Experimental Eye Research 145 (2016) 93–99

Contents lists available at ScienceDirect

Experimental Eye Research

journal homepage: www.elsevier.com/locate/yexer

Research article

Next-generation sequencing of *ABCA4*: High frequency of complex alleles and novel mutations in patients with retinal dystrophies from Central Europe

Aneta Ścieżyńska [a], Dominika Oziębło [b], Anna M. Ambroziak [c, d], Magdalena Korwin [c], Kamil Szulborski [c], Maciej Krawczyński [e, f], Piotr Stawiński [b], Jerzy Szaflik [c], Jacek P. Szaflik [c], Rafał Płoski [g, **], Monika Ołdak [a, b, *]

[a] Department of Histology and Embryology, Medical University of Warsaw, Warsaw, Poland
[b] Department of Genetics, Institute of Physiology and Pathology of Hearing, Warsaw/Kajetany, Poland
[c] Department of Ophthalmology, Medical University of Warsaw, Warsaw, Poland
[d] Faculty of Physics, University of Warsaw, Warsaw, Poland
[e] Department of Medical Genetics, Poznan University of Medical Sciences, Poznan, Poland
[f] Center for Medical Genetics GENESIS, Poznan, Poland
[g] Department of Medical Genetics, Medical University of Warsaw, Warsaw, Poland

ARTICLE INFO

Article history:
Received 21 May 2015
Received in revised form
4 September 2015
Accepted in revised form 11 November 2015
Available online 22 November 2015

Keywords:
Stargardt disease
Cone-rod dystrophy
ABCA4
Next-generation sequencing
Complex allele
Mutation

ABSTRACT

Variation in the *ABCA4* locus has emerged as the most prevalent cause of monogenic retinal diseases. The study aimed to discover causative *ABCA4* mutations in a large but not previously investigated cohort with *ABCA4*-related diseases originating from Central Europe and to refine the genetic relevance of all identified variants based on population evidence. Comprehensive clinical studies were performed to identify patients with Stargardt disease (STGD, n = 76) and cone-rod dystrophy (CRD, n = 16). Next-generation sequencing targeting *ABCA4* was applied for a widespread screening of the gene. The results were analyzed in the context of exome data from a corresponding population (n = 594) and other large genomic databases. Our data disprove the pathogenic status of p.V552I and provide more evidence against a causal role of four further *ABCA4* variants as drivers of the phenotype under a recessive paradigm. The study identifies 12 novel potentially pathogenic mutations (four of them recurrent) and a novel complex allele p.[(R152*; V2050L)]. In one third (31/92) of our cohort we detected the p.[(L541P; A1038V)] complex allele, which represents an unusually high level of genetic homogeneity for *ABCA4*-related diseases. Causative *ABCA4* mutations account for 79% of STGD and 31% of CRD cases. A combination of p.[(L541P; A1038V)] and/or a truncating *ABCA4* mutation always resulted in an early disease onset. Identification of *ABCA4* retinopathies provides a specific molecular diagnosis and justifies a prompt introduction of simple precautions that may slow disease progression. The comprehensive, population-specific study expands our knowledge on the genetic landscape of retinal diseases.

1. Introduction

ABCA4 gene (OMIM:601691; GenBank:NG_009073.1) mutations are among the leading causes of monogenic retinal diseases. More than 95% of cases of Stargardt's disease (STGD), the most common inherited juvenile macular degeneration, 30% of cases of cone-rod dystrophy (CRD) and about 8% of autosomal recessive retinitis pigmentosa (arRP) are caused by mutations in the *ABCA4* gene (Sheffield and Stone, 2011), which encodes an ATP-binding cassette transporter. The function of ABCA4 is to translocate retinoid

Abbreviations: STGD, Stargardt disease; CRD, con-rode dystrophy; RPE, retinal pigment epithelium; arRP, autosomal recessive retinitis pigmentosa; NGS, next-generation sequencing; SNPs, single nucleotide polymorphisms; AO, age of onset; FA, fluorescein angiography; ERG, electroretinography; OCT, optical coherence tomography; PCR, polymerase chain reaction; ESP, exome sequencing project; ExAC, exome aggregation consortium; HGMD, human gene mutation database; DM, disease-causing mutations; OR, odds ratio; CI, confidence interval; NBD, nucleotide-binding domain; ECD, extracellular domain; MSD, membrane-spanning domain; H, hydrophobic domain; SE, standard error.

* Corresponding author. Department of Genetics, Institute of Physiology and Pathology of Hearing, Mokra 17, Kajetany, 05-830, Nadarzyn, Poland.
** Corresponding author. Department of Medical Genetics, Center of Biostructure Research, Medical University of Warsaw, Pawińskiego 3c, 02-106, Warsaw, Poland.
E-mail addresses: rploski@wp.pl (R. Płoski), m.oldak@ifps.org.pl (M. Ołdak).

http://dx.doi.org/10.1016/j.exer.2015.11.011
0014-4835/

intermediates of the visual cycle across the photoreceptor outer segment disc membranes. Reduced activity of the ABCA4 protein results in the accumulation of retinoid intermediates, which are deleterious to photoreceptors and retinal pigment epithelium (RPE) cells (Tsybovsky et al., 2010).

The *ABCA4* gene has been difficult to investigate due to its size (50 exons, over 6800-bp open reading frame) (www.ensembl.org; accessed 12/2014) and a large allelic diversity. *ABCA4* is almost exclusively expressed in the retina (Allikmets et al., 1997) and absence of an established functional assay makes the assessment of the biological role of different *ABCA4* variants a difficult task. Introduction of next-generation sequencing (NGS) allows an efficient and cost-effective screening of the sizable gene. An indispensable part of NGS data analysis is access to genome sequencing data with a large number of samples, which provides a basic tool in determining pathogenicity of *ABCA4* variants. Identification of causal *ABCA4* mutations is of practical relevance for affected families as precautions that may reduce disease progression, such as protection from excessive light exposure and avoidance of vitamin A supplementation, should be followed until therapy options become available (Haji Abdollahi and Hirose, 2013; Paskowitz et al., 2006; Radu et al., 2008; Teussink et al., 2015; Weng et al., 1999).

Currently, more than 800 different *ABCA4* disease-associated variants are reported in the Human Gene Mutation Database (http://www.hgmd.cf.ac.uk/ac/index.php; accessed 08/2015). Some of them have a high prevalence in certain ethnic groups, such as p.C1490Y in South Africans (September et al., 2004), p.A1773V in Mexicans (Chacon-Camacho et al., 2013) and a number of other mutations specific for different European populations (Rivera et al., 2000; Valverde et al., 2006; Maugeri et al., 1999; Rosenberg et al., 2007), which underlines the need of genotyping patients of various ethnicities. As the prevalence of *ABCA4* variants has not yet been determined for Central European patients with STGD or CRD, we have applied an amplicon-based deep sequencing strategy to investigate the *ABCA4* gene. All potentially pathogenic variants identified using NGS were carefully compared with the data from several, large genomic databases to get a better insight into the prevalence of *ABCA4* mutations in different populations.

2. Patients and methods

2.1. Subjects and clinical evaluation

A total of 92 unrelated individuals with STGD (n = 76) or CRD (n = 16) from the Polish population with a mean age at disease onset (AO) of 17.3 years (CI 95%: 14.8–19.8) were recruited for the study (Supplementary Table S1). Patients and examined family members gave written consent. The study was approved by the local Ethics Committees and followed the tenets of the Declaration of Helsinki. AO was defined as the age at which visual deterioration was first observed by the patient. Next, the patients were divided into the groups with an early (childhood) and late (adult) AO (Fujinami et al., 2014). The cutoff between early and late AO were 18 years, as this age is generally regarded as the beginning of adulthood. The diagnosis of STGD was based on the presence of bilateral impairment of central vision together with irregular hyperfluorescent lesions at the macula and a dark-choroid background in fluorescein angiography (FA). In flash electroretinography (ERG, RETIscan, Roland Consult, Germany) full-field rod responses were normal and full-field cone responses were either normal or slightly reduced. Multifocal ERG was abnormal in every patient in whom it was measured. Optical coherence tomography (OCT) scans (Cirrus HD-OCT, Carl Zeiss Meditec, Dublin, CA) showed a decreased thickness of the retina, most notably in the foveola. The diagnosis of CRD was based on the presence of bilateral, progressive visual loss with concomitant nyctalopia, color vision abnormalities and photophobia. The eye fundus examination revealed panretinal degeneration affecting the macula more severely. In flash ERG full-field rod responses were moderately reduced, as well as cone-rod responses, whereas full-field cone responses were severely reduced.

2.2. Next-generation DNA sequencing (NGS)

DNA was isolated from blood samples using a standard salting out procedure or extracted from buccal swabs with the AxyPrep MAG Tissue-Blood gDNA Kit (Axygen, Union City, CA, USA). DNA amplicon libraries were prepared by amplifying the entire coding region of the *ABCA4* gene (GenBank:NM_000350.2) with primers located in the noncoding sequences and designed using the Primer3Plus software (http://www.bioinformatics.nl/cgi-bin/primer3plus/primer3plus.cgi/). The first set of primers contained a template specific sequence designed separately for each *ABCA4* exon and a universal M13 sequence tag at the 5′ ends (forward 5′-TGTAAAACGACGGCCAGT-3′ and reverse 5′-CAGGAAACAGCTATGACC-3′). The second set of primers contained the sequence of the GS FLX Titanium Primer (forward 5′-CGTATCGCCTCCCTCGCGCCA-3′ or reverse 5′-CTATGCGCCTTGCCAGCCCGC-3′) followed by the key sequence (5′-TCAG-3′), one of ten different multiplex identifiers (MIDs), according to the manufacturer's guideline (GS Junior system, Roche, Switzerland) and the M13 sequence. The second set of primers was used for reamplification of the first-round PCR products (primer sequences are available upon request). PCR products (400–500 bp) were examined on 1.5% agarose gels and purified with AMPure XP beads (New England Biolabs, UK). Amplicon profiles were confirmed using the Agilent 2100 Bioanalyzer (Agilent Technologies, Palo Alto, CA). Next, amplicon libraries were sequenced on the Roche 454 GS Junior Sequencer (Roche) and rigorous quality control was performed. On average 90,000 reads were generated per run and an average coverage was above 50 reads with a balanced ratio of forward and reverse reads. The data was analyzed using an accompanying GS Data Analysis Software package (Roche). Confirmation of the presence of novel and pathogenic variants as well as variants' segregation within families was performed by bidirectional DNA sequencing on the ABI PRISM 3500XL Genetic Analyzer (Applied Biosystems, Foster City, CA). The results were entirely consistent with those from NGS.

2.3. Bioinformatic and statistical analysis

Population frequencies of the *ABCA4* variants were obtained from the database of the 1000 Genomes Project (Abecasis et al., 2012), the NHLBI GO Exome Sequencing Project (ESP) (https://esp.gs.washington.edu/drupal/), the Exome Aggregation Consortium (ExAC) (http://exac.broadinstitute.org) (all accessed 12/2014) and the whole exome sequencing data from the Polish population (n = 594) (ZGM, R. Płoski, unpublished results).

Possible functional consequences of the novel *ABCA4* missense variants were evaluated using PolyPhen-2 (Adzhubei et al., 2013) and SIFT (Kumar et al., 2009) and were considered pathogenic when at least one of the programs predicted their deleterious effect on the protein structure and function. Novel *ABCA4* splice site and missense variants were analyzed using the Alamut Visual Software (Interactive Biosoftware, Rouen, France). The prediction was considered to be strong when at least two Alamut algorithms indicated a decrease by at least 10% in the score of natural splice site (Thery et al., 2011).

The two-sided Fisher exact test and chi-square statistics were applied to compare *ABCA4* allele frequencies between patients and