DNA sequence variation in *ACVR1C* encoding the activin-receptor like kinase 7 influences body fat distribution and protects against type 2 diabetes

Connor A. Emdin, DPhil, Amit V. Khera, MD, Krishna Aragam, MD, Mary Haas, PhD, Mark Chaffin, MS, Derek Klarin, MD, Pradeep Natarajan, MD, Alexander Bick, MD, Seyedeh M. Zekavat, BSc, Akihiro Nomura, MD, Diego Ardissino, MD, James G. Wilson, MD, Heribert Schunkert, MD, Ruth McPherson, MD, Hugh Watkins, MD PhD, Roberto Elosua, MD PhD, Matthew J Bown, FRCS, Nilesh J Samani, FRCP, Usman Baber, MD, Jeanette Erdmann, PhD, Namrata Gupta, PhD, John Danesh FRCP DPhil, Danish Saleheen, MBBS PhD, Stacey Gabriel, PhD, Sekar Kathiresan, MD

Corresponding Author:
Sekar Kathiresan, MD
Center for Genomic Medicine
Massachusetts General Hospital
185 Cambridge Street, CPZN 5.830
Boston, MA 02114
Email: skathiresan1@mgh.harvard.edu
Phone: 617 643 6120

Author Affiliations
1) Center for Genomic Medicine, Massachusetts General Hospital, Harvard Medical School, Boston, MA 02114 (Emdin, Khera, Klarin, Natarajan, Aragam, Haas, Bick, Zekavat, Nomura, Kathiresan)
2) Cardiology Division Department of Medicine, Massachusetts General Hospital, Harvard Medical School, Boston, MA 02114 (Khera, Natarajan, Aragam, Haas, Nomura, Emdin, Kathiresan)
3) Program in Medical and Population Genetics, Broad Institute, Cambridge, MA, 02142 (Emdin, Khera, Chaffin, Natarajan, Klarin, Aragam, Haas, Bick, Nomura, Zekavat, Gupta, Gabriel, Kathiresan)
4) Yale Medical School, New Haven, MA; Department of Computational Biology & Bioinformatics, Yale University, New Haven, MA 06510 (Zekavat)
5) Department of Surgery, Massachusetts General Hospital, Harvard Medical School, Boston, MA 02114 (Klarin)
6) Division of Cardiology, Azienda Ospedaliero–Universitaria di Parma, Parma, Italy 43121 (Ardissino)
7) Associazione per lo Studio Della Trombosi in Cardiologia, Pavia, Italy 27100 (Ardissino)
8) Department of Physiology and Biophysics, University of Mississippi Medical Center, Jackson, MS 39216 (Wilson)
9) Deutsches Herzzentrum München, Technische Universität München, Deutsches Zentrum für Herz-Kreislauf-Forschung, München, Germany 80333 (Schunkert)
10) University of Ottawa Heart Institute, Ottawa, Ontario, Canada K1Y4W7 (McPherson)

11) Division of Cardiovascular Medicine, Radcliffe Department of Medicine, University of Oxford, Oxford, United Kingdom OX12JD (Watkins)

12) Wellcome Trust Centre for Human Genetics, University of Oxford, Oxford, United Kingdom OX12JD (Watkins)

13) Cardiovascular Epidemiology and Genetics, Hospital del Mar Research Institute, Barcelona, Spain 08003 (Elosua)

14) CIBER Enfermedades Cardiovasculares (CIBERCV), Barcelona, Spain 28029 (Elosua)

15) Facultat de Medicina, Universitat de Vic-Central de Cataluña, Vic, Spain 08500 (Elosua)

16) Department of Cardiovascular Sciences, University of Leicester, and NIHR Leicester Biomedical Research Centre, Leicester, United Kingdom LE17RH (Samani, Brown)

17) The Zena and Michael A. Wiener Cardiovascular Institute, Icahn School of Medicine at Mount Sinai, New York, New York 10029 (Baber)

18) Institute for Cardiogenetics, University of Lübeck, Lübeck, Germany 23562 (Erdmann)

19) DZHK (German Research Centre for Cardiovascular Research), partner site Hamburg/Lübeck/Kiel, 23562 Lübeck, Germany (Erdmann)

20) Department of Biostatistics and Epidemiology, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA 19104 (Saleheen)

21) Center for Non-Communicable Diseases, Karachi, Pakistan (Saleheen)

22) Cardiovascular Epidemiology Unit, Department of Public Health and Primary Care, University of Cambridge, Cambridge, UK; Wellcome Trust Sanger Institute, Hinxton, Cambridge, CB10 1SA, UK; National Institute of Health Research Blood and Transplant; Research Unit in Donor Health and Genomics, University of Cambridge (Danesh)

Word Count: 3834

Abstract

A genetic predisposition to higher waist-to-hip ratio adjusted for body mass index (WHRadjBMI), a measure of body fat distribution, associates with increased risk for type 2 diabetes. We conducted an exome-wide association study of coding variation in UK Biobank (405569 individuals) to identify variants that lower WHRadjBMI and protect against type 2 diabetes. We identified four variants in the gene *ACVR1C*, encoding the activin-receptor like kinase 7 receptor expressed on adipocytes and pancreatic beta cells, which independently associated with reduced WHRadjBMI: Asn150His (-0.09 standard deviations, p=$3.4*10^{-17}$), Ile195Thr (-0.15 SD, p=$1.0*10^{-9}$), Ile482Val (-0.019 SD, p=$1.6*10^{-5}$) and rs72927479 (-0.035 SD, p=$2.6*10^{-12}$). Carriers of these variants exhibited reduced percent abdominal fat in dual energy X-ray imaging. Pooling across all four variants, a 0.2 SD decrease in WHRadjBMI through *ACVR1C* was associated with a 30% lower risk of type 2 diabetes (OR 0.70, CI 0.63, 0.77; p = $5.6*10^{-13}$). In an analysis of exome sequences from 55516 individuals, carriers of predicted damaging variants in *ACVR1C* were at 54% lower risk of type 2 diabetes (OR 0.46 CI 0.27, 0.81; p=0.006). These findings indicate that variants predicted to lead to loss of *ACVR1C* gene function influence body fat distribution and protect from type 2 diabetes.

Introduction

Discovery of genetic variants that protect against disease can identify novel mechanisms of disease and novel therapeutic targets.(1) For example, the discovery of low frequency coding variants in *PCSK9*, *ANGPTL3* and *APOC3* that lower blood lipid levels and protect against coronary artery disease catalyzed the development of novel therapeutics for coronary artery disease. (2-10) *PCSK9* inhibitors are now approved for treatment of coronary artery disease(4) while inhibitors of *ANGPTL3*(7) and *APOC3*(11) are in clinical development.

Body fat distribution strongly influences the development of type 2 diabetes.(12-14) In a Mendelian randomization study of 296 291 individuals, we found that a genetic predisposition to increased abdominal fat distribution was associated with elevated triglyceride levels, elevated blood pressure and an increased risk of coronary artery disease, independent of overall adiposity.(12) Furthermore, a genetic predisposition to increased abdominal fat distribution was strongly associated with the development of type 2 diabetes. For each one standard deviation genetic increase in waist-to-hip ratio adjusted for body mass index (WHRadjBMI, a measure of body fat distribution), risk of type 2 diabetes increased by 77%.(12) These findings were replicated in a separate Mendelian randomization study.(14)

These results suggest the hypothesis that genetic variants that influence body fat distribution may also influence the risk of type 2 diabetes.(12) Here, we test this hypothesis by analyzing genetic variation in more than 400000 individuals in UK Biobank to identify novel genetic variants that lower WHRadjBMI and protect against type 2 diabetes. Below, we demonstrate that variants predicted to lead to loss of function of the gene *ACVR1C*, encoding the activin-receptor like kinase 7, influence body fat distribution and protect against type 2 diabetes.

Methods

Study Design

In our discovery analysis, we analyzed the association of 614042 coding variants with WHRadjBMI in 405569 individuals in UK Biobank. Coding variants were defined as missense variants or variants predicted to result in loss of function of the protein: (1) nonsense mutations that resulted in early termination of a protein; (2) frameshift mutations due to insertions or deletions of DNA; or (3) splice-site mutations which result in an incorrectly spliced protein. Only variants imputed with a quality score (info score) > 0.3 were analysed. Threshold for significance was defined as $p < 5*10^{-8}$ (genome wide significance) and analysis was performed using PLINK2 software.[15] We reported novel variants as those located more than 1 megabase away from prior identified loci in the GIANT consortium.[16] We reported variants located outside of the MHC locus separately from those within the MHC locus, as variants within the MHC locus typically tag HLA risk alleles and are thus associated with phenotypes due to linkage disequilibrium with HLA alleles.[17] We attempted to replicate the association of novel variants with WHRadjBMI using data from the GIANT consortium, when the variant was available in the GIANT consortium.[16]

Upon identification of variants in *ACVR1C* as significantly associated with WHRadjBMI, a conditional analysis was conducted to identify additional variants significantly associated ($p < 0.0001$) with WHRadjBMI within the *ACVR1C* locus ($\pm$ 250 kb of the lead variant rs55920843). To replicate observed associations of *ACVR1C* variants with WHRadjBMI, we examined whether carriers of these variants had reduced WHRadjBMI in a meta-analysis of the GIANT consortium and two independent cohorts (Atherosclerosis Risk in Communities and Framingham Heart studies). We also examined whether these variants were associated with direct imaging-based measurements of abdominal fat in 4215 participants who underwent dual X-ray absorptiometry imaging in UK Biobank.

To test whether identified *ACVR1C* variants were also associated with risk of type 2 diabetes, we pooled data from the DIAGRAM Consortium (ExTexT2D exome chip analysis[18]) with UK Biobank. To test whether variation leading to loss of *ACVR1C* function protects against type 2 diabetes, we analyzed the sequences of the 9 exons of *ACVR1C* in 55516 participants [31672 from the Myocardial Infarction Genetics Consortium (MIGen)[19,20], 5388 from the Atherosclerosis Risk in Communities study and 18456 from the T2D Genes Consortium[21]] and examined if predicted damaging variants in the gene associate with risk of type 2 diabetes.

A phenome-wide association study of *ACVR1C* in UK Biobank was performed using an *ACVR1C* gene risk score.[22] Definitions for 31 different diseases analyzed in the phenome-wide association study are provided (Supp. Table 1). Three metabolic traits available in UK Biobank (urinary albumin-to-creatinine ratio, systolic blood pressure and diastolic blood pressure) were also analyzed. A p-value of 0.001 (0.05/34) was used for significance in this analysis.

Data Sources

For the analysis of WHRadjBMI, individual-level data from 405569 unrelated individuals from the UK Biobank (335,660 individuals of European ancestry and 69,909 individuals of Non-European ancestry) was analyzed. UK Biobank received ethical approval from the Research Ethics Committee (reference number 11/NW/0382). Analysis of UK Biobank was approved by the Partners Health Care Institutional Review Board (protocol 2013P001840). Informed consent was obtained from all participants by UK Biobank. For replication, data for WHRadjBMI from the GIANT consortium (in which the Ile482Val variant was available) was pooled with data from the Atherosclerosis Risk in Communities and Framingham Heart Study datasets (in which Asn150His, Ile195Thr and rs72927479 variants were available). The GIANT consortium consisted of 224459 participants (210088 of European ancestry and 14371 of non-European ancestry) genotyped using the MetaboChip.(16) Atherosclerosis Risk in Communities study is a community based study of 15792 white and black participants, aged 45 to 64 years.(23) The Atherosclerosis Risk in Communities dataset consisted of 10122 individuals (8015 of European ancestry and 2107 of Non-European ancestry) who were genotyped and imputed, as previously described.(24) For 5388 participants, exome sequences were also available for analysis. In the Framingham Heart Study, a community based study of 10092 individuals of predominantly European ancestry, genotyped data was available from 6073 individuals of European ancestry.

For the analysis of type 2 diabetes, estimates from UK Biobank were pooled using inverse variance weighted fixed effects meta-analysis with estimates from the DIAGRAM ExTexT2D exome chip analysis of 452244 participants (81412 diabetes cases and 370832 controls).(18) In UK Biobank, type 2 diabetes was defined as (1) self-report of type 2 diabetes, followed by a verbal interview with a trained nurse to confirm the diagnosis or (2) hospitalization for ICD code E11. Because the ExTexT2D analysis included 120286 participants from UK Biobank(18), these individuals were excluded from the analysis of type 2 diabetes in UK Biobank to prevent analysis of overlapping samples.

Sequence data for *ACVR1C* were extracted from exome sequencing performed in the MIGen Consortium as previously described.(19,20) The Burrows–Wheeler Aligner algorithm was used to align reads from participants to the reference genome (hg19). The GATK HaploTypeCaller was used to jointly call variants. Metrics including Variant Quality Score Recalibration (VQSR), quality over depth, and strand bias were then used to filter variants. The Jackson Heart Study was excluded from analysis of MIGen as it was included in the T2D Genes consortium. Exome sequences from 5388 participants in ARIC were analysed as previously described.(25) Phenotype and genotype data were retrieved from the National Center for Biotechnology Information dbGAP server (accession: phs000090.v3.p1 and phs000572.v6.p4). Exome sequencing was performed in the T2D Genes consortium as previously described.(21) To analyse exome sequences from the T2D Genes consortium, the online Genetic Association Interactive Test in the T2D Knowledge portal was used.(21)

Studies included in the MIGen consortium were: 1) the Italian Atherosclerosis Thrombosis and Vascular Biology (ATVB) study (dbGaP Study Accession phs000814.v1.p1); 2) the Exome Sequencing Project Early-Onset Myocardial Infarction (ESP-EOMI) study(9); 3) a nested case-control cohort from the Jackson Heart Study (JHS); 4) the South German Myocardial Infarction study (dbGaP Study Accession phs000916.v1.p1); 5) the Ottawa Heart Study (OHS) (dbGaP Study Accession phs000806.v1.p1); 6) the Precocious Coronary Artery Disease (PROCARDIS) study (dbGaP Study Accession phs000883.v1.p1) ; 7) the Pakistan Risk of Myocardial Infarction Study (PROMIS) (dbGaP Study Accession phs000917.v1.p1); 8) the Registre Gironi del COR (Gerona Heart Registry or REGICOR) study (dbGaP Study Accession phs000902.v1.p1); 9) the Leicester Myocardial Infarction study (dbGaP Study Accession phs001000.v1.p1); 10) the BioImage study (dbGaP Study Accession phs001058.v1.p1); 11) and the North German Myocardial Infarction study (dbGaP Study Accession phs000990.v1.p1).

Predicted damaging *ACVR1C* variants in the exome sequencing analysis were defined as those which resulted in loss-of-function of the protein (nonsense mutations that resulted in early termination of *ACVR1C*, frameshift mutations due to insertions or deletions of DNA, or splice-site mutations which result in an incorrectly spliced protein) or those labeled as damaging by each of five different algorithms (LRT score, MutationTaster, PolyPhen-2 HumDiv, PolyPhen-2 HumVar, and SIFT), as previously described.(19,26) The Variant Effect Predictor algorithm was used to annotate predicted damaging variants.(27)

Statistical Analysis

In UK Biobank, WHRadjBMI was derived through inverse normal transformation of waist-to-hip ratio after adjustment for age, sex and BMI (as in the GIANT collaboration(16)). In UK Biobank, linear regression was used to estimate the association of variants with WHRadjBMI. All UK Biobank analyses included adjustment for age, sex, 10 principal components of ancestry, and a dummy variable for the array type used in genotyping. Logistic regression was used to estimate the association of variants with type 2 diabetes. Estimates of the association of each variant with type 2 diabetes in UK Biobank were pooled with estimates from the ExTexT2D consortium using inverse-variant weighted fixed effects meta-analysis.

In the primary analysis of 405569 individuals for WHRadjBMI, we had 80% power to detect a minimum effect of 0.05 SD with a minor allele frequency of 5% at genome wide significance ($p<5*10^{-8}$). In the primary analysis of 95 978 type 2 diabetes cases and 64 6985 controls, we had 80% power to detect a minimum odds ratio of 1.05 at $p<5*10^{-8}$. With a minor allele frequency of 1%, we had 80% power to detect a minimum effect of 0.1 SD for WHRadjBMI and an OR of 1.10 for type 2 diabetes at $p<5*10^{-8}$.

To estimate the overall association of variation in *ACVR1C* with WHRadjBMI, we pooled across all variants using a gene risk score, weighted by the square root of allele frequency (estimating a weighted mean effect of *ACVR1C* variants on WHRadjBMI).(28) To estimate the overall association of variation in *ACVR1C* with type 2 diabetes, we

pooled across all variants in a gene risk score, weighted by the association of each variant with WHRadjBMI.(29)

For analysis of exome sequencing data, logistic regression was performed with adjustment for sex, five principal components of ancestry and a dummy variable for each cohort to estimate the association of predicted damaging variants with type 2 diabetes. Estimates from the MIGen consortium were pooled with estimates from the T2D Genes consortium using inverse-variance weighted fixed effects meta-analysis.

For the phenome-wide association study, all four *ACVR1C* variants were pooled in a gene-risk score in UK Biobank, as previously described.(25,30) For each individual in UK Biobank, the *ACVR1C* variants associated with lower WHRadjBMI were weighted by their effect on WHRadjBMI and summed. The association of this gene-risk score with 31 different diseases in UK Biobank and three metabolic traits was tested using logistic regression with adjustment for age, sex, 10 principal components of ancestry, and a dummy variable for the array type used in genotyping. Although exploratory due to the low number of cases for certain diseases (e.g. 1707 cases for cervical cancer), we conducted this analysis to detect possible adverse associations of *ACVR1C* with diseases that could allow for prediction of adverse effects of pharmacologic inhibition of ACVR1C.

Analyses were performed using R version 3.2.3 (R Project for Statistical Computing).

Results

Exome wide association study of body fat distribution in UK Biobank
Among 405569 participants in UK Biobank, 54% were female, the median age was 57 and the median waist-to-hip ratio, measured at enrollment, was 0.87 (Table 1). One standard deviation in waist-to-hip ratio corresponded to an absolute change of 0.09. In an analysis of 614012 coding variants in UK Biobank, no evidence of genomic inflation was observed (lambda 1.08, Supp. Figure 1).

We identified 16 low frequency variants (<5%) associated with WHRadjBMI outside of known loci (Supp. Table 2).(16) We identified an additional 43 novel common variants (frequency >5%) associated with WHRadjBMI (Supp. Table 3). We identified 94 independent variants in total, including variants at known loci (Online Data). Of 59 novel coding variants, 34 variants were available in the GIANT consortium for replication. (16) A strong correlation in the effect sizes of the association of variants with WHRadjBMI in UK Biobank and GIANT was observed (R^2=0.87). 24 variants were independently associated with WHRadjBMI in GIANT (p<0.05, Supp. Table 4). Of the 10 variants that were not significantly associated with WHRadjBMI in GIANT (p>0.05), only three variants exhibited evidence of heterogeneity between estimates in UK Biobank and GIANT, with the remaining 7 variants showing similar estimates of association in UK Biobank and GIANT (although not reaching significance in GIANT, Supp. Table 4).

The lead novel low frequency variant was a missense variant in *PNPLA2* (rs140201358, Asn252Lys) that associated with elevated WHRadjBMI (0.09 SD, p=$3.2*10^{-19}$). *PNPLA2* encodes adipocyte triglyceride lipase which hydrolyzes triglycerides in adipose tissue to mobilize fat stores.(31) Two low frequency missense variants in *ABHD15*, encoding alpha/beta-hydrolase domain containing protein 15, were found to be associated with elevated WHRadjBMI. *ABHD15* is also highly expressed in adipocytes and has been reported to mediate insulin-induced suppression of lipolysis in adipocytes.(32) We also identified a low frequency variant in a known locus (*CALCRL* Leu87Pro, 0.1% frequency) associated with lower WHRadjBMI (-0.14 SD, p=$1.9*10^{-11}$). A common non-coding variant in the *CALCRL* locus, encoding calcitonin receptor-like receptor, was previously identified in the GIANT consortium as associated with WHRadjBMI.(16) The identification of an independent low frequency missense variant in the gene suggests that *CALCRL* may be the causal gene at this locus.

Variation leading to lower WHRadjBMI: the ACVR1C locus
We next focused on novel variant alleles leading to lower WHRadjBMI. The lead novel low frequency variant associated with lower WHRadjBMI lay within the gene *ACVR1C*. *ACVR1C* Asn150His (allele frequency 1.1%), associated with 0.09 standard deviations (SD) lower WHRadjBMI (p=$3.4*10^{-17}$). An independent missense variant in ACVR1C, Ile195Thr (AF 0.4%), also associated with lower WHRadjBMI (0.15 SD, p=$1.0*10^{-9}$).

Upon conditioning on these two variants, we identified an additional coding variant: Ile482Val (AF 7%), which associated with 0.019 SD lower WHRadjBMI (p=$1.6*10^{-5}$) and rs72927479, a non-coding variant, for which the minor G allele (AF 5%) was associated with 0.035 SD lower WHRadjBMI (p=$2.6*10^{-12}$, Table 2). Despite being an

independent signal for WHRadjBMI, the non-coding variant rs72927479 is nominally correlated with Ile482Val ($r^2 = 0.06$ in UK Biobank). No other variants were correlated with one another in UK Biobank (all $r^2 < 0.001$).

Pooling across all four variants with weighting by square root of allele frequency, *ACVR1C* variation was associated with lower WHRadjBMI in UK Biobank (-0.07 SD, $p=2.6*10^{-35}$). To replicate this finding, we pooled data from the GIANT consortium (in which the Ile482Val variant was available, n=224156) with data from the Atherosclerosis Risk in Communities and Framingham Heart Study datasets (in which Asn150His, Ile195Thr and rs72927479 variants were available, n=13704). Pooling across all four variants in these three replication studies, variation in *ACVR1C* was associated with reduced WHRadjBMI (-0.07 SD, p=0.0005, Supp. Table 5).

Examining other anthropometric traits in UK Biobank, variation in *ACVR1C* was associated with elevated hip circumference (0.035 SD, $p=3*10^{-8}$), nominally elevated BMI (0.02 SD, p=0.002) and was unassociated with waist circumference (-0.007 SD, p=0.20) or height (0.005 SD, p=0.23). In the MAGIC consortium analysis of insulin resistance (HOMA-IR), Ile482Val (the only *ACVR1C* variant available for analysis), was unassociated with HOMA-IR (-1.1%, p=0.17).(33)

We examined whether *ACVR1C* variation also associated with direct imaging measurement of abdominal obesity. 4215 participants in UK Biobank underwent dual energy X-ray absorptiometry (DEXA) imaging to estimate abdominal fat mass. Carriers of *ACVR1C* variants had lower percent abdominal fat (p=0.008, Figure 1). When one outlying individual with four *ACVR1C* variants was excluded (Figure 1), carriers of *ACVR1C* continued to have significantly lower percent abdominal fat (p=0.009).

Association of variants in ACVR1C with type 2 diabetes
Genetic predisposition to increased WHRadjBMI strongly predisposes to type 2 diabetes.(12) We therefore examined whether variants in *ACVR1C* that lower WHRadjBMI adiposity protect against type 2 diabetes. In a combined analysis of UK Biobank and the DIAGRAM consortium, all four *ACVR1C* variants were found to independently protect against type 2 diabetes (OR 0.88, $p=8.7*10^{-5}$ for Asn150His; OR 0.79, p=0.005 for Ile195Thr; OR 0.95, $p=4.8*10^{-6}$ for Ile482Val; OR 0.93, p=0.0006 for rs72927479; Table 1). Pooling across all four variants, a 0.2 SD decrease in WHRadjBMI through *ACVR1C* was associated with a 30% lower risk of type 2 diabetes (OR 0.70 CI 0.63, 0.77; $p = 5.6*10^{-13}$, Figure 2). When we excluded the non-coding variant rs72927479 which is nominally correlated with Ile482Val ($r^2 = 0.06$ in UK Biobank), the *ACVR1C* gene risk score remained associated with risk of type 2 diabetes (OR 0.71 CI 0.64, 0.79, $p=1.8*10^{-10}$)

In three independent datasets with exome sequence data – the MIGen Consortium, ARIC and the T2D Genes Consortium, we examined whether variants that lead to loss of *ACVR1C* gene function protect against type 2 diabetes. The 9 exons of the *ACVR1C* gene were sequenced in 55516 individuals. 105 predicted damaging variants were identified (Supp. Tables 6-8). Among 16452 type 2 diabetes cases, the frequency of predicted damaging variants in *ACVR1C* was 0.1% (17) compared to 0.2% (88) among 39064

diabetes-free controls. Overall, carrying a predicted damaging variant in *ACVR1C* was associated with 54% lower risk of type 2 diabetes (OR 0.46 CI 0.27, 0.81; p=0.006; Figure 3). When we excluded 47 carriers of I195T (annotated as a predicted damaging variant by five of five algorithms) in the exome sequencing analysis, carriers of predicted damaging variants in *ACVR1C* remained protected from type 2 diabetes (OR 0.48 CI 0.24, 0.97, p=0.04).

To further examine whether loss of *ACVR1C* function lowers WHRadjBMI, we examined whether the non-coding variant rs72927479 associates with *ACVR1C* expression. The minor allele of rs72927479 (G, frequency 5%) associated with lower WHRadjBMI (beta -0.035, p=$2.6*10^{-12}$) and type 2 diabetes (OR 0.93, p=0.0006). In the Genotype-Tissue Expression dataset(34), the minor allele of rs72927479 associated with reduced expression of *ACVR1C* in subcutaneous adipose tissue (p=0.02) and pancreas (p=0.02, Supp. Figure 2).

Phenome-wide association study of ACVR1C in UK Biobank.

To anticipate whether *ACVR1C* inhibition may be associated with on-target adverse effects, we conducted a phenome-wide association study of 31 disease phenotypes in UK Biobank. We did not observe any significant associations between the *ACVR1C* gene risk score and the 31 diseases analyzed (Figure 4). The *ACVR1C* gene risk score was unassociated with coronary artery disease (OR 1.01, p=0.86). We also examined whether *ACVR1C* variation associates with three metabolic traits currently available in UK Biobank: urinary albumin, systolic blood pressure and diastolic blood pressure. While the *ACVR1C* gene risk score did not significantly associate with urinary albumin levels, it associated with significantly lower diastolic blood pressure (-0.6 mm Hg, p=0.0004) and nominally lower systolic blood pressure (-0.6 mm Hg, p=0.03).

Discussion

In this study, four genetic variants in *ACVR1C,* ranging in frequency from 0.2% to 7.2%, independently associated with lower WHRadjBMI and protected against type 2 diabetes. Furthermore, damaging variants in *ACVR1C* protected against type 2 diabetes in an analysis of exome sequences from 55516 individuals. An *ACVR1C* gene risk score did not associate with any of 31 additional diseases in UK Biobank but did nominally associate with lower blood pressure.

These results permit several conclusions. First, pharmaceutical inhibition of *ACVR1C* may be useful in the treatment of type 2 diabetes. *ACVR1C* encodes the activin-receptor like kinase 7 (ALK7), a transforming growth factor-beta family receptor highly expressed on pancreatic islet cells(35,36) and adipocytes(37). Overexpression of AKL7 induces growth inhibition and apoptosis of pancreatic beta cells(36,38), suggesting that it is a negative regulator of beta cell mass. A number of findings from model systems also suggest that ALK7 may be useful as a therapeutic target for abdominal obesity, type 2 diabetes and other metabolic diseases. *ACVR1C* deficient mice have been reported to have reduced body fat when fed a high-fat-diet(37,39) and have improved glucose tolerance and insulin sensitivity when obese.(37) Chemical inhibition of *ACVR1C* has also been shown to reduce fat accumulation and increase lipolysis in mice.(40) Rats with streptozocin-induced diabetes show elevated ALK7 expression and shRNA knockdown of *ACVR1C* reduces arterial stiffness in this model.(41) In combination with the human genetic results presented here, these findings suggest that inhibition of *ACVR1C* may prove useful to modify body fat distribution and lower risk for type 2 diabetes.

Second, the lack of association of the *ACVR1C* gene risk score with 31 different diseases in UK Biobank suggests that therapeutic *ACVR1C* inhibition may not have adverse on-target effects. In the Exome Aggregation Consortium, *ACVR1C* is tolerant of loss of function variants, with 18 of 66720 individuals of European ancestry carrying an early stop codon (Leu32Ter) in *ACVR1C*. (42) In combination with the phenome-wide association study presented here, these findings suggest that *ACVR1C* could be safely inhibited. However, due to the small number of cases for many of the analyzed diseases and the multiple diseases tested in the phenome-wide association study, modest associations of *ACVR1C* variation with the analyzed phenotypes cannot excluded. Furthermore, many common adverse effects of therapeutics, such as elevations in liver function enzymes, could not be analysed due to a lack of available data in UK Biobank. In particular, ALK7 deficient mice have been reported to have lengthened QT intervals(43), a phenotype that was unavailable for analysis in UK Biobank. The association of the *ACVR1C* gene risk score with nominally lower diastolic and systolic blood pressure suggests that *ACVR1C* inhibition may have the additional benefit of lowering blood pressure. However, this finding, which did not reach genome-wide significance, requires replication in independent datasets.

The primary strength of the analysis is the use of multiple data source to replicate the association of *ACVR1C* variation with WHRadjBMI and type 2 diabetes. We demonstrated associations of *ACVR1C* with WHRadjBMI and type 2 diabetes to greater than genome-wide significance (p=2.6*10^{-35} for WHRadjBMI and p=5.6*10^{-13} for type 2

diabetes). We further showed that variants leading to loss of *ACVR1C* function protected against diabetes in an analysis of predicted damaging variants in exome sequences of 55516 individuals. A primary limitation of the analysis is that we did not experimentally characterize the analyzed variants. The consistent protective associations observed with three different *ACVR1C* missense variants, with a non-coding variant that reduces *ACVR1C* expression, and with variants predicted to truncate the *ACVR1C* protein or predicted to damage *ACVR1C* function by five different algorithms suggest that variants leading to loss of *ACVR1C* function protect against type 2 diabetes. However, experimental demonstration that *ACVR1C* variants that lower WHRadjBMI and protect against type 2 diabetes actually reduce ALK7 receptor function is necessary before it can be concluded that *ACVR1C* deficiency will protect against type 2 diabetes. A second limitation is that pharmaceutical *ACVR1C* inhibition may be associated with off-target effects that cannot be characterized in a human genetic study.

In summary, variants predicted to damage *ACVR1C* gene function lower WHRadjBMI and protect against type 2 diabetes. These findings provide human genetic validation for the *ACVR1C* gene as a therapeutic target for type 2 diabetes.

References

1. Edwards SL, Beesley J, French JD, Dunning AM. Beyond GWASs: illuminating the dark road from association to function. Am J Hum Genet. 2013 Nov 7;93(5):779–97.

2. Abifadel M, Varret M, Rabès J-P, Allard D, Ouguerram K, Devillers M, et al. Mutations in PCSK9 cause autosomal dominant hypercholesterolemia. Nat Genet. 2003 Jun;34(2):154–6.

3. Cohen JC, Boerwinkle E, Mosley TH, Hobbs HH. Sequence variations in PCSK9, low LDL, and protection against coronary heart disease. N Engl J Med. Massachusetts Medical Society; 2006 Mar 23;354(12):1264–72.

4. Sabatine MS, Giugliano RP, Keech AC, Honarpour N, Wiviott SD, Murphy SA, et al. Evolocumab and Clinical Outcomes in Patients with Cardiovascular Disease. N Engl J Med. Massachusetts Medical Society; 2017 Mar 17;:NEJMoa1615664.

5. Stitziel NO, Khera AV, Wang X, Bierhals AJ, Vourakis AC, Sperry AE, et al. ANGPTL3 Deficiency and Protection Against Coronary Artery Disease. J Am Coll Cardiol. 2017 Apr 25;69(16):2054–63.

6. Dewey FE, Gusarova V, Dunbar RL, O'Dushlaine C, Schurmann C, Gottesman O, et al. Genetic and Pharmacologic Inactivation of ANGPTL3 and Cardiovascular Disease. N Engl J Med. Massachusetts Medical Society; 2017 Jul 20;377(3):211–21.

7. Graham MJ, Lee RG, Brandt TA, Tai L-J, Fu W, Peralta R, et al. Cardiovascular and Metabolic Effects of ANGPTL3 Antisense Oligonucleotides. N Engl J Med. Massachusetts Medical Society; 2017 Jul 20;377(3):222–32.

8. Pollin TI, Damcott CM, Shen H, Ott SH, Shelton J, Horenstein RB, et al. A null mutation in human APOC3 confers a favorable plasma lipid profile and apparent cardioprotection. Science. American Association for the Advancement of Science; 2008 Dec 12;322(5908):1702–5.

9. TG and HDL Working Group of the Exome Sequencing Project, National Heart, Lung, and Blood Institute, Crosby J, Peloso GM, Auer PL, Crosslin DR, Stitziel NO, et al. Loss-of-function mutations in APOC3, triglycerides, and coronary disease. N Engl J Med. 2014 Jul 3;371(1):22–31.

10. Jørgensen AB, Frikke-Schmidt R, Nordestgaard BG, Tybjærg-Hansen A. Loss-of-function mutations in APOC3 and risk of ischemic vascular disease. N Engl J Med. Massachusetts Medical Society; 2014 Jul 3;371(1):32–41.

11. Gaudet D, Alexander VJ, Baker BF, Brisson D, Tremblay K, Singleton W, et al. Antisense Inhibition of Apolipoprotein C-III in Patients with

Hypertriglyceridemia. N Engl J Med. Massachusetts Medical Society; 2015 Jul 30;373(5):438–47.

12. Emdin C, Khera AV, Natarajan P, Klarin D, Zekavat SM, Hsiao AJ, et al. Genetic Association of Waist-to-Hip Ratio With Cardiometabolic Traits, Type 2 Diabetes, and Coronary Heart Disease. JAMA : the journal of the American Medical Association. American Medical Association; 2017 Feb 14;317(6):626–34.

13. Yusuf S, Hawken S, Ounpuu S, Bautista L, Franzosi MG, Commerford P, et al. Obesity and the risk of myocardial infarction in 27,000 participants from 52 countries: a case-control study. 2005 Nov 5;366(9497):1640–9.

14. Dale CE, Fatemifar G, Palmer TM, White J, Prieto-Merino D, Zabaneh D, et al. Causal Associations of Adiposity and Body Fat Distribution With Coronary Heart Disease, Stroke Subtypes, and Type 2 Diabetes Mellitus: A Mendelian Randomization Analysis. Circulation. American Heart Association, Inc; 2017 Jun 13;135(24):2373–88.

15. Chang CC, Chow CC, Tellier LC, Vattikuti S, Purcell SM, Lee JJ. Second-generation PLINK: rising to the challenge of larger and richer datasets. Gigascience. 2015;4(1):7.

16. Shungin D, Winkler TW, Croteau-Chonka DC, Ferreira T, Locke AE, Mägi R, et al. New genetic loci link adipose and insulin biology to body fat distribution. Nature. Nature Publishing Group; 2015 Feb 12;518(7538):187–96.

17. DeBoever C, Tanigawa Y, Lindholm ME, McInnes G, Lavertu A, Ingelsson E, et al. Medical relevance of protein-truncating variants across 337,205 individuals in the UK Biobank study. Nat Commun. Nature Publishing Group; 2018 Apr 24;9(1):1612.

18. Mahajan A, Wessel J, Willems SM, Zhao W, Robertson NR, Chu AY, et al. Refining the accuracy of validated target identification through coding variant fine-mapping in type 2 diabetes. Nat Genet. Nature Publishing Group; 2018 Apr;50(4):559–71.

19. Do R, Stitziel NO, Won H-H, Jørgensen AB, Duga S, Angelica Merlini P, et al. Exome sequencing identifies rare LDLR and APOA5 alleles conferring risk for myocardial infarction. Nature. Nature Publishing Group; 2015 Feb 5;518(7537):102–6.

20. Khera AV, Won H-H, Peloso GM, Lawson KS, Bartz TM, Deng X, et al. Diagnostic Yield of Sequencing Familial Hypercholesterolemia Genes in Patients with Severe Hypercholesterolemia. J Am Coll Cardiol. 2016 Mar 28.

21. Fuchsberger C, Flannick J, Teslovich TM, Mahajan A, Agarwala V, Gaulton KJ, et al. The genetic architecture of type 2 diabetes. Nature. 2016 Aug 4;536(7614):41–7.

22. Emdin C, Khera AV, Klarin D, Natarajan P, Zekavat SM, Nomura A, et al. Phenotypic Consequences of a Genetic Predisposition to Enhanced Nitric Oxide Signaling. Circulation. American Heart Association, Inc; 2018 Jan 16;137(3):222–32.

23. The Atherosclerosis Risk in Communities (ARIC) Study: design and objectives. The ARIC investigators. American journal of epidemiology. 1989 Apr;129(4):687–702.

24. Khera AV, Chaffin M, Aragam K, Emdin C, Klarin D, Haas M, et al. Genome-wide polygenic score to identify a monogenic risk-equivalent for coronary disease. bioRxiv. Cold Spring Harbor Laboratory; 2017 Nov 15;:218388.

25. Emdin C, Khera AV, Natarajan P, Klarin D, Won H-H, Peloso GM, et al. Phenotypic Characterization of Genetically Lowered Human Lipoprotein(a) Levels. J Am Coll Cardiol. 2016 Dec 27;68(25):2761–72.

26. Khera AV, Won H-H, Peloso GM, O'Dushlaine C, Liu D, Stitziel NO, et al. Association of Rare and Common Variation in the Lipoprotein Lipase Gene With Coronary Artery Disease. JAMA : the journal of the American Medical Association. American Medical Association; 2017 Mar 7;317(9):937–46.

27. McLaren W, Pritchard B, Rios D, Chen Y, Flicek P, Cunningham F. Deriving the consequences of genomic variants with the Ensembl API and SNP Effect Predictor. Bioinformatics. Oxford University Press; 2010 Aug 15;26(16):2069–70.

28. Park J-H, Gail MH, Weinberg CR, Carroll RJ, Chung CC, Wang Z, et al. Distribution of allele frequencies and effect sizes and their interrelationships for common genetic susceptibility variants. Proc Natl Acad Sci USA. National Acad Sciences; 2011 Nov 1;108(44):18026–31.

29. Liu DJ, Peloso GM, Yu H, Butterworth AS, Wang X, Mahajan A, et al. Exome-wide association study of plasma lipids in >300,000 individuals. Nat Genet. Nature Publishing Group; 2017 Dec;49(12):1758–66.

30. Emdin C, Klarin D, Natarajan P, CARDIOGRAM Exome Consortium, Florez JC, Kathiresan S, et al. Genetic Variation at the Sulfonylurea Receptor, Type 2 Diabetes, and Coronary Heart Disease. Diabetes. American Diabetes Association; 2017 Aug;66(8):2310–5.

31. Zimmermann R, Strauss JG, Haemmerle G, Schoiswohl G, Birner-Gruenberger R, Riederer M, et al. Fat mobilization in adipose tissue is promoted by adipose triglyceride lipase. Science. American Association for the Advancement of Science; 2004 Nov 19;306(5700):1383–6.

32. Xia W, Pessentheiner AR, Hofer DC, Amor M, Schreiber R, Schoiswohl G, et al. Loss of ABHD15 Impairs the Anti-lipolytic Action of Insulin by Altering PDE3B Stability and Contributes to Insulin Resistance. Cell Rep. 2018 May

15;23(7):1948–61.

33. Dupuis J, Langenberg C, Prokopenko I, Saxena R, Soranzo N, Jackson AU, et al. New genetic loci implicated in fasting glucose homeostasis and their impact on type 2 diabetes risk. Nat Genet. Nature Publishing Group; 2010 Feb;42(2):105–16.

34. GTEx Consortium. Human genomics. The Genotype-Tissue Expression (GTEx) pilot analysis: multitissue gene regulation in humans. Science. American Association for the Advancement of Science; 2015 May 8;348(6235):648–60.

35. Bondestam J, Huotari MA, Morén A, Ustinov J, Kaivo-Oja N, Kallio J, et al. cDNA cloning, expression studies and chromosome mapping of human type I serine/threonine kinase receptor ALK7 (ACVR1C). Cytogenet Cell Genet. Karger Publishers; 2001;95(3-4):157–62.

36. Zhang N, Kumar M, Xu G, Ju W, Yoon T, Xu E, et al. Activin receptor-like kinase 7 induces apoptosis of pancreatic beta cells and beta cell lines. Diabetologia. Springer-Verlag; 2006 Mar;49(3):506–18.

37. Yogosawa S, Mizutani S, Ogawa Y, Izumi T. Activin receptor-like kinase 7 suppresses lipolysis to accumulate fat in obesity through downregulation of peroxisome proliferator-activated receptor γ and C/EBPα. Diabetes. American Diabetes Association; 2013 Jan;62(1):115–23.

38. Zhao F, Huang F, Tang M, Li X, Zhang N, Amfilochiadis A, et al. Nodal induces apoptosis through activation of the ALK7 signaling pathway in pancreatic INS-1 β-cells. American Journal of Physiology - Endocrinology and Metabolism. American Physiological Society Bethesda, MD; 2012 Jul 1;303(1):E132–43.

39. Andersson O, Korach-Andre M, Reissmann E, Ibáñez CF, Bertolino P. Growth/differentiation factor 3 signals through ALK7 and regulates accumulation of adipose tissue and diet-induced obesity. Proc Natl Acad Sci USA. National Acad Sciences; 2008 May 20;105(20):7252–6.

40. Guo T, Marmol P, Moliner A, Björnholm M, Zhang C, Shokat KM, et al. Adipocyte ALK7 links nutrient overload to catecholamine resistance in obesity. Elife. eLife Sciences Publications, Ltd; 2014 Aug 25;3:e03245.

41. Li W-B, Zhao J, Liu L, Wang Z-H, Han L, Zhong M, et al. Silencing of activin receptor-like kinase 7 alleviates aortic stiffness in type 2 diabetic rats. Acta Diabetol. Springer Milan; 2015 Aug;52(4):717–26.

42. Lek M, Karczewski KJ, Minikel EV, Samocha KE, Banks E, Fennell T, et al. Analysis of protein-coding genetic variation in 60,706 humans. Nature. Nature Research; 2016 Aug 18;536(7616):285–91.

43. Ying S, Cao H, Hu H, Wang X, Tang Y, Huang C. Alk7 Depleted Mice Exhibit Prolonged Cardiac Repolarization and Are Predisposed to Ventricular Arrhythmia.

Bondarenko VE, editor. PLoS ONE. Public Library of Science; 2016;11(2):e0149205.

Contributors: Concept and design: C.A.E., S.K. Acquisition, analysis, or interpretation of data: C.A.E., A.V.K., M.C., D.S., S.G., S.K. Drafting of the manuscript: C.A.E., A.V.K., S.K. Critical revision of the manuscript for important intellectual content: All authors. Administrative, technical, or material support: S.K.

Table 1. Baseline characteristics of participants in UK Biobank.

N Individuals	All Participants (405569)	Type 2 Diabetes Cases (20458)	Type 2 Diabetes Free Controls (385111)
Age ± SD, years	57 ± 8.1	61 ± 6.9	57 ± 8.1
Female, n (%)	218376 (54%)	7801 (38%)	210575 (55%)
UK BiLEVE Array, n (%)	48625 (12.0%)	17361 (15%)	3097 (12%)
Body Mass Index ± SD, kg/m2	27 ± 4.8 kg/m2	32 ± 5.9 kg/m2	27 ± 4.6 kg/m2
Waist-to-Hip Ratio ± SD	0.87 ± 0.09	0.95 ± 0.08	0.87 ± 0.09

Abbreviations: SD, standard deviation;

Table 2. Association of variants in *ACVR1C* with waist-to-hip ratio adjusted for body mass index and with type 2 diabetes. Estimates for WHRadjBMI were derived through linear regression analysis in UK Biobank. Estimates for type 2 diabetes were derived through meta-analysis of UK Biobank and the DIAGRAM ExTexT2D consortium.

Variant	Minor Allele Frequency (%)	*WHRadjBMI*		*Type 2 Diabetes*	
		Beta (CI)	p-value	OR (CI)	p-value
Asn150His	1.1	-0.089 (-0.11, -0.067)	$3.4*10^{-17}$	0.88 (0.83, 0.94)	$8.7*10^{-5}$
Ile195Thr	0.2	-0.15 (-0.09, 0.19)	$1.0*10^{-9}$	0.79 (0.67, 0.93)	0.005
Ile482Val	7.2	-0.019 (-0.01, -0.027)	$1.6*10^{-5}$	0.95 (0.93, 0.97)	$4.8*10^{-6}$
rs72927479	5.1	-0.035 (-0.045, -0.025)	$2.6*10^{-12}$	0.93 (0.89, 0.97)	$6.0*10^{-4}$

Figure 1. Across four *ACVR1C* genetic variants, association of number of WHRadjBMI-lowering alleles with mean directly measured abdominal fat (percent abdominal fat of total body fat) in 4215 participants who underwent dual X-ray absorptiometry scan in UK Biobank, adjusted for age, sex, ten principal components of ancestry and array type. Number of participants in each group is displayed in white for each bar.

Figure 2. Association of four variants in *ACVR1C* with waist-to-hip ratio adjusted for body mass index (WHRadjBMI, x-axis) and type 2 diabetes (T2D, y-axis).

Figure 3. Association of predicted damaging variants in *ACVR1C* with type 2 diabetes from sequences in the Myocardial Infarction Genetics Consortium (MIGen), the Atherosclerosis Risk in Communities study (ARIC) and the Type 2 Diabetes Genes Consortium (T2D Genes).

Figure 4. Association of *ACVR1C* gene risk score with 31 disease phenotypes in UK Biobank.