A genome-wide association study using a custom genotyping array identifies variants in *GPR158* associated with reduced energy expenditure in American Indians

Paolo Piaggi[1,2], Ivica Masindova[1,3], Yunhua L. Muller[1], Josep Mercader[4,5,6], Gregory B. Wiessner[1], Peng Chen[1], SIGMA Type 2 Diabetes Consortium[6], Sayuko Kobes[1], Wen-Chi Hsueh[1], Milliejoan Mongalo[1], William C. Knowler[1], Jonathan Krakoff[1], Robert L. Hanson[1], Clifton Bogardus[1], Leslie J. Baier[1]

[1]: *Phoenix Epidemiology and Clinical Research Branch, National Institute of Diabetes and Digestive and Kidney Disease, National Institutes of Health, Phoenix, AZ, USA.*
[2]: *Obesity Research Center, Endocrinology Unit, University Hospital of Pisa, Pisa, Italy*.
[3]: *Laboratory of Diabetes and Metabolic Disorders, Institute of Experimental Endocrinology, Biomedical Research Center, Slovak Academy of Sciences, Bratislava, Slovakia*.
[4]: *Diabetes Unit and Center for Genomic Medicine, Massachusetts General Hospital, Boston, Massachusetts, USA.*
[5]: *Programs in Metabolism and Medical & Population Genetics, Broad Institute of Harvard and MIT, Cambridge, Massachusetts, USA.*
[6]: *Members of the SIGMA Type 2 Diabetes Consortium are provided in Appendix S1*.

Last Names: Piaggi, Masindova, Muller, Mercader, Wiessner, Chen, Kobes, Hsueh, Mongalo, Knowler, Krakoff, Hanson, Bogardus, Baier.

Disclosure Statement: The authors have nothing to disclose.

Abbreviated title: Variants in *GPR158* associate with EE

Keywords: energy expenditure, GWAS, BMI, body fat.

Word count: 4000

Number of figures and tables: 8

References: 46

Corresponding author: Paolo Piaggi, Ph.D., Diabetes Molecular Genetics Section, Phoenix Epidemiology and Clinical Research Branch, National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK), National Institutes of Health (NIH), 445 N. 5th St., Phoenix, AZ 85004.

E-mail: paolo.piaggi@gmail.com, paolo.piaggi@nih.gov

ClinicalTrials.gov identifier: NCT00340132.

Funding: This study was supported by the Intramural Research Program of NIDDK, NIH.

Abbreviations:

1000G: 1000 Genomes Project

EE: energy expenditure

FM: fat mass

FFM: fat free mass

GWAS: genome-wide association study

MAF: minor allele frequency

OGTT: oral glucose tolerance test

RAF: risk allele frequency

RMR: resting metabolic rate

SNP: single nucleotide polymorphism

WGS: whole-genome sequence

Abstract

Pima Indians living in Arizona suffer from a high prevalence of obesity, and we have previously shown that a relatively lower energy expenditure (EE) predicts weight and fat mass gain in this population. Energy expenditure (EE) is a familial trait (heritability=0.52); therefore, in the current study, we aimed to identify genetic variants that affect EE and thereby influence BMI and body fatness in Pima Indians.

Genotypic data from 491,265 variants were analyzed for association with resting metabolic rate (RMR) and 24-h EE assessed in a whole-room calorimeter in 507 and 419 Pima Indians, respectively. Variants associated with both measures of EE were analyzed for association with maximum BMI and percent body fat (PFAT) in 5870 and 912 Pima Indians, respectively.

Rs11014566 nominally associated with both measures of EE and both measures of adiposity in Pima Indians, where the G-allele (frequency: Pima Indians=0.60, Europeans<0.01) associated with lower 24-h EE (β=−33 kcal/day per copy), lower RMR (β=−31 kcal/day), higher BMI (β=+0.6 kg/m^2) and higher PFAT (β=+0.9%). However, the association of rs11014566 with BMI did not directionally replicate when assessed in other ethnic groups. Rs11014566 tags rs144895904, which affected promoter function in an *in vitro* luciferase assay. These variants map to *GPR158*, which is highly expressed in the brain and interacts with two other genes (*RGS7* and *CACNA1B*) known to affect obesity in knock-out mice.

Our results suggest that common ethnic-specific variation in *GPR158* may influence EE; however, its role in weight gain remains controversial since it either had no association with BMI, or associated with BMI but in the opposite direction, in other ethnic groups.

INTRODUCTION

Obesity often aggregates in families. Household members typically share lifestyle factors including food choices, daily habits, and cultural views which may affect body weight; however, studies in twins reared apart have provided evidence that approximately ⅔ of the variability of BMI is attributable solely to genetics (1; 2). BMI is influenced by both energy intake and energy expenditure (EE). In Pima Indians living in Arizona we have shown that a relatively low EE predicts increases in body weight (3-5) and fat mass (4) over time; however, this inverse relationship is not observed in all populations (6-8) whereas a positive relationship has been reported in an African population (9). EE varies by age and sex, but its largest determinant is body size and composition, particularly fat free mass (FFM) as an indicator of the metabolically active tissue, which accounts for ~80% of the variance in resting metabolic rate (RMR) and 24-h EE (10). However, in addition to age, sex and body composition, twin studies have also demonstrated that genetics contributes to the inter-individual variance in EE (11; 12). Taken together, these prior studies indicate that genetics has a small but measurable effect on EE, and in Pima Indians a lower EE predicts weight gain. Therefore, identification of genetic variants that influence EE may uncover new metabolic pathways that affect body weight/fatness. The aim of the current study is to estimate the heritable portion of EE and BMI in a family-based sample of Pima Indians, and then perform a genome-wide analysis of variants in non-diabetic Pima Indians to identify genetic variation that associates with EE and BMI/body fatness in a fashion consistent with the putative mechanistic relationship (e.g., low EE and high BMI).

RESEARCH DESIGN AND METHODS

Population-based subjects with outpatient longitudinal measures of BMI

The study subjects reside in an American Indian community near Phoenix, Arizona, where most individuals are of Pima Indian heritage. From 1965 to 2007, volunteers from this community participated in a longitudinal study of type 2 diabetes where anyone age ≥5 years was invited for

biennial health examinations (13). Subjects were asked to fast prior to these exams, and glucose tolerance was assessed by a 75g oral glucose tolerance test (OGTT). Height and weight were also measured to calculate BMI. Data on maximum BMI, defined as the highest BMI recorded at a medical exam when the subject was ≥15 years of age and was determined to be free from diabetes according to ADA diagnostic criteria (14), was available for 5870 subjects (Table 1). Among these subjects, 2920 were full-heritage Pima Indian (defined as $8/8^{th}$ Pima Indian heritage by self-report), and the remaining 2950 were mixed-heritage, on average, $6/8^{th}$ American Indian (typically $4/8^{th}$ Pima Indian and an additional $2/8^{th}$ from other related tribes). Before participation, volunteers were fully informed of the nature and purpose of the studies, and written informed consent was obtained. The protocols were approved by the Institutional Review Board of the National Institute of Diabetes and Digestive and Kidney Diseases.

Inpatient subjects with measures of body composition and EE

Among the community members from the longitudinal study, a subset of 917 adults, who were confirmed to be non-diabetic by the OGTT, also participated in inpatient studies in our Clinical Research Center and had undergone detailed measures of body composition. Among these inpatient volunteers, 509 also underwent a measurement of RMR by a ventilated hood system and 419 underwent a 24-h session in a whole-room indirect calorimeter (352 subjects underwent both measures of RMR and 24-h EE during the same admission).

Following admission to the Clinical Research Center, subjects were given a standard weight-maintaining diet (50% carbohydrates, 30% fats and 20% proteins) for three days before any metabolic test was performed (15; 16). Subjects were weighed daily and food intake was adjusted to maintain body weight within ±1% of the weight measured the second day of admission. Percent body fat (PFAT), fat mass (FM) and fat free mass (FFM) were estimated by underwater weighing until August 1993 and thereafter by total body dual energy X-ray absorptiometry (DPX-1; Lunar Radiation Corp.,

Madison, WI, USA). A conversion equation was used to make measurements of body composition comparable between the two methods (17).

RMR was measured upon awakening after an overnight fast using a respiratory hood system, as previously described (18). After 10 min of acclimation to the plastic hood, subject's EE was calculated every 5 min using the equations of Lusk (19) and RMR was calculated as the average EE over 40 minutes while the subject was instructed to stay awake and motionless, and then extrapolated to 24 hours.

Twenty-four hour EE and substrate oxidation were measured in a whole-room calorimeter (respiratory chamber), as previously described (10). The volunteers entered the chamber at 08:00 and remained in the chamber for 23 hours and 15 min. The rate of EE was measured continuously, calculated for each 15-min interval, averaged and then extrapolated to the 24-h interval (24-h EE). Four meals were provided at 08:00, 11:00, 16:00, and 19:00 and the total energy content was calculated using previously described equations (20). Spontaneous physical activity (SPA) was detected by radar sensors and expressed as the percentage of time over the 15-min interval in which activity was detected (21). The EE in the inactive awake state ($EE_{0\ activity}$) was calculated as the intercept of the regression line between EE and SPA between 11:00 and 01:00 (22). Sleeping metabolic rate (SMR) was defined as the average EE of all 15-min nightly periods between 01:00 and 05:00 AM during which SPA was less than 1.5% (23). The "awake and fed" thermogenesis (AFT) was calculated as the difference between $EE_{0\ activity}$ and SMR (23).

Genotypic data for genome-wide association analysis

Genotypes for association analyses were generated using a custom Pima Indian Axiom genome-wide array (Affymetrix, Santa Clara, CA) in 7701 Pima Indian samples. This array was designed to capture common variation (minor allele frequency, MAF $\geq$0.05, or $\geq$0.01 for coding variants) detected in whole-genome sequence (WGS) data of 266 full-heritage Pima Indians from different nuclear families. We estimated that genotypes for the 491,265 array markers that passed quality control

metrics (i.e., call rate ≥90%, discrepant rate ≤2 pairs among 100 blind duplicate pairs, and lack of deviation from Hardy-Weinberg equilibrium with a $p>10^{-4}$) tag 92% of the 4.9M common variants with a MAF ≥0.05 detected in the genomes of full-heritage Pima Indians (tag defined as $r^2 \geq 0.85$ within 300 kb).

Functional analysis of *GRP158* variants

DNA fragments containing each allele homozygous at rs11014566, rs144895904, rs34673593 and rs16925884 were PCR amplified (rs11014566, primers forward 5'-ACAGGTACCATTTGTGTTAACGGCTAGA-3' and reverse 5'-TCACTCGAGGTATAAACAATTTTGCCAT-3'; rs144895904, forward 5'-ATAGGTACCAGAGATAACCGCTGTTCA-3 and reverse 5'-TCACTCGAGAGGCACAAATTACATAAC-3'; rs16925884/rs34673593, forward 5'-ACGGTACCTACTATTTGTTGTGAG-3' and reverse 5'-AGCTCGAGATATAAATGAATGAATTG-3'). The amplicons were inserted at *KpnI* and *XhoI* sites (underlined, respectively) upstream of the pGL3-Promoter firefly luciferase reporter vector (Promega, Madison, WI, USA). DNA constructs were sequenced to confirm the nucleotide variants. Murine N-42 hypothalamus cell line (Cellutions Biosystems, Inc., Burlington, ON, Canada) was maintained in DMEM medium supplemented with 10% FBS, 1% penicillin-streptomycin (ATCC) at 37°C, 5% CO_2 and 95% air atmosphere. One µg of DNA construct and 125 ng of pGL4-*Renilla* luciferase reporter vector (Promega, Madison, WI, USA) were transiently transfected into the cells with Lipofectamine LTX (Invitrogen, Life Technologies, Carlsbad, CA). At 48 hours post-transfection, cells were harvested and a dual-luciferase reporter assay was performed using a standard protocol (Promega, Madison, WI, USA). Three separate transfections were conducted and each transfection was repeated two to three times (total=8) and data were averaged. Firefly luciferase activity was normalized to *Renilla* luciferase activity and further normalized to pGL3-Promoter luciferase activity.

Statistical analysis

The variances in 24-h EE and maximum BMI attributable to family membership were estimated in families with at least two siblings by mixed models analysis and quantified by the root-mean-square deviation (RMSD) and by the intraclass correlation coefficient (ICC). Heritability was estimated in a linear mixed model from a random effect that utilized the empirical genetic relatedness matrix (see below). Linear mixed effects analysis using the maximum likelihood method was conducted to assess association of genotypes with 24-h EE, RMR, maximum BMI and PFAT with covariates of age, sex, body composition measures (FM and FFM, only for 24-h EE and RMR analyses), SPA (only for 24-h EE analysis), birth year (only for BMI analysis) and the first 5 genetic principal components calculated from 19,991 variants randomly selected from 200-kb windows across the genome (1 variant per window). Genotype was included as a fixed effect and analyzed as a numeric variable representing 0, 1, or 2 copies of a given allele (additive model) and effects were expressed per allele copy. Missing genotypes were imputed using WGS data of 266 full-heritage Pima Indians. The models were fitted using a variance components covariance structure to account for genetic relatedness among individuals. The genetic relatedness matrix was estimated as the proportion of the genome shared identical by descent between each pair of individuals who had been genotyped (a total of 29,648,850 pairs). Genomic segments shared identical by descent were identified with the *fastIBD* function of Beagle package (24) using 482,616 autosomal markers with MAF>0.05. Mixed models were fit using the SOLAR package (25). Values of BMI were log-transformed before analysis to approximate a Gaussian distribution. Linkage disequilibrium was determined using the Haploview program (version 4.2, Broad Institute, Cambridge, MA, USA). Tag variants were selected based on the sequence data of 266 Pima genomes using the Tagger algorithm (Haploview) with a pair-wise $r^2 \geq 0.85$ taken as indicative of redundancy. The statistical difference in mean luciferase activity detected in the functional study was analyzed by Student's unpaired *t*-test.

To estimate the statistical power to detect a physiologically meaningful effect of a common genetic variant on 24-h EE that meets genome-wide statistical significance, we estimated power for a sample size of 419 unrelated individuals, a 2-sided alpha=5×10^{-8}, a clinically significant effect size $\beta = -50$

kcal/day per risk allele copy, and a residual SD=143 kcal/day (after adjustment for age, sex, FM, FFM and SPA, Table 1). Using these parameters, we estimated the power to be 0.03 or 0.38 to detect a risk allele frequency (RAF, defined for the allele associated with lower EE) of 0.15 or 0.50, respectively. To reduce the chance of spurious findings (type 1 error) without undue reliance on a single EE measure in a setting of low statistical power, we selected variants with consistent evidence of association in two separate EE assessments, namely, 24-h EE and RMR, where each *p*-value was <0.01 and the direction of risk was consistent (i.e., the risk allele associated with lower 24-h EE being associated with lower RMR). Variants meeting these criteria were then analyzed for association with BMI and PFAT.

Replication Cohort

Replication of selected variants for their association with standardized values of BMI was done in non-diabetic individuals of the Slim Initiative in Genomic Medicine for the Americas (SIGMA) consortium (26) after adjustment for age, sex and first two genetic principal components. This replication sample consisted in four Studies from Mexico or Mexicans living in the US comprising a total of 4,364 non diabetic individuals. All participants provided informed consent for conducting this study. Their respective local ethics committees approved all contributing studies.

RESULTS

Estimates of familial effect on 24-h EE and BMI in Pima Indians

In 248 siblings from 98 Pima Indian families, family membership explained 41% of variance in unadjusted 24-h EE (RMSD=250 kcal/day, $p<0.001$). After adjustment for subject's age, sex, FM, FFM and SPA, family membership was still an independent determinant of 24-h EE (RMSD=77 kcal/day, $p<0.001$), accounting for 34% of the unexplained variance in 24-h EE (Figure 1A). In 3298 siblings from 1131 Pima Indian families, family membership was the largest determinant of maximum BMI, accounting for ⅓ of BMI variance among individuals (RMSD=4.4 kg/m^2, $p<0.001$,

Figure 1B). Inclusion of age, sex, birth year and Pima heritage did not alter the estimate of family variance and slightly increased the explained variance of maximum BMI from 30% to 36% ($p<0.001$). Heritability of 24-h EE and maximum BMI was 0.52 (95% CI: 0.20-0.80, $p=1.5\times10^{-3}$, adjusted for age, sex, FM, FFM, SPA and first 5 principal components) and 0.55 (95% CI: 0.51-0.60, $p=8.2\times10^{-148}$, adjusted for age, sex, birth year and first 5 principal components), respectively.

Association analysis for two independent measures of EE

The results of the genome-wide association analyses for 24-h EE and RMR are shown in Figure 2. The lists of variants with $p<0.01$ for each EE measure are reported in Supplemental Tables 1 and 2 (effect size for 24-h EE and RMR ranging from −96 to −21 kcal/day and from −132 to −25 kcal/day, respectively). As anticipated, no variant achieved genome-wide statistical significance ($p<5\times10^{-8}$) with either EE measure. However, 138 variants had nominal ($p<0.01$), directionally consistent associations with both 24-h EE and RMR (Supplemental Table 3).

Association analysis for two measures of body adiposity

The 138 variants that associated with both 24-h EE and RMR in a directionally consistent manner were further analyzed for association with maximum BMI (defined as the highest BMI recorded at a longitudinal outpatient exam) in a population-based sample of 5870 individuals, and with PFAT in a 917 subjects who had undergone metabolic testing as inpatients. While seven variants had an allele that associated with a reduced EE and were nominally associated with higher maximum BMI (Table 2) and five variants had an allele that associated with reduced EE and higher PFAT, only the variant with the strongest association with maximum BMI (rs11014566 in *GPR158*, $p=4.7\times10^{-4}$) also associated with increased PFAT ($p=2.9\times10^{-3}$).

The G allele at rs11014566 (frequency in full-heritage Pima Indians=0.60) associated with a lower 24-h EE (β=−33 kcal/day, 95% CI: −54 to −12, Figure 3B) and a lower RMR (β=−31 kcal/day, 95% CI: −55 to −7, Figure 3C). Compared with subjects homozygous for the A-allele, subjects carrying

two copies of the G-allele had lower EE over the course of 24 hours inside the metabolic chamber (Δ=−2.6 kcal/h, p=7.7×10^{-3}) and this was more evident in the sleeping state (Figure 3A). Accordingly, SNP rs11014566 was associated with sleeping EE (β=−24 kcal/day, p=1.4×10^{-2}, Figure 3D) but not with AFT (p=0.59, Figure 3E) or SPA (p=0.59). In the larger population-based sample of 5870 Pima Indians, the G-allele also associated with a higher maximum BMI (β =+1.7%≈0.6 kg/m^2 per copy, 95% CI: 0.7 to 2.6, p=4.7×10^{-4}, Figure 4A), and this association was observed both in the 2920 individuals who were full-heritage Pima Indians (β=+1.4%, p=2.7×10^{-2}, Figure 4B) and in the 2950 individuals who were mixed-heritage American Indians (β=+2.0%, p=3.4×10^{-3}, Figure 4C) in this population. This variant also associated with PFAT in 917 subjects with body composition measures (β=+0.9%, 95% CI: 0.3 to 1.5, p=2.9×10^{-3}, Figure 4D) with no difference between sexes (p=0.39). Specifically, the G-allele was associated with higher FM (β=+2.4 kg, p=2.5×10^{-4} adjusted for age, sex and height, Figure 4E) and, to a lesser extent, with higher FFM (β=+1.3 kg, p=1.7×10^{-2}, Figure 4F).

Analysis of additional variation at the *GPR158* locus

Analyses of WGS data from 266 full-heritage Pima Indians showed that rs11014566 tags (r^2>0.85) three other variants: rs144895904 (C/T, frequency T=0.61, r^2=0.99), rs34673593 (−/AT, frequency AT=0.57, r^2=0.87), and rs16925884 (C/T, frequency T=0.60, r^2=0.91), all in intron 4 of *GPR158*. Analysis of 74 tagging variants with a MAF≥0.05 across the *GPR158* gene (50 kb flanking each side, chr10:25,414,290-25,941,157) determined that rs11014566 (and its 3 tags) had the strongest association with maximum BMI, and conditional analyses demonstrated no variant in this region associated with maximum BMI after conditioning on rs11014566 (all conditioned p>0.05).

The BMI risk alleles for rs11014566 and its tags show large differences among our data and populations in the 1000 Genomes Project (1000G). For example, in our data the G-allele at rs11014566 attains the highest frequency of 0.60 in full-heritage Pima Indians, 0.48 in mixed-heritage

American Indians, while in the 1000G its frequency was 0.23 in Americans, 0.11 in Africans and <0.01 in Europeans (Figure 5).

Replication analysis in SIGMA

Since no other datasets exist with genotypic data on individuals with measures of EE, we sought to replicate our modest association with BMI in the SIGMA consortium. In the BMI meta-analysis of the SIGMA consortium including 4364 non-diabetic Mexican individuals (mean±SD, age: 57.9±8.4 yr, BMI: 27.5±4.2 kg/m^2, 1755 males), the BMI-risk alleles at both rs11014566 (β=−0.05 SD units per copy of the G-allele, frequency=0.27, p=0.04) and rs144895904 (β=−0.05 SD units per copy of the T-allele, frequency=0.28, p=0.01) were associated with lower BMI in this cohort. Similar results were obtained in sensitivity analyses including only subjects with a BMI greater than the median value of this cohort (=27 kg/m^2) or including only obese subjects with a BMI>30 kg/m^2 (data not shown).

***In vitro* functional analyses of *GPR158* variants**

To determine whether rs11014566 in *GPR158* and the three variants tagged by rs11014566 had a functional impact on promoter activity, these variants were analyzed in an *in vitro* luciferase reporter assay. DNA regions containing either the risk or the non-risk allele for each variant were PCR amplified. Due to the proximity of rs16925884 (chr10:25,740,897) and rs34673593 (chr10:25,741,140), single PCR products containing either the risk alleles or the non-risk alleles for both variants were amplified. The effect of the cloned *GPR158* variants on promoter activity was assessed in a murine hypothalamus cell line, since *GPR158* is endogenously expressed in human hypothalamus at high levels (Supplemental Figure 1). The largest difference in luciferase activity was observed when comparing constructs which differed for alleles at rs144895904 (Figure 6), where the BMI risk allele T had on average 48% higher activity as compared to the non-risk allele C

(mean±SEM, 1.21±0.09 vs. 0.82±0.07, p=0.004). There was no significant difference in luciferase activity between alleles at rs11014566 (p=0.52) or rs16925884/rs34673593 (p=0.74).

DISCUSSION

The current study was conducted in a geographically confined population of Pima Indians and, among these community members, we estimated that family membership accounts for approximately ⅓ of BMI variance among siblings from different families. Siblings share, on average, half of their genes; therefore, nearly approximately 60% of BMI variance in this Pima Indian population is genetically determined, as confirmed by our empirical heritability estimate (h^2=0.55). This estimate is consistent with that reported in other ethnic groups (27) as well as prior studies in twins (1; 2; 28), which similarly estimated that 60% of the variability in BMI among individuals of a given population living in the same environment is genetically determined and potentially ascribable to the additive effects of genetic variants. We furthered showed in Pima Indians that 24-h EE, a determinant of BMI in this population, is also an inherited characteristic (h^2=0.52). After adjustment for differences in body composition, family membership accounted for 34% of the variance in 24-h EE among siblings from different families, which is consistent with previous calculations done in much smaller cohorts of Pima Indians (3; 18).

Given that BMI, body fat (29) and EE (3; 18) are genetically determined, and body weight and FM gain in Pima Indians is at least partially attributable to a relatively lower EE (3; 4; 23), we sought to identify genetic variants that influence body fatness and BMI in adulthood via a modest but life-long effect on EE. We performed a genome-wide association study (GWAS) for EE utilizing genotypic data from our custom Pima Indian-specific array. Although our sample of 419 Pima Indians with measures of 24-h EE and genotypes represents one of the largest existing samples, it was nonetheless underpowered to detect the modest effect sizes typically observed in GWAS at genome-wide statistical significance (p=5×10^{-8}). Therefore, rather than rely solely on statistical significance to discern true from false positives, we prioritized variants that showed physiologically supportive

associations with reduced EE (assessed by separate measurements of 24-h EE and RMR) and increased body adiposity (assessed by independent measurements of BMI and body fatness). This strategy led us to identify common variation in the *GPR158* gene that satisfied these criteria. Specifically, despite higher FFM (+2.6 kg) which would generally confer higher EE due to the well documented positive association with FFM (10), Pima individuals carrying two copies of the G-allele at rs11014566 in *GPR158* had instead on average roughly a 70-kcal deficit in daily EE (of which 48 kcals were ascribable solely to sleeping EE) and approximately 5 kg more FM and a BMI increase of 1.2 kg/m^2 as compared to subjects homozygous for the A-allele. The effect of rs11014566 on BMI in Pima Indians (1.2 kg/m^2 difference between individuals homozygous for the risk vs. non-risk allele, RAF=0.60) is comparable to the effects exerted by other variants near well-established obesity genes including rs8050136 in *FTO* (1.6 kg/m^2, RAF≈0.15) (30), rs74861148 near *MC4R* (1.36 kg/m^2, RAF≈0.15) (31) and rs2025804 in *LEPR* (1.0-1.9 kg/m^2, RAF≈0.70) (32) in this same population. Similarly, the effect of rs11014566 in *GPR158* on 24-h EE (−33 kcal/day per allele copy) is comparable to that of rs11208654 in *LEPR* (−28 kcal/day, Supplemental Table 1), which tags rs2025804 previously shown to affect 24-h EE in Pima Indians (32).

Given our low statistical power due to the modest sample size, our GWAS results for EE must be interpreted with caution. Nevertheless, the high heritability of 24-h EE in the Pima Indian population increases the likelihood that variants exerting true effects on EE could be uncovered using a GWAS strategy. To identify true from false positives among variants that did not achieve genome-wide significance, we considered variants that showed directional consistency for their associations with two independent assessments of EE, including precise and reproducible measures obtained at rest while fasting (18) and over 24 hours during energy balance (10), assuming that true genetic associations with EE will display weak but consistent results in both settings.

Although the strength of our study is that it provides the first genome-wide screen for genetic variants that affect EE, it also has a major weakness in that there are no other genetic databases available for

EE to directly assess replication. Since EE has a modest but measurable effect on weight gain in Pima Indians (4), and in Pima Indians variants in *GPR158* nominally associated with adiposity, we sought replication for the association of *GPR158* with BMI, as a surrogate of EE, in other ethnic groups. The G allele at rs11014566, which predicts lower EE and higher FM and BMI, has a frequency of 0.60 in full-heritage Pima Indians, which is higher than the frequency for this allele observed among any of the 1000G populations. Notably, this allele is uncommon in Europeans (MAF<0.01) and thus its assessment for association with BMI in the GIANT (Genetic Investigation of ANthropometric Traits) consortium (43) is not optimal. Therefore, we assessed association with BMI in datasets collected from Asians (44) and Africans (45) and Hispanics (26). Rs11014566 did not associate with BMI in Asians or Africans (personal communications); however, modest associations with BMI were observed in the SIGMA meta-analysis of BMI whose Hispanic population more closely resembles the Pima Indians from an environmental perspective, although from a genetic perspective the frequency of the G (rs11014566) and T (rs144895904) alleles are 0.27 and 0.28 in Hispanics as compared to 0.60 in Pima Indians. However, in the SIGMA sample, the direction of the association with BMI was opposite to that observed in the Pima Indians (β=−0.05 SD units per copy of the G-allele at rs11014566, p=0.04) and (β=−0.05 SD units per copy of the T-allele at rs144895904, p=0.01). Given that a relationship between EE and BMI has not been shown in the SIGMA sample, while metabolic studies in other ethnic groups have reported no relationship (6-8) and even a positive relationship between EE and future weight gain (9), as opposed to the inverse association observed in American Indians (3-5), it is unclear whether an association with BMI in the opposite direction indicates that this SNP has no role in EE (i.e., our result is a false positive) or whether it indicates the complexity of feedback loops between EE and food consumption (33-35), where an imbalance predicts either weight gain or weight loss, among individuals with different body habitus and living in different environments. As additional data for genotype and EE become available in other populations, meta-analyses may be helpful to assess the extent to which our findings transfer to other populations and to boost power to detect additional variants.

The GWAS lead SNP rs11014566 maps to an intron of the *GPR158* gene that encodes the G protein-coupled receptor 158, a transmembrane protein highly expressed in brain cells. Our tissue expression profiling confirmed that *GPR158* is highly expressed in the brain. *In vitro* functional analysis of rs11014566 and three variants tagged by rs11014566 (rs144895904, rs16925884 and rs34673593) in murine hypothalamic cells showed that intronic SNP rs144895904 had a statistically significant effect on promoter activity. Human *GPR158* is involved in neurotransmitter signaling and regulation of neuronal excitability (36; 37). Although there is no direct evidence that human *GPR158* is involved in the pathophysiology of obesity, a previous study of mouse Gpr158 has demonstrated its role in the regulation of energy balance (38). GPR158 binds to the regulator of G-protein signaling 7 (RGS7) in the nervous system (39; 40). RGS7-deficient mice are protected from obesity (41) and previous studies have provided evidence that *RGS7* may constitute an obesity locus in humans (42; 43). In addition to *RGS7*, GPR158 also binds to an N-type voltage-gated calcium channel (CACNA1B) in the rat brain (44). Homozygous CACNA1B-deficient mice gain less weight during 8 weeks of high-fat diet despite similar food intake of wild-type mice (45), implying a compensatory increase in EE that may mitigate weight gain during high-fat feeding. Since *GPR158, RGS7* and *CACNA1B* are all expressed in the central nervous system, as are the well-established human obesity genes *FTO*, *MC4R*, and *TMEM18* (46), one could speculate that they too exert an effect on hypothalamic signaling to regulate energy balance. However, additional mechanistic studies are needed to clarify the physiologic pathway whereby *GPR158* may affect EE and obesity in humans.

In conclusion, analysis of genotypes from a custom Pima Indian array identified a novel genetic locus in *GPR158* affecting EE and predisposing Pima Indians to weight gain. The frequency of the risk allele is higher in Pima Indians as compared to other populations studied as part of the 1000G, and the risk allele demonstrated increased promoter activity in *in vitro* experiments. Results of this study support the hypothesis that Pima Indians may carry some genetic variants affecting EE that are enriched in this particular ethnic group; however, the effect of these variants on higher rates of weight gain remains controversial since an association with BMI was only observed in Pima Indians. We

propose that studies of *GPR158*, as well as other EE-associated genes which will be identified in the future, may shed light into the pathophysiologic mechanisms that affect EE which could eventually lead to prevention and/or possible treatments of human obesity.

ACKNOWLEDGMENTS

The authors would like to thank Kari North, Maggie Ng and Misa Graff for providing look up data in their African American cohort and Xiao Ou Shu for providing look up data in an East Asian cohort. We would also like to thank the nursing, clinical and dietary staff and laboratory technicians of the Phoenix Epidemiology and Clinical Research Branch for conducting this study.

Funding. This research was supported by the Intramural Research Program of the National Institutes of Health, National Institute of Diabetes and Digestive and Kidney Diseases. This study utilized the computational resources of the Biowulf system at the National Institutes of Health, Bethesda, MD. None of the authors reported a potential conflict of interest.

Author Contributions. P.P. designed the study, analyzed and interpreted data and wrote the manuscript. I.M. and Y.L.M. performed the *in vitro* experiments and wrote the related procedures in the manuscript. J.M. performed statistical analyses for replication in the SIGMA Type 2 Diabetes Consortium. G.B.W. and M.M. performed the *in vitro* experiments. P.C., S.K., W.H. and R.L.H. analyzed data and provided statistical advice. W.C.K, J.K., R.L.H. and C.B. obtained the clinical and physiologic data, reviewed the manuscript and assisted with the interpretation of the data. L.J.B. designed the study, assisted with the interpretation of the data, edited and reviewed the manuscript. All authors critically revised the draft and approved the final manuscript. P.P. is the guarantor of this work and, as such, had full access to all the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

Prior Presentation. Parts of this study were presented in abstract form at the 2015 American Society of Human Genetics (ASHG) Annual Meeting, Baltimore, MD, 6–10 October 2015.

REFERENCES

1. Stunkard AJ, Harris JR, Pedersen NL, McClearn GE: The body-mass index of twins who have been reared apart. New England Journal of Medicine 1990;322:1483-1487

2. Price RA, Gottesman, II: Body fat in identical twins reared apart: roles for genes and environment. Behav Genet 1991;21:1-7

3. Ravussin E, Lillioja S, Knowler WC, Christin L, Freymond D, Abbott WG, Boyce V, Howard BV, Bogardus C: Reduced rate of energy expenditure as a risk factor for body-weight gain. The New England journal of medicine 1988;318:467-472

4. Piaggi P, Thearle MS, Bogardus C, Krakoff J: Lower energy expenditure predicts long-term increases in weight and fat mass. The Journal of clinical endocrinology and metabolism 2013;98:E703-707

5. Tataranni PA, Harper IT, Snitker S, Del Parigi A, Vozarova B, Bunt J, Bogardus C, Ravussin E: Body weight gain in free-living Pima Indians: effect of energy intake vs expenditure. International journal of obesity and related metabolic disorders : journal of the International Association for the Study of Obesity 2003;27:1578-1583

6. Anthanont P, Jensen MD: Does basal metabolic rate predict weight gain? The American journal of clinical nutrition 2016;104:959-963

7. Seidell JC, Muller DC, Sorkin JD, Andres R: Fasting respiratory exchange ratio and resting metabolic rate as predictors of weight gain: the Baltimore Longitudinal Study on Aging. International journal of obesity and related metabolic disorders : journal of the International Association for the Study of Obesity 1992;16:667-674

8. Weinsier RL, Nelson KM, Hensrud DD, Darnell BE, Hunter GR, Schutz Y: Metabolic predictors of obesity. Contribution of resting energy expenditure, thermic effect of food, and fuel utilization to four-year weight gain of post-obese and never-obese women. The Journal of clinical investigation 1995;95:980-985

9. Luke A, Durazo-Arvizu R, Cao G, Adeyemo A, Tayo B, Cooper R: Positive association between resting energy expenditure and weight gain in a lean adult population. The American journal of clinical nutrition 2006;83:1076-1081

10. Ravussin E, Lillioja S, Anderson TE, Christin L, Bogardus C: Determinants of 24-hour energy expenditure in man. Methods and results using a respiratory chamber. The Journal of clinical investigation 1986;78:1568-1578

11. Bouchard C, Tremblay A, Nadeau A, Despres JP, Theriault G, Boulay MR, Lortie G, Leblanc C, Fournier G: Genetic effect in resting and exercise metabolic rates. Metabolism: clinical and experimental 1989;38:364-370

12. Fontaine E, Savard R, Tremblay A, Despres JP, Poehlman E, Bouchard C: Resting metabolic rate in monozygotic and dizygotic twins. Acta geneticae medicae et gemellologiae 1985;34:41-47

13. Knowler WC, Pettitt DJ, Saad MF, Charles MA, Nelson RG, Howard BV, Bogardus C, Bennett PH: Obesity in the Pima Indians: its magnitude and relationship with diabetes. The American journal of clinical nutrition 1991;53:1543S-1551S

14. Report of the expert committee on the diagnosis and classification of diabetes mellitus. In *Diabetes care*, 2002/12/28 ed., 2003, p. S5-20

15. Pannacciulli N, Salbe AD, Ortega E, Venti CA, Bogardus C, Krakoff J: The 24-h carbohydrate oxidation rate in a human respiratory chamber predicts ad libitum food intake. The American journal of clinical nutrition 2007;86:625-632

16. Penesova A, Venti CA, Bunt JC, Bonfiglio SM, Votruba SB, Krakoff J: Short-term isocaloric manipulation of carbohydrate intake: effect on subsequent ad libitum energy intake. European journal of nutrition 2011;50:455-463

17. Tataranni PA, Ravussin E: Use of dual-energy X-ray absorptiometry in obese individuals. The American journal of clinical nutrition 1995;62:730-734

18. Bogardus C, Lillioja S, Ravussin E, Abbott W, Zawadzki JK, Young A, Knowler WC, Jacobowitz R, Moll PP: Familial dependence of the resting metabolic rate. The New England journal of medicine 1986;315:96-100

19. Lusk G: Animal calorimetry: analysis of oxidation of mixtures of carbohydrates and fat. J Biol Chem 1924;59:41-42

20. Abbott WG, Howard BV, Christin L, Freymond D, Lillioja S, Boyce VL, Anderson TE, Bogardus C, Ravussin E: Short-term energy balance: relationship with protein, carbohydrate, and fat balances. The American journal of physiology 1988;255:E332-337

21. Schutz Y, Ravussin E, Diethelm R, Jequier E: Spontaneous physical activity measured by radar in obese and control subject studied in a respiration chamber. International journal of obesity 1982;6:23-28

22. Schutz Y, Bessard T, Jequier E: Diet-induced thermogenesis measured over a whole day in obese and nonobese women. The American journal of clinical nutrition 1984;40:542-552

23. Piaggi P, Krakoff J, Bogardus C, Thearle MS: Lower "awake and fed thermogenesis" predicts future weight gain in subjects with abdominal adiposity. Diabetes 2013;62:4043-4051

24. Browning BL, Browning SR: A fast, powerful method for detecting identity by descent. Am J Hum Genet 2011;88:173-182

25. Almasy L, Blangero J: Multipoint quantitative-trait linkage analysis in general pedigrees. Am J Hum Genet 1998;62:1198-1211

26. Consortium STD, Williams AL, Jacobs SB, Moreno-Macias H, Huerta-Chagoya A, Churchhouse C, Marquez-Luna C, Garcia-Ortiz H, Gomez-Vazquez MJ, Burtt NP, Aguilar-Salinas CA, Gonzalez-Villalpando C, Florez JC, Orozco L, Haiman CA, Tusie-Luna T, Altshuler D: Sequence variants in SLC16A11 are a common risk factor for type 2 diabetes in Mexico. Nature 2014;506:97-101

27. Allison DB, Kaprio J, Korkeila M, Koskenvuo M, Neale MC, Hayakawa K: The heritability of body mass index among an international sample of monozygotic twins reared apart. International journal of obesity and related metabolic disorders : journal of the International Association for the Study of Obesity 1996;20:501-506

28. Haworth CM, Plomin R, Carnell S, Wardle J: Childhood obesity: genetic and environmental overlap with normal-range BMI. Obesity (Silver Spring) 2008;16:1585-1590

29. 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

30. Rong R, Hanson RL, Ortiz D, Wiedrich C, Kobes S, Knowler WC, Bogardus C, Baier LJ: Association analysis of variation in/near FTO, CDKAL1, SLC30A8, HHEX, EXT2, IGF2BP2, LOC387761, and CDKN2B with type 2 diabetes and related quantitative traits in Pima Indians. Diabetes 2009;58:478-488

31. Muller YL, Thearle MS, Piaggi P, Hanson RL, Hoffman D, Gene B, Mahkee D, Huang K, Kobes S, Votruba S, Knowler WC, Bogardus C, Baier LJ: Common genetic variation in and near the melanocortin 4 receptor gene (MC4R) is associated with body mass index in American Indian adults and children. Human genetics 2014;133:1431-1441

32. Traurig MT, Perez JM, Ma L, Bian L, Kobes S, Hanson RL, Knowler WC, Krakoff JA, Bogardus C, Baier LJ: Variants in the LEPR gene are nominally associated with higher BMI and lower 24-h energy expenditure in Pima Indians. Obesity (Silver Spring) 2012;20:2426-2430

33. Piaggi P, Thearle MS, Krakoff J, Votruba SB: Higher Daily Energy Expenditure and Respiratory Quotient, Rather Than Fat-Free Mass, Independently Determine Greater ad Libitum Overeating. The Journal of clinical endocrinology and metabolism 2015;100:3011-3020

34. Blundell JE, Caudwell P, Gibbons C, Hopkins M, Naslund E, King N, Finlayson G: Role of resting metabolic rate and energy expenditure in hunger and appetite control: a new formulation. Disease models & mechanisms 2012;5:608-613

35. Dulloo AG, Jacquet J, Miles-Chan JL, Schutz Y: Passive and active roles of fat-free mass in the control of energy intake and body composition regulation. Eur J Clin Nutr 2017;71:353-357

36. Wettschureck N, Offermanns S: Mammalian G proteins and their cell type specific functions. Physiol Rev 2005;85:1159-1204

37. Neer EJ: Heterotrimeric G proteins: organizers of transmembrane signals. Cell 1995;80:249-257

38. Wagner F, Bernard R, Derst C, French L, Veh RW: Microarray analysis of transcripts with elevated expressions in the rat medial or lateral habenula suggest fast GABAergic excitation in the medial habenula and habenular involvement in the regulation of feeding and energy balance. Brain structure & function 2016;

39. Orlandi C, Xie K, Masuho I, Fajardo-Serrano A, Lujan R, Martemyanov KA: Orphan Receptor GPR158 Is an Allosteric Modulator of RGS7 Catalytic Activity with an Essential Role in Dictating Its Expression and Localization in the Brain. J Biol Chem 2015;290:13622-13639

40. Orlandi C, Posokhova E, Masuho I, Ray TA, Hasan N, Gregg RG, Martemyanov KA: GPR158/179 regulate G protein signaling by controlling localization and activity of the RGS7 complexes. The Journal of cell biology 2012;197:711-719

41. Shim H, Wang CT, Chen YL, Chau VQ, Fu KG, Yang J, McQuiston AR, Fisher RA, Chen CK: Defective retinal depolarizing bipolar cells in regulators of G protein signaling (RGS) 7 and 11 double null mice. J Biol Chem 2012;287:14873-14879

42. Aissani B, Perusse L, Lapointe G, Chagnon YC, Bouchard L, Walts B, Bouchard C: A quantitative trait locus for body fat on chromosome 1q43 in French Canadians: linkage and association studies. Obesity (Silver Spring) 2006;14:1605-1615

43. Aissani B, Wiener HW, Zhang K: Fine Mapping of the Body Fat QTL on Human Chromosome 1q43. PloS one 2016;11:e0153794

44. Muller CS, Haupt A, Bildl W, Schindler J, Knaus HG, Meissner M, Rammner B, Striessnig J, Flockerzi V, Fakler B, Schulte U: Quantitative proteomics of the Cav2 channel nano-environments in the mammalian brain. Proceedings of the National Academy of Sciences of the United States of America 2010;107:14950-14957

45. Takahashi E, Ito M, Miyamoto N, Nagasu T, Ino M, Tanaka I: Increased glucose tolerance in N-type Ca2+ channel alpha(1B)-subunit gene-deficient mice. International journal of molecular medicine 2005;15:937-944

46. Locke AE, Kahali B, Berndt SI, Justice AE, Pers TH, Day FR, Powell C, Vedantam S, Buchkovich ML, Yang J, Croteau-Chonka DC, Esko T, Fall T, Ferreira T, Gustafsson S, Kutalik Z, Luan J, Magi R, Randall JC, Winkler TW, Wood AR, Workalemahu T, Faul JD, Smith JA, Hua Zhao J, Zhao W, Chen J, Fehrmann R, Hedman AK, Karjalainen J, Schmidt EM, Absher D, Amin N, Anderson D, Beekman M, Bolton JL, Bragg-Gresham JL, Buyske S, Demirkan A, Deng G, Ehret GB, Feenstra B, Feitosa MF, Fischer K, Goel A, Gong J, Jackson AU, Kanoni S, Kleber ME, Kristiansson K, Lim U, Lotay V, Mangino M, Mateo Leach I, Medina-Gomez C, Medland SE, Nalls MA, Palmer CD, Pasko D, Pechlivanis S, Peters MJ, Prokopenko I, Shungin D, Stancakova A, Strawbridge RJ, Ju Sung Y, Tanaka T, Teumer A, Trompet S, van der Laan SW, van Setten J, Van Vliet-Ostaptchouk JV, Wang Z, Yengo L, Zhang W, Isaacs A, Albrecht E, Arnlov J, Arscott GM, Attwood AP, Bandinelli S, Barrett A, Bas IN, Bellis C, Bennett AJ, Berne C, Blagieva R,

Bluher M, Bohringer S, Bonnycastle LL, Bottcher Y, Boyd HA, Bruinenberg M, Caspersen IH, Ida Chen YD, Clarke R, Daw EW, de Craen AJ, Delgado G, Dimitriou M, Doney AS, Eklund N, Estrada K, Eury E, Folkersen L, Fraser RM, Garcia ME, Geller F, Giedraitis V, Gigante B, Go AS, Golay A, Goodall AH, Gordon SD, Gorski M, Grabe HJ, Grallert H, Grammer TB, Grassler J, Gronberg H, Groves CJ, Gusto G, Haessler J, Hall P, Haller T, Hallmans G, Hartman CA, Hassinen M, Hayward C, Heard-Costa NL, Helmer Q, Hengstenberg C, Holmen O, Hottenga JJ, James AL, Jeff JM, Johansson A, Jolley J, Juliusdottir T, Kinnunen L, Koenig W, Koskenvuo M, Kratzer W, Laitinen J, Lamina C, Leander K, Lee NR, Lichtner P, Lind L, Lindstrom J, Sin Lo K, Lobbens S, Lorbeer R, Lu Y, Mach F, Magnusson PK, Mahajan A, McArdle WL, McLachlan S, Menni C, Merger S, Mihailov E, Milani L, Moayyeri A, Monda KL, Morken MA, Mulas A, Muller G, Muller-Nurasyid M, Musk AW, Nagaraja R, Nothen MM, Nolte IM, Pilz S, Rayner NW, Renstrom F, Rettig R, Ried JS, Ripke S, Robertson NR, Rose LM, Sanna S, Scharnagl H, Scholtens S, Schumacher FR, Scott WR, Seufferlein T, Shi J, Vernon Smith A, Smolonska J, Stanton AV, Steinthorsdottir V, Stirrups K, Stringham HM, Sundstrom J, Swertz MA, Swift AJ, Syvanen AC, Tan ST, Tayo BO, Thorand B, Thorleifsson G, Tyrer JP, Uh HW, Vandenput L, Verhulst FC, Vermeulen SH, Verweij N, Vonk JM, Waite LL, Warren HR, Waterworth D, Weedon MN, Wilkens LR, Willenborg C, Wilsgaard T, Wojczynski MK, Wong A, Wright AF, Zhang Q, Brennan EP, Choi M, Dastani Z, Drong AW, Eriksson P, Franco-Cereceda A, Gadin JR, Gharavi AG, Goddard ME, Handsaker RE, Huang J, Karpe F, Kathiresan S, Keildson S, Kiryluk K, Kubo M, Lee JY, Liang L, Lifton RP, Ma B, McCarroll SA, McKnight AJ, Min JL, Moffatt MF, Montgomery GW, Murabito JM, Nicholson G, Nyholt DR, Okada Y, Perry JR, Dorajoo R, Reinmaa E, Salem RM, Sandholm N, Scott RA, Stolk L, Takahashi A, Tanaka T, Van't Hooft FM, Vinkhuyzen AA, Westra HJ, Zheng W, Zondervan KT, Heath AC, Arveiler D, Bakker SJ, Beilby J, Bergman RN, Blangero J, Bovet P, Campbell H, Caulfield MJ, Cesana G, Chakravarti A, Chasman DI, Chines PS, Collins FS, Crawford DC, Cupples LA, Cusi D, Danesh J, de Faire U, den Ruijter HM, Dominiczak AF, Erbel R, Erdmann J, Eriksson JG, Farrall M, Felix SB, Ferrannini E, Ferrieres J, Ford I, Forouhi NG, Forrester T, Franco OH, Gansevoort RT, Gejman PV, Gieger C, Gottesman O, Gudnason V, Gyllensten U, Hall AS, Harris TB, Hattersley AT, Hicks AA, Hindorff LA, Hingorani AD, Hofman A, Homuth G, Hovingh GK, Humphries SE, Hunt SC, Hypponen E, Illig T, Jacobs KB, Jarvelin MR, Jockel KH, Johansen B, Jousilahti P, Jukema JW, Jula AM, Kaprio J, Kastelein JJ, Keinanen-Kiukaanniemi SM, Kiemeney LA, Knekt P,

Kooner JS, Kooperberg C, Kovacs P, Kraja AT, Kumari M, Kuusisto J, Lakka TA, Langenberg C, Le Marchand L, Lehtimaki T, Lyssenko V, Mannisto S, Marette A, Matise TC, McKenzie CA, McKnight B, Moll FL, Morris AD, Morris AP, Murray JC, Nelis M, Ohlsson C, Oldehinkel AJ, Ong KK, Madden PA, Pasterkamp G, Peden JF, Peters A, Postma DS, Pramstaller PP, Price JF, Qi L, Raitakari OT, Rankinen T, Rao DC, Rice TK, Ridker PM, Rioux JD, Ritchie MD, Rudan I, Salomaa V, Samani NJ, Saramies J, Sarzynski MA, Schunkert H, Schwarz PE, Sever P, Shuldiner AR, Sinisalo J, Stolk RP, Strauch K, Tonjes A, Tregouet DA, Tremblay A, Tremoli E, Virtamo J, Vohl MC, Volker U, Waeber G, Willemsen G, Witteman JC, Zillikens MC, Adair LS, Amouyel P, Asselbergs FW, Assimes TL, Bochud M, Boehm BO, Boerwinkle E, Bornstein SR, Bottinger EP, Bouchard C, Cauchi S, Chambers JC, Chanock SJ, Cooper RS, de Bakker PI, Dedoussis G, Ferrucci L, Franks PW, Froguel P, Groop LC, Haiman CA, Hamsten A, Hui J, Hunter DJ, Hveem K, Kaplan RC, Kivimaki M, Kuh D, Laakso M, Liu Y, Martin NG, Marz W, Melbye M, Metspalu A, Moebus S, Munroe PB, Njolstad I, Oostra BA, Palmer CN, Pedersen NL, Perola M, Perusse L, Peters U, Power C, Quertermous T, Rauramaa R, Rivadeneira F, Saaristo TE, Saleheen D, Sattar N, Schadt EE, Schlessinger D, Slagboom PE, Snieder H, Spector TD, Thorsteinsdottir U, Stumvoll M, Tuomilehto J, Uitterlinden AG, Uusitupa M, van der Harst P, Walker M, Wallaschofski H, Wareham NJ, Watkins H, Weir DR, Wichmann HE, Wilson JF, Zanen P, Borecki IB, Deloukas P, Fox CS, Heid IM, O'Connell JR, Strachan DP, Stefansson K, van Duijn CM, Abecasis GR, Franke L, Frayling TM, McCarthy MI, Visscher PM, Scherag A, Willer CJ, Boehnke M, Mohlke KL, Lindgren CM, Beckmann JS, Barroso I, North KE, Ingelsson E, Hirschhorn JN, Loos RJ, Speliotes EK: Genetic studies of body mass index yield new insights for obesity biology. Nature 2015;518:197-206

TABLES

Table 1. Anthropometric and metabolic measures of the study groups.

	All	**Males**	**Females**
Population-based longitudinal outpatient study			
N	5870	2572	3298
Birth year	1966 ± 16	1967 ± 16	1966 ± 16
Maximum BMI (kg/m²)	35.2 ± 8.4	33.9 ± 8.1	36.1 ± 8.5
Age (years)	29.6 ± 11.4	28.9 ± 11.3	30.1 ± 11.4
Body composition inpatient study			
N	917	506	411
Age (years)	28.0 ± 8.0	28.1 ± 8.3	28.0 ± 7.6
Body fat (%)	33.4 ± 8.5	28.4 ± 7.0	39.7 ± 5.7
Fat mass (kg)	33.0 ± 14.6	29.3 ± 14.0	37.6 ± 13.9
Fat free mass (kg)	62.5 ± 14.2	68.8 ± 13.4	54.7 ± 10.9
Height (cm)	166.6 ± 8.4	172.1 ± 6.2	159.9 ± 5.2
Respiratory chamber inpatient study			
N	419	254	165
Age (years)	27.8 ± 6.4	27.8 ± 6.6	27.8 ± 6.2
Body weight (kg)	95.3 ± 22.3	98.6 ± 22.7	90.2 ± 20.7
BMI (kg/m²)	34.2 ± 7.5	33.4 ± 7.3	35.3 ± 7.7
Body fat (%)	32.6 ± 8.2	28.7 ± 6.9	38.8 ± 6.0
Fat mass (kg)	32.0 ± 13.0	29.5 ± 12.9	35.8 ± 12.2
Fat free mass (kg)	63.3 ± 12.6	69.0 ± 10.9	54.4 ± 9.5
Fasting plasma glucose concentration (mg/dL)	88.8 ± 10.0	87.3 ± 9.9	91.2 ± 9.7
2-h plasma glucose concentration (mg/dL)	123.0 ± 30.5	115.8 ± 30.0	134.0 ± 28.0
24-h EE (kcal/day)	2354 ± 396	2531 ± 347	2083 ± 303
Adjusted 24-h EE (kcal/day) [#]	0 ± 142.9	0 ± 147.0	0 ± 136.9
SPA (%)	7.5 ± 2.5	7.7 ± 2.5	7.1 ± 2.5
Sleeping EE (kcal/day)	1672 ± 284	1776 ± 271	1513 ± 223
Adjusted sleeping EE (kcal/day) [#]	0 ± 135.3	0 ± 144.4	0 ± 120.3
AFT (kcal/14·hrs)	263 ± 122	288 ± 129	223 ± 99
Adjusted AFT (kcal/14·hrs) [#]	0 ± 114.8	0 ± 124.2	0 ± 98.6

Ventilated hood inpatient study			
N	509	301	208
Age (years)	26.9 ± 6.1	26.8 ± 6.4	26.9 ± 5.8
Body weight (kg)	93.4 ± 23.0	97.3 ± 24.2	87.8 ± 19.9
BMI (kg/m²)	33.5 ± 7.6	32.9 ± 7.6	34.5 ± 7.4
Body fat (%)	32.3 ± 8.5	28.2 ± 7.4	38.2 ± 6.4
Fat mass (kg)	31.2 ± 13.5	28.9 ± 13.9	34.5 ± 12.2
Fat free mass (kg)	62.2 ± 12.9	68.4 ± 11.6	53.3 ± 8.8
Fasting plasma glucose concentration (mg/dL)	89.3 ± 10.0	87.4 ± 9.5	92.1 ± 10.0
2-h plasma glucose concentration (mg/dL)	122.8 ± 30.3	115.7 ± 28.1	132.9 ± 30.5
RMR (kcal/day)	1758 ± 326	1878 ± 322	1587 ± 247
Adjusted RMR (kcal/day) [#]	0 ± 189.6	0 ± 212.4	0 ± 151.4

Abbreviations: AFT, awake and fed thermogenesis; EE, energy expenditure; RMR: resting metabolic rate; SPA, spontaneous physical activity.

Values in each cell are reported as mean ± SD.

[#]: All four EE measures (24-h EE, sleeping EE, AFT and RMR) are adjusted for age, sex, FM, FFM by linear regression analysis; 24-h EE and AFT are further adjusted for SPA and for fasting glucose levels, respectively.

Table 2. Variants associated with lower 24-h EE, lower RMR and higher maximum BMI in American Indians.

						24-h EE (n=419)		RMR (n=509)		Maximum BMI (n=5870)	
Chr.	**Position**	**Variant ID**	**Gene(s)**	**R/NR allele**	**RAF**	**β (SE)**	***p***	**β (SE)**	***p***	**β (SE)**	***p***
10	25735858	rs11014566	*GPR158*	G/A	0.60	−33.0 (10.5)	1.7×10^{-3}	−31.1 (12.1)	9.9×10^{-3}	0.017 (0.005)	4.7×10^{-4}
21	30662480	---	*LINC00189, BACH1*	-/AAG	0.30	−30.6 (11.1)	6.0×10^{-3}	−34.5 (13.3)	9.4×10^{-3}	0.014 (0.005)	3.9×10^{-3}
12	97806525	rs12424131	*NEDD1, RMST*	A/G	0.88	−42.8 (15.5)	5.6×10^{-3}	−55.8 (19.2)	3.6×10^{-3}	0.018 (0.007)	9.9×10^{-3}
12	97810465	rs17026922	*NEDD1, RMST*	G/A	0.92	−46.8 (17.8)	8.5×10^{-3}	−65.7 (22.6)	3.6×10^{-3}	0.020 (0.008)	1.2×10^{-2}
14	21113898	rs56069351	*OR6S1, ANG*	T/C	0.25	−33.0 (11.7)	4.9×10^{-3}	−42.1 (14.1)	2.7×10^{-3}	0.012 (0.005)	1.5×10^{-2}
20	4157857	rs1538072	*SMOX*	G/C	0.77	−35.9 (12.7)	4.6×10^{-3}	−52.1 (15.0)	5.4×10^{-4}	0.012 (0.006)	3.4×10^{-2}
3	77645020	rs80275771	*ROBO2*	A/T	0.16	−38.8 (14.8)	8.9×10^{-3}	−45.9 (16.9)	6.7×10^{-3}	0.014 (0.007)	3.9×10^{-2}

Variant position is based on human genome Build 37. Variant identification number (rs#) is based on dbSNP version 141. Risk allele is defined as the allele associated with lower 24-h EE, lower RMR and higher maximum BMI. RAF: Risk allele frequency as calculated in full-heritage Pima Indians. SE: Standard Error. Beta coefficient (β) is expressed per copy of the risk allele in kcal/day (24-h EE and RMR) and $\log_e$(BMI) units. Results are adjusted for age, sex, FM and FFM (only for 24-h EE and RMR analyses), SPA (only for 24-h EE), birth year (only for BMI analysis) and the first 5 genetic principal components in a mixed model that accounted for genetic relationships among individuals. Variants are sorted by their *p*-value for the association with maximum BMI.

FIGURES

Figure 1. Familial effects on 24-h EE and maximum BMI in American Indians.

Individual values of 24-h EE (adjusted for age, sex, FM, FFM and SPA, Panel A) and maximum BMI (adjusted for age, sex, birth year and self-reported Pima heritage, Panel B) of family members from the American Indian community. Individual siblings are shown as dots while families are depicted as vertical rectangles identified by the highest and lowest sibling value. Each family includes at least two siblings and families are ranked according to mean value of 24-h EE (Panel A) and BMI (Panel B) of siblings, e.g., families with relatively low mean EE/BMI are located on the left side of x-axis. Heritability (h^2) was estimated in a linear mixed model from a random effect that utilized the empirical genetic relatedness matrix estimated as the proportion of the genome shared identical by descent between each pair of siblings who had been genotyped. Mixed models for estimating heritability included age, sex, birth year (only BMI analysis), FM (only 24-h EE analysis), FFM (only 24-h EE analysis), SPA (only 24-h EE analysis) and first 5 principal components as fixed effects. The intraclass correlation coefficient (ICC) was calculated as the ratio of the variance accounted by family membership divided the total variance, and express as the percent of total variance in 24-h EE and BMI attributable to family membership after adjustment for covariates.

Figure 2. "Manhattan" plots of genome-wide association results for 24-h EE and RMR in American Indians.

The negative base-10 logarithm of the *p*-value for the association of each genetic variant (n=491265, MAF≥0.05) with 24-h EE (Panel A) and RMR (Panel B) after adjustment for age, sex, FM, FFM, SPA (only 24-h EE) and the first 5 principal components in a mixed model that accounted for genetic relationships among individuals, is plotted against chromosome and position according to Build 37.

Figure 3. EE measures by genotypes of SNP rs11014566 in *GPR158*.

Average time courses of 24-h EE in the respiratory chamber for subjects carrying AA (closed line) and GG (dotted line) genotypes of rs11014566 (Panel A). Values of EE measured every 15 min are adjusted for age, sex, FM, FFM, SPA, and accounting for repeated measures using an AR(1) covariance structure by mixed model analysis.

Mean 24-h EE inside the metabolic chamber (Panel B), resting metabolic rate (RMR) assessed by ventilated hood system (Panel C), sleeping EE (Panel D) and "awake and fed" thermogenesis (AFT, Panel E) by genotypes of rs11014566. All EE measures are adjusted for age, sex, FM, FFM and the first 5 principal components in a mixed model that accounted for genetic relationships among individuals; 24-h EE and AFT are further adjusted for SPA and fasting glucose levels, respectively. Beta coefficients are expressed per copy of the G allele. Error bars represent mean with 95% confidence intervals.

Figure 4. Anthropometric measures by genotypes of SNP rs11014566 in *GPR158*.

Maximum BMI as derived from the longitudinal data from outpatient visits (Panel A: entire population; B: full-heritage Pima Indians; and C: mixed-heritage American Indians) and percent body fat (Panel D), fat mass (Panel E) and fat free mass (Panel F) in 917 subjects who had a body composition measure in inpatient visits by genotypes of rs11014566. BMI, percent body fat, fat mass and fat free mass values are adjusted for age, sex and the first 5 principal components in a mixed model that accounted for genetic relationships among individuals. BMI is further adjusted for subject's birth year to account for secular changes in obesity prevalence during the course of the longitudinal study in the Pima Indian community. Fat mass and fat free mass are further adjusted for subject's height.

Beta coefficients are expressed per copy of the G allele and expressed as percent body fat (PFAT) and as kg (FM and FFM). For maximum BMI, beta coefficients are expressed per $\log_e$ BMI units as percentage and, as an approximation, are converted to kg/m^2 units by multiplication with the average population BMI (i.e., a beta coefficient of 0.017 is equivalent to $1.7\% \times 35.2\ kg/m^2 = 0.60\ kg/m^2$). Error bars represent mean with 95% confidence intervals.

Figure 5. Frequencies of the G allele at rs11014566 and of the T allele at rs144895904 in American Indians and in the 1000 Genomes Project populations.

Pima: full-heritage Pima Indians; EAS: East Asians; AMR: Americans; AFR: Africans; SAS: South Asians; EUR: Europeans.

Population allele frequencies are based on 1000 Genomes Project Phase 3.

Figure 6. *In vitro* functional analyses of *GPR158* variants in murine hypothalamus cells.

Relative luciferase activity (fold change) was expressed as a ratio of firefly luciferase activity to *Renilla* luciferase activity, and further normalized to pGL3-Promoter luciferase activity. Raw data are presented along with mean and 95% confidence intervals. The statistical difference in the averaged activity was analyzed by Student's unpaired *t*-test.