Meta Gene 16 (2018) 39–49

Contents lists available at ScienceDirect

Meta Gene

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

# A multifactor dimensionality reduction model of gene polymorphisms and an environmental interaction analysis in type 2 diabetes mellitus study among Punjabi, a North India population

Basanti Barna, Badaruddoza*, Manpreet Kaur, A.J.S. Bhanwer

*Department of Human Genetics, Guru Nanak Dev University, Amritsar, Punjab, India*

ARTICLE INFO

*Keywords:*
Type 2 diabetes mellitus
Multifactor dimensionality reduction (MDR) analysis
Punjab
North-West India
Gene-gene interactions
Gene-environment interactions

ABSTRACT

The effects of any single genetic variation for a common complex disease such as T2DM may be dependent on other genetic variations (gene-gene interactions) and environmental factors (gene-environment interactions). Multifactor Dimensionality Reduction (MDR) method helps in detection and characterization of susceptibility in common complex multifactor disorders like Type 2 Diabetes Mellitus (T2DM). The negligible studies are available with this model approach to detect the status of T2DM in north Indian population. Hence, the major objectives of the present study were to investigate the association of *ENPP1* K121Q (rs1044498), *TCF7L2* G > T (rs12255372) and *GYS1 Xba*I (A1 > A2) (rs8103451) gene variants with T2DM in the north Indian population; and to determine whether significant gene-gene and gene-environment (risk factors related to obesity and cardiovascular diseases) interactions exist between these selected genes in affecting type 2 diabetes mellitus using MDR analysis. A total of 500 participants consisting of 250 type 2 diabetes mellitus (T2DM) and 250 healthy subjects were recruited for this study. Genotyping was performed by PCR-RFLP method. Anthropometric and physiometric variables such as height, weight, waist circumference (WC), hip circumference (HC), SBP and DBP, were measured using standard protocol. The odds ratio and Hardy- Weinberg equilibrium deviation analyses were performed. The gene-gene and gene-environment interactions were performed by multifactor dimensionality reduction (MDR) analysis. In results it was observed that two genes *ENPP1* and *TCF7L2* are associated with T2DM. However, an insignificant association of the *Xba*I (A1 > A2) polymorphism in *GYS1* gene with T2DM was demonstrated. The gene-gene interactions revealed that all the three SNPs have a synergistic effect with each other. The MDR method for gene-environment interactions showed all interaction models first to ninth order interactions for T2DM patients as significant for susceptibility of obesity. The results showed that both the genes *ENPP1* and *TCF7L2* interacting with WHR and WC increase the susceptibility of obesity many folds among T2DM patients and non-diabetic controls. In conclusion, it is suggested that pathogenesis of T2DM, obesity and hypertension involves interplay of a variety of susceptibility alleles and environment. The gene-gene and gene-environment interactions are not only possible, but, are probably ubiquitous in determining the susceptibility of complex human diseases. Further studies on epistatic interactions are warranted to elucidate their possible underlying role in pathogenesis of T2DM.

## 1. Introduction

Diabetes mellitus is a worldwide epidemic, causing serious physical harm and economic burden. This disease has emerged as a major public health problem of 21st century and is the fifth leading cause of death globally (Sanghera et al., 2011; Sanghera and Blackett, 2012). The prevalence of diabetes and pre-diabetes in Punjab representing north India is 13.6% and 14.6%, respectively which indicates that there are 0.12 million diabetic and 0.13 million pre-diabetic individuals in Punjab (Anjana et al., 2011).

Despite numerous literature and evidences, the genetics of T2DM is still a puzzle and it became evident that the disease is much more

*Abbreviations:* $\chi^2$, chi-square; ADA, American Diabetes Association; BMI, body mass index; DNA, deoxyribonucleic acid; dNTP, deoxynucleotide triphophates; HC, hip circumference; HWE, Hardy-Weinberg equilibrium; LADA, latent autoimmune diabetes in adults; MODY, maturity-onset diabetes in the young; MDR, multifactor dimensionality reduction (MDR) analysis; PCR, polymerase chain reaction; RFLP, restriction fragment length polymorphism; SD, standard deviation; T2DM, type 2 diabetes mellitus; WC, waist circumference; WHR, waist to hip ratio

* Corresponding author.
*E-mail address:* doza13@yahoo.co.in (Badaruddoza).

https://doi.org/10.1016/j.mgene.2018.01.009
Received 25 July 2017; Received in revised form 25 January 2018; Accepted 27 January 2018

heterogeneous and has a strong genetic interaction between type 1 diabetes and type 2 diabetes with an affluent environment. Evidence has been accumulated that multiple genetic environmental factors may play important role in determining susceptibility to T2DM. Although the different variants of candidate genes have become prime targets of genetic analysis, few studies have considered their interactions (Neuman et al., 2010; Ramu et al., 2011; Zhou et al., 2012). The gene-gene interaction which is called epistasis might explain a large portion of genetic variability for such complex diseases. Most of the studies for genetic dissection of T2DM have focused on estimating effects of individual genes and excluded potential epistasis effects in their analytic models.

The effects of any single genetic variation for a common complex disease such as T2DM may be dependent on other genetic variations (gene-gene interactions) and environmental factors (gene-environment interactions). To address this issue, in recent years, a rapid progress has been made in the development of statistical methods for analyzing gene-gene and gene-environment interactions such as logistic regression models, stratified analysis, cross over analysis, general relative risk models, composite linkage and disequilibrium methods (Basu et al., 2011; Wu et al., 2011). However, most of these methods require a large sample size to model high order interactions (Moore and Williams, 2002; Moore et al., 2006). Hence, for moderate sample size data, one method of detecting and characterizing susceptibility in common complex multifactor disorders is the Multifactor Dimensionality Reduction (MDR) method (Ritchie et al., 2001). This method detects and characterizes a high order gene-gene and gene-environment interaction in case-control studies. Using this method, multi-locus genotype is classified into high-risk and low-risk and it effectively reduces genotype predictors from n dimensions to one dimension. The new one dimensional multi-locus genotype variable is evaluated for its ability to classify and predict the disease status through cross-validation. The MDR method is model free and it does not assume any particular genetic model.

Almost no study is available with this model approach to detect the status of T2DM in north Indian population. Hence, the present study has adopted this highly important statistical tool to assess both the main effects of single locus and multi-locus interactions to test the hypothesis that T2DM related genes may contribute to etiology of obesity and cardiovascular disease. The major objectives of the present study were to investigate the association of *ENPP1* K121Q (rs1044498), *TCF7L2* G > T (rs12255372) and *GYS1 Xba*I (A1 > A2) (rs8103451) gene variants with type 2 diabetes mellitus in the north Indian population; and to determine whether significant gene-gene and gene-environment (risk factors related to obesity and cardiovascular diseases) interactions exist between these selected genes in affecting type 2 diabetes mellitus using MDR analysis.

## 2. Materials and methodology

### 2.1. Subjects

The detailed protocol of the study was approved by the Ethical Committee of Guru Nanak Dev University. A total of 500 subjects (250 T2DM patients, 250 controls) were recruited for this study. T2DM subjects were recruited from different clinical centres in Amritsar district in Punjab, whereas unrelated healthy individuals were collected randomly. All participants provided written informed consent. Diabetes was diagnosed based on the American Diabetes Association's protocol (ADA, 2009).

### 2.2. Inclusion/exclusion criteria

Inclusion: During the sample collection, the individuals over 40 years old were recruited for the study and only one member from one family was taken. Individuals with diabetes in whom the diagnosis was made by a physician, who had fasting ($10-12h$) plasma glucose of 7 mmol/L or higher 126 mg/dL were included in the study. The control group of the subjects who has fasting plasma glucose of < 5.6 mmol/L (100 mg/dL) and 2 h plasma glucose of 7.8 mmol/L or less (140 mg/dL) were included in the study.

Exclusion: The subjects with following conditions were excluded for the study: pregnant or nursing mother, subjects with any illness/chronic disease interfering with patients' ability to comply with the diagnostic protocol, individuals exposed to any medication and possessing rare forms of diabetes such as maturity-onset diabetes in the young (MODY) and latent autoimmune diabetes in adults (LADA), patients having hypertension or undergoing immunosuppressive transplant therapy and those who were unwilling and/or not able to give written informed consent.

### 2.3. Anthropometric measurements

Actual age and age on the onset of the disease were recorded from the subject's health card provided by the clinical centres. Height, weight, circumferences of waist (WC) and hip (HC) were taken from each individual using standard anthropometric techniques and tools (Singh and Bhasin, 1968). Height and weight were measured to the nearest 0.5 cm and 0.1 kg, respectively. Body mass index (BMI) was calculated for an estimate of overall adiposity using the formula: BMI = weight (kg)/height ($m^2$). Waist and hip circumference (WC and HC) for an estimate of central obesity were measured to the nearest 0.5 cm with a steel tape. Waist to hip ratio was calculated using the standard formula: WHR = WC (cm)/HC (cm). Two subsequent measurements were taken and averages were used in the analysis.

### 2.4. Physiometric measurements

Left arm blood pressures (first-phase systolic and fifth-phase diastolic) were taken from each participant with standard mercury sphygmomanometer after a 5-min rest. The average of the two subsequent measurements was used for analysis. All efforts were made to minimize the factors which affect the blood pressure like anxiety, fear, stress, laughing and recent activity (Badaruddoza and Afzal, 1999). The radial artery at the wrist was used to count the pulse. It was counted over 1 min. The difference of SBP and DBP was used as pulse pressure.

### 2.5. Molecular analysis and reason for selection of the polymorphisms

The genes involved in the pathways controlling insulin secretion and regulating pancreatic β-cell mass in the progression of diabetes mellitus have been studied using different advanced techniques. The current approaches to root out the role of some potential genetic variants present in the pathway of glucose metabolism have reached at the genome-wide association study (GWAS) level. However, there are many approaches through which the variants of many genes can be validated for their role in causing the disease. In the present investigation case-control design was followed for analysis of single nucleotide polymorphisms (SNPs) in the *ENPP1*, *GYS1* and *TCF7L* genes to assess the association of these markers with T2DM. The SNPs were analyzed and were selected after critical perusal and evaluation of the literature and reports presented in other population.

The three SNPs selected were: (i) *ENPP1* K121Q (rs1044498) polymorphism, Lysine-K to Glutamine-Q substitution of A > C base in exon 4. It is located on chromosome 6q22–23; (ii) *TCF7L2* G > T (rs12255372) which involves G > T change in intron 4. It is located on chromosome 10q25.3; (iii) *GYS1 Xba*I (A1 > A2) (rs8103451) polymorphism, change of a single base (C > T) in intron 14. It is located on chromosome 19q13.3.

Several human studies have been performed to find out the association between these polymorphisms and T2DM in different ethnic populations. In many studies, the investigators found strong association