European Journal of Human Genetics (2008) 16, 652–656

www.nature.com/ejhg

ARTICLE

HHEX gene polymorphisms are associated with type 2 diabetes in the Dutch Breda cohort

Jana V van Vliet-Ostaptchouk[1,2], N Charlotte Onland-Moret[3,4], Timon W van Haeften[5], Lude Franke[3], Clara C Elbers[3,4], Ronit Shiri-Sverdlov[2], Yvonne T van der Schouw[4], Marten H Hofker*[,1,2] and Cisca Wijmenga[3,6]

[1]*Department of Pathology and Laboratory Medicine, University Medical Center Groningen, University of Groningen, Groningen, The Netherlands;* [2]*Department of Molecular Genetics, Maastricht University, Maastricht, The Netherlands;* [3]*The Complex Genetics Section, Department of Biomedical Genetics, University Medical Center Utrecht, Utrecht, The Netherlands;* [4]*Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht, Utrecht, The Netherlands;* [5]*Department of Internal Medicine, University Medical Center Utrecht, Utrecht, The Netherlands;* [6]*Department of Genetics, University Medical Center Groningen and University of Groningen, Groningen, The Netherlands*

Recently, the hematopoietically expressed homeobox (*HHEX*) gene, encoding a transcription factor, was identified in a large genome-wide scan in French individuals as a type 2 diabetes (T2D)-susceptibility locus. We aimed to check whether this finding could be replicated in a Dutch T2D cohort. Two common variants (rs7923837 and rs1111875) located near the *HHEX* gene were genotyped in 501 unrelated T2D patients and in 920 healthy controls. The major alleles of both variants were overrepresented in T2D cases compared with controls (66.7 *vs* 64.1%, $P=0.16$ for rs7923837 and 64.6 *vs* 60.4%, $P=0.027$ for rs1111875). For both polymorphisms, the risk for T2D was significantly increased in carriers of the major alleles (rs7923837: odds ratio (OR): 1.57, 95% confidence interval (CI): 1.08–2.27, $P=0.017$ and rs1111875: OR: 1.68, 95% CI: 1.19–2.35, $P=0.003$). The haplotype analysis did not reveal a risk haplotype that provided stronger evidence for association with T2D than each variant individually. Assuming a dominant genetic model, the population-attributable risks for diabetes due to the at-risk alleles of rs7923837 and rs1111875 were estimated to be 33 and 36%, respectively. These data provide evidence that variants near the *HHEX* gene contribute to the risk of T2D in a Dutch population.
European Journal of Human Genetics (2008) **16,** 652–656; doi:10.1038/sj.ejhg.5202008; published online 30 January 2008

Keywords: association; Breda cohort; HHEX; SNP; susceptibility; type 2 diabetes

Introduction

Until recently, the progress in searching for genetic variants that predispose to type 2 diabetes (T2D) was rather slow. Only the P12A variant of the *PPARG* gene and the E23K variant of the *KCNJ11* gene had been associated, and confirmed, with this disease in various large-scale studies performed in different populations.[1] Recent identification of the *TCF7L2* (transcription factor 7-like 2) gene polymorphisms as major determinants of T2D risk[2] became a long-awaited breakthrough in finding true T2D-susceptibility genes. Indeed, the original observation of Grant *et al*[2] has been successfully replicated in more than 20 studies to date. *TCF7L2* has therefore been proposed as 'the biggest story in diabetes genetics since HLA.'[3]

*Correspondence: Professor MH Hofker, Department of Pathology and Laboratory Medicine, Center for Integrative Genomics, University Medical Center Groningen, Antonius Deusinglaan 1, Building 3215, 9713 AV Groningen, The Netherlands.
Tel: +31 50 363 2676; Fax: +31 50 363 2676;
E-mail: m.h.hofker@med.umcg.nl
Received 7 June 2007; revised 27 November 2007; accepted 20 December 2007; published online 30 January 2008

Recently, Sladek *et al*[4] reported the results of a large genome-wide association (GWA) study that identified three novel loci that contribute substantially to susceptibility for T2D, and they also confirmed the association of *TCF7L2* with T2D. One of the new loci located on chromosome 10q contains genes known to have roles in the development (the hematopoietically expressed homeobox, *HHEX*)) and function (insulin-degrading enzyme (*IDE*)) of the pancreas.[5,6] Interestingly, the *HHEX* gene also encodes a transcription factor that is involved in Wnt signaling, a fundamental pathway for cell growth and development,[7,8] as does *TCF7L2*. As replication is always a critical issue in confirming the validity of genetic association studies, we have examined whether the variants near the *HHEX* gene contribute to the risk of T2D in a Dutch population.

Subjects and methods

Breda study

DNA was available from 501 T2D patients from the Breda study (Table 1).[9] All patients were diagnosed according to the WHO criteria (random plasma glucose level >11.1 mmol/l or a fasting plasma glucose level >7.0 mmol/l). The clinical characteristics of the patients (HbA_{1c}, plasma cholesterol, HDL-cholesterol and triglycerides) were available, as well as the level of obesity in each individual, as represented by the BMI (defined as weight in kg divided by height in m squared). The control cohort of 920 subjects comprised of healthy blood bank donors of Dutch Caucasian origin.[10] All participants gave their written informed consent, and the Breda study was approved by the Medical Ethics Committee of the University Medical Center Utrecht.

Genotyping

Single-nucleotide polymorphisms (SNPs) rs7923837 and rs1111875 were genotyped using TaqMan assay-on-demand (C_31982553_10 and C_11214581_10 assays; Applied Biosystems, Nieuwerkerk a/d IJssel, The Netherlands).

Table 1 Clinical characteristics of the Dutch T2D case and control samples

Trait	*Breda study group*	*Control group*
N (female/male)	501 (270/230)[a]	920 (354/557)[b]
Age at study (years)	70.7 ± 9.9	47.8 ± 12.7
Age at diagnosis (years)	57.6 ± 14.4	—
BMI (kg/m^2)	27.6 ± 4.9	NA
HbA_{1c} (%)	6.6 ± 2.4	NA
HDL-cholesterol (mmol/l)	1.1 ± 0.5	NA
Total cholesterol (mmol/l)	4.9 ± 1.8	NA
Triacylglycerol (mmol/l)	1.7 ± 1.1	NA

BMI, body mass index; HbA_{1c}, hemoglobin A_{1c} (glucose bound to hemoglobin); HDL, high-density lipoprotein; NA, not available.
The data are presented as mean ± SD.
[a]Not available for one subject.
[b]Not available for nine subjects.

Assays were performed according to the manufacturer's specifications and the genotypes were analyzed using a TaqMan 7900HT (Applied Biosystems). The DNA samples were processed in 384-well plates. Each plate contained 8 negative controls and 16 genotyping controls (4 duplicates of 4 different samples obtained from the Centre d'Etude du Polymorphisme Humain (CEPH)). The genotype success rates were 97.5% for rs7923837 and 98.4% for rs1111875. There were no discordances in the genotypes of any of the CEPH samples. The controls were in the Hardy–Weinberg equilibrium ($\chi^2=0.05$, $P=0.83$ for rs7923837 and $\chi^2=0.53$, $P=0.46$ for rs1111875).

Statistical analysis

The genotype frequencies were tested for the Hardy–Weinberg equilibrium by χ^2 analysis. To test for association of genotypes and T2D, genotype-based odds ratios (ORs) with 95% confidence intervals (CIs) were calculated using a logistic regression model. For both SNPs, the allele that showed an increased frequency in the T2D patients was taken as the risk allele and the other allele as the reference. Differences in haplotype distribution in cases and controls were tested for significance using a two-sided χ^2 test. Haplotypes of the two SNPs were estimated using the COCAPHASE package of the UNPHASED program.[11] The D' and r^2 between the *HHEX* SNPs were calculated with the same package. Power calculation was completed using Quanto software[12] (http://hydra.usc.edu/gxe/). The population-attributable risk was calculated for both SNPs using the corresponding allele frequencies for this marker as explained by Greenland and Rothman.[13] The 'Prioritizer' method[14] was used to investigate the T2D-susceptibility loci reported by Sladek *et al*.[4] All statistical analyses were performed using the SPSS program, version 13.0 for Windows (SPSS Inc., Chicago, IL, USA).

Results

In the *HHEX* gene, we genotyped two SNPs, rs7923837 and rs1111875, that were reported by Sladek *et al*.[4] to be strongly associated with T2D. We found significant association between T2D and the two SNP variants in our case–control study, with the same risk allele as reported in the original study (Table 2). The major alleles of both polymorphisms were overrepresented among T2D patients compared with controls (66.7 *vs* 64.1%, $P=0.16$ for rs7923837 and 64.6 *vs* 60.4%, $P=0.027$ for rs1111875). For both rs7923837 and rs1111875, both heterozygous and homozygous carriers for the major allele were associated with an increased risk for T2D compared with non-carriers (OR: 1.60, 95% CI: 1.09–2.35, $P=0.016$ and OR: 1.53, 95% CI: 1.04–2.25, $P=0.033$, respectively, for rs7923837, and OR: 1.68, 95% CI: 1.18–2.39, $P=0.004$ and OR: 1.68, 95% CI: 1.16–2.41, $P=0.005$, respectively, for rs1111875). For both rs7923837 and rs1111875, we see an equal risk for

Table 2 Association study of the rs7923837 and rs1111875 SNPs with T2D

Genotype	*T2D patients (n = 501)*	*Control subjects (n = 920)*	*OR*	*95% CI*	P-*value*
Rs7923837					
AA (%)	43 (8.8)	117 (13.1)	—	—	—
GA (%)	241 (49.1)	409 (45.7)	1.60	1.09–2.35	0.016
GG (%)	207 (42.2)	369 (41.2)	1.53	1.04–2.25	0.033
A allele (%)	327 (33.3)	643 (35.9)	1.13	0.95–1.33	0.160
G allele (%)	655 (66.7)	1147 (64.1)			
Dominant model					
AA	43	117	—	—	—
GG+GA	448	778	1.57	1.08–2.27	0.017
Rs1111875					
AA (%)	51 (10.4)	148 (16.3)	—	—	—
GA (%)	245 (50.0)	424 (46.7)	1.68	1.18–2.39	0.004
GG (%)	194 (39.6)	336 (37.0)	1.68	1.16–2.41	0.005
A allele (%)	347 (35.4)	720 (39.7)	1.20	1.02–1.41	0.027
G allele (%)	633 (64.6)	1096 (60.3)			
Dominant model					
AA	51	148	—	—	—
GG+GA	439	760	1.68	1.19–2.35	0.003

T2D, type 2 diabetes.

T2D from one or two copies of the risk alleles in the *HHEX* gene, implying a dominant effect.

In our control population, the two SNPs were found to be in strong linkage disequilibrium (LD) with each other ($D' = 0.93$, $r^2 = 0.74$). The haplotype analysis did not reveal a risk haplotype that provided stronger evidence for association with T2D than each SNP individually (data not shown).

Discussion

Our data confirm the association of the common variants of rs7923837 and rs1111875 near the *HHEX* gene, identified by Sladek *et al*,[4] with an increased risk of T2D in a Dutch population. The risk allele frequencies for both SNPs were similar to those reported in the original study. Both SNPs were significantly associated with T2D risk under a dominant genetic model. The population-attributable risks due to the at-risk alleles of rs7923837 and rs1111875 were estimated to be 33 and 36%, respectively.

Our study had 60% power to detect the ORs reported by Sladek *et al*[4] for rs11117875 with a significance level of 0.05 and assuming an allele frequency of the risk allele of 0.60 and a log-additive model, as was reported by Sladek *et al*.[4] Similarly, for rs7923837, we had 71% power to detect the reported ORs by Sladek *et al*.[4] It needs to be noted that in the current study we used control subjects randomly selected from blood bank donors. Thus, we cannot exclude the possibility that some of these individuals may have diabetes or will develop this condition in later life. However, this would result in an underestimation of the 'true' effect size of *HHEX* on susceptibility to T2D in our study.

Although the *HHEX* gene is an attractive diabetes gene from the locus on chromosome 10q, we cannot exclude the possibility that one of the other genes within the locus – *IDE* of *KIF11* – is the causal one because of the extended LD. We reasoned that it is likely that the causal genes would be functionally related in some way and we used the Prioritizer program[14] to investigate how genes within these loci interact with each other and also with the known T2D genes *PPARG* and *TCF7L2*. In short, Prioritizer is a bioinformatics tool that is specifically designed to prioritize candidate genes from associated regions in the genome based on the function of the genes. Prioritizer is based on the assumption that disease genes in a specific disorder are usually functionally related, and it uses various resources, such as Reactome, Gene Ontology (GO), KEGG, etc, to identify interacting genes (a detailed description of the methods used by this tool is described by Franke *et al*[14]). Results from Prioritizer indicate that *HHEX* is predicted to interact with *PPARG* and *TCF7L2*, whereas *KIF11* and *IDE* cannot be easily related to any of these genes, which suggests that *HHEX* is the most likely causal candidate (Figure 1). For each interaction, evidence for the interaction can be looked up at the accompanying website, http://www.genenetwork.nl, along with the sources where this evidence came from. Inspection of these interactions reveals that the predicted interactions between *PPARG*, *TCF7L2* and *HHEX* are based on the fact that GO terms are shared between these genes. As this program uses existing databases and does not generate experimental evidence, the results should be interpreted with caution. With this

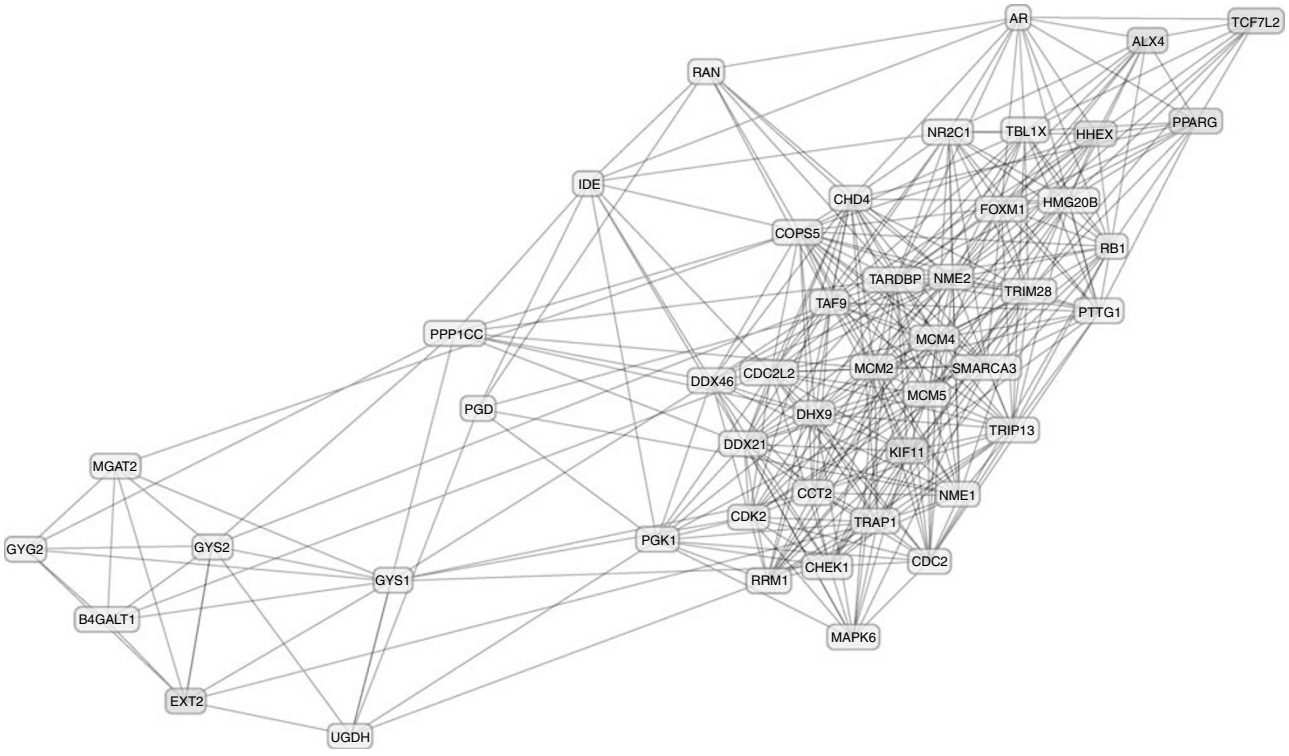

Figure 1 Prioritizer analysis of the T2D-susceptibility loci identified by Sladek *et al*.[4] Two loci (one on chromosome 10q and one on chromosome 11q) reported by Sladek *et al*[4] to be significantly associated with T2D contained multiple genes. Owing to LD, we could not unequivocally determine which genes (*IDE*, *KIF11* and *HHEX* from 10q and *EXT2* and *ALX4* from 11q) were the causal ones. We reasoned that it is likely that the causal genes would be functionally related in some way and we used the Prioritizer[14] to investigate how genes within these loci interact with each other and also with the known T2D genes *PPARG* and *TCF7L2*. The results indicated that *HHEX* and *ALX4* are predicted to interact with each other and with *PPARG* and *TCF7L2*, whereas the other positional candidate genes, *IDE*, *KIF11* and *EXT2*, cannot be easily related to any of these genes. This suggests that *HHEX* and *ALX4* are the most likely causal candidate genes from a biological perspective.

analysis, we merely wish to emphasize that among the possible candidate genes, given the association of the two SNPs and the LD patterns in this genomic region, *HHEX* is the most likely candidate.

We have previously reported the association of a *TCF7L2* variant with T2D in the same study sample.[15] *TCF7L2* has recently been established as a major determinant of diabetes risk.[3] Involvement in the Wnt signaling pathway affecting β-cell development and/or function has been proposed as the most likely mechanism for the role of *TCF7L2* in the pathogenesis of diabetes.[7] The *HHEX* gene encodes a transcription factor that is also involved in the Wnt signaling pathway[8] and that has been shown to be essential for pancreas development.[5] It might well be interesting to study a possible epistatic interaction between these two genes. As interaction studies require large study populations, it is not possible to answer this question with the Breda cohort. Functional studies and studies in larger cohorts will be required to investigate the possible epistasis between *HHEX* and *TCF7L2*, and to understand the role both genes may play in the pathogenesis of T2D.

In conclusion, we have confirmed that variants of the *HHEX* gene contribute to the risk of T2D in a Dutch population. The association of the rs7923837 and rs1111875 variants at *HHEX* with T2D risk has recently been confirmed in three large GWA studies.[16–19] These replications of the association reported by Sladek *et al*[3] support the view that such studies can be successful in identifying novel pathways and susceptibility genes for T2D.

Acknowledgements

We thank the Dutch Diabetes Foundation, the European Vascular Genomics Network and SenterNovem (IOP Genomics IGE05012) for financial support. We would like to thank Jackie Senior for critically reading the paper.

Duality of interest

The authors confirm that there is no duality of interest.

References

1 O'Rahilly S, Barroso I, Wareham NJ: Genetic factors in type 2 diabetes: the end of the beginning? *Science* 2005; **307**: 370–373.

2 Grant SF, Thorleifsson G, Reynisdottir I *et al*: Variant of transcription factor 7-like 2 (TCF7L2) gene confers risk of type 2 diabetes. *Nat Genet* 2006; **38**: 320–323.
3 Zeggini E, McCarthy MI: TCF7L2: the biggest story in diabetes genetics since HLA? *Diabetologia* 2007; **50**: 1–4.
4 Sladek R, Rocheleau G, Rung J *et al*: A genome-wide association study identifies novel risk loci for type 2 diabetes. *Nature* 2007; **445**: 881–885.
5 Bort R, Martinez-Barbera JP, Beddington RS, Zaret KS: Hex homeobox gene-dependent tissue positioning is required for organogenesis of the ventral pancreas. *Development* 2004; **131**: 797–806.
6 Farris W, Mansourian S, Chang Y *et al*: Insulin-degrading enzyme regulates the levels of insulin, amyloid beta-protein, and the beta-amyloid precursor protein intracellular domain *in vivo*. *Proc Natl Acad Sci USA* 2003; **100**: 4162–4167.
7 Smith U: TCF7L2 and type 2 diabetes – we WNT to know. *Diabetologia* 2007; **50**: 5–7.
8 Foley AC, Mercola M: Heart induction by Wnt antagonists depends on the homeodomain transcription factor Hex. *Genes Dev* 2005; **19**: 387–396.
9 van Tilburg JH, Sandkuijl LA, Strengman E *et al*: A genome-wide scan in type 2 diabetes mellitus provides independent replication of a susceptibility locus on 18p11 and suggests the existence of novel loci on 2q12 and 19q13. *J Clin Endocrinol Metab* 2003; **88**: 2223–2230.
10 Monsuur AJ, de Bakker PI, Alizadeh BZ *et al*: Myosin IXB variant increases the risk of celiac disease and points toward a primary intestinal barrier defect. *Nat Genet* 2005; **37**: 1341–1344.
11 Dudbridge F: Pedigree disequilibrium tests for multilocus haplotypes. *Genet Epidemiol* 2003; **25**: 115–121.
12 Gauderman WJ: Sample size requirements for matched case–control studies of gene–environment interaction. *Stat Med* 2002; **21**: 35–50.
13 Greenland S, Rothman RK: Measures of effect and measures of association; in Greenland S, Rothman RK (eds): *Modern Epidemiology*. Philadelphia: Lippincott-Raven, 1998, 2nd edn, pp 47–67.
14 Franke L, Bakel H, Fokkens L, de Jong ED, Egmont-Petersen M, Wijmenga C: Reconstruction of a functional human gene network, with an application for prioritizing positional candidate genes. *Am J Hum Genet* 2006; **78**: 1011–1025.
15 van Vliet-Ostaptchouk JV, Shiri-Sverdlov R, Zhernakova A *et al*: Association of variants of transcription factor 7-like 2 (TCF7L2) with susceptibility to type 2 diabetes in the Dutch Breda cohort. *Diabetologia* 2007; **50**: 59–62.
16 Steinthorsdottir V, Thorleifsson G, Reynisdottir I *et al*: A variant in CDKAL1 influences insulin response and risk of type 2 diabetes. *Nat Genet* 2007; **39**: 770–775.
17 Saxena R, Voight BF, Lyssenko V *et al*: Genome-wide association analysis identifies loci for type 2 diabetes and triglyceride levels. *Science* 2007; **316**: 1331–1336.
18 Scott LJ, Mohlke KL, Bonnycastle LL *et al*: A genome-wide association study of type 2 diabetes in Finns detects multiple susceptibility variants. *Science* 2007; **316**: 1341–1345.
19 Zeggini E, Weedon MN, Lindgren CM *et al*: Replication of genome-wide association signals in UK samples reveals risk loci for type 2 diabetes. *Science* 2007; **316**: 1336–1341.