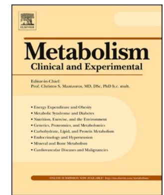
Metabolism 137 (2022) 155329

Contents lists available at ScienceDirect

Metabolism

journal homepage: www.journals.elsevier.com/metabolism



Genome-wide association and Mendelian randomization study of fibroblast growth factor 21 reveals causal associations with hyperlipidemia and possibly NASH

Susanna C. Larsson [a,b,*,1], Karl Michaëlsson [a,1], Marina Mola-Caminal [a], Jonas Höijer [a], Christos S. Mantzoros [c,d]

[a] Unit of Medical Epidemiology, Department of Surgical Sciences, Uppsala University, Uppsala, Sweden
[b] Unit of Cardiovascular and Nutritional Epidemiology, Institute of Environmental Medicine, Karolinska Institutet, Stockholm, Sweden
[c] Department of Medicine, Beth Israel Deaconess Medical Center, Harvard Medical School, Boston, MA, USA
[d] Section of Endocrinology, VA Boston Healthcare System, Harvard Medical School, Boston, MA, USA

ARTICLE INFO

Keywords:
Fibroblast growth factor 21
Alkaline phosphatase
Genome-wide association study
Hepatokine
Mendelian randomization
Single-nucleotide polymorphisms

ABSTRACT

Background: Fibroblast growth factor 21 (FGF21) is a hepatokine that produces metabolic benefits, such as improvements of lipid profile. We performed a genome-wide association study (GWAS) to identify genetic variants associated with circulating FGF21 and investigated the causal effects of FGF21 on pertinent outcomes using Mendelian randomization (MR).

Methods: We conducted a GWAS testing ~7.8 million DNA sequence variants with circulating FGF21 in a discovery cohort of 6259 Swedish adults with replication in 4483 Swedish women. We then performed two-sample MR analyses of genetically predicted circulating FGF21 in relation to alcohol and nutrient intake, cardiovascular and metabolic biomarkers and diseases, and liver function biomarkers using publicly available GWAS summary statistics data.

Results: Our GWAS identified multiple single-nucleotide polymorphisms with genome-wide significant associations ($P < 5 \times 10^{-8}$) with circulating FGF21 on chromosomes 2 and 19 in or near the *GCKR* and *FGF21* genes, respectively. The strongest signal at the *FGF21* locus (rs2548957, $\beta = 0.181$, $P < 2.18 \times 10^{-42}$) displayed in two-sample MR analyses robust associations with lower alcohol intake, lower circulating low-density lipoprotein cholesterol, apolipoprotein B, C-reactive protein, gamma-glutamyl transferase, and galectin-3 concentrations, and higher circulating insulin-like growth factor-I and alkaline phosphatase concentrations after correcting for multiple testing ($P < 0.0018$) whereas associations with fat mass, type 2 diabetes, and cardiovascular disease were largely null.

Conclusions: We identified robust associations of certain genetic variants in or near the *GCKR* and *FGF21* genes with circulating FGF21 concentrations. Furthermore, our results support a strong causal effect of FGF21 on improved lipid profile, reduced alcohol consumption and C-reactive protein concentrations, and liver function biomarkers including fibrosis. We found largely null or weak positive associations with fat mass, diabetes, and cardiovascular disease as well as higher insulin-like growth factor-I concentrations, which could indicate a compensatory increase to regulate the above FGF21 resistant states in humans.

1. Introduction

Fibroblast growth factor 21 (FGF21) is a hepatokine discovered in 2000 [1–4]. Administration of FGF21 to rodents or monkeys produces metabolic benefits, such as reduced fat mass and improved glycemic control, insulin sensitivity, and lipid profiles (i.e., decreased

Abbreviations: ALP, alkaline phosphatase; FGF21, fibroblast growth factor 21; GCKR, glucokinase regulatory protein; GWAS, genome-wide association study; IGF-I, insulin-like growth factor-I; MR, Mendelian randomization; SNP, single-nucleotide polymorphism.

* Corresponding author at: Unit of Medical Epidemiology, Department of Surgical Sciences, Uppsala University, 75185 Uppsala, Sweden.
E-mail address: susanna.larsson@surgsci.uu.se (S.C. Larsson).

[1] Equally shared first authorship.

https://doi.org/10.1016/j.metabol.2022.155329
Received 19 August 2022; Accepted 30 September 2022
Available online 5 October 2022
0026-0495/© 2022 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

triglycerides and low-density lipoprotein cholesterol, and increased high-density lipoprotein cholesterol) [1,4–7]. Several studies in rodents have also demonstrated that FGF21 alleviates inflammation, reduces circulating markers of liver fibrosis and injury, and protects against non-alcoholic liver disease [5]. Based on preclinical observations, FGF21 analogues designed to expand the half-lives of FGF21 have been assessed as potential therapeutic agents for metabolic disorders in individuals with obesity, type 2 diabetes, or non-alcoholic steatohepatitis. In humans these conditions appear to be FGF21 resistant states and most but not all [8] clinical trials in humans have demonstrated no effect of FGF21 analogues on the primary end points of glycemic control but significant improvements in lipid profile and markers of hepatocellular damage [5,7,9–13]. Moreover, FGF21 analogue treatment has been shown to increase circulating concentrations of insulin-like growth factor-I (IGF-I) in humans [7] and to reduce alcohol and sweet food intake in both mice and humans [14–17]. Given differences in FGF21 biology between mice and humans, the normal physiological role of FGF21 in humans needs a better understanding by use of long-lasting experiments, longitudinal data analyses, and instrumental variable approaches such as Mendelian randomization (MR) but suitable genetic instruments are presently not existent.

In addition to nutritional and hormonal regulation of FGF21 production [4], circulating FGF21 concentration may be genetically modified. However, a genome-wide association study (GWAS) in 4201 participants of the Taiwan Biobank identified no significant associations with circulating FGF21 [18]. Other GWAS analyses with a focus on macronutrient intake found that genetic variants in or near the *FGF21* locus (rs838133 and rs838145) were associated with increased relative carbohydrate intake but decreased relative protein and fat intake [8,19,20].

Here, we first performed a GWAS on circulating FGF21 concentrations in individuals of European ancestry. We then applied the two-sample MR design to better understand the potential therapeutic effects of FGF21 on relevant phenotypes, including alcohol and nutrient intake, cardiovascular and metabolic biomarkers and diseases, and liver function biomarkers. With respect to a drug target such as FGF21, the MR design can utilize one or multiple genetic variants in a relevant gene and associated with the targeted drug as instrument(s) to assess the effect of perturbing that drug target [21]. Through the random separation of genetic alleles at conception, MR can mimic the design of randomized clinical trials and reduce reverse causality and confounding. MR studies have commonly demonstrated to produce similar results as randomized clinical trials but in a more cost-effective and relatively inexpensive manner. Considering that genotype is fixed at birth, MR studies reflect the long-term (life-long) effect of higher, but still physiological in their range, concentrations of the exposure on health outcomes whereas clinical trials in humans investigate relative short-term effects of either physiological or high pharmacological concentrations of the exposure of interest.

2. Materials and methods

2.1. Study population

The study population for the discovery and replication cohorts included participants of clinical subcohorts of the Swedish Mammography Cohort and Cohort of Swedish Men which are embedded in the Swedish Infrastructure for Medical Population-Based Life-Course and Environmental Research (SIMPLER) (https://www.simpler4health.se/). These population-based longitudinal cohorts were established in 1987–1990 (Swedish Mammography Cohort) and 1997 (Cohort of Swedish Men) when all women who were born between 1914 and 1948 and lived in Västmanland and Uppsala counties and all men who were born between 1918 and 1952 and lived in Västmanland and Örebro counties were invited to participate in these cohorts. Participants for the two clinical subcohorts were randomly selected from cohort participants living in Västmanland and Uppsala counties. The discovery cohort sample for the present study comprised 7384 women and men who participated in a health examination in Västerås (city in Västmanland County) between 2010 and 2019. The replication cohort comprised 5022 women in Uppsala who participated in a health examination and provided a blood sample between 2003 and 2009.

At the health examination, a fasting blood sample was drawn and other data such as weight and height were collected. Participants also completed questionnaires about their diet and other lifestyle factors and health status. After GWAS and protein data quality control (see below), 6259 and 4483 participants remained for analysis in the discovery and replication cohort, respectively.

The study was approved by the Ethical Review Authority, Sweden. All participants gave written informed consent.

2.2. FGF21 measurement

Blood samples were collected in the morning following an overnight fast. For lithium-heparin plasma preparation, blood samples were light-protected and after a delay of 15–20 min at room temperature, samples were spun in a centrifuge at 1615 *g* for 11 min at 4 °C. The plasma was then frozen in multiple tubes and stored at −80 °C until analysis. The relative concentration of FGF21 in plasma was measured by a high-throughput multiplex immunoassay (Olink Proseek Multiplex CVD II; Olink Bioscience, Uppsala, Sweden), which provides normalized protein expression values on a $\log^2$ scale standardized per analysis plate. We used bridging samples from the two cohorts to standardize concentrations. For data analysis Olink NPX Manager software was applied, and one-unit higher NPX represents an approximate doubling of measured FGF21 concentration. The within and between run precision coefficient of variation was 4 % and 5 %, respectively.

2.3. Genotyping and GWAS quality control

Genotyping was performed for the same individuals for whom the protein analysis was done. For the discovery cohort, staff at Eurofins Genomics (Ebersberg, Germany) extracted DNA from 4 ml EDTA whole blood with use of QIAamp DNA Blood Midi Kit" (Cat. No. 51185) from Qiagen (Hilden, Germany). Subsequently, samples were genotyped at Eurofins Genomics with the Illumina Infinium Global Screening Array version 3 (GSAv3; Illumina, San Diego, CA, USA) including 654,027 markers. Sample exclusion filters applied were: (1) samples with discordant sex information when comparing reported sex and sex determined by the X-chromosome; (2) non-European ancestry; (3) heterozygosity outliers -/+3*IQR from Q1/Q3; (4) low sample call rate (<98 %); (5) HWE exact *p*-value $<1 \times 10^{-7}$ (–hwe midp); (6) minor allele count <20; (7) markers not present in 1000G/HRC with matching alleles; (8) allele frequency difference > 0.15 compared with 1000G/HRC. In total, 6903 samples passed quality control, and 533,425 marker sites. We imputed data by use of chr1-22,X: Michigan Imputation Server v1.2.4 using Eagle v2.4 + minimac v4 and both 1000G phase3 (v5) in tgp.ph 3/ and HRC v1.1 in in hrc1.1/ as reference panels. The final genetic dataset included approximately 7.8 million markers.

For the replication cohort, staff at the Biobank at Karolinska Institutet extracted DNA from 400 μl EDTA whole blood with the Chemagen STAR DNA Blood 400 kit (Perkin Elmer, Waltham, MA, USA) using a ChemagicStar-robot (Hamilton, Reno, NV, USA) based on magnetic bead separation. Subsequently, samples were genotyped at the SNP&SEQ Technology Platform, Science for Life Laboratory, Uppsala University with the Illumina Infinium Global Screening Array Multiple Disease version 1 (GSAv1; Illumina, San Diego, CA, USA) including 642,824 markers. In the replication cohort, we used the same exclusion filters for quality control, leaving 535,329 sites and 4724 samples. For imputation of chr1-22 the Michigan Imputation Server v1.0.4 using Eagle v2.3 + minimac v3 was used and for chrX the Michigan Imputation Server v1.2.4 using Eagle v2.4 + minimac v4, with reference panels 1000G

phase3 (v5) in tgp.ph 3/ and HRC v1.1 in in hrc1.1/. The final genetic dataset included approximately 7.8 million markers.

2.4. GWAS analysis

Analyses of the associations between genetic variants and circulating FGF21 concentrations were performed using linear regression and an additive genetic model, as implemented in SNPTEST [22]. Adjustments were made for age, sex (in the discovery cohort), and the first five genetic principal components, to correct for possible population stratification. Genetic variants associated with circulating FGF21 at the genome-wide significance threshold (i.e., $P < 5 \times 10^{-8}$) in the discovery cohort were taken forward for replication analysis. Associations were considered replicated at an a priori P value <0.05 with the same direction in the replication cohort. All P values were from two-sided statistical tests.

2.5. MR analysis

A two-sample MR design was used to assess the potential causal effects of increased circulating FGF21 concentrations on relevant phenotypes that have been previously found to associate with circulating FGF21 in animals or humans [1,4–7,9–17]. The phenotypes studied in the current MR study included intake of alcohol and nutrients (including total sugar, carbohydrate, protein, and fat), cardiovascular biomarkers and diseases (including triglycerides, low-density lipoprotein cholesterol, apolipoprotein B, high-density lipoprotein cholesterol, apolipoprotein A-I, C-reactive protein, coronary artery disease, atrial fibrillation, heart failure, and ischemic stroke), metabolic biomarkers and diseases (including whole body fat mass, fasting glucose, fasting insulin, IGF-I, and type 2 diabetes), and liver function biomarkers (including alanine aminotransferase, alkaline phosphatase [ALP], aspartate aminotransferase, gamma-glutamyl transferase, galectin-3, liver fat content, and non-alcoholic fatty liver disease). For all but three phenotypes, MR analysis to obtain the Wald ratio estimate of the association between genetically predicted circulating FGF21 and the phenotype (i.e., the ratio of the beta coefficient for the outcome divided by the beta coefficient for circulating FGF21) was conducted using the MR-Base platform [23] and the largest pertinent GWAS dataset available that contained the FGF21-associated single-nucleotide polymorphism (SNP) [23–31]. For type 2 diabetes, liver fat, content and non-alcoholic fatty liver disease, summary statistics data were obtained from the DIAMANTE consortium [32] and recent GWAS meta-analyses [33,34]. Associations with a P value <0.0018 ($P = 0.05/28$ tested phenotypes) were regarded strong associations, whereas associations with a P value between 0.0018 and 0.05 were regarded suggestive associations.

2.6. MR assumptions

The three core instrumental variable assumptions for an MR analysis are: 1) the genetic variant must be strongly associated with the exposure (the relevance assumption); 2) the genetic variant should not be associated with potential confounders (the independence assumption); and 3) the genetic variant must affect the outcome through the exposure and not through other pathways (the exclusion restriction assumption). A single genetic variant could satisfy these assumptions if it has a genome-wide significant association with the exposure and is located in or near the gene encoding the exposure (e.g., FGF21), and with no pleiotropic associations with other phenotypes. In this MR study, we considered these conditions when selecting the instrumental variable for circulating FGF21 concentrations.

3. Results

3.1. GWAS

Characteristics of the discovery and replication cohorts are presented in Table 1. The GWAS analysis identified multiple SNPs with genome-wide significant associations with plasma FGF21 concentrations on chromosome 2 in the *GCKR* (glucokinase regulatory protein) gene and on chromosome 19 in or near the *FGF21* gene (Fig. 1). All SNPs were replicated (Supplemental Table 1). The strongest signals at the *GCKR* and *FGF21* loci were rs1260326 ($\beta = 0.174$, $P = 3.98 \times 10^{-34}$) and rs2548957 ($\beta = 0.181$, $P = 2.18 \times 10^{-42}$), respectively (Table 2). The latter SNP explained 1.3 % of the variance in FGF21 concentration in the discovery cohort and 1.9 % in the replication cohort.

3.2. Two-sample MR analysis

For the two-sample MR analyses, rs2548957 near the *FGF21* locus was used as instrumental variable for circulating FGF21 concentrations. There were no other independent SNPs (linkage disequilibrium $R^2 > 0.01$) associated with circulating FGF21 in this locus. SNPs from the *GCKR* gene were not considered as instrumental variable because of the well-established strong pleiotropic associations of the *GCKR* locus with metabolic phenotypes. Details of the GWAS datasets used for the studied outcomes are shown in Table 3.

Our MR analyses showed that higher circulating FGF21 instrumented by rs2548957 was strongly associated with lower alcohol consumption, lower circulating low-density lipoprotein cholesterol, apolipoprotein B, C-reactive protein, gamma-glutamyl transferase, and galectin-3 concentrations, and higher circulating IGF-I and ALP concentrations ($P < 0.0018$; Table 3). There were suggestive associations between genetically predicted higher circulating FGF21 and whole-body fat mass, circulating alanine aminotransferase, and type 2 diabetes risk, but no association with the other studied outcomes (Table 3).

4. Discussion

FGF21 is a hormone that regulates several aspects of energy homeostasis and metabolism in animals and humans. The magnitude and pleotropic character of its beneficial action on many, if not all, metabolic abnormalities in animals have led to extensive exploration of their biology in animals and later in humans and have revealed that the biology of FGF21 is different from that in animals from several points of view. Even though observational studies in humans project a picture consistent with FGF21 resistance in diabetes, obesity, and

Table 1
Characteristics of participants in the discovery and replication cohorts.

Characteristic[a]	Discovery cohort (N = 6259)	Replication cohort (N = 4483)
Age, years	73.8 (5.3)	67.1 (6.8)
Women	2170 (34.7 %)	4483 (100.0 %)
Body mass index, kg/m^2	26.6 (4.0)	25.9 (4.3)
Alcohol intake, g/day	7.6 (3.4–13.8)	4.3 (1.5–8.0)
Plasma FGF21, NPX	7.19 (1.24)	6.80 (1.23)
Plasma triglycerides, mmol/L	1.25 (0.98–1.65)	1.12 (0.86–1.52)
Plasma LDL cholesterol, mmol/L	3.13 (1.03)	3.53 (0.91)
Plasma HDL cholesterol, mmol/L	1.51 (0.38)	1.58 (0.37)
Plasma C-reactive protein, mg/L	1.60 (0.85–3.10)	1.50 (0.77–2.95)
Plasma alanine aminotransferase, μkat/L	0.20 (0.15–0.27)	0.20 (0.15–0.27)

[a] Data are mean ± standard deviation, median (interquartile range), or number (%), as appropriate. HDL, high-density lipoprotein; LDL, low-density lipoprotein.

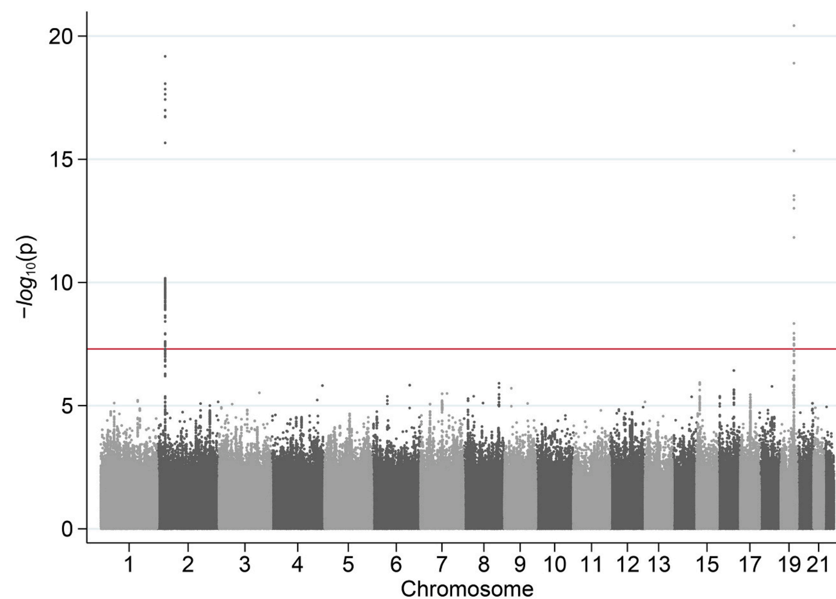


Fig. 1. Manhattan plot of GWAS analysis of circulating FGF21 in the discovery cohort of 6259 women and men. The X-axis represents the genomic position. The Y-axis represents the -$\log_{10}$ (P) of SNPs. The solid line is the genome wide significance threshold ($P < 5 \times 10^{-8}$).

Table 2
Lead SNPs associated with circulating FGF21 concentrations in GWAS analysis of the discovery and replication cohorts and in meta-analysis of both cohorts.

SNP	Chr	Position	Nearest gene	EA/OA	Discovery cohort (N = 6259)		Replication cohort (N = 4483)		Meta-analysis (N = 10,742)	
					β (SE)[a]	*P* value	β (SE)[a]	*P* value	β (SE)[a]	*P* value
rs1260326	2	27,730,940	*GCKR*	C/T	0.170 (0.019)	6.93×10^{-20}	0.180 (0.022)	3.97×10^{-16}	0.174 (0.014)	3.98×10^{-34}
rs2548957	19	49,262,500	*FGF21*	A/G	0.167 (0.018)	3.77×10^{-21}	0.201 (0.021)	9.20×10^{-22}	0.181 (0.013)	2.18×10^{-42}

Chr, chromosome; EA, effect allele; OA, other allele; SE, standard error; SNP, single-nucleotide polymorphism.
[a] β coefficients represent the change in FGF21 concentrations per additional effect allele. For data analysis Olink NPX Manager software was applied, and one-unit higher NPX represents an approximate doubling of the measured FGF21 concentration.

cardiovascular disease [35], novel FGF21-analogues have been designed and are now tested for therapeutic purposes although their success cannot be fully predicted based on existing experimental data.

Several areas of FGF21 biology, including its association with alcohol or carbohydrate (sweets) intake in humans remain to be fully elucidated and initial attempts to develop novel medications to address obesity and hyperglycemia in type 2 diabetes patients resulted have not revealed very promising results. In contrast, robust, consistent, and durable effects on lipid metabolism gradually shifted emphasis toward their use for hyperlipidemias and the efficacy of FGF21 analogues for cardiovascular disease remains to be shown. Whether FGF21 analogues may have effects on alcohol or carbohydrate intake and whether they can provide a favorable efficacy vs safety balance for non-alcoholic steatohepatitis, an unmet clinical need, or other conditions remains currently an active area of investigation. In that respect, methodologies that confer causal inferences and are not expensive or time consuming, such as MR studies, could provide invaluable information which could elucidate future randomized trial design and development of FGF21 analogues as potential pharmacotherapies in the future.

In the present study, the two-sample MR design was applied to study the causal effects of targeting FGF21, instrumented by a reliable genetic variant in the *FGF21* gene region, on a wide range of health outcomes. Firstly, we conducted a GWAS analysis in nearly 11,000 adults of European ancestry and identified genome-wide significant associations of genetic variants in and near the *GCKR* and *FGF21* gene regions with FGF21 concentrations. Secondly, we used the two-sample MR design and found evidence in support that higher circulating FGF21 concentrations instrumented by the strongest signal near the *FGF21* gene improves lipid profile and reduces alcohol consumption, C-reactive protein, gamma-glutamyl transferase, and galectin-3 concentrations, and increases IGF-I and ALP concentrations.

The observed associations of genetically predicted higher FGF21 concentrations with lower alcohol consumption and lower concentrations of gamma-glutamyl transferase, a widely used biomarker of heavy alcohol drinking, are consistent with a physiological role of FGF21 to regulate alcohol intake and provide support to results of previous studies, which have demonstrated that exogenous FGF21 treatment decreases alcohol consumption [14,16,17]. FGF21 treatment has also been reported to reduce sugar and sweet food consumption [14–16]. The lack of association between genetically predicted FGF21 concentrations and total sugar intake may be related to misclassification of sugar intake in the GWAS analysis in the UK Biobank. We also observed no significant association between genetically predicted FGF21 concentrations and macronutrient intake. These negative findings may also be due to misclassification of nutrient intake or to the true lack of associations of FGF21 with macronutrient preference in humans and need to be explored further.

Despite reduced low-density lipoprotein cholesterol and apolipoprotein B concentrations with higher FGF21 concentrations, we found no evidence of a causal effect of FGF21 on coronary artery disease and ischemic stroke risk. These findings validate hyperlipidemia as a major pharmacotherapy target of FGF21 analogues and suggest that other risk factors, such as greater fat mass and prevalence of type 2 diabetes which are associated with increasing FGF21 concentrations, indicating a degree of FGF21 resistance [35], may counteract the beneficial effect of improved lipid profile. Both fat mass and type 2 diabetes are causal risk factors for coronary artery disease and ischemic stroke [36–38], and both tend to increase with increasing FGF21 levels herein indicating a

Table 3
Associations of circulating FGF21 concentrations instrumented by rs2548957 with relevant phenotypes in two-sample MR analysis.

Phenotype	Details of the phenotype GWAS used in two-sample MR analysis				Two-sample MR analysis result[a]	
	Phenotype ID in the MR-Base platform	Study or consortium	First author's last name, year [reference]	Sample size or cases/controls	β-coefficient or odds ratio (95 % CI)	*P* value
Alcohol and nutrients						
Alcoholic drinks per week	ieu-b-73	GSCAN	Liu, 2020 [24]	335,394	−0.062 (−0.083, −0.041)	6.77×10^{-9} [b]
Total sugar intake	ukb-b-17,079	UKBB	Elsworth, 2018 [23]	64,979	0.012 (−0.048, 0.071)	0.700
Carbohydrate intake	ukb-b-7244	UKBB	Elsworth, 2018 [23]	64,979	0.034 (−0.025, 0.093)	0.261
Protein intake	ukb-b-12,043	UKBB	Elsworth, 2018 [23]	64,979	0.058 (0.000, 0.117)	0.051
Fat intake	ukb-b-12,379	UKBB	Elsworth, 2018 [23]	64,979	0.051 (−0.008, 0.110)	0.088
Cardiovascular						
Triglycerides	ieu-b-111	UKBB	Richardson, 2020 [25]	441,016	−0.019 (−0.040, 0.003)	0.094
LDL cholesterol	ieu-b-110	UKBB	Richardson, 2020 [25]	440,546	−0.078 (−0.100, −0.055)	1.92×10^{-11} [b]
Apolipoprotein B	ieu-b-108	UKBB	Richardson, 2020 [25]	439,214	−0.078 (−0.100, −0.055)	1.25×10^{-11} [b]
HDL cholesterol	ieu-b-109	UKBB	Richardson, 2020 [25]	403,943	−0.021 (−0.042, 0.000)	0.051
Apolipoprotein A-I	ieu-b-107	UKBB	Richardson, 2020 [25]	393,193	−0.019 (−0.040, 0.003)	0.088
C-reactive protein	ukb-d-30710_irnt	UKBB	Neale, 2018 [23]	343,524	−0.045 (−0.071, −0.019)	7.77×10^{-4} [b]
Coronary artery disease	ebi-a-GCST005195	CARDIoGRAMplusC4D	van der Harst, 2017 [26]	122,733/424528	0.988 (0.916, 1.066)	0.758
Atrial fibrillation	ebi-a-GCST006414	Six consortia/studies[c]	Nielsen, 2018 [27]	60,620/970216	0.969 (0.893, 1.051)	0.447
Heart failure	ebi-a-GCST009541	HERMES	Shah, 2020 [28]	47,309/930014	0.988 (0.906, 1.101)	0.973
Ischemic stroke	ebi-a-GCST006908	MEGASTROKE	Malik, 2018 [29]	34,217/406111	1.065 (0.931, 1.218)	0.358
Metabolic						
Whole body fat mass	ukb-b-19,393	UKBB	Elsworth, 2018 [23]	454,137	0.022 (0.001, 0.004)	0.040
Fasting glucose	ebi-a-GCST90002232	MAGIC	Chen, 2021 [30]	200,622	−0.022 (−0.053, 0.010)	0.179
Fasting insulin	ebi-a-GCST90002238	MAGIC	Chen, 2021 [30]	151,013	−0.003 (−0.038, 0.031)	0.851
Leptin	ebi-a-GCST90012076	SCALLOP	Folkersen, 2020 [31]	21,758	0.065 (−0.053, 0.183)	0.283
Insulin-like growth factor-I	ukb-d-30770_irnt	UKBB	Neale, 2018 [23]	NA	0.065 (0.039, 0.090)	5.85×10^{-7} [b]
Type 2 diabetes	NA	DIAMANTE	Mahajan [32]	74,124/824006	1.123 (1.043, 1.209)	0.002
Liver disease						
Alanine aminotransferase	ukb-d-30620_irnt	UKBB	Neale, 2018 [23]	NA	0.027 (0.003, 0.051)	0.030
Alkaline phosphatase	ukb-d-30610_irnt	UKBB	Neale, 2018 [23]	NA	0.236 (0.211, 0.262)	3.25×10^{-74} [b]
Aspartate aminotransferase	ukb-d-30650_irnt	UKBB	Neale, 2018 [23]	NA	−0.022 (−0.047, 0.003)	0.078
Gamma-glutamyl transferase	ukb-d-30730_irnt	UKBB	Neale, 2018 [23]	NA	−0.076 (−0.100, −0.052)	7.71×10^{-10} [b]
Galectin-3	ebi-a-GCST90012009	SCALLOP	Folkersen, 2020 [31]	21,758	−0.247 (−0.357, −0.137)	1.17×10^{-5} [b]
Liver fat	NA	UKBB	Haas, 2021 [33]	32,192	0.022 (−0.063, 0.107)	0.640
NAFLD	NA	Three studies	Anstee, 2020 [34]	1484/17781	1.337 (0.954, 1.874)	0.092

CARDIoGRAMplusC4D, Coronary ARtery DIsease Genome wide Replication and Meta-analysis plus Coronary Artery Disease Genetics; CI, confidence interval; DIAMANTE, DIAbetes Meta-ANalysis of Trans-Ethnic association studies; GIANT, Genetic Investigation of ANthropometric Traits; GSCAN, GWAS and Sequencing Consortium of Alcohol and Nicotine use; HDL, high-density lipoprotein; HERMES, Heart Failure Molecular Epidemiology for Therapeutic Targets; irnt, inverse rank-normal transformed; LDL, low-density lipoprotein; MAGIC, Meta-Analyses of Glucose and Insulin-related traits Consortium; NA, not available; NAFLD, non-alcoholic fatty liver disease; OR, odds ratio; UKBB, UK Biobank.

[a] Beta coefficients (standard deviation unit) and odds ratios are scaled per approximate doubling of genetically predicted circulating FGF21 concentrations.

[b] Significant at the Bonferroni corrected significance threshold ($P < 0.0018$), accounting for multiple comparisons.

[c] Atrial Fibrillation Consortium, UKBB, Nord-Trøndelag Health Study, Diabetes Epidemiology: Collaborative analysis of Diagnostic criteria in Europe, Michigan Genomics Initiative, and DiscovEHR.

level of FGF21 resistance. It remains to be shown whether very potent FGF21 analogues could potentially overcome the demonstrated FGF21 resistance in obesity and type 2 diabetes [35], but our data support no role for long-term higher FGF21 concentrations within the physiological range to regulate energy homeostasis, obesity, and type 2 diabetes in contrast to hyperlipidemia.

Our MR analysis confirmed the lack of association of circulating FGF21 with glycemic traits reported by previous clinical trials evaluating the effects of FGF21 analogues on glycemic control [5,7,9–13]. The observed positive association between genetically predicted circulating FGF21 and IGF-I concentrations is in line with the result of a clinical trial, which showed that treatment with an FGF21 analogue increased IGF-I concentrations [7]. We have previously found that genetically predicted higher IGF-I concentrations is associated with an increased risk of type 2 diabetes and several lines of evidence indicate that increasing IGF-1 can compensate for insulin resistance by acting through its own or the insulin receptor [39]. Hence, the observed suggestive association between genetically predicted FGF21 concentrations and type 2 diabetes might be mediated by FGF21 resistance and may reflect the existence of an underlying insulin resistance, which may be compensated, at least in part, through increased IGF-I which may act at the level of the insulin receptor in a compensatory manner [40,41].

In our study, we also found an inverse association between genetically predicted FGF21 and C-reactive protein concentrations, a marker of systemic inflammation, supporting a protective role of FGF21 in inflammation. We demonstrate, however, that improvements in inflammation and hyperlipidemia are not sufficient to result in improved cardiovascular disease outcomes in the presence of resistance to it effects in obesity and type 2 diabetes.

The prominent influence of FGF21 on circulating ALP is a novel finding that needs further exploration. The vast majority of circulating ALP activity is normally derived from liver and bone, and it is well known that serum ALP elevations occur in cholestatic liver disease and bone disorders [42]. The normal physiology of the enzyme is less known, and this study cannot differentiate between liver vs. bone as the source of ALP but the concomitant association with increased gamma-glutamyl transferase may be pointing to a liver etiology. The physiological activity of ALP involves the provision of phosphate to tissue (mainly bone) or the activation or inactivation of substrates via dephosphorylation. There seems to be no evidence of hyperphosphatemia or other deleterious effects of elevated serum ALP concentration but ALP knockout mice display accelerated fat absorption, suggesting that ALP inhibits fat absorption [43] – adding to our observation that FGF21 influences blood lipids. Multiple studies have suggested that intestinal ALP might also diminish gut inflammation, possibly via dephosphorylation of toxic bacterial lipopolysaccharides or enhancement of the barrier function of the intestinal epithelium [44,45]. The axis between FGF21, ALP activity, blood lipids, and inflammation needs further exploration.

Among other liver disease biomarkers, very weak or no association was observed between genetically predicted FGF21 and alanine aminotransferase, liver fat, and non-alcoholic fatty liver disease indicating that FGF21 in physiological concentrations may have a minor or no effect on early stages of the non-alcoholic fatty liver disease process [46,47]. Nevertheless, we observed a strong inverse association of genetically predicted FGF21 with galectin-3 concentrations, which have previously been reported to be higher in patients with liver fibrosis including liver cancer, liver failure, and liver cirrhosis compared with healthy volunteers [48]. Thus, FGF21 in physiological concentrations may have more significant effects on late stages of nonalcoholic steatohepatitis including fibrosis as well as cirrhosis and progression to liver failure and cancer.

In addition to SNPs at the *FGF21* locus, SNPs in the *GCKR* gene was strongly associated with circulating FGF21 in the present GWAS. The *GCKR* locus is widely reported to have pleiotropic effects on metabolic phenotypes, such as triglycerides [49–51], but primarily fasting glucose [51], and type 2 diabetes [32], and is also strongly associated with alcohol [24,52] and macronutrient [53] intake. It is plausible that genetic variations in *GCKR* do not directly affect FGF21 concentrations but is otherwise associated with other traits, such as alcohol consumption and nutritional factors, and most important genetic forms of diabetes and hyperglycemia. Previous evidence indicates that sugar ingestion and high-carbohydrate diets provoke changes in circulating FGF21 concentrations and hepatic expression in mice and humans [4]. Given that the *GCKR* gene, which encodes the glucokinase regulator protein, is a highly pleiotropic locus that associates with multiple phenotypes and unspecific for FGF21, we did not include genetic variants from this locus as instrument for circulating FGF21 in the MR analysis.

We are not aware of any previous GWAS for circulating FGF21 in populations of European ancestry, as done herein, but there is a GWAS for this hormone conducted in Taiwanese adults [18]. That GWAS found no genome-wide significant associations with circulating FGF21 which may be related to insufficient power due to the small sample size (n = 4201 participants).

The strongest SNP near the *FGF21* locus that was identified in our GWAS is in modest linkage disequilibrium ($R^2 \approx 0.3$) with SNPs (rs838133 and rs838145) previously found to associate with relative intake of macronutrients in GWAS analyses [19,20]. The SNP rs838133 has also been reported to be strongly associated with alcohol consumption, fat mass, lipids, liver enzymes (gamma-glutamyl transferase, ALP, and AST), and CRP concentrations [8,54], suggestively associated with major cardiovascular diseases (except for strong association with venous thromboembolism), Alzheimer's disease, and lifespan [54], but not associated with glycemic traits or type 2 diabetes [8,54]. A weak association was found between rs838145 and FGF21 protein concentrations in 377 individuals ($P = 0.01$) [20]. We observed moderate associations of both rs838145 and rs838133 with circulating FGF21 in the present discovery ($P = 1.48 \times 10^{-6}$ for rs838145, $P = 0.046$ for rs838133) and replication ($P = 1.25 \times 10^{-6}$ for rs838145, $P = 0.002$ for rs838133) cohorts.

A strength of our investigation is the relatively large sample size of the GWAS and the consistency of the results in two independent study samples. Another strength is that we could examine the potential causal associations of long-term higher FGF21 concentrations within the physiological range on a wide range of phenotypes in large study samples and in a relatively inexpensive way compared with a clinical trial and can thus inform clinical trials and rationale drug development programs in a very cost-effective manner. An additional advantage is that there was no sample overlap between the GWAS for FGF21 and the GWAS outcome datasets and therefore our MR findings could not be biased toward the observational association.

A limitation is that effect of the SNPs on protein function is not known. Another possible shortcoming is that FGF21 concentrations were measured on a relative scale and the values cannot be converted to absolute concentrations in plasma, but this would not have influenced the conclusions and inferences of this study. A further limitation is that we were unable to assess the associations of genetically predicted FGF21 concentrations with insulin resistance as neither rs2548957 nor a suitable proxy SNP was available in publicly available GWAS datasets of insulin resistance but provide the data herein that can form the basis for more detailed studies in the future. A possible limitation is that the study populations comprised individuals of European ancestry and the general populations. Hence, our findings may not be generalizable to non-European populations or to specific patients, such as diabetes patients. Finally, we cannot assess any associations or effects of clearly supra-physiological FGF21 concentrations that are expected to be reached by the several compounds currently in clinical trials in humans which could overcome FGF21 resistance in states such as obesity and diabetes [35]. We can only draw robust inferences for FGF21 concentrations within the normal physiological range in relation to hyperlipidemias, liver diseases, type 2 diabetes, and cardiovascular disease in addition to physiological associations with alcohol consumption and carbohydrate intake which could also be drug development targets in the future.

5. Conclusions

This GWAS identified robust associations of genetic variants in or near the *GCKR* and *FGF21* genes with circulating FGF21 concentrations. Furthermore, our findings support a causal role of FGF21 within the physiological range on alcohol consumption, blood lipids, systemic inflammation, IGF-I, and liver function in humans.

CRediT authorship contribution statement

SCL designed the study, contributed to data acquisition, methodology, data analysis, data interpretation, and wrote the first draft of the paper. KM designed the study, contributed to data acquisition, methodology, data interpretation, and revision of the paper. MM-C contributed to methodology, data analysis, data interpretation, and revision of the paper. JH contributed to methodology, data analysis, data interpretation, and revision of the paper. CSM designed the study and contributed to data interpretation and revision of the paper.

Declaration of competing interest

The authors declare that they have no competing interests.

Data availability

The data analyzed in this manuscript can be applied for at the homepage for SIMPLER https://www.simpler4health.se/.

Acknowledgement

We acknowledge SIMPLER for provisioning of facilities and experimental support. SIMPLER receives funding through the Swedish Research Council (grant numbers 2021-00160, 2017-00644 and 2017-06100) and Stiftelsen Olle Engkvist Byggmästare. The GWAS analyses were performed on resources provided by SNIC-SENS through the Uppsala Multidisciplinary Center for Advanced Computational Science (UPPMAX) under Project SIMP2021005.

Funding

This work was supported by research grants from the Swedish Cancer Society (Cancerfonden), the Swedish Heart Lung Foundation (Hjärt-Lungfonden, grant number 20210351), the Swedish Research Council for Health, Working Life and Welfare (Forte, grant number 2018-00123), and the Swedish Research Council (Vetenskapsrådet, grant number 2019-00977).

References

[1] Kharitonenkov A, Shiyanova TL, Koester A, Ford AM, Micanovic R, Galbreath EJ, et al. FGF-21 as a novel metabolic regulator. J Clin Invest 2005;115(6):1627–35.
[2] Fisher FM, Maratos-Flier E. Understanding the physiology of FGF21. Annu Rev Physiol 2016;78:223–41.
[3] Flippo KH, Potthoff MJ. Metabolic messengers: FGF21. Nat Metab 2021;3(3): 309–17.
[4] Staiger H, Keuper M, Berti L, Hrabe de Angelis M, Haring HU. Fibroblast growth factor 21-metabolic role in mice and men. Endocr Rev 2017;38(5):468–88.
[5] Geng L, Lam KSL, Xu A. The therapeutic potential of FGF21 in metabolic diseases: from bench to clinic. Nat Rev Endocrinol 2020;16(11):654–67.
[6] Stefan N, Haring HU. The role of hepatokines in metabolism. Nat Rev Endocrinol 2013;9(3):144–52.
[7] Talukdar S, Zhou Y, Li D, Rossulek M, Dong J, Somayaji V, et al. A long-acting FGF21 molecule, PF-05231023, decreases body weight and improves lipid profile in non-human primates and type 2 diabetic subjects. Cell Metab 2016;23(3): 427–40.
[8] Frayling TM, Beaumont RN, Jones SE, Yaghootkar H, Tuke MA, Ruth KS, et al. A common allele in FGF21 associated with sugar intake is associated with body shape, lower total body-fat percentage, and higher blood pressure. Cell Rep 2018; 23(2):327–36.
[9] Gaich G, Chien JY, Fu H, Glass LC, Deeg MA, Holland WL, et al. The effects of LY2405319, an FGF21 analog, in obese human subjects with type 2 diabetes. Cell Metab 2013;18(3):333–40.
[10] Charles ED, Neuschwander-Tetri BA, Pablo Frias J, Kundu S, Luo Y, Tirucherai GS, et al. Pegbelfermin (BMS-986036), PEGylated FGF21, in patients with obesity and type 2 diabetes: results from a randomized phase 2 study. Obesity 2019;27(1):41–9 (Silver Spring).
[11] Sanyal A, Charles ED, Neuschwander-Tetri BA, Loomba R, Harrison SA, Abdelmalek MF, et al. Pegbelfermin (BMS-986036), a PEGylated fibroblast growth factor 21 analogue, in patients with non-alcoholic steatohepatitis: a randomised, double-blind, placebo-controlled, phase 2a trial. Lancet 2019;392(10165): 2705–17.
[12] Harrison SA, Ruane PJ, Freilich BL, Neff G, Patil R, Behling CA, et al. Efruxifermin in non-alcoholic steatohepatitis: a randomized, double-blind, placebo-controlled, phase 2a trial. Nat Med 2021;27(7):1262–71.
[13] Rader DJ, Maratos-Flier E, Nguyen A, Hom D, Ferriere M, Li Y, et al. LLF580, an FGF21 analog, reduces triglycerides and hepatic fat in obese adults with modest hypertriglyceridemia. J Clin Endocrinol Metab 2022;107(1):e57–70.
[14] Hill CM, Qualls-Creekmore E, Berthoud HR, Soto P, Yu S, McDougal DH, et al. FGF21 and the physiological regulation of macronutrient preference. Endocrinology 2020;161(3).
[15] von Holstein-Rathlou S, BonDurant LD, Peltekian L, Naber MC, Yin TC, Claflin KE, et al. FGF21 mediates endocrine control of simple sugar intake and sweet taste preference by the liver. Cell Metab 2016;23(2):335–43.
[16] Talukdar S, Owen BM, Song P, Hernandez G, Zhang Y, Zhou Y, et al. FGF21 regulates sweet and alcohol preference. Cell Metab 2016;23(2):344–9.
[17] Flippo KH, Trammell SAJ, Gillum MP, Aklan I, Perez MB, Yavuz Y, et al. FGF21 suppresses alcohol consumption through an amygdalo-striatal circuit. Cell Metab 2022;34(2):317–28 e6.
[18] Chuang GT, Liu PH, Chyan TW, Huang CH, Huang YY, Lin CH, et al. Genome-wide association study for circulating fibroblast growth factor 21 and 23. Sci Rep 2020; 10(1):14578.
[19] Chu AY, Workalemahu T, Paynter NP, Rose LM, Giulianini F, Tanaka T, et al. Novel locus including FGF21 is associated with dietary macronutrient intake. Hum Mol Genet 2013;22(9):1895–902.
[20] Tanaka T, Ngwa JS, van Rooij FJ, Zillikens MC, Wojczynski MK, Frazier-Wood AC, et al. Genome-wide meta-analysis of observational studies shows common genetic variants associated with macronutrient intake. Am J Clin Nutr 2013;97(6): 1395–402.
[21] Walker VM, Davey Smith G, Davies NM, Martin RM. Mendelian randomization: a novel approach for the prediction of adverse drug events and drug repurposing opportunities. Int J Epidemiol 2017;46(6):2078–89.
[22] Marchini J, Howie B. Genotype imputation for genome-wide association studies. Nat Rev Genet 2010;11(7):499–511.
[23] Hemani G, Zheng J, Elsworth B, Wade KH, Haberland V, Baird D, et al. The MR-base platform supports systematic causal inference across the human phenome. Elife 2018;7. pii: e34408.
[24] Liu M, Jiang Y, Wedow R, Li Y, Brazel DM, Chen F, et al. Association studies of up to 1.2 million individuals yield new insights into the genetic etiology of tobacco and alcohol use. Nat Genet 2019;51(2):237–44.
[25] Richardson TG, Sanderson E, Palmer TM, Ala-Korpela M, Ference BA, Davey Smith G, et al. Evaluating the relationship between circulating lipoprotein lipids and apolipoproteins with risk of coronary heart disease: a multivariable mendelian randomisation analysis. PLoS Med 2020;17(3):e1003062.
[26] van der Harst P, Verweij N. Identification of 64 novel genetic loci provides an expanded view on the genetic architecture of coronary artery disease. Circ Res 2018;122(3):433–43.
[27] Nielsen JB, Thorolfsdottir RB, Fritsche LG, Zhou W, Skov MW, Graham SE, et al. Biobank-driven genomic discovery yields new insight into atrial fibrillation biology. Nat Genet 2018;50(9):1234–9.
[28] Shah S, Henry A, Roselli C, Lin H, Sveinbjornsson G, Fatemifar G, et al. Genome-wide association and Mendelian randomisation analysis provide insights into the pathogenesis of heart failure. Nat Commun 2020;11(1):163.
[29] Malik R, Chauhan G, Traylor M, Sargurupremraj M, Okada Y, Mishra A, et al. Multiancestry genome-wide association study of 520,000 subjects identifies 32 loci associated with stroke and stroke subtypes. Nat Genet 2018;50(4):524–37.
[30] Chen J, Spracklen CN, Marenne G, Varshney A, Corbin LJ, Luan J, et al. The trans-ancestral genomic architecture of glycemic traits. Nat Genet 2021;53(6):840–60.
[31] Folkersen L, Gustafsson S, Wang Q, Hansen DH, Hedman AK, Schork A, et al. Genomic and drug target evaluation of 90 cardiovascular proteins in 30,931 individuals. Nat Metab 2020;2(10):1135–48.
[32] Mahajan A, Taliun D, Thurner M, Robertson NR, Torres JM, Rayner NW, et al. Fine-mapping type 2 diabetes loci to single-variant resolution using high-density imputation and islet-specific epigenome maps. Nat Genet 2018;50(11):1505–13.
[33] Haas ME, Pirruccello JP, Friedman SN, Wang M, Emdin CA, Ajmera VH, et al. Machine learning enables new insights into genetic contributions to liver fat accumulation. Cell Genom 2021;1(3).
[34] Anstee QM, Darlay R, Cockell S, Meroni M, Govaere O, Tiniakos D, et al. Genome-wide association study of non-alcoholic fatty liver and steatohepatitis in a histologically characterised cohort☆. J Hepatol 2020;73(3):505–15.
[35] Pilitsi E, Farr OM, Polyzos SA, Perakakis N, Nolen-Doerr E, Papathanasiou AE, et al. Pharmacotherapy of obesity: available medications and drugs under investigation. Metabolism 2019;92:170–92.
[36] Larsson SC, Burgess S. Fat mass and fat-free mass in relation to cardiometabolic diseases: a two-sample mendelian randomization study. J Intern Med 2020;288(2): 260–2.
[37] Larsson SC, Scott RA, Traylor M, Langenberg CC, Hindy G, Melander O, et al. Type 2 diabetes, glucose, insulin, BMI, and ischemic stroke subtypes: mendelian randomization study. Neurology 2017;89(5):454–60.
[38] Yuan S, Mason AM, Burgess S, Larsson SC. Differentiating associations of glycemic traits with atherosclerotic and thrombotic outcomes: mendelian randomization investigation. Diabetes 2022;71(10):2222–32.
[39] Larsson SC, Michaelsson K, Burgess S. IGF-1 and cardiometabolic diseases: a mendelian randomisation study. Diabetologia 2020;63(9):1775–82.
[40] Mantzoros CS, Flier JS. Insulin resistance: the clinical spectrum. Adv Endocrinol Metab 1995;6:193–232.
[41] Tritos NA, Mantzoros CS. Clinical review 97: syndromes of severe insulin resistance. J Clin Endocrinol Metab 1998;83(9):3025–30.
[42] Kaplan MM. Alkaline phosphatase. N Engl J Med 1972;286(4):200–2.
[43] Narisawa S, Huang L, Iwasaki A, Hasegawa H, Alpers DH, Millan JL. Accelerated fat absorption in intestinal alkaline phosphatase knockout mice. Mol Cell Biol 2003;23(21):7525–30.
[44] Rader BA. Alkaline phosphatase, an unconventional immune protein. Front Immunol 2017;8.
[45] Levitt MD, Hapak SM, Levitt DG. Alkaline phosphatase pathophysiology with emphasis on the seldom-discussed role of defective elimination in unexplained elevations of serum ALP - a case report and literature review. Clin Exp Gastroenterol 2022;15:41–9.
[46] Polyzos SA, Kountouras J, Mantzoros CS. Adipose tissue, obesity and non-alcoholic fatty liver disease. Minerva Endocrinol 2017;42(2):92–108.

[47] Boutari C, Perakakis N, Mantzoros CS. Association of adipokines with development and progression of nonalcoholic fatty liver disease. Endocrinol Metab 2018;33(1):33–43 (Seoul).
[48] An Y, Xu S, Liu Y, Xu X, Philips CA, Chen J, et al. Role of galectins in the liver diseases: a systematic review and meta-analysis. Front Med 2021;8:744518 (Lausanne).
[49] Yeh KH, Hsu LA, Teng MS, Wu S, Chou HH, Ko YL. Pleiotropic effects of common and rare GCKR exonic mutations on cardiometabolic traits. Genes 2022;13(3) (Basel).
[50] Willer CJ, Schmidt EM, Sengupta S, Peloso GM, Gustafsson S, Kanoni S, et al. Discovery and refinement of loci associated with lipid levels. Nat Genet 2013;45(11):1274–83.
[51] Orho-Melander M, Melander O, Guiducci C, Perez-Martinez P, Corella D, Roos C, et al. Common missense variant in the glucokinase regulatory protein gene is associated with increased plasma triglyceride and C-reactive protein but lower fasting glucose concentrations. Diabetes 2008;57(11):3112–21.
[52] Sanchez-Roige S, Palmer AA, Clarke TK. Recent efforts to dissect the genetic basis of alcohol use and abuse. Biol Psychiatry 2020;87(7):609–18.
[53] Meddens SFW, de Vlaming R, Bowers P, Burik CAP, Linner RK, Lee C, et al. Genomic analysis of diet composition finds novel loci and associations with health and lifestyle. Mol Psychiatry 2020;26(6):2056–69.
[54] Larsson SC, Gill D. Genetic evidence supporting fibroblast growth factor 21 signalling as a pharmacological target for cardiometabolic outcomes and Alzheimer's disease. Nutrients 2021;13(5):1504.