Preprints.org

Article

Contribution of Circulating Mfge8 to Human T2DM and Cardiovascular Disease

Madhusmita Rout , Megan W. Malone-Perez , Gilseung Park , Megan R. Lerner , J. Kimble Frazer , Blair Apple , April Vaughan , Marvin Payton , Stavros Stavrakis , Evgeny V. Sidorov , Kar-Ming A. Fung ,

Dharambir K. Sanghera *

Article

Contribution of Circulating Mfge8 to Human T2DM and Cardiovascular Disease

Madhusmita Rout [1], Megan W. Malone-Perez [2], Gilseung Park [2], Megan Lerner [3], J. Kimble Frazer [2], Blair Apple [4], April Vaughn [4], Marvin Payton [3,#], Stavros Stavrakis [6], Evgeny Sidorov [4], KarMing A. Fung [5] and Dharambir K. Sanghera [1,7,8,9,10,*]

[1] Department of Pediatrics, Section of Genetics, College of Medicine, University of Oklahoma Health Sciences Center, Oklahoma City, OK 73104, USA

[2] Department of Pediatrics, Section of Hematology and Oncology, College of Medicine, University of Oklahoma Health Sciences Center, Oklahoma City, OK 73104, USA

[3] Department of Surgery, Oklahoma University of Health Sciences Center, Oklahoma City, OK, USA

[4] Department of Neurology, College of Medicine, University of Oklahoma Health Sciences Center, Oklahoma City, OK 73104, USA

[5] Department of Pathology, University of Oklahoma Health Sciences Center, Oklahoma City, OK, USA

[6] Department of Cardiology, University of Oklahoma Health Sciences Center, Oklahoma City, OK, USA

[7] Harold Hamm Diabetes Center, University of Oklahoma Health Sciences Center, Oklahoma City, OK, USA

[8] Department of Pharmaceutical Sciences, University of Oklahoma Health Sciences Center, Oklahoma City, OK, USA

[9] Department of Physiology, College of Medicine, University of Oklahoma Health Sciences Center, Oklahoma City, OK, USA

[10] Oklahoma Center for Neuroscience, University of Oklahoma Health Sciences Center, Oklahoma City, OK, USA

[*] Correspondence: Dharambir K. Sanghera, PhD, FAHA, Department of Pediatrics, College of Medicine, University of Oklahoma Health Sciences Center, 940 Stanton L. Young Blvd., Rm 317 BMSB, Oklahoma City, OK 73104, USA; Email: Dharambir-sanghera@ouhsc.edu; Phone: 405-271-8001; Fax: 405-271-6027

[#] Deceased.

Abstract: MFGE8 is a major exosome (EV) protein known to mediate inflammation and atherosclerosis in type 2 diabetes mellitus (T2DM) in animal studies. However, its role in the pathophysiology of human T2DM, obesity, and cardiovascular disease has been controversial and less explored. Earlier, we reported a rare Asian Indian population-specific missense variant (rs371227978; Arg148His) in the MFGE8 associated with increased circulating Mfge8 and T2DM. Here, we have further investigated the role of Mfge8 with T2DM risk in additional Asian Indians (n=4917) and Europeans and other multiethnic cohorts from UK Biobank (UKBB) (n= 455,808) and the US (n=1150). We also evaluated the exposure of Mfge8-enriched human EVs (with and without mutation) in zebrafish (ZF) for their impact on the cardiometabolic organ system. Most individual carriers of Arg148His variant not only had high circulating Mfge8 but also showed a positive significant correlation with glucose (r= 0.42; p = 4.9×10^{-04}), while the non-carriers showed a negative correlation of Mfge8 with glucose (r = -0.38; p = 0.001) in Asian Indians. However, the same variant was monomorphic in Europeans and other ethnic groups of UKBB and US cohorts. Even in the absence of Arg148His variant, serum Mfge8 correlated significantly with blood glucose both in Caucasians (r = 0.56; p = 8.4×10^{-10}) and individuals from other ethnicities (African Americans, Hispanics, and others) (r = 0.31; p = 5.8×10^{-4}). Since Mfge8 is an EV marker, the exposure of Mfge8-enriched human EVs to ZF larvae rapidly developed fatty liver disease and heart hypertrophy and exhibited redundant growth with poor muscular architecture with and without the high-fat diet (HFD), while the control group fishes developed fatty liver disease and heart hypertrophy only after the HFD feeding. With strong evidence from animal studies supporting the role of Mfge8 in obesity, insulin resistance, and the development of atherosclerosis in T2DM, the current research suggests that circulating Mfge8 may be a potential marker for preventing or reducing the risk of T2DM and cardiovascular disease.

Keywords: circulating Mfge8; rare variants; exosomes; zebrafish; cardiometabolic disease

1. Introduction

Type 2 diabetes mellitus (T2DM) is a multifactorial disease characterized by the combinatorial effect of insulin resistance and inadequate insulin secretion from β-cell of pancreatic islets [1]. The comorbidities associated with the disease range from the development of stroke [2], coronary heart disease [3], vascular diseases, renal failure, blindness, delayed wound healing [4], and amputation. T2DM, if not diagnosed or untreated for more extended periods, can affect the organs, leading to hyperglycemia or hyperinsulinemia [5]. Both environmental and genetic factors play a significant role in T2DM susceptibility [6]. The heritability estimates of T2DM range from 40% to 70%; among them, only 10-20% is explained by common variants identified in extensive genome-wide association studies (GWAS) [7,8]. Thus, efforts are needed to account for missing heritability associated with T2DM, using more biologically related explanations.

The milk fat globule-epidermal growth factor 8 (Mfge8), also known as lactadherin, is a secreted integrin ligand that binds to $\alpha v \beta 3/5$ and $\alpha 8 \beta 1$ integrin receptors [9,10]. It has a multifaceted role in regulating biological and physiological processes such as inflammation, apoptosis, atherosclerosis, wound healing [4], and tissue remodeling [11]. The anti-inflammatory nature of Mfge8 is enhanced by the clearance of apoptotic cells, and failure can lead to autoimmune diseases [9,12]. The increase in Mfge8 levels facilitates the proinflammatory phenotypic shift of vascular smooth muscle cells (VSMCs), rendering aged arteries at risk of cardiovascular diseases [13]. An inframe insertion and splice acceptor variant in the MFGE8 gene in the Finnish population was protective against coronary atherosclerosis [14]. Studies have reported increased circulating Mfge8 concentrations in diabetic patients [15]. Mfge8 concentrations in animal models increased triglyceride hydrolase activity regulating lipid storage [16]. However, the role of Mfge8 in the pathophysiology of human T2DM, obesity, and cardiovascular disease has been less explored and controversial. Using exome sequencing, we earlier identified a rare gain-of-function missense variant (rs371227978: Arg148His) associated with increased circulating Mfge8 levels and increased risk of T2DM in Punjabi Sikhs [5]. In this investigation, we have further evaluated if the increased blood levels of Mfge8 increase the risk of T2DM and cardiovascular complications in additional Asian Indians and Europeans from UK Biobank (UKBB) and non-Asian Indian populations from Oklahoma. Since Mfge8 is an exosomal (EV) marker protein, we also evaluated the fate of MFGE8-enriched human EVs (with and without mutation) in zebrafish (ZF) for their impact on the cardiometabolic organ system.

2. Methods and Materials

2.1. Study Subjects

2.1.1. Asian Indian Diabetic Heart Study (AIDHS)/Sikh Diabetes Study (SDS)

A total of 4917 individuals (2735 T2DM cases and 2162 controls) were included in this study from the Asian Indian Diabetic Heart Study (AIDHS)/Sikh Diabetes Study (SDS) [17–20]. The Sikh population is a relatively homogenous endogamous community from India. They are primarily non-smokers with the current study population having only 1.4% of smokers, and ~ 50% of them are vegetarian. However, the incidence of T2DM and cardiovascular diseases is very high [21]. Diabetic subjects (N = 2735; male/female = 1544/1191) were identified based on their medical records for symptoms and use of diabetic medications and were defined as people with diabetes based on fasting glucose levels following the American Diabetes Association guidelines as described earlier [18–20,22]. Non-diabetic controls (N = 2162; male/female = 1174/988) were selected based on a fasting glycemia < 100.8 mg/dl (5.6 mmol/l) or 2-hour glucose < 141.0 mg/dl (7.8 mmol/l). All blood samples were obtained at the baseline visit. Subjects with type 1 diabetes, or those with a family member with type 1 diabetes, or rare forms of T2DM subtypes (maturity-onset diabetes of the young [MODY][23]) or secondary diabetes (from, e.g., hemochromatosis or pancreatitis) were excluded from the study based on clinical reports, as previously described [20]. All participants in this study were from the northern part of India and provided written informed consent following procedures approved by the institutional review boards (IRBs). All AIDHS/SDS protocols and consent documents were reviewed and approved by the University of Oklahoma Health Science Center (OUHSC)'s IRB as

well as the Human Subject Protection (Ethics) committees at the participating hospitals and institutes in India as described previously [24,25].

Body mass index (BMI) was calculated as [weight (kg)/height (m^2)]. A tape measure was used to measure the waist and hip circumferences at the abdomen and the hip, respectively. The World Health Organization's (WHO) new guidelines for the BMI thresholds for Asians were followed (WHO Expert Panel, 2004) [26]. Blood pressure (BP) was measured twice after a 5-minute seated rest period with the participant's feet flat on the floor. Serum lipids [total cholesterol, fasting serum triglycerides (TG), high-density lipoprotein cholesterol (HDL-C), and low-density lipoprotein cholesterol (LDL-C)] were measured using standard enzymatic methods (Roche, Basel, Switzerland) as described previously [27]. Coronary heart disease (CHD) was considered if there was use of nitrate medication (nitroglycerine), electrocardiographic evidence of angina pain, coronary angiographic evidence of severe (greater than 50%) stenosis, or echocardiographic evidence of myocardial infarction. The diagnosis was based on the date of coronary artery bypass graft (CABG) or angioplasty and medication usage obtained from patient records as described previously [19,28]. The Mfge8 measures were performed in a selected subset of 143 individuals, including 69 T2DM cases and 73 controls who were matched for age, gender, and BMI and were of the same ancestry. All were available with MFGE8 genotypes.

2.1.2. MISS and OLIVER Cohorts

The MISS (Metabolome in Ischemic Stroke Study) is an ongoing prospective study that began in 2017, investigating ischemic stroke biomarker predictors using genomics, metabolomics, and omics technologies. The OLIVER (Oklahoma Multiethnic CV Health Disparity Study) is a multiethnic population-based study that aims to investigate the disparity of cardiovascular disease health disparity in Oklahoma populations residing in rural and urban areas and patients from Oklahoma University Medical Center (OUMC) [29,30]. Enrollment includes children and adults and recording data of anthropometry, medical history, and collection of blood samples for measurement of glucose, lipid profile, and genotyping as described previously [29,31,32]. Serum Mfge8 levels were measured in 218 multi-ethnic individuals from MISS-OLIVER studies (n=1150), including 101 Caucasians, 61 African Americans, 30 Hispanics, and 26 (Native Americans or Others). The inclusion criteria for this study involved subjects with and without T2DM, heart disease or dyslipidemia. Among the 218 individuals, 114 T2DM cases were selected based on medical records, and 104 non-T2D controls were selected based on medical records and matched for age, gender, and ethnicity. Excluded were the individuals in control who had cardiovascular diseases. All study protocols were approved by the University of Oklahoma Health Sciences Center's Institutional Review Board.

2.1.3. UK Biobank (UKBB) Study Participants

Of the 455,808 participants from UKBB, 9372 individuals (1943 T2DM cases and 7429 controls) were of South Asian (SA) ancestry, including Indians, Pakistani, Bangladeshi, and any other Asian background, and were analyzed separately. T2DM was characterized based on doctor-diagnosed disease phenotype and glycated hemoglobin levels (HbA1c) as described earlier [33]. For identifying carriers for the rare variant (rs371227978) of MFGE8, we used imputed data released by the UKBB for SA and European subjects. Participants with inconsistent reports and genotypic inferred sex inconsistencies or withdrawn consent were removed, as explained previously [29,34].

2.2. Genotyping

The missense variant in the MFGE8 (rs371227978 C/T: Arg148His) discovered using exome sequencing [5] was genotyped in a total of 4917 individuals of the AIDHS/SDS using Applied Biosystems' Custom TaqMan Genotyping Assay (Life Technologies Corp., Grand Island, NY) following the manufacturer's instructions. Genomic DNA was extracted from buffy coats or whole blood using QIAamp kits (Qiagen, Chatsworth, CA) or the salting-out method as described previously [35,36]. The entire genotyping was conducted on the samples blinded for phenotypes using TaqMan assay using Applied

Biosystems QuantStudio 6 as described elsewhere [25,37,38]. We ran 40 replicative controls and 20 negative controls for quality control to match the concordance. Genotypes were scored by analyzing data in both real-time as well as allele discrimination analysis using QuantStudio Software v1.3. Samples with poor calls or poor clusters were repeated to achieve the call rate > 95%.

2.3. Mfge8 Measurements Using ELISA

Circulating concentrations of Mfge8 levels were quantified using frozen serum/plasma aliquots by enzyme-linked immunosorbent assay (ELISA) kits from Boster Biological Technology (Pleasanton, CA, USA) using the standard protocols following the manufacturer's instructions. (www.bosterbio.com). Briefly, the Boster's Human Mfge8/Lactadherin ELISA Kit, based on classic sandwich ELISA technology, was used to measure serum Mfge8. Test samples were added to the monoclonal antibody (specific for Mfge8/cadherin) pre-coated 96-well plates. A biotinylated detection polyclonal antibody was subsequently added and then washed with 1x phosphate buffer saline (PBS). Avidin–biotin–peroxidase complex was added, and unbound conjugates were washed away with 1xPBS buffer. Horseradish peroxidase (HRP) substrate 3,3',5,5'-tetramethylbenzidine (TMB) was used to visualize HRP enzymatic reaction. HRP catalyzed TMB to produce a blue color product that changed into yellow after adding an acidic stop solution. The density of yellow is proportional to the Human Mfge8/Lactadherin amount of sample captured in the plate, and the optical density (OD) absorbance was measured at 450nm by SmartReader (ACCURIS Instruments) microplate reader. A standard curve was generated to determine the Mfge8 level in the tested serum/plasma samples following the manufacturer's instructions described earlier [5].

2.4. EVs Extraction

Sera of mutation carriers from the AIDHS/SDS were used to extract EVs. Heterozygous and homozygous carriers of this mutation exhibited high Mfge8 levels, most of which were diabetic. Following the manufacturer's protocol, EVs were isolated from patient sera using miRCURY Serum/Plasma Exosome Kit (Qiagen Inc., Chatsworth, CA, USA). We have characterized the particle size morphology of the EVs extracted from the sera of patients using transmission electron microscopy and nanoparticle tracking analysis (NTA) using Nanosight 300 (Figure 1). The EV pellet was resuspended in a resuspension buffer.
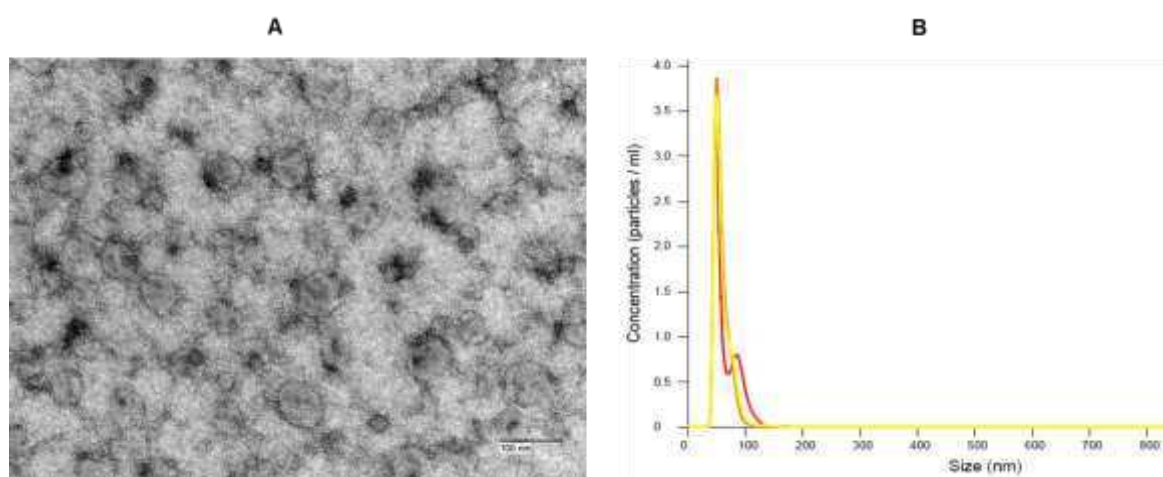
Figure 1. (**A**) Characterisation of extracellular vesicles (exosomes) isolated from human serum. The electron microscopic image by NANOSPRT12 shows the morphology of bright and round-shaped vesicles (exosome) from the serum sample. (**B**) Nanoparticle tracking analysis (NTA) Nanosight 300 particle tracking profile of harvested exosomes showing average particle size range between 50 to 135nm (n=3).

2.5. ZF Studies Using Human EVs

To evaluate the phenotypic effects of the rare variant (rs371227978) and high Mfge8 levels *in vivo*, we used TL (Tübingen long fin) strain, a commonly used strain of ZF (*Danio rerio*) lines. TL strain of ZF was set up for breeding, and the next day (Day 1 post-fertilization (1dpf), 500 embryos were collected, and dead embryos were sorted out. EVs were injected on Day 2 (2dpf) embryos using a microinjector and kept in an incubator [39] and segregated into two groups: test group injected with 10nL of pooled Mfge8 EVs injected/embryo (High Mfge8 concentrations= 20.76μg/μL) and the control group injected with 10nL of resuspension PBS buffer. Our protocol for ZF studies was approved by the Institutional Animal Care and Use Committee (IACUC) (Protocol #: 22-037-ECHI) and Institutional Bio-safety Committee (IBC) (100602A) of the University of Oklahoma Health Sciences Center.

2.5.1. Diet Experiments

The feeding experiments began on Day 6 on the 6dpf larvae. All test and control group fish were distributed in 3L tanks (20 fish per tank) and fed on defined diets for 14 days. All fish were fed with rotifers when kept in tanks. Animals were housed in the main aquarium of the ZF Animal Resource core facility of the University of Oklahoma and maintained on a 14-hour light, 10-hour dark cycle. All fish were fed with rotifers when kept in tanks. Animals were divided into four groups; Group 1: Control regular diet once daily (Fed with GEMMA Micro 75 (GM75) (Fat – 14%, Protein –59%) and kept in E3 water (5mM NaCl, 0.17 mM KCl, 0.33 mM CaCl2, 0.33 mM MgSO4, and 0.1% Methylene Blue)), Group 2: Control high-fat diet (HFD) 3 times/day (BioVita Starter fish feed (Fat – 18%, Protein –53%) and Egg yolk powder), Group 3: EVs with High Mfge8 levels fed with standard diet once daily (GM75 and E3 water), and Group 4: EVs with high Mfge8 levels fed with HFD 3 times/day (BioVita Starter fish feed and Egg yolk powder). On Day 14, animals in the test group were injected with 2nd dose of high Mfge8 EVs in the caudal vein of the larvae (10nl of pooled EVs/fish) and fed on defined diets for an additional 1 week. Three animal larvae from each feeding group were euthanized and sent for histology.

2.5.2. Larvae Tissue Embedding and Hematoxylin and Eosin (H&E) Staining

Following the manufacturer protocol, a Leica TP1020 tissue processor was used to process the tissue. Briefly, tissue in 10% neutral formalin buffer (NBF) is moved into labeled tissue blocks. The tissue in the blocks is progressively dehydrated with increasing concentrations of ethanol, then in xylene and imbibed with paraffin liquid. Due to the fragility of ZF larvae, they were placed in biospecimen bags, and 5 minutes in each step of processing was adapted. The paraffin-imbibed tissue is taken out and embedded according to orientation as needed using a 10X dissection microscope. The formalin-fixed paraffin-embedded tissues were sectioned at a desired thickness (4μm) and mounted on positively charged slides. The slides were dried overnight at room temperature and incubated at 60°C for 45 minutes. The Hematoxylin and Eosin were purchased from Leica biosystems, and staining was performed utilizing Leica ST5020 Automated Multistainer following the staining protocol at the SCC Tissue Pathology Shared Resource.

2.6. Statistical Analysis

Genotype allele frequencies were calculated by allele counting. Variables with skewed distributions were normalized by log transformation before statistical comparisons (e.g., blood glucose, TG, and Mfge8 concentrations), and *p* values reported are from the transformed data. Transformed variables were back converted into the original measurement scale for presentation. Pearson's correlation was performed to determine the linear association between Mfge8 levels and other traits in each ethnic group. All statistical analyses were performed using SPSS for Windows' statistical package (version 29.0) (SPSS Inc., Chicago, USA). The liver fat and heart size of ZF larvae were measured using NIH ImageJ [40].

3. Results

The clinical and physical characteristics of the study subjects segregated by T2DM status are presented in Table 1. As expected, mean blood glucose and systolic BP significantly increased in T2DM

cases compared to controls in all the cohorts. The genotype and allele frequency distributions of the rare variant rs371227978 located in MFGE8 are presented in Table 2. The AIDHS/SDS cohort (N = 4917) had 77 carriers, and 62% of them were diabetic, while UKBB SA (N = 9372) had 28 carriers, and 26% were diabetic. The minor allele frequencies (MAF) were 0.008 and 0.002 in AIDHS/SDS and UKBB SA cohorts, respectively.

Table 1. Clinical characteristics of study participants.

Trait	AIDHS/SDS (N =4917)			UKBB South Asians (N = 9372)			MISS (N =218)		
	Cases (N=2735)	Controls (N=2162)	p-value	Cases (N=1943)	Controls (N=7429)	p-value	Cases (N=114)	Controls (N=104)	p-value
Age (years)	54.80±11.43	48.66±14.40	8×10^{-57}	56.64±7.97	52.43±8.35	4×10^{-88}	60.72±12.17	57.91±15.19	0.14
Gender (Male/Female)	1544/1191	1174/988	0.13	1212/731	3855/3574	5×10^{-17}	66/48	44/60	0.02
BMI (Kg/m²)	27.19±4.85	26.20±4.81	1×10^{-12}	28.68±4.81	26.77±4.20	7×10^{-54}	34.64±14.08	29.19±6.80	4×10^{-4}
Systolic BP (mmHg)	144.18±24.00	132.55±30.72	2×10^{-44}	140.60±19.09	135.87±19.56	9×10^{-21}	151.61±30.85	139.95±21.14	0.001
Diastolic BP (mmHg)	84.68±12.29	80.98±12.22	2×10^{-24}	82.23±10.63	82.45±10.75	0.43	84.92±21.19	83.63±13.88	0.60
Glucose (mg/dL)	176.41±72.71	96.75±13.71	4×10^{-293}	113.72±71.34	76.93±32.33	5×10^{-98}	236.32±106.81	92.96±7.99	3×10^{-27}
Total Cholesterol (mg/dL)	180.99±48.29	185.67±69.08	0.02	78.60±27.72	93.84±28.40	2×10^{-97}	180.81±62.33	173.40±42.22	0.31
Triglycerides (mg/dL)	181.70±109.31	158.01±104.21	2×10^{-12}	37.49±21.14	34.81±20.94	1×10^{-6}	176.57±125.79	107.98±70.80	2×10^{-6}
HDL-C (mg/dL)	39.90±14.28	42.17±14.54	4×10^{-7}	17.95±8.55	20.09±9.63	2×10^{-21}	39.09±11.83	46.76±14.60	5×10^{-5}
LDL-C (mg/dL)	108.03±37.85	114.16±38.72	8×10^{-7}	48.00±19.43	59.33±19.69	2×10^{-110}	107.85±51.63	104.00±37.76	0.55

Values are displayed in mean±SD; AIDHS/SDS: Asian Indian Diabetic Heart Study/Sikh Diabetes Study; UKBB: UK Biobank; MISS: Metabolome in Ischemic Stroke Study; BMI: Body mass index; BP: Blood Pressure; HDL-C: High-density lipoproteins-Cholesterol; LDL-C: Low-density lipoproteins-Cholesterol.

**Table 2.** Genotype and frequency of MFGE8 rare variant (rs371227978) in AIDHS/SDS and UKBB populations.

Variant details	AIDHS/SDS (N = 4917)	UKBB South Asians (N = 9372)
Total Carriers	77	28
Heterozygous carriers (CT); T2D (%)	75; 61	27; 26
Homozygous mutant carriers (TT); T2D (%)	2; 100	1; 0
Minor allele frequency	0.008	0.002

AIDHS/SDS: Asian Indian Diabetic Heart Study/Sikh Diabetes Study; UKBB: UK Biobank.

3.1. Allelic Distribution and Genotype-Phenotype Association of Arg148His in Asian Indians and Other Ethnic Groups

The present study focused on population-specific rare missense variant rs371227978 located on the fifth exon of the MFGE8 gene (Figure 2A). The functional variant (Arg148His) was present mainly in Sikhs (77/4917) (both families and sporadic cases) (Figure 2B and F) and SAs (330/30616) from UKBB; no carriers were reported in Latinos (0/35432), African (0/24942) and East Asians (0/30616) in gnomAD (Figure 2F). Furthermore, this variant was highly conserved across species (Figure 2E). Genotypic–phenotypic association revealed that most of the 'T' allele carriers presented higher serum/plasma levels of Mfge8 (ranging from 0.06 ng/ml to 33.4 ng/ml) compared to wild type 'C' allele carriers (0.87 ng/ml to 11.1 ng/ml). Mfge8 levels and TG distribution among the rare variant genotype carriers correlated significantly ($p = 2.1 \times 10^{-3}$ and $p = 3.2 \times 10^{-3}$). The homozygous mutant carriers had higher Mfge8 and TG levels than heterozygous and non-carriers (Figure 2C and D).
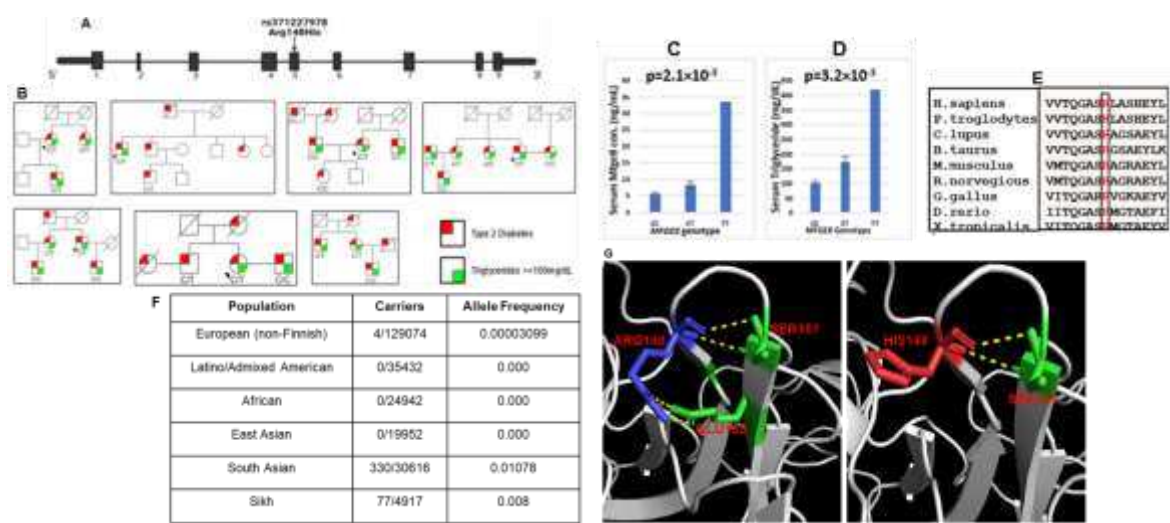
Figure 2. (A) Location of rare missense variant rs371227978 (C/T; Arg148His) on the fifth exon of the MFGE8. Black rectangles indicate exons, and the horizontal line represents non-coding sequence **(B)** Pedigrees carrying MFGE8 rare missense variant **(C)** The distribution of circulating Mfge8 concentrations by MFGE8 genotypes **(D)** The distribution of serum triglyceride levels by MFGE8 genotypes **(E)** Multiple amino acid sequence alignment (NCBI Homologene) of MFGE8 indicating conservation of arginine (shown in red) at position 148. Mutant sequence (Homo sapiens mutant) (replacement of arginine [R] by histidine [H] at 148 position), **(F)** Distribution of MFGE8 variant (rs371227978) in major global populations in Exome Aggregation Consortium, and **(G)** Rare damaging variant rs371227978 (Arg148His) in MFGE8 gene. The wild-type residue (blue) Arginine 148 forms hydrogen bonds with Serine 151 and Glutamine at position 153. However, the mutant residue (red) Histidine 148 with a pentane ring destabilizes the protein, causing of loss of hydrogen bonds with Glutamine at position 153.

*3.2. Effects of MFGE8 (Arg148His) Variant on Circulating Mfge8 Levels*

To understand the effect of genotype on the human serum Mfge8 levels underlying T2DM status, we quantified circulating Mfge8 levels in a selected subset of individuals in AIDHS/SDS and MISS cohorts. Most of the carriers had Mfge8 levels greater than 8ng/mL and had a higher number of individuals who had diabetes in each subgroup. On the other hand, non-carriers had low Mfge8 levels, mostly below 8ng/mL, and each subgroup had more nonT2DM controls than diabetic individuals (Figure 3).
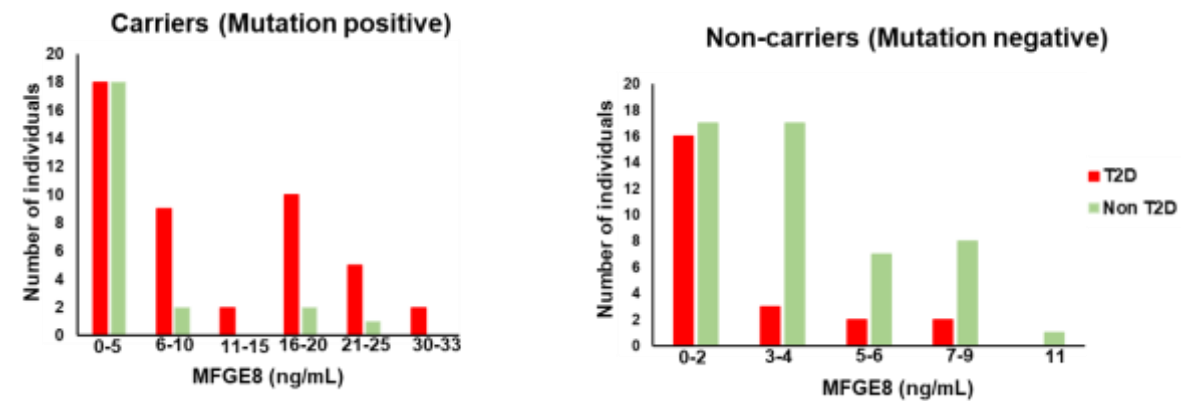
Figure 3. Bar graphs depicting the number of T2DM cases and controls divided by Mfge8 concentrations.

3.3. Association of Circulating Mfge8 Levels with Diabetes Risk Traits in Humans

The association of glucose with Mfge8 levels was marginally significant in SAs (r = 0.22; p = 0.01). Both Caucasians (r = 0.56; p = 8.4x10^{-10}) and other ethnicities (African Americans, Hispanics, and Native Americans) (r = 0.31; p = 5.8x10^{-4}) from Oklahoma showed a robust correlation of Mfge8 concentrations with glucose. The association of Mfge8 levels with glucose combining all ethnicities was also highly significant (r = 0.25; p = 2.9x10^{-6}) (Figure 4 and Table 3).
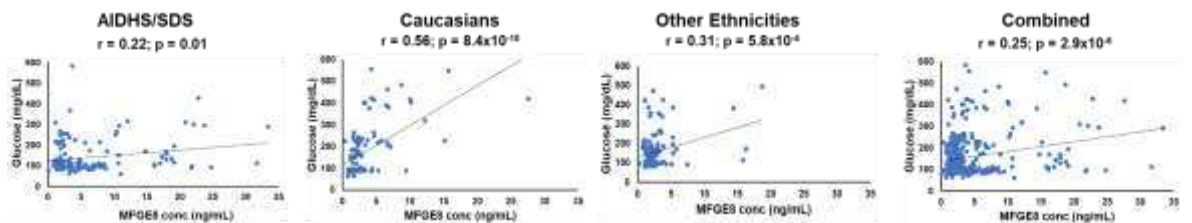
Figure 4. Scatter plots depicting the correlation of Mfge8 levels vs. glucose in different ethnicities.

**Table 3.** Correlation of Glucose with Mfge8 concentrations in different ethnicities.

	Mfge8 vs Glucose	
Ethnicity	**Pearson Correlation (r)**	**p-value**
Combined	**0.25**	**2.9 x10^{-06}**
AIDHS/SDS	0.22	0.01
Caucasians	**0.56**	**8.4x10^{-10}**
All others	**0.31**	**5.8 x10^{-04}**

AIDHS/SDS: Asian Indian Diabetic Heart Study/Sikh Diabetes Study.

We next evaluated the association of Mfge8 levels with other T2DM-associated traits. As shown in Supplementary Table 1a, BMI, systolic and diastolic BP, and TG showed a negative correlation with Mfge8 levels among wild-type (non-carriers). However, people with T2DM had a significant association with Mfge8 levels in carriers and non-carriers (r = 0.38; p = 0.001 and r = -0.29; p= 0.01) respectively in AIDHS/SDS (Supplementary Table 1a). Interestingly, no carriers of the MFGE8 variant (rs371227978) were found in the MISS-OLIVER cohort or Europeans of UKBB. To understand the effect of Mfge8 levels on diabetes and its associated traits, correlation analysis was performed by segregating the data in T2DM cases and controls. There were no significant associations of Mfge8 with traits except a marginal positive correlation with diastolic BP (r = 0.58; p = 0.02) in controls (Supplementary Table 1b).

3.4. Histopathological Examination of ZF Organs after Exposure to Human EVs with High and Low Mfge8

Based on the significant association of this rare variant with T2DM, we next evaluated the functional fate of Mfge8-enriched EVs in-vivo using the ZF model. Mfge8 is a known marker of secretory extracellular vesicles. The serum EVs from individuals having high Mfge8 levels were injected in the ZF larvae at 2dpf. The H&E images of the heart and liver for the control and high Mfge8 groups with regular diet and HFD are shown in Figure 5. Both high Mfge8 fishes with normal and HFD showed hypertrophic hearts, with the heart of ZF in the HFD group being dysmorphic with a fissure. The fat disposition in liver hepatocytes of high Mfge8 ZF was increased 3-4-fold in response to HFD. There were 80% of hepatocytes with fat deposition, and hepatocytes loaded with fat showed marked degeneration of hepatocyte nuclei with possible steatosis in the high Mfge8 HFD group. Also, the high Mfge8 fish exhibited an increase in ectopic fat in hepatocytes with a normal diet compared to the control normal diet group (Figure 5). The mean liver fat levels were significantly higher in the high Mfge8 HFD group, and cardiac hypertrophy was evident in high Mfge8 HFD fishes (Figure 6).
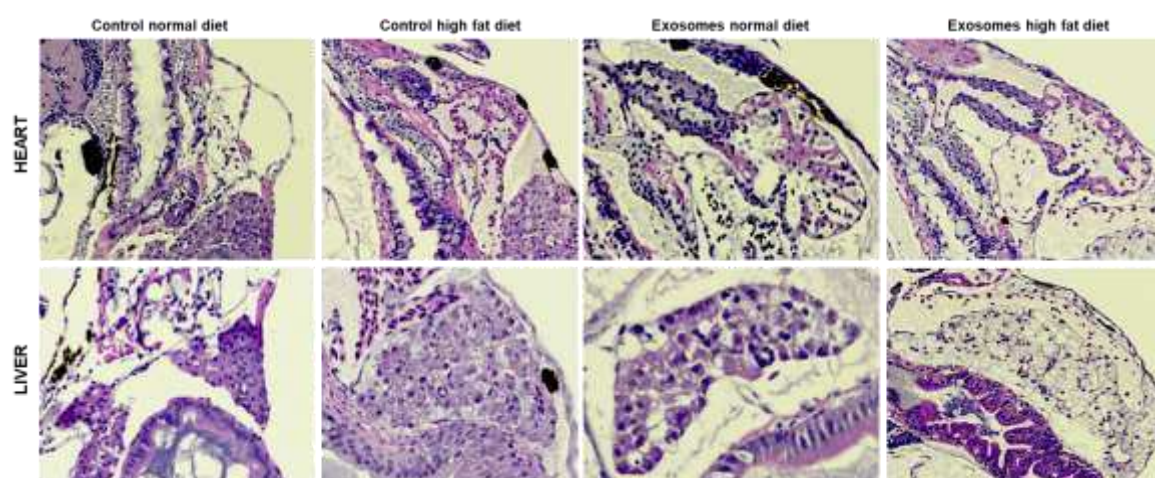
Figure 5. Composite showing the anatomical changes in the heart and liver of control and high Mfge8 and missense mutation exosomes injected zebrafish larvae fed with a normal and high-fat diet.
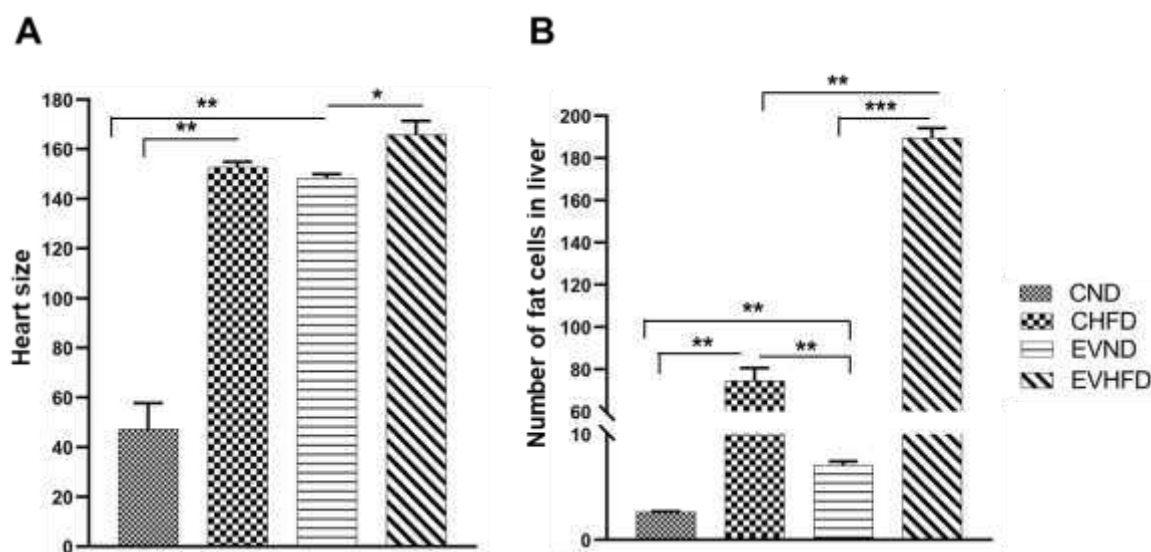
Figure 6. Bar graphs represent the liver fat measurement and heart size of zebrafish larvae using NIH ImageJ. **(A)** the heart size **(B)** the number of fat cells in the liver of Zebrafish larvae. The 9nl of serum exosome extracted from patients with high Mfge8 and missense mutation were injected in a 2-cell stage yolk sac and then in the caudal vein plexus after 14 dpf. Control fish of the corresponding birth

stage were injected with blank (PBS). Data represent the average (mean and standard errors) of three animals in each group. ***: $p < 10-4$; **: $p < 10-3$; *: $p < 10-2$.

4. Discussion

T2DM etiology is highly complex due to its multiple roots of origin. Polygenic risk scores (PRS) based on common variant-derived GWAS can only partially explain the disease risk, and the clinical effectiveness of these data for risk prediction for T2DM (as early biomarkers), or diagnostics still presents a significant challenge [33,41]. Several studies have identified the enrichment of rare (coding and non-coding) variants matching the patient's phenotype still, replication of such studies is challenging because sporadic variants tend to be population-specific due to inbreeding. Therefore, sequencing T2DM cases and their families from diverse populations is important to identify new genetic targets but it may help discover new rare sub forms of T2DM [42–45].

Our previous data suggested that elevated circulating Mfge8 may promote the development of T2DM and cardiovascular risk [5]. A novel population-specific rare missense variant (Arg148His) in MFGE8 we identified earlier [5] was associated with increased circulatory Mfge8 and TG concentrations in Asian Indians. Most individual carriers of this variant not only had high circulating Mfge8 but also showed a positive significant correlation with glucose (r=0.42; $p = 4.9 \times 10^{-04}$), while individuals without the mutation showed a negative correlation of Mfge8 with glucose (r = -0.38; p = 0.001) (Table 4). The Arg148His is situated on the C-terminal site on discoidin-like domain-1 (DD1), which has a sequence similarity to blood coagulation factor V/VIII [46]. The wild-type residue Arginine 148 forms hydrogen bonds with Serine 151 and Glutamine 153. However, the mutant residue (Histidine 148) with a pentane ring would change the configuration due to the loss of hydrogen bonds with Glutamine 153 and increase the stability of the protein [47]. Consequently, it may boost its binding affinity with avβv integrins, increasing concentration [48] (Figure 2G).

Table 4. Correlation of blood glucose with Mfge8 concentrations in the AIDHS/SDS cohort.

Model	Mfge8 vs Glucose	
	Pearson Correlation (r)	p-value
Complete	0.18	0.04
Carriers	**0.42**	**4.9x10⁻⁰⁴**
Non-Carriers	**-0.38**	**0.001**

AIDHS/SDS: Asian Indian Diabetic Heart Study/Sikh Diabetes Study.

On the other hand, we observed some non-carriers of Arg148His, who also had high Mfge8 levels, suggesting an incomplete penetrance.   Additionally, 28% of people of non-SA descent in the MISS-OLIVER cohort also had high Mfge8 levels and a strong correlation with glucose levels (r = 0.47; p = 2.2x10⁻¹³) in the absence of rs371227978 in these individuals. Moreover, not all individuals' carriers of rs371227978 (Arg/His) had T2DM, but most individuals with high Mfge8 levels were prone to develop T2DM or had high blood glucose (Figure 3). These findings suggest that other (yet unknown) variants or within the MFGE8 locus might play a role in modulating Mfge8 levels and may increase the risk for T2DM and cardiovascular complications. Notably, the serum levels of Mfge8 could only be measured in individuals from AIDHS/SDS and MISS-OLIVER cohorts and not in UKBB because the data access was limited to genomic and clinical profiles and not the specimens, which is a limitation. However, similar results by another large independent study support our findings that individuals genetically susceptible to increased Mfge8 levels have been reported to develop a higher risk for cardiovascular diseases [49]. Another study examined serum Mfge8 levels and indices of insulin resistance in a multi-ethnic cohort of White, Chinese, and Hispanic individuals living in San Francisco and found Mfge8 as an independent predictor of insulin resistance [50].

Since Mfge8 is a marker of human EVs, we further evaluated the effects of EV-associated Mfge8 using ZF. The ZF is a useful model for studying human metabolic diseases because they have a high degree of anatomical and physiological similarity to humans in terms of organs and tissues, and the

central regulation of metabolism is conserved [51]. Since EVs regulate communication between organs in pathological processes of T2DM influencing insulin signals in target tissues [52], we examined the exposure of Mfge8-enriched EVs on tissues sensitive to insulin signaling   (liver, heart, and pancreas) in ZF and compared it with the control animals. The ZF with high Mfge8 exposed EVs showed cardiac hypertrophy with and without the HFD, while the control group fishes developed hypertrophy only after the HFD. The heart enlargement in the test group fed with a regular diet may suggest the role of high Mfge8 leading to cardiovascular complications. The aortas of diabetic rats expressed high levels of Mfge8 and have been shown to contribute to the development of atherosclerotic lesions in db/db mice [53].

Further, the test group ZF with high Mfge8 injected EVs revealed a 2-to-3-fold increase in hepatocyte fat accumulation in response to regular diet. There was an over 4-fold increase in liver fat accumulation in response to HFD (Figure 5). Additionally, over 85% of hepatocytes have fat deposition with marked degeneration of hepatocyte nuclei with possible steatosis in the animals exposed to human EVs in HFD. Mfge8 regulates dietary fat absorption and storage, leading to obesity in mice [54]. From these findings, we hypothesize that the enrichment of Mfge8 in sera of T2DM patients with (Arg148His) may correlate with increased biosynthesis of EVs and affect the signaling of downstream proteins and elucidating the EV-related mechanism of glucose homeostasis associated with Mfge8. The data presented here was created without the ablation of ZF MFGE8. However, our control fish fed on HFD did not show the same phenotype as seen in the fish exposed to human EVs (Figure 5). Notably, MFGE8 was not among the list of genes associated with T2DM or risk factor traits detected in GWAS and even in extensive GWAS/meta-analysis studies comprising hundreds of thousands of T2DM patients and controls. Since the role of Mfge8 in human metabolism is unclear, more in-depth examination is needed to identify its role as a biomarker for T2DM.

In summary, the current investigation further validates our earlier findings that the increase in Mfge8 in serum may increase the risk for T2DM in general populations of the US beyond Asian Indians and even in the absence of Arg148His variant. Our data also suggests that the Mfge8-enriched human EVs in ZF larvae showed development defects (fatty liver, cardiac hypertrophy, and dysmorphology) in the insulin-sensitive organs and substantiated the phenotypic role of this protein in cardiometabolic diseases. With strong evidence from animal experiments supporting the role of Mfge8 in obesity, insulin resistance, and the development of atherosclerosis in T2DM, the current study suggests that circulating Mfge8 may be a potential marker for preventing or reducing the risk of T2DM and cardiovascular disease.

References

1. Li Y, Ran W, Zhang J, Chen S, Li Y, Luo D, et al. Elevated serum milk fat globule-epidermal growth factor 8 levels in type 2 diabetic patients are suppressed by overweight or obese status. IUBMB Life. 2017;69(2):63-71.

2. Chen R, Ovbiagele B, Feng W. Diabetes and Stroke: Epidemiology, Pathophysiology, Pharmaceuticals and Outcomes. Am J Med Sci. 2016;351(4):380-6.

3. Aronson D, Edelman ER. Coronary artery disease and diabetes mellitus. Cardiol Clin. 2014;32(3):439-55.

4. Huang W, Jiao J, Liu J, Huang M, Hu Y, Ran W, et al. MFG-E8 accelerates wound healing in diabetes by regulating "NLRP3 inflammasome-neutrophil extracellular traps" axis. Cell Death Discov. 2020;6:84.

5. Sapkota BR, Sanghera DK. A rare missense variant in the milk fat globule-EGF factor 8 (MFGE8) increases T2DM susceptibility and cardiovascular disease risk with population-specific effects. Acta Diabetol. 2020;57(6):733-41.

6. Zimmet P. Type 2 (non-insulin-dependent) diabetes--an epidemiological overview. Diabetologia. 1982;22(6):399-411.

7. Avery AR, Duncan GE. Heritability of Type 2 Diabetes in the Washington State Twin Registry. Twin Res Hum Genet. 2019;22(2):95-8.

8. Vujkovic M, Ramdas S, Lorenz KM, Guo X, Darlay R, Cordell HJ, et al. A multiancestry genome-wide association study of unexplained chronic ALT elevation as a proxy for nonalcoholic fatty liver disease with histological and radiological validation. Nat Genet. 2022;54(6):761-71.

9. Hanayama R, Tanaka M, Miwa K, Shinohara A, Iwamatsu A, Nagata S. Identification of a factor that links apoptotic cells to phagocytes. Nature. 2002;417(6885):182-7.

10. Khalifeh-Soltani A, Ha A, Podolsky MJ, McCarthy DA, McKleroy W, Azary S, et al. alpha8beta1 integrin regulates nutrient absorption through an Mfge8-PTEN dependent mechanism. Elife. 2016;5.

11. Bu HF, Zuo XL, Wang X, Ensslin MA, Koti V, Hsueh W, et al. Milk fat globule-EGF factor 8/lactadherin plays a crucial role in maintenance and repair of murine intestinal epithelium. J Clin Invest. 2007;117(12):3673-83.

12. Hanayama R, Tanaka M, Miyasaka K, Aozasa K, Koike M, Uchiyama Y, et al. Autoimmune disease and impaired uptake of apoptotic cells in MFG-E8-deficient mice. Science. 2004;304(5674):1147-50.

13. Chiang HY, Chu PH, Lee TH. MFG-E8 mediates arterial aging by promoting the proinflammatory phenotype of vascular smooth muscle cells. J Biomed Sci. 2019;26(1):61.

14. Ruotsalainen SE, Surakka I, Mars N, Karjalainen J, Kurki M, Kanai M, et al. Inframe insertion and splice site variants in MFGE8 associate with protection against coronary atherosclerosis. Commun Biol. 2022;5(1):802.

15. Cheng M, Li BY, Li XL, Wang Q, Zhang JH, Jing XJ, et al. Correlation between serum lactadherin and pulse wave velocity and cardiovascular risk factors in elderly patients with type 2 diabetes mellitus. Diabetes Res Clin Pract. 2012;95(1):125-31.

16. Khalifeh-Soltani A, Gupta D, Ha A, Iqbal J, Hussain M, Podolsky MJ, et al. Mfge8 regulates enterocyte lipid storage by promoting enterocyte triglyceride hydrolase activity. JCI Insight. 2016;1(18):e87418.

17. Sanghera DK, Bhatti JS, Bhatti GK, Ralhan SK, Wander GS, Singh JR, et al. The Khatri Sikh Diabetes Study (SDS): study design, methodology, sample collection, and initial results. Hum Biol. 2006;78(1):43-63.

18. Sanghera DK, Nath SK, Ortega L, Gambarelli M, Kim-Howard X, Singh JR, et al. TCF7L2 polymorphisms are associated with type 2 diabetes in Khatri Sikhs from North India: genetic variation affects lipid levels. Ann Hum Genet. 2008;72(Pt 4):499-509.

19. Saxena R, Bjonnes A, Prescott J, Dib P, Natt P, Lane J, et al. Genome-wide association study identifies variants in casein kinase II (CSNK2A2) to be associated with leukocyte telomere length in a Punjabi Sikh diabetic cohort. Circ Cardiovasc Genet. 2014;7(3):287-95.

20. Saxena R, Saleheen D, Been LF, Garavito ML, Braun T, Bjonnes A, et al. Genome-wide association study identifies a novel locus contributing to type 2 diabetes susceptibility in Sikhs of Punjabi origin from India. Diabetes. 2013;62(5):1746-55.

21. Sanghera DK DS. Cardiovascular disease in South Asians;Risk factors, genetics and environment.Medicine Update 2016-1. New Delhi, London, Philadelphia, Panama: The Health Sciences Publishers 2016;2.

22. American Diabetes A. Diagnosis and classification of diabetes mellitus. Diabetes Care. 2004;27 Suppl 1:S5-S10.

23. Goyal S, Tanigawa Y, Zhang WH, Chai JF, Almeida M, Sim XL, et al. APOC3 genetic variation, serum triglycerides, and risk of coronary artery disease in Asian Indians, Europeans, and other ethnic groups. Lipids Health Dis. 2021;20(1).

24. Sanghera DK, Bejar C, Sapkota B, Wander GS, Ralhan S. Frequencies of poor metabolizer alleles of 12 pharmacogenomic actionable genes in Punjabi Sikhs of Indian Origin. Sci Rep. 2018;8(1):15742.

25. Sanghera DK, Hopkins R, Malone-Perez MW, Bejar C, Tan C, Mussa H, et al. Targeted sequencing of candidate genes of dyslipidemia in Punjabi Sikhs: Population-specific rare variants in GCKR promote ectopic fat deposition. PLoS One. 2019;14(8):e0211661.

26. Consultation WHOE. Appropriate body-mass index for Asian populations and its implications for policy and intervention strategies. Lancet. 2004;363(9403):157-63.

27. Sanghera DK, Demirci FY, Been L, Ortega L, Ralhan S, Wander GS, et al. PPARG and ADIPOQ gene polymorphisms increase type 2 diabetes mellitus risk in Asian Indian Sikhs: Pro12Ala still remains as the strongest predictor. Metabolism. 2010;59(4):492-501.

28. Sapkota BR, Hopkins R, Bjonnes A, Ralhan S, Wander GS, Mehra NK, et al. Genome-wide association study of 25(OH) Vitamin D concentrations in Punjabi Sikhs: Results of the Asian Indian diabetic heart study. J Steroid Biochem Mol Biol. 2016;158:149-56.

29. Goyal S, Tanigawa Y, Zhang W, Chai JF, Almeida M, Sim X, et al. APOC3 genetic variation, serum triglycerides, and risk of coronary artery disease in Asian Indians, Europeans, and other ethnic groups. Lipids Health Dis. 2021;20(1):113.

30. Rout M, Vaughan A, Blair A, Stavrakis S, Sidorov EV, Sanghera DK. Discovery and validation of circulating stroke metabolites by NMR-based analyses using patients from the MISS and UK Biobank. Neurochem Int. 2023;169:105588.

31. Zamani P, Proto EA, Mazurek JA, Prenner SB, Margulies KB, Townsend RR, et al. Peripheral Determinants of Oxygen Utilization in Heart Failure With Preserved Ejection Fraction: Central Role of Adiposity. JACC Basic Transl Sci. 2020;5(3):211-25.

32. Sidorov E, Bejar C, Xu C, Ray B, Reddivari L, Chainakul J, et al. Potential Metabolite Biomarkers for Acute Versus Chronic Stage of Ischemic Stroke: A Pilot Study. J Stroke Cerebrovasc Dis. 2020;29(4):104618.

33. Rout M, Wander GS, Ralhan S, Singh JR, Aston CE, Blackett PR, et al. Assessing the prediction of type 2 diabetes risk using polygenic and clinical risk scores in South Asian study populations. Ther Adv Endocrinol Metab. 2023;14:20420188231220120.

34. Sudlow C, Gallacher J, Allen N, Beral V, Burton P, Danesh J, et al. UK biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age. PLoS Med. 2015;12(3):e1001779.

35. Miller SA, Dykes DD, Polesky HF. A simple salting out procedure for extracting DNA from human nucleated cells. Nucleic Acids Res. 1988;16(3):1215.

36. Sanghera DK, Manzi S, Minster RL, Shaw P, Kao A, Bontempo F, et al. Genetic variation in the paraoxonase-3 (PON3) gene is associated with serum PON1 activity. Ann Hum Genet. 2008;72(Pt 1):72-81.

37. Been LF, Hatfield JL, Shankar A, Aston CE, Ralhan S, Wander GS, et al. A low frequency variant within the GWAS locus of MTNR1B affects fasting glucose concentrations: genetic risk is modulated by obesity. Nutr Metab Cardiovasc Dis. 2012;22(11):944-51.

38. Howard EW, Been LF, Lerner M, Brackett D, Lightfoot S, Bullen EC, et al. Carriers of a novel frame-shift insertion in WNT16a possess elevated pancreatic expression of TCF7L2. BMC Genet. 2013;14:28.

39. Abdelrahman D, Hasan W, Da'as SI. Microinjection quality control in zebrafish model for genetic manipulations. MethodsX. 2021;8:101418.

40. Schneider CA, Rasband WS, Eliceiri KW. NIH Image to ImageJ: 25 years of image analysis. Nat Methods. 2012;9(7):671-5.

41. Martin AR, Kanai M, Kamatani Y, Okada Y, Neale BM, Daly MJ. Clinical use of current polygenic risk scores may exacerbate health disparities. Nat Genet. 2019;51(4):584-91.

42. Bandyopadhyay B, Chanda V, Wang Y. Finding the Sources of Missing Heritability within Rare Variants Through Simulation. Bioinform Biol Insights. 2017;11:1177932217735096.

43. Wainschtein P, Jain D, Zheng Z, Group TOAW, Consortium NT-OfPM, Cupples LA, et al. Assessing the contribution of rare variants to complex trait heritability from whole-genome sequence data. Nat Genet. 2022;54(3):263-73.

44. Zuk O, Schaffner SF, Samocha K, Do R, Hechter E, Kathiresan S, et al. Searching for missing heritability: designing rare variant association studies. Proc Natl Acad Sci U S A. 2014;111(4):E455-64.

45. Cooper RS, Tayo B, Zhu X. Genome-wide association studies: implications for multiethnic samples. Hum Mol Genet. 2008;17(R2):R151-5.

46. Aziz M, Jacob A, Matsuda A, Wang P. Review: milk fat globule-EGF factor 8 expression, function and plausible signal transduction in resolving inflammation. Apoptosis. 2011;16(11):1077-86.

47. Waterhouse A, Bertoni M, Bienert S, Studer G, Tauriello G, Gumienny R, et al. SWISS-MODEL: homology modelling of protein structures and complexes. Nucleic Acids Res. 2018;46(W1):W296-W303.

48. Datta R, Lizama CO, Soltani AK, McKleroy W, Podolsky MJ, Yang CD, et al. Autoregulation of insulin receptor signaling through MFGE8 and the alphavbeta5 integrin. Proc Natl Acad Sci U S A. 2021;118(18).

49. Soubeyrand S, Nikpay M, Turner A, Dang AT, Herfkens M, Lau P, et al. Regulation of MFGE8 by the intergenic coronary artery disease locus on 15q26.1. Atherosclerosis. 2019;284:11-7.
50. Datta R, Podolsky MJ, Yang CD, Alba DL, Singh S, Koliwad S, et al. MFGE8 inhibits insulin signaling through PTP1B. bioRxiv. 2023.
51. Gut P, Reischauer S, Stainier DYR, Arnaout R. Little Fish, Big Data: Zebrafish as a Model for Cardiovascular and Metabolic Disease. Physiol Rev. 2017;97(3):889-938.
52. Chang W, Wang J. Exosomes and Their Noncoding RNA Cargo Are Emerging as New Modulators for Diabetes Mellitus. Cells. 2019;8(8).
53. Yu F, Li BY, Li XL, Cai Q, Zhang Z, Cheng M, et al. Proteomic analysis of aorta and protective effects of grape seed procyanidin B2 in db/db mice reveal a critical role of milk fat globule epidermal growth factor-8 in diabetic arterial damage. PLoS One. 2012;7(12):e52541.
54. Khalifeh-Soltani A, McKleroy W, Sakuma S, Cheung YY, Tharp K, Qiu Y, et al. Mfge8 promotes obesity by mediating the uptake of dietary fats and serum fatty acids. Nat Med. 2014;20(2):175-83.