Meta Gene 13 (2017) 92–98

Contents lists available at ScienceDirect

Meta Gene

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

High order interaction analysis of SNPs in *PEDF* (rs12150053, rs12948385) and *EPO* (rs1617640) genes with clinical determinants of type 2 diabetic retinopathy patients from south India

Kumari Vinita[a,d], Sarangapani Sripriya[a,*], Ferdina Marie Sharmila Philomenadin[a], Kulothungan Vaitheeswaran[b], Rajiv Raman[c], Tarun Sharma[c]

[a] SN ONGC Department of Genetics and Molecular Biology, Vision Research Foundation, 18/41 College Road, Chennai 600 006, Tamil Nadu, India
[b] Sankara Nethralaya Diabetic Retinopathy Project, Medical Research Foundation, Sankara Nethralaya, 18/41 College Road, Chennai 600 006, Tamil Nadu, India
[c] Shri Bhagwan Mahavir Vitreo retinal Services, Medical Research Foundation, Sankara Nethralaya, 18/41 College Road, Chennai 600 006, Tamil Nadu, India
[d] Birla Institute of Technology & Science (BITS), Pilani, 333 031, Rajasthan, India

ARTICLE INFO

Keywords:
Diabetic retinopathy
Retinal neurodegeneration
Erythropoietin
Pigment epithelial derived factor
GMDR
CART
India

ABSTRACT

Gene environment interaction in complex diseases like type 2 diabetes retinopathy (DR) may provide valuable insight into the complexity of disease. Current study aims to explore the gene-gene and gene-environment interaction in the pathology of type 2 DR. SNPs in *PEDF* (rs12150053, rs12948385) and *EPO* (rs1617640) genes are studied for the potential interactions with various clinical risk factors in type 2 diabetes patients of south Indian origin with (N = 201) and without (N = 168) retinopathy. The interaction study is performed with the statistical tools namely generalized multifactor dimensionality reduction (GMDR) and classification and regression tree (CART) methods. GMDR analysis showed a probable interaction for insulin user status with the 3 SNP that remained significant after adjusting for various clinical factors [Testing balance accuracy of 62.41 and 61.49 respectively]. *PEDF* polymorphisms was also seen interacting with HbA1c levels ($p < 0.05$; TBA > 59). SNP rs12150053 determined the subsequent split among insulin users and poor control of HbA1c by CART. The CC genotype of rs12150053 showed an OR = 4.9 after adjusting for insulin user status. We did not find any direct disease association for the SNPs with DR in the study population. The study showed insulin user status and glycemic index as the probable interacting factors with DR, potentially modified by rs12150053. However, the direct role of these SNPs in regulating these interaction demands functional validation and replication for statistical significance.

1. Introduction

Genetic factors are strongly associated with DR as demonstrated though familial aggregation studies with heritability scores ranging from 18%–27%. Various strategies like candidate gene approaches, linkage studies etc. have suggested the association of genetic risk factors in the development of DR. Genome wide association strategies has also identified various loci for DR. Pathway based Candidate genes of specific pathways are targeted and pathway based approaches are being studied to understand not only the genetic variations but also the interacting environmental factors are being studied worldwide.

Inflammation, vascular abnormalities mediated by hyperglycemia, hyperlipidemia, hormonal/growth factor/oxidative stress and many related factors intercedes the systemic complications in diabetes of which diabetic retinopathy (DR), leads to irreversible blindness (Villarroel et al., 2010). In the early stages of DR, neurodegeneration (ND) precedes the microcirculatory changes as evident from clinically detectable/structural/functional alterations. These changes are associated with imbalance between neurotoxic like (glutamate, angiotensin II) and neuroprotective factors [pigment epithelial derived factor (PEDF), erythropoietin (EPO), neuroprotection D1 (NPD1), brain-derived neurotrophic factor (BDNF), glial cell-line-derived neurotrophic factor (GDNF)] that also shows varied levels in diabetic retina (Villarroel et al., 2010; Barber et al., 1998; Biallosterski et al., 2007;

Abbreviations: CART, classification and regression tree; CVC, cross validation consistency; DR, diabetic retinopathy; *EPO*, erythropoietin; GMDR, generalized multifactor dimensionality reduction; *PEDF*, pigment epithelial derived factor; PE, prediction error

* Corresponding author at: SN ONGC Department of Genetics and Molecular Biology, Vision Research Foundation, New No. 18/41, College Road, Chennai 600 006, India.
E-mail addresses: drss@snmail.org (S. Sripriya), sharmi_ferdi@yahoo.co.in (F.M.S. Philomenadin), vaitheeswaran81@yahoo.co.in (K. Vaitheeswaran), drrrn@snmail.org (R. Raman), drts@snmail.org (T. Sharma).

http://dx.doi.org/10.1016/j.mgene.2017.04.006
Received 1 September 2016; Received in revised form 25 March 2017; Accepted 19 April 2017
Available online 01 May 2017

Hernandez and Simo, 2012).

PEDF, a 50 kDa glycoprotein, expressed in the nervous system and retina, belongs to the serine protease family with potential anti-angiogenic, ant-oxidative and neuroprotective properties that confers its candidature as a novel therapeutic agent (Garcia-Ramirez et al., 2008). Decreased intraocular PEDF levels were observed in the mice models of proliferative diabetic retinopathy (PDR) (Whitmire et al., 2011) and in early phase of experimental DR (Cohen et al., 2008). The vitreous humour protein profiles of PDR patients also demonstrated reduced PEDF levels when compared to the normal human eyes (Wang et al., 2012).

EPO gene (7q21) encodes an angiogenic factor that also has a potent neuroprotective role in early stages of DR and protects the retinal pigment epithelium (RPE) against diabetes mediated pathological changes. An upregulated expression of the protein has been observed in the RPE (Garcia-Ramirez et al., 2008), neuroretina and vitreous of donor eyes from diabetic, PDR patients when compared to the nondiabetic controls (Garcia-Ramirez et al., 2008). Diabetic rats treated with EPO-derived peptide (McVicar et al., 2011) were protected from neuroglial and vascular degenerative pathology thus suggesting its therapeutic implications in DR. However the exact role of this protein in mediating neuroprotection remains unclear. In addition to its neuroprotective role, EPO has also been shown to act synergistically with VEGF in advanced stages of DR to mediate angiogenesis thus suggesting a pathological role for EPO in DR.

PEDF gene polymorphisms (rs12150053 [g.1664469 T > C], rs12948385 [g.1664901G > A]) were studied earlier (Iizuka et al., 2007; Balasubbu et al., 2010; Tong et al., 2008; Abhary et al., 2010) for their association with DR and identified as a significant risk factor for DR in various populations. A functional SNP, rs1617640 [g.100317298C > A], located at −1125 bp upstream to transcription start site of *EPO* gene has been associated with microvascular complications of eye and kidney. The polymorphism was also correlated with altered vitreous levels of the protein (Tong et al., 2008).

In the current study, we have tested the possible single/multi locus interaction (including various clinical risk factors of DR) for promoter polymorphisms in 2 candidate genes (*PEDF* and *EPO*) implicated in the neurodegeneration pathway and associated with DR susceptibility.

Whole genome and candidate gene association studies have identified many disease/risk predisposing loci for DR (Hampton et al., 2015). These studies however are unable to explain the complex etiology of DR. In this study, we have explored the role of gene - gene/gene-environment interaction in DR through high order interaction analysis with a rationale that candidate loci could influence the disease or phenotype outcome through its interaction with other such genetic/environmental factors (Cordell, 2009).

We have adopted two data mining approaches, the generalized multifactor dimensionality reduction (GMDR), classification and regression tree (CART) methods, recommended for gene/environment interaction studies. GMDR and CART represents non-parametric, genetic model-free analysis for detecting gene - gene/gene - environment interaction. GMDR allows adjustment for confounding parameters (discrete and quantitative) and develops a model for defining disease risk by pooling high-risk genotype combinations into one group and low-risk combinations into another (Lou et al., 2007). We further confirmed the results of GMDR and CART by logistic regression analysis.

2. Materials and methods

2.1. Study subjects and clinical methods

The study participants were enrolled between the years 2003 to 2010 from SNDREAMS (Sankara Nethralaya Diabetic Retinopathy Epidemiology and Molecular Biology Study), and epidemiology study on the prevalence of DR in south Indian population (Rani et al., 2007). The study protocol adhered to the Declaration of Helsinki and approved by Vision Research Foundation ethics committee (Ethical clearance no: 10-2003-P, 58-2007-P & 59-2007-P). A detailed history of all the patients was recorded followed by physical examination and pedigree analysis after their informed consent. Ocular examination included 45° fundus photograph using 4-field stereoscopic digital photography that were graded by 2 independent observers in a masked fashion with a higher grading agreement of $\kappa = 0.83$ (Rani et al., 2007). The diagnosis of DR was based on the modified Klein classification of the Early Treatment Diabetic Retinopathy Study scale. Type 2 diabetes patients were divided as (i) cases (DR +, with retinopathy; $N = 201$) with sight threatening DR (STDR) inclusive of PDR, non-proliferative DR (NPDR) with CSME and CSME alone with DM duration of ≥ 10 years (ii) controls (DR −, without retinopathy, $N = 168$) without any signs of DR and diabetes duration ≥ 15 years. Subjects with age related macular degeneration (AMD) and other hereditary retinal disease were excluded from the study.

2.2. SNP genotyping

DNA was extracted from peripheral blood samples by conventional phenol chloroform method (Wolff, 1997) and NucleoSpin Blood XL maxi kit method (Macherey-Nagel, Duren, Germany) according to the manufacturer's instructions. The genotype scoring for promoter SNP rs1617640 in *EPO* gene was done by polymerase chain reaction (PCR) based direct sequencing in ABI PRISM 3100 *AVANT* genetic analyzer (Applied Biosystems, Foster City California, USA). The promoter SNPs in *PEDF* gene were scored by PCR based restriction fragment length polymorphism (RFLP) method using *BsrI* (rs12150053) and *HincII* (rs12948385) enzymes (Iizuka et al., 2007; Tong et al., 2008).

2.3. Statistical analysis

Deviation from Hardy–Weinberg equilibrium (HWE) for genotypes was analyzed. Statistical analyses were performed using SPSS software (for Windows version 14.0; SPSS Science, Chicago, Illinois (IL), USA). The results were expressed as mean standard deviation for continuous and as percentage for categorical variables. The Student *t*-test was performed to compare continuous variables and proportions among groups. The distribution of the genotypes and alleles between cases and controls were analyzed by chi-square (Barber et al., 1998) tests. The specific effect of genotypes on various clinical factors was assessed by unconditional logistic regression analysis. Odds ratios (OR) and 95% confidence intervals (CI) were calculated and p value < 0.05 was considered significant.

2.4. Gene-gene and gene-environment interaction studies

Interaction analysis for SNP-SNP and SNP-other clinical covariates of DR were performed in the current study using high order, nonparametric, combinatorial approaches (GMDR and CART). GMDR analysis (version v0.9) (http://www.ssg.uab.edu/gmdr/) a score based algorithm, adopted from the MDR framework (Lou et al., 2007), predicts the best model of interaction within/between multiple loci based on the values of cross validation consistency (CVC), testing balanced accuracy (TBA) and prediction error (PE). The data split into 10 different sets, validated for the repeated occurrence of a particular set of loci under each set that was split from the original data. The maximum value of cross-validation consistency remains as 10 suggesting that the same combination of factors is being identified across all the subsets.

TBA value, measure of sample size, refers to a function of interaction when the number of cases and controls in a data set remains unequal and always ranges from 0.5 to 1.0. Prediction error implying maximum testing accuracy is being calculated based on the empirical p values calculated by employing permutation based on 10,000 shuffles. The model with minimum prediction error and/or maximum cross-