Vision Research xxx (2017) xxx–xxx

Contents lists available at ScienceDirect

Vision Research

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

Whole exome sequencing identification of novel candidate genes in patients with proliferative diabetic retinopathy

Cindy Ung [a,1], Angie V. Sanchez [b,1], Lishuang Shen [c], Samaneh Davoudi [a], Tina Ahmadi [a], Daniel Navarro-Gomez [a], Ching J. Chen [d], Heather Hancock [d], Alan Penman [e], Suzanne Hoadley [d], Mark Consugar [a], Carlos Restrepo [f], Vinay A. Shah [g], Joseph F. Arboleda-Velasquez [b,*,2], Lucia Sobrin [a,*,2], Xiaowu Gai [c,*,2], Leo A. Kim [a,*,2]

[a] Department of Ophthalmology, Massachusetts Eye and Ear Infirmary, Harvard Medical School, Boston, MA, USA
[b] Schepens Eye Research Institute, Massachusetts Eye and Ear, Department of Ophthalmology, Harvard Medical School, Boston, MA, USA
[c] Center for Personalized Medicine, Children's Hospital Los Angeles, Los Angeles, CA, USA
[d] Department of Ophthalmology, University of Mississippi Medical Center, Jackson, MS, USA
[e] Department of Medicine, University of Mississippi Medical Center, Jackson, MS, USA
[f] Basic Science Group, School of Medicine, CES University, Medellin, Colombia
[g] Department of Ophthalmology, Dean McGee Eye Institute, University of Oklahoma, OK, USA

ARTICLE INFO

Article history:
Received 10 February 2017
Received in revised form 28 March 2017
Accepted 30 March 2017
Available online xxxx

Keywords:
Whole exome sequencing
Diabetic retinopathy
Candidate genes
Aberrant angiogenesis

ABSTRACT

Rare or novel gene variants in patients with proliferative diabetic retinopathy may contribute to disease development. We performed whole exome sequencing (WES) on patients at the phenotypic extremes of diabetic retinal complications: 57 patients diagnosed with proliferative diabetic retinopathy (PDR) as cases and 13 patients with no diabetic retinopathy despite at least 10 years of type 2 diabetes as controls. Thirty-one out of the 57 cases and all 13 controls were from the African American Proliferative Diabetic Retinopathy Study (AA). The rest of the cases were of mixed ethnicities (ME). WES identified 721 candidate genes with rare or novel non-synonymous variants found in at least one case with PDR and not present in any controls. After filtering for genes with null alleles in greater than two cases, 28 candidate genes were identified in our ME cases and 16 genes were identified in our AA cases. Our analysis showed rare and novel variants within these genes that could contribute to the development of PDR, including rare non-synonymous variants in *FAM132A*, *SLC5A9*, *ZNF600*, and *TMEM217*. We also found previously unidentified variants in *VEGFB* and *APOB*. We found that *VEGFB*, *VPS13B*, *PHF21A*, *NAT1*, *ZNF600*, *PKHD1L1* expression was reduced in human retinal endothelial cells (HRECs) cultured under high glucose conditions. In an exome sequence analysis of patients with PDR, we identified variants in genes that could contribute to pathogenesis. Six of these genes were further validated and found to have reduced expression in HRECs under high glucose conditions, suggestive of an important role in the development of PDR.

1. Introduction

Diabetic retinopathy (DR) is the leading cause of blindness among U.S. working-aged adults aged 20–74 years (Ruta et al., 2013). The Diabetes Control and Complications Trial (DCCT) showed the strongest factors (duration of diabetes and hemoglobin A1c) explained 11% of the risk of developing retinopathy (Lachin et al., 2008). Similarly, the Wisconsin Epidemiologic Study of Diabetic Retinopathy (WESDR), (Klein, Klein, Moss, & Cruickshanks, 1998) a large population based study, showed that hemoglobin A1c, cholesterol and blood pressure only accounted for 10% of the risk of developing retinopathy, which suggests that other factors may influence the variation of DR. Twin studies and family studies have implicated strong genetic components in DR with heritability scores ranging from 25% to 52% for proliferative diabetic retinopathy (PDR) in either type 1 or type 2 diabetes (DM) (Hietala, Forsblom, Summanen, Groop, & FinnDiane Study, 2008).

* Corresponding authors at: 20 Staniford Street, Boston, MA 02114, USA (J.F. Arboleda-Velasquez). 243 Charles St. Boston, MA 02114, USA (L. Sobrin). 4650 Sunset Blvd, Los Angeles, CA 90027, USA (X. Gai). 20 Staniford Street, Boston, MA 02114, USA (L.A. Kim).
E-mail addresses: joseph_arboleda@meei.harvard.edu (J.F. Arboleda-Velasquez), lucia_sobrin@meei.harvard.edu (L. Sobrin), xgai@chla.usc.edu (X. Gai), leo_kim@meei.harvard.edu (L.A. Kim).
[1] Equal contributors.
[2] These authors contributed equally to this work.

http://dx.doi.org/10.1016/j.visres.2017.03.007
0042-6989/

However, previous analyses, including candidate gene and genome wide association studies (GWAS), have failed to identify genes that are reproducibly associated with DR (Abhary, Hewitt, Burdon, & Craig, 2009; Awata et al., 2014; Burdon et al., 2015; Cheung et al., 2016; Fu et al., 2010; Grassi et al., 2011, 2012; Hosseini et al., 2015; Huang et al., 2011; Peng et al., 2015; Sheu et al., 2013; Shtir et al., 2016). This failure has been attributed to small sample size, incomplete phenotyping of patients, and lack of data on rare variants in such studies. Whole exome sequencing (WES) of individuals at the phenotypic extremes of disease has previously been successful in identifying genetic factors in conditions for which genetic analyses of common variants have failed (Yuen et al., 2015). In contrast to GWAS, which employs SNPs mostly in non-coding regions to identify common markers that are in linkage disequilibrium with the functional or causal variants, studies in individuals with extreme phenotypes have often detected rare variants in coding regions with large functional effects. WES is a highly effective approach in discovering genes underlying multifactorial diseases. Only one study employing this approach has been published thus far. It found three genes that were associated with protection from DR using a gene-based approach (Shtir et al., 2016). WES of extreme phenotypes is a study design conceived and used successfully by Emond et al. in identifying *DCTN4* gene as a modifier of chronic *Pseudomonas aeruginosa* infection in patients with cystic fibrosis (Emond et al., 2012). In their study, they performed WES on 41 patients and 48 controls. The same strategy was successfully applied by the same group in discovering another two genes, *CAV2* and *TMC6*, as modifiers of cystic fibrosis (Emond et al., 2015).

Here, we hypothesized that rare or novel variants, especially null alleles, are enriched in patients with PDR and may be involved in the pathogenesis of diabetic retinopathy. We used WES to identify rare variants of large effect in individuals at the extremes of the phenotypic spectrum of diabetic retinopathy: no DR with at least 10 years of type 2 diabetes mellitus (controls) and PDR (cases). We focused especially on frameshift, nonsense, and splicing variants at the canonical splice sites as these are expected to have severe consequences on gene expression and would have a larger functional impact on the pathogenesis of DR. After candidate genes were identified, we performed functional validation studies and investigated the RNA expression of these genes in HRECs under high glucose conditions.

2. Methods

2.1. Study population

The Massachusetts Eye and Ear Infirmary (MEEI), University of Mississippi Medical Center (UMMC), Dean McGee Eye Institute and the University of Oklahoma (DMEI) Institutional Review Boards approved all research involving human subjects. Written informed consent was obtained from all patients. Patients were recruited from two sources. The first is the African American Proliferative Diabetic Retinopathy Study (AAPDR), which has been previously described (Davoudi et al., 2016; Papavasileiou et al., 2016; Penman et al., 2016). All patients from the AAPDR study had a known diagnosis of type 2 diabetes mellitus by the 2003 American Diabetes Association criteria and/or by being on anti-diabetic medication. All patients had bilateral, dilated wide-field fundus photography. Level of retinopathy was scored using the Early Treatment Diabetic Retinopathy Study (ETDRS) adaptation of the modified Airlie House classification ("Grading diabetic retinopathy from stereoscopic color fundus photographs--an extension of the modified Airlie House classification. ETDRS report number 10. Early Treatment Diabetic Retinopathy Study Research Group," 1991) and determined by masked ophthalmologist graders. The second source of patients was from MEEI and DMEI. These patients had type 1 or type 2 diabetes mellitus and consisted of mixed ethnicities. All patients had PDR and surgical treatment with vitrectomy. For the analyses, cases were defined as patients with PDR in at least one eye. Controls were patients with no diabetic retinopathy in either eye and with at least 10 years of diagnosed diabetes (Penman et al., 2016). All controls were from the AAPDR study and were all African American.

After consent was obtained, blood samples were obtained from PDR patients (n = 57) and patients with no DR (n = 13). DNA was extracted from whole blood and stored at -80 degrees Celsius until the sequencing was ready to be performed. Thirty-one out of the 57 cases and all 13 controls were from the AAPDR Study (Penman et al., 2016). The 26 remaining cases recruited from MEEI and DMEI were of different ethnicities. For the analyses, cases were divided into two groups: one group consisting of individuals from the AAPDR Study (AA cases) (Penman et al., 2016) and one group with participants from MEEI and DMEI of mixed ethnicities (ME cases). Analyses included AA cases versus controls (AA group) and ME cases versus controls (AA group). Demographic and clinical information was obtained directly as part of the AAPDR study and from the electronic medical record for MEEI and DMEI patients.

2.2. Whole-exome sequencing

Exome capture was performed using Agilent SureSelect Human All Exon V5 kit (Agilent Technologies, Santa Clara, CA) as per the manufacturer's instructions. Paired-end sequencing (2×101 base pair) was performed on an Illumina (San Diego, CA) HiSeq 2000 Next-Generation Sequencing system using v3.0 SBS chemistry with flowcell lane cluster densities of ~700–800 K/mm^2 on average. One sample was loaded per flowcell lane to obtain a minimum 10x read depth across ~96% of the target regions.

2.3. Exome data analyses

WES data was analyzed with the MEEI Bioinformatics Center standard pipeline (based on human reference genome GRCh37), as previously described (Bujakowska et al., 2014), and updated using BWA (version 0.6.2), Samtools (version 0.1.16 and 0.1.18) and latest version of ENSEMBL, (http://www.ensembl.org/index.html), 1000 Genomes Project (http://www.1000genomes.org), Exome Variant Server (EVS) (http://evs.gs.washington.edu/EVS/), SIFT (http://sift.jcvi.org/), PolyPhen-2 (http://genetics.bwh.harvard.edu/pph2/), UK10K Project (http://www.uk10k.org/) and Exome Aggregation Consortium (ExAC) (http://exac.broadinstitute.org; release 0.3). A coverage depth cutoff of 10× was then applied. Heterozygous was defined as a fraction of a variant base between 0.25–0.75 and homozygous was defined as above 0.75.

Only variants likely to alter protein function, such as missense and loss-of-function mutations, were selected for subsequent analysis. Annotations such as the phastCons score and the GERP scores were extracted from batch downloaded data files for human reference genome build hg19 from the UCSC Genome Browser. Variants were filtered further to include only those with a minor allele frequency less than 0.1% in data from the 1000 Genomes Project, EVS, and ExAC.

2.4. Cell culture and experimental protocol

HRECs (Cell Systems) were grown in endothelial cell growth medium EBM-2 and singlequots (Lonza), antibiotics (penicillin and streptomycin- Lonza), and 4% FBS (Atlanta Biologicals). Media was changed every 48 h until cells reached 80 to 90% confluency. Cells were divided and treated under normal glucose (5 mM of

Please cite this article in press as: Ung, C., et al. Whole exome sequencing identification of novel candidate genes in patients with proliferative diabetic retinopathy. *Vision Research* (2017), http://dx.doi.org/10.1016/j.visres.2017.03.007