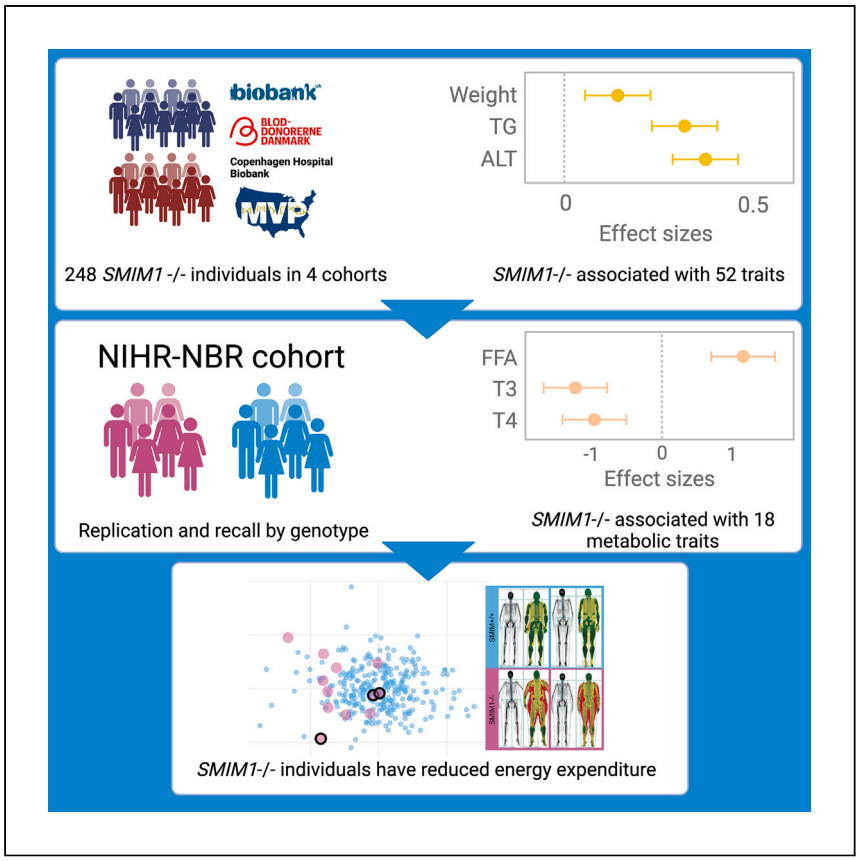
Med

CellPress
OPEN ACCESS

Case Report

SMIM1 absence is associated with reduced energy expenditure and excess weight



Stefanucci et al. show that a 17-bp loss-of-function variant in the open reading frame of *SMIM1* is associated with excess weight and other traits (dyslipidemia and insulin resistance) resembling metabolic syndrome. This is due to reduced energy expenditure, a major risk factor for obesity.

Luca Stefanucci, Camous Moslemi, Ana R. Tomé, ..., Krishna Chatterjee, Dragana Vuckovic, Mattia Frontini

m.frontini@exeter.ac.uk

Highlights

SMIM1$^{-/-}$ individuals are heavier compared with their *SMIM1*$^{+/+}$ counterparts

SMIM1$^{-/-}$ individuals have dyslipidemia and a higher likelihood of being on statins

SMIM1$^{-/-}$ individuals exhibit traits that resemble metabolic syndrome onset

Reduced energy expenditure due to mild hypothyroidism could be the underlying cause

Translation to Patients

Stefanucci et al., Med 5, 1083–1095
September 13, 2024 © 2024 The Author(s).
Published by Elsevier Inc.
https://doi.org/10.1016/j.medj.2024.05.015

Med

CellPress
OPEN ACCESS

Case Report

SMIM1 absence is associated with reduced energy expenditure and excess weight

Luca Stefanucci,[1,2,3,36] Camous Moslemi,[4,36] Ana R. Tomé,[1,2,36] Samuel Virtue,[5] Guillaume Bidault,[6] Nicholas S. Gleadall,[1,2] Laura P.E. Watson,[7] Jing E. Kwa,[8] Frances Burden,[1,2] Samantha Farrow,[1,2] DBDS Genetic Consortium, Ji Chen,[9] MAGIC, Urmo Võsa,[10] Keith Burling,[11] Lindsay Walker,[1,2] John Ord,[1,2] Peter Barker,[11] James Warner,[5] Amy Frary,[12] Karola Renhstrom,[12] Sofie E. Ashford,[12] Jo Piper,[7] Gail Biggs,[7] Wendy N. Erber,[13] Gary J. Hoffman,[14] Nadia Schoenmakers,[5] Christian Erikstrup,[15,16] Klaus Rieneck,[17] Morten H. Dziegiel,[17,18] Henrik Ullum,[19] Vian Azzu,[5,20] Michele Vacca,[5,21,22] Hugo Javier Aparicio,[23] Qin Hui,[24,25] Kelly Cho,[26,27] Yan V. Sun,[24,25] Peter W. Wilson,[24,28] Omer A. Bayraktar,[8] Antonio Vidal-Puig,[6,29] Sisse R. Ostrowski,[17,18] William J. Astle,[1,2,3,30] Martin L. Olsson,[31,32] Jill R. Storry,[31,32] Ole B. Pedersen,[4,18] Willem H. Ouwehand,[1,2,33,34] Krishna Chatterjee,[5] Dragana Vuckovic,[35] and Mattia Frontini[1,2,3,9,37,*]

SUMMARY

Background: Obesity rates have nearly tripled in the past 50 years, and by 2030 more than 1 billion individuals worldwide are projected to be obese. This creates a significant economic strain due to the associated non-communicable diseases. The root cause is an energy expenditure imbalance, owing to an interplay of lifestyle, environmental, and genetic factors. Obesity has a polygenic genetic architecture; however, single genetic variants with large effect size are etiological in a minority of cases. These variants allowed the discovery of novel genes and biology relevant to weight regulation and ultimately led to the development of novel specific treatments.

Methods: We used a case-control approach to determine metabolic differences between individuals homozygous for a loss-of-function genetic variant in the small integral membrane protein 1 (*SMIM1*) and the general population, leveraging data from five cohorts. Metabolic characterization of *SMIM1*$^{-/-}$ individuals was performed using plasma biochemistry, calorimetric chamber, and DXA scan.

Findings: We found that individuals homozygous for a loss-of-function genetic variant in *SMIM1* gene, underlying the blood group Vel, display excess body weight, dyslipidemia, altered leptin to adiponectin ratio, increased liver enzymes, and lower thyroid hormone levels. This was accompanied by a reduction in resting energy expenditure.

Conclusion: This research identified a novel genetic predisposition to being overweight or obese. It highlights the need to investigate the genetic causes of obesity to select the most appropriate treatment given the large cost disparity between them.

Funding: This work was funded by the National Institute of Health Research, British Heart Foundation, and NHS Blood and Transplant.

CONTEXT AND SIGNIFICANCE

Stefanucci et al. investigate the effects on human health of a 17-nucleotide deletion in the *SMIM1* gene, known so far for encoding the antigen of the Vel blood group. In the article, the authors show, using large population cohorts, such as the UK Biobank, that the absence of *SMIM1* is linked to excess weight, dyslipidemia, and related traits. The metabolic characterization of *SMIM1*$^{-/-}$ individuals points toward reduced energy expenditure as the main contributor to weight gain. Further research to understand the molecular mechanisms behind *SMIM1*-associated phenotypes might unlock new treatments for obesity and related conditions. This work highlights the importance of genetic and molecular insights in tackling metabolic disorders and improving public health.

INTRODUCTION

In 2013, we described a 17-bp deletion in *SMIM1* (rs566629828) that, in homozygosity, results in the absence of this protein from all tissues (hereafter *SMIM1*$^{-/-}$) and

underlies Vel-negative blood group (Figure S1).[1,2] Over the last decade, this variant has been associated with several blood traits[3] and, because other blood groups have been shown to be associated with different pathologies,[4–6] we explored the impact of the absence of *SMIM1* on human health leveraging meticulously characterized phenotypic population biobanks. We analyzed UK Biobank (UKB) data to determine if the loss-of-function (LoF) variant in *SMIM1* was associated with any additional traits other than the known blood ones. In the 488,376 participants, we identified 104 individuals with $SMIM1^{-/-}$ genotype, 90 being unrelated and of European ancestry[7] (46 females and 44 males; STAR Methods, Figure S2, and Table S1), corroborating the previously reported minor allele frequency (MAF = 0.0147) for rs566629828 deletion in this ancestry[1] and thus estimating the number of $SMIM1^{-/-}$ individuals at around 200,000 worldwide. The 17-bp deletion is in high linkage disequilibrium (D′ 0.98) with the major (A) allele of rs1175550 (MAF = 0.78), indicating that the deletion arose on the A allele. rs1175550 is a strong sentinel expression quantitative trait locus (eQTL) for *SMIM1* in the blood[8] (www.eqtlgen.org) and associated with red cell traits independently of rs566629828 (Figure S1).

RESULTS

We found that $SMIM1^{-/-}$ participants have excess weight (linear regression, Figures 1A and S2; Table S1). This analysis indicated an autosomal recessive effect; therefore, we considered only $SMIM1^{+/+}$ and $SMIM1^{-/-}$ individuals for the subsequent analyses. $SMIM1^{-/-}$ showed further association with body mass index (BMI) ($\widehat{\beta}$ = 0.27 FDR = 2.79e−2), waist circumference ($\widehat{\beta}$ = 0.27 FDR = 9.92e−3), and both arms fat mass (left: $\widehat{\beta}$ = 0.26 FDR = 3.18e−2; right: $\widehat{\beta}$ = 0.26 FDR = 3.5e−2; Figures 1B and 1C; Table S1). For weight, these differences equate to an average extra 4.6 kg in females and 2.4 kg in males (Table S1). In the UKB cohort, 26 out of the 90 $SMIM1^{-/-}$ individuals (28.8%; 15 females and 11 males), have a BMI > 30 kg/m^2, a higher percentage than the rest of the cohort (Fisher's exact test odds ratio [OR] = 1.27; p = 1.8e−1). Analysis of UKB plasma biochemistry assay results showed that $SMIM1^{-/-}$ participants had greater levels of triglycerides ($\widehat{\beta}$ = 0.3 FDR = 1.07e−2; Figures 1B and 1C). Furthermore, they exhibited greater average levels of liver enzymes with $\widehat{\beta}$ for alanine and aspartate aminotransferase of 0.50 and 0.43, and for gamma-glutamyl transferase of 0.35 (FDR = 4.10e−06, 1.60e−04, and 2.49e−03, respectively), as well as increased urate levels ($\widehat{\beta}$ = 0.35, FDR = 3.54e−04; Figures 1B and 1C; Table S1). Adjusting for the effect of BMI, removed the associations with body composition features, indicating that the higher BMI was responsible for these associations. In contrast, the associations with triglycerides, liver enzymes, and urate levels were only attenuated (Figure 1C; Table S1), suggesting that these effects were not solely dependent on BMI. Interestingly, we also identified sex-specific effects. $SMIM1^{-/-}$ female UKB participants exhibit greater average fat-free mass in arms and legs (right arm $\widehat{\beta}$ = 0.39, FDR = 2.39e−02; right leg $\widehat{\beta}$ = 0.33, FDR = 6.01e−02; Table S1) and lower average sex hormone binding globulin (SHBG) levels ($\widehat{\beta}$ = −0.41, FDR = 2.93e−2). Additional sex-specific differences were noted and are presented in Table S1. Importantly, (1) none of the above associations were detected in carriers of the 17-bp deletion in *SMIM1*, (2) none of the above associations were detected for the common eQTL variant rs1175550, suggesting that the metabolic differences were unlikely to be mediated by rs1175550-associated variation in the expression of *SMIM1* in red cells,[8] (3) even when we observed differences between $SMIM1^{+/+}$ and $SMIM1^{-/-}$ individuals, the mean values for the two groups were within the normal ranges for each measurement, and (4) no association was found between $SMIM1^{-/-}$ and fasting glucose

[1]Department of Haematology, University of Cambridge, Cambridge Biomedical Campus, Cambridge, UK

[2]National Health Service (NHS) Blood and Transplant, Cambridge Biomedical Campus, Cambridge, UK

[3]British Heart Foundation, Cambridge Centre for Research Excellence, University of Cambridge, Cambridge Biomedical Campus, Cambridge, UK

[4]Department of Clinical Immunology, Zealand University Hospital (Roskilde University), Køge, Denmark

[5]Wellcome-MRC Institute of Metabolic Science, University of Cambridge, Cambridge, UK

[6]University of Cambridge Metabolic Research Laboratories, Institute of Metabolic Science, MDU MRC, Addenbrooke's Hospital, Cambridge, UK

[7]NIHR Cambridge Clinical Research Facility, Cambridge University Hospitals, Cambridge Biomedical Campus, Cambridge, UK

[8]Wellcome Sanger Institute, Wellcome Genome Campus, Hinxton, Cambridge, UK

[9]Department of Clinical and Biomedical Sciences, University of Exeter Medical School, Faculty of Health and Life Sciences RILD Building, Barrack Road, Exeter, UK

[10]Estonian Genome Centre, Institute of Genomics, University of Tartu, Tartu, Estonia

[11]NIHR Cambridge Biomedical Research Centre Core Biochemical Assay Laboratory, Cambridge University Hospitals NHS Foundation Trust, Cambridge, UK

[12]NIHR National BioResource, Cambridge University Hospitals NHS Foundation, Cambridge Biomedical Campus, Cambridge, UK

[13]Discipline of Pathology and Laboratory Science, School of Biomedical Sciences, The University of Western Australia, Perth, WA, Australia

[14]Discipline of Pathology and Laboratory Medicine, Medical School, The University of Western Australia, Perth, WA, Australia

[15]Department of Clinical Immunology, Aarhus University Hospital, Aarhus University, Aarhus, Denmark

[16]Department of Clinical Medicine, Aarhus University, Aarhus, Denmark

[17]Department of Clinical Immunology, Rigshospitalet, University of Copenhagen, Copenhagen, Denmark

[18]Department of Clinical Medicine, University of Copenhagen, Copenhagen, Denmark

[19]Statens Serum Institut, Copenhagen, Denmark

[20]Department of Gastroenterology, Norfolk & Norwich University Hospitals NHS Foundation Trust, Norwich, UK

[21]Interdisciplinary Department of Medicine, Università degli Studi di Bari "Aldo Moro", Bari, Italy

[22]Roger Williams Institute of Hepatology, London, UK

Continued

levels in the Meta-Analyses of Glucose and Insulin-related traits Consortium (MAGIC) results.[9]

To further investigate these findings, we made home visits to obtain blood samples and health data from 25 British *SMIM1*$^{-/-}$ individuals (12 females, 13 males; not UKB participants; Figure S2) for an extensive survey of metabolism-relevant analytes, with results being compared with 180 individuals (100 females, 80 males) who carried at least one reference allele for variant rs566629828; both groups were members of the National Institute for Health and Care Research BioResource (NIHR-NBR) (all values in Table S2). We observed the same trend for *SMIM1*$^{-/-}$ individuals to be heavier but, possibly due to the small sample size, the significance threshold was not reached (Table S3). We replicated the associations between *SMIM1*$^{-/-}$ for increased average levels for alanine aminotransferase and aspartate transaminase, with the same order of magnitude as observed in UKB (Figure 2; Table S3). We also found associations between *SMIM1*$^{-/-}$ and increased leptin to adiponectin ratio (LAR) ($\widehat{\beta}$ = 0.53, FDR = 2.58e−02), and an increase in free fatty acids ($\widehat{\beta}$ = 1.18, FDR = 1.43e−06), two indices of increased fat mass and insulin resistance (Figures 2A and 2B).[10,11] LAR (a marker for obesity and metabolic state[12,13]) increase was determined by an increase in leptin ($\widehat{\beta}$ = 0.38) and a reduction in adiponectin ($\widehat{\beta}$ = −0.37), albeit with p values slightly above the defined significance level (0.06 and 0.09, respectively).

Moreover, we found that *SMIM1*$^{-/-}$ individuals have lower average levels of total triiodothyronine (T3) and thyroxine (T4) (T3: $\widehat{\beta}$ = −0.86, FDR = 9.87e−04; T4: $\widehat{\beta}$ = –0.74, FDR = 2.84e−03; Figures 2A and 2B), whereas the levels of thyroid-stimulating hormone (TSH), albeit skewed toward the bottom of the normal distribution, were not different (Table S3).

The above findings prompted us to invite 12 *SMIM1*$^{-/-}$ individuals belonging to the NIHR-NBR cohort for a 2-day metabolic assessment (Figure S2). We measured the effect of the absence of *SMIM1* on resting energy expenditure (REE) (a marker of whole-body metabolic activity) by indirect calorimetry and body mass composition by dual-energy X-ray absorptiometry (DXA), using a well-established protocol (STAR Methods).[14] These studies showed that *SMIM1*$^{-/-}$ individuals had a lower REE adjusted for lean mass (Figures 2C and x axis; Wilcoxon rank-sum test; p = 2.16e−04, Table S4), while there were no differences in average lean mass compared with 310 unselected controls (Table S4). Average free T3, but not free T4, measurements were lower in the 12 *SMIM1*$^{-/-}$ than in the control group (Tables S2 and S3). Lower circulating total thyroid hormones in SMIM1$^{-/-}$ individuals are not due to reduction in thyroid hormone binding globulin and thyroglobulin levels are not elevated, making thyroid dyshormonogenesis an unlikely cause of their altered thyroid status (Table S2). The anthropometric differences observed in the 90 *SMIM1*$^{-/-}$ UKB participants were reflected in abnormal body composition visualized by DXA scans (Figure 2D). Because of the effect on REE, T3, and T4 levels, we explored the possible involvement of *SMIM1* in the hypothalamic-pituitary-thyroid axis.[15] To gain insight into the possible molecular mechanism(s), we analyzed the single-cell RNA sequencing data in studies that dissected the transcript levels of these tissues in multiple organisms. In the mouse hypothalamus[16] (GSE113576), *Smim1* was expressed at low levels in mature oligodendrocytes and some, but not all, inhibitory neurons (Figure S3A). Its expression was largely non-overlapping with that of the thyrotropin-releasing hormone (Figure S3B). In the human anterior pituitary gland[17] (GSE142653), *SMIM1* was expressed in corticotropes, gonadotropes, and somatotropes (Figure S3C), while in human thyroid organoids and mouse thyroid[18] (GSE163818) low-level expression

[23]Department of Neurology, Boston University School of Medicine, Boston, MA, USA

[24]Atlanta VA Medical Center, Decatur, GA, USA

[25]Department of Epidemiology, Emory University Rollins School of Public Health, Atlanta, GA, USA

[26]Massachusetts Veterans Epidemiology Research and Information Center (MAVERIC), VA Boston Healthcare System, Boston, MA, USA

[27]Department of Medicine, Brigham and Women's Hospital, Harvard Medical School, Boston, MA, USA

[28]Emory University Schools of Medicine and Public Health, Atlanta, GA, USA

[29]Centro de Innvestigacion Principe Felipe, Valencia, Spain

[30]MRC Biostatistics Unit, East Forvie Building, Cambridge Biomedical Campus, University of Cambridge, Cambridge, UK

[31]Clinical Immunology and Transfusion Medicine, Office for Medical Services, Region Skåne, Lund, Sweden

[32]Department of Laboratory Medicine, Division of Hematology and Transfusion Medicine, Lund University, Lund, Sweden

[33]Department of Haematology, Cambridge University Hospitals NHS Trust, CB2 0QQ Cambridge, UK

[34]Department of Haematology, University College London Hospitals NHS Trust, NW1 2BU London, UK

[35]Department of Epidemiology and Biostatistics, School of Public Health, Faculty of Medicine, Imperial College London, London, UK

[36]These authors contributed equally

[37]Lead contact

*Correspondence: m.frontini@exeter.ac.uk

https://doi.org/10.1016/j.medj.2024.05.015

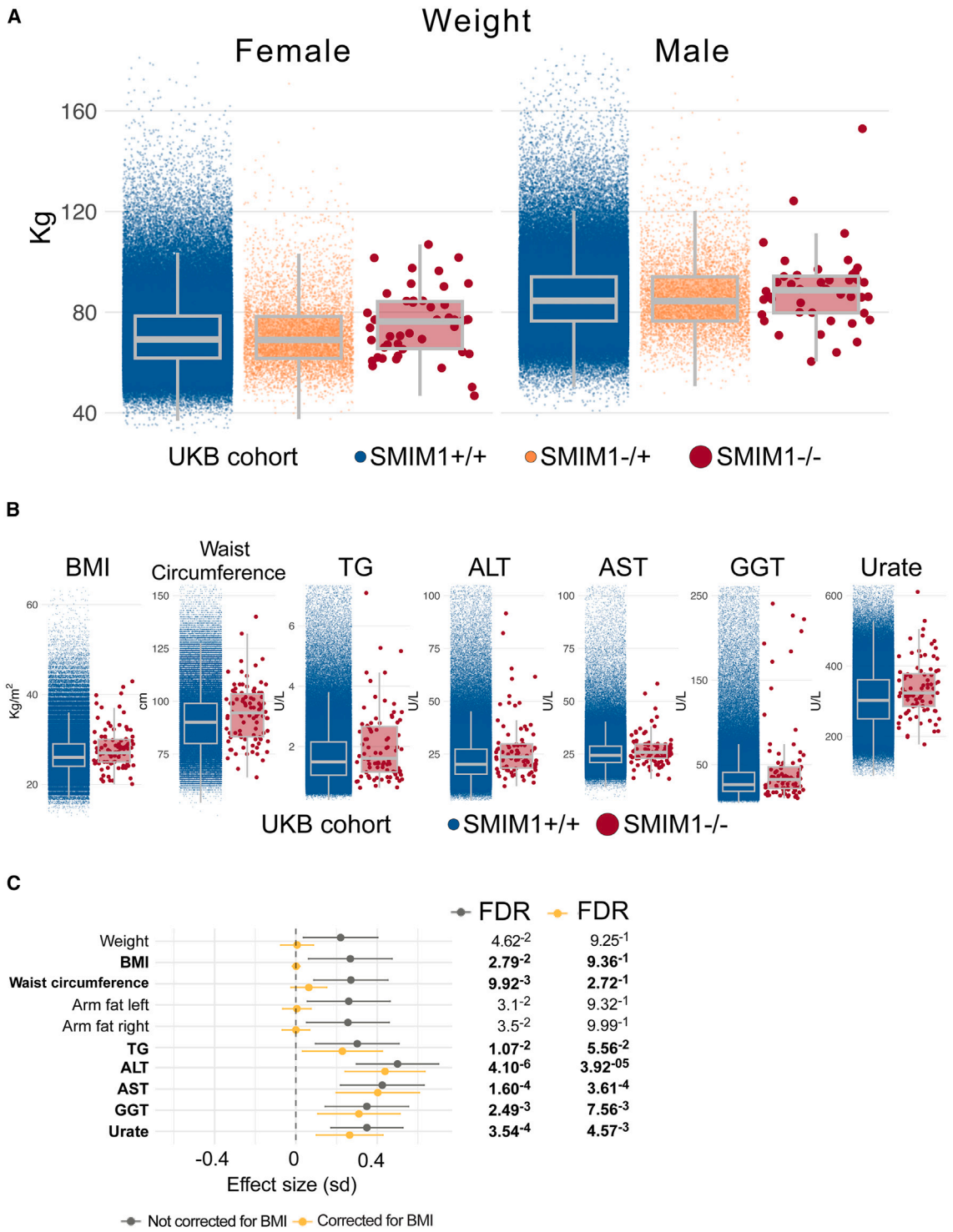


Figure 1. Differences between *SMIM1*$^{+/+}$ and *SMIM1*$^{-/-}$ individuals in the UKB cohort

(A) Boxplots for UKB participants' weight (kg) grouped according to their genotype. Sex-stratified data are shown for the three genotype groups, with females on the left and males on the right, respectively. Boxplot whiskers indicate the 95% confidence interval.

(B) Boxplots for BMI, waist circumference, and levels of triglycerides (TG), alanine aminotransferase (ALT), aspartate transaminase (AST), gamma-glutamyl transferase (GGT), and urate. Boxplot whiskers indicate the 95% confidence interval.

(C) Forest plot illustrating the effect size ($\widehat{\beta}$; percentage of standard deviation) of *SMIM1*$^{+/+}$ (blue) versus *SMIM1*$^{-/-}$ (red) for each trait. Bold characters highlight the measurements that are shown in (B). Effect sizes corrected for BMI are shown in yellow, and the non-corrected ones are in dark gray; β is represented by the dots and the 95% confidence intervals by the horizontal lines.

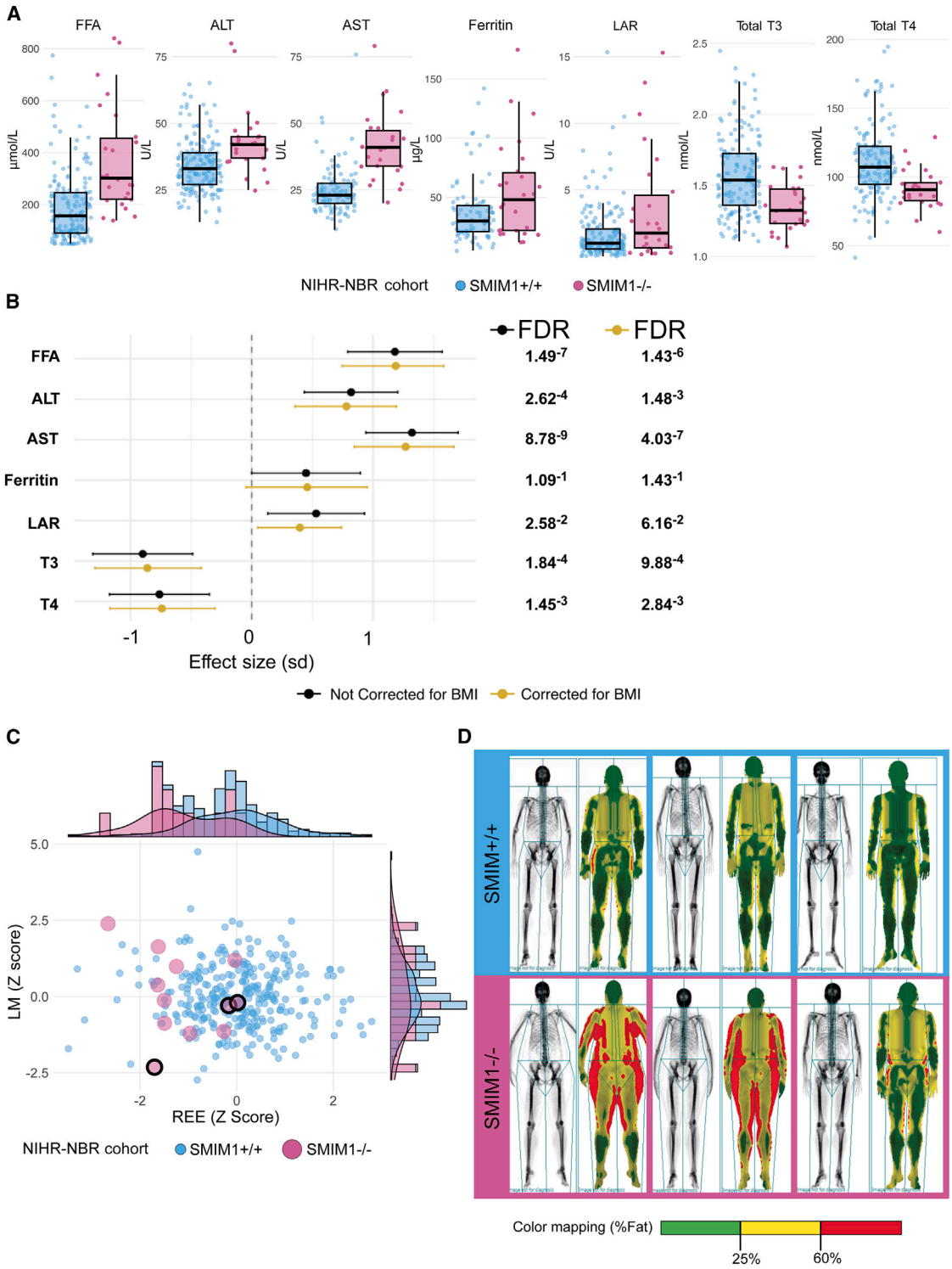


Figure 2. Differences between *SMIM1*$^{+/+}$ and *SMIM1*$^{-/-}$ individuals in the NIHR-NBR cohort and DXA body scan

(A) Boxplots for free fatty acids (FFA), alanine aminotransferase (ALT), aspartate transaminase (AST), ferritin, leptin to adiponectin ratio (LAR), total triiodothyronine (T3), and total thyroxine (T4). Boxplot whiskers indicate the 95% confidence interval.

(B) Forest plot illustrating the effect size ($\widehat{\beta}$; percentage of standard deviation) of *SMIM1*$^{+/+}$ versus *SMIM1*$^{-/-}$ for each trait. Effect sizes corrected for BMI and non-corrected ones are in yellow and dark gray, respectively. β is represented by the dots and the 95% confidence intervals by the horizontal lines.

Figure 2. *Continued*

(C) Scatterplot of *Z* scores for resting energy expenditure (REE) (x axis) and lean mass (LM) (y axis). *SMIM1*$^{+}$ individuals, light blue; *SMIM1*$^{-/-}$ individuals, pink. The three *SMIM1*$^{-/-}$ individuals shown in (D) are indicated by the pink dots with a black circumference.

(D) Representative DXA scans showing fat volume and distribution in three *SMIM1*$^{+}$ participants from the control group (top row, light blue borders) and three participants from the *SMIM1*$^{-/-}$ group (bottom row, pink borders).

was detected mainly in thyrocytes and in, as yet, uncharacterized Flt1-positive cells (Figure S3C). These analyses indicate that *SMIM1* could play one or more roles in the hypothalamic-pituitary-thyroid axis.

The associations between the genotype at rs566629828 and phenotypes observed in the UKB and NIHR-NBR cohort were orthogonally validated in 73 Danish *SMIM1*$^{-/-}$ individuals from the Danish Blood Donor Study[19] (DBDS) (blood donor, 25 females, 18 males, and 645 controls), and the Copenhagen Hospital Biobank[20] (CHB) (hospitalized or outpatients, 12 females,18 males, and 450 matched controls). Weight data, available only for the DBDS participants, showed, upon bootstrapping analysis (STAR Methods; controls matched by age, sex, and smoking status), consistent directionality for female *SMIM1*$^{-/-}$ individuals. However, the low number of Danish *SMIM1*$^{-/-}$ individuals and a less evident effect on weight in males (Figure 1B) limited the statistical power to detect differences. A meta-analysis combination of *SMIM1*$^{-/-}$ individuals from the Million Veteran Program (MVP) and the cohorts described above yielded the same directionality of effect (Figure S4). Interestingly, 20 of the 73 (27%) *SMIM1*$^{-/-}$ individuals in the Danish cohorts were diagnosed with lipoprotein metabolism disorders versus 13% in the controls (OR = 4.07, FDR = 2.09e−04). A review of all prescriptions in both cohorts showed greater use of statins in individuals lacking *SMIM1* versus controls (OR = 2.36, FDR = 2.22e−02; Table S5), indicating a higher number of individuals considered predisposed to cardiovascular events. An exploratory analysis of hospital episode statistics revealed an increased risk for cerebral events, with 5 cerebral bleeds and 5 thrombotic strokes in the 65 *SMIM1*$^{-/-}$ UKB participants for whom hospital event statistics were available (OR = 5.53 and 3.46, FDR = 6.88e−04 and 2.32e−02, respectively; Table S6).

DISCUSSION

In summary, our analysis identified a novel autosomal recessive effect for a LoF deletion (i.e., rs566629828) in *SMIM1*, a protein of yet unknown function(s), up until now only known as the antigen underlying the Vel blood group. This variant is present in at least 200,000 individuals worldwide (1 in 5,000 individuals in Great Britain and higher frequency in the Scandinavian countries [this article and Storry et al.[2]] and extremely low frequency in other ancestries). We have shown that *SMIM1*$^{-/-}$ individuals (Vel-negative blood group) exhibit a combination of metabolic features, including excess fat mass, inflammation, altered liver function, triglycerides, and altered lipoprotein metabolism. These features are due, at least in part, to reduced energy expenditure, a major risk factor in obesity.[21,22] In the most extreme cases, these effects could lead to an increased risk of insulin resistance and metabolic syndrome onset, accompanied by an increased susceptibility to cardiovascular disease, as supported by drug prescription and electronic hospital records analyses. These indicated that *SMIM1*$^{-/-}$ individuals have a higher likelihood of being prescribed statins and may be more prone to cerebral bleeds and thrombotic stroke. The effect on fat mass, and associated traits, is likely secondary, as suggested by its dependence on BMI, while we foresee a direct effect of *SMIM1* on dyslipidemia and liver function as these parameters continue to hold significance even after BMI correction. The associated phenotypes also show sexual dimorphism in their presentation. While the weight differences between the sexes likely reflect the different

distribution of this tissue between males and females,[23] association with traits such as urate and gamma-glutamyl transferase were driven by a stronger effect in one sex; whereas others, such as SHBG, were found only in females. Altogether, the observed metabolic phenotype, the increased risk for cardiovascular events, and the expression pattern of *SMIM1* are compatible with the fact that its absence results in a state of mild hypothyroidism.

The minor allele frequency of rs566629828 is at the lower end of common variations (MAF = 0.0147) and has one of the largest effects on weight ($\widehat{\beta}$ = 0.22) and BMI ($\widehat{\beta}$ = 0.27) reported so far, with the exception of extremely rare variants directly implicated in lipid metabolism.[24] For comparison, genetic variants in other well-characterized genes associated with obesity, i.e., *PCSK1* and *MC4R*, have comparable effect sizes in the general population. In particular, *MC4R* became a drug target for weight control in severe forms of genetically caused obesity.[24–26]

The rapidly growing amount of genomic data available, including blood donors typed by arrays,[27] means that more and more *SMIM1*$^{-/-}$ individuals will be identified as part of the incidental findings. Those who received a test early in life should be advised to monitor their energy intake, while individuals already overweight or obese, could be treated with a levothyroxine supplementation, an extremely cost-effective option compared with the most recent recommendations for the treatment of obesity.[28,29] While the effects have been replicated in independent cohorts, documenting central hypothyroidism with normal TSH, but low free thyroid hormone, measurements in a larger number of *SMIM1*$^{-/-}$ individuals will be necessary, before advocating a clinical trial of levothyroxine treatment as a potential therapeutic intervention.

Limitations of the study

The analyses we presented here are limited by the infrequency of *SMIM1*$^{-/-}$ individuals. To obtain sufficient statistical power, we had to use cohorts collected with different aims and with only a partial set of overlapping measurements. Differences in cohort composition and lifestyle might confound the observed effects on metabolism, and these biases could still influence the meta-analysis results. To limit confounding effects, controls were selected within the same cohort. This is exemplified in the meta-analysis (Figure S4), where only four *SMIM1*$^{-/-}$ individuals in the MVP cohort are female. We believe that the deep phenotype characterization of the recalled individuals and their consistency with the effect observed in the population helped to overcome this limitation. Future studies should investigate the function of this small transmembrane protein to identify the mechanisms by which it affects metabolism, as this could pave the way to novel therapeutic opportunities.[30–32]

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS
  - UK Biobank (UKB) cohort
  - NIHR-NBR cohort
  - Danish Blood Donor Study (DBDS) and Copenhagen Hospital biobank (CHB) cohorts

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.medj.2024.05.015.

CONSORTIA

The members of the DBDS Genetic Consortium are Karina Banasik, Jakob Bay, Jens Kjærgaard Boldsen, Thorsten Brodersen, Søren Brunak, Kristoffer Burgdorf, Mona Ameri Chalmer, Maria Didriksen, Khoa Manh Dinh, Joseph Dowsett, Christian Erikstrup, Bjarke Feenstra, Frank Geller, Daniel Gudbjartsson, Thomas Folkmann Hansen, Lotte Hindhede, Henrik Hjalgrim, Rikke Louise Jacobsen, Gregor Jemec, Bitten Aagaard Jensen, Katrine Kaspersen, Bertram Dalskov Kjerulff, Lisette Kogelman, Margit Anita Hørup Larsen, Ioannis Louloudis, Agnete Lundgaard, Susan, Christina Mikkelsen, Ioanna Nissen, Mette Nyegaard, Sisse Rye Ostrowski, Ole Birger Pedersen, Alexander Pil Henriksen, Palle Duun Rohde, Klaus Rostgaard, Michael Schwinn, Kari Stefansson, Hreinn Stefánsson, Erik Sørensen, Unnur Þorsteinsdóttir, Lise Wegner Thørner, Mie Topholm Bruun, Henrik Ullum, Thomas Werge, and David Westergaard.

The members of the MAGIC contributors are Ji Chen, Cassandra N. Spracklen, Gaëlle Marenne, Arushi Varshney, Laura J. Corbin, Jian'an Luan, Sara M. Willems, Ying Wu, Xiaoshuai Zhang, Momoko Horikoshi, Thibaud S. Boutin, Reedik Mägi, Johannes Waage, Ruifang Li-Gao, Kei Hang Katie Chan, Jie Yao, Mila D. Anasanti, Audrey Y. Chu, Annique Claringbould, Jani Heikkinen, Jaeyoung Hong, Jouke-Jan Hottenga, Shaofeng Huo, Marika A. Kaakinen, Tin Louie, Winfried März, Hortensia Moreno-Macias, Anne Ndungu, Sarah C. Nelson, Ilja M. Nolte, Kari E. North, Chelsea K. Raulerson, Debashree Ray, Rebecca Rohde, Denis Rybin, Claudia Schurmann, Xueling Sim, Loz Southam, Isobel D. Stewart, Carol A. Wang, Yujie Wang, Peitao Wu, Weihua Zhang, Tarunveer S. Ahluwalia, Emil V.R. Appel, Lawrence F. Bielak, Jennifer A. Brody, Noël P. Burtt, Claudia P. Cabrera, Brian E. Cade, Jin Fang Chai, Xiaoran Chai, Li-Ching Chang, Chien-Hsiun Chen, Brian H. Chen, Kumaraswamy Naidu Chitrala, Yen-Feng Chiu, Hugoline G. de Haan, Graciela E. Delgado, Ayse Demirkan, Qing Duan, Jorgen Engmann, Segun A. Fatumo, Javier Gayán, Franco Giulianini, Jung Ho Gong, Stefan Gustafsson, Yang Hai, Fernando P. Hartwig, Jing He, Yoriko Heianza, Tao Huang, Alicia Huerta-Chagoya, Mi Yeong Hwang, Richard A. Jensen, Takahisa Kawaguchi, Katherine A. Kentistou, Young Jin Kim, Marcus E. Kleber, Ishminder K. Kooner, Shuiqing Lai, Leslie A. Lange, Carl D. Langefeld, Marie Lauzon, Man Li, Symen Ligthart, Jun Liu, Marie Loh, Jirong Long, Valeriya Lyssenko, Massimo Mangino, Carola Marzi, May E. Montasser, Abhishek Nag, Masahiro Nakatochi, Damia Noce, Raymond Noordam, Giorgio Pistis, Michael Preuss, Laura Raffield, Laura J. Rasmussen-Torvik, Stephen S. Rich, Neil R. Robertson, Rico Rueedi, Kathleen Ryan, Serena Sanna, Richa Saxena, Katharina E. Schraut, Bengt Sennblad, Kazuya Setoh, Albert V. Smith, Lorraine Southam, Thomas Sparsø, Rona J. Strawbridge, Fumihiko Takeuchi, Jingyi Tan, Stella Trompet, Erik van den Akker, Peter J. van der Most, Niek Verweij, Mandy Vogel, Heming Wang, Chaolong Wang, Nan Wang, Helen R. Warren, Wanqing Wen, Tom Wilsgaard, Andrew Wong, Andrew R. Wood, Tian Xie, Mohammad Hadi Zafarmand, Jing-Hua Zhao, Wei Zhao,

Najaf Amin, Zorayr Arzumanyan, Arne Astrup, Stephan J.L. Bakker, Damiano Baldassarre, Marian Beekman, Richard N. Bergman, Alain Bertoni, Matthias Blüher, Lori L. Bonnycastle, Stefan R. Bornstein, Donald W. Bowden, Qiuyin Cai, Archie Campbell, Harry Campbell, Yi Cheng Chang, Eco J.C. de Geus, Abbas Dehghan, Shufa Du, Gudny Eiriksdottir, Aliki Eleni Farmaki, Mattias Frånberg, Christian Fuchsberger, Yutang Gao, Anette P. Gjesing, Anuj Goel, Sohee Han, Catharina A. Hartman, Christian Herder, Andrew A. Hicks, Chang-Hsun Hsieh, Willa A. Hsueh, Sahoko Ichihara, Michiya Igase, M. Arfan Ikram, W. Craig Johnson, Marit E. Jørgensen, Peter K. Joshi, Rita R. Kalyani, Fouad R. Kandeel, Tomohiro Katsuya, Chiea Chuen Khor, Wieland Kiess, Ivana Kolcic, Teemu Kuulasmaa, Johanna Kuusisto, Kristi Läll, Kelvin Lam, Deborah A. Lawlor, Nanette R. Lee, Rozenn N. Lemaitre, Honglan Li, Shih-Yi Lin, Jaana Lindström, Allan Linneberg, Jianjun Liu, Carlos Lorenzo, Tatsuaki Matsubara, Fumihiko Matsuda, Geltrude Mingrone, Simon Mooijaart, Sanghoon Moon, Toru Nabika, Girish N. Nadkarni, Jerry L. Nadler, Mari Nelis, Matt J. Neville, Jill M. Norris, Yasumasa Ohyagi, Annette Peters, Patricia A. Peyser, Ozren Polasek, Qibin Qi, Dennis Raven, Dermot F. Reilly, Alex Reiner, Fernando Rivideneira, Kathryn Roll, Igor Rudan, Charumathi Sabanayagam, Kevin Sandow, Naveed Sattar, Annette Schürmann, Jinxiu Shi, Heather M. Stringham, Kent D. Taylor, Tanya M. Teslovich, Betina Thuesen, Paul R.H.J. Timmers, Elena Tremoli, Michael Y. Tsai, Andre Uitterlinden, Rob M. van Dam, Diana van Heemst, Astrid van Hylckama Vlieg, Jana V. Van Vliet-Ostaptchouk, Jagadish Vangipurapu, Henrik Vestergaard, Tao Wang, Ko Willems van Dijk, Tatijana Zemunik, Goncalo R. Abecasis, Linda S. Adair, Carlos Alberto Aguilar-Salinas, Marta E. Alarcón-Riquelme, Ping An, Larissa Aviles-Santa, Diane M. Becker, Lawrence J. Beilin, Sven Bergmann, Hans Bisgaard, Corri Black, Michael Boehnke, Eric Boerwinkle, Bernhard O. Böhm, Klaus Bønnelykke, D.I. Boomsma, Erwin P. Bottinger, Thomas A. Buchanan, Mickaël Canouil, Mark J. Caulfield, John C. Chambers, Daniel I. Chasman, Yii-Der Ida Chen, Ching-Yu Cheng, Francis S. Collins, Adolfo Correa, Francesco Cucca, H. Janaka de Silva, George Dedoussis, Sölve Elmståhl, Michele K. Evans, Ele Ferrannini, Luigi Ferrucci, Jose C. Florez, Paul W. Franks, Timothy M. Frayling, Philippe Froguel, Bruna Gigante, Mark O. Goodarzi, Penny Gordon-Larsen, Harald Grallert, Niels Grarup, Sameline Grimsgaard, Leif Groop, Vilmundur Gudnason, Xiuqing Guo, Anders Hamsten, Torben Hansen, Caroline Hayward, Susan R. Heckbert, Bernardo L. Horta, Wei Huang, Erik Ingelsson, Pankow S. James, Marjo-Ritta Jarvelin, Jost B. Jonas, J. Wouter Jukema, Pontiano Kaleebu, Robert Kaplan, Sharon L.R. Kardia, Norihiro Kato, Sirkka M. Keinanen-Kiukaanniemi, Bong-Jo Kim, Mika Kivimaki, Heikki A. Koistinen, Jaspal S. Kooner, Antje Körner, Peter Kovacs, Diana Kuh, Meena Kumari, Zoltan Kutalik, Markku Laakso, Timo A. Lakka, Lenore J. Launer, Karin Leander, Huaixing Li, Xu Lin, Lars Lind, Cecilia Lindgren, Simin Liu, Ruth J.F. Loos, Patrik K.E. Magnusson, Anubha Mahajan, Andres Metspalu, Dennis O. Mook-Kanamori, Trevor A. Mori, Patricia B. Munroe, Inger Njølstad, Jeffrey R. O'Connell, Albertine J. Oldehinkel, Ken K. Ong, Sandosh Padmanabhan, Colin N.A. Palmer, Nicholette D. Palmer, Oluf Pedersen, Craig E. Pennell, David J. Porteous, Peter P. Pramstaller, Michael A. Province, Bruce M. Psaty, Lu Qi, Leslie J. Raffel, Rainer Rauramaa, Susan Redline, Paul M. Ridker, Frits R. Rosendaal, Timo E. Saaristo, Manjinder Sandhu, Jouko Saramies, Neil Schneiderman, Peter Schwarz, Laura J. Scott, Elizabeth Selvin, Peter Sever, Xiao-Ou Shu, P. Eline Slagboom, Kerrin S. Small, Blair H. Smith, Harold Snieder, Tamar Sofer, Thorkild I.A. Sørensen, Tim D. Spector, Alice Stanton, Claire J. Steves, Michael Stumvoll, Liang Sun, Yasuharu Tabara, E. Shyong Tai, Nicholas J. Timpson, Anke Tönjes, Jaakko Tuomilehto, Teresa Tusie, Matti Uusitupa, Pim van der Harst, Cornelia van Duijn, Veronique Vitart, Peter Vollenweider, Tanja G.M. Vrijkotte, Lynne E. Wagenknecht, Mark Walker, Ya X. Wang, Nick J. Wareham, Richard M. Watanabe, Hugh Watkins, Wen B. Wei, Ananda R. Wickremasinghe, Gonneke Willemsen, James F.

Wilson, Tien-Yin Wong, Jer-Yuarn Wu, Anny H. Xiang, Lisa R. Yanek, Loïc Yengo, Mitsuhiro Yokota, Eleftheria Zeggini, Wei Zheng, Alan B. Zonderman, Jerome I. Rotter, Anna L. Gloyn, Mark I. McCarthy, Josée Dupuis, James B. Meigs, Robert A. Scott, Inga Prokopenko, Aaron Leong, Ching-Ti Liu, Stephen C.J. Parker, Karen L. Mohlke, Claudia Langenberg, Eleanor Wheeler, Andrew P. Morris, and Inês Barroso.

The members of the Million Veteran Program contributors are Sumitra Muralidhar, Jennifer Moser, Jennifer E. Deen, J. Michael Gaziano, Sumitra Muralidhar, Jean Beckham, Kyong-Mi Chang, Philip S. Tsao, Shiuh-Wen Luoh, Juan P. Casas, Lori Churby, Stacey B. Whitbourne, Jessica V. Brewer, Mary T. Brophy, Luis E. Selva, Shahpoor (Alex) Shayan, Kelly Cho, Saiju Pyarajan, Philip S. Tsao, Scott L. DuVall, Todd Connor, Brady Stephens, Peter Wilson, Rachel McArdle, Louis Dellitalia, Kristin Mattocks, John Harley, Jeffrey Whittle, Frank Jacono, Jean Beckham, John Wells, Salvador Gutierrez, Kathrina Alexander, Kimberly Hammer, James Norton, Gerardo Villareal, Scott Kinlay, Junzhe Xu, Mark Hamner, Roy Mathew, Sujata Bhushan, Pran Iruvanti, Michael Godschalk, Zuhair Ballas, River Smith, Stephen Mastorides, Jonathan Moorman, Saib Gappy, Jon Klein, Nora Ratcliffe, Ana Palacio, Olaoluwa Okusaga, Maureen Murdoch, Peruvemba Sriram, Shing Shing Yeh, Neeraj Tandon, Darshana Jhala, Samuel Aguayo, David Cohen, Satish Sharma, Suthat Liangpunsakul, Kris Ann Oursler, Mary Whooley, Sunil Ahuja, Joseph Constans, Paul Meyer, Jennifer Greco, Michael Rauchman, Richard Servatius, Melinda Gaddy, Agnes Wallbom, Timothy Morgan, Todd Stapley, Peter Liang, Daryl Fujii, Philip Tsao, Patrick Strollo, Edward Boyko, Jessica Walsh, Samir Gupta, Mostaqul Huq, Joseph Fayad, Adriana Hung, Jack Lichy, Robin Hurley, Brooks Robey, and Prakash Balasubramanian.

ACKNOWLEDGMENTS

We would like to thank all participants in this study, the NIHR Cambridge Clinical Research Facility, the Wellcome Trust-MRC Institute of Metabolic Science Translational Research Facility, and the NIHR BioResource. We thank the National Institute for Health Research, NHS Blood and Transplant, and Health Data Research UK as part of the Digital Innovation Hub Program. We thank Professor Timothy M. Frayling for his helpful comments and suggestions. L.S. was a PhD student supported by the British Heart Foundation Cambridge Centre for Research Excellence (RE/18/1/34212). Landsteiner Foundation for Blood Transfusion Research supported A.R.T. during her PhD. S.V. and G.B. are supported by the British Heart Foundation (RG/18/7/33636). A.V.P. is supported by the British Heart Foundation (RG/18/7/33636) and the Medical Research Council (MRC_MC_UU_12012/5/RCUK). N.S. and K.C. are funded by the Wellcome Trust (219496/Z/19/Z and 210755/Z/18/Z). W.H.O. is an NIHR senior investigator and receives funding from the British Heart Foundation, International Society on Thrombosis and Haemostasis, MRC, NIHR, and Thermo Fisher Scientific. J.R.S. and M.L.O. are supported by the Swedish Research Council (2019-01683) and governmental university healthcare grants (ALF-4456521). M.L.O. is a Knut and Alice Wallenberg Clinical Scholar (2020.0234). D.V. is a member of the MRC Centre for Environment and Health, currently funded by the Medical Research Council (MR/S019669/1, 2019–2024). M.F. is supported by the British Heart Foundation (FS/18/53/33863) and the British Heart Foundation Cambridge Centre for Research Excellence (RE/18/1/34212). This study was supported by the National Institute for Health and Care Research Exeter Biomedical Research Centre. The views expressed are those of the author(s) and not necessarily those of the NIHR or the Department of Health and Social Care. We are grateful to all the MVP investigators; a list of MVP investigators can be found in the Consortia section. This research is supported by funding from the Department of Veterans Affairs Office

of Research and Development, Million Veteran Program Grant I01-BX003340 and I01-BX004821. This publication does not represent the views of the Department of Veterans Affairs or the United States Government. The graphical abstract was created using BioRender.com.

AUTHOR CONTRIBUTIONS

L.S. and C.M. analyzed data and wrote the manuscript. A.R.T., S.V., and G.B. collected samples and performed experimental work. N.S.G. analyzed data and reviewed the genotype calls. L.P.E.W. supervised clinical research facility data collection and analyzed data. J.E.K. analyzed single-cell datasets. F.B. and S.F. collected samples. J.C. analyzed the glucose level data. U.V. analyzed the eQTL data. K.B. and P.B. supervised the plasma biochemistry assays. J.W. performed histopathology. L.W., J.O., A.F., K. Renhstrom, S.E.A., J.P., G.B., and H.J.A. organized and performed the NIHR-NBR cohort recruitment and sample collection. W.N.E. and G.J.H. analyzed the pathology samples. N.S. provided advice for metabolic data interpretation. C.E., K. Rieneck, M.H.D., and H.U. supervised the establishment of the Danish cohort. L.P.E.W. analyzed the clinical research facility data. J.R.S. supervised the genotyping of the Danish cohorts. V.A. and M.V. provided advice with data interpretation. O.A.B. supervised single-cell RNA sequencing data analysis. A.V.-P. supervised metabolic data interpretation. S.R.O. supervised the Danish cohorts analysis. W.J.A. provided statistical analysis supervision. J.R.S. helped with the Danish cohort genotyping validation and advised on analysis. M.L.O., O.B.P., and W.H.O. provided supervision, planned analyses, contributed to data interpretation, and wrote the manuscript. K.C. supervised the clinical research facility metabolic characterization. D.V. provided statistical analysis supervision. M.F. was responsible for project organization, funds, supervision, and experimental and analysis planning and wrote the manuscript. M.F. and L.S. have unrestricted access to the data generated for this manuscript. All authors approved the final version of the manuscript.

DECLARATION OF INTERESTS

J.S. is the deputy CEO and 50% owner of BLUsang AB. He holds patents on Vel genotyping (inventors: Jill Storry, Magnus Jöud, Björn Nilsson, and Martin L. Olsson). J.S. has received speaker fees, royalties, and honoraria from the following companies: Grifols Diagnostic Solutions, QuidelOrtho Inc., and Biorad Laboratories. J.S. receives an honorarium for Section Editor work, Vox Sanguinis from John Wiley & Sons Ltd. J.S. is Vice President of the International Society of Blood Transfusion and married to Professor M.L.O. M.L.O. is CEO and 50% owner of BLUsang AB. M.L.O. holds patents on Vel genotyping (inventors: Jill Storry, Magnus Jöud, Björn Nilsson, and Martin L. Olsson). M.L.O. received speaker fees, royalties, and honoraria from the following companies: Grifols Diagnostic Solutions, QuidelOrtho Inc., and Biorad Laboratories. M.L.O. is married to Adjunct Professor J.S. W.N.E. is chair of the International Council for Standardization in Haematology. W.N.E. works as advisor for Scorpio Labs and is on the editorial board of the *Journal of Clinical Pathology*. W.H.O. is chair of the Blood Transfusion Genomics Consortium. W.H.O. is in receipt of an educational/research grant from Thermo Fisher Scientific. N.G. offers scientific consulting services to Thermo Fisher Scientific.

Received: August 7, 2023
Revised: December 6, 2023
Accepted: May 29, 2024
Published: June 20, 2024

REFERENCES

1. Cvejic, A., Haer-Wigman, L., Stephens, J.C., Kostadima, M., Smethurst, P.A., Frontini, M., van den Akker, E., Bertone, P., Bielczyk-Maczyńska, E., Farrow, S., et al. (2013). SMIM1 underlies the Vel blood group and influences red blood cell traits. Nat. Genet. *45*, 542–545.
2. Storry, J.R., Jöud, M., Christophersen, M.K., Thuresson, B., Åkerström, B., Sojka, B.N., Nilsson, B., and Olsson, M.L. (2013). Homozygosity for a null allele of SMIM1 defines the Vel-negative blood group phenotype. Nat. Genet. *45*, 537–541.
3. Vuckovic, D., Bao, E.L., Akbari, P., Lareau, C.A., Mousas, A., Jiang, T., Chen, M.-H., Raffield, L.M., Tardaguila, M., Huffman, J.E., et al. (2020). The Polygenic and Monogenic Basis of Blood Traits and Diseases. Cell *182*, 1214–1231.e11.
4. Borén, T., Falk, P., Roth, K.A., Larson, G., and Normark, S. (1993). Attachment of Helicobacter pylori to human gastric epithelium mediated by blood group antigens. Science *262*, 1892–1895.
5. Palmer, C.G.S., Turunen, J.A., Sinsheimer, J.S., Minassian, S., Paunio, T., Lönnqvist, J., Peltonen, L., and Woodward, J.A. (2002). RHD maternal-fetal genotype incompatibility increases schizophrenia susceptibility. Am. J. Hum. Genet. *71*, 1312–1319.
6. Goumidi, L., Thibord, F., Wiggins, K.L., Li-Gao, R., Brown, M.R., van Hylckama Vlieg, A., Souto, J.-C., Soria, J.-M., Ibrahim-Kosta, M., Saut, N., et al. (2021). Association between ABO haplotypes and the risk of venous thrombosis: impact on disease risk estimation. Blood *137*, 2394–2402.
7. Bycroft, C., Freeman, C., Petkova, D., Band, G., Elliott, L.T., Sharp, K., Motyer, A., Vukcevic, D., Delaneau, O., O'Connell, J., et al. (2018). The UK Biobank resource with deep phenotyping and genomic data. Nature *562*, 203–209.
8. Võsa, U., Claringbould, A., Westra, H.-J., Bonder, M.J., Deelen, P., Zeng, B., Kirsten, H., Saha, A., Kreuzhuber, R., Yazar, S., et al. (2021). Large-scale cis- and trans-eQTL analyses identify thousands of genetic loci and polygenic scores that regulate blood gene expression. Nat. Genet. *53*, 1300–1310.
9. Chen, J., Spracklen, C.N., Marenne, G., Varshney, A., Corbin, L.J., Luan, J. 'an, Willems, S.M., Wu, Y., Zhang, X., Horikoshi, M., et al. (2021). The trans-ancestral genomic architecture of glycemic traits. Nat. Genet. *53*, 840–860.
10. Finucane, F.M., Luan, J., Wareham, N.J., Sharp, S.J., O'Rahilly, S., Balkau, B., Flyvbjerg, A., Walker, M., Højlund, K., Nolan, J.J., et al. (2009). Correlation of the leptin:adiponectin ratio with measures of insulin resistance in non-diabetic individuals. Diabetologia *52*, 2345–2349.
11. Gastaldelli, A., Gaggini, M., and DeFronzo, R.A. (2017). Role of Adipose Tissue Insulin Resistance in the Natural History of Type 2 Diabetes: Results From the San Antonio Metabolism Study. Diabetes *66*, 815–822.
12. Oda, N., Imamura, S., Fujita, T., Uchida, Y., Inagaki, K., Kakizawa, H., Hayakawa, N., Suzuki, A., Takeda, J., Horikawa, Y., and Itoh, M. (2008). The ratio of leptin to adiponectin can be used as an index of insulin resistance. Metabolism *57*, 268–273.
13. Zhuo, Q., Wang, Z., Fu, P., Piao, J., Tian, Y., Xu, J., and Yang, X. (2009). Comparison of adiponectin, leptin and leptin to adiponectin ratio as diagnostic marker for metabolic syndrome in older adults of Chinese major cities. Diabetes Res. Clin. Pract. *84*, 27–33.
14. Watson, L.P.E., Raymond-Barker, P., Moran, C., Schoenmakers, N., Mitchell, C., Bluck, L., Chatterjee, V.K., Savage, D.B., and Murgatroyd, P.R. (2014). An approach to quantifying abnormalities in energy expenditure and lean mass in metabolic disease. Eur. J. Clin. Nutr. *68*, 234–240.
15. Fekete, C., and Lechan, R.M. (2014). Central regulation of hypothalamic-pituitary-thyroid axis under physiological and pathophysiological conditions. Endocr. Rev. *35*, 159–194.
16. Moffitt, J.R., Bambah-Mukku, D., Eichhorn, S.W., Vaughn, E., Shekhar, K., Perez, J.D., Rubinstein, N.D., Hao, J., Regev, A., Dulac, C., and Zhuang, X. (2018). Molecular, spatial, and functional single-cell profiling of the hypothalamic preoptic region. Science *362*, eaau5324. https://doi.org/10.1126/science.aau5324.
17. Zhang, S., Cui, Y., Ma, X., Yong, J., Yan, L., Yang, M., Ren, J., Tang, F., Wen, L., and Qiao, J. (2020). Single-cell transcriptomics identifies divergent developmental lineage trajectories during human pituitary development. Nat. Commun. *11*, 5275.
18. Romitti, M., Eski, S.E., Fonseca, B.F., Singh, S.P., and Costagliola, S. (2021). Single-cell trajectory inference guided enhancement of thyroid maturation in vitro using TGF-beta inhibition. Preprint at bioRxiv. https://doi.org/10.1101/2021.01.18.427103.
19. Hansen, T.F., Banasik, K., Erikstrup, C., Pedersen, O.B., Westergaard, D., Chmura, P.J., Nielsen, K., Thørner, L., Hjalgrim, H., Paarup, H., et al. (2019). DBDS Genomic Cohort, a prospective and comprehensive resource for integrative and temporal analysis of genetic, environmental and lifestyle factors affecting health of blood donors. BMJ Open *9*, e028401.
20. Sørensen, E., Christiansen, L., Wilkowski, B., Larsen, M.H., Burgdorf, K.S., Thørner, L.W., Nissen, J., Pedersen, O.B., Banasik, K., Brunak, S., et al. (2021). Data Resource Profile: The Copenhagen Hospital Biobank (CHB). Int. J. Epidemiol. *50*, 719–720e.
21. Roberts, S.B., Savage, J., Coward, W.A., Chew, B., and Lucas, A. (1988). Energy expenditure and intake in infants born to lean and overweight mothers. N. Engl. J. Med. *318*, 461–466.
22. Ravussin, E., Lillioja, S., Knowler, W.C., Christin, L., Freymond, D., Abbott, W.G., Boyce, V., Howard, B.V., and Bogardus, C. (1988). Reduced rate of energy expenditure as a risk factor for body-weight gain. N. Engl. J. Med. *318*, 467–472.
23. Srikanthan, P., Horwich, T.B., Calfon Press, M., Gornbein, J., and Watson, K.E. (2021). Sex Differences in the Association of Body Composition and Cardiovascular Mortality. J. Am. Heart Assoc. *10*, e017511.
24. Akbari, P., Gilani, A., Sosina, O., Kosmicki, J.A., Khrimian, L., Fang, Y.-Y., Persaud, T., Garcia, V., Sun, D., Li, A., et al. (2021). Sequencing of 640,000 exomes identifies GPR75 variants associated with protection from obesity. Science *373*, eabf8683. https://doi.org/10.1126/science.abf8683.
25. Farooqi, I.S., Keogh, J.M., Yeo, G.S.H., Lank, E.J., Cheetham, T., and O'Rahilly, S. (2003). Clinical spectrum of obesity and mutations in the melanocortin 4 receptor gene. N. Engl. J. Med. *348*, 1085–1095.
26. Clément, K., van den Akker, E., Argente, J., Bahm, A., Chung, W.K., Connors, H., De Waele, K., Farooqi, I.S., Gonneau-Lejeune, J., Gordon, G., et al. (2020). Efficacy and safety of setmelanotide, an MC4R agonist, in individuals with severe obesity due to LEPR or POMC deficiency: single-arm, open-label, multicentre, phase 3 trials. Lancet Diabetes Endocrinol. *8*, 960–970.
27. Gleadall, N.S., Veldhuisen, B., Gollub, J., Butterworth, A.S., Ord, J., Penkett, C.J., Timmer, T.C., Sauer, C.M., van der Bolt, N., Brown, C., et al. (2020). Development and validation of a universal blood donor genotyping platform: a multinational prospective study. Blood Adv. *4*, 3495–3506.
28. Mahase, E. (2022). NICE recommends weight loss drug semaglutide for obese adults. BMJ *376*, o340.
29. Prime Minister's Office (2023). New Drugs Pilot to Tackle Obesity and Cut NHS Waiting Lists (GOV.UK). https://www.gov.uk/government/news/new-drugs-pilot-to-tackle-obesity-and-cut-nhs-waiting-lists.
30. Nelson, M.R., Tipney, H., Painter, J.L., Shen, J., Nicoletti, P., Shen, Y., Floratos, A., Sham, P.C., Li, M.J., Wang, J., et al. (2015). The support of human genetic evidence for approved drug indications. Nat. Genet. *47*, 856–860.
31. King, E.A., Davis, J.W., and Degner, J.F. (2019). Are drug targets with genetic support twice as likely to be approved? Revised estimates of the impact of genetic support for drug mechanisms on the probability of drug approval. PLoS Genet. *15*, e1008489.
32. Ochoa, D., Karim, M., Ghoussaini, M., Hulcoop, D.G., McDonagh, E.M., and Dunham, I. (2022). Human genetics evidence supports two-thirds of the 2021 FDA-approved drugs. Nat. Rev. Drug Discov. *21*, 551.
33. Schwarzer, G., Carpenter, J.R., and Rücker, G. (2015). Meta-Analysis with R (Springer International Publishing).
34. Hao, Y., Hao, S., Andersen-Nissen, E., Mauck, W.M., 3rd, Zheng, S., Butler, A., Lee, M.J., Wilk, A.J., Darby, C., Zager, M., et al. (2021). Integrated analysis of multimodal single-cell data. Cell *184*, 3573–3587.e29.
35. Durinck, S., Moreau, Y., Kasprzyk, A., Davis, S., De Moor, B., Brazma, A., and Huber, W. (2005).

BioMart and Bioconductor: a powerful link between biological databases and microarray data analysis. Bioinformatics *21*, 3439–3440.

36. McCaw, Z.R., Lane, J.M., Saxena, R., Redline, S., and Lin, X. (2020). Operating characteristics of the rank-based inverse normal transformation for quantitative trait analysis in genome-wide association studies. Biometrics *76*, 1262–1272.

37. Viechtbauer, W. (2010). Conducting Meta-Analyses in R with the metafor Package. J. Stat. Softw. *36*, 1–48.

38. Downey, P., and Peakman, T.C. (2008). Design and implementation of a high-throughput biological sample processing facility using modern manufacturing principles. Int. J. Epidemiol. *37*, i46–i50.

39. Sudlow, C., Gallacher, J., Allen, N., Beral, V., Burton, P., Danesh, J., Downey, P., Elliott, P., Green, J., Landray, M., et al. (2015). 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.

40. Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M.A.R., Bender, D., Maller, J., Sklar, P., de Bakker, P.I.W., Daly, M.J., and Sham, P.C. (2007). PLINK: a tool set for whole-genome association and population-based linkage analyses. Am. J. Hum. Genet. *81*, 559–575.

41. Chen, L., Ge, B., Casale, F.P., Vasquez, L., Kwan, T., Garrido-Martín, D., Watt, S., Yan, Y., Kundu, K., Ecker, S., et al. (2016). Genetic Drivers of Epigenetic and Transcriptional Variation in Human Immune Cells. Cell *167*, 1398–1414.e24.

42. Di Angelantonio, E., Thompson, S.G., Kaptoge, S., Moore, C., Walker, M., Armitage, J., Ouwehand, W.H., Roberts, D.J., and Danesh, J.; INTERVAL Trial Group (2017). Efficiency and safety of varying the frequency of whole blood donation (INTERVAL): a randomised trial of 45 000 donors. Lancet *390*, 2360–2371.

43. Schmidt, M., Schmidt, S.A.J., Sandegaard, J.L., Ehrenstein, V., Pedersen, L., and Sørensen, H.T. (2015). The Danish National Patient Registry: a review of content, data quality, and research potential. Clin. Epidemiol. *7*, 449–490.

44. Gaziano, J.M., Concato, J., Brophy, M., Fiore, L., Pyarajan, S., Breeling, J., Whitbourne, S., Deen, J., Shannon, C., Humphries, D., et al. (2016). Million Veteran Program: A mega-biobank to study genetic influences on health and disease. J. Clin. Epidemiol. *70*, 214–223.

45. Hunter-Zinck, H., Shi, Y., Li, M., Gorman, B.R., Ji, S.-G., Sun, N., Webster, T., Liem, A., Hsieh, P., Devineni, P., et al. (2020). Genotyping Array Design and Data Quality Control in the Million Veteran Program. Am. J. Hum. Genet. *106*, 535–548.

46. Manichaikul, A., Mychaleckyj, J.C., Rich, S.S., Daly, K., Sale, M., and Chen, W.-M. (2010). Robust relationship inference in genome-wide association studies. Bioinformatics *26*, 2867–2873.

47. Fang, H., Hui, Q., Lynch, J., Honerlaw, J., Assimes, T.L., Huang, J., Vujkovic, M., Damrauer, S.M., Pyarajan, S., Gaziano, J.M., et al. (2019). Harmonizing Genetic Ancestry and Self-identified Race/Ethnicity in Genome-wide Association Studies. Am. J. Hum. Genet. *105*, 763–772.

48. Huang, J., Huffman, J.E., Huang, Y., Do Valle, Í., Assimes, T.L., Raghavan, S., Voight, B.F., Liu, C., Barabási, A.-L., Huang, R.D.L., et al. (2022). Genomics and phenomics of body mass index reveals a complex disease network. Nat. Commun. *13*, 7973.

49. Van Hout, C.V., Tachmazidou, I., Backman, J.D., Hoffman, J.D., Liu, D., Pandey, A.K., Gonzaga-Jauregui, C., Khalid, S., Ye, B., Banerjee, N., et al. (2020). Exome sequencing and characterization of 49,960 individuals in the UK Biobank. Nature *586*, 749–756.

50. Elia, M., and Livesey, G. (1992). Energy expenditure and fuel selection in biological systems: the theory and practice of calculations based on indirect calorimetry and tracer methods. World Rev. Nutr. Diet. *70*, 68–131.

51. Watson, L.P.E., Carr, K.S., Venables, M.C., Acerini, C.L., Lyons, G., Moran, C., Murgatroyd, P.R., and Chatterjee, K. (2019). Quantifying energy expenditure in childhood: utility in managing pediatric metabolic disorders. Am. J. Clin. Nutr. *110*, 1186–1191.

52. Watson, L.P.E., Venables, M.C., and Murgatroyd, P.R. (2017). An Investigation Into the Differences in Bone Density and Body Composition Measurements Between 2 GE Lunar Densitometers and Their Comparison to a 4-Component Model. J. Clin. Densitom. *20*, 498–506.

53. Cheung, L.Y.M., George, A.S., McGee, S.R., Daly, A.Z., Brinkmeier, M.L., Ellsworth, B.S., and Camper, S.A. (2018). Single-Cell RNA Sequencing Reveals Novel Markers of Male Pituitary Stem Cells and Hormone-Producing Cell Types. Endocrinology *159*, 3910–3924.

54. Ho, Y., Hu, P., Peel, M.T., Chen, S., Camara, P.G., Epstein, D.J., Wu, H., and Liebhaber, S.A. (2020). Single-cell transcriptomic analysis of adult mouse pituitary reveals sexual dimorphism and physiologic demand-induced cellular plasticity. Protein Cell *11*, 565–583.

55. Fletcher, P.A., Smiljanic, K., Maso Prévide, R., Iben, J.R., Li, T., Rokic, M.B., Sherman, A., Coon, S.L., and Stojilkovic, S.S. (2019). Cell Type- and Sex-Dependent Transcriptome Profiles of Rat Anterior Pituitary Cells. Front. Endocrinol. *10*, 623.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Biological samples		
Vel-negative blood samples	NIHR Clinical Research Facility	https://www.cambridgecrf.nihr.ac.uk/
NIHR-NBR cohort	This article	https://bioresource.nihr.ac.uk/
Critical commercial assays		
Atellica IM Free Thyroxine lite reagent	Siemens Healthineers	10995589
Immulite 2000 TBG assay	Siemens Healthineers	L2KTB2
Total cholesterol	Siemens Healthineers	DF 27
High-density lipoprotein	Siemens Healthineers	DF 48B
Aspartate aminotransferase	Siemens Healthineers	DF 41A
Alanine transaminase	Siemens Healthineers	DF 143
High-sensitivity C-reactive protein	Siemens Healthineers	RF 434
Free Fatty Acids	Sigma Aldrich (Roche)	11 383 175 001
Thyroid-stimulating hormone	DiaSorin	311211
Total T3	DiaSorin	311311
Total T4	DiaSorin	311411
Free T4	DiaSorin	311611
Ferritin	DiaSorin	313551
Leptin (Not a commercially available assay)	DELFIA®	https://www.perkinelmer.com/uk/lab-products-and-services/application-support-knowledgebase/delfia/delfia-products-catalog.html
Adiponectin (Not a commercially available assay)	DELFIA®	https://www.perkinelmer.com/uk/lab-products-and-services/application-support-knowledgebase/delfia/delfia-products-catalog.html
Deposited data		
NIHR-NBR cohort metabolites - raw data	This article, and Zenodo	https://zenodo.org/records/10685501
NIHR-NBR cohort resting energy expenditure - raw data	This article, and Zenodo	https://zenodo.org/records/10685501
UK Biobank	Downey et al.[34]	https://www.ukbiobank.ac.uk/
Million Veteran Program	Gaziano et al.[44]	https://www.research.va.gov/mvp/
Danish Blood Donor Study	Hansen et al.[19]	https://bloddonor.dk/
Copenhagen Hospital Biobank	Sørensen et al.[20]	https://www.regionh.dk/blodbanken/afdelingen/enheder-paa-rigshospitalet/Sider/biobank.aspx
Mouse hypothalamus	Moffitt et al.[16]	GEO: GSE113576
Human fetal pituitary	Zhang et al.[17]	GEO: GSE142653
Mouse pituitary	Cheung et al.[53]	GEO: GSE120410
Mouse pituitary	Ho et al.[54]	GEO: GSE146619
Rat pituitary	Fletcher et al.[55]	GEO: GSE132224
Mouse thyroid organoids	Romitti et al.[18]	GEO: GSE163818
Oligonucleotides		
Primer: *SMIM1* deletion Forward: ACAGCCTGGCCACCTGTCTTG Reverse: CTGCGGCAGCGTGAGGC	This article	N/A
Software and algorithms		
R 4.0.5	R Core Team	https://www.r-project.org/
Tidyverse 1.3.1	RStudio Team	https://tidyverse.tidyverse.org
meta 5.2–0	Schwarzer et al.[54]	N/A
Seurat (version 4.0.0)	Hao et al.[33]	https://satijalab.org/seurat/
biomaRt (version 2.46.3)	Durinck et al.[35]	https://bioconductor.org/packages/release/bioc/html/biomaRt.html
RNOmni	McCaw et al.[45]	https://cran.r-project.org/web/packages/RNOmni/index.html
metafor 3.0–2	Viechtbauer et al.[37]	https://github.com/wviechtb/metafor
GE Lunar iDXA - Encore version 18	GE Healthcare	210500GA

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
GE Lunar Prodigy - Encore version 16	GE Healthcare	N/A
Gas Exchange Measurement	GEMNutrition Ltd, Daresburysbury	17040–181
PLINK v1.9	PLINK working group	https://www.cog-genomics.org/plink/1.9/
Other		
Dimension EXL	Siemens Healthineers	https://www.siemens-healthineers.com/
Liaison XL	DiaSorin	https://int.diasorin.com/en

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources should be directed to and will be fulfilled by the lead contact, Mattia Frontini (m.frontini@exeter.ac.uk).

Materials availability

This study did not generate new unique reagents.

Data and code availability

Participants' phenotypes and *SMIM1* locus genotypes are accessible via the relevant cohort environments: UK Biobank (https://www.ukbiobank.ac.uk/) and MVP (https://www.mvp.va.gov/pwa/discover-mvp-data). Access to these cohorts requires an active project application. All the data generated for this study are available in an anonymized version in supplementary tables or in the Zenodo repository at: https://zenodo.org/records/10685501. The code used to analyze the cohorts is available at https://github.com/stefanucci-luca/vel_ko_analysis. Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS

A total of 248 *SMIM1*$^{-/-}$ unrelated European individuals (105 females, 143 males) from four different cohorts, described below were included in the study. For an overview of the cohorts, see Figure S2. Details about the recall by genotype study and the individual cohorts are reported below in the relevant sections.

UK Biobank (UKB) cohort

The UKB analyses have been conducted under application number 13745. This cohort consists of 502,682 participants, aged between 40 and 69 years of age on enrollment, recruited at 22 assessment centers across the UK between 2006 and 2010.[38,39] DNA samples were taken from participants and genotyped using the UK Biobank Axiom Array on the GeneTitan (Affymetrix, Santa Clara, CA). Genotype calling and quality control of the UKB dataset have been extensively documented elsewhere.[7]

The UKB Axiom Array contains DNA probes for direct genotyping of the variant underlying the 17-bp deletion in *SMIM1* (NC_000001.11:g.3775437_3775453del, rs566629828). The specific DNA probes used (Probeset ID: AX-86577342, Variant ID: Affx-80267180) for detection of the deletion have shown high specificity and the rare variant can be reliably called.[27] For this study, only directly measured genotypes for variant rs566629828 were used to identify UKB participants homozygous for the 17-bp deletion in *SMIM1*, as opposed to imputed genotypes. Additionally, manual inspection of genotype call plots for the deletion probeset (AX-86577342) was performed for each of the 106 genotyping batches of 4,700 UKB samples.

rs566629828 is in Hardy-Weinberg equilibrium (X^2 test, *p-value* = 0.92) with an allele distribution, in UKB, of +/+ = 396559, +/− = 11849, −/− = 90 and a theoretical expected distribution of +/+ = 396558, +/− = 11852, −/− = 89. The linkage disequilibrium score D′ value between variants rs1175550 and rs566629828 was calculated using PLINK software.[40] The clinical phenotypes have been defined according to the Hospital Episode Statistics (HES) recorded for the majority of UKB participants. The list of ICD-10 codes and fields used to select the cases and traits is in Table S7.

NIHR-NBR cohort

All NIHR-NBR participants for this study were recruited under approval 12/EE/0040 by the Research Ethics Committee East of England. Initially, the *SMIM1*$^{-/-}$ individuals were identified by the testing approach outlined in an earlier publication.[1] In short, the red cells of a small portion of the 1.4 million donations collected annually are tested by haemagglutination for the presence of the Vel blood group antigen with a specific typing reagent. Individuals with a negative result and part of the NIHR-NBR underwent a confirmatory polymerase chain reaction test for the 17-bp deletion. Of these, 25 participated in this study by providing samples and relevant health information obtained during a home visit. Twelve of the 25 attended the NIHR Clinical Research Facility at Cambridge University Hospitals (Cambridge, UK) for a 2-day metabolic assessment, no exclusion criteria were applied. The measurement results of the *SMIM1*$^{-/-}$ individuals were compared with the results obtained for 180 NIHR-NBR participants (100 females, 80 males) with a reference/reference or reference/alternate genotype for variant rs566629828 as previously determined by whole-genome sequencing.[41]

Danish Blood Donor Study (DBDS) and Copenhagen Hospital biobank (CHB) cohorts

The CHB participants were recruited under the NVK-1708829, P-2019-93 approval and the DBDS participants were recruited under the 1-10-72-95-13, NVK-1700407, P-2019-99 approval. The CHB participants have been enrolled at the Copenhagen University Hospital and general hospitals in the region of greater Copenhagen. Inclusion in the study is limited to patients attending to these hospitals and from whom a blood sample is drawn for ABO and D grouping and/or red cell antibody screening. The cohort is therefore strongly skewed toward patients with medical conditions associated with a high likelihood of requiring transfusion (e.g., surgery, chemotherapy and pregnancy).[20] The DBDS cohort of Danish blood donors is in demographics similar to the NIHR-NBR cohort from whom 25 British *SMIM1*$^{-/-}$ individuals were drawn.[19,42]

The DNA samples from the 90,000 DBDS and 90,700 CHB participants were genotyped at deCODE Genetics (Reykjavik, Iceland) using the Infinium Global Screening Array (Probeset ID: GSA-24v1-0_C2, v1.0). Imputation of the 17-bp deletion rs566629828 was performed by deCODE Genetics using their North European sequencing panel of 15,576 individuals (including 8,429 Danes) as reference. Based on these two imputed datasets, 49 and 34 individuals were identified in DBDS and CHB, with a high likelihood of being homozygous for the 17-bp deletion in *SMIM1*, respectively. The DBDS and CHB participant and genotyping data are linked to the Danish Laboratory Database (DLD), the Danish National Patient Registry[43] (NPR) and the Danish Prescription Database[43] (DPD). These linked databases were used for the association analysis performed for this study.

Million Veteran Program (MVP) cohort

The design of MVP has been previously described.[44] Veterans were recruited from over 60 Veterans Health Administration medical centers nationwide since 2011. A unique feature of MVP is the linkage of a large biobank to an extensive, national, database from 2003 onward that integrates multiple elements such as diagnosis codes, procedure codes, laboratory values, and imaging reports, which permits detailed phenotyping of this large cohort. MVP has received ethical and study protocol approval by the Veterans Affairs Central Institutional Review Board in accordance with the principles outlined in the Declaration of Helsinki.

DNA extracted from participants' blood was genotyped using a customized Affymetrix Axiom biobank array, the MVP 1.0 Genotyping Array. The array was enriched for both common and rare genetic variants of clinical significance in different ethnic backgrounds. Quality-control procedures used to assign ancestry, remove low-quality samples and variants, and perform genotype imputation were previously described.[45] We excluded: duplicate samples, samples with more heterozygosity than expected, an excess (>2.5%) of missing genotype calls, or discordance between genetically inferred sex and phenotypic gender.[45] In addition, one individual from each pair of related individuals (more than second-degree relatedness as measured by the KING software were removed.[46] SNP rs566629828 (*SMIM1*) was directly genotyped on the MVP array. The MVP participants were assigned to mutually exclusive racial/ethnic groups using HARE (Harmonized Ancestry and Race/Ethnicity), a machine-learning algorithm that integrates genetically inferred ancestry with self-identified race/ethnicity.[47] The present study included non-Hispanic European Americans with both genotypic and phenotypic data for genetic association analyses. The details of the genetic association study of BMI in the MVP were previously described.[48]

METHOD DETAILS

Confirmation of rs566629828 genotype status

Considering the limited accuracy of imputation to determine the genotype of low-frequency variants, and particularly of indels,[49] the genotype at rs566629828 was confirmed by an orthogonal test[2] using DNA extracted from 49 DBDS and 34 CHB blood samples, which were retrieved from the respective sample repositories. In short, DNA was amplified by primers flanking the 17-bp deletion in *SMIM1* exon 3 (Figure S1). The amplicons were resolved by agarose gel electrophoresis and visual inspection of the amplicon length (reference and alternate alleles being 178bp and 161bp in length, respectively). Discordant results between the genotype inferred by imputation and the PCR-genotyping test results were observed for 10 DNA samples (DBDS, $n = 6$; CHB, $n = 4$). These ambiguities were resolved by Sanger sequencing of the *SMIM1* coding exons 3 and 4 confirming that all 10 discordances were caused by erroneous imputation results. All together 43 and 30 confirmed *SMIM1*$^{-/-}$ individuals were identified in the DBDS and CHB cohorts, respectively (Figure S2). Controls are drawn from the same cohorts in a ratio of 15:1, gender and age-matched (DBDS, $n = 645$; CHB, $n = 450$). The genotype of the controls was imputed and it was either reference/reference or reference/alternate for the variant rs566629828.

Metabolic characterization

The Cambridge Central East of England Research Ethics Committee approved the study protocol for participants' metabolic characterization (06/Q0108/84). NIHR-NBR participants were asked to refrain from exercise, consume alcohol and caffeine for 24 h before arrival. Each of the 12 participants arrived at the NIHR Clinical Research Facility at Cambridge University Hospitals at 14:00 h on day 0 and

remained until noon on day 1. Resting energy expenditure was measured upon waking after an overnight fast by indirect calorimetry (GEM Nutrition) using a ventilated hood. Gas analysis exchange measurements were converted into energy equivalents using calculations by Elia and Livesey.[50] The procedure and precision values of the indirect calorimetry method have been previously described.[51] Whole-body fat, lean and bone mass body composition measurements were performed by Dual Energy X-ray Absorptiometry (DXA). For the volunteers homozygous for the 17-bp deletion in *SMIM1*, GE Lunar iDXA (Encore version 18) was used for fat mass, lean mass and bone mineral content (BMC) measurements. For the controls, there was a combination of GE Lunar iDXA measurements and GE Lunar Prodigy measurements (Encore version 16). Therefore, all relevant measurements were converted by cross-calibration equations[52] to comparable iDXA values before collating and using regression modeling. Lean mass and resting energy expenditure Z scores were derived by multiple regression modeling.[14] The coefficients were updated in line with an upgrade in DXA scanner (resting energy expenditure (kJ/min) = age; −0.015, fat mass (kg); 0.019, lean mass (kg); 0.063, intercept; 1.580, lean mass (kg) = gender (0; male, 1; female); −6.272, height2 (m2); 6.684, bone mass (kg); 10.458, fat mass (kg); 0.166, intercept; 0.888).

Single-cell RNA-seq analyses

We analyzed single-cell RNA-sequencing data from the following sources:

Mouse hypothalamus (GSE113576)[16]

Human fetal pituitary (GSE142653)[17] human in Figure S3E.

Mouse pituitary (GSE120410)[53] mouse in Figure S3E.

Mouse pituitary (GSE146619)[54] mouse in Figure S3E.

Rat pituitary (GSE132224)[55] rat in Figure S3E.

Mouse thyroid organoids (GSE163818)[18]

Normalization, visualization, and standard processing of datasets was done through Seurat.[34] For label transfer of mouse and rat pituitary datasets from the human pituitary reference: mouse and rat genes were first converted to their human homologs (as obtained via BioMart[35]), and ambiguously annotated genes were filtered out, prior to cross-species integration.

QUANTIFICATION AND STATISTICAL ANALYSIS

Linear regression was performed to estimate the effect of continuous variables. The statistical model used as covariates age, sex and BMI. Sex and BMI were removed from the equation when considering sex-stratified data or the effect of the variant on BMI, respectively. Similarly, for categorical variables, the explanatory effect of variant rs566629828 was estimated by logistic regression. For continuous traits, inverse normal transformation (R package RNOmni) was adopted to normalize the measurements.[36] In the Danish cohorts, the logistic regression analysis was performed with the response variable defined as the presence of an abnormal Nomenclature for Properties and Units (NPU)-code measurement in the DLD dataset, presence of a given ICD10/ICD8 record in the NPR dataset, or presence of a specific prescription in the DPD dataset. The explaining variables used were variant

rs566629828 genotype, age of the individuals, genetically inferred sex of the individuals (unless cohort was sex-stratified), and in case of mixed cohort analysis, the cohort of a given individual (DBDS/CHB). Since weight data does not follow a normal distribution, a Wilcoxon signed-rank test was used to assess differences in mean weight-based on variant rs566629828 genotype after sex stratification. Bootstrapping was used to assess directionality in mean weights based on the rs566629828 genotype. For each *SMIM1*$^{-/-}$ DBDS case, 100 alternate age, sex and smoking status-matched control groups were selected at random. The mean weight of each of these 100 alternate control groups was compared to the case group's mean weight. Directionality of the difference in mean weights was then assessed for each sex separately. Statistical tests have *p* values corrected with Benjamini–Hochberg procedure with alpha set at 0.05. The meta-analysis pooling across the cohorts was performed with inverse variance and the R package meta (version 5.2–0).