

# Sequence variation in *IGF1R* is associated with differences in insulin levels in nondiabetic Old Order Amish

Adam C. Naj<sup>1,2</sup>  
Wen-Hong L. Kao<sup>1,2</sup>  
Jeffrey R. O'Connell<sup>3</sup>  
Braxton D. Mitchell<sup>3</sup>  
Kristi D. Silver<sup>3\*</sup>

<sup>1</sup>Department of Epidemiology, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD, USA

<sup>2</sup>Welch Center for Epidemiology, Prevention, and Clinical Research, Baltimore, MD, USA

<sup>3</sup>Division of Endocrinology, Diabetes, and Nutrition, Department of Medicine, University of Maryland School of Medicine, Baltimore, MD, USA

\*Correspondence to: Kristi D. Silver, Division of Endocrinology, Diabetes, and Nutrition, Department of Medicine, University of Maryland School of Medicine, 660 West Redwood Street, Room 498, Baltimore, MD 21201, USA. E-mail: ksilver@medicine.umaryland.edu

## Abstract

**Background** Insulin growth factor-1 receptor (*IGF1R*) encodes the insulin-like growth factor 1 receptor, a transmembrane tyrosine kinase receptor located on chromosome 15q26.3, in a region of linkage (LOD = 2.53,  $P = 0.00032$ ) to Insulin30 on an OGTT in the Old Order Amish. Mouse models with beta-cell-specific deficiency of *IGF1R* demonstrate defects in glucose-stimulated insulin secretion.

**Methods** To test the hypothesis that genetic variation in *IGF1R* is associated with impaired insulin secretion, we genotyped 54 SNPs in 778 nondiabetic subjects from the AFDS who had undergone OGTTs and tested them for association with ln Insulin30 and ISI.

**Results** No individual SNPs were significantly associated with ln Insulin30 or ISI using a multiple hypothesis testing adjusted  $P < 0.002$ . Tests of association of 4-SNP haplotypes constructed by a windowing approach revealed an association of the CTTG-variant of a 4-SNP haplotype found in intron 20 (rs1784195–rs2715439–rs8034284–rs12440962) with lower ISI levels ( $\beta = 0.18$ ,  $SE(\beta) = 0.05$ ,  $P = 0.001$ ).

**Conclusions** Sequence variation in *IGF1R* may influence insulin secretory function, although further studies in other populations will be needed to confirm these findings. Copyright © 2009 John Wiley & Sons, Ltd.

**Keywords** type 2 diabetes mellitus; insulin secretion; candidate gene; Old Order Amish; *IGF1R*

**Abbreviations** AFDS – Amish Family Diabetes Study; BMI – body mass index; Glucose0 – glucose level at 0 min on an oral glucose tolerance test; HWE – Hardy–Weinberg equilibrium; IGT – impaired glucose tolerance; Insulin30 – insulin level at 30 min on an oral glucose tolerance test; Insulin0 – insulin level at 0 min on an oral glucose tolerance test; *IGF1R* – insulin growth factor-1 receptor; ISI – insulin secretion index (ln Insulin30 – ln Insulin0); LD – linkage disequilibrium; LOD – likelihood of odds; MAF – minor allele frequency; OGTT – oral glucose tolerance test; SNP – single nucleotide polymorphism; T2DM – type 2 diabetes mellitus.

## Introduction

Abnormal insulin secretion is one of the features that is critical to the development of T2DM [1]. While influenced by multiple environmental factors [2],

Received: 5 March 2009

Revised: 16 August 2009

Accepted: 13 September 2009

insulin secretion is highly heritable ( $h^2 = 0.50\text{--}0.58$ ). Recent genome-wide association studies have identified several diabetes susceptibility loci. Accumulating data suggest that these loci influence diabetes risk through their effects on beta cell function [3–5]. In a genome-wide linkage analysis of insulin response to an OGTT in the AFDS, suggestive linkage was observed on chromosome 15q13-26 to Insulin30 (LOD = 2.53,  $P = 0.00032$ ) [6], highlighting this region as a potential site of a candidate gene for insulin secretion and T2DM.

One strong biological candidate gene in this region is insulin growth factor-1 receptor (*IGF1R*; MIM: 147370). The *IGF1R* gene is composed of 21 exons spanning 308 kb on chromosome 15q26.3. *IGF1R* is a transmembrane tyrosine kinase receptor widely expressed in human tissues with 50% amino acid homology with the insulin receptor. In beta-cell-specific *IGF1R* null mice, beta cells develop normally, but *GLUT2* and glucokinase expression are reduced, leading to diminished glucose-stimulated insulin secretion and IGT [7]. In a combined *IGF1R* and insulin receptor beta-cell-specific knockout, beta cells again develop normally, but undergo rapid apoptosis [8], suggesting that insulin and IGF1 signaling are critical in regulating beta cell mass, and defective signaling may contribute to T2DM development.

Based on these and other studies [9] and the linkage peak on chromosome 15q, *IGF1R* was selected as a candidate gene for further study of insulin secretion traits. We genotyped a dense panel of SNPs in nondiabetic AFDS participants and performed association analyses with Insulin30 and ISI phenotypes from the OGTT.

## Materials and methods

### Subjects

The design and methods applied in the AFDS have been described elsewhere [10]. Briefly, the AFDS includes 1402 subjects from 58 pedigrees, consisting of probands with T2DM and all willing first and second degree relatives of probands and spouses over the age of 18. OGTTs were performed on 883 subjects without a previous diagnosis of diabetes. The absence of diabetes was confirmed in 778 subjects. Table 1 summarizes characteristics of the nondiabetic participants. Insulin30 and ISI ( $ISI = \ln \text{Insulin30} - \ln \text{Insulin0}$ ) were used as proxy measures for insulin secretory function. The study protocol was approved by the institutional review board at the University of Maryland School of Medicine, and informed consent was obtained from each participant.

### SNP selection and genotyping

SNPs in *IGF1R* and 10 kb flanking sequences were selected for genotyping by first identifying 53 tagging SNPs. Criteria for tagging SNP selection included a  $MAF \geq 0.1$  and a minimum  $r^2 \geq 0.8$  between SNPs, and

**Table 1. Selected traits in 778 nondiabetic AFDS participants**

Trait	Mean $\pm$ standard deviation
Age (years)	43.9 $\pm$ 14.8
Female, $n$ (%)	411 (52.8)
IFG and/or IGT, $n$ (%)	159 (20.4)
BMI ( $\text{kg}/\text{m}^2$ )	27.0 $\pm$ 4.8
% body fat	23.2 $\pm$ 10.2
Fasting glucose (mmol/L)	5.0 $\pm$ 0.5
Glucose30 (mmol/L)	8.2 $\pm$ 1.6
Insulin0 (pmol/L)	65.6 $\pm$ 32.1
Insulin30 (pmol/L)	323.5 $\pm$ 213.7
ISI <sup>a</sup>	1.50 $\pm$ 0.50

$$^a ISI = \ln[\text{Insulin30 (pmol/L)}] - \ln[\text{Insulin0 (pmol/L)}].$$

then forced inclusion of four nonsynonymous coding SNPs for a total of 53 SNPs. Tagging SNPs were identified using the tag-SNP selection program Tagger [11], as implemented in Haploview [12]. The tagging SNPs were supplemented by an additional 46 SNPs from dbSNP in regions with  $>10$  kb between tagging SNPs. SNP genotyping was performed either using Ultra-High Throughput Genotyping (SNPStream, Beckman Coulter; Fullerton, CA), Pyrosequencing (PSQ HS96A System, Biotage; Westborough, MA), or TaqMan SNP Genotyping Assay (ABI PRISM<sup>®</sup> 7900HT Sequence Detection System, Applied Biosystems Inc., Foster City, CA).

Prior to analysis, genotype data was analyzed in PEDCHECK [13] to identify Mendelian inconsistencies. One hundred and two blind replicates were included to determine replication rates. Markers with replication rates  $<95\%$  (five markers) and call rates  $<85\%$  (four markers) were excluded from analyses. Of the remaining SNPs, 36 were monomorphic or had a  $MAF < 0.05$  leaving a total of 54 SNPs for further analysis. HWE was determined for each SNP using a threshold of  $P > 0.001$  to account for multiple comparisons.

### Analysis

We estimated the effects of genotype on  $\ln$  Insulin30 and ISI levels in 778 subjects confirmed to have no diabetes. These analyses were adjusted for sex, age, age<sup>2</sup>, BMI and Insulin0, all independently associated with both phenotypes. Because subjects were recruited by family, estimates were obtained for mean  $\ln$  Insulin30 and ISI levels by genotype conditioning on the pedigree structure in association via a 'measured genotype' approach, and statistical significance for the association was determined using a likelihood ratio test under an additive model with one degree of freedom [14]. To account for multiple testing, a principal components-based approach was used correcting for LD between proximal SNPs [15] adjusting our threshold for statistical significance to  $P < 0.002$ .

As the average LD block length was 3.5 SNPs, haplotypes were chosen to be four SNPs long to minimize multiple hypothesis testing. Four-SNP haplotypes (Table S3, Supporting Information) were inferred under an assumption of zero recombinants, as implemented in ZAPLO [16],

with extended pedigrees subdivided into 165 smaller, nuclear pedigrees due to computational complexity in inferring haplotypes for the larger pedigrees. The most probable inferred haplotype pairs were used in analyses. Haplotypic association tests followed a similar approach to genotypic association analysis. To determine the clinical consequences of insulin and ISI-associated SNPs and haplotypes, we then compared SNP and haplotype frequencies between subjects with ( $n = 139$ ) and without ( $n = 468$ ) diabetes. The subjects without diabetes were chosen on the basis of their age being 38 years or older. All analyses were performed using SOLAR [17].

Joint variance components linkage and association analyses of ln Insulin30 and ISI were then performed in 468 nondiabetic AFDS participants for whom both microsatellite marker and *IGF1R* SNP genotype data were available in order to determine if the observed association was due to linkage to Insulin30 or ISI on chromosome 15 [6]. These analyses incorporated each SNP or haplotype by adjusting for the genotype as a covariate, along with adjusting for sex, age, age<sup>2</sup>, BMI, and Insulin0. Insulin30 levels were transformed by their natural logarithm prior to analysis to remove skewness.

## Results

Fifty-four SNPs (average distance = 5.7 kb), including 50 intronic SNPs, one 3'-untranslated-region SNP and three synonymous coding SNPs were polymorphic and had call rates >85% (Table S1, Supporting Information). Genotype frequencies for the 54 SNPs did not differ significantly from those predicted under HWE ( $P > 0.001$ ). LD patterns were similar to or higher than those estimated from the Caucasian founders in HapMap (Figure 1).

Of the 54 polymorphic SNPs, none were significantly associated ( $P < 0.002$ ) with variation in Insulin30 (Figure 2A). However, six SNPs with  $P < 0.05$  ( $P = 0.011$ – $0.047$ ) were identified (Intron 1: rs874305, rs4966009, rs4966012, rs1319859; Intron 2: rs2670504, rs2715440) (Table S2, Supporting Information). As with Insulin30, none of the SNPs were significantly associated with ISI at  $P < 0.002$  (Figure 2B). Associations for ISI of  $P < 0.05$  were observed for three clusters of SNPs in intron 1 (rs874305, rs4966012, rs1319859), intron 2 (rs2670504, rs2684778, rs2715440, rs7165181, rs4966035), and intron 5 (rs3743259, rs4966038) (Table S2, Supporting Information).

While no 4-SNP haplotypes demonstrated association with Insulin30 (data not shown), the CTTG haplotype of SNPs rs17847195, rs2715439, rs8034284, and rs12440962 (haplotype 48) in intron 20 was associated with lower ISI levels ( $\beta = 0.18$ ; 95% CI: 0.08, 0.29; genotypic means: CTTG/CTTG  $1.37 \pm 0.09$  versus CTTG/CTCG  $1.54 \pm 0.09$  versus CTCG/CTCG  $1.73 \pm 0.11$ ) (Figure 3), exceeding our multiple-testing-adjusted significance threshold of  $\alpha = 0.002$  with  $P = 0.001$ . Haplotypes 46 (rs8038056–rs41497346–rs17847195–rs2715439), 47

(rs41497346–rs17847195–rs2715439–rs8034284), 49 (rs2715439–rs8034284–rs12440962–rs2593053), and 50 (rs8034284–rs12440962–rs2593053–rs17847203), which flank haplotype 48, displayed associations with lower ISI consistent with haplotype 48 at  $P < 0.05$  (Table S4, Supporting Information).

We generated posterior estimates of power to detect association of Haplotype 48 with ISI levels in the sample of 468 AFDS members with complete haplotype and phenotypic data using QUANTO [18] and found that we had 86% power to detect a statistically significant difference of 0.18 in the ISI levels between genotype at  $\alpha = 0.002$  under an additive model for a haplotype with frequency of 0.16.

We tested the association of the SNPs in haplotype 48 with T2DM in 139 AFDS T2DM cases and 468 AFDS controls. While none of the SNPs were individually associated with T2DM, the CTTG haplotype was modestly associated with elevated T2DM risk ( $P = 0.02$ ) in a model adjusted for sex, age, age<sup>2</sup>, BMI, and Insulin0, although not at the gene-wide significance threshold of  $\alpha = 0.002$ .

To determine if the linkage on chromosome 15q could be attributed to associations with *IGF1R* SNPs, we performed joint linkage and association analyses of ln Insulin30 and ISI in 468 nondiabetic AFDS subjects with chromosome 15 microsatellite data [6]. Linkage analyses without adjustment for *IGF1R* genotypes demonstrated peak locus-specific LOD scores (ln Insulin30: LOD = 2.23; ISI: LOD = 3.13) in the same region of chromosome 15 (53–54 cM). Adjustment for associated variants, including haplotype 48, in linkage analyses did not substantially reduce Insulin30 and ISI peak locus-specific LOD scores (ln Insulin30: LOD = 1.72–2.17; ISI: LOD = 1.61–3.02), suggesting that these markers do not account for the previously observed linkage on chromosome 15q [6].

## Discussion

Previous studies in *IGF1R* knockout mice showing abnormal insulin secretion and the linkage peak to Insulin30 on chromosome 15q in the Amish under which *IGF1R* sits make *IGF1R* a strong positional candidate gene for insulin secretion and T2DM. Based on our detailed examination of *IGF1R*, we found evidence that *IGF1R* may influence insulin secretion in humans and thus may play a role in T2DM pathology. Here, we report an association of a haplotype in intron 20 of *IGF1R*, the CTTG-variant (frequency = 0.13) of SNPs rs17847195, rs2715439, rs8034284, and rs12440962, with lower ISI ( $P = 0.001$ ) and elevated risk of T2DM ( $P = 0.02$ ).

While some associations of genes with diabetes may be spurious [19], studies in mice suggest that the association of *IGF1R* variants with insulin secretion may underlie a true relationship. As described previously, mice with beta cells lacking *IGF1R* demonstrate defective glucose-driven insulin secretion [7], suggesting *IGF1R* is critical

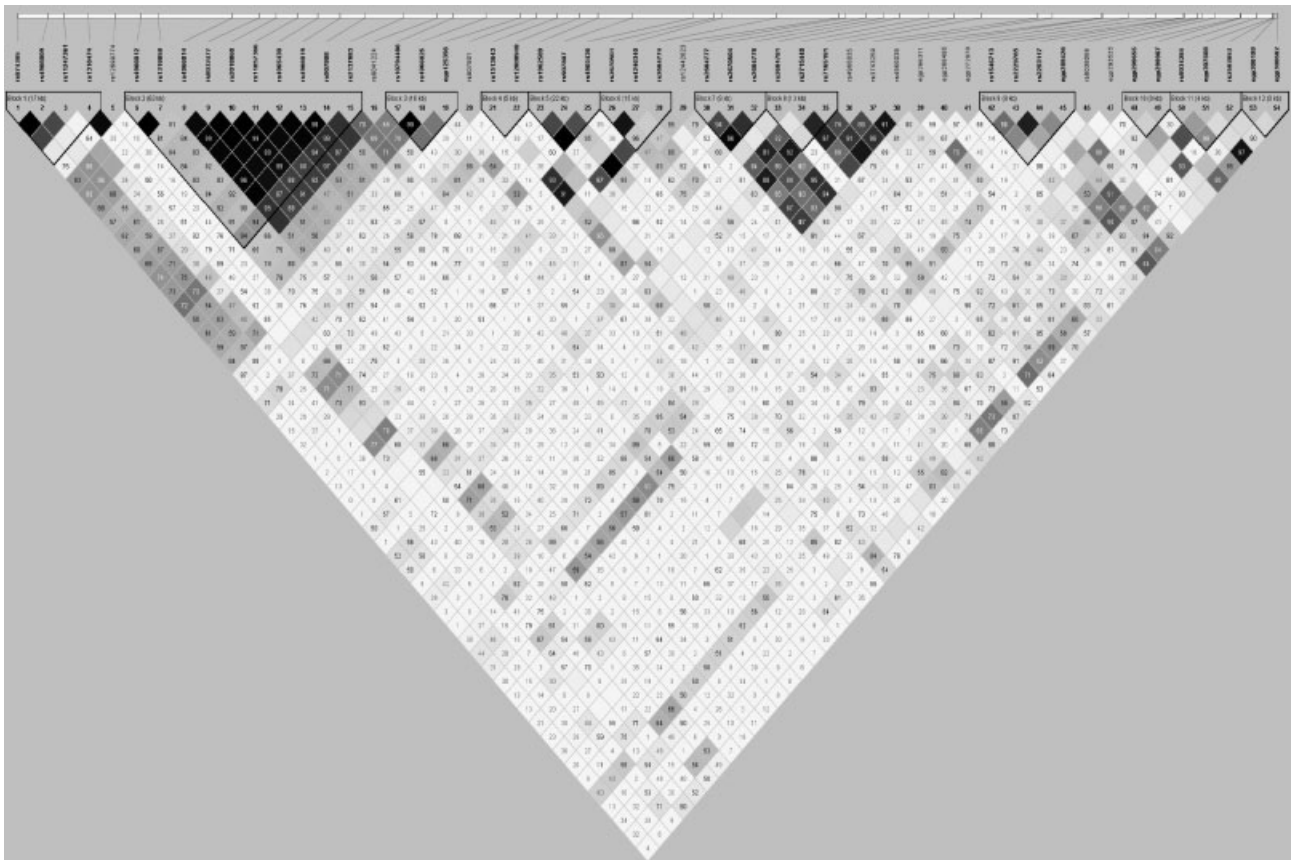


Figure 1. LD ( $r^2$ ) between 54 SNPs genotyped in *IGF1R*

for proper beta cell function. Similarly, mice with beta cell-specific knockouts of both insulin receptor and *IGF1R* are born with a normal complement of beta cells, but glucose-stimulated insulin secretion is blunted. As a result, by 3 weeks, the mice developed diabetes [8].

Previous studies of *IGF1R* sequence variation and T2DM risk are conflicting. Rasmussen *et al.* [18] analyzed *IGF1R* coding sequence in 82 Danish subjects with T2DM, and found no variants associated with T2DM or related endophenotypes including reduced birth weight and insulin sensitivity index. In contrast, in the Finnish Diabetes Prevention Study [19], a common synonymous coding variation in *IGF1R* (GAG1043GAA) (rs2229765) was significantly associated with differences in conversion rates from IGT to T2DM in subjects participating in a weight loss program, suggesting that this polymorphism increases T2DM risk among those with IGT. In our study, rs2229765 was not associated with differences in Insulin30 ( $P = 0.9$ ) or ISI ( $P = 0.9$ ) and was marginally associated with diabetes ( $P = 0.008$ ) using  $\alpha = 0.002$  for statistical significance. rs2229765, however, is present in a haplotype associated with ISI (haplotype 42), and is in strong LD with several of the SNPs in the associated haplotype in intron 20. One possible explanation for the lack of association at rs2229765 in our study may be the differences in the LD patterns underlying *IGF1R* among different populations.

Several features of our study may have improved our power to detect these associations with *IGF1R*. First,

our study was performed in subjects from a genetic isolate, thereby, minimizing confounding effects from population stratification. Second, our background LD is likely higher than in most Northern European samples [20], such that we may have better gene coverage than other outbred populations. Third, AFDS participants have similar occupational and environmental exposures [21] as they are predominantly farmers and homemakers from Lancaster, Pennsylvania, thus decreasing confounding effects from these factors. Fourth, our power to detect association with *IGF1R* may have been increased by analyzing T2DM subphenotypes (Insulin30, ISI) since these subphenotypes are likely to be less genetically complex. Finally, the existence of linkage to insulin levels and T2DM in more than one population near *IGF1R* [6,22,23] may have reduced the likelihood of a false-positive error due to multiple testing [24].

Despite the strengths of our study, the study was subject to several limitations. While we found our strongest association with a haplotype, type I error may be present in our analyses as a result of inferring haplotypes with only nuclear pedigrees. Another potential source of error may be the use of the most probable inferred haplotypes, although the average posterior probability for inferred haplotypes was 0.66, indicating that it is unlikely to be a major source of error. To minimize these limitations, an adjusted threshold for statistically significant association was used. Among the genotyped

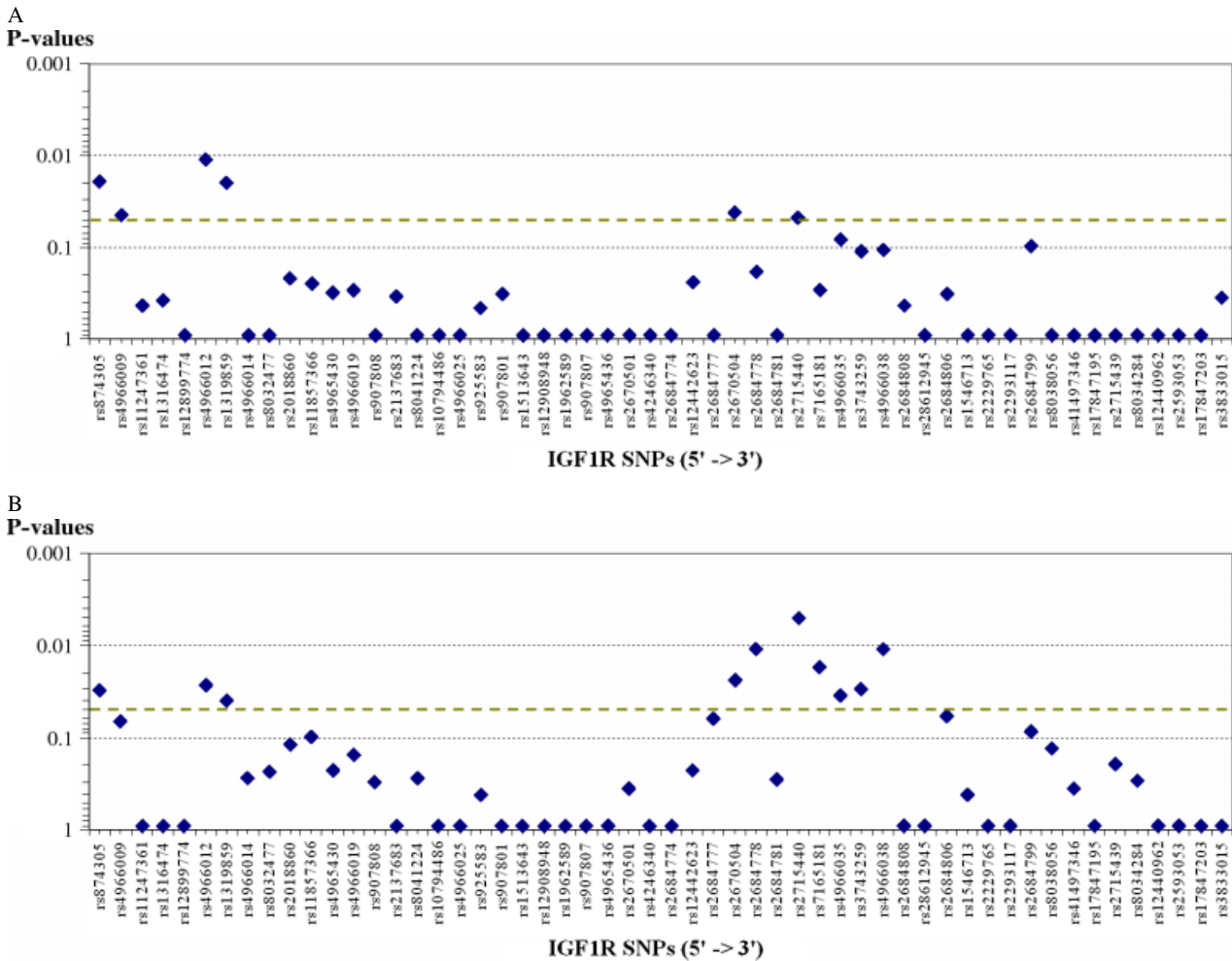


Figure 2. P-values for single SNP associations with ln Insulin30 (A) and ISI (B) in nondiabetic AFDS participants. P-values are estimated under an additive model. Associations are adjusted for age, age<sup>2</sup>, sex, BMI, and Insulin0. The dashed line marks P = 0.05. This figure is available in colour online at [www.interscience.wiley.com/journal/dmrr](http://www.interscience.wiley.com/journal/dmrr)

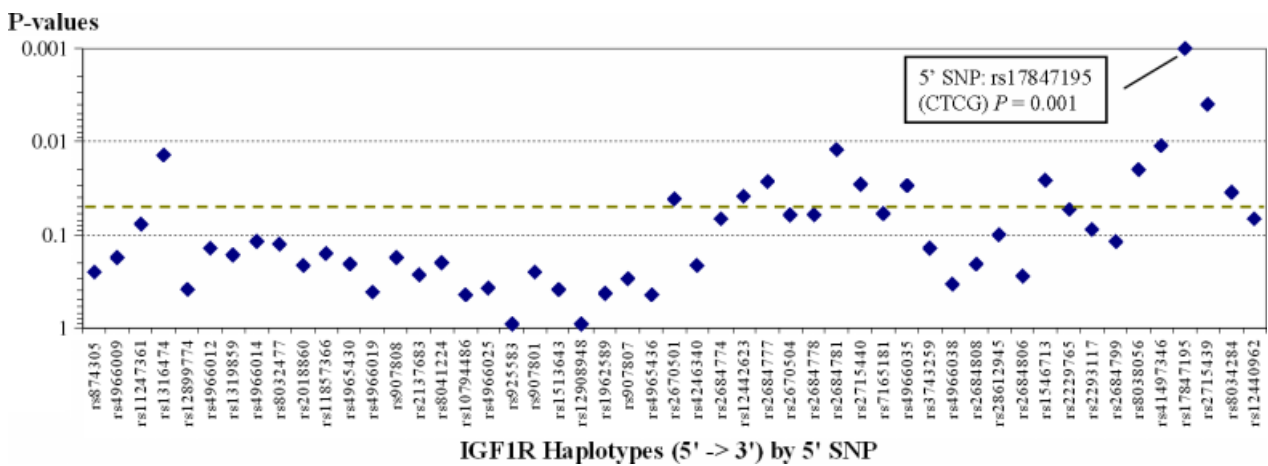


Figure 3. P-values for 4-SNP haplotype associations with ISI in nondiabetic AFDS participants. P-values were estimated under an additive model. Associations are adjusted for age, age<sup>2</sup>, sex, BMI, and Insulin0. The dashed line marks P = 0.05. Haplotypes are named for the most 5' SNP in the haplotype. This figure is available in colour online at [www.interscience.wiley.com/journal/dmrr](http://www.interscience.wiley.com/journal/dmrr)

SNPs, we observed high replication rates; however, our call rates were lower than anticipated (average 89%). Similar replication and call rates have been reported by others using SNPStream Ultra-High Throughput for

genotyping [25]. To ensure that no systematic error was present, we examined genotyping plates and found the proportion of uncalled genotypes to be consistent across plates. Moreover, the fact that all SNPs were in HWE

provided no strong evidence for under-calling or over-calling of heterozygosity. Furthermore, no statistically significant differences in trait distributions among those with and without genotype calls were found. Another potential limitation was that 45 of 99 SNPs were monomorphic, had a MAF <0.05, or were unsuccessfully genotyped. Despite this problem, the remaining 54 SNPs demonstrated a high degree of pairwise LD (average  $r^2 = 0.94$ ) and achieved a high degree of coverage of the gene. Finally, as the Amish are genetically isolated, it is possible that our association of *IGF1R* with ISI may not be generalizable to other populations; however, previous genetic studies of T2DM and related traits in the Amish have found results concordant with other Caucasian populations [26,27], suggesting that our findings in the Amish will likely be relevant to more outbred populations.

While we did not observe statistically significant single-SNP associations, there are several reasons to believe that the observed haplotypic association is real. Haplotypic association may have increased power to detect genetic effects at ungenotyped causal SNPs through indirect association, as haplotypes may be better than individual SNPs at tagging causal SNPs via LD, as exemplified by Martin *et al.* [28] for Apolipoprotein E and Alzheimer's disease and Veal *et al.* [29] for *PSORS1* and psoriasis, among others [30,31]. Haplotype frequencies may approximate more closely the allele frequency of a causative SNP [32], which may be the case with haplotype 48. Additionally, haplotypes themselves may contribute to risk for disease, as several studies have identified important haplotypic risk variants for disease such as the *APOE*  $\epsilon 3/\epsilon 4$  alleles for Alzheimer's disease [33],  $\beta 2$ -adrenergic receptor ( *$\beta 2AR$* ) in bronchodilator response [34], and complement factor H in age-related macular degeneration [35]. However, given that the haplotype is only modestly associated with ISI after adjustment for multiple hypothesis testing, this association needs to be replicated in an independent population.

The experience of our study mirrors that of many others insofar as it illustrates the challenges of going from a positive linkage result to identifying a causative SNP. One possibility (among several) is that the effect detected through linkage analysis may be attributable to one or more rare variants, while our strategy for identifying associated SNPs was predicated on a tagging SNP approach geared towards detecting common SNPs. Identifying the functional SNPs whose frequencies are rare remains a very challenging enterprise that may require both additional approaches for SNP discovery (e.g., sequencing) as well as functional assessment of associated SNPs.

In summary, our study suggests that *IGF1R* variants may influence insulin secretion. However, further studies are needed to confirm these findings.

## Supporting information

Supporting information may be found in the online version of this article.

## Acknowledgements

This work was supported by National Institutes of Health research grants R01 DK068495 (K. D. S), K01 DK067207 (W. H. L. K.), R01 DK54261, U01 HL084756-02 (J. R. O), and training grant T32 HL07024 (A. C. N); the University of Maryland General Clinical Research Center (M01 RR16500); Hopkins Bayview General Clinical Research Center (M01 RR02719); the Maryland Clinical Nutrition Research Unit (P30 DK072488); and the Baltimore Veterans Administration Geriatric Research and Education Clinical Center. We gratefully acknowledge our Amish liaisons and fieldworkers and the extraordinary cooperation and support of the Amish community, without whom these studies would not be possible.

## Conflict of Interest

None declared.

## References

- DeFronzo RA. Pathogenesis of type 2 (non-insulin dependent) diabetes mellitus: a balanced overview. *Diabetologia* 1992; **35**: 389–397.
- Chang AM, Halter JB. Aging and insulin secretion. *Am J Physiol Endocrinol Metab* 2003; **284**: E7–E12.
- Frayling TM. Genome-wide association studies provide new insights into type 2 diabetes aetiology. *Nat Rev Genet* 2007; **8**: 657–662.
- Grarup N, Rose CS, Andersson EA, *et al.* Studies of association of variants near the HHEX, CDKN2A/B and IGF2BP2 genes with type 2 diabetes and impaired insulin release in 10,705 Danish subjects validation and extension of genome-wide association studies. *Diabetes* 2007; **56**: 3105–3111.
- Pascoe L, Tura A, Patel SK, *et al.* Common variants of the novel type 2 diabetes genes, CDKAL1 and HHEX/IDE, are associated with decreased pancreatic  $\beta$ -cell function. *Diabetes* 2007; **56**: 3101–3104.
- Hsueh WC, Silver KD, Pollin TI, *et al.* A genome-wide linkage scan of insulin level derived traits: the Amish Family Diabetes Study. *Diabetes* 2007; **56**: 2643–2648.
- Kulkarni RN, Holzenberger M, Shih DQ, *et al.* Beta-cell-specific deletion of the Igf1 receptor leads to hyperinsulinemia and glucose intolerance but does not alter beta-cell mass. *Nat Genet* 2002; **31**: 111–115.
- Ueki K, Okada T, Hu J, *et al.* Total insulin and IGF-I resistance in pancreatic beta cells causes overt diabetes. *Nat Genet* 2006; **38**: 583–588.
- Xuan S, Kitamura T, Nakae J, *et al.* Defective insulin secretion in pancreatic beta cells lacking type 1 IGF receptor. *J Clin Invest* 2002; **110**: 1011–1019.
- Hsueh WC, Mitchell BD, Aburomia R, *et al.* Diabetes in the Old Order Amish: characterization and heritability analysis of the Amish Family Diabetes Study. *Diabetes Care* 2000; **23**: 595–601.
- de Bakker PI, Yelensky R, Pe'er I, Gabriel SB, Daly MJ, Altshuler D. Efficiency and power in genetic association studies. *Nat Genet* 2005; **37**: 1217–1223.
- Barrett JC, Fry B, Maller J, Daly MJ. Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics* 2005; **21**: 263–265.

13. O'Connell JR, Weeks DE. PedCheck: a program for identification of genotype incompatibilities in linkage analysis. *Am J Hum Genet* 1998; **63**: 259–266.
14. Edwards AWF. *Likelihood (Expanded Edition)*. Baltimore, MD, The Johns Hopkins University Press: 1992.
15. Nyholt DR. A simple correction for multiple testing for single-nucleotide polymorphisms in linkage disequilibrium with each other. *Am J Hum Genet* 2004; **74**: 765–769.
16. O'Connell JR. Zero-recombinant haplotyping: applications to fine mapping using SNPs. *Genet Epidemiol* 2000; **19**(Suppl. 1): S64–S70.
17. Almasy L, Blangero J. Multipoint quantitative-trait linkage analysis in general pedigrees. *Am J Hum Genet* 1998; **62**: 1198–1211.
18. Gauderman WJ, Morrison J. QUANTO 1.1: a computer program for power and sample size calculations for genetic-epidemiology studies. 2006.
19. Jafar-Mohammadi B, McCarthy MI. Genetics of type 2 diabetes mellitus and obesity – a review. *Ann Med* 2008; **40**: 2–10.
20. Service SK, Ophoff RA, Freimer NB. The genome-wide distribution of background linkage disequilibrium in a population isolate. *Hum Mol Genet* 2001; **10**: 545–551.
21. Sheffield VC, Stone EM, Carmi R. Use of isolated inbred human populations for identification of disease genes. *Trends Genet* 1998; **14**: 391–396.
22. Mori Y, Otabe S, Dina C, *et al*. Genome-wide search for type 2 diabetes in Japanese affected sib-pairs confirms susceptibility genes on 3q, 15q, and 20q and identifies two new candidate loci on 7p and 11p. *Diabetes* 2002; **51**: 1247–1255.
23. Hanis CL, Boerwinkle E, Chakraborty R, *et al*. A genome-wide search for human non-insulin-dependent (type 2) diabetes genes reveals a major susceptibility locus on chromosome 2. *Nat Genet* 1996; **13**: 161–166.
24. Blangero J. Localization and identification of human quantitative trait loci: king harvest has surely come. *Curr Opin Genet Dev* 2004; **14**: 233–240.
25. Bell PA, Chaturvedi S, Gelfand CA, *et al*. SNPstream UHT: ultra-high throughput SNP genotyping for pharmacogenomics and drug discovery. *Biotechniques* 2002; Suppl: 70–72, 74, 76–77.
26. Damcott CM, Pollin TI, Reinhart LJ, *et al*. Polymorphisms in the transcription factor 7-like 2 (TCF7L2) gene are associated with type 2 diabetes in the Amish: replication and evidence for a role in both insulin secretion and insulin resistance. *Diabetes* 2006; **55**: 2654–2659.
27. Damcott CM, Hoppman N, Ott SH, *et al*. Polymorphisms in both promoters of hepatocyte nuclear factor 4-alpha are associated with type 2 diabetes in the Amish. *Diabetes* 2004; **53**: 3337–3341.
28. Martin ER, Lai EH, Gilbert JR, *et al*. SNPping away at complex diseases: analysis of single-nucleotide polymorphisms around APOE in Alzheimer disease. *Am J Hum Genet* 2000; **67**: 383–394.
29. Veal CD, Capon F, Allen MH, *et al*. Family-based analysis using a dense single-nucleotide polymorphism-based map defines genetic variation at PSORS1, the major psoriasis-susceptibility locus. *Am J Hum Genet* 2002; **71**: 554–564.
30. Hennah W, Varilo T, Kestila M, *et al*. Haplotype transmission analysis provides evidence of association for DISC1 to schizophrenia and suggests sex-dependent effects. *Hum Mol Genet* 2003; **12**: 3151–3159.
31. Rioux JD, Daly MJ, Silverberg MS, *et al*. Genetic variation in the 5q31 cytokine gene cluster confers susceptibility to Crohn disease. *Nat Genet* 2001; **29**: 223–228.
32. Akey J, Jin L, Xiong M. Haplotypes vs single marker linkage disequilibrium tests: what do we gain? *Eur J Hum Genet* 2001; **9**: 291–300.
33. Saunders AM, Strittmatter WJ, Schmechel D, *et al*. Association of apolipoprotein E allele epsilon 4 with late-onset familial and sporadic Alzheimer's disease. *Neurology* 1993; **43**: 1467–1472.
34. Drysdale CM, McGraw DW, Stack CB, *et al*. Complex promoter and coding region beta 2-adrenergic receptor haplotypes alter receptor expression and predict in vivo responsiveness. *Proc Natl Acad Sci U S A* 2000; **97**: 10483–10488.
35. Li M, Atmaca-Sonmez P, Othman M, *et al*. CFH haplotypes without the Y402H coding variant show strong association with susceptibility to age-related macular degeneration. *Nat Genet* 2006; **38**: 1049–1054.