



Cohort profile update: Tehran cardiometabolic genetic study

Maryam S. Daneshpour¹ · Mahdi Akbarzadeh¹ · Hossein Lanjanian¹ · Bahar Sedaghati-khayat^{1,2} · Kamran Guity^{1,3} · Sajedah Masjoudi¹ · Asiyeh Sadat Zahedi¹ · Maryam Moazzam-Jazi¹ · Leila Najd Hassan Bonab¹ · Bita Shalbafan⁴ · Sara Asgarian¹ · Goodarz Koli Farhood¹ · Niloofer Javanrooh¹ · Maryam Zarkesh¹ · Parisa Riahi¹ · Mohammad Reza Moghaddas¹ · Parvaneh Arbab Dehkordi¹ · Azar Delbarpour Ahmadi¹ · Firoozeh Hosseini⁵ · Sara Jalali Farahani⁶ · Farzad Hadaegh⁷ · Parvin Mirmiran⁵ · Fahimeh Ramezani Tehrani⁸ · Arash Ghanbarian⁷ · Mohammad Sadegh Fallah Mahboob Pasand⁹ · Parisa Amiri⁶ · Majid Valizadeh¹⁰ · Farhad Hosseipناه¹⁰ · Maryam Tohidi⁷ · Asghar Ghasemi¹¹ · Azita Zadeh-Vakili¹¹ · Mohammad Piryaee¹ · Shahram Alamdari¹² · Davood Khalili⁷ · Amirabbas Momenan⁷ · Maryam Barzin¹⁰ · Siroos Zeinali⁹ · Mehdi Hedayati¹ · Fereidoun Azizi¹²

Received: 20 November 2022 / Accepted: 13 April 2023 / Published online: 12 May 2023
© Springer Nature B.V. 2023

Abstract

The Tehran cardiometabolic genetic study (TCGS) is a large population-based cohort study that conducts periodic follow-ups. TCGS has created a comprehensive database comprising 20,367 participants born between 1911 and 2015 selected from four main ongoing studies in a family-based longitudinal framework. The study's primary goal is to identify the potential targets for prevention and intervention for non-communicable diseases that may develop in mid-life and late life. TCGS cohort focuses on cardiovascular, endocrine, metabolic abnormalities, cancers, and some inherited diseases. Since 2017, the TCGS cohort has augmented by encoding all health-related complications, including hospitalization outcomes and self-reports according to ICD11 coding, and verifying consanguineous marriage using genetic markers. This research provides an update on the rationale and design of the study, summarizes its findings, and outlines the objectives for precision medicine.

Keywords Tehran cardiometabolic genetic study (TCGS) · Precision medicine · Cardiometabolic biomarkers · Endocrine disease · Nutrition · Lifestyle · Consanguineous marriage · Cohort study

Abbreviations

AGT	Angiotensinogen	FIDs	Family identification codes
AIKTIP	Kernel machine SNP set analysis	FROH	Fraction of runs of homozygosity
BMI	Body mass index	FTO	Fat Mass And Obesity Associated
BP	Blood pressure	GLMM	General linear model
C16orf95	Chromosome 16 Open Reading Frame 95	GQLSM	Generalized quasi-likelihood scoring method
CETP	Cholesteryl Ester Transfer Protein	GRM	Genomic relationship matrix
CFS	Clinical follow-up study	GWAS	Genome-wide association study
CHD	Coronary heart disease	HDL-C	High-density lipoprotein cholesterol
CRF	Chronic renal failure	HTN	Hypertension
CVD	Cardiovascular disease	IBD	Identity-by-descent
DBP	Diastolic blood pressure	ICD-11	International Classification of Diseases 11th Revision
ESRP2	Epithelial Splicing Regulatory Protein 2	IGENT	Efficient entropy based algorithm for genome-wide gene–gene interaction analysis
FH	Familial hypercholesterolemia	LDL-C	Low-density lipoprotein cholesterol
		MC4R	Melanocortin 4 Receptor
		MH	Monogenic hypertension
		MIA3	MIA SH3 Domain ER Export Factor 3
		MMP2	Matrix Metalloproteinase 2

✉ Maryam S. Daneshpour
daneshpour@sbmu.ac.ir

✉ Fereidoun Azizi
azizi@sbmu.ac.ir

Extended author information available on the last page of the article

MODY	Maturity-onset diabetes of the young
MUO	Metabolically unhealthy obesity
NLGN1	Neuroigin-1
NLRC5	NLR Family CARD Domain Containing 5
PRS	Polygenic risk score
RPGRIP1L	Protein Phosphatase 1, Regulatory Subunit 134
SBP	Systolic blood pressure
SLC12A3	Solute Carrier Family 12 Member 3
SNPs	Single nucleotide polymorphisms
T1D	Type 1 diabetes
T2D	Type 2 diabetes
TC	Total cholesterol
TCGS	Tehran cardiometabolic genetic study
TCS	Thyroid cancer study
TFAs	Trans-fatty acids
TG	Triglycerides
TLGS	Tehran Lipid Glucose Study
TNXB	Tenascin XB
TOTS	Tehran Obesity Treatment Study
WC	Waist circumference
WDP	Western dietary patterns
WGR	Whole-genome resequencing
WHR	Waist-to-hip ratio
ZBED9	Zinc Finger BED-Type Containing 9

Introduction

Although the Iranian population is likely to have extensive genetic diversity, there is a significant lack of corresponding data in public genomic repositories. Therefore, the TCGS was initiated in 2011 to investigate the genetic makeup of Iranians. TCGS is a longitudinal, transgenerational study based on families, providing the demographic, biochemical, and genetic data of 20,367 participants over 22 years [1] which allows early screening to identify high-risk individuals who can be targeted for preventive interventions in primary and secondary healthcare systems.

Healthcare systems strive to improve efficiency and therapeutic benefit by utilizing genetic or molecular profiling for groups of patients. As a result, precision medicine, by incorporating established biomarkers, functional tests, imaging, and new genomics and omics developments for each population, aims to provide the “right treatment to the right patient at the right time” [1]. TCGS aims to identify key genetic factors that contribute to the development of cardiometabolic disease in Iranians, as the effects of omics layers, particularly genetics, are well-known to vary across ethnicities. By combining genetic findings with additional non-genetic data types available in the TCGS cohort, it will be possible to accurately identify high-risk individuals and provide timely intervention and effective treatment. This

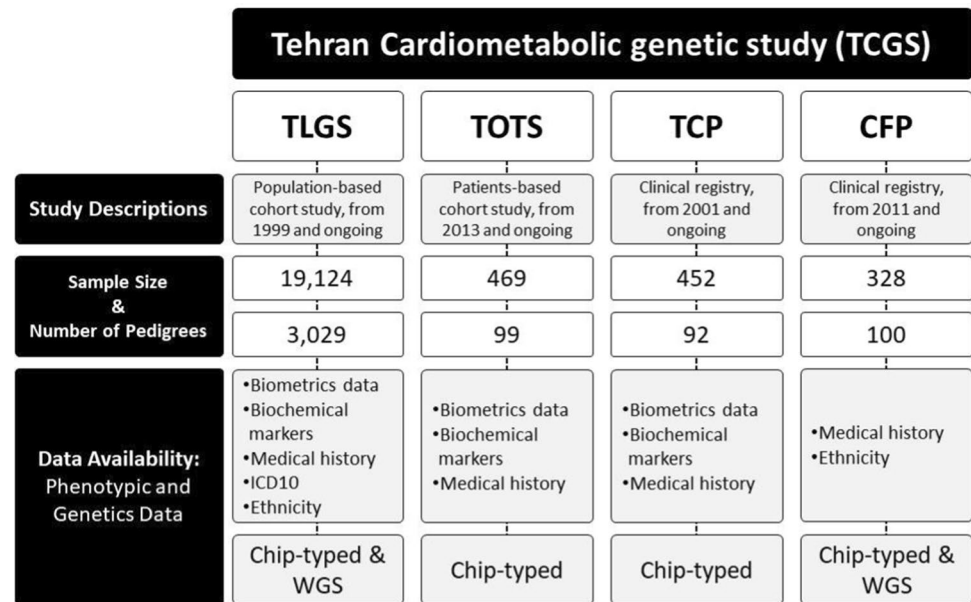
project provides significant information and data relevant to the Iranian population that contributes to global knowledge and serves as the first step toward implementing precision medicine in Iran.

TCGS design and characterization

TCGS study originated from the TLGS [2] that has followed up over 15,000 participants for at least 22 years without any pre-specified exclusion criteria. All individuals older than three years living in the area were invited to participate. The ongoing cohort project involves extensive surveys at baseline and regular follow-ups every three years with trained staff recording the development of cardiometabolic diseases and corresponding biochemical factors such as high cholesterol, low level of HDL-C, high TG, and behavior patterns such as smoking and physical activity [2]. In 2017 [3], we included the TOTS [4], the TCS [5], and CFS, which further focused on obesity intervention, thyroid cancer, and clinical genetic disorders. The TCGS project comprises 20,367 participants selected from four longitudinal, ongoing, and family-based studies (Fig. 1). Demographic characteristics, ethnicity, blood test results, medical history, drug history, gynecology-related information, diet, physical activity, and smoking information were gathered for each participant.

Between 1999 and 2022, data from all original projects, except for TOTS, was collected and digitized from paper-based questionnaires. However, in the TOTS study, physicians used computer-based questionnaires to record participants' medical conditions. Medical information was coded according to the International Statistical Classification of Diseases and Related Health Problems (ICD) version 11 (Supplementary file 2). All medical information was categorized as a self-report (SR) or obtained through follow-up during hospitalization (H). A detailed description of the TCGS population, including sex, age, ethnicity, and disease frequency, is provided in Table 1. Ethnicity status for 6,177 participants was collected through self-reporting and questionnaires on the birthplaces of the past three generations. Information on the genotypes of 3851 out of 6177 individuals was also recorded. The inclusion of various ethnic groups in the TCGS project provides insight into the diversity of the Iranian population, with the most frequent groups being Persian (76.6%), Turk (12.1%), and Gilak (3.8%) (Table 1).

The genotyping of 16,226 TCGS participants was performed using Illumina Human OmniExpress-24-v1-0 bead chip, which contains 652,919 SNP loci, at the deCODE genetics/Amgen company (Iceland) according to the manufacturer's specifications (Illumina Inc., San Diego, CA, USA) [6] (Tables S2, S3, S4, and S5). IBD information was derived from genotypes and used to confirm all familial relationships. Around 660 couples met the threshold of 0.05 for consanguineous marriage relatedness, representing

Fig. 1 Description of projects which attended TCGS

a consanguineous marriage rate of 28.50% (Figure S7). Among the couples with no IBD information, 543 declared consanguinity, resulting in an overall rate of 28.15%. The study used genotyping information to correct familial relationships and assigned FID (Figure S5). The number of the cluster was 4,428 (mean \pm standard deviation (sd)): 3.87 ± 1.86 , (Min, Max: 1, 16)), and after correction, the number of FIDs was 3,320 (mean \pm sd: 5.51 ± 5.41 , (Min, Max: 3, 56)) with 1132 individuals. There were 6,368 nuclear families with 26,392 parents/offspring and 5,051 sibships after correction. The relationship distribution in TCGS families is shown in Figure S6.

The TCGS project has used the ICD-11 classification to create a comprehensive and standard system for diagnosing health conditions. A total of 24 ICD-11 chapters were used to identify 873 specific health conditions in TCGS participants [7], as detailed in Supplementary file 2. To examine the most commonly reported mortality and morbidity statistics, we focused on the top five ICD-11 groups presented in Table S6. The findings indicate that a significant number of participants in the study had diabetes mellitus in pregnancy, goiter, and hypothyroidism as their top health conditions. These findings could be attributed to the study's initial focus on cardiovascular disorders, diabetes, and thyroid disease.

Risk factors and outcomes assessment

Here we present the key results from analyzing the most extensively studied traits in this population. Our goal is to identify genetic markers and their specific effects on this population, which can be used in more targeted risk management toward precision medicine. We outline the objectives of studying these traits and summarize the findings.

Blood pressure

Objective

Elevated BP is the most common risk factor for cardiovascular and renal diseases and is responsible for a significant proportion of morbidity and mortality worldwide[8]. Both genetic and environmental factors influence BP. To investigate the genomic basis of BP, we conducted various analyses, including familial aggregation analysis, heritability analysis [9], family-based linkage study [10], GWAS, epistasis analysis, and estimation of PRS[11]

Key methods and data collection

The primary outcome of interest was the incidence of HTN, as well as corresponding SBP and DBP levels, across three age groups: children (1–9 years), adolescents (10–17 years), and adults (ages 18 and older). For adults, HTN was defined as SBP \geq 140 mmHg, DBP \geq 90 mmHg, or if the participant was taking antihypertensive medication [13]. All participants with their average SBP and DBP during follow-up visits were included in the subsequent analysis. Missing BMI and WC were imputed.

Main results

Our research on Iranian families revealed that SBP, DBP, BMI, and WC were highly correlated in mother–offspring and sister–sister relationships, with heritability estimates of 25% and 30% for SBP and DBP, respectively, during a follow-up of 15 years. Members of index families with higher familial BMI or WC were found to have a significantly increased risk

Table 1 Baseline characteristics of participants in the TCGS

Participants, n*	20,367	
Female, n (%)	10,558 (51.8)	
Age at baseline mean (s.d.)	46.86 (21.53)	
Iranian ethnicity, n	Persian, n	4729
	Turk, n	746
	Gilak, n	255
	Lur, n	133
	Arab, n	84
	Kurd, n	69
	Balouch, n	28
	Turkman, n	14
	Tat, n	13
	Qashqai, n	3
	Mixed, n	103
Metabolic disease	Obesity	1580
	Diabetes mellitus	2280
	Dyslipidemia	1852
	Polycystic ovary	3520
	Rickets	15
Cardiovascular disease	Hypertensive heart disease	2282
	Heart failure	209
	Acute myocardial infarction	1205
	Coronary artery disease	1573
	Chronic ischemic heart disease	1024
	Heart valve diseases	114
Liver disease	Inflammatory liver disease	174
	Non-alcoholic fatty liver disease	280
Malignancy	Breast Cancer	128
	Pancreatic cancer	19
	Gastrointestinal cancer	129
	Prostate Cancer	34
	Lung cancer	18
Inflammatory disease	Psoriasis	52
	Rheumatoid arthritis	277
Neurological disorder	Cerebrovascular accident	572
	Epilepsy	372
	Parkinsonism	42
Kidney disease	kidney failure	47
	Chronic kidney disease	59
Respiratory disorder	Chronic obstructive pulmonary disease	75
	Asthma	333

*These participant counts are based on either the data from the participant's self-reported questionnaire or the information obtained from their hospitalizations. These numbers could shift depending on the results of the laboratory tests

of hypertension. We observed consistent and strong *AGT* gene signals linked with SBP and DBP, *MLGN1* gene linked with SBP and HTN, and epistasis of *TNXXB* gene and known genetic variants linked with all BP traits. In the GWAS analysis, we

identified consistent signals on the *ZBED9* gene associated with HTN in the genome-wide borderline threshold after adjusting for different environmental predictors. Our finding on *ZBED9* gene was confirmed by linkage analysis in an independent sample for all BP traits. Furthermore, single-locus analysis identified two missense variants in *ZBED9* (rs450630) and *AGT* (rs4762) associated with hypertension. Interestingly, the G allele of rs450630 exhibited an antagonistic effect on hypertension, but IGENE analysis revealed significant epistasis effects for different combinations of *ZBED9*, *AGT*, and *TNXXB* loci in the further analysis [12].

Future direction

In the upcoming research, we plan to concentrate on two areas of investigation. Firstly, we aim to study individuals with MH, commonly observed in childhood or adolescence and characterized by elevated blood pressure levels and resistance to standard treatment. Our population will be explored for mutations that enhance renal sodium reabsorption through mineralocorticoid-dependent or independent mechanisms, resulting in fluid retention and increased blood pressure. The carriers will be identified and excluded from the case group while doing the association analysis for complex hypertension.

Second, we intend to calculate the PRS for hypertension in our population. The PRS is a promising tool for predicting the risk of complex diseases. However, it has been shown that the performance of PRS may vary across different ethnic populations due to differences in the underlying genetic architecture of these populations. Therefore, we aim to identify the associated genetic variants and their effects on hypertension in our study population. Then, we will calculate the PRS for each individual using a standardized approach that considers the population-specific genetic variants and their effect sizes. We will also assess the performance of the PRS in our study population and compare it to other populations to evaluate the generalizability of the PRS across different ethnicities. This finding will allow us to determine the clinical utility of the PRS for hypertension prediction in our population and potentially other populations with similar genetic backgrounds. Calculating the PRS for hypertension will provide valuable insights into the genetic basis of hypertension and facilitate the development of personalized prevention and treatment strategies for this complex disease.

Diabetes

Objective

T2D is a serious and widespread disease associated with increased mortality and the development of CVD [13] on

a global scale. Given the prevalence of T2D in the Iranian population, our study seeks to examine the genetic architecture underlying the disease. Specifically, we aim to identify and characterize genetic variants associated with an increased risk of T2D in the Iranian population and to understand better the mechanisms underlying this disease in this specific population. By elucidating the genetic underpinnings of T2D in the Iranian population, we hope to develop more targeted and effective interventions for preventing and treating this debilitating disease. Our study may also contribute to a broader understanding of the genetic basis of T2D and inform the development of clinical strategies for managing this disease in other populations.

Key methods and data collection

As an initial step in finding the genetic architecture of T2D in Iranians, we evaluated the segregation, aggregation, and family-based heritability among TCGS participants. Additionally, we calculated a restricted weighted PRS ($p \leq 5e-8$) for TCGS participants and assessed its association with T2D incidence.

Main results

Familial aggregation and heritability of T2D were estimated by utilizing the TCGS family structure [14]. We constructed 2,594 constituent pedigrees based on 13,741 individuals aged over 20 (mean \pm sd: 39.71 ± 16.56 years), where the familial aggregation of T2D was found to be significantly significant ($p < 0.05$), and family-based heritability indicated that genetic variation accounted for 65% of T2D development and expression ($SE = 0.034$). Complex segregation analysis showed that the polygenic model was a good fit for illustrating the mode of inheritance of T2D among the TCGS participants. The risk of parental effect was higher than that of siblings within first-degree relatives ($OR = 4.11$ vs. $OR = 1.65$), and family history of T2D among first-degree relatives was more significant than the second-degree relatives ($OR = 3.84$ vs. $OR = 0.59$).

Later, as a first step toward predicting T2D development using a person-specific genetic profile, Moazzam-Jazi et al. identified multiple T2D-associated SNPs that were significantly enriched in the TCGS cohort compared to the global population [15]. These results could partly explain the differences in drug response and subsequent treatment efficiency among cases with diverse ancestries. The cumulative effect of enriched risk SNPs was assessed by computing the PRS for adults aged 20 years or older. A significant association was found between the PRS and T2D incidence in the TCGS cohort. Hence, the high genetic burden of T2D across the Iranian population could contribute to the enhanced prevalence of the disease in this population. They

also demonstrated a high hazard of T2D development in the genetically high-risk individuals compared to the genetically low-risk individuals in the model adjusted for age, sex, BMI, and other biochemical T2D risk factors [15].

Future direction

Shortly, we intend to focus on the genetic characterization of different types of diabetes, including T1D, T2D, and MODY. It is essential to differentiate between these three types of diabetes for effective treatment and management at different healthcare system levels. For example, sulfonylureas, which stimulate insulin secretion, can be effective for treating T2D and MODY 3 subtypes but may be less effective or even harmful in patients with T1D due to the absence of functional beta cells. Therefore, accurate diagnosis and classification of diabetes are critical steps toward the personalized treatment and management of this disease. It can be feasible by characterizing the key and novel genetic variants in diabetic patients in the Iranian population.

Furthermore, making the PRS based on associated genetic variants and their weight in the Iranian population is necessary. So far, this population has not been part of the global effort for GWASs, and we do not know which variants with what weight are best for creating a PRS for the Iranian population. As such, conducting a large-scale GWAS in the Iranian population is crucial to identify and characterize the genetic variants associated with T2D and other types of diabetes. By doing so, we can develop more accurate and effective PRS models for predicting the risk of diabetes in this population. Additionally, integrating genetic data with clinical and lifestyle factors can further enhance the precision of personalized diabetes management and prevention strategies. Ultimately, this may lead to better health outcomes and a reduced burden of diabetes on the Iranian healthcare system.

Lipid profile

Objective

Due to the central role of lipid profile in the development of cardiometabolic disease, we performed candidate gene analyses to identify variants associated with lipids in the Iranians. We also investigated the potential of genomic prediction for lipid profile traits.

Key methods and data collection

In the TCGS population, we utilized a nested case–control approach to investigate two SNPs related to CHD incidence. In addition, we proposed a novel strategy to identify the optimal number of SNPs with the most contribution to the

explanation of genomic phenotypic variation. This strategy employed a tenfold-10-repeat cross-validation method that utilized both WGR and GWAS. By implementing this approach, we aimed to enhance the computational efficiency of the making GRM in gBLUP [16]. Furthermore, we evaluated the strategy on lipid traits, including HDL-C, LDL-C, TG, and TC among TCGS participants.

Main results

Our findings demonstrated that certain genetic variations, specifically rs2048327-G (SLC22A3) and rs17465637-C (*MIA3*), increase the risk of CHD. The risk is around twice higher in males with the rs2048327-G allele and females with the rs17465637-C allele. We have observed that in male carriers of the rs2048327-G allele, HDL-C levels can significantly increase the likelihood of developing CHD in the future [17]. Additionally, our study has shown a significant association between risk allele of rs7865618 and CHD development in the TCGS population ($p=0.03$, OR = 1.73, CI95%:1.04–2.88) [18].

In addition, we aimed to investigate whether the *CETP* gene polymorphisms of rs5882 and rs3764261 impact the relationship between diet and changes in serum lipid profile. For this purpose, we selected 4700 individuals aged 18 and above from the TCGS participants and assessed changes in their serum lipid profile after 3.6 year follow-up period. Our results showed that carriers of the rs3764261-A allele had a greater reduction in TC levels when consuming a higher quartile of fish intake than those with the CC genotype. Conversely, the carriers of the rs5882-G allele showed an ascending trend in TG levels across quartiles of total fat, monounsaturated, and saturated fat consumption compared to those with the AA genotype. In contrast, carriers of the rs5882-G showed a declining trend in mean changes in TG concentrations across quartiles of carbohydrate intake compared to those with the AA genotype [19].

We conducted a study to identify informative SNPs that could explain the genotypic heritability of lipid traits. These findings showed that the highest prediction accuracy in predicting these traits was achieved when considering all SNPs. In contrast, only subsets of SNPs associated with these traits, as obtained from previous GWAS, resulted in the lowest prediction accuracy. However, the subset of SNPs referred to as "truly influential SNPs" showed interesting results in capturing significant genotypic variance and contributing to heritability [16].

Additionally, Sung's two-step method [20] was employed to identify pleiotropic genetic variants that exhibited a significant association with the longitudinal data of HDL-C, LDL-C, TC, and TG. Initially, a three-level GLMM was fitted for each longitudinal trait as a response variable, followed by a simultaneous genetic association

test via the GQLSM for each SNP. The results indicated that twenty variants from six genes, including *C16orf95*, *SLC12A3*, *CETP*, *NLRC5*, *ESRP2*, and *C16orf95* genes, were strongly associated ($p\text{-value} < 6.6 \times 10^{-5}$) with HDL-C, TC, and TG [21].

Future direction

The TCGS population presents a valuable opportunity to study monogenic lipid disorders, as it features a high rate of consanguineous marriages and extensive familial data. This opportunity allows for exploring the genetic basis of FH and other monogenic lipid disorders and investigating underlying molecular mechanisms. The genomic study on lipids in the TCGS population may uncover potential novel targets for therapeutic interventions.

In addition, the development of PRS can aid in comprehending the intricate genetic basis of diseases and their occurrence in specific populations. PRS can identify high-risk individuals for lipid metabolism disorders, including FH, and provide personalized preventive and therapeutic strategies. Integrating PRS with monogenic lipid disorder studies may also help explain the significant lipid abnormalities in the Iranian population. The studies could provide valuable insights into the Iranian population's genetic and environmental determinants of lipid metabolism and cardiovascular disease.

Obesity

Objective

The absence of studies on the genetic factors contributing to obesity in the Iranian population is a significant limitation for advancing personalized medicine. Therefore, to address this knowledge gap, our objective was to identify genetic variants associated with obesity and related traits, including WC, WHR, TG, TC, LDL-C, and HDL-C, and to assess their aggregated effects on the incidence of obesity among Iranians.

Key methods and data collection

After implementing quality control measures on the data, we utilized various regression tests to perform association analyses. All models were adjusted for the relevant covariates, such as age and sex. For PRS calculation, we used weighted PRS. We applied a false discovery rate (FDR) correction at the 5% significance level to account for multiple testing.

Main results

We conducted a family-based linkage and linkage disequilibrium analysis of 3,109 pedigrees in the first comprehensive study on Iranian pedigrees. Our results showed that *RPGRIP1L* is the key gene within the 16q12.2 region, and its polymorphisms could be associated with obesity risk factors among TCGS participants [22]. Moreover, we found that different SNP clusters composed of rare and common SNPs within the 16q12.2 region significantly increased BMI among Iranians. These clusters were randomly distributed across the region, with a higher density around *FTO*, *AIK-TIP*, and *MMP2* genes [23].

In a separate study, we discovered that nine correlated SNPs located upstream of the *PPARG* gene are significantly involved in the occurrence of long-term and persistent obesity [24]. Four SNPs in the *MC4R* gene are also significantly associated with the percentage of excess weight loss (EWL%) and BMI weight loss (EBMIL%), particularly after bariatric surgery lasting 12 months [25]. Moreover, rs13107325 was significantly associated with the increased likelihood of persistent metabolically healthy obesity in menopausal women [26].

FTO is represented as one of the central genes involved in obesity and its corresponding traits. Several *FTO* variants, including rs1421085, rs1558902, rs1121980, and rs8050136) were significantly associated with the MUO phenotype even after adjusting for lipid profile. However, no significant association was observed between those SNPs and metabolically healthy obesity [27]. Another study investigated the interaction between dietary patterns and *FTO* polymorphisms regarding changes in BMI and WC over 3·6 year follow-up period [28]. Six common SNPs (rs1421085, rs1121980, rs17817449, rs8050136, rs9939973, and rs3751812) within the *FTO* gene region were examined. The study revealed that individuals with the risk alleles and higher WDP scores had nearly a two-fold higher BMI than those without the risk alleles.

In individuals with a higher PRS, BMI, and WC tend to increase along with increasing WDP score [29]. This finding suggests that individuals with a genetic predisposition to obesity are more susceptible to the detrimental effects of an unhealthy diet and emphasizes the importance of reducing the consumption of unhealthy foods to prevent obesity. Additionally, WC increased with increasing WDP score in carriers of the risk alleles of rs1121980 and rs3751812 but not those without any risk alleles. A higher intake of TFAs in adults carrying the *FTO* rs8050136 risk allele was also found to significantly increase BMI and WC over an average follow-up of 3.6 years [29]. However, no significant interaction was found between combined *FTO* variants (rs1121980, rs14211085, and rs8050136) and dietary

diversity score concerning general obesity, indicating that dietary diversity patterns may play a mediatory role in the presentation of obesity-related factors [30].

A healthy dietary pattern could modify the impact of *MC4R* rs17782313 on general obesity. The interaction between the risk allele of rs17782313 and a higher healthy dietary pattern score results in a lower risk of prevalent obesity than those without the risk allele [31]. In a recent study, we identified that eight common SNPs in or near the *MC4R* gene are significantly associated with increased BMI, WC, and WHR over a lifetime. Interestingly, they found that the aggregated effect of these SNPs significantly influences increased BMI and WC only in early adulthood, not during the middle or early adulthood stages. Therefore, the effect of *MC4R* risk SNP may not remain constant during the lifetime [32].

Future direction

The etiology of monogenic obesity differs substantially from the more prevalent form of obesity. Our initial approach involves screening known and suggested mutations associated with monogenic obesity to advance our understanding of this complex trait. In the further step, we aim to conduct a GWAS that encompasses various subtypes of obesity, such as normal weight obese, metabolically obese normal weight, metabolically healthy obese, and metabolically unhealthy obese. By doing so, we hope to identify potential genetic markers associated with the various forms of obesity and gain insight into the underlying mechanisms driving this condition.

Metabolic syndrome

Objective

Metabolic syndrome (MetS) is a complex disease characterized by metabolic disorders such as abdominal obesity, dyslipidemia, hyperglycemia, and HTN. The development of MetS is primarily influenced by environmental factors such as an inappropriate diet and physical inactivity [33, 34]. Also, genetic factors play a role, as suggested by familial aggregation and heritability studies [35, 36]. Researchers have identified several locations associated with an increased risk of this syndrome [37]. By examining the interaction between these genetic variants and dietary factors associated with MetS, new strategies for preventing and treating MetS can be developed. We investigated the role of genetic variations, advanced statistical models, and gene-nutrient interactions in the risk of MetS.

Key methods and data collection

The first study involved a retrospective cohort study of 5,666 participants from the TCGS. The aim was to examine the relationship between MetS and its components with three *GCKR* polymorphisms (rs780093, rs780094, and rs1260326) using linear and logistic regression analyses in an additive genetic model. Moreover, the Cox regression analysis was performed to evaluate the association of these variants with the development of MetS over time [38].

Furthermore, several studies with smaller sample sizes have investigated gene-diet interactions in TCGS participants to predict the risk of MetS. One candidate gene study analyzed *GCKR* gene variants and clinical and demographic information on 4,756 eligible TCGS participants to develop an optimal prediction model (s) for MetS. Then, predictive models were compared (logistic regression (LR), Random Forest (RF), decision tree (DT), support vector machines (SVM), and discriminant analyses) [42]. Continuing to find optimal statistical models for predicting MetS, we designed a study that evaluated the association of MetS and three genes, namely *BUD13*, *ZPR1*, and *APOA5*, with 18 SNPs in 5,421 TCGS participants. This study was a cross-sectional study that employed two models for data analysis. The first model examined the association between variants and MetS, while the second model (HTG-MetS) evaluated the associations between genetic variants and MetS patients with high plasma TG levels. Four-gamete rules were also used to form SNP sets from correlated SNPs. The kernel machine regression models and single SNP regression were employed to estimate the association between SNP sets and MetS [43]. Two studies subsequently examined the association of *CETP* gene polymorphisms (rs5882 and rs3764261) in 441 MetS cases and 844 matched controls and *TCF7L2* gene variants (rs7903146 and rs12255372) in 1,423 individuals with dietary intakes to predict the risk of MetS. [39, 40].

Main results

The study showed that functional *GCKR* variants were associated with higher TG and lower fasting blood sugar (FBS) levels. Moreover, the results of Cox-adjusted model regression revealed that carriers of rs780094, rs780093, and rs1260326 TT genotypes had a higher risk of MetS incidence [38]. The logistic regression model showed a significant association of MetS with age, gender, schooling years, BMI, and physical activity, rs780094 and rs780093. Random Forest analysis revealed that BMI, physical activity, and age are the most influential model features. Decision tree analysis showed that individuals with BMI < 24 and physical activity < 8.8 had a lower risk of developing MetS [41]. In another study, the kernel machine analysis showed that two sets of over three sets of correlated SNPs have a significant

joint effect on MetS and HTG MetS models. Moreover, a single SNP regression analysis indicated that the highest OR in the HTG MetS model was for the G allele in rs2266788 (MetS: OR = 1.3, HTG MetS: OR = 1.4) and the T allele in rs651821 (MetS: OR = 1.3, HTG MetS: OR = 1.4). Although both models had the same ORs, the p-values in the HTG MetS model were marginally more significant [42].

Another study investigated the potential relationship between a specific genetic variant (rs5882) and dietary macronutrient intake concerning metabolic syndrome (MetS) risk. The results indicate that this genetic variant does not interact with macronutrient intake concerning MetS risk. However, the study found that individuals carrying the G allele and consuming monounsaturated fatty acids and total fat in the lowest quartile had a reduced risk of low HDL-C. Conversely, those carrying the G allele and consuming higher levels of trans-fatty acids had an increased risk of high blood pressure [39]. We also observed that consuming nuts in the highest tertile was associated with a reduced risk of MetS among T allele carriers of rs12255372, resulting in a 34% reduction of MetS risk [40].

Future direction

Metabolic syndrome is a complex condition characterized by three or more risk factors out of five. Each risk factor has unique genetic variations. Therefore, it may be beneficial to cluster individuals based on their combination of risk factors and track them longitudinally. We can identify individuals with similar genetic profiles and risk factor combinations by doing so. This information can then be used to perform genetic analyses specific to each group, allowing for more targeted and personalized treatment options. This approach may improve the effectiveness of treatments by considering the specific genetic and metabolic factors present in each individual.

Others

Run of homozygosity (ROH)

Objective

Consanguineous marriages have traditionally been favored in certain cultures. However, this practice may increase the inbreeding coefficient and the run of homozygosity (ROH). Inbreeding can negatively affect various fitness-related characteristics in humans, leading to inbreeding depression. Consanguineous families are likelier to have children with major abnormalities, including severe genetic diseases and mutant phenotypes that can be lethal in early life. Hence, it is essential to understand the genetic basis of these effects.

The initial aim was to estimate the effect of the fraction of ROH and the FROH on a handful of quantitative and binary traits of TCGS participants along with 118 other studies and to compare the ROH pattern between the nations [43].

Key methods and data collection

The TCGS cohort was one of the 119 independent genetic epidemiological study cohorts participating in the ROH-gen consortium. For this analysis, 11,760 participants were included. These participants were classified into Western Asian/Persian groups, excluding ethnic outliers, duplicates, gender mismatches, and unresolvable pedigree mismatches. The study identified ROH greater than 1.5 Mb in length for 18 traits, categorized into ten groups (Table 7 S).

Main results

The analysis of the TCGS cohort revealed a mean of FROH greater than 1.5 Mb of 0.017, with a standard error of 0.026. Additionally, We calculated the fixation index (FIS), which measures inbreeding resulting from non-random mating in the most recent generation. The study found a correlation between FROH and other measures of inbreeding, including FSNP (correlation coefficient of 0.98 FGRM (correlation coefficient of 0.98), and FSN_OutsideROH (correlation coefficient of 0.091). This consortium data analysis suggests that the TCGS cohorts exhibit high consanguinity compared with other studies.

COVID-19

Objective

Our main goal was to investigate the molecular processes underlying the development of SARS-CoV-2 infection from a genetic perspective, considering both human and viral genetic factors. To achieve this goal, we conducted three distinct studies to examine the impact of genetic variations in the *ACE2* gene, which encodes one of the main SARS-CoV-2 receptors and explore the role of human long non-coding RNAs (lncRNAs) in recognizing and responding to the virus. Additionally, we used computer simulations to model and analyze the essential molecular interactions between coronaviruses and the human genome.

Key methods and data collection

In the first study, we utilized structural bioinformatics simulation techniques to identify and analyze all the genetic polymorphisms present in the *ACE2* and *TMPRSS2* genes within TCGS participants. We then assessed how these polymorphisms influenced the ability of the virus affinity. The

second study examined the interaction between human long non-coding RNAs and the SARS-CoV-2 genome by analyzing publicly available RNA sequencing data using various in silico approaches. Additionally, we identified the specific regions of the SARS-CoV-2 genome that physically interact with lncRNAs responsive to infection. Lastly, we utilized a high-throughput computational approach in the third study to investigate the likelihood of interactions between host RNA-viral and viral RNA-host proteins.

Main results

In the first study, we identified 570 genetic variations, including SNP and INDEL, located near or within the *ACE2* gene among TCGS participants. Interestingly, two missense variants, K26R and S331F, were found to reduce the receptor's affinity for the viral spike protein, with only K26R previously reported. We also investigated the important details of *ACE2*-Spike and *ACE2*-*TMPRSS2* interactions, emphasizing the essential role of Arg652 of *ACE2* in the protease function of *TMPRSS2* [44].

In the second study, we found that three trans-acting lncRNAs (*HOTAIRM1*, *PVT1*, and *AL392172.1*) were responsible for establishing more than half of the interactions between lncRNAs and protein-coding genes (PCGs) in bronchoalveolar lavage fluid samples. These lncRNAs also exhibited a high affinity for binding to the SARS-CoV-2 genome, suggesting their significant regulatory role during SARS-CoV-2 infection. Additionally, we identified the involvement of *MALAT1* and *NEAT1* in the development of inflammation in SARS-CoV-2 infected cells. Our analysis also revealed that the 5' part of the SARS-CoV-2 genome interacts with many human lncRNAs, while the 3' part does not [45].

Finally, our RNA-protein interaction study demonstrated that key viral proteins involved in viral genome replication and transcription are conserved during evolution and have restricted interaction abilities to maintain their functions. Undesired interactions do not perturb the functions of these proteins. In contrast, the non-structural protein 3 (*NSP3*) of SARS-CoV-2 exhibits a hypermutation rate, which endows it with the affinity to interact with diverse host cell RNAs [46].

Main strengths and weaknesses

The TCGS project has numerous strengths, including the longitudinal tracking of participants from an early age that has allowed for identifying risk factors for complex cardiovascular outcomes over time. Additionally, the project has provided insights into the familial aggregation and genetic determinants of cardiometabolic risk factors across multiple generations. The study's measurements were performed by trained technicians, not self-reported, which adds credibility. Moreover, the project collected phenotypic data from

participants of all ethnicities and age groups, and these data were coded based on the last version of ICDs. Nonetheless, our works have some limitations. First, some genotyped participants do not have sufficient information about CRF, or there might be differences in phenotype measurement and definition between parents' studies. Hence, some information will be lost during genetic analysis. Also, due to the small sample size in TCGS compared to the Iranian population and possible ethnic differences, the findings may not be generalizable to all Iranians, indicating the need for national and international replication. The last limitation is the lack of diversity in the clinical phenotypes, which reduces the possibility of examining and evaluating all disease categories. We lack a sufficient sample size to evaluate all disease categories. In the future, it may be possible to validate the self-reports via patient visits in the clinic and perform additional examinations to improve measurement accuracy.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s10654-023-01008-1>.

Acknowledgements The authors would like to express gratitude to the patients participating in the Tehran lipid and glucose study, Tehran Obesity Treatment Study, Thyroid cancer study, and clinical follow-up study. Also, special thanks to the deCODE genetic company for doing the genetic screening.

Author contributions All authors contributed to the study's conception and design. Material preparation, data collection, and analysis were performed by MD, BS-K, KG, MH, MA, HL, MM, SM, AZ, LN, MZ, NJ, GF, and FA. SA and BS did ICD11 coding and clinical evaluation. SZ, MB, and MSF added their study samples to TCGS. PM, FH, AG, PA, FH, MV, DK, FR, AG, AAM, MP, AZV, MT, and SA qualified data from questionnaires and add their scientific support. SJ helped organize socio-determinant information, and PR helped evaluate family relations. FH evaluates nutrition-related data. AD and PA performed lab analysis, and RM, an IT man, helped with the project. MD wrote the first draft of the manuscript, and all authors commented on previous versions. All authors read and approved the final manuscript.

Funding This study was funded by the Research Institute for Endocrine Sciences, Shahid Beheshti University of Medical Sciences (Tehran, Iran), and the scientific and financial support of the deCODE genetic company (Reykjavik, Iceland). This work was supported by the Ministry of Health and Medical Education (MOHME:700/1736). Shahid Beheshti University for Medical Sciences was granted this project [grant no: 31178]. Iranian molecular medicine network supported the genomic bank.

Data availability The data underlying this article cannot be shared publicly due to the privacy of individuals participating in the study. The data will be shared on reasonable request with the corresponding author.

Declarations

Conflict of interest The authors have no relevant financial or non-financial interests to disclose.

Ethical approval The ethics committee has approved the TCGS project on human subject research at the Research Institute for Endocrine Sci-

ences, Shahid Beheshti University of Medical Sciences (code of "IR.SBMU.ENDOCRINE.REC.1395.366"), which were under the 1964 Helsinki Declaration and its later amendments or comparable ethical standards.

Consent to participate Written informed consent was obtained from all participants.

Informed consent This study obtained approval from the National Committee for Ethics in biomedical research in December 2012. Informed written consent was obtained from all participants.

References

- Haldorsen EMH. The right treatment to the right patient at the right time. *Occup Environ Med.* 2003;60:235–6.
- Azizi F, Ghanbarian A, Momenan AA, Hadaegh F, Mirmiran P, Hedayati M, et al. Prevention of non-communicable disease in a population in nutrition transition: Tehran Lipid and Glucose Study phase II. *Trials.* BioMed Central; 2009;10:5.
- Azizi F, Zadeh-Vakili A, Takyar M. Review of rationale, design, and initial findings: Tehran Lipid and Glucose Study. *Int J Endocrinol Metab.* 2018;16:e84777.
- Barzin M, Hosseiniapanah F, Motamedi MA, Shapoori P, Arian P, Daneshpour MA, et al. Bariatric surgery for morbid obesity: Tehran Obesity Treatment Study (TOTS) rationale and study design. *JMIR Res Protoc.* 2016;5:e8.
- Hedayati M, Zarif Yeganeh M, Sheikholeslami S, Afsari F. Diversity of mutations in the RET proto-oncogene and its oncogenic mechanism in medullary thyroid cancer. *Crit Rev Clin Lab Sci.* 2016;53:217–27. <https://doi.org/10.3109/10408363.2015.1129529>.
- Daneshpour MS, Fallah M-S, Sedaghati-Khayat B, Guity K, Khalili D, Hedayati M, et al. Rationale and design of a Genetic Study on Cardiometabolic Risk Factors: Protocol for the Tehran Cardiometabolic Genetic Study (TCGS). *JMIR Res Protoc* 2017;6:e28. Available from: <https://pubmed.ncbi.nlm.nih.gov/28232301/>
- ICD-11 for Mortality and Morbidity Statistics. Available from: <https://icd.who.int/browse11/l-m/en>
- Mills KT, Stefanescu A, He J. The global epidemiology of hypertension. *Nat Rev Nephrol.* 2020;16:223.
- Kolifarhood G, Daneshpour MS, Khayat BS, Saadati HM, Guity K, Khosravi N, et al. Generality of genomic findings on blood pressure traits and its usefulness in precision medicine in diverse populations: a systematic review. *Clin Genet.* 2019;96.
- Kolifarhood G, Daneshpour MS, Zahedi AS, Khosravi N, Sedaghati-Khayat B, Guity K, et al. Familial genetic and environmental risk profile and high blood pressure event: a prospective cohort of cardio-metabolic and genetic study. *Blood Press Taylor and Francis Ltd.* 2021;30:196–204.
- Kolifarhood G, Sabour S, Akbarzadeh M, Sedaghati-khayat B, Guity K, Rasekhi Dehkordi S, et al. Genome-wide association study on blood pressure traits in the Iranian population suggests ZBED9 as a new locus for hypertension. *Sci Rep* 2021;11.
- Akbarzadeh M, Riahi P, Kolifarhood G, Lanjanian H, Alipour N, Bonab LNH, et al. The AGT epistasis pattern proposed a novel role for ZBED9 in regulating blood pressure: Tehran Cardiometabolic genetic study (TCGS). *Gene.* 2022;831:146560.
- Martín-Timón I, Sevillano-Collantes C, Segura-Galindo A, del Cañizo-Gómez FJ. Type 2 diabetes and cardiovascular disease: Have all risk factors the same strength? *World J Diabetes.* 2014;5:444.

14. Akbarzadeh M, Riahi P, Ramezankhani A, Dehkordi SR, Roudbar MA, Zarkesh M, et al. Parental transmission plays the major role in the high aggregation of type 2 Diabetes in Iranian families: Tehran Lipid and Glucose Study (TLGS). *Can J Diabetes*. 2021.
15. Moazzam-Jazi M, Najd Hassan Bonab L, Zahedi AS, Daneshpour MS. High genetic burden of type 2 diabetes can promote the high prevalence of disease: a longitudinal cohort study in Iran. *Sci Rep*. 2020;10:1–11.
16. Akbarzadeh M, Dehkordi SR, Roudbar MA, Sargolzaei M, Guity K, Sedaghati-Khayat B, et al. GWAS findings improved genomic prediction accuracy of lipid profile traits: Tehran Cardiometabolic Genetic Study. *Sci Rep*. 2021;11:1–9.
17. Najd Hassan Bonab L, Moazzam-Jazi M, Miri Moosavi RS, Fallah MS, Lanjanian H, Masjoudi S, et al. Low HDL concentration in rs2048327-G carriers can predispose men to develop coronary heart disease: Tehran Cardiometabolic genetic study (TCGS). *Gene*. 2021;778:145485.
18. Matoo S, Fallah MS, Daneshpour MS, Mousavi R, Khayat BS, Hasanzad M, et al. Increased risk of CHD in the presence of Rs7865618 (A allele): Tehran lipid and glucose study. *Arch Iran Med*. 2017;20:153–7.
19. Hosseini-Esfahani F, Esfandiari Z, Mirmiran P, Daneshpour MS, Ghanbarian A, Azizi F. The interaction of cholesteryl ester transfer protein gene variations and diet on changes in serum lipid profiles. *Eur J Clin Nutr*. 2019;73:1291–8.
20. Sung Y, Feng Z, Subedi S. A genome-wide association study of multiple longitudinal traits with related subjects. *Stat*. 2016;5:22–44.
21. Hosseinzadeh N, Mehrabi Y, Daneshpour MS, Zayeri F, Guity K, Azizi F. Identifying new associated pleiotropic SNPs with lipids by simultaneous test of multiple longitudinal traits: an Iranian family-based study. *Gene*. 2019;692:156–69.
22. Javanrouh N, Soltanian AR, Tapak L, Azizi F, Ott J, Daneshpour MS. A novel association of rs13334070 in the RRGRI1L gene with adiposity factors discovered by joint linkage and linkage disequilibrium analysis in Iranian pedigrees: Tehran Cardiometabolic Genetic Study (TCGS). *Genet Epidemiol*. 2019;43:342–51.
23. Javanrouh N, Daneshpour MS, Soltanian AR, Tapak L. Kernel machine SNP set analysis provides new insight into the association between obesity and polymorphisms located on the chromosomal 16q.12.2 region: Tehran Lipid and Glucose Study. *Gene*. 2018;658:146–51.
24. Javanrouh N, Khalaj A, Guity K, Sedaghati-Khayat B, Valizadeh M, Barzin M, et al. Presence of CC genotype for rs17773430 could affect the percentage of excess weight loss 1 year after bariatric surgery: Tehran Obesity Treatment Study (TOTS). *Obes Surg United States*. 2020;30:537–44.
25. Javanrouh N, Khalaj A, Guity K, Sedaghati-khayat B, Valizadeh M, Barzin M, et al. Presence of CC genotype for rs17773430 could affect the percentage of excess weight loss 1 year after bariatric surgery: Tehran Obesity Treatment Study (TOTS). *Obes Surg*. 2020;30:537–44.
26. Gharooi Ahangar O, Javanrouh N, Daneshpour MS, Barzin M, Valizadeh M, Azizi F, et al. Genetic markers and continuity of healthy metabolic status: Tehran cardio-metabolic genetic study (TCGS). *Sci Rep*. 2020;10.
27. Sedaghati-khayat B, Barzin M, Akbarzadeh M, Guity K, Fallah M-S, Pourhassan H, et al. Lack of association between FTO gene variations and metabolic healthy obese (MHO) phenotype: Tehran Cardio-metabolic Genetic Study (TCGS). *Eating and Weight Disorders*. 2020;25.
28. Hosseini-Esfahani F, Koochakpour G, Mirmiran P, Daneshpour MS, Azizi F. Dietary patterns modify the association between fat mass and obesity-associated genetic variants and changes in obesity phenotypes. *Br J Nutr*. 2019;121:1247–54.
29. Koochakpour G, Esfandiari Z, Hosseini-Esfahani F, Mirmiran P, Daneshpour MS, Sedaghati-Khayat B, et al. Evaluating the interaction of common FTO genetic variants, added sugar, and trans-fatty acid intakes in altering obesity phenotypes. *Nutr Metab Cardiovasc Dis*. 2019;29:474–80.
30. Goodarzi G, Hosseini-Esfahani F, Ataie-Jafari A, Haji-Hosseini-gazestani N, Daneshpour MS, Keshavarz SA, et al. Dietary diversity modifies the association between FTO polymorphisms and obesity phenotypes. *Int J Food Sci Nutr*. 2021;72:997–1007.
31. Mousavizadeh Z, Hosseini-Esfahani F, Javadi A, Daneshpour MS, Akbarzadeh M, Javadi M, et al. The interaction between dietary patterns and melanocortin-4 receptor polymorphisms in relation to obesity phenotypes. *Obes Res Clin Pract*. 2020;14:249–56.
32. Moazzam-Jazi M, Sadat Zahedi A, Akbarzadeh M, Azizi F, Daneshpour MS. Diverse effect of MC4R risk alleles on obesity-related traits over a lifetime: Evidence from a well-designed cohort study. *Gene*. 2022;807.
33. Perez-Martinez P, M Phillips C, Delgado-Lista J, Garcia-Rios A, Lopez-Miranda J, Perez-Jimenez F. Nutrigenetics, metabolic syndrome risk and personalized nutrition. *Curr Vasc Pharmacol*. 2013;11:946–53.
34. Farhadnejad H, Parastouei K, Rostami H, Mirmiran P, Azizi F. Dietary and lifestyle inflammatory scores are associated with increased risk of metabolic syndrome in Iranian adults. *Diabetol Metab Syndr*. 2021;13:1–10.
35. Zarkesh M, Daneshpour MS, Faam B, Fallah MS, Hosseinzadeh N, Guity K, et al. Heritability of the metabolic syndrome and its components in the Tehran Lipid and Glucose Study (TLGS). *Genet Res (Camb)*. 2012;94:331–7.
36. Naseri P, Khodakarim S, Guity K, Daneshpour MS. Familial aggregation and linkage analysis with covariates for metabolic syndrome risk factors. *Gene*. 2018;659:118–22.
37. Yamada Y, Kato K, Oguri M, Horibe H, Fujimaki T, Yasukochi Y, et al. Identification of four genes as novel susceptibility loci for early-onset type 2 diabetes mellitus, metabolic syndrome, or hyperuricemia. *Biomed Rep*. 2018;9:21–36.
38. Zahedi AS, Akbarzadeh M, Sedaghati-Khayat B, Seyedhamzehzadeh A, Daneshpour MS. GCKR common functional polymorphisms are associated with metabolic syndrome and its components: a 10-year retrospective cohort study in Iranian adults. *Diabetol Metab Syndr BioMed Central*. 2021;13:1–10.
39. Esfandiari Z, Hosseini-Esfahani F, Daneshpour MS, Zand H, Mirmiran P, Azizi F. Cholesteryl ester transfer protein gene variations and macronutrient intakes interaction in relation to metabolic syndrome: Tehran lipid and glucose study. *Iran J Basic Med Sci*. 2018;21:586–92.
40. Hosseinpour-Niazi S, Bakhshi B, Zahedi A-S, Akbarzadeh M, Daneshpour MS, Mirmiran P, et al. TCF7L2 polymorphisms, nut consumption, and the risk of metabolic syndrome: a prospective population based study. *Nutr Metab*. 2021;18:1–11.
41. Akbarzadeh M, Alipour N, Moheimani H, Zahedi AS, Hosseini-Esfahani F, Lanjanian H, et al. Evaluating machine learning-powered classification algorithms which utilize variants in the GCKR gene to predict metabolic syndrome: Tehran Cardio-metabolic Genetics Study. *J Transl Med*. 2022;20:1–12.
42. Masjoudi S, Sedaghati-khayat B, Givi NJ, Bonab LNH, Azizi F, Daneshpour MS. Kernel machine SNP set analysis finds the association of BUD13, ZPR1, and APOA5 variants with metabolic syndrome in Tehran Cardio-metabolic Genetics Study. *Sci Rep*. 2021;11.
43. Clark DW, Okada Y, Moore KHS, Mason D, Pirastu N, Gandin I, et al. Associations of autozygosity with a broad range of human phenotypes. *Nat Commun*. 2019;10.
44. Lanjanian H, Moazzam-Jazi M, Hedayati M, Akbarzadeh M, Guity K, Sedaghati-khayat B, et al. SARS-CoV-2 infection

- susceptibility influenced by ACE2 genetic polymorphisms: insights from Tehran Cardio-Metabolic Genetic Study. *Sci Rep.* 2021;11.
45. Moazzam-Jazi M, Lanjanian H, Maleknia S, Hedayati M, Daneshpour MS. Interplay between SARS-CoV-2 and human long non-coding RNAs. *J Cell Mol Med.* 2021;25:5823–7.
46. Lanjanian H, Nematzadeh S, Hosseini S, Torkamanian-Afshar M, Kiani F, Moazzam-Jazi M, et al. High-throughput analysis of the interactions between viral proteins and host cell RNAs. *Comput Biol Med.* 2021;135:104611.

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

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.

Authors and Affiliations

Maryam S. Daneshpour¹  · Mahdi Akbarzadeh¹ · Hossein Lanjanian¹ · Bahar Sedaghati-khayat^{1,2} · Kamran Guity^{1,3} · Sajedah Masjoudi¹ · Asiyeh Sadat Zahedi¹ · Maryam Moazzam-Jazi¹ · Leila Najd Hassan Bonab¹ · Bita Shalbafan⁴ · Sara Asgarian¹ · Goodarz Koli Farhood¹ · Niloofar Javanrooh¹ · Maryam Zarkesh¹ · Parisa Riahi¹ · Mohammad Reza Moghaddas¹ · Parvaneh Arbab Dehkordi¹ · Azar Delbarpour Ahmadi¹ · Firoozeh Hosseini⁵ · Sara Jalali Farahani⁶ · Farzad Hadaegh⁷ · Parvin Mirmiran⁵ · Fahimeh Ramezani Tehrani⁸ · Arash Ghanbarian⁷ · Mohammad Sadegh Fallah Mahboob Pasand⁹ · Parisa Amiri⁶ · Majid Valizadeh¹⁰ · Farhad Hosseipanah¹⁰ · Maryam Tohidi⁷ · Asghar Ghasemi¹¹ · Azita Zadeh-Vakili¹¹ · Mohammad Piryaee¹ · Shahram Alamdari¹² · Davood Khalili⁷ · Amirabbas Momenan⁷ · Maryam Barzin¹⁰ · Sirous Zeinali⁹ · Mehdi Hedayati¹ · Fereidoun Azizi¹²

Mahdi Akbarzadeh
akbarzadehms@sbm.ac.ir

Hossein Lanjanian
H.Lanjanian@ut.ac.ir

Bahar Sedaghati-khayat
b.sedaghatikhayat@erasmusmc.nl

Kamran Guity
kag78@hi.is

Sajedah Masjoudi
sajedahmasjoudi@gmail.com

Asiyeh Sadat Zahedi
asie_zahedi@yahoo.com

Maryam Moazzam-Jazi
setareh227@gmail.com

Leila Najd Hassan Bonab
leila_nhb26@yahoo.com

Bita Shalbafan
b-shalbafan@alumnus.rums.ac.ir

Sara Asgarian
jf_sara@yahoo.com

Goodarz Koli Farhood
gfarhood@gmail.com

Niloofar Javanrooh
javanrouhn@gmail.com

Maryam Zarkesh
zarkesh@endocrine.ac.ir

Parisa Riahi
parisaariyahii@gmail.com

Mohammad Reza Moghaddas
mrezamoghaddas@gmail.com

Parvaneh Arbab Dehkordi
parbab2006@gmail.com

Azar Delbarpour Ahmadi
azar.delbar@yahoo.com

Firoozeh Hosseini
f.hosseini@sbm.ac.ir

Sara Jalali Farahani
sara.asgn@gmail.com

Farzad Hadaegh
fzhadaegh@endocrine.ac.ir

Parvin Mirmiran
mirmiran@endocrine.ac.ir

Fahimeh Ramezani Tehrani
ramezani@endocrine.ac.ir

Arash Ghanbarian
ghanbarian@endocrine.ac.ir

Mohammad Sadegh Fallah Mahboob Pasand
Sadegh.fallah@gmail.com

Parisa Amiri
amiri@endocrine.ac.ir

Majid Valizadeh
valizadeh@endocrine.ac.ir

Farhad Hosseipanah
fhospanah@endocrine.ac.ir

Maryam Tohidi
tohidi@endocrine.ac.ir

Asghar Ghasemi
ghasemi@endocrine.ac.ir

Azita Zadeh-Vakili
azitavakili@endocrine.ac.ir

Mohammad Piryaee
m.piryaee@gmail.com

Shahram Alamdari
alamdari@sbm.ac.ir

- Davood Khalili
ndavood@yahoo.com
- Amirabbas Momenan
momenan@endocrine.ac.ir
- Maryam Barzin
m.barzin7@gmail.com
- Sirous Zeinali
zeinalipasteur@yahoo.com
- Mehdi Hedayati
hedayati@endocrine.ac.ir
- ¹ Cellular and Molecular Endocrine Research Center, Research Institute for Endocrine Sciences, Shahid Beheshti University of Medical Sciences, P.O. Box 19195-4763, Tehran, Iran
 - ² Department of Internal Medicine, Erasmus MC, University Medical Center Rotterdam, Rotterdam, The Netherlands
 - ³ Faculty of Medicine, University of Iceland, School of Health Sciences, Reykjavik, Iceland
 - ⁴ Clinical Research Development Center of Labbafinejad Hospital, Shahid Beheshti University of Medical Sciences, Tehran, Iran
 - ⁵ Nutrition and Endocrine Research Center, Research Institute for Endocrine Sciences, Shahid Beheshti University of Medical Sciences, Tehran, Iran
 - ⁶ Research Center for Social Determinants of Health, Research Institute for Endocrine Sciences, Shahid Beheshti University of Medical Sciences, Tehran, Iran
 - ⁷ Prevention of Metabolic Disorders Research Center, Research Institute for Endocrine Sciences, Shahid Beheshti University of Medical Sciences, Tehran, Iran
 - ⁸ Reproductive Endocrinology Research Center, Research Institute for Endocrine Sciences, Shahid Beheshti University of Medical Sciences, Tehran, Iran
 - ⁹ Dr. Zeinali's Medical Genetics Lab, Kawsar Human Genetics Research Center (KHGRC), Tehran, Iran
 - ¹⁰ Obesity Research Center, Research Institute for Endocrine Sciences, Shahid Beheshti University of Medical Sciences, Tehran, Iran
 - ¹¹ Endocrine Physiology Research Center, Research Institute for Endocrine Sciences, Shahid Beheshti University of Medical Sciences, Tehran, Iran
 - ¹² Endocrine Research Center, Research Institute for Endocrine Sciences, Shahid Beheshti University of Medical Sciences, P.O. Box 19195-4763, Tehran, Iran