Genome-wide association studies and multi-omics integrative analysis reveal novel loci and their molecular mechanisms for circulating polyunsaturated, monounsaturated, and saturated fatty acids

Yitang Sun[1], Huifang Xu[1], Kaixiong Ye[1,2,*]

[1]Department of Genetics, University of Georgia, Athens, GA, USA
[2]Institute of Bioinformatics, University of Georgia, Athens, GA, USA

[*] Correspondence: Dr. Kaixiong Ye (Kaixiong.Ye@uga.edu)

Abstract

Previous genome-wide association studies (GWAS) have identified genetic loci associated with the circulating levels of FAs, but the biological mechanisms of these genetic associations remain largely unexplored. Here, we conducted GWAS to identify additional genetic loci for 19 circulating fatty acid (FA) traits in UK Biobank participants of European ancestry (N = 239,268) and five other ancestries (N = 508 – 4,663). We leveraged the GWAS findings to characterize genetic correlations and colocalized regions among FAs, explore sex differences, examine FA loci influenced by lipoprotein metabolism, and apply statistical fine-mapping to pinpoint putative causal variants. We integrated GWAS signals with multi-omics quantitative trait loci (QTL) to reveal intermediate molecular phenotypes mediating the associations between the genetic loci and FA levels. Altogether, we identified 215 significant loci for polyunsaturated fatty acids (PUFAs)-related traits in European participants, 163 loci for monounsaturated fatty acids (MUFAs)-related traits, and 119 loci for saturated fatty acids (SFAs)-related traits, including 70, 61, and 54 novel loci, respectively. A novel locus for total FAs, the percentage of omega-6 PUFAs in total FAs, and total MUFAs (around genes *GSTT1/2/2B*) overlapped with QTL signals for all six molecular phenotypes examined, including gene expression, protein abundance, DNA methylation, splicing, histone modification, and chromatin accessibility. Across 19 FA traits, 65% of GWAS loci overlapped with QTL signals for at least one molecular phenotype. Our study identifies novel genetic loci for circulating FA levels and systematically uncovers their underlying molecular mechanisms.

Keywords: GWAS, polyunsaturated fatty acids, fatty acids, xQTL, multi-omics

Introduction

Polyunsaturated fatty acids (PUFAs), monounsaturated fatty acids (MUFAs), and saturated fatty acids (SFAs) have been implicated in various diseases.[1-3] Evidence from epidemiological and genetic studies indicates that PUFAs, especially omega-3 PUFAs, are associated with reduced risks of various diseases or mortality by modifying metabolism and inflammation.[4-7] Circulating fatty acid (FA) levels are influenced by environmental and genetic factors and characterized by sex- and ancestry-specific patterns.[8-11] Dietary intake, socioeconomic status, physical inactivity, cigarette smoking, and alcohol consumption are common environmental factors that influence FA levels.[12] Although hundreds of genetic loci associated with FAs have been identified by genome-wide association studies (GWAS),[13-38] the biological mechanisms of these genetic associations remain largely unexplored. Moreover, these genetic loci were mainly discovered in the European (EUR) population and were focused on autosomes. Extending GWAS to non-EUR populations and sex chromosomes will offer a comprehensive understanding of the genetic architecture of circulating FAs.

In this study, we perform the GWAS for 19 FA traits related to PUFAs, MUFAs, and SFAs to identify more novel genetic loci. We conducted overall and sex-specific GWAS using 250,101 individuals of EUR ancestry and five other ancestries from the UK Biobank (UKB). Leveraging the GWAS summary statistics, we characterized heritability for these FA traits and examined their shared genetic basis with genetic correlation and colocalization analysis. Candidate causal variants were pinpointed with statistical fine-mapping. Finally, we integrated GWAS signals with multi-omics data to identify the mediating molecular phenotypes and provide mechanistic insights into the molecular mechanisms underlying the GWAS signals.

Methods

Participants

The UKB cohort is a prospective population-based study of ~500,000 participants from across the United Kingdom, aged between 37 and 73 years at recruitment from 2006 to 2010.[39] Of the participants with both phenotype and genotype data, we excluded those who have withdrawn consent, mismatched information between self-reported and genetic sex, sex chromosome aneuploidy, or are outliers for heterogeneity and missing genotype rate. Genetic ancestry groups have been previously defined in the Pan-ancestry genetic analysis of the UK Biobank (Pan-UKB).[40] Together, we included up to 239,268 participants of EUR ancestry in this study. We also included participants of African (AFR) (N = 3,352), admixed American (AMR) (N = 508), Central/South Asian (CSA) (N = 4,663), East Asian (EAS) (N = 1,445), and Middle Eastern (MID) (N = 865) ancestries. Table S1 provides the overall and sex-stratified characteristics of participants in each ancestry group. This study received ethical approval from the National Health Service North West Centre for Research Ethics Committee, and all participants provided informed consent. Data from the UKB resource were accessed under application number 48818.

Phenotypes – circulating FAs

Plasma FAs were measured in a randomly selected subset of 274,020 participants by Nightingale Health using nuclear magnetic resonance (NMR) spectroscopy.[41] All FA levels and their ratios were normalized within each ancestry group using the rank-based inverse normal transformation in overall GWAS.[42] This inverse normal transformation was applied separately for males and females to account for sex-specific differences in FA distributions. In total, 20 metabolic measurements were initially analyzed in the GWAS, including total FAs, 5 PUFA

absolute concentrations (i.e., total PUFAs, omega-3 PUFAs, omega-6 PUFAs, docosahexaenoic acid (DHA), and linoleic acid (LA)), their relative percentages in total FAs (i.e., PUFAs%, omega-3%, omega-6%, DHA%, and LA%), the ratio of omega-3 to omega-6 PUFAs (omega-3/omega-6), the ratio of omega-6 to omega-3 PUFAs (omega-6/omega-3), monounsaturated fatty acids (MUFAs), saturated fatty acids (SFAs), MUFAs%, SFAs%, PUFAs/MUFAs, PUFAs/SFAs, and degree of unsaturation. Both the omega-3/omega-6 ratio and omega-6/omega-3 ratio were included in our GWAS due to different preferences in the field, but they revealed the same genetic loci in opposite effect directions. In additional analyses, we focused on omega-3/omega-6 and did not include omega-6/omega-3, resulting in a total of 19 plasma metabolic measurements. FA traits were categorized into PUFA-related traits (total PUFAs, omega-3 PUFAs, omega-6 PUFAs, DHA, LA, PUFAs%, omega-3%, omega-6%, DHA%, LA%, omega-3/omega-6, PUFAs/MUFAs, PUFAs/SFAs, and degree of unsaturation), MUFA-related traits (total MUFAs, MUFAs%, and PUFAs/MUFAs), and SFA-related traits (total SFAs, SFAs%, and PUFAs/SFAs).

Genotyping and quality control

Full details about genotyping, imputation, and genotype-based quality control (QC) have been described elsewhere.[43] Briefly, we excluded variants with imputation quality score < 0.3, minor allele frequency < 0.001, genotype missingness per variant > 0.05, and Hardy-Weinberg test *P*-value < 1×10^{-8}. We also confirmed that no individuals have genotype missingness > 0.05. For the non-pseudoautosomal region of the X chromosome (GRCh37; chrX:2699520-154931044), males were treated as homozygotes of the reference or effect allele, and coded as 0 or 2, respectively.

Genome-wide association studies (GWAS)

Overall GWAS were conducted for each ancestry group and followed by sex-stratified GWAS. We carried out GWAS in participants of EUR ancestry using fastGWA from the GCTA toolbox (v.1.94.1), which controls for familial relatedness using a sparse genetic relationship matrix (GRM) with a default threshold of 0.05 based on slightly linkage disequilibrium (LD)-pruned HapMap3 variants.[44,45] Considering the relatively small sample size for other ancestry groups (AFR, AMR, CSA, EAS, and MID), the mixed-linear-model association (MLMA) method was used to perform GWAS using the full GRM.[46] Additionally, sex-specific GWAS were conducted for each ancestry group using fastGWA for EUR participants and MLMA for other ancestry groups, excluding sex as a covariate from all models.

To explore the influences of potential covariates, we performed GWAS in the EUR ancestry participants using three models. In model 1, we included age, sex, age × sex, and the top 10 genetic principal components derived from the Pan-UKB as covariates. Model 2 added additional covariates, including body mass index, Townsend deprivation index, smoking status, alcohol status, physical activity, and statin use.[47,48] In model 3, we adjusted for fish oil supplementation status and lipoprotein lipids including chylomicron, very-low-density lipoprotein (VLDL), intermediate-density lipoprotein (IDL), low-density lipoprotein (LDL), high-density lipoprotein (HDL) cholesterols as additional covariates. For GWAS in the non-EUR population, we applied models 1 and 2.

Identification of novel loci

To identify novel loci, we compared our independent loci to the previously reported FA loci in the GWAS Catalog and in other relevant publications identified by literature review up until June 2024.[49] Independent loci were identified in our results or previous GWAS based on genome-wide significance threshold (P-value $< 5 \times 10^{-8}$) and LD-based clumping ($r^2 = 0.1$,

window size = 250 kb) using PLINK.[50,51] Independent loci were merged based on proximity (±250 kb). We annotated the lead variants for their target genes and functional consequences using the Ensemble GRCh37 Variant Effect Predictor (VEP).[52]

Heritability and genetic correlation analyses

Heritability and genetic correlations were estimated with linkage disequilibrium score regression (LDSC) in the EUR ancestry subset.[53,54] In-sample LD scores were derived from s3://broad-alkesgroup-ukbb-ld/UKBB_LD/.[55] As recommended, we checked for inflation using the attenuation ratio in the EUR population and genomic inflation factor (λ_{GC}) in other populations.[56]

Colocalization with FA traits

We performed colocalization analysis with HyPrColoc (v.1.0) to assess whether two or more FA traits share a putative causal variant.[57] Pairwise and multi-trait colocalization analyses were conducted for each pair or all FA traits altogether. We used the default priors, including the probability of a variant being associated with a single trait (prior.1 = 1×10^{-4}) and the conditional probability of association with an additional trait (prior.c = 0.02). A posterior probability (PP) > 0.8 was considered as evidence for a shared causal variant.

Statistical fine-mapping of GWAS loci

To statistically fine-map the candidate causal variants for genome-wide significant loci associated with each FA trait, we used SuSiE (v.0.12.27).[58] The analysis was performed for the 1 Mb region surrounding the lead variant of each genetic locus. We utilized in-sample LD information to estimate the correlations among SNPs. The 95% credible set identified potential causal variants based on their PP. Only variants with a PP > 0.8 were reported, with higher PP values indicating more likely to be causal variants.

Gene-based and gene-set analyses

MAGMA v.1.08, implemented in FUMA, was used to perform gene-based and gene-set analyses.[59,60] In the gene-based analysis, a total of 19,665 protein-coding genes was tested, and Bonferroni correction was applied to establish the significance threshold ($P < 0.05/19{,}665 = 2.54 \times 10^{-6}$). We determined the enrichment of candidate FAs-associated genes in specific biological pathways, cellular components, or molecular functions. A total of 15,481 gene sets (5,497 curated gene sets and 9,984 gene ontology terms) obtained from MsigDB were tested, with the Bonferroni-corrected significance threshold set at $P < 3.23 \times 10^{-6}$.[61]

Omics PlEiotRopic Association (OPERA) analysis

To provide mechanistic interpretations of FAs-associated variants, we performed an integrative analysis of six types of multi-omics quantitative trait loci (QTL) summary statistics with GWAS signals using the Bayesian method, Omics PlEiotRopic Association (OPERA).[62] First, we obtained available molecular quantitative trait loci (xQTL) data, including gene expression QTL (eQTL), protein QTL (pQTL), DNA methylation QTL (mQTL), splicing QTL (sQTL), histone modification QTL (hQTL), and chromatin accessibility QTL (caQTL). We downloaded blood *cis*-eQTLs summary statistics generated from the eQTLGen study (N = 31,684).[63] The plasma pQTL summary statistics were retrieved from the UKB Pharma Proteomics Project (UKB-PPP) using UKB participants (N = 54,219).[64] The peripheral blood mQTL data were generated by McRae *et al.* (N = 1,980).[65,66] We derived summary-level sQTL data from whole blood samples in the Genotype Tissue Expression (GTEx) Project v.8 (N = 670).[67] The H3K27ac and H3K4me1 hQTL in monocytes were generated from the BLUEPRINT project (N = 200).[68] The caQTL data in lymphoblastoid cell line samples were estimated by Kumasaka *et al.* (N = 100).[69] We included the molecular phenotypes with at least one xQTL

with $P < 5 \times 10^{-8}$ and excluded the major histocompatibility complex (MHC) region due to its structural complexity. After filtering molecular phenotypes, we retained 15,785 genes, 2,190 proteins, 94,338 DNA methylation sites, 6,639 RNA splicing events, 18,152 histone marks, and 13,873 chromatin accessibility peaks. The posterior probability of association (PPA) threshold of 0.9 and the multi-exposure heterogeneity in dependent instruments (HEIDI) test threshold of 0.01 were used to determine if a GWAS signal overlaps with the QTL signals for one or multiple molecular phenotypes.

Results

Sample characteristics

The primary GWAS comprised up to 239,268 individuals of EUR ancestry. GWAS were also conducted on five additional ancestry groups, including up to 3,352 AFR, 508 AMR, 4,663 CSA, 1,445 EAS, and 865 MID individuals. Among EUR participants, the average age was 57 years and females were more likely to have higher PUFA levels (Table S1). The percentage of female participants in EUR, AFR, AMR, CSA, EAS, and MID populations was 53.9%, 59.7%, 66.9%, 44.8%, 64.6%, and 41.8%, respectively. Approximately 31.6% of individuals reported regular use of fish oil supplements and the ratio of omega-3/omega-6 was around 1:8.

Identification of genome-wide significant loci for circulating FA traits

We first performed EUR ancestry GWAS for each FA trait using three models, respectively. Overall, we identified between 37 and 124 loci for 19 FA traits in model 2, and between 26 and 48 loci in model 3 (Table S2). Approximately 13% of the loci (ranging from 0–22%) identified in model 2 were not significant in model 1 (Table S3). However, on average, 62% of the loci (range: 19–79%) from model 2 were not significant in model 3, and 31% (range: 18–43%) of the loci from model 3 were not identified in model 2 (Table S3). Figure 1A shows an example of the Manhattan plot of GWAS for the absolute concentration of total PUFAs in model 2. Of the 122 significant loci identified in model 2, 30 for total PUFAs were also significant in model 3 and a total of 92 loci identified in model 2 were no longer significant after adjusting for fish oil intake and lipoprotein lipids in model 3 (Figure 1B). We observed similar effect estimates between models 1 and 2, whereas model 3 yielded systematically smaller effect estimates than those from the other two models (Figure S1). To further pinpoint the covariates driving the distinction between models 3 and 2, we conducted GWAS based on model 2 by

additionally adjusting for fish oil supplementation in model 2.1 and for lipoprotein lipids in model 2.2. The differences in GWAS results between model 3 and the others were attributed to the additional adjustment for lipoprotein lipids (Table S2). For example, variants in the *LDLR* and *MTTP* genes, known for their strong correlations with lipoprotein lipids,[70,71] were significant in model 2 but not in model 3 for total PUFAs (Figure 1).

Regarding SNP-based heritability of FA traits in the EUR population, we found that SNPs explained 5–19% of the phenotypic variance for all FA traits across all three models (Figure S2A; Table S4). For instance, the SNP-based heritability estimates for the absolute concentrations of total PUFAs, MUFAs, and SFAs were 19%, 19%, and 16%, respectively, in model 2. The genomic inflation factor (λ_{GC}) for all GWAS ranged from 1.07 to 1.29 (Table S5). The attenuation ratio ranged from 0.03 to 0.18, and the LDSC intercept ranged from 1.01 to 1.12, suggesting that the genome-wide elevation of association statistics was primarily due to true additive polygenic effects rather than confounders such as population stratification.

To identify novel loci, we searched the literature and GWAS Catalog to compile known loci for FAs-related traits, making it the most comprehensive collection to date (Table S6). When categorizing FA traits into PUFAs-related, MUFAs-related, and SFAs-related, we identified 70, 61, and 54 novel loci among 215, 163, and 119 significant loci, respectively, with model 2 (Table S7). The numbers of novel loci for each FA trait are provided in Table S3. Compared with all previously reported FAs-related loci, the top 10 strongest novel loci identified across all FA traits in model 2 were mapped to candidate genes including *PEPD*, *SBNO1*, *IL1RN*, *INSIG2*, *VIM*, *PGS1*, *ARID5B*, *NF1*, *FAM96A*, and *PEX6* (Table S8). We also compared independent loci to those previously reported in GWAS for the same traits to identify loci that have not been previously associated with the corresponding trait. For example, a novel locus around *VEGFA*

was identified for total FAs, LA%, total SFAs, PUFAs/SFAs, and the degree of unsaturation in our study, while it was consistently associated with total MUFAs, MUFAs%, and PUFAs/MUFAs in both our and previous GWAS.[30,33,34,38] Variants in *GCKR* were found to be significantly associated with DHA% for the first time, while they had been associated with other FA traits (Table S8).

Additionally, we explored genetic associations on the X-chromosome in both overall and sex-specific GWAS. Previously, only one GWAS identified FA loci on the X-chromosome.[30] That GWAS reported two X-chromosomal loci associated with FA traits; however, no significant variants were identified in our overall GWAS. When evaluating males separately, one significant association with omega-3% was found on the X-chromosome.

Sex-specific FA GWAS

We then stratified the total sample by sex to perform sex-specific GWAS separately for 128,922 females and 110,346 males in the EUR population. Approximately 4.2% of the significant loci from sex-specific GWAS were not identified in the overall GWAS (Table S3). The SNP-based heritability ranged from 5% to 21% in females and 4% to 20% in males, and there was no evidence of genomic inflation (Figures S2B and S2C; Table S4). In model 2, we identified 128 independent loci associated with PUFAs-related traits in females and 106 loci in males, including 27 and 8 novel loci, respectively, which have not been reported in previous GWAS of PUFAs-related traits (Table S7). GWAS results for both known and novel loci associated with each FA trait are provided in Table S9 for females and Table S10 for males. Some novel loci were only identified in one sex but not in the other. For example, in the model 2 GWAS for total PUFAs, there were three novel loci in females but not in males, when compared to previous GWAS of all FA traits. Two of them were mapped to genes *RNU6-1180P* and

KRT18P55, while the third locus, located on Chromosome 10, was not mapped to any genes (Figure 2). Similarly, in the model 3 GWAS for total PUFAs, there were three novel loci, around genes *UNC5CL*, *RP11-328J6.1*, and *SIPA1L3*, which were identified only in females (Figure S3). No novel loci were identified in male-specific GWAS for total PUFAs in either model 2 or 3. In the sex-specific GWAS, we found that approximately 43% of the genome-wide significant loci identified in models 1 and 2 were not identified in model 3 (Table S3). On the other hand, about 31% of the loci identified in model 3 were not identified in the other two models (Table S3). Notably, one X-chromosomal locus associated with omega-3%, whose lead SNP is rs147828433, was identified from GWAS using model 2 in males. Our study is the first to explore loci in sex-specific GWAS of FA traits and to include the X chromosome.

GWAS for FA traits in individuals of non-EUR ancestries

Across all FA traits, we identified 38 significant loci in ancestry-specific GWAS, of which 23 are novel compared to previous studies in EUR and non-EUR populations (Table S11). The genomic inflation factor (λ_{GC}) was also reported for each non-EUR ancestry, with average λ_{GC} values of 0.97, 0.92, 0.99, 0.90, and 0.95 for AFR, AMR, CSA, EAS, and MID, respectively (Tables S5). In the non-EUR ancestry GWAS, the most significant locus was around *FADS2*, identified in four ancestry groups except AFR, which is a well-known locus in the EUR population (Table S11). Among non-EUR populations, CSA had the largest sample size, leading to more loci being identified compared to other groups, with 1–4 loci in model 1 and 0–2 in model 2 (Table S2). The GWAS in CSA male participants identified a novel locus *EDN1* associated with total PUFAs, omega-6 PUFAs, and LA, while this locus did not reach the genome-wide significance in the EUR population. In addition, variants in *VPS39*, which had not

been reported in previous GWAS, were associated with total PUFAs, omega-6 PUFAs, and LA in both total and male-only CSA participants.

Genetic correlations among FA traits

We examined genetic correlations (r_g) between FA traits using EUR GWAS summary statistics from three models (Table S12). Broadly, comparisons of genetic correlations between models revealed highly consistent estimates between models 1 and 2 but more distinct estimates in model 3 (Figure S4). Among the 171 pairs of genetic correlations, 14 pairs that had significantly negative genetic correlations in models 1 and 2 became significantly positive in model 3, while four pairs that were significantly positive in models 1 and 2 became significantly negative in model 3. The strongest correlation was a negative correlation between total MUFAs and PUFAs/MUFAs (r_g = -0.99) observed based on GWAS in both models 1 and 2 (Figures S5 and S6). Notably, genetic correlations between omega-6 PUFAs and omega-6% were negative in models 1 (r_g = -0.38) and 2 (r_g = -0.51), but they were positively correlated in model 3 (r_g = 0.34) after further adjustment for lipoprotein lipids (Figure S7). Within omega-3-related traits, genetic correlations were strong (~0.9) across all three models, with even stronger correlations found in model 3. We also estimated pairwise phenotypic correlations using Pearson's correlation coefficient (Table S13). Among significant genetic correlations, all pairs had consistent directions compared to phenotypic correlations in models 1 and 2, while 83% of pairs had consistent directions in model 3, suggesting that genetic effects on FA traits were altered after adding lipoprotein lipids as covariates.

Colocalization of GWAS signals across FA traits

To assess the probability that two FA traits share the same causal variants, we conducted pairwise colocalization analyses. For the 171 pairs of FA traits, we identified 9–231 colocalized

signals in model 2 and 6–114 signals in model 3, respectively. We found that pairs with a larger number of colocalized signals also had stronger genetic correlations (Figure S8). In model 2, potential causal variants in *DOCK7* (rs2934744), *GCKR* (rs1260326), *ZNF259* (rs964184), *FADS1* (rs174564), *CPT1A* (rs2229738), *LIPC* (rs2070895), *LIPG* (rs77960347), *LDLR* (rs73015024), *TM6SF2* (rs58542926), and *APOE* (rs7412) showed robust evidence of colocalization, indicating shared genetic variants for pairs of FA traits (Table S14). After accounting for lipoprotein lipids in model 3, *FADS1* (rs174564), *CPT1A* (rs2229738), and *BUD13* (chr11:116623213:TA:T) were identified as colocalized for FA traits, consistent with results from model 2 (Table S15). Next, we performed multi-trait colocalization analyses to evaluate the posterior probability of a shared genetic signal across all FA traits. Overall, 168 and 159 colocalized signals were identified in models 2 and 3, respectively. We identified *GCKR* (rs1260326), *TRIB1* (rs28601761), *FADS1* (rs174564), *ZNF259* (rs964184), and *APOE* (rs7412) as shared among multiple FA traits in model 2 (Table S16), and *GCKR* (rs1260326), *TRIB1* (rs28601761), *FADS1* (rs174564), and *APOE* (rs1065853) from model 3 as shared signals (Table S17).

Statistical fine-mapping for candidate causal variants

We applied the Bayesian fine-mapping method to identify putative causal variants for FA traits. We identified 477 unique potential causal variants in model 1, 422 in model 2, and 193 in model 3 (Table S18). Consistent with colocalization analysis, we confirmed that the genetic variants at the loci *TRIB1* (rs28601761), *FADS1* (rs174564), *CPT1A* (rs2229738), *LIPC* (rs2070895), *LIPG* (rs77960347), *LDLR* (rs72658867), *TM6SF2* (rs58542926), and *APOE* (rs7412) are likely to be causally associated with FA traits. Notably, we found that several loci contain multiple 95% credible sets, indicating the existence of multiple potential causal variants

within those significant loci. Among the novel loci for FA traits, we found that *ZMIZ1* (rs1782652) was associated with total FAs, total PUFAs, and omega-3 PUFAs, suggesting that it is a candidate shared causal variant for these FAs. Additional potential causal variants for novel loci were identified, including *CPS1* (rs1047891), *ATXN7L1* (rs118061830), *ARID5B* (rs77044968), *SBF2* (rs12789941), *DGKZ* (rs149903077), *CYFIP1* (rs199854211), *FAM96A* (rs62023393), *MIR122* (rs41292412), *BCL2* (rs12454712), *INSR* (rs112630404), and *PEPD* (rs62102718).

Gene-based and gene-set enrichment analyses

In the genome-wide gene-based association study, 139–541 genes from model 2 and 110–255 genes from model 3 were found to be significantly associated with 19 FA traits after Bonferroni correction (Table S19). The top gene across 19 FA traits identified in model 2 was *ALDH1A2* (chr15:58245622-58790065), which overlaps with another identified gene *LIPC* (chr15:58702768-58861151). In the GWAS of the omega-3/omega-6 ratio using model 3, *STH* was the top gene, which might be related to neurodegenerative diseases such as Alzheimer's disease.[72] Gene-tissue expression analysis revealed that the liver was the most significantly enriched tissue, followed by whole blood, kidney, small intestine, spleen, nerve, and adipose tissue (Tables S20 and 21). Additionally, gene-set enrichment analyses were performed to investigate potential gene sets and pathways with enriched GWAS signals (Tables S22 and 23). We found that genes associated with total PUFAs tend to be enriched in biological processes related to lipid homeostasis, secondary alcohol metabolism, organic hydroxy compound transport, regulations of plasma lipoprotein particle levels and lipid metabolism, and fatty acid biosynthesis (Figure 3). Fatty acid metabolic and biosynthetic processes were significantly enriched in the gene-set analysis results from GWAS of total PUFAs using model 3 (Figure S9).

OPERA identifies intermediate molecular phenotypes underlying FAs-associated loci

To better understand the biological mechanisms of our GWAS findings, we applied OPERA to explore the overlap of GWAS signals for FA traits and QTL signals for six molecular phenotypes, including eQTL, pQTL, mQTL, sQTL, hQTL, and caQTL. With a PPA threshold of 0.9 and a P_{HEIDI} threshold of 0.01, we identified 976 FAs-associated variants that overlapped with QTL signals for 2,106 unique molecular phenotype measures, including the expression levels of 242 genes, the abundance levels of 60 proteins, the methylation levels at 1,254 DNA sites, the RNA splicing events at 141 sites, the status of 233 histone marks, and the openness at 176 chromatin accessibility peaks (Figure 4A; Table S24). After adjusting for lipoprotein lipids in model 3, we found 481 variants overlapping with xQTLs (Figure 4B; Table S25). Associations with DNA methylation were more frequent than with other molecular phenotypes, and these findings were consistent with previous results in other complex traits.[62] The number of loci associated with eQTL across 19 traits ranged from 7 for the degree of unsaturation to 68 for total PUFAs in model 2, and from 2 to 41 in model 3. Using the pQTL data from UKB-PPP, we observed that omega-3 PUFAs had the highest number of significant associations with proteins, identifying 35 loci in model 2 and 36 loci in model 3.

We found that approximately 65% of GWAS loci for 19 FA traits were shared with at least one molecular phenotype. Across FA traits such as total FAs, total PUFAs, omega-6 PUFAs, and total SFAs, the potential causal variant rs72658867, annotated at the *LDLR* gene, was associated with histone modification (chr19:11105519-11214483) and chromatin accessibility (peak 264429). Among the novel loci for omega-6%, rs149903077 (*DGKZ*) was jointly associated with the expression of nearby gene *NR1H3* (ENSG00000025434), histone

modification (chr11:47234446-47272168), and chromatin accessibility (peak 200213). It was also marginally associated with methylation (cg01183595).

Notably, all six molecular phenotypes had colocalized QTL signals with omega-6%, total FAs, and total MUFAs at the novel locus *GSTTP1* or nearby genes on chromosome 22 (Figure 5). We observed that the gene *EML3* was associated with the omega-3/omega-6 ratio jointly with five other types of molecular phenotypes, except for histone modification (Figure S10). The nearby variants rs12786457 and rs184864731 were consistently associated with these five phenotypes across GWAS related to omega-3 PUFAs and omega-3%. Overall, OPERA integrates six types of xQTL data to offer mechanistic insights into genetic loci and their downstream molecular phenotypes.

Discussion

We performed overall and sex-specific GWAS for 19 FA traits in 250,101 UKB participants of EUR descent and five other ancestry groups, separately. We identified 215, 163, and 119 significant loci for PUFAs-, MUFAs-, and SFAs-related traits, including 70, 61, and 54 novel loci, respectively, in our primary model (i.e., model 2). Further adjustment for lipoprotein lipids in model 3 resulted in significant changes in GWAS signals. On average, 62% of the loci (range: 19–79%) from model 2 were not significant in model 3, while 31% (range: 18–43%) of the loci from model 3 were not identified in model 2. Our analyses of genetic correlations and colocalization revealed the levels of shared genetic basis and specific candidate shared variants across FA traits. Using statistical fine-mapping, we identified 422 putative causal variants across all FA traits. Gene-based and gene-set enrichment analyses pinpointed the liver as the most relevant tissue and highlighted biological pathways underlying FA loci. In addition, we integrated GWAS signals with QTL signals for six molecular phenotypes and revealed that approximately 65% of GWAS loci for 19 FA traits were shared with at least one molecular phenotype, offering novel mechanistic insights.

In EUR GWAS, we analyzed 19 plasma NMR measures of FAs in UKB 239,268 participants. Although previous GWAS of FA traits in the UKB reported many loci, we identified approximately 90% more loci with a doubled sample size.[30,33,34] In addition, we compared our results with all previous GWAS of FAs-related traits to identify novel loci, including the second-largest published GWAS to date, which included 136,016 participants from 33 cohorts.[38] We calculated the omega-3/omega-6 and PUFA/SFA ratios for GWAS for the first time. Although the omega-6 to omega-3 ratio is commonly used in studies, the omega-3/omega-6 ratio may better capture the benefits of omega-3. This is particularly relevant given the ongoing

382 debate over the potential harmful or beneficial effects of omega-6 PUFAs.[73] Our large GWAS of
383 FAs showed that the phenotypic variance explained by all SNPs ranged from 5–19%, while
384 heritability estimates from twin studies were approximately 25–62%.[15] These results suggest that
385 future studies with rare or structural variants from sequencing and more diverse ancestry groups
386 could be valuable for identifying additional loci for FA traits.

387 Integrating comprehensive multi-omics data facilitates understanding the mechanisms
388 behind genetic loci and their downstream molecular phenotypes. Our study shows over half of
389 GWAS loci (~65%) for 19 FA traits are associated with at least one molecular phenotype. For
390 example, the *LDLR* gene variant rs72658867 is linked to histone modification
391 (19:11105519:11214483) and chromatin accessibility (264429), impacting lipid metabolism and
392 cardiovascular health.[74] The variant rs149903077 in the novel *DGKZ* locus has been associated
393 with the expression of nearby gene *NR1H3*, chromatin accessibility, histone modifications, and
394 DNA methylation, highlighting their combined roles in nuclear receptor regulation, lipid
395 metabolism, and immune responses.[75] The novel *GSTTP1* locus from the glutathione *S*-
396 transferase (GST) family and nearby genes on chromosome 22 were identified for omega-6%,
397 total FAs, and total MUFAs, overlapping with all six molecular phenotypes and have been linked
398 to key biological processes such as detoxification, cancer susceptibility, and cellular responses to
399 oxidative stress.[76] These genes in the GST family have previously been reported to be associated
400 with prostate cancer risk, and a recent study found evidence of a nonlinear relationship between
401 omega-6% and prostate cancer.[77-80] Future studies are needed to validate our findings on the
402 molecular mechanisms underlying FA-relevant variants. These insights could aid in patient
403 stratification for precision nutrition and in identifying novel therapeutic targets related to FAs.

Gene-set enrichment analysis findings highlight biological pathways underlying FA loci, implicating lipid homeostasis, transport, and metabolism, and providing genetic support for the current understanding of fatty acid regulation.[81] Our previous study suggested that lower omega-6% may reduce alcohol consumption, with the secondary alcohol metabolism pathway potentially explaining this effect on alcohol-related behaviors.[6] In model 3, which adjusts for lipoprotein lipid levels, fatty acid metabolic and biosynthetic processes were significantly more enriched in the gene-set analysis of GWAS of total PUFAs. The investigation of the disparities between models 2 and 3 suggests that lipoprotein-related biomarkers play a critical role in FA loci. Future studies are needed to thoroughly dissect these roles, and further analyses, such as Mendelian randomization, should properly infer results using genetic variants with or without adjustment for lipoprotein-related biomarkers.[82]

We investigated the sex-specific genetic architecture of FA traits. A novel locus, *KRT18P55*, was identified in the GWAS for females, and previous research indicates that this locus is highly expressed in patients with gastric cancer.[83] Notably, only one X-chromosomal locus was identified in the male-only GWAS for omega-3%, tagged by the lead variants rs147828433 (closest protein-coding gene: *ACOT9*). A study on *ACOT9* in mice suggests that it regulates both fatty acid and amino acid metabolism in mitochondria.[84] For the female-only GWAS, the heritability was higher, and larger numbers of loci were identified. Our results highlight the importance of tracking sex differences in genetics and may motivate future studies on gene-sex interactions. We found higher plasma PUFA levels in females, consistent with previous studies, which were influenced by genetic effects, sex hormones, and conversion rates in FA metabolism.[9] Randomized controlled trials show sex- and race-specific differences in the

benefits of fish oil supplementation and omega-3 PUFAs for preventing cardiovascular events and cognitive decline.[85]

By performing ancestry-specific and sex-specific GWAS in non-EUR samples, we identified additional novel loci, suggesting the presence of ancestry-specific genetic loci for FA traits. In male CSA participants, the novel locus *EDN1*, associated with omega-6 PUFA traits, encodes the protein endothelin 1, which has been linked to cardiovascular events and prognosis.[86] Variants in *VPS39*, associated with omega-6 PUFAs in the CSA population, have been previously identified as important regulators of myoblast differentiation and muscle glucose uptake in patients with type 2 diabetes.[87] Besides these ancestry-specific signals, we also found that several loci identified from large EUR GWAS were transferable to other ancestry groups, including *FADS2*, *ZNF259*, *MYRF*, *APOA5*, *GCKR*, *APOE*, and other loci. We found significantly fewer loci in the other five non-EUR ancestry GWAS due to small sample sizes. These results suggest that future studies with larger and more diverse ancestry samples are needed to identify more loci and confirm their consistency across populations.

To the best of our knowledge, our study is the most comprehensive GWAS on FA traits in a very large cohort. First, this is the first study to integrate six types of xQTL data with FA GWAS to explore the possible molecular mechanisms of FAs-associated genetic variants. Second, we used a large number of participants of EUR ancestry with plasma NMR measures for 19 FA traits and included participants from five other ancestry groups. Third, the inclusion of the omega-3/omega-6 ratio in the GWAS provides better insights into the benefits of omega-3 PUFAs. Fourth, our sex-specific GWAS on EUR and non-EUR populations, as well as the inclusion of X-chromosomal, uncover sex differences in FA genetic basis. Fifth, we comprehensively compared our findings to all previous GWAS of FAs-related traits. Detailed

information on all novel and known loci has been provided (Tables S3 and S6), which offers a valuable resource for the FA research community. Finally, we adjusted for various covariates across three models, with a particular focus on genetic architecture adjusted for lipoprotein metabolism.

The study has limitations. The numbers of participants in non-EUR populations were small, and future studies with larger sample sizes are necessary. While we explored the molecular mechanisms of FA loci, we are eager to extend OPERA analysis to other tissue-specific molecular phenotypes, particularly the liver, and even specific cell types, to identify cell type-specific disease mechanisms.

In conclusion, our findings reveal novel genetic loci and provide the first evidence of molecular mechanisms underlying FAs-relevant variants. These insights could aid population stratification for precision nutrition and the identification of novel therapeutic targets for FAs-related diseases.

Data availability

The full summary statistics for GWAS of fatty acids are publicly available in the GWAS Catalog database (https://www.ebi.ac.uk/gwas/, GCP ID: TBA after peer review). The blood eQTL data of eQTLGen Consortium used in the analyses were downloaded from https://molgenis26.gcc.rug.nl/downloads/eqtlgen/cis-eqtl/SMR_formatted/cis-eQTL-https://zenodo.org/records/7951839/files/LBC_BSGS_meta.tar.gz?download=1SMR_20191212.tar.gz. UKBB pQTL data are available at https://metabolomips.org/ukbbpgwas/. The mQTL data are available for download from https://zenodo.org/records/7951839/files/LBC_BSGS_meta.tar.gz?download=1. Blood sQTL data can be downloaded at https://yanglab.westlake.edu.cn/data/SMR/GTEx_V8_cis_sqtl_summary/sQTL_Whole_Blood.zip. The hQTL data can be assessed from http://ftp.ebi.ac.uk/pub/databases/blueprint/blueprint_Epivar/qtl_as/QTL_RESULTS/. Summary statistics of caQTL can be found at https://zenodo.org/records/1405945/files/.

Supplemental information

Supplemental information can be found online at https://doi.org/.

Acknowledgments

This research was approved by the University of Georgia Institutional Review Board and the UK Biobank consortium (application no. 48818). This work was funded by the University of Georgia Research Foundation and by the National Institute of General Medical Sciences of the National Institutes of Health under Award Number R35GM143060. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Author contributions

Y.S. and K.Y. designed the study. Y.S. and H.X. performed analyses. Y.S. and K.Y. drafted the manuscript. All authors contributed to the review and critical revision of the manuscript.

Declaration of interests

The authors declare no competing interests.

References

1. Worthmann, A., Ridder, J., Piel, S.Y.L., Evangelakos, I., Musfeldt, M., Voss, H., O'Farrell, M., Fischer, A.W., Adak, S., Sundd, M., et al. (2024). Fatty acid synthesis suppresses dietary polyunsaturated fatty acid use. Nat Commun *15*, 45. 10.1038/s41467-023-44364-y.
2. Baum, S.J., Kris-Etherton, P.M., Willett, W.C., Lichtenstein, A.H., Rudel, L.L., Maki, K.C., Whelan, J., Ramsden, C.E., and Block, R.C. (2012). Fatty acids in cardiovascular health and disease: a comprehensive update. J Clin Lipidol *6*, 216-234. 10.1016/j.jacl.2012.04.077.
3. Astrup, A., Magkos, F., Bier, D.M., Brenna, J.T., de Oliveira Otto, M.C., Hill, J.O., King, J.C., Mente, A., Ordovas, J.M., Volek, J.S., et al. (2020). Saturated Fats and Health: A Reassessment and Proposal for Food-Based Recommendations: JACC State-of-the-Art Review. J Am Coll Cardiol *76*, 844-857. 10.1016/j.jacc.2020.05.077.
4. Scorletti, E., and Byrne, C.D. (2013). Omega-3 fatty acids, hepatic lipid metabolism, and nonalcoholic fatty liver disease. Annu Rev Nutr *33*, 231-248. 10.1146/annurev-nutr-071812-161230.
5. Zhang, Y., Sun, Y., Yu, Q., Song, S., Brenna, J.T., Shen, Y., and Ye, K. (2024). Higher ratio of plasma omega-6/omega-3 fatty acids is associated with greater risk of all-cause, cancer, and cardiovascular mortality: A population-based cohort study in UK Biobank. Elife *12*. 10.7554/eLife.90132.
6. Xu, H., Sun, Y., Francis, M., Cheng, C.F., Modulla, N.T.R., Brenna, J.T., Chiang, C.W.K., and Ye, K. (2023). Shared genetic basis informs the roles of polyunsaturated fatty acids in brain disorders. medRxiv. 10.1101/2023.10.03.23296500.
7. Sun, Y., Chatterjee, R., Ronanki, A., and Ye, K. (2022). Circulating Polyunsaturated Fatty Acids and COVID-19: A Prospective Cohort Study and Mendelian Randomization Analysis. Front Med (Lausanne) *9*, 923746. 10.3389/fmed.2022.923746.
8. Hodson, L., Skeaff, C.M., and Fielding, B.A. (2008). Fatty acid composition of adipose tissue and blood in humans and its use as a biomarker of dietary intake. Prog Lipid Res *47*, 348-380. 10.1016/j.plipres.2008.03.003.
9. Decsi, T., and Kennedy, K. (2011). Sex-specific differences in essential fatty acid metabolism. Am J Clin Nutr *94*, 1914S-1919S. 10.3945/ajcn.110.000893.
10. Mathias, R.A., Pani, V., and Chilton, F.H. (2014). Genetic Variants in the FADS Gene: Implications for Dietary Recommendations for Fatty Acid Intake. Curr Nutr Rep *3*, 139-148. 10.1007/s13668-014-0079-1.
11. Francis, M., Li, C., Sun, Y., Zhou, J., Li, X., Brenna, J.T., and Ye, K. (2021). Genome-wide association study of fish oil supplementation on lipid traits in 81,246 individuals reveals new gene-diet interaction loci. PLoS Genet *17*, e1009431. 10.1371/journal.pgen.1009431.
12. Zhuang, P., Liu, X., Li, Y., Wu, Y., Li, H., Wan, X., Zhang, L., Xu, C., Jiao, J., and Zhang, Y. (2023). Circulating fatty acids, genetic risk, and incident coronary artery disease: A prospective, longitudinal cohort study. Sci Adv *9*, eadf9037. 10.1126/sciadv.adf9037.
13. Tanaka, T., Shen, J., Abecasis, G.R., Kisialiou, A., Ordovas, J.M., Guralnik, J.M., Singleton, A., Bandinelli, S., Cherubini, A., Arnett, D., et al. (2009). Genome-wide association study

of plasma polyunsaturated fatty acids in the InCHIANTI Study. PLoS Genet *5*, e1000338. 10.1371/journal.pgen.1000338.

14. Lemaitre, R.N., Tanaka, T., Tang, W., Manichaikul, A., Foy, M., Kabagambe, E.K., Nettleton, J.A., King, I.B., Weng, L.C., Bhattacharya, S., et al. (2011). Genetic loci associated with plasma phospholipid n-3 fatty acids: a meta-analysis of genome-wide association studies from the CHARGE Consortium. PLoS Genet *7*, e1002193. 10.1371/journal.pgen.1002193.

15. Kettunen, J., Tukiainen, T., Sarin, A.P., Ortega-Alonso, A., Tikkanen, E., Lyytikainen, L.P., Kangas, A.J., Soininen, P., Wurtz, P., Silander, K., et al. (2012). Genome-wide association study identifies multiple loci influencing human serum metabolite levels. Nat Genet *44*, 269-276. 10.1038/ng.1073.

16. Wu, J.H., Lemaitre, R.N., Manichaikul, A., Guan, W., Tanaka, T., Foy, M., Kabagambe, E.K., Djousse, L., Siscovick, D., Fretts, A.M., et al. (2013). Genome-wide association study identifies novel loci associated with concentrations of four plasma phospholipid fatty acids in the de novo lipogenesis pathway: results from the Cohorts for Heart and Aging Research in Genomic Epidemiology (CHARGE) consortium. Circ Cardiovasc Genet *6*, 171-183. 10.1161/CIRCGENETICS.112.964619.

17. Guan, W., Steffen, B.T., Lemaitre, R.N., Wu, J.H.Y., Tanaka, T., Manichaikul, A., Foy, M., Rich, S.S., Wang, L., Nettleton, J.A., et al. (2014). Genome-wide association study of plasma N6 polyunsaturated fatty acids within the cohorts for heart and aging research in genomic epidemiology consortium. Circ Cardiovasc Genet *7*, 321-331. 10.1161/CIRCGENETICS.113.000208.

18. Lemaitre, R.N., King, I.B., Kabagambe, E.K., Wu, J.H., McKnight, B., Manichaikul, A., Guan, W., Sun, Q., Chasman, D.I., Foy, M., et al. (2015). Genetic loci associated with circulating levels of very long-chain saturated fatty acids. J Lipid Res *56*, 176-184. 10.1194/jlr.M052456.

19. Tintle, N.L., Pottala, J.V., Lacey, S., Ramachandran, V., Westra, J., Rogers, A., Clark, J., Olthoff, B., Larson, M., Harris, W., and Shearer, G.C. (2015). A genome-wide association study of saturated, mono- and polyunsaturated red blood cell fatty acids in the Framingham Heart Offspring Study. Prostaglandins Leukot Essent Fatty Acids *94*, 65-72. 10.1016/j.plefa.2014.11.007.

20. Dorajoo, R., Sun, Y., Han, Y., Ke, T., Burger, A., Chang, X., Low, H.Q., Guan, W., Lemaitre, R.N., Khor, C.C., et al. (2015). A genome-wide association study of n-3 and n-6 plasma fatty acids in a Singaporean Chinese population. Genes Nutr *10*, 53. 10.1007/s12263-015-0502-2.

21. Hu, Y., Li, H., Lu, L., Manichaikul, A., Zhu, J., Chen, Y.D., Sun, L., Liang, S., Siscovick, D.S., Steffen, L.M., et al. (2016). Genome-wide meta-analyses identify novel loci associated with n-3 and n-6 polyunsaturated fatty acid levels in Chinese and European-ancestry populations. Hum Mol Genet *25*, 1215-1224. 10.1093/hmg/ddw002.

22. Kettunen, J., Demirkan, A., Wurtz, P., Draisma, H.H., Haller, T., Rawal, R., Vaarhorst, A., Kangas, A.J., Lyytikainen, L.P., Pirinen, M., et al. (2016). Genome-wide study for circulating metabolites identifies 62 loci and reveals novel systemic effects of LPA. Nat Commun *7*, 11122. 10.1038/ncomms11122.

23. Kalsbeek, A., Veenstra, J., Westra, J., Disselkoen, C., Koch, K., McKenzie, K.A., O'Bott, J., Vander Woude, J., Fischer, K., Shearer, G.C., et al. (2018). A genome-wide association study of red-blood cell fatty acids and ratios incorporating dietary covariates: Framingham Heart Study Offspring Cohort. PLoS One *13*, e0194882. 10.1371/journal.pone.0194882.
24. de Oliveira Otto, M.C., Lemaitre, R.N., Sun, Q., King, I.B., Wu, J.H.Y., Manichaikul, A., Rich, S.S., Tsai, M.Y., Chen, Y.D., Fornage, M., et al. (2018). Genome-wide association meta-analysis of circulating odd-numbered chain saturated fatty acids: Results from the CHARGE Consortium. PLoS One *13*, e0196951. 10.1371/journal.pone.0196951.
25. Marklund, M., Morris, A.P., Mahajan, A., Ingelsson, E., Lindgren, C.M., Lind, L., and Riserus, U. (2018). Genome-Wide Association Studies of Estimated Fatty Acid Desaturase Activity in Serum and Adipose Tissue in Elderly Individuals: Associations with Insulin Sensitivity. Nutrients *10*. 10.3390/nu10111791.
26. Locke, A.E., Steinberg, K.M., Chiang, C.W.K., Service, S.K., Havulinna, A.S., Stell, L., Pirinen, M., Abel, H.J., Chiang, C.C., Fulton, R.S., et al. (2019). Exome sequencing of Finnish isolates enhances rare-variant association power. Nature *572*, 323-328. 10.1038/s41586-019-1457-z.
27. Coltell, O., Sorli, J.V., Asensio, E.M., Barragan, R., Gonzalez, J.I., Gimenez-Alba, I.M., Zanon-Moreno, V., Estruch, R., Ramirez-Sabio, J.B., Pascual, E.C., et al. (2020). Genome-Wide Association Study for Serum Omega-3 and Omega-6 Polyunsaturated Fatty Acids: Exploratory Analysis of the Sex-Specific Effects and Dietary Modulation in Mediterranean Subjects with Metabolic Syndrome. Nutrients *12*. 10.3390/nu12020310.
28. Riveros-Mckay, F., Oliver-Williams, C., Karthikeyan, S., Walter, K., Kundu, K., Ouwehand, W.H., Roberts, D., Di Angelantonio, E., Soranzo, N., Danesh, J., et al. (2020). The influence of rare variants in circulating metabolic biomarkers. PLoS Genet *16*, e1008605. 10.1371/journal.pgen.1008605.
29. Li-Gao, R., Hughes, D.A., van Klinken, J.B., de Mutsert, R., Rosendaal, F.R., Mook-Kanamori, D.O., Timpson, N.J., and Willems van Dijk, K. (2021). Genetic Studies of Metabolomics Change After a Liquid Meal Illuminate Novel Pathways for Glucose and Lipid Metabolism. Diabetes *70*, 2932-2946. 10.2337/db21-0397.
30. Richardson, T.G., Leyden, G.M., Wang, Q., Bell, J.A., Elsworth, B., Davey Smith, G., and Holmes, M.V. (2022). Characterising metabolomic signatures of lipid-modifying therapies through drug target mendelian randomisation. PLoS Biol *20*, e3001547. 10.1371/journal.pbio.3001547.
31. Yin, X., Chan, L.S., Bose, D., Jackson, A.U., VandeHaar, P., Locke, A.E., Fuchsberger, C., Stringham, H.M., Welch, R., Yu, K., et al. (2022). Genome-wide association studies of metabolites in Finnish men identify disease-relevant loci. Nat Commun *13*, 1644. 10.1038/s41467-022-29143-5.
32. Montasser, M.E., Aslibekyan, S., Srinivasasainagendra, V., Tiwari, H.K., Patki, A., Bagheri, M., Kind, T., Barupal, D.K., Fan, S., Perry, J., et al. (2022). An Amish founder population reveals rare-population genetic determinants of the human lipidome. Commun Biol *5*, 334. 10.1038/s42003-022-03291-2.

33. Francis, M., Sun, Y., Xu, H., Brenna, J.T., and Ye, K. (2022). Fifty-one novel and replicated GWAS loci for polyunsaturated and monounsaturated fatty acids in 124,024 Europeans. medRxiv, 2022.2005.2027.22275343. 10.1101/2022.05.27.22275343.
34. Borges, M.C., Haycock, P.C., Zheng, J., Hemani, G., Holmes, M.V., Davey Smith, G., Hingorani, A.D., and Lawlor, D.A. (2022). Role of circulating polyunsaturated fatty acids on cardiovascular diseases risk: analysis using Mendelian randomization and fatty acid genetic association data from over 114,000 UK Biobank participants. BMC Med *20*, 210. 10.1186/s12916-022-02399-w.
35. Liu, C., Wang, Z., Hui, Q., Chiang, Y., Chen, J., Brijkumar, J., Edwards, J.A., Ordonez, C.E., Dudgeon, M.R., Sunpath, H., et al. (2022). Crosstalk between Host Genome and Metabolome among People with HIV in South Africa. Metabolites *12*. 10.3390/metabo12070624.
36. Davyson, E., Shen, X., Gadd, D.A., Bernabeu, E., Hillary, R.F., McCartney, D.L., Adams, M., Marioni, R., and McIntosh, A.M. (2023). Metabolomic Investigation of Major Depressive Disorder Identifies a Potentially Causal Association With Polyunsaturated Fatty Acids. Biol Psychiatry *94*, 630-639. 10.1016/j.biopsych.2023.01.027.
37. Fuller, H., Iles, M.M., Moore, J.B., and Zulyniak, M.A. (2023). Metabolic drivers of dysglycemia in pregnancy: ethnic-specific GWAS of 146 metabolites and 1-sample Mendelian randomization analyses in a UK multi-ethnic birth cohort. Front Endocrinol (Lausanne) *14*, 1157416. 10.3389/fendo.2023.1157416.
38. Karjalainen, M.K., Karthikeyan, S., Oliver-Williams, C., Sliz, E., Allara, E., Fung, W.T., Surendran, P., Zhang, W., Jousilahti, P., Kristiansson, K., et al. (2024). Genome-wide characterization of circulating metabolic biomarkers. Nature *628*, 130-138. 10.1038/s41586-024-07148-y.
39. Sudlow, C., Gallacher, J., Allen, N., Beral, V., Burton, P., Danesh, J., Downey, P., Elliott, P., Green, J., Landray, M., et al. (2015). UK biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age. PLoS Med *12*, e1001779. 10.1371/journal.pmed.1001779.
40. Pan-ancestry genetic analysis of the UK Biobank. https://pan.ukbb.broadinstitute.org/.
41. Julkunen, H., Cichonska, A., Slagboom, P.E., Wurtz, P., and Nightingale Health, U.K.B.I. (2021). Metabolic biomarker profiling for identification of susceptibility to severe pneumonia and COVID-19 in the general population. Elife *10*. 10.7554/eLife.63033.
42. McCaw, Z.R., Lane, J.M., Saxena, R., Redline, S., and Lin, X. (2020). Operating characteristics of the rank-based inverse normal transformation for quantitative trait analysis in genome-wide association studies. Biometrics *76*, 1262-1272. 10.1111/biom.13214.
43. Bycroft, C., Freeman, C., Petkova, D., Band, G., Elliott, L.T., Sharp, K., Motyer, A., Vukcevic, D., Delaneau, O., O'Connell, J., et al. (2018). The UK Biobank resource with deep phenotyping and genomic data. Nature *562*, 203-209. 10.1038/s41586-018-0579-z.
44. Jiang, L., Zheng, Z., Qi, T., Kemper, K.E., Wray, N.R., Visscher, P.M., and Yang, J. (2019). A resource-efficient tool for mixed model association analysis of large-scale data. Nat Genet *51*, 1749-1755. 10.1038/s41588-019-0530-8.
45. International HapMap, C., Altshuler, D.M., Gibbs, R.A., Peltonen, L., Altshuler, D.M., Gibbs, R.A., Peltonen, L., Dermitzakis, E., Schaffner, S.F., Yu, F., et al. (2010). Integrating

common and rare genetic variation in diverse human populations. Nature *467*, 52-58. 10.1038/nature09298.

46. Yang, J., Zaitlen, N.A., Goddard, M.E., Visscher, P.M., and Price, A.L. (2014). Advantages and pitfalls in the application of mixed-model association methods. Nat Genet *46*, 100-106. 10.1038/ng.2876.
47. Jordan, H., Roderick, P., and Martin, D. (2004). The Index of Multiple Deprivation 2000 and accessibility effects on health. J Epidemiol Community Health *58*, 250-257. 10.1136/jech.2003.013011.
48. Cleland, C., Ferguson, S., Ellis, G., and Hunter, R.F. (2018). Validity of the International Physical Activity Questionnaire (IPAQ) for assessing moderate-to-vigorous physical activity and sedentary behaviour of older adults in the United Kingdom. BMC Med Res Methodol *18*, 176. 10.1186/s12874-018-0642-3.
49. Welter, D., MacArthur, J., Morales, J., Burdett, T., Hall, P., Junkins, H., Klemm, A., Flicek, P., Manolio, T., Hindorff, L., and Parkinson, H. (2014). The NHGRI GWAS Catalog, a curated resource of SNP-trait associations. Nucleic Acids Res *42*, D1001-1006. 10.1093/nar/gkt1229.
50. Manolio, T.A., Collins, F.S., Cox, N.J., Goldstein, D.B., Hindorff, L.A., Hunter, D.J., McCarthy, M.I., Ramos, E.M., Cardon, L.R., Chakravarti, A., et al. (2009). Finding the missing heritability of complex diseases. Nature *461*, 747-753. 10.1038/nature08494.
51. Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M.A., Bender, D., Maller, J., Sklar, P., de Bakker, P.I., Daly, M.J., and Sham, P.C. (2007). PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet *81*, 559-575. 10.1086/519795.
52. McLaren, W., Pritchard, B., Rios, D., Chen, Y., Flicek, P., and Cunningham, F. (2010). Deriving the consequences of genomic variants with the Ensembl API and SNP Effect Predictor. Bioinformatics *26*, 2069-2070. 10.1093/bioinformatics/btq330.
53. Bulik-Sullivan, B.K., Loh, P.R., Finucane, H.K., Ripke, S., Yang, J., Schizophrenia Working Group of the Psychiatric Genomics, C., Patterson, N., Daly, M.J., Price, A.L., and Neale, B.M. (2015). LD Score regression distinguishes confounding from polygenicity in genome-wide association studies. Nat Genet *47*, 291-295. 10.1038/ng.3211.
54. Bulik-Sullivan, B., Finucane, H.K., Anttila, V., Gusev, A., Day, F.R., Loh, P.R., ReproGen, C., Psychiatric Genomics, C., Genetic Consortium for Anorexia Nervosa of the Wellcome Trust Case Control, C., Duncan, L., et al. (2015). An atlas of genetic correlations across human diseases and traits. Nat Genet *47*, 1236-1241. 10.1038/ng.3406.
55. Weissbrod, O., Hormozdiari, F., Benner, C., Cui, R., Ulirsch, J., Gazal, S., Schoech, A.P., van de Geijn, B., Reshef, Y., Marquez-Luna, C., et al. (2020). Functionally informed fine-mapping and polygenic localization of complex trait heritability. Nat Genet *52*, 1355-1363. 10.1038/s41588-020-00735-5.
56. Loh, P.R., Kichaev, G., Gazal, S., Schoech, A.P., and Price, A.L. (2018). Mixed-model association for biobank-scale datasets. Nat Genet *50*, 906-908. 10.1038/s41588-018-0144-6.
57. Foley, C.N., Staley, J.R., Breen, P.G., Sun, B.B., Kirk, P.D.W., Burgess, S., and Howson, J.M.M. (2021). A fast and efficient colocalization algorithm for identifying shared genetic risk factors across multiple traits. Nat Commun *12*, 764. 10.1038/s41467-020-20885-8.

58. Wang, G., Sarkar, A., Carbonetto, P., and Stephens, M. (2020). A simple new approach to variable selection in regression, with application to genetic fine mapping. J R Stat Soc Series B Stat Methodol *82*, 1273-1300. 10.1111/rssb.12388.
59. Watanabe, K., Taskesen, E., van Bochoven, A., and Posthuma, D. (2017). Functional mapping and annotation of genetic associations with FUMA. Nat Commun *8*, 1826. 10.1038/s41467-017-01261-5.
60. Lee, S.H., Wray, N.R., Goddard, M.E., and Visscher, P.M. (2011). Estimating missing heritability for disease from genome-wide association studies. Am J Hum Genet *88*, 294-305. 10.1016/j.ajhg.2011.02.002.
61. Liberzon, A., Subramanian, A., Pinchback, R., Thorvaldsdóttir, H., Tamayo, P., and Mesirov, J.P. (2011). Molecular signatures database (MSigDB) 3.0. Bioinformatics *27*, 1739-1740. 10.1093/bioinformatics/btr260.
62. Wu, Y., Qi, T., Wray, N.R., Visscher, P.M., Zeng, J., and Yang, J. (2023). Joint analysis of GWAS and multi-omics QTL summary statistics reveals a large fraction of GWAS signals shared with molecular phenotypes. Cell Genom *3*, 100344. 10.1016/j.xgen.2023.100344.
63. Vosa, U., Claringbould, A., Westra, H.J., Bonder, M.J., Deelen, P., Zeng, B., Kirsten, H., Saha, A., Kreuzhuber, R., Yazar, S., et al. (2021). Large-scale cis- and trans-eQTL analyses identify thousands of genetic loci and polygenic scores that regulate blood gene expression. Nat Genet *53*, 1300-1310. 10.1038/s41588-021-00913-z.
64. Sun, B.B., Chiou, J., Traylor, M., Benner, C., Hsu, Y.H., Richardson, T.G., Surendran, P., Mahajan, A., Robins, C., Vasquez-Grinnell, S.G., et al. (2023). Plasma proteomic associations with genetics and health in the UK Biobank. Nature *622*, 329-338. 10.1038/s41586-023-06592-6.
65. Wu, Y., Zeng, J., Zhang, F., Zhu, Z., Qi, T., Zheng, Z., Lloyd-Jones, L.R., Marioni, R.E., Martin, N.G., Montgomery, G.W., et al. (2018). Integrative analysis of omics summary data reveals putative mechanisms underlying complex traits. Nat Commun *9*, 918. 10.1038/s41467-018-03371-0.
66. McRae, A.F., Marioni, R.E., Shah, S., Yang, J., Powell, J.E., Harris, S.E., Gibson, J., Henders, A.K., Bowdler, L., Painter, J.N., et al. (2018). Identification of 55,000 Replicated DNA Methylation QTL. Sci Rep *8*, 17605. 10.1038/s41598-018-35871-w.
67. Consortium, G.T. (2020). The GTEx Consortium atlas of genetic regulatory effects across human tissues. Science *369*, 1318-1330. 10.1126/science.aaz1776.
68. Chen, L., Ge, B., Casale, F.P., Vasquez, L., Kwan, T., Garrido-Martin, D., Watt, S., Yan, Y., Kundu, K., Ecker, S., et al. (2016). Genetic Drivers of Epigenetic and Transcriptional Variation in Human Immune Cells. Cell *167*, 1398-1414 e1324. 10.1016/j.cell.2016.10.026.
69. Kumasaka, N., Knights, A.J., and Gaffney, D.J. (2019). High-resolution genetic mapping of putative causal interactions between regions of open chromatin. Nat Genet *51*, 128-137. 10.1038/s41588-018-0278-6.
70. Goldstein, J.L., and Brown, M.S. (2009). The LDL receptor. Arterioscler Thromb Vasc Biol *29*, 431-438. 10.1161/ATVBAHA.108.179564.
71. Naik, A., Kosir, R., and Rozman, D. (2013). Genomic aspects of NAFLD pathogenesis. Genomics *102*, 84-95. 10.1016/j.ygeno.2013.03.007.

72. Gao, L., Tse, S.W., Conrad, C., and Andreadis, A. (2005). Saitohin, which is nested in the tau locus and confers allele-specific susceptibility to several neurodegenerative diseases, interacts with peroxiredoxin 6. J Biol Chem *280*, 39268-39272. 10.1074/jbc.M506116200.
73. Schulze, M.B., Minihane, A.M., Saleh, R.N.M., and Riserus, U. (2020). Intake and metabolism of omega-3 and omega-6 polyunsaturated fatty acids: nutritional implications for cardiometabolic diseases. Lancet Diabetes Endocrinol *8*, 915-930. 10.1016/S2213-8587(20)30148-0.
74. Gretarsdottir, S., Helgason, H., Helgadottir, A., Sigurdsson, A., Thorleifsson, G., Magnusdottir, A., Oddsson, A., Steinthorsdottir, V., Rafnar, T., de Graaf, J., et al. (2015). A Splice Region Variant in LDLR Lowers Non-high Density Lipoprotein Cholesterol and Protects against Coronary Artery Disease. PLoS Genet *11*, e1005379. 10.1371/journal.pgen.1005379.
75. Raben, D.M., and Tu-Sekine, B. (2008). Nuclear diacylglycerol kinases: regulation and roles. Front Biosci *13*, 590-597. 10.2741/2704.
76. Sadafi, S., Choubsaz, P., Kazemeini, S.M.M., Imani, M.M., and Sadeghi, M. (2024). Glutathione S-transferase theta 1 (GSTT1) deletion polymorphism and susceptibility to head and neck carcinoma: a systematic review with five analyses. BMC Cancer *24*, 885. 10.1186/s12885-024-12618-7.
77. Ashida, S., Orloff, M.S., Bebek, G., Zhang, L., Zheng, P., Peehl, D.M., and Eng, C. (2012). Integrated analysis reveals critical genomic regions in prostate tumor microenvironment associated with clinicopathologic phenotypes. Clin Cancer Res *18*, 1578-1587. 10.1158/1078-0432.CCR-11-2535.
78. Zhou, T.B., Drummen, G.P., Jiang, Z.P., and Qin, Y.H. (2014). GSTT1 polymorphism and the risk of developing prostate cancer. Am J Epidemiol *180*, 1-10. 10.1093/aje/kwu112.
79. Liu, D., Che, B., Chen, P., He, J., Mu, Y., Chen, K., Zhang, W., Xu, S., and Tang, K. (2022). GSTT1, an increased risk factor for prostate cancer in patients with metabolic syndrome. J Clin Lab Anal *36*, e24352. 10.1002/jcla.24352.
80. Zhang, Y., Sun, Y., Song, S., Khankari, N.K., Brenna, J.T., Shen, Y., and Ye, K. (2024). Associations of plasma omega-6 and omega-3 fatty acids with overall and 19 site-specific cancers: A population-based cohort study in UK Biobank. Int J Cancer. 10.1002/ijc.35226.
81. Li, H., Herrmann, T., Seessle, J., Liebisch, G., Merle, U., Stremmel, W., and Chamulitrat, W. (2022). Role of fatty acid transport protein 4 in metabolic tissues: insights into obesity and fatty liver disease. Biosci Rep *42*. 10.1042/BSR20211854.
82. Pingault, J.B., O'Reilly, P.F., Schoeler, T., Ploubidis, G.B., Rijsdijk, F., and Dudbridge, F. (2018). Using genetic data to strengthen causal inference in observational research. Nat Rev Genet *19*, 566-580. 10.1038/s41576-018-0020-3.
83. Tayefeh-Gholami, S., Ghanbari, M., Aghazadeh, A., Rajabi, A., Saber, A., Hussen, B.M., Farsad-Akhtar, N., and Safaralizadeh, R. (2022). Prognostic Value of LncRNA KRT18P55 in Patients with Intestinal Type of Gastric Cancer. J Gastrointest Cancer *53*, 1014-1019. 10.1007/s12029-021-00744-5.
84. Tillander, V., Arvidsson Nordstrom, E., Reilly, J., Strozyk, M., Van Veldhoven, P.P., Hunt, M.C., and Alexson, S.E. (2014). Acyl-CoA thioesterase 9 (ACOT9) in mouse may provide a

novel link between fatty acid and amino acid metabolism in mitochondria. Cell Mol Life Sci *71*, 933-948. 10.1007/s00018-013-1422-1.

85. Welty, F.K., Daher, R., and Garelnabi, M. (2024). Fish and Omega-3 Fatty Acids: Sex and Racial Differences in Cardiovascular Outcomes and Cognitive Function. Arterioscler Thromb Vasc Biol *44*, 89-107. 10.1161/ATVBAHA.122.318125.

86. Jankowich, M., and Choudhary, G. (2020). Endothelin-1 levels and cardiovascular events. Trends Cardiovasc Med *30*, 1-8. 10.1016/j.tcm.2019.01.007.

87. Davegardh, C., Sall, J., Benrick, A., Broholm, C., Volkov, P., Perfilyev, A., Henriksen, T.I., Wu, Y., Hjort, L., Brons, C., et al. (2021). VPS39-deficiency observed in type 2 diabetes impairs muscle stem cell differentiation via altered autophagy and epigenetics. Nat Commun *12*, 2431. 10.1038/s41467-021-22068-5.

Figure Captions

Main Figures

Figure 1. Manhattan plot of the genome-wide association studies for the absolute concentration of total polyunsaturated fatty acids. (A and B) The association of each variant with polyunsaturated fatty acids was obtained from genome-wide association studies using models 2 (A) and 3 (B). A variant with a *P*-value below 5×10^{-8} was considered statistically significant. The nearest gene at each genome-wide significant locus is annotated, in black for known loci and in red for novel loci.

Figure 2. Miami plot of the absolute concentration of total polyunsaturated fatty acids from genome-wide association studies using model 2. The top panel shows the GWAS results in females, while the bottom panel shows the GWAS results in males. The $-\log_{10}$(*P*-value) is plotted on the y-axis and chromosomal location is plotted on the x-axis. The genome-wide significance threshold (*P*-value $< 5 \times 10^{-8}$) is indicated by the red dashed line. The nearest gene at each genome-wide significant locus is annotated, in black for known loci and in red for novel loci. Note that the lead variant (rs165527) from the novel locus on chromosome 10 (chr10:17259567-17325281) is not shown in the figure because it is an intergenic variant without an associated gene.

Figure 3. Pathway enrichment analysis of the absolute concentration of total polyunsaturated fatty acids from genome-wide association studies using model 2. Treemap depicting significantly enriched pathways at an adjusted *P*-value threshold of 0.05. Gene ontology terms were clustered based on semantic similarity, with terms displayed as individual

rectangles. The color indicates cluster membership, and thick border lines differentiate clusters. The size of each rectangle represents the enrichment significance, and the most significantly enriched term in each cluster is highlighted in white text as the representative term.

Figure 4. Number of loci for fatty acid traits associated with different molecular phenotype combinations. (A and B) The number of loci associated with 19 fatty acid traits and molecular quantitative trait loci (xQTLs) was determined based on genome-wide association studies using models 2 (A) and 3 (B). The loci numbers are based on significant OPERA association results that pass a posterior probability of association threshold of 0.9 and the multi-exposure heterogeneity in dependent instruments (HEIDI) test ($P_{HEIDI} > 0.01$). The x-axis represents the association hypotheses for different phenotype combinations, while the y-axis shows the number of loci associated with these combinations across 19 fatty acids. Each violin plot displays the distribution of loci numbers by width, with lines indicating the 25th percentile, median, and 75th percentile.

Figure 5. Prioritization of a locus near the *GSTT1/2/2B* genes for the absolute concentration of total fatty acids. The top track displays the -$\log_{10}$(*P*-value) of GWAS SNPs (gray dots) for total fatty acids. Red diamonds indicate OPERA marginal PPA for gene associations using eQTL data, while blue circles show OPERA marginal PPA for associations with protein abundance, DNA methylation, RNA splicing, histone modification, and chromatin accessibility, respectively. The bottom track presents 14 chromatin state annotations inferred from the 127 samples of the Roadmap Epigenomics Mapping Consortium.

Figure 1.

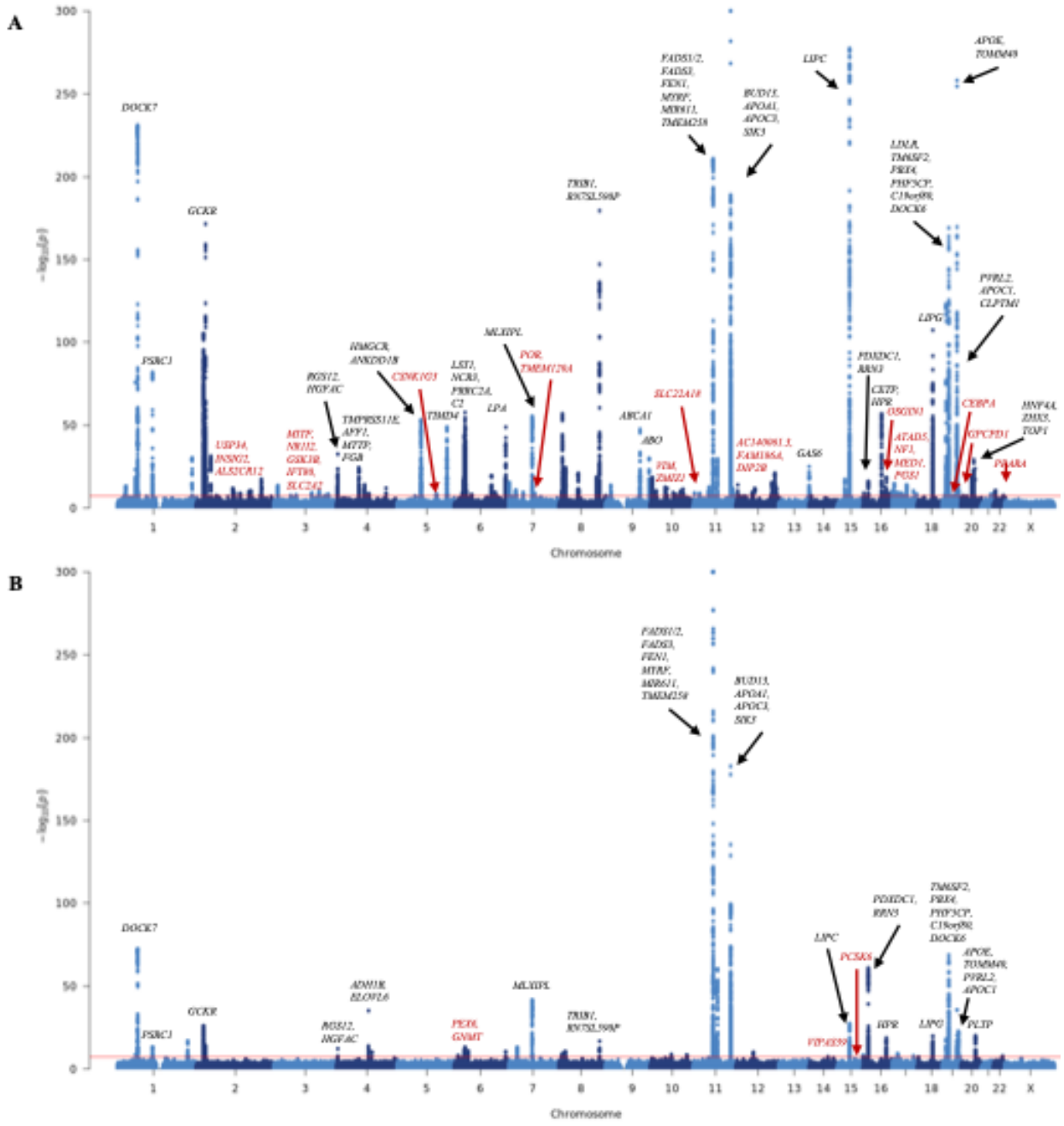

Figure 2.

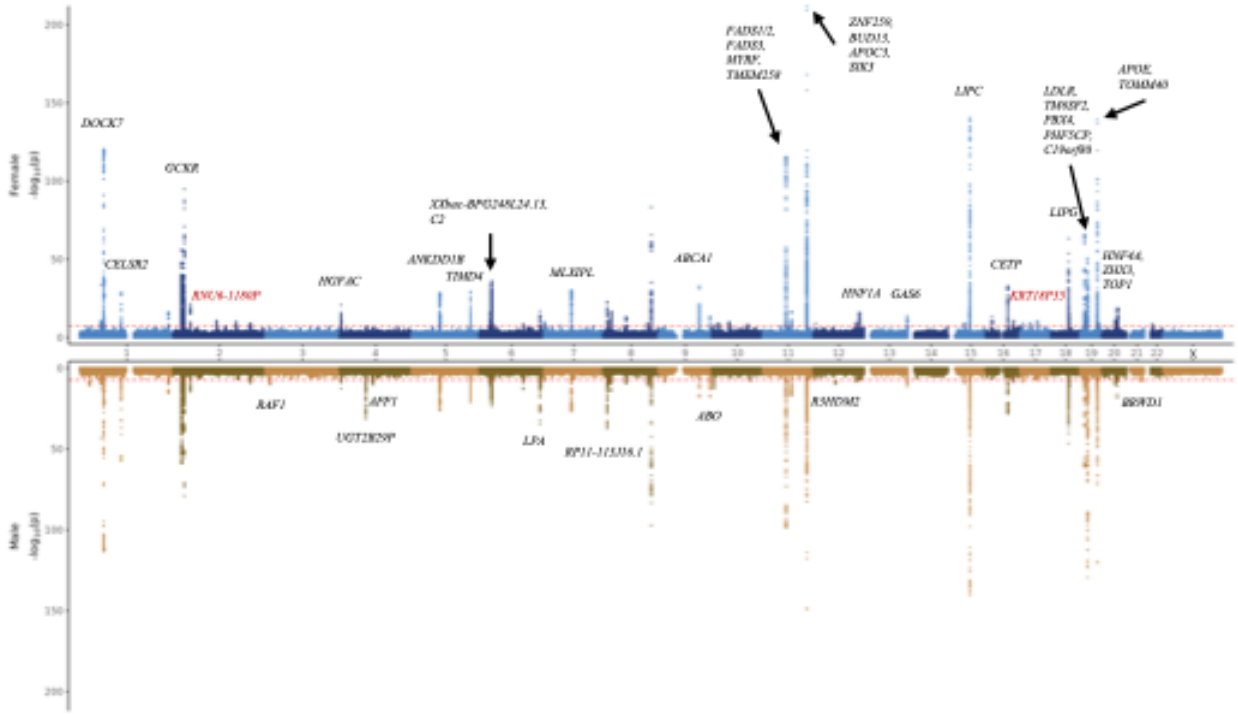

Figure 3.

Figure 4.

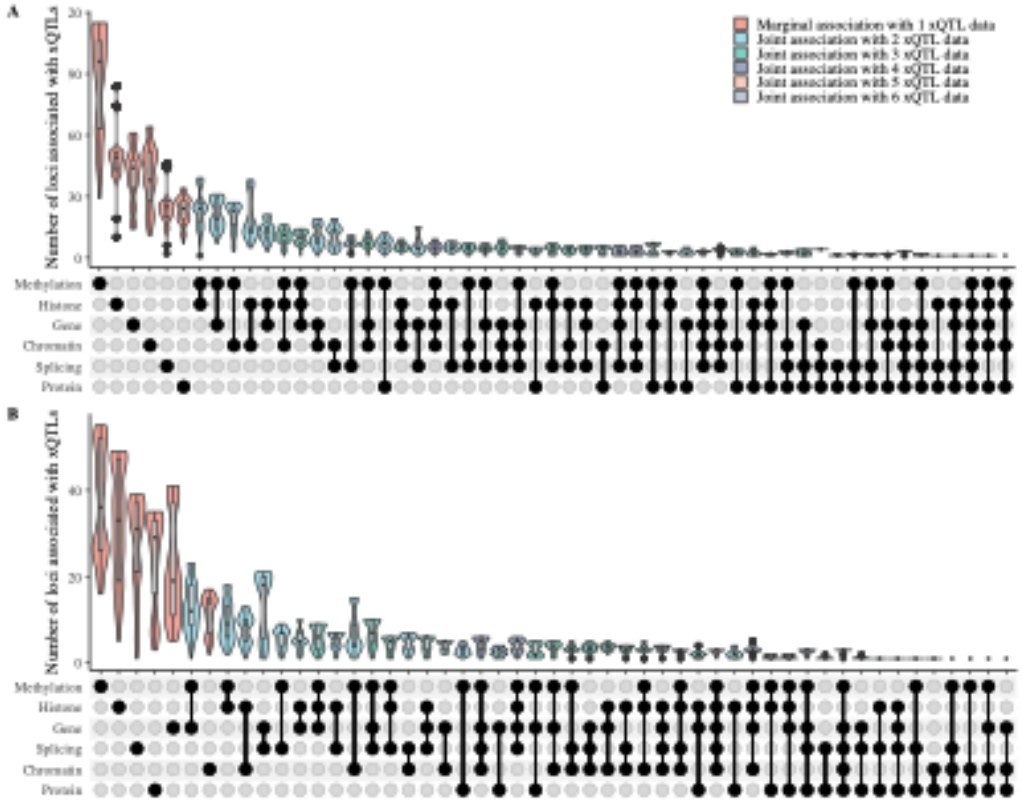

Figure 5.