DOI: 10.53469/jcmp.2024.06(12).06

Causal Association between Saturated Fatty Acid Foods and Osteoarthritis and Identification of Potential Targets

Kaige Chen[1,2], Xingqiang Li[2], Xueqin Zeng[1], Yindi Sun[1,*]

[1]Pain ward of Orthopedics of Traditional Chinese Medicine, Honghui Hospital, Xi'an Jiaotong University, Xi'an 710054, Shaanxi, China
[2]Shaanxi University of Chinese Medicine, Xianyang 712046, Shaanxi, China
**Correspondence Author*

Abstract: ***Background:*** ***Numerous studies suggest a link between the consumption of foods rich in saturated fatty acids and the development of osteoarthritis. However, the causal relationships are still unclear. Moreover, the complex pathogenesis poses challenges in developing targeted drugs and identifying biomarkers.*** ***Objective:*** ***To investigate the potential causal association between the consumption of foods in saturated fatty acids and osteoarthritis, as well as to identify potential therapeutic targets and biomarkers using colocalization analysis.*** ***Methods:*** ***A two-sample Mendelian randomization (MR) analysis based on publicly available genome-wide association studies was employed to infer the causal relationship. The effect estimates were calculated using the random-effects inverse-variance-weighted method. Bayesian colocalization analysis was conducted to identify potential therapeutic targets and metabolic products.*** ***Results:*** ***Cheese intake per standard deviation increase causally reduced the risks of knee osteoarthritis (OR = 0.605; 95% CI, 0.479-0.764; p < 0.001), osteoarthritis of the hip or knee (OR = 0.676; 95% CI, 0.553-0.826; p < 0.001), and osteoarthritis self-reported (OR = 0.720; 95% CI, 0.521-0.995; p = 0.047). Cheesy biscuits intake per standard deviation increase causally reduced the risks of knee osteoarthritis (OR = 0.485; 95% CI, 0.239-0.987; p = 0.046), osteoarthritis of the hip or knee (OR = 0.535; 95% CI, 0.315-0.909; p = 0.021). Pork intake per standard deviation increase causally reduced the level of basophil (OR = 0.738; 95% CI,0.579-0.941; p = 0.014); each standard deviation increase in mutton intake reduced the level of CRP (OR = 0.763; 95% CI. 0.594-0.979; p = 0.034), each additional standard deviation of whole milk intake reduced the level of CRP (OR = 0.079; 95% CI, 0.008-0.737; p = 0.026), and each additional standard deviation of fried potato intake reduced the level of the neutrophil count (OR = 0.944; 95% CI, 0.892-0.999; p = 0.048). Three single nucleotide polymorphisms (SNPs), namely rs143384, rs66989638, and rs8053839, associated with knee or hip osteoarthritis, as well as one SNP (rs112635299) associated with osteoarthritis metabolism, were identified through colocalization analysis.*** ***Conclusions:*** ***This two sample MR analysis found a causal negative association between foods containing saturated fatty acids and both osteoarthritis and biomarkers. Colocalization analysis identified three potential drug targets and one metabolite that could serve as a diagnostic marker.***

Keywords: Osteoarthritis, Saturated fatty acid, Mendelian randomization, Colocalization analysis, Causal association**.**

1. Introduction

Weight-bearing joints like the knee and hip are primary sites of osteoarthritis (OA), and are also a common disabling disease [1,2]. OA not only limits physical activities but also negatively affects mental well-being [3]. The prevalence of OA is increasing due to global aging and rising obesity rates [4], compounded by the impact of COVID-19 [5]. Current estimates indicate that 250 million individuals are affected, and by 2032, the proportion of individuals aged 45 and above with osteoarthritis is projected to rise from 26.6% to 29.5% for any body part, from 13.8% to 15.7% for the knee, and from 5.8% to 6.9% for the hip [2,4]. This presents significant challenges for healthcare systems and imposes substantial socioeconomic costs. Numerous studies have investigated risk factors for osteoarthritis, including age, gender, obesity, and comorbiddities [6,7]. However, the impact of the intake of different foods containing fatty acids on OA has not been clearly defined (Table 1).

Table 1: Representative study of the association between saturated fatty acid intake and different types of arthritis.

Author	Type	Time	Results	Note
Wang, Y	Review	2023	GPR120 receptor has beneficial effects on bone metabolism, promotes osteoblast formation and differentiation, and inhibits osteoclast apoptosis.	GPR120 (G -protein coupled receptor 120, GPR120) is also known as free fatty acid receptor 4 (FFAR4).
Wu, C L	Research Article	2017	Levels of pentanoic acid (C15:0, odd chain SFA) and palmitic acid are negatively correlated with joint degeneration.	Serum levels of ω-3 PUFAs were negatively correlated with osteoarthritis and wound size
Felson, D T	Research Article	2021	No clear relationship was found between the levels of Saturated FAs, n-3 FAs or n-6 FAs and the risk of developing OA or OA-related outcomes.	Only the relationship between blood fatty acids and OA was observed
Sekar, S	Research Article	2020	Dietary intervention did not alter the severity of cartilage damage in OA	Consumption of tallow, palmitic or stearic acids may exacerbate pain symptoms, whereas consumption of lauric or myristic acids may improve pain symptoms
Jin, X	Research Article	2022	Consumption of diets rich in SFAs or n-6 PUFAs may exacerbate osteoarthritis, whereas consumption of diets rich in n-3 PUFAs may have a protective effect.	The effects of MUFAs are not consistent
Tan, L	Research Article	2021	Palmitic acid promotes endoplasmic reticulum stress and cartilage damage in the mouse knee.	Both saturated and unsaturated fatty acids affect knee synovial cartilage.

Saturated fatty acids are one of the major forms of lipids in plasma and exist as triglycerides in the body, usually in

mammalian adipose tissue, and are an important source of energy in mammals [8]. Besides red meat, saturated fatty acids can be found in dairy products and certain vegeTable oils within our daily diet. The EAT-Lancet guidelines recommend an ideal daily diet with 84 grams of animal-based food, including 7 grams each of beef and pork, along with 250 grams of dairy products [4]. Apart from providing energy, saturated fatty acids contribute to membrane fluidity and lipid storage in tissues while activating pro-inflammatory pathways, thereby playing a significant role in various chronic diseases such as cardiovascular disease, metabolic syndrome, and obesity [8-10]. While most studies reported that saturated fatty acids have adverse effects on bone and joint health [11,12], there are still some studies finding no definitive correlation [13,14], and even suggesting a negative association between saturated fatty acids and osteoarthritis [15,16] (see Table 1). However, there is a dearth of high-level evidence studies, particularly randomized controlled trials, that address the issue of causal inference. Consequently, it is imperative to investigate the causal relationship between the consumption of different foods containing saturated fatty acids and osteoarthritis.

Randomized controlled studies on the intake of a specific food are difficult to conduct because restricting subjects to a specific food is difficult and unethical. Mendelian randomization (MR) is a solution to this problem. Compared with traditional observational studies, MR analyses can overcome the interference of a variety of irrelevant confounding factors, since randomization of alleles always precedes disease onset. In addition, the random pairing of genes at the time of conception and the independent classification of polymorphisms in inherited genes allow MR analyses to minimize the effects of confounding factors by using genetic markers as instrumental variables (IVs) for exposure [17]. The availability of large-scale genome-wide association studies (GWASs) has enabled further exploration of causal relationships.

Moreover, Bayesian method co-localization analysis can calculate the probability of two traits sharing a causal genetic variance [18]. Nowadays, large-scale quantitative trait loci (QTL) data are generated to establish associations between genotypes and protein abundance (pQTL) as well as gene expression (eQTL) [19, 20]. By associating GWAS data with these multidimensional QTL data, specific pathways and candidate genes can be prioritized, leading to the identification of potential genes involved in osteoarthritis pathogenesis. Consequently, we can determine whether there is shared causal genetic variation between the two traits, contributing to a better understanding of osteoarthritis's underlying mechanisms. This comprehensive analysis helps to strengthen the knowledge of the pathogenesis and potential genes of osteoarthritis and provides important clues for further research.

Therefore, the aim of this study is to address the following key research questions: (1) What is the association between the intake of different foods containing saturated fatty acids and osteoarthritis - negative, neutral, or positive? (2) How does the intake of different foods containing saturated fatty acids impact biomarkers associated with osteoarthritis? (3) Which specific single nucleotide polymorphism (SNP) regulates the expression of osteoarthritis, which genes does it regulate, and what are the metabolites associated with osteoarthritis?

2. Materials and Methods

2.1 Study Design

A Mendelian randomization study with a two-sample design was first used. The schematic of the study design and the three key hypotheses of MR are shown in **Figure 1**: (A) single nucleotide polymorphisms (SNPs) are strongly correlated with intake of saturated fatty acid-containing foods; (B) SNPs are independent of known confounders; and (C) SNPs affect osteoarthritis only through intake of saturated fatty acid-containing foods and Biomarkers.

For colocalization analysis, Five mutually exclusive hypotheses was tested: (1) no causal SNP is found for either trait (Ho); (2) only trait 1 has a causal SNP (H_1); (3) only trait 2 has a causal SNP (H_2); (4) both traits have a causal SNP, but the two causal SNPs are different (H_3); (5) both traits have a causal SNP, and share the same SNP (H_4) [21] (**Figure 2**).

This study was conducted according to the guidelines of Burgess [21] and Emdin [22] and reported according to the STROBE-MR statement [23]. Data for this study were analyzed from April 18, 2023 to July 26, 2023.

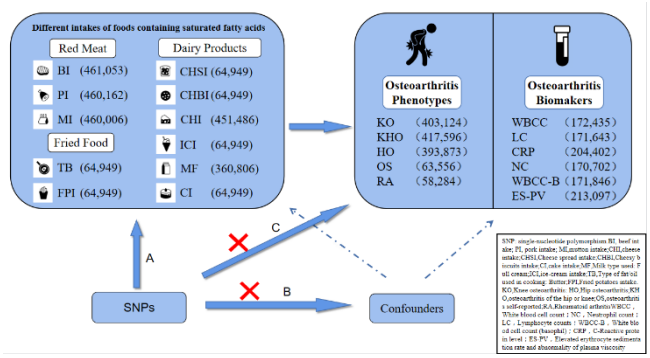

Figure 1: Three key assumptions of the Mendelian randomization study. (A) SNPs are strongly associated with different intake of foods containing saturated fatty acids; (B) SNPs are independent of confounders; (C) SNPs must only affect osteoarthritis and biomarkers via different intake of foods containing saturated fatty acids.

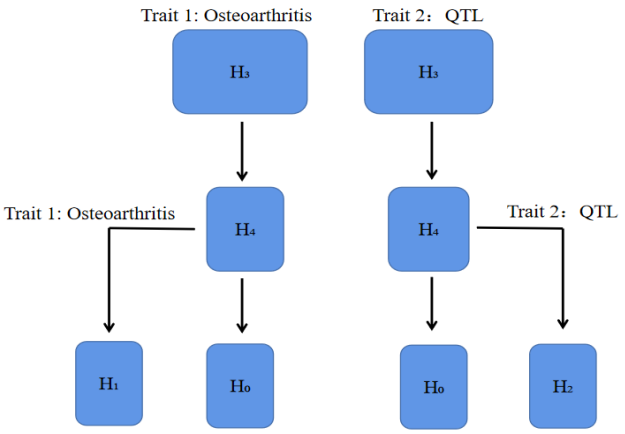

Figure 2: Five hypotheses for the colocalization analysis. Ho: Represents no associations with either of the two features. H_1: Indicates a significant association with Trait 1 (Osteoarthritis), but not with Trait 2 (eQTL/pQTL). H_2: Indicates a significant

association with Trait 2 (eQTL/pQTL), but not with Trait 1 (Osteoarthritis). H_3: Indicates independent SNP associations with both features. H_4: Indicates shared SNP associations between Trait 1 (Osteoarthritis) and Trait 2 (eQTL/pQTL).

2.2 Data Sources

All data used in the two-sample MR were at the pooled level and were derived from the public GWAS, with the main population being Europeans (both men and women). The GWAS pooled statistics for different food intakes were derived from the UK BioBank study, which evaluated the relationship between different food intakes and SNPs, including: beef intake (BI) (n = 461,053); pork intake (PI) (n = 460,162); mutton intake (MI) (n = 460,006); cheese intake (CHI) (n = 451,486); Cheese spread intake (CHSI) (n = 64,949); Cheesy biscuits intake (CHBI) (n = 64,949); cake intake (CI) (n = 64,949); full cream milk used (MF) (n = 360,806); ice-cream intake (ICI) (n = 64,949); Use butter when cooking (TB) (n = 64,949); Fried potatoes intake (FPI) (n = 64,949) [24]. Knee osteoarthritis (KO) (n = 403,124); Hip osteoarthritis (HO) (n = 393,873) and osteoarthritis of the hip or knee (KHO) (n = 417,596) was obtained from the results reported by Tachmazidou et al. [19]. Osteoarthritis self-reported (OS) (n = 63,556) was obtained from the results reported by Zengini et al. [25]. Rheumatoid arthritis (RA) (n = 58,284) was obtained from the results reported by Ha et al. [26]. White blood cell count (WBCC) (n = 172,435); Neutrophil count (NC) (n = 170,702); Lymphocyte counts (LC) (n = 171,643) and white blood cell count (basophil) (WBCC-B) (n = 171,846) was obtained from the results reported by Astle et al. [27]. C-Reactive protein level (CRP) (n = 204,402) was obtained from the results reported by Ligthart et al. [28]. Elevated erythrocyte sedimentation rate and abnormality of plasma viscosity (ES-PV) (n = 213,097) was obtained from FinnGen Consortium. Details of all data are in Supplementary Table S1,S2,S3.

All data for QTL were obtained from the public R package, "MR Instruments" on GitHub (https://github.com/mrcieu/mrinstruments).

Ethical approval was not required for this study, as all GWAS data included in the study were publicly available and ethical clearance was obtained from the respective institutional review boards. Neither patients nor the public were involved in this study.

2.3 Selection and Validation of Data

For the selection of suiTable SNPs for two-sample MR, we performed three steps. First, we selected SNPs associated with beef intake, pork intake, lamb intake, and cheese intake at a genome-wide significance threshold ($p < 5 \times 10\text{-}8$). For other exposure data where the number of SNPs was too low at the default threshold, a method of using SNPs with relaxed thresholds was used [29,30] to obtain more exposed SNPs and to include data that met the relaxation threshold ($p < 5 \times 10\text{-}6$) in the study. Second, the independence between the selected SNP is assessed on the basis of pairwise chained disequilibria [31]. SNPs associated with more SNPs or with higher p-values were deleted when $r^2 > 0.001$ (clumping window of 10,000 kb). Third, a data harmonization step was performed prior to MR analyses because the effect of a SNP on exposure and outcome must correspond to the same allele.

For colocalization analysis, We assigned the default prior probabilities for a SNP being associated with osteoarthritis ($P_1 = 1 \times 10^{-4}$), a SNP is a significant QTL ($P_2 = 1 \times 10^{-4}$) and for a SNP being associated with both traits ($P_{12} = 1 \times 10^{-5}$) [21]. We mainly focused on the last hypothesis H4 and posterior probability (PP) was used to quantify support for H4 (denoted as PPH4). We defined strong evidence of colocalization at $PPH4 \geq 0.75$ [32].

2.4 Statistical Analysis

2.4.1 Two-sample Mendelian randomization

When exposure SNP data collection was complete, we first manually screened all IVs for potential pleiotropy by searching the PhenoScanner GWAS database (version 2; http://phenoscanner.medschl.cam.ac.uk) for each IV used in the analysis for potential confounders associated with existing associations; then, we removed these SNPs to control for multiple effects and to see if the primary outcome would be reversed. An inverse-variance weighted (IVW) meta-analysis under a random-effects model was used as the main analysis. The following two methods, including weighted median and MR-Egger, will be used as sensitivity analyses. The weighted median approach provides a valid estimate if more than 50% of the information comes from valid IVs [33]. Cochrane's Q-value could indicate the heterogeneity among the selected IVs, and if $p < 0.05$, the multiplicative random-effect IVW method was used. MR-Egger was employed for assessing the horizontal multiplicity of the selected IVs, and if $p < 0.05$, horizontal multiplicity was suggested [34]. In addition, leave-one-out sensitivity analyses were performed to determine whether the overall estimates were disproportionately affected by individual SNPs.

2.4.2 Colocalization analysis

Colocalization analysis of SNPS shared from QTL and osteoarthritis datasets was performed using the "coloc.abf" function of the coloc R software package (version 3.2.1).

3. Results

3.1 SNP Selection and Validation

In summary, the included studies were published between 2016 and 2021 and were mainly based on European populations. A total of 256 instrumental variables were included in the studies as exposures (**Supplementary Table S4**); a total of 51 instrumental variables in the direction of different food intake to different osteoarthritis phenotypes were excluded by manually determining that they were potentially pleiotropic; a total of 4 instrumental variables in the direction of different food intake to osteoarthritis biomarkers were excluded by manually determining that they were potentially pleiotropic (**Supplementary Table S5**).

3.2 Osteoarthritis

According to the IVW analysis, there were 5 positives out of 55 data. The genes predicted that each standard deviation increase in cheese intake was negatively associated with three

osteoarthritis phenotypes, including: KO (OR = 0.605; 95% CI, 0.479-0.764; p < 0.001); KHO (OR = 0.676; 95% CI, 0.553-0.826; p < 0.001); OS (OR = 0.720; 95% CI, 0.521-0.995; p = 0.047). Each standard deviation increase in cheese cracker intake was negatively associated with two osteoarthritis phenotypes, including: KO (OR = 0.485; 95% CI, 0.239-0.987; p = 0.046); KHO (OR = 0.535;95% CI, 0.315-0.909; p = 0.021) (**Figure 3A**). The results of all IVW data analyses are presented in **Supplementary Table S6**. Weighted median and MR-Egger analyses produced consistent estimates but with lower precision (**Table 2**). No evidence of directed pleiotropy was found. Certain osteoarthritis phenotypes showed high heterogeneity. Therefore, IVW meta-analyses using a random-effects model to mitigate the effects of heterogeneity yielded consistent results with the original analyses.

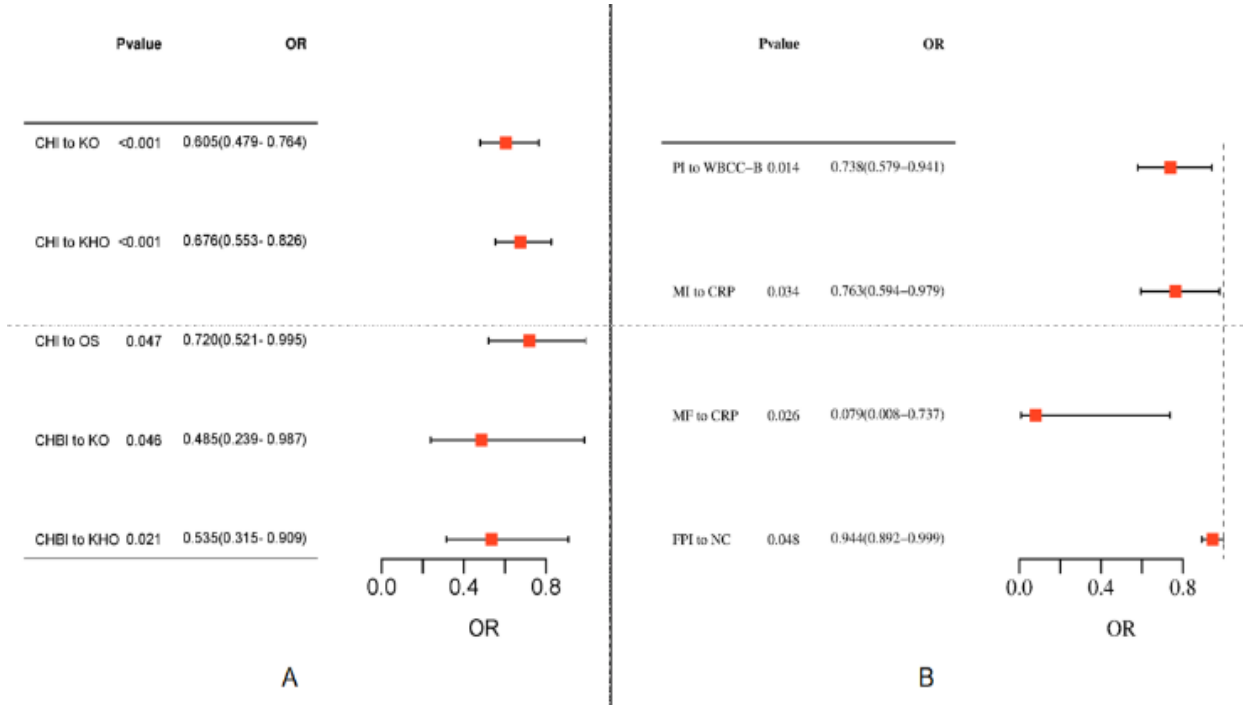

Figure 3: (A) Associations of genetically predicted different food intake with osteoarthritis. (B) Associations of genetically predicted different food intake with osteoarthritis biomarkers. (Only positive results). CI, confidence interval; OR, odds ratio; SNP, single-nucleotide polymorphism.

Table 2: Associations between genetically predicted different food intake and osteoarthritis in sensitivity analyses using the weighted-median and MR-Egger methods.CI, confidence interval; MR, Mendelian randomization; OR, odds ratio.

Direction	Weighted Median		MR-Egger		Pleiotropy		Heterogeneity	
	OR (95% CI)	p	OR (95% CI)	p	Intercept	p	Q	p
CHI to KO	0.608(0.452-0.816)	0.001	0.587(0.226-1.527)	0.281	0.001	0.952	66	0.022
CHI to KHO	0.842(0.654-1.082)	0.179	0.840(0.370-1.906)	0.679	-0.004	0.593	73	0.005
CHI to OS	0.689(0.438-1.085)	0.108	0.538(0.147-1.970)	0.355	0.005	0.652	50	0.225
CHBI to KO	0.618(0.255-1.495)	0.286	0.392(0.116-1.317)	0.143	0.003	0.670	36	0.056
CHBI to KHO	0.671(0.342-1.319)	0.247	0.393(0.161-0.961)	0.052	0.005	0.408	30	0.192

The remaining data for the positive results are shown in the Supplementary Figures. Scatter plots between different food intakes and different OA phenotypes are shown in **Supplementary Figure S1**, and a forest plot is shown in **Supplementary Figure S2**, both of which show similar results. Exclusion-by-exclusion sensitivity analyses in **Supplementary Figure S3** indicate that the overall estimates are not unduly influenced by any individual SNPs excessively. The funnel plot in **Supplementary Figure S4** also shows no evidence of horizontal pleiotropy.

3.3 Osteoarthritis Biomarkers

According to the results of IVW analysis, there were 4 positives out of 66 data, of which the TB to CRP direction could not be analyzed because there were not enough SNPs. Genetically predicted increased PI per standard deviation was negatively associated with WBCC-B (OR = 0.738; 95% CI, 0.579-0.941; p =0.014); MI per standard deviation increase was negatively correlated with CRP (OR = 0.763; 95% CI, 0.594-0.979; p = 0.034); MF per standard deviation increase was negatively correlated with CRP (OR = 0.079; 95% CI, 0.008-0.737; p = 0.026); FPI per standard deviation increase was negatively correlated with NC (OR = 0.944; 95% CI, 0.892-0.999; p = 0.048)(Figure 3B). None of the remaining significant associations were found, and the results of all IVW data analyses are presented in **Supplementary Table S7**. For the positive data on osteoarthritis biomarker outcomes, weighted median and MR-Egger analyses yielded consistent estimates with low precision (**Table 3**). No evidence of targeted pleiotropy was found. Some osteoarthritis biomarker phenotypes showed high heterogeneity. Therefore, IVW meta-analysis with random effects modeling was used to mitigate the effect of heterogeneity, and the results of the analysis remained consistent with the original.

Table 3: Associations between genetically predicted different food intake and biomarkers in sensitivity analyses using the weighted-median and MR-Egger methods.

Direction	Weighted Median		MR-Egger		Pleiotropy		Heterogeneity	
	OR (95% CI)	p	OR (95% CI)	p	Intercept	p	Q	p
PI to WBCC-B	0.851(0.634-1.143)	0.284	0.453(0.081-2.519)	0.383	0.006	0.283	19	0.134
MI to CRP	0.813(0.584-1.130)	0.217	01.496(0.435-5.145)	0.540	-0.007	0.307	10	0.385
MF to CRP	0.551(0.143-2.119)	0.386	<0.001(<0.001-16.160)	0.175	0.026	0.287	46	<0.001
FPI to NC	0.986(0.914-1.064)	0.720	0.986(0.857-1.135)	0.851	-0.002	0.513	20	0.3432

In the accompanying Figures, the remaining data on the association of positive results between different food intakes and markers of osteoarthritis are presented. Scatter plots, forest plots, results of leave-one-out sensitivity analyses, and funnel plots are shown in **Supplementary Figure S5**, **Supplementary Figure S6**, **Supplementary Figure S7**, and **Supplementary Figure S8**, respectively, in which similar results can be observed.

3.4 Colocalization Analysis

In the colocalization analysis, PPH4 ≥ 0.75 was taken as the standard [32]. In OA with eQTL, KO to eQTL (rs143384, PPH4 = 1.000), HO to eQTL (rs66989638, PPH4 = 1), KHO to eQTL (rs8053839, PPH4 = 1) (**Figure 4A**). In contrast, in the KO and KHO co-localization analyses with metabonomics-pQTL (metab-pQTL), the results were surprisingly consistent, both pointing to rs112635299 (PPH4 = 1) (**Figure 4B**). Among them, HO and metab-pQTL, and KO/HO/KHO and proteomic-pQTL could not be analyzed because there were not enough SNPs. Detailed data of co-localization analysis are shown in Supplementary Table S8.

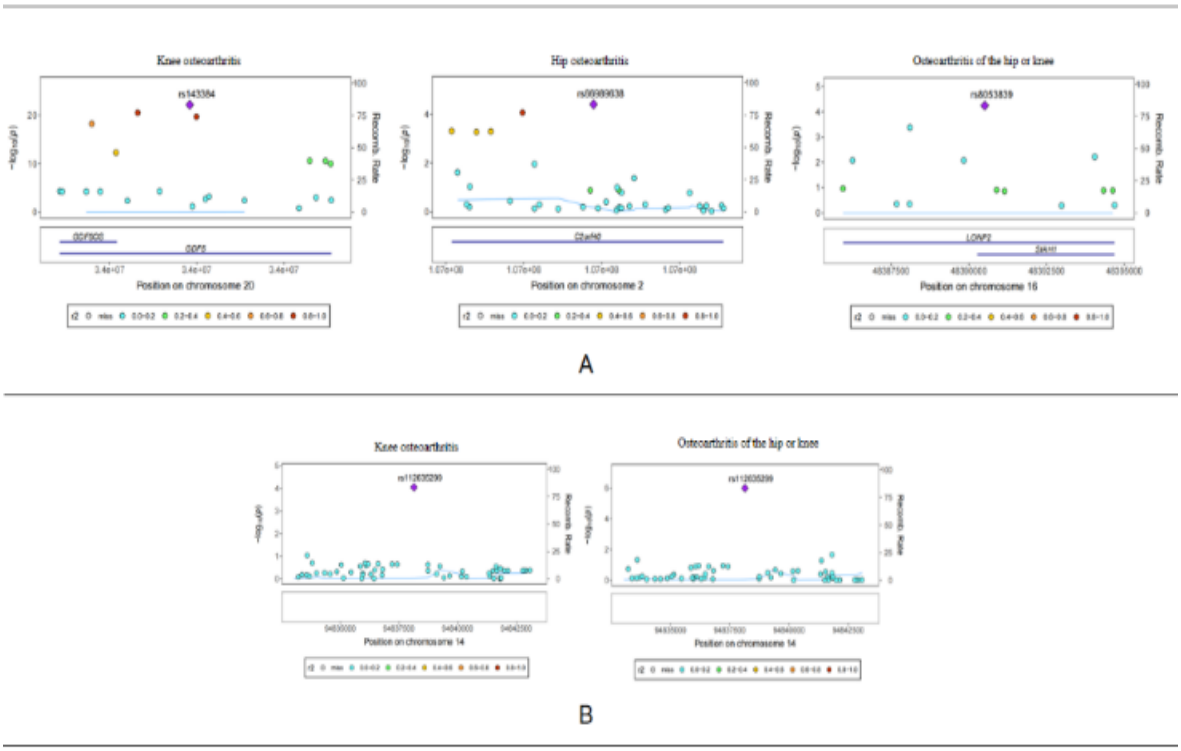

Figure 4: The co-localization analysis of KO, HO, KHO with eQTL and pQTL. (A)Results of colocalization analysis between different osteoarthritis phenotypes and eQTL. (B)Results of colocalization analysis between different osteoarthritis phenotypes and metab-pQTL.

4. Discussion

4.1 MR Analysis

To our knowledge, this is the first MR study to explore the causal relationship between the intake of foods containing saturated fatty acids and osteoarthritis and related markers. In this two-sample MR study, we comprehensively evaluated the relationship between the intake of foods containing saturated fatty acids and osteoarthritis and osteoarthritis biomarkers. Surprisingly, our results showed that the intake of cheese and cheese crackers was negatively associated with osteoarthritis. For osteoarthritis biomarkers, pork, lamb, whole milk and fried potatoes were all negatively associated. We performed a visual analysis and assessment of the content of saturated fatty acids in different food sources in relation to their respective odds ratios (OR)(Figure 5), Data on the levels of saturated fatty acids in various foods were obtained from the USDA Food and Nutritional Database [35], and the OR value represent the ratio of the odds of success to the odds of failure between two events. Detailed data on the saturated fatty acid content of the foods involved in this study are presented in

Supplementary Table S9.

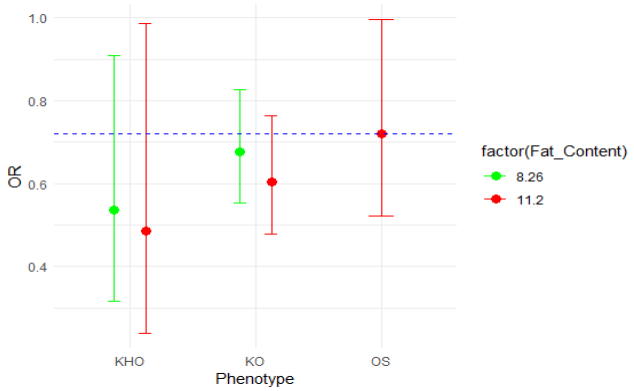

Figure 5: Comparison of SFA Content and OR Values among Different Phenotypes. (Only positive results)

We observed an increasing protective effect of saturated fatty acids against the development of osteoarthritis as the dietary intake of saturated fatty acids increased. However, it is important to note that not all foods containing saturated fatty acids exhibit an association with osteoarthritis. This discrepancy is likely attribuTable to variations in the specific types of saturated fatty acids present in different foods.

Table 4: MR Results of the effect of pork intake on different osteoarthritis phenotypes. [a]The outlier was rs838133.

Direction Without outlier[a]	OR (95% CI)	p	OR (95% CI) With outlier	p
PI to KO	2.874(1.013-8.159)	0.047	2.642(1.008 - 6.919)	0.048
PI to HO	2.036(1.001-4.143)	0.049	2.358(1.201 - 4.629)	0.013
PI to KHO	2.597(1.389-4.857)	0.003	1.479(0.074 - 258.601)	0.001

Meanwhile, significant data results were obtained when potential pleiotropy was not excluded in the direction of pork intake to KO/HO/KHO, and the data analysis still yielded significant results even after excluding rs838133 (Table 4); whereas, when rs7641973, rs9379832, rs1355171 and rs12721051 were deleted, the data analysis results were no longer significant ($P > 0.05$). We conducted a search on the NCBI website and found that rs7641973 is located on the LOC105377178 gene on chromosome 3, which belongs to the ncRNA class, and that rs9379832 is a SNP located on chromosome 6, which we did not find any significant information about. We did not search any valuable information about it; rs1355171 is a SNP located on the CCDC171 gene on chromosome 9, which was first found to be associated with cholesterol intake in 2018 [36], and its presence was later found in chronic kidney disease [37]; rs12721051 is a SNP located on the APOC1 gene on chromosome 19, the APOC1 gene is mainly expressed in the liver, the protein encoded by this gene plays a crucial role in the metabolism of high-density lipoprotein (HDL) and very-low-density lipoprotein (VLDL) [38].

4.2 Colocalization Analysis

By conducting colocalization analysis, we identified three SNPs associated with the regulation of OA development, rs143384, rs66989638 and rs8053839. rs143384 is located on chromosome 20 within the GDF5 gene, which has been implicated in OA, especially in the knee joint [39]. Recent research has shown that GDF5 expression in joints plays a crucial role in maintaining joint stability [40]. Furthermore, ERG, a transcription factor regulated by downstream signaling of GDF5, contributes to increased tolerance and structural stability after articular cartilage formation [41]. Regarding rs66989638, it is situated on chromosome 2 within the ECRG4 gene. While ECRG4 is primarily associated with esophageal cancer [42], it has also been found to be abundantly expressed in cartilage [43], where it plays a significant role in wound healing [44]. As for rs8053839, it is located on chromosome 16 within the LONP2 gene. LONP2 is essential for removing oxidatively modified proteins, preventing their accumulation, and ensuring proper peroxisome function [45]. We hypothesize that the expression of LONP2 in OA may involve the removal of oxidatively modified proteins, thereby alleviating intracellular oxidative stress and reducing the severity of inflammatory responses. However, further in-depth studies are required to substantiate this hypothesis.

Through colocalization analysis between KO, KHO, and metab-pQTL, we identified the same SNP, rs112635299, located on chromosome 14. However, we did not find any metabolites or genes associated with this SNP. It has been previously linked to low-density lipoprotein cholesterol (LDL) based on the results of a database query using MR Instruments. The query revealed its relationship to various LDL-related measures, including LDL cholesterol (LDL.C), medium-sized LDL cholesterol ester (M.LDL.CE), medium-sized LDL cholesterol (M.LDL.C), medium-sized LDL length (M.LDL.L), medium-sized LDL particle (M.LDL.P), small dense LDL cholesterol (S.LDL.C), and small dense LDL length (S.LDL.L). In the future, we may be able to diagnose osteoarthritis more quickly by using these measurements.

4.3 Comparisons with Previous Traditional Studies

The effect of saturated fatty acids on osteoarthritis remains a topic with conflicting findings. Certain studies indicate a detrimental impact of saturated fatty acid intake on osteoarthritis [11,12]. Conversely, some studies have suggested that saturated fatty acid intake does not directly cause osteoarthritis, but only aggravates osteoarthritis pain [13,14]. Notably, there are also studies suggesting a negative association between saturated fatty acid intake and osteoarthritis [15], aligning with our own findings. However, it is important to consider that these conclusions primarily stem from animal studies, while human dietary patterns are influenced by various factors such as personal habits, regional customs, and economic circumstances.

MR is capable of emulating the design of a randomized controlled trial [46]. Nonetheless, it possesses distinct advantages. Randomized controlled trials can present ethical and health challenges when enforcing strict dietary interventions on subjects. In contrast, Mendelian randomization studies circumvent these issues by utilizing naturally occurring genetic variations, obviating the need for specific interventions on individuals. Moreover, since MR is grounded in extensive data derived from large-scale human cohorts observed within real populations, the outcomes achieved are highly representative, externally valid, and

applicable to real-world scenarios. Consequently, MR better informs the management of population-wide dietary health.

Conducting co-localization analysis enables the identification of genes regulating OA and its associated protein products, effectively minimizing interference from confounding factors. This approach enhances result reliability and diminishes the impact of confounding factors that are challenging to exclude in experimental settings. Additionally, the identification of genetic variants within the same gene locus for both genes and protein products offers robust evidence supporting the regulatory role of genes on protein products. Moreover, leveraging genetic variation as a means to mimic randomized experiments yields a closer approximation of causality.

4.4 Limitations

This study possesses several limitations that warrant attention. Firstly, completely excluding the impact of potential directional pleiotropy in any Mendelian randomization (MR) study proves highly challenging. Secondly, the dietary intake studies and osteoarthritis outcome data utilized in this research originate from Europe, potentially resulting in underrepresentation of other ethnic groups. Insufficient sample availability from diverse regions within the publicly accessible GWAS database necessitated the omission of multiple-region validation. Thirdly, variations exist in the types of fatty acids found in food. The original GWAS study solely addressed diverse food types without providing comprehensive information on the primary nutrient composition or co-consumption of food items, thus limiting further analyses. Furthermore, our approach employed the intake of different foods containing saturated fatty acids as a phenotype instead of considering specific subtypes of saturated fatty acids. This decision was prompted by the fact that only saturated fatty acids are characterized as phenotypes in the GWAS database, while other databases lack sufficiently detailed data on saturated fatty acid subtypes. Hence, typical foods containing saturated fatty acids were used as representative phenotypes.

Regarding co-localization analyses, several aspects require attention. Firstly, the dataset encompassing genes detected in OA, eQTL, and pQTL was relatively small, resulting in a limited number of genes identified in both analyses. Expanding the dataset to include a larger collection of osteoarthritis samples can effectively address this concern. Secondly, gene expression and regulation represent intricate processes that can yield diverse outcomes depending on the circumstances. Future studies should consider incorporating the effects of other tissues, such as intestinal flora and renal abnormalities, into the analysis. Thirdly, our analysis was restricted to a single KO sample, necessitating a more extensive dataset to facilitate further validation.

4.5 Importance

Despite the inherent limitations, this MR study offers novel insights into evaluating causal relationships between saturated fatty acids and osteoarthritis while circumventing the common challenges encountered in observational studies. Should an inverse causal relationship between saturated fatty acid consumption and osteoarthritis be confirmed, it would warrant a reevaluation of the existing dietary guidelines for osteoarthritis patients.

The colocalization analysis yielded the identification of potential genes suiTable for targeting in drug development and metabolites that could serve as diagnostic markers for osteoarthritis.

5. Conclusions

This two-sample Mendelian randomization (MR) analysis revealed a causal, inverse association between cheese intake and knee osteoarthritis, hip or knee osteoarthritis, and self-reported osteoarthritis. Furthermore, a causal, inverse association between cheese cracker intake and hip or knee osteoarthritis was observed. Additionally, a causal, inverse association was found between pork intake and basophil count. Negative causal associations also existed between lamb and whole milk intake and CRP levels, as well as between fried potato intake and Neutrophil count. Genes linked to the development of osteoarthritis were identified: GDF5, ECRG4, LONP2. Potential blood markers for diagnosing osteoarthritis were identified, including LDL-C, and various metabolites such as M.LDL.CE, M.LDL.C, M.LDL.L, M.LDL.P, S.LDL.C, and S.LDL.L.

Supplementary Materials:

The following supporting information can be downloaded at: www.mdpi.com/xxx/s1, Figure S1: Scatter plot of the association of different food intake with osteoarthritis; Figure S2: Forest plot of the association of different food intake with osteoarthritis; Figure S3: Leave-one-out sensitivity analysis of the association of different food intake with osteoarthritis; Figure S4: Funnel plot of the association of different food intake with osteoarthritis; Figure S5: Scatter plot of the association of different food intake with osteoarthritis biomarkers; Figure S6: Forest plot of the association of different food intake with osteoarthritis biomarkers; Figure S7: Leave-one-out sensitivity analysis of the association of different food intake with osteoarthritis biomarkers; Figure S8: Funnel plot of the association of different food intake with osteoarthritis biomarkers; Table S1: 11 GWAS data for foods containing saturated fatty acids; Table S2: 5 types of GWAS data for arthritis; Table S3: 6 types of GWAS data for osteoarthritis biomarkers; Table S4: Single nucleotide polymorphisms used as instrumental variables in the Mendelian randomization analyses of cheese intake; Table S5: Manually detected potential pleiotropy in the Phenoscanner database; Table S6: Combined analysis of IVW effect sizes and 95%CI interval widths in dietary intake cohorts with varying levels of saturated fatty acid content for distinct osteoarthritis phenotypes; Table S7: Combined analysis of IVW effect sizes and 95%CI interval widths in dietary intake cohorts with varying levels of saturated fatty acid content for distinct osteoarthritis biomarkers; Table S8: Results of colocalization analysis of osteoarthritis with eQTL and pQTL (positive results only); Table S9: The amount of saturated fatty acids in different foods.

Author Contributions:

Conceptualization, K.-G.C., X.-L.G. and Y.-D.S.; methodology, K.-G.C. and Y.-F.L.; software, X.-L.G.; validation, B.D. and L.L.; formal analysis, K.-G.C. and X.-Q.Z.; investigation, K.-G.C., Y.-D.S., Y.-F.L, X.-Q.Z., X.-L.G., B.D. and L.L.; resources, Y.-D.S.; data curation, X.-Q.Z. and Y.-F.L.; writing—original draft preparation, K.-G.C.; writing—review and editing, Y.-D.S., Y.-F.L, X.-Q.Z., X.-L.G., B.D. and L.L.; visualization, K.-G.C.; supervision, Y.-D.S.; project administration, Y.-D.S. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding

Institutional Review Board Statement: Not applicable

Informed Consent Statement: Not applicable.

Data Availability Statement: Data Availability Statement: Data available in a publicly accessible repository that does not issue DOIs. Publicly available datasets were analyzed in this study. This data can be found here: [https://gwas.mrcieu.ac.uk/datasets/]; [http://www.phenoscanner.medschl.cam.ac.uk/].

Acknowledgments:

The authors would like to express their heartfelt appreciation for the technical support provided by Dr. YF. WANG from Paris-Saclay University.

Conflicts of Interest:

All authors declare no conflict of interest.

References

[1] Reyes C, Leyland KM, Peat G, Cooper C, Arden NK, Prieto-Alhambra D: Association Between Overweight and Obesity and Risk of Clinically Diagnosed Knee, Hip, and Hand Osteoarthritis: A Population-Based Cohort Study. ARTHRITIS RHEUMATOL 2016, 68(8): 1869-1875.

[2] Turkiewicz A, Petersson IF, Bjork J, Hawker G, Dahlberg LE, Lohmander LS, Englund M: Current and future impact of osteoarthritis on health care: a population-based study with projections to year 2032. OSTEOARTHR CARTILAGE 2014, 22(11): 1826-1832.

[3] Vina ER, Kwoh CK: Epidemiology of osteoarthritis: literature update. CURR OPIN RHEUMATOL 2018, 30(2):160-167.

[4] Willett W, Rockstrom J, Loken B, Springmann M, Lang T, Vermeulen S, Garnett T, Tilman D, DeClerck F, Wood A et al: Food in the Anthropocene: the EAT-Lancet Commission on healthy diets from sustainable food systems. LANCET 2019, 393(10170): 447-492.

[5] Castro DRF, Melo L, Berenbaum F: Tackling osteoarthritis during COVID-19 pandemic. ANN RHEUM DIS 2021, 80(2):151-153.

[6] Zhang Y, Jordan JM: Epidemiology of osteoarthritis. RHEUM DIS CLIN N AM 2008, 34(3):515-529.

[7] Silverwood V, Blagojevic-Bucknall M, Jinks C, Jordan JL, Protheroe J, Jordan KP: Current evidence on risk factors for knee osteoarthritis in older adults: a systematic review and meta-analysis. OSTEOARTHR CARTILAGE 2015, 23(4):507-515.

[8] Boren J, Taskinen MR, Bjornson E, Packard CJ: Metabolism of triglyceride-rich lipoproteins in health and dyslipidaemia. NAT REV CARDIOL 2022, 19(9):577-592.

[9] Boden G, Shulman GI: Free fatty acids in obesity and type 2 diabetes: defining their role in the development of insulin resistance and beta-cell dysfunction. EUR J CLIN INVEST 2002, 32 Suppl 3:14-23.

[10] Vessby B: Dietary fat, fatty acid composition in plasma and the metabolic syndrome. CURR OPIN LIPIDOL 2003, 14(1):15-19.

[11] Jin X, Dong X, Sun Y, Liu Z, Liu L, Gu H: Dietary Fatty Acid Regulation of the NLRP3 Inflammasome via the TLR4/NF-κB Signaling Pathway Affects Chondrocyte Pyroptosis. OXID MED CELL LONGEV 2022, 2022:3711371.

[12] Tan L, Harper LR, Armstrong A, Carlson CS, Yammani RR: Dietary saturated fatty acid palmitate promotes cartilage lesions and activates the unfolded protein response pathway in mouse knee joints. PLOS ONE 2021, 16(2):e247237.

[13] Felson DT, Misra D, LaValley M, Clancy M, Chen X, Lichtenstein A, Matthan N, Torner J, Lewis CE, Nevitt MC: Fatty acids and osteoarthritis: the MOST study. OSTEOARTHR CARTILAGE 2021, 29(7):973-978.

[14] Sekar S, Panchal SK, Ghattamaneni NK, Brown L, Crawford R, Xiao Y, Prasadam I: Dietary Saturated Fatty Acids Modulate Pain Behaviour in Trauma-Induced Osteoarthritis in Rats. NUTRIENTS 2020, 12(2).

[15] Wu CL, Kimmerling KA, Little D, Guilak F: Serum and synovial fluid lipidomic profiles predict obesity-associated osteoarthritis, synovitis, and wound repair. SCI REP-UK 2017, 7:44315.

[16] Wang Y, Liu H, Zhang Z: Recent Advance in Regulatory Effect of GRP120 on Bone Metabolism. AGING DIS 2023.

[17] Smith GD, Ebrahim S: 'Mendelian randomization': can genetic epidemiology contribute to understanding environmental determinants of disease? INT J EPIDEMIOL 2003, 32(1):1-22.

[18] Giambartolomei C, Vukcevic D, Schadt EE, Franke L, Hingorani AD, Wallace C, Plagnol V: Bayesian test for colocalisation between pairs of genetic association studies using summary statistics. PLOS GENET 2014, 10(5):e1004383.

[19] Tachmazidou I, Hatzikotoulas K, Southam L, Esparza-Gordillo J, Haberland V, Zheng J, Johnson T, Koprulu M, Zengini E, Steinberg J et al: Identification of new therapeutic targets for osteoarthritis through genome-wide analyses of UK Biobank data. NAT GENET 2019, 51(2):230-236.

[20] Jaeger K, Selent C, Jaehme W, Mahr S, Goebel U, Ibrahim S, Vollmar B, Mueller-Hilke B: The genetics of osteoarthritis in STR/ort mice. OSTEOARTHR CARTILAGE 2008, 16(5):607-614.

[21] Burgess S, Davey SG, Davies NM, Dudbridge F, Gill D, Glymour MM, Hartwig FP, Holmes MV, Minelli C, Relton CL et al: Guidelines for performing Mendelian randomization investigations. Wellcome Open Res 2019, 4:186.

[22] Emdin CA, Khera AV, Kathiresan S: Mendelian Randomization. JAMA-J AM MED ASSOC 2017, 318(19):1925-1926.

[23] Skrivankova VW, Richmond RC, Woolf B, Yarmolinsky J, Davies NM, Swanson SA, VanderWeele TJ, Higgins J, Timpson NJ, Dimou N et al: Strengthening the Reporting of Observational Studies in Epidemiology Using Mendelian Randomization: The STROBE-MR Statement. JAMA-J AM MED ASSOC 2021, 326(16):1614-1621.

[24] Sudlow C, Gallacher J, Allen N, Beral V, Burton P, Danesh J, Downey P, Elliott P, Green J, Landray M et al: UK biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age. PLOS MED 2015, 12(3):e1001779.

[25] Zengini E, Hatzikotoulas K, Tachmazidou I, Steinberg J, Hartwig FP, Southam L, Hackinger S, Boer CG, Styrkarsdottir U, Gilly A et al: Genome-wide analyses using UK Biobank data provide insights into the genetic architecture of osteoarthritis. NAT GENET 2018, 50(4):549-558.

[26] Ha E, Bae SC, Kim K: Large-scale meta-analysis across East Asian and European populations updated genetic architecture and variant-driven biology of rheumatoid arthritis, identifying 11 novel susceptibility loci. ANN RHEUM DIS 2021, 80(5):558-565.

[27] Astle WJ, Elding H, Jiang T, Allen D, Ruklisa D, Mann AL, Mead D, Bouman H, Riveros-Mckay F, Kostadima MA et al: The Allelic Landscape of Human Blood Cell Trait Variation and Links to Common Complex Disease. CELL 2016, 167(5):1415-1429.

[28] Ligthart S, Vaez A, Vosa U, Stathopoulou MG, de Vries PS, Prins BP, Van der Most PJ, Tanaka T, Naderi E, Rose LM et al: Genome Analyses of >200,000 Individuals Identify 58 Loci for Chronic Inflammation and Highlight Pathways that Link Inflammation and Complex Disorders. AM J HUM GENET 2018, 103(5):691-706.

[29] Hartwig FP, Borges MC, Horta BL, Bowden J, Davey SG: Inflammatory Biomarkers and Risk of Schizophrenia: A 2-Sample Mendelian Randomization Study. JAMA PSYCHIAT 2017, 74(12):1226-1233.

[30] Gao S, Zhou H, Luo S, Cai X, Ye F, He Q, Huang C, Zheng X, Li Y, Du Z et al: Investigating the Causal Relationship Between Physical Activity and Chronic Back Pain: A Bidirectional Two-Sample Mendelian Randomization Study. FRONT GENET 2021, 12:758639.

[31] Machiela MJ, Chanock SJ: LDlink: a web-based application for exploring population-specific haplotype structure and linking correlated alleles of possible functional variants. BIOINFORMATICS 2015, 31(21):3555-3557.

[32] Wu BS, Chen SF, Huang SY, Ou YN, Deng YT, Chen SD, Dong Q, Yu JT: Identifying causal genes for stroke via integrating the proteome and transcriptome from brain and blood. J TRANSL MED 2022, 20(1):181.

[33] Burgess S, Bowden J, Fall T, Ingelsson E, Thompson SG: Sensitivity Analyses for Robust Causal Inference from Mendelian Randomization Analyses with Multiple Genetic Variants. EPIDEMIOLOGY 2017, 28(1):30-42.

[34] Bowden J, Davey SG, Burgess S: Mendelian randomization with invalid instruments: effect estimation and bias detection through Egger regression. INT J EPIDEMIOL 2015, 44(2):512-525.

[35] Ahuja JKC, Moshfegh AJ, Holden JM, Harris E: USDA Food and Nutrient Databases Provide the Infrastructure for Food and Nutrition Research, Policy, and Practice. The Journal of Nutrition 2013, 143(2):241S-249S.

[36] Chang X, Dorajoo R, Sun Y, Han Y, Wang L, Khor CC, Sim X, Tai ES, Liu J, Yuan JM et al: Gene-diet interaction effects on BMI levels in the Singapore Chinese population. NUTR J 2018, 17(1):31.

[37] Romanova Y, Laikov A, Markelova M, Khadiullina R, Makseev A, Hasanova M, Rizvanov A, Khaiboullina S, Salafutdinov I: Proteomic Analysis of Human Serum from Patients with Chronic Kidney Disease. BIOMOLECULES 2020, 10(2).

[38] Rouland A, Masson D, Lagrost L, Verges B, Gautier T, Bouillet B: Role of apolipoprotein C1 in lipoprotein metabolism, atherosclerosis and diabetes: a systematic review. CARDIOVASC DIABETOL 2022, 21(1):272.

[39] Zhang R, Yao J, Xu P, Ji B, Luck JV, Chin B, Lu S, Kelsoe JR, Ma J: A comprehensive meta-analysis of association between genetic variants of GDF5 and osteoarthritis of the knee, hip and hand. INFLAMM RES 2015, 64(6):405-414.

[40] Takahata Y, Hagino H, Kimura A, Urushizaki M, Yamamoto S, Wakamori K, Murakami T, Hata K, Nishimura R: Regulatory Mechanisms of Prg4 and Gdf5 Expression in Articular Cartilage and Functions in Osteoarthritis. INT J MOL SCI 2022, 23(9).

[41] Ohta Y, Okabe T, Larmour C, Di Rocco A, Maijenburg MW, Phillips A, Speck NA, Wakitani S, Nakamura T, Yamada Y et al: Articular cartilage endurance and resistance to osteoarthritic changes require transcription factor Erg. ARTHRITIS RHEUMATOL 2015, 67(10):2679-2690.

[42] Qin X, Zhang P: ECRG4: a new potential target in precision medicine. FRONT MED-PRC 2019, 13(5):540-546.

[43] Steck E, Breit S, Breusch SJ, Axt M, Richter W: Enhanced expression of the human chitinase 3-like 2 gene (YKL-39) but not chitinase 3-like 1 gene (YKL-40) in osteoarthritic cartilage. BIOCHEM BIOPH RES CO 2002, 299(1):109-115.

[44] Dorschner RA, Lee J, Cohen O, Costantini T, Baird A, Eliceiri BP: ECRG4 regulates neutrophil recruitment and CD44 expression during the inflammatory response to injury. SCI ADV 2020, 6(11):eaay518.

[45] Pomatto LC, Raynes R, Davies KJ: The peroxisomal Lon protease LonP2 in aging and disease: functions and comparisons with mitochondrial Lon protease LonP1. BIOL REV 2017, 92(2):739-753.

[46] Hemani G, Bowden J, Davey SG: Evaluating the potential role of pleiotropy in Mendelian randomization studies. HUM MOL GENET 2018, 27(R2):R195-R208.