Interaction of an S100A9 gene variant with saturated fat and carbohydrates to modulate insulin resistance in 3 populations of different ancestries1–3

Ruth Blanco-Rojo,4–6,12 Javier Delgado-Lista,4,5,12 Yu-Chi Lee,6 Chao-Qiang Lai,7 Pablo Perez-Martinez,4,5 Oriol Rangel-Zuñiga,4,5 Caren E Smith,6 Bertha Hidalgo,8 Juan F Alcalá-Díaz,4,5 Francisco Gomez-Delgado,4,5 Laurence D Parnell,7 Donna K Arnett,8 Katherine L Tucker,9 Jose Lopez-Miranda,4,5,13* and Jose M Ordovas6,10,11,13*  

4Lipids and Atherosclerosis Unit, Reina Sofia University Hospital, Maimonides Institute for Biomedical Research at Cordoba, University of Cordoba, Cordoba, Spain; 5Centro de Investigación Biomédica en Red (CIBER) Fisiopatología Obesidad y Nutrición, Instituto de Salud Carlos III, Madrid, Spain; 6Nutrition and Genomics Laboratory and 7Agricultural Research Service, USDA, Jean Mayer USDA Human Nutrition Research Center on Aging, Tufts University, Boston, MA; 8Department of Epidemiology, University of Alabama at Birmingham, Birmingham, AL; 9Department of Clinical Laboratory and Nutritional Sciences, University of Massachusetts, Lowell, MA; 10Department of Epidemiology, Spanish National Center for Cardiovascular Research (CNIC), Madrid, Spain; and 11Madrid Institute for Advanced Studies (IMDEA) Food Institute, Madrid, Spain

ABSTRACT

Background: S100 calcium-binding protein A9 (S100A9) has previously been identified as a type 2 diabetes (T2D) gene. However, this finding requires independent validation and more in-depth analyses in other populations and ancestries.

Objectives: We aimed to replicate the associations between an S100A9 variant and insulin resistance and T2D and to initiate an investigation of potential interactions with the habitual diet in several independent populations.

Design: We investigated the association of the S100A9 variant rs3014866 with insulin resistance and T2D risk and its interactions with diet in 3 diverse populations as follows: the CORDIOPREV (Coronary Diet Intervention with Olive Oil and Cardiovascular Prevention) study, which consisted of Spanish white adults; the GOLDN (Genetics of Lipids Lowering Drugs and Diet Network) study, which involved North American non-Hispanic white adults; and Hispanic adults who participated in the BPRHS (Boston Puerto Rican Health Study).

Results: Meta-analysis indicated that T carriers presented a lower risk of T2D than CC carriers (pooled OR: 0.714; 95% CI: 0.584, 0.845; $P = 0.002$). In all 3 populations (CORDIOPREV, GOLDN, and BPRHS), we showed a significant interaction between the rs3014866 single nucleotide polymorphism and dietary SFA:carbohydrate ratio intake for the homeostasis model assessment of insulin resistance (HOMA-IR) ($P = 0.028$, $P = 0.017$, and $P = 0.026$, respectively). CC carriers had a significantly higher HOMA-IR only when SFA:carbohydrate intake was low ($P = 0.045$ for the CORDIOPREV, $P = 0.033$ for the GOLDN, and $P = 0.046$ for the BPRHS), but not when SFA:carbohydrate ratio intake was low.

Conclusions: The minor allele (T) of the S100A9 variant rs3014866 is associated with lower T2D risk in 3 populations of different ancestries. Note that individuals with the high-risk CC genotype may be more likely to benefit from a low SFA:carbohydrate ratio intake to improve insulin resistance as evaluated with the use of the HOMA-IR. These trials were registered at clinicaltrials.gov as NCT00083369 (GOLDN), and NCT01231958 (BPRHS). Am J Clin Nutr 2016;104:508–17.

1Supported by an Instituto de Salud Carlos III (ISCIII) postdoctoral research contract [Sara Borrell (Spain); to RB-R]. CES is supported by grant K08HL112845 from the National Heart, Lung, and Blood Institute, NIH. The CORDIOPREV (Coronary Diet Intervention with Olive Oil and Cardiovascular Prevention) study was supported by the Ministerio de Economía y Competitividad, Spain [grant AGL2012/29615; to JL-M] and the Ministerio de Ciencia e Innovación, Spain [grants PIE14/00005 (to JL-M) and PI13/00023 (to JD-L)] and was also partly supported by a research grant from the European Commission (NUTRITECH European Integrated Project-299511). The CORDIOPREV study was also supported by the Fundacion Patrimonio Comunarial Olivarero. Additional funding was received from Centro Tecnológico del Olivar y del Aceite, Centro de Excelencia en Investigación sobre Oleoaceitado y Salud, Junta de Andalucía (Consejería de Salud, Consejería de Agricultura y Pesca, Consejería de Innovación, Ciencia y Empresa), the Diputaciones de Jaen and Cordoba, the Centro de Excelencia en Investigación sobre Oleoaceitado y Salud and Ministerio de Medio Ambiente, the Medio Rural y Marino, and the Spanish Government. The study was also co-financed by the Fondo Europeo de Desarrollo Regional. The CIBER Fisiopatología Obesidad y Nutrición is an initiative of the ISCIII, Madrid, Spain. The Genetics of Lipids Lowering Drugs and Diet Network study was supported by the National Heart, Lung, and Blood Institute (NHBLI) [grants U01HL072524 (Genetic and Environmental Determinants of Triglycerides), NHLBI R01 HL091357 (Genomewide Association Study of Lipid Response to Fenofibrate and Diet Fat), HL54776, and HL078885 and by contracts K5-3066-5-10 and 58-1950-9-001 from the USDA, Agriculture Research Service]. The Boston Puerto Rican Health Study was supported by the NIH (grants P01 AG023394 and P50 HL105185).

2 Any opinions, findings, conclusions, or recommendations expressed in this publication are those of the authors and do not necessarily reflect the view of the USDA. Mention of trade names or commercial products in this publication is solely for the purpose of providing specific information and does not imply the recommendation or endorsement by the USDA.

3 SupplementalTables 1–4 are available from the “Online Supporting Material” link in the online posting of the article and from the same link in the online table of contents at http://ajcn.nutrition.org.

4These authors contributed equally to this work.

5These authors contributed equally to this work.

6To whom correspondence should be addressed. E-mail: jose.ordovas@tufts.edu (JM Ordovas), jlopezm@uco.es (J Lopez-Miranda).

Received February 9, 2016. Accepted for publication May 19, 2016. First published online July 20, 2016; doi: 10.3945/ajcn.116.130898.
INTRODUCTION

Type 2 diabetes (T2D) is associated with increased morbidity and mortality, and its growing global prevalence represents a major public health concern (1). The International Diabetes Federation has estimated that 382 million people worldwide are currently presenting with T2D and has projected that this number will rise to 592 million people by 2035 (2).

T2D is a metabolic disease that is characterized by hyperglycemia that results from dysregulation in insulin secretion, insulin action, or both (3). It has been established that the development of T2D is the result of a combination of genetics and the exposome, including dietary intake, exercise, sedentary lifestyle, sleep, and stress (4). A growing number of genetic loci related to T2D have been identified through genome-wide association studies (GWASs), but the cumulative effect of the currently identified loci explains only ~10% of the estimated T2D heritability (5). Therefore, it has been proposed that interactions between genetic and environmental factors may contribute to the identification of some of the missing heritability. Thus, the identification of these gene-environment interactions could provide a major step in the prevention of this disease.

S100 proteins belong to the EF-hand superfamily of Ca2+-binding proteins and have diverse intracellular and extracellular functions (6). Of them, the S100 calcium-binding protein S100A9 (calgranulin B) and S100A8 (calgranulin A) form a heterocomplex called calprotectin, which has been shown to be involved in chronic inflammation and risk of several other metabolic diseases (7, 8). However, calgranulin monomers have pleiotropic functions, some of which may reduce excessive tissue damage or facilitate the resolution of inflammation (9). Indeed, an S100A9 variant (rs3014866) was associated with protection against T2D development in one European population (10). However, this finding requires independent validation in other populations and ancestries.

Therefore, the aims of the present study were as follows: 1) to examine both the associations between the S100A9 variant rs3014866 and the HOMA-IR as a marker of insulin resistance and T2D and 2) to evaluate possible interactions of this variant with diet in the following 3 populations of different ancestries and geographical locations: the CORDIOPREV (Coronary Diet Intervention with Olive Oil and Cardiovascular Prevention) (clinicaltrials.gov; NCT00924937) population, the GOLDN (Genetics of Lipids Lowering Drugs and Diet Network) (clinicaltrials.gov; NCT00083369) population, and the BPRHS (Boston Puerto Rican Health Study) (clinicaltrials.gov; NCT01231958) population.

METHODS

Study populations

The CORDIOPREV study is a long-term secondary cardiovascular prevention intervention trial. The main objective of the study is to evaluate the efficacy of a Mediterranean diet compared with that of a low-fat diet to prevent clinical events and mortality in patients with a history of cardiovascular disease. The eligibility criteria, design, and methods of the CORDIOPREV clinical trial have been previously reported (11, 12), and the protocol is available at clinicaltrials.gov. CORDIOPREV participants were predominantly of European ancestry and were recruited in only one center in a specific region in the south of Spain (Cordoba, Andalucia, Spain). In the current study, we included 711 participants (596 men and 115 women) who had complete genotype and dietary and biochemical baseline data. The local ethics committees approved the trial protocol and all amendments, which followed the Helsinki Declaration and good clinical practices. The experimental protocol conformed to international ethical standards (13), and written informed consent was obtained from all individuals.

GOLDN participants were predominantly of European ancestry and were recruited from 2 genetically homogeneous centers (Minneapolis, Minnesota, and Salt Lake City, Utah). In the current study, 818 participants (404 men and 414 women) of European ancestry were included in our analyses. The primary aim of the GOLDN was to examine the influence of genetic and dietary factors on the response of individuals to fenofibrate. Baseline data obtained from participants before they entered the intervention were selected for this analysis. The study details and related methodology of the GOLDN have been previously described (14) and are also shown at clinicaltrials.gov. The protocol was approved by the institutional review boards at the University of Alabama, the University of Minnesota, the University of Utah, and Tufts University.

The BPRHS is a longitudinal cohort study of stress, nutrition, health, and aging for which study participants were self-identified as Puerto Rican and living in Boston and its metropolitan area (15). The ancestry composition of the BPRHS population is 57.2% European, 27.4% African, and 15.4% Native American (16). For this study, we included 1155 participants (339 men and 816 women) with complete genotype and dietary data. The study protocol was approved by the institutional review boards at Tufts University, Northeastern University, and the University of Massachusetts Lowell.

Biochemical, anthropometric, physical activity, and dietary intake measurements and T2D diagnosis

Blood samples were drawn after an overnight fast. In the CORDIOPREV, lipid variables and glucose were measured with the use of Architect c-16000 analyzers (Abbott) with the following spectrophotometric techniques (enzymatic colorimetric methods): a hexokinase method for glucose and oxidation peroxidation for total cholesterol, HDL cholesterol, and triglycerides. LDL cholesterol was calculated with the use of the Friedewald formula (provided that the triglyceride concentration was <400 mg/dL). The fasting insulin concentration was measured with the use of a chemiluminescent microparticle immunoassay and an i-2000 analyzer (Abbott Architect). The plasma
concentration of high-sensitivity C-reactive protein was determined with the use of a high-sensitivity ELISA (BioCheck Inc.). In the GOLDN, fasting insulin was obtained with the use of a radioimmunoassay and a commercial kit (Linco Research), and fasting glucose was measured with the use of a hexokinase-mediated reaction in the Hitachi commercial kit (Roche Diagnostics). Measurements of blood lipids, including triglycerides and HDL cholesterol, have been described (17). Plasma high-sensitivity C-reactive protein was measured with the use of the K-Assay CRP calibrator (2), a latex particle–enhanced-immunoturbidimetric assay on the Hitachi 911 analyzer (Kamiya Biomedical Co.). In the BPRHS, fasting insulin was measured with the use of an Immulite 1000 Insulin Kit (LKN1; Siemens Medical Solution Diagnostics) on the Immulite 1000 system (Siemens Medical Solution Diagnostics), and Olympus AU400e with Olympus glucose reagents (Olympus America Inc.) were used to measure fasting glucose. Fasting triglycerides and HDL cholesterol were measured with the use of Olympus HDL cholesterol reagents (OSR6195; Olympus America Inc.) and Olympus triglyceride reagents (OSR6033; Olympus America Inc.). C-reactive protein in serum was analyzed with the use of an immunoturbidimetric reaction in a Cobas Fara II Centrifugal Analyzer (Roche) with a DiaSorin CRP SPQ test-system antibody reagent set II (AM-0039; Atlantic Antibodies).

In the CORDIOPREV, subjects also underwent an oral glucose tolerance test at baseline. Therefore, they were considered to have T2D if they had a medical diagnosis of diabetes, were receiving glucose-lowering treatment, or had any of the American Diabetes Association criteria for the diagnosis of T2D (fasting glucose concentration ≥7 mmol/L, 2-h glucose concentration ≥11.1 mmol/L, or glycated hemoglobin ≥6.5%) (18). In the GOLDN and BPRHS, T2D was defined as reported diabetes, a fasting glucose concentration ≥7 mmol/L, or the use of a glucose-lowering medication. For the 3 populations, the HOMA-IR, which was calculated as

$$\frac{\text{fasting glucose (mmol/L)} \times \text{fasting insulin (pmol/L)}}{135}$$

was used to represent insulin resistance.

Weight, height, and waist circumference were measured according to standardized protocols in the 3 populations, and BMI (in kg/m²) was calculated as weight divided by height square.

Physical activity in the CORDIOPREV was assessed by the calculation of the validated physical activity factor the leisure time physical activity (19) by multiplying the metabolic equivalent value (20) by the number of hours per week of each reported physical activity and summed across activities to estimate the overall leisure time physical activity. In the GOLDN, a questionnaire was used to estimate the number of hours per day that were dedicated to different activities (21). In the BPRHS, a physical activity score that was based on the Paffenbarger questionnaire of the Harvard Alumni Activity Survey (22) was estimated.

Dietary intake was assessed in the CORDIOPREV with the use of a 137-item semiquantitative food-frequency questionnaire (FFQ) that was validated in Spain (23). Energy and nutrient intakes were calculated with the use of Spanish food-composition tables (24). In the GOLDN, dietary intake was collected with the use of a diet-history questionnaire, which was developed by the National Cancer Institute and was validated in 2 studies (25, 26). The food list and nutrient database used with the diet-history questionnaire were based on US dietary data (27). In the BPRHS, dietary intake was assessed with the use of an FFQ that was designed for and validated in this population (28). Energy and nutrient intakes were calculated with the use of Nutrition Data System for Research software (version 2007; Nutrition Coordinating Center, University of Minnesota). To normalize dietary intake in cohorts, we evaluated protein, carbohydrate, and total and type of fat intakes as percentages of total energy. In addition, we also examined specific SFAs and types of carbohydrates as percentages of total energy.

DNA isolation and S100A9 genotyping

The single nucleotide polymorphism (SNP) rs3014866 (4811713T>C), which is located ∼1.2-kbp upstream of the S100A9 messenger RNA start site, was selected according to previous evidence (6, 10). DNA samples were isolated from blood samples with the use of the salting-out method in the CORDIOPREV, with Gentra Puregene Blood Kits (Gentra Systems) in the GOLDN, and with QIAamp DNA Blood Mini Kits (Qiagen) in the BPRHS.

In the CORDIOPREV, genotyping was performed on an OpenArray SNP Genotyping System of Applied Biosystems (Thermo Fisher Scientific) with the use of assays that were available from the same company. Data were analyzed with the Applied Biosystems TaqManGenotyper Software 2.1 (Thermo Fisher Scientific). For the GOLDN, the S100A9 genotype was imputed with the use of MACH (29) on the basis of the original genome genotype data that was generated from the Genome-Wide Human SNP Array 6.0 (Affymetrix) and the reference panel of CEU-phased haplotypes from HapMap (30); and for the BPRHS, the S100A9 variant was genotyped with the use of Axiom Genome-Wide LAT Array (Affymetrix), which was designed especially for Hispanic populations and contains probe sets to genotype 817,810 SNPs.

Statistical analysis

The statistical analysis was carried out with the use of SPSS software (version 21.0 for Windows; SPSS Inc.), and differences were considered statistically significant at $P < 0.05$. The normal distribution was evaluated for all measured variables, and skewed variables (triglycerides, fasting insulin, HOMA-IR, and C-reactive protein) were normalized by log10. Nonadjusted data are presented as mean ± SDs for continuous variables and as frequencies for categorical variables except as indicated. The Hardy-Weinberg equilibrium for the S100A9 variant was assessed with the use of chi-square tests in each population. Because of the limited sample size, a dominant genetic model for the rs3014866 (4811713T>C) was considered for the analysis to increase the statistical power after confirmation that the model was valid in all studied populations. Thus, CT and TT participants were grouped and compared with CC participants. We also tested the statistical homogeneity by sex, and men and women were analyzed together. Nutrient intake expressed as the percentage of total energy intake (except as indicated) was...
dichotomized on the basis of the median intake of each population for the interaction analysis.

We performed a linear regression analysis to examine the associations between SNP rs3014866 and glucose metabolism variables (fasting glucose, fasting insulin, and HOMA-IR). To avoid potential confounding across cohorts, we adjusted the analyses for sex, age, current smoking, alcohol consumption, physical activity, diabetes medication use, and BMI. In addition, we adjusted the model for the study center and family relationships in the GOLDN and for population structure in the BPRHS. In the GOLDN, family relationships within the population were estimated with the use of the SAS statistical software (version 9.1 for Windows; SAS Institute Inc.) GENMOD procedure with the assumption of an exchangeable correlation structure. Admixture was estimated with the use of the first component of principal component analysis (16). The principal component analyses for sex, age, current smoking, alcohol consumption, physical activity, diabetes medication use, and BMI, kg/m² 31.0

RESULTS

Characteristics of the populations and S100A9 variant

We included 2755 individuals from 3 independent cohorts; one cohort was from Spain (the CORDIOPREV), and 2 cohorts were from the United States (the GOLDN and the BPRHS). Table 1

<table>
<thead>
<tr>
<th>TABLE 1</th>
<th>General characteristics by S100A9 polymorphism (rs3014866) genotypes in the studied populations¹</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>CORDIOPREV</td>
</tr>
<tr>
<td></td>
<td>GOLDN</td>
</tr>
<tr>
<td></td>
<td>BPRHS</td>
</tr>
<tr>
<td>n (%) ²</td>
<td>TT+CT ➔ CC</td>
</tr>
<tr>
<td>Men/women, n</td>
<td>431/78 ➔ 165/37</td>
</tr>
<tr>
<td>Age, y</td>
<td>59.3 ± 9.2 ² ➔ 60.0 ± 7.95</td>
</tr>
<tr>
<td>Type 2 diabetes, n (%)</td>
<td>187 (36.7) ➔ 91 (45.3) ³</td>
</tr>
<tr>
<td>BMI, kg/m²</td>
<td>31.0 ± 4.2 ➔ 31.1 ± 4.8</td>
</tr>
<tr>
<td>Waist circumference, cm</td>
<td>105.1 ± 11.5 ➔ 105.1 ± 11.6</td>
</tr>
<tr>
<td>Total cholesterol, mmol/L</td>
<td>4.11 ± 0.78 ➔ 4.09 ± 0.83</td>
</tr>
<tr>
<td>HDL cholesterol, mmol/L</td>
<td>1.09 ± 0.23 ➔ 1.07 ± 0.27</td>
</tr>
<tr>
<td>LDL cholesterol, mmol/L</td>
<td>2.31 ± 0.63 ➔ 2.25 ± 0.66</td>
</tr>
<tr>
<td>Serum triglycerides, mmol/L</td>
<td>1.46 ± 0.69 ➔ 1.56 ± 0.85</td>
</tr>
<tr>
<td>C-reactive protein,³ mmol/L</td>
<td>22.9 ± 18.9 ➔ 24.6 ± 21.1</td>
</tr>
<tr>
<td>Fasting glucose (mmol/L)</td>
<td>6.27 ± 2.11 ➔ 6.68 ± 2.54 ³</td>
</tr>
<tr>
<td>Fasting insulin (pmol/L)</td>
<td>78.0 ± 57.1 ➔ 83.7 ± 61.6</td>
</tr>
<tr>
<td>HOMA-IR</td>
<td>3.83 ± 4.04 ➔ 4.56 ± 6.98</td>
</tr>
<tr>
<td>Physical activity ³</td>
<td>22.4 ± 5.4 ➔ 24.3 ± 8.8</td>
</tr>
<tr>
<td>Current smoking, n (%)</td>
<td>48 (9.4) ➔ 17 (8.4)</td>
</tr>
<tr>
<td>Current drinking, n (%)</td>
<td>301 (59.0) ➔ 116 (57.7)</td>
</tr>
<tr>
<td>Diabetes medication use, n (%)</td>
<td>167 (32.7) ➔ 84 (41.8) ³</td>
</tr>
<tr>
<td>Lipid medication use, n (%)</td>
<td>435 (85.3) ➔ 175 (87.1)</td>
</tr>
</tbody>
</table>

¹Genotype frequencies in the CORDIOPREV were as follows: TT, n = 158; CT, n = 352; and CC, n = 201. Genotype frequencies in the GOLDN were as follows: TT, n = 266; CT, n = 418; and CC, n = 134. Genotype frequencies in the BPRHS were as follows: TT, n = 217; CT, n = 570; and CC, n = 368. *P < 0.05 between carriers of the T allele and CC genotype for the corresponding variable in each population (univariate general linear model). BPRHS, Boston Puerto Rican Health Study; CORDIOPREV, Coronary Diet Intervention with Olive Oil and Cardiovascular Prevention; GOLDN, Genetics of Lipids Lowering Drugs and Diet Network; S100A9, S100 calcium-binding protein A9 gene.

²For all categorical variables.

³Mean ± SD (all such values for continuous variables).

*Measured with the use of different techniques in the CORDIOPREV, GOLDN, and BPRHS as described in Methods.

*Measured with the use of different physical activity scores in the CORDIOPREV, GOLDN, and BPRHS as described in Methods.
shows the demographic, anthropometric, clinical, biochemical, and lifestyle characteristics of the participants according to S100A9 polymorphism (rs3014866) for each population. Genotype frequencies were consistent with Hardy-Weinberg equilibrium in all populations. In the CORDIOPREV and BPRHS, the prevalence of T2D (P = 0.040 and P = 0.031, respectively) and the prevalence of diabetes medication use (P = 0.026 and P = 0.042, respectively) were greater in CC than in T carriers but not in the GOLDN. In the CORDIOPREV, CC subjects also had higher fasting glucose concentrations than T carriers (P = 0.020). No significant differences were shown between carriers of the T allele and CC genotype for the rest of the measured variables in all 3 populations.

Table 2 shows dietary intakes of the populations. In the CORDIOPREV, no differences were observed between CC and TT+CT genotypes. In the GOLDN, compared with CC carriers, T carriers reported higher intakes (percentages of total energy intake) of total and complex carbohydrates and lower intakes of total fat, saturated fat (SFA), and myristic, palmitic, and stearic fatty acids and a lower SFA-to-carbohydrate (SFA:carbohydrate) ratio (P < 0.01). In contrast, BPRHS participants with the TT+CT genotype reported higher intakes of SFA and myristic and palmitic fatty acids and a higher SFA:carbohydrate ratio (P < 0.05) than did participants with the CC genotype.

**Table 2** Dietary intake by S100A9 polymorphism (rs3014866) genotype in the studied populations

<table>
<thead>
<tr>
<th></th>
<th>CORDIOPREV</th>
<th>GOLDN</th>
<th>BPRHS</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>CC</td>
<td>TT+CT</td>
<td>CC</td>
</tr>
<tr>
<td>Total energy, kcal</td>
<td>2283 ± 529</td>
<td>2214 ± 462</td>
<td>2129 ± 1248</td>
</tr>
<tr>
<td>Protein, E%</td>
<td>18.5 ± 2.7</td>
<td>18.4 ± 2.7</td>
<td>15.8 ± 2.8</td>
</tr>
<tr>
<td>Carbohydrate, E%</td>
<td>41.8 ± 6.8</td>
<td>41.4 ± 6.1</td>
<td>49.2 ± 8.6</td>
</tr>
<tr>
<td>Simple sugars, E%</td>
<td>15.9 ± 5.4</td>
<td>16.0 ± 4.8</td>
<td>11.7 ± 6.4</td>
</tr>
<tr>
<td>Complex carbohydrate, E%</td>
<td>23.1 ± 5.7</td>
<td>22.7 ± 5.7</td>
<td>34.4 ± 7.0</td>
</tr>
<tr>
<td>Total fiber, g/d</td>
<td>25.2 ± 7.7</td>
<td>25.1 ± 7.8</td>
<td>19.2 ± 10.4</td>
</tr>
<tr>
<td>Total fat, E%</td>
<td>36.9 ± 6.0</td>
<td>37.4 ± 5.6</td>
<td>35.2 ± 6.8</td>
</tr>
<tr>
<td>MUFA, E%</td>
<td>18.0 ± 3.6</td>
<td>18.2 ± 3.3</td>
<td>13.2 ± 2.8</td>
</tr>
<tr>
<td>PUFA, E%</td>
<td>6.3 ± 1.7</td>
<td>6.5 ± 1.8</td>
<td>7.6 ± 2.2</td>
</tr>
<tr>
<td>SFA, E%</td>
<td>9.0 ± 1.9</td>
<td>9.0 ± 1.8</td>
<td>11.7 ± 2.6</td>
</tr>
<tr>
<td>Myristic acid (14:0), E%</td>
<td>0.7 ± 0.4</td>
<td>0.7 ± 0.3</td>
<td>1.0 ± 0.4</td>
</tr>
<tr>
<td>Palmitic acid (16:0), E%</td>
<td>4.1 ± 1.2</td>
<td>4.1 ± 1.1</td>
<td>6.3 ± 1.3</td>
</tr>
<tr>
<td>Stearic acid (18:0), E%</td>
<td>1.6 ± 0.5</td>
<td>1.6 ± 0.5</td>
<td>2.9 ± 0.6</td>
</tr>
<tr>
<td>SFA:carbohydrate ratio</td>
<td>0.23 ± 0.10</td>
<td>0.23 ± 0.07</td>
<td>0.25 ± 0.10</td>
</tr>
</tbody>
</table>

1ORs (95% CIs) of T2D risk were estimated by linear regression analysis. P values were adjusted for age, sex, BMI, smoking status, alcohol use, and physical activity. In addition, we adjusted the model for the study center and family relationships in the GOLDN and for the population structure in the BPRHS. BPRHS, Boston Puerto Rican Health Study; CORDIOPREV, Coronary Diet Intervention with Olive Oil and Cardiovascular Prevention; GOLDN, Genetics of Lipids Lowering Drugs and Diet Network; S100A9, S100 calcium-binding protein A9 gene; T2D, type 2 diabetes.

**Table 3** Association of S100A9 variant (rs3014866) with T2D

<table>
<thead>
<tr>
<th></th>
<th>CORDIOPREV</th>
<th>GOLDN</th>
<th>BPRHS</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>No</td>
<td>Yes</td>
<td>OR (95% CI)</td>
</tr>
<tr>
<td>TT+CT CC</td>
<td>323 (63.3)</td>
<td>187 (36.7)</td>
<td>0.662 (0.450, 0.974)</td>
</tr>
<tr>
<td>CC</td>
<td>110 (54.7)</td>
<td>91 (45.3)</td>
<td>1 (reference)</td>
</tr>
<tr>
<td>TT+CT CC</td>
<td>635 (92.8)</td>
<td>49 (7.2)</td>
<td>0.757 (0.364, 1.574)</td>
</tr>
<tr>
<td>CC</td>
<td>123 (91.8)</td>
<td>11 (8.2)</td>
<td>1 (reference)</td>
</tr>
<tr>
<td>TT+CT CC</td>
<td>475 (60.4)</td>
<td>312 (39.6)</td>
<td>0.744 (0.562, 0.986)</td>
</tr>
<tr>
<td>CC</td>
<td>200 (54.3)</td>
<td>168 (45.7)</td>
<td>1 (reference)</td>
</tr>
</tbody>
</table>

P values were adjusted for age, sex, BMI, smoking status, alcohol use, and physical activity. In addition, we adjusted the model for the study center and family relationships in the GOLDN and for the population structure in the BPRHS. BPRHS, Boston Puerto Rican Health Study; CORDIOPREV, Coronary Diet Intervention with Olive Oil and Cardiovascular Prevention; GOLDN, Genetics of Lipids Lowering Drugs and Diet Network; S100A9, S100 calcium-binding protein A9 gene; T2D, type 2 diabetes.
Meta-analysis of S100A9 variant with glucose metabolism variables and T2D

Meta-analyzed T carriers had significantly lower fasting glucose ($\beta = -2.58$; 95% CI: $-4.94, -0.23$; $P = 0.03$) than CC carriers. No differences were shown with respect to fasting insulin values between genotypes, but participants with the TT+CT genotype tended to have a lower HOMA-IR than participants with the CC genotype ($\beta = -0.31$; 95% CI: $-0.67, 0.04$; $P = 0.08$). No significant heterogeneity was observed for both variables ($Q = 0.72$, $I^2 = 0\%$, and $P$-heterogeneity = 0.695 for fasting glucose; $Q = 0.90$, $I^2 = 0\%$, and $P$-heterogeneity = 0.634 for HOMA-IR) (Supplemental Table 1).

When the CORDIOPREV and BPRHS populations were analyzed separately, subjects with the TT+CT genotype showed a lower OR of T2D than participants with the CC genotype (OR for CORDIOPREV: 0.662; 95% CI: 0.450, 0.974; $P = 0.036$ and OR for BPRHS: 0.744; 95% CI: 0.562, 0.986; $P = 0.039$). Although the tendency was the same, we did not find significant results in the GOLDN population (Table 3). Nevertheless, the overall meta-analysis for all 3 cohorts showed that T carriers presented lower risk of T2D than CC carriers (pooled OR: 0.714; 95% CI: 0.584, 0.845; $P = 0.002$). No significant heterogeneity was observed ($Q = 0.25$, $I^2 = 0\%$, and $P$-heterogeneity = 0.881) (Figure 1).

Interaction of S100A9 variant with dietary intake for glucose metabolism variables

To perform the SNP-by-diet interaction tests, dietary factors were dichotomized at median intake in each population. In the CORDIOPREV and BPRHS, rs3014866 interacted significantly with dietary SFA intake ($P = 0.007$ and $P = 0.027$, respectively) for the HOMA-IR. In these 2 populations, CC carriers, compared with T carriers, had a significantly higher HOMA-IR when they had high SFA intake ($P = 0.029$ in the CORDIOPREV and $P = 0.048$ in the BPRHS) but not when they had low SFA intake (Figure 2).

In each of all 3 populations (in the CORDIOPREV, GOLDN, and BPRHS), we showed a significant interaction between the rs3014866 SNP and dietary SFA:carbohydrate intake for the HOMA-IR ($P = 0.028$, $P = 0.017$, and $P = 0.026$, respectively). CC carriers, compared with T carriers, had a significantly higher HOMA-IR when their SFA:carbohydrate intake was high ($P = 0.045$ in the CORDIOPREV, $P = 0.033$ in the GOLDN, and $P = 0.046$ in the BPRHS) but not when their SFA:carbohydrate intake was low (Figure 3).

We did not observe significant interactions between the rs3014866 genotype and SFA and SFA:carbohydrate intakes for fasting glucose or fasting insulin. Likewise, interactions between the genetic variant and carbohydrate intake for fasting glucose, fasting insulin, and the HOMA-IR were not identified (Supplemental Table 2). Moreover, no interactions reached significance with specific SFAs (myristic, palmitic, and stearic fatty acids) for mean comparisons of HOMA-IR values between CC and TT+CT genotypes with high SFA intake (univariate model analysis). BPRHS, Boston Puerto Rican Health Study; CORDIOPREV, Coronary Diet Intervention with Olive Oil and Cardiovascular Prevention; GOLDN, Genetics of Lipids Lowering Drugs and Diet Network; S100A9, S100 calcium-binding protein A9 gene.
acids) or with simple sugars, complex carbohydrates, or dietary fiber (Supplemental Tables 3 and 4).

**DISCUSSION**

The current study showed that the minor allele of the *S100A9* rs3014866 variant was associated with lower T2D risk in 3 independent populations. This finding is consistent with that of a previous study in a European population (10). In addition, to our knowledge, we have provided novel evidence that showed an interaction between rs3014866 and saturated fat and SFA-to-carbohydrate ratio intakes that appeared to influence the HOMA-IR in these populations. This result suggests that CC homozygous individuals may benefit from low SFA:carbohydrate ratio intake to decrease insulin resistance. These findings were robust because we were able to replicate them in populations of different ancestries and diverse clinical, demographic, and lifestyle characteristics.

*S100A9* is a T2D gene that was identified in a GWAS. The gene is located at 1q21, which is a known susceptibility region for T2D (31), and codes for S100A9 protein, also known as calgranulin B. The rs3014866 SNP is located in the 5’-upstream region of the *S100A9* gene, and it has been reported to increase gene expression and S100A9 concentrations in humans (10). The protein S100A9 has been associated with a range of acute and chronic inflammatory responses (32), latter-involved in the progression of insulin resistance and the development of T2D (33). However, we did not observe differences in values of C-reactive protein between genotypes in any of the studied populations. Therefore, our results support the hypothesis that S100A9 is not merely a marker of inflammation but, in addition, may be related to increased insulin-stimulated glucose uptake in some tissues (10). The calprotectin protein (S100A8/S100A9 complex) has been described as a transporter that shuttles arachidonic acid into target cells (34), and there has been evidence that this fatty acid could stimulate glucose uptake in adipocytes by increasing the glucose transporter 1 and the glucose transporter 4 concentrations (35). Moreover, rs3014866 has been associated with the expression of several genes in adipose and lymphoblastic tissues in humans. Specifically, the rs3014866 T allele was associated with significantly decreased expression of interleukin enhancer binding factor 2 in both tissues (36). Interleukin enhancer binding factor 2 is a nuclear factor that modulates IL-2, and it is overexpressed in the peripheral blood of T2D patients (37). Additional studies are needed to clarify the mechanism for the association between this *S100A9* variant and T2D risk.

In contrast, rs3014866 was associated with fasting glucose but not with fasting insulin or the HOMA-IR. Similarly, the Meta-Analyses of Glucose and Insulin-related traits Consortium

![FIGURE 3 Estimated mean ± SEM interactions between the S100A9 variant (rs3014866) and SFA/CHO ratio intake on HOMA-IR. (A) For the CORDIOPREV, low and high SFA/CHO ratios indicate median intakes (≤0.21 and >0.22). (B) For the GOLDN, low and high SFA/CHO ratios indicate median intakes (≤0.24 and >0.25). (C) In the BPRHS, low and high SFA/CHO ratios indicate median intakes (≤0.19 and >0.20). The number inside each bar indicates the number of subjects in that group with data available for HOMA-IR. Analyses were adjusted for age, sex, BMI, smoking status, alcohol use, diabetes medication use, and physical activity. In addition, we adjusted the model for the study center and family relationships in the GOLDN and for the population structure in the BPRHS. *P*-interactions between the SFA/CHO ratio (as dichotomous) and the S100A9 variant in each population were obtained with the use of a univariate model analysis. *P* < 0.05 for mean comparisons of HOMA-IR values between CC and TT+CT genotypes with high SFA/CHO ratio intake (univariate model analysis). BPBRHS, Boston Puerto Rican Health Study; CHO, carbohydrate; CORDIOPREV, Coronary Diet Intervention with Olive Oil and Cardiovascular Prevention; GOLDN, Genetics of Lipids Lowering Drugs and Diet Network; S100A9, S100 calcium-binding protein A9 gene.
showed no relation between this SNP and glycemic variables (38). This lack of an association could be explained by the gene-diet interactions that we have reported here. These novel and consistent interactions indicated that rs3014866 CC carriers showed a less HOMA-IR when their SFA and SFA:carbohydrate intakes were low, which suggested a protective effect of a diet that was low in saturated fat in relation to carbohydrate in these participants. Dietary SFAs are thought to compromise insulin sensitivity by affecting cell-membrane fluidity and functions, including insulin receptor binding and affinity (39). In addition, some studies have suggested that lipolytic products of SFAs may exert unfavorable effects on the expression of genes encoding lipogenic and lipid-transport proteins (40). Despite these reports, the relation between total SFA intake and T2D risk remains unclear (41), and it has been suggested that the SFA-chain length could be associated with different biological effects (42). However, we did not find significant interactions of the genotype with myristic, palmitic, and stearic fatty acids on measures of the HOMA-IR.

Current dietary recommendations that are aimed to prevent diabetes agree on the importance of reducing dietary saturated fat intake (18, 43) although debate remains as to which macronutrient should optimally replace the saturated fat and what dietary pattern should be recommended (41). MUFAs, PUFAs and complex carbohydrates have been shown to have beneficial effects in glycemic control and insulin sensitivity when substituted for SFAs in the diet (44, 45). In this context, gene-diet interactions may play a role, and several studies have shown that various dietary approaches could provide differing effects on glucose metabolism depending on the genetic background of the participants (46). Therefore, our findings reinforce the hypothesis that a single dietary model may not be optimal for all individuals (47), and dietary advice that takes the genotype into account may be the key for a more effective prevention of T2D.

The potential mechanisms for the interactions observed should be explored in future studies. Nevertheless, an in silico analysis (http://snpinfo.niehs.nih.gov/) of the S100A9 variant rs3014866 suggested that the major allele (C) might create a specific binding site for the transcription factor peroxisome proliferator-activated receptor γ (PPARG), which might be eliminated by the presence of the minor allele (T). This transcription factor is known to regulate adipocyte differentiation and has been implicated in lipid and glucose metabolism (48). A reduction in PPARG activity has been shown to be associated with insulin sensitivity (49). Several studies have suggested that PPARG may play a role in the development of insulin resistance in response to a Western-type high-fat diet (50, 51). Therefore, these findings provide support for the observed interaction by which rs3014866 CC carriers have a lower HOMA-IR with lower SFA:carbohydrate ratio intake.

Some limitations of the current study should be mentioned. The CORDIOPREV, GOLDN, and BPRHS populations represent individuals of different ancestries, lifestyles, and clinical and baseline characteristics. For example, GOLDN participants are younger with lower BMI and a lower prevalence of T2D, which could explain the lack of a direct association between the S100A9 variant and T2D risk in this population. Dietary intakes also differ significantly in the 3 populations, and it was not possible to find a cutoff for SFA or SFA:carbohydrate ratio intakes that was common to all subjects for the interaction with rs3014866 on insulin resistance. Moreover, different FFQs and methodologies were used to evaluate dietary data. In addition, although the HOMA-IR is a validated index for estimating insulin resistance in human studies (52) and was the only method available in the 3 populations, we recognize that the use of this single measure was a limitation. Despite these shortcomings, we showed consistent trends across the 3 populations in terms of the main genetic associations and gene-diet interactions. In addition, moderate sample sizes of the studied populations may have limited the statistical power. Nevertheless, to our knowledge, this is the first study to replicate the association between rs3014866 and T2D risk, and novel interactions between this SNP and dietary factors to modulate insulin resistance were identified.

In conclusion, the minor allele (T) of the S100A9 variant rs3014866 is associated with decreased T2D risk in 3 populations of different ancestries and diverse clinical and demographic characteristics. Note that individuals who lack the protective (T) allele (CC carriers) may benefit from low SFA:carbohydrate ratio intake to improve insulin sensitivity. Our study adds support to accumulating evidence that T2D arises from the complex interplay between genetic and environmental factors and, if replicated, may contribute to the development of effective individualized dietary recommendations to prevent T2D.

The authors’ responsibilities were as follows—RB-R: designed the study hypothesis, analyzed the data, and wrote the manuscript; RB-R and JMO: had primary responsibility for the final content of the manuscript; JD-L, PP-M, BH, JFA-D, FG-D, LDP, and KLT: contributed to the editing of the manuscript; Y-CL, C-QL, OR-Z, and CES: conducted the research; DKA, KLT, and JL-M: provided essential data; JD-L, JL-M, and JMO: designed the research and provided study oversight; and all authors: read and approved the final manuscript. None of the authors reported a conflict of interest related to the study.

REFERENCES


