

# Ancestry-informative markers on chromosomes 2, 8 and 15 are associated with insulin-related traits in a racially diverse sample of children

Yann C. Klimentidis,<sup>1\*</sup> Jasmin Divers,<sup>2</sup> Krista Casazza,<sup>3</sup> T. Mark Beasley,<sup>1</sup> David B. Allison<sup>1,3</sup> and Jose R. Fernandez<sup>1,3</sup>

<sup>1</sup>Section on Statistical Genetics, Department of Biostatistics, University of Alabama at Birmingham, Birmingham, AL, USA

<sup>2</sup>Department of Biostatistical Sciences, Wake Forest University Health Sciences, Winston-Salem, NC, USA

<sup>3</sup>Department of Nutrition Sciences, University of Alabama at Birmingham, Birmingham, AL, USA

\*Correspondence to: Tel: +1 205 975 9273; Fax: +1 205 975 2540; E-mail: yann@uab.edu

Date received (in revised form): 29th September 2010

## Abstract

Type 2 diabetes represents an increasing health burden. Its prevalence is rising among younger age groups and differs among racial/ethnic groups. Little is known about its genetic basis, including whether there is a genetic basis for racial/ethnic disparities. We examined a multi-ethnic sample of 253 healthy children to evaluate associations between insulin-related phenotypes and 142 ancestry-informative markers (AIMs), while adjusting for sex, age, Tanner stage, genetic admixture, total body fat, height and socio-economic status. We also evaluated the effect of measurement errors in the estimation of the individual ancestry proportions on the regression results. We found that European genetic admixture is positively associated with insulin sensitivity ( $S_I$ ), and negatively associated with the acute insulin response to glucose, fasting insulin levels and the homeostasis model assessment of insulin resistance. Our analysis revealed associations between individual AIMs on chromosomes 2, 8 and 15 and these phenotypes. Most notably, marker rs3287 at chromosome 2p21 was found to be associated with  $S_I$  ( $p = 5.8 \times 10^{-5}$ ). This marker may be in admixture linkage disequilibrium with nearby loci (*THADA* and *BCL11A*) that previously have been reported to be associated with diabetes and diabetes-related phenotypes in several genome-wide association and linkage studies. Our results provide further evidence that variation in the 2p21 region containing *THADA* and *BCL11A* is associated with type 2 diabetes. Importantly, we have implicated this region in the early development of diabetes-related phenotypes, and in the genetic aetiology of population differences in these phenotypes.

**Keywords:** insulin sensitivity, genetic admixture, type 2 diabetes, genetic association, ancestry-informative marker

## Introduction

Type 2 diabetes prevalence in the paediatric population is increasing, while age at onset is decreasing.<sup>1,2</sup> Type 2 diabetes also disproportionately affects racial/ethnic minorities in the USA.<sup>3</sup> Twin and familial studies have shown a substantial genetic component to the disease, as well as its related phenotypes.<sup>4–10</sup> Although much is known about

environmental contributions to type 2 diabetes, only a very small proportion of the variation due to genetic factors is currently explainable by identified genetic polymorphisms.<sup>11–13</sup> Similarly, little is known about the specific genetic factors that may contribute to population differences in diabetes prevalence. Because the origins of type 2 diabetes are likely to be rooted in childhood, a better

understanding of genetic determinants among paediatric populations could lead to a better insight into the aetiology of type 2 diabetes and eventually improved prediction and prevention of the disease.

Endo-phenotypes can be useful in closely dissecting the genetic basis of eventual disease status.<sup>14</sup> For type 2 diabetes, several such measurable phenotypes exist, typically examining measures of glucose and insulin homeostasis. These measures serve as indicators of reduced insulin response and action that may presage type 2 diabetes.<sup>15,16</sup> Furthermore, previous studies have suggested a genetic basis for racial/ethnic differences in insulin dynamics.<sup>17,18</sup> Examining the genetic basis for these detailed phenotypes therefore allows for a much better understanding of the link between the genetic and metabolic pathways that underlie the development of type 2 diabetes.

Since the loci recently identified by genome-wide association studies (GWAS)<sup>19–21</sup> occurred predominantly among individuals of European descent, there is considerable uncertainty as to whether these associations translate to other populations. Hispanic Americans (HAs) and African Americans (AAs) suffer from higher rates of type 2 diabetes than European Americans (EAs),<sup>22,23</sup> and similar differences exist for endo-phenotypes among children, regardless of disease status.<sup>24–28</sup> Admixed populations — such as HAs, AAs and EAs — can be examined to determine if there is a genetic basis for these population differences and to identify specific genetic regions associated with both ancestry and insulin-related outcomes.<sup>29</sup> Previous studies have shown that admixture is a strong predictor of diabetes and insulin-related traits but also that this relationship may be mediated through environmental factors such as income or educational level that are correlated with genetic admixture.<sup>17, 30–32</sup>

Genetic mapping methods such as admixture mapping capitalise on population differences in a trait and the extended blocks of linkage disequilibrium (LD) created after the admixture process.<sup>33</sup> They are, therefore, most applicable to recently admixed populations such as HAs and AAs. To our knowledge, only one previous study has used this type of method to localise genetic variants associated with type 2 diabetes.<sup>34</sup> Given the large

differences in type 2 diabetes and the plausible genetic origins of these differences, there is a great need for additional studies that attempt to pinpoint population-specific genetic risk factors.

In this study, we also capitalise on population differences in a phenotype and the extended LD blocks resulting from admixture in an attempt to identify specific genetic regions that might be involved in the aetiology of population differences in insulin-related traits. Specifically, we examine the association between genetic admixture and four insulin-related outcomes: insulin sensitivity ( $S_I$ ), acute insulin response to glucose (AIR<sub>g</sub>), fasting insulin (FI) levels and the homeostasis model assessment of insulin resistance (HOMA-IR). We then examine the association between 142 ancestry-informative markers (AIMs) and each of these traits to identify genetic regions that potentially account for ethnic/racial differences. We performed this study among a multi-ethnic sample of children, adjusting for known influential covariates, including genetic admixture as a control for genetic background, and socio-economic status. We place our findings in the context of other studies that have found associations in similar regions.

## Methods

### Study participants

A total of 253 children between the ages of seven and 12 (52 per cent male) were recruited as part of a cross-sectional cohort study examining population differences in metabolic phenotypes among children with no major illnesses or medical diagnoses. The children were classified by parental report as AA ( $n = 87$ ), EA ( $n = 108$ ), HAs ( $n = 52$ ) and bi-racial ( $n = 6$ ). All children were pubertal stage  $\leq 3$ , as assessed by a paediatrician according to the criteria of Marshall and Tanner.<sup>35</sup> Written informed assent and consent were obtained from children and parents, respectively, as approved by the University of Alabama at Birmingham Institutional Review Board. All measurements were taken between 2004 and 2008 at the University of Alabama at Birmingham General Clinical Research Center (GCRC) and Department of Nutrition Sciences.

### Anthropometric measurements

In the first of two sessions completed by participants, pubertal status, anthropometric measurements and body composition were assessed. Height was measured without shoes to the nearest centimetre using a stadiometre (Heightronic 235; Measurement Concepts, Snoqualmie, WA, USA). Body composition was assessed by dual-energy X-ray absorptiometry (DXA) using a GE Lunar Prodigy densitometer (GE LUNAR Radiation Corp., Madison, WI, USA), as previously described.<sup>36</sup> Participants were measured while lying flat on their backs with arms at their sides, wearing light clothing. Analysis of DXA scans was performed using paediatric software (Encore 2002, version 6.10.029).

### Insulin-related measurements

At the second visit (which took place within 30 days of the first visit), participants were admitted to the GCRC for an overnight visit. Following an overnight fast, blood samples were obtained to establish the basal levels of glucose and insulin, and a frequently sampled intravenous glucose tolerance test (FSIGTT) was performed as described elsewhere.<sup>37–39</sup>  $S_I$  (the increase in fractional glucose disappearance per unit of insulin increase) and AIR<sub>g</sub> (the area above baseline insulin concentration during ten minutes following exposure to glucose) were estimated from the FSIGTT using minimal modelling.<sup>40</sup> HOMA-IR, a surrogate measure of insulin resistance, derived as: fasting glucose (mg/dl) X fasting insulin ( $\mu$ U/ml)/405<sup>41</sup> was also calculated.

### Socio-economic status (SES)

SES was measured using the Hollingshead 4-factor index of social class, which combines information on the education and occupational prestige of parents.<sup>42</sup> Scores range from 8 to 66, with higher scores representing higher status.

### Genetic analysis

DNA from blood was obtained from all study participants and was typed at 142 AIMs by Prevention

Genetics (Marshfield, WI, USA). Each marker (single nucleotide polymorphism [SNP]) was genotyped using a fluorescent allele-specific polymerase chain reaction (AS-PCR)-based assay.<sup>43</sup> Reaction components were assembled on an array tape platform (<http://www.douglasscientific.com>) using nanolitre volumes (500–1000 nl). PCRs were carried out in a water bath thermocycler using a standard three-stage parameter (denaturation, primer annealing, primer extension). The specific parameters of each PCR vary depending on the nature of the primers and the SNP being genotyped. The array tape was scanned post-PCR and the ratio of fluorescent signals was used to determine the genotype (homozygous for one allele, or heterozygous). A subset of these AIMs is described elsewhere.<sup>44</sup>

These markers were chosen because they exhibit large frequency differences between ancestral West African, Amerindian and West European populations. Individual West African, Amerindian and European genetic admixture estimates were obtained by maximum likelihood estimation,<sup>45</sup> using the genotypes at each AIM and an estimate of the allele frequencies of these AIMs in the three ancestral parental populations (see Table S1).

### Statistical analyses

Differences in mean values for phenotypes between racial/ethnic groups were examined using analysis of variance. Multiple linear regression analyses were used to test the association between European admixture and total fat and the four insulin-related phenotypes, and to examine the association between each of 142 SNPs and four insulin-related phenotypes. For  $S_I$ , FI and HOMA-IR, the model was defined by age, Tanner stage, sex, SES, European admixture, Amerindian admixture, total fat and height. By controlling for two of three admixture estimates, we prevented the introduction of co-linearity in the statistical models, since the three admixture estimates add up to 1. For AIR<sub>g</sub>, the model was additionally adjusted for  $S_I$ . To conform to the assumptions of regression, all models were evaluated for residual normality;

logarithmic transformation was performed when appropriate. Outliers were removed based on whether residuals were greater than three standard deviations away from the mean.

Genotyped SNPs were tested for association with the four insulin-related phenotypes using linear regression under additive, dominant, recessive and two-degrees-of-freedom genotypic models. Considering each phenotype and each genetic model separately, we applied a Bonferroni multiple correction to the marker association tests; a  $p$ -value cut-off of  $3.6 \times 10^{-4}$  keeps the nominal type I error rate at 0.05. To determine the extent to which measurement error in admixture estimates could skew the results, we applied the method described by Divers *et al.*<sup>46</sup> Basically, we obtained an estimate of the measurement error covariance and applied the simulation extrapolation (SimEx) algorithm<sup>47</sup> to retest for association between each marker and phenotype, for each mode of inheritance model. Analyses were carried out with PLINK,<sup>48</sup> SAS 9.1 software (SAS Institute, Cary, NC, USA) and R.<sup>49</sup>

## Results

### Descriptive characteristics

Table 1 shows the descriptive characteristics of the sample. Differences in total fat, AIRg, FI and HOMA-IR were statistically significant between racial/ethnic groups (all at  $p < 0.01$ ). HA had higher total fat, FI and HOMA-IR than both EA

and AA. AA had higher AIRg, and lower  $S_I$  than other groups. EA had the highest  $S_I$  values ( $p < 0.0001$ ).

### Association between genetic admixture and insulin-related phenotypes

The associations between European genetic admixture and total fat and all insulin-related phenotypes were statistically significant ( $p < 0.05$ ; Table 2). Individuals with higher European admixture had less total fat, higher  $S_I$  and lower AIRg, FI, and HOMA-IR. In these models, there were significant associations of Tanner stage and total fat with  $S_I$ , FI and HOMA-IR ( $p < 0.01$ ). Total fat was also associated with AIRg ( $p = 0.04$ ). Upon analysing these associations within racial/ethnic groups, we found significant associations of European genetic admixture only among HA for  $S_I$ , FI and AIRg ( $p < 0.05$ ). These results suggest that specific genetic variants may exist, contributing to population differences for these phenotypes.

### Association between single markers and insulin-related phenotypes

The results of single SNP analyses are presented in Table 3. Marker rs3287, located at 2p21, is significantly associated with  $S_I$  under the recessive ( $p = 5.8 \times 10^{-5}$ ) and genotypic ( $p = 1.9 \times 10^{-4}$ ) models. Allele G at this SNP is associated with decreased  $S_I$ , and is at a higher frequency in the West African parental population (0.75) compared with both the European (0.18) and Amerindian

Table 1. Population characteristics for total sample and for each racial/ethnic group (mean  $\pm$  standard deviation)

	Total sample	European Americans	African Americans	Hispanic Americans	Other
Female/male (n)	122/131	51/57	40/47	28/24	3/3
Total fat (kg)	8.76 $\pm$ 5.49	8.35 $\pm$ 5.11	8.13 $\pm$ 6.02	10.77 $\pm$ 5.13	7.22 $\pm$ 4.52
AIRg ( $\mu$ IU/ml $\times$ 10 min)	903.9 $\pm$ 708.4	619.1 $\pm$ 425.4	1235.4 $\pm$ 827.5	940.8 $\pm$ 748.6	768.0 $\pm$ 328.8
$S_I$ ( $\times 10^{-4}$ /min/( $\mu$ IU/ml))	5.65 $\pm$ 3.05	6.69 $\pm$ 2.55	4.59 $\pm$ 2.61	5.32 $\pm$ 3.88	6.07 $\pm$ 3.24
FI ( $\mu$ IU/ml)	12.51 $\pm$ 5.89	10.91 $\pm$ 4.14	13.27 $\pm$ 5.74	14.58 $\pm$ 7.92	10.5 $\pm$ 5.13
HOMA-IR	3.02 $\pm$ 1.51	2.64 $\pm$ 1.06	3.12 $\pm$ 1.41	3.64 $\pm$ 2.11	2.47 $\pm$ 1.25

EA, European American; AA, African American; HA, Hispanic American; AIRg, acute insulin response to glucose,  $S_I$ , insulin sensitivity, FI, fasting insulin; HOMA-IR, Homeostasis model assessment of insulin resistance.

Table 2. Association between European admixture and phenotypes (standardised parameter estimate and *p*-value shown, respectively)

	Entire sample	Hispanic Americans	African Americans	European Americans
Total fat	0.13 ( <i>p</i> = 0.037*)	-0.12 ( <i>p</i> = 0.40)	0.17 ( <i>p</i> = 0.090)	0.02 ( <i>p</i> = 0.80)
S <sub>I</sub>	0.34 ( <i>p</i> < 0.0001*)	0.25 ( <i>p</i> = 0.031*)	-0.01 ( <i>p</i> = 0.93)	0.13 ( <i>p</i> = 0.18)
AIRg	-0.28 ( <i>p</i> < 0.0001*)	-0.23 ( <i>p</i> = 0.038*)	-0.13 ( <i>p</i> = 0.19)	-0.09 ( <i>p</i> = 0.37)
FI	-0.20 ( <i>p</i> = 0.0013*)	-0.27 ( <i>p</i> = 0.045*)	-0.07 ( <i>p</i> = 0.45)	-0.02 ( <i>p</i> = 0.79)
HOMA-IR	-0.16 ( <i>p</i> = 0.0095*)	-0.26 ( <i>p</i> = 0.064)	-0.11 ( <i>p</i> = 0.25)	-0.005 ( <i>p</i> = 0.96)

Covariates: age, sex, Tanner stage, SES, total fat, height (for AIR, we also controlled for S<sub>I</sub>).

\**p* < 0.05.

S<sub>I</sub>, insulin sensitivity; AIRg, acute insulin response to glucose; FI, fasting insulin; HOMA-IR, Homeostasis model assessment of insulin resistance.

(0.18) parental populations. In our sample, the frequency of the G allele is 0.58 among AAs, 0.25 among EAs and 0.21 among HAs. We found no associations that withstand Bonferroni correction (*p* < 3.6 × 10<sup>-4</sup>) within each racial/ethnic group.

Marker rs1373302, located at 8q13, is significantly associated with AIRg under the dominant model (*p* = 9.7 × 10<sup>-5</sup>). Allele T at this marker is associated with increased AIRg and is at a higher frequency among the West African (0.65) and European (0.73) parental populations than among the Amerindian parental population (0.08). We found no significant associations within racial/ethnic groups for this marker after adjusting for Bonferroni correction. Marker rs2671110, located

at 7q32.3, is significantly associated with AIRg among AAs under the recessive model (*p* = 1.4 × 10<sup>-4</sup>). Allele A at this marker is associated with increased AIRg and is at a higher frequency among the West African parental population (0.94) than among the Amerindian (0.0) and European (0.13) parental populations.

Marker rs12439722 is significantly associated with FI and HOMA-IR under the recessive and genotypic models (*p* = 2.8 × 10<sup>-4</sup> and 1.4 × 10<sup>-4</sup>, respectively). Allele A at this marker is associated with increased FI, and is at a higher frequency in the West African (1.0) and European (0.97) parental populations than in the Amerindian (0.17) parental population.

Table 3. Summary of strongest associations by phenotype (*p*-value cut-off with Bonferroni correction: 3.6 × 10<sup>-4</sup>)

Gene	Chr. location (mB)	SNP	Phenotype	Model	Group	Uncorrected <i>p</i> -value
<i>SPTBN1</i>	2; 54.7M	rs3287	S <sub>I</sub>	Recessive	All	5.8 × 10 <sup>-5</sup>
<i>SPTBN1</i>	2; 54.7M	rs3287	S <sub>I</sub>	Genotypic	All	1.9 × 10 <sup>-4</sup>
<i>HERC1</i>	15; 61.7M	rs12439722	S <sub>I</sub>	Genotypic	All	5.4 × 10 <sup>-4</sup>
<i>TRPA1</i>	8; 73.1M	rs1373302	AIRg	Dominant	All	9.7 × 10 <sup>-5</sup>
<i>PLXNA4</i>	7; 131.5M	rs2671110	AIRg	Recessive	AA	1.4 × 10 <sup>-4</sup>
<i>HERC1</i>	15; 61.7M	rs12439722	FI	Recessive	All	2.8 × 10 <sup>-4</sup>
<i>HERC1</i>	15; 61.7M	rs12439722	FI	Genotypic	All	1.4 × 10 <sup>-4</sup>
<i>HERC1</i>	15; 61.7M	rs12439722	HOMA-IR	Recessive	All	8.0 × 10 <sup>-4</sup>
<i>HERC1</i>	15; 61.7M	rs12439722	HOMA-IR	Genotypic	All	4.8 × 10 <sup>-4</sup>

Covariates: age, sex, Tanner stage, socioeconomic status, European admixture, Amerindian admixture, height, total fat (for acute insulin response to glucose, control for S<sub>I</sub> as well). SNP, single nucleotide polymorphism; *SPTBN1*, spectrin, beta, non-erythrocytic 1; *PLXNA4*, plexin A4.

The application of the measurement error correction methods did not yield results that were significantly different to those we observed with the naïve analyses. For example, we observed a  $p$ -value of  $7.68 \times 10^{-5}$  for the association between rs3287 and  $S_I$  under the recessive model after accounting for the measurement error versus  $5.9 \times 10^{-5}$  without the adjustment. This result confirms that individual ancestry proportions are very well measured. Therefore, with 142 AIMs, we can be confident that our results are not driven by measurement errors in the estimation of individual ancestry.

## Discussion

We sought to examine the potential genetic basis for population differences in insulin-related phenotypes in a racially/ethnically diverse sample of children. We found that European genetic admixture is associated with insulin-related phenotypes. Next, we determined whether any of the individual 142 AIMs scattered throughout the genome were associated with any of the insulin-related phenotypes. We found a strong association between  $S_I$  and an AIM at chromosome 2p21 (rs3287), explaining 4.14 per cent of the variance of the trait. Although this effect size may appear large compared with other genetic association studies, our use of refined phenotypes, the inclusion of many covariates and the use of admixed individuals is likely to have increased our ability to detect an effect size of this magnitude. We also found weaker, but statistically significant associations between AIRg and an AIM at chromosome 8q13 (rs1373302) located in the transient receptor potential cation channel, subfamily A, member 1 (*TRPA1*) gene, and between FI and HOMA-IR and an AIM at chromosome 15q22 (rs12439722) in the gene for the hect (homologous to the E6-AP[UBE3A] carboxyl terminus) domain and RCC1 (CHC1)-like domain RLD) 1 (*HERC1*). It should be noted that, although we have used a multiple correction for the 142 markers tested, we have not corrected for each of the genetic models tested. If we were to use a Bonferroni correction

for all markers and models tested, the  $p$ -value threshold would be  $8.8 \times 10^{-5}$ . In this case, only the association between rs3287 and  $S_I$  ( $p = 5.8 \times 10^{-5}$ ) would be considered statistically significant. The four genetic models are likely to be correlated, however, making such a correction overly conservative.

Our finding that insulin-related phenotypes are associated with European admixture is in agreement with previous findings.<sup>17</sup> European admixture is positively associated with favourable insulin-related phenotypes (higher  $S_I$  and lower AIRg, FI and HOMA-IR). When we examine the association of European admixture within racial/ethnic groups, they are only statistically significant among HAs. It is difficult to interpret these results, however, because of the different sample sizes and different admixture proportion distributions by racial/ethnic group. Among the other covariates examined, we find that total body fat and Tanner stage are the strongest risk factors associated with these insulin-related traits. These results are consistent with those of other studies that have shown that adiposity is a major risk factor for these traits.<sup>50,51</sup> Insulin-related traits have also previously been found to be associated with pubertal stage.<sup>52,53</sup>

The 2p21 chromosomal region previously has been identified as being associated with type 2 diabetes and related traits via both linkage scans and GWAS. Marker rs3287 is located between two loci, thyroid adenoma associated (*THADA*) and B-cell chronic lymphocytic leukaemia/lymphoma 11A (*BCL11A*) at 2p21. These loci have been identified previously in two recent GWAS meta-analyses of type 2 diabetes.<sup>21,54</sup> This region was also identified in linkage scans for insulin- and diabetes-related traits.<sup>55,56</sup> It is plausible that, through their effects on cell apoptosis<sup>57</sup> and/or nutrient transport,<sup>58</sup> these loci may be associated with the progression of type 2 diabetes and/or that different pathways may be involved across populations. Given that the occurrence of the rs3287 risk allele is higher in the West African parental population than in the European and Amerindian parental populations, and that AAs tend to have lower  $S_I$ , this or another nearby variant that is in admixture LD may explain

part of the observed differences in type 2 diabetes susceptibility between AAs and EAs. The markers that we have found to be associated with FI and HOMA-IR are in a region on chromosome 15 that previously has been found to be associated with insulin-related traits in a linkage scan.<sup>59</sup>

Unlike other association studies, we have identified these associations relatively early in the lifespan. It could be that the children with unfavourable insulin-related phenotypes are already on the path towards developing type 2 diabetes. In the long-term, these markers could inform prediction and treatment strategies for early-onset type 2 diabetes and explain population differences. The fact that none of the AIMs showed any significant association with any of the insulin-related phenotypes when performing analyses within racial/ethnic groups may be related to reduced power due to a smaller sample size. Furthermore, our ability to find significant associations by race is strongly influenced by the frequency of the variants and the fact that AIM alleles tend to have a low frequency in one group and a high frequency in another group. By using a multi-ethnic approach, we have the advantage of having more intermediate allele frequencies represented, thus increasing the power to detect associations.

This study had several strengths. First, the use of several endo-phenotypes that are likely to be proximal to the development of type 2 diabetes may have pinpointed more effectively the genetic factors that eventually lead to disease phenotypes. Secondly, the inclusion of individuals from different racial/ethnic backgrounds, and the use of markers that differ in frequency between populations, can lead to a better understanding of the genetic basis for population differences in insulin-related phenotypes and the prevalence of diabetes. Thirdly, the inclusion of environmental and phenotypic measurements enhances the ability to pinpoint the genetic regions that directly influence the disease-causing phenotype.

The study also had some limitations. A relatively small number of genetic markers were used, reducing our ability to provide a high level of resolution with regard to the precise location of potential risk

variants. The main weakness of the study was in the small sample size, which raises concerns about the statistical power of the study. Of all the reported associations, the one between marker rs3287 and  $S_1$  will require the greatest level of statistical power in order to reject the null hypothesis. A power calculation for this association model reveals that at a  $p$ -value of  $6 \times 10^{-5}$ , our data provides 67 per cent power to estimate the R-squared effect of 0.45 that we obtained for the full model, with a semi-partial R-squared for the marker of 0.04. Although this level of power might not seem sufficient, the concerns of this association being a type 1 error are dissipated by the fact that it represents a form of replication of several previously reported findings at chromosome 2p21. Evidently, this level of detection with a small sample size was aided by the use of precise phenotyping, the consideration of physiological parameters and the inclusion of admixture estimates, as previously discussed.

In conclusion, we have shown that regions on chromosomes 2, 8 and 15 were associated with insulin-related traits in this sample. These results suggest that these regions may harbour causal variants that may also explain population differences in the insulin-related phenotypes and, ultimately, type 2 diabetes prevalence, since the markers tested exhibit large frequency differences between groups. Future studies must combine detailed phenotypic, environmental and genetic measures on similarly diverse but larger sample sizes. The inclusion of different populations is of paramount importance if we are to understand the genetic basis for population differences and fairly implement effective prevention, intervention and treatment strategies.

## Acknowledgments

We thank the participants of the AMERICO study and their families, and Nicholas Pajewski for help with statistical analyses. Funded by NIH R01-DK067426 and NIH T32HL007457 from the National Heart, Lung, and Blood Institute, Nutrition Obesity Research Center Grant P30-DK56336, UAB Diabetes Research and Training Center Grant P60DK079626, and by General Clinical Research Center Grant M01RR00032.

## References

- Pinhas-Hamiel, O. and Zeitler, P. (2005), 'The global spread of type 2 diabetes mellitus in children and adolescents', *J. Pediatr.* Vol. 146, pp. 693–700.
- Alberti, G., Zimmet, P., Shaw, J., Bloomgarden, Z. et al. (2004), 'Type 2 diabetes in the young: The evolving epidemic. The international diabetes federation consensus workshop', *Diabetes Care* Vol. 27, pp. 1798–1811.
- Narayan, K.M., Boyle, J.P., Thompson, T.J., Sorensen, S.W. et al. (2003), 'Lifetime risk for diabetes mellitus in the United States', *JAMA* Vol. 290, pp. 1884–1890.
- Austin, M.A., Edwards, K.L., McNeely, M.J., Chandler, W.L. et al. (2004), 'Heritability of multivariate factors of the metabolic syndrome in nondiabetic Japanese Americans', *Diabetes* Vol. 53, pp. 1166–1169.
- Beck-Nielsen, H., Vaag, A., Poulsen, P. and Gaster, M. (2003), 'Metabolic and genetic influence on glucose metabolism in type 2 diabetic subjects — Experiences from relatives and twin studies', *Best Pract. Res. Clin. Endocrinol. Metab.* Vol. 17, pp. 445–467.
- Hanson, R.L., Imperatore, G., Narayan, K.M., Roumain, J. et al. (2001), 'Family and genetic studies of indices of insulin sensitivity and insulin secretion in Pima Indians', *Diabetes Metab. Res. Rev.* Vol. 17, pp. 296–303.
- Henkin, L., Bergman, R.N., Bowden, D.W., Ellsworth, D.L. et al. (2003), 'Genetic epidemiology of insulin resistance and visceral adiposity. The IRAS Family Study design, methods', *Ann. Epidemiol.* Vol. 13, pp. 211–217.
- Kaprio, J., Tuomilehto, J., Koskenvuo, M., Romanov, K. et al. (1992), 'Concordance for type 1 (insulin-dependent) and type 2 (non-insulin-dependent) diabetes mellitus in a population-based cohort of twins in Finland', *Diabetologia* Vol. 35, pp. 1060–1067.
- Newman, B., Selby, J.V., King, M.C., Slemenda, C. et al. (1987), 'Concordance for type 2 (non-insulin-dependent) diabetes mellitus in male twins', *Diabetologia* Vol. 30, pp. 763–768.
- Poulsen, P., Kyvik, K.O., Vaag, A. and Beck-Nielsen, H. (1999), 'Heritability of type II (non-insulin-dependent) diabetes mellitus and abnormal glucose tolerance — A population-based twin study', *Diabetologia* Vol. 42, pp. 139–145.
- Lyssenko, V., Jonsson, A., Almgren, P., Pulizzi, N. et al. (2008), 'Clinical risk factors, DNA variants, and the development of type 2 diabetes', *N. Engl. J. Med.* Vol. 359, pp. 2220–2232.
- van Hoek, M., Dehghan, A., Witteman, J.C., van Duijn, C.M. et al. (2008), 'Predicting type 2 diabetes based on polymorphisms from genome-wide association studies: A population-based study', *Diabetes* Vol. 57, pp. 3122–3128.
- Lillioja, S. and Wilton, A. (2009), 'Agreement among type 2 diabetes linkage studies but a poor correlation with results from genome-wide association studies', *Diabetologia* Vol. 52, pp. 1061–1074.
- Gregersen, P.K. (2009), 'Closing the gap between genotype and phenotype', *Nat. Genet.* Vol. 41, pp. 958–959.
- Goran, M.I., Lane, C., Toledo-Corral, C. and Weigensberg, M.J. (2008), 'Persistence of pre-diabetes in overweight and obese Hispanic children: Association with progressive insulin resistance, poor beta-cell function, and increasing visceral fat', *Diabetes* Vol. 57, pp. 3007–3012.
- Ventura, E.E., Lane, C.J., Weigensberg, M.J., Toledo-Corral, C.M. et al. (2009), 'Persistence of the metabolic syndrome over three annual visits in overweight hispanic children: Association with progressive risk for type 2 diabetes', *J. Pediatr.* Vol. 155, pp. 535–541.
- Gower, B.A., Fernandez, J.R., Beasley, T.M., Shriver, M.D. et al. (2003), 'Using genetic admixture to explain racial differences in insulin-related phenotypes', *Diabetes* Vol. 52, pp. 1047–1051.
- Hyatt, T.C., Phadke, R.P., Hunter, G.R., Bush, N.C. et al. (2009), 'Insulin sensitivity in African-American and white women: Association with inflammation', *Obesity* Vol. 17, pp. 276–82.
- Saxena, R., Voight, B.F., Lyssenko, V., Burtt, N.P. et al. (2007), 'Genome-wide association analysis identifies loci for type 2 diabetes and triglyceride levels', *Science* Vol. 316, pp. 1331–1336.
- Sladek, R., Rocheleau, G., Rung, J., Dina, C. et al. (2007), 'A genome-wide association study identifies novel risk loci for type 2 diabetes', *Nature* Vol. 445, pp. 881–885.
- Zeggini, E., Weedon, M.N., Lindgren, C.M., Frayling, T.M. et al. (2007), 'Replication of genome-wide association signals in UK samples reveals risk loci for type 2 diabetes', *Science* Vol. 316, pp. 1336–1341.
- Harris, M.I., Flegal, K.M., Cowie, C.C., Eberhardt, M.S. et al. (1998), 'Prevalence of diabetes, impaired fasting glucose, and impaired glucose tolerance in U.S. adults. The Third National Health and Nutrition Examination Survey, 1988–1994', *Diabetes Care* Vol. 21, p. 518–524.
- Permutt, M.A., Wasson, J. and Cox, N. (2005), 'Genetic epidemiology of diabetes', *J. Clin. Invest.* Vol. 115, pp. 1431–1439.
- Haffner, S.M., D'Agostino, R., Saad, M.F., Rewers, M. et al. (1996), 'Increased insulin resistance and insulin secretion in nondiabetic African-Americans and Hispanics compared with non-Hispanic whites. The Insulin Resistance Atherosclerosis Study', *Diabetes* Vol. 45, pp. 742–748.
- Weiss, R., Dufour, S., Taksali, S.E., Tamborlane, W.V. et al. (2003), 'Prediabetes in obese youth: A syndrome of impaired glucose tolerance, severe insulin resistance, and altered myocellular and abdominal fat partitioning', *Lancet* Vol. 362, pp. 951–957.
- Svec, F., Nastasi, K., Hilton, C., Bao, W. et al. (1992), 'Black-white contrasts in insulin levels during pubertal development. The Bogalusa Heart Study', *Diabetes* Vol. 41, pp. 313–317.
- Arslanian, S., Suprasongsin, C. and Janosky, J.E. (1997), 'Insulin secretion and sensitivity in black versus white prepubertal healthy children', *J. Clin. Endocrinol. Metab.* Vol. 82, pp. 1923–1927.
- Gower, B.A., Granger, W.M., Franklin, E., Shewchuk, R.M. et al. (2002), 'Contribution of insulin secretion and clearance to glucose-induced insulin concentration in African-American and Caucasian children', *J. Clin. Endocrinol. Metab.* Vol. 87, pp. 2218–2224.
- Halder, I. and Shriver, M.D. (2003), 'Measuring and using admixture to study the genetics of complex diseases', *Hum. Genomics* Vol. 1, pp. 52–62.
- Williams, R.C., Long, J.C., Hanson, R.L., Sievers, M.L. et al. (2000), 'Individual estimates of European genetic admixture associated with lower body-mass index, plasma glucose, and prevalence of type 2 diabetes in Pima Indians', *Am. J. Hum. Genet.* Vol. 66, pp. 527–538.
- Parra, E.J., Hoggart, C.J., Bonilla, C., Dios, S. et al. (2004), 'Relation of type 2 diabetes to individual admixture and candidate gene polymorphisms in the Hispanic American population of San Luis Valley, Colorado', *J. Med. Genet.* Vol. 41, p. e116.
- Florez, J.C., Price, A.L., Campbell, D., Riba, L. et al. (2009), 'Strong association of socioeconomic status with genetic ancestry in Latinos: Implications for admixture studies of type 2 diabetes', *Diabetologia* Vol. 52, pp. 1528–1536.
- McKeigue, P.M. (2005), 'Prospects for admixture mapping of complex traits', *Am. J. Hum. Genet.* Vol. 76, pp. 1–7.
- Elbein, S.C., Das, S.K., Hallman, D.M., Hais, C.L. et al. (2009), 'Genome-wide linkage and admixture mapping of type 2 diabetes in African American families from the American Diabetes Association GENNID (Genetics of NIDDM) Study Cohort', *Diabetes* Vol. 58, pp. 268–274.
- Marshall, W.A. and Tanner, J.M. (1968), 'Growth and physiological development during adolescence', *Annu. Rev. Med.* Vol. 19, pp. 283–300.
- Goran, M.I., Driscoll, P., Johnson, R., Nagy, T.R. et al. (1996), 'Cross-calibration of body-composition techniques against dual-energy X-ray absorptiometry in young children', *Am. J. Clin. Nutr.* Vol. 63, pp. 299–305.
- Matthews, D.R., Edge, J.A. and Dunger, D.B. (1990), 'An unbiased glucose clamp method using a variable insulin infusion: Its application in diabetic adolescents', *Diabet. Med.* Vol. 7, pp. 246–251.
- Pacini, G. and Bergman, R.N. (1986), 'MINMOD: A computer program to calculate insulin sensitivity and pancreatic responsiveness from the frequently sampled intravenous glucose tolerance test', *Comput. Methods Programs Biomed.* Vol. 23, pp. 113–122.
- Casazza, K., Phadke, R.P., Fernandez, J.R., Watanabe, R.M. et al. (2009), 'Obesity attenuates the contribution of African admixture to the insulin secretory profile in peripubertal children: A longitudinal analysis', *Obesity* Vol. 17, pp. 1318–1325.



40. Bergman, R.N., Ider, Y.Z., Bowden, C.R. and Cobelli, C. (1979), 'Quantitative estimation of insulin sensitivity', *Am. J. Physiol.* Vol. 236, pp. E667–E677.
41. Matthews, D.R., Hosker, J.P., Rudenski, A.S., Naylor, B.A. *et al.* (1985), 'Homeostasis model assessment: Insulin resistance and beta-cell function from fasting plasma glucose and insulin concentrations in man', *Diabetologia* Vol. 28, pp. 412–419.
42. Hollingshead, A.B. (1975), *Four Factor Index of Social Class*, Department of Sociology, Yale University, New Haven, CT.
43. Hawkins, J.R., Khripin, Y., Valdes, A.M. and Weaver, T.A. (2002), 'Miniaturized sealed-tube allele-specific PCR', *Hum. Mutat.* Vol. 19, pp. 543–553.
44. Halder, I., Shriver, M., Thomas, M., Fernandez, J.R. *et al.* (2008), 'A panel of ancestry informative markers for estimating individual biogeographical ancestry and admixture from four continents: Utility and applications', *Hum. Mutat.* Vol. 29, pp. 648–658.
45. Hanis, C.L., Chakraborty, R., Ferrell, R.E. and Schull, W.J. (1986), 'Individual admixture estimates: Disease associations and individual risk of diabetes and gallbladder disease among Mexican-Americans in Starr County, Texas', *Am. J. Phys. Anthropol.* Vol. 70, pp. 433–441.
46. Divers, J., Vaughan, L.K., Padilla, M.A., Fernandez, J.R. *et al.* (2007), 'Correcting for measurement error in individual ancestry estimates in structured association tests', *Genetics* Vol. 176, pp. 1823–1833.
47. Carroll, R.J., Rupper, D., Stefanski, L.A. and Crainiceanu, C.M. (2006), *Measurement Error in Nonlinear Models: A Modern Perspective* (2nd edn), Chapman & Hall/CRC, Boca Raton, FL.
48. Purcell, S., Neale, B., Todd-Brown, K., Thomas, L. *et al.* (2007), 'PLINK: A tool set for whole-genome association and population-based linkage analyses', *Am. J. Hum. Genet.* Vol. 81, pp. 559–575.
49. R Foundation for Statistical Computing (2009), *R: A Language and Environment for Statistical Computing*, Vienna, Austria.
50. Shaibi, G.Q. and Goran, M.I. (2008), 'Examining metabolic syndrome definitions in overweight Hispanic youth: A focus on insulin resistance', *J. Pediatr.* Vol. 152, pp. 171–176.
51. Levy-Marchal, C., Arslanian, S., Cutfield, W., Sinaiko, A. *et al.* (2010), 'Insulin resistance in children: Consensus, perspective, and future directions', *J. Clin. Endocrinol. Metab.*, Sept. 8, epub ahead of print, DOI: 10.1210/jc.2010.1047.
52. Moran, A., Jacobs, D.R., Jr, Steinberger, J., Hong, C.P. *et al.* (1999), 'Insulin resistance during puberty: Results from clamp studies in 357 children', *Diabetes* Vol. 48, pp. 2039–2044.
53. Ball, G.D., Huang, T.T., Gower, B.A., Cruz, M.L. *et al.* (2006), 'Longitudinal changes in insulin sensitivity, insulin secretion, and beta-cell function during puberty', *J. Pediatr.* Vol. 148, pp. 16–22.
54. Voight, B.F., Scott, L.J., Steinthorsdottir, V., Morris, A.P. *et al.* (2010), 'Twelve type 2 diabetes susceptibility loci identified through large-scale association analysis', *Nat. Genet.* Vol. 42, pp. 579–589.
55. Diego, V.P., Goring, H.H., Cole, S.A., Almasy, L. *et al.* (2006), 'Fasting insulin and obesity-related phenotypes are linked to chromosome 2p: The Strong Heart Family Study', *Diabetes* Vol. 55, pp. 1874–1878.
56. An, P., Teran-Garcia, M., Rice, T., Rankinen, T. *et al.* (2005), 'Genome-wide linkage scans for prediabetes phenotypes in response to 20 weeks of endurance exercise training in non-diabetic whites and blacks: The HERITAGE Family Study', *Diabetologia* Vol. 48, pp. 1142–1149.
57. Drieschner, N., Kerschling, S., Soller, J.T., Rippe, V. *et al.* (2007), 'A domain of the thyroid adenoma associated gene (THADA) conserved in vertebrates becomes destroyed by chromosomal rearrangements observed in thyroid adenomas', *Gene* Vol. 403, pp. 110–117.
58. Lettre, G., Sankaran, V.G., Bezerra, M.A., Araujo, A.S. *et al.* (2008), 'DNA polymorphisms at the BCL11A, HBS1L-MYB, and beta-globin loci associate with fetal hemoglobin levels and pain crises in sickle cell disease', *Proc. Natl. Acad. Sci. USA* Vol. 105, pp. 11869–11874.
59. Hsueh, W.C., Silver, K.D., Pollin, T.I., Bell, C.J. *et al.* (2007), 'A genome-wide linkage scan of insulin level derived traits: The Amish Family Diabetes Study', *Diabetes* Vol. 56, pp. 2643–2648.

Table S1. Parental population allele frequencies of ancestry-informative markers

rs#	West African	European	Amerindian
285	0.971	0.494	0.442
2161	0.430	0.295	0.645
2695	0.805	0.845	0.132
2752	0.878	0.417	0.626
2763	0.129	0.185	0.631
2816	1.000	0.518	0.965
2862	0.386	0.201	0.740
3287	0.751	0.180	0.181
3309	0.401	0.300	0.707
3340	0.054	0.197	0.696
4646	0.332	0.287	0.714
6003	0.704	0.104	0.017
16383	0.261	0.780	0.052
17203	0.785	0.172	0.805
140864	0.124	0.007	0.586
174548	0.792	0.765	0.040
222541	0.036	0.743	0.982
275183	0.882	0.065	0.052
292932	1.000	1.000	0.273
344461	0.000	0.867	0.969
345162	0.102	0.866	0.810
386569	0.105	0.188	1.000
463717	0.894	0.360	0.765
594689	0.087	0.440	0.106
602662	0.418	0.408	0.050
624328	0.805	0.038	0.019
638169	0.851	0.040	0.000
717091	0.208	0.832	0.688
717962	0.868	0.825	0.125

Continued

Table S1. Continued

rs#	West African	European	Amerindian
718092	0.764	0.213	0.820
719776	0.974	0.150	0.250
723822	0.800	0.906	0.132
724729	0.107	0.842	0.974
725416	0.868	0.050	0.075
725667	0.816	0.075	0.000
726391	0.947	0.222	0.632
732381	0.908	0.714	0.357
734780	0.291	0.930	0.146
764679	0.789	0.975	0.263
768324	0.816	0.925	0.063
832173	0.889	0.179	0.207
878825	0.563	0.909	0.136
905595	0.750	0.192	0.000
938431	0.400	0.200	0.947
951554	0.711	0.850	0.200
983271	0.588	0.050	0.833
1003645	0.955	0.199	0.172
1008984	0.693	0.119	0.724
1042602	0.005	0.485	0.029
1044498	0.052	0.900	0.837
1058115	0.404	0.360	0.406
1074075	0.895	0.175	0.025
1112828	0.940	0.775	0.125
1229984	0.011	0.073	0.021
1316579	0.763	0.250	0.975
1320149	0.111	0.910	0.983
1320892	0.842	0.789	0.000
1327805	0.632	0.200	0.917
1352158	0.716	0.195	0.897

Continued

Table S1. Continued

rs#	West African	European	Amerindian
1369290	0.921	0.100	0.000
1373302	0.278	0.188	1.000
1403454	1.000	0.868	0.150
1415878	1.000	0.917	0.025
1435090	0.204	0.318	0.847
1461227	0.395	0.111	0.825
1465648	0.079	0.900	0.875
1487214	0.833	0.125	0.050
1506069	0.026	0.950	0.909
1526028	0.184	1.000	0.975
1594335	0.801	0.251	0.236
1780349	0.530	1.000	1.000
1800498	0.135	0.630	0.077
1808089	0.395	0.450	1.000
1861498	0.053	0.861	1.000
1881826	0.206	0.125	0.784
1891760	0.803	0.563	0.066
1935946	0.105	0.725	0.000
1980888	0.804	0.892	0.052
1986586	0.423	1.000	0.000
2045431	0.828	0.164	0.603
2077681	0.105	0.925	0.933
2078588	0.974	0.075	0.125
2118350	0.359	0.825	0.276
2157257	0.915	0.367	0.300
2188457	0.268	0.861	0.943
2207782	0.029	0.750	0.000
2216595	0.775	0.184	0.808
2225251	0.026	0.778	0.211
2228478	0.513	0.108	0.035

Continued

Table S1. Continued

rs#	West African	European	Amerindian
2284708	0.071	0.840	0.643
2292549	0.740	0.028	0.000
2340727	0.891	0.147	0.034
2341823	0.875	0.079	0.400
2351254	0.667	0.053	0.778
2396676	0.947	0.250	0.275
2426512	0.045	0.740	0.914
2542481	0.429	0.918	0.612
2549156	0.957	0.806	0.633
2671110	0.938	0.133	0.000
2710663	0.073	0.214	0.890
2776937	0.092	0.792	0.810
2789823	0.104	1.000	1.000
2806424	0.842	0.292	0.690
2814778	0.001	0.999	0.999
2877967	0.885	0.143	0.210
3123101	0.612	0.685	0.219
3751631	0.888	0.135	0.582
3760281	0.902	0.109	0.655
3782972	0.947	0.780	0.323
3827760	0.000	0.000	0.927
4676223	0.061	0.714	0.120
4717627	0.681	0.694	0.160
7163836	0.480	0.610	0.040
7211872	0.122	0.990	1.000

Table S1.

rs#	West African	European	Amerindian
7246566	0.000	0.576	0.030
7531501	0.978	0.092	0.660
7786541	0.391	0.000	0.010
7795585	0.082	0.920	0.970
7854707	0.969	0.110	0.031
8007610	0.204	0.940	0.760
8072587	0.969	0.214	0.720
9542741	0.053	0.840	0.460
9809114	0.500	0.714	0.173
9847307	0.083	0.000	0.865
10483251	0.854	0.833	0.112
10752631	0.842	0.373	0.000
10756789	1.000	0.327	0.920
11004105	0.181	0.240	0.310
11160369	0.071	0.030	0.560
12277775	0.677	0.870	0.240
12439722	1.000	0.970	0.167
12459172	0.112	0.200	0.850
12598094	0.939	0.940	0.260
12647878	0.990	0.950	0.550
12692384	0.776	0.720	0.030
12737539	0.120	0.600	0.000
16834947	0.628	0.010	0.587
16891982	1.000	0.062	0.980
17822931	1.000	0.888	0.770