Molecular Genetics and Metabolism 99 (2010) 80–89

Contents lists available at ScienceDirect

Molecular Genetics and Metabolism

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

Influence of hepatocyte nuclear factor 4α (*HNF4α*) gene variants on the risk of type 2 diabetes: A meta-analysis in 49,577 individuals

Silvia Sookoian, Carolina Gemma, Carlos J. Pirola *

Molecular Genetics and Biology of Complex Diseases Department, Institute of Medical Research A. Lanari, University of Buenos Aires – National Council of Scientific and Technological Research (CONICET), Combatientes de Malvinas 3150, Buenos Aires (1427), Argentina

ARTICLE INFO

Article history:
Received 28 June 2009
Received in revised form 14 August 2009
Accepted 14 August 2009
Available online 21 August 2009

Keywords:
Hepatocyte nuclear factor
HNF4α
Liver
Meta-analysis
Systematic review
Polymorphisms
Type 2 diabetes

ABSTRACT

Background: The nuclear receptor hepatocyte nuclear factor 4α (*HNF4α*) contributes to the regulation of a large fraction of liver and pancreatic islet transcriptomes.
Aim: To evaluate the influence of *HNF4α* polymorphisms across the entire locus on the occurrence of type 2 diabetes (T2D) by means of a meta-analysis.
Methods: We evaluated haplotype block structure of *HNF4α* variants owing to linkage disequilibrium (LD). From 1455 reports, we evaluated 21 observational studies.
Results: Six haplotype blocks of LD were constructed with SNPs with $r^2 > 0.8$; there were also 14 unlinked SNPs. Overall, we included 22,920 cases and 26.657 controls. Among 17 heterogeneous studies (21,881 cases and 24,915 controls), including 3 SNPs of P2 promoter region in block 1, we observed a significant association with T2D in fixed (OR 0.94, 95%CI: 0.905–0.975, $p = 0.001$) and random (OR 0.988, 95%CI: 0.880–0.948, $p = 0.000012$) model. Three homogeneous studies were evaluated in block 2 (2684 cases and 2059 controls), and a significant association with T2D was also observed: OR: 1.121, 95%CI 1.013–1.241, $p = 0.027$. Three additional variants were associated with T2D: two intronic SNPs (rs4810424: OR: 1.080, 95%CI: 1.010–1.154, $p < 0.03$ and rs3212183: OR: 0.843, 95%CI: 0.774–0.918, $p < 0.00009$) and one missense variant (rs1800961: OR: 0.770, 95%CI: 0.595–0.995, $p < 0.05$, 6562 cases and 6723 controls).
Conclusions: In addition to *HNF4α* variants in the promoter region, other SNPs may be involved on the occurrence of T2D.

Introduction

Type 2 diabetes (T2D) is a common, multifactorial disease, with many known and unknown genetic and environmental factors influencing risk.

The discovery of novel T2D genes using whole-genome association studies (GWAS) has provided insight into the genetic architecture of T2D and approximately 20 common variants are now robustly implicated in T2D susceptibility [1].

Additional variants have emerged from meta-analysis of GWAS [2] or combined analysis of biological plausibility with publicly available data from GWAS [3].

Some other candidate genes related to the susceptibility to T2D were explored in the past on the basis of the pathophysiological mechanisms involved in the development of the disease. The nuclear receptor hepatocyte nuclear factor alpha (*HNF4α)* is an example of that. In fact, the protein encoded by *HNF4α* coordinates the expression of several genes required for glucose transport and gluconeogenesis [4,5]. Moreover, *HNF4α* also plays a role in regulating the secretion of insulin by direct activation of the insulin gene promoter [6].

HNF4α, located in chromosome 20: 42,417,855–42,493,444, potentially encodes nine distinct isoforms (*HNF4α1–HNF4α9)* that result from both alternate promoter usage and alternative splicing. Isoforms *HNF4α1–HNF4α6* are coded from the P1 (hepatic) promoter, isoforms *HNF4α7–HNF4α9* are transcribed from the P2 (pancreatic) promoter [7].

Based on both the previous knowledge about mutations of this gene causing the maturity-onset diabetes of the young (MODY) type 1 and results from linkage studies in T2D that showed suggestive linkage peaks in the region of *HNF4α* [8,9], it was speculated that *HNF4α* gene variants may be associated with T2D [10–19]. Most of the gene variants associated with increased T2D risk were identified upstream of the *HNF4α* coding region, in the alternative β-cell promoter P2, which is 46 kb upstream to the P1 promoter of the human gene [20]. However, the evidence for the association of

* Corresponding author. Address: Instituto de Investigaciones Médicas A. Lanari, Combatientes de Malvinas 3150, Buenos Aires (1427), Argentina. Fax: +54 11 4523 8947.
E-mail addresses: cpirola@ciudad.com.ar, pirola.carlos@lanari.fmed.uba.ar (C.J. Pirola).

1096-7192/$ - see front matter © 2009 Elsevier Inc. All rights reserved.
doi:10.1016/j.ymgme.2009.08.004

polymorphisms in or near P2 region with the risk of T2D was not replicated across all the published studies.

The *HNF4α* gene spans ~76 kb with 10 exons on chromosome 20q13.1–13.2, and in addition to the regulatory region, some other gene variants were also surveyed supporting a role in the T2D predisposition [17,21]. However, a study that evaluated a high density of SNPs spanning the *HNF4α* region failed to replicate any sign beyond the P2 region [22].

Therefore, the significance of the *HNF4α* variants, not only in the P2 promoter haplotype but also along the entire gene locus, in the risk of type 2 diabetes varies in the literature showing that in some populations the *HNF4α* variants do have an effect on disease risk while in other do not. On the hypothesis that any of the *HNF4α* variants could be a putative causal variant or proxy of some others, and to address discrepancies in *HNF4α* genetic association studies, we decided to evaluate the influence of the gene polymorphisms across the entire locus on the occurrence of type 2 diabetes by means of a meta-analysis of all available publications by combining SNPs according to haplotype blocks using the HapMap database (http://www.hapmap.org) whenever possible.

Materials and methods

Data sources and study selection

For the electronic searches, published studies were found through Pubmed at the National Library of Medicine (http://ncbi.nlm.nih.gov/entrez/query) and in Medline databases for the query "(*HNF4α* OR hepatocyte nuclear factor alpha) AND (gene OR variants OR polymorphism OR alleles) AND (type 2 diabetes OR T2D OR non-insulin-dependent diabetes mellitus)". Reference lists in relevant publications were also examined. The literature search was limited to human and was done on studies up to January 2009 and availability of an English-language abstract or paper for review. There were not country restrictions. We first conducted a literature-based systematic review of all relevant studies (irrespective of the SNP typed).

We evaluated 1455 citations identifying 21 studies that met the selection criteria: population-based or hospital-based case-control, cross-sectional studies concerning the relationship between *HNF4α* variants and T2D, in which information about number of subjects in each category, sufficient data to calculate outcomes, and genotyping performed with a validated molecular method could be extracted.

In the case of cohorts, we included variables before any intervention. Data from one further study that fulfilled the eligibility criteria were excluded from the study because data about genotype distribution in cases and controls is disclosed in such a way that precluded further analysis (for instance, calculation of multipoint identities by descent) [16].

An evaluation of study quality of the reviewed articles using the median of the impact factor of journals, in which they had been published, was included [23].

Data collection

All odd ratios (OR) were calculated against healthy control subjects. For each study, information was collected concerning the following characteristics of the subjects: demographic information (age, sex, and ethnicity) and T2D defined as currently taking medication for diabetes or medical record information conforming to World Health Organization (WHO) criteria (World Health Organization: *Diabetes Mellitus: Report of a WHO Study Group* Geneva World Health Org., 1985 (Tech. Rep. Ser., No. 727) or by the oral glucose tolerance test (OGTT) using criteria of the American Diabetes Association (2-h glucose >11.1 mmol/l or fasting glucose >7 mmol/l), as described in each paper. Control subjects were defined either as those who had normal glucose tolerance by (OGTT) or who were normoglycemic.

Statistical analysis

Summary effects, odds ratio (OR) and corresponding 95%CI were estimated by both fixed and random effects meta-analysis using the Mantel–Haenszel method. Heterogeneity was evaluated with *Q* statistic and the I2 statistic, a transformation of *Q* that estimates the percentage of the variation in effect sizes caused by heterogeneity.

Regarding heterogeneity, we identified study characteristics that stratify the studies into subsets with homogeneous effects. We considered possible sources of heterogeneity and stratified the studies by ethnicity, age and gender and repeated the analysis separately for each group. It is worth to mention that because of the conflictive evidence among studies concerning Scandinavian vs Caucasian non-Scandinavian we decided to group these studies in two separated categories if possible. If heterogeneity continued, we ranked the studies according to their individual χ^2, removed the studies with the higher χ^2, and repeated the process until homogeneity was achieved. If the association became homogeneous after stratification or after removing the outliers studies, we recalculated the overall effect and 95%CI, and no further action was taken. Sensitivity of the findings was examined by recalculating the pooled association sizes and joint values of *p* in homogeneous subgroups as well as after excluding one by one study at a time.

All calculations were performed using the Comprehensive Meta-Analysis computer program (Biostat, Englewood, NJ, USA). To check for publication bias, we used a visual inspection of funnel plots, the Begg and Mazumdar's rank correlation test [24] and the Egger's regression intercept method [25], but we only show results from the later, as it is the most powerful approach for detecting publication bias.

A *p* value lower than 0.05 was considered statistically significant.

According to the procedure described by Han et al. [26], an *a priori* global estimation of the power for detecting association with the list of analyzed SNPs as possible causal variants was performed assuming an OR of 1.3 for a prevalence of the disease of 10% and on average 3000 cases and 3000 controls, numbers well below the ones including in the meta-analysis. This method gave us an estimated global power greater than 90%.

HNF4α linkage disequilibrium (LD) and haplotype block structure

We evaluated the patterns of LD and haplotype block structure of the *HNF4α* gene variants included in all the published studies evaluated in this meta-analysis. Haplotype block structure was inferred by Haploview available at http://www.broad.mit.edu/mpg/haploview/; in this program, the extent of LD was measured in terms of D', and r^2. To be conservative we constructed blocks of LD with SNPs with $r^2 > 0.8$. SNPs in these blocks are interchangeable and then they can be combined in the meta-analysis irrespective of which one has been genotyped in a particular study [15].

The LD plot, shown in Fig. 1, illustrates six haplotype blocks (B1–B6) in Caucasian population according data from the HapMap project (www.hapmap.org). As available in the HapMap, similar analysis was performed for Japanese subjects (data not shown).

Besides the six LD blocks, we included single SNP that are not in high LD with neighboring variants. Then, we evaluated all the SNPs included in more than one study either belonging to any of the above-mentioned LD blocks or being singleton blocks.