Pima Indians living in Arizona have 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. 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 5,870 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/m²), 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 knockout mice. Our results suggest that common ethnic-specific variation in GPR158 may influence EE; however, its role in weight gain remains controversial, as it either had no association with BMI or associated with BMI but in the opposite direction in other ethnic groups.

Obesity often aggregates in families. Household members typically share lifestyle factors including food choices, daily habits, and cultural views that may affect body weight; however, studies in twins reared apart have provided evidence that approximately two-thirds 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 (FM) (4) over time. However, this inverse relationship is not observed in all populations (6–8), and 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 interindividual 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 Pima Indians without diabetes 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 aged ≥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 75-g oral glucose tolerance test. 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 American Diabetes Association diagnostic criteria (14), was available for 5,870 subjects (Table 1). Among these subjects, 2,920 were full-heritage Pima Indian (defined as eight-eighths Pima Indian heritage by self-report) and the remaining 2,950 were mixed heritage, on average, six-eighths American Indian (typically four-eighths Pima Indian and an additional two-eighths 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 not have diabetes by the oral glucose tolerance test also participated in inpatient studies in our Clinical Research Section 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 Section, subjects were given a standard weight-maintaining diet (50% carbohydrates, 30% fats, and 20% proteins) for 3 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), FM, and 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). 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, the subject’s EE was calculated every 5 min using the equations of Lusk (19), and RMR was calculated as the average EE over 40 min while the subject was instructed to stay awake and motionless and then extrapolated to 24 h.

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 h 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 was calculated as the intercept of the regression line between EE and SPA between 11:00 and 01:00 (22). Sleeping metabolic rate was defined as the average EE of all 15-min nightly periods between 01:00 and 05:00 during which SPA was <1.5% (23). The “awake and fed” thermogenesis (AFT) was calculated as the difference between the EE in the inactive awake state and the sleeping metabolic rate (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 7,701 Pima Indian samples. This array was designed to capture common variation (minor allele frequency [MAF] ≥ 0.05, or ≥ 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,
Table 1 – Anthropometric and metabolic measures of the study groups

|                             | All          | Males        | Females       |
|------------------------------|--------------|--------------|---------------|
| Population-based longitudinal outpatient study |              |              |               |
| N                            | 5,870        | 2,572        | 3,298         |
| Birth year                   | 1,966 ± 16   | 1,967 ± 16   | 1,966 ± 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    |
| PFAT (%)                     | 33.4 ± 8.5   | 28.4 ± 7.0   | 39.7 ± 5.7    |
| FM (kg)                      | 33.0 ± 14.6  | 29.3 ± 14.0  | 37.6 ± 13.9   |
| FFM (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    |
| PFAT (%)                     | 32.6 ± 8.2   | 28.7 ± 6.9   | 38.8 ± 6.0    |
| FM (kg)                      | 32.0 ± 13.0  | 29.5 ± 12.9  | 35.8 ± 12.2   |
| FFM (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    |
| 24-h plasma glucose concentration (mg/dL) | 123.0 ± 30.5 | 115.8 ± 30.0 | 134.0 ± 28.0  |
| Adjusted 24-h EE (kcal/day)# | 0 ± 142.9    | 0 ± 147.0    | 0 ± 136.9     |
| Adjusted sleeping EE (kcal/day)# | 0 ± 135.3    | 0 ± 144.4    | 0 ± 120.3     |
| AFT (kcal/14 h)              | 263 ± 122    | 288 ± 129    | 223 ± 99      |
| AFT (kcal/14 h)#             | 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    |
| PFAT (%)                     | 32.3 ± 8.5   | 28.2 ± 7.4   | 38.2 ± 6.4    |
| FM (kg)                      | 31.2 ± 13.5  | 28.9 ± 13.9  | 34.5 ± 12.2   |
| FFM (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   |
| Adjusted RMR (kcal/day)#     | 0 ± 189.6    | 0 ± 212.4    | 0 ± 151.4     |

Data are mean ± SD unless otherwise noted. #All four EE measures (24-h EE, sleeping EE, AFT, and RMR) are adjusted for age, sex, FM, and FFM by linear regression analysis; 24-h EE and AFT are further adjusted for SPA and for fasting glucose levels, respectively.

and lack of deviation from Hardy-Weinberg equilibrium with a $P > 10^{-4}$ tag 92% of the 4.9 million common variants with a MAF ≥ 0.05 detected in the genomes of full-heritage Pima Indians (tag defined as $r^2 ≥ 0.85$ within 300 kb).

Functional Analysis of GPR158 Variants

DNA fragments containing each allele homozygous at rs11014566, rs144895904, rs34673593, and rs16925884 were PCR amplified (rs11014566, primers forward 5’-ACAGTTCTAATTAGTATGCCTGAGA-3’ and reverse 5’-TCACGAGA-GGCCGACAAATTACATAAC-3’; rs144895904, forward 5’-TAATCGCTTACGAGATCAAAGCTGTTCA-3 and reverse 5’-TCACTGAGACGCGCCACACAAATTACATAAC-3’; rs16925884/rs34673593, forward 5’-ACAGTTCTAATTAGTATGCCTGAGA-3’ and reverse 5’-AGCTGAGATAAATGGAATCTTAGATT-3’). The amplicons were inserted at KpnI and XhoI sites (underlined, respectively) upstream of the pGL3 promoter firefly luciferase reporter vector (Promega, Madison, WI). 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% CO2, and 95% air atmosphere. One microgram of DNA construct and 125 ng of pGL4 Renilla luciferase reporter vector (Promega, Madison, WI) were transiently transfected into the cells with Lipofectamine LTX (Invitrogen, Life Technologies, Carlsbad, CA). At 48 h posttransfection, cells were harvested and a
dual-luciferase reporter assay was performed using a standard protocol (Promega, Madison, WI). Three separate transfections were conducted, each transfection was repeated two to three times (for a total of eight), 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-model 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 five genetic principal components calculated from 19,991 variants randomly selected from 200-kb windows across the genome (one 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 using 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 Haplovew program (version 4.2, Broad Institute, Cambridge, MA). Tag variants were selected based on the sequence data of 266 Pima genomes using the Tagger algorithm (Haplovew) with a pairwise $r^2 \geq 0.85$ taken as indicative of redundancy. The statistical difference in mean luciferase activity detected in the functional study was analyzed by unpaired Student 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 two-sided $\alpha = 5 \times 10^{-8}$, 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 individuals without diabetes 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 of four studies from Mexico or Mexicans living in the U.S. comprising a total of 4,364 individuals without diabetes. 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-Hour 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 each 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 (Fig. 1A). In 3,298 siblings from 1,131 Pima Indian families, family membership was the largest determinant of maximum BMI, accounting for one-third of BMI variance among individuals (RMSD = 4.4 kg/m², $P < 0.001$) (Fig. 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, $P = 1.5 \times 10^{-3}$ adjusted for age, sex, FM, FFM, SPA, and first five principal components) and 0.55 (95% CI 0.51–0.60, $P = 8.2 \times 10^{-4}$ adjusted for age, sex, birth year, and first five 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 Fig. 2. The lists of variants with $P < 0.01$ for each EE measure are reported in Supplementary 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 ($P < 5 \times 10^{-8}$) with either EE measure. However, 138 variants had nominal ($P < 0.01$), directionally consistent associations with both 24-h EE and RMR (Supplementary Table 3).
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 5,870 individuals and with PFAT in 917 subjects who had undergone metabolic testing as inpatients. Although 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, \( P = 4.7 \times 10^{-24} \)) also associated with increased PFAT (\( P = 2.9 \times 10^{-3} \)).

The G allele at rs11014566 (frequency in full-heritage Pima Indians = 0.60) associated with a lower 24-h EE (\( \beta = -33 \text{ kcal/day}, \text{95\% CI} -54 \text{ to} -12 \)) (Fig. 3B) and a lower RMR (\( \beta = -31 \text{ kcal/day}, \text{95\% CI} -55 \text{ to} -7 \)) (Fig. 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 h inside the metabolic chamber (\( \Delta = -2.6 \text{ kcal/h}, \text{95\% CI} 7.7 \times 10^{-3} \)), and this was more evident in the sleeping state (Fig. 3A). Accordingly, single nucleotide polymorphism (SNP) rs11014566 was associated with higher BMI (\( \beta = +1.7\% \approx 0.6 \text{ kg/m}^2 \text{ per copy, 95\% CI} 0.7 \text{ to} 2.6, \text{95\% CI} 4.7 \times 10^{-3} \)) (Fig. 4A), and this association was observed both in the 2,920 individuals who were full-heritage Pima Indians (\( \beta = +1.4\%, \text{95\% CI} 0.7 \text{ to} 2.6, \text{95\% CI} 4.7 \times 10^{-3} \)) (Fig. 4A), and in the 2,950 individuals who were mixed-heritage American Indians (\( \beta = +2.0\%, \text{95\% CI} 0.3 \text{ to} 1.5, \text{95\% CI} 2.9 \times 10^{-3} \)) (Fig. 4B) in this population. This variant also associated with PFAT in 917 subjects with body composition measures (\( \beta = +0.9\%, \text{95\% CI} 0.3 \text{ to} 1.5, \text{95\% CI} 2.9 \times 10^{-3} \)) (Fig. 4D) without difference between sexes (\( P = 0.39 \)). Specifically, the G allele was associated with higher FM (\( \beta = +2.4 \text{ kg}, \text{95\% CI} 2.5 \times 10^{-4} \), adjusted for age, sex, and height) (Fig. 4E) and, to a lesser extent, with higher FFM (\( \beta = +1.3 \text{ kg}, \text{95\% CI} 1.7 \times 10^{-2} \)) (Fig. 4F).
Analysis of Additional Variation at the \textit{GPR158} Locus

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

The BMI risk alleles for \(\text{rs11014566} \) 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 \(\text{rs11014566} \) attains the highest frequency of 0.60 in full-heritage Pima Indians and 0.48 in mixed-heritage American Indians, whereas in the 1000G its frequency was 0.23 in Americans, 0.11 in Africans, and <0.01 in Europeans (Fig. 5).

Replication Analysis in SIGMA

As no other data sets 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 4,364 Mexican individuals without diabetes (mean ± SD age 57.9 ± 8.4 years, BMI 27.5 ± 4.2 kg/m², 1,755 males), the BMI risk alleles at both \(\text{rs11014566} \) (\( \beta = -0.05 \) SD units per copy of the \( G \) allele, frequency = 0.27, \( P = 0.04 \))
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 one-third 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 and rs144895904 ($\beta = -0.05$ 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 nonrisk allele for each variant were PCR amplified. Because of the proximity of rs16925884 (chr10:25,740,897) and rs34673593 (chr10:25,741,140), single PCR products containing either the risk alleles or the nonrisk alleles for both variants were amplified. The effect of the cloned GPR158 variants on promoter activity was assessed in a murine hypothalamus cell line, as GPR158 is endogenously expressed in the human hypothalamus at high levels (Supplementary Fig. 1). The largest difference in luciferase activity was observed when comparing constructs that differed for alleles at rs144895904 (Fig. 6), where the BMI risk allele T had on average 48% higher activity as compared with the nonrisk allele C (mean $\pm$ SEM, 1.21 $\pm$ 0.09 vs. 0.82 $\pm$ 0.07, $P = 0.004$). There was no significant difference in luciferase activity between alleles at rs11014566 ($P = 0.52$) or rs16925884/rs34673593 ($P = 0.74$).
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 ($P = 5 \times 10^{-8}$). 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 because of 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/m² as compared with subjects homozygous for the A allele. The effect of rs11014566 on BMI in Pima Indians (1.2 kg/m² difference between individuals homozygous for the risk vs. nonrisk 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/m², RAF ≈ 0.15) (30), rs74861148 near MC4R (1.36 kg/m², RAF ≈ 0.15) (31), and rs2025804 in LEPR (1.0–1.9 kg/m², 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) (Supplementary 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 community 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 h 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. Because 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 (33) is not optimal. Therefore, we assessed association with BMI in data sets collected from Asians (34), Africans (35) and Hispanics (26). rs11014566 did not associate with BMI in Asians or Africans (K.E. North, M.C.Y. Ng, M. Graff, and X.O. Shu, personal communication); 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, respectively, in Hispanics as compared with 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 and that 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 (36–38), 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, rs169255884, 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 (39,40). 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 (41). GPR158 binds to the regulator of G-protein signaling 7 (RGS7) in the nervous system (42,43). RGS7-deficient mice are protected from obesity (44), and previous studies have provided evidence that RGS7 may constitute an obesity locus in humans (33,45). In addition to RGS7, GPR158 also binds to an N-type voltage-gated calcium channel (CACNA1B) in the rat brain (34). Homozygous CACNA1B-deficient mice gain less weight during 8 weeks of high-fat diet despite similar food intake of wild-type mice (35), implying a compensatory increase in EE that may mitigate weight gain during high-fat feeding. Because GPR158, RGS7, and CACNA1B are all expressed in the central nervous system, as are the welldocumented 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 physiological 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 with other populations studied as part of the 1000G, and the risk allele demonstrated increased promoter activity 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 because an association with BMI was only observed in Pima Indians. We propose that studies of GPR158, as well as other EE-associated genes that will be identified in the future, may shed light into the pathophysiological mechanisms that affect EE, which could eventually lead to prevention and/or possible treatments of human obesity.
13. Knower WC, Pettitt DJ, Saad MF, et al. Obesity in the Pima Indians: its magnitude and relationship with diabetes. Am J Clin Nutr 1991;53(Suppl.):1543S–1551S
14. The Expert Committee on the Diagnosis and Classification of Diabetes Mellitus. Report of The Expert Committee on the Diagnosis and Classification of Diabetes Mellitus. Diabetes Care 2003;26(Suppl. 1):S5–S20
15. Panaccione R, Sabbe AD, Ortega E, Ventri CA, Bogardus C, Krakoff J. The 24-h carbohydrate oxidation rate in a human respiratory chamber predicts ad libitum food intake. Am J Clin Nutr 2007;86:625–632
16. Penedes A, Ventri CA, Bunt JC, Bonfiglio SM, Votruba SB, Krakoff J. Short-term isocaloric manipulation of carbohydrate intake: effect on subsequent ad libitum energy intake. Eur J Nutr 2011;50:455–463
17. Tataranni PA, Ravussin E. Use of dual-energy X-ray absorptiometry in obese individuals. Am J Clin Nutr 1995;62:730–734
18. Bogardus C, Lilligo S, Ravussin E, et al. Familial dependence of the resting metabolic rate. N Engl J Med 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, et al. Short-term energy balance: relationship with protein, carbohydrate, and fat balances. Am J Physiol 1988;255:E332–E337
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. Int J Obes 1982;6:23–28
22. Schutz Y, Bessard T, Jéquier E. Diet-induced thermogenesis measured over a whole day in obese and nonobese women. Am J Clin Nutr 1991;53(Suppl.):1543S–1551S
23. Piaggi P, Krakoff J, Bogardus C, Thearle MS. Lower carbohydrate oxidation rate in a human respiratory chamber predicts ad libitum food intake. Am J Clin Nutr 2007;86:625–632
24. Williams AL, Jacobs SB, Moreno-Macías H, et al.; SIGMA Type 2 Diabetes Consortium; ReproGen Consortium; GLGC; ICBP; MAGIC Investigators; MuTHER Consortium; MIGen Consortium; AGEN-BMI Working Group; CARDIOGRAMplusC4D Consortium; CKDGen Consortium; GLGC; ICBP; MAGIC Investigators; MuTHER Consortium; MIGen Consortium; PAGE Consortium; ReproGen Consortium; GENIE Consortium; International Endocrine Consortium. Genetic studies of body mass index yield new insights for obesity biology. Nature 2015;518:197–206