Low-Frequency Variants in *HMGA1* Are Not Associated With Type 2 Diabetes Risk

Marcel Marquez,[1] Marlène Huyvaert,[1] John R.B. Perry,[2,3] Richard D. Pearson,[2] Mario Falchi,[4]
Andrew P. Morris,[2] Sidonie Vivequin,[1] Stéphane Lobbens,[1] Loïc Yengo,[1] Stefan Gaget,[1]
Francois Pattou,[5] Odile Poulain-Godefroy,[1] Guillaume Charpentier,[6] Lena M.S. Carlsson,[7]
Peter Jacobson,[7] Lars Sjöström,[7] Olivier Lantieri,[8] Barbara Heude,[9] Andrew Walley,[4]
Beverley Balkau,[9] Michel Marre,[10] the DIAGRAM Consortium,* Philippe Froguel,[1,4] and
Stéphane Cauchi[1]

It has recently been suggested that the low-frequency c.136–14_136–13insC variant in high-mobility group A1 (*HMGA1*) may strongly contribute to insulin resistance and type 2 diabetes risk. In our study, we attempted to confirm that *HMGA1* is a novel type 2 diabetes locus in French Caucasians. The gene was sequenced in 368 type 2 diabetic case subjects with a family history of type 2 diabetes and 372 normoglycemic control subjects without a family history of type 2 diabetes. None of the 41 genetic variations identified were associated with type 2 diabetes. The lack of association between the c.136–14_136–13insC variant and type 2 diabetes was confirmed in an independent French group of 4,538 case subjects and 4,015 control subjects and in a large meta-analysis of 16,605 case subjects and 46,179 control subjects. Finally, this variant had no effects on metabolic traits and was not involved in variations of *HMGA1* and insulin receptor (*INSR*) expressions. The c.136–14_136–13insC variant was not associated with type 2 diabetes in individuals of European descent. Our study emphasizes the need to analyze a large number of subjects to reliably assess the association of low-frequency variants with the disease. *Diabetes* 61:524–530, 2012

Insulin resistance and pancreatic β-cell dysfunction are the two physiologic hallmarks of type 2 diabetes, and the expectation was that the majority of susceptibility loci identified in genome-wide association studies (GWASs) would map to either of these diabetes-related phenotypes. Yet, the majority of these loci can be directly mapped to the pancreatic β-cell or are of unknown biologic consequence (1). Few loci were associated with insulin resistance (2,3). Furthermore, the known common variants account for <10% of the overall estimated genetic contribution to type 2 diabetes predisposition (3). There is obviously a "dark matter" that remains to be discovered (4). A growing number of researchers are turning to rare genetic changes with strong effects as important contributors (5–8). Because whole-genome sequencing requires computational power and remains costly for large case-control studies, candidate gene sequencing may be an effective approach to identify rare and low-frequency variants associated with common diseases.

In 2005, Foti et al. (9) observed that lack of *HMGA1* causes insulin resistance and diabetes in humans. This gene encodes a nonhistone protein that binds on AT-rich segments in the minor groove of DNA and regulates gene transcription (10,11). *HMGA1* is a positive regulator of both insulin receptor (*INSR*) and insulin-like growth factor-I receptor (*IGF-IR*) genes and is induced by the Wnt/β-catenin pathway, which is known to play a key role in type 2 diabetes risk (12–16). In humans, two rare Mendelian mutations were identified only in diabetic patients and associated with impaired *HMGA1* expression: a hemizygous deletion of *HMGA1* (two carriers from the same family) and a heterozygous single-nucleotide deletion in the 3′ untranslated region of *HMGA1* (c.*369del) (9). Chiefari et al. (17) recently observed that a low-frequency variant in *HMGA1* (c.136–14_136–13insC) was strongly associated with type 2 diabetes risk and decreased expression of *HMGA1* and *INSR* genes. All these elements suggest that rare variants located in this gene can be strongly associated with type 2 diabetes through their effects on insulin resistance. However, extensive analysis of the *HMGA1* gene in additional populations is necessary to validate previous findings and identify novel variants potentially associated with type 2 diabetes risk (18).

RESEARCH DESIGN AND METHODS

Study populations. For the sequencing of the *HMGA1* gene, a screening group of 368 French type 2 diabetic case subjects and 372 normoglycemic control subjects was selected. The case subjects were composed of those with at least two first-degree relatives with type 2 diabetes, collected by the UMR CNRS 8199 laboratory. French control subjects were composed of those from the general population study Data from an Epidemiological Prospective Study on the Insulin Resistance Syndrome (DESIR), who were selected because they remained normoglycemic after 9 years of follow-up and had no family history of diabetes (19).

The c.136–14_136–13insC variant was also analyzed in a replication group of 4,538 French type 2 diabetic case subjects and 4,015 French normoglycemic control subjects (Table 1). Type 2 diabetic case subjects were composed of those from the UMR CNRS 8199 laboratory (*n* = 377), the Non-Insulin-Dependent Diabetes, Hypertension, Microalbuminuria or Proteinuria, Cardiovascular Events,

TABLE 1
Association between the c.136–14_136–13insC variant and type 2 diabetes in French individuals

Study	Glycemic status	n	$-/-$	$-/C$	C/C	MAF (%)	H-W	OR (95% CI)	P
Screening	Normoglycemic	372	351	21	0	2.82	0.99	0.91 (0.24–3.36)	0.88
	Type 2 diabetes	368	357	11	0	1.49	0.99		
Replication	Normoglycemic	4,015	3,808	206	1	2.59	0.52	1.12 (0.87–1.44)	0.38
	Type 2 diabetes	4,538	4,282	256	0	2.82	0.06		
Screening + replication	Normoglycemic	4,387	4,159	227	1	2.61	0.37	1.07 (0.84–1.37)	0.57
	Type 2 diabetes	4,906	4,639	267	0	2.72	>0.05		

ORs (95% CI) were estimated using logistic regression models adjusted for age, sex, and BMI. H-W, deviation from Hardy-Weinberg equilibrium (P value).

and Ramipril (DIABHYCAR) study ($n = 2,095$), the DESIR study ($n = 295$), the Biological Atlas of Severe Obesity (ABOS) study ($n = 117$), and the Corbeil-Essonnes Hospital ($n = 1,654$) (20). Control subjects were composed of subjects from the UMR CNRS 8199 laboratory ($n = 1,306$), the DESIR study ($n = 2,291$), the ABOS study ($n = 207$), and the Fleurbaix-Laventie Ville Santé study ($n = 211$). The DIABHYCAR study design and results have been reported (20,21). Participants were selected on the basis of type 2 diabetes, treatment with oral antidiabetic agents on enrollment, high urinary albumin concentration, age ≥50 years, and serum creatinine concentration ≤150 μmol/L. French patients were selected by their general practitioners, and ≥98% were Caucasian (20). The ABOS study was first designed to analyze the association between glycemic or ponderal status and tissue gene expression and has been extended to provide longitudinal follow-up and samplings of a cohort of 900 morbid obese patients subjected to bariatric (weight loss) surgery. The Fleurbaix-Laventie Ville Santé study is a longitudinal epidemiologic study on the determinants of weight gain (22).

All type 2 diabetic case subjects and normoglycemic control subjects of this study were defined according to the 2003 American Diabetes Association criteria (type 2 diabetes: fasting plasma glucose ≥7.0 mmol/L or treated for diabetes; normoglycemic: fasting glucose <5.6 mmol/L). Only French Caucasians were analyzed. Informed consent was obtained from each subject, and our study protocol was approved by local ethics committees.

GWAS data were made available from the DIAGRAM consortium (3). All sample characteristics and statistics remain the same as previously described (3). However, any overlapping samples from the Diabetes Gene Discovery Group (DGDG) study (French samples) were excluded. Total sample size in the seven-study meta-analysis remained at 7,451 type 2 diabetic case subjects and 38,290 control subjects. Correlations between polymorphisms and the c.136–14_136–13insC variant were estimated using pilot release data from the CEU population of the 1000 Genomes project (23).

Sequencing. Overall, 4.5 kb of the *HMGA1* gene were sequenced, including the entire promoter region (-591 base pairs [bp] before the transcription start site 1 to +980 bp after), 8 exons, and 16 intron-exon junctions and untranslated regions. *HMGA1* gene structure was based on RefSeq transcripts (National Center for Biotechnology Information, Bethesda, MD) (NM_002131, NM_145899, NM_145901, NM_145902, NM_145903, NM_145905), genomic sequence (Gen-Bank NT_007592.15). Primers were designed using Primer-BLAST (National Center for Biotechnology Information) and were validated not to amplify the pseudogene regions using In-Silico PCR (University of California, Santa Cruz, CA) (Supplementary Table 1). PCR amplifications were purified using Agencourt AMPure (Agencourt Bioscience Corporation, Beverly, MA) on Biomex NX (Beckman Coulter, Fullerton, CA) after validation of a single band of excepted size on agarose gel. Purified PCR amplification products were then used for sequence reaction with Big Dye Terminator Cycle Sequencing Ready Reaction Kit 3.1 (Applied Biosystems, Foster City, CA). Sequence products were purified with Agencourt CleanSEQ (Agencourt Bioscience Corporation) on Biomek NX read on an ABI Prism 3730XL DNA Sequencer and assembled and analyzed with Variant Reporter, Seqscape (Applied Biosystems), and Sequencher software (Gene Codes Corporation, Ann Arbor, MI). All variants were validated by resequencing of forward and reverse strands.

Genotyping. Genotyping was performed using the High Resolution Melting (HRM) method using MeltDoctor HRM Master Mix (Applied Biosystems) on a Lightcycler 480 (Roche, Indianapolis, IN). HRM curve analysis was performed on LightCycler 480 software, and positive signals were confirmed by sequencing. There was a concordance rate of >99% between genotyping and sequencing results for the 368 type 2 diabetic case subjects and 372 normoglycemic control subjects of the screening study.

Intermediate metabolic traits. BMI is the weight in kilograms divided by the square of the height in meters. The homeostasis model assessment (HOMA)-2–insulin resistance and HOMA2-B indices were calculated using the HOMA calculator based on fasting glucose and fasting insulin levels in subjects in

a steady-state situation (fasting plasma glucose: 3–25 mmol/L; fasting plasma insulin: 20–400 pmol/L; available at http://www.dtu.ox.ac.uk/homacalculator). HbA$_{1c}$ was only available in a subset of the DESIR and UMR CNRS 8199 samples. The insulinogenic index (IGI) was calculated on the basis of the oral glucose tolerance test (OGTT) measures as IGI = (30 min plasma insulin $-$ fasting plasma insulin)/30 min plasma glucose, where plasma insulin is in pmol/L and plasma glucose is in mmol/L (24). The insulin sensitivity index (ISI) was calculated on the basis of OGTT measures as ISI = $10,000/\sqrt{}$(fasting plasma glucose $\times$ fasting plasma insulin $\times$ mean OGTT glucose $\times$ mean OGTT insulin), where plasma insulin and OGTT insulin are in mU/L and plasma glucose and OGTT glucose are in mg/dL (25). The disposition index (DI) was calculated on the basis of OGTT measures in subjects with a fasting plasma glucose ≥70 mg/dL as DI = ISI $\times$ 100 $\times$ 30 min plasma insulin $\times$ (30 min plasma glucose $\times$ [30 min plasma glucose – 70 mg/dL]), where plasma insulin is in mU/L and plasma glucose is in mg/dL (26). Indices based on OGTT measures were only available for a subset of UMR CNRS 8199 samples. Waist-to-hip ratio was calculated as waist circumference in centimeters divided by hip circumference in centimeters and was only available in DESIR and UMR CNRS 8199 samples. Only normoglycemic subjects were analyzed to avoid possible modifying effects of type 2 diabetes and treatments on metabolic traits.

Expression study. The expression study was conducted in 149 Swedish obesity-discordant siblings (included in the Swedish Obese Subjects [SOS] Sib Pair study) using Merlin under a dominant genetic model (27,28). Average family size was 4.34, and median age (first to third quartiles) was 44 years (36–62). Gene expression data were available from 347 siblings, and peripheral blood was available from all subjects. Gene expression was measured in subcutaneous adipose tissue using the Affymetrix Human U133PLUS2.0 microarrays for *HMGA1* (probe-sets 206074_s_at and 210457_x_at) and *INSR* (probe-sets 213792_s_at, 207851_s_at, 226212_s_at, and 226216_at). A Genetic Power Calculator was used to estimate the power of our study in identifying association between the c.136–14_136–13insC variant and *HMGA1* and *INSR* gene expression levels (29). To this aim, we assessed the heritability of the *HMGA1* and *INSR* transcription levels on the Swedish sample and used observations from the study by Chiefari et al. (17) to estimate both the minor allele frequency (MAF) of the c.136–14_136–13insC variant and its effects on the gene expression levels. The association study was carried out by fitting a linear regression model to each probe-set within a variance component framework to account for correlation between gene expression levels within each family (30).

Statistical analyses. Given the low allele frequency of the c.136–14_136–13insC variant, only the dominant genetic model of inheritance was considered. In our screening analysis, the statistical power of our study design was sufficient to detect low-frequency variants with strong effect sizes (type I error rate of 0.05) (Supplementary Fig. 1). MAFs between case subjects and control subjects were compared using the χ^2 test and Monte Carlo simulations ($n = 2,000$). Associations of the c.136–14_136–13insC variant with type 2 diabetes and age at diagnosis were analyzed using logistic regression and linear regression models, respectively. No associations between the c.136–14_136–13insC variant and the confounding factors (i.e., age at examination, sex, and BMI) were detected using adjusted linear and logistic regression models ($P = 0.39$, $P = 0.42$, $P = 0.21$, respectively). Furthermore, these confounding factors had no influence on the association between the c.136–14_136–13insC variant and type 2 diabetes ($P = 0.81$, $P = 0.87$, $P = 0.59$, respectively). No between-cohort heterogeneity in the genotype distributions of the c.136–14_136–13insC variant and no deviation from Hardy-Weinberg equilibrium were detected (Supplementary Table 2). Linear regression models were used to compare quantitative intermediate traits between normoglycemic and type 2 diabetic subjects. All statistical models were adjusted for age, sex, and BMI (if appropriate). When necessary, metabolic traits were log-transformed to satisfy the assumption of normality. Heterogeneity between studies was assessed by the Woolf's test and quantified by the Higgins statistic (I^2) (31,32). Because of

strong heterogeneity between previous and current findings, we applied a random-effects meta-analysis to estimate summary odds ratios (ORs) 1, 3, and 4 (33). However, a fixed-effects meta-analysis was performed to estimate the summary OR 2 given that no heterogeneity was observed between summary OR 1 and DIAGRAM data. Statistical power was assessed using Quanto (version 1.2.4). All *P* values were two-sided. The Bonferroni correction was calculated to address the problem of multiple comparisons. SPSS (v. 14.0.2; SPSS, Chicago, IL) and R statistics (v. 2.10.1) were used for general statistics.

RESULTS

Screening of the *HMGA1* gene. The *HMGA1* gene was sequenced in 368 type 2 diabetic case subjects having at least two first-degree relatives with type 2 diabetes and 372 normoglycemic control subjects prospectively followed during 9 years and without family history of diabetes. Overall, 41 genetic variants were detected, including four variants (c.136–14_136–13insC, rs41269028, rs2780219, and rs17038619) previously identified by Chiefari et al. (17) (Fig. 1 and Supplementary Table 3). None of them were differentially distributed between normoglycemic control subjects and type 2 diabetic case subjects. However, the c.136–14_136–13insC C allele tended to be more frequent in control subjects compared with type 2 diabetic case subjects (MAF 2.82 vs. 1.49%, *P* = 0.09) (Table 1, Supplementary Table 3, and Supplementary Fig. 2). Given that this low-frequency variant was supposed to be strongly associated with type 2 diabetes, we verified this hypothesis by analyzing larger groups of case and control subjects.

Association of the c.136–14_136–13insC variant and type 2 diabetes. An additional 4,538 type 2 diabetic case subjects and 4,015 normoglycemic control subjects were genotyped for the c.136–14_136–13insC variant. Genotype distributions and ORs are reported in Table 1. No association between the c.136–14_136–13insC variant and type 2 diabetes risk was observed (screening + replication groups: OR 1.07, CI 0.84–1.37, *P* = 0.57). Furthermore, no association between the c.136–14_136.13insC variant and the age at diagnosis was observed (*P* = 0.61). Given that Chiefari et al. (17) analyzed two independent groups of Italian control subjects and one group of Italian type 2 diabetic patients, we performed two meta-analyses on the association of the

c.136–14_136–13insC variant and type 2 diabetes risk (dominant model of inheritance) including either control group 1 (*n* = 2,544, personal interviews to determine the lack of family history of type 2 diabetes and related conditions) or control group 2 (*n* = 784, self-reported medical questionnaire).

By including the Italian control group 1, we detected a strong heterogeneity between studies (Woolf test: χ^2 = 67.08, df = 2, *P* = 2.66 × 10^{-15}; Higgins test: I^2 = 97% [94%; 98.5%]). In a total of 9,154 case subjects and 7,889 control subjects of European descent (Italian, U.S., and French), no association between the c.136–14_136–13insC variant and type 2 diabetes was detected in our meta-analysis (summary OR 1 = 2.95, 95% CI 0.76–11.54, *P* = 0.12) (Fig. 2). By using CEU data generated from the 1000 Genomes pilot project, we also identified 10 polymorphisms that were perfectly correlated (r^2 = 1) with the c.136–14_136–13insC variant (Supplementary Table 4). Pairwise correlations (r^2) between these 10 polymorphisms were high in HapMap 3 (0.77–1.00). Disease association statistics were available for these polymorphisms from a recently published DIAGRAM consortium meta-analysis of 8,130 type 2 diabetic case subjects and 38,987 control subjects (3). After exclusion of samples already tested in this present study, the total sample size was 7,451 case subjects and 38,290 control subjects. Disease association test statistics ranged from *P* = 0.31 to *P* = 0.89 for these 10 polymorphisms. The rare T allele (HapMap CEU MAF = 2.7%) of the polymorphism closest to the gene (rs9394200: distance ~5 kb), which tags the risk-increasing insertion (OR 0.94, 95% CI 0.82–1.07, *P* = 0.32), was selected to assess the significance of this signal in further meta-analysis. In a total of 16,605 case subjects and 46,179 control subjects of European descent (Italian, U.S., French, DIAGRAM consortium), no association with type 2 diabetes was detected (summary OR 2 = 0.95, 95% CI 0.83–1.08, *P* = 0.44) (Fig. 2).

By including the Italian control group 2 (MAF 1.66%) with another meta-analysis comprising 9,154 case subjects and 6,129 control subjects of European descent (Italian, U.S., and French), a between-study heterogeneity was also detected (Woolf test: χ^2 = 8.29, df = 2, *P* = 0.02; Higgins
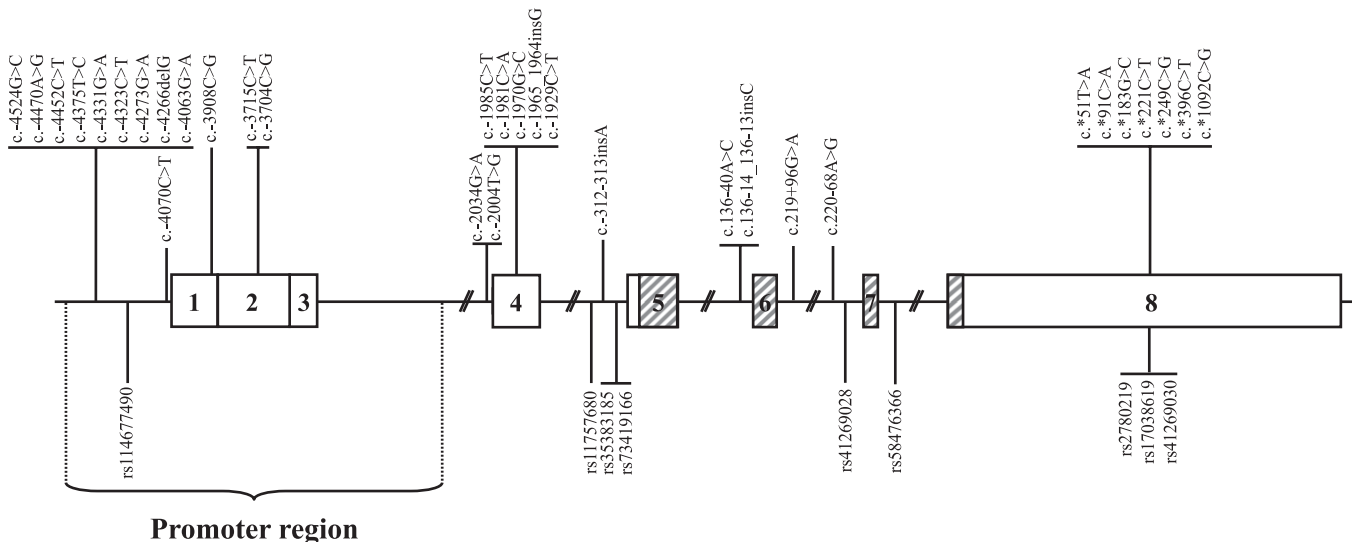

FIG. 1. Genetic variants identified in the *HMGA1* gene by sequencing. A total of 41 genetic variants were identified when sequencing the *HMGA1* gene in 368 type 2 diabetic subjects and 372 control subjects from our screening group of samples. Coding regions are represented by hatched boxes.
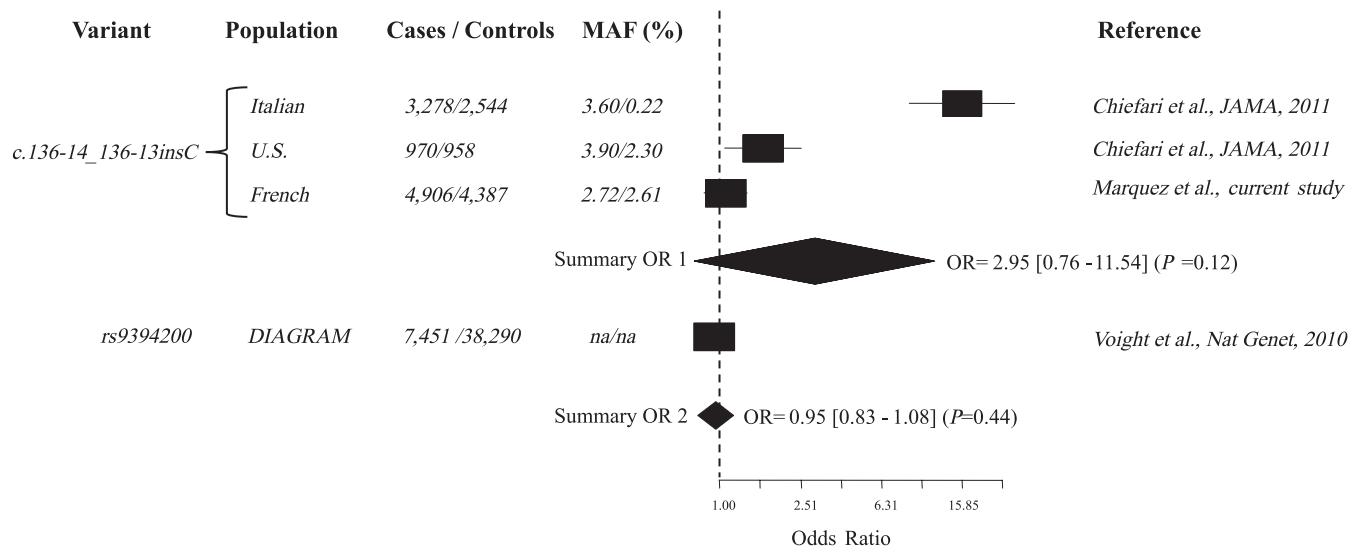
FIG. 2. Meta-analysis on the association between the c.136–14_136–13insC variant (or correlated polymorphism) and type 2 diabetes. na, not applicable; summary OR 1, Italian + U.S. + French data; summary OR 2, Italian + U.S. + French + DIAGRAM data.

test: I^2 = 76% [20.6%; 92.7%]). Again, no association was observed between the c.136–14_136–13insC variant and type 2 diabetes (summary OR 3 = 1.49, 95% CI 0.98–2.26, P = 0.06). Similar conclusions were reached when this meta-analysis was extended to the DIAGRAM data (Italian, U.S., French, DIAGRAM consortium) in 16,605 case subjects and 44,419 control subjects of European origin (summary OR 4 = 1.13, 95% CI 0.73–1.77, P = 0.58).

Effects on metabolic traits in normoglycemic control subjects. Normoglycemic control subjects were then analyzed for type 2 diabetes–related quantitative traits (Table 2). Carrying a c.136–14_136–13insC C allele was not found to influence fasting glucose, fasting insulin, HOMA2 indices, IGI, ISI, DI, HbA_{1c}, waist-to-hip ratio, and BMI.

Effects on *HMGA1* and *INSR* expressions. The *HMGA1* and *INSR* expressions were assessed in the subcutaneous adipose tissue of 347 siblings from the SOS Sib Pair study. The expression levels of both *INSR* and *HMGA1* in the subcutaneous adipose tissue were above average (74 and 71% of the most expressed gene, respectively). Maximum likelihood estimate of the minor allele frequency for the c.136–14_136–13insC variant was 3.30% (47 heterozygous carriers). For the power calculation, we assumed MAF of 1% for c.136–14_136–13insC. Heritabilities, estimated through Merlin in the Swedish sample and averaged among all alternate transcripts from the same gene, were 70 and 27% for *HMGA1* and *INSR*, respectively. Effect sizes (r^2) were estimated as $d^2/(d^2 + 4)$ from Cohen's d (34), calculated using the gene transcript mean levels and CIs as observed between carriers and noncarriers of the c.136–14_136–13insC variant in the study by Chiefari et al. (17). Average r^2 values between type 2 diabetic carriers and both type 2 diabetic and healthy noncarriers were 0.66 for *HMGA1* and 0.74 for *INSR*. We observed that in the study by Chiefari et al. (17), the type 2 diabetic noncarriers showed slightly lower expression levels of *HMGA1* and *INSR* than the healthy noncarriers control subjects (P < 0.05). By assuming a conservative lower QTL effect size of 30% in our sample (where no type 2 diabetic carriers were included), we observed a statistical power >80% at α = 0.001 to identify

a dominant effect for the c.136–14_136–13insC on the *HMGA1* and *INSR* gene expression levels in our sample. The association study showed that the c.136–14_136–13insC variant was not associated with changes in *INSR* and *HMGA1* expressions in the Swedish sample (P > 0.12) (Supplementary Table 5).

DISCUSSION

It has been suggested that the candidate gene approach may be a practical and logical alternative to GWAS of complex genetic traits (35). In fact, the study of genes with functional polymorphisms and coding for proteins involved in disease-related pathways may increase statistical efficiency by minimizing multiple comparisons. Furthermore, the GWAS approach is still unable to reliably assess the association between rare variants and complex diseases such as type 2 diabetes (6,35–37). However, candidate gene studies have brought a majority of false-positive results in the past, and a stringent study design is necessary to reach a statistical power that allows valid conclusions.

Low-frequency variants with strong effects may contribute to type 2 diabetes risk through their impact on insulin resistance (9,17,38–41). Although the *HMGA1* gene was not part of the GWAS signals, there was some evidence for its role in insulin resistance and type 2 diabetes development (3,9,13,17,42–44).

However, our data suggest that low-frequency variants, including c.136–14_136–13insC, are not associated with type 2 diabetes in the French population. Furthermore, the c.136–14_136–13insC variant had no significant effects on intermediate quantitative traits. Given the low statistical power of our screening step, we cannot exclude that other rare variants increasing type 2 diabetes susceptibility may exist in the *HMGA1* gene. Therefore, we emphasize the need to sequence many more individuals to reliably identify rare variants associated with common diseases.

Our data are not in line with those recently reported by Chiefari et al. (17), who observed that the c.136–14_136–13insC variant was strongly associated with type 2 diabetes

TABLE 2
Effects of the c.136–14_136–13insC variant on metabolic traits in normoglycemic control subjects

Metabolic trait	−/−	−/C+C/C	P	Detectable variation
BMI (kg/m^2)*	27.1 ± 7.7 (n = 4,159)	27.7 ± 7.7 (n = 228)	0.36	1.5
Fasting plasma glucose (mmol/L)*	5.1 ± 0.5 (n = 4,159)	5.1 ± 0.5 (n = 228)	0.53	0.1
Fasting plasma insulin (mU/L)*	7.5 ± 5.9 (n = 4,115)	8.1 ± 7.9 (n = 227)	0.45	1.2
HOMA2-IR*	1.0 ± 0.6 (n = 3,731)	1.1 ± 0.6 (n = 205)	0.61	0.1
HOMA2-B (%)*	90.8 ± 35.5 (n = 3,731)	92.7 ± 35.4 (n = 205)	0.74	7.2
HbA$_{1c}$ (%)*	5.2 ± 0.4 (n = 2,529)	5.2 ± 0.4 (n = 141)	0.81	0.1
IGI*	39.5 ± 30.6 (n = 781)	41.1 ± 32.1 (n = 40)	0.35	14
ISI*	8.4 ± 7.1 (n = 804)	7.5 ± 4.6 (n = 41)	0.27	3.2
DI*	5.7 ± 15.4 (n = 804)	4.7 ± 8.9 (n = 41)	0.28	6.9
Waist-to-hip ratio	0.9 ± 0.1 (n = 3,377)	0.9 ± 1.0 (n = 183)	0.66	0.02

Quantitative traits were compared using linear regression models adjusted for age, sex, and BMI (when appropriate). Detectable variation: α = 0.05, 80% power. *Log-transformed to satisfy the assumption of normality.

risk in populations of European descent (OR 1.64–15.77). In our study (4,906 type 2 diabetic subjects vs. 4,387 normoglycemic subjects), we had a statistical power of 80% to detect a dominant OR of 1.29 for a MAF of 2.61% (type I error = 0.05). Furthermore, large meta-analyses in Europeans (including Italian control group 1 or 2) detected no association between the c.136–14_136–13insC variant (or correlated polymorphism) and type 2 diabetes. Of note, strong between-study heterogeneity was observed, mainly driven by the low allele frequencies in both groups of Italian control subjects (Fig. 2). This heterogeneity may be due to different inclusion criteria in the Italian, U.S., and French studies (questionnaire, personal interview, biological measures). In the largest Italian control group (control group 1), Chiefari et al. (17) conducted personal interviews to select those without a family history of type 2 diabetes or related conditions to reduce the number of people with insulin resistance. Apart from the fact that interviews are not always reliable, it seems unlikely that the low allele frequency in the Italian control group (MAF 0.22%) is only due to an absence of family history of type 2 diabetes. In fact, in our screening group of samples, a 13-fold higher allele frequency was observed in French control subjects with similar features (MAF 2.82%). Furthermore, the allele frequency of the c.136–14_136–13insC variant tended to be even higher in normoglycemic control subjects without a family history of type 2 diabetes than in type 2 diabetic case subjects with a strong family history of type 2 diabetes (MAF 1.49%). Finally, the effect size was not different between the screening group (ascertained for their family history of type 2 diabetes) and the replication group (not ascertained for their family history of type 2 diabetes) of French individuals (Woolf test: χ^2 = 0.1, df = 1, P = 0.76), contrary to what was observed in Italians by Chiefari et al. (17).

We observed nominal deviation from the Hardy-Weinberg equilibrium when case and control subjects were analyzed together (P = 0.02), reflecting an underrepresentation of C/C homozygotes. Chiefari et al. (17) reported that the absence of homozygous individuals for this variant in their study was consistent with previous observations indicating that mice homozygous for the *Hmga1* null allele have reduced fertility and die prenatally. However, the only homozygous carrier of the c.136–14_136–13insC C allele was aged 47 years, normoglycemic, without a family history of type 2 diabetes, and the mother of three children. Therefore, we hypothesize that their underrepresentation is rather due to sampling effects (modest sample size given the low allele frequency).

Although Chiefari et al. (17) reported strong associations between the c.136–14_136–13insC variant and type 2 diabetes (1.64 < OR < 15.77), no strong effects on type 2 diabetes–related intermediate traits were detected. These results are consistent with those of Chiefari et al. (17), who did not observe any variation in fasting insulin levels. They also showed a decreased expression of *HMGA1* and *INSR* in blood monocytes of type 2 diabetic subjects carrying the c.136–14_136–13insC variant compared with both type 2 diabetic and healthy noncarriers, whereas no association was identified in our study using subcutaneous adipose tissue of normoglycemic individuals. Chiefari et al. (17) compared a subset of case subjects carrying the c.136–14_136–13insC variant with subsets of wild-type case and control subjects. Their selection criteria were not specified, and no comparison between control subjects (carriers vs. noncarriers) was reported. Therefore, the expression variation detected in monocytes may have been biased. Furthermore, point mutations affecting *INSR* expression were previously described but not sequenced in these samples (45–47). Finally, the regulation of *INSR* and *HMGA1* expressions may not be the same in monocytes and in key tissues involved in insulin resistance (e.g., adipose tissue, liver, and skeletal muscle). However, additional European individuals need to be analyzed to confirm the lack of functional effects.

In conclusion, our data suggest that the c.136–14_136–13insC variant in the *HMGA1* gene is not associated with type 2 diabetes risk in Europeans, contrary to what was reported by Chiefari et al. (17). The rapid development of cost-efficient next-generation sequencing should allow reliable detection of associations between rare mutations and type 2 diabetes in large populations (48–50).

REFERENCES

1. Perry JR, Frayling TM. New gene variants alter type 2 diabetes risk predominantly through reduced beta-cell function. Curr Opin Clin Nutr Metab Care 2008;11:371–377
2. Watanabe RM. The genetics of insulin resistance: where's Waldo? Curr Diab Rep 2010;10:476–484
3. Voight BF, Scott LJ, Steinthorsdottir V, et al.; MAGIC investigators; GIANT Consortium. Twelve type 2 diabetes susceptibility loci identified through large-scale association analysis. Nat Genet 2010;42:579–589
4. Manolio TA, Collins FS, Cox NJ, et al. Finding the missing heritability of complex diseases. Nature 2009;461:747–753
5. Cirulli ET, Goldstein DB. Uncovering the roles of rare variants in common disease through whole-genome sequencing. Nat Rev Genet 2010;11: 415–425
6. Bodmer W, Bonilla C. Common and rare variants in multifactorial susceptibility to common diseases. Nat Genet 2008;40:695–701
7. Schork NJ, Murray SS, Frazer KA, Topol EJ. Common vs. rare allele hypotheses for complex diseases. Curr Opin Genet Dev 2009;19:212–219
8. Pritchard JK. Are rare variants responsible for susceptibility to complex diseases? Am J Hum Genet 2001;69:124–137
9. Foti D, Chiefari E, Fedele M, et al. Lack of the architectural factor HMGA1 causes insulin resistance and diabetes in humans and mice. Nat Med 2005; 11:765–773
10. Bustin M. Regulation of DNA-dependent activities by the functional motifs of the high-mobility-group chromosomal proteins. Mol Cell Biol 1999;19: 5237–5246
11. Reeves R, Beckerbauer L. HMGI/Y proteins: flexible regulators of transcription and chromatin structure. Biochim Biophys Acta 2001; 1519:13–29
12. Brunetti A, Manfioletti G, Chiefari E, Goldfine ID, Foti D. Transcriptional regulation of human insulin receptor gene by the high-mobility group protein HMGI(Y). FASEB J 2001;15:492–500
13. Foti D, Iuliano R, Chiefari E, Brunetti A. A nucleoprotein complex containing Sp1, C/EBP beta, and HMGI-Y controls human insulin receptor gene transcription. Mol Cell Biol 2003;23:2720–2732
14. Aiello A, Pandini G, Sarfstein R, et al. HMGA1 protein is a positive regulator of the insulin-like growth factor-I receptor gene. Eur J Cancer 2010; 46:1919–1926
15. Akaboshi S, Watanabe S, Hino Y, et al. HMGA1 is induced by Wnt/beta-catenin pathway and maintains cell proliferation in gastric cancer. Am J Pathol 2009;175:1675–1685
16. Perry JR, McCarthy MI, Hattersley AT, et al. Interrogating type 2 diabetes genome-wide association data using a biological pathway-based approach. Diabetes 2009;58:1463–1467
17. Chiefari E, Tanyolaç S, Paonessa F, et al. Functional variants of the HMGA1 gene and type 2 diabetes mellitus. JAMA 2011;305:903–912
18. Garg A. HMGA1, a novel locus for type 2 diabetes mellitus. JAMA 2011;305: 938–939
19. Sladek R, Rocheleau G, Rung J, et al. A genome-wide association study identifies novel risk loci for type 2 diabetes. Nature 2007;445:881–885
20. Lièvre M, Marre M, Chatellier G, et al.; DIABHYCAR Study Group. The non-insulin-dependent diabetes, hypertension, microalbuminuria or proteinuria, cardiovascular events, and ramipril (DIABHYCAR) study: design, organization, and patient recruitment. Control Clin Trials 2000;21:383–396
21. Marre M, Lievre M, Chatellier G, Mann JF, Passa P, Ménard J; DIABHYCAR Study Investigators. Effects of low dose ramipril on cardiovascular and renal outcomes in patients with type 2 diabetes and raised excretion of urinary albumin: randomised, double blind, placebo controlled trial (the DIABHYCAR study). BMJ 2004;328:495
22. Lafay L, Basdevant A, Charles MA, et al. Determinants and nature of dietary underreporting in a free-living population: the Fleurbaix Laventie Ville Santé (FLVS) Study. Int J Obes Relat Metab Disord 1997;21:567–573
23. Patterson K. 1000 genomes: a world of variation. Circ Res 2011;108:534–536
24. Wareham NJ, Phillips DI, Byrne CD, Hales CN. The 30 minute insulin incremental response in an oral glucose tolerance test as a measure of insulin secretion. Diabet Med 1995;12:931
25. Matsuda M, DeFronzo RA. Insulin sensitivity indices obtained from oral glucose tolerance testing: comparison with the euglycemic insulin clamp. Diabetes Care 1999;22:1462–1470
26. Bergman RN, Ader M, Huecking K, Van Citters G. Accurate assessment of beta-cell function: the hyperbolic correction. Diabetes 2002;51(Suppl. 1): S212–S220
27. Walley AJ, Jacobson P, Falchi M, et al. Differential co-expression analysis of obesity-associated networks in human subcutaneous adipose tissue. Int J Obes. 22 March 2011 [Epub ahead of print]
28. Abecasis GR, Cherny SS, Cookson WO, Cardon LR. Merlin: rapid analysis of dense genetic maps using sparse gene flow trees. Nat Genet 2002;30:97–101
29. Purcell S, Cherny SS, Sham PC. Genetic Power Calculator: design of linkage and association genetic mapping studies of complex traits. Bioinformatics 2003;19:149–150
30. Chen WM, Abecasis GR. Family-based association tests for genomewide association scans. Am J Hum Genet 2007;81:913–926

31. Higgins JP, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. BMJ 2003;327:557–560
32. Agresti A. *Categorical Data Analysis.* Hoboken, NJ, Wiley, 2002
33. DerSimonian R, Kacker R. Random-effects model for meta-analysis of clinical trials: an update. Contemp Clin Trials 2007;28:105–114
34. Cohen J. *Power Analysis for the Behavioral Sciences.* 2nd ed. Hillsdale, NJ, Lawrence Erlbaum, 1988
35. Tabor HK, Risch NJ, Myers RM. Candidate-gene approaches for studying complex genetic traits: practical considerations. Nat Rev Genet 2002;3:391–397
36. McCarthy MI, Abecasis GR, Cardon LR, et al. Genome-wide association studies for complex traits: consensus, uncertainty and challenges. Nat Rev Genet 2008;9:356–369
37. Barrett JC, Cardon LR. Evaluating coverage of genome-wide association studies. Nat Genet 2006;38:659–662
38. Sakul H, Pratley R, Cardon L, Ravussin E, Mott D, Bogardus C. Familiality of physical and metabolic characteristics that predict the development of non-insulin-dependent diabetes mellitus in Pima Indians. Am J Hum Genet 1997;60:651–656
39. Elbein SC, Hasstedt SJ, Wegner K, Kahn SE. Heritability of pancreatic beta-cell function among nondiabetic members of Caucasian familial type 2 diabetic kindreds. J Clin Endocrinol Metab 1999;84:1398–1403
40. Watanabe RM, Valle T, Hauser ER, et al. Familiality of quantitative metabolic traits in Finnish families with non-insulin-dependent diabetes mellitus. Finland-United States Investigation of NIDDM Genetics (FUSION) Study investigators. Hum Hered 1999;49:159–168
41. Lehtovirta M, Kaprio J, Forsblom C, Eriksson J, Tuomilehto J, Groop L. Insulin sensitivity and insulin secretion in monozygotic and dizygotic twins. Diabetologia 2000;43:285–293
42. Chiefari E, Iiritano S, Paonessa F, et al. Pseudogene-mediated post-transcriptional silencing of HMGA1 can result in insulin resistance and type 2 diabetes. Nat Commun 2010;1:40
43. Chiefari E, Paonessa F, Iiritano S, et al. The cAMP-HMGA1-RBP4 system: a novel biochemical pathway for modulating glucose homeostasis. BMC Biol 2009;7:24
44. Darville MI, Terryn S, Eizirik DL. An octamer motif is required for activation of the inducible nitric oxide synthase promoter in pancreatic beta-cells. Endocrinology 2004;145:1130–1136
45. Taylor SI, Accili D, Imai Y. Insulin resistance or insulin deficiency: which is the primary cause of NIDDM? Diabetes 1994;43:735–740
46. Goldfine ID. The insulin receptor: molecular biology and transmembrane signaling. Endocr Rev 1987;8:235–255
47. Taylor SI, Cama A, Accili D, et al. Mutations in the insulin receptor gene. Endocr Rev 1992;13:566–595
48. Mardis ER. Next-generation DNA sequencing methods. Annu Rev Genomics Hum Genet 2008;9:387–402
49. Shendure J, Ji H. Next-generation DNA sequencing. Nat Biotechnol 2008; 26:1135–1145
50. Guey LT, Kravic J, Melander O, et al. Power in the phenotypic extremes: a simulation study of power in discovery and replication of rare variants. Genet Epidemiol. 2011 Feb 9 [Epub ahead of print]