Polymorphisms of the Scavenger Receptor Class B Member 1 Are Associated with Insulin Resistance with Evidence of Gene by Sex Interaction

Jeanette J. McCarthy, Aleefia Somji, Lauren A. Weiss, Brian Steffy, Raquel Vega, Elizabeth Barrett-Connor, Gregory Talavera, and Richard Glynne

Graduate School of Public Health (J.J.M., G.T.), San Diego State University, San Diego, California 92182; Institute for Genome Sciences and Policy (J.J.M., A.S.), Duke University Medical Center, Durham, North Carolina 27710; and Department of Family and Preventive Medicine (L.A.W., E.B.-C.), University of California at San Diego, and Genomics Institute of the Novartis Research Foundation (B.S., R.V., R.G.), San Diego, California 92121

Background: Genetic variation in diabetes-associated genes cumulatively explain little of the overall heritability of this trait. We sought to determine whether polymorphisms of the scavenger receptor class B, member I (SCARB1), an estrogen-regulated chromosome 12q24 positional candidate diabetes gene, were associated with type 2 diabetes or insulin resistance in a sex-specific fashion.

Methods: We evaluated 34 haplotype-tagged single-nucleotide polymorphisms (SNPs) of SCARB1 for their association with type 2 diabetes and measures of insulin resistance in two populations: a clinic-based sample of 444 Mexican-American women from Proyecto SALSA and a community-based sample of 830 white women from the Rancho Bernardo Study.

Results: We identified significant associations between a tagged SNP in intron 9, rs9919713, and fasting glucose in the SALSA population ($P = 2.3 \times 10^{-4}$). In the Rancho Bernardo Study, the same SNP also showed significant association with the related traits homeostasis model assessment for insulin resistance ($P = 3.0 \times 10^{-1}$), fasting glucose ($P = 1.1 \times 10^{-3}$), and type 2 diabetes ($P = 9.0 \times 10^{-3}$). In men from the Rancho Bernardo population, the opposite effect was found (genotype by sex interaction in the Rancho Bernardo population $P < 10^{-3}$ for insulin resistance).

Conclusions: Our data support an association between SCARB1 variants and insulin resistance, especially in women, with evidence of significant gene by sex interaction. These findings warrant further investigation in additional populations and prompt exploration of a role for SR-BI in the development of insulin resistance. (J Clin Endocrinol Metab 94: 1789–1796, 2009)

Type 2 diabetes mellitus has emerged as one of the most prevalent chronic diseases worldwide. In the United States alone, approximately 16 million people have clinically manifest type 2 diabetes, whereas another 13 million have impaired glucose tolerance (1). Mexican-Americans are disproportionately affected, with twice the prevalence observed in non-Hispanic whites (10.4 and 5.2%, respectively) (1). Genome-wide association studies have uncovered a number of loci associated with type 2 diabetes, but together they account for little of the overall variance and provide little predictive power over traditional risk factors (2). It is likely that many loci still remain to be discovered.

Genes located in chromosomal regions showing linkage to type 2 diabetes in family-based studies are rational candidates for more detailed investigation. One such gene, the scavenger receptor class B, member I (SR-BI; gene name SCARB1) lies in a region on chromosome 12q24 that has been linked to type 2 diabetes in numerous studies (3) and more recently to abdominal obesity, a risk factor for type 2 diabetes, in three studies (4–6). SR-BI is a high-density lipoprotein (HDL) receptor involved in reverse cho-

Abbreviations: BMI, Body mass index; HDL, high-density lipoprotein; HOMA-IR, homeostasis model assessment of insulin resistance; ht, haplotype tagged; SNP, single-nucleotide polymorphism; SR-BI, scavenger receptor class B type 1.
lesterol transport. Direct evidence for association of SCARB1 polymorphisms related to diabetes traits comes from one small intervention study that found insulin sensitivity to dietary fat associated with a SCARB1 variant in exon 1 (7). Genetic variation in the SCARB1 gene has also been associated with increased risk of coronary artery disease (8), obesity (9, 10), triglycerides (10–13), and HDL-cholesterol (10–18), all facets of the metabolic syndrome. Furthermore, there is evidence that diabetes status may modify the SCARB1 association with HDL-cholesterol (19). A striking feature of many of these reported associations is the presence of gene by sex interaction, whereby the genetic variants behave very differently in men and women, often showing the complete opposite effect and only becoming apparent when analyzing the sexes separately (10–14, 17, 18).

Although genome-wide significance has not been demonstrated in published whole-genome association studies of type 2 diabetes for SCARB1, none have examined sex-specific effects. The purpose of this study was to test the association between SCARB1 polymorphisms and type 2 diabetes and insulin resistance in two cohorts of women from San Diego County: a clinic-based cohort of Mexican-Americans from Proyecto SALSA and a community-based white cohort from the Rancho Bernardo Study.

### Subjects and Methods

#### Mexican-American study population

Subjects were drawn from Proyecto SALSA, a cross-sectional clinic-based study of gene-environment interaction in the development of the metabolic syndrome.
metabolic syndrome among Latinos from San Diego County, conducted from 2003–2006 and described elsewhere (20). Because not all subjects had extensive medical records available, diabetes status was determined primarily using patient self-report at enrollment. However, subjects reporting no history of type 2 diabetes but having elevated fasting plasma glucose (≥7.0 mmol/liter) measured from fasting blood taken on the day of enrollment were also classified as having type 2 diabetes. Of the 742 subjects enrolled between April 2004 and April 2006, 646 had sufficient quality and quantity of DNA available for analysis and were genotyped for SCARB1 polymorphisms. Five people were removed for unknown diabetes status, and 37 subjects became available only near the end of our investigation. Because sex-specific effects of SCARB1 alleles were noted previously, the present study was limited to the 444 women who comprised about 75% of the study population. Genetic analysis of fasting glucose was further limited to subjects without a history of diabetes and not taking diabetes medication.

Rancho Bernardo Study

The Rancho Bernardo Study is a population-based cohort of white men and women from Rancho Bernardo, CA, described in detail elsewhere (21). The present study uses clinical data and laboratory measures collected between 1984 and 1987 during a follow-up visit focusing on diabetes and cardiovascular disease (visit 4). Morning blood samples were obtained after an overnight fast and again 2 h after a standard 75-g oral glucose tolerance test. Based on 1999 World Health Organization (WHO) criteria, individuals were classified as having type 2 diabetes if they met one of the following: fasting plasma glucose 7.0 mmol/liter (126 mg/dl) or higher, postchallenge glucose 11.1 mmol/liter (200 mg/dl) or higher, taking oral medication for diabetes in the last 2 wk, or having been told by a physician that they had diabetes. We also included in our analysis any incident cases of diabetes based on physician diagnosis and or fasting hyperglycemia that developed between visit 4 (1984–1987) and visit 7 (1992–1995). The homeostasis model assessment for insulin resistance (HOMA-IR) was calculated as: [fasting glucose (millimoles per liter) × fasting insulin (microunits per milliliter)]/22.5 from measures taken at visit 4. For the analysis of HOMA-IR and fasting glucose, subjects with extreme outlier measures of glucose, insulin, or HOMA-IR, those taking antidiabetic agents or insulin or diagnosed with type 2 diabetes at the

### Table 2. Baseline characteristics of Mexican-American women from Proyecto SALSA and non-Hispanic white men and women from the Rancho Bernardo Study

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Rancho Bernardo men (n = 552)</th>
<th>Rancho Bernardo women (n = 830)</th>
<th>Proyecto SALSA women (n = 444)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Type 2 diabetes mellitus (%)</td>
<td>18</td>
<td>14</td>
<td>36</td>
</tr>
<tr>
<td>Age (yr)</td>
<td>71.5 ± 11.0</td>
<td>69.4 ± 8.9</td>
<td>53.6 ± 13.4</td>
</tr>
<tr>
<td>Age range (yr)</td>
<td>42–91</td>
<td>38–89</td>
<td>18–87</td>
</tr>
<tr>
<td>Waist circumference (cm)</td>
<td>95.0 ± 11.0</td>
<td>77.9 ± 9.2</td>
<td>98.2 ± 13.7</td>
</tr>
<tr>
<td>BMI (kg/m²)</td>
<td>26.4 ± 3.7</td>
<td>24.0 ± 3.6</td>
<td>30.6 ± 5.9</td>
</tr>
<tr>
<td>HDL-cholesterol (mmol/liter)</td>
<td>1.25 ± 0.34</td>
<td>1.81 ± 0.49</td>
<td>1.39 ± 0.32</td>
</tr>
<tr>
<td>Triglycerides (mmol/liter)a</td>
<td>1.19 (0.79–1.72)</td>
<td>1.12 (0.80–1.63)</td>
<td>1.47 (1.11–2.08)</td>
</tr>
<tr>
<td>Fasting plasma glucose (mmol/liter) a</td>
<td>5.38 (5.06–5.82)</td>
<td>5.12 (4.85–5.47)</td>
<td>5.16 (4.77–5.49)</td>
</tr>
<tr>
<td>HOMA-IRa</td>
<td>2.33 (1.67–3.33)</td>
<td>2.09 (1.51–2.85)</td>
<td>Not available</td>
</tr>
</tbody>
</table>

Results are mean ± SD unless noted otherwise.

---

**FIG. 1.** Association between SCARB1 ht-SNPs and fasting glucose, HOMA, and type 2 diabetes (T2DM) in women from Proyecto SALSA (top) and the Rancho Bernardo Study (bottom). P values are for testing association between each SNP in an additive, age-adjusted model. Arrows indicate association with rs9919713 SNP in each population.
time were removed and the variables natural log (ln)-transformed before analysis. Stored DNA was used to genotype SCARB1 variants in 869 women. We excluded two participants with type 1 diabetes and 35 others with indeterminate diabetes status, leaving a total of 830 women, 118 with type 2 diabetes (74 prevalent, 44 incident) and 712 without diabetes, for these analyses. In follow-up analyses, we included data from 552 men drawn from this same population (115 with type 2 diabetes and 437 without) and analyzed in the same manner as the women.

Selection of single-nucleotide polymorphisms (SNPs)

The gene encoding SR-BI (SCARB1; alias CD36L1, SRB1, CLA-1, and CLA1) is located on chromosome 12q24.1. The coordinates (123,785,056–123,878,214) are based on NCBI Build 35 and were used to denote nucleotide positions throughout the manuscript. The Tagger program in Haplovew software version 3.2 (Broad Institute, Cambridge, MA) was used to select 34 haplotype-tagged (ht)-SNPs in the SCARB1 gene having minor allele frequencies greater than 0.05 and capturing all alleles with correlations ($r^2$) greater than 0.80 among Caucasians. To control for admixture, we genotyped an additional 30 ancestry-informative markers specific for Mexican-Americans, spread throughout the genome and selected from among those published in Choudhry et al. (22).

Genotyping

Genotyping of individual DNAs was carried out using either the Sequenom iPLEX multiplex mass spectrometry genotyping system (Sequenom, Inc., La Jolla, CA) or using the 5′-nuclease assay with allele-specific TaqMan probes. The percent missing genotypes (supplemental Table S1, published as supplemental data on The Endocrine Society’s Journals Online web site at http://jcem.endojournals.org) and Hardy-Weinberg equilibrium were evaluated using Haplovew software as a means of quality control. Haplovew software (version 4.0) was used to generate linkage disequilibrium plots using the $r^2$ statistic and defining blocks by the Gabriel method (23). Table 1 summarizes the SCARB1 ht-SNPs evaluated in this study.

Statistical analysis

SAS statistical software, version 9.1 (SAS Institute, Cary, NC) was used to carry out linear and logistic regression analysis of SCARB1 genotypes associated with diabetes traits. Genotypes were coded to test an additive (one degree of freedom) model, but a dominant model was also tested (by pooling homozygous variant and heterozygous genotypes) in subsequent analyses where the variant allele frequency was low (<5%). The current sample sizes were determined to have sufficient (80%) power to detect differences in mean natural log (ln) traits of 0.07 (SALSA) and 0.055 (Rancho Bernardo) and odds ratios of 1.4 (SALSA) and 1.7 (Rancho Bernardo) for the association with diabetes for our top SNP. All diabetes trait association models included age as a covariate. Additional models included body mass index (BMI) and HDL-cholesterol as covariates, both expressed as continuous variables and their interaction terms introduced one at a time. To assess gene by sex interaction, we genotyped men from the Rancho Bernardo Study and ran a regression model in the aggregate population of Rancho Bernardo men and women, without diabetes using a $t$ test and included the proportion Caucasian ancestry, generated from the STRUCTURE program, as a continuous covariate in multiple regression analysis to control for population stratification.

This study was approved by the Institutional Review Boards at both San Diego State University and Duke University.

Results

Characteristics of the study populations are shown in Table 2. The overall prevalence of type 2 diabetes was 36% in women from the SALSA population and in the Rancho Bernardo Study was 21% in men and 14% in women, including cases existing at baseline as well as incident cases occurring over approximately the next 20 yr. SALSA women had higher BMI, waist circumference, and triglycerides and lower HDL-cholesterol levels than Rancho Bernardo women, despite being an average of 16 yr younger. These differences are not unexpected, given that the Rancho Bernardo and SALSA studies differed substantially with respect to study design (community-based vs. clinic-based), time period of ascertainment (1980s vs. 2000s), and demographics of the subjects (relatively affluent, white, older women vs. younger, low-socioeconomic-status Hispanic women).

Single locus allelic association

Of the 34 ht-SNPs selected for this study (Table 1), several SNPs had nominally significant ($0.01 < P < 0.05$) deviations from Hardy-Weinberg equilibrium, and these were flagged. Figure 1 shows the $-\log (10)$ $P$ values for age-adjusted SCARB1 allelic associations with the discrete trait of type 2 diabetes and fasting plasma glucose (A) and HOMA-IR (B) and proportion of subjects with type 2 diabetes mellitus (C) by SCARB1 rs9919713 minor allele status (0, 1, or 2 copies) in females from the Rancho Bernardo (RB) and SALSA studies.
the quantitative traits of fasting plasma glucose and HOMA-IR in women from both populations. Several significant associations were found, the most robust being with SNP rs9919713, which showed a significant association in the same direction with fasting glucose in the SALSA study (P = 0.0002) and with HOMA-IR (P = 0.0003), fasting glucose (P = 0.001), and WHO-defined type 2 diabetes (P = 0.002) in the Rancho Bernardo Study (Fig. 2). All but the association with diabetes in the Rancho Bernardo Study remained significant after applying a Bonferroni correction for multiple testing considering the 34 ht-SNPs (corrected P = 0.007 for fasting glucose in SALSA; P = 0.03 and P = 0.01 for fasting glucose and HOMA-IR, respectively, in the Rancho Bernardo Study). In addition, several other loci showed significant associations, but no other SNP was consistently replicated across populations and traits. Based on the consistency across traits and studies, we chose to pursue the locus tagged by rs9919713 further and plotted the associated interval around rs9919713, using linkage disequilibrium measures (r²) generated from the HapMap Caucasian population data. This SNP lies on an approximately 6-kb block from NCBI Build 35 nucleotide positions 123,787,056–123,873,214, spanning intron 9 to intron 11 of the SCARB1 gene (Fig. 3).

Analysis of additional SNPs in the associated interval

Five additional SNPs in and around this block were selected from the dbSNP database and typed in the Rancho Bernardo and SALSA women. Several SNPs in the interval were strongly associated with fasting glucose in the SALSA population and with fasting glucose, HOMA-IR, and type 2 diabetes in the Rancho Bernardo women (Fig. 4). In both populations, presence of the minor allele was associated with increased fasting plasma glucose and in the Rancho Bernardo Study with increased HOMA-IR as well. The total variance in fasting glucose and HOMA-IR explained by this locus was less than 5%. The variant allele of rs9919713 was also associated with type 2 diabetes in the Rancho Bernardo Study, conferring a 3-fold increased odds (3.04; 95% confidence interval = 3.53–6.06), with minor allele frequencies of 0.078 in cases and 0.036 in controls.

Analysis of confounding and interaction

Because Hispanic populations are highly admixed, we considered the possibility of a spurious association due to population stratification in the SALSA population. Analysis of ancestry-informative markers using the STRUCTURE program found the most parsimonious model to consist of two underlying population clusters. The proportion of Caucasian ancestry was only slightly lower in diabetes cases compared with controls, but this difference was not statistically significant (P = 0.38). After controlling for proportion of Caucasian ancestry in the association analysis of SCARB1 SNPs and fasting glucose, the associations remained virtually unchanged. To determine whether the associations observed in the Rancho Bernardo Study could be attributed to Mexican-American admixture within that population, we considered five other ht-SNPs within the SCARB1 gene whose allele frequencies differed by more than 2-fold between the Rancho Bernardo Caucasians and SALSA Mexican-Americans. Three of these were at least nominally significantly associated with either fasting glucose or diabetes in the Rancho Bernardo Study but did not alter the association with rs9919713 when entered into the same model.

We also investigated whether associations with type 2 diabetes traits were mediated through other factors previously associated with polymorphisms in SCARB1: HDL-cholesterol levels and BMI. Controlling for these two covariates resulted in a partial attenuation of effect for most SNPs, but all SNPs that had been significant previously remained significant (not shown). For example, the association between rs9919713 and fasting glucose remained significant in SALSA women (P = 1 × 10⁻³) and Rancho Bernardo women (P = 2 × 10⁻³).

Finally, we sought to determine whether the association between rs9919713 and insulin resistance and type 2 diabetes was also observed in men from the Rancho Bernardo study and found a nonsignificant association in the opposite direction (fasting glucose P = 0.06; HOMA-IR P = 0.19; type 2 diabetes P = 0.64; Fig. 5), which became significant for glucose and HOMA-IR after controlling for BMI and HDL (fasting glu-
cose $P = 0.01$; HOMA-IR $P = 0.05$). In the aggregate population, there was no evidence of association with fasting glucose ($P = 0.31$) or HOMA-IR ($P = 0.07$) and only a nominally significant association with type 2 diabetes ($P = 0.04$), reflecting the contribution of the 60% women in the cohort. However, there was strong evidence of genotype by sex interaction for fasting glucose ($P = 2 \times 10^{-4}$), HOMA-IR ($P = 6 \times 10^{-4}$), and type 2 diabetes ($P = 0.025$). There were too few men available in the SALSA population to do a meaningful analysis ($n = 68$). We found no evidence of interaction between rs9919713 and BMI or HDL-cholesterol in women from any population (all $P > 0.10$).

**Discussion**

The present study provides evidence for an association between genetic variation in SCARB1 and surrogate measures of insulin resistance, observed most strongly in women and characterized by significant gene-sex interaction. Consistent associations with several polymorphisms in a block spanning intron 9 to intron 11 of the SCARB1 gene were observed across two diverse study populations: a clinic-based population of Mexican-American women and a population-based sample of white women from the Rancho Bernardo Study. These associations were identified using a candidate gene approach applying ht-SNPs across the SCARB1 gene, with careful consideration of previously described sex-specific effects of this locus. Previous candidate gene studies of the SCARB1 gene have mostly been limited to a few candidate SNPs in exon 1, intron 5, and exon 8; none have taken a comprehensive ht-SNP approach.

The possibility that the observed association is spurious is minimized when considering the highly significant $P$ values, the replication across one Caucasian and one Hispanic population, and the consistency across traits. Both studies showed significant association between the SCARB1 SNP and glucose levels with the same allele and in the same direction. Although the Rancho Bernardo study additionally showed a 3-fold increased odds of diabetes associated with the SCARB1 risk allele, this was not the case for the SALSA study. This difference in the risk of diabetes despite elevated glucose might reflect variance in other genetic factors interacting with SCARB1. Alternatively, because of the

**FIG. 4.** Association between SCARB1 SNPs in the rs9919713 associated interval and fasting glucose, HOMA, and type 2 diabetes (T2DM) in the SALSA and Rancho Bernardo populations. Due to a low minor allele frequency, the SNP was analyzed in a dominant model (pooled homozygous variant and heterozygous genotypes) in the Rancho Bernardo population.

**FIG. 5.** Mean (±SE, bars) age-adjusted values of natural log (ln)-transformed fasting plasma glucose by SCARB1 rs9919713 minor allele carrier status in males and females from the Rancho Bernardo Study.
high prevalence of metabolic syndrome, it could simply reflect misclassification of diabetes status in the SALSA study. Large differences in allele frequencies between ethnic groups make genetic associations with SCARB1 SNPs particularly susceptible to confounding by population stratification. Although controlling ethnicity with ancestry-informative markers had a negligible effect on the association of SCARB1 and insulin resistance in the SALSA population, cryptic stratification could remain.

SCARB1 has failed to emerge as a diabetes locus in any published genome-wide association study to date, perhaps a consequence of not taking into account gene by sex interaction or more directly the interaction between genes and sex hormones, which vary not only by sex but also by age. It has become increasingly evident over the past year that discoveries made in genome-wide association studies fail to explain more than a small fraction of known heritability, nor have such studies consistently rediscovered loci that have been previously found through family studies for reasons that remain unclear (24). We propose that consideration of gene by sex (and age) interaction may reveal additional loci relevant to diabetes.

It is noteworthy that in the Rancho Bernardo study, the associations were found in the opposite direction in men and women. The decision to limit our initial analysis to women was based on previous studies showing gene by sex interaction for SCARB1 polymorphisms with several metabolic traits (12, 13, 18). Moreover, we confirmed significant gene by sex interaction for SCARB1 genotype on insulin resistance in the Rancho Bernardo Study (the SALSA population was underpowered to detect genotype by sex interaction). This effect could not be attributed to differences in HDL-cholesterol levels or BMI between men and women. It is possible that the sex differences reflect interaction between the gene and sex hormones; both estrogen and testosterone have been shown to regulate SR-BI expression (25, 26), and one study found that gene variants influencing HDL-cholesterol in women may interact with estrogen therapy (13). Alternatively, sex differences could reflect the influence of other risk factors that differ in prevalence between men and women. The nature of the observed genetic associations in this study as they relate to possible sex-specific effects remains to be determined, but our findings highlight the importance of considering gene by sex interaction to uncover SNPs for complex traits.

SR-BI is a cell membrane protein, predominantly localized to the calveolae, and is highly expressed in the liver, steroidogenic tissues, and macrophages. SR-BI has been studied extensively as the calveolae, and is highly expressed in the liver, steroidogenic traits.

Our results lend further support to the possible involvement of members of the CD36 family of genes in the development of type 2 diabetes, providing compelling evidence for SCARB1 as a candidate gene in the previously linked chromosome 12q24 region. Further evaluation of this locus to identify underlying causative variants is warranted, as are studies to understand the basis of the observed sex-specific effects.

Acknowledgments

We thank Ed Martinez, Chief Executive Officer of the San Ysidro Health Center, and the rest of the San Ysidro Health Center staff and patients who contributed their time and support to this research, along with the participants from the Rancho Bernardo Study. We also thank Tim Wiltshire, Enrique Saez, and Nicholas Gekakis for helpful discussion.

Address all correspondence and requests for reprints to: Jeanette McCarthy, Ph.D., Associate Professor, Community and Family Medicine and the Institute for Genome Sciences and Policy, 450 Research Drive, Duke University Medical Center, Box 91009, Durham North Carolina 27708. E-mail: jeannette.mccarthy@duke.edu.

This study was supported in part by Grant RO1 HL085191-01 from the National Heart, Lung, and Blood Institute (J.J.M.), Grant P60 MD00220 from the National Center on Minority Health and Health Disparities for the San Diego EXPORT Center (G.T.), Grant DK31801 from the National Institute of Diabetes and Digestive and Kidney Diseases (P.I. Barrett-Connor), and Grant AG07181 (E.B.-C.) from the National Institute on Aging.

Disclosure Summary: The authors have nothing to disclose.

References


