Faculdade de Medicina
Universidade de Coimbra

Comprehensive analysis of *MYOC* gene in Primary Open-Angle Glaucoma

Dissertação apresentada à Faculdade de Medicina da Universidade de Coimbra, para cumprimento dos requisitos necessários à obtenção do grau de Mestre em Investigação Biomédica, realizada sob a orientação científica do Doutor Filipe Silva e da Professora Doutora Manuela Grazina (Faculdade de Medicina, Universidade de Coimbra)

Filipa Ferreira, 2014

Copyright© Filipa Ferreira, Filipe Silva, Manuela Grazina 2014

Esta cópia da tese é fornecida na condição de que quem a consulta reconhece que os direitos de autor são pertença do autor da tese e do orientador científico e que nenhuma citação ou informação obtida a partir dela pode se publicada sem a referência e autorização.
This copy of the thesis has been supplied on condition that anyone who consults it is understood to recognize that its copyright rests with its author and scientific supervisor and that no quotation from the thesis and no information derived from it may be published without proper acknowledgment and authorization.

Agradecimentos

A realização desta dissertação não era possível sem todo o apoio, paciência, incentivo e colaboração de um grupo de pessoas pelo qual tenho um apreço inestimável. A todas elas um muito obrigado por terem colaborado e tornado possível esta caminhada.

Em primeiro lugar quero agradecer ao Doutor Filipe Silva por ter feito com que gostasse tanto de genética após uma Rotação Laboratorial que tanto me motivou. A forma como me ensinou a pensar, incentivou a pesquisar, exigiu rigor e a motivação com que se dedica às circunstâncias fizeram com que o meu crescimento a nível não só académico mas também pessoal evoluísse. Por tudo isso e toda atenção, paciência e preocupação não tenho palavras para demonstrar a minha gratidão e apreço, e mais uma vez obrigada.

À Professora e Co-orientadora Manuela Grazina obrigado por me ter dado a conhecer um outro lado da genética neste longo caminho desde as aulas até ao trabalho práctico.

Aos meus pais e irmão muito obrigado por terem tornado possível realizar mais esta etapa da minha vida e pelo apoio e compreensão que tiveram sempre, mesmo nos momentos mais difíceis.

Ao José Albuquerque não poderia deixar de agradecer tudo o que me ensinou, principalmente a não desistir e a mostrar que somos sempre mais fortes do que aquilo que nos enfrenta.

A todos os meus colegas de laboratório mas principalmente ao Nelson Monteiro obrigado pela companhia nas “noitadas” de laboratório, disponibilidade, paciência e amizade até ao fim. Sem ti teria sido muito mais difícil.

Ao Doutor João Pratas o meu agradecimento pela dedicação e empenho que teve neste projecto e para comigo.

À equipa e colaboradores do laboratório de Bioquímica e Genética pelo apoio mostrado e experiência transmitida na realização do trabalho experimental.

A todas as minhas amigas, especialmente Catarina, Jéssica, Susana e Melanie, obrigada por estarem sempre ali quando precisei “We are what We are”.

Aos doentes e familiares agradeço a contribuição e disponibilidade que tiveram para participar neste estudo.

À Isabel Marques agradeço a disponibilidade que teve em fornecer as amostras controlo e respectivos dados.

Ao Doutor Pedro Faria e José Moura Pereira obrigado pela disponibilidade e conhecimento que me transmitiram.

Aos Coordenadores do Mestrado, Doutor Paulo Pereira e Doutor Henrique Girão, por me terem aceite neste Mestrado e por todo o conhecimento que me transmitiram.

A todos vós muito obrigado!

Filipa Ferreira

“Energy and persistence conquer all things”.

Benjamin Franklin

Index

Tables index……ii
Figures index……iii
Abbreviations……iv
Abstract……vi
1. Introduction……1
1.1. Purpose……8
2. Methodology……9
2.1. Human subjects……10
2.2. DNA extraction and quantification……11
2.3. Polymerase chain reaction……11
2.4. Electrophoresis……12
2.5. Purification of PCR products……13
2.6. Sequencing reaction……13
2.7. Purification of sequencing reaction products……13
2.8. Capillary electrophoresis……14
2.9. SNaPshot analysis……14
2.10. Statistical analysis……15
2.11. In silico analysis……15
3. Results……17
3.1. Sequencing analysis of MYOC gene……18
3.2. Missense variants analysis……21
3.3. SNaPshot analysis……23
3.4. Case-control study with MYOC variants……26
3.5. In silico analysis……29
3.5.1. Evolutionary conservation study……30
3.5.2. Pathogenicity prediction……36
3.5.3. Transcription factor study……37
4. Discussion……39
5. Conclusion……46
6. References……48

Tables index

Table I. POAG Chromosomal loci……………………………………………………………..5

Table II. Subjects characteristics…………………………………………………….......11

Table III. Primers used for amplification of MYOC gene exons…………………………..12

Table IV. Primers designed for the SNaPshot analysis………………………………......16

Table V. Sequence variations identified in MYOC gene…………………………………19

Table VI. Clinical features of Glaucoma patients with MYOC variants……………….......21

Table VII. Control subjects genotypes……………………………………………………….26

Table VIII. Genotypes statistical analysis…………………………………………………….27

Table IX. Alleles statistical analysis……………………………………………………………27

Table X. Males case-control study…………………………………………………………….28

Table XI. Females case-control study…………………………………………………………29

Table XII. Transcription factor study for variant c.-190G>T………………………………..38

Table XIII. Transcription factor study for variant c.-83G>A……….......……………….......38

Figures Index

Figure 1. POAG classification…………………………………………………………….....3
Figure 2. Schematic representation of human myocilin protein……………………….......7
Figure 3. Schematic representation of MYOC gene with all the identified genetic variations in the Portuguese population……………………………………..……20
Figure 4. Electropherograms of the region surrounding the two variants candidates to Glaucoma causing disease mutations………………………………….……..22
Figure 5. Electropherograms of the multiplex SNaPshot analysis for the selected genetic variants found in POAG patients………………………………………..……..23
Figure 6. Evolutionary conservation study for c.-190G>T variant……………………..……31
Figure 7. Evolutionary conservation study for c.-83G>A variant……………………….....32
Figure 8. Evolutionary conservation study for c.227G>A (Arg76Lys) variant…………...33
Figure 9. Evolutionary conservation study for c.878C>A (Thr293Lys) variant…...........34
Figure 10. Evolutionary conservation study for c.1334C>T (Ala445Val) variant…..……..35
Figure 11. Pathogenicity analysis of Arg76Lys, Thr293Lys and Ala445Val genetic variants..……………………………………………………………………….36

Abbreviations

A	Adenine
AH	Aqueous Humor
AhR/AR	Complex aryl hidrocarbon receptor
Ala	Alanine
Arg	Arginine
Bp	Bases pair
C	Cytosine
Cys	Cysteine
DNA	Deoxyribonucleic acid
DNAse	Desoxirribonuclease
ECM	Extra Celullar Matrix
EDTA	Ethylenediamine tetraacetic acid
ER	Endoplasmic Reticulum
F	Forward
Fig.	Figure
FISH	Fluorescence in situ Hybridization
G	Guanine
GLC	Glaucoma Locus
Gln	Glutamine
H	Hour
Het	Heterozygous
HTM	Human Trabecular Meshwork
HWE	Hardy-Weinberg equilibrium
IOP	Intra Ocular Pressure
JOAG	Juvenil Open Angle Glaucoma
kDa	Kilodalton
Leu	Leucine
Lys	Lysine
Max	Maximum
Min	Minimum
Min	minute(s)
Mut	Mutant
MYOC	Myocilin
n	Number of subjects
NCBI	National Center for Biotechnology Information

nGRE	negative Glucocorticoid response element
NTG	Normal Tension Glaucoma
OLF	Olfactomedin like domain
OR	Odds Ration
P	p-value
PCR	Polymerase Chain Reaction
POAG	Primary Open Angle Glaucoma
Pro	Proline
R	Reverse
RGCs	Retinal Ganglionar Cells
RNAse	Ribonuclease
RT	Room Temperature
SC	Score
SD	Standard Deviation
Sec	Second(s)
T	Timine
TBE	Tris Borate EDTA
TF	Transcription Factor
Thr	Threonine
TIGR	Trabecular Meshwork-Inducible Glucocorticoid Response
Tyr	Tyrosine
USF	Upstream regulatory factor
UTR	Untranslated Region
UV	Ultraviolet
Val	Valine
Wt	Wild-type
X	STOP codon
Yo	Years old

Abstract

Glaucoma is an optic neuropathy characterized by irreversible optic nerve degeneration, progressive visual field defects, and cupping of glaucomatous optic disc. This condition is a leading cause of blindness worldwide affecting ~67 million people with 6.8 million bilaterally blind. Primary Open-Angle Glaucoma (POAG) is the most common form of Glaucoma accounting for 70% of all cases in Caucasians. This type of Glaucoma comprises three subtypes: POAG with adult age of onset, Juvenile Open-Angle Glaucoma (JOAG; onset <35 years old) and Normal Tension Glaucoma (NTG; normal intraocular pressure).

MYOC gene was the first gene identified as POAG causative and ever since approximately 99 sequence variations were considered Glaucoma causing disease mutations.

The purpose of this study was to identify sequence variations in *MYOC* gene that may be responsible for the phenotype in a group of Portuguese POAG patients.

In order to fulfil the objective of this work, 104 POAG patients were analyzed by a PCR-sequencing method. Additionally, 54 controls subjects were studied by a SNaPshot approach, followed by statistical analysis with GraphPad Prism v.6. Statistically significant differences were assumed when $p < 0.05$. Finally, *in silico* analyses were performed to investigate the evolutionary conservation, functional effects and transcription factor search in the most relevant variants identified.

The sequencing results of the 3 coding exons and proximal splicing junctions of *MYOC* gene in 104 POAG patients allowed identifying 27 variants. From these, two variants are likely Glaucoma causing disease mutations and were identified in two NTG patients. A statistically significant association was found among females suggesting a possible protective effect (OR=3.29). Further studies are required to clarify the possible involvement of the identified variants in POAG etiology.

1. Introduction

Glaucoma is a group of optic neuropathies essentially characterized by a progressive degeneration of retinal ganglion cells (RGCs) and their axons, leading to excavation in the optic nerve head and, consequently, a specific, progressive and irreversible visual field defects (Fuse 2010). At the beginning, the slightly peripheral vision loss will not interfere with the daily routine and remains undetected (Janssen et al. 2013). Thus, until reach an advanced stage of the disease that usually consists on central vision loss, the majority of the patients are unaware that they have the disease and, consequently, remain undiagnosed and untreated (Fan & Wiggs 2010). These are the reason why some authors call Glaucoma as the “silent blinder” (Coleman 2000).

This ocular pathology is a leading cause of irreversible blindness worldwide affecting approximately 67 million people with 6.8 million bilaterally blind (Pasutto et al. 2012). It is estimated that this number will increase to 79.6 million in 2020 (Quigley & Broman 2006) and that it will become an even greater public health concern due to the impact that the symptoms have in the quality of life of the patients (Fingert et al. 2002).

Glaucoma may be classified according to the origin of the disease, the age-of-onset and the anatomy of the anterior chamber angle (Faucher et al. 2002; Fan et al. 2006; Ray & Mookherjee 2009).

The primary glaucomas share some clinical features including specific abnormal appearance of the optic nerve head, specific loss of the visual field and chronic painless progression (Sarfarazi, 1997). Primary Open-Angle Glaucoma (POAG), Glaucoma with primary etiology and open anterior chamber angle, is the most frequent type of Glaucoma and it may be subdivided in 3 subtypes. According to the age of onset, patients with an age of onset usually ranging between 3 and 35 years old (yo) are designated as Juvenil Open-Angle Glaucoma (JOAG), while patients with adulthood onset (usually more than 35 yo) are classified as Primary Open-Angle Glaucoma (POAG). The intraocular pressure (IOP) also contribute for the classification of POAG patients distiguishing the high pressure Glaucoma (POAG and JOAG), with an IOP equal or higher than 21mmHg, from the low IOP (less than 21mmHg), known as Normal Tension Glaucoma (NTG) (Angius

1998; Faucher et al. 2002; Fan et al. 2006; Fuse 2010; Gemenetzi et al. 2012) (Fig. 1). Moreover, JOAG patients usually have a more severe phenotype with considerably higher IOP (Fingert et al. 2002).

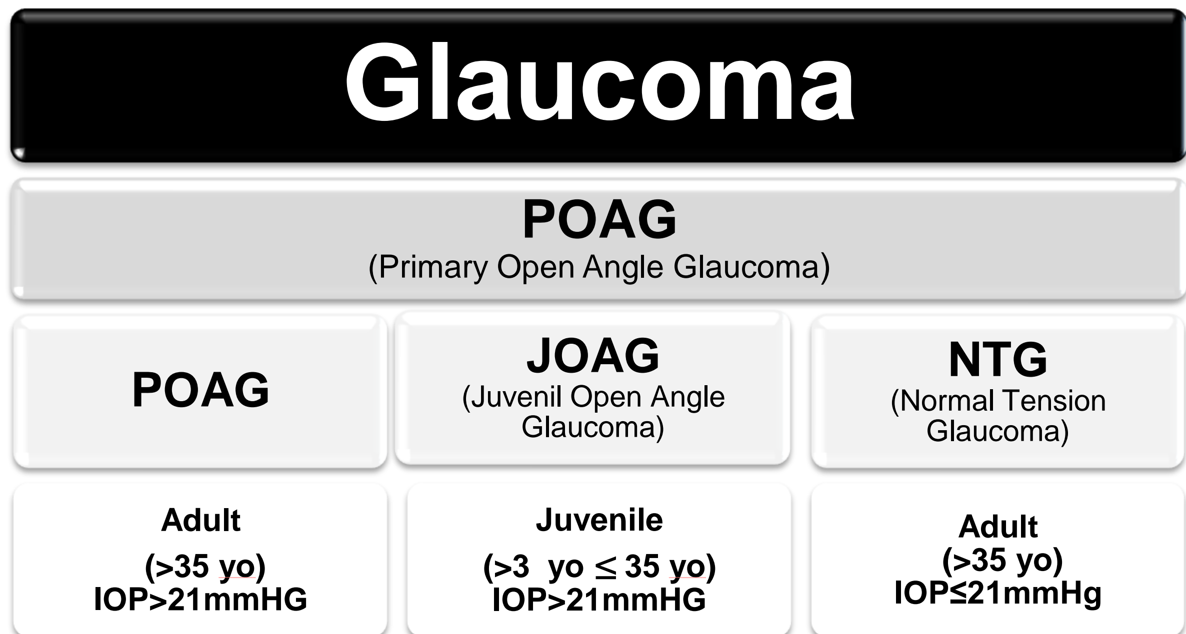


Figure 1. POAG classification

POAG accounts for 70% of all the Glaucoma cases in the Caucasian population (Bayat et al. 2008), and adult POAG accounts for approximately 2% of the same population in industrialized countries (Faucher et al. 2002). The variable prevalence among population is remarkable since in Asian surrounds 1-4% whereas in African countries may increase to 2-8%, suggesting a significant genetic contribution to the disease (Janssen et al. 2013). Alward and collaborators also mentioned the highest prevalence in African and African-descendents populations when compared with the Caucasian populations (Alward et al. 1998). The prevalence of NTG patients accounts for 20-50% of the POAG cases (Weisschuh et al. 2005).

The elevated IOP has been considered a major risk factor for Glaucoma but it is necessary to take in mind that the onset and progression of Glaucoma is independent of IOP (Sarfarazi 1997). Ageing (rising 4% from 40 to 80 years old) (Ennis et al. 2010), race (African-Americans have a 4-fold increased risk of developing POAG) (Kanagavalli et al. 2004), myopia (Janssen et al. 2013), systemic hypertension (Fuse 2010; Gemenetzi et al. 2012), demographic factors (Menaa et al. 2011), diabetes mellitus (Fuse 2010; Gemenetzi et al. 2012) and family history (first-degree relatives of affected individuals have 3-9 fold more likely to develop the disease) (Whigham et al. 2011), are other recognized risk factors.

Over the years has been demonstrated that POAG has a genetic etiology (Fingert, 2011). Several studies have been performed and POAG families carrying mutations with a Mendelian autosomal dominant fashion have been detected (Stone et al. 1997; Sarfarazi, 1997; Angius, 1998). Nevertheless, it is likely that multiple genes and/or environmental factors contribute to the complex inheritance of POAG (Fan & Wiggs 2010).

The importance of discovering disease causing genes has increased over the years. This information could be a very usefull knowledge about the pathogenesis of heritable eye diseases at the most basic level (Fingert, 2011).

To date, 22 *loci* have been found to be linked with POAG (Table I), but only 5 genes were already identified including: *Myocilin* (*MYOC*, *GLC1A*), *Optineurin* (*OPTIN, GLC1E*), *Ankyrin Repeat and SOCS Box Containing 10* (*ASB10, GLC1F*), *WD-Repeat Domains 36* (*WDR36*, *GLC1G*) and *Neurotrophin-4* (*NTF4, GLC1O*) (Gray et al. 2013).

Table I. POAG Chromosomal *loci*			
Locus	**Location**	**Gene**	**Type of glaucoma**
GLC1A	**1q23-q24**	***MYOC***	**JOAG**
GLC1B	2cen-q13		POAG
GLC1C	3q21-q24		POAG
GLC1D	8q23		POAG
GLC1E	**10p14**	***OPTN***	**POAG/NTG**
GLC1F	**7q35-q36**	***ASB10***	**POAG**
GLC1G	**5q22.2**	***WDR36***	**POAG**
GLC1H	2p16,3-p15		POAG
GLC1I	15q11-q13		POAG
GLC1J	9q22		JOAG
GLC1K	20p12		JOAG
GLC1L	3p22-p21		POAG
GLC1M	5q		JOAG
GLC1N	15q22-q24		JOAG
GLC1O	**19q13.3**	***NTF4***	**POAG/JOAG/NTG**
GLC1P	12q14		NTG
GLC1Q	**4q35.1-q35.2**		**POAG**

MYOC was the first gene identified and associated with POAG. Previously to this gene identification, Sheffield and colleagues mapped the GLC1A *locus* (*GLC* stands for Glaucoma, "*1*" designates Primary Open-Angle and *A* stands for the first linkage to the phenotype) comprising an interval on the long arm of chromosome 1 (1q21-q31) that was associated with JOAG (Stone et al. 1997). Later, this region was limited to 1q23—q25 region (Angius et al 1998). The defective gene at this locus was finally identified as *TIGR* (*Trabecular meshwork-Inducibe Glucocorticoid Response*) (Stone et al. 1997). Meanwhile, Japanese researchers mapped this gene by FISH within chromosome 1q23-1q24 region and due to its homologous regions with myosin, *TIGR* was named *MYOC* (Kubota et al. 1997).

MYOC is abundantly expressed in ocular tissues such as ciliary body, iris, trabecular meshwork (TM) and trabecular beams, connective tissue and abundantly detected in the aqueous humor. Also the heart, skeletal muscle, stomach, thyroid,

trachea, bone marrow, thymus, prostate, small intestine and colon show *MYOC* expression (Kanagavalli et al. 2004; Gobeil et al. 2006; Lopez-Martinez et al. 2007; Menaa et al. 2011).

The human *MYOC* gene is composed by three exons of 604, 126, and 782 base pairs (bp) which encodes a 55 to 57 kDa acidic glicoprotein with a total of 504 amino acids (Tamm, 2002; Fan et al. 2006). The region of the gene that encodes the promoter contains several DNA sequences that are essential for its expression at the basal transcriptional level (Fingert et al. 2002). Also a upstream regulatory factor (USF) that binds to E-box its critical for basal transcription of *MYOC* in TM cells and astrocytes (Kirstein et al. 2000). Besides that, multiple steroid response elements were also identified and studied in this region. This data could explain the fact of myocilin has an elevated expression in TM cells induced by glucocorticoids (Nguyen, 1998).

Myocilin protein is formed by three known domains as demostrated in Fig. 2. The N-terminal region of the gene contains a signal peptide that targets myocilin to secretion via secretory pathway, and an α-helical coiled coil domain (also known as myosin-like domain), which contains therein a leucine zipper motif, possibly involved in myocilin-myocilin interactions. There are some evidences that myocilin may form dimers/oligomers through the leucine zipper motif (Fingert et al. 2002; Kanagavalli et al. 2004; Qiu et al. 2014). Lastly, this protein is largely composed by an olfactomedin-like domain (OLF) at the C-terminal region which is structurally organized in β-sheets (Kanagavalli et al. 2004). A functional analysis of myocilin domains demonstrated that changes in coiled-coil/leucine zipper regions prevented the adhesion of the mutated myocilin to the Extracellular Matrix (ECM) or/and to the cell surface, and that OLF integrity is critical for myocilin folding and failure in this process cause myocilin sequestration (Adam et al. 1997; Gobeil et al. 2006).

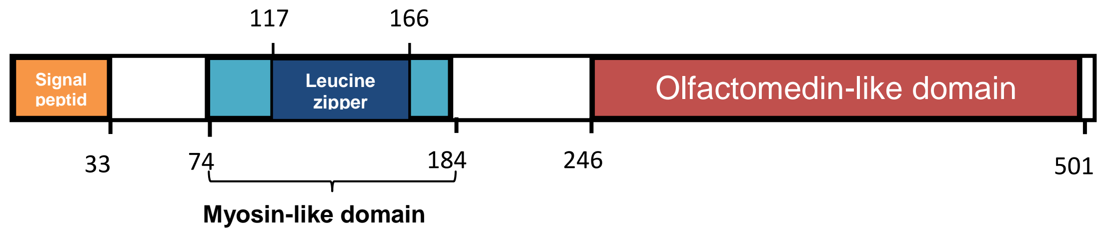


Figure 2. Schematic representation of human myocilin protein. Colored areas mark the position of the signal peptide, the leucine zipper-like motif and the myosin-like domain (N-terminal region), as well as the olfactomedin (OLF)-like domain (C-terminal region). Adapted from (Tamm, 2002) and (Menaa et al. 2011).

Despite *MYOC* has been the first gene associated with POAG its functions are still unclear. However, some data suggest two possible mechanisms of myocilin function in the pathophysiology of Glaucoma, the insufficient levels of secreted myocilion and/or an impaired function of Human Trabecular Meshwork (HTM) cells resulting from the congestioned secretory pathway (Jacobson et al. 2001). These facts lead to an increased resistance of aqueous humor (AH) outflow, while the maintenance of *MYOC* mutant forms within the HTM cells cause citotoxicity and programmed cell death (Jacobson et al. 2001; Sohn et al. 2002; Fingert, 2011).

Several studies have been performed to identify mutations in *MYOC* gene that may be considered responsible for the POAG phenotype. It is estimated that 2-4% of the identified mutations in this gene are responsible for POAG and up to 22-36% were found in JOAG cases. Furthermore, *MYOC* gene mutations have been found in Caucasian, Asian and African-American populations of POAG patients (Takamoto & Araie 2014). Moreover, *MYOC* gene mutations were found in German and Spanish NTG patients (Weisschuh et al. 2005; Lopez-Martinez et al. 2007). Additionally, the prevalence of NTG was previously reported as higher in the Japanese population than in other populations, which may indicate the existence of specific *MYOC* mutations in the Japanese patients with NTG phenotype (Mabuchi et al. 2001; Takamoto & Araie 2014).

Some missense variants within the *MYOC* gene exon 3 were deeply studied and revealed a change in the protein function. It is already known that TM cells participate in trabecular extracelular matrix turnover, and that this tissue is responsible for AH outflow regulation (Ueda et al. 2000; Sohn et al. 2002). *MYOC* mutations have been implicated on the increased resistance of HTM outflow which, consequently, increases IOP.

Myocilin is a glicoprotein localized in multiple sites intra and extracellularly of TM cells (Ueda et al. 2000). Several studies have been performed in trabecular tissue and cultured cells demonstrating that myocilin is an extracellularly secreted protein, but a substancial part is present in the endoplasmic reticulum (ER) (Sohn et al. 2009). Otherwise, mutant forms of *MYOC,* such as Pro370Leu (responsible for the most severe phenotype, as much as it is known) and the most common Gln368X, are not secreted, keeping up retained in the cells and affecting the ER (Adam et al. 1997; Fingert et al. 1999; Qiu et al. 2014). The authors concluded that the truncated protein also form oligomers with wyld-type form of myocilin supressing its secretion (Jacobson et al. 2001).

Some promoter variants such as -8C>T, -126T>G, -78T>G and -77G>A were found in POAG patients (Saura et al. 2005) These variations of sequence could change the association of the transcription factors with *MYOC* consensus sequence, changing the transcription rate and, consequently, altering the regulation and the expression of the gene.

1.1. Purpose

The aim of the present study was to identify sequence variations in *MYOC* gene that could be responsible for the phenotype in a group of Portuguese POAG patients. Additionally, determine the most relevant mutation in such population using bioinformatic tools and statistical analysis was also a goal.

2.Methodology

2.1 Human subjects

Clinical characterization of the 104 POAG patients' was performed at the Ophthalmology Department of the Centro Hospitalar e Universitário de Coimbra, E.P.E. by Dr. Moura Pereira and Dr. Pedro Faria. This study was approved by the Faculty of Medicine, University of Coimbra Ethics Committee, following the Tenets of the Declaration of Helsinki and a written consent for genetic testing was obtained from adult probands or parents of minor patients.

All patients underwent a detailed ophthalmologic examination to ensure the diagnosis of POAG including: 1) exclusion of secondary causes, 2) open drainage angle on gonioscopy (Shaffer's grading III-IV), 3) presence of typical glaucomatous optic disc damage (excavation) and 4) visual field defects detected by automated perimetry (with Humphrey's perimeter). The IOP was also evaluated since ocular hypertension is a major risk factor for Glaucoma and allows distinguishing POAG subtypes. Accordingly, patients with POAG or JOAG subtypes have IOP above 21mmHg while NTG patients present with IOP less than 21mmHg. Likewise, the age of onset contributed to distinguish adult POAG patients from the JOAG that presented an onset comprised between 3 and 35 yo.

The patients group consisted of 52 males and 52 females with an average age of 73 ± 10 years, ranging from 36 to 88 yo (Table II). The control samples were obtained from the Laboratório de Hematologia do Centro Hospitalar e Universitário de Coimbra E.P.E. from the general population. The control subjects included 28 males and 26 females with an average age of 48 ± 17 years, ranging from 21 to 80 yo (Table II).

Table II. Subjects characteristics			
	n	**Age**	**Gender**
Patients	104	Average = 73 SD = 10 Min = 36 Max = 88	52 M 52 F
Controls	54	Average = 48 SD = 17 Min = 21 Max = 80	28M 26F

2.2 DNA Extraction and quantification

The POAG patients and controls peripheral blood was collected in EDTA Tubes SARSETED and the DNA was extracted using a standard phenol–chloroform method followed by ethanol precipitation (Ausubel et al. 2003).

The extracted DNA was quantified by spectrophotometry at optical density of 260nm using the NanoDrop™ 1000 (Thermo Fisher Scientific, Inc.) and the GeneQuant™ II (Pharmacia Biotech).

2.3 Polymerase chain reaction

Individual exons and adjacent regions of *MYOC* gene were amplified by Polymerase Chain Reaction (PCR) technique using the Primers designed with Primer3 software (http://bioinfo.ut.ee/primer3-0.4.0/primer3/) (Table III) and the reference *MYOC* gene sequence obtained from NCBI database (http://www.ncbi.nlm.nih.gov/gene/). The PCR technique was performed on a Veriti® Thermal Cycler (Applied Biosystems). The PCR reactions were performed using 50-200ng of genomic DNA mixed with the following reagents: 1X Taq Buffer 10X (with $(NH_4)_2SO_4$) (Fermentas) 0.2µM of Forward and Reverse

Primers (Sigma-Aldrich), 1.5mM $MgCl_2$ (Fermentas), 0.2mM dNTPs (5PRIME), 1U Taq Polymerase (Fermentas) and RNAse/DNase free Water (AccuGENE) to a final volume of 10µl. The reaction mixture were subjected to a specific PCR program with an initial denaturation step of 5 minute (min) at 95ºC followed by 35 cycles, each with denaturation at 95°C for 30 seconds (sec), annealing at 59-63°C for 30 sec, and extension for 1 min at 72°C, with a final elongation step of 10 min at 72°C.

Table III. Primers used for amplification of *MYOC* gene exons

	Primer sequence (5' - 3')
Exon 1	F: GGTGCATAAATTGGGATGTTC R: GAATTAAGGAAAGCACAGCGA
Exon 2	F: AACCAGCCCTTCTAGTGGAAT R: CTATCACCCACACTTGGGAG
Exon 3	F:TGTCAATAACATGAAACACAGATTGAT R: GGAATTGTAGTCTGAGGGCGT

2.4 Electrophoresis

PCR products underwent an electrophoresis on an agarose gel containing 1% agarose SeaKem LE (Lonza) and 1% ethidium bromide (Acros/Fisher Bioreagents) in 1x Tris Borate EDTA (TBE) solution (National diagnotics). The products with a Loading Buffer (BIORAD) migrated under an electric current of 200 volts for 15 minutes. A *NZY DNA Ladder* VI (NZYTech) was also used to estimate the size of the fragments and respective quantification by comparison with the Molecular Weight Marker. The fragments were visualized through UV light in the ChemiDoc™ XRS System and respective System with Image Lab™ Software (BIORAD).

2.5 Purification of PCR products

The amplified PCR products were purified using 1µl of ExoSAP-IT® (USB)/4µl of amplified PCR product, incubated at 37ºC for 1 hour (h) followed by 15 min at 75ºC on a Veriti® Thermal Cycler (Applied Biosystems).

2.6 Sequencing reaction

Sequencing reactions were performed with 10ng of PCR products, 1X BigDye® Terminator 5X Sequencing Buffer (Applied Biosystems), 0.16µM of Forward or Reverse Primer (Sigma-Aldrich), BigDye® Terminator v3.1 according to manufacturer recommendations (Applied Biosystems) and RNAse/DNAse free Water for a final volume of 10µl. The Primers used were the same as for PCR products amplification (Table III). Sequencing reaction was run on a Veriti® Thermal Cycler (Applied Biosystems) using the following conditions: an initial denaturation step of 2 min at 96ºC followed by 25 cycles each of a denaturation at 96ºC for 5 sec, annealing at 59-63ºC for 10 sec, followed by one last step of elongation at 60ºC for 4 min.

2.7 Purification of Sequencing reaction products

A total of 8µl of RNAse/DNAse free Water and 32µl of 95% ethanol (Merck) were independently added to each Sequencing reaction product followed by capping and inversion of the plate for mixing the reagents. After 15 min of rest at room temperature (RT), the solution was subjected to a centrifugation at 2500 × g for 30 min). The supernatants were discarded in a paper towel and the samples were washed with 75µl 70% ethanol (Merck). The plate was capped and inverted a few times, followed by centrifugation at 2000 x g during

10 min. The supernatants were discarded as previously and the inverted plate on a paper towel was centrifuged at 700 × g for 1 min. Purified Sequencing reaction products were resuspended with 10 µl of HiDi™ (Applied Biosystems) and denatured at 96ºC for 5 min on a Veriti® Thermal Cycler (Applied Biosystems) previously to be analyzed by Capillary electrophoresis.

2.8 Capillary electrophoresis

Automated capillary DNA sequencing was performed on a Genetic Analyzer 3130 (Applied Biosystems) equipment. DNA sequencing data obtained was analyzed using Sequencing Analysis Software v.5.4® (Applied Biosystems) and SeqScape v2.5® (Applied Biosystems), that allowed the identification of *MYOC* gene variants by comparison with gene reference sequence (http://www.ncbi.nlm.nih.gov/sites/entrez).

2.9 SNaPshot analysis

SNaPshot reactions were performed with 1µl of previously purified PCR product added to SNaPshot Ready Multiplex Reaction Mix (Applied Biosystems) according to manufacturer recommendations, 0.4 µM of each Primer (Table IV) and RNAse/DNAse free Water up to a final reaction volume of 5µL. Extension reactions were achieved on a Veriti® Thermal Cycler (Applied Biosystems) for 25 cycles under the following conditions: 96ºC for 10 sec, 50°C for 5 sec and 60°C for 30 sec. After the SNaPshot reactions the products were treated with 1µl of Shrimp Alkaline Phosphatase ® (SAP) (USB) at 37ºC for 60 min followed by 15 min at 75ºC. Subsequently, 1µl of purified SNaPshot reaction product was mixed with 0.125µl of GeneScan™-120 LIZ™ (Applied Biosystems) size standard and 8.875µl of HiDi™ (Applied Biosystems), and denaturated for 5 min at 95ºC. Afterwards, the denatured

SNaPshot reaction products were subjected to a capillary electrophoresis on a Genetic Analyzer 3130 (Applied Biosystems) and the results were analyzed with the GeneMapper v.4.1 Software (Applied Biosystems).

2.10 Statistical analysis

The SNaPshot results of the polymorphic sequence variations were analyzed with GraphPad Prism v.6. The standard χ^2 test with contingency tables was used to test for deviation from the Hardy-Weinberg Equilibrium and to compare the genotypic and allelic frequencies in affected and control groups. Statistically significant differences were assumed when $p < 0.05$.

2.11 In silico analysis

The evolutionary conservation study, both in nucleotides and in amino acids sequences, included primate species that have several similarities with humans (*Homo sapiens*), such as Chimpanzee (*Pan troglodytes*), Gorilla (*Gorilla gorilla*), Macaque (*Macaca mulatta*) and Orangutan (*Pongo abelli*), and non-primates species such as Mouse (*Mus musculus*), Rat (*Rattus norvegicus*) and Zebrafish (*Danio rerio*). The www.ensembl.org was the genome database used for all necessary species sequences.

The prediction of functional effects of the sequence variations on proteins was performed using the PolyPhen-2® *software* (genetics.bwh.harvard.edu/pph2/) and the PROVEAN® software (provean.jcvi.org/seq_submit.php).

The evaluation of transcription factors binding site was performed using TFSEARCH® (Searching Transcription Factor Binding Sites (ver 1.3)) online software (www.cbrc.jp/research/db/TFSEARCH.html).

Table IV. Primers designed for the SNaPshot analysis	
Genetic variations	**Primers (length/direction)**
c.-224T>C	30F: CTTTTTAAAAAGAAACTCCAAACAGACTTC
c.-126T>C	35R: AGAGAGGTTTATATATACTGGGGAGCCAGCCCTTC
c.-190G>T	40R: GTGAGGCTGGGTGGGGCTGTGCACAGGGGGGTTGCCTTCA
c.1193A>G (Lys398Arg)	45F: TGGATGAAGCAGGCCTCTGGGTCATTTACAGCACCGATGAGGCCA
c.-83G>A	50R: CTCTGCTGTGCTGAGAGGTGCCTGGATGGGTGGCCTTGCTGGCTCATGCC
c.1334C>T (Ala445Val)	55R: CAGGGTCTTGCTGATACCTGTGCCTGTGTCATAAGCAAAGTTGACGGTAGCATCT
c.227G>A (Arg76Lys)	60R: TCCAGGGAGCTGAGTCGAGCTTTGGTGGCCTCCAGGTCTAAGCGTTGGGTGCTGCTGTCT
c.877G>T (Thr293Lys	65R: CCTTAGAAGGGTAGCCCTGCATAAACTGGCTGATGAGGTCATACTCAAAAACCTGGCGGACATCC

The nomenclature used in **Table IV** for sequence variations is according to den Dunnen and Antonarakis, 2000.

3. Results

3.1 Sequencing analysis of *MYOC* gene

After PCR-sequencing analysis of 104 POAG Portuguese patients it was possible to identify at least one genetic variant in *MYOC* gene exons or their adjacent regions. In total, 27 different sequence variations were found, including 11 coding and 16 non-coding variants (Table V). From the coding variants 4 were missense alterations [c.227G>A (Arg76Lys), c.878C>A (Thr193Lys), c.1193A>G (Lys198Arg) and c.1334C>T (Ala445Val)] and 7 were synonymous variants [c.39T>G (Pro13Pro), c.141C>T (Cys47Cys), c.477A>G (Leu159Leu), c.855G>T (Thr285Thr), c.975G>A (Thr325Thr) and c.1041T>C (Tyr347Tyr)]. Four sequence variations were found in the promoter region (c.-224T>C, c.-190G>T, c.-126T>C and c.-83G>A). It is noticeable that the promoter variant c.-83G>A and the exon 1 alteration c.227G>A were always found simultaneously in the same patients, even sharing the same genotype. Finally, the adjacent regions of *MYOC* gene also revealed several non-coding variants. Intron 1 presented six variants (c.604+177G>A, c.605-604G>A, c.605-374G>C, c.605-332G>A, c.605-302C>G and c.605-280G>T), while intron 2 had six alterations (c.731-205 A>C, c.731-192 G>A, c.731-73 C>T, c.730+35A>G, c.730+138G>A and c.730+176delCT) (Figure 3).

Table V. Sequence variations identified in *MYOC* gene					
SNP	**Nucleotide position**	**Amino acid change**	**n = 104**		
			wt	**Het**	**Mut**
rs235920	c.-224T>C	-	55 TT	44 TC	5 CC
rs76745622	c.-190G>T	-	103 GG	1GT	-
rs34928744	c.-126T>C	-	100 TT	4 TC	-
rs2075648	c. -83G>A	-	86 GG	17 GA	1AA
rs12082573	**c.39T>G**	**Pro13Pro**	**101 TT**	**3 TG**	**-**
-	**c.141C>T**	**Cys47Cys**	**103 CC**	**1 CT**	**-**
rs2234926	**c.227G>A**	**Arg76Lys**	**86 GG**	**17 GA**	**1 AA**
rs61730977	**c.477A>G**	**Leu159Leu**	**102 AA**	**2 AG**	**-**
rs604864	c.605-604G>A	-	103 GG	1 GA	-
rs603930	c.605-374G>C	-	97 GG	7 GA	-
rs41263718	c.605-332G>A	-	84 GG	20 GA	1AA
rs41263716	c.605-302C>G	-	96 CC	8 CG	-
rs603490	c.605-280G>T	-	26 GG	55 GT	23 TT
rs113416006	c.604+177G>A	-	102 GG	2 GA	-
rs57824969	**c.612G>T**	**Thr204Thr**	**103 GG**	**1 GT**	
rs2032555	c.730+35A>G	-	50 AA	46 AG	8 AA
-	c.730+138G>A	-	103 GG	1 GA	-
rs144871239	c.730+176delCT	-	103	1del CT	-
rs12076134	c.731-205A>C	-	74 AA	30 AC	-
-	c.731-192G>A	-	103 GG	1GA	-
rs79255460	c.731-73C>T	-	102 CC	2CT	-
rs146606638	**c.855G>T**	**Thr285Thr**	**103 GG**	**1GT**	**-**
rs139122673	**c.878C>A**	**Thr293Lys**	**103 CC**	**1 CA**	**-**
rs61730976	**c.975G>A**	**Thr325Thr**	**101 GG**	**3 GA**	**-**
rs61730974	**c.1041T>C**	**Tyr347Tyr**	**98 TT**	**6 TC**	**-**
rs56314834	**c.1193A>G**	**Lys398Arg**	**103 AA**	**1 AG**	**-**
rs140967767	**c.1334C>T**	**Ala445Val**	**103 CC**	**1 CT**	**-**

Amino acid coding variants are in bold. Abbreviations: wt - wild type; Het - heterozygotes; Mut - mutant
The nomenclature used in **Table V** for sequence variations is according to den Dunnen and Antonarakis, 2000

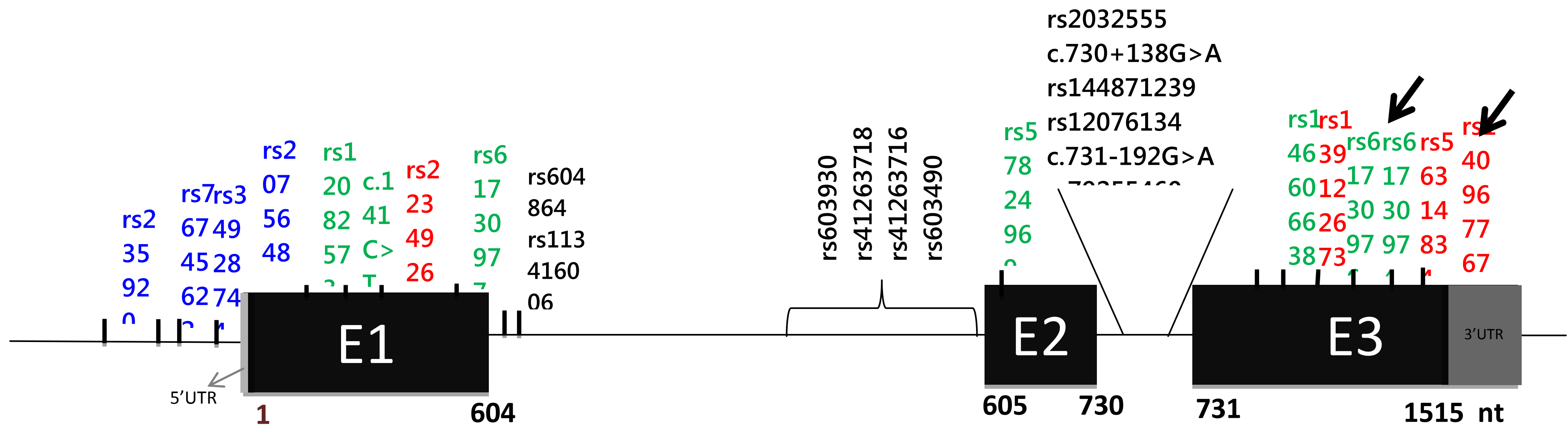


Figure 3. Schematic representation of *MYOC* gene with all the identified genetic variations in the Portuguese population. The non-synonymous sequence variations are represented in red while the synonymous are represented in green. In blue are represented the variants located in the promoter region. In black are represented the variants located in intron 1 and 2. The two black arrows point to the variations that are more likely to be responsible for the phenotype. Abbreviations: E1 - Exon 1; E2 - Exon 2; E3 - Exon 3; UTR – Untranslated region. The nomenclature used in **Figure 3** for sequence variations is according to den Dunnen & Antonarakis 2000.

3.2 Missense variants analysis

Initially, all four missense sequence variations could be considered as candidates to Glaucoma causing disease mutations. However, after database search, it was possible to determine that the c.227G>A (Arg76Lys) and the c.1193A>G (Lys398Arg) variants were previously described as neutral polymorphism, since their presence was also found in individuals without the Glaucoma phenotype, while the c.878C>A (Thr293Lys) and the c.1334C>T (Ala445Val) variants were described as Glaucoma causing disease mutations (http://www.myocilin.com/variants.php). These last two sequence variations are located in *MYOC* gene exon 3 and result from an alteration at the second nucleotide of both codon 293 (c.878C>A), changing a threonine by a lysine (Thr293Lys), and codon 445 (c.1334C>T), changing an alanine by a valine (Ala445Val) (Fig. 4). These missense variants were identified in two male patients with an average age of diagnosis of 76 years old and presenting the NTG subtype (Table VI).

Table VI. Clinical features of Glaucoma patients with *MYOC* variants

Glaucoma causing disease alteration	Subtype of Glaucoma	Age of diagnosis	Gender
c.878 C>A T293K	**NTG**	**75**	**M**
c.1334C>T A445V	**NTG**	**77**	**M**

The nomenclature used in **Table VI** for sequence variations is according to den Dunnen and Antonarakis, 2000

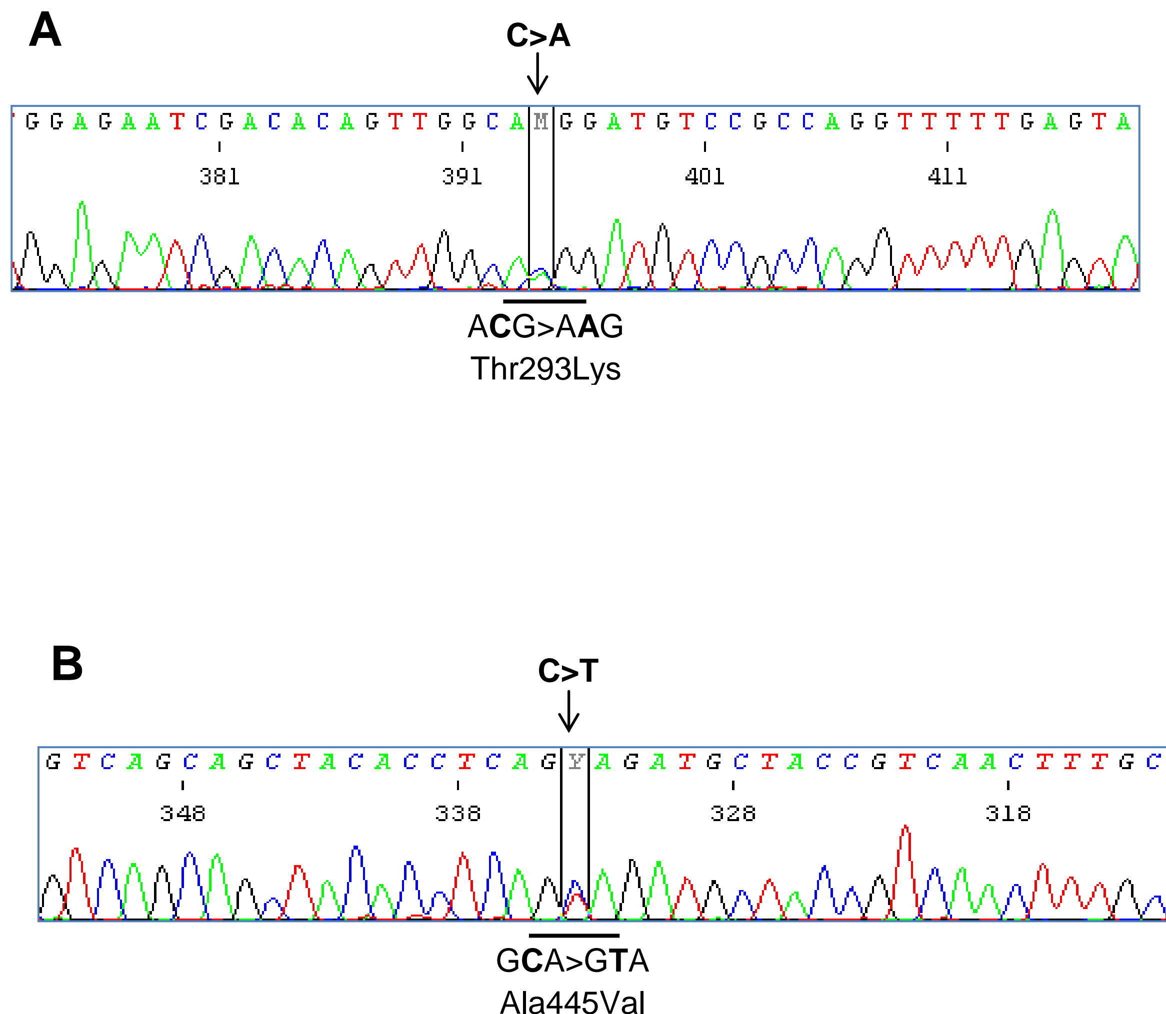


Figure 4. Electropherograms of the region surrounding the two variants candidates to Glaucoma causing disease mutations. A. Heterozygous transversion of a C to an A at nucleotide 878, changing the codon ACG to AAG and causing a Threonine to Lysine amino acid substitution at position 293. **B**. Heterozygous transition of a C to a T at nucleotide 1334, changing the codon GCA to GTA and causing an Alanine to Valine amino acid substitution at position 445. The nomenclature used in **Figure 4** for sequence variations is according to den Dunnen and Antonarakis, 2000

3.3 SNaPshot analysis

In order to understand if a variant is a mutation or a polymorphism it is necessary to determine its frequency in a group of normal individuals. Accordingly, a group of variants that were likely responsible for the POAG phenotype were selected and analyzed in a cohort of control subjects using a SNaPshot approach. Therefore, the missense variants in exon 1 (c.227G>A) and exon 3 (c.878C>A, c.1193A>G and c.1334C>T), and the promoter variants (c.-224T>C, c.-190G>T, c.-126T>C and c.-83G>A) were all selected and analyzed. The following electropherograms (Fig. 5) show the different genotypes (wild type, heterozygous and mutant) identified for each sequence variation selected. The complete genotyping results are shown in table VII. It is noticeable that control samples only presented the wild type genotype for variants c.-190G>T, c.878G>T and c.1334C>T, and that variants c.-224T>C, c.-83G>A and c.227G>A showed a reasonable heterozygosity.

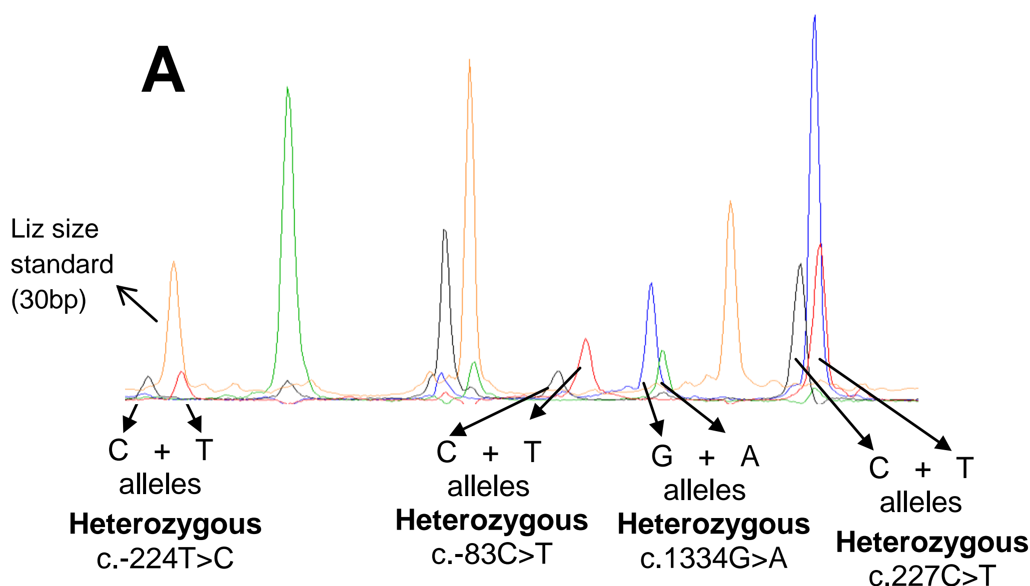
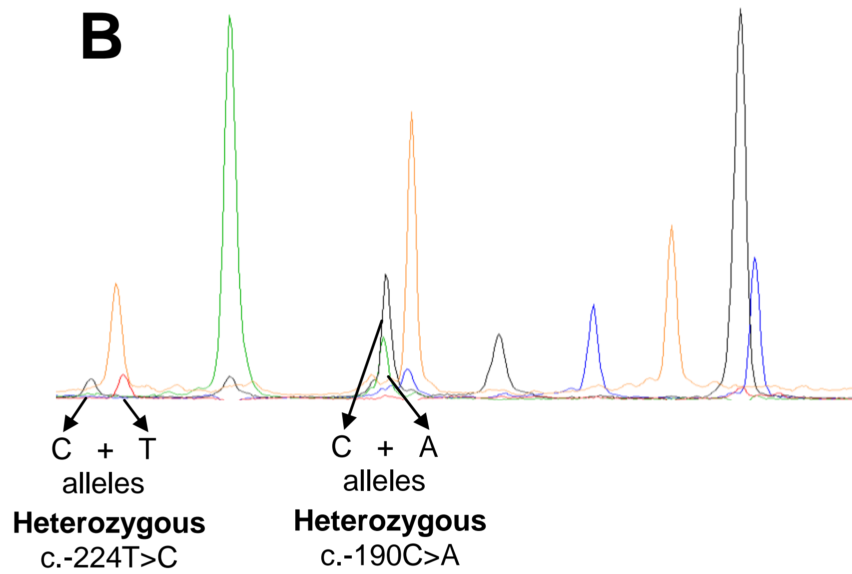


Figure 5. continues in the next page



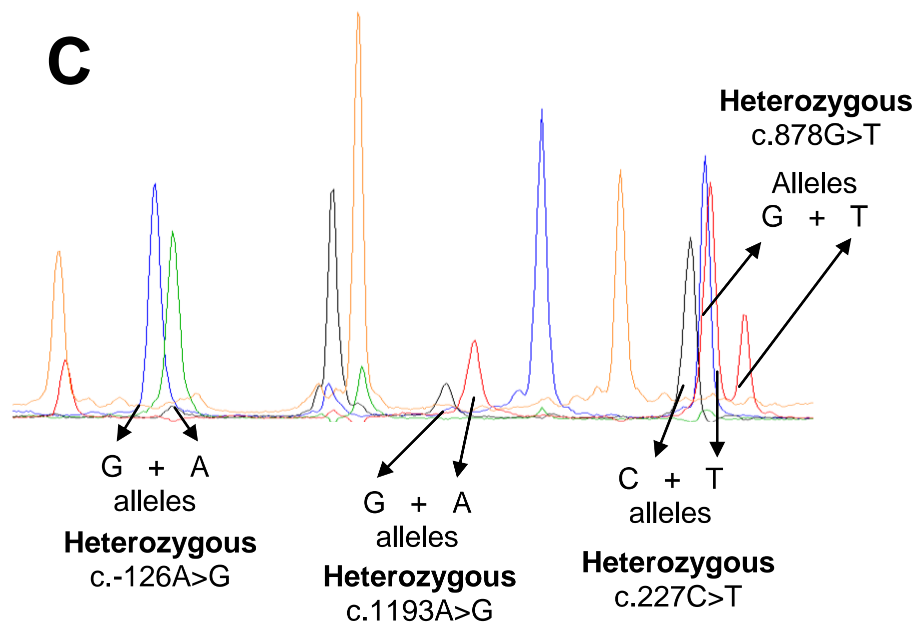


Figure 5. continues in the next page

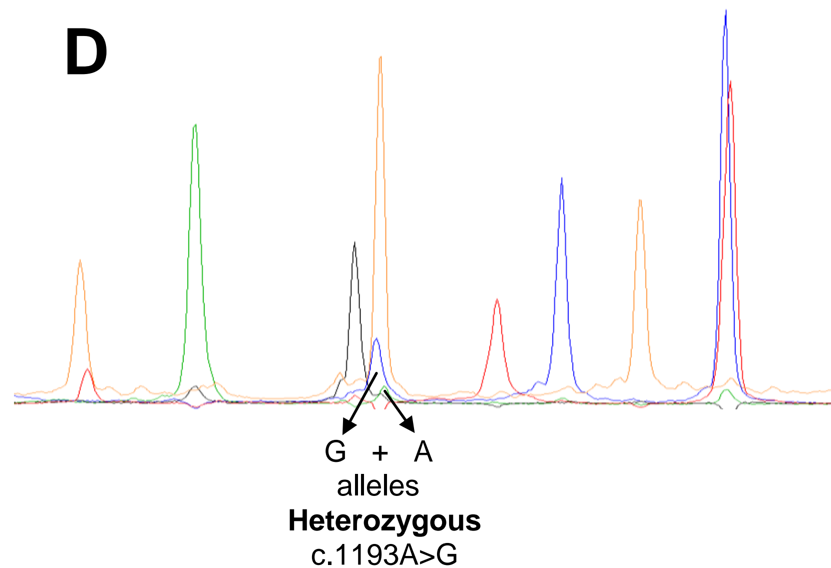


Figure 5. Electropherograms of the multiplex SNaPshot analysis for the selected genetic variants found in POAG patients. For SNaPshot analysis, variants were organized accordingly to the following order: c.-224T>C, c.-126T>C, c.-190G>T, c.1193A>G (Lys398Arg), c.-83G>A, c.1334C>T (Ala445Val), c.227G>A (Arg76Lys), and c.878G>T (Thr293Lys). To each SNaPshot reaction product was added a LIZ standard size (orange lines in electropherograms). **A.** Electropherogram with heterozygous variants c-224T>C, c.-83C>T, c1334G>A and c.227C>T, while the remaining variants are homozygous; **B.** Electropherogram with heterozygous variants c.-224T>C and c.-190G>T, while the remaining variants are homozygous; **C.** Electropherogram with heterozygous variants c.-126T>C, c.1193A>G (Lys398Arg), c.227G>A (Arg76Lys), and c.878G>T (Thr293Lys), while the remaining variants are homozygous; **D.** Electropherogram with heterozygous variant c.1193A>G (Lys398Arg), while the remaining variants are homozygous. The nomenclature used in **Figue 5** for sequence variations is according to den Dunnen and Antonarakis, 2000

Table VII. Control subjects genotypes			
Genetic variants	n = 54		
	wt	Het	Mut
c.-224T>C	31	18	5
c.-190G>T	54	-	-
c.-126T>C	48	6	-
c.-83G>A	38	15	1
c.227G>A (Arg76Lys)	38	15	1
c.878G>T (Thr293Lys)	54	-	-
c.1193A>G (Lys398Arg)	53	1	-
c.1334C>T (Ala445Val)	54	-	-

Abbreviations: wt - wild type; Het - heterozygotes; Mut - mutant
The nomenclature used in **Table VII** for sequence variations is according to den Dunnen and Antonarakis, 2000.

3.4 Case-control study with *MYOC* variants

Considering the previous SNaPshot results, a case-control study was performed aiming to evaluate if it is possible to establish an association between variants c.-224T>C, c.-126T>C, c.-83G>A and c.227G>A genotypes and/or alleles, and the disease phenotype. Tables VIII and IX summarize respectively, the genotypic and allelic frequencies, both for patients and controls, and the p value for the respective statistical analysis. Hardy-Weinberg equilibrium (HWE) was determined for all the variants groups of POAG and control samples and all the p values were not significant. None of the variants genotypes or alleles present a significant difference between patients and controls. Again, the promoter variant rs2075648 and the exon 1 alteration rs2234926 were always found simultaneously in the same control samples, sharing the same genotype, as previously seen in the POAG patients. This suggests that both variants are in linkage disequilibrium segregating the mutant allele always together.

Table VIII. Genotypes statistical analysis					
Variants	**Genotypes**	**Patients (n=104)**	**Controls (n=54)**	**p-value**	**HWE p-value Patients/Controls**
rs235920 c.-224T>C	TT TC CC	0.53 0.42 0.05	0.57 0.33 0.09	0.37	0.76/0.74
rs34928744 c.-126T>C	TT TC	0.96 0.04	0.89 0.11	0.08	1/1
rs2075648 c.-83G>A **rs2234926** c.227G>A Arg76Lys	GG GA AA	0.83 0.16 0.01	0.70 0.28 0.02	0.20	1/1

Abbreviation: HWE – Hardy-Weinberg Equilibrium
The nomenclature used in **Table VIII** for sequence variations is according to den Dunnen and Antonarakis, 2000.

Table IX. Alleles statistical analysis				
Variants	**Alleles**	**Patients (n=104)**	**Controls (n=54)**	**p**
rs235920 c.-224T>C	T C	0.74 0.26	0,74 0,26	0.99
rs34928744 c.-126T>C	T C	0.98 0.02	0,94 0,06	0.08
rs2075648 c.-83G>A **rs2234926** c.227G>A Arg76Lys	G A	0.91 0.09	0,84 0,16	0.08

The nomenclature used in **Table IX** for sequence variations is according to den Dunnen and Antonarakis, 2000.

Further analysis was performed taking into consideration the gender of the POAG patients and the controls. HWE was determined and no significant differences were found between patients and controls in both males and females for any of the variants. In the males group, the analyzed variants did not reveal any significant difference between patients and controls (Table X). However, in the females group, a significant difference was identified for the linkage disequilibrium variants rs2075648 and rs2234926 alleles, also presenting a borderline significant result with p value equal to 0.05 for the genotypes. The remaining variants did not present any significant difference (Table XI).

Table X. Males case-control study

Variants	Genotypes frequency	χ^2	HWE	Alleles frequency	χ^2
	Patients (n=52)/ Controls (n=28)	p-value	Patients/ Controls	Patients/ Controls	p-value
rs235920 c.-224T>C	TT 0.52/0.64 TC 0.44/0.25 CC 0.04/0.11	0.16	0.64/0.46	T 0.74/0.77 C 0.26/0.33	0.70
rs34928744 c.-126T>C	TT 0.98/0.86 TC 0.02/0.14	0.19	1/1	T 0.97/0.93 C 0.03/0.07	0.21
rs2075648 c.-83G>A **rs2234926** c.224G>A Arg76Lys	GG 0.77/0.75 GA 0.23/0.25	0.85	0.81/1	G 0.88/0.875 A 0.12/0.125	0.86

The nomenclature used in **Table X** for sequence variations is according to den Dunnen and Antonarakis, 2000

Table XI. Females case-control study					
Variants	**Genotypes frequency**	χ^2	**HWE**	**Alleles frequency**	χ^2
	Patients (n=52)/ Controls (n=26)	p-value	Patients/ Controls	Patients/ Controls	p-value
rs235920	TT 0.54/0.50 TC 0.40/0.42 CC 0.06/0.08	0.92	0.92/1	T 0.74/0.71 C 0.26/0.29	0.70
rs34928744	TT 0.98/0.92 TC 0.02/0.08	0.21	1/1	T 0.99/0.96 C 0.01/0.04	0.21
rs2075648 c.-83G>A **rs2234926** c.224G>A Arg76Lys	GG 0.88/0.65 GA 0.10/0.31 AA 0.02/0.04	0.05	0.51/1	G 0.93/0.81 A 0.07/0.19	0.02* OR=3.29

* Marks statistically significant result.
Abreaviatures: OR- Odds Ratio
The nomenclature used in **Table XI** for sequence variations is according to den Dunnen and Antonarakis, 2000

3.5 *In silico* analysis

Computational analysis is a strong tool frequently used to predict sequence variations pathogenicity. In this study the *in silico* analysis were performed for c.878G>T (Thr293Lys) and c.1334C>T (Ala445Val) genetic variants described as possible Glaucoma causing mutations and absent in controls, variant c.-190T>C because it also was not found in controls, and variants c.-83G>A and c.224G>A (Arg76Lys) since a statistically significant association was identified between Controls and the mutant alleles in female subjects.

3.5.1 Evolutionary conservation study

The evolutionary conservation study of the promoter variants (c.-190G>T and c.-83G>A) and missense variants [c.227G>A (Arg76Lys), c.878G>T (Thr293Lys) and c.1334C>T (Ala445)] comparing the primate and non-primate species with the human nucleotide and amino acid sequences surrounding those *loci* was determined.

Concerning the promoter variant c.-190G>T all the species share the same nucleotide (Fig. 6). On the opposite, for the c.-83G>A promoter variant all the non-primates species and the *Pan troglodytes* present an A instead of a G (Fig. 7), while all the other primate species share the G, the same nucleotide as wild type humans (Fig. 7A).

The evolutionary conservation study of the missense variant c.227G>A (Arg76Lys) allowed to determine that all the species share the same nucleotide and amino acid with the exception of *Macaca mulata* that instead of an A presents a G causing an amino acid change from an arginine to a lysine (Fig. 8A), and *Danio rerio* that does not show any nucleotide or amino acid sequence for the region surrounding that *locus* (Fig. 8B).

The evolutionary conservation study of the missense variant c.878G>T (Thr293Lys) allowed to determine that all the species share the same nucleotide and the same amino acid with the exception of *Danio rerio* that instead of an amino acid threonine presents a serine (Fig. 9).

The evolutionary conservation study of the missense variant c.1334C>T (Ala445Val) determined that all the species share the same nucleotide and the same amino acid with the exception of *Rattus norvegicus* that instead of a C presents a T causing an amino acid change from an alanine to a valine (Fig. 10A) and *Danio rerio* that instead of an amino acid alanine presents a proline (Fig. 10B).

A

```
Patient(c.-190G>T)          AATCTTGCTGGCAGCTTGAAGGCAACCCCCCTG
Homo sapiens                AATCTTGCTGGCAGCGTGAAGGCAACCCCCCTG
Pan troglodytes             AATCTTGCTGGCAGAGTGAAGGCAACCCCCCTG
Gorilla gorilla             AATCTTGCTGGCAGCGTGAAGGCAACCCCCCTG
Macaca mulatta              AATCTTGCTGGCAGTGTGAAGGCAACCCCCCTG
Pongo abelii                AATCTTGCTGGCAGCGTGAAGGCAACCCCCCTG
```

B

```
Patient(c.-190G>T)          AATCTTGCTGGCAGCTTGA-AGG--CAACCCCCCTG
Homo sapiens                AATCTTGCTGGCAGCGTGA-AGG--CAACCCCCCTG
Mus musculus                --TCTTGCTGGCAGTGTGAG---TGTAATCCTCCTA
Ratus norvegicus            AATCTTGCTGGCAGTGTGAG---TCCAATCCTCCTA
Danio rerio                 A------CT--CAG-GTGAGAGGTC--AT-----TA
```

Figure 6. Evolutionary conservation study for c.-190G>T variant. A. Evolutionary conservation of c.-190G>T variant in primates. **B.** Evolutionary conservation of c.-190G>T variant in non-primates. The nomenclature used in **Figure 6** for sequence variation is according to den Dunnen and Antonarakis, 2000.

A

```
Patient (c.-83G>A)        ACCTCTCTGGAGCTCAGGCATGAGCCAGCAAGG
Homo sapiens              ACCTCTCTGGAGCTCGGGCATGAGCCAGCAAGG
Pan Troglodytes           ACCTCTCTGGAGCTCAGGCATGAGCCAGCAAGG
Gorilla gorilla           GCCTCTCTGGAGCTCGGGCATGAGCCAGCAAGG
Macaca mulatta            ACCTCTCTGGAGCTCGGGCATGAGCCAGCAAGG
Pongo abelii              ACCTCTCTGGAGCTCGGGCACGAGCCAGCAAGG
```

B

```
Patient (c.-83G>A)        ACCTCTCTGGAGCTCA-GGCA---TGAGCCAG-CAA-GG
Homo sapiens              ACCTCTCTGGAGCTCG-GGCA---TGAGCCAG-CAA-GG
Mus musculus              ATGTCTTTGGACTTCA-GGC---TTGAGCCAG-CAG-GG
Ratus norvegicus          ACCTCTTTGGATTTCA-GGC---TTGAGCCAG-CAG-GG
Danio rerio               ACCT---TGGA---CAAGG-AGCTT----CAGGCAACAG
```

Figure 7. Evolutionary conservation study for c.-83G>A variant. A. Evolutionary conservation of c.-83G>A variant in primates. **B.** Evolutionary conservation of c.-83G>A variant in non-primates. The nomenclature used in **Figure 7** for sequence variation is according to den Dunnen and Antonarakis, 2000.

A

Patient(c.227G>A)	ATCCATAACTTACAGA**A**AGACAGCAGCACCCAA	IHNLQ**K**DSSTQ
Homo sapiens	ATCCATAACTTACAGA**G**AGACAGCAGCACCCAA	IHNLQ**R**DSSTQ
Pan troglodytes	ATCCATAACCTACAGA**G**AGACAGCAGCACCCAA	IHNLQ**R**DSSTQ
Gorilla gorilla	ATCCATAACCTACAGA**G**AGACAGCAGCACCCAA	IHNLQ**R**DSSTQ
Macaca mulatta	ATCCATAACCTACAGA**A**AGACAGCAGCACCCAG	IHNLQ**K**DSSTQ
Pongo abelii	ATCCGTAACCTACAGA**G**AGACAGCAGCACCCAA	IRNLQ**R**DSSTQ

B

Patient(c.227G>A)	ATCCATAACTTACAGA**A**AGACAGCAGCACCCAA	IHNLQ**K**DSSTQ
Homo sapiens	ATCCATAACTTACAGA**G**AGACAGCAGCACCCAA	IHNLQ**R**DSSTQ
Mus musculus	ATCCAAGACCTTCAGA**G**AGACAGCAGCATCCAG	IQDLQ**R**DSSIQ
Rattus norvegicus	ATCCAGGACCTTCAGA**G**AGATAGCAGCATCCAG	IQDLQ**R**DSSIQ
Danio rerio	---------------------------------	-----------

Figure 8. Evolutionary conservation study for c.227G>A (Arg76Lys) variant. A. Evolutionary conservation of c.227G>A (Arg76Lys) variant in primates. **B.** Evolutionary conservation of c.227G>A (Arg76Lys) variant in non-primates. The nomenclature used in **Figure 8** for sequence variation is according to den Dunnen and Antonarakis, 2000.

.

A

Patient (c.878C>A)	ATCGACACAGTTGGCA**A**GGATGTCCGCCAGGTT	IDTVG**K**DVRQV
Homo sapiens	ATCGACACAGTTGGCA**C**GGATGTCCGCCAGGTT	IDTVG**T**DVRQV
Pan troglodytes	ATCGACACAGTTGGCA**C**AGATGTCCGCCAGGTT	IDTVG**T**DVRQV
Gorilla gorilla	ATCGACACAGTTGGCA**C**AGATGTCCGCCAGGTT	IDTVG**T**DVRQV
Macaca mulatta	ATTGACACAGTTGGCA**C**AGATGTCCGCCAGGTT	IDTVG**T**DVRQV
Pongo abelii	ATTGACACAGTTGGCA**C**AGATGTCCGCCAGGTT	IDTVG**T**DVRQV

B

Patient (c.878C>A)	ATCGACACAGTTGGCA**A**GGATGTCCGCCAGGTT	IDTVG**K**DVRQV
Homo sapiens	ATCGACACAGTTGGCA**C**GGATGTCCGCCAGGTT	IDTVG**T**DVRQV
Mus musculus	ATTGACACGGTTGGCA**C**AGAGATCCGCCAGGTG	IDTVG**T**EIRQV
Rattus norvegicus	ATTGACACGGTTGGCA**C**AGGCATCCGCCAGGTG	IDTVG**T**GIRQV
Danio rerio	ATTGATTCTGTCGGTT**C**TGAAGTGCGTCAACTC	IDSVG**S**EVRQL

Figure 9. Evolutionary conservation study for c.878C>A (Thr293Lys) variant. A. Evolutionary conservation of c.878C>A (Thr293Lys) variant in primates. **B.** Evolutionary conservation of c.878C>A (Thr293Lys) variant in non-primates. The nomenclature used in **Figure 9** for sequence variation is according to den Dunnen and Antonarakis, 2000.

A

Patient (c.1334C>T)	AGCAGCTACACCTCAG**T**AGATGCTACCGTCAAC	SSYTS**V**DATVN
Homo sapiens	AGCAGCTACACCTCAG**C**AGATGCTACCGTCAAC	SSYTS**A**DATVN
Pan troglodytes	AGCAGCTACTCCTCAG**C**AGATGCTACCGTCAAC	SSYSS**A**DATVN
Gorilla gorilla	AGCAGCTACTCCTCAG**C**AGATGCTACCATCAAC	SSYSS**A**DATIN
Macaca mulatta	AGCAGCTACTCCTCAG**C**AGATGCAACCGTCAAC	SSYSS**A**DATVN
Pongo abelii	AGCAGCTACTCCTCAG**C**AGATGCTACCATCAAC	SSYSS**A**DATIN

B

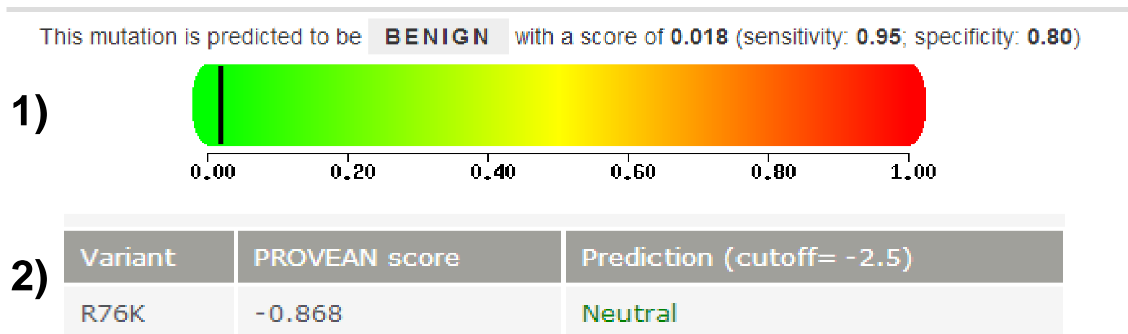
Patient (c.1334C>T)	AGCAGCTACACCTCAG**T**AGATGCTACCGTCAAC	SSYTS**V**DATVN
Homo sapiens	AGCAGCTACACCTCAG**C**AGATGCTACCGTCAAC	SSYTS**A**DATVN
Mus musculus	AGCAGCTACTCTTCAG**C**CCATGCAACCGTCAAC	SSYSS**A**HATVN
Rattus norvegicus	AGCAGCTACTCTTCAG**T**CCATGCAACCATCAAC	SSYSS**V**HATIN
Danio rerio	GCTAGTTACACCTCAC**C**AAACACCACGGTTAAT	ASYTS**P**NTTVN

Figure 10. Evolutionary conservation study for c.1334C>T (Ala445Val) variant. A. Evolutionary conservation of c.1334C>T (Ala445Val) variant in primates. **B.** Evolutionary conservation of c.1334C>T (Ala445Val) variant in non-primates. The nomenclature used in **Figure 10** for sequence variation is according to den Dunnen and Antonarakis, 2000.

3.5.2 Pathogenicity prediction

The pathogenicity prediction indicated that variants c.227G>A (Arg76Lys), c.878C>A (Thr293Lys) and c.1334C>T (Ala445Val) are all benign, according to Polyphen-2® with scores 0.018, 0.251 and 0.005, respectively (Fig. 11), and neutral, according to PROVEAN® with scores -0.868, -0.410 and -1.605, respectively (Fig. 11).

A



B

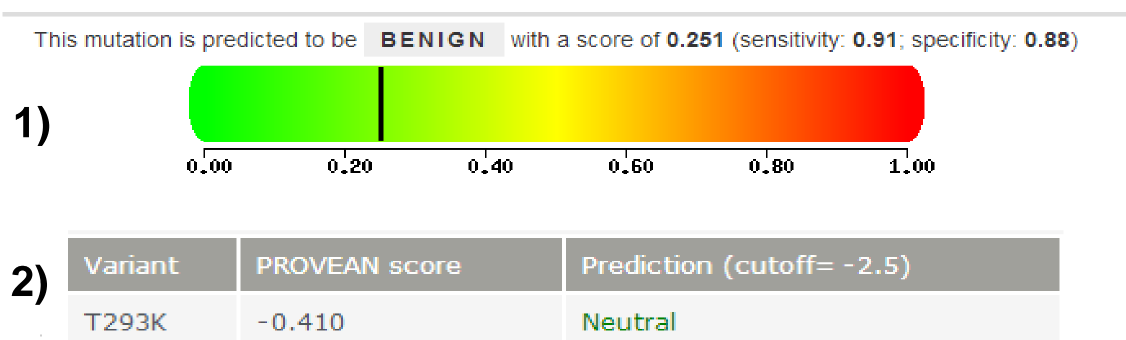
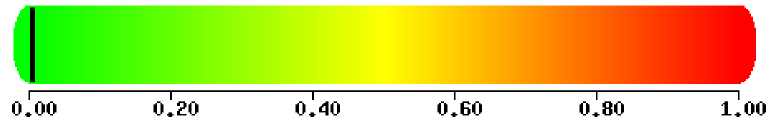


Figure 11. continues in the next page

C

This mutation is predicted to be **BENIGN** with a score of **0.005** (sensitivity: **0.97**; specificity: **0.74**)

1)



2)

Variant	PROVEAN score	Prediction (cutoff= -2.5)
A445V	-1.605	Neutral

Figure 11. Pathogenicity analysis of Arg76Lys, Thr293Lys and Ala445Val genetic variants. Polyphen-2® analysis for variants Arg76Lys (**A.1**), Thr293Lys (**B.1**) and Ala445Val (**C.1**); PROVEAN analysis for variant Arg76Lys (**A.2**), Thr293Lys (**B.2**) and Ala445Val (**C.2**).

3.5.3 Transcription factor study

Since variants c.-190T>C (rs76745622) and c.-83G>A (rs2075648) are located in *MYOC* gene promoter, it seemed to be important to evaluate if any transcription factor binding site could be altered by the mutant allele, changing the affinity of the transcription factor for its promoter binding site and modifying the *MYOC* gene expression regulation. The complex AhR/Ar transcription factor is predicted to bind *MYOC* gene promoter at the *locus* surrounding the c.-190G>T with a score of 79.0 for allele G, but with a lower score (68.3) when it is present the allele T (Table XII). The GATA-1 transcription factor is predicted to bind *MYOC* gene promoter at the *locus* surrounding the c.-83G>A with a score of 71.8 for allele G, but with a lower score (62.9) when it is present the allele A (Table XIII).

Table XII. Transcription factor study for variant c.-190G>T

Nucleotide sequence	TF	Sc
ATTTTCTAAGAATCTTGCTGGCAGCGTGAAGGCAACCCCCCTGTGCACA --------------------	AhR/Ar	79.0
ATTTTCTAAGAATCTTGCTGGCAGCTTGAAGGCAACCCCCCTGTGCACA --------------------	AhR/Ar	68.3

Abbreviations: TF= Transcription factor; SC=Score

The nomenclature used in **Table XII** for sequence variations is according to den Dunnen and Antonarakis, 2000.

Table XIII. Transcription factor study for variant c.-83G>A

Nucleotide sequence	TF	Sc
GTATATATAAACCTCTCTGGAGCTCGGGCATGAGCCAGCAAGGCCACCCA ----------	GATA-1	71.8
GTATATATAAACCTCTCTGGAGCTCAGGCATGAGCCAGCAAGGCCACCCA ----------	GATA-1	62.9

Abbreviations: TF= Transcription factor; SC=Score

The nomenclature used in **Table XIII** for sequence variations is according to den Dunnen and Antonarakis, 2000.

4. Discussion

Over the years, more than 80 mutations in *MYOC* gene have been identified as responsible for primary open angle glaucoma (Zhuo et al. 2008). The purpose of this study was to identify *MYOC* gene sequence variations that could be responsible for POAG phenotype in a group of Portuguese patients.

The analysis of 104 POAG patients and 54 controls was consistent with previous studies that determined a worldwide frequency of 2-4% of POAG patients with *MYOC* mutations (Weisschuh et al. 2005), and that the majority of them (>90%) are clustered in olfactomedin-like domain, the most conserved myocilin protein domain, encoded by genes' exon (Kanagavalli et al. 2003; Lopez-Martinez et al. 2007; Braghini et al. 2013).The present study allowed to identify two Glaucoma causing disease mutations in two NTG patients (approximately 2% of the POAG Portuguese patients analyzed) (Table VI), absent in the controls group (Table VII) and both in *MYOC* gene exon 3 (Fig. 2), probably changing the myocilin protein olfactomedin-like domain (Thr293Lys and Ala445Val). Such mutations were previously reported in two NTG patients carrying each one of them (Weisschuh et al. 2005; Lopez-Martinez et al. 2007). The evolutionary conservation study of c.878C>A (Thr293Lys) revealed that the nucleotide sequence as well as the adjacent regions and correspondent amino acids are highly conserved among primates and non-primates (Fig. 9). This highly conserved polypeptide chain suggests that it is required its full preservation for maintaining the normal protein structure or function, meaning that changes on this region likely have a negative impact on myocilin protein and may contribute for POAG phenotype. Additionally, the substitution of a threonine for a lysine may also have some deleterious effect on protein features considering that the second amino acid is basic polar and have a strong positive charge while the first is polar uncharged (Alward et al. 1998). Concerning the secondary structure of the protein, the alteration of a threonine (β-former) to a lysine (β-breaker) may influence the protein to adopt a helical conformation instead of their normal β-sheet (Chou & Fasman 1978), possibly contributing for the POAG phenotype.

The nucleotide and the amino acid *locus* surrounding the c.1334C>T (Ala445Val) are highly conserved among primates but moderately (nucleotide) and weakly (amino acid) conserved among non-primates (Fig. 10) meaning that, among primates, changes in this *locus* likely cause changes on the protein structure and function that may be associated with the POAG etiology. However, the substitution of an alanine by a valine unlikely cause alterations on the protein properties since both are non-polar and hydrophobic amino acids (Alward et al. 1998). Additionally, considering the protein secondary structure, the Ala445Val mutation changes an alanine, which is a strong former α and an indifferent β amino acid, to a valine, which is a former α and a strong former β amino acid, possibly causing an increased preference for a β-sheet conformation (Chou & Fasman 1978), but without any clear influence on the POAG phenotype.

The bioinformatics study with Polyphen-2® and PROVEAN software considered both c.878C>A (Thr293Lys) and c.1334C>T (Ala445Val) variants benign and neutral polymorphisms' respectively (Fig. 11). However, these *MYOC* variants may alter the affinity of the protein for its extracellular ligands changing the extracellular functions of the protein (Gobeil et al. 2006). Also some authors reported that mutations in olfactomedin-like domain cause misfolded accumulated myocilin and homodimers formation, as well as heterodimers with wt myocilin, avoiding its secretion from the endoplasmic reticulum (ER), and leading the HTM cells death (Joe et al. 2003; Gobeil et al. 2004; Liu & Vollrath 2004). These facts suggest that these two Glaucoma causing disease mutations could be crucial in a so far unknown physiological role of myocilin. Nevertheless, further investigations are required to evaluate the pathogenicity of these sequence variations.

Since 1997, mutations in *MYOC* gene have been mainly associated with JOAG, a subtype of Glaucoma with very high IOP (Ojha et al. 2013). However, the two Glaucoma causing disease mutations [c.878C>A (Thr293Lys) and c.1334C>T (Ala445Val)] so far identified in the Portuguese POAG population were found in two patients clinically classified with the NTG subtype of Glaucoma, a similar result to previous publications in German and Spanish populations (Weisschuh et al. 2005; Lopez-Martinez et al. 2007).

This fact could indicate that mutations in *MYOC* gene alone may not be enough to cause ocular hypertension, requiring other genetic, biological or environmental factors to promote the IOP raise. This may suggest that *MYOC* gene mutations are not responsible for Glaucoma by raising the IOP, with corresponding compression of the optic nerve at lamina crivosa, damaging adjacent neuronal cells and leading to RGC apoptosis, but instead through another unidentified mechanism IOP independent and associate to NTG etiology.

Additionally, a less reported promoter variant, the c.-190G>A, was identified in this study merely in one POAG patient and it was not found in the control subjects (Table VII). This result is analogous to what was observed in an Indian population (Banerjee et al. 2012). However, another study reported the c.-190G>A variant in one control subject (Lopez-Martinez et al., 2007), suggesting that it has no relevance for the disease establishment. Nevertheless, the evolutionary conservation study shows a high conservation of this nucleotide among all the studied species but a low conservation of the adjacent regions in the non-primate species (Fig. 6), likely meaning that this variant may have a significant role in the regulation of *MYOC* gene expression. Accordingly, specific transcription factors binding to that promoter *locus* may be compromised by this variant. Bioinformatic analysis allowed to determine that the TF complex aryl hydrocarbon receptor (AhR/AR) has high affinity to bind to the c.-190G>A *locus* (Table XII). This transcription factor is responsible for regulate several developmental and physiological events, which could be important provided that c.-190T>C variant belongs to a negative glucocorticoid response element (nGRE) (Pocar et al. 2005; Banerjee et al. 2012). Another interesting feature of this TF consists on its ability to regulate enzymes responsible for metabolizing xenobiotics, such as CYP1B1, a protein expressed by another Glaucoma associated gene. The affinity of this TF decrease in the presence of the mutant allele of this variant, which may negatively influence the transcription of the *MYOC* gene with possible consequences to the disease etiology.

It was also found a promoter variant (c.-83G>A) and an exon 1 alteration [c.227G>A (Arg76Lys)] sharing the same genotype simultaneously in patients and controls samples and suggesting that both variants are in linkage disequilibrium, segregating the mutant allele always together. This phenomena was also identified in Japanese, Chinese and Spanish populations (Mabuchi et al. 2001; Pang et al. 2002; Lopez-Martinez et al. 2007). So far, no association of these two variants in linkage disequilibrium with POAG phenotype has been identified. Nonetheless, in this study was observed a higher frequency of the mutant allele in the female controls than in the POAG female patients (Table XI). This data suggest some protective effect of this allele in the development of the disease, but further analyses are necessary to confirm or refuse this hypothesis. The evolutionary conservation study regarding the c.-83T>C variant shows a moderate conservation among primates and a high conservation among non-primates, but no conservation when the primates are compared with the non-primates, since almost all the studied primate species share a G while all the non-primate species share an A (Fig. 7). Considering this, it is likely that changes in this *locus* alone are not associated with the affected phenotype. *In silico* analysis revealed a GATA-1 transcription factor (TF) with affinity to bind at the *locus* including the c.-83T>C variant (Table XIII). The affinity of this TF decrease in the presence of the mutant allele of this variant (Table XIII), which may negatively influence the transcription of the *MYOC* gene with possible consequences to the disease etiology.

The evolutionary conservation study for variant c.227G>A (Arg76Lys) revealed a moderate conservation in both primate and non-primate species (Fig. 8). Nevertheless, it is noticeable that *Danio rerio* does not present any nucleotide or amino acid sequence for the analyzed *locus* (Fig. 8B). This may indicate that this *locus* have an important role for structural and functional myocilin protein in mammalians, with possible implications for the disease etiology in human, and it is completely unnecessary for fishes myocilin protein. Considering the bioinformatics study with Polyphen-2® and PROVEAN software the c.227G>A (Arg76Lys) variant was classified as benign and neutral polymorphism,

respectively (Fig. 11). Furthermore, the substitution of an arginine by a lysine unlikely cause alterations on the protein properties since both amino acids are positively charged, basic polar and have hydrophilic features. Additionally, considering the protein secondary structure, the Arg76Lys variant changes an arginine, which is an indifferent α-helix and β-sheet amino acid, to a lysine, which is an indifferent α and a β breaker amino acid (Chou & Fasman 1978), possibly without any consequences or, in the worse scenario if it is considered that the arginine is localized in a β-sheet, probably destabilizing the late conformation. Nevertheless, it is always possible that this alteration acts as a protective variant for POAG due to an unknown mechanism.

A further missense variant [c.1193G>A (Lys398Arg)] was found in both patients and controls (Table V and VII). This may explain the neutral polymorphism classification attributed to this variant (http://www.myocilin.com/variants.php). Accordingly, both lysine and arginine are positively charged, basic polar amino acids with hydrophilic features. Additionally, considering the protein secondary structure, the Lys398Arg variant changes a lysine to an arginine, the same amino acids as for Arg76Lys, with a similar situation that lysine inhibits the β-sheet conformation and enables the α-helix, which is structurally maintained by arginine amino acid. Considering this it is likely that variant c.1193G>A (Lys398Arg) is a polymorphism without any association with POAG.

Two other variants (c.-224T>C and c.-126T>C) were found in the *MYOC* gene promoter in both patients and controls (Table V and VII) but no statistically significant association was identified with the phenotype (Table VIII and IX), which leads to the conclusion that are polymorphisms without any influence on POAG etiology.

The sequencing analysis also allowed to find 7 synonymous variants, three of them located in exon 1 [c.39T>G (Pro13Pro), c.141C>T (Cys47Cys) and c.477A>G (Leu159Leu)], a single one in exon 2 [c.612G>T (Thr204Thr)] and the remaining in exon 3 [c.855G>T (Thr285Thr), c.975G>A (Thr325Thr) and c.1041T>C (Tyr347Tyr)] (Table V). These variants were also previously reported both in POAG patients and controls (Alward et al. 1998; Fingert et al. 1999; Lopez-Martinez et al. 2007) with the exception of

c.141C>T (Cys47Cys) that, as much as it is known, was identified for the first time in this work. These variants frequencies do not seem to have any significant difference from the online free databases for Caucasian populations (http://www.ncbi.nlm.nih.gov/snp/). Altogether, these data suggest that these variants are very much unlikely responsible for the POAG phenotype.

Several variations of sequence were detected within intron 1 (605-604G>A, c.604+177G>A, c.605-374G>C, c.605-332G>A, c.605-302C>G and c.605-280G>T) and intron 2 (c.730+35A>G, c.730+138G>A, c.730+176delCT, c.731-205 A>C, c.731-192 G>A and c.731-73 C>T). From a total of 12 intronic variants, two were novel (c.730+138G>A and c.731-192 G>A) (Table V). As these variants are located in non-coding regions (Fig. 2) it is unlikely that an association with the phenotype may occur.

5.Conclusion

This study allows the identification of genetic variants in *MYOC* gene in a cohort of POAG patients from the Portuguese population.

Two variants are likely Glaucoma causing disease mutations [c.878C>A (Thr293Lys) and c.1334C>T (Ala445Val)] and were identified in two male patients who clinically present as NTG. None of the control subject has any of these variants, supporting the opinion that these variants are responsible for POAG in those patients.

For the first time a statistically significant association was found among females and the disease. In this study, a control-case study shows that females have an increased protection (OR=3.29) against POAG due to variants c.-83G>A and c.227G>A (Arg76Lys).

In this study, a previously reported sequence variation (Lopez-Martinez et al. 2007; Banerjee et al. 2012) located in promoter region of the gene (c.-190T>C) was present only in one POAG patient and none controls, indicating a possible association with the phenotype in the Portuguese population.

Further studies are required to clarify the possible involvement of the identified variants in POAG etiology. Also, functional evaluation of the impact of these sequence variations on myocilin may be important to understand the role of this protein in Glaucoma.

6. References

Adam, M.F. et al., 1997. Recurrent mutations in a single exon encoding the evolutionarily conserved olfactomedin-homology domain of TIGR in familial open-angle glaucoma. *Human molecular genetics*, 6(12), pp.2091–7. Available at: http://www.ncbi.nlm.nih.gov/pubmed/9328473.

Alward, W.L. et al., 1998. Clinical features associated with mutations in the chromosome 1 open-angle glaucoma gene (GLC1A). *The New England journal of medicine*, 338(15), pp.1022–7. Available at: http://www.ncbi.nlm.nih.gov/pubmed/9535666.

Angius, A. et al, 1998. A Novel Mutation in the GLC1A Gene Causes Juvenile Open-angle Glaucoma in 4 Families From the Italian Region of Puglia. , 116(June), pp.793–797.

Ausubel, F.M. et al., Preparations and analyis of DNA. In: *Current Protocols in Molecular Biology*, John Wiley & Sons, Inc 2003; Chapter 2: 1–12.

Banerjee, D. et al., 2012. Comprehensive analysis of myocilin variants in east Indian POAG patients. *Molecular vision*, 18(March), pp.1548–57. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3380904&tool=pmcentrez&rendertype=abstract.

Bayat, B. et al., 2008. Contributions of MYOC and CYP1B1 mutations to JOAG. *Molecular vision*, 14(December 2007), pp.508–17. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2268862&tool=pmcentrez&rendertype=abstract.

Braghini, C.A. et al., 2013. New mutation in the myocilin gene segregates with juvenile-onset open-angle glaucoma in a Brazilian family. *Gene*, 523(1), pp.50–7. Available at: http://www.ncbi.nlm.nih.gov/pubmed/23566828.

Chou, P.Y. & Fasman, G.D., 1978. Empirical predictions of protein conformation. *Annual review of biochemistry*, 47, pp.251–76. Available at: http://www.ncbi.nlm.nih.gov/pubmed/354496.

Coleman, A.L., 2000. Glaucoma. , 354, pp.1803–1810.

Den Dunnen, J.T. & Antonarakis, S.E., 2000. Mutation nomenclature extensions and suggestions to describe complex mutations: a discussion. *Human mutation*, 15(1), pp.7–12. Available at: http://www.ncbi.nlm.nih.gov/pubmed/10612815.

Ennis, S. et al., 2010. Prevalence of myocilin gene mutations in a novel UK cohort of POAG patients. *Eye (London, England)*, 24(2), pp.328–33. Available at: http://www.ncbi.nlm.nih.gov/pubmed/19407846.

Fan, B.J. et al., 2006. Novel myocilin mutation in a Chinese family with juvenile-onset open-angle glaucoma. *Archives of ophthalmology*, 124(1), pp.102–6. Available at: http://www.ncbi.nlm.nih.gov/pubmed/16401791.

Fan, B.J. & Wiggs, J.L., 2010. Glaucoma: genes , phenotypes , and new directions for therapy. , 120(9).

Faucher, M. et al., 2002. Founder TIGR/myocilin mutations for glaucoma in the Québec population. *Human molecular genetics*, 11(18), pp.2077–90. Available at: http://www.ncbi.nlm.nih.gov/pubmed/12189160.

Fingert, J.H. et al., 1999. Analysis of myocilin mutations in 1703 glaucoma patients from five different populations. *Human molecular genetics*, 8(5), pp.899–905. Available at: http://www.ncbi.nlm.nih.gov/pubmed/10196380.

Fingert, J.H. et al., 2002. Myocilin Glaucoma. , 47(6), pp.20–22.

Fingert, J.H., 2011. Primary open-angle glaucoma genes. *Eye (London, England)*, 25(5), pp.587–95. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3171270&tool=pmcentrez&rendertype=abstract.

Fuse, N., 2010. Genetic bases for glaucoma. *Tohoku journal of experimental medicine*, 37, pp.1–10. Available at: http://scholar.google.com/scholar?hl=en&btnG=Search&q=intitle:Genetic+Bases+for+Glaucoma#0.

Gemenetzi, M., Yang, Y. & Lotery, a J., 2012. Current concepts on primary open-angle glaucoma genetics: a contribution to disease pathophysiology and future treatment. *Eye (London, England)*, 26(3), pp.355–69. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3298987&tool=pmcentrez&rendertype=abstract.

Gobeil, S. et al., 2004. Intracellular sequestration of hetero-oligomers formed by wild-type and glaucoma-causing myocilin mutants. *Investigative ophthalmology & visual science*, 45(10), pp.3560–7. Available at: http://www.ncbi.nlm.nih.gov/pubmed/15452063.

Gobeil, S., Letartre, L. & Raymond, V., 2006. Functional analysis of the glaucoma-causing TIGR/myocilin protein: integrity of amino-terminal coiled-coil regions and olfactomedin homology domain is essential for extracellular adhesion and secretion. *Experimental eye research*, 82(6), pp.1017–29. Available at: http://www.ncbi.nlm.nih.gov/pubmed/16466712.

Gray, K. a et al., 2013. Genenames.org: the HGNC resources in 2013. *Nucleic acids research*, 41(Database issue), pp.D545–52. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3531211&tool=pmcentrez&rendertype=abstract.

Jacobson, N. et al., 2001. Non-secretion of mutant proteins of the glaucoma gene myocilin in cultured trabecular meshwork cells and in aqueous humor. *Human molecular genetics*, 10(2), pp.117–25. Available at: http://www.ncbi.nlm.nih.gov/pubmed/11152659.

Janssen, S.F. et al., 2013. The vast complexity of primary open angle glaucoma: disease genes, risks, molecular mechanisms and pathobiology. *Progress in retinal and eye research*, 37, pp.31–67. Available at: http://www.ncbi.nlm.nih.gov/pubmed/24055863.

Joe, M.K. et al., 2003. Accumulation of mutant myocilins in ER leads to ER stress and potential cytotoxicity in human trabecular meshwork cells. *Biochemical and biophysical research communications*, 312(3), pp.592–600. Available at: http://www.ncbi.nlm.nih.gov/pubmed/14680806.

Kanagavalli, J. et al., 2004. A review of genetic and structural understanding of the role of myocilin in primary open angle glaucoma. *Indian journal of ophthalmology*, 52(4), pp.271–80. Available at: http://www.ncbi.nlm.nih.gov/pubmed/15693317.

Kanagavalli, J. et al., 2003. Evaluation and understanding of myocilin mutations in Indian primary open angle glaucoma patients. *Molecular vision*, 9(August), pp.606–14. Available at: http://www.ncbi.nlm.nih.gov/pubmed/14627955.

Kirstein, L. et al., 2000. Regulation of human myocilin/TIGR gene transcription in trabecular meshwork cells and astrocytes: role of upstream stimulatory factor. *Genes to cells : devoted to molecular & cellular mechanisms*, 5(8), pp.661–76. Available at: http://www.ncbi.nlm.nih.gov/pubmed/10947851.

Kubota, R. et al., 1997. A novel myosin-like protein (myocilin) expressed in the connecting cilium of the photoreceptor: molecular cloning, tissue expression, and chromosomal mapping. *Genomics*, 41(3), pp.360–9. Available at: http://www.ncbi.nlm.nih.gov/pubmed/9169133.

Liu, Y. & Vollrath, D., 2004. Reversal of mutant myocilin non-secretion and cell killing: implications for glaucoma. *Human molecular genetics*, 13(11), pp.1193–204. Available at: http://www.ncbi.nlm.nih.gov/pubmed/15069026.

Lopez-Martinez, F. et al., 2007. Role of MYOC and OPTN sequence variations in Spanish patients with primary open-angle glaucoma. *Molecular vision*, 13(March), pp.862–72. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2770203&tool=pmcentrez&rendertype=abstract.

Mabuchi, F. et al., 2001. Analysis of myocilin gene mutations in Japanese patients with normal tension glaucoma and primary open-angle glaucoma. *Clinical genetics*, 59(4), pp.263–8. Available at: http://www.ncbi.nlm.nih.gov/pubmed/11298682.

Menaa, F. et al., 2011. Keeping an eye on myocilin: a complex molecule associated with primary open-angle glaucoma susceptibility. *Molecules (Basel, Switzerland)*, 16(7), pp.5402–21. Available at: http://www.ncbi.nlm.nih.gov/pubmed/21709622.

Nguyen, T.D., 1998. Gene Structure and Properties of TIGR, an Olfactomedin-related Glycoprotein Cloned from Glucocorticoid-induced Trabecular Meshwork Cells. *Journal of Biological Chemistry*, 273(11), pp.6341–6350. Available at: http://www.jbc.org/cgi/doi/10.1074/jbc.273.11.6341.

Ojha, P., Wiggs, J.L. & Pasquale, L.R., 2013. The genetics of intraocular pressure. *Seminars in ophthalmology*, 28(5-6), pp.301–5. Available at: http://www.ncbi.nlm.nih.gov/pubmed/24138038.

Pang, C.P. et al., 2002. TIGR/MYOC gene sequence alterations in individuals with and without primary open-angle glaucoma. *Investigative ophthalmology & visual science*, 43(10), pp.3231–5. Available at: http://www.ncbi.nlm.nih.gov/pubmed/12356829.

Pasutto, F. et al., 2012. Variants in ASB10 are associated with open-angle glaucoma. *Human molecular genetics*, 21(6), pp.1336–49. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3284122&tool=pmcentrez&rendertype=abstract.

Pocar, P. et al., 2005. Molecular interactions of the aryl hydrocarbon receptor and its biological and toxicological relevance for reproduction. *Reproduction (Cambridge, England)*, 129(4), pp.379–89. Available at: http://www.ncbi.nlm.nih.gov/pubmed/15798013.

Qiu, Y. et al., 2014. Cellular processing of myocilin. *PloS one*, 9(4), p.e92845. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3986080&tool=pmcentrez&rendertype=abstract.

Quigley, H. a & Broman, a T., 2006. The number of people with glaucoma worldwide in 2010 and 2020. *The British journal of ophthalmology*, 90(3), pp.262–7. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1856963&tool=pmcentrez&rendertype=abstract.

Ray, K. & Mookherjee, S., 2009. Molecular complexity of primary open angle glaucoma: current concepts. *Journal of genetics*, 88(4), pp.451–67. Available at: http://www.ncbi.nlm.nih.gov/pubmed/20090207.

Sarfarazi, M., 1997. Recent advances in molecular genetics of glaucomas. *Human molecular genetics*, 6(10), pp.1667–77. Available at: http://www.ncbi.nlm.nih.gov/pubmed/9300658.

Saura, M. et al., 2005. Mutations including the promoter region of myocilin/TIGR gene. *European journal of human genetics : EJHG*, 13(3), pp.384–7. Available at: http://www.ncbi.nlm.nih.gov/pubmed/15483649.

Sohn, S. et al., 2009. Dual localization of wild-type myocilin in the endoplasmic reticulum and extracellular compartment likely occurs due to its incomplete secretion. *Molecular vision*, 15(December 2008), pp.545–56. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2654785&tool=pmcentrez&rendertype=abstract.

Sohn, S. et al., 2002. Expression of wild-type and truncated myocilins in trabecular meshwork cells: their subcellular localizations and cytotoxicities. *Investigative ophthalmology & visual science*, 43(12), pp.3680–5. Available at: http://www.ncbi.nlm.nih.gov/pubmed/12454036.

Stone, E.M. et al., 1997. Identification of a gene that causes primary open angle glaucoma. *Science (New York, N.Y.)*, 275(5300), pp.668–70. Available at: http://www.ncbi.nlm.nih.gov/pubmed/9005853.

Takamoto, M. & Araie, M., 2014. Genetics of primary open angle glaucoma. *Japanese journal of ophthalmology*, 58(1), pp.1–15. Available at: http://www.ncbi.nlm.nih.gov/pubmed/24258795.

Tamm, E.R., 2002. Myocilin and glaucoma: facts and ideas. *Progress in retinal and eye research*, 21(4), pp.395–428. Available at: http://www.ncbi.nlm.nih.gov/pubmed/12150989.

Ueda, J. et al., 2000. Ultrastructural Localization of Myocilin in Human Trabecular Meshwork Cells and Tissues. *Journal of Histochemistry & Cytochemistry*, 48(10), pp.1321–1329. Available at: http://jhc.sagepub.com/lookup/doi/10.1177/002215540004801003

Weisschuh, N. et al., 2005. Prevalence of myocilin and optineurin sequence variants in German normal tension glaucoma patients. *Molecular vision*, 11(April), pp.284–7. Available at: http://www.ncbi.nlm.nih.gov/pubmed/15851979.

Whigham, B.T. et al., 2011. Myocilin mutations in black South Africans with POAG. *Molecular vision*, 17(August 2010), pp.1064–9. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3086605&tool=pmcentrez&rendertype=abstract.

Zhuo, Y.-H. et al., 2008. Pro370Leu MYOC gene mutation in a large Chinese family with juvenile-onset open angle glaucoma: correlation between genotype and phenotype. *Molecular vision*, 14(August), pp.1533–9. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2518531&tool=pmcentrez&rendertype=abstract.