UNIVERSITÀ
DEGLI STUDI
DI PADOVA

Sede Amministrativa: Università degli Studi di Padova

Dipartimento di Medicina Molecolare

CORSO DI DOTTORATO DI RICERCA IN MEDICINA MOLECOLARE
CURRICOLO: BIOMEDICINA
CICLO XXXI

"Retinal Disorders: advanced methods for the genetic diagnosis and genotype-phenotype correlations"

Coordinatore: Ch.mo Prof. Stefano Piccolo
Supervisore: Ch.mo Dr. Enzo Di Iorio

Dottorando: Patrizia Nespeca

TABLE OF CONTENTS

Appendix A

Paper: “Novel variants of RPGR in X-linked retinitis pigmentosa families and genotype-phenotype correlation” by Parmeggiani et al. (Eur J Ophthalmol 2017).

Appendix B

Paper: “Generation of a transgene-free human induced pluripotent stem cell line (UNIPDi001-A) from oral mucosa epithelial stem cells” by Alvisi et al. (Stem Cell Res. 2018).

Appendix C

Paper: “Generation of a transgene-free induced pluripotent stem cells line (UNIPDi002-A) from oral mucosa epithelial stem cells carrying the R304Q mutation in TP63 gene” by Trevisan et al. (Stem Cell Res. 2018).

Appendix D

Paper: “Personalized stem cell therapy to correct corneal defects due to a unique homozygous-heterozygous mosaicism of Ectrodactyly-Ectodermal dysplasia-Clefting syndrome” by Barbaro et al. (Stem Cells Transl Med. 2016).

Appendix E

Paper: “Correction of mutant p63 in EEC syndrome using siRNA mediated allele-specific silencing restores defective stem cell function” by Barbaro et al. (Stem Cells 2016).

Abbreviation Index

1000g: 1000 genome databases
ACHR: Achromatopsia
AMD: Age related Macular Degeneration
AAV: Adeno-Associated Virus
BBS: Bardet-Biedl Syndrome
BCVA: Best-Corrected Visual Acuity
BLAST: Basic Local Alignment Search Tool
C: Conserved
CD: Cone Distrophy
CDS: Codifying Sequence
cGMP: cyclic Guanosine Monophosphate
Chr: Chromosome
CMV: Cytomegalovirus
Cov: Coverage
CRD: Cone-Rode Distrophy
CRISPR: Clustered Regularly Interspaced Short Palindromic Repeats
CSNB: Congenital Stationary Night Blindness
dbSNP138: Single Nucleotide Polymorphism 138 database
ddNTPs: dideoxynucleotides
dNTPs: deoxynucleotides
DSB: Double Strand Break
EDTA: Ethylene Diamine Triacetic Acid
EEC: Ectrodactyly-Ectodermal dysplasia-Clefting-syndrome
ERG: Electroretinography
ESP6500: Exome Sequencing Project v. 6500 databases
ExAC: Exome Aggregation Consortium
FAF: Fundus Autofluorescence
GERP++: Genomic Evolutionary Rate Profiling
HDR: Homology directed repair
HET: Heterozygous
HGMD: Human Gene Mutation Database
HOM: Homozygous
iPSCs: induced Pluripotent Stem Cells
IRD: Inherited Retinal Disorders
JS: Joubert Syndrome
LCA: Leber Congenital Amaurosis
MAF: Minor Allele Frequency
MD: Macular Degeneration (MD)
NA: Not Available
NB: Night Blindness (NB)
NC: Not Conserved
NGS: Next Generation Sequencing
NHEJ: Non-homologous end joining
OCT: Optical Coherence Tomography
OMD: Occult Macular Dystrophy
P: Pathogenic
PP: Probably Pathogenic
PCR: Polymerase Chain Reaction
Polyphen-2: Polymorphism Phenotyping v2 software
Qual: Quality
Ref: Reference
RPE: Retinal Pigmented Epithelium
RP: Retinitis Pigmentosa
RSC: Retinal Stem Cells
SIFT: Sorting Intolerant from Tolerant software
SNP: Single Nucleotide Polymorphism
STR: Short Tandem Repeat
STD: Stargardt Disease
TBE: Tris-borate-EDTA buffer

US: Usher syndrome
Var: Variant
VUS: Variant of Uncertain Significance
WHO: World Health Organization

Abstract

Background: According to the World Health Organization, 39 million people is completely blind. Between the genetic disorders causing visual impairment, by affecting 1 person to 3000-4000, Inherited Retinal Dystrophies, Occult Macular Dystrophy and Age-related Macular Degeneration are the most diffuse. Inherited Retinal Disorders (IRD) are retinal degenerations caused by mutations in at least 280 genes and more loci. IRD can be both syndromic or not-syndromic, inherited or sporadic, and all the inheritance patterns are possible. In fact, in addition to the autosomal dominant, autosomal recessive and the X-linked forms, sporadic, digenic and mitochondrial ones are possible. One of the most significant characteristics is their great genetic heterogeneity, due to the high number of genes and possible mutations involved. These genes transduce proteins involved in phototransduction and visual cycle processes, expressed in retinal pigmented epithelium or photoreceptors cells. The diseases are progressive and characterized by difficulties in night vision, loss of peripheral vision and photophobia, until complete blindness. Besides the genetic heterogeneity, these pathologies are characterized by allelic and clinical heterogeneity, which make the clinical diagnosis and the genotype-phenotype correlations complex, whenever the genetic cause is known. If genetic heterogeneity is determined by the presence of different mutations that induce the same clinical manifestation, clinical (or allelic) heterogeneity occurs in the presence of mutations on the same gene that cause different phenotypes.

While Occult Macular Dystrophy (OMD) is an inherited macular dystrophy caused by mutation in RP1L1 gene with an autosomal dominant pattern of inheritance, Age-related Macular Degeneration (AMD) is a complex disorder caused by an association of genetic, environmental and advanced age susceptibility.

Currently no diagnostic tests or definitive treatment for these disorders are available.

Aim of the study: The objective of this project is the development of diagnostic tests valid for the genetic characterization of patients affected by IRD or OMD. For this purpose, 620 patients and affected family members were clinically characterized, and a genetic biobank was created from the collection of their DNA, available for possible genetic testing. Based on the pathology they are affected, diagnostic tests were performed using Sanger sequencing or Next Generation Sequencing (NGS), following which correlations between the genotype and the phenotype were made.

Materials and methods: IRD patients were recruited and DNA extracted from blood samples. Retinal tissue and biopsies were collected too, in accordance with the current legal regulations. Genetic screening of 190 patients was realized using NGS targeted technology by MiSeq Illumina and identified mutations were then confirmed using Sanger sequencing. Genotype-phenotype correlations

were elaborated, and risk factors determination was performed in 118 patients by means TaqMan PCR for 7 polymorphisms. OMD patients were screened for RP1L1 gene and 60 proband suspected to be affected by X-linked retinitis pigmentosa for RP2 and RPGR genes using Sanger sequencing.

After a clinical anamnesis and in-depth visual examination, the patients were subjected to blood sampling from which the DNA was extracted and stored in the biobank. In parallel, a retinal biobank was generated from the collection of donors' retinas, collected according to the regulations in force. Genetic screening was performed on 190 patients affected by different autosomal recessive or dominant forms of IRD, using the targeted NGS technology, with panels of 25 genes. After selecting the variants or mutations identified by the analysis, Sanger sequencing was performed to confirm them. The correct attribution of the results to the analyzed individuals was achieved by comparing the genetic profiles obtained from the biobank DNA and the residual one from the NGS analysis. Genetic screening of patients with X-linked forms of IRD (n = 60) or OMD (n = 5) was obtained by means of Sanger sequencing of RPGR and RP2 (IRD) or RP1L1 (OMD) genes. A genetic susceptibility tool was developed and applied to 118 patients (39 of whom previously screened with NGS), using real time PCR, to evaluate the expression of 7 polymorphisms associated with AMD in subjects affected by IRD.

Results: With NGS screening 256 variants/mutations were identified in 126 patients. Among these, 49% is represented by novel alterations. The most mutated genes are USH2A (15%) and ABCA4 (14%), while no alteration was found in the BEST1, CRX, LRAT and RLBP1 genes. 27% of cases were completely resolved, while 39% only partially, due to the identification of only one recessive mutation or of variants of uncertain significance. Missense variants cover 80% of the total, followed by nonsense mutations (9%), frameshift insertions/deletions (8%), in-frame insertions/deletions (2%) and splicing alterations (1%). The correlation of the identified genotype with the diagnosed phenotype was possible in about half of genetically characterized patients. From the comparison it was possible to identify triallelic forms, often with the missense G1961E mutation of ABCA4 as an aggravating allele, and a new gene-disease association for the PRPF31 gene that in three patients causes Usher's syndrome. The analysis of the probands with X-linked forms of Retinitis Pigmentosa allowed the identification of a frameshift insertion in the ORF15 of the RPGR gene, a nonsense mutation in RPGR gene and an already known missense mutation in RP2, in three distinct families. Among the patients with OMD, the missense S1199F mutation in RP1L1 was identified for two of them. The susceptibility test for the 7 SNPs showed a high genetic high of AMD in 38 subjects, medium high in 16, medium in 36, and low in 28.

Conclusions: NGS technology has proved to be a useful diagnostic tool for the autosomal dominant or autosomal recessive forms of retinal dystrophies, allowing massive and parallel sequencing of

many genes and patients. The percentage of genetically characterized patients is consistent with the data in the literature. The use of a panel of genes turns out to be an acceptable compromise between the results obtained and the criticality related to the large amount of output data, compared to whole genome or whole exome sequencing. Much of the genetic data has been correlated with the clinical manifestations of the patients. Sanger technology, gold standard of sequencing, is optimal for the identification of the genetic cause in all pathologies where one or a few genes are involved, such as the X-linked forms of IRD or the OMD. The expression of risk factors, predisposing to macular degeneration, while not correlating with the severity of the confirmed mutations, could provide important relationships with the pathologies of which the subjects are affected.

The identification of the genetic causes of these pathologies will bring a new thrust to the development of gene and cell therapies, fundamental for the treatment of hereditary retinal diseases.

Riassunto

Presupposti dello studio: Secondo l'organizzazione mondiale della sanità, 39 milioni di persone sono ciechi. Tra principali cause genetiche che inducono compromissione visiva, con una prevalenza di 1:2500-4000, le Distrofie Retiniche Ereditarie (i.e. Inherited Retinal Dystrophy - IRD), la Distrofia Maculare Occulta (OMD) e la Degenerazione Maculare legate all'Età (AMD) sono le più diffuse. Le prime sono un gruppo di patologie che inducono degenerazioni retiniche indotte da mutazione in almeno 280 geni e molti loci. Le forme di IRD possono essere sindromiche o non sindromiche, sporadiche o ereditarie, con tutti i pattern di eredità possibili. Infatti, oltre alle forme autosomiche dominati e recessive e le forme legate al cromosoma X, sono possibili forme sporadiche, digeniche e mitocondriali. Una delle principali caratteristiche di queste patologie è la loro elevata eterogeneità, legata all'elevato numero di geni e di mutazioni coinvolti. Questi geni, infatti, codificano proteine coinvolte nei processi di foto-trasduzione e del ciclo visivo, espressi principalmente nelle cellule dell'epitelio pigmentato retinico o nei fotorecettori. Queste malattie hanno un andamento progressivo che inizia, generalmente, con difficoltà nella visione notturna, perdita della visione periferica e fotofobia, fino alla cecità assoluta. Oltre alla eterogeneità genetica, queste patologie sono caratterizzate da eterogeneità allelica e clinica, che rendono complessa la diagnosi clinica ed anche le correlazioni genotipo-fenotipo, qualora si conosca la causa genetica del paziente. Se l'eterogeneità genetica è determinata dalla presenza di differenti mutazioni che inducono la stessa manifestazione clinica, l'eterogeneità clinica (o allelica) si ha in presenza di mutazioni sullo stesso gene che causano fenotipi differenti.

Mentre la Degenerazione Maculare Occulta è una distrofia maculare ereditaria causata da mutazioni nel gene RP1L1, con una trasmissione autosomica dominate, la Degenerazione Maculare legata all'Età è una patologia complessa causata dall'associazione di fattori genetici, ambientali e dall'età avanzata.

Attualmente per queste patologie non sono disponibili né test diagnostici né cure.

Scopo dello studio: L'obiettivo di questo progetto è lo sviluppo di test diagnostici validi per la caratterizzazione genetica dei pazienti affetti da IRD o OMD. Per tale scopo, 620 pazienti e familiari affetti sono stati clinicamente caratterizzati e dalla collezione dei loro DNA è stata creata una biobanca genetica disponibile per eventuali test genetici. In base alla patologia di cui sono affetti, i test diagnostici sono stati eseguiti mediante utilizzo di Sanger sequencing o Next Generation Sequencing (NGS) al seguito del quale sono stati realizzate delle correlazioni tra il genotipo e il fenotipo.

Materiali e metodi: Dopo anamnesi clinica e approfondito esame visivo, i pazienti sono stati sottoposti a prelievo di sangue da cui è stato estratto il DNA stoccato all'interno della biobanca. In parallelo una

biobanca retinica è stata generata dalla collezione di retine di domatori, raccolte in accordo alle normative vigenti. Lo screening genetico è stato eseguito su 190 pazienti affetti da differenti forme autosomiche recessive o dominanti di IRD, mediante la tecnologia targeted NGS, con pannelli di 25 geni. Dopo la selezione delle varianti o mutazioni identificate dall'analisi, è stato eseguito il sequenziamento Sanger per la loro conferma. La corretta attribuzione dei risultati agli individui analizzati è stata realizzata grazie alla comparazione di profili genetici ottenuti dal DNA della biobanca e quello residuo dall'analisi NGS. Lo screening genetico, dei pazienti affetti da forme X-linked di IRD (n=60) o da OMD (n=5) è stato ottenuto da sequenziamento Sanger dei geni RPGR e RP2 (IRD) o RP1L1 (OMD). Un test di suscettibilità genetica è stato sviluppato e applicato a 118 pazienti (39 dei quali precedentemente screenati con NGS), mediante real time PCR, per valutare l'espressione di 7 polimorfismi associati all'AMD nei soggetti affetti da IRD.

Risultati: Con lo screening NGS 256 varianti/mutazioni sono state identificate in 126 pazienti. Tra queste il 49% è rappresentato da alterazioni non note in letteratura. I geni maggiormente mutati sono USH2A (15%) e ABCA4 (14%), mentre nessuna alterazione è stata riscontrata nei geni BEST1, CRX, LRAT and RLBP1. Il 27% dei casi è stato completamente risolto, mentre il 39% solo parzialmente, a causa dell'identificazioni di una sola mutazione recessiva o di varianti di significato incerto. Le varianti missenso coprono l'80% del totale, seguite dalle mutazioni nonsenso (9%), dalle inserzioni/delezioni frameshift (8%), dalle inserzioni/delezioni in-frame (2%) e dalle alterazioni di splicing (1%). La correlazione del genotipo identificato con il fenotipo diagnosticato è stata possibile in quasi la metà dei pazienti geneticamente caratterizzati. Dalla comparazione è stato possibile identificare forme trialleliche, con la mutazione missenso G1961E di ABCA4 come allele aggravante, e una nuova associazione gene-malattia per il gene PRPF31 che in tre pazienti causa la sindrome di Usher. L'analisi dei probandi affetti da forme X-linked di Retinite Pigmentosa ha permesso l'identificazione di un'inserzione framshift nell'ORF15 del gene RPGR, una mutazione nonsenso in RPGR gene e una mutazione missenso nota in RP2, in tre famiglie distinte. Tra i pazienti affetti da OMD, è stata identificato per due di essi la mutazione missenso S1199F in RP1L1. Il test di suscettibilità per i 7 SNP ha evidenziato un elevato alto genetico di AMD in 38 soggetti, medio alto in 16, medio in 36, e basso in 28.

Conclusioni: La tecnologia NGS si è rivelata un utile strumento diagnostico per le forme autosomiche dominanti o recessive di distrofie retiniche, permettendo il sequenziamento massivo e parallelo di molti geni e pazienti. La percentuale di pazienti geneticamente caratterizzati è coerente con i dati presenti in letteratura. L'uso di un pannello di geni risulta essere un accettabile compromesso tra la i risultati ottenuti e le criticità legate alla grande quantità di dati output, rispetto alle analisi di whole genome o whole exome sequencing. Buona parte dei dati genetici sono stati correlati alle

manifestazioni cliniche dei pazienti. La tecnologia Sanger, gold standard del sequenziamento, è ottimale per l'identificazione della causa genetica in tutte le patologie ove uno o pochi geni sono coinvolti, come le forme X-linked di IRD o la OMD. L'espressione di fattori di rischio, predisponenti alla degenerazione maculare, pur non correlando con la severità delle mutazioni confermate, potrebbe fornire importanti relazioni con le patologie di cui i soggetti sono affetti.

L'identificazione delle cause genetiche di queste patologie porterà una nuova spinta nello sviluppo di nuove terapie geniche e cellulari, fondamentali per il trattamento delle patologie retiniche ereditarie.

Introduction

Looking around amazing images, noises, colours and sensations permeate our body and our mind. All these sensations are arisen by our sense: sight, smell, taste, hearing and touch. Although all of them are fundamental, the considered essential sense is the sight.

An image is the result of a complex plot where the principle actors are the eyes. These two specialized organs, protected inside the orbit, are sensorial receptors that work as a camera. The eye is a hallow sphere divided in two compartments (the anterior and the posterior segment) by a lens. The anterior segment of the ocular bulbs is the visible district and comprises cornea, conjunctiva, iris, aqueous humor, lens and ciliary body. The posterior segment, instead, occupy the region from posterior surface of crystalline to retina consisting of sclera, choroid, Bruch's membrane, RPE, vitreous and neural retina.

Light rays deviated by cornea and crystalline pass into the retina and excited cones and rods cells, that originate transmitted to cerebral cortex through optic nerve. The retina consists of different layers, separated by synaptic plexiform layers containing neuronal processes and synaptic contacts. Different neuronal cells are present in these layers: cones and rods in the outer nuclear layer, the bipolar, horizontal and amacrine cells in the inner nuclear layer and the ganglion cells in the ganglion cell layer. Cones and rods are photoreceptors and specialized neurons with a polarized structure. If rods are stimulated by dim light and situated in the peripheral retina, cones are excited by bright light and are located in the central retina, including macula. There are approximatively 120 million of rods and 6 million of cones. The photoreceptors consist of inner and outer segments, cell body and synaptic terminal. The outer segment disc, photosensitive organelle, is in contact with RPE cells. Each disc contains visual pigments that consist of opsin molecules, rhodopsin in rods while retinal and cone opsin in cones. Different opsins are expressed in the cones, distinguish them in L-cones (sensible to 560 nm wavelength), M-cones (sensible to 530 nm wavelength) and S-opsin (sensible to 420 nm wavelength). Phototransduction starts when retinal, inside the binding pocket of the opsin, absorbs a photon and changes its conformation from 11-cis retinal to trans retinal. These modifications induce the activation of signalling cascade, with changes in cGMP and Ca^{2+} levels, second messengers that generate electrical signal.

The retinal pigmented epithelium (RPE) is a hexagonal monolayer of cells, fundamental for renewal of outer segment of photoreceptors, feeding of photoreceptors, phagocytosis of toxic products and metabolism of retinoids, maintenance of blood-retinal barrier and reduction of phototoxic damage (Dias et al. 2018; Broadgate et al. 2017) (Figure 1).

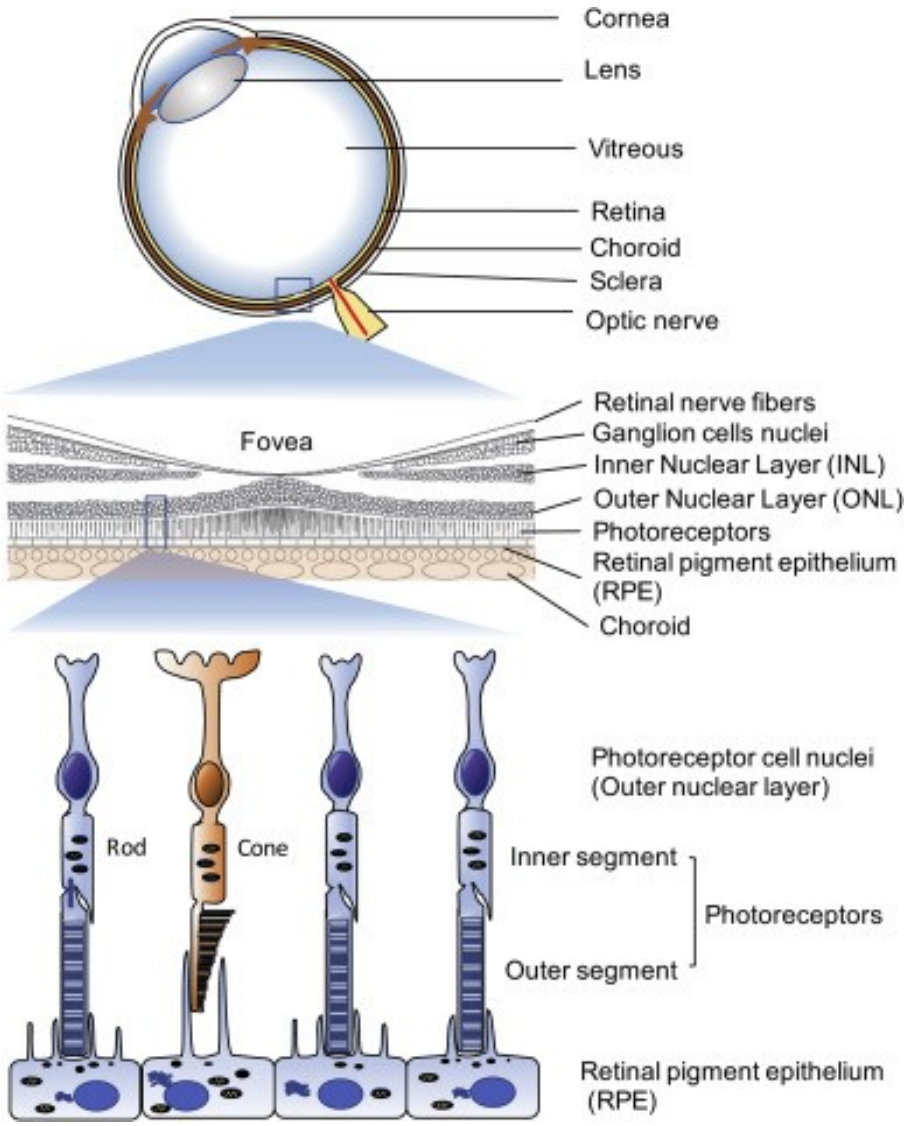

Figure 1. Representation of ocular and retinal structures. Retina, choroid and sclera constitute the posterior segment of the bulbs. Retina is composed of different layers, including retinal pigmented epithelium and photoreceptors. [From Dias et al. 2018].

Visual cycle

The absorption of light by rhodopsin begins the process of light transduction (Figure 2). During the light absorption, 11-cis-retinal, complexed with opsin in the rhodopsin, changes its conformation in all-transretinal, activating the rhodopsin in meta-rhodopsin, that can in this state activate transducin. When active rhodopsin is inactivated (through phosphorylation) by rhodopsin receptor kinase, it releases all-trans-retinal and binds 11-cis-retinal again, repeating the step. The consequent accumulation of the all-trans-retinal activates the ATP-binding cassette protein (ABCA4) that facilitates their transport out the intradiscal space, inside rods and cones cytosol where all-trans retinal complexes with N-retinylidine-PE. Here, after a hydrolysis step that releases N-retinylidine-PE, all-trans-retinal is reduced to all-trans-retinol by RDH protein, a membrane-bound NADPH-dependent retinol dehydrogenase. All-trans retinol is the translocated to RPE on the IRBP carrier, located in the

interphotoreceptor matrix. Into the RPE, after the bond to CRBP, all-trans retinol is reisomerized by an esterification catalysed by the lecithin-retinol transferase LRAT aided by the chaperon RPE65, important for the adaptation of the visual cycle to dark and light. This adaptation is guarantees by the concentration of 11-cis retinol and by the activity of LRAT, that regulates the conformational switch of RPE65. The mRPE65 (bound to membrane) having affinity for the all-trans retinylesters and donating an acyl group to LRAT, guarantees the production of 11-cis retinal during light. Instead, the soluble RPE65 (sRPE65) induces a reduction in the production of 11-cis retinal in the dark, due to its affinity to vitamin A. 11-cis retinol is oxidized by 11-cis-retinal dehydrogenase or RDH5, to 11-cis retinal. An alternative pathway for the regeneration of retinal involves the activity of the RGR receptor (RPE-retinal G protein-coupled receptor), that acts in an inverted light-induced rhodopsin isomerization converting all-trans-retinal into 11-cis-retinal. RGR function is to maintain 11-cis retinal constant independently to levels of environmental light. The newly generated 11-cis retinal gets through the subretinal space and enters the photoreceptors, helped by IRBP that protect it from isomerization en-route to the photoreceptors. After its entry into the outer segment, 11-cis retinal binds opsin and the light transduction goes back (Crouch et al. 1992; Strauss 2005; Ala-Laurila et al. 2006; Dias et al. 2018).

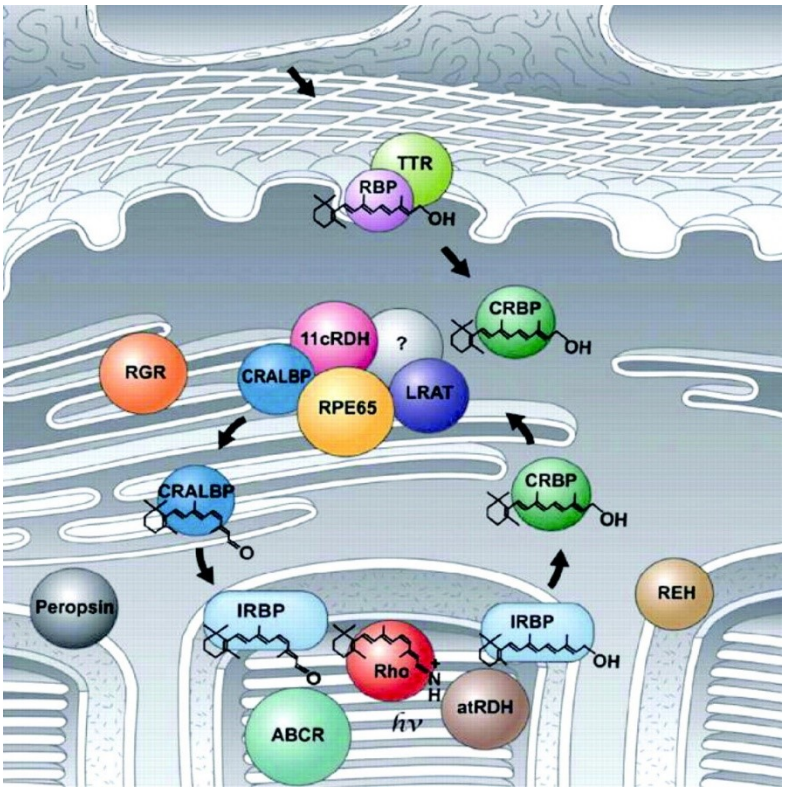

Figure 2. Simplified representation of visual cycle. Light transduction starts with photon absorption by rhodopsin. This induces the stereochemical change of 11-cis-retinal into all-trans-retinal. All-trans-retinal is processed into all-trans-retinol and translocated to the RPE. In the RPE retinol is reisomerized to 11-cis-retinal and then transported into the photoreceptors. The regeneration process can occur in two different pathways: one through as an isomerohydrolase, the other, a light-dependent pathway, through RPE-retinal G protein-coupled receptor which isomerizes all-trans-retinal to

11-cis-retinal after light absorption. IRBP, interstitial retinal binding protein; CRBP, cellular retinol binding protein; CRALBP, cellular retinaldehyde binding protein; atRDH, all-trans-retinol dehydrogenase; 11cRDH, 11-cisretinol dehydrogenase; LRAT, lecithin retinol acyltransferase; RPE65, RPE specific protein with molecular mass of 65 kDa; RGR, RPE-retinal G protein-coupled receptor; RBP/TTR, retinol-binding protein/transthyretin complex. [From Dias et al. 2018].

Ocular disorders

According to the World Health Organization, 39 out of the 285 million visually impaired patients are completely blind. It has also been recorded that the major causes of these visual impairments are uncorrected refractive errors (42%) and cataract (33%), affecting older aged group. These two causes of blindness could be overpassed ameliorating health care standards, consisting also in medical surgery, while infective causes like *Chlamidya trachomatis* (3% in worldwide) could be eradicated by increasing hygienic conditions in poor and rural countries.

At the contrary, genetic disorders, affecting retina or cornea, not always could be prevented or treated. These disorders can be monogenic, like inherited retinal dystrophies (IRD), or polygenic like age-related macular degeneration (AMD), that affect 1in 3 people over the age of 85. Principle disorders affecting retina will be discussed in depth below.

Inherited Retinal Dystrophies

With the acronym IRD (Inherited Retinal Disorders) is indicated a wide group of complex retinal disorders, characterized by a progressive degeneration of the photoreceptors of the retina (cones and rods) and of retinal pigment epithelium (RPE). Typical symptoms start to be evident during the second-third decade of the patients' life, and includes firstly night blindness, caused by the loss of rod function, that predominates in all the clinical course of the disease, nyctalopia then followed by central vision loss and changes in the retinal pigment epithelium (RPE), that causes an abnormal electroretinogram (ERG) response.

The development and progression of these diseases, that affect collectively 1 in 2000 to one in 3000, are mutable and unpredictable too, because they are correlated both to hereditary (i.e. expressiveness and genetic penetrance) and environmental factors (like exposure to dangerous light radiations) (Chaing et al. 2015; Parmeggiani et al. 2002; Yuong et al. 1994). Otherwise they can be classified on the base of clinical symptoms, as the age of onset of the first signs and the progression of the disorder, or on the base of involved cell types or, if known, considering the genetic defect (Di Iorio et al. 2017). Thus, because IRD are a group of heterogenous disorders, from the genetic and the clinical point of view. The genetic heterogeneity is due to the presence of multiple allele in a single locus (allelic heterogeneity). The principle effect is the difficulty to a genotype-phenotype correlation. The situation, where clinically similar and indistinguishable disorders may arise from mutation in different loci in different patients, take the name of locus heterogeneity, that in some cases could be

demonstrated through pedigrees analysis. In some patients, the discrimination of the forms could be complex due to the overlap of some symptoms or the casual involvement of other tissues (phenotypic heterogeneity). However, when the unique organ involved is the retina, the principle phenotype of non-syndromic forms of IRD are related to Stargardt disease (STD), Cone Distrophy (CD), Cone-Rode Distrophy (CRD), Achromatopsia (ACHR) that affected the central region of the retina; Retinitis Pigmentosa (RP) that involves the peripheral zones and both in the Leber Congenital Amaurosis (LCA). In the case of syndromic IRD (20-30%) is possible to distinguish principally Usher syndrome (US), Bardet-Biedl syndrome (BBS), and Joubert syndrome (JS) (Figure 3). Regarding the affected cells, the IRD can be classified in 1) rod dominant abnormalities as rod-cone dystrophies–Retinitis Pigmentosa, 2) cone dominant abnormality as cone or cone-rod dystrophy and achromatopsia, 3) macular dystrophy as Stargardt disease, Best macular dystrophy, Pattern dystrophy etc., 4) photoreceptors and bipolar cells abnormalities as X-linked retinoschisis and congenital stationary night blindness, 5) vitreoretinopathies as Wagner syndrome and Stickler syndrome, 6) hereditary choroidal diseases as choroideremia (Dias et al. 2018).

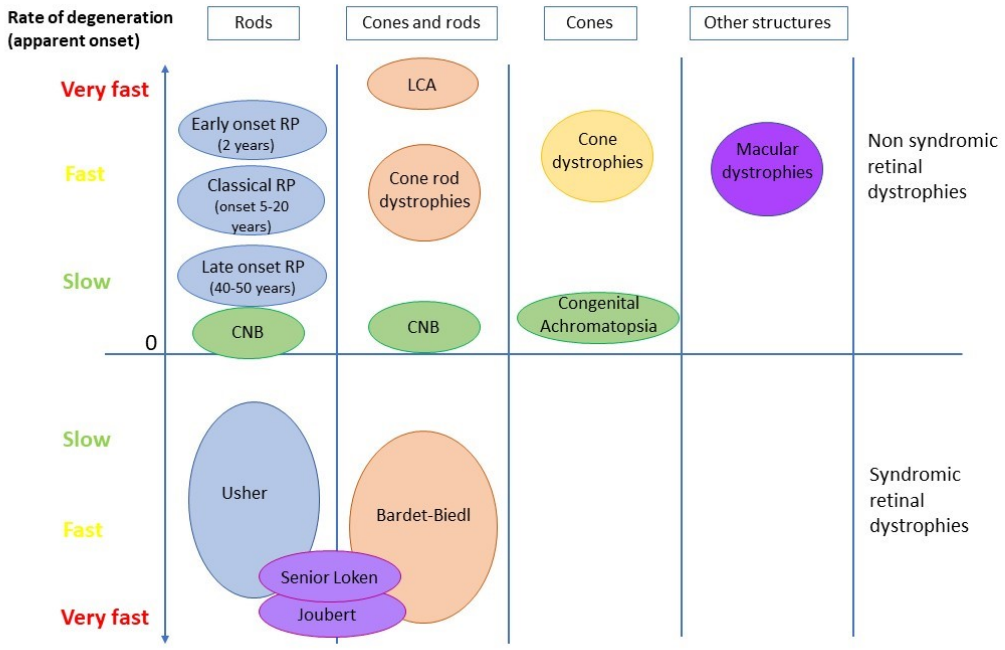

Figure 3. Classification of IRD based on the type of involved photoreceptors, rate of degeneration and syndromic/non-syndromic forms. CNB: Congenital Night Blindness; LCA: Leber Congenital Amaurosis.

Despite IRD don't show ethnic specificity, the spectrum of pathogenic variants within a given gene

may vary between population, as in certain isolated populations or those with a high rate of consanguinity.

Despite all these diseases have different genetic and biochemical mechanisms, they share the morphological and functional retinal cells damage, which is also related to chronic immune-inflammatory disorders (Kauppinen et al. 2016; Parmeggiani et al. 2012; Anderson et al. 2010; Xu et al. 2009). As for RP, the degeneration of photoreceptors is characterized by bone spicules, pigmented deposits, in the peripheral retina. The initial degeneration of the rods, due for mutations in gene specifically expressed in these cells, extend to cones cells in the subsequent phase of the disease, highlighting the co-dependence between these photoreceptors. The initial stage of the disorder is characterized by nyctalopia, due to rod alteration until their death for apoptosis. These symptoms, detectable only in dim light, don't let to a diagnosis, possible instead in the intermediate stage, where fundus examination reveals pigment deposits and the visual test underline correlation of impairment degree and photoreceptors degeneration. Pigmented debris, in the advanced stage, are spread on the peripheral and central retina, associated to retinal vessel attenuation, atrophy of RPE optic nerve head pallor. Further clinical examination can confirm the clinical diagnosis and monitor the progression, as full-field electroretinography, fundus autofluorescence (FAF), optical coherence tomography (OCT).

Genetics of IRD

Although IRDs are monogenic diseases, their molecular genetics is articulated. Mutations can affect many genes, that to date rely almost 280 genes associated with different retinal disorders and additional novel genes remain to be discovered (RetNet; Broadgate et al. 2017; Figure 4). As mentioned above, the complex nature of IRD can be included under the definition of heterogeneity: genetic (different genes cause the same phenotype), allelic (different pathogenic variants drop in the same gene) and phenotypic (different mutations in the same gene may cause different diseases or the same mutation in different people can show variable clinical consequences). For example, mutation in ABCA4 are associated to Cone or Cone-Rod Dystrophies (CD or CRD), Stargardt (STD), Retinitis Pigmentosa (RP) and Macular Degeneration (MD), while alterations in PDE6B can cause Night Blindness (NB), Congenital Stationary Night Blindness (CSNB) and RP (Dias et al. 2018; Verbakel et al. 2018).

So, many genotype-phenotype correlations could be not recognized due to the complex genetic architectures, based on genetic modifiers and multigenic patterns of inheritance that could obscure the molecular diagnosis (Dias et al. 2018; Chiang et al. 2015). Indeed, all the inherited pattern is

possible: the non-syndromic forms can be inherited in autosomal dominant, autosomal recessive, or X-linked manner. Sporadic forms can be caused by mutations occurred during embryonic development of retina, and rare digenic, forms occur in individuals who are heterozygous for both a ROM1 and a PRPH2 pathogenic variant (Stuck et al. 2016; Kajiwara et al. 1994). Triallelic and mitochondrial forms and multiple independent mutations in large families have been reported (Jones et al. 2017).

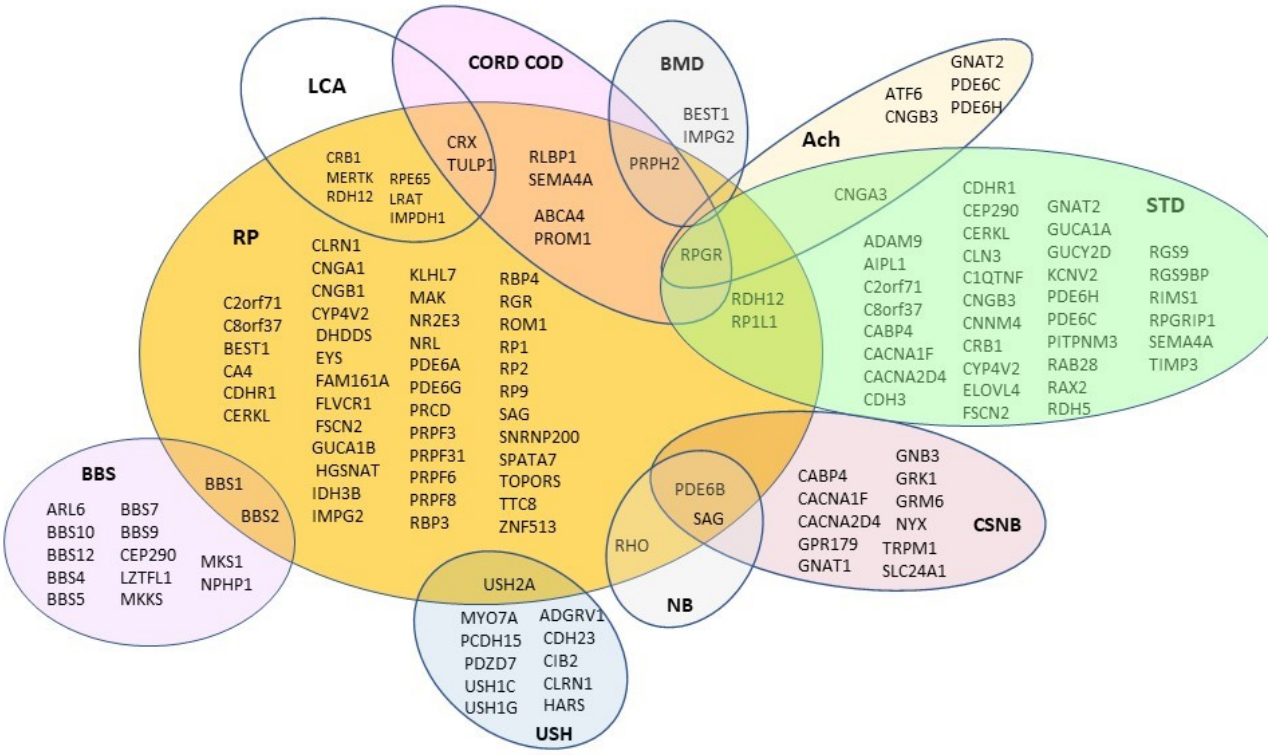

Figure 4. Graphic representation of some of the 280 involved genes in IRD, subdivided by the phenotype they cause. Ach: Achromatopsia; BBS: Bardet-Biedl Syndrome; BMD: Best Macular Dystrophy; CORD COD: Cone Rod and Cone Dystrophies; CSNB: Congenital Stationary Night Blindness; NB: Night Blindness; RP: Retinitis Pigmentosa; STD: Stargardt Disease; USH: Usher Syndrome.

Occult Macular Dystrophy

Described for the first time in 1989 by Miyake (Miyake et al. 1989), the Occult Macular Dystrophy (OMD) is an inherited macular dystrophy characterized by a progressive visual impairment, without visible fundus abnormalities (Piermarocchi et al. 2015). This rare disease affects one person to 4000

in an autosomal dominant pattern of inheritance, although sporadic forms are possible (Takahashi et al. 2014). Despite the pathophysiology is unclear, typical OMD is characterized by normal ophthalmoscopy, normal fluorescein angiography and normal full-field electroretinograms (ERGs) (Akahori et al. 2010), leading to a possible misdiagnosis with amblyopia, optic nerve disease or nonorganic visual disorder (Hanazono et al. 2010). The genetic cause of the disease is attributed to mutations in RP1-like protein 1 (RP1L1) gene, whose expression is restricted to the retina and specifically in photoreceptors (Bowne et al. 2003). This gene, composed of 4 exons, translate a protein with variable length due to the presence of 48bp polymorphic coding repeats in the mRNA. The shortest protein is 2,400 amino acid long, with only one repetition which contains a high percentage of glutamine and glutamic acid residues. Located in chromosome 8p23.1, RP1L1 encodes a large retinal-specific transcript with high sequence similarity to RP1. Considering their genomic organization, these genes are both composed by three coding exons with comparable size, proceeded by a non-coding one, probably linked to a common origin as a result of the duplication of an ancestor gene (Conte et al. 2002). Considering the translated protein, the similarity is mainly found in the two doublecortin domains in the N-terminal. Doublecortin domains are involved in interaction with microtubes, function already attributed to RP1 and RP1L1 genes (Taylor et al. 2000; Yoshiura et al. 2000). Moreover, these two genes play an active and synergic role in the retina, being involved in the photosensitivity and in the outer segment morphogenesis of rod photoreceptors.

Age related Macular Degeneration

One of the prominent aging-associated neurodegenerative diseases, with Alzheimer's and Parkinson's diseases, the Age-related Macular Degeneration (AMD) is a genetically complex disorder affecting the photoreceptors-RPE-Bruch's membrane-choriocapillaris complex (Scholl et al. 2007). Its blindness prevalence is approximatively 8.7% in world affecting around 170 million people (Pennington et al. 2016; Wong et al. 2014; WHO). Early signs of AMD comprise the formation in the Brunch's membrane of deposits, discernible in basal laminar or in basal linear deposits (drusen). Other relevant signs identified by clinicians through OCT are RPE defects, depigmented areas, subretinal haemorrhage, fluid and fibrosis, macular thickness and scar. Different forms of the disease can be distinguished: the "dry" or atrophic AMD (also known as geographic atrophy), characterized by progressive visual impairment due to the atrophy of retinal pigment epithelium (RPE) and the "wet" or neovascular AMD, due to the intraretinal invasion of vessels from the choroid, cause of bleed and dense macular scars (Fraccaro et al. 2015).

The complexity of the AMD depends on different risk factors: environmental, advanced age and genetic susceptibility. One of the principal environmental factors is smocking: with an unknown

mechanism, it increases in about two-fold the risk to develop the disorder. Additional risk factors are sunlight exposure, alcohol and infections. The strongest risk factor, the aging, is a physiological process genetically regulated. The aging induces in the organism molecular, morphological and functional alterations in response to damage. In the eyes, the degeneration in retina and RPE and the accumulation of lipofuscin and cholesterol in RPE are some of the aging signs.

Since 1990, genetic risk factors have been documented, through twin studies and genetic linkage analysis, revealing an apparently autosomal dominant inheritance. Several large population studies documented the presence of susceptibility loci in complement pathway genes, which result in a pathogenic mechanism closely related to immune-inflammatory process (Eandi et al. 2016; Nagasaka et al. 2016; Karlstetter et al. 2015; Murakami et al. 2015). The first described common genetic variant associated with AMD was the rs1061170 followed by rs4038846 and rs1410996 in the CFH gene. Being a regulator of complement activation, alterations in this gene induce multifaceted inflammation phenomena inside retina-RPE anatomo-functional unit. Other complement genes strongly associated with AMD are C2 (rs9332739), CFB (rs641153), C3 (rs2230199) and ARM2 (rs10490924) (Klein et al. 2013; Francis et al. 2011; Swaroop et al. 2009). The assessment of algorithm able to identify subjects at risk for advanced AMD is appealing, principally for the possibility to predict it after the drusen appearance but before the vision loss.

Role of inflammation in retinal cells degeneration

Both progression of IRD such as also OMD (Eandi et al. 2016; Nagasaka et al. 2016; Karlstetter et al. 2015; Murakami et al. 2015) and onset/evolution of AMD (Kauppinen et al. 2016; Parmeggiani et al. 2012; Anderson et al. 2010; Xu et al. 2009) are determined by genetic and epigenetic factors involving in complex para-inflammatory chronic processes. Although normal para-inflammation preserve the tissues homeostasis and their functionality, after the exposition to prolonged stress induces the initiation and the progression of retinal damages secondary to the overload of reactive oxygen or nitrogen species (ROS or RNS), with consequent loss of rod-related vision and/or loss of cone-related vision. In the human retina the main sources of dangerous ROS, such as high concentration of hydrogen peroxide and superoxide anion, are represented by: i. light exposure (in particular secondary to harmful blue, violet and ultraviolet radiations); ii. metabolic reaction like phagocytosis and lipofuscin phototoxicity (triggering RPE damages via activation of NLRP3 inflammasome); iii. mitochondrial respiration (Brandstetter et al. 2016; Xu et al. 2009). The interaction of RNS and ROS produces peroxynitrite, that induces nitration od proteins residues and the oxidation of metal centres, DNA lipids and proteins increasing the oxidative cell damage (Xu et al. 2009). The photo-oxidative damage and activation of the NLRP3 inflammasome have been

identified as key contributing factors to both RPE cell death and oxidative modifications in photoreceptor outer segments (POS), initiating a deleterious pathogenic cycle in which NLRP3 inflammasome particles (released from macrophages) and inflammasome priming by IL-1 and/or C5 complement activation alter the cell death mechanism from apoptosis to pyroptosis. This alteration critically increases that phototoxic effect on the retina able to trigger those inflammaging phenomena plausibly involving in the phenotypic heterogeneity observed from time to time in pedigrees with RP as well as in patients with other IRD, and also in heterozygous female carriers of X-linked RP (Parmeggiani et al. 2017; Parmeggiani et al. 2016). In addition, inflammaging phenomena significantly increase the individual susceptibility to AMD in patients with specific genotypic background mainly implicating in the immuno-inflammatory pathways of the complement system that is also able to negatively influence the grade of AMD severity and the incidence of AMD complications (Parmeggiani et al. 2013; Parmeggiani et al. 2012; Campa et al. 2010).

Biobank

Born around 1959 after the death of a young man for unknown causes, after identified as HIV infection, biobank became during decades an important topic in the research and clinical sphere, not only for the institution of these collections of samples but also in its definition. The term biobank refers to an organized collection of human biological material and associated information stored for one or more research purposes (Fransson et al. 2015). Today, generally, hospitals, oncologic and research centres collect routinely samples of various origin that could became useful for future research. In a simplistic view it can have a structure (place and instrument), the organized collection of biological samples, the data associated to these samples (clinical, epidemiological, biological data, etc.) (Caenazzo 2012).

Biobanks have arisen great expectations, such as the identification of genetic defects, the susceptibility to certain diseases, the possibility of developing personalized therapies based on the genetic background of the patient, identification of new biomarkers and the development of new drugs. For all these reasons the interest on this theme is growing exponentially and many institutions are organizing their own biobanks.

First and Next Generation Sequencing

For long time, the genetic and clinical heterogeneity of IRD have hampered the elaboration of molecular tests specific for the diagnosis of these diseases. These tests are expensive and, moreover, not always with exhaustive results, since not all the genes involved in IRD have been discovered yet. For this reason, and in light of the presence of large number of affected patients, waiting to genetic

diagnosis, the use of NGS technology seems to be the most suitable approach, since many different genes and patients can be analysed in one single run.

DNA sequencing has a long history, since the days of Maxam-Gilbert chemical sequencing and Sanger dideoxynucleotides based sequencing techniques in the 70's. By refining the "plus and minus" system, developed on 1975 with Coulson (Sanger and Coulson 1975), Sanger and colleagues elaborated the chain termination method, based on the use of dideoxynucleotides (ddNTPs). The incorporation of radiolabelled ddNTPs, analogues of dNTPs that lack the 3' hydroxyl group, induces the termination of the sequence and the exposure to X rays leads their lecture on a polyacrilemide gel. In 1986, the radiolabels were substituted with fluorescent dyes, permitting the automatization of the technique and so its capacity to sequence 1000bp. Sanger sequencing is still the gold standard for accuracy and error rate less than 1%, even though it is expensive and time consuming (Broadgate et al. 2017).

Used during the realization of the "Human Genome Project", which had the purpose of sequencing the first whole human genome, the excessive costs and effort (2,7 billion dollars and 13 years of work) drove the creation of innovative methods, based on the massive parallel sequencing, known as next generation sequencing technologies. These techniques are in rapid evolution and from their advent, in 2010, many aspects have changed. Nowadays, different platforms are present on the market and their costs are getting lower year by year (Table 1). All the platforms share one common revolutionary feature that distinguish them from Sanger method and identifies them as "Next Generation": the short read, massively parallel sequencing of clonal amplified clusters of DNA molecules or of single molecules physically separated in flow cells.

Massively parallel sequencing allows sequencing of hundreds of Mb to Gb of DNA in one single run.

	Maximum Throughput/run	**Mean Length (nucleotide)**	**Error rate**	**Applications**	**Source of errors**
454 FLX	700Mb	Up to 1000bp	0.003%	*De novo* genome sequencing and resequencing, target	Intensity cut-off, homopolymers, signal cross-talk interference among neighbours,

	Maximum Throughput/run	Mean Length (nucleotide)	Error rate	Applications	Source of errors
				resequencing, genotyping, metagenomics	amplification, mixed beads
Illumina	600Gb	~100bp or 300bp	<0.01%	Genome resequencing, quantitative transcriptomics, genotyping, metagenomics	Signal interference among neighbouring clusters, homopolymers, phasing, nucleotide labelling, amplification, low coverage of AT rich regions
SOLiD	90Gb	75bp	0.001%	Genome resequencing, quantitative transcriptomics, genotyping	Signal interference among neighbours, phasing, nucleotide labelling, signal degradation, mixed beads, low coverage of AT rich regions
Helicos	37Gb	~35	2-7% to 0.2-1%	Non-amplifiable samples, PCR free and unbiased quantitative analyses	Polymerase employed, molecule loss, low intensities
Ion Torrent	10Gb	Up to 400bp	<1%	De novo genome sequencing and resequencing, target resequencing, genotyping, RNA-seq on low-	Homopolymers, amplification

	Maximum Throughput/run	Mean Length (nucleotide)	Error rate	Applications	Source of errors
				complexity transcriptome, metagenomics	
PacBio RSII	230Mb	Up to 30000bp	<0.001%	Target resequencing (amplicons), genotyping	Intensity cut-off, homopolymers, signal cross-talk interference among neighbours, amplification, mixed beads
GridlON™ Nanopore	-	>2kb	2%	identification and monitoring of viral pathogens and antibiotic resistance, environmental and food safety monitoring, human and plant genome sequencing, haplotyping and other applications	-

Table 1. Features of "next-generation" sequencing (NGS) platforms (Alvarez-Cubero et al. 2017; Barzon et al. 2011).

For the genetic diagnosis of IRD patient, object of this work, Illumina sequencing by synthesis was used since it is the most adopted for these types pf analysis and guarantees the highest accuracy, yield of error-free reads and percentage of base calls above quality score 30 (Nakazato et al. 2013, Ross et al. 2013). In recent years, many works on massive NGS screening of IRD patients, principally using Illumina technologies, have been published with interesting results in terms of diagnosis rate and genotype-phenotype correlation, highlighting novel mutation and novel genetic associations (Di

Resta et al. 2018; Wawrocka et al. 2018; Jones et al. 2017; Schulz et al. 2017; Perez-Carro et al. 2016; Chiang et al. 2015; Ge et al. 2015).

These shotgun sequencing require an initial step of DNA fragmentation (from 50 to 500 bp) through mechanical, enzymatic or insertion mechanisms (Figure 5).

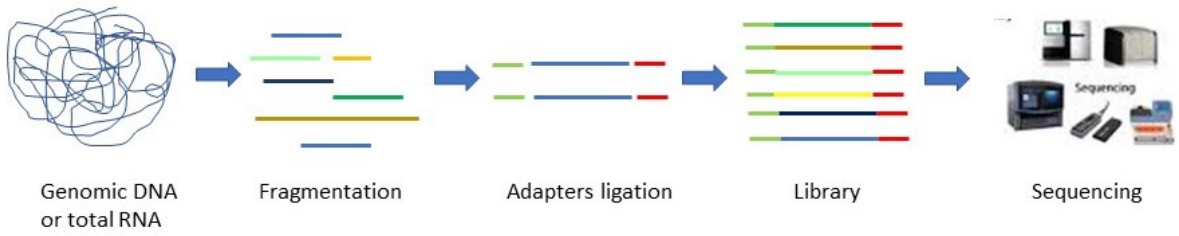

Figure 5. Schematic workflow of the NGS library preparation: the template is fragmented in small sequences and then ligated to specific adapters to generate the library, fundamental for the sequencing.

Each fragment is used to create a library. In the library the fragments are ligated to specific sequences, 4-8 nucleotides long, called adapters, and to 6-10ntp long barcode, specific for the sample. The library preparation is the critical step of the experiment. By mixing multiple libraries, a library pool is generated and sequenced as target for the amplification step. After the hybridization of the single DNA molecules to primers fixed on a solid surface, the clonal amplification begins. NGS chemistry is similar to capillary electrophoresis (Sanger) sequencing, since DNA polymerase catalyses the incorporation of dNTPs marked with fluorophores into a DNA template strand during several cycles of DNA synthesis in both cases. At every cycle the incorporated nucleotide is identified by fluorophore excitation (Børsting et al. 2015). The significant difference is that while Sanger technique analyses fragments one by one, after they have been divided for length by capillary electrophoresis, NGS sequences millions of fragments using a massive parallel technique.

In figure 6 (from Alvarez-Cubero et al. 2017) are represented all the NGS sequencing technologies actually available, but the Illumina workflow is set out in more detail below (www.illumina.com):

1- Library preparation: the template is randomly and mechanically or enzymatically fragmented in order to form the library. Adapter and index oligonucleotide sequences are ligated both to 5' and 3' ends. The indexes are sequences specific for each cluster, essential for DNA fragment identification. The adapters are necessary for the hybridization to sequences fixed on the matrix, thus permitting cluster formation. Adapter-ligated fragments are then amplified through PCR and purified using specific filters.

2- Cluster generation: the created library is loaded into a flow cell, where for sequence complementary the fragments are flanked on adapter sequences fixed on the crystal surface. At this point the library is amplified by bridge PCR, thus creating many clonal clusters, everyone different from the other one and deriving from one single molecule. So, the templates are ready for sequencing.

3- Sequencing by synthesis: after the primer hybridization to the fragments bound to the chip, the amplification with reversible terminator-bound fluorescently labelled nucleotides starts, thus incorporating one base per round (base by base technology). Each of the four bases, once light-excited, has a unique emission, and, after each round, the detector records which base has just been added. As all 4 reversible terminator-bound dNTPs are present during each sequencing cycle, natural competition minimizes incorporation bias and greatly reduces error rates compared to other technologies (Ross et al. 2013). The result is an accurate sequencing that virtually eliminates errors due to particular sequences, such as repetitive regions and homopolymers.

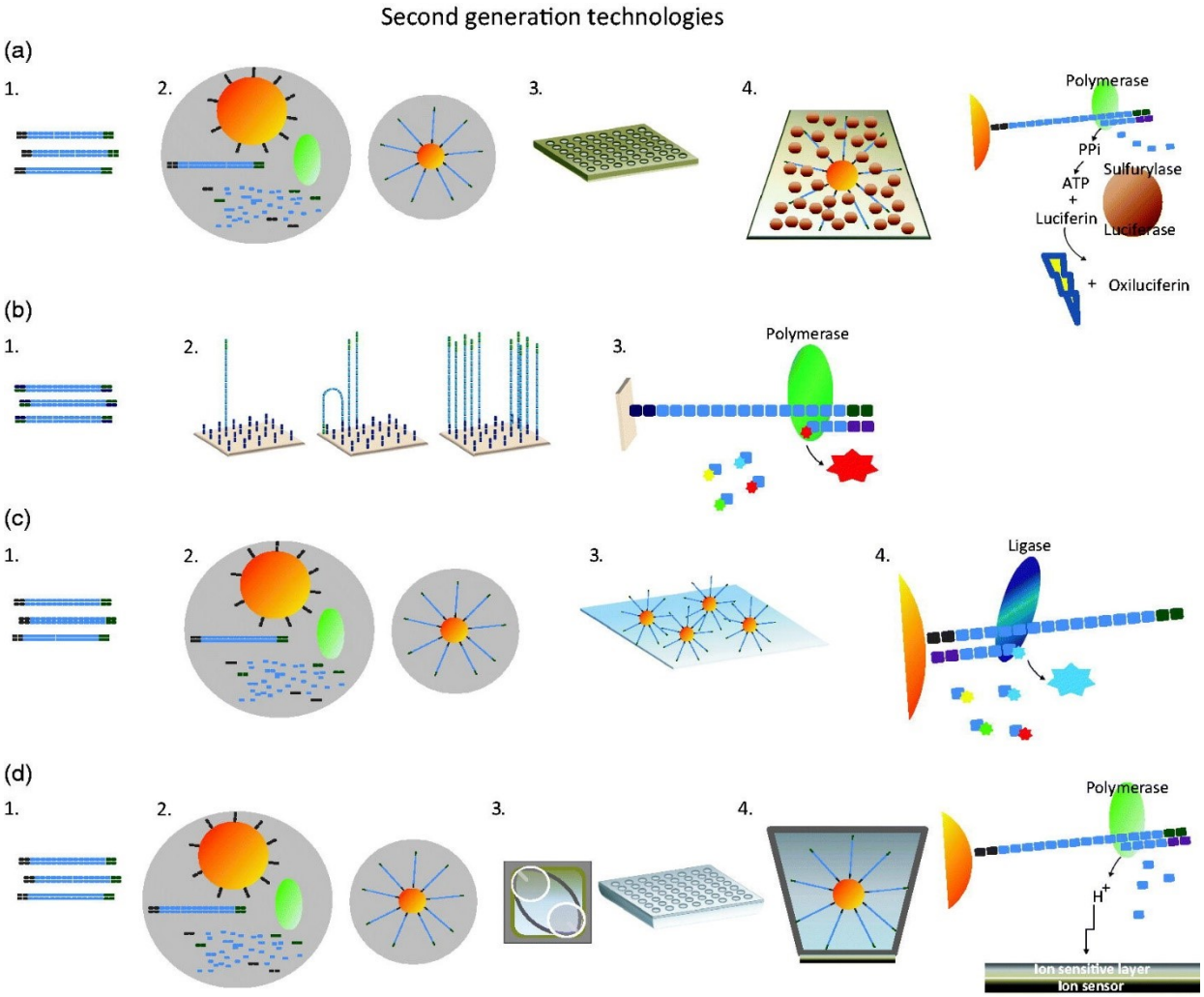
Second generation technologies
(a)
1.
2.
3.
4.
Polymerase
PPi
Sulfurylase
ATP
+
Luciferin
Luciferase
+ Oxiluciferin
(b)
1.
2.
3.
Polymerase
(c)
1.
2.
3.
4.
Ligase
(d)
1.
2.
3.
4.
Polymerase
H+
Ion sensitive layer
Ion sensor
Third generation technologies
(e)
1.
2.
3.
Polymerase
(f)
1.
2.
3.
Inmobilized polymerase
Fourth generation technologies
(g)
1.
2.

Figure 6. Representation of the workflow of second (a-d), third (e-f) and fourth (g) generation NGS technologies: (a) roche 454; (b) Illumina; (c) AB SOLiD 5500; (d) Ion PGM & Ion Proton; (e) Heliscope; (f) PacBio; (g) Nanopore Sequencing GridION. In the second-generation platforms, the process starts with (1) the DNA fragmentation and ligation to specific adapters, followed by (2) the amplification step by means emulsion-PCR (a; d) or bridge PCR (b). For pyrosequencing (a), the DNA-bounded beads are loaded to multiwall chip (3), in which are present enzymes necessary for the sequencing (DNA polymerase, ATP sulfurylase and luciferase). For Illumina technology (b), the sequencing (3) occurs for synthesis by the addition of fluorescently labelled reversible terminators. In the AB SOLiD 5500 platform the DNA library is fixed and sequenced, while the Ion PGM & Ion Proton sequencing the incorporation of nucleotides induces the increase of H+ ions that changes the pH in the wall, detected and converted in the corresponding added base. The Heliscope workflow (e) includes the fragmentation and the poly-A tailing of the template (1), then hybridized in a glass surface (2) and sequenced adding terminator nucleotides. After the signal capture the fluorophore is cut away to continue the process (3). In the PacBio system (f), the hairpin adapters bind the double-strand DNA fragments (1) and the hairpin-DNA structures create consensus sequence (2) from which starts the sequencing (3). In the GridION (g) when a charged molecule across the pore (1), it generates a resistance in the ions flow that induce a change in the electrical current measurable by the detection system (2). The differences in signal are related to the size of molecules [from Alvarez-Cubero et al. 2017].

After the sequencing, the sequence reads are aligned to a reference wild type genome, to reconstitute the proper order of the fragments. In this step many analyses are possible, as single nucleotide polymorphism (SNP) or insertion-deletion (indel) identification, phylogenetic or metagenomic analysis and many others.

To obtain a more accurate analysis paired-end technology can be used: both ends of the DNA fragment are sequenced, in this way twice the number of reads is produced at the same time and effort and the resulting alignment is more accurate, thus permitting the identification of indels, which is not possible with single-read data.

The genetic characterization of monogenic diseases can be performed on genes related to the disorder using targeted sequencing, which allows the isolation and then sequencing of a subset of regions or genes of interest in the genome. For this type of test, two different available methods are the target enriched workflow and the amplicons generation workflow (www.illumina.com). In the first one, biotinylated probes capture fragments of interest that then are isolated using magnetic supports. In the second one, highly multiplexed PCR oligo sets assure the amplification and the purification of regions of interest.

The advantages of this NGS method are the less waste of time and money, the higher coverage levels (where coverage means the average number of sequenced bases that align to each base of the reference DNA). In fact, typically whole exome sequencing reaches coverages of about 30x-50x per genome, while targeted sequencing can easily reach 500x-1000x coverages.

Materials and methods

Approval of Ethical Committee

All applicable institutional and governmental regulations concerning the ethical involvement of human volunteers were observed during this research. As required by the local ethical committee, the documentation submitted to the approval request has been drawn up and includes the study protocol, the study synopsis, the informed consent, the privacy notice, the bibliography, the list of the centers involved, the curriculum vitae of the local study manager and the data collection form.

Informed consent was collected from all the patients or their legal guardians.

Human Subjects recruitment

Patients cohort comprised individuals with diagnosed IRD, OMD and their families, recruited in the Center for Retinitis Pigmentosa of Veneto Region (Camposampiero Hospital, Italy). After detailed ophthalmic examinations, including electroretinogram (ERG), fundus, visual field test and best-corrected visual acuity (BCVA) measurements, clinical anamnesis and family history were harvested. Clinical diagnosis was obtained considering the functional signs (night blindness, nystagmus, photophobia, reduction in color discrimination, reduction in visual acuity), the fundus appearance (pigment deposits, bone-spicule pigment migration, foveal reflex, frank atrophy in RPE in the fovea), the ERG response (decrease in a and b wave amplitude, subnormal or non-detectable scotopic or photopic response), the OCT (foveal hypoplasia, loss of inner/outer photoreceptor segment junction) and the age at diagnosis.

In the cohort were included (n°3) patients with a previous recognized variant (benign polymorphism or pathological variant), used as validation control.

Genetic biobank creation

Blood samples from each patient were collected and the genomic DNA was extracted using an automated DNA extractor (Roche, Mannheim, Germany), following the manufacturer's instructions. DNA concentration was measured by Qubit Fluorimeter (Thermo Fisher Scientific, Waltham, MA, USA) or Nanodrop One (Thermo Fisher Scientific, Waltham, MA, USA). All the obtained genetic material was stored with a specific number code which connects to patient's information. Being essential in order to make sure that each sample is connectable to the personal, genealogical, clinical and genetic data of the belonging subject, without violating the privacy laws in force (Authorization n. 9/2014 regarding the treatment of personal and genetic data published by the Italian data protection

authority), a proper organization was created; thus making possible the elaboration of new analyses and studies of new therapies also in the following years.

Next Generation Sequencing

190 samples from the DNA biobank underwent targeted Next Generation Sequencing in two different panel designed by Illumina to capture IRD-associated genes (Table 2) following the manufacturers' protocols. In the cohort were included (n°3) patients with a previous recognized variant (benign polymorphism or pathological variant), used as validation control.

Panel a			**Panel b**		
Gene	**OMIM**	**Inheritance pattern**	**Gene**	**OMIM**	**Inheritance pattern**
ABCA4	* 601691	AR	*ABCA4*	* 601691	AR
BEST1	* 607854	AD	*ADGRV1*	*602851	AR
CNGA1	* 123825	AR	*BBS1*	*209901	DR/AR
CNGB1	* 600724	AR	*BBS10*	*610148	AR
CRB1	* 604210	AR	*BEST1*	* 607854	AD
CRX	* 602225	AD	*CDH23*	*605516	AR/DR
EYS	* 612424	AR	*CNGA1*	* 123825	AR
IMPDH1	* 146690	AD	*CNGB1*	* 600724	AR
LRAT	* 604863	AR	*CRB1*	* 604210	AR
MERTK	* 604705	AR	*CRX*	* 602225	AD
NR2E3	*604485	AD	*EYS*	* 612424	AR
PDE6A	* 180071	AR	*IMPDH1*	* 146690	AD
PDE6B	* 180072	AR	*MYO7A*	*276903	AR/AD
PROM1	*604365	AR	*NR2E3*	*604485	AD
PRPF3	* 607301	AD	*PDE6A*	* 180071	AR
PRPF31	*606419	AD	*PDE6B*	* 180072	AR
PRPF8	*607300	AD	*PRPF3*	* 607301	AD
PRPH2	* 179605	AD	*PRPF31*	*606419	AD
RDH12	* 608830	AD	*PRPF8*	*607300	AD
RHO	* 180380	AD/AR	*PRPH2*	* 179605	AD
RLBP1	* 180090	AR	*RHO*	* 180380	AD/AR
RPE65	* 180069	AD/AR	*RPE65*	* 180069	AD/AR
ROM1	* 180721	AD	*RP1*	* 603937	AR/AD
RP1	* 603937	AR/AD	*USH1C*	*605242	AR
TULP1	* 602280	AR	*USH2A*	*608404	AR

Table 2. Representation of the gene panels applied in the research.

Briefly, the qualified genomic DNA sample was randomly fragmented by Covaris through mechanical fragmentation and the size of the library fragments was mainly distributed between 250 bp. Then upstream and downstream oligos (containing the adapter) specific for the targeted region of interest

were hybridized to the resulting fragments. Unbound oligos were then removed using a size-selection filter. Extracted DNA was amplified by ligation-mediated PCR (LM-PCR). Extension-ligation products were amplified again using new oligos formed by an adapter and an index sequence and then purified by the Agencourt AMPure XP beads. Each captured library was then loaded on Miseq platform (Illumina, San Diego, CA, USA), and high-throughput sequencing for each captured library was performed. Raw image files were processed for base-calling with default parameters and the sequences of each individual were generated. All variations were filtered using dbSNP138, the 1000 Genomes Project and ESP6500 databases.

Variant Prioritization and Pathogenicity Assessment

Identified variants were filtered by applying the following prioritization criteria:

1. Variants in the coding sequence, excluding intergenic, 5′ and 3′ untranslated regions or deep intronic variants, were selected.
2. Variants were filtered by minor allele frequency (MAF) ≤ 0.01 or without MAF value in the 1000 Genomes Project and Exome Aggregation Consortium (ExAC).
3. Variants were checked in dbSNP132: Variants with *rs* number and present in Human Gene Mutation Data (HGMD®) were considered as known pathogenic mutations after all the references were checked, while variants without *rs* number were considered as novel.
4. Nonsense and frameshift variants were considered to be pathogenic.
5. To assess the pathogenesis of the missense variants, different predictive software programs were used including Sorting Intolerant from Tolerant (SIFT; sift.jcvi.org), Polymorphism Phenotyping v2 (Polyphen-2; genetics.bwh.harvard.edu/pph2), and Mutation Taster (www.mutationtaster.org). Those variants predicted as damaging by at least two different prediction software were initially considered potentially pathogenic.

PCR

To confirm some of the mutations identified by Next Generation Sequencing and to proceed in the RP1L1, RP2 and RPGR analysis, the classical and more reliable Sanger sequencing was performed. PCR was used for the amplification of the DNA regions in which NGS assesses the presence of an uncertain mutation. Forward and reverse primers were designed for each interesting mutation and

tested to set the reaction program (Table 3). Primer3 Plus software (http://primer3plus.com/) was used to design PCR primers, with the following settings: minimum size of 18bp and maximum size of 27 bp, melting temperature between 57°C and 63°C. Buffer 10X with $MgCl_2$ 25mM, primers mix 10µM, dNTPs mix 0,4µM, AmpliTaq Gold® DNA Polymerase (ThermoFisher Scientific, Waltham, MA) and 100-150ng of DNA were used (final reaction volume of 25µl). The success of the PCR reaction was verified using gel electrophoresis.

Primer name	Forward Sequence	Reverse Sequence	Gene	Mutation/ exon	Tm (°C)	Frag. Lenght
RP #1B	ggggcagaaaagacacaacc	ctccccatcctccaaccc	ABCA4	H423R	59	224
RP #2	cctcctagtgctggtagttttc	ctagtggaagatttaaaattac	ABCA4	G1961E	52	317
RP #3	cagaaggaactgtagatgc	catccatcccccaggttac	MERKT	R466K	55	317
RP #4	ggctctaaactatgatggc	agactatttcaggtaatatagtaaac	EYS	Y3135X	58	320
RP #5B	aggagaccaatgaaggagaaact	caaacctgagatttgcgggt	RP1	R1595Q	59	322
RP #6	tcactgccatgtccccag	ccattcctcctcagccctg	IMPDH1	K62X	59	300
RP #7	caaacttctccccaccaagc	gagcaggtccagacggtatt	IMPDH1	D226V	59	222
RP #8	gtcaaagtggccagcatcc	aaaaggtcagggatgggagg	LRAT	E146V/K147X	59	300
RP # 9	cagcatggagttgaatgagaca	acacacacttctgggaggtc	ABCA4	K397E	59	299
RP # 10	ttctggaggcgtgagatctg	tttccccaacccaagagacc	ABCA4	K2056X/K2076X	59	299
RP # 11	gagtttgactgacagcccca	tgtgtcgcttactggtggaa	ABCA4	P981fs	59	269
RP # 12	cctgtgagagcatctgggc	cctgccacagtctgatgca	ABCA4	C1840X	60	300
RP # 13	tggggccatgtaattaggca	ccaggaaatgacaggctagc	ABCA4	R943Q	59	283
RP # 14	ctggccagtgacatcccc	ctcccctccagggaccac	ABCA4	D155V	59	284
RP # 15	ggtggaatggaatgtggaact	catgtacgatgcttgcccc	ABCA4	L1815X	58	291
RP # 16	gcccaaaagccatgaaccaa	tgagaaagccaggatccaca	MERKT	V272E	59	299
RP # 17	cccctgacaaccactcatct	aaagtctgggtgaaaacctca	MERKT	K642X	58	385
RP # 18	caaatgctggcagtgatcct	tgaccagcttttcggttctatg	CNGA1	L654fs	59	250
RP # 19	ctagagcgcggcagactag	gcacagccttgggttacatt	CRB1	N732fs	59	284
RP # 20	gctccatctgaatcagccaa	agtcccaaggcttgttgagg	PDE6A	Y700fs	58	293
RP # 21	acgtctctccagtgatgaca	ggcgaggtcatatctaggtca	PDE6A	V738fs	58	300
RP # 22	tgtacaccatggaatgttctact	tgttgtccaatattgtacttgcc	PROM1	L116fs	58	299
RP # 23	ttcttcctttctccccacgg	ataaatacagacccatgcacc	PRPF3	E93G/I113fs	58	280
RP # 24	aaatctcctctcccaccacc	cacgtttcttggagtgcact	PRPH2	E321fs	58	293
RP # 25	cgaagcagtaatcaagagggc	tccccactttccctttcttca	RP1	D433fs	59	293
RP # 26	ggaggacattcactttacttccg	tttcagaagaggacagattggt	RPE65	N194fs	58	204
RP # 27	ctgcttggctgtgtggaag	cagcacgttggtttcctgg	TULP1	N342fs	59	286
RP # 28	acagaaccaaacatcctacatct	atgatgagcacaccaaggga	EYS	G2287fs	58	271
RP # 29	actctcccgtccttcttgtg	gtgtagtcattgtggttcctgt	ABCA4	R1098S	59	278
RP # 30	tcctgtcgccatgtaacaca	acgttggcatacagagtgtc	ABCA4	R212C	58	277
RP # 31	gacagacaccccaggactc	tgctgagttataacccatgcc	ABCA4	A1357V	59	297

RP # 33	gggagtggataagacccctc	ctgagaaggaggaagaggca	RP1L1	ex2	58	460
RP # 43	gcctccttgcatttgaccat	cgagcccttcctgaaacatc	ABCA4	R537H/L541P	59	349
RP#43bis	gcctccttgcatttgaccat	tttctccaaagtcagtgggaa	ABCA4	R537H/L541P	60	375
RP # 44	accacagtgagcatcttgattg	tgtctccatccagctctgag	ABCA4	A1038	59	283
RP # 45	aggaggctgcaggataaatgt	tctaagcagcatgtgaccca	ABCA4	F1397	59	298
RP # 46	tttctcagcagctcaatcca	aaatctactgccctgatcataca	ABCA4	F1533I	57	319
RP # 47	gcgaagacaaagggaggaac	gaccaggatgttagctccgt	LRAT	E52G	59	281
RP # 48	tcaaaattccaaggttcaaggac	tttgccctggttgtagcatg	RP1	F833L	58	384
RP # 49	tcttccagccttttcctccc	ttggcaggtgagagagaagg	IMPDH1	L174P	59	364
RP # 50	aaaggagtgctgaggttcct	cacagcccttcccactctg	BEST1	W127R	59	295
RP # 51	attccttgggtgcctgagat	accgagagatagtttcctgca	NR2E3	A256V	58	273
RP # 52	acttctacgtgcccttctcc	atgcagagaggtgtagaggg	RHO	V61I	59	278
RP # 53	gacaactaaggcccggagag	ggagaccctaactgcccc	CNGB1	Q290L	59	287
RP # 54	gggaggtgtgagtgagctat	tctctctaaagacatggaagcct	ABCA4	Q625X	58	344
RP # 55	tagggcagaatgagaaggcg	cacagccttcttcagctcag	ABCA4	H776L/L751P	59	288
RP # 56	ggcatctgggtgctgttg	acttcaaatgctggttctgaca	ABCA4	L819P	58	394
RP # 57	gggagcctgagaatagccat	cacagagcccaaggagttg	ABCA4	G863A	58	275
RP # 58	cctttgggctctgtctctgt	aacttttgatcctgtccgct	MERTK	N211H/L216H	58	296
RP # 59	aatgcatttgtttggtagctgt	ccctccatgcccatccttg	MERTK	R293H	59	290
RP # 60	actgagttctgaaagctttgact	caggaaaggcataatcacttggt	MERTK	K533E	58	328
RP # 61	tgcatgatcatcagtgtttaagg	taacagcagtccctttcccc	MERTK	N806D	58	347
RP # 62	ccctgactctataattgcctcct	aggacttctgagccttctga	MERTK	D940G	58	300
RP # 63	ctcaccttcgcttctcccc	agactgacataaaagaggatggc	RPE65	P8H	59	299
RP # 64	tcaaaatgtgtttctttgcctgt	acaagaatcaacagtgcctaca	RPE65	C231Y	58	334
RP # 65	ccctttcattcacaagccca	tcagcaatatgaagccaaacct	RPE65	Y275C	58	299
RP # 66	tggattgcacaaagagatggg	tgaacttctgctgtgaccac	MERTK	V870I	58	334
RP # 67	ctggggactataggcaggac	cttgtcggcctccccaga	TULP1	N130Y	59	289
RP # 68	gcaatcaagctgttcccca	ccagtggacaaggtgatgga	EYS	T120M	59	276
RP # 69	ctgtcctccacctgaaccag	gcagttctccagcagcattc	BEST1	S583P	59	288
RP # 70	actgccttggctcactacaa	cttattgcccctctgagcga	ROM1	G173E/F178S	59	272
RP # 71	caggtggtcagatcagggag	gtccctgagccacggaag	PDE6B	E254X	59	284
RP#72	ggaactgactcaacccctttg	tgcaaatcctcggtcttcct	RP1	G1875A	59	391
RP #73	tggaggatgtcaagggaagg	ttagagaagattgggcttggag	CNGB1	A961V	59	299
RP#74	tctctgcgcatttctgtattct	accagttagagtgcaaacaacc	EYS	H2599R	58	385
RP#75	tgctgaagtgattggatggg	gctgagggaaagagagagca	PRPF3	T494M	58	385
RP#77	actctgattttaattgggcttcc	gttgcagggctagaagaaca	PRPF8	R1617X	58	289
RP#78	cgaatcaagagcaacgtgga	ggttcaagcccagactgatc	PRPH2	M262I	58	398
RP#79	aggagttggtaatggcagtagt	tgtcgcaacttaactggtgag	CRB1	C790X	58	392
RP#80	accttccctggcccagtc	agaaggacagctggttcgg	CRX	Q291K	60	295
RP#81	caccaggtgcctctcaaatt	ctggaagcacactcatcgt	CRB1	K97T	58	372
RP#82	cgggacacagacacccct	accctcacatacgcagcaat	PRPH2	S189P	60	300
RP#83	tcaccagccaaaatcaccac	cgcggccagttctttgag	CNGB1	V1096M	59	297
RP#84	tgaaagagtcactgccgaga	tcaaagaaccgccacttctg	ABCA4	T1253M	59	397

RP#85	acactgttcaaacttcaccca	gatctagtgaagggaagatgcc	EYS	S1809P	58	392
RP#86	gcctgacagttgacaatgaca	ccaggaaaaccagagtgttgt	PRPF3	P474L	59	393
RP#87	aagtgaggggagtgactgtg	gaatcctctcctgctcacct	ABCA4	N2111S	59	286
RP#88	agacctgctccattgggc	acaggagtcataaccaataatgct	EYS	S1271C	59	346
RP#89	cctccttgtcccacaggtg	accaatcatgctacccagatg	PRPF8	D931N	58	295
RP#90	tcctgaaaagtacttggcctt	agcattctcccaactagcaga	RP1	T373I	58	400
RP#91	agagcaaagtgggtaggtgg	ggtggggctctctgtagga	ABCA4	G2266X	60	285
RP#92	tcattgctcctgtcccttgg	tggactttatgctcacctgc	CRB1	Y1206H	59	285
RP#93	aaaatttgtttcactcctgggaa	accattcaaaacagtcctgagc	PRPF3	R61X	58	293
RP#94	ctccagggcacacaggaa	ccactgagaggaaggacacg	IMPDH1	E98K	59	294
RP#95	ttctgtgcagtgacgagact	gcatatactcaaagcactgttca	RPE65	E417K	58	300
RP#96	acagtgtgaacgtggcatca	ccttaggtagtgcagccgg	PRPF8	K2034X	60	296
RP#97	ggtccctgtgtgaaatcgc	tacctccacgaaatgccact	CRB1	L522F	59	287
RP#98	tggctgaagaacaagaccaa	tcagtggggaaaatgaaagct	ABCA4	R511H	57	384
RP#99	cttgtttcagaagggcaggc	actgggttcattttgaggcc	TULP1	K91E	59	389
RP#100	actcctggttgactgtgagg	gggagaggcagagaaaggtt	NR2E3	P339T	59	387
RP#101	actacacttggcaagcatgg	gcatacttggctgggtcca	PRPH2	I32V	59	291
RP#102	gatccgtgtgtcaatggagg	ccagagatctaaaatgaatcaaggc	CRB1	R1331H	58	300
RP#103	gcctgtgatgacccatgttc	ctttcccaggcttaatccaaaag	CNGB1	T854A	58	294
RP#104	acctcagcctcccaaagtg	atgaacacatgcgccatacc	PROM1	E590K	59	295
RP#105	gcctgagatttatcccagtgc	gacaatttgtgctgggtcac	EYS	P1318L	58	289
RP#106	ggtgggatggttttgggatt	cctcaactcacctgtgcaag	PRPF8	W1661G	58	296
RP#107	tggccggtatgtatgtctgg	aagatcatctgccggtcaca	CNGB1	A961V	59	280
RP#108	ggggttctgatcacgttttaaga	ggctccatcactccccatta	PRPF8	Y2127N	59	290
RP#109	gctcagagtgtgtccctgat	cctcccatgcacacaggt	RDH12	F254S	59	282
RP#110	ccttctctctcacctgccaa	agaaggatggcaggagttgt	IMPDH1	R203C	59	299
RP#111	ggtctttcttgttgcccagg	ggagaggaggagctggaaga	ABCA4	D2177N	60	399
RP#112	atcacttgaacccgggaagt	ctcagtggctatgtcctgga	PRPF8	Y1043F	59	377
RP#113	tggacctggacaaagaacca	acagataacaccagtctacgct	PROM1	R182Q	59	242
RP#121	cgtcttctgaaggtcaagttcc	ccgtgggaacataggtagga	RP1	D68Y	60	283
RP#122	tcactgacaggagcttaccc	aaaagggccaaattccaaat	PRPF8	K50E	58	300
RP#125	accaagtgtgcgagtgaatg	acgcacccacaggttgag	PRPH2	I196N	60	277
RP#127	ccattgcttttgtcagtgtga	tggaacgaggtggctataca	PRPF8	K892E	59	300
RP#130	ataaataaagcgggcggtga	gctcccattaaaatgcatca	ABCA4	L1201R	60	273
RP#133	catgcagaaatggaggtgaa	tgaactggaggtttctcattct	EYS	D2930G	59	291
RP#140	ggcagctacttcggtgagat	cctggtttgcaggaggtcta	CNGA1	E583G/L585fs/ E596V	60	300
RP#141	cagagaaatgcatcttcaatgtc	cccatgccctctctctcat	PDE6A	D123Y	59	287
RP#147	tcagagtcccctccaaacat	gcccaagagaagcataaaagg	CNGB1	T461M	60	300
RP#154	tcctcacctacctgctcctg	tgacttcccaagttgctgtg	RDH12	S203R	60	276
RP#159	cctgccaacatggaggtatt	cagaataattgtggggacaga	CRB1	165_167del	59	300
RP#161	ggctggtggtgtctctgtgt	gctgtgtacataccacatcactttt	MERTK	746_746del	60	297
RP#162	ggggtgacattgtagaacctg	caaggggttgagtcagttcc	RP1	N1766H	59	294
RP#163	ggccagattggccatatttta	agagcacccggtgaccat	EYS	I1872T	61	295

RP#164	tttccccataactaaaacttatcca	tggaggccttttctgttacatt	EYS	C2749Y	60	298
RP#165	accactccccaccctgct	cccacagtgctggagtacg	PDE6B	D600N	62	300
RP#167	gaacccattcattacctttgaa	ttaagtaaaagttagggttaaaaccag	EYS	N382I	58	299
RP#168	tcctgcaaatgttcaaattca	ttctggacttgtcattttcgtc	EYS	G2348D	59	298
RP#169	ttgtatcttggtaactatccccttt	cctctctttaccacctgcagta	CNGA1	W180X	58	284
RP#170	ccactggaatgcatgtgtgt	aaattctgctctggctgcat	CNGA1	V378A	60	298
RP#171	tcaaatcagcctccaaatcc	tgaagcaaaaacaaagaaacga	EYS	S2052P	60	348
RP#173	ctgggcaacagagcaagact	cctcagagctccttgcaaac	PDE6B	c.2130-2A>T	60	279
RP#174	gtgcacctgagcttgtgtgt	ctcctctcccctgagcaat	PDE6B	c.622-2A>G/ c.711+1G>A	60	398
RP#175	gaatgttatatttttatttgcctctc	agttttgacagctttaatgagaaac	PROM1	c.1051-2A>G	58	300
RP#176	ccaaatgactttctcagtttcca	gggattacaggcgtgagcta	PRPF8	c.3299+1G>A	60	250
RP#177	agcttggcaagggtatatcg	tcttcaccaaggtgatgttca	ABCA4	c.442+1G>A	59	296
RP#178	cagccaggaatggaccatag	agcctctaggtgagcccttc	BEST1	R218H	60	285
RP#200	gacagcagcagaaacagcag	ataggggcggaggtcag	CRX	Y142C	60	281
RP#201	gccaaccagcagagtctacc	tccactcctctccatcgtct	PRPF31	l125fs	60	269
RP#202	tttcagtaccactaaaacagaaggtc	gccatcatagtttagagccaca	EYS	Y3034C	60	295
RP#203	agcttgttccctgaggctgt	ccctgtccctagcctggta	MYO7A	R933H	60	298
RP#204	cagacgtatacgggctccat	cttgatggcatttccgacag	MYO7A	R150Q	60	299
RP#205	gcaacaggagaggctgactt	gtggctaggagggcttgtg	MYO7A	P2123L	60	287
RP#206	cgtgcttagtggaggcagt	aacttcccaaggggtaggg	MYO7A	L370P	60	287
RP#207	caaaggaccgacagctgaa	gatatctgttcaataagcaagtcctc	USH2A	N4856	60	286
RP#208	gcttggtttcctttcagacg	aatgccattgggagattctg	USH2A	W3955X	60	288
RP#209	agtgatgtgatacattctgcattt	tcccaccattaaagcgtttc	ADGRV1	T6063I	59	289
RP#210	gcttatggctgtgtaaacatgg	cgttaggagagcctgggaaa	BBS10	C91fs	60	290
RP#211	tttctgccttatgtcctgaag	gtgcaaacaatcagctcgaa	USH2A	P4632S	58	387
RP#212	gcaatttaaacttgatgttctgtca	cggccaaattcaggatcatt	CNGA1	R280C	61	300
RP#213	gagccagaggcgctctta	cccacacttccaacctgttt	ABCA4	ex1	60	220
RP#214	actgcacacatgggatctga	agcagatctgatccccacac	ABCA4	ex2	60	298
RP#215	cctaggtctgcatcctgctt	gcatttcagcacgtgaagg	ABCA4	ex3	60	272
RP#216	tgggtgacagagcgagataa	tctccataggtgagggaaatg	ABCA4	ex4	60	298
RP#217	taggacgtgggtgtctttcc	ctggagcttttgcaaggatt	ABCA4	ex6	60	381
RP#218	aggagatcagactgtgcctatg	cttttccttgaacaggtttgaa	ABCA4	ex7	58	275
RP#219	tggcctcacagcagattatg	cccaggtttggtttcaccta	ABCA4	ex8	60	393
RP#220	aagtgataggggcagaaaaga	tgggaaaaggaaacaagtgg	ABCA4	ex10	59	288
RP#221	caaacatttattctgcctctacca	ggtggccacatgactaatcc	ABCA4	ex14	60	386
RP#222	agcacatggagtgtgcgtag	ccttccaaggaaacgtgttg	ABCA4	ex15	60	398
RP#223	gtcaaaggctggctaggaaa	cccttgccatgagatgtttt	ABCA4	ex18	60	292
RP#224	tctcactgaacctggtgtgg	gctgggctctagagaagtgtg	ABCA4	ex20	60	295
RP#225	tttaccaacaatgcccaacc	tgctgtgtgctccttctcac	ABCA4	ex23	60	308
RP#226	cccaacaaaacagagcttgg	ctcccaaattgctgggatta	ABCA4	ex26	60	370
RP#227	ccagcctggtatttcattgc	taacccatgcctgaaggaac	ABCA4	ex27	60	485
RP#228	cgagtgttcaggcagcctac	acatggaagtgaccagcaga	ABCA4	ex29	60	393
RP#229	cagaaaccagcctccaagag	ggactcccctagtcctctg	ABCA4	ex30	60	391

RP#230	atcggaaaatgcaaaaatgc	tgtcttgtgacttcctgagca	ABCA4	ex32	60	291
RP#231	tccctacaaaacccgagaga	ggtaaatttttagctccagagcag	ABCA4	ex33-34	60	435
RP#232	attccatccctcctcctgct	tgccacataaatgttggggaga	ABCA4	ex35	60	595
RP#233	gcaacccgcaaatacccatt	acagcagatacacaaggccc	ABCA4	ex36	60	543
RP#234	actaaaggttcccttggcagg	cagaagccagacagatccctg	ABCA4	ex37	60	509
RP#235	gctcttgctcagttcccaca	aaacattgtggagtggggct	ABCA4	ex40	60	358
RP#236	ctggaacccaaggctgactc	agatgctcttggaagacgagt	ABCA4	ex41	60	515
RP#237	gctcgtggcctctgatgg	tgggtctatgtgttgggctt	ABCA4	ex43	60	437
RP#238	tggaaacatggaagaatctgaca	acttggcacacagtagacaca	ABCA4	ex44	59	510
RP#239	gtgcccccaaccttctctc	tcgaccacaggaaattaggcc	ABCA4	ex47	60	509
RP#240	gctcaagtgatcctcccacc	gcctccctcttatggcaattc	ABCA4	ex48	59	503
RP#241	tgtatgcacacatctcgtatgt	ttctggggtctggagaagga	ABCA4	ex50	58	513
RP#242	aggcagcattaggtgctgat	gctccttcatcttgcggta	PRPF31	K327fs	60	371
RP#243	ggtgtacacacgcacacaca	tcccaaggtcacagtgtcag	PRPF31	E185X	60	348
RP#244	ttggtggggtacagcctatt	gtgctaccaggaaggcacat	ABCA4	ex 10	59	288
RP#245	aattgctgggatttcaggtg	cccaataaacagagggcaag	ABCA4	ex48	60	518
RP1L1 1	cctccatcttccagctgct	tgtcccatgcccaacttttg	RP1L1	ex 1	60	498
RP1L1 2	cgcgtgcctctctccttt	tgtctccatgctgttcacct	RP1L1	ex 1	60	541
RP1L1 3	gctctgagaagccccaagg	ttcgctggatccctcatgag	RP1L1	ex3	60	478
RP1L1 4	gacacgctcctatggtccc	gccctgctccatctatgc	RP1L1	ex3	60	532
RP1L1 5	gataggggctgagcggaaa	ctgggtgaggacattgcact	RP1L1	ex3	60	492
RP1L1 6	ctcttccacccccttccacc	cagccccagattttgagcag	RP1L1	ex3	60	497
RP1L1 7	ccactcggattttgtttctgga	gtcttctcctcggacagcc	RP1L1	ex3	60	549
RP1L1 8	gtccccaggccccaattc	ggcctctctgtctgctcc	RP1L1	ex3	60	496
RP1L1 9	ctggggagtagtgacactgg	gtcttccccaacgtcacatc	RP1L1	ex3	60	414
RP1L1 10	gccagcaaagtgaggttcaa	cctctgctcctcactgtctc	RP1L1	ex3	60	447
RP1L1 11	ccctgaaaacctccaaccag	ttcaagagcctctccttgca	RP1L1	ex3	60	482
RP1L1 12	gacagcaagaagaagaggcg	gccaccgaagagctcctc	RP1L1	ex3	60	478
RP1L1 13	cttgagcctggtttggaaaag	cagggtccgctcagagaag	RP1L1	ex3	60	456
RP1L1 14	gggtaggatggtgctggag	gacagtcctagtgctcgtgg	RP1L1	ex3	60	421
RP1L1 15	agaggtcccatcaaagaggc	cccatcactctgtcctggat	RP1L1	ex3	60	456
RP1L1 16	caggagggagagggcataag	ccgcctcttctgcctctt	RP1L1	ex3	60	466
RP1L1 17	gccctggagactgaaggg	cctcttgcatctcccccttca	RP1L1	ex3	60	265
RP1L1 18	agttatagagtcccaggagg	cctcctgtgcctcctctt	RP1L1	ex3	60	323
RP1L1 19	gaagacagaaggggatgc	ggtctctacgccttctgg	RP1L1	ex3	60	565
RP1L1 20	gcccaggatgcagaaggg	tcagaagcctcctcagattgg	RP1L1	ex3	60	353
RP1L1 21	gccagagtcagaaggagaaac	gcatcagggctccttgtgt	RP1L1	ex3	60	270

RP1L1 22	gagacactccccaccaaagg	agtggactgaacgttgctca	RP1L1	ex3	60	439
RP1L1 11bis	ggaggctcaggggagagtag	gtggacgcttccttcttctg	RP1L1	ex3	60	597
RP1L1 19bis	gtgtagaggcccagccaaa	gctccccttttctcacctt	RP1L1	ex3	60	692
RPGR 1	accgtcctctacagcctcct	gggcctgtccccctacag	RPGR	ex 1	61	178
RPGR 2	gagacaaggggtttgtatgga	acacagcagcatatctataacaaaata	RPGR	ex 2	58	300
RPGR 3	tgctttgtggtgacctcatct	tgacattaaagaactacacagtcaaca	RPGR	ex 3	60	239
RPGR 4	gaatcagacaagcccgctat	agccacgttactggaatgaga	RPGR	ex 4	60	247
RPGR 5	tgtctcataaaaaggggggactc	atgctcccttcggtttactg	RPGR	ex 5	59	299
RPGR 6	ttcagagcctggctaccttt	tttaataacaacatagaagtgggaga	RPGR	ex 6	58	261
RPGR 7	tttgacggtaagaccagcttt	aaaatttcattagccaccacaga	RPGR	ex 7	59	285
RPGR 8	ccccagaggcacttaaccttt	ttccatttttctcagccatt	RPGR	ex 8	59	394
RPGR 9	gcagaatttgaggggggataa	tgtctgacatttggcttttagg	RPGR	ex 9	59	396
RPGR 10	agagagattcaccaagccagtc	tcttctagttttctttgcagtgct	RPGR	ex 10	59	378
RPGR 11	tgttggcatactttgaactttct	gctctaaccagggagagaacaa	RPGR	ex 11	59	400
RPGR 12	ttctgtccagttgcctttca	cacacattcacatacacaatgaact	RPGR	ex 12	59	228
RPGR 13	tctgaactgttttcccttcattt	ttctttccaaaaaccatccaa	RPGR	ex 13	59	290
RPGR 14	gcaaaatatctggtaaaacttgc	cctctttttgtaaaccctctcc	RPGR	ex 14	58	386
RPGR 15	ggaaattgattgaacaaggaaa	taccttcaggcccatcctct	RPGR	ex 15	59	300
RPGR 16	cagcaatatcaaatccctcgt	tttctgatcccaaaggcaac	RPGR	ex 16	60	356
RPGR 17	gaacctgcttaaagattcagactt	cccaattatggcatacatacaca	RPGR	ex 17	58	219
RPGR 18	cccacaactcagggttcata	ggggccatgaaaactcatatt	RPGR	ex 18	59	299
RPGR 19	ggtgccccataaagagatga	tggtcacacttatagaagctgaat	RPGR	ex 19	58	475
ORF15-1	aggaaggagcagaggattca	ccctcttcttccattcttcc	RPGR	ORF15	59	/
ORF15-2	ggggagaaagacaagggtag	tcctttcccctcctctactt	RPGR	ORF15	59	/
ORF15-3	ggaagaaggagaccaaggag	cccatttccctgtgtgttag	RPGR	ORF15	59	/
ORF15-4	gcaggatggagaggagtaca	gagagaggccaaaatttacca	RPGR	ORF15	59	/
ORF15-3 ALT	gggagaacgtgaaaaggagga	cttcttcgcctgtctcctga	RPGR	ORF15	59	/
ORF15-3 bis	ggagaacgtgaaaaggagga	tgccatatttcacagatccttt	RPGR	ORF15	59	/
RP2-1	ttcacgccacactctaggaa	ctcagggaggctgagacg	RP2	ex 1	59	208
RP2-2	ctggcagccaatagtccttt	gggttgcctgtatttccaat	RP2	ex 2	59	850
RP2-3	ggaaatcattgttttagtctcagaa	ggcaaaagtaaatcaaatcagca	RP2	ex 3	59	247
RP2-4	ccctcaaagaagatgctgga	aaggaaaaggagatgaaatgaaaa	RP2	ex 4	60	247
RP2-5	ctgtggagaacatgggcttt	tcacaagttgggaaaggctta	RP2	ex 5	60	208
RP2-5bis	tgcttccaaaagagcttaattt	tttttgaacctcaagtaaatgatgtaa	RP2	ex 5	60	467

Table 3. Primers used for Sanger confirmation of NGS identified variants and analysis of RP2, RPGR and RP1L1.

Different thermal profiles were developed and optimized. The first one, used for shorter amplicons, includes 10 minutes at 95°C, followed by forty cycles of denaturation for 20 seconds at 95°C, annealing for 20 seconds at 60°C and extension for 30 seconds at 72°C while in the second, employed for the longer amplicons, the times of denaturation (30s), annealing (30s) and elongation (1min) were extended. For the analysis of RP2 and RPGR PCR was performed in a reaction volume of 31µl containing ReadyMixTM REDTaqR PCR (Sigma Seelze, Germany), water, primers forward and, respectively, reverse 10µM with an appropriate thermal profile: initial denaturation for 1 min at 95° C, 16 cycle of amplification with 95° C for 30 s, 58 or 60° C for 1 min, 68° C for 11 min and, then, 94° C for 1 min, 55° C for 1 min and 72°C final extension for 10 min.

The quality of PCR products was tested by agarose gel electrophoresis 1%, using TBE buffer (Tris-boric acid buffer, EDTA 0.5 M pH 8.0). In the gels were loaded 10µl of each PCR product and 2µl of loading dye (6X). The staining used for the DNA visualization was Gel Red (0.25µg/µl). The DNA bands were visualized using the Gel Doc 1000 (BioRad system, Wien, A) (Figure 7).

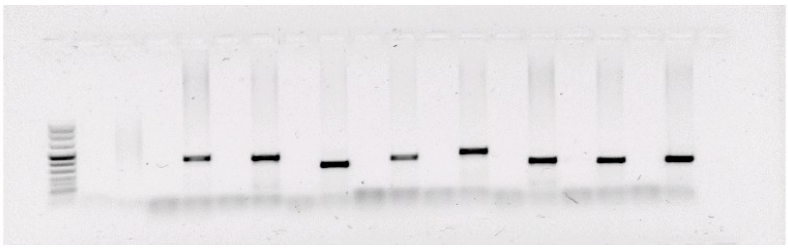

Figure 7. DNA electrophoresis on agarose gel of seventeen DNA samples from different patients. Seventeen different couples of primers were used. Order of loading: blank/DNA sample for every couple of primer.

Sanger Sequencing

PCR amplicons were treated with ExoSAP-IT® (ThermoFisher Scientific, Waltham, MA) to clean up the PCR product as manufacture's protocol. The purified fragments were then submitted to a phase of Big Dye Terminator. The Big Dye mix contains 4µl of Buffer 5X, 2µl of Big Dye terminators, 3,2µl of primer 1µM. From 3 to 5µl of PCR products were added (final reaction volume of 20µl). For the amplification, the following thermal profile was followed: a first step of 95°C for 5 minutes, continued by thirty-five cycles of 96°C for 10 seconds, 50°C for seconds, 60°C for 4 minutes. The amplification products were then purified and sequenced on an AB 3130xl Automated Capillary DNA

Sequencer, according to the manufacturer's protocols (Applied Biosystems, Carlsbad, CA), using the BigDye^{R} Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems, Carlsbad, CA).

For the mutation screening of RP2 and RPGR genes lastly, three-dimensional models of wild-type and mutant proteins were generated and analyzed through the Swiss-PDBViewer (version 4.1) and PyMOL softwares starting from the crystal structure of RPGR protein (RPGR RCC1-like domain available in the Protein Data Bank-PDB ID code 4JHN).

DNA Profiling test

To avoid mistakes in the identification of the samples analyzed with NGS technology by IGA and the correspondent DNA in the biobank, a DNA profiling test was carried out after NGS analysis, on the remaining DNA. The result was compared with the genetic profile of the corresponding sample in the biobank. AmpFlSTR® Identifiler® Plus kit was used for the amplification of the samples through multiplex PCR. This kit co-amplifies 15 STRs (tetranucleotide repeats), and a chromosomal locus specific for the discrimination of male-female. The STR in analysis are the described in the table above (Table 4):

Polymorhism	Cromosomic Location	Repeated Sequence	Bp Range	Allele Nomenclature
D3S1358	3p	TCTA $(TCTG)_{1-3}$ $(TCTA)_n$	114-142	From 12 to 19
vWA	12p12-pter	TCTA $(TCTG)_{3-4}$ $(TCTA)_n$	157-197	From 11 to 21
FGA	4q28	$(TTTC)_3$TTTT TTCT $(CTTT)_n$ CTCC $(TTCC)_2$	219-267	From 18 to 30
Amelogenin	Xp22.1-22.3 Yp11.2	-	107 113	X Y
D8S1179	8	$(TCTR)_n$	128-168	From 8 to 19
D21S11	21	$(TCTA)_n$ $(TCTG)_n$[$(TCTA)_3$TA$(TCTA)_3$ TCA$(TCTA)_2$TCCATA]$(TCTA)_n$	189-243	From 25 to 36
D18S51	18q21.3	$(AGAA)_n$	273-341	From 9 to 26
D5S818	5q21-31	$(AGAT)_n$	135-171	From 7 to 16
D13S317	13q22-31	$(GATA)_n$	206-234	From 8 to 15
D7S820	7q11.21-22	$(GATA)_n$	258-294	From 6 to 15
CSF1PO	5q33.1	(AGAT)n	304-341	From 6 to 15

TH01	11p15.1	$(AATG)_n$	162-201	From 4 to 13.3
D16S539	16q24.1	$(GATA)_n$	252-292	From 5 to 15
D2S1338	2q35	$(TGCC)_n(TTCC)_n$	307-358	From 15 to 28
D19S433	19q12	$(AAGG)_n(AAAG)_n(AAGG)_n(TAGG)_n(AAGG)_n$	101-135	From 9 to 17.2
TPOX	2p25.3	(AATG)n	221-249	From 6 to 13

Table 4. Representation of STR polymorphisms, sequences and chromosomic localizations.

For each DNA sample, the PCR-multiplex reaction mix was composed by 2,5µl of AmpFlSTR® Identifiler® Plus Master Mix, 1,25µl of AmpFlSTR® Identifiler® Plus Primer Set and 2,5µl of DNA. For every PCR reaction, a negative control was utilized. For amplification GeneAmpR PCR System 9700 (Applied Biosystems, Carlsabad, CA) thermocycler was used, with the following program: 11 minutes of initial incubation at 95°C, 30 cycles of denaturation (20 seconds at 94°C) and annealing (3 minutes at 59°C) and a final extension phase of 10 minutes at 60°C.

Amplification products were then analyzed by means of capillary electrophoresis on ABI PRISM 3130 Genetic Analyzer (Applied Biosystems, Carlsabad, CA) using the GeneMapperID ver. 3.2 software, GeneScan 500 LIZ internal standard and the allelic ladder. Before sequencing, amplification products were prepared with a mix of formamide (15µl for every sample) and internal standard (0,5µl for each sample) and 1µl of sample or 1µl of allelic ladder.

The result consists of an electropherogram with some picks (one or two for each STR) that were automatically compared with a reference ladder. This comparison is necessary for the determination of the genetic profile of each subject. Allelic determination took place after sequencing by the software GeneScan Analysis 3.7.

qPCR

TaqMan real-time polymerase chain reaction was applied to determine the genotypes of 106 patients with IRD (53 males and 53 females) and 12 control subjects (5 males and 7 females with already known genotype for analysed SNP). TaqMan real-time polymerase chain reaction was performed with ViiA7 instrument (Thermo Fisher Scientific) to genotype rs10490924, rs1061170, rs2230199, rs403846, rs641153, rs9332739 and rs1410996 polymorphisms, following the manufacturer's instructions. Based on scientific literature, the genotypes at risk are GT and TT for rs10490924; TC and CC for rs1061170; CG and GG for rs2230199; GA and AA for rs403846; GA and AA for

rs641153; CG and CC for rs9332739 and CT and CC for rs1410996. The elaboration of the genotyping data was analysed by AMD Genetic Test Software.

Results

Patients recruitments and clinical evaluation

Patients involved in this study were clinically examined at the Center for Retinitis Pigmentosa of the Veneto Region, in the ULSS 15 Regional Hospital of Camposampiero (Italy), with a diagnosis of different forms of Inherited Retinal Disorders and Occult Macular Dystrophy on the base of clinical symptoms.

The IRD-NGS cohort, made up 190 unrelated patients, 100 females and 90 males, was selected in the database on the base of the pattern of inheritance, considering the autosomal dominant and recessive one. According to clinical data, considering symptoms, their onset and the results of instrumental examinations, in the cohort selected subjects are affected by Retinitis Pigmentosa (RP), inverse RP, RP associated to Usher Syndrome (type I, II and/or III), RP plus Laurence-Moon-Biedl and plus Senior Syndrome, RP associated to hypoacusis Best Syndrome, Stargardt disease, and Cones Dystrophy. Patients affected by the same disease show variable symptoms due to clinical heterogeneity.

X-linked probands (n=60) were selected for the analysis of RP2 and RPGR genes, while 5 patients underwent RP1L1 sequencing.

Before to be part of this project, all the subjects were informed of the purpose and the technique used and were invited to give their consent for the study.

Genetic biobank

To proceed with the genetic characterization of the diseases, a genetic biobank was organized. The creation of a biobank requires some critical step (Figure 8), starting from the acquisition of data and informed consent, until the obtainment of biological samples and their storage and management. The most important requisite for the creation of this sample collection is the traceability of the samples, registering all the variables during the collection and the storage processing (Caenazzo, 2012).

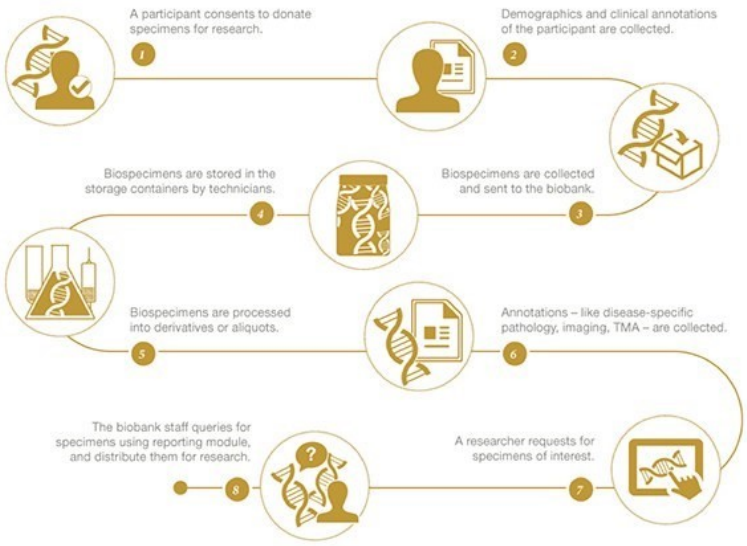

Figure 8. Typical workflow for the creation of a biobank, regardless of the type of specimen stored.

The genetic biobank consists of a collection of genetic samples from every patient presenting an Inherited Retinal Degeneration, which comprises not only Retinitis Pigmentosa, but also Leber Congenital Amaurosis, Stargardt disease, Usher syndrome and many other disorders that damage retinal tissue.
Blood samples were collected from every donor and conserved for the subsequent extraction of genomic DNA. Then, the extracted DNA was used for the genetic diagnosis of causative mutations in each patient. The IRD and OMD genetic biobank stored at the University of Padua consists of a huge number of samples and data: 620 different DNAs each with a storing number, which connects every specimen to the data of its owner (name and surname, sampling date, informed consent, birth date, phenotype, inheritance pattern and all the clinical notes of the ophthalmologist) following internal traceability criteria.

Panel optimization and mutation identification

To identify the pathogenic mutations in the selected patients affected by IRD (n=190), a targeted NGS that covers all coding exons of 25 known retinal disease genes (Table 2) was performed. The patients were divided in two cohorts of 96 run analyzed in two different runs. The two runs differ for the analyzed genes: the first one, identifiable as panel a (Table 2), covers the exome of ABCA4, BEST1, CNGA1, CNGB1, CRB1, CRX, EYS, IMPDH1, LRAT, MERTK, NR2E3, PDE6A, PDE6B, PROM1, PRPF3, PRPF31, PRPF8, PRPH2, RDH12, RHO, RLBP1, RPE65, ROM1, RP1, TULP1 genes reaching a depth of coverage of 341,4x; while the panel b, made by a new design of the panel and the run, sequenced the exons of ABCA4, ADGRV1, BBS1, BBS10, BEST1, CDH23, CNGA1, CNGB1, CRB1, CRX, EYS, IMPDH1, MYO7A, NR2E3, PDE6A, PDE6B, PRPF3, PRPF31, PRPF8, PRPH2, RHO, RPE65, RP1, USH1C, USH2A, with a depth of coverage of 382,1x (Table 5). The depth of coverage represents the mean deep of the sequencing, a parameter indicating the number of times in which a single nucleotide of the genetic panel is sequenced. The coverage, instead, is

the estimation of sequence percentage obtained with respect to the sequence selected during the preparation of the libraries or at the design stage of the genetic panel. For the panel a, the percentage of sequences that were read more than 50 times is 50,74%, while for panel b, after the optimization of the panel and run design, this value reached 88,72%.

	Panel a	**Panel b**
N. of analized genes	25	25
Total bases of target (CDS)	68306	109582
Mean depth of coverage (DoC)	341,4x	382,1x
Coverage% >10x	69,48%	95,73%
Coverage% >20x	62,32%	93,18%
Coverage% >50x	50,74%	88,72%

Table 5. The optimization of experiment parameters and panel has increased the yield of the screening, as visible from the increase in coverage values.

NGS results were bioinformatically filtered, as described in the Material and Methods chapter. Data output supplied by IGA services were organized in excel files, in which many information about the identified variations were categorized, thus providing details about the position in the gene where the variants have been identified, their frequency in the population, the prediction of pathogenicity by means predictor software and quality information of the run (Table 6):

OUTPUT	**MEANING**
Chr; Start;End	The chromosome and the position give the contig on which the variant occurs
Ref; Var	The reference and alternative base that vary in the samples

Region	Variant hits exon or intergenic region, or intron, or a non-coding RNA gene.
Gene	Gene name
GeneDetail.refGene	The transcript identifier and the sequence change in the corresponding transcript
Exon_function	The functional consequences of the exonic variant
Annotation	Gene name, the transcript identifier and the sequence change
ESP6500_ALL	MAF in US Exome Sequencing Project dataset (6500 exomes) for all populations
1000g2012apr_all	MAF in 1000Genomes April 2012 release.
dbSNP138	rs# code from the dbSNP version 138
AVSIFT_score	Assigns a "functional importance" score to SNPs
LJB_PhyloP_Pred	Pathogenicity score from dbNSFP (conserved/not conserved)
LJB_SIFT_Pred	Pathogenicity call from dbNSFP: T -tolerated, D -deleterious.
LJB_PolyPhen2_Pred	Pathogenicity call: D-probably damaging, P-possibly damaging, B-benign.
LJB_LRT_Pred	Pathogenicity score from dbNSFP: D – deleterious
LJB_MutationTaster_Pred	Pathogenicity score from dbNSFP: ‘disease_causing', 'polymorphism'
LJB_GERP++	Nucleotide conservation score from dbNFSP GERP
clinvar_20140929	Archive of reports of the relationships among human variations and phenotypes
Cosmic65	Catalogue of somatic mutations in cancer ID.
Filter flag	Variant reliability
Status	Heterozygous/homozygous
Cov	Reference allele, variant allele and Total depth of reads
Qual	Genotype quality encoded as the probability that the genotype call is wrong.

Table 6. Legend of information obtained by NGS results. Chr: chromosome; Ref: reference; Var: variant; Cov: coverage; Qual: quality.

To identify the pathogenic variant a stepwise mutation identification strategy was applied, as previously described. Sanger confirmation was performed for the putative mutations' confirmation (Figure 9).

Figure 9. Electropherograms of three diverse types of identified mutation: A) missense; B) nonsense and C) frameshift deletion. The mutated base is underlined by the red arrow.

The already known mutations were recognized by consulting the public databases, such as the HGMD database (Stenson et al. 2014), while novel variants were annotated for their impact on protein coding. Pathogenicity was determined by searching in databases and scientific literature, and with the assistance of prediction programs such as Mutation Taster and PolyP2. Interspecies conservation was taken in account too, working with Basic Local Alignment Search Tool (BLAST), thus understanding the criticality of the involved amino acids for the right functioning of the protein encoded by that gene.

Pathogenic mutations were found in 126 patients with a solving rate of 66% (tables 7-8).

Patient ID	**rs ID**	**NM ID**	**Gene**	**Nucleotidic change**	**Protein change**	**Status**	**Pathogenicity**
41	rs137853190	NM_001142800	EYS	exon43:c.T9405A	Y3135X	Het	Pathogenic
44		NM_001289395	TULP1	exon4:c.212_235delCCTCCTCTTCCTCGTCCTCCTCGT	71_79del	Het	Pathogenic
	rs145525174	NM_000350	ABCA4	exon12:c.G1654A	V552I	Het	lp
45		NM_006269	RP1	exon4:c.1298_1299insTATCATCC	D433fs	Hom	Pathogenic

46	rs140482171	NM_000350	ABCA4	exon11:c.G1532A	R511H	Het	Pathogenic
52		NM_201253	CRB1	exon2:c.493_501delGATGGAATT	165_167del	Het	Pathogenic
58	rs1800553	NM_000350	ABCA4	exon42:c.G5882A	G1961E	Het	Pathogenic
		NM_000283	PDE6B	exon4:c.G760T	E254X	Het	Pathogenic
60	rs1800553	NM_000350	ABCA4	exon42:c.G5882A	G1961E	Het	Pathogenic
129		NM_001297	CNGB1	exon23:c.G2266C	G756R	Het	P
171		NM_000539	RHO	exon3:c.C538T	P180S	Het	PP
	rs374847529	NM_000322	PRPH2	exon3:c.T923A	L308Q	Het	PP
173		NM_000322	PRPH2	exon2:c.T587A	I196N	Het	P
174		NM_000087	CNGA1	exon11:c.1755_1758delAGTT	L585fs	Hom	Pathogenic
176		NM_206933	USH2A	exon6:c.923_924insGCCA	H308fs	Het	Pathogenic
		NM_206933	USH2A	exon11:c.C1876T	R626X	Het	Pathogenic
177		NM_006343	MERTK	exon17:c.2236_2238delAAG	746_746del	Het	Pathogenic
178	rs1800553	NM_000350	ABCA4	exon42:c.G5882A	G1961E	Het	Pathogenic
179		NM_001145847	PROM1	exon4:c.347dupT	L116fs	Het	Pathogenic
180	rs28937873	NM_014249	NR2E3	exon6:c.G932A	R311Q	Het	Pathogenic
		NM_206933	USH2A	exon27:c.T5498A	V1833E	Het	Pathogenic
		NM_206933	USH2A	exon11:c.C1876T	R626X	Het	Pathogenic
182		NM_201253	CRB1	exon7:c.2531_2537delATGGTGG	N844fs	Het	Pathogenic

183	rs62645748	NM_201253	CRB1	exon9:c.G2843A	C948Y	Het	Pathogenic
	rs143275144	NM_206933	USH2A	exon66:c.T14519C	L4840P	Het	PP
		NM_201253	CRB1	exon2:c.G254T	C85F	Het	PP
187		NM_000350	ABCA4	exon46:c.6282+1G>C		Hom	P
189	rs76157638	NM_000350	ABCA4	exon17:c.G2588C	G863A	Het	Pathogenic
191		NM_153676	USH1C	exon12:c.C880T	R294W	Het	PP
195	rs1800553	NM_000350	ABCA4	exon42:c.G5882A	G1961E	Het	Pathogenic
196	rs147593839	NM_001297	CNGB1	exon28:c.G2824A	T461M	Het	PP
210		NM_015629	PRPF31	exon10:c.982dupT	K327fs	Het	Pathogenic
212		NM_015629	PRPF31	exon5:c.373delC	L125fs	Het	Pathogenic
213		NM_000322	PRPH2	exon2:c.T587A	I196N	Het	Pathogenic
214		NM_206933	USH2A	exon13:c.C2720T	T907I	Het	Pathogenic
		NM_000539	RHO	exon2:c.C491T	A164V	Het	Pathogenic
228		NM_000350	ABCA4	exon27:c.C4070T	A1357V	Het	Pathogenic
	rs61750200	NM_000350	ABCA4	exon6:c.C634T	R212C	Het	Pathogenic
229	rs374847529	NM_000322	PRPH2	exon3:c.T923A	L308Q	Het	PP
230		NM_000283	PDE6B	exon14:c.G1798A	D600N	Hom	Pathogenic
		NM_201253	CRB1	exon2:c.T614C	I205T	Het	PP
251		NM_024685	BBS10	exon2:c.C389T	S130F	Het	PP

253	rs61750126	NM_000350	ABCA4	exon24:c.T3602G	L1201R	Het	Pathogenic
264		NM_024685	BBS10	exon2:c.271dupT	C91fs	Het	Pathogenic
		NM_206933	USH2A	exon64:c.C13894T	P4632S	Het	P
	rs367772790	NM_000087	CNGA1	exon11:c.C838T	R280C	Het	P
		NM_000087	CNGA1	exon9:c.540delG	W180X	Het	Pathogenic
265	rs367772790	NM_000087	CNGA1	exon11:c.C838T	R280C	Het	PP
		NM_000087	CNGA1	exon9:c.540delG	W180X	Het	Pathogenic
281	rs61750155	NM_000350	ABCA4	exon34:c.C4793A	A1598D	Het	Pathogenic
286		NM_001142800	EYS	exon26:c.G5511A	W1837X	Hom	Pathogenic
296		NM_206933	USH2A	exon54:c.10699delC	L3567X	Het	Pathogenic
301		NM_000539	RHO	exon1:c.A83G	Q28R	Het	Pathogenic
316		NM_000350	ABCA4	exon13:c.C1912T	P638S	Het	P
		NM_206933	USH2A	exon10:c.A1772G	D591G	Het	P
		NM_206933	USH2A	exon5:c.T798A	F266L	Het	P
317		NM_015629	PRPF31	exon7:c.G553T	E185X	Het	Pathogenic
318	rs77775126	NM_006269	RP1	exon4:c.C1118T	T373I	Het	Pathogenic
320	rs61751384	NM_000350	ABCA4	exon47:c.G6449A	C2150Y	Het	Pathogenic
	rs1800553	NM_000350	ABCA4	exon42:c.G5882A	G1961E	Het	Pathogenic
330		NM_206933	USH2A	exon6:c.923_924insGCCA	H308fs	Het	Pathogenic

338		NM_206933	USH2A	exon61:c.11918delC	A3973fs	Het	Pathogenic
		NM_206933	USH2A	exon2:c.C187T	R63X	Het	Pathogenic
344	rs61749440	NM_000350	ABCA4	exon18:c.C2690T	T897I	Het	Pathogenic
347	rs147593839	NM_001297	CNGB1	exon17:c.C1382T	T461M	Het	P
		NM_152443	RDH12	exon7:c.C609A	S203R	Het	P
353		NM_206933	USH2A	exon6:c.G1001A	R334Q	Het	Pathogenic
355	rs1800553	NM_000350	ABCA4	exon42:c.G5882A	G1961E	Het	Pathogenic
357		NM_000539	RHO	exon1:c.G181A	V61I	Het	PP
	rs1800553	NM_000350	ABCA4	exon42:c.G5882A	G1961E	Het	Pathogenic
		NM_201253	CRB1	exon3:c.C694A	P232T	Het	P
		NM_001289395	TULP1	exon11:c.C980T	T327M	Het	PP
		NM_000350	ABCA4	exon46:c.6282+1G>C		Het	P
364		NM_000283	PDE6B	exon1:c.G313A	E105K	Het	P
		NM_000283	PDE6B	exon4:c.T739A	F247I	Het	P
369	rs147593839	NM_001297	CNGB1	exon17:c.C1382T	T461M	Het	P
370		NM_000350	ABCA4	exon6:c.T688A	C230S	Hom	Pathogenic
371		NM_000322	PRPH2	exon2:c.T587A	I196N	Het	P
374	rs1800552	NM_000350	ABCA4	exon40:c.G5693A	R1898H	Het	Pathogenic
377	rs1800553	NM_000350	ABCA4	exon42:c.G5882A	G1961E	Het	Pathogenic

379		NM_001289395	TULP1	exon4:c.212_235delCCTCCTCTTCCTCGTCCTCCTCGT	71_79del	Het	Pathogenic
		NM_000350	ABCA4	exon48:c.C6647T	A2216V	Het	Pathogenic
381		NM_000087	CNGA1	exon11:c.T1133C	V378A	Het	P
		NM_000350	ABCA4	exon16:c.T2456A	L819Q	Het	P
383	rs61750126	NM_000350	ABCA4	exon24:c.T3602G	L1201R	Hom	Pathogenic
385	rs121434241	NM_004698	PRPF3	exon11:c.C1481T	T494M	Het	Pathogenic
386	rs77775126	NM_006269	RP1	exon4:c.C1118T	T373I	Het	Pathogenic
387		NM_000260	MYO7A	exon44:c.C6049T	Q2017X	Het	Pathogenic
389		NM_000322	PRPH2	exon2:c.T587A	I196N	Het	P
391		NM_000350	ABCA4	exon9:c.T1140A	N380K	Het	Pathogenic
		NM_000322	PRPH2	exon2:c.T587A	I196N	Het	P
392		NM_001142800	EYS	exon43:c.C9177G	Y3059X	Hom	Pathogenic
393		NM_206933	USH2A	exon22:c.G4628T	G1543V	Het	P
400		NM_201253	CRB1	exon6:c.G1732C	V578L	Het	Pathogenic
	rs111033364	NM_206933	USH2A	exon61:c.G11864A	W3955X	Hom	Pathogenic
		NM_014249	NR2E3	exon2:c.G166A	G56R	Het	Pathogenic
		NM_006269	RP1	exon4:c.A2628T	K876N	Het	
405		NM_000322	PRPH2	exon2:c.T587A	I196N	Het	P
434	rs1800553	NM_000350	ABCA4	exon42:c.G5882A	G1961E	Het	Pathogenic

436	rs35520756	NM_024649	BBS1	exon8:c.G700A	E234K	Het	P
451		NM_000350	ABCA4	exon45:c.6221delG	G2074fs	Het	Pathogenic
		NM_201253	CRB1	exon2:c.493_501delGATGGAATT	165_167del	Het	Pathogenic
452	rs1800553	NM_000350	ABCA4	exon42:c.G5882A	G1961E	Het	Pathogenic
455	rs1800553	NM_000350	ABCA4	exon42:c.G5882A	G1961E	Het	Pathogenic
	rs80338902	NM_206933	USH2A	exon13:c.G2276T	C759F	Het	Pathogenic
459	rs1800553	NM_000350	ABCA4	exon42:c.G5882A	G1961E	Het	Pathogenic
460		NM_206933	USH2A	exon54:c.10699delC	L3567X	Het	Pathogenic
		NM_006269	RP1	exon4:c.4453delG	G1485fs	Het	Pathogenic
466	rs17544552	NM_032119	ADGRV1	exon28:c.C5780T	T1927M	Het	Pathogenic
472	rs104894082	NM_006269	RP1	exon4:c.C2029T	R677X	Het	Pathogenic
473		NM_032119	ADGRV1	exon31:c.C6778T	R2260X	Het	Pathogenic
476	rs61754030	NM_000350	ABCA4	exon18:c.A2701G	T901A	Het	Pathogenic
		NM_000283	PDE6B	exon1:c.G313A	E105K	Het	P
		NM_000283	PDE6B	exon4:c.T739A	F247I	Het	P
	rs17544552	NM_032119	ADGRV1	exon28:c.C5780T	T1927M	Het	
480	rs61751418	NM_000350	ABCA4	exon8:c.G982T	E328X	Het	Pathogenic
487	rs41281314	NM_022124	CDH23	exon30:c.A3625G	T1209A	Het	Pathogenic
	rs121434236	NM_006445	PRPF8	exon43:c.A6926G	H2309R	Het	Pathogenic

489		NM_006269	RP1	exon2:c.G148C	G50R	Het	Pathogenic
491	rs28939720	NM_201253	CRB1	exon7:c.C2234T	T745M	Het	Pathogenic
493		NM_000440	PDE6A	exon17:c.2100_2122delTCCGTGTCTGCTCCAGCATCATG	Y700fs*700	Het	Pathogenic
519		NM_000260	MYO7A	exon4:c.C268T	R90W	Het	
520		NM_000329	RPE65	exon14:c.C1579T	H527Y	Hom	P
528	rs139799843	NM_206933	USH2A	exon66:c.A14567G	N4856S	Het	PP
530	rs61751392	NM_000350	ABCA4	exon12:c.T1622C	L541P	Het	Pathogenic
	rs1800553	NM_000350	ABCA4	exon42:c.G5882A	G1961E	Het	Pathogenic
572	rs76157638	NM_000350	ABCA4	exon17:c.G2588C	G863A	Het	Pathogenic
574		NM_000260	MYO7A	exon29:c.C3659T	P1220L	Het	Pathogenic
576	rs199636364	NM_000087	CNGA1	exon4:c.C94T	R32X	Hom	Pathogenic
578		NM_000350	ABCA4	exon10:c.1302delA	V434fs	Het	Pathogenic
		NM_000087	CNGA1	exon11:c.1755_1758delAGTT	L585fs	Hom	Pathogenic
580	rs181255269	NM_022124	CDH23	exon20:c.C2263T	H755Y	Het	Pathogenic
582		NM_206933	USH2A	exon59:c.11411delC	P3804fs	Hom	Pathogenic
	rs202205304	NM_024649	BBS1	exon15:c.G1535A	R512H	Het	Pathogenic

Table 7. Pathogenic mutations identified in 86 patients. For each patient are indicated the gene in which the mutation occurs and its nucleotide and protein change. The pathogenicity assessment is referred to reference (Pathogenetic) or to prediction software (P, PP). Abbreviations: P: pathogenetic for prediction software; PP: probably pathogenetic for prediction software; C: conserved; NC: not conserved.

	rs ID	NM ID	Gene	Nucleotidic change	Protein change	Status	Cons
14	rs112002818	NM_001297	CNGB1	exon28:c.C2882T	A961V	Het	C
25	rs376128303	NM_001297	CNGB1	exon32:c.G3286A	V1096M	Het	C
53		NM_000327	ROM1	exon1:c.G47A	R16H	Het	C
61		NM_000440	PDE6A	exon1:c.G367T	D123Y	Het	NC
129		NM_022124	CDH23	exon15:c.G1672A	V558M	Hom	C
170	rs112002818	NM_001297	CNGB1	exon28:c.C2882T	A961V	Hom	C
		NM_001298	CNGB1	exon28:c.A2857C	N953H	Het	NC
171	rs61749435	NM_000350	ABCA4	exon16:c.T2546C	V849A	Het	NC
		NM_032119	ADGRV1	exon21:c.G4666A	E1556K	Het	C
	rs77469944	NM_000440	PDE6A	exon1:c.G103A	D35N	Het	C
173		NM_000322	PRPH2	exon2:c.T587A	I196N	Het	C
	rs112002818	NM_001297	CNGB1	exon28:c.C2882T	A961V	Het	C
		NM_001298	CNGB1	exon28:c.A2857C	N953H	Het	NC
178	rs112002818	NM_001297	CNGB1	exon28:c.C2882T	A961V	Het	C
		NM_001298	CNGB1	exon28:c.A2857C	N953H	Het	NC
185	rs201333192	NM_001142800	EYS	exon26:c.T5615C	I1872T	Het	C
190		NM_001145847	PROM1	exon16:c.G1768A	E590K	Het	C
191		NM_032119	ADGRV1	exon21:c.G4433A	G1478D	Het	C
		NM_032119	ADGRV1	exon68:c.13655dupT	V4552fs	Het	C
		NM_153676	USH1C	exon17:c.G1414T	V472L	Het	C
192	rs112002818	NM_001297	CNGB1	exon28:c.C2882T	A961V	Het	C
		NM_001298	CNGB1	exon28:c.A2857C	N953H	Hom	NC
195		NM_001142573	IMPDH1	exon5:c.A514G:p.T172A	T172A	Het	C
205		NM_001142800	EYS	exon7:c.A1069C	I357L	Het	C
		NM_006269	RP1	exon4:c.A3781G	M1261V	Het	NC
206	rs144929093	NM_001142573	IMPDH1	exon6:c.C607T	R203C	Het	C
211	rs112002818	NM_001297	CNGB1	exon28:c.C2882T	A961V	Het	C
		NM_001298	CNGB1	exon28:c.A2857C	N953H	Hom	NC
213	rs112002818	NM_001297	CNGB1	exon28:c.C2882T	A961V	Het	C
		NM_001298	CNGB1	exon28:c.A2857C	N953H	Hom	NC

229		NM_000322	PRPH2	exon3:c.T923A	L308Q	Het	C
237		NM_024685	BBS10	exon2:c.C1307T	T436I	Het	NC
246	rs112002818	NM_001297	CNGB1	exon28:c.C2882T	A961V	Het	C
		NM_001298	CNGB1	exon28:c.A2857C	N953H	Het	NC
296		NM_206933	USH2A	exon56:c.T10996G	C3666G	Het	C
298	rs376128303	NM_001297	CNGB1	exon32:c.G3286A	V1096M	Het	C
299	rs202245413	NM_000260	MYO7A	exon5:c.G449A	R150Q	Het	C
310		NM_000260	MYO7A	exon16:c.G1817A	R606H	Het	C
313	rs144612129	NM_000329	RPE65	exon8:c.A845G	N282S	Het	C
		NM_022124	CDH23	exon56:c.C8285T	A2762V	Het	C
320	rs41305898	NM_032119	ADGRV1	exon20:c.C4214T	S1405F	Het	C
		NM_000260	MYO7A	exon3:c.G121T	D41Y	Het	C
321	rs112002818	NM_001297	CNGB1	exon28:c.C2882T	A961V	Het	C
	rs12927214	NM_001297	CNGB1	exon18:c.T1604C	V535A	Het	C
323	rs112002818	NM_001297	CNGB1	exon28:c.C2882T	A961V	Het	C
324	rs201031527	NM_000087	CNGA1	exon5:c.G191T	G64V	Het	NC
		NM_000260	MYO7A	exon47:c.C6368T	P2123L	Het	C
329	rs375321470	NM_001142800	EYS	exon25:c.A3753C	Q1251H	Het	C
338	rs200635365	NM_022124	CDH23	exon34:c.A4405G	I1469V	Het	C
	rs61749435	NM_000350	ABCA4	exon16:c.T2546C	V849A	Het	NC
340		NM_000322	PRPH2	exon1:c.G137C	R46P	Het	C
	rs374483122	NM_014249	NR2E3	exon8:c.G1001A	R334Q	Het	C
344		NM_206933	USH2A	exon13:c.T2672C	I891T	Het	NC
	rs41311343	NM_032119	ADGRV1	exon51:c.T10577C	M3526T	Het	NC
353		NM_001142800	EYS	exon22:c.G3419A	G1140E	Het	NC
	rs41311343	NM_032119	ADGRV1	exon51:c.T10577C	M3526T	Het	NC
368	rs141134147	NM_206933	USH2A	exon41:c.A7951G	N2651D	Het	C
369		NM_000350	ABCA4	exon25:c.A3755T	E1252V	Het	C
374		NM_032119	ADGRV1	exon74:c.A15505T	I5169F	Het	C
382		NM_206933	USH2A	exon11:c.G1909A	V637I	Het	C
387		NM_000260	MYO7A	exon9:c.A913G	M305V	Het	C

389		NM_001142800	EYS	exon26:c.C4108T	H1370Y	Het	C
390		NM_001142800	EYS	exon15:c.A2309C	Q770P	Het	C
397	rs56136489	NM_206933	USH2A	exon67:c.C14753T	T4918M	Het	NC
	rs41308846	NM_032119	ADGRV1	exon28:c.G6133A	G2045R	Het	C
		NM_022124	CDH23	exon53:c.G7807A	D2603N	Het	C
398	rs140872693	NM_001145847	PROM1	exon5:c.C577G	R193G	Het	C
400		NM_032119	ADGRV1	exon65:c.A13145G	N4382S	Het	NC
434	rs41311343	NM_032119	ADGRV1	exon51:c.T10577C	M3526T	Het	NC
		NM_032119	ADGRV1	exon74:c.G15299A	R5100K	Het	NC
436		NM_024685	BBS10	exon1:c.A124G	T42A	Het	C
		NM_001142800	EYS	exon26:c.A4843G	T1615A	Het	NC
466		NM_006269	RP1	exon4:c.T1510C	S504P	Het	C
471	rs201008674	NM_006269	RP1	exon4:c.G1726A	V576M	Het	C
473		NM_032119	ADGRV1	exon9:c.G1581T	K527N	Het	C
476	rs188498736	NM_022124	CDH23	exon22:c.C2568G	I856M	Het	C
477		NM_032119	ADGRV1	exon39:c.C8760G	F2920L	Het	C
	rs192843629	NM_001297	CNGB1	exon22:c.C2209T	R737C	Het	C
483		NM_014249	NR2E3	exon2:c.A170G	K57R	Het	C
	rs201489714	NM_000260	MYO7A	exon23:c.G2798A	R933H	Het	C
484		NM_001142801	EYS	exon12:c.T1772G	I591S	Het	NC
487	rs61752901	NM_000329	RPE65	exon9:c.A881C	K294T	Het	C
		NM_000260	MYO7A	exon8:c.A800G	K267R	Het	C
491		NM_201253	CRB1	exon7:c.A2423G	Y808C	Het	C
494	rs147010346	NM_000440	PDE6A	exon1:c.G367T	D123Y	Het	C
520	rs200848641	NM_000260	MYO7A	exon36:c.C4916T	T1639M	Het	C
533		NM_022124	CDH23	exon24:c.A2926G	S976G	Het	C
	rs112002818	NM_001297	CNGB1	exon28:c.C2882T	A961V	Het	C
572		NM_000260	MYO7A	exon9:c.A913G	M305V	Het	C
576	rs112002818	NM_001297	CNGB1	exon28:c.C2882T	A961V	Het	C
578	rs199508694	NM_022124	CDH23	exon15:c.G1637A	R546Q	Het	C
579		NM_206933	USH2A	exon41:c.G7955C	G2652A	Het	C

	rs112002818	NM_001297	CNGB1	exon28:c.C2882T	A961V	Het	C
580	rs200649500	NM_022124	CDH23	exon20:c.C2239T	R747C	Het	C
582	rs111033529	NM_206933	USH2A	exon42:c.C8431A	P2811T	Hom	C

Table 8. Variants with uncertain significance (VUS) identified in 65 patients. For each patient are indicated the gene, nucleotide and protein change, status and conservation of the involved residues. Abbreviations: C: conserved; NC: not conserved.

A total of 169 variations have been identified and confirmed in this study (Figure 4). Among them, missense mutations account for 80%, followed by nonsense mutations (10%), frameshift (8%), in-frame deletions (2%) and splicing alteration (1%) (Figure 10A). Interestingly, about 49% of the variant alleles have not been previously reported (82/169).

The mutations are distributed across 28 retinal disease genes with USH2A and ABCA4 as the most frequently mutated genes, accounting for 26 and 24 variants (15% and 14%). In addition, mutations have been found in ADGRV1 (13), MYO7A (11), CDH23 (10), EYS (10), CRB1 (9), RP1 (9), CNGB1 (6), CNGA1 (7), BBS10 (4), NR2E3 (4), PDE6B (4), RHO (4), PDE6A (3), PROM1 (3), PRPF31 (3), PRPH2 (3), RPE65 (3), BBS1 (2), IMPDH1 (2), TULP1 (2), USH1C (2), and the rest of variants in MERTK, PRPF3, PRPF8, RDH12 and ROM1 genes with one case for each one (Figure 10C). No mutations were found in 4 genes: BEST1, CRX, LRAT and RLBP1.

Further analyzing the identified variants, we realized from previous literature and prediction software the pathogenicity of some of those variants, thus allowing us to divide the screened patients in three categories (Figure 10B): (i) solved cases (patients for which a certain disease causing mutation has been identified and confirmed), (ii) partially solved cases (those cases in which only one pathogenic autosomal recessive alteration has been identified those individuals in which no information is present in the literature about the identified variants) and (iii) unsolved cases (those cases in which no candidate variants were found).

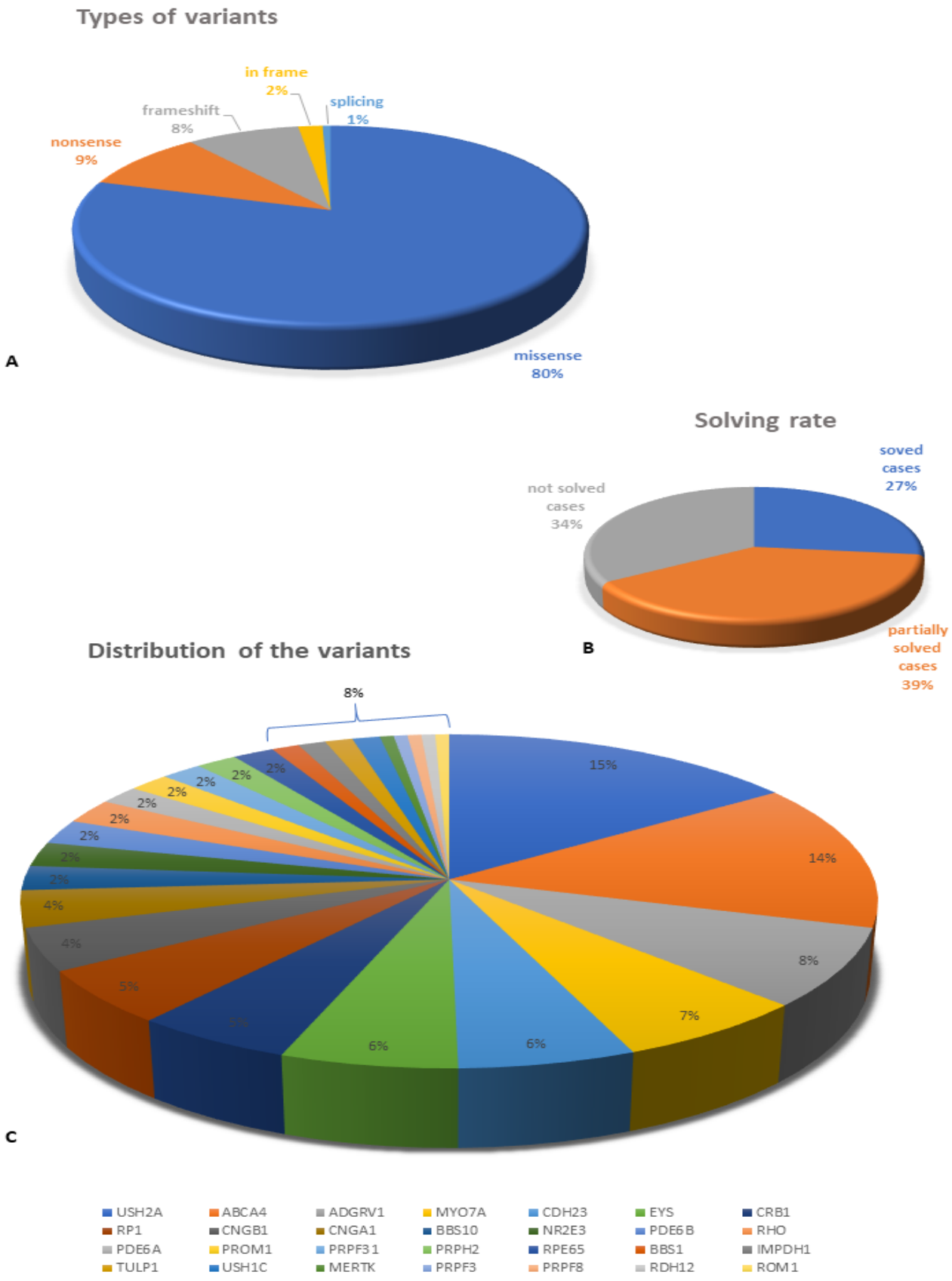

Figure 10. (A) All the identified mutations along with their several types. (B) Representation of the patients divided in three categories on the base of obtained results. (C) 28 retinal disease genes were assigned 169 putative causative mutations.

Validation of the NGS panel

To verify the reliability of our NGS approach, we used, as controls, a validation cohort of 3 out of the total 190 patients analyzed, who had two different mutations previously identified by other certified laboratories. Our NGS strategy allowed us to redetect all the variants found and previously checked by Sanger sequencing giving us a sensitivity of 100% for those validation variants.

Personal Identification

A genetic profile comparison between DNA samples residual from NGS analysis performed by IGA Technologies srl and the corresponding DNA from the biobank has been held to avoid mistakes in sample identification. The locus Amelogenin (AME), used for sex determination, and 15 autosomal STRs (D8S1179, D21S11, D7S820, CSF1P0, D3S1358, TH01, D13S317, D16S539, D2S1338, D19S433, vWA, TPOX, D18S51, D5S818, FGA) were analyzed by means capillary electrophoresis, after a multiplex amplification.

After the setting of the validated threshold, to avoid three most common artifacts ("stutters", "pull ups" and "allelic imbalance"), the comparison of the genetic profiles (see an example in Figure 11) allowed to state that the NGS analysis has been held correctly, without identity exchanges, since all the comparisons between the corresponding electropherograms matched positively.

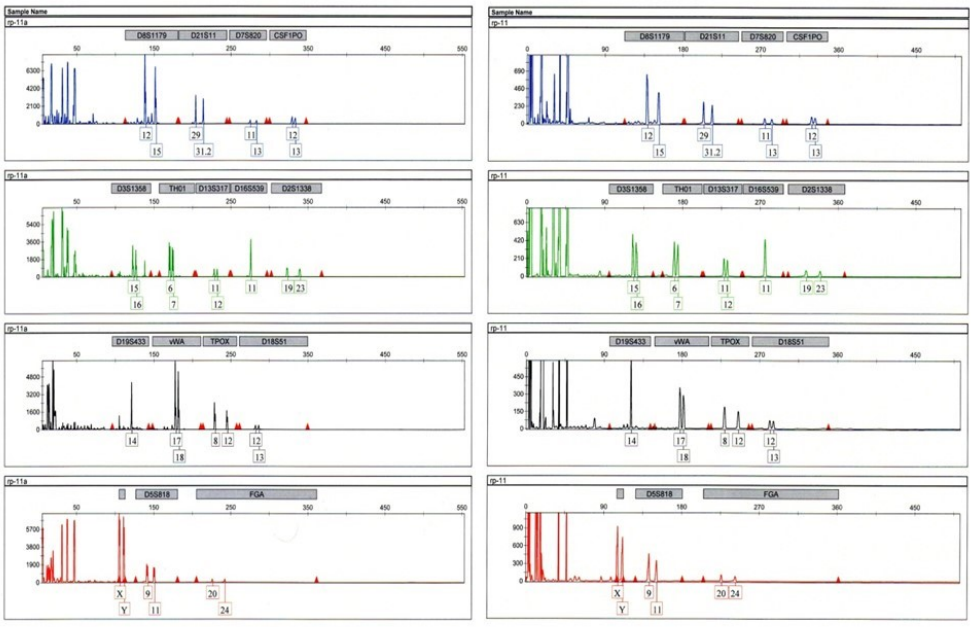

Figure 11. Comparison between the electropherograms of the DNA samples from the biobank (rp-11) and the corresponding remaining specimens after NGS analysis (rp-11a).

Genotype-phenotype correlations

By linking the genetic finding with the phenotypical characteristics, a genotype-phenotype correlation was performed. In this way for 57 patients the genotype was related to phenotype: 6 patients are affected by Stargardt disease, 7 affected by inverse RP, 24 by RP plus hearing loss and 20 by RP (Figure 12). The cases of doubtful correlation are due to the identification of genetic variants in genes so far not related to the specific phenotype of the patient or of variants of uncertain significance. The cases of mismatch in the correlation between the genetic finding and the symptoms of the disease are collected in the unsolved group.

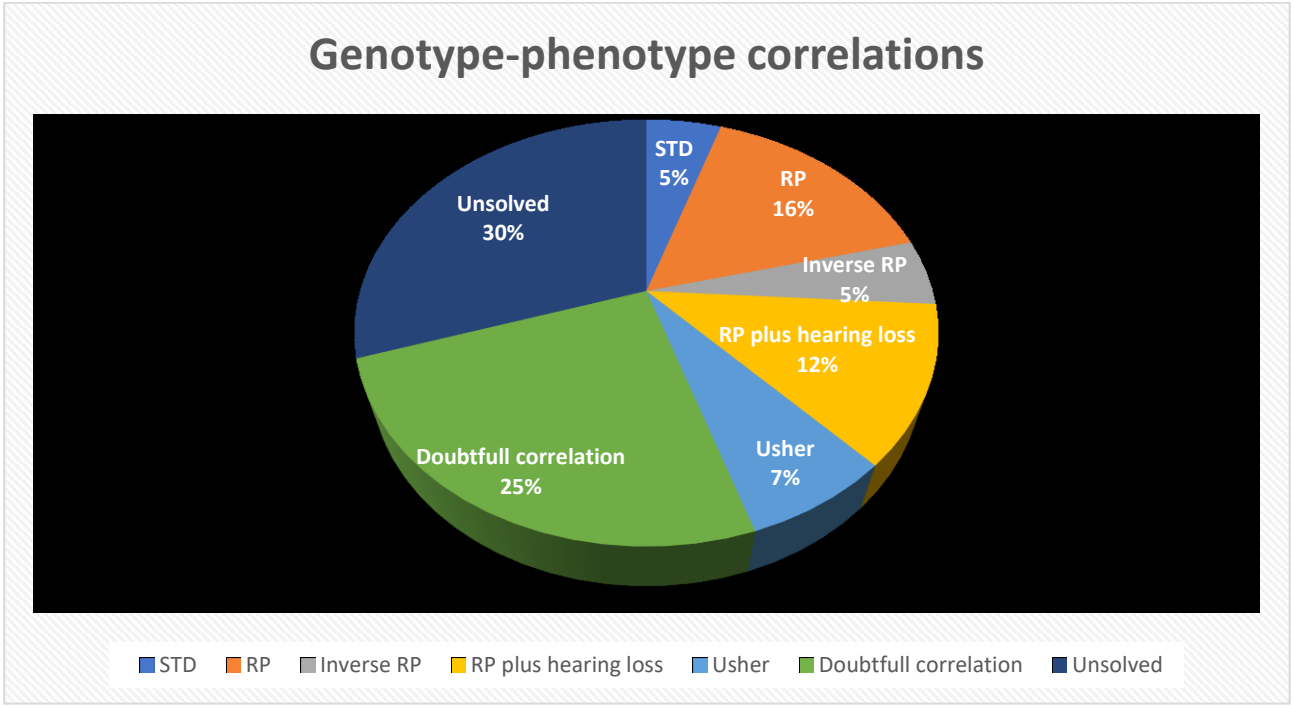

Figure 12. Graphic elaboration of the correlations' results in our cohort: of the 126 genetically characterized, only 55 obtained a successful genotype-phenotype correlation, with a rate of 44%.

STD: 4/6 STD patients show compound heterozygous mutations, while in one patient the mutation C230S is in homozygous state. In one case the genetic screening reveals only one mutation in heterozygosis, leading a partial genotyping. Principle signs of the disease, including reduction of central vision, dysmomatopsia, central scotomas and photophobia, were detected in all the patients.

Inverse RP: Inverse Retinitis Pigmentosa is characterized by areas of choroidal degeneration with pigment migration and bony spicule formation in the macular area. At the contrary of more typical forms of RP, it destroys central vision, leaving peripheral vision intact.

In two patients (187 and 383) were identified homozygous mutation in ABCA4 (L1201R and c.6282+1G>C, respectively). The I196N heterozygous alteration in PRPH2 appears to be the genetic cause in a patient, while an alteration of uncertain significance (N282S) in RPE65 gene could induce

the disease in patient 313. Other patients affected by inverse RP show a complex genetic profile: patient 451 reveals a multigenic pattern, including the frameshift deletion (G2074fs) in ABCA4 and an in-frame deletion (p.165_167del) in CRB1. As reported in scientific literature, the G1961E mutation in ABCA4 seems to aggravate the effects of the compound heterozygous mutations in ADGRV1 (M3526T and R5100K) in the patient 434. In another subject, patient 379, an in-frame deletion (p.71_79del) in TULP1 gene is combined to a A2216V missense mutation in ABCA4 gene that induces the deletion A2216_Q2243 for an alternative splicing. Doubtful are instead the genetic finding in patient 381 (the V378A alteration in CNGA1 and the L819Q one in ABCA4) that doesn't correlate with the phenotypic characteristics, and in patients 46, 60, 310, 390, 459 and 494 in which genetic screening reveals a partial characterization, not completely matching with the phenotypic tracts. No alteration has been found in two inverse RP patients (18 and 373).

RP plus hearing loss: The association of visual impairment and deafness results from mutations in gene express in the hearing and visual organs and tissues. 25 patients of these genetic analysis reveal phenotypic signs of RP associated to hearing impairment. In some patients, mutations in USH2A, PDE6B, CDH23, MYO7A and ADGRV1 induce Usher syndrome. The frameshift insertion H308fs in USH2A gene is present in compound heterozygous state with the R626X mutation in patients 176, affected by Usher Syndrome, arisen in the first symptoms of hemeralopia at the age of 12 years. Compound heterozygous mutation in USH2A, causing Usher Syndrome, have been identified in patients 296 (nonsense L3567X and missense C3666G mutations), 316 (F266L and D591G missense mutations) and 338 (R63X nonsense mutation and A3973fs frameshift deletion). Patients 174 and 578 are affected by Usher syndrome with early onset (in childhood) caused by homozygous frameshift deletion in the exon 11 of CNGA1, that for patient 578 is aggravated by the frameshift deletion V434fs in ABCA4; the genetic cause in patient 129 is a homozygous alteration in CDH23 and in patient 191 are compound heterozygous mutations in ADGRV1 (G1478D and V4552fs). Classic RP plus hypoacusis, with onset in the second and third decade of life, was diagnosed in patients 473 and 580 and genetically confirmed by compound heterozygous mutations in ADGRV1 (K527N and R2260X) and in CDH23 (R747C and H755Y) respectively, while early onset mutations (E105K and F247I) were confirmed in patients 476 and 364 in PDE6B gene. Two patients with different deafness affections are associated for the presence of the G1961E mutation in ABCA4 gene and for alterations in USH2A (C759F mutation in patient 455 and A2713S and V1702M variants in patient 452). Compound heterozygous mutations (M305V and Q2017X) in MYO7A are related to classic retinitis pigmentosa associated to hearing alteration in patient 387. Dominant mutations in NR2E3 gene cause RP plus hearing impairment with late onset in patient 483. This gene is also involved in triallelic alterations in patients 400 and 180, with mutation in USH2A gene (W3955X

homozygous nonsense mutation and compound heterozygous R626X and V1833E ones, respectively). Auditive deficit, addicted to RP, is caused by the W1837X homozygous nonsense mutation in EYS gene in patient 286 and by compound heterozygous alterations (C948Y and C85F) in CRB1 in patient 183.

For three patients (210, 212 and 317), the genetic screening revealed mutations in a gene not yet related to hearing involvement: PRPF31. This gene was mutated with loss of function due to frameshift insertion and deletion and nonsense mutation.

Partial are instead the genetic finding in 7 patients in which only one mutation has been identified in USH2A or in MYO7A genes.

RP: Of 19 RP clinically and genetically solved patients, 8 are affected by autosomal dominant forms. Missense alterations in RHO, PRPF3, PRPF8 and RDH12 were confirmed in patients 171, 301, 385, 487 and 347 respectively, while in two patients, heterozygous mutations in RP1 seem to induce the RP phenotypic effects (T373I in patient 318 and R677X in patient 472). The autosomal recessive nature of the gene RP1 is visible in patients 45 and 460, affected by frameshift insertion in homozygosity (pat. 45) or by frameshift deletion in digenic form with nonsense mutation in USH2A (L3567X in pat. 460). Homozygous missense mutations affect 4 patients, involving the following autosomal recessive genes: PDE6B (D600N in pat. 230), EYS (Y3059X in pat. 392), RPE65 (H527Y in pat. 520) and CNGA1 (R32X in pat. 576). Compound heterozygous alterations have been identified in CNGA1 gene in two sisters (pat. 264 and 265), in which the W180X nonsense mutation and the R280C missense alteration are aggravated by the frameshift insertion in BBS1 gene (C91fs). Even if they correspond to the symptoms of the patients 178 (aggravated by G1961E mutation in ABCA4), 192 and 321, the pathogenicity of the two variants, A961V and N953H in CNGB1 gene, is doubtful, given their frequency in the population. In proband 491, the already known T745M missense mutation in CRB1, in double heterozygosis with the Y808C missense alteration, induces the classic RP symptoms, while in proband 58, the E254X nonsense mutation in PDE6B is combined with the most diffuse missense mutation in ABCA4, the G1961E. Partial or doubtful genetic characterization was obtained in 19 patients, instead no phenotype-genotype correlation was feasible for 13 affected subjects.

Risk factor determination

A total of 118 patients, 12 of them as quality controls and 58 screened by NGS, were analyzed to genotype rs10490924, rs1061170, rs2230199, rs403846, rs641153, rs9332739 and rs1410996 polymorphisms, known to be genetic risk factors for the AMD. As showed in Table 9, in the cohort

of analyzed patients (58 of them previously analyzed in the NGS screenings), the results revealed a high genetic predisposition to macular degeneration in 38 subjects, middle high in 16, medium in 36, while to low risk are predisposed 28 patients (17 of them are associated to mild risk). In 11 of the 58 IRD patients screened for AMD SNPs are present mutations in ABCA4 gene.

Patient ID	*rs 10490924*	*rs 1061170*	*rs 2230199*	*rs 403846*	*rs 641153*	*rs 9332739*	*rs 1410996*	**Genetic Risk**	**ABCA4-related Risk**
2	G/G	C/T	C/C	G/A	G/G	G/G	T/C	medium	NA
3	G/G	C/T	C/C	G/A	G/G	G/G	T/C	medium	NA
26	G/G	C/T	C/C	G/A	G/G	G/G	T/C	medium	/
31	G/T	T/T	C/C	G/G	G/G	G/G	T/C	medium	NA
34	G/T	C/T	C/C	G/A	G/G	G/G	C/C	high	NA
36	G/T	C/T	C/C	G/A	G/G	C/C	C/C	medium	NA
84	G/G	T/T	C/C	G/G	G/G	G/G	T/T	mild	NA
85	G/T	T/T	C/C	G/G	G/G	G/G	T/C	medium	NA
94	G/T	T/T	C/C	G/G	G/A	G/G	T/C	mild	NA
108	G/G	T/T	C/G	G/G	G/G	C/C	T/T	low	NA
131	G/T	C/T	C/G	A/A	G/A	G/G	C/C	high	NA
133	G/T	C/C	C/C	G/G	G/G	G/G	T/C	high	NA
161	G/T	T/T	C/C	G/G	G/G	G/G	T/T	medium	NA
168	G/G	T/T	C/C	G/G	A/A	G/G	T/T	low	NA
170	G/G	T/T	C/G	G/G	G/G	G/G	T/C	mild	/
171	G/T	T/C	C/C	A/A	G/G	G/G	C/C	high	/
173	G/G	C/T	C/C	G/A	G/A	G/G	T/C	medium	/
175	G/G	C/T	C/C	G/A	G/G	G/G	T/C	medium	NA
176	G/G	T/C	C/C	G/A	G/G	G/G	C/T	medium	/
178	G/G	C/T	C/G	A/A	A/A	G/G	C/C	medium	exon42:c.G5882A; p.G1961E (Het)
186	G/G	C/C	C/G	A/A	G/G	G/G	C/C	high	/
189	T/T	C/T	C/C	G/A	G/G	G/G	T/C	high	exon17:c.G2588C; p.G863V (Het)
193	G/G	C/C	C/C	A/A	G/G	G/G	C/C	high	/
194	G/G	T/T	C/C	G/G	G/G	C/C	T/T	low	/
207	G/T	T/T	C/G	G/G	G/G	G/G	T/C	middle-high	/
210	T/T	C/C	C/G	A/A	G/G	G/G	C/C	high	/
211	G/T	C/T	C/C	G/A	G/G	G/G	T/C	high	/
212	T/T	T/T	C/G	G/A	G/G	G/G	C/T	high	/
213	G/G	T/T	C/C	G/G	G/A	G/G	T/T	low	/
219	T/T	C/C	C/C	G/A	G/G	G/G	C/C	high	NA
220	G/G	T/T	C/C	G/G	G/G	G/G	T/C	mild	NA

232	T/T	C/C	C/G	A/A	G/G	G/G	C/C	high	NA
238	G/G	C/T	C/C	G/A	G/G	G/G	T/C	medium	/
246	G/G	T/T	C/C	G/G	G/G	G/G	T/C	mild	/
256	G/G	C/T	C/G	G/A	G/G	G/G	C/C	middle-high	NA
257	G/G	C/T	C/G	G/A	G/A	G/G	T/C	medium	NA
264	G/G	C/T	C/G	A/A	G/G	G/G	C/C	high	/
265	G/G	C/T	C/C	A/A	G/G	G/G	C/C	middle-high	/
269	G/G	C/C	C/G	G/A	G/G	C/C	T/C	medium	NA
287	G/G	T/T	C/C	G/G	G/G	G/G	C/C	mild	/
293	G/G	C/T	G/G	G/A	G/G	G/G	T/C	middle-high	/
294	G/G	C/T	G/G	G/A	G/G	G/G	T/C	middle-high	NA
296	G/G	T/T	G/G	G/G	G/G	G/G	C/C	medium	/
297	G/G	T/T	G/G	G/G	A/A	G/G	T/C	low	NA
308	G/G	T/T	C/G	G/A	G/G	G/G	T/C	medium	NA
310	G/T	T/T	C/C	G/G	G/G	G/G	T/C	medium	/
313	G/G	T/T	C/C	G/G	G/G	G/G	T/T	mild	/
316	G/G	T/C	C/C	G/A	G/G	G/G	C/T	medium	exon13:c.C1912T; p.P638S (Het)
317	G/G	T/C	C/C	G/A	G/G	G/G	C/T	medium	/
321	G/G	T/T	C/C	G/G	A/A	G/G	T/T	low	/
322	G/T	T/T	C/C	G/G	G/G	G/G	T/C	medium	NA
323	G/G	C/T	C/C	G/A	G/G	G/G	T/C	medium	/
324	G/T	C/T	C/C	G/A	G/G	C/C	T/C	medium	/
327	G/T	C/C	C/C	G/A	G/G	G/G	T/C	high	NA
332	G/G	C/C	C/G	A/A	G/A	G/G	C/C	high	NA
337	G/G	C/C	C/C	A/A	G/G	G/G	C/C	high	NA
338	G/G	C/T	C/C	G/A	A/A	G/G	T/C	mild	exon16:c.T2546C; p.V849A (Het)
342	G/T	T/T	C/C	G/G	G/G	G/G	T/T	medium	NA
343	G/T	T/T	C/G	G/G	G/G	G/G	C/C	middle-high	NA
346	G/T	C/T	C/C	G/A	G/G	G/G	T/C	high	NA
357	G/G	C/C	C/G	A/A	G/A	G/G	C/C	high	exon42:c.G5882A; p.G1961E (Het); exon46:c.6282+1G>C (Het)
360	G/T	C/T	C/G	G/A	G/G	C/C	T/C	medium	NA
362	T/T	C/C	C/C	G/A	G/G	G/G	T/C	high	NA
371	G/T	C/T	C/C	G/A	G/G	G/G	T/C	high	/
372	G/G	T/T	G/G	G/G	A/A	G/G	C/C	low	/
384	G/G	C/T	C/G	G/A	G/G	G/G	T/C	middle-high	/
387	G/G	T/C	G/C	A/A	G/G	G/G	C/C	high	/
389	G/T	T/T	C/C	G/G	G/G	G/G	T/T	medium	/

391	G/T	C/T	C/C	G/A	G/G	C/G	T/C	middle-high	exon9:c.T1140A; p.N380K (Het)
393	G/G	C/C	C/G	A/A	G/G	G/G	C/C	high	/
400	G/G	C/C	C/C	G/A	G/G	G/G	C/T	middle-high	/
405	G/T	C/C	C/C	A/A	G/G	G/G	C/C	high	/
423	T/T	C/T	C/C	G/A	G/G	G/G	T/C	high	NA
433	G/G	T/T	C/C	G/A	G/A	G/G	C/C	mild	/
445	G/T	C/T	C/G	G/A	G/A	G/G	T/C	high	NA
452	G/T	T/T	C/C	G/A	G/G	G/G	C/T	middle-high	exon42:c.G5882A; p.G1961E (Het)
453	G/G	C/T	C/G	G/A	G/G	G/G	C/C	middle-high	NA
454	G/G	C/T	C/C	G/A	G/G	G/G	T/C	medium	NA
455	G/G	C/C	C/C	G/A	G/G	C/G	C/T	medium	exon42:c.G5882A; p.G1961E (Het)
459	G/G	T/T	C/G	G/G	G/G	G/G	T/C	mild	exon42:c.G5882A; p.G1961E (Het)
460	G/G	T/C	G/C	G/A	G/G	G/C	C/T	medium	/
461	G/G	C/C	C/C	A/A	G/G	G/G	C/C	high	NA
462	G/G	C/T	C/C	G/A	G/A	G/G	T/C	medium	NA
463	T/T	T/T	C/C	G/G	G/A	G/G	T/C	middle-high	NA
468	G/T	C/T	C/G	G/A	G/A	G/G	T/C	high	/
472	G/G	T/T	C/C	G/G	G/G	G/G	T/T	mild	
473	G/T	T/T	C/C	G/G	G/G	G/G	T/T	medium	
476	G/G	T/T	C/C	G/G	G/G	G/G	T/T	mild	exon18:c.A2701G; p.T901A (Het)
483	G/G	C/C	C/C	G/A	G/G	G/G	C/T	middle-high	/
486	G/T	T/T	C/C	G/G	G/G	C/G	T/T	mild	/
490	G/G	C/T	C/G	G/A	G/G	G/G	C/C	middle-high	NA
504	G/T	T/T	C/C	G/G	G/G	G/G	T/T	medium	/
511	G/G	T/T	C/C	G/G	G/A	G/G	T/T	low	NA
517	G/G	C/T	C/C	G/A	G/G	G/G	T/C	medium	NA
520	G/G	T/C	G/C	A/A	G/G	G/G	C/T	high	/
532	G/G	T/T	C/G	G/G	G/A	G/G	T/T	low	NA
534	G/T	C/C	C/C	G/A	G/G	G/G	T/C	high	NA
537	G/G	C/C	C/C	A/A	G/A	G/G	C/C	middle-high	NA
545	G/G	T/T	C/G	G/G	A/A	G/G	T/T	low	NA
546	G/G	C/T	C/C	G/A	G/G	G/G	C/C	medium	NA
560	G/T	C/C	C/G	A/A	G/G	G/G	C/C	high	NA
576	G/T	T/C	C/C	G/A	G/G	G/G	C/C	high	
578	G/G	T/T	C/G	G/G	G/G	G/G	C/T	mild	exon10:c.1302delA; p.V434fs (Het)
580	T/T	C/C	C/C	G/A	G/G	G/G	C/T	high	/
581	G/G	T/T	C/C	G/G	G/A	G/G	T/T	low	/
582	G/G	C/C	G/C	A/A	G/G	G/G	C/C	high	/

583	G/T	C/T	C/C	G/A	G/G	G/G	T/C	high	/
587	G/G	T/T	C/C	G/G	G/G	G/G	T/T	mild	NA
604	G/T	C/T	C/C	G/A	G/G	G/G	T/C	high	NA
605	G/T	C/T	C/G	G/A	G/A	G/G	C/C	high	NA
606	G/G	C/T	C/C	G/G	G/A	G/G	T/C	mild	NA
607	G/G	C/T	C/C	G/A	G/G	G/G	C/C	medium	NA
608	G/G	T/T	C/C	G/G	G/G	G/G	T/C	mild	NA
609	G/T	T/T	C/C	G/G	G/G	G/G	T/T	medium	NA
610	G/G	C/C	C/C	A/A	G/A	G/G	C/C	middle-high	NA
611	G/T	C/T	C/G	G/A	G/A	G/G	C/C	high	NA
612	G/G	C/T	C/C	G/A	G/G	G/G	T/C	medium	NA
613	G/T	C/T	C/G	G/A	G/G	G/G	T/C	high	NA

Table 9. SNP genotyping results of 118 patients. The SNP in ARMS2 (rs10490924), in CFH (rs1061170, rs1410996 and rs403846) and in C3 (rs2230199) are related to high risk to develop AMD, while SNP in CFB (rs641153) and in C2 (rs9332739) have a protective effect. Abbreviations: NA= not available; het= heterozygous; hom= homozygous; /= no mutation identified by NGS.

X-linked genetic characterization

The analysis of a cohort of 60 patients with X-linked pattern of inheritance were performed by means a complete Sanger sequencing of the RPGR and RP2 genes. The Sanger sequencing of RP2 gene showed in an affected subject a missense mutation in the exon 2 (c.413A>G; p.Glu138Gly), previously described by Miano (Miano et al. 2001) (Figure 13). The mutation causes the substitution of a glutamic acid residue at the position 138 with a glycine one.

The analysis of RPGR gene bared three novel variants, two mutation and a polymorphism in two different families. In one family, in exon 14, the nucleotide change of a guanine with a thymine generates a premature stop codon in position 531 (c.1591G>T; p.Glu531*). In the segregation analysis, the nonsense mutation, that lead to a truncated protein lacking the C-terminal, was detected as hemizygous mutation in two affected brothers of the family and in the mother as heterozygous mutation by originating a mild form of RP.

In the second family, were found a missense substitution (c.1105C>T; p.Arg369Cys) and a novel causative frameshift mutation in exon ORF15 (c.2091_2092insA [g.ORF15+338_339insA]; p.A697fs [p.ORF15+A112fs]) of RPGR gene. Although in exon 10 of RPGR gene the replacement of Arg369 with a cysteine resulted in the substitution of the positive charged residue with neutral polar one, its detection in a healthy family member strongly suggests that it was not a disease-causing variant, suspect confirmed by conservation analysis (proximity to the highly conserved RCC1-like domain), the high allele frequency and the three-dimensional model of RPGR protein (Parmeggiani et al. 2017; appendix A).

The novel frameshift mutation in ORF15, caused by the insertion of an alanine nucleotide, induces an alteration in the reading frame, thus comporting the loss of the C terminal domain of the protein. Discovered in hemizygous status in the affected subject and in heterozygous status in the mother and other three female, that are obligate carriers, it coherently co-segregated with the phenotypic condition, confirming the X-linked inheritance pattern (Figure 13) (Parmeggiani et al. 2016).

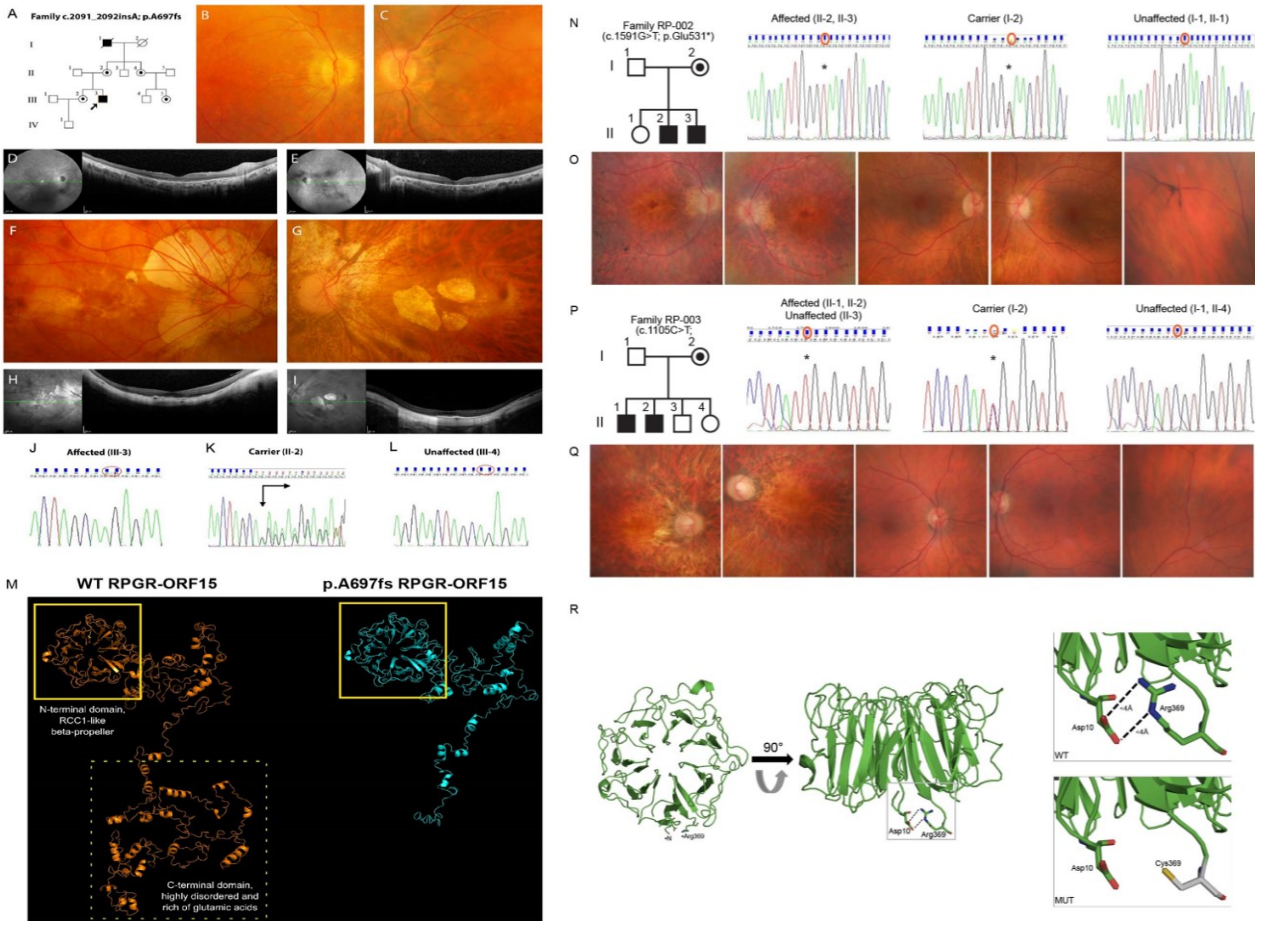

Figure 13. Clinical confirmations of the phenotypical tracts characterizing RP probands (from C to I and Q). familiar pedigree and Sanger identifications of Ala697fs (A, J, K, L), Glu531* (N) and R369C (P) in RPGR gene. Protein structure study of the missense variation for Ala697fs (M) and Arg369Cys (R) in RPGR gene.

Missense mutation causative of OMD

Molecular analysis, by means Sanger sequencing, in two probands with OMD diagnosis, revealed an already known heterozygous mutation in the RP1L1 gene, previously described by Takahashi (Takahashi et al. 2014). In exon 3, the mutation induces the change c. 3596 C>T that causes the substitution of a serine residue (TCT) with a phenylalanine one (TTT) at the amino acid position 1199 (Figure 14). No other pathological variants were detected in this gene. Thanks to the availability of the paternal blood of one of the patients, molecular analysis was extended showing the paternal segregation of the mutation for one of the two affected proband.

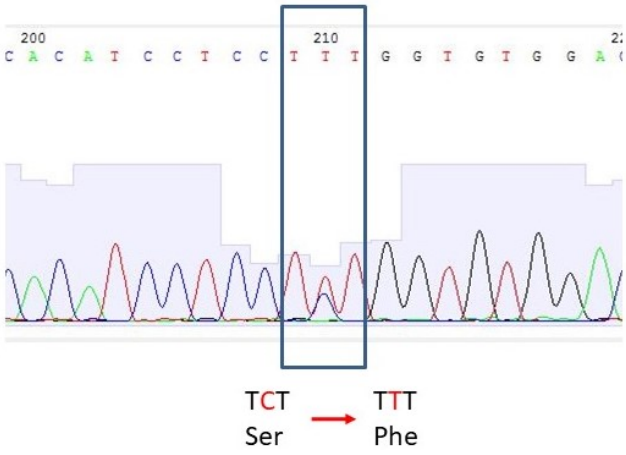

Figure 14. Identification of the heterozygous mutation in RP1L1 gene in two patients. The mutation induces the substitution of Ser1199 residue with a Phe one.

Discussion and Conclusion

Inherited Retinal Disorders (IRD) are a complex group of inherited degenerative and disabling diseases, characterized by a progressive visual impairment. First symptoms appear during the second or third decade of patients' life, starting with difficulties in night vision and dark-light adaptation. The diseases become more severe, until complete vision loss. IRD are progressive diseases and no molecular diagnostic tests for the identification of the cause of the pathology is available nowadays. Occult macular dystrophy is inherited macular dystrophy clinically difficult to diagnose due to normal response of the patients to ophthalmoscopy, fluorescein angiography and full-field electroretinogram (ERG). Respect to IRD, it is caused by alteration in only one gene, RP1L1, with autosomal dominant pattern of inheritance.

The difficulty of the diagnosis for affected people drove to the setting-up of a large-scale project, focused on the genetic diagnosis of these diseases. This project is born with the aim of developing effective diagnostic techniques for IRD and for OMD patients. The study of a complex group of disease like these ones requires the comparison of a vast number of affected subjects. The collections of biological materials, or biobank, together with clinical anamnesis of individuals represent a fundamental instrument to elucidate genetic causes and to traduce the biomedical research in development of treatments.

For this reason, this large-scale work started from the recruitment and clinical characterization of a large cohort of subjects (620) affected by Inherited Retinal Disorders (IRD) or Occult Macular Degeneration (OMD), followed by the elaboration of both genetic and retinal biobanks, comprising all the patient's DNA and donors' retinas. Although there is no legislation regarding biobanks, donors' rights are guaranteed by a reminder of the privacy (DL 196/2003) and confidentiality of the information contained in the collection through the informed consent. However, the WHO guide lines and the UNESCO declaration on bioethics can be applied on biobanks (Caenazzo, 2012).

In the second stage of the project, mutations for 190 of those IRD individuals were investigated by means of Next Generation Sequencing (NGS), using a targeted approach for 32 genes associated with autosomal IRD outcome. IRD genetics is complex, considering the involvement of at least 280 genes and all pattern of inheritance (Dias et al. 2018; Broadgate et al. 2017). No X-linked related genes were analyzed with the targeted sequencing methods. Considering that 80% of X-linked forms are mainly due to mutations in RP2 (10-15%) and RPGR (70-75%) genes, the best approach, for X-linked patients, to identify the causative mutation is the direct Sanger sequencing of these two genes (Parmeggiani et al. 2016), that was applied to a cohort of 60 probands with suspected X-linked

Retinitis Pigmentosa, revealing a novel frameshift insertion in ORF15, a novel nonsense mutation in RPGR gene and an already known missense mutation in RP2 gene (Parmeggiani et al. 2017; Parmeggiani et al. 2016).

If on the one hand, the use of NGS allows to analyze in parallel more samples and more genes, on the other hand it produces a huge quantity of output data which must then be opportunely studied and skimmed. Of great importance, even before the analysis phase, is the design of the gene panel, the selection of the genes of interest and of the experimental parameters. It is therefore not surprising how the optimization of these aspects is reflected in an increase in the yield of the experiment. Considering the number of genes related to IRD at present, the gene panel used in this project appears limited. However, it must be considered that during the design phase of the project and of the genes panel, only 50 genes were associated with the pathologies under study and only later, also thanks to the use of whole exome sequencing approaches, other gene-disease associations are come to light.

The elaboration of a filtering workflow become a critical point, prior to the confirmation of identified variants by means Sanger sequencing. The prioritization process, designed during the project, provided the selection of all the mutations in the coding regions, particularly missense, nonsense, in frame or frameshift variants with an adequate coverage. Variants were selected also considering the Minor Allelic Frequency, checked in 1000 Genomes Project and Exome Aggregation Consortium (ExAC), and they were considered pathogenic if nonsense or frameshift mutations, otherwise predictive software and PubMed references were utilized. It has been difficult to establish an approximate coverage threshold underneath which the identified mutations were considered system errors. After many tests on low coverage mutations with Sanger sequencing, we assessed that the proper coverage threshold is around 15 folds.

Inheritance pattern of the identified mutations has not been completely determined yet, due to the complexity to recontact patients and their family, since it is important for them, especially the members at risk, such as the offspring, and for us to understand the sporadic or inherited nature of the mutations and their pattern of inheritance, in order to be able to enlarge the knowledge of these complex diseases.

A turning point of this work is the use of Illumina Next Generation Sequencing, innovative technology that allows sequencing of considerable amounts of samples per run and is still developing year by year, leading to the progressive increase of data output and decrease of sequencing costs. This technique has proved to be essential for this project, since IRD are heterogeneous diseases caused by many different genes (at least 280) and a lot of different mutations.

The total solving rate reached for the genetic screening in this work is 66%: only 27% of cases obtained a complete genetic characterization while for the remaining cases (39%) the genetic analysis

results partial or not definable, for the identification of novel or known variants with uncertain significance. Indeed, the prediction of the functional effect of novel variants is easy for frameshift mutations, nonsense or splice-site alterations that result in a truncated protein, while it is complex for variants in the exon-intron junction or missense alterations (Jonsson et al. 2018). The latter correspond to the type most commonly found within our cohort, representing 80% of the variants on the total. The distribution of the variants sees the genes USH2A and ABCA4 among the most mutated, with their 15% and 14% respectively. The first one is expressed in many different tissues, especially retina and inner ear, hence USH2A mutations cause Usher syndrome or autosomal recessive Retinitis Pigmentosa. Located in the basement membrane, usherin interacts with other proteins such as EYS and USH1C to fulfill its maintenance role of development and homeostasis of otic and retinal cells. ABCA4 translate a retina-specific ABC transporter that uses N-retinylidene-PE as a substrate. It is expressed only in the outer segment disk of photoreceptors indicating that it mediates transport of an essential molecule of the visual cycle across the photoreceptor cell membrane, fundamental for retinal metabolism (Cella et al. 2009). ABCA4 mutations are responsible of Stargardt disease, but also Retinitis Pigmentosa, Cone-Rod Dystrophy, Fundus Flavimaculatus, and Age-Related Macular Degeneration.

Considering that we studied 32 of at least 280 genes involved IRD, we can be satisfied of the percentage of solved cases obtained. Moreover, the complex heterogeneity of these pathologies often makes the correct correlation between identified genotype and manifest phenotype difficult. Only 57 patients, in fact, were completely solved in these terms. The identification of some variants of uncertain significance (VUS), the analysis of only 32 genes and the absence of pedigree can explain this low rate of resolution.

Once the genetic cause of the disease has been identified, it is important to define whether it correlates with the patient's clinical manifestation. In fact, the genotype-phenotype correlation is not always immediate, due to clinical and allelic heterogeneity and incomplete penetrance. As demonstrated in the literature (Jones et al. 2017), the IRDs overturn the canonical genetic rules, in which the gene implication is attributable to a single candidate gene, as not only digenic but also triallelic forms are possible. In two subjects, affected by different forms of RP (inverse RP and RP with neurosensory hypoacusis), compound heterozygous mutations identified in USH2A gene in one individual and in ADGRV1 gene in the other seem to be aggravated by the G1961E missense mutation in ABCA4 gene, in contrast to what Lee asserted, for which the mutation appears as hypomorphic allele, with an unknown mechanism (Lee et al. 2016). Our consideration is instead in line with the research of Cella and colleagues, according to which the G1961E mutation would aggravate the phenotype (being considered as “mild to moderate” allele), reducing the ATPase activity of the ABCA4 protein (Cella

et al. 2009).

Within our patients' cohort, some geno-phenotypic correlations were direct and simple, as for patients with STD in whom mutations in the ABCA4 gene were identified. Nonetheless, one of the STD affected subjects is partially correlated because the screening has detected only one mutation in ABCA4 gene. Highly probable the other variant could be found in intronic regions not included in the genetic analysis (Jonsson et al. 2018). The correlation was more circuitous for the individuals affected by RP, RP plus hearing loss and Inverse RP: not only for the high presence of compound heterozygous mutations, but also for the phenotypic heterogeneity shown by the same genetic alteration. It is the case of two patients that manifest the I196N genetic missense alteration in PRPH2 gene that seems responsible to Retinitis Pigmentosa in subject 405 and Inverse RP in subject 173. This variant, however, is among the VUS, although it has been found to be pathogenic by the prediction analysis and yet to correspond with the manifest phenotypes. The analysis of 5 (MYO7A, CDH23, USH1C, USH2A and ADGVR1) of the 11 genes associated with Usher Syndrome allowed to solve 6 of the 24 subjects affected by RP and hearing loss. In addition to these, three probands with Usher Syndrome have frameshift (deletion and duplication) and nonsense mutations in the PRPF31 gene. These findings are of considerable relevance since up to now the PRPF31 gene has only been associated with Retinitis Pigmentosa, with an haploinsufficient pathogenic mechanism (Ellingford et al. 2018). Thus, we could be looking at a new gene-disease association that will have considerable relevance for the studies of these pathologies. Therefore, although the whole genome or whole exome sequencing would guarantee the more complete genetic diagnosis (in particular the whole genome approach) and also the possible discovery of new disease-genes, the targeted NGS allows an optimal genetic characterization, if the panel design is consistent with the clinical manifestations of enrolled individuals, but also the identification of new gene-disease associations, reducing the criticality related to the high quantity of output data and to the ethical aspects (linked to patients' rights and researchers' duties on the dissemination of obtained information not related to the object of the study). The heterogeneity of IRD induced us to use NGS technologies to reduce time of analysis and increase efficiency and number of analyzed patients. While the clinical diagnosis of OMD let to indagate the principle actor in the pathology, represented by RP1L1 gene. Molecular testing on OMD patients, allowed to the light a missense mutation in two patients, S1199F, previously published by Takahashi (Takahashi et al. 2014). The reduced number of patients and the small size of the gene make Sanger sequencing a good compromise between time and costs.

The identification of the causative mutation will allow the outgrowth of regenerative therapy approaches, focused on the correction of the genetic defect, through gene augmentation (by means viral vectors) or gene editing (by means CRISPR/Cas9 system), and the reconstruction of the retinal

tissue, starting from RSCs or iPSCs. The virus-based or CRISPR/Cas9-based gene therapy approach can be applied for patients in early stage of the disease, with residual retinal stem cells, or on induced pluripotent stem cells (iPSCs), for patients in late-stages of IRD that have already lost all their RSCs, as recently demonstrated on animal models by Latella (Latella et al. 2016) and Burnight (Burnight et al. 2018). In the near future, these therapeutic methods will be the subject of new projects of our research group, which already possesses knowledge in the field of cell reprogramming and gene therapy, through their application on a rare genetic pathology (see appendix B, C, D and E). This will be important to pave the way for all those IRD patients with partial or complete loss of retinal tissue.

Genetic predisposition to Macular Degeneration

The Age-related Macular Degeneration is the main cause of legal blindness in the industrialized countries. This complex and multifactorial disorder is the result of the combination of genetic and environmental factors. Genetic susceptibility interacts with aging and improper dietary and smoking habits increasing the AMD occurrence. Striking links between AMD and polymorphisms in complement genes have been demonstrated underling the implication of inflammatory processes in the disease pathogenesis. Indeed, in AMD patients are elevated the levels of C3a, C5a and C5b-9 products. The principle genes associated to AMD predisposition are CFH (rs1061170, rs1410996 and rs403846), ARMS2 (rs10490924), C3 (rs2230199), CFB (rs641153) and C2 (rs9332739) genes. Being associated with a degenerative retinal degeneration, these polymorphisms could show implications also in IRD. The elaboration of a genetic test of susceptibility applicable to IRD affected people, aimed to delineate a relevant correlation with these SNP and retinal dystrophies. A high or middle-high genetic predisposition has been highlighted 54 patients (38 high and 16 middle-high), medium risk in 36 subjects and low genetic predisposition in 28 individuals (17 of them are associated to mild risk). Of the 118 individuals screened, 58 were previously analyzed by NGS to identify the genetic mutation inducing the IRD phenotype. Nevertheless, the genetic predisposition to AMD, corresponding to high and middle-high risk, seems to not correlate with the severity of the confirmed mutations in the IRD patients. One of the IRD-genes screened by the targeted NGS approach, ABCA4, seems able to influence the macular phenotype. For this reason, an ABCA4-genetic risk was considered evaluating a correlation between the mutations in ABCA4 and the genetic predisposition to AMD. Although the number of samples is too low to make statistically significant considerations, some small but intriguing observations can be raised. At first glance, no correspondence would seem to exist between the severity of the ABCA4 mutations and the genetic risk of developing AMD. Observing more carefully, however, one notices how the G1961E mutation, whose pathogenicity turns out to be contradictory in many studies, is associated with a variable susceptibility risk. In one

of the patients, indeed, the mild risk of AMD is related to a frameshift deletion in ABCA4. This dissonant association could help in the interpretation of the result of genetic analysis going to define a lower degree of severity than the other frameshift deletion detected in homozygosity on another gene. In order to avoid forcing any interpretation of these preliminary data, it is necessary to increase the number of examined patients.

Further elaborations of the results combined with the clinical data could furthermore delineate the right association of these risk factors with the disease that affects patients.

Conclusions

Retinal dystrophies are a real challenge for molecular diagnosis due to the high heterogeneity that makes difficult a clinical and molecular diagnosis, for the overlap of symptoms, the unclear inheritance patterns, variable onset and progression and for the huge amount of gene and allelic loci. This project demonstrates how Next Generation Sequencing approaches are fundamental for the determination of the genetic cause in individuals affected by autosomal recessive or autosomal dominant forms of IRD. This approach, although based on a reduced panel of genes, allowed the genetic characterization of a consistent part of the enrolled subjects, reducing the difficulties of data analysis and ethical issues that would incur using whole genome or whole exome sequencing approaches. The collaboration between researchers and clinicians has also allowed to correlate the genotype and phenotype in many of the analyzed cases, bringing to the light even a new gene-disease association for PRPF31, before now unrelated to Usher Syndrome. The X-linked forms of RP, mainly caused by alterations in the RPGR and RP2 genes, and the OMD, induced by mutations in RP1L1, can be genetically characterized by Sanger sequencing, in a simple and reliable way. The development of susceptibility tests, such as that focused on 7 polymorphisms in complement-pathway genes associated with AMD, could demonstrate useful implications in the overall assessment of individuals with IRD.

Given the high number of patients and the reduced knowledge of the genetics at the base of these pathologies, it would be useful to extend the genetic analyses to the whole cohort of Italian patients from which the biobank was generated, not only to identify the causative mutation of the phenotype which patients are affected, but also to delineate mechanisms, genes-disease associations and to define the relationship between these diseases and the genes involved in the inflammatory processes responsible for the retinal degeneration process underlying these pathologies.

This research has proved useful by defining valid genetic diagnostic tools and laying the basis for the development of gene- or mutation-specific cell and gene therapies approaches fundamental for the treatment of hereditary retinal diseases.

The validation of an effective diagnostic tool, based on NGS approach, is a fundamental start point for the identification of the genetic basis of IRD. Knowing the causative mutation of the disease, not only provides a psychological help to the patient, which identifies the cause of his pathological condition, but allows clinicians to improve the counselling and the management of the disease. The correct clinical diagnosis, in fact, permits the clinician to outline the critical steps in the therapeutic and / or preventive path to limit or slow down the clinical course. For researchers, on the other hand, it provides the basis for the development of therapeutic approaches focused on gene and cell therapy, which, once validated, would allow the complete recovery of the visual capacity and the conduct of a healthy life.

References

- Akahori M., Tsunoda K., Miyake Y., Fukuda Y., Ishiura H., Tsuji S., Usui T., Hatase T., Nakamura M., Ohde H., Itabashi T., Okamoto H., Takada Y. and Iwata T. (2010). "Dominant Mutations in RP1L1 Are Responsible for Occult Macular Dystrophy". *Am J Hum Gen*, 2010;87(3), pp.424-429.
- Ala-Laurila P., Kolesnikov A.V., Crouch R.K., Tsina E., Shukolyukov S.A., Govardovskii V.I., Koutalos Y., Wiggert B., Estevez M.E., Cornwall M.C. (2006). "Visual cycle: Dependence of retinol production and removal on photoproduct decay and cell morphology". *J Gen Physiol*. 2006 Aug;128(2):153-69. Epub 2006 Jul 17.
- Alvarez-Cubero M.J., Saiz M., Martínez-García B., Sayalero S.M., Entrala C., Lorente J.A., Martinez-Gonzalez L.J. (2017). "Next generation sequencing: an application in forensic sciences?". *Ann Hum Biol*. 2017 Nov;44(7):581-592. doi: 10.1080/03014460.2017.1375155.
- Alvisi G., Trevisan M., Masi G., Canel V., Caenazzo L., Nespeca P., Barzon L., Di Iorio E., Barbaro V., Palù G. (2018). "Generation of a transgene-free human induced pluripotent stem cell line (UNIPDi001-A) from oral mucosa epithelial stem cells". *Stem Cell Res*. 2018 Apr;28:177-180. doi: 10.1016/j.scr.2018.02.007.
- Anderson D.H., Radeke M.J., Gallo N.B., Chapin E.A., Johnson P.T., Curletti C.R., Hancox L.S., Hu J., Ebright J.N., Malek G., Hauser M.A., Rickman C.B., Bok D., Hageman G.S., Johnson L.V. (2010). "The pivotal role of the complement system in aging and age-related macular degeneration: hypothesis re-visited". *Prog Retin Eye Res*. 2010 Mar;29(2):95-112.
- Barzon L., Lavezzo E., Militello V., Toppo S., Palù G. (2011). "Applications of Next-Generation Sequencing Technologies to Diagnostic Virology". *Int. J. Mol. Sci*. 2011, 12, 7861-7884.
- Børsting C. and Morling N. (2015). "Next generation sequencing and its application in forensic genetics". *Forensic Sci Int Genet*. 2015 Sep;18:78-89. doi: 10.1016/j.fsigen.2015.02.002.
- Bowne S.J., Daiger S.P., Malone K.A., Heckenlively J.R., Kennan A., Humphries P., Hughbanks-Wheaton D., Birch D.G., Liu Q., Pierce E.A., Zuo J., Huang Q., Donovan D.D., Sullivan L.S. (2003). "Characterization of RP1L1, a highly polymorphic paralog of the retinitis pigmentosa 1 (RP1) gene". *Mol Vis*, 2003;24;9:129-37.
- Brandstetter C., Patt J., Holz F.G., Krohne T.U. (2016). "Inflammasome priming increases retinal pigment epithelial cell susceptibility to lipofuscin phototoxicity by changing the cell death mechanism from apoptosis to pyroptosis". *J Photochem Photobiol B*. 2016 Aug;161:177-83. doi: 10.1016/j.jphotobiol.2016.05.018.
- Broadgate S., Yu J., Downes S., Halford S. (2017). "Unravelling the Genetics of Inherited Retinal Dystrophies: Past, Present and Future". *Progr Retin Eye Res*, vol. 59, 2017, pp. 53–96., doi:10.1016/j.preteyeres.2017.03.003.
- Burnight E.R., Giacalone J.C., Cooke J.A., Thompson J.R., Bohrer L.R., Chirco K.R., Drack A.V., Fingert J.H., Worthington K.S., Wiley L.A., Mullins R.F., Stone E.M., Tucker B.A. (2018). "CRISPR-Cas9 genome engineering: Treating inherited retinal degeneration". *Prog Retin Eye Res*. 2018 Jul;65:28-49. doi: 10.1016/j.preteyeres.2018.03.003.
- Caenazzo L.: Biobanche. (2012) Limena: libreriauniversitaria.it.
- Campa C., Costagliola C., Incorvaia C., Sheridan C., Semeraro F., De Nadai K., Sebastiani A., Parmeggiani F. (2010). "Inflammatory mediators and angiogenic factors in choroidal

neovascularization: pathogenetic interactions and therapeutic implications". *Mediators Inflamm*. 2010;2010. pii: 546826. doi: 10.1155/2010/546826.

- Cella W., Greenstein V.C., Zernant-Rajang J., Smith T.R., Barile G., Allikmets R., Tsang S.H. (2009). "G1961E mutant allele in the Stargardt disease gene ABCA4 causes bull's eye maculopathy". *Exp Eye Res*. 2009 Jun 15; 89(1): 16–24.
- Chiang J., Lamey T., McLaren T., Thompson J., Montgomery H., & De Roach, J. (2015). "Progress and prospects of next-generation sequencing testing for inherited retinal dystrophy". *Expert Rev Mol Diagn* 2015; 15(10), 1269-1275.
- Coles B.L.K., Angénieux B., Inoue T., Del Rio-Tsonis K., Spence J.R., McInnes R.R., Arsenijevic Y., Van der Kooy D. (2004). "Facile isolation and the characterization of human retinal stem cells". *PNAS* 2004; 101(44), 15772–15777.
- Conte I., Lestingi M., den Hollander A., Alfano G., Ziviello C., Pugliese M., Circolo D., Caccioppoli C., Ciccodicola A. and Banfi S. (2003). "Identification and characterisation of the retinitis pigmentosa 1-like1 gene (RP1L1): a novel candidate for retinal degenerations". *Eur J Hum Gen*, 2003;11(2), pp.155-162.
- Crouch R.K., Hazard E.S., Lind T., Wiggert B., Chader G., Corson D.W. (1992). "Interphotoreceptor retinoid-binding protein and alpha-tocopherol preserve the isomeric and oxidation state of retinol". *Photochem Photobiol*. 1992 Aug;56(2):251-5.
- Dias M.F., Joo K., Kemp J.A., Fialho S.L., da Silva Cunha A. Jr, Woo S.J., Kwon Y.J. (2018). "Molecular genetics and emerging therapies for retinitis pigmentosa: Basic research and clinical perspectives". *Prog Retin Eye Res*. 2018 Mar;63:107-131. doi: 10.1016/j.preteyeres.2017.10.004.
- Di Iorio V., Karali M., Brunetti-Pierri R., Filippelli M., Di Fruscio G., Pizzo M., Mutarelli M., Nigro V., Testa F., Banfi S., Simonelli F. (2017). "Clinical and Genetic Evaluation of a Cohort of Pediatric Patients with Severe Inherited Retinal Dystrophies". *Genes* (Basel). 2017 Oct 20;8(10). pii: E280. doi: 10.3390/genes8100280.
- Di Resta C., Ivana Spiga I., Silvia Presi S., Stefania Merella S., Giovanni Battista Pipitone G.B., Manitto M.P., Querques G., Battaglia Parodi M., Ferrari M., Carrera P. (2018). "Integration of multigene panels for the diagnosis of hereditary retinal disorders using Next Generation Sequencing and bioinformatics approaches". *EJIFCC*. 2018 Apr; 29(1): 15–25. PMCID: PMC5949615.
- Eandi C.M., Charles Messance H., Augustin S., Dominguez E., Lavalette S., Forster V., Hu S.J., Siquieros L., Craft C.M., Sahel J.A., Tadayoni R., Paques M., Guillonneau X., Sennlaub F. (2016). "Subretinal mononuclear phagocytes induce cone segment loss via IL-1β". *Elife*. 2016 Jul 20;5. pii: e16490. doi: 10.7554/eLife.16490.
- Ellingford J.M., Horn B., Campbell C., Arno G., Barton S., Tate C., Bhaskar S., Sergouniotis P.I., Taylor R.L., Carss K.J., Raymond L.F.L., Michaelides M., Ramsden S.C., Webster A.R., Black G.C.M. (2018). "assessment of the incorporation of CNV surveillance into gene panel next-generation sequencing testing for inherited retinal diseases". *J Med Genet* 2018;55:114-121.
- Foltz L.P., Clegg D.O. (2018). "Patient-derived induced pluripotent stem cells for modelling genetic retinal dystrophies". *Prog Retin Eye Res*. 2018 Sep 11. pii: S1350-9462(18)30009-0. doi: 10.1016/j.preteyeres.2018.09.002.
- Fraccaro P., Nicolo M., Bonetto M., Giacomini M., Weller P., Traverso C.E., Prosperi M., OSullivan D. (2015). "Combining Macula Clinical Signs and Patient Characteristics for Age-Related Macular Degeneration Diagnosis: a Machine Learning Approach". BMC Ophthalmology, 2015 vol. 15, no. 1, doi:10.1186/1471-2415-15-10.

- Francis P.J., Klein M.L. (2011). "Update on the Role of Genetics in the Onset of Age-Related Macular Degeneration." *Clin Ophthalmol*. 2011;5:1127-33. doi: 10.2147/OPTH.S11627.
- Fransson M.N., Rial-Sebbag E., Brochhausen M., Litton J.E. (2015). "Toward a common language for biobanking". *Eur J Hum Genet.* 2015 Jan;23(1):22-8. doi: 10.1038/ejhg.2014.45.
- Ge Z., Bowles K., Goetz K., Scholl H., Wang F., Wang X., Xu S., Wang K., Wang H. & Chen R. (2015). "NGS-based Molecular diagnosis of 105 eyeGENE® probands with Retinitis Pigmentosa". *Sci Rep* 2015; 5, 18287.
- Hanazono G., Ohde H., Shinoda K., Tsunoda K., Tsubota, K., Miyake Y. (2010). "Pattern-reversal visual-evoked potential in patients with occult macular dystrophy". *Clin Ophthalmol*, 2010 Dec 10;4:1515-20. doi: 10.2147/OPTH.S15088.
- Kabuto T., Takahashi H., Goto-Fukuura Y., Igarashi T., Akahori M., Kameya S., Iwata T., Mizota A., Yamaki K., Miyake Y., Takahashi H. (2012). "A new mutation in the RP1L1 gene in a patient with occult macular dystrophy associated with a depolarizing pattern of focal macular electroretinograms". *Mol Vis*. 2012;18:1031-9.
- Kajiwara K., Berson E.L., Dryja T.P. (1994). "Digenic retinitis pigmentosa due to mutations at the unlinked peripherin/RDS and ROM1 loci". *Science*. 1994 Jun 10;264(5165):1604-8.
- Karlstetter M., Scholz R., Rutar M., Wong W.T., Provis J.M., Langmann T. (2015). "Retinal microglia: just bystander or target for therapy?". *Prog Retin Eye Res*. 2015 Mar;45:30-57. doi: 10.1016/j.preteyeres.2014.11.004.
- Kauppinen A., Paterno J.J., Blasiak J., Salminen A., Kaarniranta K. (2016). "Inflammation and its role in age-related macular degeneration". *Cell Mol Life Sci*. 2016 May;73(9):1765-86. doi: 10.1007/s00018-016-2147-8.
- Klein L.M., Francis P.J., Ferris F.L., Hamon S.C., Clemons T.E. (2011). "Risk Assessment Model for Development of Advanced Age-Related Macular Degeneration". *Arch Ophthalmol*. 2011;129(12):1543-1550.
- Klein R., Myers C.E., Meuer S.M., Gangnon R.E., Sivakumaran T.A., Iyengar S.K., Lee K.E., Klein B.E. (2013). "Risk Alleles in CFH and ARMS2 and the Long-Term Natural History of Age-Related Macular Degeneration: The Beaver Dam Eye Study". *JAMA Ophthalmol*. 2013 Mar;131(3):383-92. doi: 10.1001/jamaophthalmol.2013.713.
- Jones K.D., Wheaton D.K., Bowne S.J.., Sullivan LS., Birch D.G., Chen R., Daiger S.P. (2017). "Next-generation sequencing to solve complex inherited retinal dystrophy: A case series of multiple genes contributing to disease in extended families". *Mol Vis*. 2017 Jul 20;23:470-481.
- Jonsson F., Westin I.M., Österman L., Sandgren O., Burstedt M., Holmberg M., Golovleva I. (2018). "ATP-binding cassette subfamily A, member 4 intronic variants c.4773+3A>G and c.5461-10T>C cause Stargardt disease due to defective splicing". *Acta Ophthalmol*. 2018 Feb 20. doi: 10.1111/aos.13676.
- Lamba D.A., McUsic A., Hirata R.K., Wang P., Russell D., Reh A.T. (2010). "Generation, purification and transplantation of photoreceptors derived from human induced pluripotent stem cells". *PlosONE* 2010; 5(1), e8763.
- Latella M., Di Salvo M., Cocchiarella F., Benati D., Grisendi G., Comitato A., Marigo V., Recchia A. (2016). "In vivo Editing of the Human Mutant Rhodopsin Gene by Electroporation of Plasmid-based CRISPR/Cas9 in the Mouse Retina". *Mol Ther Nucleic Acids*. 2016 Nov 22;5(11):e389. doi: 10.1038/mtna.2016.92.
- Lee W., Xie Y., Zernant J., Yuan B., Bearelly S., Tsang S.H., Lupski J.R., Allikmets R. (2016). "Complex inheritance of ABCA4 disease: four mutations in a family with multiple macular phenotypes". *Hum Genet*. 2016 Jan;135(1):9-19. doi: 10.1007/s00439-015-1605-y.

- Miano M.G., Testa F., Filippini F., Trujillo M., Conte I., Lanzara C., Millán J.M., De Bernardo C., Grammatico B., Mangino M., Torrente I., Carrozzo R., Simonelli F., Rinaldi E., Ventruto V., D'Urso M., Ayuso C., Ciccodicola A. (2001). "Identification of novel RP2 mutations in a subset of X-linked retinitis pigmentosa families and prediction of new domains". Hum Mutat. 2001 Aug;18(2):109-19.
- Miyake Y., Ichikawa K., Shiose Y., Kawase Y. (1989). "Hereditary macular dystrophy without visible fundus abnormality". *Am J Ophthalmol*. 1989;108:292–9.
- Murakami Y., Yoshida N., Ikeda Y., Nakatake S., Fujiwara K., Notomi S., Nabeshima T., Nakao S., Hisatomi T., Enaida H., Ishibashi T. (2015). "Relationship between aqueous flare and visual function in retinitis pigmentosa". *Am J Ophthalmol*. 2015 May;159(5):958-63.e1. doi: 10.1016/j.ajo.2015.02.001.
- Nagasaka Y., Ito Y., Ueno S., Terasaki H. (2016). "Increased aqueous flare is associated with thickening of inner retinal layers in eyes with retinitis pigmentosa". *Sci Rep*. 2016 Sep 22;6:33921. doi: 10.1038/srep33921.
- Nakazato T., Ohta T., Bono H. (2013). "Experimental design-based functional mining and characterization of high-throughput sequencing data in the sequence read archive". *PLoS One*. 2013 Oct 22;8(10):e77910. doi: 10.1371/journal.pone.0077910.
- Parmeggiani F., Barbaro V., De Nadai K., Lavezzo E., Toppo S., Chizzolini M., Perri P., Palù G., Parolin C., Di Iorio E. (2016). "Identification of X-linked dominant-negative inheritance of novel RPGR-ORF15 mutation in Italian family with retinitis pigmentosa and pathologic myopia". *Sci Rep*. 2016 Dec 20;6:39179. doi: 10.1038/srep39179.
- Parmeggiani F., Barbaro V., Migliorati A., Raffa P., Nespeca P., De Nadai K., Del Vecchio C., Palù G., Parolin C., Di Iorio E. (2017). "Novel variants of RPGR in X-linked Retinitis Pigmentosa families and genotype-phenotype correlation". *Eur J Ophthalmol*. 2017 Mar 10;27(2):240-248. doi: 10.5301/ejo.5000879.
- Parmeggiani F., Milan E., Steindler P. (2002). "Retinopatia pigmentosa: inquadramento clinico e terapeutico". In: La gestione clinica e riabilitativa del paziente ipovedente. Eds Fabiano Editore, Canelli (AT), pp 81-88.
- Parmeggiani F., Romano M.R., Costagliola C., Semeraro F., Incorvaia C., D'Angelo S., Perri P., De Palma P., De Nadai K., Sebastiani A. (2012). "Mechanism of inflammation in age-related macular degeneration". *Mediators Inflamm*. 2012; 2012:546786. doi: 10.1155/2012/546786.
- Parmeggiani F., Sorrentino F.S., Romano M.R., Costagliola C., Semeraro F., Incorvaia C., D'Angelo S., Perri P., De Nadai K., Bonomo Roversi E., Franceschelli P., Sebastiani A., Rubini M. (2013). "Mechanism of inflammation in age-related macular degeneration: an up-to-date on genetic landmarks". *Mediators Inflamm*. 2013;2013:435607. doi: 10.1155/2013/435607.
- Peng Y.Q., Tang L.S, Yoshida S., Zhou Y.D. (2017). "Applications of CRISPR/Cas9 in retinal degenerative diseases". *Int J Ophthalmol*. 2017 Apr 18;10(4):646-651. doi: 10.18240/ijo.2017.04.23.
- Pennington K.L. and Deangelis M.M. (2016). "Epidemiology of Age-Related Macular Degeneration (AMD): Associations with Cardiovascular Disease Phenotypes and Lipid Factors". *Eye Vis (Lond).* 2016 Dec 22;3:34. doi: 10.1186/s40662-016-0063-5.
- Perez-Carro R., Corton M., Sánchez-Navarro I., Zurita O., Sanchez-Bolivar N., & Sánchez-Alcudia R., Lelieveld H., Aller E., Lopez-Martinez M.A., Lòpez-Molina I., Fernandez-San Jose, Blanco-Kelly F., Riveiro-Alvarez R., Gilissen C., Millan J.M., Avila-Fernandez A.,

Ayus C. (2016). “Panel-based NGS reveals novel pathogenic mutations in autosomal recessive retinitis pigmentosa”. *Sci Rep*, 2016; 6, 19531.

- Piermarocchi S., Segato T., Leon A., Colavito D., Miotto S. (2015). “Occult macular dystrophy in an Italian family carryng a mutation in RP1L1 gene”. *Mol Med Rep*. 2016 Mar;13(3):2308-12. doi: 10.3892/mmr.2016.4784.
- Ross M.G., Russ C., Costello M., Hollinger A., Lennon N.J., Hegarty R., Nusbaum C., Jaffe D.B. (2013). “Characterizing and measuring bias in sequence data”. *Genome Biol.* 2013 May 29;14(5):R51. doi: 10.1186/gb-2013-14-5-r51.
- Sanger F., Coulson A.R. (1975). “A rapid method for determining sequences in DNA by primed synthesis with DNA polymerase”. *J Mol Biol*. 1975 May 25;94(3):441-8. No abstract available.
- Scholl H.P.N., Fleckenstein M., Issa P.C., Keilhauer C., Holz F.G., Weber B.H.F. (2007). “Genetics of Age-Related Macular Degeneration: Update”. *Essentials in Ophthalmology Medical Retina*, 2007, pp. 35–52., doi:10.1007/978-3-540-33672-3_3.
- Scholl H.P.N., Fleckenstein M., Charbel Issa P.C., Claudia Keilhauer C., Holz F.G., Bernhard B.H.F. (2007). “An update on the genetics of age-related macular degeneration”. *Mol Vis*.; 2007;13: 196–205.
- Schulz H.L., Grassmann F., Kellner U., Spital G., Rüther K., Jägle H., Hufendiek K., Rating P., Huchzermeyer C., Baier M.J., Weber B.H., Stöhr H. (2017). “ Mutation Spectrum of the ABCA4 Gene in 335 Stargardt Disease Patients From a Multicenter German Cohort-Impact of Selected Deep Intronic Variants and Common SNPs”. *Invest Ophthalmol Vis Sci*. 2017 Jan 1;58(1):394-403. doi: 10.1167/iovs.16-19936.
- Stenson P., Mort M., Ball E., Shaw K., Phillips A., Cooper D. (2014). “The Human Gene Mutation Database: building a comprehensive mutation repository for clinical and molecular genetics, diagnostic testing and personalized genomic medicine”. *Hum Genet*. 2014;133(1), 1-9.
- Strauss O. (2005). “The retinal pigment epithelium in visual function”. *Physiological Reviews*, 2005;85, no. 3, pp. 845–881., doi:10.1152/physrev.00021.2004.
- Stuck M., Conley S., Naash M. (2016). “PRPH2/RDS and ROM-1: Historical context, current views and future considerations”. *Prog Retin Eye Res*. 2016 May;52:47-63. doi: 10.1016/j.preteyeres.2015.12.002.
- Swaroop A., Chew E.Y., Bowes Rickman C., Abecasis G.R. (2009). “Unraveling a Multifactorial Late-Onset Disease: From Genetic Susceptibility to Disease Mechanisms for Age-Related Macular Degeneration”. *Ann Rev Genomics Hum Genet*, 2009;10:19-43. doi: 10.1146/annurev.genom.9.081307.164350.
- Takahashi H., Hayashi T., Tsuneoka H., Nakano T., Yamada H., Katagiri S., Fujino Y., Noda Y., Yoshimoto M., Kawashima H. (2014). “Occult macular dystrophy with bilateral chronic subfoveal serous retinal detachment associated with a novel RP1L1 mutation (p.S1199P)”. *Doc Ophthalmol*, 2014;129:49–56.
- Takahashi K., Yamanaka S. (2006) “Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors”. *Cell* 2006; 126, 663–676.
- Taylor K., Holzer A., Bazan J., Walsh C., Gleeson J. (2000). “Patient Mutations in Doublecortin Define a Repeated Tubulin-binding Domain”. *J Biol Chem*, 2000 Nov 3;275(44):34442-50.
- Trevisan M., Barbaro V., Riccetti S., Masi G., Barzon L., Nespeca P., Alvisi G., Di Iorio E., Palù G. (2018). "Generation of a transgene-free induced pluripotent stem cells line (UNIPDi002-A) from oral mucosa epithelial stem cells carrying the R304Q mutation in TP63

gene. *Stem Cell Res*. 2018 Apr;28:149-152. doi: 10.1016/j.scr.2018.02.006. Epub 2018 Feb 16.

- Verbakel S.K., van Huet R.A.C., Boon C.J.F., den Hollander A.I., Collin R.W.J., Klaver C.C.W., Hoyng C.B., Roepman R., Klevering B.J. (2018). "Non-syndromic retinitis pigmentosa". *Prog Retin Eye Res*. 2018 Sep;66:157-186. doi: 10.1016/j.preteyeres.2018.03.005.
- Wang F., Wang H., Tuan H., Nguyen D., Sun V., Keser V., Bowne S.J., Sullivan L.S., Luo H., Zhao L., Wang X., Zaneveld J.E., Salvo J.S., Siddiqui S., Mao L., Wheaton D.K., Birch D.G., Branham K.E., Heckenlively J.R., Wen C., Flagg K., Ferreyra H., Pei J., Khan A., Ren H., Wang K., Lopez I., Qamar R., Zenteno J.C., Ayala-Ramirez R., Buentello-Volante B., Fu Q., Simpson D.A., Li Y., Sui R., Silvestri G., Daiger S.P., Koenekoop R.K., Zhang K., Chen R. (2014). "Next generation sequencing-based molecular diagnosis of retinitis pigmentosa: identification of a novel genotype-phenotype correlation and clinical refinements". *Hum Genet* 2014; 133(3), 331-345.
- Wawrocka A., Skorczyk-Werner A., Wicher K., Niedziela Z., Ploski R., Rydzanicz M., Sykulski M., Kociecki J., Weisschuh N., Kohl S., Biskup S., Wissinger B., Krawczynski M.R. (2018). "Novel variants identified with next-generation sequencing in Polish patients with cone-rod dystrophy". *Mol Vis*. 2018;24: 326–339.
- Wong W.L., Su X., Li X., Cheung C.M., Klein R., Cheng C.Y., Wong T.Y. (2014). "Global Prevalence of Age-Related Macular Degeneration and Disease Burden Projection for 2020 and 2040: a Systematic Review and Meta-Analysis". *The Lancet Global Health*, 2014;2, no. 2, doi:10.1016/s2214-109x(13)70145-1.
- Xu H., Chen M., Forrester JV. (2009). "Para-inflammation in the aging retina". *Prog Retin Eye Res*. 2009 Sep;28(5):348-68.
- Yoshiura K., Noda Y., Kinoshita A., Niikawa, N. (2000). "Colocalization of doublecortin with the microtubules: An ex vivo colocalization study of mutant doublecortin". *J Neurobiol*, 2000 May;43(2):132-9.
- Yuong R.W.(1994). "The family of the sunlight-related eye diseases". *Optom Vis Sci* 1994; 71, 125-144.
- Zhao L., Wang F., Wang H., Li Y., Alexander S., Wang K., Willoughby C.E., Zaneveld J.E., Jiang L., Soens Z.T., Earle P., Simpson D., Silvestri G., Chen R. (2015). "Next-generation sequencing-based molecular diagnosis of 82 retinitis pigmentosa probands from Northern Ireland". *Hum Genet* 2015; 134(2), 217-230.