RESEARCH ARTICLE SUMMARY

HUMAN GENOMICS

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

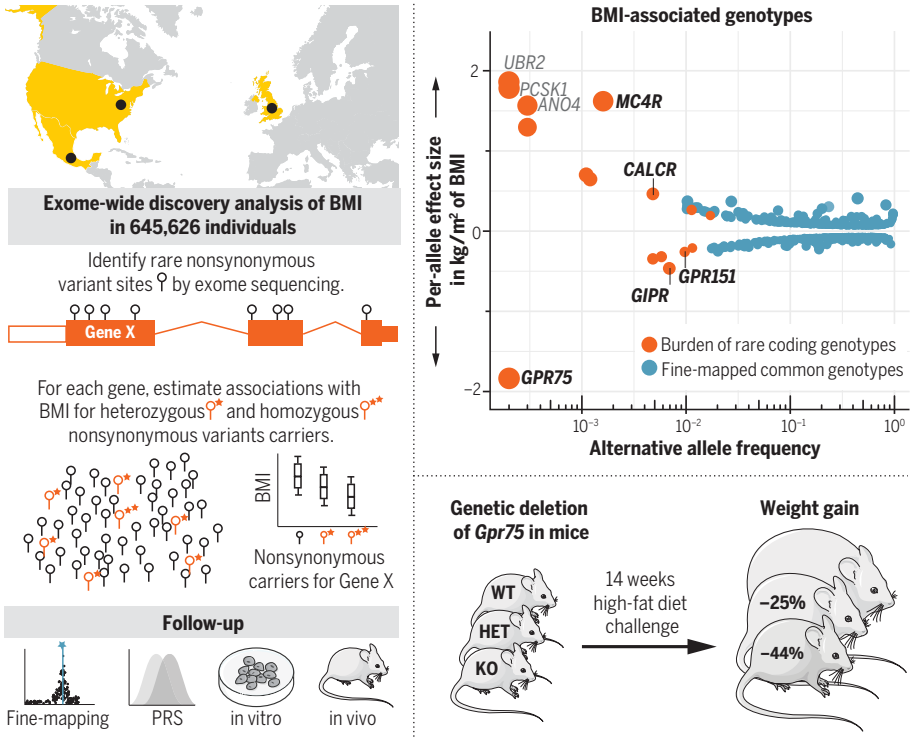
Parsa Akbari†, Ankit Gilani†, Olukayode Sosina†, Jack A. Kosmicki, Lori Khrimian, Yi-Ya Fang, Trikaldarshi Persaud, Victor Garcia, Dylan Sun, Alexander Li, Joelle Mbatchou, Adam E. Locke, Christian Benner, Niek Verweij, Nan Lin, Sakib Hossain, Kevin Agostinucci, Jonathan V. Pascale, Ercument Dirice, 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. Emberson, Rory 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, Mark W. Sleeman, 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 $P < 3.6 \times 10^{-7}$), 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/m^2 lower BMI, 5.3 kg lower bodyweight, and 54% lower odds of obesity in heterozygous carriers. Knock out 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 [Arg^{190}→Gln (Arg190Gln), Glu288Gly], 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. ■



Exome sequencing–based discovery of BMI-associated genes. (Left) Design for the discovery gene-burden analysis, with a depiction of follow-up analyses along the bottom. (Top right) Relationship between allele frequency and effect-size estimates for BMI-associated genotypes. (Bottom right) Weight gain for *Gpr75*$^{+/+}$ (wild type, WT), *Gpr75*$^{-/+}$ (heterozygous, HET), and *Gpr75*$^{-/-}$ (knockout, KO) mice during a high-fat diet challenge. PRS, polygenic risk score.

The list of author affiliations is available in the full article online.
*Corresponding author. Email: luca.lotta@regeneron.com (L.A.L.); aris.baras@regeneron.com (A.B.)
†These authors contributed equally to this work.
‡These authors contributed equally to this work.
Cite this article as P. Akbari *et al.*, *Science* 373, eabf8683 (2021). DOI: 10.1126/science.abf8683

READ THE FULL ARTICLE AT
https://doi.org/10.1126/science.abf8683

Downloaded from http://science.sciencemag.org/ on July 2, 2021

RESEARCH ARTICLE

HUMAN GENOMICS

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

Parsa Akbari[1]†, Ankit Gilani[2]†, Olukayode Sosina[1]†, Jack A. Kosmicki[1], Lori Khrimian[3], Yi-Ya Fang[3], Trikaldarshi Persaud[1], Victor Garcia[2], Dylan Sun[1], Alexander Li[1], Joelle Mbatchou[1], Adam E. Locke[1], Christian Benner[1], Niek Verweij[1], Nan Lin[1], Sakib Hossain[2], Kevin Agostinucci[2], Jonathan V. Pascale[2], Ercument Dirice[2], Michael Dunn[3], Regeneron Genetics Center‡, DiscovEHR Collaboration‡, William E. Kraus[4,5], Svati H. Shah[6,7], Yii-Der I. Chen[8], Jerome I. Rotter[8], Daniel J. Rader[9], Olle Melander[10,11], Christopher D. Still[12], Tooraj Mirshahi[12], David J. Carey[12], Jaime Berumen-Campos[13], Pablo Kuri-Morales[13], Jesus Alegre-Díaz[13], Jason M. Torres[14], Jonathan R. Emberson[14], Rory Collins[14], Suganthi Balasubramanian[1], Alicia Hawes[1], Marcus Jones[1], Brian Zambrowicz[3], Andrew J. Murphy[3], Charles Paulding[1], Giovanni Coppola[1], John D. Overton[1], Jeffrey G. Reid[1], Alan R. Shuldiner[1], Michael Cantor[1], Hyun M. Kang[1], Goncalo R. Abecasis[1], Katia Karalis[1], Aris N. Economides[1,3], Jonathan Marchini[1], George D. Yancopoulos[3], Mark W. Sleeman[3], Judith Altarejos[3], Giusy Della Gatta[1], Roberto Tapia-Conyer[13]§, Michal L. Schwartzman[2]§, Aris Baras[1]§*, Manuel A. R. Ferreira[1]§, Luca A. Lotta[1]§*

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 glycemic control 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. Studies in mouse models (*1*–*3*) and human forms of extreme, early-onset obesity (*4*–*10*) have uncovered the influence of the leptin-melanocortin system on appetite regulation. Additionally, genome-wide association studies (GWASs) of body mass index (BMI) have highlighted the polygenic contribution to the inherited basis of adiposity, identifying thousands of common genetic variants, each with small effect size, and reaffirming the broad influence of the central nervous system on body mass regulation (*11*–*15*).

Studies of rare protein-coding variants have helped translate genetic associations into biological and therapeutic insights (*4*, *6*–*10*, *16*–*24*). 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 multiethnic 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 [inverse-variance weighted (IVW) meta-analysis $P < 3.6 \times 10^{-7}$, 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), *CALCR* (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).

Downloaded from http://science.sciencemag.org/ on July 2, 2021

[1]Regeneron Genetics Center, Regeneron Pharmaceuticals Inc., Tarrytown, NY 10591, USA. [2]Department of Pharmacology and Medicine, New York Medical College School of Medicine, Valhalla, NY 10595, USA. [3]Regeneron Pharmaceuticals Inc., Tarrytown, NY 10591, USA. [4]Division of Cardiology, Duke University Medical Center, Durham, NC 27710, USA. [5]Duke Center for Living, Duke University Medical Center, Durham, NC 27705, USA. [6]Department of Medicine, Duke University Medical Center, Durham, NC 27710, USA. [7]Duke Molecular Physiology Institute, Duke University Medical Center, Durham, NC 27701, USA. [8]Institute for Translational Genomics and Population Sciences, The Lundquist Institute for Biomedical Innovation, and Department of Pediatrics, Harbor-UCLA Medical Center, Torrance, CA 90502, USA. [9]Department of Genetics, Perelman School of Medicine, University of Pennsylvania, PA 19104, USA. [10]Department of Clinical Sciences Malmö, Lund University, 221 00 Malmö, Sweden. [11]Department of Emergency and Internal Medicine, Skåne University Hospital, 214 28, Malmö, Sweden. [12]Geisinger Obesity Institute, Geisinger Health System, Danville, PA 17882, USA. [13]Faculty of Medicine, National Autonomous University of Mexico, Copilco Universidad, Coyoacán, 4360 Ciudad de México, Mexico. [14]Nuffield Department of Population Health, University of Oxford, Oxford OX3 7LF, England, UK.
*Corresponding author. Email: luca.lotta@regeneron.com (L.A.L.); aris.baras@regeneron.com (A.B.)
†These authors contributed equally to this work. ‡A complete list of investigators contributing to the Regeneron Genetics Center and DiscovEHR Collaboration is provided at the end of this paper.
§These authors contributed equally to this work.

Table 1. Associations with body mass index in the exome-wide gene-burden analysis. The table reports genes for which the gene burden of rare nonsynonymous variants was associated with body mass index at the exome-wide level of statistical significance ($P < 3.6 \times 10^{-7}$). Analyses were performed in 645,626 participants from the UKB, GHS, and MCPS studies. Genomic coordinates reflect chromosome and physical position in base pairs according to Genome Reference Consortium Human Build 38. Abbreviations: CI, confidence interval; SD, standard deviation; BMI, body mass index; AAF, alternative allele frequency; RR, reference-reference genotype; RA, reference-alternative heterozygous genotype; AA, alternative-alternative homozygous genotype; pLOF, predicted loss of function; missense (1/5), missense variant predicted to be deleterious by at least 1 out of 5 in silico prediction algorithms; missense (5/5), missense variant predicted to be deleterious by 5 out of 5 in silico prediction algorithms.

Gene	Genomic coordinates	Genetic exposure, variant type; frequency cutoff in %	Beta (95% CI) per allele in SD units of BMI	*P*	AAF (fraction of 1)	Genotype counts (RR|RA|AA genotypes)	Beta (95% CI) per allele in kg/m^2 units of BMI	Beta (95% CI) per allele in kg of body weight	Beta (95% CI) per allele in lb of body weight
UHMK1	1: 162497250	pLOF plus missense (5/5); AAF < 1%	−0.06 (−0.09, −0.04)	4.9×10^{-8}	0.0048	639,499|6,117|10	−0.3 (−0.5, −0.2)	−1.0 (−1.4, −0.6)	−2.2 (−3.0, −1.4)
GPR75	2: 53853133	pLOF; AAF < 1%	−0.34 (−0.46, −0.22)	2.6×10^{-8}	0.0002	645,398|228|0	−1.8 (−2.5, −1.2)	−5.3 (−7.1, −3.4)	−11.6 (−15.7, −7.5)
ROBO1	3: 78598912	pLOF; AAF < 1%	0.24 (0.15, 0.32)	5.6×10^{-8}	0.0003	645,182|444|0	1.3 (0.8, 1.7)	3.7 (2.4, 5.0)	8.1 (5.2, 11.1)
KIAA1109	4: 122170647	pLOF; AAF < 1%	0.13 (0.09 0.18)	5.1×10^{-8}	0.0011	644,238|1,388|0	0.7 (0.5, 1.0)	2.1 (1.3, 2.9)	4.6 (3.0, 6.3)
PCSK1	5: 96393000	pLOF; AAF < 1%	0.33 (0.21, 0.44)	1.0×10^{-8}	0.0002	645,368|258|0	1.8 (1.2, 2.4)	5.1 (3.3, 6.8)	11.2 (7.4, 15.0)
GPR151	5: 146514853	pLOF plus missense (5/5); AAF < 1%	−0.05 (−0.06, −0.03)	5.7×10^{-9}	0.0097	633,209|12,372|45	−0.3 (−0.4, −0.2)	−0.7 (−1.0, −0.5)	−1.7 (−2.2, −1.1)
SPARC	5: 151663570	pLOF plus missense (1/5); AAF < 1%	0.05 (0.03, 0.06)	1.2×10^{-10}	0.0113	631,094|14,494|38	0.3 (0.2, 0.4)	0.8 (0.5, 1.0)	1.7 (1.2, 2.2)
UBR2	6: 42564319	pLOF; AAF < 1%	0.34 (0.23, 0.45)	2.4×10^{-9}	0.0002	645,363|263|0	1.8 (1.2, 2.4)	5.2 (3.5, 7.0)	11.5 (7.8, 15.3)
CALCR	7: 93426355	pLOF plus missense (1/5); AAF < 0.1%	0.09 (0.06, 0.11)	1.4×10^{-13}	0.0048	639,416|6,208|2	0.5 (0.3, 0.6)	1.3 (1.0, 1.7)	3.0 (2.2, 3.8)
PDE3B	11: 14644075	pLOF; AAF < 1%	0.12 (0.08 0.17)	1.7×10^{-7}	0.0012	644,074|1,552|0	0.7 (0.4, 0.9)	1.9 (1.2, 2.6)	4.2 (2.6, 5.7)
ANO4	12: 100901785	pLOF; AAF < 1%	0.29 (0.19, 0.38)	1.1×10^{-9}	0.0003	645,248|378|0	1.6 (1.1, 2.1)	4.5 (3.0, 5.9)	9.9 (6.7, 13.0)
KIAA0586	14: 58427628	pLOF plus missense (5/5); AAF < 1%	−0.04 (−0.05, −0.02)	3.1×10^{-7}	0.0115	630,769|14,844|13	−0.2 (−0.3, −0.1)	−0.6 (−0.8, −0.4)	−1.3 (−1.8, −0.8)
MC4R	18: 60371350	pLOF plus missense (5/5); AAF < 1%	0.30 (0.26, 0.34)	4.0×10^{-48}	0.0016	643,589|2,036|1	1.6 (1.4, 1.8)	4.6 (4.0, 5.3)	10.2 (8.8, 11.6)
DPP9	19: 4676563	pLOF plus missense (5/5); AAF < 1%	−0.06 (−0.08, −0.04)	2.0×10^{-8}	0.0058	638,141|7,469|16	−0.3 (−0.4, −0.2)	−0.9 (−1.3, −0.6)	−2.0 (−2.8, −1.3)
ANKRD27	19: 32598144	pLOF plus missense (1/5); AAF < 1%	0.04 (0.02, 0.05)	5.9×10^{-9}	0.0171	623,572|22,016|38	0.2 (0.1, 0.3)	0.6 (0.4, 0.8)	1.3 (0.8, 1.7)
GIPR	19: 45669520	pLOF plus missense (1/5); AAF < 1%	−0.09 (−0.11, −0.07)	9.3×10^{-19}	0.0069	636,680|8,937|9	−0.5 (−0.6, −0.4)	−1.3 (−1.6, −1.0)	−3.0 (−3.6, −2.3)

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/m^2 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 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; $P = 6.9 \times 10^{-9}$] (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; $P = 6.9 \times 10^{-5}$], 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 underweight

Downloaded from http://science.sciencemag.org/ on July 2, 2021

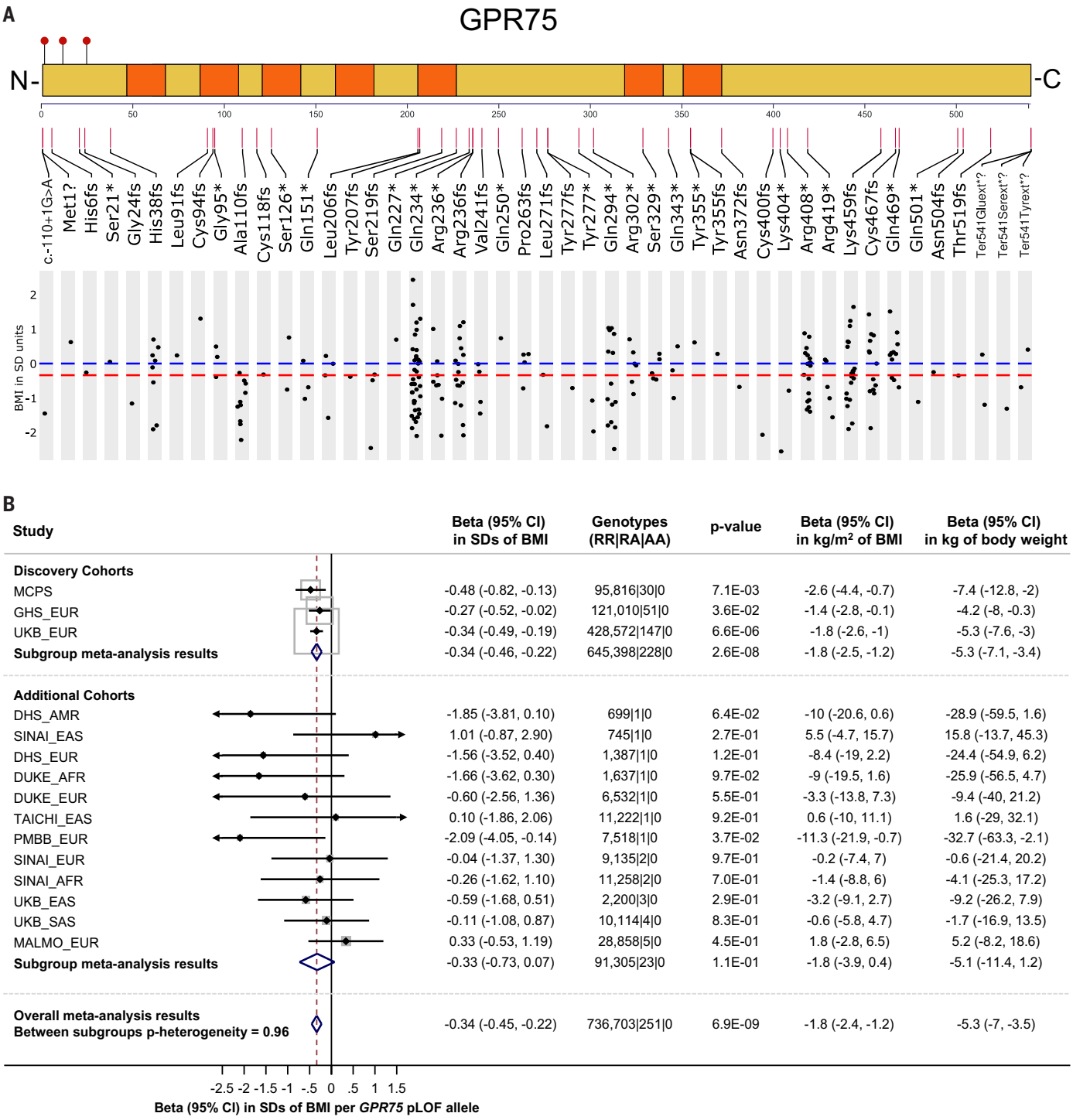


Fig. 1. Protein-truncating variants in *GPR75* associated with lower body mass index in humans. (**A**) Linear model of the GPR75 protein and its domains (top; intra- and extracellular domains in yellow, transmembrane domains in orange), the distribution on the GPR75 protein of 46 pLOF variants found 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 Metabochip 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.

Downloaded from http://science.sciencemag.org/ on July 2, 2021

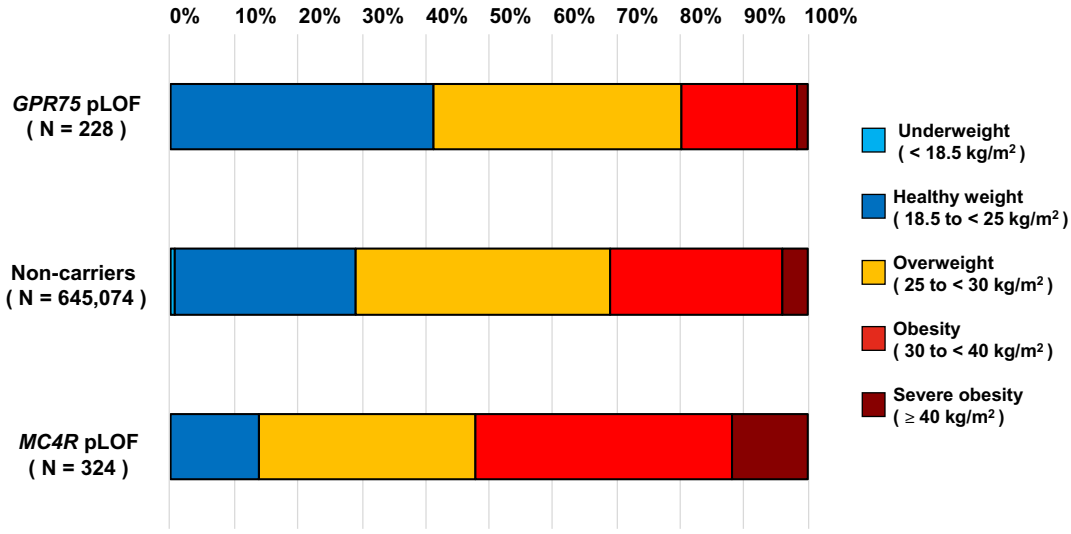


Fig. 2. Distribution in body mass index categories for carriers and noncarriers of predicted loss-of-function variants in *GPR75* or *MC4R*. Distribution of heterozygous carriers of pLOF genetic variants in *GPR75* (top), noncarriers (middle), and heterozygous carriers of pLOF genetic variants in *MC4R* (bottom) in body mass index categories according to the World Health Organization's classification in the UKB, GHS, and MCPS cohorts.

(Fig. 2). 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 ankyrin repeat and SOCS box containing 3 (*ASB3*) (fig. S3). 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. S3). 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 coding variants in *ASB3* or in the *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 *GPR75-ASB3* across multiple statistical models with different variant annotation and allele frequency inclusion criteria (table S7), nor was there an association for pLOF variants in either *ASB3* or *GPR75-ASB3* (table S7). Finally, we estimated the association with BMI for the burden of rare pLOF variants in *GPR75* conditional upon *ASB3* and *GPR75-ASB3* genotypes. The association of *GPR75* pLOF variants with lower BMI was unaffected by adjusting for *ASB3* and *GPR75-ASB3* genotypes (table S8). 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 genome-wide 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; $P = 1.3 \times 10^{-9}$), 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 at 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. 1A 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 (Ala110fs, Ser219fs, Gln234*, Cys400fs, and Lys404*) 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

Downloaded from http://science.sciencemag.org/ on July 2, 2021

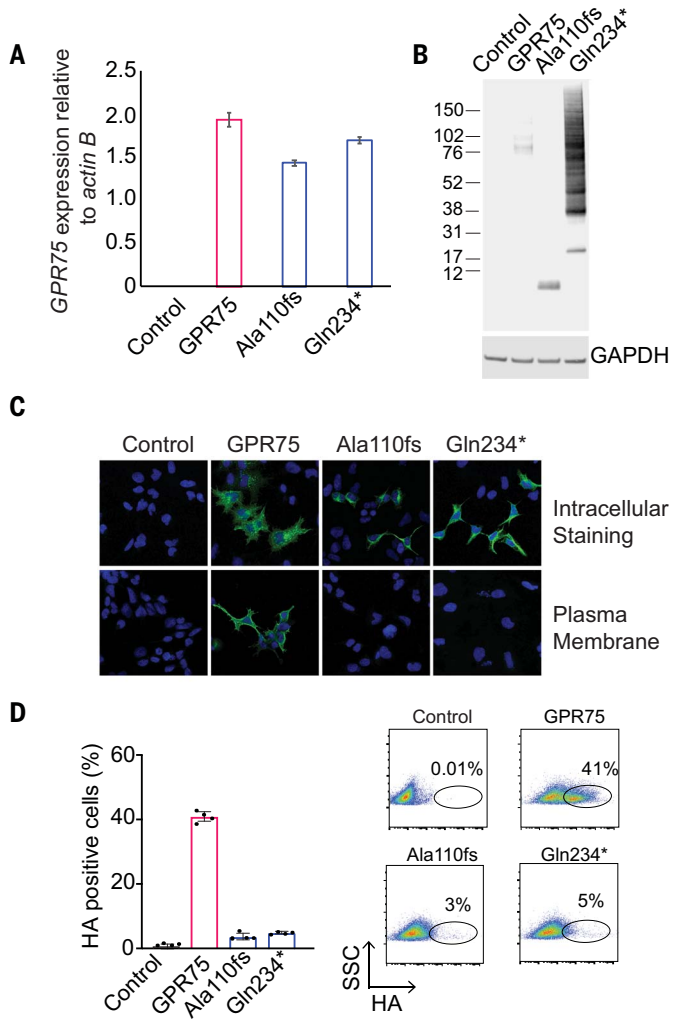


Fig. 3. In vitro expression studies of two predicted loss-of-function genetic variants in *GPR75*. (**A**) Results of quantitative reverse transcription polymerase chain reaction experiments which measured *GPR75* mRNA levels. Expression of *GPR75* was calculated relative to the beta-actin gene. Values 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/m^2 lower BMI (IVW meta-analysis, $P = 4.1 \times 10^{-7}$) as compared with a −1.4 kg/m^2 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 bioimpedance 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 phenotypes tested; 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- and 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.

Downloaded from http://science.sciencemag.org/ on July 2, 2021

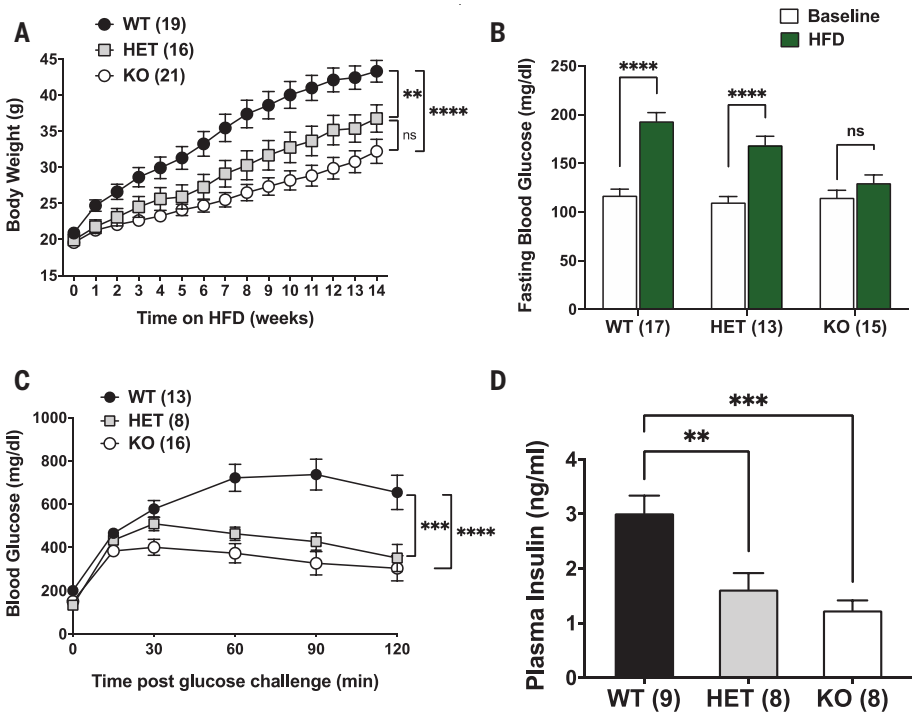


Fig. 4. Weight gain during high-fat diet and metabolic phenotype in mice with a genetic deletion of *Gpr75*. (**A**) Weekly body weight gain during a 14-week high-fat diet (HFD) challenge. (**B**) Changes in fasting blood glucose before and after the high-fat diet challenge. (**C**) Results of a glucose tolerance test at the end of the 14-week high-fat diet challenge. (**D**) Plasma insulin at the end of the 14-week high-fat diet challenge. Each panel shows results in $Gpr75^{+/+}$ (WT), $Gpr75^{+/-}$ (HET), and $Gpr75^{-/-}$ (KO) mice. Number of mice included in each group and analysis is shown in parentheses. Results are presented as mean ± standard error. Abbreviations: ns, not statistically significant; $^{**}P < 0.01$, $^{***}P < 0.001$, $^{****}P < 0.0001$ by two-way analysis of variance with Tukey's multiple comparisons test.

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/m^2) 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 Arg^{190}→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-recognin 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

Downloaded from http://science.sciencemag.org/ on July 2, 2021

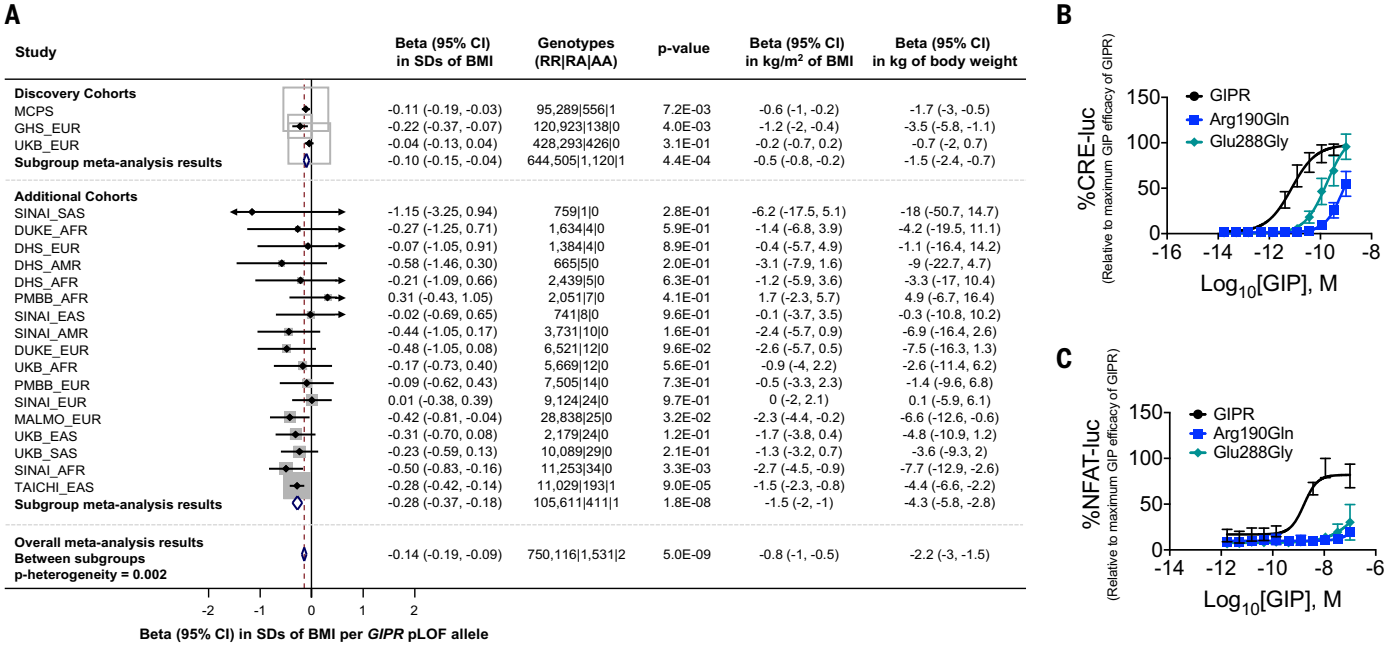


Fig. 5. Association of protein-truncating genetic variants in *GIPR* with lower body mass index. (**A**) Association with BMI for pLOF genetic variants in *GIPR* across both discovery and additional cohorts. Abbreviations: CI, confidence interval; RR, reference-reference genotype; RA, reference-alternative 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 Metabochip 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. (**B**) In vitro expression results using the CRE-reporter assay for the Arg190Gln and Glu288Gly missense variants in *GIPR*, which were associated with lower BMI. (**C**) In vitro expression results using the nuclear factor of activated T cells (NFAT)–dependent assay for the variants. Both missense variants showed lower glucose-dependent insulinotropic polypeptide (GIP)–induced Gs signaling and lower GIP-induced Gq signaling than wild type.

Inheritance in Man (MIM): 614962], *LEPR* [leptin receptor deficiency, MIM: 614963], *POMC* [proopiomelanocortin deficiency, MIM: 609734], *PCSK1* [proprotein convertase 1/3 deficiency, MIM: 600955], 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 *SOS2* and *SRRM2* genes

In addition to the gene-burden association analyses, we also conducted single-variant association analyses, looking for rare nonsynonymous 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 cysteine 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 Pro191Arg in *SOS2* (SOS Ras/Rho guanine nucleotide exchange factor 2); table S17]. The missense variant in *SOS2* was associated with 0.05 SDs (0.27 kg/m^2) lower BMI per allele. Notably, mutations in *SOS2* are associated with autosomal dominant forms of Noonan syndrome (MIM: 616559), a condition that has been associated with a lower 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 nonsynonymous 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

Downloaded from http://science.sciencemag.org/ on July 2, 2021

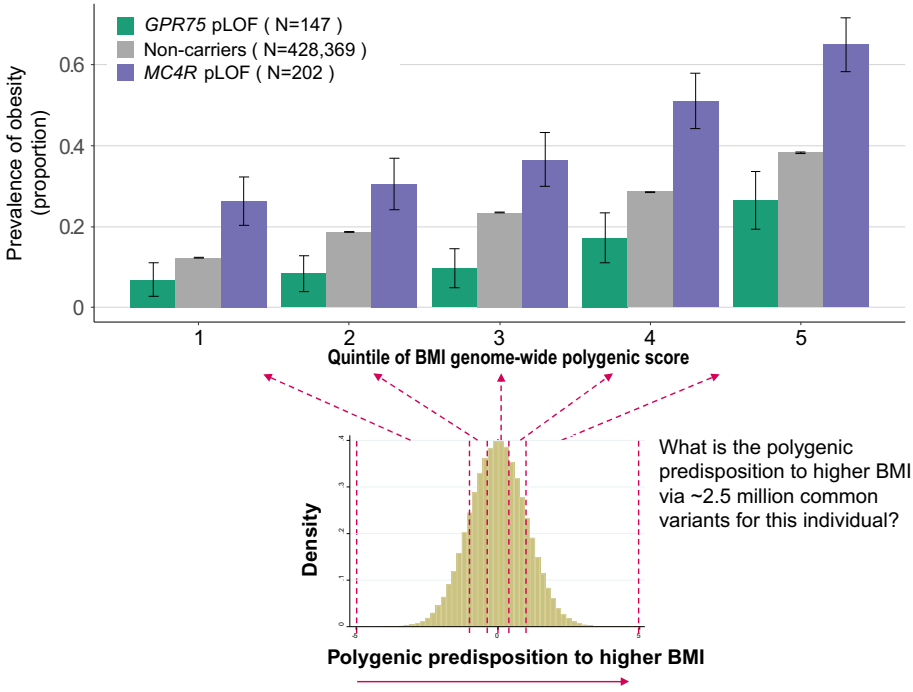


Fig. 6. Interplay of a genome-wide polygenic score for higher body mass index with rare genetic mutations in *GPR75* and *MC4R*. This figure shows the prevalence of obesity (defined as BMI ≥ 30 kg/m^2) in heterozygous carriers of pLOF variants in *GPR75*, noncarriers, or heterozygous carriers of pLOF 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

Downloaded from http://science.sciencemag.org/ on July 2, 2021

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-wide 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. GIPR 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 (*48*). 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 (*49*–*51*). Ablation of *Calcr*-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 involved in addictive behavior and differential food-intake response to nicotine (*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 (*33*) is tiny (−0.3 kg/m^2 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 accounting for population stratification and relatedness 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 $P < 3.6 \times 10^{-7}$, 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 insulinemic phenotypes of knockout*Gpr75*$^{-/-}$, heterozygous *Gpr75*$^{-/+}$, 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 homologue. *Nature* **372**, 425–432 (1994). doi: 10.1038/372425a0; pmid: 7984236
2. W. Fan, B. A. Boston, R. A. Kesterson, V. J. Hruby, R. D. Cone, Role of melanocortinergic neurons in feeding and the agouti obesity syndrome. *Nature* **385**, 165–168 (1997). doi: 10.1038/385165a0; pmid: 8990120
3. D. Huszar *et al*., Targeted disruption of the melanocortin-4 receptor results in obesity in mice. *Cell* **88**, 131–141 (1997). doi: 10.1016/S0092-8674(00)81865-6; pmid: 9019399
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. O'Rahilly *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/NEJM199511233332104; pmid: 7477119
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. Clément *et al*., A mutation in the human leptin receptor gene causes obesity and pituitary dysfunction. *Nature* **392**, 398–401 (1998). doi: 10.1038/32911; pmid: 9537324
8. G. S. Yeo *et al*., A frameshift mutation in *MC4R* associated with dominantly inherited human obesity. *Nat. Genet.* **20**, 111–112 (1998). doi: 10.1038/2404; pmid: 9771698
9. C. Vaisse, K. 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: 9771699
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: 10.1038/509; 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/science.1141634; 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: 25673413
14. L. Yengo *et al*., Meta-analysis of genome-wide association studies for height and body mass index in ~700000 individuals of European ancestry. *Hum. Mol. Genet.* **27**, 3641–3649 (2018). doi: 10.1093/hmg/ddy271; pmid: 30124842
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.cell.2019.03.028; pmid: 31002795
16. V. Turcot *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: 29273807
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.cell.2019.03.044; pmid: 31002796

Downloaded from http://science.sciencemag.org/ on July 2, 2021

18. K. Musunuru *et al*., Exome sequencing, *ANGPTL3* mutations, and familial combined hypolipidemia. *N. Engl. J. Med.* **363**, 2220–2227 (2010). doi: 10.1056/NEJMoa1002926; pmid: 20942659
19. F. E. Dewey *et al*., Genetic and pharmacologic inactivation of ANGPTL3 and cardiovascular disease. *N. Engl. J. Med.* **377**, 211–221 (2017). doi: 10.1056/NEJMoa1612790; pmid: 28538136
20. N. S. Abul-Husn *et al*., A protein-truncating *HSD17B13* variant and protection from chronic liver disease. *N. Engl. J. Med.* **378**, 1096–1106 (2018). doi: 10.1056/NEJMoa1712191; pmid: 29562163
21. A. B. Jørgensen, R. Frikke-Schmidt, B. G. Nordestgaard, A. Tybjærg-Hansen, Loss-of-function mutations in *APOC3* and risk of ischemic vascular disease. *N. Engl. J. Med.* **371**, 32–41 (2014). doi: 10.1056/NEJMoa1308027; pmid: 24941082
22. J. Crosby *et al*., Loss-of-function mutations in *APOC3*, triglycerides, and coronary disease. *N. Engl. J. Med.* **371**, 22–31 (2014). doi: 10.1056/NEJMoa1307095; pmid: 24941081
23. J. Flannick *et al*., Loss-of-function mutations in *SLC30A8* protect against type 2 diabetes. *Nat. Genet.* **46**, 357–363 (2014). doi: 10.1038/ng.2915; pmid: 24584071
24. G. Marenne *et al*., Exome sequencing identifies genes and gene sets contributing to severe childhood obesity, linking *PHIP* variants to repressed *POMC* transcription. *Cell Metab.* **31**, 1107–1119.e12 (2020). doi: 10.1016/j.cmet.2020.05.007; pmid: 32492392
25. F. E. Dewey *et al*., Distribution and clinical impact of functional variants in 50,726 whole-exome sequences from the DiscovEHR study. *Science* **354**, aaf6814 (2016). doi: 10.1126/science.aaf6814; pmid: 28008009
26. C. V. Van Hout *et al*., Exome sequencing and characterization of 49,960 individuals in the UK Biobank. *Nature* **586**, 749–756 (2020). doi: 10.1038/s41586-020-2853-0; pmid: 33087929
27. J. Flannick *et al*., Exome sequencing of 20,791 cases of type 2 diabetes and 24,440 controls. *Nature* **570**, 71–76 (2019). doi: 10.1038/s41586-019-1231-2; pmid: 31118516
28. R. Do *et al*., Exome sequencing identifies rare *LDLR* and *APOA5* alleles conferring risk for myocardial infarction. *Nature* **518**, 102–106 (2015). doi: 10.1038/nature13917; pmid: 25487149
29. Materials and methods are available as supplementary materials.
30. C. Sudlow *et al*., UK Biobank: An open access resource for identifying the causes of a wide range of complex diseases of middle and old age. *PLOS Med.* **12**, e1001779 (2015). doi: 10.1371/journal.pmed.1001779; pmid: 25826379
31. D. J. Carey *et al*., The Geisinger MyCode community health initiative: An electronic health record-linked biobank for precision medicine research. *Genet. Med.* **18**, 906–913 (2016). doi: 10.1038/gim.2015.187; pmid: 26866580
32. R. Tapia-Conyer *et al*., Cohort profile: The Mexico City Prospective Study. *Int. J. Epidemiol.* **35**, 243–249 (2006). doi: 10.1093/ije/dyl042; pmid: 16556648
33. C. A. Emdin *et al*., Analysis of predicted loss-of-function variants in UK Biobank identifies variants protective for disease. *Nat. Commun.* **9**, 1613 (2018). doi: 10.1038/s41467-018-03911-8; pmid: 29691411
34. A. S. Hauser, M. M. Attwood, M. Rask-Andersen, H. B. Schiöth, D. E. Gloriam, Trends in GPCR drug discovery: New agents, targets and indications. *Nat. Rev. Drug Discov.* **16**, 829–842 (2017). doi: 10.1038/nrd.2017.178; pmid: 29075003
35. M. Vázquez-Moreno *et al*., The melanocortin 4 receptor p.Ile269Asn mutation is associated with childhood and adult obesity in Mexicans. *J. Clin. Endocrinol. Metab.* **105**, e1468–e1477 (2020). doi: 10.1210/clinem/dgz276; pmid: 31841602
36. I. S. Farooqi *et al*., Clinical spectrum of obesity and mutations in the melanocortin 4 receptor gene. *N. Engl. J. Med.* **348**, 1085–1095 (2003). doi: 10.1056/NEJMoa022050; pmid: 12646665
37. G. Binder *et al*., Health and quality of life in adults with Noonan syndrome. *J. Pediatr.* **161**, 501–505.e1 (2012). doi: 10.1016/j.jpeds.2012.02.043; pmid: 22494877
38. A. P. Morris, Transethnic meta-analysis of genomewide association studies. *Genet. Epidemiol.* **35**, 809–822 (2011). doi: 10.1002/gepi.20630; pmid: 22125221
39. M. V. Holmes *et al*., Association between alcohol and cardiovascular disease: Mendelian randomisation analysis based on individual participant data. *BMJ* **349**, g4164 (2014). doi: 10.1136/bmj.g4164; pmid: 25011450
40. M. Claussnitzer *et al*., FTO obesity variant circuitry and adipocyte browning in humans. *N. Engl. J. Med.* **373**, 895–907 (2015). doi: 10.1056/NEJMoa1502214; pmid: 26287746
41. D. Cook *et al*., Lessons learned from the fate of AstraZeneca's drug pipeline: A five-dimensional framework. *Nat. Rev. Drug Discov.* **13**, 419–431 (2014). doi: 10.1038/nrd4309; pmid: 24833294
42. M. R. Nelson *et al*., The support of human genetic evidence for approved drug indications. *Nat. Genet.* **47**, 856–860 (2015). doi: 10.1038/ng.3314; pmid: 26121088
43. N. O. Stitziel *et al*., ANGPTL3 deficiency and protection against coronary artery disease. *J. Am. Coll. Cardiol.* **69**, 2054–2063 (2017). doi: 10.1016/j.jacc.2017.02.030; pmid: 28385496
44. T. I. Pollin *et al*., A null mutation in human *APOC3* confers a favorable plasma lipid profile and apparent cardioprotection. *Science* **322**, 1702–1705 (2008). doi: 10.1126/science.1161524; pmid: 19074352
45. J. C. Cohen, E. Boerwinkle, T. H. Mosley Jr., H. H. Hobbs, Sequence variations in *PCSK9*, low LDL, and protection against coronary heart disease. *N. Engl. J. Med.* **354**, 1264–1272 (2006). doi: 10.1056/NEJMoa054013; pmid: 16554528
46. V. Garcia *et al*., 20-HETE signals through G-protein–coupled receptor GPR75 (G_q) to affect vascular function and trigger hypertension. *Circ. Res.* **120**, 1776–1788 (2017). doi: 10.1161/CIRCRESAHA.116.310525; pmid: 28325781
47. S. Dedoni, L. A. Campbell, B. K. Harvey, V. Avdoshina, I. Mocchetti, The orphan G-protein-coupled receptor 75 signaling is activated by the chemokine CCL5. *J. Neurochem.* **146**, 526–539 (2018). doi: 10.1111/jnc.14463; pmid: 29772059
48. J. J. Holst, M. M. Rosenkilde, Recent advances of GIP and future horizons. *Peptides* **125**, 170230 (2020). doi: 10.1016/j.peptides.2019.170230; pmid: 31838219
49. P. A. Hollander *et al*., Pramlintide as an adjunct to insulin therapy improves long-term glycemic and weight control in patients with type 2 diabetes: A 1-year randomized controlled trial. *Diabetes Care* **26**, 784–790 (2003). doi: 10.2337/diacare.26.3.784; pmid: 12610038
50. P. Hollander *et al*., Effect of pramlintide on weight in overweight and obese insulin-treated type 2 diabetes patients. *Obes. Res.* **12**, 661–668 (2004). doi: 10.1038/oby.2004.76; pmid: 15090634
51. O. Schmitz, B. Brock, J. Rungby, Amylin agonists: A novel approach in the treatment of diabetes. *Diabetes* **53** (suppl. 3), S233–S238 (2004). doi: 10.2337/diabetes.53.suppl_3.S233; pmid: 15561917
52. W. Cheng *et al*., Calcitonin receptor neurons in the mouse nucleus tractus solitarius control energy balance via the non-aversive suppression of feeding. *Cell Metab.* **31**, 301–312.e5 (2020). doi: 10.1016/j.cmet.2019.12.012; pmid: 31955990
53. B. Antolin-Fontes *et al*., The habenular G-protein–coupled receptor 151 regulates synaptic plasticity and nicotine intake. *Proc. Natl. Acad. Sci. U.S.A.* **117**, 5502–5509 (2020). doi: 10.1073/pnas.1916132117; pmid: 32098843
54. P. R. Loh, G. Kichaev, S. Gazal, A. P. Schoech, A. L. Price, Mixed-model association for biobank-scale datasets. *Nat. Genet.* **50**, 906–908 (2018). doi: 10.1038/s41588-018-0144-6; pmid: 29892013
55. J. Mbatchou *et al*., Computationally efficient whole-genome regression for quantitative and binary traits. *Nat. Genet.* 10.1038/s41588-021-00870-7 (2021). doi: 10.1038/s41588-021-00870-7; pmid: 34017140
56. D. M. Valenzuela *et al*., High-throughput engineering of the mouse genome coupled with high-resolution expression analysis. *Nat. Biotechnol.* **21**, 652–659 (2003). doi: 10.1038/nbt822; pmid: 12730667
57. W. T. Poueymirou *et al*., F0 generation mice fully derived from gene-targeted embryonic stem cells allowing immediate phenotypic analyses. *Nat. Biotechnol.* **25**, 91–99 (2007). doi: 10.1038/nbt1263; pmid: 17187059

ACKNOWLEDGMENTS

The authors thank J. C. Cohen, University of Texas Southwestern, Dallas (USA). **Funding:** This research received funding from Regeneron Pharmaceuticals. New York Medical College investigators received funding from the National Institutes of Health grants HL034300 (M.L.S.) and HL139793 (M.L.S.) and the diversity supplement HL139793-1S (V.G.). The Mexico City Prospective Study is funded by a core grant from the UK Medical Research Council to the MRC Population Health Research Unit at the University of Oxford and has previously received funding from the Mexican Health Ministry, the Mexican National Council of Science and Technology, the Wellcome Trust, the British Heart Foundation, Cancer Research UK, and the Nuffield Department of Population Health at the University of Oxford. Y.I.C. and J.I.R. were supported by the National Center for Advancing Translational Sciences, CTSI grant UL1TR001881, and the National Institute of Diabetes and Digestive and Kidney Disease Diabetes Research Center (DRC) grant DK063491 to the Southern California Diabetes Endocrinology Research Center. **Author contributions:** All authors contributed to the collection of data, interpretation of results, critical review of the manuscript for important intellectual content, and final approval of submission of the manuscript for publication. Project conceptualization: A.B., M.A.R.F., and L.A.L. Development or application of methods: P.A., A.G., O.S., J.A.K., L.K., Y.-Y.F., T.P., V.G., D.S., A.Li, J.Mb., A.E.L., C.B., N.V., N.L., S.H., K.A., J.V.P., E.D., M.D., S.B., A.H., M.J., B.Z., A.J.M., C.P., G.C., J.D.O., J.G.R., A.R.S., M.C., H.M.K., G.R.A., K.K., A.N.E., J.Ma., G.D.Y., M.W.S., J.A., G.D.G., M.L.S., A.B., M.A.R.F., and L.A.L. Paper drafting: P.A., O.S., and L.A.L. Data analysis: P.A., O.S., J.A.K., D.S., A.Li, J.Mb., A.E.L., C.B., N.V., N.L., S.B., A.H., H.M.K., G.R.A., J.Ma., M.A.R.F., and L.A.L. In vitro experiments design, analysis, and interpretation: L.K., Y.-Y.F., T.P., M.W.S., J.A., and G.D.G. Design, analysis, interpretation of mouse models and experiments: A.G., V.G., S.H., K.A., J.V.P., E.D., M.D., and M.L.S. Design, data and biological sample collection and organization, securing funding, implementation and coordination for the Mexico City Prospective Study; J.B.-C., P.K.-M., J.A., J.M.T., J.R.E., R.C., and R.T.-C. Design, data and biological sample collection and organization, securing funding, implementation and coordination for the Geisinger Health System study: C.D.S., T.M., and D.J.C. Design, data and biological sample collection and organization, securing funding, implementation and coordination for the Malmö Diet and Cancer study: O.M. Design, data and biological sample collection and organization, securing funding, implementation and coordination for the Taiwan metabochip consortium: Y.I.C. and J.I.R. Design, data and biological sample collection and organization, securing funding, implementation and coordination for the University of Pennsylvania Penn Medicine BioBank: D.J.R. Design, data and biological sample collection and organization, securing funding, implementation and coordination for the Duke Catheterization Genetics cohort: W.E.K. and S.H.S. **Competing interests:** Regeneron authors receive salary from and own options and/or stock of the company. J.R.E. reports grant funding from Boehringer Ingelheim to the University of Oxford for the EMPA-KIDNEY trial of empagliflozin in patients with chronic kidney disease (clinical trials registration number NCT03594110). R.C. has obtained grants to the University of Oxford for the independent design, conduct, analysis, and reporting of randomized trials of lipid modifying therapies from Merck & Co. and from Medco/Novartis. In addition, R.C. received a research prize from Pfizer. This work has been described in one or more pending provisional patent applications. R.C. is an inventor on a patent application, for which he has waived personal royalties in favor of the Nuffield Department of Public Health, held/submitted by University of Oxford that covers a statin-related myopathy genetic test. Emma Link, Rory Collins, Sarah Parish, and Mark Lathrop are inventors on patent application EP2857525A1 held/submitted by Oxford University Innovation Ltd. that covers the method of determining the susceptibility to statin-induced myopathy. M.A.R.F. and L.A.L. are inventors on a provisional patent application (no. 63/019,589) submitted by RGC relating to PCSK1 genetics. P.A., O.S., A.B., M.A.R.F., and L.A.L. are inventors on a provisional patent application (no. 63/042,327) submitted by RGC relating to GPR75 genetics. P.A., O.S., A.B., M.A.R.F., and L.A.L. are inventors on a provisional patent application (no. 63/066,182) submitted by RGC relating to CALCR genetics. G.D.Y. is the chief scientific officer and member of the board of directors at Regeneron Pharmaceuticals. A.J.M. is an executive officer of Regeneron Pharmaceuticals. **Data and materials availability:** The data supporting the findings of this manuscript are reported in the main text, in the figures, or in the supplementary materials and are tabulated in Table 1 and tables S1 to S21. UKB individual-level genotypic and phenotypic data are available to approved investigators via the UK Biobank study (www.ukbiobank.ac.uk/). Additional information about registration for access to the data are available at www.ukbiobank.ac.uk/register-apply/. Data access for approved applications requires a data transfer agreement between the researcher's institution and UK Biobank, the terms of which are available on the UK Biobank website (www.ukbiobank.ac.uk/media/ezrderzw/applicant-mta.pdf). GHS individual-level data are available to qualified academic noncommercial researchers through the portal at https://regeneron.envisionpharma.com/vt_regeneron/ under a data access agreement. MCPS individual-level data are available to qualified noncommercial researchers by emailing mcps-access@ndph.ox.ac.uk. The data access policy is outlined at www.ctsu.ox.ac.uk/files/mcps-data-access-policy-v2-1-english.pdf. Summary association statistics for GHS and MCPS are available as online supplemental files or at https://rgc-community.regeneron.com subject to a license. eQTL summary statistics were downloaded from the GTEx portal (https://gtexportal.org/). T2D Knowledge Portal summary statistics for associations with type 2 diabetes were accessed via the web portal (https://t2d.hugeamp.org/).

Downloaded from http://science.sciencemag.org/ on July 2, 2021

Regeneron Genetics Center
Goncalo Abecasis[1], Aris Baras[1], Michael Cantor[1], Giovanni Coppola[1], Andrew Deubler[1], Aris Economides[1], Katia Karalis[1], Luca A. Lotta[1], John D. Overton[1], Jeffrey G. Reid[1], Alan R. Shuldiner[1], Nilanjana Banerjee[1], Dadong Li[1], Deepika Sharma[1], Xiaodong Bai[1], Suganthi Balasubramanian[1], Andrew Blumenfeld[1], Gisu Eom[1], Lukas Habegger[1], Alicia Hawes[1], Shareef Khalid[1], Evan K. Maxwell[1], William Salerno[1], Jeffrey C. Staples[1], Josh Backman[1], Mathew Barber[1], Christian Benner[1], Shan Chen[1], Amy Damask[1], Lee Dobbyn[1], Manuel A. R. Ferreira[1], Arkopravo Ghosh[1], Lauren Gurski[1], Eric Jorgenson[1], Bindu Kalesan[1], Jack A. Kosmicki[1], Hyun Min Kang[1], Alexander Li[1], Nan Lin[1], Daren Liu[1], Adam E. Locke[1], Jonathan Marchini[1], Anthony Marcketta[1], Joelle Mbatchou[1], Arden Moscati[1], Colm O'Dushlaine[1], Charles Paulding[1], Jonathan Ross[1], Eli Stahl[1], Dylan Sun[1], Cristopher Van Hout[1], Kyoko Watanabe[1], Bin Ye[1], Andrey Ziyatdinov[1], Marcus B. Jones[1], Michelle G. LeBlanc[1], Jason A. Mighty[1], Lyndon J. Mitnaul[1], Ariane Ayer[1], Kavita Praveen[1]

DiscovEHR Collaboration
Lance J. Adams[2], Jackie Blank[2], Dale Bodian[2], Derek Boris[2], Adam Buchanan[2], David J. Carey[2], Ryan D. Colonie[2], F. Daniel Davis[2], Dustin N. Hartzel[2], Melissa Kelly[2], H. Lester Kirchner[2], Joseph B. Leader[2], David H. Ledbetter[2], J. Neil Manus[2], Christa L. Martin[2], Michelle Meyer[2], Tooraj Mirshahi[2], Matthew Oetjens[2], Thomas Nate Person[2], Christopher D. Still[2], Natasha Strande[2], Amy Sturm[2], Jen Wagner[2], Marc Williams[2]

[1]Regeneron Genetics Center, Regeneron Pharmaceuticals Inc., Tarrytown, NY 10591, USA. [2]Geisinger Health System, Danville, PA 17882, USA.

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/373/6550/eabf8683/suppl/DC1
Materials and Methods
Figs. S1 to S9
Tables S1 to S21
References (*58*–*79*)
MDAR Reproducibility Checklist

View/request a protocol for this paper from *Bio-protocol*.

8 December 2020; accepted 17 May 2021
10.1126/science.abf8683

Downloaded from http://science.sciencemag.org/ on July 2, 2021

Science

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. Emberson, Rory 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, 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.abf8683

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, abf8683, this issue p. eabf8683; see also abh3556, p. 30

ARTICLE TOOLS http://science.sciencemag.org/content/373/6550/eabf8683

SUPPLEMENTARY MATERIALS http://science.sciencemag.org/content/suppl/2021/06/30/373.6550.eabf8683.DC1

RELATED CONTENT http://science.sciencemag.org/content/sci/373/6550/30.full

REFERENCES This article cites 79 articles, 11 of which you can access for free
http://science.sciencemag.org/content/373/6550/eabf8683#BIBL

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

Downloaded from http://science.sciencemag.org/ on July 2, 2021

PERMISSIONS http://www.sciencemag.org/help/reprints-and-permissions

Downloaded from http://science.sciencemag.org/ on July 2, 2021

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