HUMAN GENOMICS

Sequencing of 640,000 exomes identifies GPR75 variants associated with protection from obesity

Parsa Akbari‡, Ankit Giliani‡, Olukyode Ososina‡, Jack A. Kosmicki, Lori Khrimian, Yi-Ya Fang, Trikaldarshi Persaud, Victor Garcia, Dylan Sun, Alexander Li, Joelle Mbatchou, Adam E. Locke, Christian Benner, Niek Verweij, Nan Lin, Sakib Hossain, Kevin Agostinucci, Jonathan V. Pascale, Ercument Dirice, Michael Dunn, Regeneron Genetics Center, DiscovEHR Collaboration, William E. Kraus, Svati H. Shah, Yili Der I. Chen, Jerome I. Rotter, Daniel J. Rader, Olle Melander, Christopher D. Still, Tooraj Mirshahi, David J. Carey, Jaime Berumen-Campos, Pablo Kuri-Morales, Jesus Alegre-Diaz, Jason M. Torres, Jonathan R. Emerson, Roger Collins, Suganthi Balasubramanian, Alicia Hawes, Marcus Jones, Brian Zambrowicz, Andrew J. Murphy, Charles Paulding, Giovanni Coppola, John D. Overton, Jeffrey G. Reid, Alan R. Shuldiner, Michael Cantor, Hyun M. Kang, Goncalo R. Abecasis, Katia Karalis, Aris N. Economides, Jonathan Marchini, George D. Yancopoulos, Alicia Hawes, Marcus Jones, Brian Zambrowicz, Andrew J. Murphy, Charles Paulding, Giovanni Coppola, John D. Overton, Jeffrey G. Reid, Alan R. Shuldiner, Michael Cantor, Hyun M. Kang, Goncalo R. Abecasis, Katia Karalis, Aris N. Economides, Jonathan Marchini, George D. Yancopoulos, Mark W. Slemman, Judith Altarejos, Giusy Della Gatta, Roberto Tapia-Conyer‡, Michal L. Schwartzman‡, Aris Baras‡*, Manuel A. R. Ferreira‡, Luca A. Lotta‡*

INTRODUCTION: Obesity accounts for a substantial and growing burden of disease globally. Body adiposity is highly heritable, and human genetic studies can lead to biological and therapeutic insights.

RATIONALE: Whole-exome sequencing of hundreds of thousands of individuals is complementary to approaches used to date in obesity genetics and has the potential to identify rare protein-coding variants with large phenotypic impact. We sequenced the exomes of 645,626 individuals from the UK, the US, and Mexico and estimated associations of rare coding variants with body mass index (BMI), a measure of overall adiposity used to define obesity in clinical practice. We complemented exome sequencing with fine-mapping of common alleles, polygenic score analysis, and in vitro and in vivo modeling work.

RESULTS: We identified 16 genes for which the burden of rare nonsynonymous variants was associated with BMI at exome-wide statistical significance (inverse-variance-weighted meta-analysis $P < 3.6 \times 10^{-5}$), including associations at five brain-expressed G protein-coupled receptors (CALCR, MC4R, GIPR, GPR151, and GPR75). We observed an overrepresentation of genes highly expressed in the hypothalamus, a key center for the neuroendocrine regulation of energy balance. Protein-truncating variants in GPR75 were found in ~4/10,000 sequenced people and were associated with 1.8 kg/m² lower BMI, 5.3 kg lower bodyweight, and 54% lower odds of obesity in heterozygous carriers. Knockout of Gpr75 in mice resulted in resistance to weight gain in a high-fat diet model, which was allele-dose dependent (25% and 44% lower weight gain, respectively, for heterozygous Gpr75+/− mice and knockout Gpr75−/− mice compared with wild type) and accompanied by improved glycemic control and insulin sensitivity. Protein-truncating variants in CALCR were associated with higher BMI and obesity risk, whereas protein-truncating variants in GIPR and two missense alleles [Arg190→Gln (Arg190Gln), Gln288Gly], which we show result in loss of function in vitro, were associated with lower adiposity. Among monogenic obesity genes in the leptin-melanocortin pathway, heterozygous predicted loss-of-function variants in LEP, POMC, PCSK1, and MC4R (but not LEPR) were associated with higher BMI. Rare protein-truncating variants in UBR2, ANO4, and PCSK1 were associated with more than twofold higher odds of obesity in heterozygous carriers, similar to predicted-deleterious nonsynonymous variants in MC4R, which are considered the most common cause of monogenic obesity. Polygenic predisposition due to >2 million common genetic variants influenced the penetrance of obesity in rare variant carriers in an additive fashion.

CONCLUSION: These results suggest that inhibition of GPR75 may be a therapeutic strategy for obesity and illustrate the power of massive-scale exome sequencing for the identification of large-effect coding variant associations and drug targets for complex traits.

The list of author affiliations is available in the full article online.

†These authors contributed equally to this work.
‡These authors contributed equally to this work.
*L. A. L.; aris.baras@regeneron.com (A. B. B.)

READ THE FULL ARTICLE AT

https://doi.org/10.1126/science.abf8683

Akbari et al., Science 373, 73 (2021) 2 July 2021

1 of 1
Sequencing of 640,000 exomes identifies GPR75 variants associated with protection from obesity

Parsa Akbari1*, Ankit Gilani2†, Olukayode Sosina3‡, Jack A. Kosmicki4, Lori Khrimian5, Yi-Ya Fang2, Trikalardushi Persaud6, Victor Garcia7, Dylan Sun8, Alexander Li9, Joelle Mbatchou9, Adam E. Locke10, Christian Benner11, Niek Verweij12, Sakib Hossain13, Kevin Agostinucci14, Jonathan V. Pascale15, Ercument Dirice16, Michael Dunn17, Regeneron Genetics Center17, DiscoveEHR Collaboration17, William E. Kraus18, Svati H. Shah19,20, Yi-Der I. Chen21, Jerome I. Rotter22, Danila J. Rader22, Jason M. Torres23, William E. Kraus4,5, Svati H. Shah6,7, Yii-Der I. Chen8, Jerome I. Rotter8, Daniel J. Rader9, and Michael Dunn3

In mice resulted in resistance to weight gain and improved glycerol consumption in a high-fat diet model. Inhibition of GPR75 may provide a therapeutic strategy for obesity.

Obesity and its health complications account for a substantial and growing burden of global disease. Understanding the genetic and molecular underpinnings of body adiposity can be a pathway to the development of safe and effective therapeutic strategies. Body fat is a highly heritable trait, and genetic studies have revealed biological pathways that regulate energy balance. Analyses of coding variation in human obesity have focused on (i) candidate gene or exome sequencing in pedigrees or case collections with extreme phenotypes or (ii) array-genotyping of cataloged variant sites in large cohorts. Whole-exome sequencing of hundreds of thousands of individuals from population or health system-based studies is a complementary approach that may identify large-effect coding variants influencing the propensity to become obese or protection against obesity (25–28). Here, we report a multietnic exome-sequencing association study for BMI in more than 640,000 individuals across three distinct cohorts and regions of the world (the United Kingdom, the United States, and Mexico).

Exome-wide gene-burden association of rare coding alleles with body mass index

We performed high-coverage whole-exome sequencing in 645,626 individuals (29), including 428,719 individuals of European ancestry from the UK Biobank cohort (UKB; table S1) (30), 121,061 individuals of European ancestry from the MyCode Community Health Initiative cohort from the US-based Geisinger Health System (GHS; table S1) (31), and 95,846 individuals of admixed American ancestry from the Mexico City Prospective Study (MCPS; table S1) (32).

In an exome-wide meta-analysis across these three cohorts, there were 16 genes for which the burden of rare nonsynonymous genetic variants was associated with BMI at the exome-wide level of statistical significance [invers variance weighted (IVW) meta-analysis P < 3.6 × 10−7, a Bonferroni correction for 20,000 genes, and seven variant selection models (29); Table 1 and fig. S1]. These associations were conditionally-independent of BMI-associated common variants identified by fine-mapping genome-wide association signals (29) and were consistent across the constituent datasets of the meta-analysis (table S2).

Among the 16 genes, our analysis identified two for which rare mutations are known to cause monogenic obesity [MC4R (melanocortin 4 receptor) (8, 9) and PCSK1 (proprotein convertase subtilisin/kexin type 1) (6)] and two genes where rare coding variants have been associated with BMI [GPR151 (G protein-coupled receptor 151) (33) and GIPR (gastric inhibitory polypeptide receptor) (16)]. For the other 12 genes, our study provides genetic evidence linking rare coding variation to BMI and obesity-related phenotypes. Five of the 16 genes encode G protein-coupled receptors (GPCRs; the largest class of drug targets in the human genome) (34) expressed in the brain and central nervous system [GPR75 (G protein-coupled receptor 75), CALCRL (calcitonin receptor), GIPR, GPR151, and MC4R].

A tissue expression analysis for BMI-associated genes in our results revealed an overrepresentation of genes that are highly and specifically expressed in the hypothalamus, a key center for the neuroendocrine regulation of energy balance (fig. S2).

Sequencing of 640,000 exomes identifies GPR75 variants associated with protection from obesity

Large-scale human exome sequencing can identify rare protein-coding variants with a large impact on complex traits such as body adiposity. We sequenced the exomes of 645,626 individuals from the United Kingdom, the United States, and Mexico and estimated associations of rare coding variants with body mass index (BMI). We identified 16 genes with an exome-wide significant association with BMI, including those encoding five brain-expressed G protein-coupled receptors (CALCR, MC4R, GIPR, GPR151, and GPR75). Protein-truncating variants in GPR75 were observed in ~4/10,000 sequenced individuals and were associated with 1.8 kilograms per square meter lower BMI and 54% lower odds of obesity in the heterozygous state. Knock out of Gpr75 in mice resulted in resistance to weight gain and improved glycerol consumption in a high-fat diet model. Inhibition of GPR75 may provide a therapeutic strategy for obesity.

© 2021 The Authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original U.S. Government Works. Distributed under a Creative Commons Attribution NonCommercial License 4.0 (CC BY-NC).
Protein-truncating GPR75 variants associated with leanness and protection against obesity in humans

We explored in depth the association for rare predicted loss-of-function (pLOF) variants in the GPR75 gene, as this gene encodes a GPCR highly expressed in the brain across species, and this association was the largest effect-size association with lower BMI in our exome-wide analysis. Predicted loss-of-function variants in GPR75 were observed in ~4 out of every 10,000 sequenced people, with similar frequency across populations (Table S2), and carrier status was associated with 0.34 standard deviations lower BMI, corresponding to 1.8 kg/m² lower BMI or about 5.3 kg, or 12 lb, lower body weight (Table 1 and Fig. 1A).

The association with lower BMI was directionally consistent and statistically significant in each of the constituent cohorts of our discovery meta-analysis (Table S2) as well as within age and sex subgroups (table S3). We further corroborated the association of GPR75 pLOF variants with lower BMI in a combined discovery analysis including an additional 91,328 individuals not included in the discovery set [per-allele beta in standard deviation (SD) units of BMI in the meta-analysis of discovery and additional cohorts, −0.34; 95% confidence interval (CI), −0.45, −0.22; \( P = 6.9 \times 10^{-7} \)] (Fig. 1B). This strong association with lower BMI was accompanied by a corresponding association with protection against obesity. Heterozygous carriers of GPR75 pLOF variants had 54% lower odds of obesity compared with noncarriers in a meta-analysis of the UKB, GHS, and MCPS cohorts [Table S4; per-allele odds ratio (OR), 0.46; 95% CI, 0.31, 0.67; \( P = 6.9 \times 10^{-7} \)], and their distribution across BMI categories was drastically shifted toward lower BMI categories (Fig. 2). None of 228 heterozygous carriers of GPR75 pLOF variants were overweight...
**Fig. 1.** Protein-truncating variants in GPR75 associated with lower body mass index in humans. (A) Linear model of the GPR75 protein and its domains (top: intra- and extracellular domains in yellow, transmembrane domains in orange), the distribution on the GPR75 protein of 46 pLOF variants found by exome sequencing (middle), and the distribution of BMI in standardized units among heterozygous carriers of each variant (bottom). In the bottom subpanel, horizontal blue bars show the mean BMI in noncarriers, while horizontal red bars show the overall covariates-adjusted mean BMI in carriers of any pLOF genetic variant in GPR75. (B) Meta-analysis of the association with BMI of pLOF variants in GPR75 in discovery and additional cohorts. Abbreviations: CI, confidence interval; RR, reference-reference genotype; RA, reference-alternative heterozygous genotype; AA, alternative-alternative homozygous genotype; DHS, Dallas Heart Study; SINAI, Mount Sinai BioMe cohort; DUKE, Duke Catheterization Genetics cohort; TAICHI, Taiwanese Chinese persons from the Taiwan Metabolochip Consortium; PMBB, University of Pennsylvania Medicine BioBank; MALMO, Malmö Diet and Cancer Study; AFR, African ancestry; AMR, American ancestry; EAS, East Asian ancestry; EUR, European ancestry; SAS, South Asian ancestry. MCPS included individuals of admixed American ancestry.
In the UKB cohort, GPR75 pLOF carriers were more likely than noncarriers to self-report a thinner-than-average comparative body size at age 10 (table S5).

We examined the genomic context of the BMI association for pLOF variants in GPR75. The first and smallest exon of GPR75, containing untranslated sequence, is included in both GPR75 and in a putative GPR75-ASB3 readthrough gene with the nearby ankryin repeat and SOCS box containing 3 (ASB3) (fig. S8). The second and final GPR75 exon (containing the entire translated region of GPR75) is not shared with any other gene or transcript (fig. S3). We conducted a number of analyses to ensure that the association of pLOF variants could be firmly attributed to the GPR75 gene. First, 45 of the 46 pLOF variants in GPR75 that contributed to the association with lower BMI were located in exon 2 (table S6), which is exclusive to the GPR75 gene (fig. S8). Accordingly, the burden genotypes for pLOF variants in GPR75 had no linkage disequilibrium (LD; squared Pearson correlation coefficient ($R^2$) < 0.0001) with the burden genotype for pLOF variants affecting the GPR75-ASB3 readthrough gene or the ASB3 gene. Second, we estimated the association with BMI of the burden of rare pLOF variants predicted to truncate both ASB3 and GPR75-ASB3 readthrough gene in our large exome sequencing meta-analysis. There was no association with BMI for the burden of rare nonsynonymous variants in ASB3 or in the GPR75-ASB3 readthrough gene or ASB3 gene. Therefore, the association with lower BMI for rare pLOF variants in GPR75 can be confidently attributed to the GPR75 gene.

We also explored whether there were common variant associations in the locus. In the 1-Mb window surrounding GPR75 (500 kb to either side of the gene), there were 26 common variants associated with BMI at the genomewide level of statistical significance (IVW meta-analysis, $P < 5 \times 10^{-8}$) in our GWAS of imputed common variants in Europeans (table S9 and fig. S4), whereas there were no genome-wide significant associations in admixed Americans (fig. S4). These 26 variants all fine-mapped to a signal led by rs59428052 (G-allele frequency, 14.7%; posterior probability of causal association, 30.4%; per-allele beta in SD units of BMI, $-0.015$; 95% CI, $-0.020$ to $-0.010$; $P = 1.3 \times 10^{-6}$), which is an intergenic variant nearest to ASB3 and ~200 kb downstream of GPR75. The rs59428052 variant did not colocalize with any expression quantitative trait locus (eQTL) signal nor were any of the additional 25 variants at the locus in LD ($R^2 > 0.8$) with any sentinel eQTLs in Genotype-Tissue Expression (GTEx) Portal v8 (table S9). Two of the 26 variants were in LD with a missense variant in ASB3 and GPR75-ASB3 (rs36020289), which does not affect the GPR75 transcript (table S9).

We performed a formal conditional analysis adjusting for the 26 common variants associated with BMI in the region and identified that the association with lower BMI for pLOF variants in GPR75 remains unchanged (table S8). Therefore, the association with lower BMI for rare pLOF variants in GPR75 is conditionally independent of any of the 26 common variants associated with BMI in the locus in Europeans.

In summary, our human genetic analysis at the locus indicates that: (i) rare pLOF variants in GPR75 are associated with lower BMI with a large effect association, (ii) the pLOF association is attributed to GPR75 and not to other nearby transcripts, (iii) the signal is independent of BMI-associated common variants in the region, and (iv) the small-effect intergenic common variant signal found in that region by GWAS fine-mapping in Europeans has no apparent link with GPR75.

The association with lower BMI for pLOF variants in GPR75 was due to multiple independent rare pLOF variants predicted to truncate GPR75 at different locations (fig. IA and table S6). Because of their rarity, none of the 46 rare pLOF variants found by exome sequencing in our analysis were well ascertained by array-genotyping or imputation (table S6). Leave-one-out analyses showed that the burden signal was robust to the exclusion of one pLOF variant at a time (table S10). Out of 46 rare pLOF variants in GPR75, five (Ala1106s, Ser219fs, Gln234fs, Cys400fs, and Lys404fs) were individually associated with lower BMI at a nominal level of statistical significance (IVW meta-analysis $P < 0.05$; table S11), whereas none were associated with higher BMI. When excluding all five of these variant sites from analysis, the remaining set of pLOF variants was still associated with lower BMI (table S10).

We expressed in vitro the two most frequent (minor allele count ≥ 10) among the pLOF variants individually associated with BMI and show that they result in cellular retention of a truncated receptor likely leading to a complete loss of function (Fig. 3). We predict that the loss of a functional copy (i.e., haploinsufficiency) or production of a truncated protein that disrupts receptor multimers (i.e., dominant negative effects) may explain the association of GPR75 truncation with lower BMI. We hypothesized that in the case of haploinsufficiency, the earlier N-terminal truncation of GPR75 would result in greater phenotypic impact than a C-terminal truncation within...
Fig. 3. In vitro expression studies of two predicted loss-of-function genetic variants in GPR75. (A) Results of quantitative reverse transcription polymerase chain reaction experiments which measured GPR75 mRNA levels. Expression of GPR75 was calculated relative to the beta-actin gene. Values represent the mean and standard deviation of three technical replicates representative of one of three biological replicate experiments performed for each condition. (B) Western blotting analysis of GPR75 protein levels. GPR75 Ala110fs and Gln234* protein products correspond to the predicted molecular weight of 14 and 25 kDa, respectively. The results are representative of three biological replicates. (C) Immunofluorescence staining experiments describing the cellular localization of GPR75. The top images show intracellular staining achieved by membrane permeabilization, while the bottom images show plasma membrane localization (nonpermeabilized cellular membrane). (D) Flow cytometry analysis of the cell surface expression of GPR75. Identified cell populations are presented in percent (%) of live HA-TAG GPR75 positive cells. Values represent the mean of four biological replicates per condition and their standard deviation. All experiments were performed in HEK293 cells that were transfected with green fluorescent protein control plasmids (Control), GPR75 wild type (GPR75), GPR75-Ala110fs, or GPR75-Gln234* plasmids. Abbreviations: GAPDH, glyceraldehyde-3-phosphate dehydrogenase; SSC, side scatter; HA, hemagglutinin tag.

the last intracellular domain. Genetic variants resulting in truncation of GPR75 before the final intracellular domain were associated with a -2.1 kg/m² lower BMI (IVW meta-analysis, \( P = 4.1 \times 10^{-7} \)) as compared with a -1.4 kg/m² lower BMI (IVW meta-analysis, \( P = 0.012 \)) for variants resulting in truncation within the final domain (table S12). This difference was even more pronounced for truncations within the last 100 amino acids of the final C-terminal domain (table S12).

Body composition analysis with bi impedance in the UKB cohort showed that the association with lower BMI was driven by an association with lower overall body fat mass and lower body fat percentage (fig. S5). In an agnostic phenome-wide analysis of GPR75 pLOF variants (29), we did not observe statistically significant associations with common diagnoses or measured continuous traits after correction for the number of statistical tests performed (2173 phenotype tests; Bonferroni-corrected \( P \) value threshold, \( P < 2.3 \times 10^{-5} \)), reflecting the rarity of these variants and the stringent multiple test correction.

A detailed analysis of metabolic traits revealed a nominally-significant association (IVW meta-analysis \( P < 0.05 \)) with higher high-density lipoprotein cholesterol, which is consistent with a favorable metabolic profile (table S13). Carriers of pLOF in GPR75 had lower odds of type 2 diabetes than did noncarriers (63,492 cases and 549,961 controls; per-allele OR, 0.92; 95% CI, 0.59, 1.45; \( P = 0.73 \); table S13), but the difference was not statistically significant. We interrogated exome sequencing association statistics from up to 20,791 type 2 diabetes cases and 24,440 controls included in the Type 2 Diabetes (T2D) Knowledge Portal (https://t2d.hugeamp.org/; accessed 8 January 2021) and similarly observed numerically lower odds of type 2 diabetes in carriers of GPR75 pLOF variants (OR for type 2 diabetes, 0.52; 95% CI, 0.14 to 1.97; \( P = 0.30 \); alternative allele frequency, 0.03%). Owing to the rarity of pLOF variants in GPR75 and given the genetic relationship between BMI and type 2 diabetes, we estimate that millions of people would need to be sequenced to detect an association at \( P < 0.05 \) (table S13). An analysis for HbA1c, a continuous biomarker of glycemic levels, led to similar results (table S13).

Gpr75 deletion confers resistance to high-fat diet–induced obesity in mice

In a mouse model of high-fat diet (HFD)–induced obesity, experimental deletion of Gpr75 protected against weight gain and its associated abnormalities in glucose and insulin metabolism (Fig. 4). When placed on HFD for 14 weeks, Gpr75+/− mice approximately doubled their weight. Body weight changed from an average (standard deviation) of 20.9 (2.1) to 43.3 (6.5) grams (body weight change, +22.4 g). In contrast, mice with a genetic deletion of Gpr75 gained less weight in an allele-dose–dependent fashion (body weight change +16.9 g, difference in weight change compared with wild type -5.5 g or -25% for Gpr75+/− mice; body weight change +12.6 g, difference in weight change compared with wild type -9.8 g or -44% for the Gpr75−/− mice; Fig. 4A). Increases in fasting blood glucose seen with HFD in Gpr75+/− mice were reduced in an allele-dose–dependent manner in Gpr75+/− and Gpr75−/− mice (Fig. 4B). Mice with a genetic deletion in Gpr75 were also resistant to HFD-induced impairments in glucose tolerance and insulin sensitivity (Fig. 4, C and D). At the end of 14 weeks of HFD, plasma leptin levels were lower in Gpr75+/− and Gpr75−/− mice compared with wild-type mice (fig. S6), whereas adiponectin levels were higher, resulting in a 2- to 10-fold lower leptin-to-adiponectin ratio in Gpr75+/− and Gpr75−/− mice compared with wild type (fig. S6).

Genomic insights at other BMI-associated GPCRs and known monogenic obesity genes from exome sequencing

Gene-burden associations at other GPCRs illustrate the complementarity of large-scale exome sequencing analyses to common variants GWAS in identifying effector genes, establishing directionality of association (i.e., whether LOF in a gene is associated with higher or lower BMI levels), and identifying variants whose functional follow-up can provide biological insights.
In our GWAS fine-mapping analysis, we identified four distinct signals in the 1-Mb region around the CALCR gene (Table S14). Although these common variant associations point to CALCR as a possible effector gene, they do not on their own inform whether reduced CALCR function would be associated with higher or lower BMI. In the exome analysis, we identified a significant association for the burden of rare [alternative allele frequency (AAF) of <0.1%] pLOF and predicted-deleterious missense variants in CALCR with 0.09 SDs (~0.5 kg/m²) higher BMI and 20% higher odds of obesity (OR, 1.20; 95% CI, 1.12, 1.29; \( P = 8.9 \times 10^{-7} \); Table 1 and table S4). In addition, the burden of CALCR pLOF genetic variants alone (i.e., excluding missense variants) was associated with higher BMI (Table S15), indicating that loss of function in CALCR is associated with higher adiposity and obesity risk in humans.

At GIPR, a known BMI locus (16), we identified an exome-wide significant association with lower BMI for the burden of pLOF and predicted deleterious missense variants (Table 1). This association remained statistically significant, albeit attenuated, after accounting for Arg190→Gln (Arg190Gln) or Glu288Gly (table S15), two rare missense variants with uncertain functional consequences previously associated with lower BMI (16).

In our exome sequencing meta-analysis, rare protein-truncating variants in GIPR were associated with lower BMI, with a similar effect size to that of Arg190Gln or Glu288Gly (Fig. 5A and table S15), suggesting that these missense variants may result in a loss of function. We tested this hypothesis in cell-based expression experiments, showing that both Arg190Gln-GIPR and Glu288Gly-GIPR result in a near-complete loss of function with respect to Gs and Gq signaling when agonized with recombinant glucose-dependent insulinotropic polypeptide (Fig. 5, B and C, and fig. S7) as compared with wild type. These results indicate that heterozygous loss of function in GIPR results in lower BMI and obesity risk in humans.

Using exome sequencing, we confirmed the association of pLOF and deleterious missense variants in MC4R with higher BMI (Table 1) and that of gain-of-function variants in MC4R (Val103Ile and Ile251Leu) with lower BMI (table S15). In cell-based experiments, we also show that the Val103Ile variant is resistant to the agouti-related peptide–mediated inhibition of MC4R signaling and confirm the signaling preference toward β-arrestin pathway of this gain-of-function variant (17) (fig. S8).

In the MCPS cohort, we identified an MC4R signal that was fine-mapped to a single likely-causal missense variant (Ile269Asn, 18:60371544: A:T in table S14). This variant was previously associated with childhood and adult obesity in Mexico (35) and experimentally shown to result in a complete LOF of both cyclic adenosine monophosphate signaling and β-arrestin recruitment (17). The variant is extremely rare in non-Latino ancestries, but the allele frequency was 1% in admixed American individuals from the MCPS cohort. Heterozygous MC4R deficiency is considered the most common monogenic form of human obesity (prevalence in cases of severe early-onset obesity, ~6%) (36). In obese individuals of European ancestry from the UKB cohort, the prevalence of pLOF or rare (AAF < 1%) deleterious missense variants in MC4R was 0.4%, indicating that MC4R deficiency is a rare genetic contributor to general obesity in that population. In obese admixed American participants from the MCPS cohort, the prevalence of MC4R deficiency due to the Ile269Asn variant was ~3%. Therefore, the prevalence of heterozygous MC4R deficiency in general obesity is more than sevenfold greater in admixed Americans from Mexico than in Europeans from the UK.

Homozygous MC4R deficiency has been described in only a handful of cases of severe early-onset obesity (36), but the penetrance of obesity in this condition is unknown. In the MCPS cohort, there were 17 homozygous carriers of Ile269Asn, 12 of whom were obese and 5 overweight. These data suggest that homozygous MC4R deficiency might be incompatible with the maintenance of a healthy weight in adulthood.

Our exome-wide study also identified additional large effect-size associations with higher BMI and obesity risk for rare coding variants in other genes, including associations with more than twofold higher odds of obesity for protein-truncating variants in PCSK1 (in a heterozygous state), UBR2 (ubiquitin protein ligase E3 component n-recogin 2), and ANO4 (anoctamin 4) (Table 1 and table S4). These associations were similar in effect size to those for rare (AAF < 1%) pLOF or deleterious missense variants in MC4R, but the frequency of these genotypes was lower than that of MC4R pLOF or deleterious missense alleles (Table 1). Hence, these rare mutations occur in only a small number of people in the general population. However, they may have a large impact on obesity risk in those small groups of carriers.

We used our large exome-sequencing dataset to investigate the association with BMI of heterozygous pLOF variants in five leptin-melanocortin pathway genes where pLOF variants have been previously linked to monogenic obesity (LEP [leptin deficiency], Mendelian...
Inheritance in Man (MIM): 614962, LEPR [leptin receptor deficiency, MIM: 614963], POMC [proopiomelanocortin deficiency, MIM: 609734], PCSK1 [proprotein convertase 1/3 deficiency, MIM: 609955], and MC4R [melanocortin 4 receptor deficiency, MIM: 618406]. Of these, only MC4R deficiency is considered to have autosomal dominant inheritance, while all other deficiencies are considered to be autosomal recessive. Heterozygous carrier status for pLOF variants in LEP, POMC, PCSK1, or MC4R was associated with higher BMI (IVW meta-analysis $P < 0.05$; table S16), whereas heterozygous carrier status for pLOF variants in LEPR was not associated with BMI, suggesting a pure autosomal recessive inheritance for leptin receptor deficiency.

**Rare single-variant exome-wide analysis reveals additional signals at the SOST2 and SRRM2 genes**

In addition to the gene-burden association analyses, we also conducted single-variant association analyses, looking for rare non-synonymous alleles individually associated with BMI (IVW meta-analysis $P < 5 \times 10^{-8}$) conditional on GWAS fine-mapped common variants (29). We identified seven such variants, including five occurring in genes identified in the primary gene-burden analysis (GPR151, SPARC (secreted protein acidic and cystine rich), MC4R, GIPR, and ANKRD27 (ankyrin repeat domain 27); table S17) and two variants in genes not identified in other analyses [Arg2033Pro in SRRM2 (serine/arginine repetitive matrix 2) and Pro91Arg in SOST2 (SOS Ras/Rho guanine nucleotide exchange factor 2); table S17]. The missense variant in SOST2 was associated with 0.05 SDs (0.27 kg/m$^2$) lower BMI per allele. Notably, mutations in SOST2 are associated with autosomal dominant forms of Noonan syndrome (MIM: 616559), a condition that has a higher prevalence of being overweight (37).

Combined use of common variant GWAS fine-mapping and exome-sequencing gene-burden associations to prioritize likely effector genes for BMI

Fine-mapping of GWAS associations in the discovery cohorts of our analysis identified 1905 independent signals led by sentinel common variants (29) (AAF > 1%; table S18), which had a median 95% credible set size of 36 likely-causal variants (interquartile range, 12 to 119). We used MANTRA (Meta-Analysis of Transethnic Association) studies (38), a Bayesian transethnic meta-analysis approach, to estimate associations of fine-mapped signals across ancestries and observed strong evidence of association across datasets (median log$_{10}$ Bayes factor, 7.0; interquartile range, 4.9 to 10.4; table S18).

Of the 1905 signals, 13 (0.7%) fine-mapped to a single non-synonymous sentinel variant with >95% posterior probability of causal association (12 missense variants and 1 splice site variant in 12 genes; table S19). These included a gene identified in our gene-burden analysis (MC4R) and an additional 11 genes. We investigated whether there were associations for the burden of rare (AAF < 1%) pLOF or pLOF plus predicted deleterious (5/5 in silico prediction algorithms) missense variants in these genes prioritized by GWAS fine-mapping. We found evidence of association (IVW meta-analysis $P < 0.05$) for 4 of these 11 genes (36%: table S19). These common and rare variants association pairs included the His48Arg missense variant and the burden of rare pLOF or predicted deleterious missense variants in ADH1B, encoding the key ethanol metabolism enzyme alcohol dehydrogenase 1B. His48Arg is known to associate with higher BMI via increased alcohol consumption (39), and our rare variant analysis corroborates the causal nature of that association and the importance of alcohol consumption in weight regulation.

We used exome sequencing results to investigate whether there were associations for rare
At 6 of the 10 regions, the gene identified in variant associations (identify likely effector genes for the common European or all-ancestry GTEx v8 datasets to physical proximity, common nonsynonymous loci with common variant signals, we used common sentinel variants (table S14). At the six regions there were no nearby fine-mapped GWAS fine-mapping, while at the remaining (sentinel AAF > 1%) signals identified by burden analysis, there were common variant (table S20).

dicted deleterious (5/5 in silico prediction al-
for the burden of rare coding variants in the this analysis can only exclude large-effect asso-
ciations for this individual?

This figure shows the prevalence of obesity (defined as BMI ≥ 30 kg/m²) in heterozygous carriers of pLOF variants in GPR75, noncarriers, or heterozygous carriers of pLOF variants in MC4R within quintiles of a genome-wide polygenic score for higher BMI. P values for interaction between the polygenic score and rare pLOF variants on body mass index were 0.09 and 0.85 for GPR75 and MC4R, respectively. Error bars represent 95% confidence intervals.

nonsynonymous variants in genes at the FTO locus, where common variants have been associated with BMI in early GWASs (11, 12) and where experimental studies have suggested the distant IRX3 and IRX5 as likely effector genes (40). We did not observe an association with BMI for rare coding variants in FTO, IRX3, or IRX5 (table S20). Because pLOF variants in these genes were rare (AAF ≤ 0.02%; table S20), this analysis can only exclude large-effect associations with BMI for heterozygous pLOF variants in those genes. The strongest associations for the burden of rare coding variants in the region was for rare (AAF < 1%) pLOF or predicted deleterious (5/5 in silico prediction algorithms) missense variants in the CHD9 gene (table S20).

At 10 of the 16 1-Mb regions around the genes identified in our exome-wide gene-burden analysis, there were common variant (sentinel AAF > 1%) signals identified by GWAS fine-mapping, while at the remaining six regions there were no nearby fine-mapped common sentinel variants (table S14). At the 10 loci with common variant signals, we used physical proximity, common nonsynonymous variants, and eQTL colocalization in either European or all-ancestry GTEx v8 datasets to identify likely effector genes for the common variant associations (29) (tables S14 and S21). At 6 of the 10 regions, the gene identified in the exome-wide analysis was also one of the genes that were prioritized by common variants associations (table S14). However, at four of those six loci, other genes were also prioritized, resulting in an uncertain effector gene attribution on the basis of common variants alone. In these four loci, gene-burden associations for the gene identified in the exome analysis had an average of 19 orders of magnitude stronger statistical association than any of the other possible effector genes prioritized by common variants [average difference in −log10(p), 19.3; range, 6.3 to 46.6; table S14], suggesting that gene-burden analyses may considerably help to prioritize effector genes at some GWAS-associated loci. At the remaining loci, common variant associations prioritized different genes than the ones identified in the exome-wide analysis (table S14). These results suggest that, at those loci, common variant associations act via different genes than the ones found by exome sequencing or that the gene prioritization from common variant associations did not identify the correct effector gene for those common variant signals.

Polygenic burden influences the penetrance of obesity in carriers of high-impact rare coding alleles

Both common and rare alleles contribute to the risk of general obesity, but their interplay in shaping obesity risk has been understudied because of the lack of datasets with both common and rare variant ascertainment. Here, we generated a genome-wide polygenic score capturing genetic predisposition to higher BMI due to more than 2.5 million common alleles (29) and studied its interplay with rare, large effect-size coding variants in shaping risk for obesity in the population-based UKB cohort. This analysis suggests that polygenic burden influences the penetrance of obesity and the level of BMI in carriers of GPR75 (large-effect protective association) or MC4R (large-effect risk-increasing association) pLOF variants in a linearly additive manner (Fig. 6 and fig. S9). The penetrance of obesity in individuals carrying protein-truncating variants in MC4R varied from less than 30% to more than 60% in people at the bottom versus top quintile for the distribution of the polygenic score (Fig. 6). There was a nearly 60% absolute difference in the prevalence of obesity between extremes of genetic predisposition, that is, GPR75 pLOF carriers in the bottom quintile of polygenic burden and MC4R pLOF carriers in the top quintile (Fig. 6).

Discussion

By conducting a large exome-sequencing study on the influences of rare coding variation on body adiposity, we made a number of observations that advance our understanding of the genes and pathways involved in propensity for and resistance to obesity in humans. We discovered an association for rare protein-truncating genetic variants in GPR75 with lower adiposity and substantial protection against obesity. We validated this human genetic association in a high-fat diet model of obesity in mice, where genetic ablation of Gpr75 was associated with resistance to weight gain, greater insulin sensitivity, and improved glycemic control. In our analysis, the association for pLOF variants in GPR75 showed the largest effect-size genetic association with lower adiposity and protection against obesity at the genome-wide level. The estimated effect size for this association appears to be three to four times larger than the largest effect-size associations for common genetic variants at the FTO locus (11, 13) or for low-frequency gain-of-function variants in MC4R (17). The observation of a consistent association across cohorts from different regions of the world and with different study design as well as multiple ancestries highlights the generalizability of this association in people with various genetic backgrounds and environmental exposures.

Human genetics validation is a predictor of the likely success of drug development programs (41, 42) where the identification of naturally occurring protective alleles has catalyzed the translation from genetic association to therapeutic drug development in a growing number
of examples (17–22, 43–45). Therefore, our findings suggest that GPR75 inhibition could be a therapeutic approach for obesity. The expression of GPR75 in the hypothalamic nuclei and other brain regions, protection against weight gain for Gpr75 knockout mice under high-fat diet challenge, and previous evidence of the role of brain GPCRs in energy balance regulation suggest that this receptor may be implicated in the brain-mediated regulation of energy balance, providing an important direction for future mechanistic research. It will also be important to clarify whether 20-hydroxyeicosatetraenoic acid (20-HETE), an eicosanoid metabolite of arachidonic acid, or C-C motif chemokine ligand 5 (CCL5), a chemokine, which have been previously proposed as putative ligands for GPR75 (46, 47) or other yet-undiscovered ligands are responsible for the link between GPR75 loss of function and body weight regulation.

In addition to GPR75, our agnostic exome analysis identified four other GPCRs expressed in the brain and previously implicated in energy balance regulation. This highlights once again the importance of neurological pathways in obesity risk in humans first shown in family studies of extreme obesity and more recently in genome-wide association analyses of common variants. GPR is expressed in adipose, brain, bone, and other metabolically-active tissues where it acts as the receptor for glucose-dependent insulinotropic polypeptide, an incretin hormone involved in the regulation of insulin secretion, gastric emptying, and other metabolic processes (46). CALCR is the receptor of calcitonin and amylin, a peptide hormone secreted by pancreatic beta cells, which has been shown to promote satiety, delayed gastric emptying, and weight control in patients with type 2 diabetes (48–51). Ablation of CalcL-expressing neurons in the nucleus tractus solitarius has been recently implicated in the disruption of a leptin-independent pathway of appetite regulation in murine models (52). GPR151 encodes a habenular receptor expressing neurons in the nucleus accumbens (53). The association effect size for pLOF and predicted deleterious missense variants in our analysis as well as the previously reported association for the Arg95* allele (39) is tiny (~0.3 kg/m² per allele), suggesting that it may be secondary to other behavioral or neurological phenotypes possibly related to addiction. In addition to these GPCRs, our exome-wide analysis revealed associations with body adiposity for rare coding variants in several other genes. Although in this study we primarily focused on the associations in GPR75 and other GPCRs owing to their more immediate translational potential, the identification of these additional associations in our large study provides an initial foundation for understanding the role of these genes in the regulation of body fat and obesity risk.

Our study illustrates the power and versatility of massive-scale exome sequencing in population- or health system–based cohort studies as a genetic discovery approach complementary to common variant GWASs and pedigree-based studies. We show the utility of this approach in (i) discovering large-effect protective or risk-increasing associations for genes and individual rare coding variants, (ii) identifying functional variants that can be studied in vitro for biological insight, (iii) prioritizing effector genes for common variant signals identified by fine-mapping, and (iv) understanding the impact of rare variants on complex traits such as obesity across populations and their interplay with common variation in shaping disease risk or resistance. The results from this exome-sequencing analysis in more than 640,000 people suggest that inhibition of GPR75 may be a therapeutic strategy for obesity and illustrate the power of massive-scale exome sequencing for the identification of large-effect coding variant associations and drug targets for complex traits.

Methods summary

Detailed materials and methods are provided in the supplementary materials (29). Briefly, we performed high-coverage whole-exome sequencing in 645,626 individuals, including 428,719 individuals of European ancestry from the UK Biobank cohort, 121,061 individuals of European ancestry from the MyCode Community Health Initiative cohort from the US-based Geisinger Health System, and 95,846 individuals of admixed American ancestry from the Mexico City Prospective Study. The outcome measure was BMI, calculated as weight in kilograms divided by the square of standing height in meters. We estimated associations with BMI for the burden of rare nonsynonymous variants in each sequenced gene by fitting mixed-effects regression models using BOLT-LMM v2.3.4 (54) or REGENIE v1.0 (55). Results across studies were pooled by inverse-variance weighted meta-analysis. Consistent with previous literature (27, 28), for our primary gene-burden association analysis, we considered a threshold of exome-wide statistical significance of P < 3.6 × 10⁻⁶, a Bonferroni correction for 20,000 genes, and seven variant selection models. In parallel, to leverage evidence from common variants, we performed genome-wide association studies of imputed common alleles in the same discovery dataset used for our exome sequencing analysis. We leveraged fine-mapping of common alleles, formal conditional analyses, physical proximity mapping, linkage disequilibrium with common nonsynonymous alleles, eQTL colocalization, and polygenic score analysis to highlight the complementarity of evidence from exome sequencing and GWASs of common variants. To illustrate the translational value of exome sequencing–identified rare coding associations, we performed targeted in vitro and in vivo experiments. We performed in vitro expression studies for the functional characterization of naturally-occurring variants in GPR75, MC4R, and GIPR. We also developed genetically engineered knockout Gpr75⁻/⁻ mouse strains using the VelociGene technology (56, 57) and studied the weight gain, glycemic, and insulinnemic phenotypes of knockout Gpr75⁻/⁻, heterozygous Gpr75⁻/+ mice, and wild-type Gpr75⁺/+ mice in a high-fat diet model.

REFERENCES AND NOTES

1. Y. Zhang et al., Positional cloning of the mouse obese gene and its human homolog. Nature 372, 425–432 (1994). doi: 10.1038/372425a0; pmid: 9784236
2. W. Fan, B. A. Boston, R. A. Kesterson, V. J. Hruby, R. D. Cone. Role of melancortinergic neurons in feeding and the agouti obesity syndrome. Nature 385, 165–168 (1997). doi: 10.1038/385165a0; pmid: 8990020
3. D. Huzar et al., Targeted disruption of the melanocortin-4 receptor results in obesity in mice. Cell 88, 131–141 (1997). doi: 10.1016/S0092-8674(08)8856-6; pmid: 9013999
4. C. T. Montague et al., Congenital leptin deficiency is associated with severe early-onset obesity in humans. Nature 387, 903–908 (1997). doi: 10.1038/43185; pmid: 9202122
5. S. D’Ovidio et al., Brief report: Impaired processing of prohormones associated with abnormalities of glucose homeostasis and adrenal function. N. Engl. J. Med. 333, 1386–1390 (1995). doi: 10.1056/NEJM199512333332104; pmid: 747109
6. R. S. Jackson et al., Obesity and impaired prohormone processing associated with mutations in the human prohormone convertase 1 gene. Nat. Genet. 16, 303–306 (1997). doi: 10.1038/ng0797-303; pmid: 9207799
7. K. D’Clement et al., A mutation in the human leptin receptor gene causes obesity and pituitary dysfunction. Nature 392, 398–401 (1998). doi: 10.1038/392391a1; pmid: 9537324
8. G. S. Yoo et al., A frameshift mutation in MC4R associated with dominantly inherited human obesity. Nat. Genet. 20, 111–112 (1998). doi: 10.1038/2404; pmid: 9716688
9. C. Vaisse, K. D’Clement, B. Guy-Grand, P. Froguel. A frameshift mutation in human MC4R is associated with a dominant form of obesity. Nat. Genet. 20, 113–114 (1998). doi: 10.1038/2407; pmid: 9716699
10. H. Krude et al., Severe early-onset obesity, adrenal insufficiency and red hair pigmentation caused by POMC mutations in humans. Nat. Genet. 19, 155–157 (1998). doi: 10.1038/1059; pmid: 9620771
11. T. M. Frayling et al., A common variant in the FTO gene is associated with body mass index and predisposes to childhood and adult obesity. Science 316, 889–894 (2007). doi: 10.1126/ science1141634; pmid: 17434869
12. R. J. Loos et al., Common variants near MC4R are associated with fat mass, weight and risk of obesity. Nat. Genet. 40, 768–775 (2008). doi: 10.1038/ng.140; pmid: 18454148
13. A. E. Locke et al., Genetic studies of body mass index yield new insights for obesity biology. Nature 518, 197–206 (2015). doi: 10.1038/nature14177; pmid: 25674413
14. L. Yengo et al., Meta-analysis of genome-wide association studies for height and body mass index in ~700,000 individuals of European ancestry. Hum. Mol. Genet. 27, 3641–3649 (2018). doi: 10.1093/hmg/ddy271; pmid: 30234842
15. A. V. Khera et al., Polygenic prediction of weight and obesity trajectories from birth to adulthood. Cell 177, 587–596.e9 (2019). doi: 10.1016/j.cellet.2019.03.028; pmid: 31022795
16. V. Turcotti et al., Protein-altering variants associated with body mass index implicate pathways that control energy intake and expenditure in obesity. Nat. Genet. 50, 26–41 (2018). doi: 10.1038/s41588-017-0011-x; pmid: 29073807
17. L. A. Lotta et al., Human gain-of-function MC4R variants show signaling bias and protect against obesity. Cell 177, 597–607.e9 (2019). doi: 10.1016/j.cellet.2019.03.044; pmid: 31002796
Sequencing of 640,000 exomes identifies GPR75 variants associated with protection from obesity

Parsa Akbari, Ankit Gilani, Olukayode Sosina, Jack A. Kosmicki, Lori Khrimian, Yi-Ya Fang, Trikaldarshi Persaud, Victor Garcia, Dylan Sun, Alexander Li, Joelle Mbatchou, Adam E. Locke, Christian Benner, Niek Verweij, Nan Lin, Sakib Hossain, Kevin Agostinucci, Jonathan V. Pascale, Ercument Dirice, Michael Dunn, Regeneron Genetics Center, DiscovEHR Collaboration, William E. Kraus, Svati H. Shah, Yii-Der I. Chen, Jerome I. Rotter, Daniel J. Rader, Olle Melander, Christopher D. Still, Tooraj Mirshahi, David J. Carey, Jaime Berumen-Campos, Pablo Kuri-Morales, Jesus Alegre-Díaz, Jason M. Torres, Jonathan R. Embersen, Rory Collins, Suguanthi Balasubramanian, Alicia Hawes, Marcus Jones, Brian Zambrowicz, Andrew J. Murphy, Charles Paulding, Giovanni Coppola, John D. Overton, Jeffrey G. Reid, Alan R. Shuldiner, Michael Cantor, Hyun M. Kang, Goncalo R. Abecasis, Katia Karalis, Aris N. Economides, Jonathan Marchini, George D. Yancopoulos, Mark W. Sleeman, Judith Altarejos, Giusy Della Gatta, Roberto Tapia-Conyer, Michal L. Schwartzman, Aris Baras, Manuel A. R. Ferreira and Luca A. Lotta

Science 373 (6550), eabf8683.
DOI: 10.1126/science.eabf8683

How genes affect human obesity

Obesity is linked to many human diseases, including diabetes, cancer, and heart disease. There is thus great interest in understanding how genes predispose individuals to, or protect individuals from, obesity. Akbari et al. sequenced more than 600,000 exomes from the United Kingdom, the United States, and Mexico and identified 16 rare coding variants (see the Perspective by Yeo and O’Rahilly). Some of the alleles associated with body mass index (BMI) were brain-expressed G protein-coupled receptors. One variant allele was found in Mexican populations at low frequency and was associated with lower BMI. Deletion of this gene in mice resulted in a resistance to weight gain, suggesting that this gene provides an avenue of study for the prevention or treatment of obesity.

Science, eabf8683, this issue p. eabf8683; see also abh3556, p. 30

Use of this article is subject to the Terms of Service

Science (print ISSN 0036-8075; online ISSN 1095-9203) is published by the American Association for the Advancement of Science, 1200 New York Avenue NW, Washington, DC 20005. The title Science is a registered trademark of AAAS.

Copyright © 2021 The Authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original U.S. Government Works
