ARTICLE

Functional characterization of a novel p.Ser76Thr variant in *IGFBP4* that associates with body mass index in American Indians

Yunhua L. Muller[1]✉, Michael Saporito[1], Samantha Day[1], Khushdeep Bandesh[1], Cigdem Koroglu[1], Sayuko Kobes[1], William C. Knowler[1], Robert L. Hanson[1], Regeneron Genetics Center*, Cristopher V. Van Hout[2,3], Alan R. Shuldiner[2], Clifton Bogardus[1] and Leslie J. Baier[1]

Insulin-like growth factor binding protein 4 (IGFBP4) is involved in adipogenesis, and IGFBP4 null mice have decreased body fat through decreased PPAR-γ expression. In the current study, we assessed whether variation in the *IGFBP4* coding region influences body mass index (BMI) in American Indians who are disproportionately affected by obesity. Whole exome sequence data from a population-based sample of 6779 American Indians with longitudinal measures of BMI were used to identify variation in *IGFBP4* that associated with BMI. A novel variant that predicts a p.Ser76Thr in IGFBP4 (Thr-allele frequency = 0.02) was identified which associated with the maximum BMI measured during adulthood (BMI 39.8 kg/m^2 for Thr-allele homozygotes combined with heterozygotes vs. 36.2 kg/m^2 for Ser-allele homozygotes, $\beta = 6.7\%$ per Thr-allele, $p = 8.0 \times 10^{-5}$, adjusted for age, sex, birth-year and the first five genetic principal components) and the maximum age- and sex-adjusted BMI *z*-score measured during childhood/adolescence (*z*-score 0.70 SD for Thr-allele heterozygotes vs. 0.32 SD for Ser-allele homozygotes, $\beta = 0.37$ SD per Thr-allele, $p = 8.8 \times 10^{-6}$). In vitro functional studies showed that IGFBP4 with the Thr-allele (BMI-increasing) had a 55% decrease ($p = 0.0007$) in FOXO-induced transcriptional activity, reflecting increased activation of the PI3K/AKT pathway mediated through increased IGF signaling. Over-expression and knock-down of IGFBP4 in OP9 cells during differentiation showed that IGFBP4 upregulates adipogenesis through PPARγ, CEBPα, AGPAT2 and SREBP1 expression. We propose that this American Indian specific variant in *IGFBP4* affects obesity via an increase of IGF signaling.

European Journal of Human Genetics (2022) 30:1159–1166; https://doi.org/10.1038/s41431-022-01129-3

INTRODUCTION

Insulin-like growth factor (IGF) is an important regulator of adipose tissue development [1, 2]. In an adipose-specific double knockout of the insulin receptor and the IGF receptor, mice had reduced adiposity [3]. It is well established that the activity of IGF is regulated by IGF-binding proteins (IGFBPs). IGFBP4 is highly expressed in adipose tissue, and recent studies have shown that IGFBP4 is required for adipogenesis and influences the distribution of adipose depots in mice [4]. IGFBP4 null mice have decreased adipogenesis and fat expansion. Furthermore, IGFBP4 knockout mice have a blunted IGF-mediated of AKT phosphorylation, suggesting that IGFBP4 exerts its role in adiposity through IGF-mediated PI3K/AKT activation [4].

American Indians living on a reservation in South-Central Arizona have a high prevalence of obesity [5]. Since heritable factors are estimated to explain 40–70% of inter-individual variability in body mass index (BMI) [6], we have previously directly sequenced physiologic candidate genes for obesity in this population, such as those in the leptin-melanocortin pathway; these studies have identified variation in *MC4R, SIM1* and *LEPR* that associated with higher BMI in this population [7–11]. Based on the potential role of IGFBP4 in adiposity, in the current study we sought to determine whether coding variation in *IGFBP4* also contributes to obesity in this longitudinally studied population. We analyzed whole exome sequence data from 6779 American Indians to identify *IGFBP4* variation that associates with BMI in childhood and adulthood; and assessed the potential functionality of associated variants on IGF signaling in vitro. We also assessed the role of IGFBP4 in adipogenesis.

MATERIALS (SUBJECTS) AND METHODS

Participants

Study participants were derived from a longitudinal study of the etiology of type 2 diabetes among American Indians living on a reservation in Arizona. The study protocols were approved by the Institutional Review Board of the National Institute of Diabetes and Digestive and Kidney Diseases. All community residents (age ≥5 years) were invited to participate in outpatient biennial exams, which included measures of height and weight for calculation of

[1]Phoenix Epidemiology and Clinical Research Branch, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Phoenix, AZ, USA. [2]Regeneron Genetics Center, Regeneron Pharmaceuticals, Tarrytown, NY, USA. [3]Laboratorio Internacional de Investigation sobre el Genoma Humano, Campus Juriquilla de la Universidad Nacional Autonoma de Mexico, Queretaro, QRO, Mexico. *A list of authors and their affiliations appears at the end of the paper. ✉email: ymuller@mail.nih.gov

Received: 5 November 2021 Revised: 18 April 2022 Accepted: 26 May 2022
Published online: 10 June 2022

BMI and a 75-g oral glucose tolerance test to determine diabetes according to the criteria of American Diabetes Association [12]. Participants were examined in years 1965–2007, and 6779 participants were included in the exome sequencing study. These participants had provided written consent for DNA sequencing; and informed consent was obtained from all subjects. The consent was approved by the Institutional Review Board of the NIDDK and/or by the Phoenix Area Indian Health Service Institutional Review Board. Among these participants, the median (25th percentile, 75th percentile) number of exams per person was 4 [2, 8]. Participation duration was 13.9 (4.9, 28.2) years. Maximum BMI in adulthood was defined as the highest recorded BMI from an exam at age ≥15 years. Since diabetes duration and treatment can affect bodyweight, BMI was also analyzed as the highest BMI recorded at an exam at age ≥15 years when the subject was non-diabetic as determined by an oral glucose tolerance test. Among 6002 subjects who had a measure of maximum BMI; 5268 subjects also had a measure of maximum BMI recorded at an exam when the subject was non-diabetic (734 individuals who did not have a measure of BMI from a non-diabetic exam were excluded from this analysis). To assess susceptibility to obesity in childhood, we analyzed the maximum age- and sex-adjusted BMI *Z*-score in 4882 American Indian children or adolescents (between the ages of 5–20 years). Median BMI for girls and boys (50th percentile) in this community is approximately at the 95th percentile of the US population at every age [8]. Since the distribution of BMI in these children is very different from that in other populations, we used a population-specific *z*-score which was calculated by subtracting the mean within categories of age (1 year) and sex from BMI and dividing by its standard deviation within categories in all research participants. Given the longitudinal nature of the study, many individuals were analyzed as both children and adults. Thus, the analyses of childhood BMI *z*-score and adulthood BMI were not independent. A subset of non-diabetic subjects was also characterized in our Clinical Research Center for % body fat ($N = 836$) and insulin sensitivity ($N = 488$) as described previously [13].

Replication of the association of the p.Ser76Thr variant with BMI was assessed in 847 individuals who live in urban Phoenix and are at least half American Indian (any tribe) and have a cross sectional measure of BMI.

Whole exome sequencing

Whole exome sequencing (≥20× coverage in >85% of target regions) was performed by the Regeneron Genetics Center (Tarrytown, NY, USA) [9].

Sequence data were aligned to the reference human genome (hg38). Variant calling and joint genotyping were performed using the Genome Analysis Toolkit (GATK 3.5) [14]. A total of 1.6 M high quality, bi-allelic variants passed quality control metrics including: variant call rate ≥90%, no deviation from Hardy–Weinberg equilibrium ($p \geq 1 \times 10^{-4}$) and discrepant rate ≤2.5% pairs among 111 blind duplicate pairs.

Statistical analysis

Statistical analyses were performed using the software of the SAS Institute (Cary, NC, USA). The associations of genotypes with BMI were analyzed by linear regression using a model fitted with a variance components covariance structure to account for genetic relatedness among individuals as previously described [15]. The natural logarithm of BMI or the childhood *z*-score; was taken as the dependent variable. Results were adjusted for age, sex, birth-year and the first five genetic principal components.

In vitro functional analyses of the nonsynonymous p.Ser76Thr variant in the IGF signaling pathway

Expression vectors containing the full-length human wild type *IGFBP4, IGF1R* and *FOXO1A* were obtained from Origene (Rockville, MD, USA). The mutant *IGFBP4* construct of p.Ser76Thr was generated using the QuikChange Lightning site-directed mutagenesis kit (Agilent technologies, Santa Clara, CA, USA) and confirmed by Sanger sequencing.

COS7 and HEK-293 (ATCC, Manassas, VA, USA) were both maintained in Dulbecco's modified Eagle medium (Gibco, Amarillo, TX, USA) containing 4.5 g/l glucose (high glucose optimized cell growth), 1.5 g/l sodium bicarbonate, 10% FBS (containing IGF, insulin and other growth factors, Gibco) and 1% penicillin-streptomycin (Gibco) in 5% CO_2 and 95% air atmosphere at 37 °C. A FOXO reporter assay is designed or well established for measuring PI3K/AKT activity mediated by IGF or insulin signaling (Qiagen, Hilden, Germany). Therefore, this assay was used to assess the effects of the IGFBP4 p.Ser76Thr on the IGF signaling pathway. COS7 cells were co-transfected with 500 ng wild type or mutant *IGFBP4* construct, 500 ng *FOXO1A*, 500 ng *IGF1R*, and 250 ng reporter mixture of FOXO-induced firefly luciferase construct and constitutively expressed Renilla luciferase construct (40:1) (Qiagen) using the transfection reagent lipofectamine LTX (Invitrogen, Waltham, MA, USA). Forty-eight hours post-transfection, cells were harvested, and dual-luciferase activity (Promega, Madison, WI, USA) was measured on the TD-20/20 luminometer (Turner BioSystems, Sunnyvale, CA, USA). Relative luciferase activity was calculated as a ratio of the firefly to renilla luciferase activity to control for transfection efficiency; activity was further normalized to the *IGFBP4* wild type construct. Data from four separate transfections in COS7 were averaged and the difference between wild type and mutant allele was compared using a *t*-test. This transfection study was repeated in HEK293 cells using the same method.

In vitro analyses of the role of IGFBP4 in adipogenesis

To examine the role of IGFBP4 in adipogenesis, OP9 cells (ATCC), which are stromal cells derived from mouse bone marrow that accumulate large triglyceride droplets after only 2–3 days of adipogenic stimuli, were used [16]. To knock-down IGFBP4, OP9 cells were transfected with 100 nM siRNA targeting IGFBP4 (Silencer Select Pre-designed siRNA, Thermo Fisher Scientific, Waltham, MA, USA) or 100 nM Silencer Negative Control #1 siRNA (Thermo Fisher Scientific) using the transfection reagent RNAmax (Invitrogen). In parallel, to over-express IGFBP4, OP9 cells were transfected with 1 ug wild type *IGFBP4* DNA expression vector or a negative control PCMV vector with the Optifect transfection reagent (Invitrogen). Twenty-four hours post-transfection, adipogenesis was induced using Insulin Oleate media (MEM-α [Gibco], 0.2% FBS, 175 nM insulin [Sigma, St. Louis, MO, USA], 900 μM oleate:albumin [Sigma] and 500 U penicillin-streptomycin). Forty-eight hours post-induction, RNA was extracted from OP9 cells knocking-down or over-expressing IGFBP4 using the RNAeasy kit (Qiagen). cDNA was generated using the advantage RT for PCR kit (Takara, Kusatsu, Shiga, Japan). Quantitative PCR assays ($N = 4$) were performed to assess the efficiency of knocking-down and over-expressing IGFBP4; this assay was also used to measure expression of ten known adipogenic marker genes [16–18]: *PPARγ*, *C/EBPα*, *AGPAT2*, *SREBP1*, *GPAM*, *FABP4*, *PLIN1*, *FASN*, *LPL* and *ADIPOQ* in response to the knock-down and over-expression of IGFBP4. In parallel, 48 h post-induction, cells were lysed for measurement of the lipid content using a Triglyceride Assay kit according to the manufacturer's protocol (Biovision, Milpitas, CA, USA). The triglyceride levels were normalized to the total protein content and data from five separate experiments were averaged. A *t*-test was used to determine the significant difference in triglyceride levels between cells over-expressing and knocking-down IGFBP4.

In silico analyses of the structural effects of the p.Ser76Thr variant

The crystal structure of IGFBP4 was downloaded as a Protein Data Bank (PDB) file from the RCSB PDB (https://www.rcsb.org) (PDB ID: 1WQJ) [19]. In silico mutagenesis of the wild type or mutant protein structure at p.Ser76Thr was performed using the PyMOL

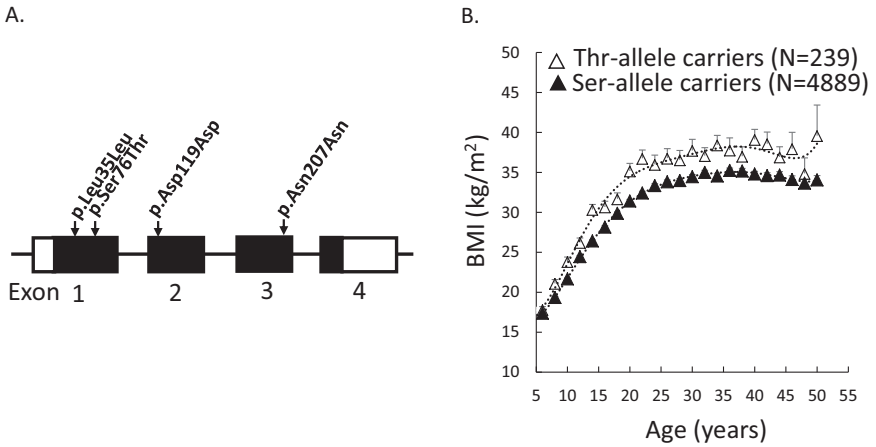

Fig. 1 The p.Ser76Thr variant in *IGFBP4* associated with lifetime BMI in American Indians. A Location of four coding variants in the *IGFBP4* gene. The black boxes represent the coding exons and white boxes represent the untranslated regions (UTR). **B** Effect of p.Ser76Thr in *IGFBP4* on lifetime BMI trajectory (age 5–50 years). Mean BMI for the Thr-allele carriers ($N = 239$ including 1 Thr/Thr carrier) was compared to the Ser/Ser carriers ($N = 4889$). BMI is given as an unadjusted mean ± SEM at each age group at 2 year intervals by genotype. Subjects with at least 2 BMI exams (median number of exams = 4) were included in the analyses.

Molecular Graphics System (Version 1.2r3pre, Schrödinger, Inc., New York, NY, USA). Energy minimization was carried out by the steepest descent (100 steps) and conjugate gradient (500 steps) method in the SWISS-PdbViewer [20], and wild type and mutant protein structures were compared. Protein stability was calculated using Site Directed Mutator, University of Cambridge [21] and CUPSAT, Cologne University [22].

RESULTS

Association of a novel p.Ser76Thr variant in IGFBP4 with BMI in American Indians

Four variants identified in whole exome sequence data across the *IGFBP4* coding region are shown in Fig. 1A. One of these variants, rs3403704 which predicts a synonymous Asn207Asn, was rare (3 carriers) and thus was excluded from single SNP statistical analyses due to inadequate power. Among the other three variants (Table 1), one predicted a Serine to a Threonine at codon 76 (Ser76Thr) and had a Combined Annotation Dependent Depletion (CADD) score = 9.7 (a CADD of ten equates to the top 10% most likely to be deleterious [23]). This novel variant had a minor allele frequency (mAF) of 0.02 in our American Indian cohort and has never been detected in non-American Indian populations. The other two variants (rs79982277 and rs598892) had been previously identified in other ethnic groups and both predict synonymous changes (Asp119Asp and Leu35Leu, respectively).

The Ser76Thr variant associated with maximum BMI in adult American Indians where the Thr-allele carriers ($N = 263$) had a mean BMI 39.8 kg/m^2 vs. 36.2 kg/m^2 for the Ser/Ser carriers ($N = 5739$) (Table 1, $\beta = 6.7\%$ per Thr-allele, $p = 8.0 \times 10^{-5}$ in an additive model; $\beta = 7.3\%$ per Thr-allele, $p = 4.6 \times 10^{-5}$ in a dominant model, adjusted for age, sex, birth-year and the first five genetic principal components). The Ser76Thr also associated with the maximum BMI prior to the development of diabetes, where the Thr-allele carriers had a mean BMI 38.9 kg/m^2 vs. 35.2 kg/m^2 for the Ser/Ser carriers (Table 1, $\beta = 7.2\%$ per Thr-allele, $p = 4.6 \times 10^{-5}$ in an additive model; $\beta = 7.3\%$ per Thr-allele, $p = 2.1 \times 10^{-5}$ in a dominant model). Furthermore, this variant associated with the age-, sex-specific *z*-score during childhood/adolescence, where the Ser76Thr heterozygotes ($N = 225$) had BMI *z*-score of 0.7 SD compared to 0.32 SD for the Ser76Ser homozygotes ($N = 4657$) (Table 1, $\beta = 0.37$ SD per Thr-allele, $p = 8.8 \times 10^{-6}$). The Δ*z*-score between carriers and non-carriers of the Thr-allele was 0.38, corresponding to a ΔBMI of 2.3 kg/m^2. There was no significant interaction between genotype and sex.

The association of p.Ser76Thr with BMI replicated among 847 American Indians who live in urban Phoenix. Eight individuals heterozygous for Ser76Thr had a mean BMI of 40.2 kg/m^2 as compared to 32.3 kg/m^2 for individuals homozygous for the Ser-allele ($\beta = 18.0$ % per Thr-allele, $p = 0.05$; Table 1).

Since p.Ser76Thr associated with BMI in both childhood and adulthood, the change in BMI between the ages of 5–50 years, at 2-year intervals, was evaluated. Among 5128 subjects who had at least two exams where BMI was measured (the median [25th percentile, 75th percentile] number of exams for these subjects = 4 [3, 6]) before developing diabetes; 14% had only childhood exams, 24% had only adult exams, and 62% had exams in both childhood and adulthood. The lifetime BMI trajectory of the Thr-allele carriers ($N = 239$) vs. the Ser/Ser carriers ($N = 4889$) is shown in Fig. 1B. Subjects carrying a Thr-allele had consistently higher BMI than subjects homozygous for Ser-allele at almost every age.

Most subjects had additional anthropometric measures available from the same visit when the maximum BMI was measured, and data for these are shown in Table 2. Individuals with Thr/Ser, when compared to Ser/Ser, had a greater waist circumference (122.3 vs. 110.9 cm), greater hip circumference (125.3 vs. 118.4 cm), and greater thigh circumference (67.2 vs. 64.6 cm). The waist-hip and waist-thigh ratios were also greater in the Thr/Ser carriers. All of the measurements which included the waist circumference (waist, waist to hip, and waist to thigh) remained significant after adjustment of BMI indicative of central obesity. A much smaller subset of individuals ($N = 836$) had been analyzed for % body fat. The Thr/Ser carriers had a higher % body fat as compared to the Ser/Ser carriers (35.5% vs. 33.6%), but this difference did not reach statistical significance.

The p.Ser76Thr variant was also analyzed for association with various measures of glucose homeostasis (Table 2). Among 4887 individuals there was a slight trend for those with Thr/Ser to have higher glucose levels at fasting and at 2-h during an OGTT. However, among 488 non-diabetic American Indians who had insulin sensitivity assessed by a hyperinsulinemic–euglycemic clamp to measure rates of insulin-stimulated glucose disappearance, the p.Ser76Thr did not associate with the rate of glucose disposal under physiological concentrations of insulin stimulation.

Table 1. Associations of the coding variants in *IGFBP4* with maximum adult BMI at any exam ($N = 6002$), maximum BMI at an exam when the subject was non-diabetic ($N = 5268$) and childhood BMI *z*-score ($N = 4882$) in American Indians.

Coding variant	Ref /Alt	Alt Allele Freq	Trait	Mean BMI Alt (*N*)	Mean BMI Het (*N*)	Mean BMI Ref (*N*)	Beta	*p*	CADD
Chr17:40443961 c.T226A p.Ser76Thr	T/A	0.02	MaxBMI	33.0[a] (2)	39.8 (261)	36.2 (5739)	6.7%	8.0×10^{-5}	9.7
			MaxBMI non-db	33.0[a] (2)	38.9 (227)	35.2 (5039)	7.2%	4.6×10^{-5}	
			MaxBMI *z*-score	– (0)	0.70 (225)	0.32 (4657)	0.37 SD	8.8×10^{-6}	
		0.005	Replication cohort BMI[b]	–	40.2 (8)	32.3 (839)	18.0%	0.05	
rs79982277 c.C357T p.Asp119Asp	C/T	0.002	MaxBMI	– (0)	25.1 (8)	36.4 (5994)	−17.0%	0.02	–
			MaxBMI non-db	– (0)	24.0 (7)	35.4 (5261)	−19.0%	0.01	
			MaxBMI *z*-score	– (0)	−0.11 (10)	0.34 (4872)	−0.31 SD	0.35	
rs598892 c.G105A p.Leu35Leu	G/A	0.78	MaxBMI	36.8 (3902)	35.6 (1857)	35.5 (243)	0.4%	0.55	–
			MaxBMI non-db	35.7 (3424)	34.7 (1629)	34.1 (215)	0.6%	0.31	
			MaxBMI *z*-score	0.36 (3226)	0.31 (1460)	0.24 (196)	0.02 SD	0.54	

Data are given for the variants in *IGFBP4* associated with maximum BMI in adulthood (MaxBMI), maximum BMI before individuals developed diabetes (MaxBMI non-db) and maximum BMI *z*-score in childhood (MaxBMI *z*-score) in all American Indians. BMI is $\log_e$-transformed before analyses to approximate a normal distribution, and the regression coefficient (β) was exponentiated to obtain the effect estimate for each risk allele, expressed as a multiplier. For presentation in the table, a multiplier was converted to the effect size in % increase per allele. Beta for childhood BMI (maximum) was expressed as *z*-score (SD unit) per copy of the risk. Mean BMI is the unadjusted raw value. Beta and *p* (using an additive model) were adjusted for age, sex, birth-year and the first five genetic principal components.

[a]These two individuals had MaxBMIs of 25.9 and 40 kg/m^2.

[b]Replication cohort described in Methods.

Functional analyses of the IGFBP4 p.Ser76Thr variant in altering PI3K/AKT activity through IGF signaling

The p.Ser76Thr had a CADD of 9.7, predicting that it may impact the function of IGFBP4. To test this, an in vitro FOXO-reporter assay which detects PI3K/AKT mediated transcriptional activity was used to assess the effect of the p.Ser76Thr on IGF signaling (Fig. 2). Cells were cultured in media with 10% fetal bovine serum containing IGF, insulin or other growth factors to stimulate IGF signaling [24]. COS7 cells expressing the mutant IGFBP4 (Thr at codon 76, BMI-risk allele) showed a 55% decrease (Fig. 2A, $N = 4$ experiments, $p = 0.0007$) in FOXO1-induced transcriptional activity as compared with cells expressing the wild type IGFBP4 (Ser at codon 76, BMI non-risk allele). This decrease in FOXO1-induced transcriptional activity reflected increased FOXO1 nuclear exclusion mediated through increased activation of the PI3K/AKT by IGF signaling [25, 26]. Therefore, our data implied that cells expressing the BMI-risk Thr-allele may have higher IGF signaling than those expressing the non-risk Ser-allele. When compared to the PCMV empty vector, cells expressing the wild type IGFBP4 had a 39% decrease in FOXO1-induced transcriptional activity. These results suggested that the wild type IGFBP4 might serve as an activator of the PI3K/AKT pathway, and the mutant IGFBP4 further activated this pathway. This luciferase reporter assay was repeated in HEK293 cells. Similarly, but to a less extent, cells expressing the BMI-risk Thr-allele had a 13.5 % decrease compared to the non-risk Ser-allele (Fig. 2B, $N = 4$ experiments, $p = 0.02$) in FOXO1-induced transcriptional activity. Our data support the finding that IGFBP4 exerts its role in adiposity through IGF-mediated PI3K/AKT activation in mice [4]. The mutant IGFBP4 may further gain activity in the IGF-mediated PI3K/AKT pathway which is consistent with that high IGF signaling leads to obesity [27].

Prediction of the structural effects of the p.Ser76Thr variant

A partial X-ray crystal structure of the IGFBP4 protein consisting of amino acid residues 24 to 103 has previously been determined at a good resolution of 1.60 Å (PDB id: 1WQJ). In silico protein analyses predicted very similar structures between the wild type Ser76 and the mutant Thr76 (root mean square deviation = 0.002 Å). Using the PyMol tool, a polar contact was detected between the wild type Ser76 and Glu46 (Supplementary Fig. 1). However, the mutant Thr76 generated an additional Hydrogen (H)-bond with Glu46 and a new H-bond with the neighboring residue Gly77 (Supplementary Fig. 1). Protein stability was further calculated by the change of free energy ($\Delta\Delta G$); and the Thr-allele had a moderate increase as compared with the Ser-allele (predicted a positive $\Delta\Delta G = 0.08$ kcal/mol by Site Directed Mutator, $\Delta\Delta G =$ 2.11 kcal/mol by CUPSAT).

Roles of IGFBP4 in adipogenesis

To further evaluate the role of IGFBP4 in adipogenesis, we assessed gene expression levels for adipogenic markers during differentiation in OP9 cells over-expressing or knocking-down IGFBP4. Forty-eight hours after cell differentiation, OP9 cells over-expressing IGFBP4 had an average 635-fold increase ($N = 4$) in the *IGFBP4* RNA levels compared to the negative control (PCMV empty vector); whereas the *IGFBP4* RNA expression levels had an average 95% suppression in cells transfected with siRNA to knock-down IGFBP4 ($N = 4$). Quantitative RT-PCR was performed on ten known adipogenic marker genes [16–18]: *PPARγ*, *C/EBPα*, *AGPAT2*, *SREBP1*, *GPAM*, *FABP4*, *PLIN1*, *FASN*, *LPL* and *ADIPOQ*. RNA expression levels of these adipogenic markers in OP9 cells without transfection were also compared before and after differentiation, and all genes were upregulated after differentiation (Fig. 3A, $p \le 0.05$). Genes with the highest fold change (>5 fold) were *PPARγ*, *AGPAT2* and *FASN*. Endogenous *IGFBP4* was highly expressed but only upregulated by 1.2 folds during differentiation (Fig. 3A, $p =$ 0.02). During differentiation, in OP9 cells over-expressing IGFBP4, RNA expression levels of *PPARγ*, *C/EBPα*, *AGPAT2* and *SREBP1* were

Table 2. Associations of the Ser76Thr variant in *IGFBP4* with other anthropometric and glucose homeostasis traits in American Indians.

Trait	Ser/Thr Mean ± SD (*N*)	Ser/Ser Mean ± SD (*N*)	Beta ± SE	*p*	*p* adjusted for BMI
Waist (cm)	122.3 ± 62.2 (224)	110.9 ± 23.1 (5128)	8.2 ± 1.9	1.6×10^{-5}	0.008
Hip (cm)	125.3 ± 19.4 (216)	118.4 ± 18.4 (4990)	4.9 ± 1.4	2.4×10^{-4}	0.89
Thigh (cm)	67.2 ± 9.5 (224)	64.6 ± 9.8 (5127)	2.0 ± 0.7	0.005	0.64
Waist/Hip	0.98 ± 0.54 (216)	0.94 ± 0.19 (4990)	0.03 ± 0.02	0.02	0.04
Waist/Thigh	1.87 ± 1.57 (224)	1.73 ± 0.39 (5127)	0.10 ± 0.03	0.003	0.02
Body fat (%)	35.5 ± 7.5 (35)	33.6 ± 8.3 (801)	0.8 ± 1.1	0.46	
Fasting glucose (mg/dl)	96.8 ± 10.7 (210[a])	94.3 ± 10.4 (4677)	1.6 ± 0.8	0.04	
2-h glucose (mg/dl)	122.7 ± 36.1 (210[a])	115.5 ± 32.5 (4677)	4.3 ± 2.4	0.07	
Rate of glucose disposal (log_{10}mg [kgEMBS]$^{-1}$ min^{-1})	0.54 ± 0.10 (27)	0.56 ± 0.13 (461)	−0.003 ± 0.02	0.89	

Data are presented as unadjusted Mean ± SD. *N* is the number of the subjects. The *p* values and betas for waist, hip, thing, waist/hip and waist/thigh were adjusted for age, sex, birth-year and the first five genetic principal components. These traits were then additionally adjusted for BMI. The *p* value and beta for body fat (%) were adjusted for age, sex and the first five genetic principal components. The *p* values and betas for fasting and 2-h glucose levels in response to an OGTT were adjusted for age, sex and the first five genetic principal components. The *p* value and beta for the rate of glucose disposal under insulin stimulation were adjusted for age, sex, % body fat, (% body fat)^2 and the first five genetic principal components.
EMBS estimated metabolic body size.
[a]One Thr/Thr homozygote had this measurement and was included with the 209 heterozygotes for analysis.

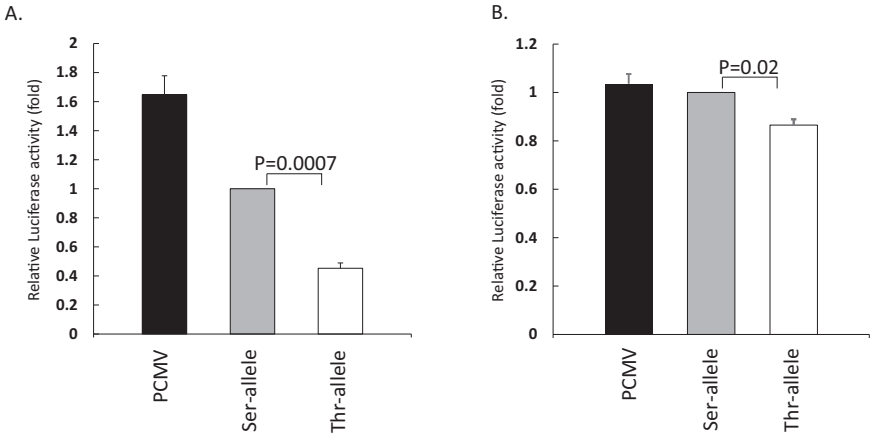

Fig. 2 The p.Ser76Thr variant in *IGFBP4* altered FOXO-induced transcription through PI3K/AKT signaling. The effect of p.Ser76Thr in *IGFBP4* on FOXO-induced transcription through PI3K/AKT signaling in COS7 cells (**A**) and HEK293 cells (**B**) as assessed using an in vitro luciferase activity assay. Each bar represents the averaged luciferase activity (mean ± SD) of cells expressing the wild type IGFBP4 Ser-allele, mutant Thr-allele and negative control PCMV vector from independent experiments ($N = 4$). Statistical differences in activity were determined using a *t*-test.

upregulated compared to cells transfected with the negative control vector PCMV6 (Fig. 3B, gray bar: over-expression; black bar: negative control, 1.2–1.6 fold, $p < 0.05$); whereas in cells knocking-down IGFBP4, expression levels of these same genes: *PPARγ*, *C/EBPα*, *AGPAT2* and *SREBP1* were down-regulated compared to cells transfected with negative control siRNA (Fig. 3B, 0.6–0.8 fold, $p < 0.05$). These results indicate that IGFBP4 upregulates RNA expression of key regulators *PPARγ*, *C/EBPα*, *AGPAT2* and *SREBP1* during OP9 adipogenesis. In contrast, no changes in RNA expression levels were detected for the other markers, namely *GPAM, FABP4, PLIN1, FASN, LPL* and *ADIPOQ* (Fig. 3B, $p > 0.05$).

Oil Red-O staining was also performed in differentiated OP9 cells over-expressing or knocking-down IGFBP4. An increase in Oil Red-O staining was observed in cells over-expressing IGFBP4 as compared to PCMV negative control; whereas reduced Oil Red-O staining was observed in cells knocking-down IGFBP4 as compared to the negative control siRNA (Fig. 4A). Lipid content was also assessed in differentiated OP9 cells over-expressing or knocking-down IGFBP4. In five separate experiments, cells over-expressing IGFBP4 had an increased triglyceride content (18.9%) as compared to the negative control; whereas cells knocking-down IGFBP4 had a decreased triglyceride content (9.6%) as compared to the negative control. The difference between over-expressing and knocking-down IGFBP4 was 28.5% (Fig. 4B, $p = 0.05$). There was a marked 25-fold increase in triglyceride content when preadipocytes differentiated to adipocytes (Fig. 4B, $p = 0.0001$).

DISCUSSION

In this candidate gene study of *IGFBP4* in American Indians, we identified a novel (not reported in any public database), uncommon (mAF = 0.02) p.Ser76Thr variant in *IGFBP4* that significantly associated with BMI in adulthood ($p = 8.0 \times 10^{-5}$)

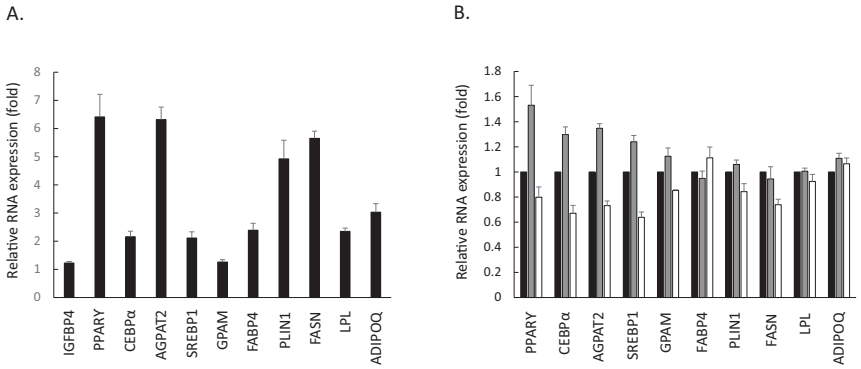

Fig. 3 Over-expressing and knocking-down IGFBP4 during OP9 cell differentiation affected RNA expression of adipogenic markers. **A** RNA expression of adipogenic markers in OP9 cells (non-transfected) during differentiation. Relative RNA expression is presented as fold change after and before OP9 cell differentiation. **B** RNA expression of adipogenic markers in OP9 cells over-expressing and knocking-down IGFBP4 during differentiation. Relative RNA expression is presented as fold change to the negative control. Gray bar: over-expression; white bar: knock-down; black bar: negative control. $N = 4$ for repeated experiments.

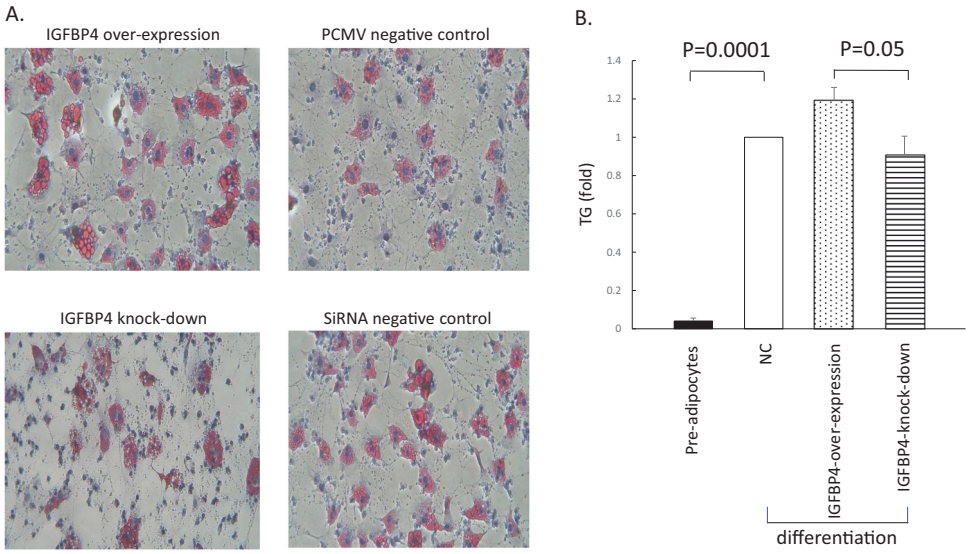

Fig. 4 Effects of IGFBP4 on adipogenesis. **A** After differentiation, OP9 cells over-expressing and knocking-down IGFBP4, and negative controls, were stained with Oil Red-O. **B** Triglyceride content (fold) in OP9 cells over-expressing and knocking-down IGFBP4 was normalized to the negative control. Preadipocytes were undifferentiated OP9 cells. $N = 5$ for repeated experiments.

and BMI *z*-score in childhood ($p = 8.9 \times 10^{-6}$). A lifetime BMI trajectory (age 5–50 years) showed that Thr-allele carriers have a consistently higher BMI at almost all ages as compared to Ser-allele carriers. At most adult ages, the difference was ~3 kg/m^2. Carriers of the Thr-allele also had higher waist circumference ($p = 1.6 \times 10^{-5}$), hip circumference ($p = 2.4 \times 10^{-4}$), thigh circumference ($p = 0.005$), waist-hip ratio ($p = 0.03$) and waist-thigh ratio ($p = 0.002$). These differences remained significant even after adjusting for BMI which is consistent with central obesity. However, the Thr-allele carriers had only a nominal increase in fasting glucose levels and had no evidence of reduced insulin sensitivity based on a hyperinsulinemic, euglycemic clamp.

IGFBP4 is required for adipogenesis and influences the distribution of adipose depots [4]. IGFBP4 null mice (IGFBP4$^{-/-}$) had decreased adiposity at 2–4 months of age. Inguinal and gonadal white adipose tissues from IGFBP4$^{-/-}$ mice had decreased weight and PPARγ expression [4]. There were sexual differences in adipocyte expansion of IGFBP4$^{-/-}$ mice under a high-fat-diet. IGFBP4$^{-/-}$ female mice were resistant to diet-induced weight gain. IGFBP4 is required for fat expansion in female mice but not in male mice [4]. This sex dimorphism during fat expansion is thought due to the sex hormone estrogen [4]. In our study, the effect size of the BMI association was 2.46 kg/m^2 per risk allele in females and 1.99 kg/m^2 per risk allele in males, however, an interaction of genotype by sex did not achieve statistical significance.

IGF plays an important role in adipose tissue development. Various actions of IGF are regulated by IGFBP in a tissue specific manner. IGFBP4 is highly expressed in adipose tissue, cervix, ovary and uterus [28]. The levels of IGF1, IGF2 and IGFBP4 were markedly increased when the fat explants were treated with growth hormone in vitro [28]. It is known that IGFBP4 regulates the binding of IGF with IGF1R, thus stimulating the adipogenesis and cell growth through PI3K/AKT activation. In IGFBP4$^{-/-}$ mice, the increase in phosphorylated AKT, a downstream target of IGF signaling was blunted in cultured primary ear mesenchymal stem cells during adipogenesis [4]. Our in vitro luciferase assay indicated that the Thr-allele at p.Ser76Thr, which associated with an increase in BMI, had a decrease in FOXO-induced transcriptional activity, presumably due to increased FOXO1 nuclear exclusion resulted from increased activation of the PI3K/AKT pathway. This increased PI3K/AKT reflected an increase in IGF activity which may be regulated by IGFBP4. The BMI-increasing

Thr-allele may gain activity in the IGF-mediated PI3K/AKT pathway that is consistent with the stimulatory role of IGF signaling in development of obesity in humans [27].

Our in silico analyses of the structural effects of p.Ser76Thr in *IGFBP4* predicted a stronger H-bonding for the Thr-allele as compared to the Ser-allele which may lead to a moderate increase in the protein stability. Therefore, we propose that the gain of function of IGFBP4 containing the Thr-allele observed in our in vitro functional studies may be the consequence of increased protein stability via stronger H-bonding. We further propose that the gain of function of IGFBP4 upregulates IGF activity, thereby increasing PI3K/AKT activity, and increasing adipogenesis and obesity. While this variant may affect adipogenesis which could lead to obesity, it is not yet clear whether additional mechanisms were involved such as it could lead to altered energy balance through adipokine(s) or fatty acid metabolisms thereby altered body energy stores/obesity; the same is true for mechanisms by which the Pro12Ala variant in PPARγ associated with obesity [29, 30].

A prior study in mice has shown that IGFBP4 plays a role in adipogenesis and regulates adipose PPARγ expression. However, the involvement of other adipogenic markers is still unclear. In the present study, we used stromal cells (OP9) derived from mouse bone marrow that accumulate large triglyceride droplets after only 2–3 days of adipogenic stimuli. OP9 cells over-expressing and knocking-down IGFBP4 had opposite effects on PPARγ, C/EBPα, AGPAT2 and SREBP1 RNA expression levels, indicating that IGFBP4 upregulated the expression of PPARγ, C/EBPα, AGPAT2 and SREBP1 during adipogenesis. While C/EBPα and PPARγ play pivotal roles in adipogenesis, AGPAT2 (1-Acylglycerol-3-Phosphate O-Acyltransferase 2) regulates lipid metabolism via activation of PI3K/AKT and PPARγ signaling in early stages of adipogenesis. The transcription factor SREBP1 (sterol-regulatory-element-binding protein 1) plays a central role in adipocyte development by inducing expression of several genes critical for lipid biosynthesis including lipoprotein lipase, fatty acid synthase and glycerol phosphate acyltransferase [31].

Lipid accumulation in OP9 cells over-expressing or knocking-down IGFBP4 was assessed by measuring triglyceride levels; this allowed us to evaluate the direct effect of IGFBP4 on adipogenesis. Consistent with a lipogenic role of IGFBP4, our results showed a moderate increase in lipid accumulation in cells over-expressing IGFBP4, and a moderate decrease in lipid accumulation in cells knocking-down IGFBP4. However, cells expressing the mutant Thr-allele in IGFBP4 only had 7% increase in triglyceride content as compared to cells expressing wild type IGFBP4 and did not achieve a significant difference ($N = 6$, $p = 0.14$).

In summary, we identified an American Indian specific p.Ser76Thr variant in *IGFBP4* that may increase obesity risk in this population. Carriers of the Thr-allele had higher mean BMIs as compared to non-carriers across all ages (assessed in individuals 5–50 years). Our in vitro studies support that the mechanism whereby this variant affects BMI is likely via an effect on IGF signaling during adipogenesis.

DATA AVAILABILITY

Materials are available upon request. Individual-level data are not publicly available due to privacy concerns, but may be made available upon reasonable request—see dbGAP accession number phs002490.v1.p1 for details.

CODE AVAILABILITY

Analyses were conducted using available software applications, as described in the Methods.

REFERENCES

1. Laron Z. Lessons from 50 years of study of Laron syndrome. Endocr Pr. 2015;21:1395–402.
2. Janecka A, Kołodziej-Rzepa M, Biesaga B. Clinical and molecular features of laron syndrome, a genetic disorder protecting from cancer. Vivo. 2016;30:375–81.
3. Boucher J, Mori MA, Lee KY, Smyth G, Liew CW, Macotela Y, et al. Impaired thermogenesis and adipose tissue development in mice with fat-specific disruption of insulin and IGF-1 signalling. Nat Commun. 2012;3:902.
4. Maridas DE, DeMambro VE, Le PT, Mohan S, Rosen CJ. IGFBP4 is required for adipogenesis and influences the distribution of adipose depots. Endocrinology. 2017;158:3488–500.
5. Knowler WC, Pettitt DJ, Saad MF, Bennett PH. Diabetes mellitus in the Pima Indians: incidence, risk factors and pathogenesis. Diabetes Metab Rev. 1990;6:1–27.
6. Barsh GS, Farooqi IS, O'Rahilly S. Genetics of body-weight regulation. Nature. 2000;404:644–51.
7. Muller YL, Thearle MS, Piaggi P, Hanson RL, Hoffman D, Gene B, et al. Common genetic variation in and near the melanocortin 4 receptor gene (MC4R) is associated with body mass index in American Indian adults and children. Hum Genet. 2014;133:1431–41.
8. Thearle MS, Muller YL, Hanson RL, Mullins M, Abdussamad M, Tran J, et al. Greater impact of melanocortin-4 receptor deficiency on rates of growth and risk of type 2 diabetes during childhood compared with adulthood in Pima Indians. Diabetes. 2012;61:250–7.
9. Kim HI, Ye B, Gosalia N, Regeneron Genetics Center, Köroğlu C, Hanson RL, et al. Characterization of exome variants and their metabolic impact in 6,716 American Indians from the Southwest US. Am J Hum Genet. 2020;107:251–64.
10. Traurig M, Mack J, Hanson RL, Ghoussaini M, Meyre D, Knowler WC, et al. Common variation in SIM1 is reproducibly associated with BMI in Pima Indians. Diabetes. 2009;58:1682–9.
11. Traurig MT, Perez JM, Ma L, Bian L, Kobes S, Hanson RL, et al. Variants in the LEPR gene are nominally associated with higher BMI and lower 24-h energy expenditure in Pima Indians. Obesity. 2012;20:2426–30.
12. The Expert Committee on the Diagnosis and Classification of Diabetes Mellitus. The Expert Committee on the Diagnosis and Classification of Diabetes Mellitus: report of the Expert Committee on the diagnosis and classification of diabetes mellitus. Diabetes Care. 1997;20:1183–97.
13. Lillioja S, Mott DM, Spraul M, Ferraro R, Foley JE, Ravussin E, et al. Insulin resistance and insulin secretory dysfunction as precursors of non-insulin-dependent diabetes mellitus. Prospective studies of Pima Indians. N Engl J Med. 1993;329:1988–92.
14. McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A, et al. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res. 2010;20:1297–303.
15. Hsueh WC, Nair AK, Kobes S, Chen P, Göring HHH, Pollin TI, et al. Identity-by-descent mapping identifies major locus for serum triglycerides in Amerindians largely explained by an APOC3 founder mutation. Circ Cardiovasc Genet. 2017;10:e001809.
16. Lane JM, Doyle JR, Fortin JP, Kopin AS, Ordovás JM. Development of an OP9 derived cell line as a robust model to rapidly study adipocyte differentiation. PLoS ONE. 2014;9:e112123.
17. Albert JS, Yerges-Armstrong LM, Horenstein RB, Pollin TI, Sreenivasan UT, Chai S, et al. Null mutation in hormone-sensitive lipase gene and risk of type 2 diabetes. N Engl J Med. 2014;370:2307–15.
18. Cautivo KM, Lizama CO, Tapia PJ, Agarwal AK, Garg A, Horton JD, et al. AGPAT2 is essential for postnatal development and maintenance of white and brown adipose tissue. Mol Metab. 2016;5:491–505.
19. Siwanowicz I, Popowicz GM, Wisniewska M, Huber R, Kuenkele KP, Lang K, et al. Structural basis for the regulation of insulin-like growth factors by IGF binding proteins. Structure. 2005;13:155–67.
20. Guex N, Peitsch MC. SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling. Electrophoresis. 1997;18:2714–23.
21. Pandurangan AP, Ochoa-Montaño B, Ascher DB, Blundell TL. SDM: a server for predicting effects of mutations on protein stability. Nucleic Acids Res. 2017;45:W229–35.
22. Parthiban V, Gromiha MM, Schomburg D. CUPSAT: prediction of protein stability upon point mutations. Nucleic Acids Res. 2006;34:W239–42.
23. Kircher M, Witten DM, Jain P, O'Roak BJ, Cooper GM, Shendure J. A general framework for estimating the relative pathogenicity of human genetic variants. Nat Genet. 2014;46:310–5.
24. Burke TW, Vuk-Pavlović S. Insulin-like growth factor-I is a serum component stimulating growth of human neuroblastoma. Vitr Cell Dev Biol Anim. 1993;29:391–4.
25. Guo S. Molecular basis of insulin resistance: the role of IRS and Foxo1 in the control of diabetes mellitus and its complications. Drug Discov Today Dis Mech. 2013;10:e27–33.

26. Moll L, Schubert M. The role of insulin and insulin-like growth factor-1/FoxO-mediated transcription for the pathogenesis of obesity-associated dementia. Curr Gerontol Geriatr Res. 2012. https://doi.org/10.1155/2012/384094.
27. Berryman DE, Glad CAM, List EO, Johannsson G. The GH/IGF-1 axis in obesity: pathophysiology and therapeutic considerations. Nat Rev Endocrinol.2013; 9:346–56.
28. Gude MF, Hjortebjerg R, Oxvig C, Thyø AA, Magnusson NE, Bjerre M, et al. PAPP-A, IGFBP-4 and IGF-II are secreted by human adipose tissue cultures in a depot-specific manner. Eur J Endocrinol. 2016;175:509–19.
29. Yao Y, Li J, Jin Y, Chen Y, He L. Association between PPAR-γ2 Pro12Ala polymorphism and obesity: a meta-analysis. Mol Biol Rep. 2015;42:1029–38.
30. Mansoori A, Amini M, Kolahdooz F, Seyedrezazadeh E. Obesity and Pro12Ala polymorphism of peroxisome proliferator-activated receptor-gamma gene in healthy adults: a systematic review and meta-analysis. Ann Nutr Metab. 2015;67:104–18.
31. Moseti D, Regassa A, Kim W. Molecular regulation of adipogenesis and potential anti-adipogenic bioactive molecules. Int J Mol Sci. 2016;17:124.

ACKNOWLEDGEMENTS

We thank the study participants as well as the clinical staff of the Phoenix Epidemiology and Clinical Research Branch for collecting phenotypes for the study. The opinions expressed in this paper are those of the authors, and do not necessarily reflect the views of the Indian Health Service.

AUTHOR CONTRIBUTIONS

YLM and LJB contributed to the study design. YLM, MS, SD, KB, WCK, CVVH, Regeneron Genetics Center, ARS, RLH, CB and LJB contributed to the data acquisition. CK, SK, WCK, CVVH and RLH contributed to the data analysis. All authors contributed to data interpretation and manuscript drafting; and approved the final version.

FUNDING

This work was supported by the Intramural Research Programme of National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health.

COMPETING INTERESTS

The authors declare that there is no conflict of interest. ARS or CVVH is or was full-time employees of the Regeneron Genetics Center from Regeneron Pharmaceuticals Inc. and receive or received stock options and restricted stock units as compensation.

ETHICAL APPROVAL

These participants had provided written consent for DNA sequencing; and informed consent was obtained from all subjects. The consent was approved by the Institutional Review Board of the NIDDK and/or by the Phoenix Area Indian Health Service Institutional Review Board.

ADDITIONAL INFORMATION

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41431-022-01129-3.

Correspondence and requests for materials should be addressed to Yunhua L. Muller.

Reprints and permission information is available at http://www.nature.com/reprints

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

REGENERON GENETICS CENTER

Cristopher V. Van Hout[2,3] and Alan R. Shuldiner[2]

Lists of members and their affiliations appears in the Supplementary Information.