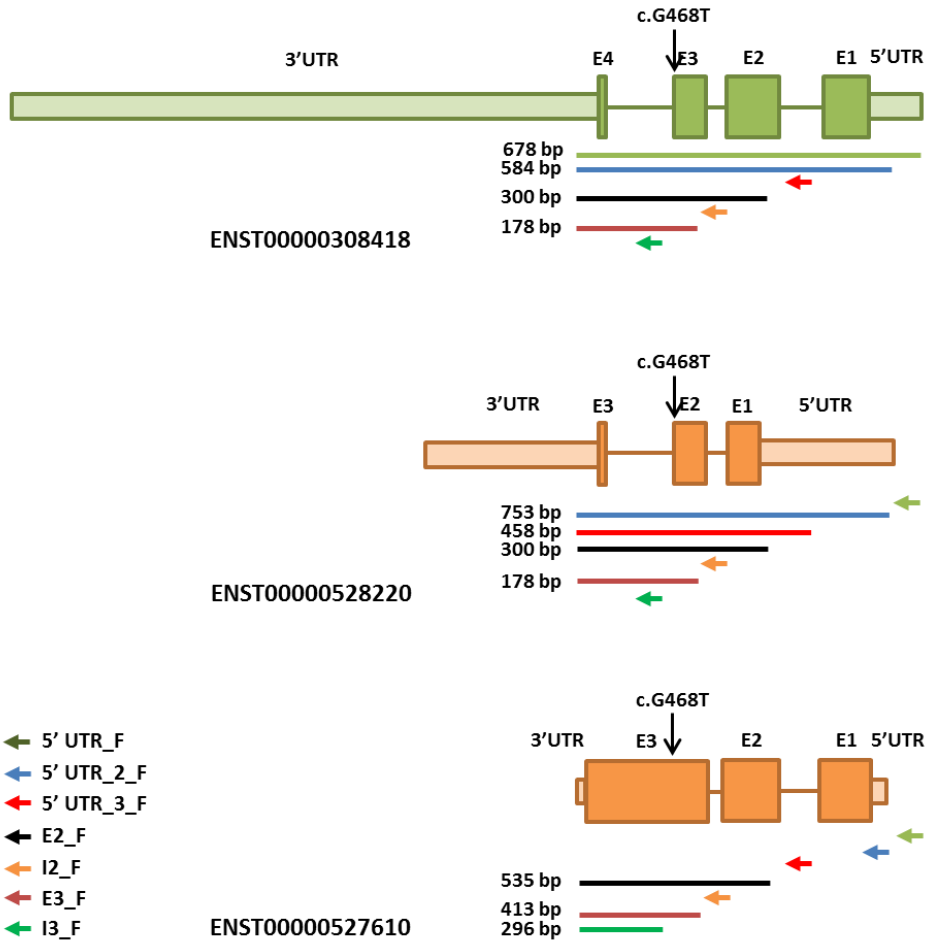


Figure 4.10 – Predicted amplicons expected for *RNASEH2C* transcripts. The schemes represent the *RNAEH2C* transcripts, with light rectangles representing the UTRs, rectangles indicating the exons, E1-E4, and the thin line the introns. The colored lines bellow the schemes indicate the position of expected amplicons for the wild-type, with the corresponding size in bp. The small arrows signal the location of primers within the introns or amplicons where no amplification was expected. Amplification with these primers should not produce any bands.

The results obtained with some of the primers suggest that the variant does not change the expected transcript splicing. However, the amplification of introns suggests the presence of transcripts not described in databases.

Results obtained show the absence of significant differences between the homozygous patient 31, his heterozygous mother and the wild-type control individual (Figure 4.9). If the variant had an impact on *RNASEH2C* transcription, patient 31 would have different transcripts from the control individual, and his mother would have both transcript forms. The absence of differences implies the variant has no functional impact in splicing.

As small weight differences would not be detected by agarose gel electrophoresis, the variant was validated in the amplicons produced by primers 5'UTR_2_F and E3_F in patient 31 and the wild-type control. Sanger sequencing confirmed the absence of abnormal splicing caused by the variant in the patient sample for both amplicons. Patient 31 had the same sequence

without intron 3 as the wild-type control individual, except for the changed nucleotide (Figure 4.11).

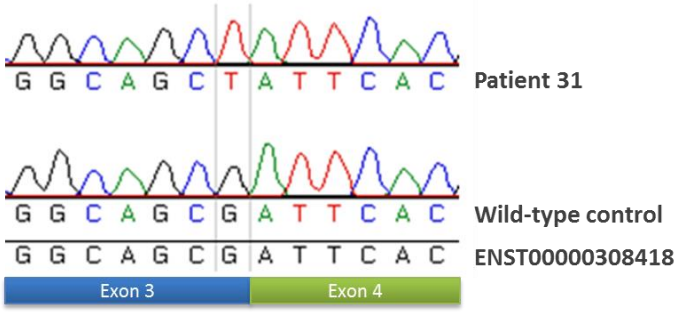


Figure 4.11 – Electropherogram of the 5'UTR_2_F amplified fragment of patient 31 and the wild-type control. ENST00000308418 transcript sequence remains the same, except for the c.G468T variant at the last nucleotide of exon 3. The results of the E3_F amplified fragment were similar.

4.4. Discussion

Experimental validation of the pathogenicity of a variant is a critical step of variant characterization. In Chapter 3, WES led to the identification of two interesting variants in two PCG patients. One was a missense variant in the known candidate gene *TEK* and the other a splice site variant in *RNASEH2C*, a gene associated with Aicardi-Goutières syndrome. Different nature of the variants (missense vs. splice site variant) required different validation approaches.

TEK encodes the tyrosine kinase transmembrane receptor Tie2 involved in developmental angiogenesis and is a key regulator of adult vascular homeostasis (Augustin et al., 2009). Its expression is abundant in vascular and lymphatic endothelial cells and recently was identified in the endothelial cells of SC (Kizhatil et al., 2014). Souma and collaborators found pathogenic variants in PCG patients. Primary congenital glaucoma mutations were related to decreased expression, aggregate formation, reduced phosphorylation or absence of cell junctions traffic (Souma et al., 2016).

Similarly to the results obtained by Souma and collaborators, variant p.V188G found in patient 4 led to Tie2 expression reduction and also to the absence of auto-phosphorylation. Proteasomal degradation was observed with a significant accumulation of Tie2 in the insoluble fraction of cellular lysates when inhibited. Integrating these experimental results with the *in silico* predictions, p.V188G appears to cause a significant protein structural change that will promote protein aggregation. Posteriorly, these aggregates are destroyed by the proteasome after ubiquitin ligation, resulting in the reduction of Tie2 at the cellular membrane. The protein structural change caused by p.V188G appears to disturb the tyrosine kinase domain preventing Tie2 autophosphorylation.

p.C233Y identified in a PCG patient by Souma and collaborators revealed similar protein consequences despite being located in the EGF domain. Expression was reduced such as phosphorylation and the test using proteasome inhibitors also showed proteasome-mediated degradation.

Experimental studies in mice demonstrated that Tie2 is essential for vasculature development, leading to the intrauterine death of knockout mice due to vascular defects. When *Angpt1* is conditionally deleted after E13.5 the severe vasculature phenotype is not present and mice survive (Jeansson et al., 2011). After this study, Thomson and collaborators demonstrated that genetic disruption of Ang1, Ang2 and Tie2 using Ang1/Ang2 and Tie2 knockout mice, results in defects in aqueous humor drainage system due to absence of Schlemm's canal, leading to high IOP, buphthalmos, and other classic features of glaucoma, including retinal ganglion degeneration and vision loss (Thomson et al., 2014). The study of hemizygous mice did not show the ocular phenotype of knockout mice, only an increase of 25% in IOP, but revealed the presence of hypomorphic SC with convolutions and focal narrowing, not only proving Tie2 necessity for SC formation, but also gene dosage sensitivity during canal development (Souma et al., 2016). A strong phenotype absence in hemizygous mice, contrarily to the strong human PCG phenotype, can be explained by the presence of alternative mechanisms that compensate the lack of enough functional Tie2 that humans do not have. Also, *Cyp1b1* knockout mice do not show a very severe phenotype (Libby et al., 2003), suggesting that human and mice anterior ocular structures are somehow different.

In patient 4, p.V188G variant revealed to be a LoF mutation with a decreased expression. This reduced Tie2 expression may lead to endothelial SC cell stress, resulting in a hypomorphic SC with consequent reduced aqueous humor drainage, which leads to IOP increase and other PCG symptoms. Moreover, defects on lymphatic vessel involved in the uveoscleral pathway may also contribute to reducing aqueous humor drainage. The apparent lack of vascular phenotypes in patient 4 and the PCG patients studied by Souma and collaborators suggests that a reduction of Tie2 does not have a strong effect in vasculature development. Other compensatory mechanisms may be activated in these cells. However, these alternative mechanisms appear to be present in SC, leading to the disease.

The presence of *TEK* mutations in PCG patients parents, without glaucoma, reported by Souma and collaborators suggests an inheritance model with variable expressivity. As suggested in Chapter 3, the presence of modifiers may be the cause of this phenotype variability. One possible candidate for patient 4 is *MYOC*. For this reason, the study of patient 4 parents is extremely important to better understand the impact of *TEK* and *MYOC* variants in this patient.

The second experimental validation was performed for the c.G468T variant of *RNASEH2C*. This gene encodes a non-catalytic subunit of RNase H2 complex necessary to degrade the RNA strand of the RNA:DNA hybrids (Reijns et al., 2011). The hypothesis in study was that the donor splice site variant c.G468T would create an alternative splicing leading to the disruption of the C-terminal of the RNASEH2C. This region of the protein is required for interaction with the catalytic subunit of the RNase H2 complex and is essential to the stability and enzymatic function of the complex (Reijns et al., 2011). Analysis of the *RNASEH2C* transcripts on blood cells did not reveal significant differences in patient 31, arguing against the pathogenicity of the variant.

RNASEH2C was expressed in the HTM cell line and blood, in accordance with the described ubiquitous expression of this gene (Crow et al., 2006). However, the type of transcripts identified in this study was somehow different from the ones outlined in the databases, suggesting expression of different transcript forms. The beginning of the 5'UTR of the gene appears to be different in HTM cells and blood, as indicated by the amplification profiles with the primers 5'UTR_F and 5'UTR_2_F. The difference may reside in nucleotide differences in the primer region or transcript beginning. Transcript ENST00000527610 retains intron 3. This transcript explains the unpredicted amplification in all samples when using I3_F primer. However, this transcript must be present in a small proportion in comparison to the other transcripts in the HTM cell line and blood, because when amplified using other forward primers the retention of intron 3 is not detected. Despite the description of several protein-coding and nonsense-mediated decay transcripts, none of these transcripts can explain the amplification obtained when using the I2_F primer. Further experiments are necessary to elucidate the expression patterns of *RNASEH2C* in blood and other tissues. However such investigation is outside the scope of this study.

The c.G468T variant has already been reported in a patient with systemic lupus erythematosus in the heterozygous state. In this study, the authors cloned the *RNASEH2C* cDNA from the peripheral blood mononuclear cells of the patient and observed the retention of intron 3 in 7 out of 20 clones, and predicted an insertion of 51 amino acids between p.A156 and p.I157 (Gunther et al., 2015). This result is in disagreement with the results obtained for patient 31. A possible reason could be the presence of the transcript retaining intron 3 in a small proportion of transcripts amplifiable on blood in the presence of other *RNASEH2C* transcripts without intron 3 but being preferentially cloned instead of other transcripts.

At the time of c.G468T variant identification in patient 31, only the 1000 G and ESP projects data were available. This variant had a low frequency for the T allele: 0.003 and 0.002 for the European population of the respective projects and no homozygous individuals described.

More recently, the availability of data from the ExAC project allowed for a re-evaluation of the variant frequency. The number of individuals in the heterozygous state increased to 242 (MAF(A) of 0.0037) in the European population, and one homozygous individual is also reported. The lack of significative differences between the *RNASEH2C* transcripts in patient 31, his mother and the wild-type control, and the presence of a homozygous individual described in a public database without pediatric disorders suggest that c.G468T alone, even if leading to the retention of intron 3, may be benign.

In conclusion, p.V188G revealed to be a LoF mutation, leading to the creation of a nonfunctional protein. In the future, the genetic study of the parents of this patient should be performed, and the effect of *MYOC* variant correlated with these results to evaluate the possible role of *MYOC* p.A445V as a modulator. According to the experimental validation of c.G468T variant in *RNASEH2C*, the phenotype of patient 31 probably is not related to this gene and variant. The developmental delay may have an independent cause of unilateral glaucoma. In fact, this delay may be a consequence of the neonatal sepsis and this patient may be a true PCG case.

CHAPTER 5

Search of variants in genes associated with other types of glaucoma

5.1. Introduction

Childhood glaucoma is a heterogeneous group of optic neuropathies responsible for 3-6 % of blindness cases in the pediatric population. This percentage increases in developing countries due to later diagnosis and treatment initiation (Kong et al., 2012; Hoguet et al., 2016).

Childhood glaucoma can be classified as primary glaucoma in the case of isolated trabeculodysgenesis, or secondary glaucoma when it develops associated with another condition (Marchini et al., 2014). Primary congenital glaucoma is the most common childhood glaucoma (Tamcelik et al., 2014; Hoguet et al., 2016). Secondary glaucomas encompass several conditions and can be classified in glaucoma associated with a non-acquired ocular anomaly, glaucoma associated with a non-acquired systemic disease or syndrome, glaucoma associated with acquired conditions, and glaucoma following congenital cataract surgery. Despite the higher number of PCG cases over secondary glaucomas, the possibility of the secondary type should be considered during diagnosis. In conditions such as Axenfeld-Rieger syndrome (ARS), Peters syndrome, neurofibromatosis type I or Rubinstein-Taybi syndrome congenital glaucoma may occur as a disease symptom (Kaur et al., 2014; Marchini et al., 2014).

In some patients, congenital glaucoma may appear early in life even before the other symptoms associated with conditions. In these cases, genetic information can be an advantage for accurate disease diagnosis. As an example, half of the ARS patients have congenital glaucoma (Chang et al., 2012) and some cases initially classified as PCG revealed to be ARS. For instance, an Italian patient diagnosed with PCG had his exome sequenced to identify the genetic cause of the disease, and a novel *de novo* mutation in *FOXC1* was identified. As *FOXC1* has been associated mainly with ARS, the re-evaluation of the patient's clinical data led to a final diagnosis of ARS with developmental glaucoma (Pasutto et al., 2015). In another case, a patient was diagnosed with unilateral PCG at one month of age. Three years later the patient developed neurofibromatosis type I symptoms (Li et al., 2015).

Early and correct diagnosis is crucial to guarantee the appropriate treatment for the patient and genetic counseling to the family. A differential diagnosis is essential and, as previously reported, in some cases the clinical features may not be sufficient. The genetic study of the patients can be used not only to confirm the previous clinical diagnosis as also for helping to make a new one.

Although in some cases genetic information can contribute to confirming the diagnosis in certain cases the results may be difficult to interpret. Broad sequence interrogation approaches like a panel or exome sequencing produce extensive reports of known and new variants that have not been interpreted together before. Cases of PCG with variants in genes also causing other diseases have been reported. For example, a few patients had mutations

simultaneously in *CYP1B1* and *MYOC* (Kaur et al., 2005), in *CYP1B1* and *FOXC1* (Medina-Trillo et al., 2016) or just in *FOXC1* (Medina-Trillo et al., 2015). *MYOC* mutations are characteristic of POAG, while *FOXC1* has been mainly associated with ARS. Identification of the role of each gene in the disease mechanism may help understand the mechanisms behind PCG.

As such a group of genes was selected and evaluated for possible pathogenic variants in the *CYP1B1* negative patients under study (Chapter 3). Gene selection took into account disorders with secondary glaucoma, disorders with some of the PCG features but without glaucoma and other glaucoma disorders. These genes were studied for the 14 PCG samples and patient 37, presenting other symptoms in addition to congenital glaucoma.

5.2. Material and methods

5.2.1.Patient description

All negative *CYP1B1* patients previously sequenced for the whole exome continued for the analysis of variants in potential genes associated with congenital glaucoma. Another patient was added to this analysis, patient 37. His exome was sequenced as previously described in section 3.2.2 using Ion AmpliSeq™ RDY Exome Kit with the Ion Proton™ System (Thermo Fisher Scientific, Carlsbad, USA).

Patient 37 was diagnosed at 14 months with bilateral congenital glaucoma. Posterior embryotoxon and iridocorneal extensions were present in both eyes. Corneal diameter, central corneal thickness, anterior chamber and axial length enlargement were increased in both eyes. (Table 5.1 and Figure 5.1). No anomalous anterior segment features were detected in the parents.

The presence of hypertelorism, maxillary hypoplasia, dental configuration (Figure 5.1) and short stature (length below the 3rd centile of WHO growth standards) led to suspicion of ARS. Furthermore, the young boy had frontal prominence, broad flat nasal root and unilateral testicular atrophy. No cardiac malformations, with a normal echocardiography, neither periumbilical skin redundancy were found. Growth hormone and insulin-like growth factor I (IGF1) were normal at 23 months and weight gain and neuromotor development have also been normal.

Table 5.1 – Biometric parameters revealed by detailed ophthalmological examination for patient 37. Posterior embryotoxon and iridocorneal extensions are present in both eyes. Corneal diameter, central corneal thickness, anterior chamber and axial length are increased in both eyes. High intraocular pressure was noted bilaterally while Haab striae is observed only in the right eye.

Parameters	Right Eye	Left Eye
Horizontal corneal diameter (mm)	14	13

Parameters	Right Eye	Left Eye
Vertical corneal diameter (mm)	13.5	13
Axial length (mm)	21.61	21.26
Anterior chamber (mm)	3.96	3.67
IOP (mmHg, by Perkins tonometer, under anesthesia and maleate timolol)	20	18
US central corneal thickness	656	612
Haab Striae	Yes	No
Posterior embryotoxon	Yes	Yes
Iridocorneal extensions	Yes	Yes
Cup-to-disc	Large disc and large cup	Large disc and large cup

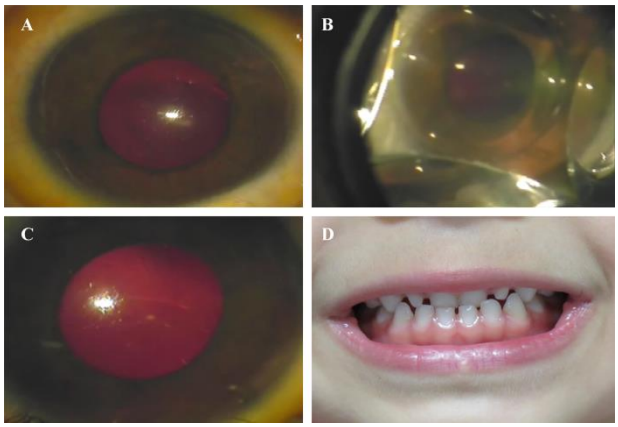


Figure 5.1 – Ocular and dental changes observed in the patient 37. (A) Embryotoxon. (B) Iridocorneal angle showing iridocorneal extensions. (C) Haab Striae. (D) Microdontia.

This patient was included in this study to test if the analysis of the selected genes could be a useful tool for leading to an accurate diagnosis. The study of this patient also aimed confirmation of the secondary cause of congenital glaucoma.

5.2.2.Gene selection

Conditions that may lead to the development of congenital glaucoma were identified through the literature using the PubMed search tool (https://www.ncbi.nlm.nih.gov/pubmed) with the following search terms: “congenital glaucoma”, “infantile glaucoma”, and “pediatric glaucoma”. The London Ophthalmic Genetics Database GENEEYE (London Medical Databases) was also searched with the “infantile onset glaucoma” and “buphthalmos” terms. Genes associated with the disorders were identified through searches in the literature, in OMIM (https://www.omim.org) and Orphanet databases (http://www.orpha.net).

5.2.3.Analysis of the variants present in the selected genes

Non-synonymous and splice site variants with a European MAF inferior to 0.01were selected for next evaluation. *In silico* classification considering the predicted functional impact (SIFT, PolyPhen-2, MutationTaster, LRT, MutationAssessor, FATHMM, RadialSVM, LR, and CADD), conservation among species (GERP++, PhiloP and alignments using Clustal Omega) and protein stability impact (i-Mutant2.0) were performed as previously described in 3.2.4. A variant was considered as pathogenic when 6 out 9 of functional impact software predicted it as pathogenic or when already described as associated with the disease.

5.2.4.Sanger sequencing

Variants found in *NF1*, *ELP4,* and *FOXC1* were confirmed by Sanger sequencing. Sanger sequencing was also used to study regions with lower coverage in *FOXC1* and *CYP1B1*, and the *FOXC1* 5' and 3' UTRs. The amplicons were generated by touchdown PCR in 50 µL reactions using the following conditions: 100 ng of DNA, 1X Phusion GC Buffer, 0.2 mM of each dNTP, 0.2 µM of each primer (Table 5.2), 6% DMSO or 1.5 M of betaine, and 0.02 U/µL of Phusion DNA Polymerase (Invitrogen, Carlsbad, CA). The touchdown PCR program comprised one denaturing step at 98°C for 3 minutes followed by another denaturing step at 99°C for 1 minute; a touchdown phase with a decrease of 1°C in the annealing temperature per cycle: 14 cycles of denaturation at 98°C for 10 seconds, specific annealing temperature plus 14°C with a decrease of 1°C per cycle for 25 seconds and extension at 72°C for 25 seconds; another generic amplification phase: 21 cycles of denaturation at 98°C for 10 seconds, specific annealing temperature for 25 seconds and extension at 72°C for 25 seconds; and a final extension step at 72°C for 5 minutes. *CYP1B1* region was amplified using the *CPY1B1* E2_2F and E2_2R primers as described in 2.2.2. Amplicons were visualized in a 1% agarose gel electrophoresis stained with ethidium bromide.

After purification, the amplicons were sequenced in both directions. Electropherograms were aligned to the reference sequences of *ELP4* (ENSG00000109911), *NF1* (ENSG00000196712), *FOXC1* (ENST00000380874), and *CYP1B1* (ENST00000610745), respectively, with ContigExpress.

Table 5.2 – Oligonucleotide primers used for amplification and sequencing of *ELP4*, *NF1* and *FOXC1* variants. Primer sequence, amplicon size and annealing temperature are listed. bp: base pair.

ID	Primer sequence (5' – 3')	Amplicon length (bp)	Annealing temperature (°C)	PCR additives
ELP4_F	ATGTTCTCCGTGGTCTTCTG	309	55	6% DMSO
ELP4_R	AGATAGGAGGAAAATTATTACCAAG			
NF1_F	GCTGGTATTCAGAATCGAGATTAG	775	58	6% DMSO
NF1_R	GCAGCTACTAGATACCGACCA			
FOXC1 Y64X_F	GCCGCGAGTTCCTGCAA	687	64	Betaine 1.5M
FOXC1 Y64X_R	CACTCGTTGAGCGAGAGGTT			
FOXC1 LC_F	AGGACAGGCTGCACCTCAA	983	61	Betaine 1.5M
FOXC1 LC_R	TGACTCGAACATCTCCCGCA			
FOXC1 3'UTR_1_F	TGAAAAATGGAGAAACCCTCTGAC	993	60	6% DMSO
FOXC1 3'UTR_1_R	GCGAATCACAGGCCACGTA			
FOXC1 3'UTR_2_F	GGCCATTTCTCCAGAGAGTC	999	60	6% DMSO
FOXC1 3'UTR_2_R	CGAAACACACGGCACTCATG			

5.3. Results

5.3.1. Genes with possible involvement in congenital glaucoma

5.3.1.1. Causative genes of disorders associated with secondary childhood glaucoma

Secondary childhood glaucomas may have a non-acquired or acquired origin. The non-acquired are commonly associated with genetic diseases. The most frequent conditions and respective genes are presented in Table 5.3.

Table 5.3 – Genetic disorders and associated genes that may lead to the development of secondary childhood glaucomas (Sampaolesi et al., 2009; Wilson et al., 2009; Yanoff and Sassani, 2009; Allingham et al., 2011; Hoyt and Taylor, 2013; Kaur et al., 2014; Marchini et al., 2014).

Classification	Anomaly or disease	Associated genes
Glaucoma associated with a non-acquired ocular anomaly	Aniridia	*PAX6*, *ELP4*, *TRIM44*
	Congenital iris ectropion syndrome	*PAX6*
	Iridotrabecular dysgenesis (iris hypoplasia)	*FOXC1*, *PITX2*
	Axenfeld-Rieger anomaly/ syndrome	*FOXC1*, *PITX2*, *COL4A1*
	Peters anomaly/ syndrome	*FOXC1*, *PAX6*, *PITX2*, *CYP1B1*
	Sclerocornea	*PXDN*, *FOXE3*

Classification	Anomaly or disease	Associated genes
	Congenital hereditary endothelial dystrophy	*SLC4A11*
	Posterior polymorphous dystrophy	*VSX1*, *COL8A2*, *ZEB1*
	Megalocornea	*CHRDL1*
	Isolated ectopia lentis	*ADAMTSL4*, *FBN1*
Glaucoma associated with a non-acquired systemic disease or syndrome	Marfan syndrome	*FBN1*, *TGFBR2*
	Stickler syndrome	*COL2A1*, *COL11A1*, *COL9A1*, *COL9A2*
	Oculocerebrorenal syndrome (Lowe syndrome)	*OCRL*
	Sturge-Weber syndrome	*GNAQ*
	Neurofibromatosis type 1	*NF1*
	Oculodentodigital dysplasia	*GJA1*
	Charcot-Marie-Tooth Neuropathy Type 4 B2	*SBF2*
	Rubinstein-Taybi syndrome	*CREBBP*, *EP300*
	Homocystinuria	*CBS*, *MTHFR*, *MTR*, *MTRR*, *MMADHC*
	Weill-Marchesani syndrome	*FBN1*, *ADAMTS17*, *LTBP2*, *ADAMTS10*
Glaucoma associated with acquired conditions	Retinoblastoma	*RB1*

5.3.1.2. Causative genes of conditions that share symptoms with PCG but are not glaucoma

Several conditions present symptoms that may resemble PCG symptoms. Some of these do not have a genetic cause as IOP increase due to diagnosis cooperation issues or Descemet tears bands from birth trauma. Others have a genetic basis and have symptoms such as buphthalmos, cornea enlargement, corneal splits, corneal edema or opacity, optic nerve abnormalities. The conditions and respective genes are presented in Table 5.4.

Table 5.4 – Genetic conditions that may resemble PCG symptoms and its respective genes (Khan, 2011).

PCG symptom	Condition	Associated genes
Buphthalmos	Overgrowth syndromes (Sotos syndrome)	*NSD1*
Cornea enlargement	Megalocornea	*CHRDL1*
	Connective tissue disorders (Marfan syndrome)	*FBN1*, *TGFBR2*
Corneal splits	Corneal ectasia (in Brittle Cornea Syndrome)	*ZNF469*, *PRDM5*
Corneal	Congenital hereditary endothelial dystrophy	*SLC4A11*

PCG symptom	Condition	Associated genes
edema or opacity	Posterior polymorphous dystrophy	*VSX1*, *COL8A2*, *ZEB1*
	Congenital stromal corneal dystrophy	*DCN*
Optic nerve abnormalities	Optic disc colobomata	*PAX6*
	Papillorenal syndrome	*PAX2*

5.3.1.3. Other type of glaucomas-related genes

Other type of glaucomas-related genes may also have some influence in PCG pathogenesis. For that reason these genes were also identified and selected (Table 5.5).

Table 5.5 – Other type of glaucomas and their respective associated genes.

Type of glaucoma	Associated genes	References
Primary open angle glaucoma	*MYOC*, *OPTN*, *WDR36*, *CYP1B1*, *NTF4*, *TBK1*, *TMCO1*, *GALC*, *GAS7*, *ATOH7*	(Kumar et al., 2016)
Primary angle closure glaucoma	*PLEKHA7*, *COL11A1*	(Vithana et al., 2012)
Normal tension glaucoma	*TLR4*	(Shibuya et al., 2008)
Pseudoexfoliation glaucoma	*LOXL1*, *CNTNAP2*	(Thorleifsson et al., 2007; Krumbiegel et al., 2011)

5.3.2. Variants present in the selected genes

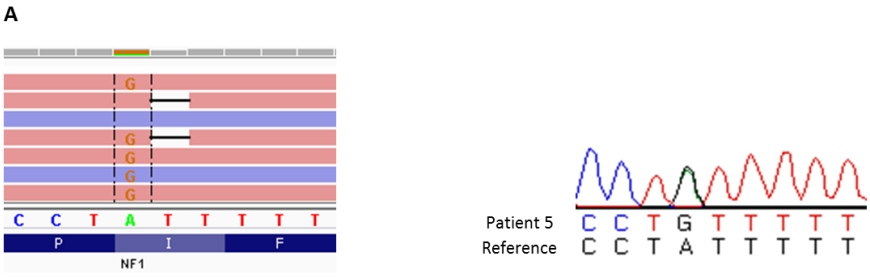
Thirty-two rare variants of the identified genes were selected in the fifteen samples (Table S5.1). A subset of eight variants were predicted as pathogenic by the *in silico* approach or were previously described as pathogenic in literature or databases (Table 5.6).

Table 5.6 – Genes and correspondent variants predicted as pathogenic and/or described as pathogenic in the literature or databases. All variants are in the heterozygous state.

Sample ID	Gene	Variants	Classification
4	*MYOC*	p.A445V	Variant analyzed in Chapter 3.
5	*NF1*	p.I1584V	Variant not described in public databases and in-house database. Predicted as benign but described as pathogenic for neurofibromatosis 1.
9	*CBS*	p.I278T	Variant predicted as pathogenic and described as pathogenic for homocystinuria. However, this disorder has a recessive inheritance pattern.
9	*ADAMTS10*	c.2403+4A>G	Variant predicted to change the donor splice site but present in the homozygous state in 1 individual of the public databases and more than 650 individuals in the heterozygous state. Gene associated with the recessive

			Weill-Marchesani syndrome.
25	*ELP4*	p.Q309H	New missense variant predicted to disturb the donor splice site. Gene that may be associated with anidiria.
37	*FOXC1*	p.Y64X	New variant that leads to the creation of a codon stop. Gene associated with Axenfeld-Rieger anomaly/ syndrome and PCG.
43	*CBS*	p.R369C	Variant predicted as pathogenic. Also described as associated with homocystinuria due to CBS deficiency in HGMD database but has 2 different entries in ClinVar database: classified as pathogenic and as uncertain significance. Found to result in an extensive unfolding with decreased stability and low residual activity. However, homocystinuria due to CBS deficiency is classified as a recessive disease and the variant was identified in the homozygous state in 4 individuals of the public databases.
45	*MTRR*	p.R718C	Variant predicted as pathogenic. Found in 43 control individuals in the heterozygous state. Classified as uncertain significance in ClinVar database. Associated with the recessive homocystinuria-megaloblastic anemia, cbl E type.

Some of the identified variants may have an impact in glaucoma development. Patient 5 harbored the *NF1* p.I1584V missense variant (c.A4750G, rs199474766) in the heterozygous state (Figure 5.2). Despite predicted as benign for the majority of the functional impact predictors (Table S5.1B) this variant was described as causative of neurofibromatosis type 1 (Fahsold et al., 2000; Welti et al., 2011). It locates in a conserved region across species at the beginning of the Sec14 domain, a phospholipids binding domain. Sanger sequencing validated the presence of the p.I1584V variant (Figure 5.2B).



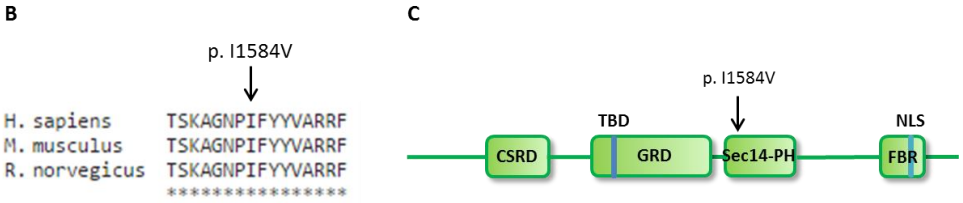


Figure 5.2 – Characterization of the *NF1* p.I1584V variant. (A) An adenine changes to a guanine at position 17:29592335. Sanger sequencing validated the result obtained by WES. (B) The amino acid 1584 is conserved across species. (C) The variant is located at the beginning of the Sec14-PH domain. CSRD: cysteine/serine-rich domain, TBD: tubulin-binding domain, GRD: GAP-related domain, Sec14-PH: Sec14-homologous domain and pleckstrin homology domain, FBR: focal adhesion kinase binding region, NLS: nuclear localization signal. The left side of the protein scheme represents the protein N-terminal.

Patient 25 presented the new missense *ELP4* p.Q309H variant (c.G927T). This variant was pathogenic according to five predictors and as benign by the other four (Table S5.1B). It was also predicted to disturb the donor splice site of exon 7 (Table 5.7). Sanger sequencing confirmed the presence of the heterozygous variant (Figure 5.3).

Table 5.7 – Scores obtained with different splicing effect software for the wild-type and the c.G927T alleles. A higher score indicates a higher degree of similarity to the consensus sequence or a higher probability or confidence of a site being a true splice site.

Software	Wild-type score	Mutated Score
Human Splicing Finder	88.86	77.99
NNSplice	0.99	disappeared
SplicePort	2.17	0.34
MaxEntScan	10.03	3.06

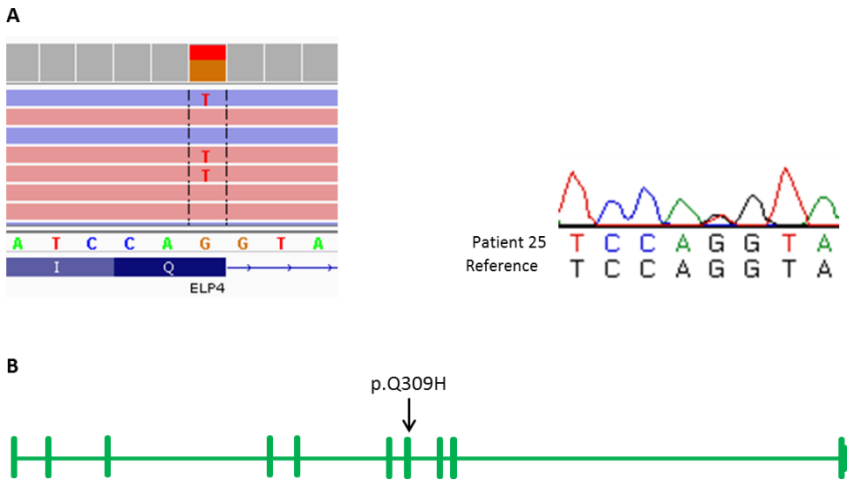


Figure 5.3 – Characterization of the *ELP4* p.Q309H variant. (A) A guanine changes to a thymine at position 11:31653952. Sanger sequencing validated the result obtained by WES. (B) The variant is located at the last position of exon 7, predicting to disturb the donor splice site.

Patient 37 presented a novel heterozygous nonsense variant in *FOXC1* (c.192 C>G, p.Y64X) (Figure 5.4). Sanger sequencing confirmed the variant in the patient and its absence from both unaffected parents, indicating the presence of a *de novo* variant. The *in silico* functional impact evaluated by MutationTaster classified the variant as "disease causing", and CADD scored the variant with 29.3, suggesting the deleterious effect of the mutation. This p.Y64X mutation is located before the DNA-binding forkhead domain (FHD) (Figure 5.4), resulting in a truncated protein composed only of the activation domain with the loss of 88.6% of the protein. Consequently, the FHD, the inhibition domain, and the activation domain 2 will be absent. *FOXC1* presented no other rare variants.

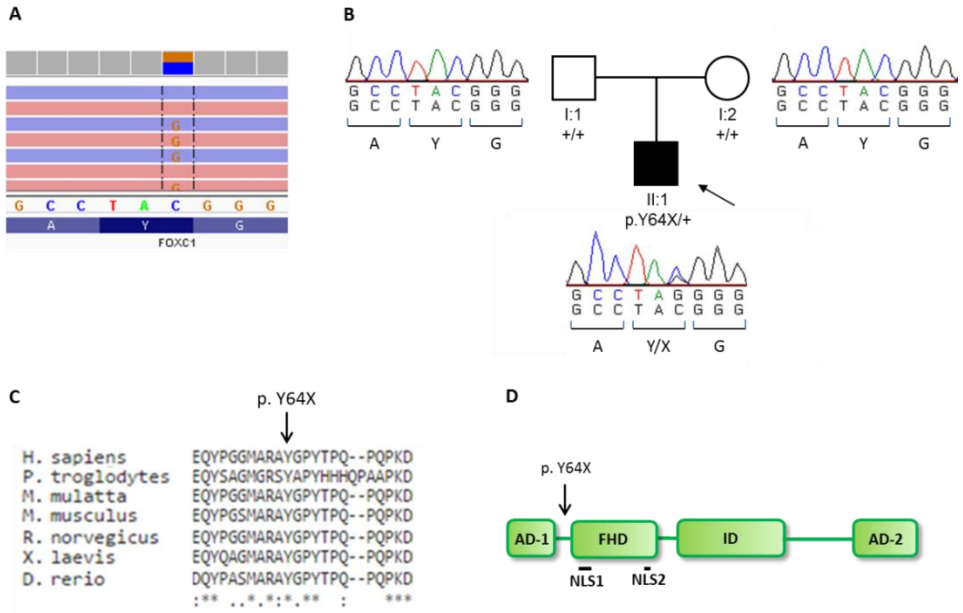


Figure 5.4 – Characterization of the *FOXC1* p.Y64X variant. (A) A cytosine changes to a guanine at position 6:1610872. (B) Sanger sequencing validated the result obtained by WES and demonstrated to be a *de novo* variant. (C) The amino acid 64 is conserved across species. (D) The variant is located before the forkhead domain, which binds to the DNA, leading to the loss of forkhead domain, inhibition domain and activation domain 2. AD-1: activation domain 1, FHD: forkhead domain, ID: inhibition domain, AD-2: activation domain 2, NLS1: nuclear localization signal 1, NLS1: nuclear localization signal 2. The left side of the protein scheme represents the protein N-terminal.

The presence of the nonsense *FOXC1* p.Y64X variant together with the clinical features of hypertelorism, maxillary hypoplasia, microdontia (Figure 5.1), short stature frontal prominence, broad flat nasal root and unilateral testicular atrophy confirmed the ARS diagnosis in patient 37.

Moreover, a common previously described variant as a susceptibility locus for primary angle closure glaucoma in *COL11A1* (Vithana et al., 2012) was identified when selecting the rare variants of this gene. The variant p.P1323L, c.3968C>T, rs3753841 was found in heterozygosity in patient 37 (Figure 5.5).

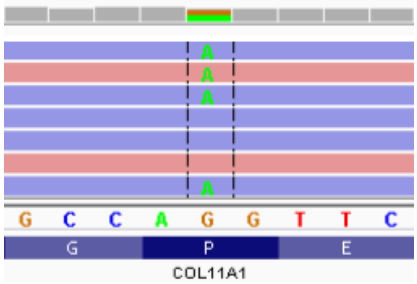


Figure 5.5 – *COL11A1* p.P1323L variant. A guanine changes to an adenine at position 1:103379918.

Two predicted pathogenic *CBS* variants in patient 9 (p.I278T) and in patient 43 (p.R369C), as well as a *MTRR* variant in patient 45 (p.R718C) were found in the heterozygous state. Both genes are associated with the development of recessive forms of homocystinuria. Patient 9 had the splicing variant c.2403+4A>G identified in *ADAMTS10* in the heterozygous state. *ADAMTS10* is a causative gene for the recessive form of Weill-Marchesani syndrome, a rare connective tissue disorder.

5.4. Discussion

A group of genes with possible involvement in congenital glaucoma development was selected and studied by WES in 15 samples. This approach allowed the genetic confirmation of ARS suspected for patient 37 through the identification of a novel nonsense variant in *FOXC,* a well-known causal ARS gene. This patient had the predicted pathogenic variant p.Y64X, which locates before the DNA-binding FHD, resulting in a truncated protein composed only of the activation domain 1 with the loss of 88.6% of the protein. Consequently, the FHD, the inhibition domain, and the activation domain 2 will be absent. In a recent study, Medina-Trillo and collaborators described an increased transactivation activity in two nonsense mutations: p.Y47X and p.Q106X. Similarly to the p.Y64X variant, p.Y47X is also located before the FHD. The authors suggested that this unexpected gain-of-function might be related to the isolated expression of the activation domain 1, without the inhibitor domain. The activation domain 1 appears to act as a potent and independent transcription up-regulator that can interact with other transcription factors beyond *FOXC1* (Medina-Trillo et al., 2015). According to this data, the *de novo* variant identified in patient 37 can create a hypermorphic allele, being responsible for the ARS phenotype observed in the patient.

Axenfeld-Rieger syndrome is a rare and autosomal dominant developmental disorder characterized by anterior segment dysgenesis, affecting the peripheral cornea and iris, and iridocorneal angle. Iris attachments to the posterior embryotoxon (prominent Schwalbe's line) in both eyes characterize ARS patients. Other iris findings might include stromal thinning or

atrophy, corectopia, iris holes and ectropion uvea. A major consequence of the anterior segment dysgenesis is increased intraocular pressure, leading to glaucoma, in approximately half of the ARS patients. Extraocular findings in ARS are usually present and may include craniofacial anomalies, cardiovascular malformations, anomalies of the pituitary gland, growth hormone deficiency, male genitourinary system anomalies (hypospadias) or periumbilical skin bridges (Tumer and Bach-Holm, 2009; Chang et al., 2012). In addition to the ophthalmological characteristics of patient 37, the presence of hypertelorism, maxillary hypoplasia. frontal prominence, broad flat nasal root, microdontia, unilateral testicular atrophy, and short stature were also suggestive of ARS.

The mechanism behind the development of glaucoma caused by *FOXC1* disruption is not clearly understood. This transcription factor is expressed during eye organogenesis in the periocular mesenchyme, trabecular meshwork, sclera, and conjunctival epithelium (Kidson et al., 1999). It has a significant role in the correct corneal development, preserving corneal transparency by regulating vascular growth and is also involved in the correct iris and trabecular meshwork formation (Kidson et al., 1999; Seo et al., 2012). Studies performed in HTM cells revealed the important role of FOXC1 in the response to oxidative stress and FOXC1 dysregulation results in TM cell death. The loss of TM cellularity will prevent aqueous humor uptake with a consequent increase in IOP, a major risk factor for the development of glaucoma (Berry et al., 2008; Ito et al., 2014).

Patient 37 also had the common *COL11A1* p.P1323L variant previously described as a susceptibility locus for PACG (Vithana et al., 2012). *COL11A1* encodes a α chain of type XI collagen and pathogenic mutations in this gene are causative of syndromes with ocular involvement, such as Stickler syndrome and Marshall syndrome (Richards et al., 2010; Khalifa et al., 2014). This protein is a component of the extracellular matrix of TM and, therefore, contributes to its structure (Acott and Kelley, 2008; Vithana et al., 2012). Despite rs3753841 being predicted to have a moderate effect on the protein, a small aberrant activity or expression, in the presence of the hypermorphic *FOXC1* allele, could contribute to a disruption of the TM and aqueous humor drainage with IOP increase. Further studies will clarify the possible impact of this variant in association with *FOXC1* in the pathogenesis of glaucoma in ARS patients.

In addition to the *FOXC1* variant confirming ARS diagnosis, other interesting variants were found in studied patients. The most relevant were p.I1584V in *NF1* and p.Q309H in *ELP4*. Patient 5 presented the heterozygous p.I1584V variant previously classified as causative of neurofibromatosis type 1 in two studies (Fahsold et al., 2000; Welti et al., 2011). Neurofibromatosis type 1 is a dominant hereditary tumor predisposition syndrome

characterized by the development of multiple benign and malignant tumors as neurofibromas of the peripheral nerves and also leukemia, pheochromocytomas, gastrointestinal stromal tumors, among others (Kresak and Walsh, 2016). One possible complication of this disease is congenital glaucoma. Over the years several cases of neurofibromatosis type I have been described with congenital glaucoma and some of them even before appearance of other typical disease symptoms, which initially suggests the presence of PCG (Payne et al., 2003; Colas-Tomas et al., 2010; Edward et al., 2012; Oystreck et al., 2012; Li et al., 2015; Mbagwu et al., 2015). The development of glaucoma in these patients can result from infiltration of the anterior chamber angle by neurofibromas, from an angle closure due to neurofibromatous thickening of the ciliary body and choroid, from fibrovascularization leading to synechial angle closure and neovascular glaucoma, or from developmental angle abnormalities (Edward et al., 2012). Weti and collaborators experimentally evaluated the functional impact of the p.I1584V in the structure and lipid binding and exchange activities. However, there was no mRNA degradation, decrease in protein stability or insolubility. The folding of the mutant and wild-type proteins was similar. Crystallographic analysis of Sec14-PH domain only revealed a small difference of the protein backbone around the spatially close residue E1699 and the mutated protein appeared not to disrupt lipid binding and exchange activities (Welti et al., 2011). These results are in accordance with the predicted classification of a benign variant by the *in silico* approach. In addition, no symptoms of neurofibromatosis type I were described for patient 5, a 14-year old male patient. According to these information a reclassification of the p.I1584V to a polymorphism could be considered. However, the possibility of incomplete penetrance can not be ruled out.

A few cases involving delections of *ELP4* have been associated with aniridia and in some patients with other ocular anomalies as cataracts and glaucoma (Cheng et al., 2011; Zhang et al., 2011). This gene had also been proposed to play a role in the optic nerve disc area (Gasten et al., 2012). *ELP4* encodes a component histone acetyltransferase complex of the RNA polymerase II holoenzyme, involved in transcriptional elongation and is postulated to regulate *PAX6*. A delection containing *ELP4* led to the absence of *PAX6* expression (Lauderdale et al., 2000). The heterozygous missense and splicing variant p.Q309H was identified in patient 25. This variant can disturb the protein after exon 7, as indicated by predictors. Patient 25 developed a cataract, a discreet macular hypoplasia, and anisometropic amblyopia in the left eye after congenital glaucoma diagnosis. Considering these characteristics, the p.Q309H variant should not be discarded as participating in the phenotype patient 25.

The other predicted pathogenic variants identified in patient 9, 43 and 45 in genes *CBS* and *ADAMTS10*, *CBS* and *MTRR*, respectively. However, all these genes are associated with

recessive disorders, suggesting that these variants may not related with the patients phenotype.

This strategy proved to be efficient for patient 37 diagnosis. However, this patient also had other symptoms in addition to the initial congenital glaucoma suggestive of ARS. The search for pathogenic variants in the other 14 PCG patients suggested the absence of genetic disorders associated with secondary congenital glaucoma and other glaucomas and a correct classification of PCG. However, the possibility of secondary congenital glaucoma due to unknown genetic causes should not be discarded. A future detailed study of the pathogenic variants that emerged from this analysis in PCG mechanism may contribute to a better understanding of the disease.

CHAPTER 6

Final conclusions and future perspectives

Since the discovery of the GLC3A, GLC3B, and GLC3C loci in PCG samples, the identification of the genetic causes behind PCG has been studied. However, only one gene is recognized by all the scientific community as causative of PCG, the *CYP1B1*. Moreover, mutations in this gene are the principal cause of PCG. Still, a significant number of patients remain without identification of the causal mutations, even in the recent discovered candidate genes.

This study focused on the identification of genetic variants in a Portuguese population of PCG patients. The main purpose of this work was not only the identification of pathogenic variants in the PCG candidate genes, but essentially the study of the patient without mutations in these genes to discovery new genes and putative pathogenic variants and understand how they are involved in the disease development.

The initial genetic study of 37 PCG Portuguese patients revealed an unexpectedly high number of patients with mutations in PCG candidate *CYP1B1* (Chapter 2). In fact, sixty-five percent of the patients were found to have *CYP1B1* mutations in exons according to the recessive inheritance model described for PCG. This number contrasts with the 22-50% reported for non-consanguineal populations (Stoilov et al., 2002; Sitorus et al., 2003; Chakrabarti et al., 2010; Fuse et al., 2010; Milla et al., 2013). A study by Cardoso and collaborators in a Portuguese PCG population identified *CYP1B1* mutations in only 28.6% of the patients (Cardoso et al., 2015). These results suggest that the geographic origin of patients within the country may be responsible for the gene mutation prevalence differences. No mutations in regulatory and intronic regions of *CYP1B1* were detected. These results agree with other studies since until date only two promoter mutations have been found in PCG patients (Chakrabarti et al., 2010; Lopez-Garrido et al., 2013). This study identified 13 patients without *CYP1B1* mutations that became the population of study (Chapter 2). .

At the beginning of this study, three additional PCG candidate genes had been reported in a reduced number of patients: *LTPB2* (Ali et al., 2009; Narooie-Nejad et al., 2009), *FOXC1* (Chakrabarti et al., 2009) and *MYOC* (Kaur et al., 2005). These findings indicated a low mutation probability in these candidate genes in the 13 PCG patients without *CYP1B1* mutations and also the existence of other genes involved in PCG pathogenesis. Whole exome sequencing is a recent high-throughput methodology that allows the sequencing of almost all coding and essential splice site regions of the genome. For that reason, the thirteen PCG patients without *CYP1B1* mutations and a possible PCG patient also without *CYP1B1* mutations were analyzed by WES (Chapter 3).

The first approach considered the search of putative pathogenic variants in genes associated with the disease. This search identified the heterozygous missense p.V188G variant in *TEK*, a PCG candidate gene reported in 2016, in patient 4 (Chapter 3). The *in vitro* expression,

solubility, and auto-phosphorylation of the p.V188G Tie2, the protein encoded by *TEK* showed decreased expression and autophosphorylation, with the formation of aggregates posteriorly degraded by the ubiquitin-proteasome pathway, confirming the initially *in silico* predictions of pathogenicity (Chapter 4). These results represent the second study showing the involvement of *TEK* in PCG pathogenesis. Interestingly, this patient also had the heterozygous *MYOC* missense p.A445V variant. Different studies associated this variant with other types of glaucoma, but a recent analysis suggests its reclassification as a polymorphism (Donegan et al., 2015). In patient 4, *MYOC* p.A445V in addition to the pathogenic *TEK* variant may be determinant for PCG phenotype. The segregation study of both variants is imperative to understand the impact of the *MYOC* variant. If each variant is inherited from a different progenitor, a detailed experimental study of the two variants simultaneously should be performed to evaluate the impact of this combination. The other PCG candidate genes did not harbor other rare and predicted pathogenic variants in the studied patients (Chapter 3). This result suggests the contribution of other mutated genes to disease pathogenesis.

As PCG has been considered a recessive disease, variant search targeted genes with rare and predicted pathogenic variants following this model (Chapter 3). No common genes with relevant variants were identified in patients. The most interesting homozygous variant was a splicing variant in *RNASEH2C* in patient 31. *RNASEH2C* is not associated with PCG, but is related to Aicardi-Goutières syndrome. Some patients with this disorder manifest congenital glaucoma, and patient 31 also presented other symptoms that could be associated with AGS. Contrary to the *in silico* predictions, analysis of the transcripts present in the blood of the mutated *RNASEH2C* patient did not reveal differences comparatively to the transcripts of a wild-type control (Chapter 4). These results do not support *RNASEH2C* variant as causative of patient 31 phenotype. However, some patients with this disorder manifest congenital glaucoma and furthermore patient 31 also presented more symptoms that could be associated with AGS. Contrary to the *in silico* predictions analysis of the transcripts present in blood of mutated *RNASEH2C* did not reveal differences comparatively to the wild-type transcript. These results invalidates the *RNASEH2C* variant as causative of the phenotype of patient 31. However, alternative splicing may have occurred in a cell subset, not being detected by PCR. Testing the presence of altered transcripts in isolated cells may help to elucidate the consequences of this variant.

Other patients had interesting candidate genes with pathogenic homozygous mutations (Chapter 3). Patient 33 had a rare and pathogenic variant in *BCOR*, a causative gene of Lenz microphthalmia syndrome and oculo-facio-cardio-dental syndrome, with a case of congenital glaucoma already reported in a patient with other ocular and non-ocular features (Zhu et al.,

2015). Patient 41 harbored two heterozygous variants in *TNC*. Tenascin-C is a component of the TM that appears to be necessary for TM maintenance and homeostasis (Keller et al., 2013). The absence of a strong PCG candidate gene for the recessive model and the identification of the heterozygous variant in *TEK* led to the analysis of genes following the dominant inheritance model (Chapter 3). As expected the number of variants for each sample was high, difficulting the study of each variant in detail. Therefore, the novel and predicted pathogenic variants were selected. Similarly to the recessive model, there was no relevant candidate gene shared by the samples. Patient 1 inherited an *S1PR3* variant from the mother and the *MMP1* and *NEDD4* variants from the father. Two of these genes are related to TM, *MMP1* is a member of the meshwork and *NEDD4* also contributes to TM organization. *S1PR3* is a gene expressed in SC and involved in the regulation of angiogenesis and vascular endothelial cell function. The pathogenicity of the variants and their inheritance pattern, and the gene correlation with TM and SC, involved in glaucoma, may suggest a digenic case of PCG. Patient 9 harbored an *HSPG2* variant. This gene encodes perlecan, a component of the TM and SC (Wirtz et al., 1997), and disruption of this protein may be related to PCG development. Patient 25 had a novel and predicted pathogenic variant in *TRPM3*, a gene proposed to have an impact in IOP and associated with other ocular conditions, such as cataracts (Bennett et al., 2014). Interestingly, patient 25 revealed the presence of a left eye cataract, reinforcing the importance of this gene in patient 25 phenotype. Patient 45 had a variant in *ABCC4*. The MRP4 protein encoded by this gene is expressed in HTM cells and was proved to have a significant role in IOP regulation. The segregation analysis in families and *in vitro* experimental validation of these variants should be evaluated to understand the impact of these genes in PCG better.

Primary congenital glaucoma is the most common type of childhood glaucoma, but other forms of secondary glaucoma may occur. In these cases, congenital glaucoma may be the first symptom to appear, leading to the incorrect diagnosis of PCG. In addition, other non-glaucoma conditions may resemble some of the PCG clinical features. In these situations, in an attempt to obtain the most accurate diagnosis, genetic information can be an advantage not only to confirm the previous clinical diagnosis as also for helping to establish a new one. Also, genes usually involved in one type of glaucoma may contribute to the development of other types. A group of genes related to the most frequent disorders associated with congenital glaucoma, related to non-glaucoma conditions that may resemble PCG, and linked to other types of glaucoma was selected and evaluated by WES for the 13 patients previously studied (Chapter 5). An additional patient with other ocular and non-ocular symptoms was also included in this evaluation to confirm the secondary cause of congenital glaucoma and to test the utility and accuracy of the selected genes in a diagnosis. A novel *de novo FOXC1* mutation confirmed the

secondary cause of congenital glaucoma in patient 37 and the diagnosis of ARS. The remaining PCG patients did not reveal a strong candidate variant associated with their phenotype, except patient 25 with a possible causative variant in *ELP4* that has been associated with aniridia and other ocular anomalies as cataracts and glaucoma. Experimental validation is required to confirm the pathogenic effect of this variant, especially the predicted splicing disruption. The lack of strong candidates may suggest the absence of genetic disorders associated with secondary congenital glaucoma, implying a correct diagnosis of PCG. The search for pathogenic variants in these genes, and also in *CYP1B1*, can be advantageous particularly in young congenital glaucoma patients that have not developed all disease symptoms yet, or in which the disease diagnosis is difficult because of the lack of cooperation. A gene panel, or a virtual version of it, involving all the genes linked to glaucoma could be used, resulting in prompt diagnosis and allowing an adequate early treatment in other to minimize the severity of the disease.

Despite the utilization of a high-throughput technology, only patient 4 revealed a pathogenic variant in a candidate PCG gene. In patients 32, 36, 42 and 43 no rare and predicted pathogenic variants were detected. Similar to this study, a high number of samples without identification of the candidate gene is frequent, between 60 and 80% of the cases (Alodhayb et al., 2013; Sawyer et al., 2016; Seaby et al., 2016). This high number of patients that remain with unclear genetic associations may be related to factors such as technical causes, for example, a low or lack of coverage in the region of the causative mutation or a technical sequencing error. A good knowledge of the type of errors inherent to each technology and a control database of exomes sequenced with the same technology are essential in the distinction between a variant and a sequencing error. Another cause may be related to the origin of the variant. Variants can be located outside the coding regions not sequenced by WES or be due to CNVs and chromosome rearrangements not easily identified by WES. Somatic pathogenic variants only present in the affected tissues and not detectable in the DNA of blood cells may also complicate variant detection. Another factor that may contribute to unresolved genetic cases is the lack of information on gene function, the expression on tissues of interest or in embryonic development. Also, the distinction of rare pathogenic variants from rare polymorphisms is a challenge. Trio analysis is one way to circumvent this issue. This approach not only allows the *de novo* identification of variants but also leads to a significant reduction of the number of candidate variants.

In this study, the *in silico* functional impact software tools helped to predict the pathogenicity of the *TEK* variant. Although these tools score variants according to a set of assumptions and advanced algorithms, prediction accuracy still needs improvements (Kumar et al., 2014) and

the pathogenicity of the variants should be validated experimentally. On the contrary, the results observed for the *RNASEH2C* splice site mutation indicate that splice prediction is not straightforward. In fact, the splicing process involves multiple sites and regulatory factors that may not be taken into consideration by the *in silico* prediction methods. Additionally, splicing regulatory proteins can compete with degenerated binding sites and have conflicting effects, which may lead to less reliable predictions (Caminsky et al., 2014). For this reason *in silico* predictors should be improved considering the maximum of factors involved in the splicing process. Until then, experimental validation of the splice site variants is mandatory to accept a variant as having splicing consequences.

In conclusion, whole exome sequencing allowed the identification of the genetic cause in two individuals and contributed to expanding the spectrum of mutations associated with the diseases. One PCG patient had a novel *TEK* mutation, another patient with ARS and secondary congenital glaucoma carried a novel *de novo* mutation in *FOXC1*. The identification of *TEK* mutation in a PCG patient confirms the involvement of *TEK* in PCG pathogenesis, recently reported by Souma and collaborators (Souma et al., 2016). This study also concludes that other mechanisms must be involved in PCG development since not every patient revealed candidate genes for PCG. However, several possible candidate genes, some of them related to the TM, the SC and IOP, were identified. Future studies may link these genes to PCG pathogenesis, either by increasing the number of studied patients with variants in these genes ans by validation of the pathogenic impact of the variants. A group of genes associated with congenital glaucoma may have a positive impact helping the clinical diagnosis process to reach a fast and accurate diagnosis. In fact, genetic diagnosis is useful confirming the clinical suspicion, allowing the promptly adequate treatment necessary to avoid disease progression. Moreover, the genetic study allows evaluation of the risk of the genetic disease, essential for genetic counseling to families or pre-natal diagnosis. Above all, the identification of new genes and variants contributes to the comprehension of the molecular mechanisms behind PCG pathogenesis. The understanding of these mechanisms is expected to contribute to the development of new non-invasive and effective PCG therapeutic approaches allowing to preserve the vision of PCG patients all their lives.

References

Abu-Amero, K., Kondkar, A.A., Chalam, K.V., 2015. An Updated Review on the Genetics of Primary Open Angle Glaucoma. Int J Mol Sci. 16**,** 28886-911.

Abu-Amero, K.K., Osman, E.A., Mousa, A., Wheeler, J., Whigham, B., Allingham, R.R., Hauser, M.A., Al-Obeidan, S.A., 2011. Screening of CYP1B1 and LTBP2 genes in Saudi families with primary congenital glaucoma: genotype-phenotype correlation. Mol Vis. 17**,** 2911-9.

Acampora, D., Mazan, S., Avantaggiato, V., Barone, P., Tuorto, F., Lallemand, Y., Brulet, P., Simeone, A., 1996. Epilepsy and brain abnormalities in mice lacking the Otx1 gene. Nat Genet. 14**,** 218-22.

Acott, T.S., Kelley, M.J., 2008. Extracellular matrix in the trabecular meshwork. Exp Eye Res. 86**,** 543-61.

Acuna-Hidalgo, R., Veltman, J.A., Hoischen, A., 2016. New insights into the generation and role of de novo mutations in health and disease. Genome Biol. 17**,** 241.

Adzhubei, I., Jordan, D.M., Sunyaev, S.R., 2013. Predicting functional effect of human missense mutations using PolyPhen-2. Curr Protoc Hum Genet. Chapter 7**,** Unit7 20.

Akarsu, A.N., Turacli, M.E., Aktan, S.G., Barsoum-Homsy, M., Chevrette, L., Sayli, B.S., Sarfarazi, M., 1996. A second locus (GLC3B) for primary congenital glaucoma (Buphthalmos) maps to the 1p36 region. Hum Mol Genet. 5**,** 1199-203.

Aklillu, E., Oscarson, M., Hidestrand, M., Leidvik, B., Otter, C., Ingelman-Sundberg, M., 2002. Functional analysis of six different polymorphic CYP1B1 enzyme variants found in an Ethiopian population. Mol Pharmacol. 61**,** 586-94.

Aklillu, E., Ovrebo, S., Botnen, I.V., Otter, C., Ingelman-Sundberg, M., 2005. Characterization of common CYP1B1 variants with different capacity for benzo[a]pyrene-7,8-dihydrodiol epoxide formation from benzo[a]pyrene. Cancer Res. 65**,** 5105-11.

Albert, D., Miller, J., Azar, D., Blodi, B., 2008. Albert & Jakobiec's Principles & Practice of Ophthalmology, Vol. 2, Saunders.

Ali, M., McKibbin, M., Booth, A., Parry, D.A., Jain, P., Riazuddin, S.A., Hejtmancik, J.F., Khan, S.N., Firasat, S., Shires, M., Gilmour, D.F., Towns, K., Murphy, A.L., Azmanov, D., Tournev, I., Cherninkova, S., Jafri, H., Raashid, Y., Toomes, C., Craig, J., Mackey, D.A., Kalaydjieva, L., Riazuddin, S., Inglehearn, C.F., 2009. Null mutations in LTBP2 cause primary congenital glaucoma. Am J Hum Genet. 84**,** 664-71.

Allingham, R.R., Damji, K.F., Freedman, S.F., Moroi, S.E., Rhee, D.J., 2011. Shields Textbook of Glaucoma, Vol., Lippincott Williams & Wilkins.

Alodhayb, S., Morales, J., Edward, D.P., Abu-Amero, K.K., 2013. Update on pediatric glaucoma. Semin Ophthalmol. 28**,** 131-43.

Alsalem, J.A., Patel, D., Susarla, R., Coca-Prados, M., Bland, R., Walker, E.A., Rauz, S., Wallace, G.R., 2014. Characterization of vitamin D production by human ocular barrier cells. Invest Ophthalmol Vis Sci. 55**,** 2140-7.

Alward, W.L., Fingert, J.H., Coote, M.A., Johnson, A.T., Lerner, S.F., Junqua, D., Durcan, F.J., McCartney, P.J., Mackey, D.A., Sheffield, V.C., Stone, E.M., 1998. Clinical features associated with mutations in the chromosome 1 open-angle glaucoma gene (GLC1A). N Engl J Med. 338**,** 1022-7.

Anderson, D.R., 1981. The development of the trabecular meshwork and its abnormality in primary infantile glaucoma. Trans Am Ophthalmol Soc. 79**,** 458-85.

Ansorge, H.L., Meng, X., Zhang, G., Veit, G., Sun, M., Klement, J.F., Beason, D.P., Soslowsky, L.J., Koch, M., Birk, D.E., 2009. Type XIV Collagen Regulates Fibrillogenesis: PREMATURE

COLLAGEN FIBRIL GROWTH AND TISSUE DYSFUNCTION IN NULL MICE. J Biol Chem. 284, 8427-38.

Augustin, H.G., Koh, G.Y., Thurston, G., Alitalo, K., 2009. Control of vascular morphogenesis and homeostasis through the angiopoietin-Tie system. Nat Rev Mol Cell Biol. 10, 165-77.

Auton, A., Brooks, L.D., Durbin, R.M., Garrison, E.P., Kang, H.M., Korbel, J.O., Marchini, J.L., McCarthy, S., McVean, G.A., Abecasis, G.R., 2015. A global reference for human genetic variation. Nature. 526, 68-74.

Bamshad, M.J., Ng, S.B., Bigham, A.W., Tabor, H.K., Emond, M.J., Nickerson, D.A., Shendure, J., 2011. Exome sequencing as a tool for Mendelian disease gene discovery. Nat Rev Genet. 12, 745-55.

Banerjee, A., Chakraborty, S., Chakraborty, A., Chakrabarti, S., Ray, K., 2016. Functional and Structural Analyses of CYP1B1 Variants Linked to Congenital and Adult-Onset Glaucoma to Investigate the Molecular Basis of These Diseases. PLoS One. 11, e0156252.

Banerjee, J., Li, A., Ting Leung, C., Peterson-Yantorno, K., Stamer, W.D., Civan, M.M., 2012. Potential Role of Anoctamin-6 in Cell Volume Regulation of Human Trabecular Meshwork Cells. Invest Ophthalmol Vis Sci. 53, 5334.

Barrett, J.C., Fry, B., Maller, J., Daly, M.J., 2005. Haploview: analysis and visualization of LD and haplotype maps. Bioinformatics. 21, 263-5.

Barton, W.A., Tzvetkova-Robev, D., Miranda, E.P., Kolev, M.V., Rajashankar, K.R., Himanen, J.P., Nikolov, D.B., 2006. Crystal structures of the Tie2 receptor ectodomain and the angiopoietin-2-Tie2 complex. Nat Struct Mol Biol. 13, 524-32.

Bayat, B., Yazdani, S., Alavi, A., Chiani, M., Chitsazian, F., Tusi, B.K., Suri, F., Narooie-Nejhad, M., Sanati, M.H., Elahi, E., 2008. Contributions of MYOC and CYP1B1 mutations to JOAG. Mol Vis. 14, 508-17.

Bejjani, B.A., Stockton, D.W., Lewis, R.A., Tomey, K.F., Dueker, D.K., Jabak, M., Astle, W.F., Lupski, J.R., 2000. Multiple CYP1B1 mutations and incomplete penetrance in an inbred population segregating primary congenital glaucoma suggest frequent de novo events and a dominant modifier locus. Hum Mol Genet. 9, 367-74.

Belmouden, A., Melki, R., Hamdani, M., Zaghloul, K., Amraoui, A., Nadifi, S., Akhayat, O., Garchon, H.J., 2002. A novel frameshift founder mutation in the cytochrome P450 1B1 (CYP1B1) gene is associated with primary congenital glaucoma in Morocco. Clin Genet. 62, 334-9.

Bennett, T.M., Mackay, D.S., Siegfried, C.J., Shiels, A., 2014. Mutation of the melastatin-related cation channel, TRPM3, underlies inherited cataract and glaucoma. PLoS One. 9, e104000.

Berraho, A., Serrou, A., Fritez, N., El Annas, A., Bencherifa, F., Gaboun, F., Hilal, L., 2015. Genotype-phenotype correlation in Moroccan patients with primary congenital glaucoma. J Glaucoma. 24, 297-305.

Berry, F.B., Skarie, J.M., Mirzayans, F., Fortin, Y., Hudson, T.J., Raymond, V., Link, B.A., Walter, M.A., 2008. FOXC1 is required for cell viability and resistance to oxidative stress in the eye through the transcriptional regulation of FOXO1A. Hum Mol Genet. 17, 490-505.

Boland, J.F., Chung, C.C., Roberson, D., Mitchell, J., Zhang, X., Im, K.M., He, J., Chanock, S.J., Yeager, M., Dean, M., 2013. The new sequencer on the block: comparison of Life Technology's Proton sequencer to an Illumina HiSeq for whole-exome sequencing. Hum Genet. 132, 1153-63.

Borras, T., Bryant, P.A., Chisolm, S.S., 2006. First look at the effect of overexpression of TIGR/MYOC on the transcriptome of the human trabecular meshwork. Exp Eye Res. 82, 1002-10.

Brendel, V., Xing, L., Zhu, W., 2004. Gene structure prediction from consensus spliced alignment of multiple ESTs matching the same genomic locus. Bioinformatics. 20, 1157-69.

Brown, R.L., Xiong, W.H., Peters, J.H., Tekmen-Clark, M., Strycharska-Orczyk, I., Reed, B.T., Morgans, C.W., Duvoisin, R.M., 2015. TRPM3 expression in mouse retina. PLoS One. 10**,** e0117615.

Burian, H.M., 1960. A case of Marfan's syndrome with bilateral glaucoma. With description of a new type of operation for developmental glaucoma (trabeculotomy ab externo). Am J Ophthalmol. 50**,** 1187-92.

Burns, J.N., Turnage, K.C., Walker, C.A., Lieberman, R.L., 2011. The stability of myocilin olfactomedin domain variants provides new insight into glaucoma as a protein misfolding disorder. Biochemistry. 50**,** 5824-33.

Caminsky, N., Mucaki, E.J., Rogan, P.K., 2014. Interpretation of mRNA splicing mutations in genetic disease: review of the literature and guidelines for information-theoretical analysis. F1000Res. 3**,** 282.

Campos-Mollo, E., Lopez-Garrido, M.P., Blanco-Marchite, C., Garcia-Feijoo, J., Peralta, J., Belmonte-Martinez, J., Ayuso, C., Escribano, J., 2009. CYP1B1 mutations in Spanish patients with primary congenital glaucoma: phenotypic and functional variability. Mol Vis. 15**,** 417-31.

Capriotti, E., Fariselli, P., Casadio, R., 2005. I-Mutant2.0: predicting stability changes upon mutation from the protein sequence or structure. Nucleic Acids Res. 33**,** W306-10.

Cardoso, M.S., Anjos, R., Vieira, L., Ferreira, C., Xavier, A., Brito, C., 2015. CYP1B1 gene analysis and phenotypic correlation in Portuguese children with primary congenital glaucoma. Eur J Ophthalmol. 25**,** 474-7.

Carreon, T.A., Edwards, G., Wang, H., Bhattacharya, S.K., 2016. Segmental outflow of aqueous humor in mouse and human. Exp Eye Res.

Chakrabarti, S., Kaur, K., Kaur, I., Mandal, A.K., Parikh, R.S., Thomas, R., Majumder, P.P., 2006. Globally, CYP1B1 mutations in primary congenital glaucoma are strongly structured by geographic and haplotype backgrounds. Invest Ophthalmol Vis Sci. 47**,** 43-7.

Chakrabarti, S., Kaur, K., Rao, K.N., Mandal, A.K., Kaur, I., Parikh, R.S., Thomas, R., 2009. The transcription factor gene FOXC1 exhibits a limited role in primary congenital glaucoma. Invest Ophthalmol Vis Sci. 50**,** 75-83.

Chakrabarti, S., Ghanekar, Y., Kaur, K., Kaur, I., Mandal, A.K., Rao, K.N., Parikh, R.S., Thomas, R., Majumder, P.P., 2010. A polymorphism in the CYP1B1 promoter is functionally associated with primary congenital glaucoma. Hum Mol Genet. 19**,** 4083-90.

Chan, C., Li, W., Pang, C., Lam, D., Yew, D., 2007. Revisiting the Trabecular Meshwork in the Eyes of the Developing Human. Neuroembryol Aging. 4**,** 13-18.

Chang, E.E., Goldberg, J.L., 2012. Glaucoma 2.0: neuroprotection, neuroregeneration, neuroenhancement. Ophthalmology. 119**,** 979-86.

Chang, T.C., Summers, C.G., Schimmenti, L.A., Grajewski, A.L., 2012. Axenfeld-Rieger syndrome: new perspectives. Br J Ophthalmol. 96**,** 318-22.

Chatterjee, A., Villarreal, G., Jr., Rhee, D.J., 2014. Matricellular proteins in the trabecular meshwork: review and update. J Ocul Pharmacol Ther. 30**,** 447-63.

Chavarria-Soley, G., Michels-Rautenstrauss, K., Pasutto, F., Flikier, D., Flikier, P., Cirak, S., Bejjani, B., Winters, D.L., Lewis, R.A., Mardin, C., Reis, A., Rautenstrauss, B., 2006. Primary congenital glaucoma and Rieger's anomaly: extended haplotypes reveal founder effects for eight distinct CYP1B1 mutations. Mol Vis. 12**,** 523-31.

Chavarria-Soley, G., Sticht, H., Aklillu, E., Ingelman-Sundberg, M., Pasutto, F., Reis, A., Rautenstrauss, B., 2008. Mutations in CYP1B1 cause primary congenital glaucoma by reduction of either activity or abundance of the enzyme. Hum Mutat. 29**,** 1147-53.

Chen, L., Huang, L., Zeng, A., He, J., 2015. CYP1B1 gene mutations with incomplete penetrance in a Chinese pedigree with primary congenital glaucoma: a case report and review of literatures. Int J Clin Exp Med. 8**,** 14538-41.

Cheng, F., Song, W., Kang, Y., Yu, S., Yuan, H., 2011. A 556 kb deletion in the downstream region of the PAX6 gene causes familial aniridia and other eye anomalies in a Chinese family. Mol Vis. 17**,** 448-55.

Cheong, S.S., Hentschel, L., Davidson, A.E., Gerrelli, D., Davie, R., Rizzo, R., Pontikos, N., Plagnol, V., Moore, A.T., Sowden, J.C., Michaelides, M., Snead, M., Tuft, S.J., Hardcastle, A.J., 2016. Mutations in CPAMD8 Cause a Unique Form of Autosomal-Recessive Anterior Segment Dysgenesis. Am J Hum Genet. 99**,** 1338-1352.

Chun, S., Fay, J.C., 2009. Identification of deleterious mutations within three human genomes. Genome Res. 19**,** 1553-61.

Chung, S., Nakashima, M., Zembutsu, H., Nakamura, Y., 2011. Possible involvement of NEDD4 in keloid formation; its critical role in fibroblast proliferation and collagen production. Proc Jpn Acad Ser B Phys Biol Sci. 87**,** 563-73.

Colas-Tomas, T., Gutierrez-Diaz, E., Tejada-Palacios, P., Barcelo-Mendiguchia, A., Mencia-Gutierrez, E., 2010. Management of congenital glaucoma in neurofibromatosis type 1: a report of two cases. Int Ophthalmol. 30**,** 211-4.

Colomb, E., Kaplan, J., Garchon, H.J., 2003. Novel cytochrome P450 1B1 (CYP1B1) mutations in patients with primary congenital glaucoma in France. Hum Mutat. 22**,** 496.

Crow, Y.J., Leitch, A., Hayward, B.E., Garner, A., Parmar, R., Griffith, E., Ali, M., Semple, C., Aicardi, J., Babul-Hirji, R., Baumann, C., Baxter, P., Bertini, E., Chandler, K.E., Chitayat, D., Cau, D., Dery, C., Fazzi, E., Goizet, C., King, M.D., Klepper, J., Lacombe, D., Lanzi, G., Lyall, H., Martinez-Frias, M.L., Mathieu, M., McKeown, C., Monier, A., Oade, Y., Quarrell, O.W., Rittey, C.D., Rogers, R.C., Sanchis, A., Stephenson, J.B., Tacke, U., Till, M., Tolmie, J.L., Tomlin, P., Voit, T., Weschke, B., Woods, C.G., Lebon, P., Bonthron, D.T., Ponting, C.P., Jackson, A.P., 2006. Mutations in genes encoding ribonuclease H2 subunits cause Aicardi-Goutieres syndrome and mimic congenital viral brain infection. Nat Genet. 38**,** 910-6.

Crow, Y.J., Chase, D.S., Lowenstein Schmidt, J., Szynkiewicz, M., Forte, G.M., Gornall, H.L., Oojageer, A., Anderson, B., Pizzino, A., Helman, G., Abdel-Hamid, M.S., Abdel-Salam, G.M., Ackroyd, S., Aeby, A., Agosta, G., Albin, C., Allon-Shalev, S., Arellano, M., Ariaudo, G., Aswani, V., Babul-Hirji, R., Baildam, E.M., Bahi-Buisson, N., Bailey, K.M., Barnerias, C., Barth, M., Battini, R., Beresford, M.W., Bernard, G., Bianchi, M., Billette de Villemeur, T., Blair, E.M., Bloom, M., Burlina, A.B., Carpanelli, M.L., Carvalho, D.R., Castro-Gago, M., Cavallini, A., Cereda, C., Chandler, K.E., Chitayat, D.A., Collins, A.E., Sierra Corcoles, C., Cordeiro, N.J., Crichiutti, G., Dabydeen, L., Dale, R.C., D'Arrigo, S., De Goede, C.G., De Laet, C., De Waele, L.M., Denzler, I., Desguerre, I., Devriendt, K., Di Rocco, M., Fahey, M.C., Fazzi, E., Ferrie, C.D., Figueiredo, A., Gener, B., Goizet, C., Gowrinathan, N.R., Gowrishankar, K., Hanrahan, D., Isidor, B., Kara, B., Khan, N., King, M.D., Kirk, E.P., Kumar, R., Lagae, L., Landrieu, P., Lauffer, H., Laugel, V., La Piana, R., Lim, M.J., Lin, J.P., Linnankivi, T., Mackay, M.T., Marom, D.R., Marques Lourenco, C., McKee, S.A., Moroni, I., Morton, J.E., Moutard, M.L., Murray, K., Nabbout, R., Nampoothiri, S., Nunez-Enamorado, N., Oades, P.J., Olivieri, I., Ostergaard, J.R., Perez-Duenas, B., Prendiville, J.S., Ramesh, V., Rasmussen, M., Regal, L., Ricci, F., Rio, M., Rodriguez, D., Roubertie, A., Salvatici, E., Segers, K.A., Sinha, G.P., Soler, D., Spiegel, R., Stodberg, T.I., Straussberg, R., Swoboda, K.J., Suri, M., Tacke, U., Tan, T.Y., te Water Naude, J., Wee Teik, K., Thomas, M.M., Till, M., Tonduti, D., Valente, E.M., Van Coster, R.N., van der Knaap, M.S., Vassallo, G., Vijzelaar, R., Vogt, J., Wallace, G.B., Wassmer, E., Webb, H.J., Whitehouse, W.P., Whitney, R.N., Zaki, M.S., Zuberi, S.M., Livingston, J.H., Rozenberg, F., Lebon, P., Vanderver, A., Orcesi, S., Rice, G.I., 2015. Characterization of human disease phenotypes associated with mutations in TREX1, RNASEH2A, RNASEH2B, RNASEH2C, SAMHD1, ADAR, and IFIH1. Am J Med Genet A. 167A**,** 296-312.

Cummings, T.J., 2013. Ophthalmic Pathology: A Concise Guide, Vol., Springer, New York.

Dautriche, C.N., Tian, Y., Xie, Y., Sharfstein, S.T., 2015. A Closer Look at Schlemm's Canal Cell Physiology: Implications for Biomimetics. J Funct Biomater. 6**,** 963-85.

Davydov, E.V., Goode, D.L., Sirota, M., Cooper, G.M., Sidow, A., Batzoglou, S., 2010. Identifying a high fraction of the human genome to be under selective constraint using GERP++. PLoS Comput Biol. 6**,** e1001025.

De Groef, L., Van Hove, I., Dekeyster, E., Stalmans, I., Moons, L., 2013. MMPs in the trabecular meshwork: promising targets for future glaucoma therapies? Invest Ophthalmol Vis Sci. 54**,** 7756-63.

de Melo, M.B., Mandal, A.K., Tavares, I.M., Ali, M.H., Kabra, M., de Vasconcellos, J.P., Senthil, S., Sallum, J.M., Kaur, I., Betinjane, A.J., Moura, C.R., Paula, J.S., Costa, K.A., Sarfarazi, M., Paolera, M.D., Finzi, S., Ferraz, V.E., Costa, V.P., Belfort, R., Jr., Chakrabarti, S., 2015. Genotype-Phenotype Correlations in CYP1B1-Associated Primary Congenital Glaucoma Patients Representing Two Large Cohorts from India and Brazil. PLoS One. 10**,** e0127147.

Della Paolera, M., de Vasconcellos, J.P., Umbelino, C.C., Kasahara, N., Rocha, M.N., Richeti, F., Costa, V.P., Tavares, A., de Melo, M.B., 2010. CYP1B1 gene analysis in primary congenital glaucoma Brazilian patients: novel mutations and association with poor prognosis. J Glaucoma. 19**,** 176-82.

deLuise, V.P., Anderson, D.R., 1983. Primary infantile glaucoma (congenital glaucoma). Surv Ophthalmol. 28**,** 1-19.

Demirdas, S., Dulfer, E., Robert, L., Kempers, M., van Beek, D., Micha, D., van Engelen, B.G., Hamel, B., Schalkwijk, J., Loeys, B., Maugeri, A., Voermans, N.C., 2016. Recognizing the tenascin-X deficient type of Ehlers-Danlos syndrome: a cross-sectional study in 17 patients. Clin Genet.

Desmet, F.O., Hamroun, D., Lalande, M., Collod-Beroud, G., Claustres, M., Beroud, C., 2009. Human Splicing Finder: an online bioinformatics tool to predict splicing signals. Nucleic Acids Res. 37**,** e67.

Diskin, S., Kumar, J., Cao, Z., Schuman, J.S., Gilmartin, T., Head, S.R., Panjwani, N., 2006. Detection of differentially expressed glycogenes in trabecular meshwork of eyes with primary open-angle glaucoma. Invest Ophthalmol Vis Sci. 47**,** 1491-9.

Dogan, R.I., Getoor, L., Wilbur, W.J., Mount, S.M., 2007. SplicePort--an interactive splice-site analysis tool. Nucleic Acids Res. 35**,** W285-91.

Donegan, R.K., Hill, S.E., Freeman, D.M., Nguyen, E., Orwig, S.D., Turnage, K.C., Lieberman, R.L., 2015. Structural basis for misfolding in myocilin-associated glaucoma. Hum Mol Genet. 24**,** 2111-24.

Dong, C., Wei, P., Jian, X., Gibbs, R., Boerwinkle, E., Wang, K., Liu, X., 2015. Comparison and integration of deleteriousness prediction methods for nonsynonymous SNVs in whole exome sequencing studies. Hum Mol Genet. 24**,** 2125-37.

Dopazo, J., Amadoz, A., Bleda, M., Garcia-Alonso, L., Aleman, A., Garcia-Garcia, F., Rodriguez, J.A., Daub, J.T., Muntane, G., Rueda, A., Vela-Boza, A., Lopez-Domingo, F.J., Florido, J.P., Arce, P., Ruiz-Ferrer, M., Mendez-Vidal, C., Arnold, T.E., Spleiss, O., Alvarez-Tejado, M., Navarro, A., Bhattacharya, S.S., Borrego, S., Santoyo-Lopez, J., Antinolo, G., 2016. 267 Spanish Exomes Reveal Population-Specific Differences in Disease-Related Genetic Variation. Mol Biol Evol. 33**,** 1205-18.

Edward, D.P., Morales, J., Bouhenni, R.A., Patil, J., Edward, P.R., Cummings, T.J., Chaudhry, I.A., Alkatan, H., 2012. Congenital ectropion uvea and mechanisms of glaucoma in neurofibromatosis type 1: new insights. Ophthalmology. 119**,** 1485-94.

Elder, M.J., 1994. Combined trabeculotomy-trabeculectomy compared with primary trabeculectomy for congenital glaucoma. Br J Ophthalmol. 78**,** 745-8.

Ely, A.L., El-Dairi, M.A., Freedman, S.F., 2014. Cupping reversal in pediatric glaucoma--evaluation of the retinal nerve fiber layer and visual field. Am J Ophthalmol. 158**,** 905-15.

Fahsold, R., Hoffmeyer, S., Mischung, C., Gille, C., Ehlers, C., Kucukceylan, N., Abdel-Nour, M., Gewies, A., Peters, H., Kaufmann, D., Buske, A., Tinschert, S., Nurnberg, P., 2000. Minor lesion mutational spectrum of the entire NF1 gene does not explain its high mutability but points to a functional domain upstream of the GAP-related domain. Am J Hum Genet. 66**,** 790-818.

Faiq, M.A., Dada, R., Qadri, R., Dada, T., 2015. CYP1B1-mediated Pathobiology of Primary Congenital Glaucoma. J Curr Glaucoma Pract. 9**,** 77-80.

Faucher, M., Anctil, J.L., Rodrigue, M.A., Duchesne, A., Bergeron, D., Blondeau, P., Cote, G., Dubois, S., Bergeron, J., Arseneault, R., Morissette, J., Raymond, V., 2002. Founder TIGR/myocilin mutations for glaucoma in the Quebec population. Hum Mol Genet. 11**,** 2077-90.

Fingert, J.H., Heon, E., Liebmann, J.M., Yamamoto, T., Craig, J.E., Rait, J., Kawase, K., Hoh, S.T., Buys, Y.M., Dickinson, J., Hockey, R.R., Williams-Lyn, D., Trope, G., Kitazawa, Y., Ritch, R., Mackey, D.A., Alward, W.L., Sheffield, V.C., Stone, E.M., 1999. Analysis of myocilin mutations in 1703 glaucoma patients from five different populations. Hum Mol Genet. 8**,** 899-905.

Foley, A.R., Hu, Y., Zou, Y., Columbus, A., Shoffner, J., Dunn, D.M., Weiss, R.B., Bonnemann, C.G., 2009. Autosomal recessive inheritance of classic Bethlem myopathy. Neuromuscul Disord. 19**,** 813-7.

Frebourg, T., 2014. The challenge for the next generation of medical geneticists. Hum Mutat. 35**,** 909-11.

Fukuhara, S., Sako, K., Minami, T., Noda, K., Kim, H.Z., Kodama, T., Shibuya, M., Takakura, N., Koh, G.Y., Mochizuki, N., 2008. Differential function of Tie2 at cell-cell contacts and cell-substratum contacts regulated by angiopoietin-1. Nat Cell Biol. 10**,** 513-26.

Fukuhara, S., Sako, K., Noda, K., Zhang, J., Minami, M., Mochizuki, N., 2010. Angiopoietin-1/Tie2 receptor signaling in vascular quiescence and angiogenesis. Histol Histopathol. 25**,** 387-96.

Fuse, N., Miyazawa, A., Takahashi, K., Noro, M., Nakazawa, T., Nishida, K., 2010. Mutation spectrum of the CYP1B1 gene for congenital glaucoma in the Japanese population. Jpn J Ophthalmol. 54**,** 1-6.

Ganesh, A., Mai, D.T., Levin, A.V., 2013. Pediatric glaucoma terminology. Am J Med Genet A. 161A**,** 3205-15.

Gao, K., Masuda, A., Matsuura, T., Ohno, K., 2008. Human branch point consensus sequence is yUnAy. Nucleic Acids Res. 36**,** 2257-67.

Gasten, A.C., Ramdas, W.D., Broer, L., van Koolwijk, L.M., Ikram, M.K., de Jong, P.T., Aulchenko, Y.S., Wolfs, R.C., Hofman, A., Rivadeneira, F., Uitterlinden, A.G., Oostra, B.A., Lemij, H.G., Klaver, C.C., Jansonius, N.M., Vingerling, J.R., van Duijn, C.M., 2012. A genetic epidemiologic study of candidate genes involved in the optic nerve head morphology. Invest Ophthalmol Vis Sci. 53**,** 1485-91.

Gearhart, M.D., Corcoran, C.M., Wamstad, J.A., Bardwell, V.J., 2006. Polycomb group and SCF ubiquitin ligases are found in a novel BCOR complex that is recruited to BCL6 targets. Mol Cell Biol. 26**,** 6880-9.

Geetha, T.S., Michealraj, K.A., Kabra, M., Kaur, G., Juyal, R.C., Thelma, B.K., 2014. Targeted deep resequencing identifies MID2 mutation for X-linked intellectual disability with varied disease severity in a large kindred from India. Hum Mutat. 35**,** 41-4.

Gerometta, R., Spiga, M.G., Borras, T., Candia, O.A., 2010. Treatment of sheep steroid-induced ocular hypertension with a glucocorticoid-inducible MMP1 gene therapy virus. Invest Ophthalmol Vis Sci. 51**,** 3042-8.

Gilissen, C., Hoischen, A., Brunner, H.G., Veltman, J.A., 2012. Disease gene identification strategies for exome sequencing. Eur J Hum Genet. 20**,** 490-7.

Gobeil, S., Rodrigue, M.A., Moisan, S., Nguyen, T.D., Polansky, J.R., Morissette, J., Raymond, V., 2004. Intracellular sequestration of hetero-oligomers formed by wild-type and glaucoma-causing myocilin mutants. Invest Ophthalmol Vis Sci. 45**,** 3560-7.

Goel, M., Picciani, R.G., Lee, R.K., Bhattacharya, S.K., 2010. Aqueous humor dynamics: a review. Open Ophthalmol J. 4**,** 52-9.

Gould, D.B., Smith, R.S., John, S.W., 2004. Anterior segment development relevant to glaucoma. Int J Dev Biol. 48**,** 1015-29.

Grimm, C., Kraft, R., Sauerbruch, S., Schultz, G., Harteneck, C., 2003. Molecular and functional characterization of the melastatin-related cation channel TRPM3. J Biol Chem. 278**,** 21493-501.

Gubbiotti, M.A., Neill, T., Iozzo, R.V., 2016. A current view of perlecan in physiology and pathology: A mosaic of functions. Matrix Biol.

Gunther, C., Kind, B., Reijns, M.A., Berndt, N., Martinez-Bueno, M., Wolf, C., Tungler, V., Chara, O., Lee, Y.A., Hubner, N., Bicknell, L., Blum, S., Krug, C., Schmidt, F., Kretschmer, S., Koss, S., Astell, K.R., Ramantani, G., Bauerfeind, A., Morris, D.L., Cunninghame Graham, D.S., Bubeck, D., Leitch, A., Ralston, S.H., Blackburn, E.A., Gahr, M., Witte, T., Vyse, T.J., Melchers, I., Mangold, E., Nothen, M.M., Aringer, M., Kuhn, A., Luthke, K., Unger, L., Bley, A., Lorenzi, A., Isaacs, J.D., Alexopoulou, D., Conrad, K., Dahl, A., Roers, A., Alarcon-Riquelme, M.E., Jackson, A.P., Lee-Kirsch, M.A., 2015. Defective removal of ribonucleotides from DNA promotes systemic autoimmunity. J Clin Invest. 125**,** 413-24.

Gupta, V., James, M.K., Singh, A., Kumar, S., Gupta, S., Sharma, A., Sihota, R., Kennedy, D.J., 2016. Differences in Optic Disc Characteristics of Primary Congenital Glaucoma, Juvenile, and Adult Onset Open Angle Glaucoma Patients. J Glaucoma. 25**,** 239-43.

Hamanaka, T., Bill, A., Ichinohasama, R., Ishida, T., 1994. The Development of Sehlemm's Canal from Aspects of the Aqueous Humor Drainage. Med Electron Microsc. 27**,** 282-85.

Hilton, E., Johnston, J., Whalen, S., Okamoto, N., Hatsukawa, Y., Nishio, J., Kohara, H., Hirano, Y., Mizuno, S., Torii, C., Kosaki, K., Manouvrier, S., Boute, O., Perveen, R., Law, C., Moore, A., Fitzpatrick, D., Lemke, J., Fellmann, F., Debray, F.G., Dastot-Le-Moal, F., Gerard, M., Martin, J., Bitoun, P., Goossens, M., Verloes, A., Schinzel, A., Bartholdi, D., Bardakjian, T., Hay, B., Jenny, K., Johnston, K., Lyons, M., Belmont, J.W., Biesecker, L.G., Giurgea, I., Black, G., 2009. BCOR analysis in patients with OFCD and Lenz microphthalmia syndromes, mental retardation with ocular anomalies, and cardiac laterality defects. Eur J Hum Genet. 17**,** 1325-35.

Hilton, E.N., Manson, F.D., Urquhart, J.E., Johnston, J.J., Slavotinek, A.M., Hedera, P., Stattin, E.L., Nordgren, A., Biesecker, L.G., Black, G.C., 2007. Left-sided embryonic expression of the BCL-6 corepressor, BCOR, is required for vertebrate laterality determination. Hum Mol Genet. 16**,** 1773-82.

Hirst, J., Lui, W.W., Bright, N.A., Totty, N., Seaman, M.N., Robinson, M.S., 2000. A family of proteins with gamma-adaptin and VHS domains that facilitate trafficking between the trans-Golgi network and the vacuole/lysosome. J Cell Biol. 149**,** 67-80.

Hochstrasser, M., 1992. Ubiquitin and intracellular protein degradation. Curr Opin Cell Biol. 4**,** 1024-31.

Hoguet, A., Grajewski, A., Hodapp, E., Chang, T.C., 2016. A retrospective survey of childhood glaucoma prevalence according to Childhood Glaucoma Research Network classification. Indian J Ophthalmol. 64**,** 118-23.

Hollander, D.A., Sarfarazi, M., Stoilov, I., Wood, I.S., Fredrick, D.R., Alvarado, J.A., 2006. Genotype and phenotype correlations in congenital glaucoma. Trans Am Ophthalmol Soc. 104**,** 183-95.

Hoyt, C., Taylor, D., 2013. Pediatric Ophthalmology and Strabismus, Vol., Saunders.

Huynh, K.D., Fischle, W., Verdin, E., Bardwell, V.J., 2000. BCoR, a novel corepressor involved in BCL-6 repression. Genes Dev. 14**,** 1810-23.

Ito, Y.A., Goping, I.S., Berry, F., Walter, M.A., 2014. Dysfunction of the stress-responsive FOXC1 transcription factor contributes to the earlier-onset glaucoma observed in Axenfeld-Rieger syndrome patients. Cell Death Dis. 5**,** e1069.

Janssen, S.F., Gorgels, T.G., Bossers, K., Ten Brink, J.B., Essing, A.H., Nagtegaal, M., van der Spek, P.J., Jansonius, N.M., Bergen, A.A., 2012. Gene expression and functional annotation of the human ciliary body epithelia. PLoS One. 7**,** e44973.

Jeansson, M., Gawlik, A., Anderson, G., Li, C., Kerjaschki, D., Henkelman, M., Quaggin, S.E., 2011. Angiopoietin-1 is essential in mouse vasculature during development and in response to injury. J Clin Invest. 121**,** 2278-89.

Johnson, M., McLaren, J.W., Overby, D.R., 2016. Unconventional aqueous humor outflow: A review. Exp Eye Res.

Karali, M., Peluso, I., Marigo, V., Banfi, S., 2007. Identification and characterization of microRNAs expressed in the mouse eye. Invest Ophthalmol Vis Sci. 48**,** 509-15.

Kaur, K., Reddy, A.B., Mukhopadhyay, A., Mandal, A.K., Hasnain, S.E., Ray, K., Thomas, R., Balasubramanian, D., Chakrabarti, S., 2005. Myocilin gene implicated in primary congenital glaucoma. Clin Genet. 67**,** 335-40.

Kaur, S., Kaushik, S., Pandav, S.S., 2014. Glaucoma in Childhood. Delhi J Ophthalmol. 24**,** 160-6.

Keller, K.E., Kelley, M.J., Acott, T.S., 2007. Extracellular matrix gene alternative splicing by trabecular meshwork cells in response to mechanical stretching. Invest Ophthalmol Vis Sci. 48**,** 1164-72.

Keller, K.E., Bradley, J.M., Kelley, M.J., Acott, T.S., 2008. Effects of modifiers of glycosaminoglycan biosynthesis on outflow facility in perfusion culture. Invest Ophthalmol Vis Sci. 49**,** 2495-505.

Keller, K.E., Acott, T.S., 2013. The Juxtacanalicular Region of Ocular Trabecular Meshwork: A Tissue with a Unique Extracellular Matrix and Specialized Function. J Ocul Biol. 1**,** 3.

Keller, K.E., Vranka, J.A., Haddadin, R.I., Kang, M.H., Oh, D.J., Rhee, D.J., Yang, Y.F., Sun, Y.Y., Kelley, M.J., Acott, T.S., 2013. The effects of tenascin C knockdown on trabecular meshwork outflow resistance. Invest Ophthalmol Vis Sci. 54**,** 5613-23.

Kennedy, K.D., AnithaChristy, S.A., Buie, L.K., Borras, T., 2012. Cystatin a, a potential common link for mutant myocilin causative glaucoma. PLoS One. 7**,** e36301.

Khalifa, O., Imtiaz, F., Ramzan, K., Allam, R., Hemidan, A.A., Faqeih, E., Abuharb, G., Balobaid, A., Sakati, N., Owain, M.A., 2014. Marshall syndrome: further evidence of a distinct phenotypic entity and report of new findings. Am J Med Genet A. 164A**,** 2601-6.

Khalil, D.H., Abdelhakim, M.A., 2016. Primary trabeculotomy compared to combined trabeculectomy-trabeculotomy in congenital glaucoma: 3-year study. Acta Ophthalmol.

Khan, A.O., 2011. Conditions that can be mistaken as early childhood glaucoma. Ophthalmic Genet. 32**,** 129-37.

Kidson, S.H., Kume, T., Deng, K., Winfrey, V., Hogan, B.L., 1999. The forkhead/winged-helix gene, Mf1, is necessary for the normal development of the cornea and formation of the anterior chamber in the mouse eye. Dev Biol. 211**,** 306-22.

Kircher, M., Witten, D.M., Jain, P., O'Roak, B.J., Cooper, G.M., Shendure, J., 2014. A general framework for estimating the relative pathogenicity of human genetic variants. Nat Genet. 46**,** 310-5.

Kizhatil, K., Ryan, M., Marchant, J.K., Henrich, S., John, S.W., 2014. Schlemm's canal is a unique vessel with a combination of blood vascular and lymphatic phenotypes that forms by a novel developmental process. PLoS Biol. 12**,** e1001912.

Kondoh, T., Okamoto, N., Norimatsu, N., Uetani, M., Nishimura, G., Moriuchi, H., 2007. A Japanese case of oto-palato-digital syndrome type II: an apparent lack of phenotype-genotype correlation. J Hum Genet. 52**,** 370-3.

Kong, L., Fry, M., Al-Samarraie, M., Gilbert, C., Steinkuller, P.G., 2012. An update on progress and the changing epidemiology of causes of childhood blindness worldwide. J AAPOS. 16**,** 501-7.

Kresak, J.L., Walsh, M., 2016. Neurofibromatosis: A Review of NF1, NF2, and Schwannomatosis. J Pediatr Genet. 5**,** 98-104.

Krumbiegel, M., Pasutto, F., Schlotzer-Schrehardt, U., Uebe, S., Zenkel, M., Mardin, C.Y., Weisschuh, N., Paoli, D., Gramer, E., Becker, C., Ekici, A.B., Weber, B.H., Nurnberg, P., Kruse, F.E., Reis, A., 2011. Genome-wide association study with DNA pooling identifies variants at CNTNAP2 associated with pseudoexfoliation syndrome. Eur J Hum Genet. 19**,** 186-93.

Kumar, A., Rajendran, V., Sethumadhavan, R., Shukla, P., Tiwari, S., Purohit, R., 2014. Computational SNP analysis: current approaches and future prospects. Cell Biochem Biophys. 68**,** 233-9.

Kumar, S., Malik, M.A., Goswami, S., Sihota, R., Kaur, J., 2016. Candidate genes involved in the susceptibility of primary open angle glaucoma. Gene. 577**,** 119-31.

Larsen, K.B., Lutterodt, M., Rath, M.F., Moller, M., 2009. Expression of the homeobox genes PAX6, OTX2, and OTX1 in the early human fetal retina. Int J Dev Neurosci. 27**,** 485-92.

Larsen, W., 1993. Human Embryology, Vol., Churchill Livingstone Inc., New York.

Lauderdale, J.D., Wilensky, J.S., Oliver, E.R., Walton, D.S., Glaser, T., 2000. 3' deletions cause aniridia by preventing PAX6 gene expression. Proc Natl Acad Sci U S A. 97**,** 13755-9.

Lek, M., Karczewski, K.J., Minikel, E.V., Samocha, K.E., Banks, E., Fennell, T., O'Donnell-Luria, A.H., Ware, J.S., Hill, A.J., Cummings, B.B., Tukiainen, T., Birnbaum, D.P., Kosmicki, J.A., Duncan, L.E., Estrada, K., Zhao, F., Zou, J., Pierce-Hoffman, E., Berghout, J., Cooper, D.N., Deflaux, N., DePristo, M., Do, R., Flannick, J., Fromer, M., Gauthier, L., Goldstein, J., Gupta, N., Howrigan, D., Kiezun, A., Kurki, M.I., Moonshine, A.L., Natarajan, P., Orozco, L., Peloso, G.M., Poplin, R., Rivas, M.A., Ruano-Rubio, V., Rose, S.A., Ruderfer, D.M., Shakir, K., Stenson, P.D., Stevens, C., Thomas, B.P., Tiao, G., Tusie-Luna, M.T., Weisburd, B., Won, H.H., Yu, D., Altshuler, D.M., Ardissino, D., Boehnke, M., Danesh, J., Donnelly, S., Elosua, R., Florez, J.C., Gabriel, S.B., Getz, G., Glatt, S.J., Hultman, C.M., Kathiresan, S., Laakso, M., McCarroll, S., McCarthy, M.I., McGovern, D., McPherson, R., Neale, B.M., Palotie, A., Purcell, S.M., Saleheen, D., Scharf, J.M., Sklar, P., Sullivan, P.F., Tuomilehto, J., Tsuang, M.T., Watkins, H.C., Wilson, J.G., Daly, M.J., MacArthur, D.G., 2016. Analysis of protein-coding genetic variation in 60,706 humans. Nature. 536**,** 285-91.

Li, H., Durbin, R., 2010. Fast and accurate long-read alignment with Burrows-Wheeler transform. Bioinformatics. 26**,** 589-95.

Li, H., Liu, T., Chen, X., Xie, L., 2015. A rare case of primary congenital glaucoma in combination with neurofibromatosis 1: a case report. BMC Ophthalmol. 15**,** 149.

Li, J., Gao, B., Xiao, X., Li, S., Jia, X., Sun, W., Guo, X., Zhang, Q., 2016. Exome sequencing identified null mutations in LOXL3 associated with early-onset high myopia. Mol Vis. 22**,** 161-7.

Li, N., Zhou, Y., Du, L., Wei, M., Chen, X., 2011. Overview of Cytochrome P450 1B1 gene mutations in patients with primary congenital glaucoma. Exp Eye Res. 93**,** 572-9.

Libby, R.T., Smith, R.S., Savinova, O.V., Zabaleta, A., Martin, J.E., Gonzalez, F.J., John, S.W., 2003. Modification of ocular defects in mouse developmental glaucoma models by tyrosinase. Science. 299**,** 1578-81.

Liliom, K., Guan, Z., Tseng, J.L., Desiderio, D.M., Tigyi, G., Watsky, M.A., 1998. Growth factor-like phospholipids generated after corneal injury. Am J Physiol. 274**,** C1065-74.

Lister, A., 1965. Technique of goniotomy. Br J Ophthalmol. 49**,** 594-9.

Liton, P.B., Luna, C., Challa, P., Epstein, D.L., Gonzalez, P., 2006. Genome-wide expression profile of human trabecular meshwork cultured cells, nonglaucomatous and primary open angle glaucoma tissue. Mol Vis. 12**,** 774-90.

Livingston, J.H., Crow, Y.J., 2016. Neurologic Phenotypes Associated with Mutations in TREX1, RNASEH2A, RNASEH2B, RNASEH2C, SAMHD1, ADAR1, and IFIH1: Aicardi-Goutieres Syndrome and Beyond. Neuropediatrics. 47**,** 355-360.

Lopez-Garrido, M.P., Medina-Trillo, C., Morales-Fernandez, L., Garcia-Feijoo, J., Martinez-de-la-Casa, J.M., Garcia-Anton, M., Escribano, J., 2013. Null CYP1B1 genotypes in primary congenital and nondominant juvenile glaucoma. Ophthalmology. 120**,** 716-23.

Lopez-Martinez, F., Lopez-Garrido, M.P., Sanchez-Sanchez, F., Campos-Mollo, E., Coca-Prados, M., Escribano, J., 2007. Role of MYOC and OPTN sequence variations in Spanish patients with primary open-angle glaucoma. Mol Vis. 13**,** 862-72.

Lutjen-Drecoll, E., 1999. Functional morphology of the trabecular meshwork in primate eyes. Prog Retin Eye Res. 18**,** 91-119.

MacArthur, D.G., Manolio, T.A., Dimmock, D.P., Rehm, H.L., Shendure, J., Abecasis, G.R., Adams, D.R., Altman, R.B., Antonarakis, S.E., Ashley, E.A., Barrett, J.C., Biesecker, L.G., Conrad, D.F., Cooper, G.M., Cox, N.J., Daly, M.J., Gerstein, M.B., Goldstein, D.B., Hirschhorn, J.N., Leal, S.M., Pennacchio, L.A., Stamatoyannopoulos, J.A., Sunyaev, S.R., Valle, D., Voight, B.F., Winckler, W., Gunter, C., 2014. Guidelines for investigating causality of sequence variants in human disease. Nature. 508**,** 469-76.

Mackay, D.S., Bennett, T.M., Shiels, A., 2015. Exome Sequencing Identifies a Missense Variant in EFEMP1 Co-Segregating in a Family with Autosomal Dominant Primary Open-Angle Glaucoma. PLoS One. 10**,** e0132529.

Mandal, A.K., Chakrabarti, D., 2011. Update on congenital glaucoma. Indian J Ophthalmol. 59 Suppl**,** S148-57.

Mantravadi, A.V., Vadhar, N., 2015. Glaucoma. Prim Care. 42**,** 437-49.

Marchini, G., Toscani, M., Chemello, F., 2014. Pediatric glaucoma: current perspectives. Pediatric Health Med Ther. 5**,** 15-27.

Maugeri, A., van Driel, M.A., van de Pol, D.J., Klevering, B.J., van Haren, F.J., Tijmes, N., Bergen, A.A., Rohrschneider, K., Blankenagel, A., Pinckers, A.J., Dahl, N., Brunner, H.G., Deutman, A.F., Hoyng, C.B., Cremers, F.P., 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. Am J Hum Genet. 64**,** 1024-35.

Maugeri, A., Flothmann, K., Hemmrich, N., Ingvast, S., Jorge, P., Paloma, E., Patel, R., Rozet, J.M., Tammur, J., Testa, F., Balcells, S., Bird, A.C., Brunner, H.G., Hoyng, C.B., Metspalu, A., Simonelli, F., Allikmets, R., Bhattacharya, S.S., D'Urso, M., Gonzalez-Duarte, R., Kaplan, J., te Meerman, G.J., Santos, R., Schwartz, M., Van Camp, G., Wadelius, C., Weber, B.H., Cremers, F.P., 2002. The ABCA4 2588G>C Stargardt mutation: single origin and increasing frequency from South-West to North-East Europe. Eur J Hum Genet. 10**,** 197-203.

Maul, E., Strozzi, L., Munoz, C., Reyes, C., 1980. The outflow pathway in congenital glaucoma. Am J Ophthalmol. 89**,** 667-73.

Maumenee, A.E., 1958. The pathogenesis of congenital glaucoma: a new theory. Trans Am Ophthalmol Soc. 56**,** 507-70.

Mauri, L., Uebe, S., Sticht, H., Vossmerbaeumer, U., Weisschuh, N., Manfredini, E., Maselli, E., Patrosso, M., Weinreb, R.N., Penco, S., Reis, A., Pasutto, F., 2016. Expanding the clinical spectrum of COL1A1 mutations in different forms of glaucoma. Orphanet J Rare Dis. 11**,** 108.

Mbagwu, M., Rahmani, B., Srivastava, A., Burrowes, D., Bryar, P.J., 2015. Unsuspected Ganglioneuroma of the Choroid Diagnosed after Enucleation. Ocul Oncol Pathol. 2**,** 48-50.

McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., Garimella, K., Altshuler, D., Gabriel, S., Daly, M., DePristo, M.A., 2010. The Genome Analysis Toolkit:

a MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res. 20**,** 1297-303.

Medina-Trillo, C., Sanchez-Sanchez, F., Aroca-Aguilar, J.D., Ferre-Fernandez, J.J., Morales, L., Mendez-Hernandez, C.D., Blanco-Kelly, F., Ayuso, C., Garcia-Feijoo, J., Escribano, J., 2015. Hypo- and hypermorphic FOXC1 mutations in dominant glaucoma: transactivation and phenotypic variability. PLoS One. 10**,** e0119272.

Medina-Trillo, C., Ferre-Fernandez, J.J., Aroca-Aguilar, J.D., Bonet-Fernandez, J.M., Escribano, J., 2016. Functional characterization of eight rare missense CYP1B1 variants involved in congenital glaucoma and their association with null genotypes. Acta Ophthalmol. 94**,** e555-e560.

Meghpara, B., Li, X., Nakamura, H., Khan, A., Bejjani, B.A., Lin, S., Edward, D.P., 2008. Human anterior chamber angle development without cell death or macrophage involvement. Mol Vis. 14**,** 2492-8.

Melani, M., Simpson, K.J., Brugge, J.S., Montell, D., 2008. Regulation of cell adhesion and collective cell migration by hindsight and its human homolog RREB1. Curr Biol. 18**,** 532-7.

Merlini, L., Martoni, E., Grumati, P., Sabatelli, P., Squarzoni, S., Urciuolo, A., Ferlini, A., Gualandi, F., Bonaldo, P., 2008. Autosomal recessive myosclerosis myopathy is a collagen VI disorder. Neurology. 71**,** 1245-53.

Mettu, P.S., Deng, P.F., Misra, U.K., Gawdi, G., Epstein, D.L., Rao, P.V., 2004. Role of lysophospholipid growth factors in the modulation of aqueous humor outflow facility. Invest Ophthalmol Vis Sci. 45**,** 2263-71.

Micheal, S., Siddiqui, S.N., Zafar, S.N., Iqbal, A., Khan, M.I., den Hollander, A.I., 2016. Identification of Novel Variants in LTBP2 and PXDN Using Whole-Exome Sequencing in Developmental and Congenital Glaucoma. PLoS One. 11**,** e0159259.

Milla, E., Mane, B., Duch, S., Hernan, I., Borras, E., Planas, E., Dias Mde, S., Carballo, M., Gamundi, M.J., 2013. Survey of familial glaucoma shows a high incidence of cytochrome P450, family 1, subfamily B, polypeptide 1 (CYP1B1) mutations in non-consanguineous congenital forms in a Spanish population. Mol Vis. 19**,** 1707-22.

Mizokami, K., Sugiura, T., San Juan Jr, R., 1994. The Development of Human Trabecular Meshwork. Med Electron Microsc. 27**,** 277-81.

Mochizuki, H., Lesley, A.G., Brandt, J.D., 2011. Shrinkage of the scleral canal during cupping reversal in children. Ophthalmology. 118**,** 2008-13.

Monemi, S., Spaeth, G., DaSilva, A., Popinchalk, S., Ilitchev, E., Liebmann, J., Ritch, R., Heon, E., Crick, R.P., Child, A., Sarfarazi, M., 2005. Identification of a novel adult-onset primary open-angle glaucoma (POAG) gene on 5q22.1. Hum Mol Genet. 14**,** 725-33.

Moore, K., Persaud, T., 2008. The Developing Human: Clinically Oriented Embryology, Vol., Saunders, Elsevier Inc, Philadelphia, PA.

Morgan, J.T., Wood, J.A., Walker, N.J., Raghunathan, V.K., Borjesson, D.L., Murphy, C.J., Russell, P., 2014. Human trabecular meshwork cells exhibit several characteristics of, but are distinct from, adipose-derived mesenchymal stem cells. J Ocul Pharmacol Ther. 30**,** 254-66.

Mukhopadhyay, N.K., Cinar, B., Mukhopadhyay, L., Lutchman, M., Ferdinand, A.S., Kim, J., Chung, L.W., Adam, R.M., Ray, S.K., Leiter, A.B., Richie, J.P., Liu, B.C., Freeman, M.R., 2007. The zinc finger protein ras-responsive element binding protein-1 is a coregulator of the androgen receptor: implications for the role of the Ras pathway in enhancing androgenic signaling in prostate cancer. Mol Endocrinol. 21**,** 2056-70.

Nakano, M., Lockhart, C.M., Kelly, E.J., Rettie, A.E., 2014. Ocular cytochrome P450s and transporters: roles in disease and endobiotic and xenobiotic disposition. Drug Metab Rev. 46**,** 247-60.

Narooie-Nejad, M., Paylakhi, S.H., Shojaee, S., Fazlali, Z., Rezaei Kanavi, M., Nilforushan, N., Yazdani, S., Babrzadeh, F., Suri, F., Ronaghi, M., Elahi, E., Paisan-Ruiz, C., 2009. Loss of

function mutations in the gene encoding latent transforming growth factor beta binding protein 2, LTBP2, cause primary congenital glaucoma. Hum Mol Genet. 18**,** 3969-77.

Ng, P.C., Henikoff, S., 2003. SIFT: Predicting amino acid changes that affect protein function. Nucleic Acids Res. 31**,** 3812-4.

Nishiguchi, K.M., Tearle, R.G., Liu, Y.P., Oh, E.C., Miyake, N., Benaglio, P., Harper, S., Koskiniemi-Kuendig, H., Venturini, G., Sharon, D., Koenekoop, R.K., Nakamura, M., Kondo, M., Ueno, S., Yasuma, T.R., Beckmann, J.S., Ikegawa, S., Matsumoto, N., Terasaki, H., Berson, E.L., Katsanis, N., Rivolta, C., 2013. Whole genome sequencing in patients with retinitis pigmentosa reveals pathogenic DNA structural changes and NEK2 as a new disease gene. Proc Natl Acad Sci U S A. 110**,** 16139-44.

Oystreck, D.T., Morales, J., Chaudhry, I., Alorainy, I.A., Elkhamary, S.M., Pasha, T.M., Bosley, T.M., 2012. Visual loss in orbitofacial neurofibromatosis type 1. Ophthalmology. 119**,** 2168-73.

Panicker, S.G., Mandal, A.K., Reddy, A.B., Gothwal, V.K., Hasnain, S.E., 2004. Correlations of genotype with phenotype in Indian patients with primary congenital glaucoma. Invest Ophthalmol Vis Sci. 45**,** 1149-56.

Pasutto, F., Mauri, L., Popp, B., Sticht, H., Ekici, A., Piozzi, E., Bonfante, A., Penco, S., Schlotzer-Schrehardt, U., Reis, A., 2015. Whole exome sequencing reveals a novel de novo FOXC1 mutation in a patient with unrecognized Axenfeld-Rieger syndrome and glaucoma. Gene. 568**,** 76-80.

Pattabiraman, P.P., Rao, P.V., 2010. Mechanistic basis of Rho GTPase-induced extracellular matrix synthesis in trabecular meshwork cells. Am J Physiol Cell Physiol. 298**,** C749-63.

Pattabiraman, P.P., Pecen, P.E., Rao, P.V., 2013. MRP4-mediated regulation of intracellular cAMP and cGMP levels in trabecular meshwork cells and homeostasis of intraocular pressure. Invest Ophthalmol Vis Sci. 54**,** 1636-49.

Pavlakis, E., Chiotaki, R., Chalepakis, G., 2011. The role of Fras1/Frem proteins in the structure and function of basement membrane. Int J Biochem Cell Biol. 43**,** 487-95.

Payne, M.S., Nadell, J.M., Lacassie, Y., Tilton, A.H., 2003. Congenital glaucoma and neurofibromatosis in a monozygotic twin: case report and review of the literature. J Child Neurol. 18**,** 504-8.

Perry, L.P., Jakobiec, F.A., Zakka, F.R., Walton, D.S., 2012. Newborn primary congenital glaucoma: histopathologic features of the anterior chamber filtration angle. J AAPOS. 16**,** 565-8.

Phng, L.K., Gerhardt, H., 2009. Angiogenesis: a team effort coordinated by notch. Dev Cell. 16**,** 196-208.

Plasilova, M., Stoilov, I., Sarfarazi, M., Kadasi, L., Ferakova, E., Ferak, V., 1999. Identification of a single ancestral CYP1B1 mutation in Slovak Gypsies (Roms) affected with primary congenital glaucoma. J Med Genet. 36**,** 290-4.

Quigley, H.A., 1977. The pathogenesis of reversible cupping in congenital glaucoma. Am J Ophthalmol. 84**,** 358-70.

Quigley, H.A., Enger, C., Katz, J., Sommer, A., Scott, R., Gilbert, D., 1994. Risk factors for the development of glaucomatous visual field loss in ocular hypertension. Arch Ophthalmol. 112**,** 644-9.

Quigley, H.A., Broman, A.T., 2006. The number of people with glaucoma worldwide in 2010 and 2020. Br J Ophthalmol. 90**,** 262-7.

Rada, J.A., Achen, V.R., Perry, C.A., Fox, P.W., 1997. Proteoglycans in the human sclera. Evidence for the presence of aggrecan. Invest Ophthalmol Vis Sci. 38**,** 1740-51.

Rada, J.A., Achen, V.R., Penugonda, S., Schmidt, R.W., Mount, B.A., 2000. Proteoglycan composition in the human sclera during growth and aging. Invest Ophthalmol Vis Sci. 41**,** 1639-48.

Reese, M.G., Eeckman, F.H., Kulp, D., Haussler, D., 1997. Improved splice site detection in Genie. J Comput Biol. 4**,** 311-23.

Reijns, M.A., Bubeck, D., Gibson, L.C., Graham, S.C., Baillie, G.S., Jones, E.Y., Jackson, A.P., 2011. The structure of the human RNase H2 complex defines key interaction interfaces relevant to enzyme function and human disease. J Biol Chem. 286**,** 10530-9.

Reis, L.M., Tyler, R.C., Muheisen, S., Raggio, V., Salviati, L., Han, D.P., Costakos, D., Yonath, H., Hall, S., Power, P., Semina, E.V., 2013. Whole exome sequencing in dominant cataract identifies a new causative factor, CRYBA2, and a variety of novel alleles in known genes. Hum Genet. 132**,** 761-70.

Remington, L.A., 2005. Clinical Anatomy and Physiology of the Visual System, Vol., Butterworth-Heinemann, St. Louis.

Reva, B., Antipin, Y., Sander, C., 2011. Predicting the functional impact of protein mutations: application to cancer genomics. Nucleic Acids Res. 39**,** e118.

Rezaie, T., Child, A., Hitchings, R., Brice, G., Miller, L., Coca-Prados, M., Heon, E., Krupin, T., Ritch, R., Kreutzer, D., Crick, R.P., Sarfarazi, M., 2002. Adult-onset primary open-angle glaucoma caused by mutations in optineurin. Science. 295**,** 1077-9.

Richards, A.J., McNinch, A., Martin, H., Oakhill, K., Rai, H., Waller, S., Treacy, B., Whittaker, J., Meredith, S., Poulson, A., Snead, M.P., 2010. Stickler syndrome and the vitreous phenotype: mutations in COL2A1 and COL11A1. Hum Mutat. 31**,** E1461-71.

Richards, S., Aziz, N., Bale, S., Bick, D., Das, S., Gastier-Foster, J., Grody, W.W., Hegde, M., Lyon, E., Spector, E., Voelkerding, K., Rehm, H.L., 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. Genet Med. 17**,** 405-24.

Rojas, B., Ramirez, A.I., de-Hoz, R., Salazar, J.J., Remirez, J.M., Trivino, A., 2006. [Structural changes of the anterior chamber angle in primary congenital glaucoma with respect to normal development]. Arch Soc Esp Oftalmol. 81**,** 65-71.

Roughley, P.J., Mort, J.S., 2014. The role of aggrecan in normal and osteoarthritic cartilage. J Exp Orthop. 1**,** 8.

Safari, I., Suri, F., Haji-Seyed-Javadi, R., Yazdani, S., Elahi, E., 2016. The p.Gly61Glu Mutation in CYP1B1 Affects the Extracellular Matrix in Glaucoma Patients. Ophthalmic Res. 56**,** 98-103.

Samant, M., Medsinge, A., Nischal, K.K., 2016. Pediatric Glaucoma: Pharmacotherapeutic Options. Paediatr Drugs. 18**,** 209-19.

Sampaolesi, R., Zarate, J., Sampaolesi, J.R., 2009. The Glaucomas Volume I - Pediatric Glaucomas, Vol., Springer.

Sarfarazi, M., Akarsu, A.N., Hossain, A., Turacli, M.E., Aktan, S.G., Barsoum-Homsy, M., Chevrette, L., Sayli, B.S., 1995. Assignment of a locus (GLC3A) for primary congenital glaucoma (Buphthalmos) to 2p21 and evidence for genetic heterogeneity. Genomics. 30**,** 171-7.

Sawyer, S.L., Hartley, T., Dyment, D.A., Beaulieu, C.L., Schwartzentruber, J., Smith, A., Bedford, H.M., Bernard, G., Bernier, F.P., Brais, B., Bulman, D.E., Warman Chardon, J., Chitayat, D., Deladoey, J., Fernandez, B.A., Frosk, P., Geraghty, M.T., Gerull, B., Gibson, W., Gow, R.M., Graham, G.E., Green, J.S., Heon, E., Horvath, G., Innes, A.M., Jabado, N., Kim, R.H., Koenekoop, R.K., Khan, A., Lehmann, O.J., Mendoza-Londono, R., Michaud, J.L., Nikkel, S.M., Penney, L.S., Polychronakos, C., Richer, J., Rouleau, G.A., Samuels, M.E., Siu, V.M., Suchowersky, O., Tarnopolsky, M.A., Yoon, G., Zahir, F.R., Majewski, J., Boycott, K.M., 2016. Utility of whole-exome sequencing for those near the end of the diagnostic odyssey: time to address gaps in care. Clin Genet. 89**,** 275-84.

Schlote, T., Rohrbach, J., Grueb, M., Mielke, J., 2006. Pocket Atlas of Ophthalmology, Vol., Clinical Sciences (Thieme), Wemding.

Schlotzer-Schrehardt, U., Lommatzsch, J., Kuchle, M., Konstas, A.G., Naumann, G.O., 2003. Matrix metalloproteinases and their inhibitors in aqueous humor of patients with pseudoexfoliation syndrome/glaucoma and primary open-angle glaucoma. Invest Ophthalmol Vis Sci. 44**,** 1117-25.

Schwarz, J.M., Cooper, D.N., Schuelke, M., Seelow, D., 2014. MutationTaster2: mutation prediction for the deep-sequencing age. Nat Methods. 11**,** 361-2.

Schymkowitz, J., Borg, J., Stricher, F., Nys, R., Rousseau, F., Serrano, L., 2005. The FoldX web server: an online force field. Nucleic Acids Res. 33**,** W382-8.

Seaby, E.G., Pengelly, R.J., Ennis, S., 2016. Exome sequencing explained: a practical guide to its clinical application. Brief Funct Genomics. 15**,** 374-84.

Seo, S., Singh, H.P., Lacal, P.M., Sasman, A., Fatima, A., Liu, T., Schultz, K.M., Losordo, D.W., Lehmann, O.J., Kume, T., 2012. Forkhead box transcription factor FoxC1 preserves corneal transparency by regulating vascular growth. Proc Natl Acad Sci U S A. 109**,** 2015-20.

Sheikh, S.A., Waryah, A.M., Narsani, A.K., Shaikh, H., Gilal, I.A., Shah, K., Qasim, M., Memon, A.I., Kewalramani, P., Shaikh, N., 2014. Mutational spectrum of the CYP1B1 gene in Pakistani patients with primary congenital glaucoma: novel variants and genotype-phenotype correlations. Mol Vis. 20**,** 991-1001.

Shibuya, E., Meguro, A., Ota, M., Kashiwagi, K., Mabuchi, F., Iijima, H., Kawase, K., Yamamoto, T., Nakamura, M., Negi, A., Sagara, T., Nishida, T., Inatani, M., Tanihara, H., Aihara, M., Araie, M., Fukuchi, T., Abe, H., Higashide, T., Sugiyama, K., Kanamoto, T., Kiuchi, Y., Iwase, A., Ohno, S., Inoko, H., Mizuki, N., 2008. Association of Toll-like receptor 4 gene polymorphisms with normal tension glaucoma. Invest Ophthalmol Vis Sci. 49**,** 4453-7.

Shihab, H.A., Gough, J., Cooper, D.N., Stenson, P.D., Barker, G.L., Edwards, K.J., Day, I.N., Gaunt, T.R., 2013. Predicting the functional, molecular, and phenotypic consequences of amino acid substitutions using hidden Markov models. Hum Mutat. 34**,** 57-65.

Siepel, A., Pollard, K.S., Haussler, D., 2006. New Methods for Detecting Lineage-Specific Selection. Proceedings of the 10th International Conference on Research in Computational Molecular Biology. 190-205.

Sievers, F., Wilm, A., Dineen, D., Gibson, T.J., Karplus, K., Li, W., Lopez, R., McWilliam, H., Remmert, M., Soding, J., Thompson, J.D., Higgins, D.G., 2011. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. Mol Syst Biol. 7**,** 539.

Sitorus, R., Ardjo, S.M., Lorenz, B., Preising, M., 2003. CYP1B1 gene analysis in primary congenital glaucoma in Indonesian and European patients. J Med Genet. 40**,** e9.

Smedley, D., Jacobsen, J.O., Jager, M., Kohler, S., Holtgrewe, M., Schubach, M., Siragusa, E., Zemojtel, T., Buske, O.J., Washington, N.L., Bone, W.P., Haendel, M.A., Robinson, P.N., 2015. Next-generation diagnostics and disease-gene discovery with the Exomiser. Nat Protoc. 10**,** 2004-15.

Souma, T., Tompson, S.W., Thomson, B.R., Siggs, O.M., Kizhatil, K., Yamaguchi, S., Feng, L., Limviphuvadh, V., Whisenhunt, K.N., Maurer-Stroh, S., Yanovitch, T.L., Kalaydjieva, L., Azmanov, D.N., Finzi, S., Mauri, L., Javadiyan, S., Souzeau, E., Zhou, T., Hewitt, A.W., Kloss, B., Burdon, K.P., Mackey, D.A., Allen, K.F., Ruddle, J.B., Lim, S.H., Rozen, S., Tran-Viet, K.N., Liu, X., John, S., Wiggs, J.L., Pasutto, F., Craig, J.E., Jin, J., Quaggin, S.E., Young, T.L., 2016. Angiopoietin receptor TEK mutations underlie primary congenital glaucoma with variable expressivity. J Clin Invest. 126**,** 2575-87.

Souzeau, E., Burdon, K.P., Dubowsky, A., Grist, S., Usher, B., Fitzgerald, J.T., Crawford, A., Hewitt, A.W., Goldberg, I., Mills, R.A., Ruddle, J.B., Landers, J., Mackey, D.A., Craig, J.E., 2013. Higher prevalence of myocilin mutations in advanced glaucoma in comparison with less advanced disease in an Australasian disease registry. Ophthalmology. 120**,** 1135-43.

Sowden, J.C., 2007. Molecular and developmental mechanisms of anterior segment dysgenesis. Eye (Lond). 21**,** 1310-8.

Spiga, M.G., Borras, T., 2010. Development of a gene therapy virus with a glucocorticoid-inducible MMP1 for the treatment of steroid glaucoma. Invest Ophthalmol Vis Sci. 51**,** 3029-41.

Srinivasan, R.S., de Erkenez, A.C., Hemenway, C.S., 2003. The mixed lineage leukemia fusion partner AF9 binds specific isoforms of the BCL-6 corepressor. Oncogene. 22**,** 3395-406.

Stamer, W.D., Read, A.T., Sumida, G.M., Ethier, C.R., 2009. Sphingosine-1-phosphate effects on the inner wall of Schlemm's canal and outflow facility in perfused human eyes. Exp Eye Res. 89**,** 980-8.

Stoilov, I., Akarsu, A.N., Sarfarazi, M., 1997. Identification of three different truncating mutations in cytochrome P4501B1 (CYP1B1) as the principal cause of primary congenital glaucoma (Buphthalmos) in families linked to the GLC3A locus on chromosome 2p21. Hum Mol Genet. 6**,** 641-7.

Stoilov, I., Akarsu, A.N., Alozie, I., Child, A., Barsoum-Homsy, M., Turacli, M.E., Or, M., Lewis, R.A., Ozdemir, N., Brice, G., Aktan, S.G., Chevrette, L., Coca-Prados, M., Sarfarazi, M., 1998. Sequence analysis and homology modeling suggest that primary congenital glaucoma on 2p21 results from mutations disrupting either the hinge region or the conserved core structures of cytochrome P4501B1. Am J Hum Genet. 62**,** 573-84.

Stoilov, I.R., Costa, V.P., Vasconcellos, J.P., Melo, M.B., Betinjane, A.J., Carani, J.C., Oltrogge, E.V., Sarfarazi, M., 2002. Molecular genetics of primary congenital glaucoma in Brazil. Invest Ophthalmol Vis Sci. 43**,** 1820-7.

Stoilov, I.R., Marfarazi, M., 2002. The third genetic locus (GLC3C) for primary congenital glaucoma (PCG) maps to chromosome 14q24.3. (Abstract). Invest Ophthal Vis Sci. 43**,** e3015.

Stone, E.M., Fingert, J.H., Alward, W.L., Nguyen, T.D., Polansky, J.R., Sunden, S.L., Nishimura, D., Clark, A.F., Nystuen, A., Nichols, B.E., Mackey, D.A., Ritch, R., Kalenak, J.W., Craven, E.R., Sheffield, V.C., 1997. Identification of a gene that causes primary open angle glaucoma. Science. 275**,** 668-70.

Sumida, G.M., Stamer, W.D., 2011. S1P(2) receptor regulation of sphingosine-1-phosphate effects on conventional outflow physiology. Am J Physiol Cell Physiol. 300**,** C1164-71.

Sunyaev, S.R., 2012. Inferring causality and functional significance of human coding DNA variants. Hum Mol Genet. 21**,** R10-7.

Suri, F., Yazdani, S., Narooie-Nejhad, M., Zargar, S.J., Paylakhi, S.H., Zeinali, S., Pakravan, M., Elahi, E., 2009. Variable expressivity and high penetrance of CYP1B1 mutations associated with primary congenital glaucoma. Ophthalmology. 116**,** 2101-9.

Suri, F., Yazdani, S., Elahi, E., 2015. Glaucoma in iran and contributions of studies in iran to the understanding of the etiology of glaucoma. J Ophthalmic Vis Res. 10**,** 68-76.

Suzuki, M., Hara, Y., Takagi, C., Yamamoto, T.S., Ueno, N., 2010. MID1 and MID2 are required for Xenopus neural tube closure through the regulation of microtubule organization. Development. 137**,** 2329-39.

Swaminathan, S.S., Oh, D.J., Kang, M.H., Rhee, D.J., 2014. Aqueous outflow: segmental and distal flow. J Cataract Refract Surg. 40**,** 1263-72.

Tamcelik, N., Atalay, E., Bolukbasi, S., Capar, O., Ozkok, A., 2014. Demographic features of subjects with congenital glaucoma. Indian J Ophthalmol. 62**,** 565-9.

Tamm, E.R., 2009. The trabecular meshwork outflow pathways: structural and functional aspects. Exp Eye Res. 88**,** 648-55.

Tanaka, K., Kato, A., Angelocci, C., Watanabe, M., Kato, Y., 2014. A potential molecular pathogenesis of cardiac/laterality defects in Oculo-Facio-Cardio-Dental syndrome. Dev Biol. 387**,** 28-36.

Tang, Y.M., Wo, Y.Y., Stewart, J., Hawkins, A.L., Griffin, C.A., Sutter, T.R., Greenlee, W.F., 1996. Isolation and characterization of the human cytochrome P450 CYP1B1 gene. J Biol Chem. 271**,** 28324-30.

Tanna, P., Strauss, R.W., Fujinami, K., Michaelides, M., 2016. Stargardt disease: clinical features, molecular genetics, animal models and therapeutic options. Br J Ophthalmol.

Tawara, A., Inomata, H., 1981. Developmental immaturity of the trabecular meshwork in congenital glaucoma. Am J Ophthalmol. 92**,** 508-25.

Tawara, A., Inomata, H., 1994. Distribution and characterization of sulfated proteoglycans in the trabecular tissue of goniodysgenetic glaucoma. Am J Ophthalmol. 117**,** 741-55.

Thiagalingam, A., De Bustros, A., Borges, M., Jasti, R., Compton, D., Diamond, L., Mabry, M., Ball, D.W., Baylin, S.B., Nelkin, B.D., 1996. RREB-1, a novel zinc finger protein, is involved in the differentiation response to Ras in human medullary thyroid carcinomas. Mol Cell Biol. 16**,** 5335-45.

Thomson, B.R., Heinen, S., Jeansson, M., Ghosh, A.K., Fatima, A., Sung, H.K., Onay, T., Chen, H., Yamaguchi, S., Economides, A.N., Flenniken, A., Gale, N.W., Hong, Y.K., Fawzi, A., Liu, X., Kume, T., Quaggin, S.E., 2014. A lymphatic defect causes ocular hypertension and glaucoma in mice. J Clin Invest. 124**,** 4320-4.

Thorleifsson, G., Magnusson, K.P., Sulem, P., Walters, G.B., Gudbjartsson, D.F., Stefansson, H., Jonsson, T., Jonasdottir, A., Stefansdottir, G., Masson, G., Hardarson, G.A., Petursson, H., Arnarsson, A., Motallebipour, M., Wallerman, O., Wadelius, C., Gulcher, J.R., Thorsteinsdottir, U., Kong, A., Jonasson, F., Stefansson, K., 2007. Common sequence variants in the LOXL1 gene confer susceptibility to exfoliation glaucoma. Science. 317**,** 1397-400.

Thyer, H.W., Wilson, P., 1972. Trabeculectomy. Br J Ophthalmol. 56**,** 37-40.

Tsuda, H., Takebayashi, K., Nakanishi, S., Kageyama, R., 1998. Structure and promoter analysis of Math3 gene, a mouse homolog of Drosophila proneural gene atonal. Neural-specific expression by dual promoter elements. J Biol Chem. 273**,** 6327-33.

Tucker, R.P., Chiquet-Ehrismann, R., 2015. Tenascin-C: Its functions as an integrin ligand. Int J Biochem Cell Biol. 65**,** 165-8.

Tumer, Z., Bach-Holm, D., 2009. Axenfeld-Rieger syndrome and spectrum of PITX2 and FOXC1 mutations. Eur J Hum Genet. 17**,** 1527-39.

Tuori, A., Uusitalo, H., Thornell, L.E., Yoshida, T., Virtanen, I., 1999. The expression of tenascin-X in developing and adult rat and human eye. Histochem J. 31**,** 245-52.

Ueda, J., Wentz-Hunter, K., Yue, B.Y., 2002. Distribution of myocilin and extracellular matrix components in the juxtacanalicular tissue of human eyes. Invest Ophthalmol Vis Sci. 43**,** 1068-76.

Vasiliou, V., Gonzalez, F.J., 2008. Role of CYP1B1 in glaucoma. Annu Rev Pharmacol Toxicol. 48**,** 333-58.

Velayutham, V., 2009. Basic Sciences in Ophthalmology, Vol., Jaypee Brothers Medical Pub, New Delhi.

Vincent, A.L., Billingsley, G., Buys, Y., Levin, A.V., Priston, M., Trope, G., Williams-Lyn, D., Heon, E., 2002. Digenic inheritance of early-onset glaucoma: CYP1B1, a potential modifier gene. Am J Hum Genet. 70**,** 448-60.

Vithana, E.N., Khor, C.C., Qiao, C., Nongpiur, M.E., George, R., Chen, L.J., Do, T., Abu-Amero, K., Huang, C.K., Low, S., Tajudin, L.S., Perera, S.A., Cheng, C.Y., Xu, L., Jia, H., Ho, C.L., Sim, K.S., Wu, R.Y., Tham, C.C., Chew, P.T., Su, D.H., Oen, F.T., Sarangapani, S., Soumittra, N., Osman, E.A., Wong, H.T., Tang, G., Fan, S., Meng, H., Huong, D.T., Wang, H., Feng, B., Baskaran, M., Shantha, B., Ramprasad, V.L., Kumaramanickavel, G., Iyengar, S.K., How, A.C., Lee, K.Y., Sivakumaran, T.A., Yong, V.H., Ting, S.M., Li, Y., Wang, Y.X., Tay, W.T., Sim, X., Lavanya, R., Cornes, B.K., Zheng, Y.F., Wong, T.T., Loon, S.C., Yong, V.K., Waseem, N., Yaakub, A., Chia, K.S., Allingham, R.R., Hauser, M.A., Lam, D.S., Hibberd, M.L., Bhattacharya, S.S., Zhang, M., Teo, Y.Y., Tan, D.T., Jonas, J.B., Tai, E.S., Saw, S.M.,

Hon, D.N., Al-Obeidan, S.A., Liu, J., Chau, T.N., Simmons, C.P., Bei, J.X., Zeng, Y.X., Foster, P.J., Vijaya, L., Wong, T.Y., Pang, C.P., Wang, N., Aung, T., 2012. Genome-wide association analyses identify three new susceptibility loci for primary angle closure glaucoma. Nat Genet. 44**,** 1142-6.

Vittal, V., Rose, A., Gregory, K.E., Kelley, M.J., Acott, T.S., 2005. Changes in gene expression by trabecular meshwork cells in response to mechanical stretching. Invest Ophthalmol Vis Sci. 46**,** 2857-68.

Vittitow, J., Borras, T., 2004. Genes expressed in the human trabecular meshwork during pressure-induced homeostatic response. J Cell Physiol. 201**,** 126-37.

Vogt, J., Agrawal, S., Ibrahim, Z., Southwood, T.R., Philip, S., Macpherson, L., Bhole, M.V., Crow, Y.J., Oley, C., 2013. Striking intrafamilial phenotypic variability in Aicardi-Goutieres syndrome associated with the recurrent Asian founder mutation in RNASEH2C. Am J Med Genet A. 161A**,** 338-42.

Weisschuh, N., Neumann, D., Wolf, C., Wissinger, B., Gramer, E., 2005. Prevalence of myocilin and optineurin sequence variants in German normal tension glaucoma patients. Mol Vis. 11**,** 284-7.

Weisschuh, N., Wolf, C., Wissinger, B., Gramer, E., 2009. A clinical and molecular genetic study of German patients with primary congenital glaucoma. Am J Ophthalmol. 147**,** 744-53.

Welti, S., Kuhn, S., D'Angelo, I., Brugger, B., Kaufmann, D., Scheffzek, K., 2011. Structural and biochemical consequences of NF1 associated nontruncating mutations in the Sec14-PH module of neurofibromin. Hum Mutat. 32**,** 191-7.

Wentz-Hunter, K., Kubota, R., Shen, X., Yue, B.Y., 2004. Extracellular myocilin affects activity of human trabecular meshwork cells. J Cell Physiol. 200**,** 45-52.

Wiggs, J.L., Damji, K.F., Haines, J.L., Pericak-Vance, M.A., Allingham, R.R., 1996. The distinction between juvenile and adult-onset primary open-angle glaucoma. Am J Hum Genet. 58**,** 243-4.

Wiggs, J.L., Langgurth, A.M., Allen, K.F., 2014. Carrier frequency of CYP1B1 mutations in the United States (an American Ophthalmological Society thesis). Trans Am Ophthalmol Soc. 112**,** 94-102.

Wilkins, M., Indar, A., Wormald, R., 2001. Intra-operative mitomycin C for glaucoma surgery. Cochrane Database Syst Rev. CD002897.

Wilson, M.E., Saunders, R.A., Trivedi, R.H., 2009. Pediatric Ophthalmology - Current Thought and Practical Guide, Vol., Springer.

Wirtz, M.K., Bradley, J.M., Xu, H., Domreis, J., Nobis, C.A., Truesdale, A.T., Samples, J.R., Van Buskirk, E.M., Acott, T.S., 1997. Proteoglycan expression by human trabecular meshworks. Curr Eye Res. 16**,** 412-21.

Wright, C., Tawfik, M.A., Waisbourd, M., Katz, L.J., 2016. Primary angle-closure glaucoma: an update. Acta Ophthalmol. 94**,** 217-25.

Yang, H., Wang, K., 2015. Genomic variant annotation and prioritization with ANNOVAR and wANNOVAR. Nat Protoc. 10**,** 1556-66.

Yang, M., Guo, X., Liu, X., Shen, H., Jia, X., Xiao, X., Li, S., Fang, S., Zhang, Q., 2009. Investigation of CYP1B1 mutations in Chinese patients with primary congenital glaucoma. Mol Vis. 15**,** 432-7.

Yanoff, M., Sassani, J.W., 2009. Ocular Pathology, Vol., Elsevier Inc.

Yu Chan, J.Y., Choy, B.N., Ng, A.L., Shum, J.W., 2015. Review on the Management of Primary Congenital Glaucoma. J Curr Glaucoma Pract. 9**,** 92-9.

Yu, X., Seegar, T.C., Dalton, A.C., Tzvetkova-Robev, D., Goldgur, Y., Rajashankar, K.R., Nikolov, D.B., Barton, W.A., 2013. Structural basis for angiopoietin-1-mediated signaling initiation. Proc Natl Acad Sci U S A. 110**,** 7205-10.

Yu, Y., Triebwasser, M.P., Wong, E.K., Schramm, E.C., Thomas, B., Reynolds, R., Mardis, E.R., Atkinson, J.P., Daly, M., Raychaudhuri, S., Kavanagh, D., Seddon, J.M., 2014. Whole-

exome sequencing identifies rare, functional CFH variants in families with macular degeneration. Hum Mol Genet. 23**,** 5283-93.

Yue, B.Y., 2011. Myocilin and Optineurin: Differential Characteristics and Functional Consequences. Taiwan J Ophthalmol. 1**,** 6-11.

Zhang, X., Zhang, Q., Tong, Y., Dai, H., Zhao, X., Bai, F., Xu, L., Li, Y., 2011. Large novel deletions detected in Chinese families with aniridia: correlation between genotype and phenotype. Mol Vis. 17**,** 548-57.

Zhang, Y.Z., Zhao, D.H., Yang, H.P., Liu, A.J., Chang, X.Z., Hong, D.J., Bonnemann, C., Yuan, Y., Wu, X.R., Xiong, H., 2014. Novel collagen VI mutations identified in Chinese patients with Ullrich congenital muscular dystrophy. World J Pediatr. 10**,** 126-32.

Zhao, Y., Wang, S., Sorenson, C.M., Teixeira, L., Dubielzig, R.R., Peters, D.M., Conway, S.J., Jefcoate, C.R., Sheibani, N., 2013. Cyp1b1 mediates periostin regulation of trabecular meshwork development by suppression of oxidative stress. Mol Cell Biol. 33**,** 4225-40.

Zhu, X., Dai, F.R., Wang, J., Zhang, Y., Tan, Z.P., 2015. Novel BCOR mutation in a boy with Lenz microphthalmia/oculo-facio-cardio-dental (OFCD) syndrome. Gene. 571**,** 142-4.

Annexes

Table S2.1 – Clinical features of the PCG patients. M: male, F: female, OU:bilateral, OD: right eye, OS: left eye, IOP: intraocular pressure, Y: yes, N: no, Surg: surgery, C/D ratio: cup-to-disc ratio, NA: not available, VA: visual acuity, NLP: no light perception, LP: light perception, VFT: visual field test.

Patient ID	Sex	Age Onset/Age present (years)	Affected eyes	IOP at diagnosis OD/OS (mm Hg)	Corneal status	Family history	Treatment	C/D ratio OD/OS	Corneal diameter OD/OS (mm)	VA OD/OS (units)	Age at first surgery (years)	Post-surg IOP OD/OS (mm Hg)	Number of surgeries	VFT
1	F	<1/49	OU	30/29	Megalocornea; Haab striae	N	Surg, drops	0.9/0.9	13/13	<0.1/0.4	1	14/19	8	Constricted
2	M	<1/13	OU	40/42	Megalocornea; Haab striae	Y	Surg, drops	0,9/0,9	13,5/13,5	<0.1/<0.1	<1	15/17	6	Constricted
3	F	<1/35	OU	35/34	Megalocornea; Haab striae	N	Surg, drops	NA/NA	13/13	0.2/0.4	<1	17/18	4	General reduced sensitivity
4	M	<1/10	OD	35/22	OD megalocornea	N	Surg, drops	1.0/0.3	14/11.5	<01/0.8	<1	30/17	3	Constricted OD
5	M	<1/14	OU	26/27	Megalocornea; Haab striae	N	Surg, drops	0.9/1.0	14/14	0.4/0.1	<1	16/16	2	Constricted
6	M	<1/11	OU	45/39	Megalocornea; Haab striae	N	Surg, drops	1.0/1.0	14/13	<0.1/<0.1	<1	19/16	4	Constricted
7	F	<1/37	OU	35/45	OD megalocornea; OS prosthesis	N	Surg, drops	NA/ -	14/ -	NLP/ -	1	21/ -	8	Zero
8	M	<1/12	OU	28/29	Megalocornea; Haab striae	N	Surg, drops	0.6/0.6	13/13	0./0.5	<1	15/14	2	Reduced sensitivity
9	M	<1/18	OD	30/13	OD megalocornea; OS no changes	Y consanguinity	Surg, drops	0.3/0.3	13/10.5	0.8/1.0	1	17/17	2	Normal
10	M	<1/21	OU	40/45	OD megalocornea; OS phthisis bulbi	N	Surg, drops	0.9/ -	13,5/ -	<0.1/NLP	1	19/ -	8	Severely constricted
11	F	<1/37	OU	40/40	OD prosthesis; OS megalocornea; Haab striae	N	Surg, drops	- /0.9	- /14	-/LP	2	- / 22	6	Severely constricted
12	M	<1/23	OU	36/37	OD phthisis bulbi; OS megalocornea	N	Surg, drops	- /0.6	5/13,5	NLP/0.6	<1	10,5/18	7	Constricted, reduced sensitivity
14	F	<1/6	OU	30/29	Megalocornea; Haab striae	N	Surg, drops	0.4/0.5	12.5/13	0.2/0.2	<1	14/13	2	NA/NA

Patient ID	Sex	Age Onset/Age present (years)	Affected eyes	IOP at diagnosis OD/OS (mm Hg)	Corneal status	Family history	Treatment	C/D ratio OD/OS	Corneal diameter OD/OS (mm)	VA OD/OS (units)	Age at first surgery (years)	Post-surg IOP OD/OS (mm Hg)	Number of surgeries	VFT
15	F	<1/17	OU	32/38	Megalocornea; Haab striae	N	Surg, drops	0.9/0.9	12.5/13.5	0.5/<0.1	<1	12,5/18	5	Constricted
16	F	<1/36	OU	28/27	Megalocornea; Haab striae	N	Surg, drops	0.7/0.7	12,5/12.5	0.3/0.4	1	15/16,5	3	Constricted
17	M	<1/27	OU	28/27	Megalocornea; Haab striae	N	Surg, drops	0.9/0.9	13/13	0.5/0.6	<1	14/14.5	2	Constricted
18	F	<1/36	OU	NA/NA	Prosthesis OD; Megalocornea OS; central corneal graft	N	Surg, drops	-/0.9	-/13	-/0.25	<1	- /9	8	NA/severely constricted
21	F	<1/33	OU	37/38	Complete opacification; megalocornea	N	Surg, drops	NA/NA	14/14	NLP/NLP	2	25/26	6	NA/NA
22	F	<1/9	OU	45/42	Megalocornea; Haab striae	N	Surg, drops	0.9/1.0	13.5/13.5	NLP/LP	<1	20/18	8	NA/NA
23	M	<1/7	OU	38/32	Megalocornea; Haab striae	N	Surg, drops	0.7/0.9	13.5/13	<0.1/<0.1	<1	17/18	4	NA/NA
25	M	<1/10	OU	24/22	Megalocornea	N	Surg, drops	0.4/0.3	12/13.5	1.0/<0.1	2	15/16.5	2	Normal
26	F	<1/43	OU	40/40	Centrally opacified megalocorneas	Y	Surg, drops	1.0/1.0	14/14	<0.1/<0.1	2	21/20	4	Severely constricted
28	F	<1/33	OU	31/30	Megalocornea; Haab striae	N	Surg, drops	0.9/.9	13.5/13	0.1/<0.1	1	15/16	4	Severely constricted
29	F	<1/24	OU	35/32	Megalocornea; Haab striae	Y	Surg, drops	1.0/1.0	13.5/13.5	<0.1/<0.1	<1	18/17	6	NA/NA
32	F	<1/9	OU	40/41	Megalocornea; Haab striae	N	Surg, drops	NA/NA	13.5/12.5	NA/NA	<1	19/13	4	NA/NA
33	M	<1/4	OD	23/15	OD megalocornea; AS dysgenesis	N	Drops	NA/0.1	15/11.5	NA/NA	-	-	0	NA/NA
34	F	<1/37	OU	33/34	OD prosthesis; OS megalocornea, Haab striae	N	Surg, drops	NA/NA	- /14.5	- /LP	1	- /17.5	8	NA/NA
35	M	<1/39	OU	40/42	Megalocornea; OD central graft transparent; OS prosthesis	N	Surg, drops	NA/NA	12/ -	<0.1/ -	1	17/ -	8	NA/NA
36	F	<1/5	OU	21/19	Megalocornea; corneal opacities	N	Surg, drops	0.5/0.5	13/13.5	0.5/0.1	<1	19/25	1	NA/NA

Patient ID	Sex	Age Onset/Age present (years)	Affected eyes	IOP at diagnosis OD/OS (mm Hg)	Corneal status	Family history	Treatment	C/D ratio OD/OS	Corneal diameter OD/OS (mm)	VA OD/OS (units)	Age at first surgery (years)	Post-surg IOP OD/OS (mm Hg)	Number of surgeries	VFT
38	F	<1/2	OU	24/18	Megalocornea; OD Haab's striae; OD corneal opacities	N	Surg, drops	0,5/0,4	12.5/12.5	0.1/0.2	<1	9/8	2	NA/NA
39	M	<1/32	OU	22/22	Megalocornea; corneal oedema	N	Drops	NA	NA	0.8/1.0	-	18/11	0	normal
40	M	<1/1	OU	25/34	corneal opacities (+OS), megalocornea (+OS)	N	surg	0,3/0,4	13/13	NA/NA	<1	11/14	1	NA/NA
41	M	<1/1	OU	22/24	corneal oedema (+OD)	N	surg	NA	12.5/13.5	NA/NA	<1	8/8	1	NA/NA
42	F	<1/3	OD	33/18	corneal opacities OD	N	Surg, drops	0,7/0,2	13/11,5	NA	<1	11/8	1	NA/NA
43	M	<1/<1	OU	28/23	corneal opacities	N	surg	0,4/0,4	11/11	NA	<1	12/12	1	NA/NA
44	F	<1/1	OU	20/29	corneal opacities	N	Surg, drops	0,8/0,5	16,5/14	NA	<1	19/16	1	NA/NA
45	M	1/1	OU	26/20	Megalocornea; Haab striae	N	Surg, drops	0,9/0,7	14/14	NA	1	35/29	1	NA/NA

Table S2.2 – Genotype of the five SNPs involved in the *CYP1B1* haplotype and the corresponding mutations for the positive *CYP1B1* patients. All the genotypes were obtained by Sanger sequencing. Patients with only one mutation have the mutation in the homozygous state, while patients with two mutations have both mutations in the heterozygous state.

Patient ID	*CYP1B1* mutations	R48G	A119S	V432L	D449D	N453S
1	-	C/G	G/T	C	C	A/G
2	p.R468_S476dup	C	G	G	T	A
3	p.E387K	G	T	C	C	A
4	-	C	G	G/C	T/C	A
5	-	C/G	G/T	G/C	T/C	A
6	p.A179RfsX18 p.E387K	C/G	G/T	G/C	T/C	A
7	p.A179RfsX18 p.R355HfsX69	C	G	G	T	A
8	p.E387K	G	T	C	C	A
9	-	G	T	C	C	A
10	p.A179RfsX18	C	G	G	T	A
11	p.A179RfsX18	C	G	G	T	A
12	p.A179RfsX18 p.R444X	C/G	G/T	G/C	T/C	A
13	p.E387K p.P437L	C/G	G/T	G/C	T/C	A
14	p.A179RfsX18 p.L378Q	C/G	G/T	G/C	T/C	A
15	p.A179RfsX18 p.T404SfsX30	C	G	G	T	A
16	p.A179RfsX18 p.R355HfsX69	C	G	G	T	A
17	-	C	G	C	C	A/G
18	p.E387K p.T404SfsX30	C/G	G/T	G/C	T/C	A
19	p.E387K	G	T	C	C	A
20	p.R355HfsX69	C	G	G	T	A
21	-	C	G	G/C	T/C	A

Patient ID	*CYP1B1* mutations	R48G	A119S	V432L	D449D	N453S
22	p.A179RfsX18 p.R355HfsX69	C	G	G	T	A
23	p.A179RfsX18 p.D449MfsX8	C	G	G/C	T/C	A/G
24	p.R468_S476dup	C	G	G	T	A
25	-	C	G	C	C	G
26	-	C	G	G/C	T/C	A
27	p.R355HfsX69	C	G	G	T	A
28	p.T404SfsX30	C	G	G	T	A
29	p.A179RfsX18 p.T404SfsX30	C	G	G	T	A
30	p.A179RfsX18 p.E387K	C/G	G/T	G/C	T/C	A

Annex

Table S3.1 – Clinical features of the PCG patients without CYP1B1 mutations. M: male, F: female, OU:bilateral, OD: right eye, OS: left eye, IOP: intraocular pressure, Y: yes, N: no, Surg: surgery, C/D ratio: cup-to-disc ratio, NA: not available, VA: visual acuity, NLP: no light perception, LP: light perception, VFT: visual field test.

Patient ID	Sex	Age Onset/Age present (years)	Affected eyes	IOP at diagnosis OD/OS (mm Hg)	Corneal status	Family history	Treatment	C/D ratio OD/OS	Corneal diameter OD/OS (mm)	VA OD/OS (units)	Age at first surgery (years)	Post-surg IOP OD/OS (mm Hg)	Number of surgeries	VFT
1	F	<1/49	OU	30/29	Megalocornea; Haab striae	N	Surg, drops	0.9/0.9	13/13	<0.1/0.4	1	14/19	8	Constricted
4	M	<1/10	OD	35/22	OD megalocornea	N	Surg, drops	1.0/0.3	14/11.5	<01/0.8	<1	30/17	3	Constricted OD
5	M	<1/14	OU	26/27	Megalocornea; Haab striae	N	Surg, drops	0.9/1.0	14/14	0.4/0.1	<1	16/16	2	Constricted
9	M	<1/18	OD	30/13	OD megalocornea; OS no changes	Y consanguinity	Surg, drops	0.3/0.3	13/10.5	0.8/1.0	1	17/17	2	Normal
18	F	<1/36	OU	NA/NA	Prosthesis OD; Megalocornea OS; central corneal graft	N	Surg, drops	-/0.9	-/13	-/0.25	<1	- /9	8	NA/severely constricted
25	M	<1/10	OU	24/22	Megalocornea	N	Surg, drops	0.4/0.3	12/13.5	1.0/<0.1	2	15/16.5	2	Normal
31	M	<1/6	OD	NA/NA	Centrally opacified megalocorneas	N	NA/NA	NA/NA	NA/NA	NA/NA	NA/NA	NA/NA	NA/NA	NA/NA
32	F	<1/9	OU	40/41	Megalocornea; Haab striae	N	Surg, drops	NA/NA	13.5/12.5	NA/NA	<1	19/13	4	NA/NA
33	M	<1/4	OD	23/15	OD megalocornea; AS dysgenesis	N	Drops	NA/0.1	15/11.5	NA/NA	-	-	0	NA/NA
36	F	<1/5	OU	21/19	Megalocornea; corneal opacities	N	Surg, drops	0.5/0.5	13/13.5	0.5/0.1	<1	19/25	1	NA/NA
41	M	<1/1	OU	22/24	corneal oedema (+OD)	N	surg	NA	12.5/13.5	NA/NA	<1	8/8	1	NA/NA
42	F	<1/3	OD	33/18	corneal opacities OD	N	Surg, drops	0,7/0,2	13/11,5	NA	<1	11/8	1	NA/NA
43	M	<1/<1	OU	28/23	corneal opacities	N	surg	0,4/0,4	11/11	NA	<1	12/12	1	NA/NA
45	M	1/1	OU	26/20	Megalocornea; Haab striae	N	Surg, drops	0,9/0,7	14/14	NA	1	35/29	1	NA/NA

Table S3.2 – WES quality control parameters for samples sequenced with Ion Proton™ and HiSeq 2000. The number of reads obtained and the percentage of reads mapping the reference human genome hg19/ GRCh37 were similar, with exception of sample 31. Exomes sequenced with Ion Proton had a higher coverage. NA: not available.

Sample ID	Total number of reads	% Reads Aligned	Mean Target Coverage	% Target Bases with 20x Coverage	Uniformity
Sample 1	47,854,200	98.50	131.2	93.55	90.92
Sample 4	43,020,016	98.82	130.7	93.85	91.42
Sample 5	41,130,428	98.97	126.1	92.51	90.08
Sample 9	57,309,401	99.58	181.7	94.61	89.24
Sample 18	51,438,861	99.57	163.0	94.78	90.46
Sample 25	49,496,138	99.57	156.9	94.90	90.91
Sample 31	36,099,240	99.10	46.5	77.26	NA
Sample 32	53,921,730	99.09	70.1	88.61	NA
Sample 33	48,914,612	99.13	62.7	85.53	NA
Sample 36	46,975,416	99.46	146.1	93.76	89.91
Sample 41	50,603,510	99.39	156.3	94.62	90.43
Sample 42	52,741,439	99.40	165.1	95.11	90.93
Sample 43	46,366,052	99.61	147.0	94.90	91.66
Sample 45	51,262,773	99.62	165.0	94.70	89.95
Mother 1	39,342,535	98.75	109.7	93.21	92.38
Father 1	38,743,817	98.44	108.2	91.06	90.14
Mother 31	41,466,681	98.47	116.4	94.03	90.12
Father 31	43,021,898	98.63	119.2	92.42	90.51
Sample 19	46,096,092	99.05	131.6	95.17	93.50

Table S3.3 – Initial number of variants and after each filter application during the prioritization step, per sample. SD: standard deviation.

Sample ID	Initial number of variants in the VCF	Selection of exonic and spice site variants	Exclusion of synonymous variants	Exclusion of variants present in the 1000 G, ESP, ExAC and in-house database with MAF > 1% (EUR)
Sample 1	53370	23337	12937	812
Sample 4	53161	23313	12768	643
Sample 5	52676	22851	12415	721
Sample 9	52410	22818	12988	735

Sample 18	53489	23441	13285	732
Sample 25	52989	23190	13236	815
Sample 31	45265	20007	10733	680
Sample 32	45494	20130	10719	521
Sample 33	45295	19933	10756	641
Sample 36	53534	23129	12682	708
Sample 41	62070	26876	14630	897
Sample 42	52958	23149	13136	799
Sample 43	52969	23169	13191	835
Sample 45	53652	23237	12591	720

Table S3.4A – Genes, variants and variant frequencies for the analysis of the recessive inheritance pattern for sample 1. All variants have a European MAF inferior to 0.01 in at least in one database. Chr: chromosome, Ref: reference nucleotide, Alt: altered nucleotide, 1000G: 1000 Genomes Project, ExAC: Exome Aggregation Consortium, ESP: NHLBI Exome Sequencing Project, EUR: European, AFR: African, AMR: American, EAS: East Asian, SAS: South Asian, NFE: Non-Finnish European, FIN: Finnish European, OTH: Other populations, EA: European American, AA: African American. Indication of two variants for the same gene represents compound heterozygosity, while a single variant indicates homozygosity. Dot: no information on databases.

Gene	Chr	Start	Ref	Alt	Variants	1000G EUR	1000G AFR	1000G AMR	1000G EAS	1000G SAS	ExAC NFE	ExAC FIN	ExAC AFR	ExAC AMR	ExAC EAS	ExAC SAS	ExAC OTH	ESP EA	ESP AA
AHNAK	11	62286165	C	G	p.G1631S	0.018	0	0.0043	0	0.002	0.0095	0.0204	0.0015	0.0021	0	0.0024	0.0044	0.0098	0.0018
		62296998	C	T	p.G5242R	0	0.033	0.0014	0	0	0.0003	0	0.0218	0.0029	0	0	0.0055	0.0002	0.021
CCDC57	17	80159827	C	T	c.-7G>A	0	0	0.0014	0	0	2.07E-5	0	0	0.0020	0.0002	0	0.0017	.	.
		80146291	C	G	p.E286Q	0.001	0	0	0	0	1.71E-5	0	0	0	0	0	0	.	.
CFAP74	1	1905526	C	A	p.E204D	.	.	.	.	.	0.0034	0.0005	0.0010	0.0022	0	6.34E-5	0.0012	0.0046	0.0014
		1890559	C	T	c.G1851A	0.005	0.0023	0.0014	0	0.0031	0.0027	0.0003	0.0020	0.0026	0.0002	0.0040	0.0011	0.0043	0.0005
EPOR	19	11494762	T	C	c.115+7A>G	0	0.0015	0.0029	0	0	0	0	0	0.0025	0	0	0	.	.
		11488877	C	T	p.R437H	0.002	0.0091	0.0014	0	0	0.0008	0	0.0068	0.0014	0	6.06E-5	0.0022	0.0007	0.0052
FSIP2	2	186610279	C	T	p.R290W	.	.	.	.	.	.	.	.	.	.	.	.	.	.
		186655636	A	G	p.D1347G	.	.	.	.	.	.	.	.	.	.	.	.	.	.
MX2	21	42748886	C	T	p.S18F	.	.	.	.	.	0.0001	0	0	0.0002	0	0	0.0011	.	.
		42754447	A	T	p.T230S	.	.	.	.	.	.	.	.	.	.	.	.	.	.
NUDCD1	8	110305614	T	C	p.Y200C	0.001	0	0	0	0	0.0005	0.0003	0.0005	0.0005	0	0.0004	0	0.0006	0
		110302130	T	C	p.I225V	.	.	.	.	.	4.75E-5	0	9,98E-5	8,95E-5	0	0	0	.	.
PPFIBP2	11	7647098	C	T	p.H268Y	0.003	0.0008	0.0014	0	0.001	0.0040	0.0015	0.0004	0.0011	0	0.0044	0.0033	0.0034	0.0007
		7656816	G	C	p.E410Q	.	.	.	.	.	7.49E-5	0	0	0	0	0	0	.	.
SRRM2	16	2815310	G	A	p.R1594Q	.	.	.	.	.	4.50E-5	0	0	0	0.0001	0	0	0.0001	0
		2815888	C	T	p.R1787C	0	0	0.0029	0	0	0.0002	0	0.0004	0.0003	0	0	0	0.0001	0.0002
STRA6	15	74472522	C	T	p.G635S	0.012	0	0.0029	0.001	0.0031	0.0104	0.0098	0.0022	0.0023	0	0.0050	0.0077	0.011	0.0027
TNS3	7	47408732	G	A	p.G316S	0.011	0.0015	0.0043	0	0.03	0.0116	0.0032	0.0029	0.0059	0.0002	0.0261	0.0122	0.015	0.0022
		47436475	C	T	p.S504L	0.001	0	0	0	0	0.0002	0	0.0002	0.0005	0	6.06E-5	0	0.0002	0

Table S3.4B – Genes, variants and predicted functional impact for the analysis of the recessive inheritance model for sample 1. Each functional impact predictor has a different classification with different cut-offs. SIFT: T – tolerated (>0.05), D – damaging (≤0.05); PolyPhen2: B – benign (<0.452), P – possibly damaging (0.453 and 0.956), D – probably damaging (>0.957); LRT: D – deleterious, N – neutral, U – unknown; MutationTaster: P – polymorphism automatic, N – polymorphism, D – disease causing, A – disease causing automatic; MutationAssessor: N – neutral, L – low, M – medium, H – high; FATHMM: T – tolerated (> -1.5), D – damaging (< -1.5); Radial SVM: T – tolerated, D – damaging; LR: T – tolerated, D – damaging; CADD: ≥15 considered pathogenic; GERP++: ≥2 considered conserved; PhyloP: ≥1; HSF: D – disturbs donor splice site; A – disturbs acceptor splice site; No – predicted to not affect splicing. Dot – no prevision was generated. The last column indicates the classification of each variant. A variant is classified as pathogenic when 6 of the 9 functional impact software predict it as damaging, otherwise is classified as benign. A variant is classified as relatively frequent if MAF is > 1% in other populations but European or if there are individuals with homozygous variants, otherwise is classified as rare.

Gene	Variants	SIFT	PolyPhen2	LRT	Mutation Taster	Mutation Assessor	FATHMM	Radial SVM	LR	CADD	GERP++	PhyloP	HSF	Classification
AHNAK	p.G1631S	T	D	U	N	M	T	T	T	15.55	2.99	1.577	.	Relatively frequent and predicted as benign.
	p.G5242R	T	P	N	N	M	T	T	T	4.996	-7.49	-0.179	.	
CCDC57	c.-7G>A	.	.	.	.	.	.	.	.	.	.	.	A	Rare and predicted as benign.
	p.E286Q	T	D	N	D	L	T	T	T	13.74	5.54	1.213	.	
CFAP74	p.E204D	T	B	N	N	L	.	T	T	7.539	2.46	1.887	.	Relatively frequent and predicted as benign.
	c.G1851A	.	.	.	D	.	.	T	T	0.010	-9.44	-2.004	.	
EPOR	c.115+7A>G	.	.	.	.	.	.	.	.	.	.	.	.	c.115+7A>G rare and predicted as pathogenic. p.R437H relatively frequent and predicted as benign.
	p.R437H	T	B	N	N	N	T	T	T	7.010	2.81	0.294	D	
FSIP2	p.R290W	D	.	.	N	L	T	T	T	9.760	1.36	0.449	.	Rare and predicted as benign.
	p.D1347G	D	.	N	N	L	T	T	T	6.266	3.05	1.013	.	
MX2	p.S18F	D	B	N	N	N	D	D	D	10.85	2.57	0.722	.	Rare and predicted as benign.
	p.T230S	T	P	U	N	M	D	D	D	12.88	-6.6	0.089	.	
NUDCD1	p.Y200C	T	P	N	D	L	T	T	T	13.60	2.07	1.248	.	Rare and predicted as benign.
	p.I225V	T	B	N	N	N	T	T	T	1.710	-6.87	-0.272	.	
PPFIBP2	p.H268Y	T	B	N	N	L	T	T	T	15.75	5.24	1.527	.	Rare and predicted as benign.
	p.E410Q	T	B	N	N	L	T	T	T	10.47	4.47	2.673	.	
SRRM2	p.R1594Q	D	B	N	N	N	T	T	T	3.243	1.2	0.422	.	Rare and predicted as benign.
	p.R1787C	T	D	D	D	N	T	T	T	7.841	5.46	0.460	.	
STRA6	p.G635S	T	B	N	N	N	T	T	T	0.015	-4.79	-0.733	.	Relatively frequent and predicted as benign.
TNS3	p.G316S	T	B	N	N	M	D	T	D	9.347	3.64	3.727	.	Relatively frequent and predicted as benign.
	p.S504L	T	B	N	N	N	T	T	T	0.028	0.729	-0.078	.	

Table S3.5A – Genes, variants and variant frequencies for the analysis of the recessive inheritance pattern for sample 4. All variants have a European MAF inferior to 0.01 in at least in one database. Chr: chromosome, Ref: reference nucleotide, Alt: altered nucleotide, 1000G: 1000 Genomes Project, ExAC: Exome Aggregation Consortium, ESP: NHLBI Exome Sequencing Project, EUR: European, AFR: African, AMR: American, EAS: East Asian, SAS: South Asian, NFE: Non-Finnish European, FIN: Finnish European, OTH: Other populations, EA: European American, AA: African American. Indication of two variants for the same gene represents compound heterozygosity, while a single variant indicates homozygosity. Dot: no information on databases.

Gene	Chr	Start	Ref	Alt	Variants	1000G EUR	1000G AFR	1000G AMR	1000G EAS	1000G SAS	ExAC NFE	ExAC FIN	ExAC AFR	ExAC AMR	ExAC EAS	ExAC SAS	ExAC OTH	ESP EA	ESP AA
FAM179B	14	45495033	A	G	p.T1096A	.	.	.	.	.	.	.	.	.	.	.	.	.	.
HTT	4	3225197	G	C	p.S2512T	.	.	.	.	.	0	0	0.0001	0	0	0	0	.	.
		3240251	C	T	p.P2990L	.	.	.	.	.	.	.	.	.	.	.	.	.	.
INTS4	11	77639382	A	G	c.1371+6T>C	0.002	0.0083	0.0086	0	0	0.0014	0.0014	0.0106	0.0037	0	0	0.0044	0.001	0.0095
KRTAP5-3	11	1629398	C	G	p.C73S	.	.	.	.	.	.	.	.	.	.	.	.	.	.
		1629156	T	A	p.S154C	.	.	.	.	.	1.50E-5	0	0	0	0	0	0	.	.
		1629149	G	C	p.S156C	.	.	.	.	.	.	.	.	.	.	.	.	.	.
		1629119	G	C	p.S166C	.	.	.	.	.	.	.	.	.	.	.	.	.	.
MUC2	11	1096504	C	G	p.L2173V	.	.	.	.	.	6.10E-5	0	0.0001	8.68E-5	0	0	0	.	.
PLEC	8	144996991	T	G	p.E2506A	0	0	0.0014	0	0	0	0	0	0.0005	0	0	0	.	.
		144994768	G	A	p.A3211V	.	.	.	.	.	7.99E-5	0	0	0	0	0	0	.	.
		144991271	C	T	p.A4377T	0	0	0.0014	0.001	0	0.0005	0	0.0001	0.0002	0.0001	0	0.0011	0.0007	0
TTN	2	179629473	G	T	p.R3257S	.	.	.	.	.	0	0	0	8.66E-5	0	0	0	0.0001	0
		179578879	G	A	p.P8836S	.	.	.	.	.	0	0	0	8.71E-5	0	0	0	.	.
ZC3H7A	16	11861374	A	G	p.I474T	0.001	0	0.0014	0	0.001	0.0010	0.0002	9.77E-5	0	0	0.0009	0	0.001	0

Table S3.5B – Genes, variants and predicted functional impact for the analysis of the recessive inheritance model for sample 4. Each functional impact predictor has a different classification with different cut-offs. SIFT: T – tolerated (>0.05), D – damaging (≤0.05); PolyPhen2: B – benign (<0.452), P – possibly damaging (0.453 and 0.956), D – probably damaging (>0.957); LRT: D – deleterious, N – neutral, U – unknown; MutationTaster: P – polymorphism automatic, N – polymorphism, D – disease causing, A – disease causing automatic; MutationAssessor: N – neutral, L – low, M – medium, H – high; FATHMM: T – tolerated (> -1.5), D – damaging (< -1.5); Radial SVM: T – tolerated, D – damaging; LR: T – tolerated, D – damaging; CADD: ≥15 considered pathogenic; GERP++: ≥2 considered conserved; PhyloP: ≥1; HSF: D – disturbs donor splice site; A – disturbs acceptor splice site; No – predicted to not affect splicing. Dot – no prevision was generated. The last column indicates the classification of each variant. A variant is classified as pathogenic when 6 of the 9 functional impact software predict it as damaging, otherwise is classified as benign. A variant is classified as relatively frequent if MAF is > 1% in other populations but European or if there are individuals with homozygous variants, otherwise is classified as rare.

Gene	Variants	SIFT	Polyphen2	LRT	Mutation Taster	Mutation Assessor	FATHMM	Radial SVM	LR	CADD	GERP++	phyloP	HSF	Classification
FAM179B	p.T1096A	T	B	N	D	N	T	T	T	9.737	2.87	0.920	.	Rare and predicted as benign.
HTT	p.S2512T	T	D	D	D	L	T	T	T	25.8	4.88	9.420	.	Rare and predicted as benign.
	p.P2990L	D	B	D	D	L	T	T	T	17.75	4.8	5.738	.	
INTS4	c.1371+6T>C	.	.	.	.	.	.	.	.	.	.	.	No	Relatively frequent and predicted not to affect splicing.
KRTAP5-3	p.C73S	T	B	.	N	N	T	T	T	2.350	-7.27	-1.799	.	Rare and predicted as benign.
	p.S154C	T	B	.	N	N	T	T	T	2.139	-3.61	.	.	
	p.S156C	T	B	.	N	N	T	T	T	0.001	-2.15	.	.	
	p.S166C	T	B	.	N	N	T	T	T	0.001	-2.46	.	.	
MUC2	p.L2173V	T	P	.	N	N	T	T	T	11.60	-0.785	0.044	.	Rare and predicted as benign.
PLEC	p.E2506A	T	B	U	D	M	D	D	D	12.15	5.15	3.295	.	Rare and predicted as benign, except p.E2506A.
	p.A3211V	T	B	U	N	N	T	T	T	6.769	3.54	1.624	.	
	p.A4377T	T	B	U	N	N	T	T	T	1.788	3.31	2.070	.	
TTN	p.R3257S	D	D	.	D	L	T	T	T	18.14	5.69	7.894	.	Rare and predicted as benign.
	p.P8836S	D	D	.	D	M	T	T	T	15.91	5.86	7.695	.	
ZC3H7A	p.I474T	D	P	N	D	L	T	T	T	15.28	5.77	3.501	.	Rare and predicted as benign.

Table S3.6A – Genes, variants and variant frequencies for the analysis of the recessive inheritance pattern for sample 5. All variants have a European MAF inferior to 0.01 in at least in one database. Chr: chromosome, Ref: reference nucleotide, Alt: altered nucleotide, 1000G: 1000 Genomes Project, ExAC: Exome Aggregation Consortium, ESP: NHLBI Exome Sequencing Project, EUR: European, AFR: African, AMR: American, EAS: East Asian, SAS: South Asian, NFE: Non-Finnish European, FIN: Finnish European, OTH: Other populations, EA: European American, AA: African American. Indication of two variants for the same gene represents compound heterozygosity, while a single variant indicates homozygosity. Dot: no information on databases.

Gene	Chr	Start	Ref	Alt	Variants	1000G EUR	1000G AFR	1000G AMR	1000G EAS	1000G SAS	ExAC NFE	ExAC FIN	ExAC AFR	ExAC AMR	ExAC EAS	ExAC SAS	ExAC OTH	ESP6500 EA	ESP6500 AA
CNTNAP4	16	76486444	C	T	p.P298S	.	.	.	.	.	1.50E-5	0	0	0	0	0	0	.	.
		76574606	G	A	p.D1048N	.	.	.	.	.	7.87E-5	0	0	0	0	0	0	.	.
COL15A1	9	101788252	C	T	c.2043+4C>T	.	.	.	.	.	5.72E-5	0	0	0	0.0001	0	0	.	.
		101810229	C	T	p.P914S	.	.	.	.	.	1.78E-5	0	0	0	0	0	0	.	.
FRMD4A	10	13699134	CGCCCCC CGCGCCC CCCG	-	p.A813_G819 del	.	.	.	.	.	0.0114	0	0.0271	0.0018	0.0022	0.0106	0.0161	.	.
GALNT7	4	174169136	C	G	p.D44E	.	.	.	.	.	.	.	.	.	.	.	.	.	.
		174169188	G	A	p.G62R	.	.	.	.	.	.	.	.	.	.	.	.	.	.
LOXHD1	18	44104707	CTT	-	p.K457del														
		44109144	C	T	p.G398E	0.011	0	0.0014	0	0	0.0049	0.0496	0	0	0	0.0008	0.0282	0.0044	0.0014
LRRIQ1	12	85546824	A	C	p.D1481A	0.008	0	0.0014	0	0	0.0036	0.0280	0.0003	0.0006	0	0.0001	0.0045	0.0031	0.0003
		85638660	G	A	p.A1704T	.	.	.	.	.	3.02E-5	0	0	0	0	0	0	.	.
NIN	14	51224761	G	A	p.A996V	0.005	0	0.0043	0	0	0.0034	0.0005	0.0010	0.0014	0	0.0004	0	0.0041	0.0011
		51224101	C	T	p.R1216Q	0.001		0.0014	0	0	0.0005	0	0.0005	0.0005	0	0.0001	0	0.0006	0.0005
OBSCN	1	228404305	G	A	p.R760Q	0.006	0.0008	0.0029	0	0.0051	0.0171	0.0225	0.0028	0.0046	0	0.0087	0.0192	0.0084	0.0007
		228467095	T	C	p.V2878A	0.006	0	0.0014	0	0.0041	0.0071	0.0093	0.0007	0.0015	0	0.0038	0.0046	0.0089	0.0007
PCNT	21	47754488	A	G	p.S149G	.	.	.	.	.	9.24E-5	0	0.0005	0	0.0002	0	0	.	.
		47754510	A	G	p.H156R	0.049	0.014	0.023	0	0.0031	0.0008	0.0003	0.0006	0.0005	0	0.0002	0	0.037	0.0073
PLEC	8	144998302	G	A	p.A2069V	0	0	0	0.001	0	1.66E-5	0	0	0	0.0002	0	0	.	.
		144995500	T	C	p.Y2967C	0.014	0.0015	0	0	0	0.0143	0.0192	0.0025	0.0018	0	0.0008	0.0156	0.012	0.0025
PPP1R32	11	61253294	G	A	p.G200S	0.002	0	0.0014	0	0.016	0.0015	0	0.0003	0.0004	0	0.0152	0.0067	0.0014	0
		61257987	C	G	p.A408G	0.006	0	0	0	0.002	0.0024	0.0012	0.0002	0.0006	0	0.0018	0.0055	0.0017	0.0002
RP1L1	8	10480683	G	A	p.A10V	.	.	.	.	.	1.65E-5	0	0	0	0	0	0	.	.

Gene	Chr	Start	Ref	Alt	Variants	1000G EUR	1000G AFR	1000G AMR	1000G EAS	1000G SAS	ExAC NFE	ExAC FIN	ExAC AFR	ExAC AMR	ExAC EAS	ExAC SAS	ExAC OTH	ESP6500 EA	ESP6500 AA
		10464895	TGGGCCT CCCCTTC AGCCTCC GGGGTCT CTACGCC TTCTGGC TCTGGCT GGGCCTC CTCTTCA GCCTCCG GGGCCTC TACACCT TCTAACT CTGGT	-	p.A2206_Q2238del	.	.	.	.	.	.	.	.	.	.	.	.	.	.
SOWAHD	X	118893328	C	T	p.P233L	0.003	.	.	.	.	0.0014	0.0048	0	0.0026	0	0	0	.	.
STON2	14	81862399	G	A	p.S71F	.	.	.	.	.	.	.	.	.	.	.	.	.	.
		81743617	C	G	p.V680LfsX5	.	.	.	.	.	.	.	.	.	.	.	.	.	.
TTN	2	179642589	C	G	p.R1441P	0.012	0.0008	0.01	0.016	0.016	0.0116	0.0180	0.0012	0.0060	0.0088	0.0174	0.0132	0.0076	0.0023
		179407152	C	G	p.R30803P	0.001	0	0.0014	0	0	0.0003	0	0	0.0010	0	0	0	0.0002	0

Table S3.6B – **Genes, variants and predicted functional impact for the analysis of the recessive inheritance model for sample 5.** Each functional impact predictor has a different classification with different cut-offs. SIFT: T – tolerated (>0.05), D – damaging (≤0.05); PolyPhen2: B – benign (<0.452), P – possibly damaging (0.453 and 0.956), D – probably damaging (>0.957); LRT: D – deleterious, N – neutral, U – unknown; MutationTaster: P – polymorphism automatic, N – polymorphism, D – disease causing, A – disease causing automatic; MutationAssessor: N – neutral, L – low, M – medium, H – high; FATHMM: T – tolerated (> -1.5), D – damaging (< -1.5); Radial SVM: T – tolerated, D – damaging; LR: T – tolerated, D – damaging; CADD: ≥15 considered pathogenic; GERP++: ≥2 considered conserved; PhyloP: ≥1; HSF: D – disturbs donor splice site; A – disturbs acceptor splice site; No – predicted to not affect splicing. Dot – no prevision was generated. The last column indicates the classification of each variant. A variant is classified as pathogenic when 6 of the 9 functional impact software predict it as damaging, otherwise is classified as benign. A variant is classified as relatively frequent if MAF is > 1% in other populations but European or if there are individuals with homozygous variants, otherwise is classified as rare.

Gene	Variants	SIFT	Polyphen 2	LRT	Mutation Taster	Mutation Assessor	FATHMM	Radial SVM	LR	CADD	GERP++	PhyloP	HSF	Classification
CNTNAP4	p.P298S	D	D	D	D	.	T	D	D	17.52	5.52	3.238	.	Rare, p.P298S predicted as pathogenic
	p.D1048N	T	B	N	N	.	T	T	T	7.949	2.48	2.110	.	and p.D1048N as benign.
COL15A1	c.2043+4C>T	.	.	.	.	.	.	.	.	.	.	.	No	Rare, c.2043+4C>T predicted not to affect
	p.P914S	T	D	D	D	M	T	T	T	14.68	5.95	3.389	.	splicing and p.P914S as benign.
FRMD4A	p.A813_G819del	.	.	.	.	.	.	.	.	.	.	.	.	Rare and predicted as pathogenic.
GALNT7	p.D44E	T	B	N	N	N	T	T	T	2.059	2.84	0.742	.	Rare and predicted as benign.
	p.G62R	T	B	U	D	L	T	T	T	14.84	4.86	4.077	.	
LOXHD1	p.K457del	.	.	.	.	.	.	.	.	.	.	.	.	p.G398E relatively frequent and despite
	p.G398E	D	D	D	D	M	T	T	T	17.04	5.14	5.810	.	software predicted as pathogenic in Clinvar is classificated as benign.
LRRIQ1	p.D1481A	D	P	.	N	N	T	T	T	8.389	4.33	2.689	.	Rare and predicted as benign.
	p.A1704T	T	B	.	N	N	T	T	T	11.56	4.77	0.818	.	
NIN	p.A996V	T	B	N	N	N	T	T	T	2.040	-10.5	-0.364	.	Rare and predicted as benign.
	p.R1216Q	T	B	N	N	M	T	T	T	12.98	4.86	0.543	.	
OBSCN	p.R760Q	T	B	N	N	N	T	T	T	7.515	-5.63	-0.013	.	Relatively frequent and predicted as benign.
	p.V2878A	T	P	D	D	M	T	T	T	15.50	4.47	7.798	.	
PCNT	p.S149G	T	P	.	N	L	T	T	T	0.108	0.428	1.190	.	p.H156R relatively frequent and both
	p.H156R	T	B	.	N	N	T	T	T	0.151	0.158	-0.670	.	predicted as benign.
PLEC	p.A2069V	T	B	N	N	N	D	T	T	9.851	-2.08	0.777	.	p.Y2967C relatively frequent and both
	p.Y2967C	T	B	U	D	M	T	T	T	8.770	3.77	1.272	.	predicted as benign.
PPP1R32	p.G200S	D	D	N	D	M	T	D	D	13.79	2.89	3.650	.	Relatively frequent. p.G200S predicted as
	p.A408G	D	B	N	N	L	T	T	T	14.60	-9.41	-1.784	.	pathogenic and p.A408G as benign.
RP1L1	p.A10V	D	P	.	N	L	T	T	T	21.90	-0.702	-0.201	.	p. A2206_Q2238del relatively frequent.
	p.A2206_Q2238del	.	.	.	.	.	.	.	.	.	.	.	.	p.A10V rare and predicted as benign.
SOWAHD	p.P233L	T	P	.	D	L	T	T	T	16.44	4.11	5.512	.	Relatively frequent and predicted as benign.

Gene	Variants	SIFT	Polyphen 2	LRT	Mutation Taster	Mutation Assessor	FATHMM	Radial SVM	LR	CADD	GERP++	PhyloP	HSF	Classification
STON2	p.S71F	D	P	N	N	L	T	T	T	21.4	4.89	2.013	.	Rare and pathogenic.
	p.V680LfsX5	T	P	N	N	L	T	T	T	11.84	5.17	1.433	.	
TTN	p.R1441P	D	D	.	D	L	T	T	T	16.67	5.45	7.584	.	p.R1441P relatively frequent and both
	p.R30803P	D	D	.	D	L	T	T	T	33	6.06	6.072	.	predicted as benign.

Table S3.7A – Genes, variants and variant frequencies for the analysis of the recessive inheritance pattern for sample 9. All variants have a European MAF inferior to 0.01 in at least in one database. Chr: chromosome, Ref: reference nucleotide, Alt: altered nucleotide, 1000G: 1000 Genomes Project, ExAC: Exome Aggregation Consortium, ESP: NHLBI Exome Sequencing Project, EUR: European, AFR: African, AMR: American, EAS: East Asian, SAS: South Asian, NFE: Non-Finnish European, FIN: Finnish European, OTH: Other populations, EA: European American, AA: African American. Indication of two variants for the same gene represents compound heterozygosity, while a single variant indicates homozygosity. Dot: no information on databases.

Gene	Chr	Start	Ref	Alt	Variants	1000G EUR	1000G AFR	1000G AMR	1000G EAS	1000G SAS	ExAC NFE	ExAC FIN	ExAC AFR	ExAC AMR	ExAC EAS	ExAC SAS	ExAC OTH	ESP EA	ESP AA
ABCA4	1	94517254	C	G	p.G863A	0.003	0.0008	0	0	0.002	0.0081	0.0014	0.0016	0.0010	0	0.0022	0.0023	0.0067	0.0023
ABCB6	2	220078565	T	C	p.T521A	.	.	.	.	.	0.0002	0	0	0.0002	0	0	0.0011	0.0002	0
		220078006	C	T	p.G588S	0.004	0.0008	0.0029	0	0	0.0071	0.0039	0.0014	0.0028	0	0.0010	0.0022	0.0066	0.0009
ADGRB3	6	69665952	A	G	p.Q411R	0.001	0.0008	0	0	0	0.0003	0	0.0013	0.0006	0	0	0.0033	0.0002	0.0011
		69728399	A	G	c.2107+8A>G	0.001	0	0	0	0	3.03E-5	0	0	0	0	0	0	.	.
CFAP46	10	134698782	A	C	p.F1151C	.	.	.	.	.	.	.	.	.	.	.	.	.	.
		134627760	T	C	c.7286-2A>G	0.008	0	0.0014	0	0	0.0034	0.0007	0.0004	0.0009	0	0.0003	0.0035	0.0045	0.0005
CYP8B1	3	42916935	G	A	p.S125L	.	.	.	.	.	.	.	.	.	.	.	.	.	.
DNAH11	7	21775289	G	C	p.R2491P	.	.	.	.	.	0.0005	0	0	0	0	9.16E-5	0	0.0004	0.0003
		21781773	A	G	p.N2715D	.	.	.	.	.	4.51E-2	0	0	8.70E-5	0.0001	0	0	.	.
EFR3B	2	25356244	G	C	c.1260-8G>C	.	.	.	.	.	.	.	.	.	.	.	.	.	.
		25366738	G	A	c.2050+7G>A	.	.	.	.	.	.	.	.	.	.	.	.	.	.
HUWE1	X	53574724	C	T	p.A3516T	.	.	.	.	.	.	.	.	.	.	.	.	.	.
IZUMO1	19	49244700	G	A	p.P264S	0.001	0	0	0	0	0.0013	0	9.63E-5	0.0003	0	0.0001	0.0011	0.0021	0.0005
LRP1B	2	141242964	C	T	p.V3125I	.	.	.	.	.	1.50E-5	0	0	0.0018	0	6.06E-5	0	.	.
		141093253	G	A	p.P4016L	0.008	0	0.0072	0	0	0.0049	0.0011	0.0008	0.0018	0	0.0012	0.0044	0.006	0.0018
MAGI1	3	65376943	T	G	p.T764P	0.003	0.0008	0.0029	0	0.0031	0.0033	0	0.0004	0.0011	0	0.0038	0.0022	0.0027	0.0005
MAP4K5	14	50952367	C	T	c.258-7G>A	0.019	0.014	0.017	0	0.002	0.0087	0.0009	0.0049	0.0176	0.0056	0.0295	0.0165	0.013	0.0085
NEURL4	17	7224519	T	A	p.E1091V	0.004	0.012	0.01	0	0.0072	0.0038	0.0003	0.0101	0.0056	0	0.0029	0.0056	0.0042	0.0076
		7222374	G	A	p.L1227F	0.005	0	0	0	0.0031	0.0032	0.0006	0.0001	0.0008	0	0.0053	0.0035	0.0028	0
PTF1A	10	23481728	G	C	p.G90A	0	0.015	0.0029	0	0	0	0	0	0	0	0	0	.	.
RASIP1	19	49242235	C	G	p.G269R	.	.	.	.	.	.	.	.	.	.	.	.	.	.
SFI1	22	31924818	C	T	p.R79X	0	0	0	0	0.002	0.0009	0.0003	0	0.0003	0	0.0035	0.0011	0.0006	0
		31969110	G	A	p.V278M	.	.	.	.	.	7.50E-5	0	0	0.0002	0	0.0001	0	.	.
ZNF541	19	48048986	G	A	p.S267F	.	.	.	.	.	0.0005	0	0	0	0	0.0003	0	.	.

Table S3.7B – Genes, variants and predicted functional impact for the analysis of the recessive inheritance model for sample 9. Each functional impact predictor has a different classification with different cut-offs. SIFT: T – tolerated (>0.05), D – damaging (≤0.05); PolyPhen2: B – benign (<0.452), P – possibly damaging (0.453 and 0.956), D – probably damaging (>0.957); LRT: D – deleterious, N – neutral, U – unknown; MutationTaster: P – polymorphism automatic, N – polymorphism, D – disease causing, A – disease causing automatic; MutationAssessor: N – neutral, L – low, M – medium, H – high; FATHMM: T – tolerated (> -1.5), D – damaging (< -1.5); Radial SVM: T – tolerated, D – damaging; LR: T – tolerated, D – damaging; CADD: ≥15 considered pathogenic; GERP++: ≥2 considered conserved; PhyloP: ≥1; HSF: D – disturbs donor splice site; A – disturbs acceptor splice site; No – predicted to not affect splicing. Dot – no prevision was generated. The last column indicates the classification of each variant. A variant is classified as pathogenic when 6 of the 9 functional impact software predict it as damaging, otherwise is classified as benign. A variant is classified as relatively frequent if MAF is > 1% in other populations but European or if there are individuals with homozygous variants, otherwise is classified as rare.

Gene	Variants	SIFT	Polyphen2	LRT	Mutation Taster	Mutation Assessor	FATHMM	Radial SVM	LR	CADD	GERP++	PhyloP	HSF	Classification
ABCA4	p.G863A	D	P	D	A	M	T	D	T	23.8	4.9	6.798	.	Rare and predicted as pathogenic. Clinvar classification as pathogenic.
ABCB6	p.T521A	T	B	D	D	N	D	T	T	12.89	5.02	6.161	.	p.T521A rare and p.G588S relatively frequent. p.T521A predicted as benign and p.G588S as pathogenic.
	p.G588S	D	D	N	D	L	D	D	D	26.2	3.83	6.673	.	
ADGRB3	p.Q411R	T	B	N	D	L	T	T	T	12.16	4.54	5.068	.	Rare, p.Q411R predicted as benign and c.2107+8A>G as affecting donor splice site.
	c.2107+8A>G	.	.	.	.	.	.	.	.	.	.	.	D	
CFAP46	p.F1151C	D	.	.	N	.	T	T	T	13.40	2.99	3.989	.	Rare, p.F1151C predicted as benign and c.7286-2A>G as affecting acceptor splice site.
	c.7286-2A>G	.	.	.	D	.	.	.	.	9.585	4.28	1.198	A	
CYP8B1	p.S125L	T	B	N	N	N	T	T	T	1.663	-1.56	-0.006	.	Rare and predicted as benign.
DNAH11	p.R2491P	D	D	D	D	.	T	T	T	20.3	5.03	4.984	.	Rare and predicted as pathogenic but gene described with a large number of variants.
	p.N2715D	D	P	D	D	.	T	T	T	18.75	5.58	6.726	.	
EFR3B	c.1260-8G>C	.	.	.	.	.	.	.	.	.	.	.	No	Rare and predicted not to affect splicing.
	c.2050+7G>A	.	.	.	.	.	.	.	.	.	.	.	No	
HUWE1	p.A3516T	T	B	N	N	N	T	T	T	9.959	3.3	0.495	.	Rare and predicted as benign.
IZUMO1	p.P264S	T	B	N	N	L	T	T	T	10.30	-7.21	-0.673	.	Rare and predicted as benign.
LRP1B	p.V3125I	T	B	D	D	N	D	T	D	14.02	5.44	2.797	.	p.V3125I rare and p.P4016L relatively frequent. Both predicted as benign.
	p.P4016L	T	B	D	D	L	D	T	D	21.2	5.49	7.333	.	
MAGI1	p.T764P	T	D	N	D	N	T	T	T	11.14	0.817	-0.215	.	Relatively frequent and predicted as benign.
MAP4K5	c.258-7G>A	.	.	.	.	.	.	.	.	.	.	.	No	Relatively frequent and predicted not to affect splicing.
NEURL4	p.E1091V	D	P	D	D	L	T	T	T	14.12	4.73	6.418	.	Relatively frequent and predicted as benign.
	p.L1227F	T	D	D	D	L	T	T	T	19.22	4.83	5.142	.	
PTF1A	p.G90A	T	B	N	N	N	D	T	T	7.083	1.94	0.923	.	Rare and predicted as benign.
RASIP1	p.G269R	T	B	N	N	N	T	T	T	0.065	3.08	0.056	.	Rare and predicted as benign.
SFI1	p.R79X	T	.	N	D	.	.	.	.	17.84	4.26	1.111	.	Rare. p.R79X predicted as pathogenic and p.V278M predicted as benign.
	p.V278M	T	D	N	N	N	T	T	T	12.16	2.64	1.191	.	

Gene	Variants	SIFT	Polyphen2	LRT	Mutation Taster	Mutation Assessor	FATHMM	Radial SVM	LR	CADD	GERP++	PhyloP	HSF	Classification
ZNF541	p.S267F	D	P	.	D	L	T	T	T	18.48	5.03	4.290	.	Rare and predicted as benign.

Table S3.8A – Genes, variants and variant frequencies for the analysis of the recessive inheritance pattern for sample 18. All variants have a European MAF inferior to 0.01 in at least in one database. Chr: chromosome, Ref: reference nucleotide, Alt: altered nucleotide, 1000G: 1000 Genomes Project, ExAC: Exome Aggregation Consortium, ESP: NHLBI Exome Sequencing Project, EUR: European, AFR: African, AMR: American, EAS: East Asian, SAS: South Asian, NFE: Non-Finnish European, FIN: Finnish European, OTH: Other populations, EA: European American, AA: African American. Indication of two variants for the same gene represents compound heterozygosity, while a single variant indicates homozygosity. Dot: no information on databases.

Gene	Chr	Start	Ref	Alt	Variants	1000G EUR	1000G AFR	1000G AMR	1000G EAS	1000G SAS	ExAC NFE	ExAC FIN	ExAC AFR	ExAC AMR	ExAC EAS	ExAC SAS	ExAC OTH	ESP EA	ESP AA
CELSR2	1	109792928	G	A	p.G76D	.	.	.	.	.	0	0	0	8.69E-5	0	0	0	.	.
		109805791	A	T	p.D1579V	0.002	.	.	.	.	0.0018	0.0009	9.67E-2	0.0012	0	0	0.0022	0.0017	0.0002
CST9	20	23586242	C	T	c.255+5G>A	0.002	0	0.0014	0	0.001	0.0014	0.0003	0.0003	0.0026	0	0.0011	0.0023	0.0028	0.0011
		23584259	T	C	p.N123S	0.001	0	0	0.001	0	0.0001	0.0002	9.61E-5	0.0015	0.0002	6.06E-5	0.0022	0	0.0002
HSPG2	1	22206977	C	T	p.V692M	0.001	0	0	0	0	0.0043	0.0014	0.0007	0.0034	0	0.0005	0.0012	0.0033	0.0009
		22205601	T	C	p.N786S	0.004	0	0.0072	0	0	0.0095	0.0157	0.0012	0.0020	0	0.0017	0.0088	0.009	0.0016
KAT8	16	31129218	GGGCGGGGCCCAGGGCTGGGGGCGGGGCGGAGCTCAGGGCCAGGGGGTG	-	c.210+8_56del	.	.	.	.	.	.	.	.	.	.	.	.	.	.
LRP2	2	170137037	C	T	c.1172-8G>A	0.02	0	0.01	0	0.001	0.0148	0.0037	0.0029	0.0059	0	0.0019	0.0046	0.016	0.0025
		170100029	G	A	p.S1145L	0.001	0	0	0	0.013	0.0006	0	0.0003	0.0003	0	0.0131	0.0033	0.0007	0.0002
LUZP1	1	23419914	G	A	p.R281C	0	0.011	0.0014	0.005	0	0	0	0.0056	8.64E-5	0. 0005	0.0007	0	0	0.0064
		23418576	T	C	p.T727A	0	0.016	0.0029	0	0	8.99E-5	0	0.0109	0.0004	0	0	0	0.0001	0.0093
MUC17	7	100682606	AG	TA	p.S2636T	.	.	.	.	.	.	.	.	.	.	.	.	.	.
		100682611	A	G	p.I2638M	.	.	.	.	.	.	.	.	.	.	.	.	.	.
		100682613	T	C	p.L2639P	0	0	0	0	0.0008	0.0002	0	0.0014	0	6.09E-5	0	0	.	.
NDRG4	16	58537777	A	G	p.I85V	0.001	0	0	0	0	0.0022	0.0009	0.0003	8.68E-5	0	0	0.0033	0.0033	0
		58540522	A	G	c.672+4A>G	0.007	0.0008	0.0029	0	0.0031	0.0060	0	0.0016	0.0025	0	0.0007	0	0.0016	0.0007
OBSCN	1	228462007	C	T	p.R2224W	.	.	.	.	.	0.0005	0.0002	0.0003	0.0004	0	6.07E-5	0	0.0002	0.0002
		228563804	T	C	p.C8585R	.	.	.	.	.	.	.	.	.	.	.	.	.	.
PCDHA12	5	140256468	G	C	p.G471R	0.004	0.0008	0	0	0	0.0025	0.0168	0.0004	0.0003	0	0.0012	0.0088	0.0015	0.0005
		140256675	G	T	p.A540S	0.004	0.014	0	0	0	0.0026	0.0168	0.0101	0.0005	0	0.0013	0.0101	0.0015	0.0082
		140257401	G	A	p.D782N	0.004	0.014	0	0	0	0.0027	0.0177	0.0111	0.0006	0	0.0012	0.0109	0.0015	0.0084

Gene	Chr	Start	Ref	Alt	Variants	1000G EUR	1000G AFR	1000G AMR	1000G EAS	1000G SAS	ExAC NFE	ExAC FIN	ExAC AFR	ExAC AMR	ExAC EAS	ExAC SAS	ExAC OTH	ESP EA	ESP AA
RYR1	19	38939153	TCAGTG GGGTTT GTGGCG CCCTCCC	GTGGG GTTTG TGGCG CCCTCC CTCA	c.957+5_957 +29del	.	.	.	.	.	0.0005	0	9.66E-2	0.0010	0	0	0	.	.
		38973979	C	T	p.D590N	.	.	.	.	.	0	0	0.0004	0	0	0	0	.	.
RYR2	1	237889591	C	G	p.R3570G	.	.	.	.	.	.	.	.	.	.	.	.	.	.
SH3TC1	4	8218840	C	T	p.S262F	.	.	.	.	.	0.0002	0	0	0.0002	0	0	0	0.0005	0
STAB1	3	52550173	G	A	p.E1355K	0.002	0	0.0014	0	0	0.0010	0	0.0003	0.0008	0.0001	0.0002	0.0012	0.0017	0
		52557693	G	C	p.R2439P	0.001	0	0.0014	0	0	0.0007	0	0.0002	0.0003	0.0001	0.0001	0	0.0012	0
WDFY3	4	85642558	A	G	c.7605+4T>C	0.001	0	0.0014	0	0	0.0029	0.0006	0.0012	0.0021	0	0.0004	0	0.0031	0.0007
		85612894	T	C	p.I3032V	0.007	0.0008	0.0043	0	0	0.0110	0.0015	0.0018	0.0051	0	0.0025	0.0066	0.016	0.003

Table S3.8B – Genes, variants and predicted functional impact for the analysis of the recessive inheritance model for sample 18. Each functional impact predictor has a different classification with different cut-offs. SIFT: T – tolerated (>0.05), D – damaging (≤0.05); PolyPhen2: B – benign (<0.452), P – possibly damaging (0.453 and 0.956), D – probably damaging (>0.957); LRT: D – deleterious, N – neutral, U – unknown; MutationTaster: P – polymorphism automatic, N – polymorphism, D – disease causing, A – disease causing automatic; MutationAssessor: N – neutral, L – low, M – medium, H – high; FATHMM: T – tolerated (> -1.5), D – damaging (< -1.5); Radial SVM: T – tolerated, D – damaging; LR: T – tolerated, D – damaging; CADD: ≥15 considered pathogenic; GERP++: ≥2 considered conserved; PhyloP: ≥1; HSF: D – disturbs donor splice site; A – disturbs acceptor splice site; No – predicted to not affect splicing. Dot – no prevision was generated. The last column indicates the classification of each variant. A variant is classified as pathogenic when 6 of the 9 functional impact software predict it as damaging, otherwise is classified as benign. A variant is classified as relatively frequent if MAF is > 1% in other populations but European or if there are individuals with homozygous variants, otherwise is classified as rare.

Gene	Variants	SIFT	Polyphen2	LRT	Mutation Taster	Mutation Assessor	FATHMM	RadialSVM	LR	CADD	GERP++	PhyloP	HSF	Classification
CELSR2	p.G76D	D	B	.	N	N	T	T	T	16.25	5.26	0.386	.	Rare and predicted as benign.
	p.D1579V	T	B	.	D	L	D	T	D	14.17	4.74	3.020	.	
CST9	c.255+5G>A	.	.	.	.	.	.	.	.	.	.	.	D	c.255+5G>A rare and p.N123S relatively frequent. c.255+5G>A predicted as affecting donor splice site and p.N123S predicted as
	p.N123S	T	D	N	N	L	T	T	T	11.73	0.731	0.593	.	

Gene	Variants	SIFT	Polyphen2	LRT	Mutation Taster	Mutation Assessor	FATHMM	RadialSVM	LR	CADD	GERP++	PhyloP	HSF	Classification
														benign.
HSPG2	p.V692M	D	B	N	D	L	T	T	T	20.5	4.53	1.238	.	Relatively frequent and predicted as benign.
	p.N786S	D	D	D	N	M	T	T	T	18.65	4.08	5.239	.	
KAT8	c.210+8_56del	.	.	.	.	.	.	.	.	.	.	.	.	Rare and predicted not to affect splicing.
LRP2	c.1172-8G>A	.	.	.	.	.	.	.	.	.	.	.	No	c.1172-8G>A relatively frequent, predicted not to affect splicing and classified as benign in ClinVar. p.S1145L relatively frequent and predicted as benign.
	p.S1145L	T	B	N	N	L	D	T	D	12.15	-3.79	3.860	.	
LUZP1	p.R281C	T	B	N	D	L	T	T	T	16.53	4.13	1.802	.	Relatively frequent and predicted as benign.
	p.T727A	T	B	N	N	N	T	T	T	0.005	3.81	0.043	.	
MUC17	p.S2636T	.	.	.	.	.	.	.	.	.	.	.	.	Rare and p.I2638M and p.L2639P predicted as benign
	p.I2638M	T	B	.	N	N	T	T	T	0.003	-0.093	-7.273	.	
	p.L2639P	T	B	.	N	N	T	T	T	0.004	-0.093	-1.682	.	
NDRG4	p.I85V	D	B	D	D	M	T	T	T	19.15	4.72	6.966	.	p.I85V relatively frequent in public databases and c.672+4A>G rare. p.I85V predicted as benign and c.672+4A>G predicted to affect donor splice site.
	c.672+4A>G	.	.	.	.	.	.	.	.	.	.	.	D	
OBSCN	p.R2224W	D	D	N	N	M	T	T	T	13.72	1.99	-0.144	.	Rare. p.R2224W predicted as benign and p.C8585R as pathogenic.
	p.C8585R	D	D	.	D	M	T	T	T	28.3	5.74	5.049	.	
PCDHA12	p.G471R	D	D	.	D	H	T	D	D	24.5	4.77	6.229	.	p.A540S and p.D782N relatively frequent and predicted as benign.
	p.A540S	T	P	.	N	N	T	T	T	5.347	0.661	0.350	.	
	p.D782N	T	B	.	N	N	T	T	T	13.59	3.91	3.599	.	
RYR1	c.957+5_957+29del	.	.	.	.	.	.	.	.	.	.	.	D	Rare. c.957+5_957+29del predicted as affecting donor splice site and p.D590N as pathogenic.
	p.D590N	T	P	U	D	L	D	D	D	15.91	5.01	7.743	.	
RYR2	p.R3570G	T	B	U	N	N	D	T	D	1.618	-0.505	-0.177	.	Rare and predicted as benign.
SH3TC1	p.S262F	T	B	N	N	L	T	T	T	9.178	2.03	0.595	.	Rare and predicted as benign.
STAB1	p.E1355K	T	B	D	D	N	D	T	T	11.88	4.07	4.111	.	Rare and predicted as benign.
	p.R2439P	T	B	N	N	N	D	T	T	8.270	-8.73	-2.517	.	

Gene	Variants	SIFT	Polyphen2	LRT	Mutation Taster	Mutation Assessor	FATHMM	RadialSVM	LR	CADD	GERP++	PhyloP	HSF	Classification
WDFY3	c.7605+4T>C p.I3032V	. T	. B	. D	. D	. N	. T	. T	. T	. 0.970	. 4.33	. 5.119	No .	Relatively frequent. c.7605+4T>C predicted not to affect splicing and p.I3032V as benign.

Table S3.9A – Genes, variants and variant frequencies for the analysis of the recessive inheritance pattern for sample 25. All variants have a European MAF inferior to 0.01 in at least in one database. Chr: chromosome, Ref: reference nucleotide, Alt: altered nucleotide, 1000G: 1000 Genomes Project, ExAC: Exome Aggregation Consortium, ESP: NHLBI Exome Sequencing Project, EUR: European, AFR: African, AMR: American, EAS: East Asian, SAS: South Asian, NFE: Non-Finnish European, FIN: Finnish European, OTH: Other populations, EA: European American, AA: African American. Indication of two variants for the same gene represents compound heterozygosity, while a single variant indicates homozygosity. Dot: no information on databases.

Gene	Chr	Start	Ref	Alt	Variants	1000G EUR	1000G AFR	1000G AMR	1000G EAS	1000G SAS	ExAC NFE	ExAC FIN	ExAC AFR	ExAC AMR	ExAC EAS	ExAC SAS	ExAC OTH	ESP EA	ESP AA
ANK3	10	61898817	G	T	p.D15E	.	.	.	.	.	0	0	0	0.0003	0	0	0	.	.
		61819543	C	A	p.D794Y	0.004	0	0.0014	0	0	0.0063	0.0021	0.0012	0.0013	0	0	0.0022	0.007	0.0018
DNAH3	16	20975647	C	T	p.V3187M	.	.	.	.	.	0.0002	0	9.66E-5	0	0	6.06E-5	0	0	0.0002
		20955853	G	C	p.P3826A	.	.	.	.	.	.	.	.	.	.	.	.	.	.
GRIK4	11	120856676	CACCCCCGC CGGCGGCG CGCCG	-	p.H860QfsX188	.	.	.	.	.	.	.	.	.	.	.	.	.	.
HELZ2	20	62198514	G	A	p.R733W	.	.	.	.	.	6.28E-5	0	0	0	0	0.0010	0	0.0002	0
		62193253	G	A	p.P2205L	.	.	.	.	.	.	.	.	.	.	.	.	.	.
HYDIN	16	71096064	A	T	c.2376+8T>A	0	0.0015	0	0	0	0.0002	0	0.0028	0	0	0	0	0.0001	0.0014
		70993542	T	C	c.6142+8A>G	0	0.0015	0	0	0	0	0	0.0007	0	0	0	0	.	.
MACF1	1	39781294	G	A	p.R1132H	0	0	0	0	0	4.50E-5	0	0	8.68E-5	0.0001	0	0	.	.
		39853640	A	C	p.Q2980H	0.004	0.0008	0.0029	0	0	0.0061	0.0050	0.0013	0.0041	0	0.0005	0.0044	0.0067	0.002
MTUS1	8	17541943	T	G	p.K911T	0.0089	0	0.0043	0.001	0.01	0.0099	0.0012	0.0021	0.0067	0.0017	0.0090	0.01	0.0094	0.0023
		17510766	C	G	p.E1105Q	0.006	0.011	0.0029	0.002	0.0082	0.0113	0.0006	0.0104	0.0076	0.0031	0.0075	0.0066	0.0078	0.0075
MYO18B	22	26388398	C	T	p.R2076W	0.008	0.0008	0	0	0.0041	0.0058	0.0068	0.0008	0.0028	0	0.0035	0.0022	0.0075	0.0007
		26264283	C	T	c.3886-4C>T	.	.	.	.	.	.	.	.	.	.	.	.	.	.
OR4M2	15	22368810	C	T	p.P79S	.	.	.	.	.	7.49E-5	0	9.61E-5	8.65E-5	0	0	0	0	0.0002
		22368813	-	AAAT	p.M81IfsX16	0.008	0	0.0043	0	0.0031	0.0080	0.0354	0.0007	0.0034	0	0.0021	0.0100	0.0063	0.0012
PIK3C2G	12	18435554	C	A	p.S180X	0	0	0.0029	0	0	0.0001	0	0	0.0035	0	0	0	.	.
		18719887	C	T	p.P1303S	0.006	0	0.0014	0	0.002	0.0051	0.0191	0.0015	0.0022	0	0.0038	0.0112	0.0048	0.0011
PPP2R3A	3	135720572	G	A	p.G78R	0.008	0	0.0014	0	0	0.0041	0.0003	0.0007	0.0021	0	0.0003	0.0044	0.0037	0.0011
		135720663	A	G	p.N108S	0.002	0	0.011	0	0.001	0.0046	0.0011	0.0009	0.0028	0	0.0017	0.0099	0.005	0.0014
SLC22A18	11	2929502	G	T	p.G62C	0.002	0	0.0029	0	0.001	0.0045	0.0011	0.0005	0.0025	0	0.0013	0.0033	0.0038	0.0005
		2946317	C	T	p.R389C	0.006	0.0015	0.0029	0	0	0.0054	0.0053	0.0014	0.0023	0.0001	0.0010	0.0033	0.0044	0.0011
TTN	2	179464331	G	A	p.A18766V	.	.	.	.	.	.	.	.	.	.	.	.	.	.
		179439719	C	T	p.E23714K	.	.	.	.	.	1.51E-5	0	0	0	0	0	0	.	.

Table S3.9B – Genes, variants and predicted functional impact for the analysis of the recessive inheritance model for sample 25. Each functional impact predictor has a different classification with different cut-offs. SIFT: T – tolerated (>0.05), D – damaging (≤0.05); PolyPhen2: B – benign (<0.452), P – possibly damaging (0.453 and 0.956), D – probably damaging (>0.957); LRT: D – deleterious, N – neutral, U – unknown; MutationTaster: P – polymorphism automatic, N – polymorphism, D – disease causing, A – disease causing automatic; MutationAssessor: N – neutral, L – low, M – medium, H – high; FATHMM: T – tolerated (> -1.5), D – damaging (< -1.5); Radial SVM: T – tolerated, D – damaging; LR: T – tolerated, D – damaging; CADD: ≥15 considered pathogenic; GERP++: ≥2 considered conserved; PhyloP: ≥1; HSF: D – disturbs donor splice site; A – disturbs acceptor splice site; No – predicted to not affect splicing. Dot – no prevision was generated. The last column indicates the classification of each variant. A variant is classified as pathogenic when 6 of the 9 functional impact software predict it as damaging, otherwise is classified as benign. A variant is classified as relatively frequent if MAF is > 1% in other populations but European or if there are individuals with homozygous variants, otherwise is classified as rare.

Gene	Variants	SIFT	Polyphen2	LRT	Mutation Taster	Mutation Assessor	FATHMM	Radial SVM	LR	CADD	GERP++	PhyloP	HSF	Classification
ANK3	p.D15E	T	B	D	D	L	T	T	T	21.2	5.12	5.443	.	p.D15E rare and p.D794Y relatively frequent. Both predicted as benign.
	p.D794Y	T	D	.	D	L	T	T	T	17.21	5.77	5.264	.	
DNAH3	p.V3187M	D	D	N	D	H	T	T	T	16.25	4.07	3.860	.	Rare and predicted as benign.
	p.P3826A	T	D	D	D	M	T	T	T	15.66	5.57	5.376	.	
GRIK4	p.H860QfsX188	.	.	.	.	.	.	.	.	.	.	.	.	Rare and predicted as pathogenic.
HELZ2	p.R733W	D	B	N	N	L	D	T	D	11.88	-1.03	0.651	.	Rare and predicted as beningn.
	p.P2205L	D	B	N	N	M	T	T	T	10.93	-2.49	-0.801	.	
HYDIN	c.2376+8T>A	.	.	.	.	.	.	.	.	.	.	.	No	Rare and predicted not to affect splicing. Gene with a large number of variants and located in a non-conserved region of the genome.
	c.6142+8A>G	.	.	.	.	.	.	.	.	.	.	.	No	
MACF1	p.R1132H	D	P	.	D	M	T	T	T	20.0	4.08	5.350	.	p.R1132H rare and p.Q2980H relatively frequent. Both predicted as pathogenic.
	p.Q2980H	D	D	D	D	M	T	T	T	14.95	2.56	2.046	.	
MTUS1	p.K911T	D	D	N	D	M	T	T	T	17.02	4.06	2.851	.	Relatively frequent and predicted as benign.
	p.E1105Q	T	D	D	D	L	T	T	T	14.54	4.5	4.668	.	
MYO18B	p.R2076W	D	D	D	D	M	D	D	D	21.2	-1.79	2.058	.	p.R2076W relatively frequent c.3886-4C>T rare. p.R2076W predicted as pathogenic and c.3886-4C>T predicted not to affect acceptor splice site.
	c.3886-4C>T	.	.	.	.	.	.	.	.	.	.	.	No	
OR4M2	p.P79S	D	P	D	D	H	T	T	T	17.50	2.5	7.479	.	p.P79S rare and p.M81IfsX16 relatively frequent. Both predicted as pathogenic.
	p.M81IfsX16	.	.	.	.	.	.	.	.	.	.	.	.	
PIK3C2G	p.S180X	T	.	N	A	.	.	.	.	16.78	4.52	2.250	.	p.S180X rare and p.P1303S relatively frequent p.S180X predicted as pathogenic and p.P1303S as benign.
	p.P1303S	T	D	D	D	L	T	T	T	17.27	4.01	2.750	.	
PPP2R3A	p.G78R	D	D	D	D	L	T	T	T	17.94	5.79	2.678	.	Relatively frequent and predicted as

Gene	Variants	SIFT	Polyphen2	LRT	Mutation Taster	Mutation Assessor	FATHMM	Radial SVM	LR	CADD	GERP++	PhyloP	HSF	Classification
	p.N108S	T	B	D	D	N	T	T	T	3.604	5.81	6.070	.	benign.
SLC22A18	p.G62C	D	D	D	D	M	T	D	D	13.97	3.34	2.882	.	Relatively frequent. p.G62C predicted as pathogenic and p.R389C as benign.
	p.R389C	T	D	N	N	M	T	T	T	13.04	2.64	1.153	.	
TTN	p.A18766V	D	D	.	D	M	T	D	D	17.12	5.55	9.793	.	Rare and p.A18766V predicted as pathogenic and p.E23714K as benign.
	p.E23714K	D	B	.	D	L	T	T	T	13.55	5.69	4.682	.	

Table S3.10A – Genes, variants and variant frequencies for the analysis of the recessive inheritance pattern for sample 31. All variants have a European MAF inferior to 0.01 in at least in one database. Chr: chromosome, Ref: reference nucleotide, Alt: altered nucleotide, 1000G: 1000 Genomes Project, ExAC: Exome Aggregation Consortium, ESP: NHLBI Exome Sequencing Project, EUR: European, AFR: African, AMR: American, EAS: East Asian, SAS: South Asian, NFE: Non-Finnish European, FIN: Finnish European, OTH: Other populations, EA: European American, AA: African American. Indication of two variants for the same gene represents compound heterozygosity, while a single variant indicates homozygosity. Dot: no information on databases.

Gene	Chr	Start	Ref	Alt	Variants	1000G EUR	1000G AFR	1000G AMR	1000G EAS	1000G SAS	ExAC NFE	ExAC FIN	ExAC AFR	ExAC AMR	ExAC EAS	ExAC SAS	ExAC OTH	ESP EA	ESP AA
COL6A2	21	47532288	G	A	p.G171R	0.002	0	0.0014	0	0	0.0013	0	0.0001	0.0005	0	0.0021	0	0.0008	0
		47549283	G	A	p.E879K	0	0	0	0	0	0.0008	0.0011	0.0002	0.0030	0.0001	0	0.0023	0.0012	0
DMD	X	31854852	C	T	p.A2387T	0	0	0	0	0.001	0.0012	0	0.0002	0.0003	0	9.88E-5	0	0.0022	0
DOCK11	X	117739235	A	G	p.Q866R	0.006	0	0	0	0	0.0096	0.0064	0.0006	0.0019	0	0.0002	0.0048	0.01	0.0008
FMR1NB	X	147063027	G	C	p.E35D	.	.	.	.	.	.	.	.	.	.	.	.	.	.
LONRF3	X	118109098	G	A	p.E119K	.	.	.	.	.	0.0010	0.0154	0	0	0	0	0	0.0007	0
MAGEC1	X	140995712	G	T	p.S841I	0.013	0	0.0043	0	0	0.0141	0.0135	0.0025	0.0058	0	0.0007	0.0127	0.016	0.0026
MID2	X	107170053	G	A	p.R653H	.	.	.	.	.	.	.	.	.	.	.	.	.	.
RNASEH2C	11	65487516	C	A	c.G468T	0.003	0	0.0014	0	0	0.0037	0.0002	0.0005	0.0027	0	0.0004	0.0055	0.0024	0.0005
TBC1D25	X	48419200	G	A	p.R635H	0.0099	0	0	0	0.0031	0.0135	0.0188	0.0015	0.0011	0	0.0042	0.0063	0.012	0.0016
TMEM151A	11	66061946	G	A	p.A77T	0.002	0	0	0	0	0.0025	0.0004	0.0001	0.0022	0	0.0006	0.0029	0.0018	0
USH1C	11	17554805	T	C	p.H34R	0	0.023	0.0043	0	0	0.0001	0	0.0176	0.0023	0	0.0001	0.0011	0.0001	0.015
		17548863	C	T	p.V135I	0.005	0	0	0	0	0.0034	0.0014	0.0002	0.0018	0.0003	0.0006	0.0033	0.003	0.0009
ZFHX3	16	72992150	C	G	p.G632A	0	0	0	0	0	.	.	.	.	.	.	.	.	0
		72831151	G	T	p.N1810K	0.004	0	0	0	0	0.0003	0	0	0.0003	0	0	0	0.0002	0

Table S3.10B – Genes, variants and predicted functional impact for the analysis of the recessive inheritance model for sample 31. Each functional impact predictor has a different classification with different cut-offs. SIFT: T – tolerated (>0.05), D – damaging (≤0.05); PolyPhen2: B – benign (<0.452), P – possibly damaging (0.453 and 0.956), D – probably damaging (>0.957); LRT: D – deleterious, N – neutral, U – unknown; MutationTaster: P – polymorphism automatic, N – polymorphism, D – disease causing, A – disease causing automatic; MutationAssessor: N – neutral, L – low, M – medium, H – high; FATHMM: T – tolerated (> -1.5), D – damaging (< -1.5); Radial SVM: T – tolerated, D – damaging; LR: T – tolerated, D – damaging; CADD: ≥15 considered pathogenic; GERP++: ≥2 considered conserved; PhyloP: ≥1; HSF: D – disturbs donor splice site; A – disturbs acceptor splice site; No – predicted to not affect splicing. Dot – no prevision was generated. The last column indicates the classification of each variant. A variant is classified as pathogenic when 6 of the 9 functional impact software predict it as damaging, otherwise is classified as benign. A variant is classified as relatively frequent if MAF is > 1% in other populations but European or if there are individuals with homozygous variants, otherwise is classified as rare.

Gene	Variants	SIFT	Polyphen2	LRT	Mutation Taster	Mutation Assessor	FATH MM	Radial SVM	LR	CADD	GERP++	Phylo P	HSF	Classification
COL6A2	p.G171R	D	D	D	D	M	T	D	D	16.59	4.34	3.490	.	Rare and predicted as pathogenic.
	p.E879K	T	P	.	D	.	D	D	D	14.57	4.26	6.366	.	
DMD	p.A2387T	T	B	U	D	L	T	T	T	9.578	4.76	1.902	.	Rare and predicted as benign.
DOCK11	p.Q866R	T	B	D	D	N	T	T	T	10.83	5.43	5.211	.	Relatively frequent and predicted as benign.
FMR1NB	p.E35D	T	B	.	N	N	T	T	T	4.827	-0.023	0.570	.	Rare and predicted as benign.
LONRF3	p.E119K	T	B	N	D	N	T	T	T	9.234	2.67	0.923	.	Rare and predicted as benign.
MAGEC1	p.S841I	D	P	.	N	N	T	T	T	10.73	0.109	-1.338	.	Relatively frequent and predicted as benign.
MID2	p.R653H	D	D	D	D	M	T	D	D	23.5	5.37	9.302	.	Rare and predicted as pathogenic.
RNASEH2C	c.G468T	.	.	.	.	.	.	.	.	.	.	.	D	Relatively frequent and predicted as pathogenic.
TBC1D25	p.R635H	D	P	N	D	L	T	T	T	14.78	4.06	4.003	.	Relatively frequent and predicted as benign.
TMEM151A	p.A77T	T	B	N	D	N	.	T	T	0,678	2.83	6.114	.	Rare and predicted as benign.
USH1C	p.H34R	D	P	D	D	L	T	T	T	24.0	5.07	7.004	.	Rare, predicted as benign and ClinVar classification as benign.
	p.V135I	T	D	D	D	N	T	T	T	27.0	5.13	7.137	.	
ZFHX3	p.G632A	T	B	D	D	L	T	T	T	13.31	5.28	5.649	.	Rare and predicted as benign.
	p.N1810K	T	B	D	D	L	T	T	T	13.61	5.21	0.621	.	

Table S3.11A – Genes, variants and variant frequencies for the analysis of the recessive inheritance pattern for sample 32. All variants have a European MAF inferior to 0.01 in at least in one database. Chr: chromosome, Ref: reference nucleotide, Alt: altered nucleotide, 1000G: 1000 Genomes Project, ExAC: Exome Aggregation Consortium, ESP: NHLBI Exome Sequencing Project, EUR: European, AFR: African, AMR: American, EAS: East Asian, SAS: South Asian, NFE: Non-Finnish European, FIN: Finnish European, OTH: Other populations, EA: European American, AA: African American. Indication of two variants for the same gene represents compound heterozygosity, while a single variant indicates homozygosity. Dot: no information on databases.

Gene	Chr	Start	Ref	Alt	Variants	1000G EUR	1000G AFR	1000G AMR	1000G EAS	1000G SAS	ExAC NFE	ExAC FIN	ExAC AFR	ExAC AMR	ExAC EAS	ExAC SAS	ExAC OTH	ESP EA	ESP AA
BOD1L1	4	13603723	C	T	p.A1601T	.	.	.	.	.	0.0005	0.0003	0	0	0	0	0	.	.
		13600993	T	G	p.T2511P	0.013	0.0008	0.01	0	0.0031	0.0137	0.0127	0.0017	0.0054	0	0.0063	0.0067	0.013	0.0034
FAM205A	9	34729338	C	A	p.W29C	0.008	0	0	0	0.001	0.0045	0.0115	0	0	0	0.0014	0.0082	0.0035	0
		34723250	G	T	p.N1329K	.	.	.	.	.	.	.	.	.	.	.	.	.	.
IGF2R	6	160469510	C	G	p.L817V	0.0099	0.0008	0.0086	0	0.019	0.0092	0.0056	0.0009	0.0062	0	0.0237	0.0154	0.0085	0.0016
		160523631	C	T	p.A2308V	0.001	0	0	0	0	0.0006	0.0030	0	0	0.0001	0	0.0012	0.0005	0
NOTCH4	6	32191658	-	AGCAGC	p.L16_C17insLL	.	.	.	.	.	.	.	.	.	.	.	.	.	.
PRKDC	8	48800245	C	G	p.V1479L	0	0	0	0	0.001	0.0032	0.0004	0.0003	0.0002	0	0.0013	0.0015	0.0015	0
		48792167	A	T	p.L1707Q	0	0	0	0	0.001	0.0063	0.0012	0.0010	0.0015	0	0.0015	0.0023	0.0035	0.0003
SCN4A	17	62038602	T	C	p.H599R	0.002	0.0008	0	0	0	0.0050	0.0024	0.0011	0.0007	0	0.0007	0.0011	0.0048	0.0007
		62026852	C	T	p.V964I	0	0	0.0014	0	0	0.0006	0.0109	0.0001	0.0010	0	7.01E-2	0.0013	0.0006	0
WAPL	10	88259879	T	C	p.Q374R	0	0	0.0014	0	0	0.0009	0.0005	0	0.0006	0	0	0.0011	0.0009	0.0002
		88213414	A	G	c.T2832C	0.002	0.0008	0	0	0	0.0030	0.0015	0.0004	0.0010	0	0.0009	0.0011	0.0036	0.0009
ZIC5	13	100622668	GGC	-	p.P421del	.	.	.	.	.	0.0028	0.0683	0.0098	0	0	0.0023	0.0333	0.02	0.016
ZSCAN29	15	43658785	C	G	p.G249R	0	0	0	0	0.002	0.0001	0	0	0	0	0.0016	0	0.0001	0
		43653413	C	T	p.R806H	0.001	0	0	0	0	0.0014	0.0029	0	0	0	0.0007	0	0.0012	0

Table S3.11B – Genes, variants and predicted functional impact for the analysis of the recessive inheritance model for sample 32. Each functional impact predictor has a different classification with different cut-offs. SIFT: T – tolerated (>0.05), D – damaging (≤0.05); PolyPhen2: B – benign (<0.452), P – possibly damaging (0.453 and 0.956), D – probably damaging (>0.957); LRT: D – deleterious, N – neutral, U – unknown; MutationTaster: P – polymorphism automatic, N – polymorphism, D – disease causing, A – disease causing automatic; MutationAssessor: N – neutral, L – low, M – medium, H – high; FATHMM: T – tolerated (> -1.5), D – damaging (< -1.5); Radial SVM: T – tolerated, D – damaging; LR: T – tolerated, D – damaging; CADD: ≥15 considered pathogenic; GERP++: ≥2 considered conserved; PhyloP: ≥1; HSF: D – disturbs donor splice site; A – disturbs acceptor splice site; No – predicted to not affect splicing. Dot – no prevision was generated. The last column indicates the classification of each variant. A variant is classified as pathogenic when 6 of the 9 functional impact software predict it as

damaging, otherwise is classified as benign. A variant is classified as relatively frequent if MAF is > 1% in other populations but European or if there are individuals with homozygous variants, otherwise is classified as rare.

Gene	Variants	SIFT	Polyphen 2	LRT	Mutation Taster	Mutation Assessor	FATHM M	Radial SVM	LR	CADD	GERP++	PhyloP	HSF	Classification
BOD1L1	p.A1601T	D	P	D	D	L	T	T	T	18.15	5.23	2.279	.	p.A1601T rare and p.T2511P relatively frequent. Both predicted as benign.
	p.T2511P	T	B	N	N	N	T	T	T	7.254	0.241	-0.930	.	
FAM205A	p.W29C	D	D	.	D	M	T	T	T	17.02	4.44	3.403	.	Rare, p.W29C predicted as pathogenic and p.N1329K as benign.
	p.N1329K	D	B	.	N	N	T	T	T	9.128	-2.32	-0.278	.	
IGF2R	p.L817V	T	P	D	D	M	T	T	T	16.02	4.31	1.403	.	p.L817V relatively frequent and p.A2308V rare. Both predicted as benign.
	p.A2308V	T	B	N	N	M	T	T	T	12.06	4.56	4.292	.	
NOTCH4	p.L16_C17insLL	.	.	.	.	.	.	.	.	.	.	.	.	Rare but variant number of triplets in this region.
PRKDC	p.V1479L	T	B	N	N	N	T	T	T	0.015	-11.1	-3.031	.	Relatively frequent. p.V1479L predicted as benign and p.L1707Q as pathogenic.
	p.L1707Q	D	D	D	D	M	T	T	T	17.03	5.87	8.826	.	
SCN4A	p.H599R	D	B	D	D	M	D	D	D	26.8	4.29	7.865	.	p.H599R relatively frequent and p.V964I rare. p.H599R predict as pathogenic and p.V964I as benign.
	p.V964I	T	B	D	D	N	D	T	T	16.79	4.53	0.565	.	
WAPL	p.Q374R	D	D	D	D	L	T	T	T	17.93	5.5	7.698	.	Rare and p.Q374R predicted as benign and c.T2832C as affecting donor splice site.
	c.T2832C	.	.	.	.	.	.	.	.	.	.	.	D	
ZIC5	p.P421del	.	.	.	.	.	.	.	.	.	.	.	.	Relatively frequent.
ZSCAN29	p.G249R	D	D	D	D	M	T	T	T	17.20	3.74	0.454	.	Rare. p.G249R predicted as pathogenic and p.R806H as benign.
	p.R806H	T	B	N	N	L	T	T	T	14.36	4.13	0.161	.	

Table S3.12A – Genes, variants and variant frequencies for the analysis of the recessive inheritance pattern for sample 33. All variants have a European MAF inferior to 0.01 in at least in one database. Chr: chromosome, Ref: reference nucleotide, Alt: altered nucleotide, 1000G: 1000 Genomes Project, ExAC: Exome Aggregation Consortium, ESP: NHLBI Exome Sequencing Project, EUR: European, AFR: African, AMR: American, EAS: East Asian, SAS: South Asian, NFE: Non-Finnish European, FIN: Finnish European, OTH: Other populations, EA: European American, AA: African American. Indication of two variants for the same gene represents compound heterozygosity, while a single variant indicates homozygosity. Dot: no information on databases.

Gene	Chr	Start	Ref	Alt	Variants	1000G EUR	1000G AFR	1000G AMR	1000G EAS	1000G SAS	ExAC NFE	ExAC FIN	ExAC AFR	ExAC AMR	ExAC EAS	ExAC SAS	ExAC OTH	ESP EA	ESP AA
ARSE	X	2867685	C	G	p.G172R	.	.	.	.	.	.	.	.	.	.	.	.	.	.
BCOR	X	39922279	G	A	p.S1298F	.	.	.	.	.	0	0	0.0008	0	0	0	0	.	.
CFAP47	X	35970001	A	G	p.I323V	.	.	.	.	.	.	.	.	.	.	.	.	.	.
COL4A5	X	107829855	G	A	p.R348K	.	.	.	.	.	0.0001	0	0	0	0	0	0	0.0001	0
CXorf67	X	51150304	C	T	p.R146C	.	.	.	.	.	.	.	.	.	.	.	.	.	.
GPC3	X	132826404	C	T	p.V413M	0.004	0.0015	0.0072	0.002	0.0041	0.0042	0.0002	0.0013	0.0241	0.0014	0.0061	0.0047	0.0037	0.001
H6PD	1	9307018	C	G	c.628-7C>G	.	.	.	.	.	0.0017	0	9.62E-5	0.0021	0.0045	6.06E-5	0.0011	0.0026	0.0009
		9323894	G	A	p.A459T	0.002	0	0.0014	0	0.001	0.0012	0	0.0003	0.0016	0.0001	0.0007	0	0.0014	0.0007
IRS1	2	227661003	C	G	p.G818R	0.015	0.0008	0.0029	0	0.002	0.0108	0.0216	0.0026	0.0042	0	0.0027	0.0166	0.013	0.0032
		227659793	C	T	p.R1221H	.	.	.	.	.	1.54E-5	0	0	0	0	0	0	.	.
LAMA5	20	60885808	G	C	p.H3453Q	0	0	0	0	0.001	0.0004	0.0003	0	0	0	0.0009	0	.	.
		60885242	C	T	p.E3576K	0	0	0.0014	0	0.001	0.0043	0.0031	0.0014	0.0016	0	0.0009	0	0.0058	0.0014
LAMC2	1	183184663	G	A	p.R115Q	0.016	0	0.0014	0	0.0031	0.0098	0.0398	0.0011	0.0064	0	0.0092	0.0198	0.009	0.0009
		183205708	G	C	p.R857P	0.007	0	0.0043	0	0	0.0064	0.0024	0.0013	0.0025	0	0.0002	0.0033	0.0049	0.0014
MAP3K15	X	19433335	G	A	p.R494C	0.006	0.0008	0	0		0.0199	0.0130	0.0018	0.0062	0.0026	0.0017	0	0.012	0.0008
MRGPRG	11	3239878	C	G	p.A56P	.	.	.	.	.	0.0108	0.0021	0.0037	0.0074	0	0.0043	0	.	.
MTMR8	X	63488525	G	T	p.N669K	0.008	0	0.0058	0	0	0.0082	0.0450	0.0008	0.0016	0	0.0007	0.0223	0.0039	0.001
		63488524	A	T	p.L670M	0.008	0	0.0058	0	0	0.0081	0.0450	0.0008	0.0016	0	0.0007	0.0223	0.0037	0.001
NLRP13	19	56443467	G	T	p.L71I	.	.	.	.	.	6.00E-2	0	0.0002	0.0008	0	0	0	0.0001	0
		56424305	T	C	p.D293G	0.006	0	0.0029	0	0.0041	0.0020	0.0006	0.0002	0.0003	0	0.0012	0	0.0019	0.0002
OGT	X	70776796	T	-	c.1167-6delT	0.008	0	0.0029	0	0.002	0.0067	0.0069	0.0013	0.0037	0	0.0028	0.0017	0.0069	0.0027
OR11L1	1	248004480	G	T	p.T240K	.	.	.	.	.	3.00E-5	0	0.0005	0	0	0	0	0	0.0002
		248004389	G	C	p.I270M	0	0.003	0	0	0	3.00E-5	0	0.0037	0	0	0	0	0	0.005

Gene	Chr	Start	Ref	Alt	Variants	1000G EUR	1000G AFR	1000G AMR	1000G EAS	1000G SAS	ExAC NFE	ExAC FIN	ExAC AFR	ExAC AMR	ExAC EAS	ExAC SAS	ExAC OTH	ESP EA	ESP AA
PHF8	X	53970659	G	C	p.L889V	.	.	.	.	.	.	.	.	.	.	.	.	.	.
RP1L1	8	10469454	G	T	p.N718K	0.001	0	0.0043	0	0	0.0007	0.0003	0	0.0005	0	0	0.0011	0.0007	0.0002
		10468756	G	A	p.S951L	0	0.0008	0	0	0	1.56E-5	0	0.0001	0	0	0	0	0.0001	0
SLC38A5	X	48318119	G	A	c.1212C>T	.	.	.	.	.	4.48E-5	0	0	0	0	0	0	.	.
TLR10	4	38776712	A	G	p.L167P	0.001	0	0.0014	0	0	0.0008	0.0003	0	0.0006	0	0.0007	0.0011	0.0012	0
		38774816	C	T	p.R799Q	0.006	0	0.0029	0	0.0092	0.0060	0.0009	0.0005	0.0047	0	0.0080	0.0088	0.0065	0.0007
TTLL5	14	76135756	T	C	c.75-3T>C	0.012	0	0	0	0.028	0.0175	0.0160	0.0025	0.0048	0	0.0363	0.0165	0.017	0.0039
		76249759	A	G	p.T958A	0.012	0	0	0	0.019	0.0177	0.0122	0.0033	0.0051	0	0.0285	0.0154	0.018	0.0043
USF3	3	113376117	TGC	-	p.Q147del	.	.	.	.	.	.	.	.	.	.	.	.	.	.
VPS52	6	33219675	A	G	p.I623T	0	0	0	0.008	0	0.0029	0.004	0.0061	0.0029	0.0005	0.0001	0.0018	0.0023	0.0079
ZIC5	13	100622668	GGC	-	p.P421del	.	.	.	.	.	.	.	.	0.0028	0.0683	0.0098	0	0	0.0023

Table S3.12B – **Genes, variants and predicted functional impact for the analysis of the recessive inheritance model for sample 33.** Each functional impact predictor has a different classification with different cut-offs. SIFT: T – tolerated (>0.05), D – damaging (≤0.05); PolyPhen2: B – benign (<0.452), P – possibly damaging (0.453 and 0.956), D – probably damaging (>0.957); LRT: D – deleterious, N – neutral, U – unknown; MutationTaster: P – polymorphism automatic, N – polymorphism, D – disease causing, A – disease causing automatic; MutationAssessor: N – neutral, L – low, M – medium, H – high; FATHMM: T – tolerated (> -1.5), D – damaging (< -1.5); Radial SVM: T – tolerated, D – damaging; LR: T – tolerated, D – damaging; CADD: ≥15 considered pathogenic; GERP++: ≥2 considered conserved; PhyloP: ≥1; HSF: D – disturbs donor splice site; A – disturbs acceptor splice site; No – predicted to not affect splicing. Dot – no prevision was generated. The last column indicates the classification of each variant. A variant is classified as pathogenic when 6 of the 9 functional impact software predict it as damaging, otherwise is classified as benign. A variant is classified as relatively frequent if MAF is > 1% in other populations but European or if there are individuals with homozygous variants, otherwise is classified as rare.

Gene	Variants	SIFT	Polyphen2	LRT	Mutation Taster	Mutation Assessor	FATHMM	Radial SVM	LR	CADD	GERP++	PhyloP	HSF	Classification
ARSE	p.G172R	D	D	D	D	H	D	D	D	14.92	3.56	6.351	.	Rare and predicted as pathogenic.
BCOR	p.S1298F	D	P	.	D	L	T	T	T	20.7	5.87	7.909	.	Rare and predicted as benign.
CFAP47	p.I323V	T	B	N	N	N	T	T	T	3.848	-2.29	0.219	.	Rare and predicted as benign.
COL4A5	p.R348K	T	B	N	N	N	D	T	T	4.467	2.28	1.527	.	Rare and predicted as benign.

Gene	Variants	SIFT	Polyphen2	LRT	Mutation Taster	Mutation Assessor	FATHMM	Radial SVM	LR	CADD	GERP++	PhyloP	HSF	Classification
CXorf67	p.R146C	D	B	N	N	.	T	T	T	11.75	-0.884	0.184	.	Rare and predicted as benign.
GPC3	p.V413M	D	D	D	D	M	T	T	T	18.09	5.07	3.465	.	Relatively frequent and predicted as pathogenic.
H6PD	c.628-7C>G	.	.	.	.	.	.	.	.	.	.	.	A	Relatively frequent. c.628-7C>G predicted as affecting acceptor splice site and p.A459T as benign.
	p.A459T	T	B	N	N	L	D	D	D	9.131	4.47	1.630	.	
IRS1	p.G818R	T	D	D	D	L	T	T	T	16.79	4.85	3.784	.	p.G818R relatively frequent and p.R1221H rare. Both predicted as benign.
	p.R1221H	T	B	N	N	N	T	T	T	14.86	4.82	2.474	.	
LAMA5	p.H3453Q	T	B	U	N	L	T	T	T	7.907	-2.36	-0.326	.	p.H3453Q rare add p.E3576K relatively frequent. Both predicted as benign.
	p.E3576K	T	B	U	N	M	T	T	T	9.725	1.5	0.574	.	
LAMC2	p.R115Q	T	B	N	D	L	T	T	T	15.65	2.99	2.045	.	Relatively frequent and predicted as benign.
	p.R857P	T	B	N	N	N	T	T	T	12.62	3.44	2.222	.	
MAP3K15	p.R494C	D	.	D	D	L	T	T	T	13.06	3.53	3.450	.	Relatively frequent and predicted as benign.
MRGPRG	p.A56P	D	.	.	N	M	T	T	D	15.81	2.02	1.709	.	Relatively frequent and predicted as benign.
MTMR8	p.N669K	D	B	.	N	N	D	T	T	10.26	-0.974	0.230	.	Relatively frequent and predicted as benign.
	p.L670M	T	B	.	N	N	D	T	T	0.026	-0.975	-0.269	.	
NLRP13	p.L71I	T	P	.	N	M	T	T	T	11.98	0.898	0.180	.	Rare and predicted as benign.
	p.D293G	T	B	.	N	N	T	T	T	8.961	1.76	-0.232	.	
OGT	c.1167-6delT	.	.	.	.	.	.	.	.	.	.	.	A	Relatively frequent and predicted as affecting acceptor splice site.
OR11L1	p.T240K	D	D	U	D	H	T	D	T	16.60	3.51	6.745	.	Rare. p.T240K predicted as pathogenic and p.I270M as benign.
	p.I270M	T	B	U	N	N	T	T	T	0	2.16	-0.021	.	
PHF8	p.L889V	T	D	N	N	N	T	T	T	6.884	4.0	0.871	.	Rare and predicted as benign.
RP1L1	p.N718K	T	B	N	N	N	T	T	T	11.41	-0.039	0.001	.	Rare and predicted as benign.
	p.S951L	D	B	N	N	N	T	T	T	12.64	0.597	3.208	.	
SLC38A5	c.1212C>T	.	.	.	.	.	.	.	.	.	.	.	D	Rare and predicted to affect donor splice site.
TLR10	p.L167P	D	D	N	D	M	T	T	T	17.20	5.23	6.785	.	p.L167P rare and p.R799Q relatively frequent.

Gene	Variants	SIFT	Polyphen2	LRT	Mutation Taster	Mutation Assessor	FATHMM	Radial SVM	LR	CADD	GERP++	PhyloP	HSF	Classification
	p.R799Q	T	B	N	N	N	T	T	T	9.637	-1.62	-0.532	.	Both predicted as benign.
TTLL5	c.75-3T>C	.	.	.	.	.	.	.	.	.	.	.	No	Relatively frequent. c.75-3T>C predicted not to affect splicing and p.T958A predicted as benign.
	p.T958A	T	B	N	N	N	T	T	T	0.404	4.03	2.672	.	
USF3	p.Q147del	.	.	.	.	.	.	.	.	.	.	.	.	Rare but located in a region with a variable number of triplets.
VPS52	p.I623T	T	B	D	D	N	.	T	T	14.30	3.67	5.795	.	Relatively frequent and predicted as benign.
ZIC5	p.P421del	.	.	.	.	.	.	.	.	.	.	.	.	Relatively frequent and may be pathogenic.

Table S3.13A – Genes, variants and variant frequencies for the analysis of the recessive inheritance pattern for sample 36. All variants have a European MAF inferior to 0.01 in at least in one database. Chr: chromosome, Ref: reference nucleotide, Alt: altered nucleotide, 1000G: 1000 Genomes Project, ExAC: Exome Aggregation Consortium, ESP: NHLBI Exome Sequencing Project, EUR: European, AFR: African, AMR: American, EAS: East Asian, SAS: South Asian, NFE: Non-Finnish European, FIN: Finnish European, OTH: Other populations, EA: European American, AA: African American. Indication of two variants for the same gene represents compound heterozygosity, while a single variant indicates homozygosity. Dot: no information on databases.

Gene	Chr	Start	Ref	Alt	Variants	1000G EUR	1000G AFR	1000G AMR	1000G EAS	1000G SAS	ExAC NFE	ExAC FIN	ExAC AFR	ExAC AMR	ExAC EAS	ExAC SAS	ExAC OTH	ESP EA	ESP AA
AFTPH	2	64778749	C	A	p.F47L	0.001	0.0015	0	0	0	0.0003	0	0.0016	0.0012	0	0	0	0.0002	0.0016
		64780257	A	G	p.N550S	0.001	0.021	0.0043	0.002	0	0.0004	0.0002	0.0148	0.0016	0.0030	6.06E-5	0	0.0003	0.017
CDH23	10	73567371	G	A	p.V2803I	0	0.0015	0	0	0	9.12E-5	0	0.0008	0.0002	0	0	0	0	0.0002
		73574728	A	C	p.D3253A	0.001	0.018	0.0014	0	0.01	0.0008	0	0.0158	0.0019	0	0.0117	0.0068	0.0001	0.0079
CLCA4	1	87041255	G	A	p.V642I	0	0	0	0	0.0061	0.0010	0.0003	0	0.0005	0	0.0121	0.0011	0.0006	0
		87045896	TCCTAC ACCTAC	-	p.P877_T880del	.	.	.	.	.	0.0012	0.0008	0.0003	0.0011	0	0.0022	0.0011	0.71	0.63
ELF4	X	129200878	G	A	p.R604C	0.003	0	0.0029	0	0	0.0045	0.0020	0.0005	0.0013	0.0023	0.0006	0.0048	0.0036	0.0008
FILIP1L	3	99569771	G	T	p.T250K	.	.	.	.	.	.	.	.	.	.	.	.	.	.
		99567954	G	A	p.P856S	.	.	.	.	.	1.50E-5	0	0	0	0	0.0014	0	.	.
GOLGB1	3	121415714	C	G	p.R1219P	0.005	0	0.0043	0	0.002	0.0077	0.0008	0.0023	0.0042	0	0.0021	0.0033	0.0098	0.0025
		121416813	G	A	c.C2557T	0.015	0	0.0072	0.001	0.002	0.0126	0.0044	0.0038	0.0061	0.0002	0.0032	0.0077	0.016	0.0045
JPH3	16	87678514	G	A	p.G345S	0.002	0	0.0014	0	0	0.0014	0	0.0009	0.0008	0.0001	6.07E-5	0.0011	0.002	0.0011
		87678545	G	A	p.R355Q	.	.	.	.	.	3.09E-5	0	0.0001	0	0	0	0	0	0.0002
KAT8	16	31129218	GGGCG GGGCCC AGGGCT GGGGG CGGGG CGGAGC TCAGGG CCAGGG GGTG	-	c.210+8_56del GGGCGGGGCCC AGGGCTGGGGG CGGGGCGGAGC TCAGGGCCAGG GGGTG	.	.	.	.	.	.	.	.	.	.	.	.	.	.
MAGEB1	X	30269187	T	G	p.W193G	.	.	.	.	.	2.31E-5	0.0002	0	0.0005	0.0015	0	0	.	.
PCDHGA8	5	140772679	C	T	p.P100L	0.0099	0	0.011	0	0.001	0.0042	0.0002	0.0014	0.0069	0	6.06E-5	0.0056	0.0054	0.0012
		140773719	G	A	p.D447N	0.0099	0	0.011	0	0	0.0043	0.0006	0.0013	0.0071	0	0	0.0056	0.0058	0.0009
		140774131	GT	CC	p.R584P	.	.	.	.	.	.	.	.	.	.	.	.	.	.
PKD1	16	2153618	C	T	p.G2814R	0.005	0.0023	0.0058	0	0	0.0075	0.0156	0.0016	0.0045	0.0006	6.06E-5	0.0080	0.0083	0.0014
		2153345	C	T	p.V2905I	0.005	0.0008	0.0014	0	0	0.0021	0.0003	0.0007	0.0026	0.0001	0.0002	0	0.0026	0.001

Gene	Chr	Start	Ref	Alt	Variants	1000G EUR	1000G AFR	1000G AMR	1000G EAS	1000G SAS	ExAC NFE	ExAC FIN	ExAC AFR	ExAC AMR	ExAC EAS	ExAC SAS	ExAC OTH	ESP EA	ESP AA
PLD2	17	4719927	C	T	p.R490C	0.001	0.015	0.0043	0	0	0.0004	0	0.0094	0.0015	0	0	0.0011	0.0003	0.0077
		4721908	C	A	c.2123+6C>A	.	.	.	.	.	6.01E-5	0	0	0.0003	0	0	0	.	.
PPL	16	4960852	C	G	c.161G>C	0.001	0	0	0	0.001	0.0016	0	0.0004	0.0010	0	0.0024	0	0.0017	0.0005
		4940928	G	A	c.1969-6C>T	0.002	0	0	0	0.001	0.0023	0	0	0.0027	0	0.0032	0	0.0014	0
PRRG1	X	37312617	C	G	p.P134A	0	0	0	0	0.0051	0.0016	0.0002	0.0002	0.0003	0.0003	0.0047	0	0.0018	0
R3HCC1L	10	99968670	A	G	p.T267A	0.005	0	0.0029	0	0.03	0.0080	0.0014	0.0011	0.0047	0	0.0205	0.0066	0.0071	0.0011
		99969076	C	T	p.A402V	0	0.023	0.0029	0	0	0	0	0.0277	0.0011	0	0	0.0011	0.0001	0.03
SIPA1L3	19	38573053	C	T	p.A283V	0.005	0.0008	0.0029	0	0	0.0097	0.0343	0.0012	0.0029	0	0.0022	0.0070	0.005	0.0009
		38610443	A	G	p.Y930C	.	.	.	.	.	0.0003	0	0	0	0	0	0	0.0002	0
SLC2A4	17	7189048	G	A	p.V383I	0.007	0	0.0043	0	0.001	0.0059	0.0002	0.0012	0.0041	0.0002	0.0031	0.0022	0.005	0.0023
TTN	2	179595372	T	C	p.E5646G	0.0099	0	0.0014	0	0	0.0112	0.0039	0.0014	0.0005	0	0.0028	0.0022	0.01	0.0024
		179583262	C	T	p.E7874K	.	.	.	.	.	6.01E-05	0	0	0	0	6.07E-05	0	.	.
		179404402	T	C	p.N31156S	0.013	0	0.01	0	0	0.0067	0.0009	0.0006	0.0042	0	0.0009	0.0033	0.0081	0.001
XPNPEP2	X	128884502	G	T	p.K232N	0.0089	0.0008	0.0014	0	0	0.0134	0.0029	0.0014	0.0036	0.0004	0.0011	0.0043	0.007	0.0016
ZNF630	X	47918590	C	T	p.R414Q	.	.	.	.	.	0.0002	0	0	0	0	0	0	.	.

Table S3.13B – Genes, variants and predicted functional impact for the analysis of the recessive inheritance model for sample 36. Each functional impact predictor has a different classification with different cut-offs. SIFT: T – tolerated (>0.05), D – damaging (≤0.05); PolyPhen2: B – benign (<0.452), P – possibly damaging (0.453 and 0.956), D – probably damaging (>0.957); LRT: D – deleterious, N – neutral, U – unknown; MutationTaster: P – polymorphism automatic, N – polymorphism, D – disease causing, A – disease causing automatic; MutationAssessor: N – neutral, L – low, M – medium, H – high; FATHMM: T – tolerated (> -1.5), D – damaging (< -1.5); Radial SVM: T – tolerated, D – damaging; LR: T – tolerated, D – damaging; CADD: ≥15 considered pathogenic; GERP++: ≥2 considered conserved; PhyloP: ≥1; HSF: D – disturbs donor splice site; A – disturbs acceptor splice site; No – predicted to not affect splicing. Dot – no prevision was generated. The last column indicates the classification of each variant. A variant is classified as pathogenic when 6 of the 9 functional impact software predict it as damaging, otherwise is classified as benign. A variant is classified as relatively frequent if MAF is > 1% in other populations but European or if there are individuals with homozygous variants, otherwise is classified as rare.

Gene	Variants	SIFT	Polyphen2	LRT	Mutation Taster	Mutation Assessor	FATHMM	Radial SVM	LR	CADD	GERP++	PhyloP	HSF	Classification
AFTPH	p.F47L	D	P	D	D	L	T	T	T	14.37	1.99	1.152	.	Rare and both predicted as benign.
	p.N550S	T	B	N	N	N	T	T	T	0.006	1.61	0.081	.	
CDH23	p.V2803I	T	B	D	D	N	T	T	T	17.51	5.6	5.532	.	Rare and predicted as benign. p.D3253A with ClinVar classification as benign.
	p.D3253A	D	B	N	D	L	T	T	T	16.31	4.15	6.289	.	

Gene	Variants	SIFT	Polyphen2	LRT	Mutation Taster	Mutation Assessor	FATHMM	Radial SVM	LR	CADD	GERP++	PhyloP	HSF	Classification
CLCA4	p.V642I	T	P	N	N	L	T	T	T	15.38	2.6	0.036	.	p.V642I relatively frequent and p.P877_T880del rare. p.V642I predicted as benign and p.P877_T880del as pathogenic.
	p.P877_T880del	.	.	.	.	.	.	.	.	.	.	.	.	
ELF4	p.R604C	T	B	N	N	N	T	T	T	12.39	-1.96	0.124	.	Relatively frequent and predicted as benign.
FILIP1L	p.T250K	T	B	N	D	L	T	T	T	15.33	6.17	6.679	.	Rare and predicted as benign.
	p.P856S	T	B	D	N	N	T	T	T	0.007	5.84	4.089	.	
GOLGB1	p.R1219P	T	B	N	N	N	T	T	T	5.811	-7.43	-2.474	.	Relatively frequent and predicted as benign.
	c.C2557T	D	P	N	N	L	T	T	T	10.66	4.3	1.479	.	
JPH3	p.G345S	T	B	N	N	N	T	T	T	2.860	4.56	4.047	.	Rare, p.G345S predicted as benign and p.R355Q as pathogenic.
	p.R355Q	D	D	D	D	M	T	T	T	31	4.43	6.520	.	
KAT8	c.210+8_56del GGGCGGGGCCC AGGGCTGGGGG CGGGGCGGAGC TCAGGGCCAGG GGGTG	.	.	.	.	.	.	.	.	.	.	.	No	Rare and predicted not to affect splicing.
MAGEB1	p.W193G	T	B	N	N	N	T	T	T	3.051	-4.42	-2.305	.	Rare and predicted as benign.
PCDHGA8	p.P100L	T	B	U	N	L	T	T	T	0.323	1.38	-0.094	.	p.P100L and p.D447N relatively frequent and p.R584P rare. p.P100L predicted as benign and p.D447N as pathogenic.
	p.D447N	D	D	U	D	H	T	D	D	17.76	5.06	9.777	.	
	p.R584P	.	.	.	.	.	.	.	.	.	.	.	.	
PKD1	p.G2814R	T	P	N	N	L	T	T	T	12.64	3.59	2.229	.	Relatively frequent and predicted as benign.
	p.V2905I	T	B	N	D	L	T	T	T	13.38	3.95	2.322	.	
PLD2	p.R490C	T	B	N	N	N	T	T	T	1.286	-0.585	-0.366	.	p.R490C relatively frequent and c.2123+6C>A rare. p.R490C predicted as benign and c.2123+6C>A not to affect splicing.
	c.2123+6C>A	.	.	.	.	.	.	.	.	.	.	.	No	
PPL	c.161G>C	T	B	D	N	L	T	T	T	10.96	1.74	1.134	D	Relatively frequent. c.161G>C predicted to affect donor splice site and c.1969-6C>T not to affect splicing.
	c.1969-6C>T	.	.	.	.	.	.	.	.	.	.	.	No	
PRRG1	p.P134A	T	B	N	D	N	D	D	D	11.67	4.34	1.215	.	Relatively frequent and predicted as benign.
R3HCC1L	p.T267A	.	B	N	N	L	T	T	T	2.291	-7.45	-1.625	.	Relatively frequent and predicted as benign.
	p.A402V	.	B	N	N	M	T	T	T	5.558	1.34	0.162	.	

Gene	Variants	SIFT	Polyphen2	LRT	Mutation Taster	Mutation Assessor	FATHMM	Radial SVM	LR	CADD	GERP++	PhyloP	HSF	Classification
SIPA1L3	p.A283V	T	B	N	N	N	T	T	T	8.368	0.013	1.198	.	p.A283V relatively frequent and p.Y930C rare. p.A283V predicted as benign and p.Y930C as pathogenic.
	p.Y930C	D	D	D	D	M	D	D	D	19.24	5.63	7.412	.	
SLC2A4	p.V383I	T	B	N	A	N	T	T	T	13.57	4.35	0.201	.	Relatively frequent and predicted as benign.
TTN	p.E5646G	D	B	.	D	N	T	T	T	11.53	5.99	5.065	.	Rare and predicted as benign.
	p.E7874K	D	P	.	D	M	T	T	T	10.31	6.16	4.646	.	
	p.N31156S	D	B	.	D	N	T	T	T	15.94	6.17	6.084	.	
XPNPEP2	p.K232N	T	B	N	N	N	T	T	T	3.496	-0.918	1.700	.	Relatively frequent and predicted as benign.
ZNF630	p.R414Q	T	B	.	N	N	T	T	T	0.014	-2.21	-3.562	.	Rare and predicted as benign.

Table S3.14A – Genes, variants and variant frequencies for the analysis of the recessive inheritance pattern for sample 41. All variants have a European MAF inferior to 0.01 in at least in one database. Chr: chromosome, Ref: reference nucleotide, Alt: altered nucleotide, 1000G: 1000 Genomes Project, ExAC: Exome Aggregation Consortium, ESP: NHLBI Exome Sequencing Project, EUR: European, AFR: African, AMR: American, EAS: East Asian, SAS: South Asian, NFE: Non-Finnish European, FIN: Finnish European, OTH: Other populations, EA: European American, AA: African American. Indication of two variants for the same gene represents compound heterozygosity, while a single variant indicates homozygosity. Dot: no information on databases.

Gene	Chr	Start	Ref	Alt	Variants	1000G EUR	1000G AFR	1000G AMR	1000G EAS	1000G SAS	ExAC NFE	ExAC FIN	ExAC AFR	ExAC AMR	ExAC EAS	ExAC SAS	ExAC OTH	ESP EA	ESP AA
ARHGEF10L	1	18023403	C	T	p.T1123I	0	0.011	0.0014	0	0	0	0	0.0122	0.0004	0	0	0	0	0.013
		18023556	A	G	p.Q1174R	0	0.0023	0	0	0	0	0	0.0010	0	0	0	0	0	0.0007
C22orf42	22	32548561	C	T	p.M120I	0.028	0.16	0.029	0.005	0.022	0.0067	0.0041	0.0546	0.0045	0.0005	0.0093	0.0067	0.0006	0.0085
CAST	5	96082087	T	A	p.D373E	0	0.0053	0	0	0	1.50E-5	0	0.0121	0.0005	0	0	0.0011	0	0.012
		96090384	G	A	p.R509H	0	0.003	0	0	0	1.50E-5	0	0.0080	0.0004	0	0.0002	0.0011	0	0.0082
CENPO	2	25022540	T	A	c.29-4T>A	.	.	.	.	.	.	.	.	.	.	.	.	.	.
		25037327	A	G	p.N94S	.	0.0098	0.0014	0	0	0	0	0.0144	0.0010	0	6.06E-5	0	0	0.013
CSTF2	X	100081707	G	A	p.A263T	0	0	0.0014	0	0	0	0	0.0004	0	0	0	0	0	0.0008
DNER	2	230377595	C	T	p.D351N	0	0.0015	0	0	0	1.50E-5	0	0.0012	8.65E-5	0	0	0	0	0.0011
		230231668	G	T	p.P675T	.	.	.	.	.	.	.	.	.	.	.	.	.	.
DSG4	18	28986090	G	A	p.E563K	.	.	.	.	.	.	.	.	.	.	.	.	.	.
		28993458	G	A	p.S1027N	0	0.0061	0.0014	0	0	3.00E-5	0	0.0062	0.0008	0	0	0.0022	0	0.0068
FCGBP	19	40412125	T	C	p.E1168G	0	0	0	0	0	0	0	0.0130	0.0009	0	6.06E-5	0.0011	0	0.011
		40364397	C	T	p.G4749S	0	0	0	0	0	3.12E-5	0	0.0066	0.0007	0	0	0	0	0.0064
FLNA	X	153579974	C	T	p.R2333H	.	.	.	.	.	0	0	0.0007	0	0	0	0	0	0.0011
FLNB	3	58062842	A	T	p.Y121F	0	0.0076	0	0	0	0	0	0.0089	0	0	0	0.0011	0.0001	0.01
		58108856	G	A	p.V1055M	0.001	0.0061	0	0	0	1.50E-5	0	0.0073	0.0003	0	0	0.0011	0.0001	0.0082
FREM3	4	144620443	T	G	p.K462N	0	0.023	0.0029	0	0	0	0	0.0132	0	0	0	0	0	0.015
		144618344	T	C	p.Q1162R	0	0.0098	0.0014	0	0	0	0	0.0112	0	0	0	0	0	0.0072
FRMPD4	X	12736844	C	A	p.T1300N	0	0	0	0	0	2.12E-5	0	0.0074	0.0001	0	0	0	0	0
ADGRF2	6	47646856	C	T	p.L85F	.	.	.	.	.	.	.	.	.	.	.	.	.	.
		47649176	G	T	p.G226V	0	0	0.0014	0	0	0	0	0.0087	0.0004	0	6.06E-5	0	0	0
		47650369	C	T	c.1863+7C>T	0	0	0.0014	0	0	0	0	0.0139	0.0004	0	6.24E-5	0	0	0

Gene	Chr	Start	Ref	Alt	Variants	1000G EUR	1000G AFR	1000G AMR	1000G EAS	1000G SAS	ExAC NFE	ExAC FIN	ExAC AFR	ExAC AMR	ExAC EAS	ExAC SAS	ExAC OTH	ESP EA	ESP AA
HEXA	15	72636487	A	G	c.1527-6T>C	0.001	0	0	0	0	0.0014	0	0.0010	0.0010	0	0	0.0056	0.0009	0.0014
		72648926	C	T	p.V96I	.	.	.	.	.	.	.	.	.	.	.	.	.	.
HYDIN	16	71007312	G	-	p.I1772FfsX15	.	.	.	.	.	.	.	.	.	.	.	.	.	.
		70937582	G	A	p.P2932L	.	.	.	.	.	.	.	.	.	.	.	.	.	.
KIAA1671	22	25425463	G	T	p.R499S	.	.	.	.	.	.	.	.	.	.	.	.	.	.
		25566913	G	A	c.4649+8G>A	0	0.0091	0	0	0	0	0	0.0093	0	0	0.0006	0	.	.
		25573309	C	G	c.4900-6C>T	0	0.0023	0	0	0	0	0	0.0028	0	0	0	0	0	0.0029
KIAA1804	1	233514908	C	T	p.T719M	.	.	.	.	.	1.50E-5	0	0.0005	0	0	6.06E-5	0	0	0.0002
		233515048	C	T	p.R766W	.	.	.	.	.	4.50E-5	0	0.0006	0	0	0.0002	0.0011	0.0001	0.0002
KRTAP10-10	21	46057634	TGT	CA A	p.C101L	.	.	.	.	.	.	.	.	.	.	.	.	.	.
		46058044	G	A	p.R237H	0	0.0098	0	0	0	0.0004	0	0.0091	0.0010	0.0005	0	0	0.0003	0.0098
KRTAP10-7	21	46020817	G	A	p.S94N	0	0.015	0.0014	0	0	1.52E-5	0	0.0067	0.0008	0	0	0.0022	0	0.0061
		46021023	C	A	p.P163T	0	0.015	0.0014	0	0	1.50E-5	0	0.0065	0.0009	0	0	0.0022	0	0.0071
		46021479	C	T	p.P315S	0	0.014	0.0014	0	0.001	1.51E-5	0	0.0062	0.0009	0	0	0.0022	0	0.0002
MAGEE2	X	75004693	G	C	p.P65R	0	0.0008	0	0	0	0	0	0.0015	0	0	0	0	0	0.0016
MARCH10	17	60879057	C	T	p.D14N	0	0.0015	0	0	0	3.00E-5	0	0.0011	0	0	0	0	0	0.0016
		60837208	C	T	p.E124K	0	0.012	0	0.001	0	0.0001	0	0.0070	0.0031	0.0014	6.16E-5	0.0033	0.0002	0.0064
		60813751	G	A	p.S493L	0	0.0083	0	0	0	4.50E-5	0	0.0122	0.0006	0	0	0	0.0002	0.01
MBOAT4	8	29989820	C	T	p.R316Q	0	0.0008	0.0014	0	0	0	0	0.0014	0.0024	0	0	0	0	0.0014
		29989463	T	G	p.N435T	0	0.0023	0.0014	0	0	0	0	0.0027	0.0026	0	0	0	0	0.0036
MKI67	10	129911869	A	G	c.1481-3T>C	0	0.003	0	0	0	6.00E-5	0	0.0084	0.0003	0.0001	0	0	0	0.01
		129903541	G	C	p.A1828G	0	0.011	0.0014	0	0	0.0001	0	0.0101	0.0008	0	0	0	0.0001	0.013
MYO18A	17	27441099	G	A	p.A843V	.	.	.	.	.	0.0002	0	0	0.0002	0	0	0	0.0001	0
		27417123	G	A	c.5389-7C>T	.	.	.	.	.	0.0025	0	0.0007	0.0025	0	0.0002	0.0042	0.0023	0.0005
MYOM3	1	24432402	C	T	c.563+8G>A	0	0.014	0	0	0	0.0002	0	0.0098	0.0002	0	7.04E-5	0	0.0005	0.01
		24394452	G	A	c.3358+8C>T	.	.	.	.	.	0.0006	0	0.0001	0.0007	0	0	0	0.0005	0

Gene	Chr	Start	Ref	Alt	Variants	1000G EUR	1000G AFR	1000G AMR	1000G EAS	1000G SAS	ExAC NFE	ExAC FIN	ExAC AFR	ExAC AMR	ExAC EAS	ExAC SAS	ExAC OTH	ESP EA	ESP AA
NEUROD4	12	55421067	C	T	p.H282Y	.	.	.	.	.	0	0	9.61E-5	0	0	0	0	.	.
NRAP	10	115400039	C	T	p.A459T	.	.	.	.	.	0	0	0	8.64E-5	0	0	0	.	.
		115389366	GC	AA	p.A674V	.	.	.	.	.	.	.	.	.	.	.	.	.	.
OR10P1	12	56031093	C	T	p.R140W	0	0.003	0	0	0	3.00E-5	0	0.0029	0	0	0	0	0.0001	0.0036
		56031264	G	A	p.E197K	0	0.0023	0	0	0	0	0	0.0022	8.64E-5	0	6.06E-5	0	0	0.0025
PCNXL2	1	233395015	G	C	p.P198R	.	.	.	.	.	.	.	.	.	.	.	.	.	.
		233120180	C	T	p.G2095E	0	0.0008	0	0	0	0	0	0.0011	0	0	0	0	0	0.0014
PDE4DIP	1	145015924	T	C	p.Y55C	.	.	.	.	.	0	0	0.0002	8.64E-5	0	0	0	0	0.0007
		144854574	T	C	p.H2193R	.	.	.	.	.	1.50E-5	0	0.0124	0.0003	0	0.0001	0	.	.
PHKA2	X	18936979	T	C	c.1964-7A>G	0	0.0023	0	0	0	0.0006	0	0.0020	0.0002	0	0.0002	0	0.0018	0.0013
SPATA31D1	9	84608992	C	A	p.Q1203K	0	0.0091	0	0	0	1.50E-5	0	0.0067	8.64E-5	0	0	0	0	0.0064
		84609818	C	G	p.A1478G	0	0.003	0	0	0	1.50E-5	0	0.0027	0	0	0	0	0	0.0023
STARD9	15	42980885	A	G	p.H2370R	.	.	.	.	.	0	0	0.0005	0	0	0	0	0	0.0014
		42988035	G	A	p.S4414N	0	0.0083	0	0	0	0	0	0.0067	0	0	0	0	.	.
STK26	X	131188895	G	A	c.273+6G>A	0	0.0038	0.0014	0	0	0	0	0.0067	0.0002	0	0	0	.	0.0047
SYNE2	14	64518503	A	C	p.E2624D	0	0.0038	0.0014	0	0	0	0	0.0057	0.0003	0	0	0	0	0.0065
		64554434	A	G	p.I3844V	0	0.0045	0.0014	0	0	0.0001	0	0.0044	0.0005	0	0	0	0	0.0036
TNC	9	117852990	C	T	p.R103H	.	.	.	.	.	0	0	0.0009	8.64E-5	0	0	0	0	0.0005
		117849127	C	T	p.V295M	.	.	.	.	.	0	0	0.0005	0	0	0	0	.	.
TTN	2	179544983	C	G	p.R11139T	0.006	0.0008	0.0014	0	0	0.0022	0.0002	0.0003	0.0017	0	0.0004	0	0.0033	0.0003
		179458912	G	C	p.A19403G	0	0.0008	0	0	0	0	0	0.0001	8.72E-5	0	0	0	.	.
UGT1A3	2	234637879	T	C	p.I36T	.	.	.	.	.	0.0001	0	0	0	0	6.06E-5	0	.	.
		234637918	G	A	p.R49Q	.	.	.	.	.	0	0	0.0002	0	0	0	0	0	0.0002
UPF3B	X	118975082	C	T	p.R255K	0	0.0076	.	0	0	0	0	0.0077	0.0007	0	0	0	0	0.0083
VILL	3	38047456	-	AA	p.W710SfsX25	0	0.0061	0.0029	0	0	4.55E-5	0	0.0068	0.0007	0	0	0	0	0.0052
		38048481	G	A	p.E836K	0	0.015	0.0043	0	0	0.0002	0	0.0132	0.0013	0	0	0.0022	0	0.013
ZIC3	X	136649457	G	A	p.A203T	.	.	.	.	.	0	0	0.0036	0	0	0	0	0	0.0029

Table S3.14B – Genes, variants and predicted functional impact for the analysis of the recessive inheritance model for sample 41. Each functional impact predictor has a different classification with different cut-offs. SIFT: T – tolerated (>0.05), D – damaging (≤0.05); PolyPhen2: B – benign (<0.452), P – possibly damaging (0.453 and 0.956), D – probably damaging (>0.957); LRT: D – deleterious, N – neutral, U – unknown; MutationTaster: P – polymorphism automatic, N – polymorphism, D – disease causing, A – disease causing automatic; MutationAssessor: N – neutral, L – low, M – medium, H – high; FATHMM: T – tolerated (> -1.5), D – damaging (< -1.5); Radial SVM: T – tolerated, D – damaging; LR: T – tolerated, D – damaging; CADD: ≥15 considered pathogenic; GERP++: ≥2 considered conserved; PhyloP: ≥1; HSF: D – disturbs donor splice site; A – disturbs acceptor splice site; No – predicted to not affect splicing. Dot – no prevision was generated. The last column indicates the classification of each variant. A variant is classified as pathogenic when 6 of the 9 functional impact software predict it as damaging, otherwise is classified as benign. A variant is classified as relatively frequent if MAF is > 1% in other populations but European or if there are individuals with homozygous variants, otherwise is classified as rare.

Gene	Variants	SIFT	Polyphen2	LRT	Mutation Taster	Mutation Assessor	FATHMM	Radial SVM	LR	CADD	GERP++	PhyloP	HSF	Classification
ARHGEF10L	p.T1123I	D	B	N	N	L	T	T	T	11.32	4.06	0.645	.	p.T1123I relatively frequent and p.Q1174R rare. Both predicted as benign.
	p.Q1174R	T	D	D	D	L	T	T	T	21.3	4.87	4.664	.	
C22orf42	p.M120I	T	B	.	N	N	T	T	T	0.028	-1.65	0.156	.	Relatively frequent and predicted as benign.
CAST	p.D373E	T	D	N	N	M	T	T	T	19.05	3.6	1.497	.	Relatively frequent. Predicted as benign.
	p.R509H	T	D	N	N	L	T	T	T	16.92	3.72	0.753	.	
CENPO	c.29-4T>A	.	.	.	.	.	.	.	.	.	.	.	No	Rare. c.29-4T>A predicted not to affect splicing. p.N94S predicted as benign.
	p.N94S	T	B	N	N	L	T	T	T	4.073	1.58	0.038	.	
CSTF2	p.A263T	T	B	N	N	L	T	T	T	6.797	-1.3	-0.059	.	Rare and predicted as benign.
DNER	p.D351N	T	B	D	D	L	T	T	T	13.59	5.76	2.359	.	Rare. p.D351N predicted as benign and p.P675T as pathogenic.
	p.P675T	T	D	D	D	N	D	D	D	15.28	5.93	9.357	.	
DSG4	p.E563K	T	B	N	N	L	T	T	T	10.28	1.02	-0.171	.	Both rare and predicted as benign.
	p.S1027N	T	B	N	N	L	T	T	T	4.385	0.35	0.083	.	
FCGBP	p.E1168G	T	D	D	D	M	D	D	D	15.52	4.39	4.897	.	p.E1168G relatively frequent and p.G4749S rare. p.E1168G predicted as pathogenic and G4749S as benign.
	p.G4749S	T	B	U	N	M	T	T	T	11.23	4.1	2.708	.	
FLNA	p.R2333H	T	P	D	D	M	T	D	D	13.71	5.27	7.773	.	Rare and predicted as pathogenic.
FLNB	p.Y121F	D	D	D	D	L	D	D	D	28.5	5.04	9.287	.	Relatively frequent and predicted as pathogenic.
	p.V1055M	D	D	D	D	M	D	D	D	18.71	5.66	9.809	.	
FREM3	p.K462N	T	.	N	D	N	T	T	T	1.418	-4.26	0.039	.	Rare and predicted as benign.
	p.Q1162R	T	.	N	N	L	T	T	T	0.006	-1.11	1.500	.	

Gene	Variants	SIFT	Polyphen2	LRT	Mutation Taster	Mutation Assessor	FATHMM	Radial SVM	LR	CADD	GERP++	PhyloP	HSF	Classification
FRMPD4	p.T1300N	D	B	D	N	N	T	T	T	13.31	5.75	5.478	.	Rare and predicted as benign.
GPR111	p.L85F	T	B	N	N	N	T	T	T	0.033	-4.76	-0.943	.	p.L85F rare and p.G226V and c.1863+7C>T relatively frequent. p.L85F and p.G226V predicted as benign and c.1863+7C>T as not to affect splicing.
	p.G226V	T	D	D	N	M	T	T	T	14.83	5.43	3.254	.	
	c.1863+7C>T	D	B	N	N	.	T	T	T	9.013	0.176	-1.522	No	
HEXA	c.1527-6T>C	.	.	.	.	.	.	.	.	.	.	.	No	Rare. c.1527-6T>C predicted not to affect splicing and p.V96I as benign.
	p.V96I	T	B	N	N	L	D	T	D	13.78	-3.19	-0.384	.	
HYDIN	p.I1772FfsX15	.	.	.	.	.	.	.	.	.	.	.	.	Rare. p.P2932L predicted as benign. Gene with a large number of variants and located in a non-conserved region of the genome.
	p.P2932L	T	D	U	D	M	T	T	T	20.6	5.25	7.831	.	
KIAA1671	p.R499S	D	P	D	N	L	.	T	T	16.47	1.53	0.518	.	Rare. p.R499S predicted as benign, c.4649+8G>A as not to affect splicing and c.4900-6C>T to affect acceptor splice site.
	c.4649+8G>A	.	.	.	.	.	.	.	.	.	.	.	No	
	c.4900-6C>T	.	.	.	.	.	.	.	.	.	.	.	A	
KIAA1804	p.T719M	D	P	N	N	L	T	T	T	14.24	3.03	1.515	.	Rare and predicted as benign.
	p.R766W	T	B	N	D	N	T	T	T	14.14	2.99	0.236	.	
KRTAP10-10	p.C101L	.	.	.	.	.	.	.	.	.	.	.	.	Rare and p.R237H predicted as benign.
	p.R237H	D	B	.	N	M	T	T	T	4.110	1.49	0.072	.	
KRTAP10-7	p.S94N	D	P	.	N	M	T	T	T	11.56	3.78	-0.242	.	Relatively frequent and predicted as benign.
	p.P163T	D	B	.	N	M	T	T	T	2.564	0.677	0.000	.	
	p.P315S	T	D	.	D	M	T	T	T	12.46	3.15	2.433	.	
MAGEE2	p.P65R	T	B	.	N	N	T	T	T	13.21	2.54	1.317	.	Rare and predicted as benign.
MARCH10	p.D14N	T	B	N	D	L	T	T	T	13.69	4.62	2.481	.	Rare and predicted as benign.
	p.E124K	T	B	N	N	N	T	T	T	7.757	4.79	2.695	.	
	p.S493L	D	B	N	N	L	T	T	T	11.92	3.39	2.193	.	
MBOAT4	p.R316Q	.	D	D	D	M	T	T	D	16.14	5.05	4.541	.	Rare. p.R316Q predicted as pathogenic and p.N435T as benign.
	p.N435T	.	B	N	D	L	T	T	T	9.145	4.69	0.893	.	
MKI67	c.1481-3T>C	.	.	.	.	.	.	.	.	.	.	.	No	Rare. c.1481-3T>C predicted not to affect splicing

Gene	Variants	SIFT	Polyphen2	LRT	Mutation Taster	Mutation Assessor	FATHMM	Radial SVM	LR	CADD	GERP++	PhyloP	HSF	Classification
	p.A1828G	T	D	N	N	L	T	T	T	13.58	1.64	0.080	.	and p.A1828G as benign.
MYO18A	p.A843V	T	P	D	D	L	D	D	D	23.2	5.72	4.460	.	Rare. p.A843V predicted as pathogenic and c.5389-7C>T not to affect splicing.
	c.5389-7C>T	.	.	.	.	.	.	.	.	.	.	.	No	
MYOM3	c.563+8G>A	.	.	.	.	.	.	.	.	.	.	.	No	Rare and predicted not to affect splicing.
	c.3358+8C>T	.	.	.	.	.	.	.	.	.	.	.	No	
NEUROD4	p.H282Y	D	P	D	D	M	D	D	D	24.5	5.85	4.875	.	Rare and predicted as pathogenic.
NRAP	p.A459T	T	D	D	D	M	T	T	T	26.2	5.17	5.384	.	Rare and p.A674V predicted as benign.
	p.A674V	.	.	.	.	.	.	.	.	.	.	.	.	
OR10P1	p.R140W	T	P	N	N	M	T	T	T	11.16	-5.36	-6.578	.	Rare and predicted as benign.
	p.E197K	D	D	D	N	M	T	T	T	11.70	4.44	1.090	.	
PCNXL2	p.P198R	T	P	.	D	N	T	T	T	11.68	4.17	4.195	.	Rare and predicted as benign.
	p.G2095E	D	B	N	N	L	T	T	T	9.982	3.97	0.248	.	
PDE4DIP	p.Y55C	T	B	.	N	N	T	T	T	7.504	3.55	1.239	.	Rare and predicted as benign.
	p.H2193R	T	B	.	N	N	T	T	T	0.004	2.08	0.253	.	
PHKA2	c.1964-7A>G	.	.	.	.	.	.	.	.	.	.	.	A	Relatively frequent and predicted to affect acceptor splice site.
SPATA31D1	p.Q1203K	T	P	.	N	N	T	T	T	9.014	2.24	0.348	.	p.Q1203K relatively frequent and p.A1478G rare. Predicted as benign.
	p.A1478G	D	P	.	N	L	T	T	T	12.36	-0.307	0.070	.	
STARD9	p.H2370R	D	.	.	N	L	T	T	T	13.41	-1.18	0.077	.	Rare. p.H2370R predicted as benign and p.S4414N as pathogenic.
	p.S4414N	D	D	.	D	M	T	D	D	25.0	5.75	8.832	.	
STK26	c.273+6G>A	.	.	.	.	.	.	.	.	.	.	.	No	Rare and predicted not to affect splicing.
SYNE2	p.E2624D	T	P	N	N	L	T	T	T	7.784	-3.08	0.101	.	Relatively frequent and predicted as benign.
	p.I3844V	T	B	N	N	N	T	T	T	7.472	-6.75	-2.691	.	
TNC	p.R103H	D	D	D	D	M	T	T	T	34	5.94	7.818	.	Rare and predicted as pathogenic.
	p.V295M	D	D	D	D	M	T	T	T	18.50	5.56	3.679	.	
TTN	p.R11139T	D	B	.	D	N	T	T	T	9.474	1.55	-0.290	.	Rare and predicted as benign.

Gene	Variants	SIFT	Polyphen2	LRT	Mutation Taster	Mutation Assessor	FATHMM	Radial SVM	LR	CADD	GERP++	PhyloP	HSF	Classification
	p.A19403G	D	B	.	N	L	T	T	T	16.35	6.17	3.098	.	
UGT1A3	p.I36T	T	B	.	N	N	T	T	T	0.021	1.97	0.325	.	Rare and predicted as benign.
	p.R49Q	T	B	.	N	N	T	T	T	7.040	-5.19	0.112	.	
UPF3B	p.R255K	T	B	D	D	L	T	T	T	8.492	4.64	2.267	.	Relatively frequent and predicted as benign.
VILL	p.W710SfsX25	.	.	.	.	.	.	.	.	.	.	.	.	p.W710SfsX25 rare and p.E836K relatively frequent. p.W710SfsX25 predicted as pathogenic and p.E836K as benign.
	p.E836K	T	P	N	N	L	T	T	T	13.35	3.04	0.370	.	
ZIC3	p.A203T	T	B	D	D	N	T	T	T	6.226	3.71	2.797	.	Rare and predicted as benign.

Table S3.15A – Genes, variants and variant frequencies for the analysis of the recessive inheritance pattern for sample 42. All variants have a European MAF inferior to 0.01 in at least in one database. Chr: chromosome, Ref: reference nucleotide, Alt: altered nucleotide, 1000G: 1000 Genomes Project, ExAC: Exome Aggregation Consortium, ESP: NHLBI Exome Sequencing Project, EUR: European, AFR: African, AMR: American, EAS: East Asian, SAS: South Asian, NFE: Non-Finnish European, FIN: Finnish European, OTH: Other populations, EA: European American, AA: African American. Indication of two variants for the same gene represents compound heterozygosity, while a single variant indicates homozygosity. Dot: no information on databases.

Gene	Chr	Start	Ref	Alt	Variants	1000G EUR	1000G AFR	1000G AMR	1000G EAS	1000G SAS	ExAC NFE	ExAC FIN	ExAC AFR	ExAC AMR	ExAC EAS	ExAC SAS	ExAC OTH	ESP EA	ESP AA
BBX	3	107466840	C	T	p.A260V	0.002	0	0.0014	0	0.0031	0.0017	0.0008	0.0005	0.0008	0	0.0110	0.0055	0.0017	0.0002
		107493689	A	T	p.K344X	.	.	.	.	.	.	.	.	.	.	.	.	.	.
FMO2	1	171168521	G	A	p.R174H	0.004	0.0015	0.0029	0	0.0061	0.0060	0.0024	0.0019	0.0020	0	0.0044	0.0045	0.0069	0.0011
		171168557	G	A	p.A186H	.	.	.	.	.	0	0	0.0004	0	0.0001	0	0	0	0.0002
GREB1	2	11727568	G	A	p.R407Q	0.002	0.0008	0.0014	0	0	0.0011	0.0002	0.0020	0.0004	0	0.0005	0	0.0012	0.002
		11758707	G	A	p.A1236T	0.005	0	0.0014	0	0	0.0072	0.0007	0.0016	0.0026	0.0001	0.0011	0.0064	0.0079	0.0014
KAT8	16	31129218	GGGCGGGGCCCAGGGCTGGGGGCGGGGCGGAGCTCAGGGCCAGGGGGTG	-	c.210+8_56delGGGCGGGGCCCAGGGCTGGGGCGGGGCGGAGCTCAGGGCCAGGGGGTG	.	.	.	.	.	.	.	.	.	.	.	.	.	.
MEGF9	9	123476258	G	A	p.P127S	0.003	0	0	0	0	0.0008	0.0002	0.0001	0.0003	0	0.0001	0	.	.
		123476215	G	A	p.T141I	.	.	.	.	.	.	.	.	.	.	.	.	.	.
OTOF	2	26750773	C	T	p.V52M	.	.	.	.	.	0.0002	0	9.63E-5	0.0003	0	6.06E-05	0	0.0003	0
		26698877	T	C	p.M966V	.	.	.	.	.	.	.	.	.	.	.	.	.	.
RTTN	18	67864918	T	G	p.D212A	0.003	0.0008	0.0029	0	0	0.0047	0.0067	0.0015	0.0012	0.0001	0.0002	0.0045	0.0045	0.0005
UGT2B28	4	70146423	G	A	p.A69T	0.002	0.0008	0.0058	0	0	0.0042	0.0005	0.0015	0.0051	0	0.0014	0.0035	0.0054	0.0017
UMODL1	21	43531298	G	C	p.G656R	.	.	.	.	.	.	.	.	.	.	.	.	.	.
VPS13C	15	62214756	A	G	p.I2272T	.	.	.	.	.	0.0010	0	0	0.0004	0	0	0	0.0015	0
		62146742	T	C	p.I3726V	0.006	0	0	0	0	0.0062	0.0008	0.0007	0.0023	0	0.0017	0.0077	0.0056	0.0005
WRN	8	30938692	G	T	p.L383F	0.001	0	0	0	0.001	0.0030	0	0.0003	0.0009	0	0.0010	0.0044	0.0028	0.0002
		30998961	G	A	p.A995T	0.002	0	0	0	0.002	0.0034	0.0003	0.0004	0.0010	0	0.0012	0.0044	0.003	0.0002

Table S3.15B – Genes, variants and predicted functional impact for the analysis of the recessive inheritance model for sample 42. Each functional impact predictor has a different classification with different cut-offs. SIFT: T – tolerated (>0.05), D – damaging (≤0.05); PolyPhen2: B – benign (<0.452), P – possibly damaging (0.453 and 0.956), D – probably damaging (>0.957); LRT: D – deleterious, N – neutral, U – unknown; MutationTaster: P – polymorphism automatic, N – polymorphism, D – disease causing, A – disease causing automatic; MutationAssessor: N – neutral, L – low, M – medium, H – high; FATHMM: T – tolerated (> -1.5), D – damaging (< -1.5); Radial SVM: T – tolerated, D – damaging; LR: T – tolerated, D – damaging; CADD: ≥15 considered pathogenic; GERP++: ≥2 considered conserved; PhyloP: ≥1; HSF: D – disturbs donor splice site; A – disturbs acceptor splice site; No – predicted to not affect splicing. Dot – no prevision was generated. The last column indicates the classification of each variant. A variant is classified as pathogenic when 6 of the 9 functional impact software predict it as damaging, otherwise is classified as benign. A variant is classified as relatively frequent if MAF is > 1% in other populations but European or if there are individuals with homozygous variants, otherwise is classified as rare.

Gene	Variants	SIFT	Polyphen2	LRT	Mutation Taster	Mutation Assessor	FATHMM	Radial SVM	LR	CADD	GERP++	PhyloP	HSF	Classification
BBX	p.A260V	T	D	N	N	N	D	D	D	14.01	0.992	0.029	.	p.A260V relatively frequent and p.K344X rare. predicted as benign.
	p.K344X	D	D	D	D	L	T	T	T	39	5.51	8.820	.	
FMO2	p.R174H	T	D	D	D	L	T	T	T	17.10	5.26	3.368	.	R174H relatively frequent and p.A186H rare. Predicted as benign.
	p.A186H	D	B	N	D	M	T	T	T	17.48	2.93	0.020	.	
GREB1	p.R407Q	D	B	.	N	.	T	T	T	6.321	1.08	0.153	.	p.R407Q rare and p.A1236T relatively frequent. Predicted as as benign.
	p.A1236T	T	B	N	N	N	T	T	T	6.176	-7.9	-0.574	.	
KAT8	c.210+8_56delGGCGGGGCCCAGGGCTGGGGGCGGGGCGGAGCTCAGGGCCAGGGGGTG	.	.	.	.	.	.	.	.	.	.	.	No	Rare and predicted not to affect splicing.
MEGF9	p.P127S	T	B	N	N	L	T	T	T	16.74	-2.74	-0.363	.	Rare and predicted as benign.
	p.T141I	T	B	N	N	N	T	T	T	17.58	2.54	1.493		
OTOF	p.V52M	T	P	D	N	L	T	T	T	15.94	4.51	1.825	.	Rare and predicted as benign.
	p.M966V	T	B	D	D	N	T	T	T	17.37	5.41	5.044	.	
RTTN	p.D212A	D	D	D	D	M	T	T	T	11.99	3.41	6.215	.	Relatively frequent. Predicted as benign and ClinVar classification as likely benign.
UGT2B28	p.A69T	T	B	U	N	N	T	T	T	0.006	-4.36	-7.778	.	Relatively frequent and predicted as benign.
UMODL1	p.G656R	T	B	.	N	.	T	T	T	7.155	-0.153	-2.487	.	Rare and predicted as benign.
VPS13C	p.I2272T	T	B	N	N	N	T	T	T	0.009	0.083	0.392	.	p.I2272T rare and p.I3726V relatively frequent. p.I2272T predicted as benign and p.I3726V as pathogenic.
	p.I3726V	D	P	D	D	M	T	T	T	17.90	5.36	2.092	.	

Gene	Variants	SIFT	Polyphen2	LRT	Mutation Taster	Mutation Assessor	FATHMM	Radial SVM	LR	CADD	GERP++	PhyloP	HSF	Classification
WRN	p.L383F	T	B	N	N	M	T	T	T	8.744	2.78	0.169	.	Relatively frequent. Predicted as benign and ClinVar classification as benign and likely benign.
	p.A995T	T	B	N	N	N	T	T	T	9.173	-2.44	0.170	.	

Table S3.16A – Genes, variants and variant frequencies for the analysis of the recessive inheritance pattern for sample 43. All variants have a European MAF inferior to 0.01 in at least in one database. Chr: chromosome, Ref: reference nucleotide, Alt: altered nucleotide, 1000G: 1000 Genomes Project, ExAC: Exome Aggregation Consortium, ESP: NHLBI Exome Sequencing Project, EUR: European, AFR: African, AMR: American, EAS: East Asian, SAS: South Asian, NFE: Non-Finnish European, FIN: Finnish European, OTH: Other populations, EA: European American, AA: African American. Indication of two variants for the same gene represents compound heterozygosity, while a single variant indicates homozygosity. Dot: no information on databases.

Gene	Chr	Start	Ref	Alt	Variants	1000G EUR	1000G AFR	1000G AMR	1000G EAS	1000G SAS	ExAC NFE	ExAC FIN	ExAC AFR	ExAC AMR	ExAC EAS	ExAC SAS	ExAC OTH	ESP EA	ESP AA
ALMS1	2	73679005	A	C	p.K1783T	.	.	.	.	.	.	.	.	.	.	.	.	.	.
		73799632	C	G	p.T3542S	0.003	0.0008	0.0043	0	0.002	0.0099	0.0026	0.0026	0.0037	0	0.0048	0.0034	0.012	0.0032
AMBRA1	11	46564927	G	C	p.P214A	.	.	.	.	.	0.0021	0	9.64E-5	0.0003	0	0.0004	0	0.0017	0
		46564053	C	T	p.R415H	0.001	0.0008	0	0	0.0041	0.0018	0	0.0003	0.0008	0	0.0036	0	0.0013	0
ARMCX4	X	100749588	A	C	p.Q2004H	0.001	0	0.0043	0	0	0.0021	0	0	0.0072	0	0	0	.	.
ASPM	1	197112706	G	A	p.P226S	.	.	.	.	.	0.0001	0	0	0.0002	0	0	0	0.0001	0
		197073196	G	A	p.R1729W	0.016	0.0015	0.0086	0	0.0031	0.0124	0.0018	0.0018	0.0087	0.0006	0.0113	0.0177	0.013	0.0014
ADGRE5	19	14517319	C	G	p.I666M	0.003	0	0.0043	0	0	0.0047	0	0.0020	0	0	0	0.0081	0.003	0.0014
		14518914	C	T	p.A830V	0.001	0	0.0029	0	0	0.0008	0	0.0004	8.76E-5	0	6.09E-5	0	0.001	0.0002
DCHS2	4	155254540	G	C	p.L480M	0.003	0.0008	0.0014	0	0.0041	0.0068	0.0026	0.0007	0.0007	0.0001	0.0049	0.0044	0.0059	0.0009
		155411070	A	T	p.H940Q	.	.	.	.	.	.	.	.	.	.	.	.	.	.
FAM90A1	12	8377323	C	T	p.D36N	0.003	0.0015	0.0072	0	0.001	0.0024	0.0229	0.0012	0.0764	0.0006	0.0003	0.0106	0.0006	0
FRMPD4	X	12725699	G	A	p.V467I	.	.	.	.	.	8.35E-5	0	0	0	0	0	0	0.0003	0
KAT8	16	31129218	GGGCGGGG CCCAGGGCT GGGGGCGG GGCGGAGCT CAGGGCCAG GGGGTG	-	c.210+8_56delG GGCGGGGCCCA GGGCTGGGGGC GGGGCGGAGCT CAGGGCCAGGG GGTG	.	.	.	.	.	.	.	.	.	.	.	.	.	.
KDM5C	X	53222024	G	T	c.4047-5C>A	.	.	.	.	.	.	.	.	.	.	.	.	.	.
MCF2	X	138699734	C	T	p.E373K	0.002	0	0	0	0	0.0011	0.0055	0	0.0011	0	0.0008	0.0032	0.0019	0.0003
MUC4	3	195508030	G	T	p.P3474H	0.001	0.021	0.0014	0	0.002	0	0	0.0011	0	0	0.0001	0	.	.
		195507998	G	C	p.H3485D	.	.	.	.	.	0.0001	0	0.0037	0	0	0	0	.	.
MYH7B	20	33577679	G	A	p.R617H	0	0.0008	0	0	0	0.0006	0	0.0002	8.66E-5	0.0001	0	0	0.0007	0.0002
		33588185	A	T	p.K1666M	0.006	0	0.0014	0.001	0	0.0074	0.0020	0.0012	0.0015	0.0013	0.0005	0.0035	0.0092	0.0019
NDUFV3	21	44317156	A	C	c.48+7G>A	0.008	0.0008	0.01	0	0.001	0.0066	0.0009	0.0026	0.0097	0	0.0017	0.01	0.0049	0.0018
		44313498	G	A	c.168A>C, p.K56N	0.007	0.0008	0.0086	0	0.001	0.0104	0	0.0065	0.0077	0	0.0015	0.0308	0.0031	0.0003
OCRL	X	128691930	A	G	c.439+3A>G	0.003	0	0	0	0	0.0036	0.0038	0.0005	0.0008	0	0.0012	0.0016	0.0027	0

Gene	Chr	Start	Ref	Alt	Variants	1000G EUR	1000G AFR	1000G AMR	1000G EAS	1000G SAS	ExAC NFE	ExAC FIN	ExAC AFR	ExAC AMR	ExAC EAS	ExAC SAS	ExAC OTH	ESP EA	ESP AA
OTOF	2	26686872	G	A	p.T1688M	0.003	0	0	0	0	0.0017	0.0003	0.0003	0.0014	0	0	0.0022	0.0019	0
		26681081	C	T	p.D1941N	.	.	.	.	.	.	.	.	.	.	.	.	.	.
PDE4DIP	1	144994670	C	A	p.S87I	0.001	0	.0	0	0.013	0.0015	0	0.0005	0.0023	0	0.0119	0.0132	0.0009	0.0009
		144881547	G	A	p.R1173C	.	.	.	.	.	0.0017	0	0.0019	0.0023	0	0.0117	0.0132	0.001	0.0018
		144857705	G	A	p.R2011W	.	.	.	.	.	0.0022	0.0014	0.0007	0.0029	0.0001	0.0120	0.0132	0.0007	0.0007
TICRR	15	90145108	T	C	p.L823P	0	0	0.0014	0	0	0.0006	0	0.0002	0.0003	0	0	0	0.0006	0.0003
		90167043	C	A	p.P1168T	.	.	.	.	.	4.53E-5	0	0	0	0	0	0	.	.
VWF	12	6167119	G	C	p.A542G	0.001	0	0.0014	0	0	0.0007	0	0.0002	0.0004	0	0	0	0.0009	0
XPNPEP2	X	128884502	G	T	p.K232N	0.0089	0.0008	0.0014	0	0	0.0134	0.0029	0.0014	0.0036	0.0004	0.0011	0.0043	0.007	0.0016
ZFX	X	24197822	A	G	p.N194S	0	0.0038	0.0058	0	0	0.0005	0	0.0061	0.0012	0.0002	0.0003	0	0.0003	0.0047
ZNF81	X	47774613	C	T	p.H190Y	.	.	.	.	.	.	.	.	.	.	.	.	.	.

Table S3.16B – Genes, variants and predicted functional impact for the analysis of the recessive inheritance model for sample 43. Each functional impact predictor has a different classification with different cut-offs. SIFT: T – tolerated (>0.05), D – damaging (≤0.05); PolyPhen2: B – benign (<0.452), P – possibly damaging (0.453 and 0.956), D – probably damaging (>0.957); LRT: D – deleterious, N – neutral, U – unknown; MutationTaster: P – polymorphism automatic, N – polymorphism, D – disease causing, A – disease causing automatic; MutationAssessor: N – neutral, L – low, M – medium, H – high; FATHMM: T – tolerated (> -1.5), D – damaging (< -1.5); Radial SVM: T – tolerated, D – damaging; LR: T – tolerated, D – damaging; CADD: ≥15 considered pathogenic; GERP++: ≥2 considered conserved; PhyloP: ≥1; HSF: D – disturbs donor splice site; A – disturbs acceptor splice site; No – predicted to not affect splicing. Dot – no prevision was generated. The last column indicates the classification of each variant. A variant is classified as pathogenic when 6 of the 9 functional impact software predict it as damaging, otherwise is classified as benign. A variant is classified as relatively frequent if MAF is > 1% in other populations but European or if there are individuals with homozygous variants, otherwise is classified as rare.

Gene	Variants	SIFT	Polyphen2	LRT	Mutation Taster	Mutation Assessor	FATHMM	Radial SVM	LR	CADD	GERP++	PhyloP	HSF	Classification
ALMS1	p.K1783T	D	P	N	N	L	T	T	T	13.04	2.37	2.203	.	p.K1783T rare and p.T3542S relatively frequent. Predicted as benign.
	p.T3542S	T	B	N	D	L	T	T	T	15.29	3.64	1.586	.	
AMBRA1	p.P214A	D	B	D	D	L	T	T	T	13.93	5.93	6.639	.	Relatively frequent. p.P214A predicted as benign and p.R415H as pathogenic.
	p.R415H	D	D	N	D	L	T	D	D	16.07	5.12	4.347	.	
ARMCX4	p.Q2004H	.	.	.	N	.	.	T	T	0.022	2.11	0.060	.	Rare and predicted as benign.
ASPM	p.P226S	T	B	N	D	L	T	D	T	15.55	4.34	3.153	.	p.P226S rare and p.R1729W relatively frequent. p.P226S predicted as benign and p.R1729W as benign with ClinVar classification as benign.
	p.R1729W	D	D	N	D	M	T	T	T	17.34	5.98	2.434	.	
ADGRE5	p.I666M	D	D	N	D	M	T	T	T	15.23	2.63	-1.322	.	Rare and predicted as benign.

Gene	Variants	SIFT	Polyphen2	LRT	Mutation Taster	Mutation Assessor	FATHMM	Radial SVM	LR	CADD	GERP++	PhyloP	HSF	Classification
	p.A830V	T	B	.	N	N	T	T	T	9.156	2.47	0.180	.	
DCHS2	p.L480M	T	D	N	D	M	T	T	T	21.0	-6.01	-0.597	.	p.L480M rare and p.H940Q relatively frequent. Predicted as benign.
	p.H940Q	T	D	.	N	M	T	T	T	7.809	-0.454	0.805	.	
FAM90A1	p.D36N	D	D	.	N	N	T	T	T	12.58	0.706	0.250	.	Relatively frequent and predicted as benign.
FRMPD4	p.V467I	T	P	D	D	N	T	T	T	17.11	5.64	7.504	.	Rare and predicted as benign.
KAT8	c.210+8_56del GGGCGGGGCC CAGGGCTGGG GGCGGGGCG GAGCTCAGGG CCAGGGGGTG	.	.	.	.	.	.	.	.	.	.	.	No	Rare and predicted not to affect splicing.
KDM5C	c.4047-5C>A	.	.	.	.	.	.	.	.	.	.	.	No	Rare and predicted not to affect splicing.
MCF2	p.E373K	T	P	D	D	M	T	T	T	32	5.98	4.587	.	Relatively frequent and predicted as benign.
MUC4	p.P3474H	.	D	.	N	N	T	T	T	4.725	-0.725	0.510	.	Rare and predicted as benign.
	p.H3485D	.	P	.	N	N	T	T	T	0.024	-1.49	-0.405	.	
MYH7B	p.R617H	D	D	N	D	M	D	D	D	29.2	4.32	7.770	.	p.R617H rare and p.K1666M relatively frequent. Predicted as pathogenic.
	p.K1666M	D	D	N	D	H	D	D	D	21.5	4.07	9.083	.	
NDUFV3	c.48+7G>A	.	.	.	.	.	.	.	.	.	.	.	No	c.48+7G>A, p.K56N rare and c.168A>C relatively frequent. c.48+7G>A predicted not to affect splicing. c.168A>C, p.K56N predicted as benign, but as affect acceptor splice site.
	c.168A>C, p.K56N	T	D	N	N	M	.	T	T	17.52	2.42	1.417		
OCRL	c.439+3A>G	.	.	.	.	.	.	.	.	.	.	.	D	Relatively frequent. Predicted to affect donor splice site, but ClinVar classification as likely benign.
OTOF	p.T1688M	D	D	D	D	H	D	D	D	20.2	4.4	7.777	.	p.T1688M relatively frequent and p.D1941N rare. p.T1688M predicted as pathogenic but ClinVar classification as likely benign and p.D1941N predicted as benign.
	p.D1941N	T	B	.	D	.	T	T	T	21.8	4.8	7.740	.	
PDE4DIP	p.S87I	D	D	.	D	M	T	T	T	34	5.78	5.979	.	Relatively frequent and predicted as pathogenic.
	p.R1173C	D	D	.	D	L	T	T	T	22.0	6.05	2.167	.	
	p.R2011W	D	P	.	D	M	T	T	T	20.2	3.98	3.420	.	
TICRR	p.L823P	T	B	N	N	N	T	T	T	8.416	-6.72	-0.154	.	Rare and predicted as benign.
	p.P1168T	D	D	N	D	L	T	T	T	15.06	5.27	1.350	.	
VWF	p.A542G	D	B	N	D	L	T	T	T	15.59	-0.398	1.020	.	Rare and predicted as benign.
XPNPEP2	p.K232N	T	B	N	N	N	T	T	T	3.496	-0.918	1.700	.	Relatively frequent and predicted as benign.
ZFX	p.N194S	.	B	D	D	L	T	T	T	6.001	5.97	3.732	.	Relatively frequent and predicted as benign.

Gene	Variants	SIFT	Polyphen2	LRT	Mutation Taster	Mutation Assessor	FATHMM	Radial SVM	LR	CADD	GERP++	PhyloP	HSF	Classification
ZNF81	p.H190Y	T	B	N	N	N	T	T	T	0.067	2.23	0.663	.	Rare and predicted as benign.

Table S3.17A – Genes, variants and variant frequencies for the analysis of the recessive inheritance pattern for sample 45. All variants have a European MAF inferior to 0.01 in at least in one database. Chr: chromosome, Ref: reference nucleotide, Alt: altered nucleotide, 1000G: 1000 Genomes Project, ExAC: Exome Aggregation Consortium, ESP: NHLBI Exome Sequencing Project, EUR: European, AFR: African, AMR: American, EAS: East Asian, SAS: South Asian, NFE: Non-Finnish European, FIN: Finnish European, OTH: Other populations, EA: European American, AA: African American. Indication of two variants for the same gene represents compound heterozygosity, while a single variant indicates homozygosity. Dot: no information on databases.

Gene	Chr	Start	Ref	Alt	Variants	1000G EUR	1000G AFR	1000G AMR	1000G EAS	1000G SAS	ExAC NFE	ExAC FIN	ExAC AFR	ExAC AMR	ExAC EAS	ExAC SAS	ExAC OTH	ESP EA	ESP AA
AHNAK2	14	105417229	G	T	p.A1520E	0	0.0053	0.0029	0.004	0.01	0.0003	0.0002	0.0012	0.0052	0.0006	0.0016	0.0047	0.0001	0.001
		105412333	G	A	p.S3152L	0	0	0.0014	0	0	0.0003	0	0.0002	0.0003	0	0.0014	0.0013	.	.
		105409483	T	A	p.D4102V	0	0.026	0.0029	0	0	0.0002	0	0.0227	0.0016	0	0.0009	0.0022	0.0002	0.021
ATN1	12	7045892	CAGCAGCAG CAGCAG	-	p.Q488_Q492del	.	.	.	.	.	.	.	.	.	.	.	.	.	.
		7047949	C	G	p.D941E	0.002	0	0.0014	0	0.001	0.0027	0.0003	0.0013	0.0023	0	0.0003	0.0022	0.0033	0.0009
COL9A2	1	40780067	A	G	c.151-8T>C	0	0.0053	0.0014	0	0	4.5E-5	0	0.0035	0.0003	0	0	0	0.0002	0.0034
		40778463	A	G	c.250-7T>C	0	0.0083	0.0014	0	0	0.0001	0	0.0058	0	0	0	0	0.0001	0.0039
DLG3	X	69720418	T	C	p.I748T	.	.	.	.	.	.	.	.	.	.	.	.	.	.
FAAH2	X	57458438	G	A	p.A362T	0.001	0	0.0014	0	0	0.0016	0	0.0001	0.0015	0.0002	0	0.0016	0.0024	0
FAM9A	X	8764370	C	T	p.A143T	0.0099	0	0	0	0.001	0.0055	0.0049	0.0012	0.0013	0	0.0007	0.0021	0.0052	0.0005
FCGBP	19	40412179	G	A	p.P1150L	0.002	0	0.0029	0	0	0.0007	0	0.0001	0.0020	0	0.0001	0	0.0005	0
		40357398	C	G	p.L5305F	0.001	0	0	0	0	0.0013	0.0003	0.0002	0.0006	0	0.0001	0	0.0006	0.0005
FLNB	3	58090893	G	A	p.R566Q	0.003	0.0008	0.0058	0	0	0.0008	0	0.0009	0.0017	0	0.0002	0	0.001	0.0007
		58124050	G	A	p.G1666S	0	0	0.0029	0	0	0.0001	0	9.61E-5	0.0003	0	0	0	0.0001	0.0002
IDS	X	148578002	C	T	p.D252N	0.007	0	0	0	0	0.0026	0.0183	0.0001	0.0005	0.0002	0	0.0032	0.0036	0

Gene	Chr	Start	Ref	Alt	Variants	1000G EUR	1000G AFR	1000G AMR	1000G EAS	1000G SAS	ExAC NFE	ExAC FIN	ExAC AFR	ExAC AMR	ExAC EAS	ExAC SAS	ExAC OTH	ESP EA	ESP AA
KAT8	16	31129218	GGGCGGGGC CCAGGGCTG GGGGCGGG GCGGAGCTC AGGGCCAGG GGGTG	-	c.210+8_56delGG GCGGGGCCCAGG GCTGGGGGCGGG GCGGAGCTCAGG GCCAGGGGGTG	.	.	.	.	.	.	.	.	.	.	.	.	.	.
MC1R	16	89986030	G	A	p.V122M	0	0.0008	0	0	0	0.0011	0	0.0004	0.0008	0	0	0.0022	0.0001	0.0002
		89986091	G	A	p.R142H	0.0089	0.0008	0.0058	0	0	0.0073	0.0012	0.0011	0.0079	0.0001	0.0001	0.0056	0.0074	0.0007
MICALL1	22	38318367	C	T	p.R320W	0	0	0	0.001	0.002	.	.	.	.	.	.	.	.	.
		38328672	G	C	p.R710P	.	.	.	.	.	.	.	.	.	.	.	.	.	.
MYO9B	19	17212814	A	G	p.Q96R	.	.	.	.	.	0.0030	0.0026	0.0003	0.0020	0	0.0005	0.0023	0.0027	0.0007
		17322899	C	T	p.P2085L	0.004	0	0.0014	0	0.001	0.0073	0.0052	0.0005	0.0025	0.0002	0.0015	0.0024	0.0022	0.0003
POLA1	X	24745161	C	T	p.P496S	0	0	0	0.001	0	0.0005	0	0.0001	0.0003	0	0.0010	0.0016	0.0009	0.0003
SPTBN5	15	42160751	C	T	p.V1968M	0.008	0	0.0014	0	0	0.0091	0.0067	0.0009	0.0029	0	0.0012	0.0058	0.0076	0.0019
		42147463	C	T	p.A3128T	.	.	.	.	.	9.46E-5	0	0	0	0	6.21E-5	0	.	.
TAF1A	1	222734865	A	G	c.1086-5T>C	0.003	0.0015	0	0	0	0.0077	0.0050	0.0014	0.0017	0	0	0.0023	0.0048	0.0005
		222732051	C	T	p.R435Q	0.007	0	0.014	0	0.0031	0.0088	0.0002	0.0012	0.0050	0.0001	0.0057	0.0033	0.0084	0.0005
TRAK1	3	42251606	GAG	-	p.E698del	.	.	.	.	.	.	.	.	.	.	.	.	.	.
TTN	2	179469794	C	T	p.R18037Q	0	0	0.0014	0	0	1.51E-5	0	0	0.0002	0	0	0	.	.
		179453429	G	A	p.T21008I	0.017	0.0008	0.016	0	0	0.0143	0.0030	0.0023	0.0099	0	0.0006	0.0122	0.016	0.0031
UBR3	2	170734103	G	A	p.G315D	.	.	.	.	.	.	.	.	.	.	.	.	.	.
		170789520	T	C	c.2663+2T>C	.	.	.	.	.	.	.	.	.	.	.	.	.	.
UMODL1	21	43531298	G	C	p.G656R	.	.	.	.	.	.	.	.	.	.	.	.	.	.
URB1	21	33757991	A	G	c.T145C	0.004	0	0.0014	0	0	0.0008	0	0.0005	0	0	0	0	.	.
		33711126	G	A	p.T1467M	0.012	0	0.0058	0	0	0.0073	0.0025	0.0009	0.0074	0	0.0004	0.0041	0.0079	0
ZNF182	X	47842356	G	T	p.N94K	0.006	0	0.0014	0.015	0	0.0089	0.0042	0.0012	0.0011	0.0192	0.0008	0.0016	0.0095	0.0021
ZNF648	1	182026013	G	T	p.S378X	0.002	0	0	0	0.0051	0.0014	0.0023	0	0.0014	0	0.0053	0.0043	0.0007	0

Gene	Chr	Start	Ref	Alt	Variants	1000G EUR	1000G AFR	1000G AMR	1000G EAS	1000G SAS	ExAC NFE	ExAC FIN	ExAC AFR	ExAC AMR	ExAC EAS	ExAC SAS	ExAC OTH	ESP EA	ESP AA
		182025483	C	T	p.G555S	0.002	0	0	0	0	0.0002	0	0	0	0.0001	0.0011	0	.	.

Table S3.17B – Genes, variants and predicted functional impact for the analysis of the recessive inheritance model for sample 45. Each functional impact predictor has a different classification with different cut-offs. SIFT: T – tolerated (>0.05), D – damaging (≤0.05); PolyPhen2: B – benign (<0.452), P – possibly damaging (0.453 and 0.956), D – probably damaging (>0.957); LRT: D – deleterious, N – neutral, U – unknown; MutationTaster: P – polymorphism automatic, N – polymorphism, D – disease causing, A – disease causing automatic; MutationAssessor: N – neutral, L – low, M – medium, H – high; FATHMM: T – tolerated (> -1.5), D – damaging (< -1.5); Radial SVM: T – tolerated, D – damaging; LR: T – tolerated, D – damaging; CADD: ≥15 considered pathogenic; GERP++: ≥2 considered conserved; PhyloP: ≥1; HSF: D – disturbs donor splice site; A – disturbs acceptor splice site; No – predicted to not affect splicing. Dot – no prevision was generated. The last column indicates the classification of each variant. A variant is classified as pathogenic when 6 of the 9 functional impact software predict it as damaging, otherwise is classified as benign. A variant is classified as relatively frequent if MAF is > 1% in other populations but European or if there are individuals with homozygous variants, otherwise is classified as rare.

Gene	Variants	SIFT	Polyphen2	LRT	Mutation Taster	Mutation Assessor	FATHMM	Radial SVM	LR	CADD	GERP++	PhyloP	HSF	Classification
AHNAK2	p.A1520E	T	B	.	N	L	T	T	T	4.317	-7.96	-2.965	.	Relatively frequent and predicted as benign.
	p.S3152L	T	B	.	N	M	T	T	T	7.535	2.28	.	.	
	p.D4102V	D	D	U	N	H	T	T	T	14.43	4.05	3.077	.	
ATN1	p.Q488_Q492del	.	.	.	.	.	.	.	.	.	.	.	.	p.Q488_Q492del rare and p.D941E relatively frequent. p.D941E predicted as benign.
	p.D941E	T	B	U	N	N	T	T	T	0.220	0.375	0.269	.	
COL9A2	c.151-8T>C	.	.	.	.	.	.	.	.	.	.	.	No	Rare and predicted not to affect splicing.
	c.250-7T>C	.	.	.	.	.	.	.	.	.	.	.	No	
DLG3	p.I748T	.	B	D	D	L	T	T	T	16.18	4.85	7.482	.	Rare and predicted as benign.
FAAH2	p.A362T	T	B	N	N	N	T	T	T	0.243	-2.27	-0.122	.	Rare and predicted as benign.
FAM9A	p.A143T	T	P	.	N	N	.	T	T	5.235	-0.327	0.178	.	Relatively frequent and predicted as benign.
FCGBP	p.P1150L	T	D	D	D	M	D	D	D	12.16	2.27	0.479	.	Rare, p.P1150L predicted as pathogenic and p.L5305F as benign.
	p.L5305F	T	D	U	N	L	T	T	T	10.78	2.34	0.253	.	
FLNB	p.R566Q	T	B	D	D	N	D	T	T	21.3	5.14	4.485	.	Rare. p.R566Q predicted as benign and p.G1666S as pathogenic.
	p.G1666S	T	D	D	D	N	D	T	D	23.1	6.17	5.428	.	
IDS	p.D252N	T	B	N	D	N	D	D	D	9.979	4.9	4.664	.	Relatively frequent and predicted as benign.

Gene	Variants	SIFT	Polyphen2	LRT	Mutation Taster	Mutation Assessor	FATHMM	Radial SVM	LR	CADD	GERP++	PhyloP	HSF	Classification
KAT8	c.210+8_56delGG GCGGGGCCCAGG GCTGGGGGCGGG GCGGAGCTCAGG GCCAGGGGGTG												No	Rare and predicted as not affect splicing.
MC1R	p.V122M	.	B	U	N	L	T	T	T	10.38	3.84	1.591	.	Relatively frequent. p.V122M predicted as benign and p.R142H as pathogenic.
	p.R142H	.	D	U	D	H	D	D	D	35	4.81	6.453	.	
MICALL1	p.R320W	.	D	D	D	M	T	T	T	19.39	2.8	1.804	.	Rare and predicted as pathogenic.
	p.R710P	D	D	D	D	M	T	D	T	26.8	4.45	8.029	.	
MYO9B	p.Q96R	T	B	N	D	M	D	T	T	13.08	4.38	3.216	.	Rare and predicted as benign.
	p.P2085L	T	B	N	N	N	.	T	T	5.007	3.64	2.302	.	
POLA1	p.P496S	T	D	D	D	H	T	D	T	21.4	5.25	7.077	.	Relatively frequent and predicted as pathogenic.
SPTBN5	p.V1968M	T	B	N	N	N	T	T	T	7.804	0.076	0.262	.	Rare and predicted as benign.
	p.A3128T	T	D	D	D	M	T	T	T	19.26	5.27	3.796	.	
TAF1A	c.1086-5T>C	.	.	.	.	.	.	.	.	.	.	.	No	Relatively frequent. c.1086-5T>C predicted as not affecting splicing and p.R435Q predicted as benign.
	p.R435Q	T	B	N	D	N	T	T	T	10.77	-3.51	-1.046	.	
TRAK1	p.E698del	.	.	.	.	.	.	.	.	.	.	.	.	Rare, variant number of triplets in this region.
TTN	p.R18037Q	D	D	.	D	L	T	T	T	15.38	5.78	7.770	.	p.R18037Q rare and p.T21008I relatively frequent. Predicted as benign and p.R18037Q with ClinVar classification of benign.
	p.T21008I	D	B	.	N	N	T	T	T	9.129	-6.48	1.233	.	
UBR3	p.G315D	T	B	.	D	L	T	T	T	13.66	4.67	3.860	.	Rare. p.G315D predicted as benign and c.2663+2T>C as affecting donor splice site.
	c.2663+2T>C	.	.	.	D	.	.	.	.	23.4	5.13	7.968	D	
UMODL1	p.G656R	T	B	.	N	.	T	T	T	7.155	-0.153	-2.487	.	Rare and predicted as benign.
URB1	c.T145C	.	.	.	.	.	.	.	.	.	.	.	A	c.T145C rare and p.T1467M relatively frequent. c.T145C predicted as affect acceptor splice site and p.T1467M as benign.
	p.T1467M	T	B	N	N	L	T	T	T	11.32	2.11	2.184	.	
ZNF182	p.N94K	T	B	.	P	N	T	T	T	0.016	1.53	1.483	.	Relatively frequent and predicted as benign.
ZNF648	p.S378X	T	.	.	D	.	.	.	.	37	2.35	3.951	.	Rare. p.S378X predicted as pathogenic and p.G555S as benign.
	p.G555S	T	B	.	N	N	T	T	T	0.006	-1.95	-1.430	.	

Table S3.18A – Genes, variants and variant frequencies for the analysis of the dominant inheritance model for sample 1. All variants have a European MAF inferior to 0.01 in at least in one database. Chr: chromosome, Ref: reference nucleotide, Alt: altered nucleotide, 1000G: 1000 Genomes Project, ExAC: Exome Aggregation Consortium, ESP: NHLBI Exome Sequencing Project, EUR: European, AFR: African, AMR: American, EAS: East Asian, SAS: South Asian, NFE: Non-Finnish European, FIN: Finnish European, OTH: Other populations, EA: European American, AA: African American. Dot: no information on databases.

Gene	Chr	Start	Ref	Alt	Variant	1000G EUR	1000G AFR	1000G AMR	1000G EAS	1000G SAS	ExAC NFE	ExAC FIN	ExAC AFR	ExAC AMR	ExAC EAS	ExAC SAS	ExAC OTH	ESP6500 EA	ESP6500 AA
AMHR2	chr12	53818947	A	T	c.425-2A>T	.	.	.	.	.	.	.	.	.	.	.	.	.	.
C10orf2	chr10	102748527	A	G	p.D187G	.	.	.	.	.	.	.	.	.	.	.	.	.	.
CLDN9	chr16	3063472	G	C	p.G37R	.	.	.	.	.	.	.	.	.	.	.	.	.	.
GAL3ST4	chr7	99757702	T	C	p.Y437C	.	.	.	.	.	.	.	.	.	.	.	.	.	.
HECA	chr6	139487908	GGCA GT	-	p.A254_V255del	.	.	.	.	.	.	.	.	.	.	.	.	.	.
HKDC1	chr10	71005846	T	C	p.M296T	.	.	.	.	.	.	.	.	.	.	.	.	.	.
KIF3B	chr20	30898190	G	C	p.G204R	.	.	.	.	.	.	.	.	.	.	.	.	.	.
LARS	chr5	145512473	-	T	p.R795KfsX6	.	.	.	.	.	.	.	.	.	.	.	.	.	.
MMP1	chr11	102663340	G	C	p.F343L	.	.	.	.	.	.	.	.	.	.	.	.	.	.
MMP1	chr11	102663332	T	C	c.C1029G+4A>G	.	.	.	.	.	.	.	.	.	.	.	.	.	.
MMP19	chr12	56234449	A	G	c.520+2T>C	.	.	.	.	.	.	.	.	.	.	.	.	.	.
MX2	chr21	42754447	A	T	p.T230S	.	.	.	.	.	.	.	.	.	.	.	.	.	.
MYH7B	chr20	33588479	T	G	p.L1764R	.	.	.	.	.	.	.	.	.	.	.	.	.	.
NEDD4	chr15	56142772	A	C	p.F858V	.	.	.	.	.	.	.	.	.	.	.	.	.	.
NOTCH1	chr9	139410505	C	T	p.A533T	.	.	.	.	.	.	.	.	.	.	.	.	.	.
PCNT	chr21	47805747	C	T	c.C3313T	.	.	.	.	.	.	.	.	.	.	.	.	.	.

Gene	Chr	Start	Ref	Alt	Variant	1000G EUR	1000G AFR	1000G AMR	1000G EAS	1000G SAS	ExAC NFE	ExAC FIN	ExAC AFR	ExAC AMR	ExAC EAS	ExAC SAS	ExAC OTH	ESP6500 EA	ESP6500 AA
RBP3	chr10	48388255	C	T	p.G875S	.	.	.	.	.	.	.	.	.	.	.	.	.	.
RD3	chr1	211652406	G	T	p.P187H	.	.	.	.	.	.	.	.	.	.	.	.	.	.
RRH	chr4	110749271	ACTT	-	p.Y31X	.	.	.	.	.	.	.	.	.	.	.	.	.	.
S1PR3	chr9	91616553	C	A	p.Y146X	.	.	.	.	.	.	.	.	.	.	.	.	.	.
SAMD11	chr1	879103	G	A	p.G572E	.	.	.	.	.	.	.	.	.	.	.	.	.	.
SCN8A	chr12	52115391	T	C	p.F566S	.	.	.	.	.	.	.	.	.	.	.	.	.	.
SLC13A2	chr17	26818848	C	T	p.Q335X	.	.	.	.	.	.	.	.	.	.	.	.	.	.
STAB1	chr3	52551371	C	A	p.H1509N	.	.	.	.	.	.	.	.	.	.	.	.	.	.
URB1	chr21	33694197	G	A	p.Q1800X	.	.	.	.	.	.	.	.	.	.	.	.	.	.

Table S3.18B – Genes, variants and predicted functional impact for the analysis of the dominant inheritance model for sample 1. Each functional impact predictor has a different classification with different cut-offs. SIFT: T – tolerated (>0.05), D – damaging (≤0.05); PolyPhen2: B – benign (<0.452), P – possibly damaging (0.453 and 0.956), D – probably damaging (>0.957); LRT: D – deleterious, N – neutral, U – unknown; MutationTaster: P – polymorphism automatic, N – polymorphism, D – disease causing, A – disease causing automatic; MutationAssessor: N – neutral, L – low, M – medium, H – high; FATHMM: T – tolerated (> -1.5), D – damaging (< -1.5); Radial SVM: T – tolerated, D – damaging; LR: T – tolerated, D – damaging; CADD: ≥15 considered pathogenic; GERP++: ≥2 considered conserved; PhyloP: ≥1; HSF: D – disturbs donor splice site; A – disturbs acceptor splice site; No – predicted to not affect splicing. Dot – no prevision was generated. The last column indicates the gene function.

Gene	Variants	SIFT	Polyphen2	LRT	Mutation Taster	Mutation Assessor	FATHMM	Radial SVM	LR	CADD	GERP++	phyloP	HSF	Gene Function
AMHR2	c.425-2A>T	.	.	.	D	.	.	.	.	16.61	4.66	2.726	A	Receptor for the anti-Mullerian hormone which. in addition to testosterone. results in male sex differentiation.
C10orf2	p.D187G	T	P	D	D	L	D	D	D	15.46	5.71	9.175	.	DNA helicase involved in mitochondrial DNA metabolism. Is thought to play a key role in mtDNA replication.

Gene	Variants	SIFT	Polyphen2	LRT	Mutation Taster	Mutation Assessor	FATHMM	Radial SVM	LR	CADD	GERP++	phyloP	HSF	Gene Function
CLDN9	p.G37R	D	D	D	D	H	D	D	D	21.4	3.82	6.441	.	Member of the claudin family. Plays a major role in tight junction-specific obliteration of the intercellular space. through calcium-independent cell-adhesion activity.
GAL3ST4	p.Y437C	D	D	U	D	M	T	T	T	18.87	5.92	6.273	.	Member of the galactose-3-O-sulfotransferase protein family. Catalyzes the transfer of sulfate to beta-1.3-linked galactose residues in O-linked glycoproteins.
HECA	p.A254_V255del	.	.	.	.	.	.	.	.	.	.	.	.	Cytoplasmic protein that may be part of the regulatory mechanism in the development of epithelial tube networks such as the circulatory system and lungs.
HKDC1	p.M296T	D	P	D	D	M	D	D	D	19.40	4.24	7.730	.	Unknown.
KIF3B	p.G204R	D	D	D	D	L	T	D	T	18.68	3.91	9.657	.	Heterodimer with kinesin family member 3A involved in tethering the chromosomes to the spindle pole and in chromosome movement. Microtubule-based anterograde translocator for membranous organelles
LARS	p.R795KfsX6	.	.	.	.	.	.	.	.	.	.	.	.	Cytosolic leucine-tRNA synthetase that catalyzes the specific attachment of an amino acid to its cognate tRNA.
MMP1	p.F343L	D	D	N	D	H	T	T	T	22.6	6.16	1.512	.	Matrix metalloproteinase that cleaves collagens of types I. II. and III at one site in the helical domain. Also cleaves collagens of types VII and X.
MMP1	c.C1029G+4A>G	.	.	.	.	.	.	.	.	.	.	.	D	Matrix metalloproteinase that cleaves collagens of types I. II. and III at one site in the helical domain. Also cleaves collagens of types VII and X.
MMP19	c.520+2T>C	.	.	.	D	.	.	.	.	22.5	5.95	8.507	D	Matrix metalloproteinase that degrades various components of the extracellular matrix. such as aggrecan and cartilage oligomeric matrix protein. during development. haemostasis and pathological conditions.
MX2	p.T230S	T	P	U	N	M	D	D	D	12.88	-6.6	0.089	.	Interferon-induced dynamin-like GTPase with potent antiviral activity against human immunodeficiency virus type 1.
MYH7B	p.L1764R	D	D	N	D	H	T	D	D	19.39	4.24	7.856	.	A heavy chain of myosin II that is involved in muscle contraction.

Gene	Variants	SIFT	Polyphen2	LRT	Mutation Taster	Mutation Assessor	FATHMM	Radial SVM	LR	CADD	GERP++	phyloP	HSF	Gene Function
NEDD4	p.F858V	D	D	D	D	H	D	D	D	13.78	5.45	9.140	.	E3 ubiquitin-protein ligase which accepts ubiquitin from an E2 ubiquitin-conjugating enzyme in the form of a thioester and then directly transfers the ubiquitin to targeted substrates. NEDD4 upregulated expressions of fibronectin and type 1 collagen and contributed to the excessive accumulation of extracellular matrix.
NOTCH1	p.A533T	T	B	D	D	M	D	D	D	19.38	5.02	4.446	.	Transmembrane receptor that affects the implementation of differentiation. proliferation and apoptotic programs. Involved in angiogenesis; negatively regulates endothelial cell proliferation and migration and angiogenic sprouting.
PCNT	c.C3313T	.	.	.	.	.	.	.	.	.	.	.	A	Integral component of the filamentous matrix of the centrosome. Interacts with the microtubule nucleation component gamma-tubulin and is likely important to normal functioning of the centrosomes. cytoskeleton. and cell-cycle progression.
RBP3	p.G875S	D	D	D	D	M	T	D	D	17.70	5.48	7.282	.	Retinol-binding protein that shuttles 11-cis and all trans retinoids between the retinol isomerase in the pigment epithelium and the visual pigments in the photoreceptor cells of the retina.
RD3	p.P187H	D	D	D	D	M	T	T	T	17.22	4.33	8.499	.	Retinal protein that is associated with promyelocytic leukemia-gene product bodies in the nucleus.
RRH	p.Y31X	.	.	.	.	.	.	.	.	.	.	.	.	Membrane-bound G protein-coupled receptor that may play a role in retinal pigment epithelium physiology either by detecting light directly or by monitoring the concentration of retinoids or other photoreceptor-derived compounds.
S1PR3	p.Y146X	T	.	D	D	.	.	.	.	37	-2.91	-1.473	.	Receptor for sphingosine 1-phosphate and likely contributes to the regulation of angiogenesis and vascular endothelial cell function.
SAMD11	p.G572E	D	D	D	D	H	D	D	D	28.5	5.43	9.423	.	May play a role in photoreceptor development.
SCN8A	p.F566S	D	P	D	D	M	D	D	D	23.5	3.83	7.810	.	Ion pore region of the voltage-gated sodium channel that mediates the voltage-dependent sodium ion

Gene	Variants	SIFT	Polyphen2	LRT	Mutation Taster	Mutation Assessor	FATHMM	Radial SVM	LR	CADD	GERP++	phyloP	HSF	Gene Function
														permeability of excitable membranes
SLC13A2	p.Q335X	T	.	D	A	.	.	.	.	36	5.53	4.705	.	Cotransport of sodium ions and dicarboxylates such as succinate and citrate.
STAB1	p.H1509N	T	D	N	D	M	D	D	D	15.26	4.77	6.156	.	Transmembranar receptor for acetylated low density lipoprotein. Binds to both Gram-positive and Gram-negative bacteria and may play a role in defense against bacterial infection. May have a role in angiogenesis.
URB1	p.Q1800X	T	.	N	A	.	.	.	.	45	3.43	1.999	.	Unknown.

Table S3.19A – Genes, variants and variant frequencies for the analysis of the dominant inheritance model for sample 4. All variants have a European MAF inferior to 0.01 in at least in one database. Chr: chromosome, Ref: reference nucleotide, Alt: altered nucleotide, 1000G: 1000 Genomes Project, ExAC: Exome Aggregation Consortium, ESP: NHLBI Exome Sequencing Project, EUR: European, AFR: African, AMR: American, EAS: East Asian, SAS: South Asian, NFE: Non-Finnish European, FIN: Finnish European, OTH: Other populations, EA: European American, AA: African American. Dot: no information on databases.

Gene	Chr	Start	Ref	Alt	Variant	1000G EUR	1000G AFR	1000G AMR	1000G EAS	1000G SAS	ExAC NFE	ExAC FIN	ExAC AFR	ExAC AMR	ExAC EAS	ExAC SAS	ExAC OTH	ESP6500 EA	ESP6500 AA
ACACB	chr12	109696884	A	G	p.N2156S	.	.	.	.	.	.	.	.	.	.	.	.	.	.
AP1G1	chr16	71798249	T	G	c.927+4A>C	.	.	.	.	.	.	.	.	.	.	.	.	.	.
AP1M1	chr19	16308886	A	T	c.42+7A>T	.	.	.	.	.	.	.	.	.	.	.	.	.	.
AP4S1	chr14	31542180	G	T	c.290+1G>T	.	.	.	.	.	.	.	.	.	.	.	.	.	.
C1QTNF2	chr5	159781873	C	T	p.G94E	.	.	.	.	.	.	.	.	.	.	.	.	.	.
CMYA5	chr5	78985858	C	A	p.S43X	.	.	.	.	.	.	.	.	.	.	.	.	.	.
CPEB2	chr4	15055811	G	A	p.A811T	.	.	.	.	.	.	.	.	.	.	.	.	.	.
CYTH3	chr7	6213288	AGGCT	-	p.Q147HfsX18	.	.	.	.	.	.	.	.	.	.	.	.	.	.
DEFB116	chr20	29896334	G	A	p.Q19X	.	.	.	.	.	.	.	.	.	.	.	.	.	.
DNMT3L	chr21	45674590	T	A	c.605-2A>T	.	.	.	.	.	.	.	.	.	.	.	.	.	.
FASN	chr17	80041513	C	G	p.V1741L	.	.	.	.	.	.	.	.	.	.	.	.	.	.
INTS8	chr8	95839992	G	C	p.Q163H	.	.	.	.	.	.	.	.	.	.	.	.	.	.
ISLR2	chr15	74426933	GC	TT	p.C613F	.	.	.	.	.	.	.	.	.	.	.	.	.	.
KLHDC8A	chr1	205308371	G	C	p.Y236X	.	.	.	.	.	.	.	.	.	.	.	.	.	.
LHPP	chr10	126205748	G	A	c.625-1G>A	.	.	.	.	.	.	.	.	.	.	.	.	.	.
LIPK	chr10	90512445	G	A	p.G378R	.	.	.	.	.	.	.	.	.	.	.	.	.	.

Gene	Chr	Start	Ref	Alt	Variant	1000G EUR	1000G AFR	1000G AMR	1000G EAS	1000G SAS	ExAC NFE	ExAC FIN	ExAC AFR	ExAC AMR	ExAC EAS	ExAC SAS	ExAC OTH	ESP6500 EA	ESP6500 AA
MYCL	chr1	40366916	C	T	p.W64X	.	.	.	.	.	.	.	.	.	.	.	.	.	.
PDZD7	chr10	102770294	-	CTGCGG	p.S784delinsSRS	.	.	.	.	.	.	.	.	.	.	.	.	.	.
PGK2	chr6	49754306	T	C	p.K199E	.	.	.	.	.	.	.	.	.	.	.	.	.	.
PPP1R13B	chr14	104216151	G	A	p.L317F	.	.	.	.	.	.	.	.	.	.	.	.	.	.
SPIB	chr19	50931352	A	G	p.D183G	.	.	.	.	.	.	.	.	.	.	.	.	.	.
TEK	chr9	27169562	T	G	p.V188G	.	.	.	.	.	.	.	.	.	.	.	.	.	.
TMEM174	chr5	72469606	C	T	p.P179L	.	.	.	.	.	.	.	.	.	.	.	.	.	.
UGGT1	chr2	128917238	G	A	p.E836K	.	.	.	.	.	.	.	.	.	.	.	.	.	.

Table S3.19B – Genes, variants and predicted functional impact for the analysis of the dominant inheritance model for sample 4. Each functional impact predictor has a different classification with different cut-offs. SIFT: T – tolerated (>0.05), D – damaging (≤0.05); PolyPhen2: B – benign (<0.452), P – possibly damaging (0.453 and 0.956), D – probably damaging (>0.957); LRT: D – deleterious, N – neutral, U – unknown; MutationTaster: P – polymorphism automatic, N – polymorphism, D – disease causing, A – disease causing automatic; MutationAssessor: N – neutral, L – low, M – medium, H – high; FATHMM: T – tolerated (> -1.5), D – damaging (< -1.5); Radial SVM: T – tolerated, D – damaging; LR: T – tolerated, D – damaging; CADD: ≥15 considered pathogenic; GERP++: ≥2 considered conserved; PhyloP: ≥1; HSF: D – disturbs donor splice site; A – disturbs acceptor splice site; No – predicted to not affect splicing. Dot – no prevision was generated. The last column indicates the gene function.

Gene	Variants	SIFT	Polyphen2	LRT	Mutation Taster	Mutation Assessor	FATHMM	Radial SVM	LR	CADD	GERP++	phyloP	HSF	Gene Function
ACACB	p.N2156S	D	D	D	D	M	D	D	D	23.9	5.07	9.157	.	Acetyl-CoA carboxylase beta that catalyzes the ATP-dependent carboxylation of acetyl-CoA to malonyl-CoA. Carries out three functions: biotin carboxyl carrier protein. biotin carboxylase and carboxyltransferase. Involved in inhibition of fatty acid and glucose oxidation and enhancement of fat storage. May play a role in regulation of mitochondrial fatty acid oxidation through malonyl-CoA-dependent

Gene	Variants	SIFT	Polyphen2	LRT	Mutation Taster	Mutation Assessor	FATHMM	Radial SVM	LR	CADD	GERP++	phyloP	HSF	Gene Function
														inhibition of carnitine palmitoyltransferase 1.
AP1G1	c.927+4A>C	.	.	.	.	.	.	.	.	.	.	.	D	Component of clathrin-coated vesicles transporting ligand-receptor complexes from the plasma membrane or from the trans-Golgi network to lysosomes.
AP1M1	c.42+7A>T	.	.	.	.	.	.	.	.	.	.	.	D	Subunit of clathrin-associated adaptor protein complex 1 that plays a role in protein sorting in the trans-Golgi network and endosomes.
AP4S1	c.290+1G>T	.	.	.	N	.	.	T	T	24.6	5.48	8.265	D	Small subunit of adaptor protein complex-4 that is associated with both clathrin- and nonclathrin-coated vesicles.
C1QTNF2	p.G94E	D	D	D	D	H	D	D	D	20.4	5.16	7.538	.	Unknown.
CMYA5	p.S43X	T	.	N	A	.	.	.	.	18.80	2.73	0.014	.	May serve as an anchoring protein that mediates the subcellular compartmentation of PKA via binding to PRKAR2A. May function as a repressor of calcineurin-mediated transcriptional activity. May attenuate calcineurin ability to induce slow-fiber gene program in muscle and may negatively modulate skeletal muscle regeneration.
CPEB2	p.A811T	D	D	D	D	L	T	T	T	32	5.42	9.837	.	mRNA-binding protein that may play a role in translational regulation of stored mRNAs in transcriptionally inactive haploid spermatids.
CYTH3	p.Q147HfsX18	.	.	.	.	.	.	.	.	.	.	.	.	Member of the pleckstrin homology. Sec7 and coiled-coil domains family involved in the control of Golgi structure and function. and it may have a physiological role in regulating ADP-ARF6 functions. in addition to acting on ARF1.
DEFB116	p.Q19X	T	.	.	D	.	.	.	.	14.13	2.12	0.389	.	Peptide with antibacterial activity.
DNMT3L	c.605-2A>T	.	.	.	D	.	.	.	.	8.014	3.68	3.321	A	Nuclear protein that stimulates *de novo* methylation by DNA cytosine methyltransferase 3 alpha. is thought to be required for the establishment of maternal genomic imprints. mediates transcriptional repression

Gene	Variants	SIFT	Polyphen2	LRT	Mutation Taster	Mutation Assessor	FATHMM	Radial SVM	LR	CADD	GERP++	phyloP	HSF	Gene Function
														through interaction with histone deacetylase 1.
FASN	p.V1741L c.G5221C	D	D	N	D	H	T	T	T	19.68	3.59	7.548	A	Fatty acid synthetase that catalyzes the formation of long-chain fatty acids from acetyl-CoA. malonyl-CoA and NADPH.
INTS8	p.Q163H	D	P	D	D	L	.	T	T	16.85	0.197	1.027	.	Subunit of the integrator complex that is involved in the cleavage of small nuclear RNAs U1 and U2 within the nucleus.
ISLR2	p.C613F	.	.	.	.	.	.	.	.	.	.	..	.	Member of immunoglobulin superfamily containing leucine rich repeat 2 required for axon extension during neural development.
KLHDC8A	p.Y236X	T	.	D	A	.	.	.	.	37	1.91	3.391	.	Kelch domain-containing protein upregulated expression may provide an alternative pathway for tumors to maintain aggressiveness in the absence of epidermal growth factor receptor dependence.
LHPP	c.625-1G>A	.	.	.	D	.	.	.	.	21.0	4.83	7.106	A	Phosphatase that hydrolyzes imidodiphosphate. 3-phosphohistidine and 6-phospholysine.
LIPK	p.G378R	D	D	D	D	H	T	D	D	23.0	5.65	4.450	.	Lipase that plays a highly specific role in the last step of keratinocyte differentiation. May have an essential function in lipid metabolism of the most differentiated epidermal layers.
MYCL	p.W64X	T	.	.	D	.	.	.	.	18.25	4.53	2.880	.	Unknown.
PDZD7	p.S784delinsSRS	.	.	.	.	.	.	.	.	.	.	.	.	Ciliary protein.
PGK2	p.K199E	D	B	D	D	M	D	D	D	14.26	2.89	7.607	.	Phosphoglycerate kinase 2 essential for sperm motility and male fertility.
PPP1R13B	p.L317F	D	D	D	D	M	T	T	T	34	6.17	9.869	.	Regulator that plays a central role in regulation of apoptosis via its interaction with p53/TP53. Regulates TP53 by enhancing the DNA binding and transactivation function of TP53 on the promoters of proapoptotic genes *in vivo*.
SPIB	p.D183G	D	P	D	D	M	T	T	T	17.22	4.08	5.684	.	Transcriptional activator that binds to the PU-box (5'-GAGGAA-3') and acts as a lymphoid-specific enhancer.
TEK	p.V188G	T	D	D	D	L	T	D	D	21.7	5.48	5.490	.	Tyrosine-protein kinase receptor that mediates a

Gene	Variants	SIFT	Polyphen2	LRT	Mutation Taster	Mutation Assessor	FATHMM	Radial SVM	LR	CADD	GERP++	phyloP	HSF	Gene Function
														signaling pathway that functions in embryonic vascular development.
TMEM174	p.P179L	D	D	D	D	M	.	T	T	24.4	5.39	3.395	.	Unknown.
UGGT1	p.E836K	.	P	D	D	M	T	T	T	19.61	4.51	7.365	.	Glucosyltransferase that recognizes glycoproteins with minor folding defects. Reglucosylates single N-glycans near the misfolded part of the protein.

Table S3.20A – Genes, variants and variant frequencies for the analysis of the dominant inheritance model for sample 5. All variants have a European MAF inferior to 0.01 in at least in one database. Chr: chromosome, Ref: reference nucleotide, Alt: altered nucleotide, 1000G: 1000 Genomes Project, ExAC: Exome Aggregation Consortium, ESP: NHLBI Exome Sequencing Project, EUR: European, AFR: African, AMR: American, EAS: East Asian, SAS: South Asian, NFE: Non-Finnish European, FIN: Finnish European, OTH: Other populations, EA: European American, AA: African American. Dot: no information on databases.

Gene	Chr	Start	Ref	Alt	Variant	1000G EUR	1000G AFR	1000G AMR	1000G EAS	1000G SAS	ExAC NFE	ExAC FIN	ExAC AFR	ExAC AMR	ExAC EAS	ExAC SAS	ExAC OTH	ESP6500 EA	ESP6500 AA
ABCA12	chr2	215835097	G	A	p.P1864S	.	.	.	.	.	.	.	.	.	.	.	.	.	.
ABCA13	chr7	48356798	A	G	p.N3302D	.	.	.	.	.	.	.	.	.	.	.	.	.	.
ABCA2	chr9	139910861	G	C	p.L996V	.	.	.	.	.	.	.	.	.	.	.	.	.	.
ABCC1	chr16	16043661	G	A	c.48+5G>A	.	.	.	.	.	.	.	.	.	.	.	.	.	.
ALPI	chr2	233323623	G	C	p.G452R	.	.	.	.	.	.	.	.	.	.	.	.	.	.
CELF3	chr1	151688403	G	A	p.R32W	.	.	.	.	.	.	.	.	.	.	.	.	.	.
CPAMD8	chr19	17014601	G	A	p.L1487F	.	.	.	.	.	.	.	.	.	.	.	.	.	.
CRB2	chr9	126129917	C	T	p.Q336X	.	.	.	.	.	.	.	.	.	.	.	.	.	.
ENPEP	chr4	111469484	T	C	c.2151+2T>C	.	.	.	.	.	.	.	.	.	.	.	.	.	.
FAM63B	chr15	59064427	C	T	p.A278V	.	.	.	.	.	.	.	.	.	.	.	.	.	.
FRAS1	chr4	79236886	A	G	c.1817A>G	.	.	.	.	.	.	.	.	.	.	.	.	.	.
FUT11	chr10	75532804	G	C	c.709+5G>C	.	.	.	.	.	.	.	.	.	.	.	.	.	.
GAD1	chr2	171702070	C	A	p.P269Q	.	.	.	.	.	.	.	.	.	.	.	.	.	.
GUCY1B3	chr4	156715188	G	T	p.D226Y	.	.	.	.	.	.	.	.	.	.	.	.	.	.
MOB4	chr2	198415008	A	G	p.I148M	.	.	.	.	.	.	.	.	.	.	.	.	.	.
KATNAL1	chr13	30784483	T	C	p.Y415C	.	.	.	.	.	.	.	.	.	.	.	.	.	.

Gene	Chr	Start	Ref	Alt	Variant	1000G EUR	1000G AFR	1000G AMR	1000G EAS	1000G SAS	ExAC NFE	ExAC FIN	ExAC AFR	ExAC AMR	ExAC EAS	ExAC SAS	ExAC OTH	ESP6500 EA	ESP6500 AA
KCNC1	chr11	17803262	A	G	p.E580G	.	.	.	.	.	.	.	.	.	.	.	.	.	.
LOXHD1	chr18	44104707	CTT	-	p.K457del	.	.	.	.	.	.	.	.	.	.	.	.	.	.
MSH3	chr5	80071522	G	C	p.E755Q	.	.	.	.	.	.	.	.	.	.	.	.	.	.
NAB1	chr2	191535068	C	T	p.A279V	.	.	.	.	.	.	.	.	.	.	.	.	.	.
NARS	chr18	55273141	A	G	p.L400P	.	.	.	.	.	.	.	.	.	.	.	.	.	.
PAPD5	chr16	50187912	CCTCGGCGTCCTCGCCTC	-	p.S112_P118del	.	.	.	.	.	.	.	.	.	.	.	.	.	.
PBXIP1	chr1	154920800	-	CATCGTCACTGC	p.D151delinsGSDDD	.	.	.	.	.	.	.	.	.	.	.	.	.	.
PXYLP1	chr3	141011529	C	T	p.P309S	.	.	.	.	.	.	.	.	.	.	.	.	.	.
RAB32	chr6	146870719	C	A	p.L124M	.	.	.	.	.	.	.	.	.	.	.	.	.	.
SDR42E1	chr16	82034398	A	G	c.66T>C	.	.	.	.	.	.	.	.	.	.	.	.	.	.
TGS1	chr8	56708534	C	G	c.1368-3C>G	.	.	.	.	.	.	.	.	.	.	.	.	.	.
TRADD	chr16	67189277	C	G	c.429+3G>C	.	.	.	.	.	.	.	.	.	.	.	.	.	.
TRPT1	chr11	63993049	G	A	p.R27X	.	.	.	.	.	.	.	.	.	.	.	.	.	.
ZNF880	chr19	52876363	G	C	c.13-1G>C	.	.	.	.	.	.	.	.	.	.	.	.	.	.

Table S3.20B – Genes, variants and predicted functional impact for the analysis of the dominant inheritance model for sample 5. Each functional impact predictor has a different classification with different cut-offs. SIFT: T – tolerated (>0.05), D – damaging (≤0.05); PolyPhen2: B – benign (<0.452), P – possibly damaging (0.453 and 0.956), D – probably damaging (>0.957); LRT: D – deleterious, N – neutral, U – unknown; MutationTaster: P – polymorphism automatic, N – polymorphism, D – disease causing, A – disease causing automatic; MutationAssessor: N – neutral, L – low, M – medium, H – high; FATHMM: T – tolerated (> -1.5), D – damaging (< -1.5); Radial SVM: T – tolerated, D – damaging; LR: T – tolerated, D – damaging; CADD: ≥15 considered pathogenic; GERP++: ≥2 considered conserved; PhyloP: ≥1; HSF: D – disturbs donor splice site; A – disturbs acceptor splice site; No – predicted to not affect splicing. Dot – no prevision was generated. The last column indicates the gene function.

Gene	Variants	SIFT	Polyphen2	LRT	Mutation Taster	Mutation Assessor	FATHMM	Radial SVM	LR	CADD	GERP++	phyloP	HSF	Gene Function
ABCA12	p.P1864S	D	D	D	D	M	D	D	D	25.0	5.29	4.873	.	Probable transporter involved in lipid homeostasis.
ABCA13	p.N3302D	T	D	N	D	M	D	D	D	19.09	5.77	4.117	.	Member of ABC transmembrane transporters subfamily A.
ABCA2	p.L996V	T	D	D	D	L	D	D	D	20.7	4.23	9.582	.	Probable transporter. its natural substrate has not been found yet. May have a role in macrophage lipid metabolism and neural development.
ABCC1	c.48+5G>A	.	.	.	.	.	.	.	.	.	.	.	D	Transporter that mediates export of organic anions and drugs from the cytoplasm. Mediates ATP-dependent transport of glutathione and glutathione conjugates. leukotriene C4. estradiol-17-beta-o-glucuronide. methotrexate. antiviral drugs and other xenobiotics. Confers resistance to anticancer drugs. Hydrolyzes ATP with low efficiency. Not found expressed in HTM cell cultures.
ALPI	p.G452R	D	D	D	D	H	D	D	D	15.07	5.15	4.399	.	A digestive brush-border enzyme that is a component of the gut mucosal defense system and is thought to function in the detoxification of lipopolysaccharide. and in the prevention of bacterial translocation in the gut.
CELF3	p.R32W	D	D	D	D	M	T	T	T	17.17	2.59	1.904	.	RNA-binding protein involved in the regulation of pre-mRNA alternative splicing.
CPAMD8	p.L1487F	T	P	U	D	M	.	T	T	21.2	3.32	7.765	.	Member of complement component-3/alpha-2-macroglobulin family that is involved in innate immunity and damage control.
CRB2	p.Q336X	T	.	N	D	.	.	.	.	36	4.13	0.115	.	Component of the crumbs cell polarity complex that may play a role in polarized cells morphogenesis.

Gene	Variants	SIFT	Polyphen2	LRT	Mutation Taster	Mutation Assessor	FATHMM	Radial SVM	LR	CADD	GERP++	phyloP	HSF	Gene Function
ENPEP	c.2151+2T>C	.	.	.	D	.	.	.	.	19.19	5.5	6.507	D	Glutamyl aminopeptidase that appears to have a role in the catabolic pathway of the renin-angiotensin system. Probably plays a role in regulating growth and differentiation of early B-lineage cells.
FAM63B	p.A278V	D	P	D	D	N	T	T	T	34	4.95	7.664	.	Unknown.
FRAS1	c.1817A>G	.	.	.	.	.	.	.	.	.	.	.	D	Extracellular matrix protein that appears to function in the regulation of epidermal-basement membrane adhesion and organogenesis during development.
FUT11	c.709+5G>C	.	.	.	.	.	.	.	.	.	.	.	D	Probably a fucosyltransferase.
GAD1	p.P269Q	D	D	D	D	M	T	T	T	35	5.67	7.487	.	Glutamic acid decarboxylase that catalyzes the production of GABA.
GUCY1B3	p.D226Y	D	D	D	D	M	D	D	D	19.49	3.26	5.446	.	Subunit of the soluble guanylate cyclase mediates responses to nitric oxide by catalyzing the biosynthesis of the signaling molecule cGMP.
MOB4	p.I148M	T	P	D	D	M	.	T	T	15.93	5.24	5.028	.	May play a role in membrane trafficking. specifically in membrane budding reactions.
KATNAL1	p.Y415C	D	D	D	D	M	D	D	D	23.1	5.77	6.140	.	Regulates microtubule dynamics in Sertoli cells. a process that is essential for spermiogenesis and male fertility.
KCNC1	p.E580G	D	.	.	D	.	D	D	D	15.34	5.49	7.147	.	Potassium voltage-gated channel that mediates the voltage-dependent potassium ion permeability of excitable membranes.
LOXHD1	p.K457del	.	.	.	.	.	.	.	.	.	.	.	.	Lipoxygenase involved in hearing. Required for normal function of hair cells in the inner ear.
MSH3	p.E755Q	.	D	D	D	M	D	D	D	21.1	5.54	7.151	.	Component of the post-replicative DNA mismatch repair system MMR.
NAB1	p.A279V	D	B	D	D	M	.	T	T	33	5.33	7.231	.	Acts as a transcriptional repressor for zinc finger transcription factors EGR1 and EGR2.
NARS	p.L400P	D	D	D	D	H	T	D	D	22.9	5.76	8.548	.	Asparaginyl-tRNA synthetase.
PAPD5	p.S112_P118del	.	.	.	.	.	.	.	.	.	.	.	.	Catalytic subunit of a TRAMP-like complex that has a poly(A) RNA polymerase activity and is involved in a post-transcriptional quality control mechanism. Plays

Gene	Variants	SIFT	Polyphen2	LRT	Mutation Taster	Mutation Assessor	FATHMM	Radial SVM	LR	CADD	GERP++	phyloP	HSF	Gene Function
														a role in replication-dependent histone mRNA degradation. probably through terminal uridylation of mature histone mRNAs.
PBXIP1	p.D151delinsGSDDD	.	.	.	.	.	.	.	.	.	.	.	.	Interacts with the PBX1 homeodomain protein. inhibiting its transcriptional activation potential by preventing its binding to DNA. Can interact with estrogen receptors alpha and beta and promote the proliferation of breast cancer. brain tumors. and lung cancer.
PXYLP1	p.P309S	D	D	D	D	M	T	D	D	24.4	5.58	7.484	.	Phosphatase responsible for the 2-O-dephosphorylation of xylose in the glycosaminoglycan-protein linkage region of proteoglycans thereby regulating the amount of mature glycosaminoglycan chains.
RAB32	p.L124M	D	D	D	D	M	T	D	D	18.59	4.65	2.903	.	Anchors the type II regulatory subunit of protein kinase A to the mitochondrion and aids in mitochondrial fission. Appears to be involved in autophagy and melanosome secretion.
SDR42E1	c.66T>C	.	.	.	.	.	.	.	.	.	.	.	D	Member of the large short-chain dehydrogenase/reductase (SDR) family of enzymes that metabolize steroid hormones. prostaglandins. retinoids. lipids. and xenobiotics.
TGS1	c.1368-3C>G	.	.	.	.	.	.	.	.	.	.	.	A	Trimethylguanosine synthase that catalyzes the 2 serial methylation steps for the conversion of the 7-monomethylguanosine caps of snRNAs and snoRNAs to a 2.2.7-trimethylguanosine cap structure. Plays a role in transcriptional regulation.
TRADD	c.429+3G>C	.	.	.	.	.	.	.	.	.	.	.	D	The nuclear form acts as a tumor suppressor by preventing ubiquitination and degradation of isoform p19ARF/ARF of CDKN2A by TRIP12. Adapter molecule for TNFRSF1A/TNFR1 that specifically associates with the cytoplasmic domain of activated TNFRSF1A/TNFR1 mediating its interaction with FADD. Overexpression of TRADD leads to two major TNF-induced responses. apoptosis and activation of NF-kappa-B.

Gene	Variants	SIFT	Polyphen2	LRT	Mutation Taster	Mutation Assessor	FATHMM	Radial SVM	LR	CADD	GERP++	phyloP	HSF	Gene Function
TRPT1	p.R27X	T	.	N	A	.	.	.	.	34	1.05	1.103	.	Phosphotransferase that catalyzes the last step of tRNA splicing. the transfer of the splice junction 2-phosphate from ligated tRNA to NAD to produce ADP-ribose 1-2 cyclic phosphate.
ZNF880	c.13-1G>C	.	.	.	N	.	.	.	.	0.479	-1.14	0.776	A	Unknown.

Table S3.21A – Genes, variants and variant frequencies for the analysis of the dominant inheritance model for sample 9. All variants have a European MAF inferior to 0.01 in at least in one database. Chr: chromosome, Ref: reference nucleotide, Alt: altered nucleotide, 1000G: 1000 Genomes Project, ExAC: Exome Aggregation Consortium, ESP: NHLBI Exome Sequencing Project, EUR: European, AFR: African, AMR: American, EAS: East Asian, SAS: South Asian, NFE: Non-Finnish European, FIN: Finnish European, OTH: Other populations, EA: European American, AA: African American. Dot: no information on databases.

Gene	Chr	Start	Ref	Alt	Variant	1000G EUR	1000G AFR	1000G AMR	1000G EAS	1000G SAS	ExAC NFE	ExAC FIN	ExAC AFR	ExAC AMR	ExAC EAS	ExAC SAS	ExAC OTH	ESP6500 EA	ESP6500 AA
ADAMTS20	chr12	43833841	GAAAA	-	p.L773QfsX23	.	.	.	.	.	.	.	.	.	.	.	.	.	.
ANO8	chr19	17438495	C	G	c.2418+3G>C	.	.	.	.	.	.	.	.	.	.	.	.	.	.
BBS1	chr11	66294123	G	T	p.G395V	.	.	.	.	.	.	.	.	.	.	.	.	.	.
C6orf226	chr6	42858223	A	G	p.X102R	.	.	.	.	.	.	.	.	.	.	.	.	.	.
CCDC168	chr13	103401051	C	T	p.G666S	.	.	.	.	.	.	.	.	.	.	.	.	.	.
DAGLB	chr7	6470186	G	A	p.A285V	.	.	.	.	.	.	.	.	.	.	.	.	.	.
DCDC2	chr6	24205311	A	-	p.G315EfsX32	.	.	.	.	.	.	.	.	.	.	.	.	.	.
DNAH2	chr17	7701878	A	G	c.8401A>G	.	.	.	.	.	.	.	.	.	.	.	.	.	.
DOCK5	chr8	25216560	A	G	p.T978A	.	.	.	.	.	.	.	.	.	.	.	.	.	.
EXOC1	chr4	56730490	T	A	p.L118X	.	.	.	.	.	.	.	.	.	.	.	.	.	.
GCH1	chr14	55369188	T	A	p.E65V	.	.	.	.	.	.	.	.	.	.	.	.	.	.
HIBADH	chr7	27671983	CTT	-	p.E111del	.	.	.	.	.	.	.	.	.	.	.	.	.	.
HSD3B2	chr1	119964447	T	G	p.L108W	.	.	.	.	.	.	.	.	.	.	.	.	.	.
HSPG2	chr1	22206692	G	C	p.R751G	.	.	.	.	.	.	.	.	.	.	.	.	.	.
IGFBP1	chr7	45931529	A	G	c.520-2A>G	.	.	.	.	.	.	.	.	.	.	.	.	.	.
NRL	chr14	24552044	G	C	p.P5R	.	.	.	.	.	.	.	.	.	.	.	.	.	.

Gene	Chr	Start	Ref	Alt	Variant	1000G EUR	1000G AFR	1000G AMR	1000G EAS	1000G SAS	ExAC NFE	ExAC FIN	ExAC AFR	ExAC AMR	ExAC EAS	ExAC SAS	ExAC OTH	ESP6500 EA	ESP6500 AA
RNF14	chr5	141358287	C	G	p.Y242X	.	.	.	.	.	.	.	.	.	.	.	.	.	.
SOCS1	chr16	11349196	-	CCGGGA	p.A47delinsVPA	.	.	.	.	.	.	.	.	.	.	.	.	.	.
STAB1	chr3	52537012	C	G	c.584-3C>G	.	.	.	.	.	.	.	.	.	.	.	.	.	.
TNFRSF11A	chr18	60025572	C	T	c.C519T	.	.	.	.	.	.	.	.	.	.	.	.	.	.
WNT8B	chr10	102240879	G	A	c.G366A														

Table S3.21B – Genes, variants and predicted functional impact for the analysis of the dominant inheritance model for sample 9. Each functional impact predictor has a different classification with different cut-offs. SIFT: T – tolerated (>0.05), D – damaging (≤0.05); PolyPhen2: B – benign (<0.452), P – possibly damaging (0.453 and 0.956), D – probably damaging (>0.957); LRT: D – deleterious, N – neutral, U – unknown; MutationTaster: P – polymorphism automatic, N – polymorphism, D – disease causing, A – disease causing automatic; MutationAssessor: N – neutral, L – low, M – medium, H – high; FATHMM: T – tolerated (> -1.5), D – damaging (< -1.5); Radial SVM: T – tolerated, D – damaging; LR: T – tolerated, D – damaging; CADD: ≥15 considered pathogenic; GERP++: ≥2 considered conserved; PhyloP: ≥1; HSF: D – disturbs donor splice site; A – disturbs acceptor splice site; No – predicted to not affect splicing. Dot – no prevision was generated. The last column indicates the gene function.

Gene	Variants	SIFT	Polyphen2	LRT	Mutation Taster	Mutation Assessor	FATHMM	Radial SVM	LR	CADD	GERP++	phyloP	HSF	Gene Function
ADAMTS20	p.L773QfsX23	.	.	.	.	.	.	.	.	.	.	.	.	Metallopeptidase that may play a role in tissue-remodeling process occurring in both normal and pathological conditions. May have a protease-independent function in the transport from the endoplasmic reticulum to the Golgi apparatus of secretory cargos.
ANO8	c.2418+3G>C	.	.	.	.	.	.	.	.	.	.	.	D	Transmembrane protein that does not exhibit calcium-activated chloride channel CaCC activity.
BBS1	p.G395V	D	D	D	.	D	M	T	D	D	23.5	4.69	.	Required for proper BBSome complex assembly and its ciliary localization.
C6orf226	p.X102R	.	.	.	.	N	.	.	.	.	3.873	-0.118	.	Unknown.

Gene	Variants	SIFT	Polyphen2	LRT	Mutation Taster	Mutation Assessor	FATHMM	Radial SVM	LR	CADD	GERP++	phyloP	HSF	Gene Function
CCDC168	p.G666S	.	.	.	.	.	.	.	.	.	.	.	.	Unknown.
DAGLB	p.A285V	D	D	P	D	D	M	T	T	T	26.2	4.25	.	Lipase that catalyzes the hydrolysis of diacylglycerol to 2-arachidonoyl-glycerol. the most abundant endocannabinoid in tissues. Required for axonal growth during development and for retrograde synaptic signaling at mature synapses.
DCDC2	p.G315EfsX32	.	.	.	.	.	.	.	.	.	.	.	.	Plays a role in the inhibition of canonical Wnt signaling pathway. May be involved in neuronal migration during development of the cerebral neocortex. Involved in the control of ciliogenesis and ciliary length.
DNAH2	c.8401A>G	T	P	B	D	D	L	T	T	T	16.45	5.91	A	Force generating protein of respiratory cilia. Produces force towards the minus ends of microtubules. Dynein has ATPase activity; the force-producing power stroke is thought to occur on release of ADP. Involved in sperm motility; implicated in sperm flagellar assembly.
DOCK5	p.T978A	D	P	P	D	D	M	T	T	T	18.57	5.96	.	Guanine nucleotide exchange factor for Rho and Rac. Along with DOCK1. mediates CRK/CRKL regulation of epithelial and endothelial cell spreading and migration on type IV collagen.
EXOC1	p.L118X	T	.	.	D	A	.	.	.	.	40	5.28	.	Component of the exocyst complex involved in the docking of exocytic vesicles with fusion sites on the plasma membrane.
GCH1	p.E65V	D	P	B	D	D	L	D	D	D	15.86	3.39	.	GTP cyclohydrolase involved in tetrahydrobiopterin biosynthesis. catalyzing the conversion of GTP into 7.8-dihydroneopterin triphosphate. Tetrahydrobiopterin is an essential cofactor required by aromatic amino acid hydroxylases as well as nitric oxide synthases.
HIBADH	p.E111del	.	.	.	.	.	.	.	.	.	.	.	.	Involved in the mitochondrial function of spermatozoa. and maintains sperm motility. It may serve as a sperm-motility marker.
HSD3B2	p.L108W	.	D	D	D	D	H	D	D	D	16.79	4.1	.	Bifunctional enzyme that catalyzes the oxidative conversion of delta(5)-ene-3-beta-hydroxy steroid.

Gene	Variants	SIFT	Polyphen2	LRT	Mutation Taster	Mutation Assessor	FATHMM	Radial SVM	LR	CADD	GERP++	phyloP	HSF	Gene Function
														and the oxidative conversion of ketosteroids. It plays a crucial role in the biosynthesis of all classes of hormonal steroids.
HSPG2	p.R751G	D	D	D	N	D	M	T	D	D	21.1	5.38	.	Proteoglycan that binds to and cross-links many extracellular matrix components and cell-surface molecules. It has been shown that this protein interacts with laminin. prolargin. collagen type IV. FGFBP1. FBLN2. FGF7 and transthyretin. etc.. and it plays essential roles in multiple biological activities.
IGFBP1	c.520-2A>G	.	.	.	.	D	.	.	.	.	9.588	3.81	A	Member of the insulin-like growth factor binding protein family that binds both insulin-like growth factors I and II and circulates in the plasma. Binding of this protein prolongs the half-life of the insulin-like growth factors and alters their interaction with cell surface receptors.Promotes cell migration.
NRL	p.P5R	D	P	P	N	D	L	D	D	D	18.56	4.29	.	Transcription factor which regulates the expression of several rod-specific genes. including RHO and PDE6B.
RNF14	p.Y242X	T	.	.	D	A	.	.	.	.	39	4.53	.	Might act as an E3 ubiquitin-protein ligase that accepts ubiquitin from specific E2 ubiquitin-conjugating enzymes and then transfers it to substrates. which could be nuclear proteins. Could play a role as a coactivator for androgen- and. to a lesser extent. progesterone-dependent transcription.
SOCS1	p.A47delinsVPA	.	.	.	.	.	.	.	.	.	.	.	.	Member of the STAT-induced STAT inhibitor involved in negative regulation of cytokines that signal through the JAK/STAT3 pathway. Through binding to JAKs. inhibits their kinase activity. Appears to be a major regulator of signaling by interleukin 6 and leukemia inhibitory factor. Regulates interferon-gamma mediated sensory neuron survival. Appears to be a negative regulator in IGF1R signaling pathway.
STAB1	c.584-3C>G	.	.	.	.	.	.	.	.	.	.	.	D	Acts as a scavenger receptor for acetylated low density lipoprotein. Binds to both Gram-positive and Gram-negative bacteria and may play a role in defense against bacterial infection. Probable role in

Gene	Variants	SIFT	Polyphen2	LRT	Mutation Taster	Mutation Assessor	FATHMM	Radial SVM	LR	CADD	GERP++	phyloP	HSF	Gene Function
														angiogenesis. Involved in the delivery of newly synthesized CHID1/SI-CLP from the biosynthetic compartment to the endosomal/lysosomal system.
TNFRSF11A	c.C519T	.	.	.	.	.	.	.	.	.	.	.	D	Receptor for TNFSF11/RANKL/TRANCE/OPGL; essential for RANKL-mediated osteoclastogenesis. Involved in the regulation of interactions between T-cells and dendritic cells.
WNT8B	c.G366A	.	.	.	.	.	.	.	.	.	.	.	D	Ligand for members of the frizzled family of seven transmembrane receptors. May play an important role in the development and differentiation of certain forebrain structures. notably the hippocampus.

Table S3.22A – Genes, variants and variant frequencies for the analysis of the dominant inheritance model for sample 18. All variants have a European MAF inferior to 0.01 in at least in one database. Chr: chromosome, Ref: reference nucleotide, Alt: altered nucleotide, 1000G: 1000 Genomes Project, ExAC: Exome Aggregation Consortium, ESP: NHLBI Exome Sequencing Project, EUR: European, AFR: African, AMR: American, EAS: East Asian, SAS: South Asian, NFE: Non-Finnish European, FIN: Finnish European, OTH: Other populations, EA: European American, AA: African American. Dot: no information on databases.

Gene	Chr	Start	Ref	Alt	Variant	1000G EUR	1000G AFR	1000G AMR	1000G EAS	1000G SAS	ExAC NFE	ExAC FIN	ExAC AFR	ExAC AMR	ExAC EAS	ExAC SAS	ExAC OTH	ESP6500 EA	ESP6500 AA
CEP295	chr11	93433149	C	T	p.Q1691X	.	.	.	.	.	.	.	.	.	.	.	.	.	.
DGKZ	chr11	46391528	G	A	p.R368H	.	.	.	.	.	.	.	.	.	.	.	.	.	.
EPB41L3	chr18	5406797	G	C	p.I776M	.	.	.	.	.	.	.	.	.	.	.	.	.	.
FGFR3	chr4	1807174	C	T	p.P502L	.	.	.	.	.	.	.	.	.	.	.	.	.	.
KRTAP5-1	chr11	1606349	ACA	TTG	p.V44_C49del	.	.	.	.	.	.	.	.	.	.	.	.	.	.
MAP2K3	chr17	21205571	TG	-	c.516_c.516+1 delTG	.	.	.	.	.	.	.	.	.	.	.	.	.	.
OBSCN	chr1	228563804	T	C	p.C7628R	.	.	.	.	.	.	.	.	.	.	.	.	.	.
PCDH1	chr5	141243485	C	A	p.G804V	.	.	.	.	.	.	.	.	.	.	.	.	.	.
POLG	chr15	89864365	C	T	p.G909S	.	.	.	.	.	.	.	.	.	.	.	.	.	.
POLQ	chr3	121206921	T	-	p.K1619NfsX2	.	.	.	.	.	.	.	.	.	.	.	.	.	.
RGCC	chr13	42032550	T	G	p.M60R	.	.	.	.	.	.	.	.	.	.	.	.	.	.
SLC23A3	chr2	220029065	A	G	c.1168-5T>C	.	.	.	.	.	.	.	.	.	.	.	.	.	.
SYNJ2	chr6	158454579	-	CGTG TATG CCTC CCAC AAGC AGGC CAAG GCCT	p.S210CfsX15	.	.	.	.	.	.	.	.	.	.	.	.	.	.

Gene	Chr	Start	Ref	Alt	Variant	1000G EUR	1000G AFR	1000G AMR	1000G EAS	1000G SAS	ExAC NFE	ExAC FIN	ExAC AFR	ExAC AMR	ExAC EAS	ExAC SAS	ExAC OTH	ESP6500 EA	ESP6500 AA
				GCCT CGTC TCT															
TMC6	chr17	76121963	G	A	c.C274T, p.R92C	.	.	.	.	.	.	.	.	.	.	.	.	.	.
TNXB	chr6	32011576	G	A	p.S3825L	.	.	.	.	.	.	.	.	.	.	.	.	.	.
USF3	chr3	113376142	GCTG CT	-	p.Q1462_Q1463 del	.	.	.	.	.	.	.	.	.	.	.	.	.	.

Table S3.22B – Genes, variants and predicted functional impact for the analysis of the dominant inheritance model for sample 18. Each functional impact predictor has a different classification with different cut-offs. SIFT: T – tolerated (>0.05), D – damaging (≤0.05); PolyPhen2: B – benign (<0.452), P – possibly damaging (0.453 and 0.956), D – probably damaging (>0.957); LRT: D – deleterious, N – neutral, U – unknown; MutationTaster: P – polymorphism automatic, N – polymorphism, D – disease causing, A – disease causing automatic; MutationAssessor: N – neutral, L – low, M – medium, H – high; FATHMM: T – tolerated (> -1.5), D – damaging (< -1.5); Radial SVM: T – tolerated, D – damaging; LR: T – tolerated, D – damaging; CADD: ≥15 considered pathogenic; GERP++: ≥2 considered conserved; PhyloP: ≥1; HSF: D – disturbs donor splice site; A – disturbs acceptor splice site; No – predicted to not affect splicing. Dot – no prevision was generated. The last column indicates the gene function.

Gene	Variants	SIFT	Polyphen2	LRT	Mutation Taster	Mutation Assessor	FATHMM	Radial SVM	LR	CADD	GERP++	phyloP	HSF	Gene Function
CEP295	p.Q1691X	D	.	N	N	.	.	.	.	41	2.78	0.907	.	Centriole-enriched protein that mediates centriole-to-centrosome conversion at late mitosis. but is dispensable for cartwheel removal or centriole disengagement.
DGKZ	p.R368H	T	P	D	D	M	D	D	D	26.8	3.57	9.800	.	Diacylglycerol kinase that may attenuate protein kinase C activity by regulating diacylglycerol levels in intracellular signaling cascade and signal transduction.
EPB41L3	p.I776M	T	D	D	D	M	T	D	D	16.95	-4.75	-0.120	.	Tumor suppressor that inhibits cell proliferation and promotes apoptosis. Modulates the activity of protein arginine N-methyltransferases. including PRMT3 and PRMT5.
FGFR3	p.P502L	T	B	D	D	L	D	D	D	15.65	3.79	3.748	.	Tyrosine kinase that acts as cell-surface receptor for fibroblast growth factors and plays an essential role in the regulation of cell proliferation. differentiation and apoptosis. Plays an essential role in the regulation of

Gene	Variants	SIFT	Polyphen2	LRT	Mutation Taster	Mutation Assessor	FATHMM	Radial SVM	LR	CADD	GERP++	phyloP	HSF	Gene Function
														chondrocyte differentiation. proliferation and apoptosis. and is required for normal skeleton development. Regulates both osteogenesis and postnatal bone mineralization by osteoblasts. Promotes apoptosis in chondrocytes. but can also promote cancer cell proliferation. Required for normal development of the inner ear.
KRTAP5-1	p.V44_C49del	.	.	.	.	.	.	.	.	.	.	.	.	Component of interfilamentous matrix that are essential for the formation of a rigid and resistant hair shaft through their extensive disulfide bond cross-linking with abundant cysteine residues of hair keratins. The matrix proteins include the high-sulfur and high-glycine-tyrosine keratins.
MAP2K3	c.516_c.516+1 delTG	.	.	.	.	.	.	.	.	.	.	.	D	Dual specificity kinase. Is activated by cytokines and environmental stress *in vivo*. Catalyzes the concomitant phosphorylation of a threonine and a tyrosine residue in the MAP kinase p38.
OBSCN	p.C7628R	D	D	.	D	M	T	T	T	23.8	5.74	5.049	.	Giant sarcomeric signaling protein involved in myofibrillogenesis. Seems to be involved in assembly of myosin into sarcomeric A bands in striated muscle. Isoform 3 together with ANK1 isoform Mu17/Ank1.5 may provide a molecular link between the sarcoplasmic reticulum and myofibrils.
PCDH1	p.G804V	D	P	D	D	M	T	T	T	15.00	5.23	2.897	.	Membrane protein found at cell-cell boundaries. Involved in neural cell adhesion. Cells expressing the protein showed cell aggregation activity.
POLG	p.G909S	T	B	D	D	M	D	D	D	19.38	4.31	3.761	.	Catalytic subunit of mitochondrial DNA polymerase involved in the replication of mitochondrial DNA.
POLQ	p.K1619NfsX2	.	.	.	.	.	.	.	.	.	.	.	.	DNA polymerase that promotes microhomology-mediated end-joining. an alternative non-homologous end-joining machinery triggered in response to double-strand breaks in DNA.
RGCC	p.M60R	D	D	D	D	M	.	T	T	31	5.32	6.981	.	Modulates the activity of cell cycle-specific kinases. Enhances CDK1 activity. May contribute to the regulation of the cell cycle.

Gene	Variants	SIFT	Polyphen2	LRT	Mutation Taster	Mutation Assessor	FATHMM	Radial SVM	LR	CADD	GERP++	phyloP	HSF	Gene Function
SLC23A3	c.1168-5T>C	.	.	.	.	.	.	.	.	.	.	.	A	Unknown.
SYNJ2	p.S210CfsX15	.	.	.	.	.	.	.	.	.	.	.	.	Inositol 5-phosphatase which may be involved in distinct membrane trafficking and signal transduction pathways. May mediate the inhibitory effect of Rac1 on endocytosis.
TMC6	c.C274T, p.R92C	D	D	U	D	M	T	T	T	22.7	4.02	6.653	D	Probable ion channel.
TNXB	p.S3825L	D	P	D	D	M	T	T	T	22.8	5.91	5.539	.	Extracellular matrix glycoprotein that have anti-adhesive effects. as opposed to fibronectin which is adhesive. Functions in matrix maturation during wound healing.
USF3	p.Q1462_Q1463 del	.	.	.	.	.	.	.	.	.	.	.	.	Unknown.

Table S3.23A – Genes, variants and variant frequencies for the analysis of the dominant inheritance model for sample 25. All variants have a European MAF inferior to 0.01 in at least in one database. Chr: chromosome, Ref: reference nucleotide, Alt: altered nucleotide, 1000G: 1000 Genomes Project, ExAC: Exome Aggregation Consortium, ESP: NHLBI Exome Sequencing Project, EUR: European, AFR: African, AMR: American, EAS: East Asian, SAS: South Asian, NFE: Non-Finnish European, FIN: Finnish European, OTH: Other populations, EA: European American, AA: African American. Dot: no information on databases.

Gene	Chr	Start	Ref	Alt	Variant	1000G EUR	1000G AFR	1000G AMR	1000G EAS	1000G SAS	ExAC NFE	ExAC FIN	ExAC AFR	ExAC AMR	ExAC EAS	ExAC SAS	ExAC OTH	ESP6500 EA	ESP6500 AA
ARHGEF26	chr3	153870725	AGTG CACT GC	-	c.1487+4_c.1487+13delAGTGCACTGC	.	.	.	.	.	.	.	.	.	.	.	.	.	.
ATP13A4	chr3	193156832	C	T	p.D848L	.	.	.	.	.	.	.	.	.	.	.	.	.	.
C3orf30	chr3	118866204	-	TA	p.Q391fsX11	.	.	.	.	.	.	.	.	.	.	.	.	.	.
CARD9	chr9	139259011	C	T	c.G1435A; p.D479N	.	.	.	.	.	.	.	.	.	.	.	.	.	.
ELP4	chr11	31653952	G	T	c.G927T; p.Q309H	.	.	.	.	.	.	.	.	.	.	.	.	.	.
GAREM1	chr18	29867016	G	A	p.P515L	.	.	.	.	.	.	.	.	.	.	.	.	.	.
GBP6	chr1	89843768	T	-	p.L142RfsX4	.	.	.	.	.	.	.	.	.	.	.	.	.	.
GLDC	chr9	6605193	G	C	p.P267A	.	.	.	.	.	.	.	.	.	.	.	.	.	.
GRIK4	chr11	120856676	CACC CCCG CCGG CGGC GCGC CG	-	p.H860QfsX188	.	.	.	.	.	.	.	.	.	.	.	.	.	.
LAMA3	chr18	21427580	G	T	p.G1362C	.	.	.	.	.	.	.	.	.	.	.	.	.	.
LRRFIP1	chr2	238617259	C	T	p.R47W	.	.	.	.	.	.	.	.	.	.	.	.	.	.
MAPK4	chr18	48190472	GAA	-	p.K49del	.	.	.	.	.	.	.	.	.	.	.	.	.	.
NCKAP1L	chr12	54901637	G	C	c.G307C; p.D103H	.	.	.	.	.	.	.	.	.	.	.	.	.	.

Gene	Chr	Start	Ref	Alt	Variant	1000G EUR	1000G AFR	1000G AMR	1000G EAS	1000G SAS	ExAC NFE	ExAC FIN	ExAC AFR	ExAC AMR	ExAC EAS	ExAC SAS	ExAC OTH	ESP6500 EA	ESP6500 AA
NEFH	chr22	29876963	G	A	p.V238M	.	.	.	.	.	.	.	.	.	.	.	.	.	.
NELFE	chr6	31922352	TCTCGG	-	p.R240_D241del	.	.	.	.	.	.	.	.	.	.	.	.	.	.
NEURL1	chr10	105349921	C	G	p.P506R	.	.	.	.	.	.	.	.	.	.	.	.	.	.
PAFAH2	chr1	26308852	T	C	c.666+3A>G	.	.	.	.	.	.		.	.	.	.	.	.	.
PRKAR1B	chr7	720282	G	T	p.P87T	.	.	.	.	.	.	.	.	.	.	.	.	.	.
PWWP2B	chr10	134230562	A	G	c.1487-3A>G	.	.	.	.	.	.	.	.	.	.	.	.	.	.
SEMA6D	chr15	48057163	G	T	p.G446V	.	.	.	.	.	.	.	.	.	.	.	.	.	.
SPDL1	chr5	169028529	A	-	p.M524CfsX3	.	.	.	.	.	.	.	.	.	.	.	.	.	.
TRPM3	chr9	73233934	G	C	p.R596T	.	.	.	.	.	.	.	.	.	.	.	.	.	.

Table S3.23B – Genes, variants and predicted functional impact for the analysis of the dominant inheritance model for sample 25. Each functional impact predictor has a different classification with different cut-offs. SIFT: T – tolerated (>0.05), D – damaging (≤0.05); PolyPhen2: B – benign (<0.452), P – possibly damaging (0.453 and 0.956), D – probably damaging (>0.957); LRT: D – deleterious, N – neutral, U – unknown; MutationTaster: P – polymorphism automatic, N – polymorphism, D – disease causing, A – disease causing automatic; MutationAssessor: N – neutral, L – low, M – medium, H – high; FATHMM: T – tolerated (> -1.5), D – damaging (< -1.5); Radial SVM: T – tolerated, D – damaging; LR: T – tolerated, D – damaging; CADD: ≥15 considered pathogenic; GERP++: ≥2 considered conserved; PhyloP: ≥1; HSF: D – disturbs donor splice site; A – disturbs acceptor splice site; No – predicted to not affect splicing. Dot – no prevision was generated. The last column indicates the gene function.

Gene	Variants	SIFT	Polyphen2	LRT	Mutation Taster	Mutation Assessor	FATHMM	Radial SVM	LR	CADD	GERP++	phyloP	HSF	Gene Function
ARHGEF26	c.1487+4_c.1487+13delAGTGCACTGC	.	.	.	.	.	.	.	.	.	.	.	D	Regulates Rho GTPases by catalyzing the exchange of GDP for GTP. Required for the formation of cup-like structures during trans-endothelial migration of leukocytes.
ATP13A4	p.D848L	D	D	D	D	H	D	D	D	35	5.96	7.433	.	Cation-transporting P-type ATPase.

Gene	Variants	SIFT	Polyphen2	LRT	Mutation Taster	Mutation Assessor	FATHMM	Radial SVM	LR	CADD	GERP++	phyloP	HSF	Gene Function
C3orf30	p.Q391fsX11	.	.	.	.	.	.	.	.	.	.	.	.	Might play an important role in spermatogenesis.
CARD9	c.G1435A; p.D479N	T	B	N	D	L	T	T	T	14.10	2.73	0.780	A	Adapter protein that plays a key role in innate immune response to a number of intracellular pathogens. Is at the crossroads of ITAM-tyrosine kinase and the Toll-like receptors and NOD2 signaling pathways.
ELP4	c.G927T; p.Q309H	T	D	D	D	M	T	T	T	15.68	4.0	5.804	A	Unknown.
GAREM1	p.P515L	D	D	D	D	L	T	T	T	24.4	5.26	9.145	.	Adaptor protein that functions in the epidermal growth factor receptor-mediated signaling pathway.
GBP6	p.L142RfsX4	.	.	.	.	.	.	.	.	.	.	.	.	Guanylate-binding protein that hydrolyze GTP to both GDP and GMP.
GLDC	p.P267A	D	D	D	D	H	D	D	D	26.5	5.45	9.199	.	Member of glycine cleavage system that catalyzes the degradation of glycine. Binds the alpha-amino group of glycine through its pyridoxal phosphate cofactor.
GRIK4	p.H860QfsX188	.	.	.	.	.	.	.	.	.	.	.	.	Receptor for glutamate.
LAMA3	p.G1362C	D	D	.	D	H	T	D	D	19.46	5.54	9.433	.	Alpha subunit of laminin responsive to several epithelial-mesenchymal regulators including keratinocyte growth factor. epidermal growth factor and insulin-like growth factor.
LRRFIP1	p.R47W	D	D	D	D	M	T	T	T	17.09	4.69	0.712	.	Transcriptional repressor that may regulate expression of TNF. EGFR and PDGFA. May also bind double-stranded RNA. Positively regulates Toll-like receptor signaling in response to agonist probably by competing with the negative FLII regulator for MYD88-binding.
MAPK4	p.K49del	.	.	.	.	.	.	.	.	.	.	.	.	Atypical MAPK protein that phosphorylates microtubule-associated protein 2 and MAPKAPK5.
NCKAP1L	c.G307C; p.D103H	.	D	D	D	M	T	T	T	23.5	5.59	9.371	A	Transmembrane protein that is a part of the Scar/WAVE complex. Only expressed in hematopoietic cells.
NEFH	p.V238M	D	P	N	D	L	D	D	D	11.39	3.44	4.517	.	Heavy neurofilament protein involved in the maintenance of neuronal caliber. Has an important

Gene	Variants	SIFT	Polyphen2	LRT	Mutation Taster	Mutation Assessor	FATHMM	Radial SVM	LR	CADD	GERP++	phyloP	HSF	Gene Function
														function in mature axons that is not subserved by the two smaller neurofilament proteins.
NELFE	p.R240_D241del	.	.	.	.	.	.	.	.	.	.	.	.	Essential component of the NELF complex. a complex that negatively regulates the elongation of transcription by RNA polymerase II.
NEURL1	p.P506R	T	D	D	D	M	.	T	D	26.5	5.64	7.487	.	Plays a role in hippocampal-dependent synaptic plasticity. learning and memory. Involved in the formation of spines and functional synaptic contacts by modulating the translational activity of the cytoplasmic polyadenylation element-binding protein CPEB3. Promotes ubiquitination of CPEB3. and hence induces CPEB3-dependent mRNA translation activation of glutamate receptor GRIA1 and GRIA2.
PAFAH2	c.666+3A>G	.	.	.	.	.	.	.	.	.	.	.	D	Lipase that catalyzes the removal of the acetyl group at the SN-2 position of platelet-activating factor (identified as 1-O-alkyl-2-acetyl-sn-glyceryl-3-phosphorylcholine). Broader substrate specificity than simply platelet activating factor.
PRKAR1B	p.P87T	T	B	U	D	M	D	D	D	14.47	4.91	5.377	.	Regulatory subunit of the cAMP-dependent protein kinases involved in cAMP signaling in cells.
PWWP2B	c.1487-3A>G	.	.	.	.	.	.	.	.	.	.	.	A	Unknown.
SEMA6D	p.G446V	D	D	D	D	H	T	D	D	32	5.59	9.869	.	Transmembrane semaphorin that shows growth cone collapsing activity on dorsal root ganglion neurons *in vitro*. May be a stop signal for the dorsal root ganglion neurons in their target areas. and possibly also for other neurons. May also be involved in the maintenance and remodeling of neuronal connections.
SPDL1	p.M524CfsX3	.	.	.	.	.	.	.	.	.	.	.	.	Required for the localization of dynein and dynactin to the mitotic kintochore. Also required for correct spindle orientation. Does not appear to be required for the removal of spindle assembly checkpoint proteins from the kinetochore upon bipolar spindle attachment.

Gene	Variants	SIFT	Polyphen2	LRT	Mutation Taster	Mutation Assessor	FATHMM	Radial SVM	LR	CADD	GERP++	phyloP	HSF	Gene Function
TRPM3	p.R596T	D	D	D	D	M	T	T	T	32	5.65	9.618	.	Calcium channel mediating constitutive calcium ion entry.

Table S3.24A – Genes, variants and variant frequencies for the analysis of the dominant inheritance model for sample 31. All variants have a European MAF inferior to 0.01 in at least in one database. Chr: chromosome, Ref: reference nucleotide, Alt: altered nucleotide, 1000G: 1000 Genomes Project, ExAC: Exome Aggregation Consortium, ESP: NHLBI Exome Sequencing Project, EUR: European, AFR: African, AMR: American, EAS: East Asian, SAS: South Asian, NFE: Non-Finnish European, FIN: Finnish European, OTH: Other populations, EA: European American, AA: African American. Dot: no information on databases.

Gene	Chr	Start	Ref	Alt	Variant	1000G EUR	1000G AFR	1000G AMR	1000G EAS	1000G SAS	ExAC NFE	ExAC FIN	ExAC AFR	ExAC AMR	ExAC EAS	ExAC SAS	ExAC OTH	ESP6500 EA	ESP6500 AA
AKT1	14	105246532	C	T	p.R23Q	.	.	.	.	.	.	.	.	.	.	.	.	.	.
APC	5	112175787	G	T	p.G1499V	.	.	.	.	.	.	.	.	.	.	.	.	.	.
C15orf65	15	55710697	A	G	p.N99D	.	.	.	.	.	.	.	.	.	.	.	.	.	.
CACNA1A	19	13318676	CTG	-	p.Q2324del	.	.	.	.	.	.	.	.	.	.	.	.	.	.
CAPRIN2	12	30888114	C	A	p.Q199H	.	.	.	.	.	.	.	.	.	.	.	.	.	.
CHKB	22	51018819	G	C	p.P235R	.	.	.	.	.	.	.	.	.	.	.	.	.	.
CNNM3	2	97482579	C	T	p.R189C	.	.	.	.	.	.	.	.	.	.	.	.	.	.
CNOT6	5	179994190	G	C	p.E274Q	.	.	.	.	.	.	.	.	.	.	.	.	.	.
DALRD3	3	49055543	A	C	p.L125R	.	.	.	.	.	.	.	.	.	.	.	.	.	.
DNAH5	5	13841104	T	C	p.T1874A	.	.	.	.	.	.	.	.	.	.	.	.	.	.
DNAH5	5	13718996	T	C	p.Y4165C	.	.	.	.	.	.	.	.	.	.	.	.	.	.
ECM2	9	95274334	T	G	p.N377H	.	.	.	.	.	.	.	.	.	.	.	.	.	.
GADD45G	9	92220908	G	T	p.D132Y	.	.	.	.	.	.	.	.	.	.	.	.	.	.
GSTZ1	14	77793842	C	G	p.P55A	.	.	.	.	.	.	.	.	.	.	.	.	.	.
IGHMBP2	11	68703704	G	A	c.1757-1G>A	.	.	.	.	.	.	.	.	.	.	.	.	.	.
IRF2BPL	14	77493648	GCGG CGGC GGCG	-	p.A159_A162del	.	.	.	.	.	.	.	.	.	.	.	.	.	.

Gene	Chr	Start	Ref	Alt	Variant	1000G EUR	1000G AFR	1000G AMR	1000G EAS	1000G SAS	ExAC NFE	ExAC FIN	ExAC AFR	ExAC AMR	ExAC EAS	ExAC SAS	ExAC OTH	ESP6500 EA	ESP6500 AA
LRTOMT	11	71806478	T	A	p.V164D	.	.	.	.	.	.	.	.	.	.	.	.	.	.
MFGE8	15	89449129	A	G	p.F182L	.	.	.	.	.	.	.	.	.	.	.	.	.	.
MID2	X	107170053	G	A	p.R653H	.	.	.	.	.	.	.	.	.	.	.	.	.	.
MTMR2	11	95580967	G	A	p.R364X	.	.	.	.	.	.	.	.	.	.	.	.	.	.
PLEKHG5	1	6537588	C	T	c.211+1G>A	.	.	.	.	.	.	.	.	.	.	.	.	.	.
PRPH	12	49689431	G	T	p.E150X	.	.	.	.	.	.	.	.	.	.	.	.	.	.
PTAFR	1	28476553	TCAG TGAC CGTA	-	p.T324_E327del	.	.	.	.	.	.	.	.	.	.	.	.	.	.
SEMA5A	5	9197377	A	C	p.L324R	.	.	.	.	.	.	.	.	.	.	.	.	.	.
TMEM201	1	9670657	C	T	p.S520F	.	.	.	.	.	.	.	.	.	.	.	.	.	.
TRIOBP	22	38121230	-	A	p.N890KfsX12	.	.	.	.	.	.	.	.	.	.	.	.	.	.
WDR11	10	122624656	C	T	p.L271F	.	.	.	.	.	.	.	.	.	.	.	.	.	.
WRAP73	1	3552552	C	-	p.A187PfsX67	.	.	.	.	.	.	.	.	.	.	.	.	.	.
ZNF74	22	20760491	T	C	p.C390R	.	.	.	.	.	.	.	.	.	.	.	.	.	.

Table S3.24B – Genes, variants and predicted functional impact for the analysis of the dominant inheritance model for sample 31. Each functional impact predictor has a different classification with Table S3.4B different cut-offs. SIFT: T – tolerated (>0.05), D – damaging (≤0.05); PolyPhen2: B – benign (<0.452), P – possibly damaging (0.453 and 0.956), D – probably damaging (>0.957); LRT: D – deleterious, N – neutral, U – unknown; MutationTaster: P – polymorphism automatic, N – polymorphism, D – disease causing, A – disease causing automatic; MutationAssessor: N – neutral, L – low, M – medium, H – high; FATHMM: T – tolerated (> -1.5), D – damaging (< -1.5); Radial SVM: T – tolerated, D – damaging; LR: T – tolerated, D – damaging; CADD: ≥15 considered pathogenic; GERP++: ≥2 considered conserved; PhyloP: ≥1; HSF: D – disturbs donor splice site; A – disturbs acceptor splice site; No – predicted to not affect splicing. Dot – no prevision was generated. The last column indicates the gene function.

Gene	Variants	SIFT	Polyphen2	LRT	Mutation Taster	Mutation Assessor	FATHMM	Radial SVM	LR	CADD	GERP++	phyloP	HSF	Gene Function
AKT1	p.R23Q	D	D	D	D	M	T	D	D	25.3	4.61	3.683	.	Serine-threonine protein kinase activated by platelet-derived growth factor through phosphatidylinositol 3-kinase. In the developing nervous system AKT is a critical mediator of growth factor-induced neuronal survival. Survival factors can suppress apoptosis in a transcription-independent manner by activating the serine/threonine kinase AKT1. which then phosphorylates and inactivates components of the apoptotic machinery.
APC	p.G1499V	.	D	D	D	L	T	T	D	21.4	6.16	7.632	.	Tumor suppressor protein that acts as an antagonist of the Wnt signaling pathway. Also involved in other processes including cell migration and adhesion. transcriptional activation. and apoptosis.
C15orf65	p.N99D	.	.	.	.	.	.	.	.	.	.	.	.	Unknown.
CACNA1A	p.Q2324del	.	.	.	.	.	.	.	.	.	.	.	.	Voltage-sensitive calcium channels subunit that gives rise to P and/or Q-type calcium currents.
CAPRIN2	p.Q199H	D	D	N	D	N	T	D	D	23.0	3.72	2.652	.	May regulate the transport and translation of mRNAs. modulating for instance the expression of proteins involved in synaptic plasticity in neurons. Involved in regulation of growth as erythroblasts shift from a highly proliferative state towards their terminal phase of differentiation. May be involved in apoptosis.
CHKB	p.P235R	T	D	D	D	M	T	T	T	21.7	5.41	9.461	.	Kinase that has a key role in phospholipid biosynthesis. Catalyzes the first step in phosphatidylethanolamine biosynthesis. Phosphorylates ethanolamine. and can also act on choline. Has higher activity with ethanolamine. May

Gene	Variants	SIFT	Polyphen2	LRT	Mutation Taster	Mutation Assessor	FATHMM	Radial SVM	LR	CADD	GERP++	phyloP	HSF	Gene Function
														not significantly contribute to *in vivo* phosphatidylcholine biosynthesis.
CNNM3	p.R189C	D	D	U	D	M	D	D	D	14.38	3.68	0.185	.	Probable metal transporter.
CNOT6	p.E274Q	.	D	D	D	M	T	T	T	26.5	5.06	9.813	.	Catalytic component of the CCR4-NOT core transcriptional regulation complex. The encoded protein has a 3'-5' RNase activity and prefers polyadenylated substrates. The CCR4-NOT complex plays a role in many cellular processes. including miRNA-mediated repression. mRNA degradation. and transcriptional regulation.
DALRD3	p.L125R	D	D	D	D	M	T	T	T	19.71	4.92	3.779	.	Protein with a DALR anticodon binding domain similar to that of class Ia aminoacyl tRNA synthetases.
DNAH5	p.T1874A	D	P	D	D	M	T	T	T	18.36	5.77	7.997	.	Dynein that force generating protein of respiratory cilia. Produces force towards the minus ends of microtubules. Has ATPase activity; the force-producing power stroke is thought to occur on release of ADP. Required for structural and functional integrity of the cilia of ependymal cells lining the brain ventricles.
DNAH5	p.Y4165C	D	D	D	D	H	T	T	T	16.40	5.77	6.149	.	Dynein that force generating protein of respiratory cilia. Produces force towards the minus ends of microtubules. Has ATPase activity; the force-producing power stroke is thought to occur on release of ADP. Required for structural and functional integrity of the cilia of ependymal cells lining the brain ventricles.
ECM2	p.N377H	D	D	D	D	H	T	D	D	23.9	5.68	6.866	.	Matrix glycoprotein that promotes matrix assembly and cell adhesiveness.
GADD45G	p.D132Y	D	D	D	D	M	T	T	T	18.28	4.38	4.100	.	Involved in the regulation of growth and apoptosis. Mediates activation of stress-responsive MTK1/MEKK4 MAPKKK.
GSTZ1	p.P55A	D	D	D	D	M	T	T	T	20.3	4.71	3.994	.	Glutathione S-transferase that catalyzes the conversion of maleylacetoacetate to fumarylacetoacatate. one of the steps in the phenylalanine/tyrosine degradation pathway.

Gene	Variants	SIFT	Polyphen2	LRT	Mutation Taster	Mutation Assessor	FATHMM	Radial SVM	LR	CADD	GERP++	phyloP	HSF	Gene Function
IGHMBP2	c.1757-1G>A	.	.	.	D	.	.	.	.	15.54	4.25	8.953	A	Helicase that unwinds RNA and DNA duplices in an ATP-dependent reaction. Acts as a transcription regulator. Required for the transcriptional activation of the flounder liver-type antifreeze protein gene. Binds to the insulin II gene RIPE3B enhancer region. Interacts with tRNA-Tyr. Stimulates the transcription of the human neurotropic virus JCV.
IRF2BPL	p.A159_A162del	.	.	.	.	.	.	.	.	.	.	.	.	Transcription factor that may contribute to the control of female reproductive function. May play a role in gene transcription by transactivating GNRH1 promoter and repressing PENK promoter.
LRTOMT	p.V164D	D	D	D	D	M	T	D	T	14.63	5.5	5.621	.	Transmembrane O-Methyltransferase that catalyzes the O-methylation. and thereby the inactivation. of catecholamine neurotransmitters and catechol hormones. Required for auditory function.
MFGE8	p.F182L	D	D	D	D	M	D	D	D	18.56	4.97	6.911	.	A preproprotein that is proteolytically processed to form multiple protein products. The major encoded protein product. lactadherin. is a membrane glycoprotein that promotes phagocytosis of apoptotic cells. This protein has also been implicated in wound healing. autoimmune disease. and cancer. Lactadherin can be further processed to form a smaller cleavage product. medin. which comprises the major protein component of aortic medial amyloid.
MID2	p.R653H	D	D	D	D	M	T	D	D	23.5	5.37	9.302	.	May play a role in microtubule stabilization.
MTMR2	p.R364X	T	.	D	A	.	.	.	.	27.8	3.62	2.823	.	Phosphoinositide lipid phosphatase that possesses phosphatase activity towards phosphatidylinositol-3-phosphate and phosphatidylinositol-3.5-bisphosphate.
PLEKHG5	c.211+1G>A	.	.	.	D	.	.	.	.	18.26	4.52	6.528	D	Guanine nucleotide exchange factor that activates RHOA and maybe the NF-kappa-B signaling pathway. Involved in the control of neuronal cell differentiation. Plays a role in angiogenesis through regulation of endothelial cells chemotaxis.

Gene	Variants	SIFT	Polyphen2	LRT	Mutation Taster	Mutation Assessor	FATHMM	Radial SVM	LR	CADD	GERP++	phyloP	HSF	Gene Function
PRPH	p.E150X	D	.	N	A	.	.	.	.	49	3.78	6.361	.	Type III intermediate filament cytoskeletal protein found in neurons of the peripheral nervous system.
PTAFR	p.T324_E327del	.	.	.	.	.	.	.	.	.	.	.	.	Receptor for platelet activating factor. a chemotactic phospholipid mediator that possesses potent inflammatory. smooth-muscle contractile and hypotensive activity. Seems to mediate its action via a G protein that activates a phosphatidylinositol-calcium second messenger system.
SEMA5A	p.L324R	D	D	D	D	M	T	T	T	23.0	5.31	9.075	.	Membrane semaphorin protein involved in axonal guidance during neural development.
TMEM201	p.S520F	D	D	D	D	L	.	T	T	23.7	5.19	5.416	.	Transmembranar protein that may be involved in mitotic spindle.
TRIOBP	p.N890KfsX12	.	.	.	.	.	.	.	.	.	.	.		May regulate actin cytoskeletal organization. cell spreading and cell contraction by directly binding and stabilizing filamentous F-actin. The localized formation of TARA and TRIO complexes coordinates the amount of F-actin present in stress fibers. May also serve as a linker protein to recruit proteins required for F-actin formation and turnover.
WDR11	p.L271F	T	P	D	D	L	D	D	D	18.68	5.06	3.766	.	Member of the WD repeat protein family.
WRAP73	p.A187PfsX67	.	.	.	.	.	.	.	.	.	.	.	.	Member of the WD repeat protein family that regulates the spindle anchoring at the mitotic centrosome. Involved in ciliary vesicle formation at the mother centriole and required for the docking of vesicles to the basal body during ciliogenesis.
ZNF74	p.C390R	D	D	D	D	H	D	D	D	18.55	3.97	4.647	.	May play a role in RNA metabolism.

Table S3.25A – Genes, variants and variant frequencies for the analysis of the dominant inheritance model for sample 32. All variants have a European MAF inferior to 0.01 in at least in one database. Chr: chromosome, Ref: reference nucleotide, Alt: altered nucleotide, 1000G: 1000 Genomes Project, ExAC: Exome Aggregation Consortium, ESP: NHLBI Exome Sequencing Project, EUR: European, AFR: African, AMR: American, EAS: East Asian, SAS: South Asian, NFE: Non-Finnish European, FIN: Finnish European, OTH: Other populations, EA: European American, AA: African American. Dot: no information on databases.

Gene	Chr	Start	Ref	Alt	Variant	1000G EUR	1000G AFR	1000G AMR	1000G EAS	1000G SAS	ExAC NFE	ExAC FIN	ExAC AFR	ExAC AMR	ExAC EAS	ExAC SAS	ExAC OTH	ESP6500 EA	ESP6500 AA
AATK	17	79102268	A	G	c.414+2T>C	.	.	.	.	.	.	.	.	.	.	.	.	.	.
CCDC85A	2	56419910	G	A	p.S192N	.	.	.	.	.	.	.	.	.	.	.	.	.	.
CYP27B1	12	58158547	C	-	p.G318EfsX36	.	.	.	.	.	.	.	.	.	.	.	.	.	.
FARP2	2	242350479	C	A	p.L148M	.	.	.	.	.	.	.	.	.	.	.	.	.	.
FMO5	1	146672982	G	C	p.A312G	.	.	.	.	.	.	.	.	.	.	.	.	.	.
GLT6D1	9	138516179	A	G	p.Y199H	.	.	.	.	.	.	.	.	.	.	.	.	.	.
OTX1	2	63282997	C	A	p.S204X	.	.	.	.	.	.	.	.	.	.	.	.	.	.
PSMD11	17	30791074	T	A	p.L109Q	.	.	.	.	.	.	.	.	.	.	.	.	.	.
QSER1	11	32956606	G	A	p.E1139K	.	.	.	.	.	.	.	.	.	.	.	.	.	.
ZBED6	1	203767844	AAC	-	p.T399del	.	.	.	.	.	.	.	.	.	.	.	.	.	.
ZNF106	15	42737173	G	A	p.S1048F	.	.	.	.	.	.	.	.	.	.	.	.	.	.

Table S3.25B – Genes, variants and predicted functional impact for the analysis of the dominant inheritance model for sample 32. Each functional impact predictor has a different classification with different cut-offs. SIFT: T – tolerated (>0.05), D – damaging (≤0.05); PolyPhen2: B – benign (<0.452), P – possibly damaging (0.453 and 0.956), D – probably damaging (>0.957); LRT: D – deleterious, N – neutral, U – unknown; MutationTaster: P – polymorphism automatic, N – polymorphism, D – disease causing, A – disease causing automatic; MutationAssessor: N – neutral, L – low, M – medium, H – high; FATHMM: T – tolerated (> -1.5), D – damaging (< -1.5); Radial SVM: T – tolerated, D – damaging; LR: T – tolerated, D – damaging; CADD: ≥15 considered pathogenic; GERP++: ≥2 considered conserved; PhyloP: ≥1; HSF: D – disturbs donor splice site; A – disturbs acceptor splice site; No – predicted to not affect splicing. Dot – no prevision was generated. The last column indicates the gene function.

Gene	Variants	SIFT	Polyphen2	LRT	Mutation Taster	Mutation Assessor	FATHMM	Radial SVM	LR	CADD	GERP++	phyloP	HSF	Gene Function
AATK	c.414+2T>C	.	.	.	D	.	.	.	.	9.046	2.89	4.721	D	Tyrosine kinase induced during apoptosis. Its expression may be a necessary pre-requisite for the induction of growth arrest and/or apoptosis of myeloid precursor cells. Produce neuronal differentiation in a neuroblastoma cell line.
CCDC85A	p.S192N	T	D	D	D	M	.	T	T	25.0	5.18	9.470	.	Unknown.
CYP27B1	p.G318EfsX36	.	.	.	.	.	.	.	.	.	.	.	.	Monooxygenase localized to the inner mitochondrial membrane where it hydroxylates 25-hydroxyvitamin D3 at the 1alpha position. This enzyme regulates the level of biologically active vitamin D and plays an important role in calcium homeostasis.
FARP2	p.L148M	D	D	D	D	H	T	T	T	18.05	4.47	0.448	.	Functions as guanine nucleotide exchange factor that activates RAC1. May have relatively low activity. Plays a role in the response to class 3 semaphorins and remodeling of the actin cytoskeleton. Plays a role in TNFSF11-mediated osteoclast differentiation. especially in podosome rearrangement and reorganization of the actin cytoskeleton. Regulates the activation of ITGB3. integrin signaling and cell adhesion.
FMO5	p.A312G	.	D	D	D	M	T	T	T	23.1	5.26	4.563	.	Flavin-containing monooxygenase. Does not seem to be a drug-metabolizing enzyme.
GLT6D1	p.Y199H	D	D	D	D	M	T	T	T	16.33	3.49	4.374	.	GT6 glycosyltransferase.
OTX1	p.S204X	T	.	D	D	.	.	.	.	39	3.61	7.453	.	A transcription factor that probably plays a role in the development of the brain and the sense organs.
PSMD11	p.L109Q	D	D	D	D	M	T	D	D	22.2	5.61	7.512	.	Component of the lid subcomplex of the 26S

Gene	Variants	SIFT	Polyphen2	LRT	Mutation Taster	Mutation Assessor	FATHMM	Radial SVM	LR	CADD	GERP++	phyloP	HSF	Gene Function
														proteasome. Required for proteasome assembly. Plays a key role in increased proteasome activity in embryonic stem cells ESCs: its high expression in ESCs promotes enhanced assembly of the 26S proteasome. followed by higher proteasome activity.
QSER1	p.E1139K	D	D	D	D	M	T	T	T	21.2	5.58	9.476	.	Unknown.
ZBED6	p.T399del	.	.	.	.	.	.	.	.	.	.	.	.	Transcriptional repressor which binds to the consensus sequence 5-GCTCGC-3 and represses transcription of IGF2. May also regulate expression of other target genes containing this consensus binding site.
ZNF106	p.S1048F	D	D	N	D	L	T	D	D	21.3	5.42	4.181	.	Unknown.

Table S3.26A – Genes, variants and variant frequencies for the analysis of the dominant inheritance model for sample 33. All variants have a European MAF inferior to 0.01 in at least in one database. Chr: chromosome, Ref: reference nucleotide, Alt: altered nucleotide, 1000G: 1000 Genomes Project, ExAC: Exome Aggregation Consortium, ESP: NHLBI Exome Sequencing Project, EUR: European, AFR: African, AMR: American, EAS: East Asian, SAS: South Asian, NFE: Non-Finnish European, FIN: Finnish European, OTH: Other populations, EA: European American, AA: African American. Dot: no information on databases.

Gene	Chr	Start	Ref	Alt	Variant	1000G EUR	1000G AFR	1000G AMR	1000G EAS	1000G SAS	ExAC NFE	ExAC FIN	ExAC AFR	ExAC AMR	ExAC EAS	ExAC SAS	ExAC OTH	ESP6500 EA	ESP6500 AA
ACAN	15	89391176	G	C	p.G547R	.	.	.	.	.	.	.	.	.	.	.	.	.	.
AFTPH	2	64778942	A	T	p.S112C	.	.	.	.	.	.	.	.	.	.	.	.	.	.
AP4S1	14	31542180	G	T	c.294+1G>T	.	.	.	.	.	.	.	.	.	.	.	.	.	.
ARSE	X	2867685	C	G	p.G172R	.	.	.	.	.	.	.	.	.	.	.	.	.	.
C4orf22	4	81504305	A	G	p.R101G	.	.	.	.	.	.	.	.	.	.	.	.	.	.
CASP5	11	104874001	C	T	c.G543A	.	.	.	.	.	.	.	.	.	.	.	.	.	.
DCDC5	11	30953392	C	-	p.L20FfsX31	.	.	.	.	.	.	.	.	.	.	.	.	.	.
DPEP2	16	68021480	CTGAGA	-	p.V462_S463del	.	.	.	.	.	.	.	.	.	.	.	.	.	.
FKRP	19	47259236	G	A	p.E177K	.	.	.	.	.	.	.	.	.	.	.	.	.	.
GBX2	2	237076427	GGCGGC	-	p.P62_P63del	.	.	.	.	.	.	.	.	.	.	.	.	.	.
GGA1	22	38026124	G	-	p.K428RfsX88	.	.	.	.	.	.	.	.	.	.	.	.	.	.
HSF2	6	122733821	T	C	p.V81A	.	.	.	.	.	.	.	.	.	.	.	.	.	.
IPO8	12	30784898	G	C	p.L983V	.	.	.	.	.	.	.	.	.	.	.	.	.	.
MAN2B1	19	12761012	G	C	p.Q691E	.	.	.	.	.	.	.	.	.	.	.	.	.	.
MAN2C1	15	75653626	C	G	c.1314+1G>C	.	.	.	.	.	.	.	.	.	.	.	.	.	.
MUSK	9	113562640	G	A	p.G661D	.	.	.	.	.	.	.	.	.	.	.	.	.	.

Gene	Chr	Start	Ref	Alt	Variant	1000G EUR	1000G AFR	1000G AMR	1000G EAS	1000G SAS	ExAC NFE	ExAC FIN	ExAC AFR	ExAC AMR	ExAC EAS	ExAC SAS	ExAC OTH	ESP6500 EA	ESP6500 AA
PNLIP	10	118307868	A	G	c.202-4A>G	.	.	.	.	.	.	.	.	.	.	.	.	.	.
RREB1	6	7246973	-	GAG GGC	p.E1346_G1377 dup	.	.	.	.	.	.	.	.	.	.	.	.	.	.
WNK4	17	40933028	-	C	p.E107RfsX38	.	.	.	.	.	.	.	.	.	.	.	.	.	.

Table S3.26B – Genes, variants and predicted functional impact for the analysis of the dominant inheritance model for sample 33. Each functional impact predictor has a different classification with different cut-offs. SIFT: T – tolerated (>0.05), D – damaging (≤0.05); PolyPhen2: B – benign (<0.452), P – possibly damaging (0.453 and 0.956), D – probably damaging (>0.957); LRT: D – deleterious, N – neutral, U – unknown; MutationTaster: P – polymorphism automatic, N – polymorphism, D – disease causing, A – disease causing automatic; MutationAssessor: N – neutral, L – low, M – medium, H – high; FATHMM: T – tolerated (> -1.5), D – damaging (< -1.5); Radial SVM: T – tolerated, D – damaging; LR: T – tolerated, D – damaging; CADD: ≥15 considered pathogenic; GERP++: ≥2 considered conserved; PhyloP: ≥1; HSF: D – disturbs donor splice site; A – disturbs acceptor splice site; No – predicted to not affect splicing. Dot – no prevision was generated. The last column indicates the gene function.

Gene	Variants	SIFT	Polyphen2	LRT	Mutation Taster	Mutation Assessor	FATHMM	Radial SVM	LR	CADD	GERP++	phyloP	HSF	Gene Function
ACAN	p.G547R	D	D	.	D	M	T	T	T	21.3	5.35	9.044	.	Proteoglycan component of extracellular matrix of cartilagenous tissues. Resist to compression in cartilage. Binds avidly to hyaluronic acid via an N-terminal globular region.
AFTPH	p.S112C	D	D	D	D	M	T	T	T	15.70	5.85	4.835	.	May play a role in membrane trafficking.
AP4S1	c.294+1G>T	.	.	.	N	.	.	T	T	24.6	5.48	8.265	D	Subunit of novel type of clathrin- or non-clathrin-associated protein coat involved in targeting proteins from the trans-Golgi network TGN to the endosomal-lysosomal system.
ARSE	p.G172R	D	D	D	D	H	D	D	D	14.92	3.56	6.351	.	Sulfatase that may be essential for the correct composition of cartilage and bone matrix during development. Has no activity toward steroid sulfates.
C4orf22	p.R101G	D	D	D	D	M	T	T	T	13.52	4.36	4.445	.	Unknown.
CASP5	c.G543A	.	.	.	.	.	.	.	.	.	.	.	D	Caspase mediator of programmed cell death (apoptosis).
DCDC5	p.L20FfsX31	.	.	.	.	.	.	.	.	.	.	.	.	Member of the doublecortin family that may bind tubulin and enhance microtubule polymerization.
DPEP2	p.V462_S463del	.	.	.	.	.	.	.	.	.	.	.	.	Membrane-bound dipeptidase that hydrolyzes leukotriene D4 into leukotriene E4.
FKRP	p.E177K	T	B	D	D	L	D	D	D	10.39	3.52	3.115	.	Transferase involved in the biosynthesis of the phosphorylated O-mannosyl trisaccharide (N-acetylgalactosamine-beta-3-N-acetylglucosamine-beta-4-(phosphate-6-)mannose).
GBX2	p.P62_P63del	.	.	.	.	.	.	.	.	.	.	.	.	May act as a transcription factor for cell pluripotency and differentiation in the embryo.
GGA1	p.K428RfsX88	.	.	.	.	.	.	.	.	.	.	.	.	Plays a role in protein sorting and trafficking between

Gene	Variants	SIFT	Polyphen2	LRT	Mutation Taster	Mutation Assessor	FATHMM	Radial SVM	LR	CADD	GERP++	phyloP	HSF	Gene Function
														the trans-Golgi network and endosomes. Mediates the ARF-dependent recruitment of clathrin to the trans-Golgi network and binds ubiquitinated proteins and membrane cargo molecules with a cytosolic acidic cluster-dileucine motif.
HSF2	p.V81A	D	P	D	D	L	D	D	D	18.24	5.19	7.993	.	DNA-binding protein that specifically binds heat shock promoter elements and activates transcription.
IPO8	p.L983V	D	D	D	D	M	T	D	D	21.9	4.96	4.511	.	Seems to function in nuclear protein import. either by acting as autonomous nuclear transport receptor or as an adapter-like protein in association with the importin-beta subunit KPNB1.
MAN2B1	p.Q691E	D	D	D	D	M	T	D	D	27.5	4.91	7.064	.	Mannosidase necessary for the catabolism of N-linked carbohydrates released during glycoprotein turnover. Cleaves all known types of alpha-mannosidic linkages.
MAN2C1	c.1314+1G>C	.	.	.	D	.	.	.	.	19.80	5.54	7.185	D	Mannosidase.
MUSK	p.G661D	D	D	D	D	M	D	D	D	21.9	5.16	9.813	.	Muscle-specific tyrosine kinase receptor. May play a role in clustering of the acetylcholine receptor in the postsynaptic neuromuscular junction.
PNLIP	c.202-4A>G	.	.	.	.	.	.	.	.	.	.	.	A	Carboxyl esterase that hydrolyzes insoluble. emulsified triglycerides. and is essential for the efficient digestion of dietary fats.
RREB1	p.E1346_G1377 dup	.	.	.	.	.	.	.	.	.	.	.	.	Transcription factor that binds to RAS-responsive elements of gene promoters. May be involved in Ras/Raf-mediated cell differentiation by enhancing calcitonin expression. Represses the angiotensinogen gene. Negatively regulates the transcriptional activity of AR. Potentiates the transcriptional activity of NEUROD1.
WNK4	p.E107RfsX38	.	.	.	.	.	.	.	.	.	.	.	.	Serine-threonine protein kinase that plays an important role in the regulation of electrolyte homeostasis. cell signaling. survival and proliferation.

Table S3.27A – Genes, variants and variant frequencies for the analysis of the dominant inheritance model for sample 36. All variants have a European MAF inferior to 0.01 in at least in one database. Chr: chromosome, Ref: reference nucleotide, Alt: altered nucleotide, 1000G: 1000 Genomes Project, ExAC: Exome Aggregation Consortium, ESP: NHLBI Exome Sequencing Project, EUR: European, AFR: African, AMR: American, EAS: East Asian, SAS: South Asian, NFE: Non-Finnish European, FIN: Finnish European, OTH: Other populations, EA: European American, AA: African American. The last column indicates the gene function. Dot: no information on databases.

Gene	Chr	Start	Ref	Alt	Variant	1000G EUR	1000G AFR	1000G AMR	1000G EAS	1000G SAS	ExAC NFE	ExAC FIN	ExAC AFR	ExAC AMR	ExAC EAS	ExAC SAS	ExAC OTH	ESP6500 EA	ESP6500 AA
ABL2	chr1	179102513	T	C	c.158-4A>G	.	.	.	.	.	.	.	.	.	.	.	.	.	.
CCDC85C	chr14	99981624	G	A	p.R407W	.	.	.	.	.	.	.	.	.	.	.	.	.	.
CHD9	chr16	53276752	G	A	c.G2878A; p.G960R	.	.	.	.	.	.	.	.	.	.	.	.	.	.
DEF8	chr16	90023963	C	G	p.I150M	.	.	.	.	.	.	.	.	.	.	.	.	.	.
EP400	chr12	132490780	A	T	p.D1020V	.	.	.	.	.	.	.	.	.	.	.	.	.	.
ERN2	chr16	23718081	G	C	p.R161G	.	.	.	.	.	.	.	.	.	.	.	.	.	.
ITPR1	chr3	4808285	G	A	p.M1857I	.	.	.	.	.	.	.	.	.	.	.	.	.	.
KIAA1324	chr1	109714505	G	-	p.G163AfsX15	.	.	.	.	.	.	.	.	.	.	.	.	.	.
LAMC3	chr9	133944366	C	T	p.P940L	.	.	.	.	.	.	.	.	.	.	.	.	.	.
LRBA	chr4	151520198	-	C	p.T2003DfsX4	.	.	.	.	.	.	.	.	.	.	.	.	.	.
MUC12	chr7	100635933	CAAA	-	p.T698CfsX82	.	.	.	.	.	.	.	.	.	.	.	.	.	.
PDYN	chr20	1961085	T	C	p.K217E	.	.	.	.	.	.	.	.	.	.	.	.	.	.
RIMS1	chr6	72962473	G	A	p.R903Q	.	.	.	.	.	.	.	.	.	.	.	.	.	.
SLC25A38	chr3	39437973	T	G	p.Y293D	.	.	.	.	.	.	.	.	.	.	.	.	.	.
SLC41A3	chr3	125727634	G	A	p.Q399X	.	.	.	.	.	.	.	.	.	.	.	.	.	.
TBC1D9B	chr5	179318516	C	T	p.D303N	.	.	.	.	.	.	.	.	.	.	.	.	.	.

Gene	Chr	Start	Ref	Alt	Variant	1000G EUR	1000G AFR	1000G AMR	1000G EAS	1000G SAS	ExAC NFE	ExAC FIN	ExAC AFR	ExAC AMR	ExAC EAS	ExAC SAS	ExAC OTH	ESP6500 EA	ESP6500 AA
UTP3	chr4	71554417	GCG	TAAGTG GGCAGC TAAGT	p.R8LfsX30	.	.	.	.	.	.	.	.	.	.	.	.	.	.
VWDE	chr7	12443341	A	T	p.M1K	.	.	.	.	.	.	.	.	.	.	.	.	.	.

Table S3.27B – Genes, variants and predicted functional impact for the analysis of the dominant inheritance model for sample 36. Each functional impact predictor has a different classification with different cut-offs. SIFT: T – tolerated (>0.05), D – damaging (≤0.05); PolyPhen2: B – benign (<0.452), P – possibly damaging (0.453 and 0.956), D – probably damaging (>0.957); LRT: D – deleterious, N – neutral, U – unknown; MutationTaster: P – polymorphism automatic, N – polymorphism, D – disease causing, A – disease causing automatic; MutationAssessor: N – neutral, L – low, M – medium, H – high; FATHMM: T – tolerated (> -1.5), D – damaging (< -1.5); Radial SVM: T – tolerated, D – damaging; LR: T – tolerated, D – damaging; CADD: ≥15 considered pathogenic; GERP++: ≥2 considered conserved; PhyloP: ≥1; HSF: D – disturbs donor splice site; A – disturbs acceptor splice site; No – predicted to not affect splicing. Dot – no prevision was generated. The last column indicates the gene function.

Gene	Variants	SIFT	Polyphen2	LRT	Mutation Taster	Mutation Assessor	FATHMM	Radial SVM	LR	CADD	GERP++	phyloP	HSF	Gene Function
ABL2	c.158-4A>G	.	.	.	.	.	.	.	.	.	.	.	A	Nonreceptor tyrosine kinase that plays a role in cytoskeletal rearrangements through its C-terminal F-actin- and microtubule-binding sequences.
CCDC85C	p.R407W	D	D	U	D	M	.	T	T	20.4	4.95	2.899	.	May play an important role in cortical development. especially in the maintenance of radial glia.
CHD9	c.G2878A; p.G960R	.	D	D	D	M	D	D	D	34	5.53	9.358	A	Transcriptional coactivator for PPARA and possibly other nuclear receptors. Proposed to be a ATP-dependent chromatin remodeling protein. Has DNA-dependent ATPase activity and binds to A/T-rich DNA. Associates with A/T-rich regulatory regions in promoters of genes that participate in the differentiation of progenitors during osteogenesis.
DEF8	p.I150M	D	D	D	D	M	T	T	T	16.68	-7.3	-2.774	.	Unknown.
EP400	p.D1020V	D	D	D	D	M	D	D	D	11.10	5.69	8.394	.	Component of the NuA4 histone acetyltransferase complex which is involved in transcriptional activation of select genes principally by acetylation of nucleosomal histones H4 and H2A.

Gene	Variants	SIFT	Polyphen2	LRT	Mutation Taster	Mutation Assessor	FATHMM	Radial SVM	LR	CADD	GERP++	phyloP	HSF	Gene Function
ERN2	p.R161G	D	D	D	D	M	T	T	T	19.35	5.6	6.845	.	Induces translational repression through 28S ribosomal RNA cleavage in response to ER stress. Pro-apoptotic. Appears to play no role in the unfolded-protein response. unlike closely related proteins.
ITPR1	p.M1857I	T	D	D	D	M	D	D	D	32	5.76	9.740	.	Intracellular channel that mediates calcium release from the endoplasmic reticulum following stimulation by inositol 1.4.5-trisphosphate. Involved in the regulation of epithelial secretion of electrolytes and fluid through the interaction with AHCYL1. Plays a role in ER stress-induced apoptosis.
KIAA1324	p.G163AfsX15	.	.	.	.	.	.	.	.	.	.	.	.	Transmembrane protein that may protect cells from cell death by inducing cytosolic vacuolization and upregulating the autophagy pathway.
LAMC3	p.P940L	D	P	D	D	M	T	T	T	28.1	5.23	7.430	.	Gamma 3 chain isoform laminin that binding to cells via a high affinity receptor. laminin is thought to mediate the attachment. migration and organization of cells into tissues during embryonic development by interacting with other extracellular matrix components.
LRBA	p.T2003DfsX4	.	.	.	.	.	.	.	.	.	.	.	.	May be involved in coupling signal transduction and vesicle trafficking to enable polarized secretion and/or membrane deposition of immune effector molecules.
MUC12	p.T698CfsX82	.	.	.	.	.	.	.	.	.	.	.	.	Mucin involved in epithelial cell protection. adhesion modulation. and signaling. May be involved in epithelial cell growth regulation. Stimulated by both cytokine TNF-alpha and TGF-beta in intestinal epithelium.
PDYN	p.K217E	T	D	D	D	M	T	D	D	26.5	5.0	6.523	.	preproprotein that is proteolytically processed to form the secreted opioid peptides beta-neoendorphin. dynorphin. leu-enkephalin. rimorphin. and leumorphin. These peptides are ligands for the kappa-type of opioid receptor.
RIMS1	p.R903Q	D	D	N	D	M	T	T	T	32	6.02	6.281	.	Member of RAS superfamily member that regulates synaptic vesicle exocytosis. Also plays a role in the regulation of voltage-gated calcium channels during

Gene	Variants	SIFT	Polyphen2	LRT	Mutation Taster	Mutation Assessor	FATHMM	Radial SVM	LR	CADD	GERP++	phyloP	HSF	Gene Function
														neurotransmitter and insulin release.
SLC25A38	p.Y293D	D	D	D	D	H	D	D	D	28.1	5.85	7.776	.	Mitochondrial carrier required during erythropoiesis. Probably involved in the biosynthesis of heme. possibly by facilitating 5-aminolevulinate production. May act by importing glycine into mitochondria or by exchanging glycine for ALA across the mitochondrial inner membrane.
SLC41A3	p.Q399X	T	.	N	A	.	.	.	.	27.5	2.87	5.297	.	Unknown.
TBC1D9B	p.D303N	T	D	D	D	M	D	D	D	24.0	4.96	5.726	.	May act as a GTPase-activating protein for Rab family protein(s).
UTP3	p.R8LfsX30	.	.	.	.	.	.	.	.	.	.	.	.	Essential for gene silencing: has a role in the structure of silenced chromatin. Plays a role in the developing brain.
VWDE	p.M1K	D	P	N	N	.	D	T	D	19.64	3.38	3.102	.	Unknown.

Table S3.28A – Genes, variants and variant frequencies for the analysis of the dominant inheritance model for sample 41. All variants have a European MAF inferior to 0.01 in at least in one database. Chr: chromosome, Ref: reference nucleotide, Alt: altered nucleotide, 1000G: 1000 Genomes Project, ExAC: Exome Aggregation Consortium, ESP: NHLBI Exome Sequencing Project, EUR: European, AFR: African, AMR: American, EAS: East Asian, SAS: South Asian, NFE: Non-Finnish European, FIN: Finnish European, OTH: Other populations, EA: European American, AA: African American. Dot: no information on databases.

Gene	Chr	Start	Ref	Alt	Variant	1000G EUR	1000G AFR	1000G AMR	1000G EAS	1000G SAS	ExAC NFE	ExAC FIN	ExAC AFR	ExAC AMR	ExAC EAS	ExAC SAS	ExAC OTH	ESP6500 EA	ESP6500 AA
ACACA	chr17	35615193	C	T	p.V498I	.	.	.	.	.	.	.	.	.	.	.	.	.	.
ANPEP	chr15	90348451	G	T	c.758-3C>A	.	.	.	.	.	.	.	.	.	.	.	.	.	.
CCDC174	chr3	14711584	C	T	p.R366C	.	.	.	.	.	.	.	.	.	.	.	.	.	.
CFD	chr19	860905	C	T	p.S86L	.	.	.	.	.	.	.	.	.	.	.	.	.	.
CNPY4	chr7	99722428	GAG GAG	-	p.E222_E223del	.	.	.	.	.	.	.	.	.	.	.	.	.	.
DMKN	chr19	36002382	GCCA CTGC TGCC	-	p.G280_G283del	.	.	.	.	.	.	.	.	.	.	.	.	.	.
DNER	chr2	230231668	G	T	p.P675T	.	.	.	.	.	.	.	.	.	.	.	.	.	.
DSC2	chr18	28654723	C	T	p.G605D	.	.	.	.	.	.	.	.	.	.	.	.	.	.
FSD1	chr19	4306334	C	T	c.243+8C>T	.	.	.	.	.	.	.	.	.	.	.	.	.	.
HTATIP2	chr11	20385386	-	GGCG GCGG CTGC TCT	p.A10delinsAAAA AL	.	.	.	.	.	.	.	.	.	.	.	.	.	.
HYDIN	chr16	71007312	G	-	p.I1772FfsX15	.	.	.	.	.	.	.	.	.	.	.	.	.	.
INPPL1	chr11	71943381	G	C	c.1712+1G>C	.	.	.	.	.	.	.	.	.	.	.	.	.	.
ITIH5	chr10	7683909	A	C	p.F94V	.	.	.	.	.	.	.	.	.	.	.	.	.	.
KCNE3	chr11	74168608	T	C	p.M1V	.	.	.	.	.	.	.	.	.	.	.	.	.	.

Gene	Chr	Start	Ref	Alt	Variant	1000G EUR	1000G AFR	1000G AMR	1000G EAS	1000G SAS	ExAC NFE	ExAC FIN	ExAC AFR	ExAC AMR	ExAC EAS	ExAC SAS	ExAC OTH	ESP6500 EA	ESP6500 AA
KCNU1	chr8	36694408	T	G	p.L488W	.	.	.	.	.	.	.	.	.	.	.	.	.	.
KMT2C	chr7	151859758	G	A	p.P3635L	.	.	.	.	.	.	.	.	.	.	.	.	.	.
KRT76	chr12	53162773	-	CTG	p.S547delinsSS	.	.	.	.	.	.	.	.	.	.	.	.	.	.
NSMCE1	chr16	27237077	A	G	p.C231R	.	.	.	.	.	.	.	.	.	.	.	.	.	.
P3H1	chr1	43218335	C	T	c.G1346A; p.G449D	.	.	.	.	.	.	.	.	.	.	.	.	.	.
PCSK4	chr19	1483649	G	A	c.C1391T; p.T464I	.	.	.	.	.	.	.	.	.	.	.	.	.	.
PLAGL1	chr6	144263733	T	C	p.K74E	.	.	.	.	.	.	.	.	.	.	.	.	.	.
RNFT2	chr12	117271666	T	A	p.Y318N	.	.	.	.	.	.	.	.	.	.	.	.	.	.
RTKN	chr2	74657644	G	C	c.C375G; p.D125E	.	.	.	.	.	.	.	.	.	.	.	.	.	.
SDK1	chr7	4305003	C	A	p.S2210Y	.	.	.	.	.	.	.	.	.	.	.	.	.	.
SETD1B	chr12	122255505	G	A	c.G3129A	.	.	.	.	.	.	.	.	.	.	.	.	.	.
SLC13A5	chr17	6594260	C	T	c.1276-1>G>A	.	.	.	.	.	.	.	.	.	.	.	.	.	.
SPG11	chr15	44855492	GTT	-	p.Q2387del	.	.	.	.	.	.	.	.	.	.	.	.	.	.
STOX2	chr4	184930799	T	A	p.L270I	.	.	.	.	.	.	.	.	.	.	.	.	.	.
TMC1	chr9	75404103	G	A	p.R365K	.	.	.	.	.	.	.	.	.	.	.	.	.	.
ZNF493	chr19	21607611	A	-	p.E589DfsX17	.	.	.	.	.	.	.	.	.	.	.	.	.	.

Table S3.28B – Genes, variants and predicted functional impact for the analysis of the dominant inheritance model for sample 41. Each functional impact predictor has a different classification with different cut-offs. SIFT: T – tolerated (>0.05), D – damaging (≤0.05); PolyPhen2: B – benign (<0.452), P – possibly damaging (0.453 and 0.956), D – probably damaging (>0.957); LRT: D – deleterious, N – neutral, U – unknown; MutationTaster: P – polymorphism automatic, N – polymorphism, D – disease causing, A – disease causing automatic; MutationAssessor: N – neutral, L – low, M – medium, H – high; FATHMM: T – tolerated (> -1.5), D – damaging (< -1.5); Radial SVM: T – tolerated, D – damaging; LR: T – tolerated, D – damaging; CADD: ≥15 considered pathogenic; GERP++: ≥2 considered conserved; PhyloP: ≥1; HSF: D – disturbs donor splice site; A – disturbs acceptor splice site; No – predicted to not affect splicing. Dot – no prevision was generated. The last column indicates the gene function.

Gene	Variants	SIFT	Polyphen2	LRT	Mutation Taster	Mutation Assessor	FATHMM	Radial SVM	LR	CADD	GERP++	phyloP	HSF	Gene Function
ACACA	p.V498I	T	B	D	D	N	D	D	D	14.28	4.93	1.443	.	Acetyl-CoA carboxylase that catalyzes the carboxylation of acetyl-CoA to malonyl-CoA. the rate-limiting step in fatty acid synthesis. Carries out three functions: biotin carboxyl carrier protein. biotin carboxylase and carboxyltransferase.
ANPEP	c.758-3C>A	.	.	.	.	.	.	.	.	.	.	.	D	Broad specificity aminopeptidase. Plays a role in the final digestion of peptides generated from hydrolysis of proteins by gastric and pancreatic proteases.
CCDC174	p.R366C	D	D	D	D	M	T	T	T	19.82	0.936	2.029	.	Interacts with eukaryotic translation initiation factor 4A. isoform 3. Appears to be a part of the exon junction complex. which is involved in RNA processing. translation. and nonsense-mediated mRNA decay.
CFD	p.S86L	T	P	U	D	L	D	D	D	16.98	2.59	0.192	.	Serine peptidase that cleaves factor B when the latter is complexed with factor C3b. activating the C3bbb complex. which then becomes the C3 convertase of the alternate pathway.
CNPY4	p.E222_E223del	.	.	.	.	.	.	.	.	.	.	.	.	Plays a role in the regulation of the cell surface expression of TLR4.
DMKN	p.G280_G283del	.	.	.	.	.	.	.	.	.	.	.	.	May act as a soluble regulator of keratinocyte differentiation.
DNER	p.P675T	T	D	D	D	N	D	D	D	15.28	5.93	9.357	.	Transmembranar activator of the NOTCH1 pathway. May mediate neuron-glia interaction during astrocytogenesis.
DSC2	p.G605D	D	D	N	D	M	T	D	T	14.68	5.27	2.201	.	Component of intercellular desmosome junctions. Involved in the interaction of plaque proteins and intermediate filaments mediating cell-cell adhesion. May contribute to epidermal cell positioning (stratification) by mediating differential adhesiveness

Gene	Variants	SIFT	Polyphen2	LRT	Mutation Taster	Mutation Assessor	FATHMM	Radial SVM	LR	CADD	GERP++	phyloP	HSF	Gene Function
														between cells that express different isoforms.
FSD1	c.243+8C>T	.	.	.	.	.	.	.	.	.	.	.	D	May be involved in microtubule organization and stabilization.
HTATIP2	p.A10delinsAAAAAL	.	.	.	.	.	.	.	.	.	.	.	.	Oxidoreductase required for tumor suppression. NAPDH-bound form inhibits nuclear import by competing with nuclear import substrates for binding to a subset of nuclear transport receptors. May act as a redox sensor linked to transcription through regulation of nuclear import. Isoform 1 is a metastasis suppressor with proapoptotic as well as antiangiogenic properties. Isoform 2 has an antiapoptotic effect.
HYDIN	p.I1772FfsX15	.	.	.	.	.	.	.	.	.	.	.	.	Required for ciliary motility.
INPPL1	c.1712+1G>C	.	.	.	D	.	.	.	.	19.16	5.3	6.013	D	SH2-containing 5'-inositol phosphatase that is involved in the regulation of insulin function. Plays a role in the regulation of epidermal growth factor receptor turnover and actin remodelling.
ITIH5	p.F94V	D	D	D	D	.	T	T	T	16.86	5.71	8.832	.	Heavy chain component of one of the inter-alpha-trypsin inhibitor family members that are involved in extracellular matrix stabilization and in the prevention of tumor metastasis.
KCNE3	p.M1V	D	P	D	D	.	D	D	D	17.66	4.93	4.828	.	Member of the potassium channel. voltage-gated. isk-related subfamily. Assembles with a potassium channel alpha-subunit to modulate the gating kinetics and enhance stability of the multimeric complex.
KCNU1	p.L488W	D	D	U	D	M	T	T	T	20.2	2.88	5.789	.	Testis-specific potassium channel activated by both intracellular pH and membrane voltage that mediates export of potassium. Critical for fertility. May play an important role in sperm osmoregulation required for the acquisition of normal morphology and motility when faced with osmotic challenges. such as those experienced after mixing with seminal fluid and entry into the vagina.

Gene	Variants	SIFT	Polyphen2	LRT	Mutation Taster	Mutation Assessor	FATHMM	Radial SVM	LR	CADD	GERP++	phyloP	HSF	Gene Function
KMT2C	p.P3635L	D	D	D	D	M	D	D	D	13.45	5.51	9.467	.	Histone methyltransferase that methylates Lys-4 of histone H3. H3 Lys-4 methylation represents a specific tag for epigenetic transcriptional activation. Central component of the MLL2/3 complex. a coactivator complex of nuclear receptors. involved in transcriptional coactivation. KMT2C/MLL3 may be a catalytic subunit of this complex.
KRT76	p.S547delinsSS	.	.	.	.	.	.	.	.	.	.	.	.	Keratin probably contributing to terminal cornification.
NSMCE1	p.C231R	D	D	D	D	M	D	D	D	16.47	5.42	7.315	.	Component of the SMC5-SMC6 complex. a complex involved in DNA double-strand breaks by homologous recombination. The complex is required for telomere maintenance. Is involved in positive regulation of response to DNA damage stimulus.
P3H1	c.G1346A; p.G449D	T	B	N	D	L	T	T	T	17.75	4.96	4.096	A	Basement membrane-associated chondroitin sulfate proteoglycan that has prolyl 3-hydroxylase activity catalyzing the post-translational formation of 3-hydroxyproline in collagens. especially types IV and V. May be involved in the secretory pathway of cells.
PCSK4	c.C1391T; p.T464I	.	B	.	N	L	T	T	T	0.767	0.101	-0.208	D	Proprotein convertase involved in the processing of hormone and other protein precursors at sites comprised of pairs of basic amino acid residues. It plays a critical role in fertilization. fetoplacental growth. and embryonic development and processes multiple prohormones including pro-pituitary adenylate cyclase-activating protein and pro-insulin-like growth factor II.
PLAGL1	p.K74E	D	P	D	D	M	T	T	T	21.0	5.01	7.977	.	C2H2 zinc finger protein that functions as a suppressor of cell growth.
RNFT2	p.Y318N	D	D	D	D	M	T	T	T	26.0	4.98	7.655	.	Unknown.
RTKN	c.C375G; p.D125E	T	B	N	D	M	T	T	T	16.11	0.989	0.859	D	Scaffold protein that mediates Rho signaling to activate NF-kappa-B and may confer increased resistance to apoptosis to cells in gastric tumorigenesis. May play a novel role in the

Gene	Variants	SIFT	Polyphen2	LRT	Mutation Taster	Mutation Assessor	FATHMM	Radial SVM	LR	CADD	GERP++	phyloP	HSF	Gene Function
														organization of septin structures.
SDK1	p.S2210Y	D	D	D	D	L	T	D	D	23.5	4.49	6.553	.	Adhesion molecule that promotes lamina-specific synaptic connections in the retina. Expressed in specific subsets of interneurons and retinal ganglion cells and promotes synaptic connectivity via homophilic interactions.
SETD1B	c.G3129A	.	.	.	.	.	.	.	.	.	.	.	A	Histone methyltransferase that specifically methylates Lys-4 of histone H3. when part of the SET1 histone methyltransferase complex. but not if the neighboring Lys-9 residue is already methylated. H3 Lys-4 methylation represents a specific tag for epigenetic transcriptional activation.
SLC13A5	c.1276-1>G>A	.	.	.	D	.	.	.	.	21.5	5.07	7.066	A	High-affinity sodium/citrate cotransporter that mediates citrate entry into cells. May facilitate the utilization of circulating citrate for the generation of metabolic energy and for the synthesis of fatty acids and cholesterol.
SPG11	p.Q2387del	.	.	.	.	.	.	.	.	.	.	.	.	Potential transmembrane protein that is phosphorylated upon DNA damage.
STOX2	p.L270I	D	D	D	D	M	T	D	D	16.82	2.43	2.712	.	Unknown.
TMC1	p.R365K	D	D	D	D	M	D	D	D	36	6.17	9.835	.	Probable ion channel required for the normal function of cochlear hair cells.
ZNF493	p.E589DfsX17	.	.	.	.	.	.	.	.	.	.	.	.	May be involved in transcriptional regulation.

Table S3.29A – Genes, variants and variant frequencies for the analysis of the dominant inheritance model for sample 42. All variants have a European MAF inferior to 0.01 in at least in one database. Chr: chromosome, Ref: reference nucleotide, Alt: altered nucleotide, 1000G: 1000 Genomes Project, ExAC: Exome Aggregation Consortium, ESP: NHLBI Exome Sequencing Project, EUR: European, AFR: African, AMR: American, EAS: East Asian, SAS: South Asian, NFE: Non-Finnish European, FIN: Finnish European, OTH: Other populations, EA: European American, AA: African American. Dot: no information on databases.

Gene	Chr	Start	Ref	Alt	Variant	1000G EUR	1000G AFR	1000G AMR	1000G EAS	1000G SAS	ExAC NFE	ExAC FIN	ExAC AFR	ExAC AMR	ExAC EAS	ExAC SAS	ExAC OTH	ESP6500 EA	ESP6500 AA
AKAP13	chr15	86198720	C	T	p.R1483X	.	.	.	.	.	.	.	.	.	.	.	.	.	.
BBX	chr3	107493689	A	T	p.K344X	.	.	.	.	.	.	.	.	.	.	.	.	.	.
BRD7	chr16	50373999	T	C	c.592-2A>G	.	.	.	.	.	.	.	.	.	.	.	.	.	.
CCDC105	chr19	15133854	GAGCA	-	p.E475GfsX6	.	.	.	.	.	.	.	.	.	.	.	.	.	.
CCDC141	chr2	179749709	TTC	-	p.E546del	.	.	.	.	.	.	.	.	.	.	.	.	.	.
DCTPP1	chr16	30440417	AGG	-	p.L58del	.	.	.	.	.	.	.	.	.	.	.	.	.	.
DRP2	chrX	100505999	G	A	p.V598I	.	.	.	.	.	.	.	.	.	.	.	.	.	.
DUSP7	chr3	52087992	G	A	p.L306F	.	.	.	.	.	.	.	.	.	.	.	.	.	.
EIF1	chr17	39846439	G	A	c.G295A:p.E99K	.	.	.	.	.	.	.	.	.	.	.	.	.	.
KIR3DL3	chr19	55246720	G	C	c.950G>C	.	.	.	.	.	.	.	.	.	.	.	.	.	.
MTX1	chr1	155182207	A	G	p.Y268C	.	.	.	.	.	.	.	.	.	.	.	.	.	.
NT5DC3	chr12	104190672	C	A	c.G753T	.	.	.	.	.	.	.	.	.	.	.	.	.	.
PIEZO1	chr16	88798771	T	G	p.G988A	.	.	.	.	.	.	.	.	.	.	.	.	.	.
PLXNB3	chrX	153036422	C	T	c.C2139T	.	.	.	.	.	.	.	.	.	.	.	.	.	.
PRM3	chr16	11367317	AGG	-	p.S45del	.	.	.	.	.	.	.	.	.	.	.	.	.	.
SEC31A	chr4	83791516	C	T	p.D282N	.	.	.	.	.	.	.	.	.	.	.	.	.	.

Gene	Chr	Start	Ref	Alt	Variant	1000G EUR	1000G AFR	1000G AMR	1000G EAS	1000G SAS	ExAC NFE	ExAC FIN	ExAC AFR	ExAC AMR	ExAC EAS	ExAC SAS	ExAC OTH	ESP6500 EA	ESP6500 AA
SLC16A8	chr22	38478711	G	T	p.T57K	.	.	.	.	.	.	.	.	.	.	.	.	.	.
SLC17A5	chr6	74304834	C	T	p.W485X	.	.	.	.	.	.	.	.	.	.	.	.	.	.
VPS51	chr11	64878914	T	C	p.L735P	.	.	.	.	.	.	.	.	.	.	.	.	.	.

Table S3.29B – Genes, variants and predicted functional impact for the analysis of the dominant inheritance model for sample 42. Each functional impact predictor has a different classification with different cut-offs. SIFT: T – tolerated (>0.05), D – damaging (≤0.05); PolyPhen2: B – benign (<0.452), P – possibly damaging (0.453 and 0.956), D – probably damaging (>0.957); LRT: D – deleterious, N – neutral, U – unknown; MutationTaster: P – polymorphism automatic, N – polymorphism, D – disease causing, A – disease causing automatic; MutationAssessor: N – neutral, L – low, M – medium, H – high; FATHMM: T – tolerated (> -1.5), D – damaging (< -1.5); Radial SVM: T – tolerated, D – damaging; LR: T – tolerated, D – damaging; CADD: ≥15 considered pathogenic; GERP++: ≥2 considered conserved; PhyloP: ≥1; HSF: D – disturbs donor splice site; A – disturbs acceptor splice site; No – predicted to not affect splicing. Dot – no prevision was generated. The last column indicates the gene function.

Gene	Variants	SIFT	Polyphen2	LRT	Mutation Taster	Mutation Assessor	FATHMM	Radial SVM	LR	CADD	GERP++	phyloP	HSF	Gene Function
AKAP13	p.R1483X	T	.	.	D	.	.	.	.	27.2	5.9	1.804	.	Anchors PKA and acts as an adapter protein to selectively couple G alpha-13 and Rho. Activates estrogen receptor beta by a p38 MAPK-dependent pathway. Stimulates exchange activity on Rho proteins *in vitro*. but not on CDC42. Ras or Rac and may bind calcium ions.
BBX	p.K344X	D	D	D	D	L	T	T	T	39	5.51	8.820	.	Transcription factor that is necessary for cell cycle progression from G1 to S phase.
BRD7	c.592-2A>G	.	.	.	D	.	.	.	.	18.68	5.31	8.040	A	Component of one form of the SWI/SNF chromatin remodeling complex. Acts both as coactivator and as corepressor. May play a role in chromatin remodeling.
CCDC105	p.E475GfsX6	.	.	.	.	.	.	.	.	.	.	.	.	Unknown.
CCDC141	p.E546del	.	.	.	.	.	.	.	.	.	.	.	.	Plays a critical role in radial migration and centrosomal function.
DCTPP1	p.L58del	.	.	.	.	.	.	.	.	.	.	.	.	Pyrophosphatase. which converts dCTP to dCMP and inorganic pyrophosphate. Displays weak activity against

Gene	Variants	SIFT	Polyphen2	LRT	Mutation Taster	Mutation Assessor	FATHMM	Radial SVM	LR	CADD	GERP++	phyloP	HSF	Gene Function
														dTTP and dATP. but none against dGTP. This protein may be responsible for eliminating excess dCTP after DNA synthesis and may prevent overmethylation of CpG islands.
DRP2	p.V598I	D	D	D	D	L	T	T	D	28.1	6.06	7.904	.	Required for normal myelination and for normal organization of the cytoplasm and the formation of Cajal bands in myelinating Schwann cells. Required for normal PRX location at appositions between the abaxonal surface of the myelin sheath and the Schwann cell plasma membrane. Possibly involved in membrane-cytoskeleton interactions of the central nervous system.
DUSP7	p.L306F	D	D	N	D	M	D	D	D	29.0	4.97	6.692	.	Regulates the activity of the MAP kinase family in response to changes in the cellular environment.
EIF1	c.G295A: p.E99K	T	B	D	D	N	T	T	T	21.6	5.55	9.623	A	Necessary for scanning and involved in initiation site selection. Promotes the assembly of 48S ribosomal complexes at the authentic initiation codon of a conventional capped mRNA.
KIR3DL3	c.950G>C	.	B	.	N	M	T	T	T	4.518	-0.211	0.301	A	Receptor on natural killer cells that may inhibit the activity of natural killer cells thus preventing cell lysis.
MTX1	p.Y268C	D	D	D	D	M	T	T	T	17.35	4.38	8.643	.	Involved in transport of proteins into the mitochondrion. Essential for embryonic development.
NT5DC3	c.G753T	T	B	D	D	L	T	T	T	17.81	5.6	7.422	D	Unknown.
PIEZO1	p.G988A	D	D	D	D	M	T	T	T	22.3	3.87	7.166	.	Pore-forming subunit of a mechanosensitive non-specific cation channel. Generates currents characterized by a linear current-voltage relationship that are sensitive to ruthenium red and gadolinium. Plays a key role in epithelial cell adhesion. Acts as shear-stress sensor that promotes endothelial cell organization and alignment in the direction of blood flow through calpain activation. Plays a key role in blood vessel formation and vascular structure in both development and adult physiology.
PLXNB3	c.C2139T	.	B	.	D	.	T	T	T	8.556	3.5	-0.124	A	Receptor for semaphorin 5A. and plays a role in axon guidance. invasive growth and cell migration.

Gene	Variants	SIFT	Polyphen2	LRT	Mutation Taster	Mutation Assessor	FATHMM	Radial SVM	LR	CADD	GERP++	phyloP	HSF	Gene Function
PRM3	p.S45del	.	.	.	.	.	.	.	.	.	.	.	.	Protamine that substitutes for histones in the chromatin of sperm during the haploid phase of spermatogenesis. Compacts sperm DNA into a highly condensed. stable and inactive complex.
SEC31A	p.D282N	D	D	D	D	M	T	D	D	36	5.38	7.747	.	Component of the coat protein complex II that promotes the formation of transport vesicles from the endoplasmic reticulum. The coat has two main functions. the physical deformation of the endoplasmic reticulum membrane into vesicles and the selection of cargo molecules.
SLC16A8	p.T57K	D	P	U	D	M	T	D	T	34	4.23	7.774	.	Proton-linked monocarboxylate transporter. Catalyzes the rapid transport across the plasma membrane of many monocarboxylates such as lactate. pyruvate. branched-chain oxo acids derived from leucine. valine and isoleucine. and the ketone bodies acetoacetate. beta-hydroxybutyrate and acetate.
SLC17A5	p.W485X	T	.	D	D	.	.	.	.	35	5.3	6.419	.	Membrane transporter that exports glucuronic acid and free sialic acid out of the lysosome after it is cleaved from sialoglycoconjugates undergoing degradation. this is required for normal central nervous system myelination. Mediates aspartate and glutamate membrane potential-dependent uptake into synaptic vesicles and synaptic-like microvesicles. Also functions as an electrogenic 2NO(3)(-)/H(+) cotransporter in the plasma membrane of salivary gland acinar cells. mediating the physiological nitrate efflux. 25% of the circulating nitrate ions is typically removed and secreted in saliva.
VPS51	p.L735P	D	D	D	D	M	.	D	D	25.9	5.05	7.535	.	Component of the GARP complex that participates in retrograde transport of acid hydrolase receptors. likely by promoting tethering and SNARE-dependent fusion of endosome-derived carriers to the trans-Golgi network. Acts as component of the EARP complex that is involved in endocytic recycling.

Table S3.30A – Genes, variants and variant frequencies for the analysis of the dominant inheritance model for sample 43. All variants have a European MAF inferior to 0.01 in at least in one database. Chr: chromosome, Ref: reference nucleotide, Alt: altered nucleotide, 1000G: 1000 Genomes Project, ExAC: Exome Aggregation Consortium, ESP: NHLBI Exome Sequencing Project, EUR: European, AFR: African, AMR: American, EAS: East Asian, SAS: South Asian, NFE: Non-Finnish European, FIN: Finnish European, OTH: Other populations, EA: European American, AA: African American. Dot: no information on databases.

Gene	Chr	Start	Ref	Alt	Variant	1000G EUR	1000G AFR	1000G AMR	1000G EAS	1000G SAS	ExAC NFE	ExAC FIN	ExAC AFR	ExAC AMR	ExAC EAS	ExAC SAS	ExAC OTH	ESP6500 EA	ESP6500 AA
ANO4	chr12	101433784	C	T	p.R317X	.	.	.	.	.	.	.	.	.	.	.	.	.	.
C1orf95	chr1	226736634	CGGC GGCG GCGG CGGC GGCG GTGG	-	p.A11_A18del	.	.	.	.	.	.	.	.	.	.	.	.	.	.
COL14A1	chr8	121290375	A	G	p.D1080G	.	.	.	.	.	.	.	.	.	.	.	.	.	.
CUL7	chr6	43020454	G	C	p.L25V	.	.	.	.	.	.	.	.	.	.	.	.	.	.
FRG2B	chr10	135438858	C	T	p.W194X	.	.	.	.	.	.	.	.	.	.	.	.	.	.
GID4	chr17	17965216	G	A	p.G256R	.	.	.	.	.	.	.	.	.	.	.	.	.	.
KCTD20	chr6	36454727	T	G	p.Y345X	.	.	.	.	.	.	.	.	.	.	.	.	.	.
NCBP3	chr17	3721776	T	C	p.D364G	.	.	.	.	.	.	.	.	.	.	.	.	.	.
NPHP4	chr1	5927110	T	A	p.I1180F	.	.	.	.	.	.	.	.	.	.	.	.	.	.
OPRK1	chr8	54142044	C	T	p.C319Y	.	.	.	.	.	.	.	.	.	.	.	.	.	.
RASD1	chr17	17398585	-	GCCG CT	p.G234delins SGG	.	.	.	.	.	.	.	.	.	.	.	.	.	.
ROBO1	chr3	78717167	G	T	p.A611E	.	.	.	.	.	.	.	.	.	.	.	.	.	.
SLC4A3	chr2	220504412	G	A	p.V1078M	.	.	.	.	.	.	.	.	.	.	.	.	.	.
TLR6	chr4	38828784	C	T	p.G771R	.	.	.	.	.	.	.	.	.	.	.	.	.	.

Gene	Chr	Start	Ref	Alt	Variant	1000G EUR	1000G AFR	1000G AMR	1000G EAS	1000G SAS	ExAC NFE	ExAC FIN	ExAC AFR	ExAC AMR	ExAC EAS	ExAC SAS	ExAC OTH	ESP6500 EA	ESP6500 AA
TMEM249	chr8	145577451	G	A	p.S172L	.	.	.	.	.	.	.	.	.	.	.	.	.	.
TTN	chr2	179499333	G	T	p.P12418T	.	.	.	.	.	.	.	.	.	.	.	.	.	.
UTP4	chr16	69170592	T	A	c.160-7T>A	.	.	.	.	.	.	.	.	.	.	.	.	.	.
VEZF1	chr17	56056604	-	TGTTGC	p.Q349delinsQQQ	.	.	.	.	.	.	.	.	.	.	.	.	.	.

Table S3.30B – Genes, variants and predicted functional impact for the analysis of the dominant inheritance model for sample 43. Each functional impact predictor has a different classification with different cut-offs. SIFT: T – tolerated (>0.05), D – damaging (≤0.05); PolyPhen2: B – benign (<0.452), P – possibly damaging (0.453 and 0.956), D – probably damaging (>0.957); LRT: D – deleterious, N – neutral, U – unknown; MutationTaster: P – polymorphism automatic, N – polymorphism, D – disease causing, A – disease causing automatic; MutationAssessor: N – neutral, L – low, M – medium, H – high; FATHMM: T – tolerated (> -1.5), D – damaging (< -1.5); Radial SVM: T – tolerated, D – damaging; LR: T – tolerated, D – damaging; CADD: ≥15 considered pathogenic; GERP++: ≥2 considered conserved; PhyloP: ≥1; HSF: D – disturbs donor splice site; A – disturbs acceptor splice site; No – predicted to not affect splicing. Dot – no prevision was generated. The last column indicates the gene function.

Gene	Variants	SIFT	Polyphen2	LRT	Mutation Taster	Mutation Assessor	FATHMM	Radial SVM	LR	CADD	GERP++	phyloP	HSF	Gene Function
ANO4	p.R317X	T	.	D	D	.	.	.	.	40	3.8	3.243	.	Transmembrane protein that has calcium-dependent phospholipid scramblase activity; scrambles phosphatidylserine. phosphatidylcholine and galactosylceramide. Does not exhibit calcium-activated chloride channel CaCC activity.
C1orf95	p.A11_A18del	.	.	.	.	.	.	.	.	.	.	.	.	Unknown.
COL14A1	p.D1080G	T	D	D	D	M	D	D	D	27.1	5.6	9.309	.	Plays an adhesive role by integrating collagen bundles. It is probably associated with the surface of interstitial collagen fibrils via COL1. The COL2 domain may then serve as a rigid arm which sticks out from the fibril and protrudes the large N-terminal globular domain into the extracellular space. where it might interact with other matrix molecules or cell surface receptors.

Gene	Variants	SIFT	Polyphen2	LRT	Mutation Taster	Mutation Assessor	FATHMM	Radial SVM	LR	CADD	GERP++	phyloP	HSF	Gene Function
CUL7	p.L25V	D	P	D	D	L	D	T	T	19.31	1.71	1.475	.	Component of an E3 ubiquitin-protein ligase complex that interacts with TP53. CUL9. and FBXW8 proteins.
FRG2B	p.W194X	T	.	N	D	.	.	.	.	14.24	.	0.259	.	Unknown.
GID4	p.G256R	D	D	D	D	H	.	D	D	35	5.36	9.561	.	Subunit of the Mediator complex that is a coactivator required for activation of RNA polymerase II transcription by DNA bound transcription factors.
KCTD20	p.Y345X	T	.	D	D	.	.	.	.	37	-6.15	-1.534	.	Promotes the phosphorylation of AKT family members.
NCBP3	p.D364G	D	D	D	D	L	.	T	T	32	5.85	6.982	.	Associates with NCBP1/CBP80 to form an alternative cap-binding complex CBC which plays a key role in mRNA export. NCBP3 serves as adapter protein linking the capped RNAs to NCBP1/CBP80.
NPHP4	p.I1180F	D	D	D	D	M	T	T	T	16.78	2.27	1.465	.	Involved in the organization of apical junctions in kidney cells together with NPHP1 and RPGRIP1L/NPHP8. Does not seem to be strictly required for ciliogenesis.
OPRK1	p.C319Y	D	D	D	D	H	T	D	T	29.5	5.8	7.818	.	Opioid receptor that functions as a receptor for endogenous ligands. as well as for various synthetic opioids. Ligand binding results in inhibition of adenylate cyclase activity and neurotransmitter release. This opioid receptor plays a role in the perception of pain and mediating the hypolocomotor. analgesic and aversive actions of synthetic opioids.
RASD1	p.G234delinsSGG	.	.	.	.	.	.	.	.	.	.	.	.	Activator of G-protein signaling and acts as a direct nucleotide exchange factor for Gi-Go proteins. This protein interacts with the neuronal nitric oxide adaptor protein CAPON. and a nuclear adaptor protein FE65. which interacts with the Alzheimer's disease amyloid precursor protein. This gene may play a role in dexamethasone-induced alterations in cell morphology. growth and cell-extracellular matrix interactions.
ROBO1	p.A611E	.	D	D	D	M	T	T	T	34	5.84	9.771	.	Receptor for SLIT1 and SLIT2 that mediates cellular responses to molecular guidance cues in cellular

Gene	Variants	SIFT	Polyphen2	LRT	Mutation Taster	Mutation Assessor	FATHMM	Radial SVM	LR	CADD	GERP++	phyloP	HSF	Gene Function
														migration. including axonal navigation at the ventral midline of the neural tube and projection of axons to different regions during neuronal development.
SLC4A3	p.V1078M	D	D	D	D	M	D	D	D	27.5	4.38	9.443	.	Plasma membrane anion exchange protein of wide distribution. Mediates at least a part of the Cl(-)/ HCO3(-) exchange in cardiac myocytes.
TLR6	p.G771R	D	B	N	D	L	D	D	D	15.71	4.94	1.698	.	Toll-like receptor that participates in the innate immune response to Gram-positive bacteria and fungi. Specifically recognizes diacylated and. to a lesser extent. triacylated lipopeptides.
TMEM249	p.S172L	D	D	.	N	.	.	T	T	17.86	3.68	2.893	.	Unknown.
TTN	p.P12418T	D	D	.	D	M	T	T	T	17.01	6.17	9.762	.	Key component in the assembly and functioning of vertebrate striated muscles. By providing connections at the level of individual microfilaments. it contributes to the fine balance of forces between the two halves of the sarcomere. In non-muscle cells. seems to play a role in chromosome condensation and chromosome segregation during mitosis. Might link the lamina network to chromatin or nuclear actin. or both during interphase.
UTP4	c.160-7T>A	.	.	.	.	.	.	.	.	.	.	.	A	May be a transcriptional regulator. Acts as a positive regulator of HIVEP1 which specifically binds to the DNA sequence 5-GGGACTTTCC-3 found in enhancer elements of numerous viral promoters such as those of HIV-1. SV40. or CMV. Ribosome biogenesis factor involved in small subunit (SSU) pre-rRNA processing at sites A. A0. 1 and 2b.
VEZF1	p.Q349delins QQQ	.	.	.	.	.	.	.	.	.	.	.	.	Possible transcription factor that specifically binds to the CT/GC-rich region of the interleukin-3 promoter and mediates tax transactivation of IL-3.

Table S3.31A – Genes, variants and variant frequencies for the analysis of the dominant inheritance model for sample 45. All variants have a European MAF inferior to 0.01 in at least in one database. Chr: chromosome, Ref: reference nucleotide, Alt: altered nucleotide, 1000G: 1000 Genomes Project, ExAC: Exome Aggregation Consortium, ESP: NHLBI Exome Sequencing Project, EUR: European, AFR: African, AMR: American, EAS: East Asian, SAS: South Asian, NFE: Non-Finnish European, FIN: Finnish European, OTH: Other populations, EA: European American, AA: African American. Dot: no information on databases.

Gene	Chr	Start	Ref	Alt	Variant	1000G EUR	1000G AFR	1000G AMR	1000G EAS	1000G SAS	ExAC NFE	ExAC FIN	ExAC AFR	ExAC AMR	ExAC EAS	ExAC SAS	ExAC OTH	ESP6500 EA	ESP6500 AA
ABCC4	chr13	95830305	C	T	p.G529E	.	.	.	.	.	.	.	.	.	.	.	.	.	.
ACAN	chr15	89379445	C	A	p.T3N	.	.	.	.	.	.	.	.	.	.	.	.	.	.
ATN1	chr12	7045892	CAGCAG CAGCAG CAG	-	p.Q488_492Qdel	.	.	.	.	.	.	.	.	.	.	.	.	.	.
CECR2	chr22	18022348	A	G	p.Y676C	.	.	.	.	.	.	.	.	.	.	.	.	.	.
CNTNAP4	chr16	76350369	G	-	p.E24SfsX13	.	.	.	.	.	.	.	.	.	.	.	.	.	.
DFNA5	chr7	24738881	G	C	c.1258-3C>G	.	.	.	.	.	.	.	.	.	.	.	.	.	.
DOCK2	chr5	169127143	G	A	c.G1258A; p.G420R	.	.	.	.	.	.	.	.	.	.	.	.	.	.
DTL	chr1	212276215	T	G	p.I709S	.	.	.	.	.	.	.	.	.	.	.	.	.	.
LMBR1L	chr12	49500756	C	A	p.E49X	.	.	.	.	.	.	.	.	.	.	.	.	.	.
LTBP4	chr19	41119868	T	A	p.C898X	.	.	.	.	.	.	.	.	.	.	.	.	.	.
MICALL1	chr22	38328672	G	C	p.R710P	.	.	.	.	.	.	.	.	.	.	.	.	.	.
PKD1	chr16	2147901	AGCTGT CCAGCA CGGACG AGTCCA GGC	-	p.C3370_S3378del	.	.	.	.	.	.	.	.	.	.	.	.	.	.
PSEN2	chr1	227071592	C	T	p.R110C	.	.	.	.	.	.	.	.	.	.	.	.	.	.
PTPN7	chr1	202121684	G	-	p.I326MfsX2	.	.	.	.	.	.	.	.	.	.	.	.	.	.

Gene	Chr	Start	Ref	Alt	Variant	1000G EUR	1000G AFR	1000G AMR	1000G EAS	1000G SAS	ExAC NFE	ExAC FIN	ExAC AFR	ExAC AMR	ExAC EAS	ExAC SAS	ExAC OTH	ESP6500 EA	ESP6500 AA
QSER1	chr11	32956634	G	T	p.G1148V	.	.	.	.	.	.	.	.	.	.	.	.	.	.
RCVRN	chr17	9808497	T	C	p.M1V	.	.	.	.	.	.	.	.	.	.	.	.	.	.
RFX3	chr9	3330253	A	C	c.474+6T>G	.	.	.	.	.	.	.	.	.	.	.	.	.	.
SAMD9	chr7	92732446	C	T	p.A989T	.	.	.	.	.	.	.	.	.	.	.	.	.	.
STARD7	chr2	96858122	G	C	p.H276Q	.	.	.	.	.	.	.	.	.	.	.	.	.	.
TRAK1	chr3	42251606	GAG	-	p.E698del	.	.	.	.	.	.	.	.	.	.	.	.	.	.
WNT11	chr11	75902701	C	T	p.R266Q	.	.	.	.	.	.	.	.	.	.	.	.	.	.
ZKSCAN8	chr6	28121489	-	A	p.I478NfsX17	.	.	.	.	.	.	.	.	.	.	.	.	.	.

Table S3.31B – Genes, variants and predicted functional impact for the analysis of the dominant inheritance model for sample 45. Each functional impact predictor has a different classification with different cut-offs. SIFT: T – tolerated (>0.05), D – damaging (≤0.05); PolyPhen2: B – benign (<0.452), P – possibly damaging (0.453 and 0.956), D – probably damaging (>0.957); LRT: D – deleterious, N – neutral, U – unknown; MutationTaster: P – polymorphism automatic, N – polymorphism, D – disease causing, A – disease causing automatic; MutationAssessor: N – neutral, L – low, M – medium, H – high; FATHMM: T – tolerated (> -1.5), D – damaging (< -1.5); Radial SVM: T – tolerated, D – damaging; LR: T – tolerated, D – damaging; CADD: ≥15 considered pathogenic; GERP++: ≥2 considered conserved; PhyloP: ≥1; HSF: D – disturbs donor splice site; A – disturbs acceptor splice site; No – predicted to not affect splicing. Dot – no prevision was generated. The last column indicates the gene function.

Gene	Variants	SIFT	Polyphen2	LRT	Mutation Taster	Mutation Assessor	FATHMM	Radial SVM	LR	CADD	GERP++	phyloP	HSF	Gene Function
ABCC4	p.G529E	D	D	D	D	M	D	D	D	32	5.8	7.338	.	Transporter that may be an organic anion pump relevant to cellular detoxification.
ACAN	p.T3N	D	D	.	D	M	T	T	T	17.54	5.3	5.617	.	Proteoglycan component of extracellular matrix of cartilagenous tissues. Resist to compression in cartilage. Binds avidly to hyaluronic acid via an N-terminal globular region.
ATN1	p.Q488_492Qdel	.	.	.	.	.	.	.	.	.	.	.	.	Transcriptional corepressor that recruits NR2E1 to repress transcription. Promotes vascular smooth cell

Gene	Variants	SIFT	Polyphen2	LRT	Mutation Taster	Mutation Assessor	FATHMM	Radial SVM	LR	CADD	GERP++	phyloP	HSF	Gene Function
														migration and orientation. Corepressor of MTG8 transcriptional repression. Has some intrinsic repression activity which is independent of the number of poly-Asn repeats.
CECR2	p.Y676C	D	D	D	D	M	T	T	T	14.73	5.29	3.026	.	Part of the CECR2-containing-remodeling factor complex. which facilitates the perturbation of chromatin structure in an ATP-dependent manner. May be involved through its interaction with LRPPRC in the integration of cytoskeletal network with vesicular trafficking. nucleocytosolic shuttling. transcription. chromosome remodeling and cytokinesis.
CNTNAP4	p.E24SfsX13	.	.	.	.	.	.	.	.	.	.	.	.	Presynaptic protein involved in both dopaminergic synaptic transmission and GABAergic system. thereby participating in the structural maturation of inhibitory interneuron synapses. Involved in the dopaminergic synaptic transmission by attenuating dopamine release through a presynaptic mechanism. Also participates in the GABAergic system.
DFNA5	c.1258-3C>G	.	.	.	.	.	.	.	.	.	.	.	A	Involved in apoptosis and cell survival. Plays a role in the TP53-regulated cellular response to DNA damage probably by cooperating with TP53.
DOCK2	c.G1258A; p.G420R	D	P	D	D	M	T	T	T	35	5.88	7.782	D	Involved in cytoskeletal rearrangements required for lymphocyte migration in response of chemokines. Activates RAC1 and RAC2. but not CDC42. by functioning as a guanine nucleotide exchange factor. which exchanges bound GDP for free GTP. May also participate in IL2 transcriptional activation via the activation of RAC2.
DTL	p.I709S	D	D	D	D	L	T	D	D	23.3	5.72	5.156	.	Substrate-specific adapter of a DDB1-CUL4-X-box E3 ubiquitin-protein ligase complex required for cell cycle control. DNA damage response and translesion DNA synthesis.
LMBR1L	p.E49X	D	.	D	D	.	.	.	.	37	5.25	7.194	.	Probable LCN1 receptor that may mediate LCN1 endocytosis.

Gene	Variants	SIFT	Polyphen2	LRT	Mutation Taster	Mutation Assessor	FATHMM	Radial SVM	LR	CADD	GERP++	phyloP	HSF	Gene Function
LTBP4	p.C898X	T	.	D	D	.	.	.	.	40	-0.082	-0.759	.	Extracellular matrix protein that may be involved in the assembly. secretion and targeting of TGFB1 to sites at which it is stored and/or activated. May play critical roles in controlling and directing the activity of TGFB1. May have a structural role in the extracellular matrix.
MICALL1	p.R710P	D	D	D	D	M	T	D	T	26.8	4.45	8.029	.	Probable lipid-binding protein with higher affinity for phosphatidic acid. a lipid enriched in recycling endosome membranes. On endosome membranes. may act as a downstream effector of Rab proteins recruiting cytosolic proteins to regulate membrane tubulation. May be involved in a late step of receptor-mediated endocytosis regulating for instance endocytosed-EGF receptor trafficking. Alternatively. may regulate slow endocytic recycling of endocytosed proteins back to the plasma membrane. May indirectly play a role in neurite outgrowth.
PKD1	p.C3370_S3378del	.	.	.	.	.	.	.	.	.	.	.	.	Transmembrane glycoprotein that is involved in renal tubulogenesis and involved in fluid-flow mechanosensation by the primary cilium in renal epithelium. May be an ion-channel regulator. Involved in adhesive protein-protein and protein-carbohydrate interactions.
PSEN2	p.R110C	D	P	D	D	L	D	D	D	21.4	4.55	3.474	.	Probable catalytic subunit of the gamma-secretase complex. an endoprotease complex that catalyzes the intramembrane cleavage of integral membrane proteins such as Notch receptors and beta-amyloid precursor protein. Requires the other members of the gamma-secretase complex to have a protease activity. May play a role in intracellular signaling and gene expression or in linking chromatin to the nuclear membrane. May function in the cytoplasmic partitioning of proteins.
PTPN7	p.I326MfsX2	.	.	.	.	.	.	.	.	.	.	.	.	Tyrosine phosphatase that acts preferentially on tyrosine-phosphorylated MAPK1. Plays a role in the

Gene	Variants	SIFT	Polyphen2	LRT	Mutation Taster	Mutation Assessor	FATHMM	Radial SVM	LR	CADD	GERP++	phyloP	HSF	Gene Function
														regulation of T and B-lymphocyte development and signal transduction.
QSER1	p.G1148V	D	D	D	D	M	T	D	D	19.02	5.58	9.476	.	Unknown.
RCVRN	p.M1V	D	D	D	D	.	T	D	D	22.0	4.87	7.923	.	Seems to be implicated in the pathway from retinal rod guanylate cyclase to rhodopsin. May be involved in the inhibition of the phosphorylation of rhodopsin in a calcium-dependent manner. The calcium-bound recoverin prolongs the photoresponse.
RFX3	c.474+6T>G	.	.	.	.	.	.	.	.	.	.	.	D	Transcription factor required for ciliogenesis and islet cell differentiation during endocrine pancreas development. Essential for the differentiation of nodal monocilia and left-right asymmetry specification during embryogenesis. Required for the biogenesis of motile cilia by governing growth and beating efficiency of motile cells. Also required for ciliated ependymal cell differentiation. Regulates transcription by forming a heterodimer with another RFX protein and binding to the X-box in the promoter of target genes. Represses transcription of MAP1A in non-neuronal cells but not in neuronal cells.
SAMD9	p.A989T	D	D	D	D	M	T	T	T	16.50	4.88	7.240	.	May play a role in the inflammatory response to tissue injury and the control of extra-osseous calcification. acting as a downstream target of TNF-alpha signaling. Involved in the regulation of EGR1.
STARD7	p.H276Q	T	D	D	D	M	T	T	D	22.2	5.02	4.844	.	May play a protective role in mucosal tissues by preventing exaggerated allergic responses.
TRAK1	p.E698del	.	.	.	.	.	.	.	.	.	.	.	.	Involved in the regulation of endosome-to-lysosome trafficking. including endocytic trafficking of EGF-EGFR complexes and GABA-A receptors.
WNT11	p.R266Q	T	D	D	D	M	T	D	D	29.9	5.66	7.818	.	Ligand for members of the frizzled family of seven transmembrane receptors. Probable developmental protein. May be a signaling molecule which affects the development of discrete regions of tissues. Is likely to signal over only few cell diameters.

Gene	Variants	SIFT	Polyphen2	LRT	Mutation Taster	Mutation Assessor	FATHMM	Radial SVM	LR	CADD	GERP++	phyloP	HSF	Gene Function
ZKSCAN8	p.I478NfsX17	.	.	.	.	.	.	.	.	.	.	.	.	May be involved in transcriptional regulation.

Table S3.32 – Patient 31 variants without parents WES coverage. The last column indicates the reason for not continuing the study of this variants.

Gene	Variants	Reason of exclusion
ARHGEF17	p.C44Y	Predicted as benign.
CNNM3	p.R189C	Inherited from the mother.
DSPP	p.S671N	Relatively frequent.
FSCB	p.P679_A680insSEVQPPPAEEAP	Relatively frequent.
GRIA1	p.G16S	Predicted as benign.
HMCN2	p.R509C	Relatively frequent.
INPP5J	p.P151L	Predicted as benign.
KCNJ16	c.-94+7G>A	Predicted not to affect splicing.
LOC100129697	p.S403C	Predicted as benign.
MEX3C	p.L70P	Predicted as benign.
MUC19	p.T194M	Relatively frequent.
PRR18	p.L175I	Relatively frequent and predicted as benign.
PRR12	p.Q881R	Predicted as benign.
SHH	p.G290D	Relatively frequent.

Table S5.1A – Genes, variants and predicted functional impact for the analysis of the genes associated with secondary glaucoma, conditions ressambling PCG, and other type of glaucomas.
All variants have a European MAF inferior to 0.01 in at least in one database. Chr: chromosome, Ref: reference nucleotide, Alt: altered nucleotide, 1000G: 1000 Genomes Project, ExAC: Exome Aggregation Consortium, ESP: NHLBI Exome Sequencing Project, EUR: European, AFR: African, AMR: American, EAS: East Asian, SAS: South Asian, NFE: Non-Finnish European, FIN: Finnish European, OTH: Other populations, EA: European American, AA: African American. All variants are in the heterozygous state. Dot: no information on databases.

Sample ID	Gene	Chr	Start	Ref	Alt	Variant	1000G EUR	1000G AFR	1000G AMR	1000G EAS	1000G SAS	ExAC NFE	ExAC FIN	ExAC AFR	ExAC AMR	ExAC EAS	ExAC SAS	ExAC OTH	ESP6500 EA	ESP6500 AA
4	*MYOC*	1	171605246	G	A	p.A445V	0.001	0	0	0	0	0.0003	0	9.61E-5	0	0	0	0	0.0003	0.0002
5	*ZNF469*	16	88500123	G	A	p.R2054Q	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	ELP4	11	31804978	G	C	p.S394T	0.004	0	0.0029	0	0.0072	0.0060	0.0035	0.0011	0.0040	0.0005	0.0121	0.0089	0.0079	0.0011
	NF1	17	29592335	A	G	p.I1584V	.	.	.	.	.	.	.	.	.	.	.	.	.	.
9	*ZNF469*	16	88498850	T	A	p.W1630R	0	0	0.0029	0	0	0.0002	0	0	0.0028	0	0	0	.	.
	CREBBP	16	3823850	C	T	p.A789T	.	.	.	.	.	1.50E-5	0	0	0	0	0	0	.	.
	CBS	21	44483184	A	G	p.I278T	0.001	0	0	0	0	.	.	.	.	.	.	.	0.0028	0.0036
	ADAMTS10	19	8651438	T	C	c.2403+4A>G	0.008	0	0.0058	0	0.001	0.0081	0.0011	0.0009	0.0043	0	0.0032	0.0066	0.0088	0.0011
18	*ZNF469*	16	88500957	C	T	p.P2332L	0.005	0	0.0029	0	0	0.0033	0	0	0.0029	0	0.0009	0	.	.
	DCN	12	91558500	A	G	c.212-6T>C	0.004	.	.	.	.	0.0064	0.0055	0.0006	0	0	0	0.0011	0.0043	0.0007
25	*ELP4*	11	31653952	G	T	p.Q309H	.	.	.	.	.	.	.	.	.	.	.	.	.	.
31	*ADAMTSL4*	1	150530009	C	T	p.S696L	0.002	0.0098	0.0014	0	0	0.0010	0	0.0113	0.0022	0	0.0001	0.0055	0.0002	0.0066
	ADAMTSL4	1	150531616	G	A	p.R913H	0.001	0	0	0	0.001	0.0003	0	0.0001	0.0003	0.0005	0.0003	0.0025	0.0001	0
32	*COL11A1*	1	103540260	G	T	p.P189T	.	.	.	.	.	.	.	.	.	.	.	.	.	.
33	*ZNF469*	16	88504578	G	A	p.G3539E	0.001	0	0	0	0	0.0020	0	0	0	0	0.0004	0	.	.
	COL11A1	1	103379249	A	G	c.3979-3T>C	0	0.0098	0.0014	0	0	0.0001	0	0.0048	0.0005	0	0	0.0011	0.0003	0.0048
36	*COL4A1*	13	110827629	G	T	p.A1045E	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	COL11A1	1	103491103	G	A	p.P322S	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	NF1	17	29553639	A	T	p.N730Y	.	.	.	.	.	1.50E-5	0	0	0.0003	0	0	0	.	.
	OPTN	10	13160930	G	T	p.K223N	.	.	.	.	.	0	0	0	0.0003	0	0	0	.	.
37	*FOXC1*	6	1610872	C	G	p.Y64X	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	COL4A1	13	110866167	T	C	p.T81A	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	ADAMTS17	15	100871162	T	C	p.K183R	0.001	0.0023	0.0014	0	0	0.0019	0.0002	0.0014	0.0019	0	0	0.0011	0.0021	0.0016
	GALC	14	88416224	C	T	p.E435K	.	.	.	.	.	0	0	0.0001	8.78E-3	0	0	0	.	.
41	*FBN1*	15	48704843	C	T	p.E2717K	0	0.0008	0	0	0	0.0001	0	0	0.0008	0	0	0	0.0001	0
42	*WDR36*	5	110440034	G	T	p.A353S	0.001	0	0	0	0	0.0008	0.0003	9.65E-5	0.0002	0	0	0	0.0016	0

Sample ID	Gene	Chr	Start	Ref	Alt	Variant	1000G EUR	1000G AFR	1000G AMR	1000G EAS	1000G SAS	ExAC NFE	ExAC FIN	ExAC AFR	ExAC AMR	ExAC EAS	ExAC SAS	ExAC OTH	ESP6500 EA	ESP6500 AA
	NSD1	5	176639032	C	T	p.A1211V	.	.	.	.	.	1.50E-5	0	0	0	0	0	0	.	.
43	*FBN1*	15	48766748	A	G	p.I1355T	.	.	.	.	.	0	0	0	0	0.0001	6.06E-5	0	.	.
	CBS	21	44480591	G	A	p.R369C	0.003	0.0023	0	0	0	0.0053	0.0024	0.0012	0.0015	0	0	0	0.004	0.0011
	COL9A2	1	40778463	A	G	c.250-7T>C	0	0.0083	0.0014	0	0	0.0001	0	0.0058	0	0	0	0	0.0001	0.0039
45	*COL9A2*	1	40780067	A	G	c.151-8T>C	0	0.0053	0.0014	0	0	4.50E-5	0	0.0035	0.0003	0	0	0	0.0002	0.0034
	MTRR	5	7900145	C	T	p.R718C	0.002	0	0.0029	0	0	0.0004	0.0003	9.83E-5	0.0004	0.0002	0.0002	0	0.0005	0

Table S5.1B – Genes, variants and predicted functional impact for the analysis of the genes associated with secondary glaucoma, conditions ressambling PCG, and other type of glaucomas. Each functional impact predictor has a different classification with different cut-offs. SIFT: T – tolerated (>0.05), D – damaging (≤0.05); PolyPhen2: B – benign (<0.452), P – possibly damaging (0.453 and 0.956), D – probably damaging (>0.957); LRT: D – deleterious, N – neutral, U – unknown; MutationTaster: P – polymorphism automatic, N – polymorphism, D – disease causing, A – disease causing automatic; MutationAssessor: N – neutral, L – low, M – medium, H – high; FATHMM: T – tolerated (> -1.5), D – damaging (< -1.5); Radial SVM: T – tolerated, D – damaging; LR: T – tolerated, D – damaging; CADD: ≥15 considered pathogenic; GERP++: ≥2 considered conserved; PhyloP: ≥1; HSF: D – disturbs donor splice site; A – disturbs acceptor splice site; No – predicted to not affect splicing. Dot – no prevision was generated. All variants are in the heterozygous state. The last column indicates the classification of each variant. A variant is classified as pathogenic when 6 of the 9 functional impact software predict it as damaging, otherwise is classified as benign. A variant is classified as relatively frequent if MAF is > 1% in other populations but European or if there are individuals with homozygous variants, otherwise is classified as rare.

Sample ID	Gene	Variants	SIFT	Polyphen2	LRT	Mutation Taster	Mutation Assessor	FATHMM	Radial SVM	LR	CADD	GERP++	phyloP	HSF	Classification
4	*MYOC*	p.A445V	T	B	N	N	N	D	T	T	6.695	2.21	1.443	.	Rare and predicted as benign.
5	*ZNF469*	p.R2054Q	T	B	.	N	N	T	T	T	11.40	-2.41	-1.577	.	p.R2054Q and p.I1584V rare and p.S394T relatively frequent. Predicted as benign, but p.I1584V described as pathogenic.
	ELP4	p.S394T	T	P	.	D	L	T	T	T	19.60	4.83	7.420	.	
	NF1	p.I1584V	T	B	D	D	N	T	T	T	7.311	5.06	7.170	.	
9	*ZNF469*	p.W1630R	D	D	.	N	N	T	T	T	13.31	3.09	1.779	.	p.W1630R, p.A789T and p.I278T rare and c.2403+4A>G relatively frequent. p.W1630R and p.A789T predicted as benign, p.I278T predicted as pathogenic and c.2403+4A>G predicted as affect donor splice site.
	CREBBP	p.A789T	T	B	N	N	N	D	T	T	9.717	-8.31	-0.220	.	
	CBS	p.I278T	D	P	D	A	M	D	D	D	13.05	3.52	7.928	.	
	ADAMTS10	c.2403+4A>G	.	.	.	.	.	.	.	.	.	.	.	D	
18	*ZNF469*	p.P2332L	D	B	.	N	N	T	T	T	6.919	0.555	0.061	.	p.P2332L rare and c.212-6T>C relatively frequent. p.P2332L predicted as benign and c.212-6T>C as not affecting splicing.
	DCN	c.212-6T>C	.	.	.	.	.	.	.	.	.	.	.	No	
25	*ELP4*	p.Q309H	T	D	D	D	M	T	T	T	15.68	4.0	5.804	D	Rare and predicted to affect donor splice

Sample ID	Gene	Variants	SIFT	Polyphen2	LRT	Mutation Taster	Mutation Assessor	FATHMM	Radial SVM	LR	CADD	GERP++	phyloP	HSF	Classification
															site.
31	*ADAMTSL4*	p.S696L	D	D	.	D	L	T	T	T	29.0	4.76	1.889	.	Rare and predicted as benign.
	ADAMTSL4	p.R913H	T	B	.	N	N	T	T	T	11.32	2.23	1.915	.	
32	*COL11A1*	p.P189T	T	B	D	D	L	T	T	T	14.66	4.82	3.259	.	Rare and predicted as benign.
33	*ZNF469*	p.G3539E	D	B	.	N	L	T	T	T	10.07	-0.737	0.052	.	Relatively frequent. p.G3539E predicted as benign and c.3979-3T>C as not affecting splicing.
	COL11A1	c.3979-3T>C	.	.	.	.	.	.	.	.	.	.	.	No	
36	*COL4A1*	p.A1045E	T	B	N	N	N	D	T	T	0.007	3.5	0.467	.	Rare and predicted as benign.
	COL11A1	p.P322S	T	B	N	D	M	T	T	D	13.12	4.62	7.064	.	
	NF1	p.N730Y	T	B	D	D	N	T	T	T	3.696	4.43	6.888	.	
	OPTN	p.K223N	T	B	N	N	L	D	T	T	9.672	1.29	0.608	.	
37	*FOXC1*	p.Y64X	T	.	U	D	.	.	.	.	29.3	1.97	0.370	.	Rare. p.Y64X predicted as pathogenic and p.T81A, p.K183R and p.E435K as benign.
	COL4A1	p.T81A	T	B	N	N	N	D	T	D	5.673	-1.45	-0.188	.	
	ADAMTS17	p.K183R	T	B	N	N	L	T	T	T	11.08	-3.59	-0.384	.	
	GALC	p.E435K	T	B	N	N	N	D	T	T	2.489	1.45	-0.345	.	
41	*FBN1*	p.E2717K	T	B	D	D	M	D	T	T	25.2	5.38	7.651	.	Rare and predicted as benign.
42	*WDR36*	p.A353S	D	B	D	D	L	T	T	T	15.71	5.68	7.825	.	Rare and predicted as benign.
43	*NSD1*	p.A1211V	.	B	N	N	N	D	T	T	8.390	3.62	2.053	.	p.A1211V and p.I1355T rare and p.R369C relatively frequent. p.A1211V and p.I1355T predicted as benign and p.R369C as pathogenic.
	FBN1	p.I1355T	T	B	D	D	N	D	T	T	11.62	4.87	9.073	.	
	CBS	p.R369C	D	D	D	D	M	D	D	D	23.4	4.9	9.217	.	
45	*COL9A2*	c.250-7T>C	.	.	.	.	.	.	.	.	.	.	.	No	Rare. c.250-7T>C and c.151-8T>C predicted not to affect splicing. p.R718C as benign.
	COL9A2	c.151-8T>C	.	.	.	.	.	.	.	.	.	.	.	No	
	MTRR	p.R718C	D	B	D	D	M	D	D	T	17.04	5.7	4.175	.	