

Title: Genome-wide association study reveals shared and distinct genetic architecture underlying fatty acid and bioactive oxylipin metabolites in the Hispanic Community Health Study/Study of Latinos (HCHS/SOL)

Authors:

Carolina G. Downie¹, Heather M. Highland¹, Mona Alotaibi², Barrett M. Welch³, Annie Green Howard⁴, Susan Cheng⁵, Nick Miller⁶, Mohit Jain^{6,7}, Robert C. Kaplan⁸, Adam G. Lilly⁹, Tao Long⁶, Tamar Sofer¹⁰, Bharat Thyagarajan¹¹, Bing Yu¹², Kari E. North¹, Christy L. Avery¹

Affiliations:

¹Department of Epidemiology, University of North Carolina at Chapel Hill, Chapel Hill, NC

²Division of Pulmonary, Critical Care and Sleep Medicine, University of California, San Diego, San Diego, CA

³School of Public Health, University of Nevada, Reno, Reno, NV

⁴Department of Biostatistics, University of North Carolina at Chapel Hill, Chapel Hill, NC

⁵Department of Cardiology, Smidt Heart Institute, Cedars-Sinai Medical Center, Los Angeles, CA

⁶Sapient Bioanalytics, San Diego, CA

⁷Departments of Medicine and Pharmacology, University of California, San Diego, San Diego, CA

⁸Department of Epidemiology and Population Health, Albert Einstein College of Medicine, Bronx, NY; Public Health Sciences Division, Fred Hutchison Cancer Center, Seattle, WA

⁹Department of Sociology, University of North Carolina at Chapel Hill, Chapel Hill, NC

¹⁰CardioVascular Institute (CVI), Beth Israel Deaconess Medical Center, Boston, MA;
Department of Biostatistics, Harvard T.H. Chan School of Public Health, Boston, MA

¹¹Department of Laboratory Medicine and Pathology, University of Minnesota Medical Center,
Minneapolis, MN

¹²Department of Epidemiology, Human Genetics, and Environmental Sciences, The University
of Texas Health Science Center at Houston School of Public Health, Houston, TX

Abstract:

Bioactive fatty acid-derived oxylipin molecules play key roles in mediating inflammation and oxidative stress, which underlie many chronic diseases. Circulating levels of fatty acids and oxylipins are influenced by both environmental and genetic factors; characterizing the genetic architecture of bioactive lipids could yield new insights into underlying biological pathways. Thus, we performed a genome wide association study (GWAS) of $n=81$ fatty acids and oxylipins in $n=11,584$ Hispanic Community Health Study/Study of Latinos (HCHS/SOL) participants with genetic and lipidomic data measured at study baseline (58.6% female, mean age = 46.1 years, standard deviation = 13.8 years). Additionally, given the effects of central obesity on inflammation, we examined interactions with waist circumference using two-degree-of-freedom joint tests. Heritability estimates ranged from 0% to 47.9%, and 48 of the 81 oxylipins and fatty acids were significantly heritable. Moreover, 40 (49.4%) of the 81 oxylipins and fatty acids had at least one genome-wide significant ($p < 6.94E-11$) variant resulting in 19 independent genetic loci involved in fatty acid and oxylipin synthesis, as well as downstream pathways. Four loci (lead variant minor allele frequency [MAF] range: 0.08-0.50), including the desaturase-encoding *FADS* and the OATP1B1 transporter protein-encoding *SLCO1B1*, exhibited associations with four or more fatty acids and oxylipins. The majority of the 15 remaining loci (87.5%) (lead variant MAF range = 0.03-0.45, mean = 0.23) were only associated with one oxylipin or fatty acid, demonstrating evidence of distinct genetic effects. Finally, while most loci identified in two-degree-of-freedom tests were previously identified in our main effects analyses, we also identified an additional rare variant (MAF = 0.002) near *CARS2*, a locus previously implicated in inflammation. Our analyses revealed shared and distinct genetic architecture underlying fatty

acids and oxylipins, providing insights into genetic factors and motivating future multi-omics work to characterize these compounds and elucidate their roles in disease pathways.

Introduction

Chronic inflammation is a pathological feature underlying many chronic diseases,¹⁻⁵ and inflammatory responses involve a variety of so-called “inflammatory mediators” that include both signaling proteins (e.g., cytokines, chemokines) and bioactive lipids such as oxylipins.⁶ Oxylipins are derived from the oxygenation of precursor mono- and polyunsaturated fatty acids through cascading pathways mediated by cytochrome P450 (CYP), lipoxygenase (LOX), and cyclooxygenase (COX) enzymes, or through other enzymatic and non-enzymatic reactions⁷ (**Supplementary Figure 1**), and play a variety of roles mediating inflammation and oxidative stress.^{8,9} Given the role of inflammation in many chronic diseases, there is growing interest in oxylipin-targeted therapies.¹⁰ Oxylipins and their precursor fatty acids are influenced by both environmental and genetic factors.¹¹⁻¹³ Thus, characterizing the genetic architecture of oxylipins and related fatty acids could yield new insights into underlying biological pathways that could inform drug targets.¹¹

Genetic analyses of several polyunsaturated fatty acids from which many oxylipins are derived have identified multiple variants at loci encoding desaturase and elongase enzymes (e.g., *FADS*, *ELOVL2*).¹⁴⁻¹⁸ Moreover, previous associations between the LOX-encoding *ALOX12* locus and 12-HETE oxylipins have been reported, consistent with the role that LOX enzymes play in the derivation of these metabolites.^{12,19} Still, few analyses of the genetic architecture of oxylipins have been conducted on a large scale. This is due in part to historical challenges in quantifying and annotating oxylipin species: hundreds of oxylipins with distinct structures and stereochemistry can be derived from fatty acids.⁸ However, recent advancements in mass spectrometry-based metabolomics and annotation approaches have enabled the improved detection, quantification, and annotation of oxylipins.²⁰⁻²²

To address this research gap, we estimated the heritability and performed a genome-wide association study (GWAS) of n=81 oxylipins and related fatty acids quantified via directed, non-targeted liquid chromatography-mass spectrometry (LC-MS) in the Hispanic Community Health Study/Study of Latinos (HCHS/SOL). Moreover, motivated by the fact that central adiposity is associated with other markers of low-grade inflammation (e.g., C-reactive protein and interleukin-6)²³⁻²⁶ and may influence oxylipin biosynthesis,²⁷ we also extended our models to include interactions with waist circumference and performed two-degree-of-freedom joint tests, increasing our power to identify additional loci that have both main and interaction effects.^{28,29} Our findings reveal both shared and distinct genetic architecture of modest-to-moderately heritable circulating fatty acids and oxylipins, providing insights into the biological pathways underlying these metabolites.

Methods

Hispanic Community Health Study/Study of Latinos (HCHS/SOL)

HCHS/SOL is a multi-center, community-based cohort study of Hispanic/Latino adults in the United States designed to study the incidence and prevalence of, and risk factors for, multiple chronic diseases, including cardiovascular diseases, kidney, lung and liver diseases.³⁰ A total of 16,415 participants of Mexican, Puerto Rican, Dominican, Cuban, and Central and South American background aged 18-74 years were recruited from four U.S. cities between 2008 and 2011: Bronx, NY; Chicago, IL; Miami, FL; and San Diego, CA. Households were selected via two-stage sampling within census block groups.³¹ At study baseline (2008-2011), participants received standardized examinations and interviewer-administered questionnaires, with follow-up examinations in 2015-2017 and 2020-2023. Informed consent was obtained from all participants, and the study was conducted in accordance with the principles of the Declaration of Helsinki.

Approval from the University of North Carolina at Chapel Hill Institutional Review Board was obtained for this work.

Oxylipin profiling, annotation, and quality control

Plasma oxylipin profiling was conducted by Sapient Bioanalytics, using a rapid liquid chromatography-mass spectrometry (rLC-MS) methodology that has been described in detail previously²² (**Supplementary Methods**). Oxylipins and fatty acids were annotated by aligning experimental accurate mass-to-charge ratio (m/z), retention time (RT) and MS2 spectra with an in-house curated library of commercial standards as well as alignment with previously reported structural isomers.¹ These annotations often indicated multiple potential chemically related compound mappings, due to compounds with similar chemical formulas and mass-to-charge ratios co-eluting at the same time. Furthermore, oxylipin metabolites with chemical structures that aligned with previously reported structural isomers for which biological pathways were unknown were annotated as “putative known oxylipins.” Likewise, oxylipins with chemical structures that did not align to any previously reported standards or structural isomers and for which biological pathways were unknown were annotated as “putative novel oxylipins.”

In this study, we restricted our analysis to n=81 metabolites annotated as oxylipins and fatty acids based on the procedure described above. To facilitate interpretation of analyses, for the oxylipin metabolites, we defined precursor fatty acid groupings based on literature searches describing the biosynthetic pathways of oxylipin compounds as well as online resources such as PubChem (<https://pubchem.ncbi.nlm.nih.gov>), the Human Metabolome Database (<https://hmdb.ca>), and Cayman Chemical (<https://www.caymanchem.com>). When annotations indicated multiple potential compound mappings due to compounds with similar chemical formulas and mass-to-charge ratios co-eluting at the same time, we grouped them based on any

of the precursor fatty acids from which they may have been derived (e.g., “arachidonic acid or eicosapentaenoic acid”). Moreover, fatty acids were grouped based on the class of fatty acids to which they belonged: polyunsaturated fatty acids, monounsaturated fatty acids, or saturated fatty acids.

The percent missingness across all 81 metabolites ranged from 0% -23.4%, with the majority of metabolites (93.8%) missing less than 2% of all observations. Metabolite observations with missing data were imputed from a random uniform distribution between the minimum measured metabolite value/10 and the minimum measured metabolite values. All metabolite values were subsequently \log_2 -transformed to reduce the impact of extreme outlier values.³²

Genotyping and quality control

HCHS/SOL participants were genotyped on the Infinium Expanded Multi-Ethnic Genotyping Array (MEGA)³³, and imputed to the TOPMed R2 (Freeze 8) imputation panel using IMPUTE v2.3.2 (**Supplementary Methods**). Variants with an imputation quality score < 0.4 or effective sample size < 20 , calculated as $2 \times \text{MAF} \times (1 - \text{MAF}) \times N \times Q$, where MAF is minor allele frequency, Q is imputation quality, and N is sample size, were excluded from main effect analyses. For joint effect analyses, an effective sample size filter of 30 was applied, where the N used in effective sample size calculations was the sex-stratified sample size ($N_{\text{females}} = 6725$, $N_{\text{males}} = 4753$). All genomic coordinates are reported on GRCh38.

Spearman correlations

Because Spearman correlations are more robust to outliers and non-normal distributions than Pearson correlations,³⁴ we computed Spearman correlations of the n=81 oxylipins and fatty acids to determine their correlation structure.

Heritability analysis

We estimated SNP-based heritability, or the proportion of variance explained for each metabolite due to the additive effect of genotyped common variants, and corresponding 95% confidence intervals, using Haseman-Elston regression, as previously described for the HCHS/SOL study.³⁵ For this analysis, we used a kinship matrix calculated using autosomal variants genotyped on an Illumina custom array (SOL HCHS Custom 15041502 B3), consisting of the Illumina Omni 2.5M array (HumanOmni2.5-8v1-1) and approximately 150,000 custom variants,³⁶ and then imputed to the 1000 Genomes Phase 3 reference panel.

Main effects GWAS

The \log_2 -transformed metabolite values were regressed on age, age², sex, study recruitment site, and self-identified background, and residuals from these models were subsequently inverse rank normalized to yield asymptotically normal marginal distributions.³⁷ Genetic association analyses of the inverse rank normalized residuals were conducted using generalized estimating equations via SUGEN,³⁸ adjusting for 10 ancestral principal components, which were calculated using Plink.

To account for multiple testing of 81 correlated oxylipins and fatty acids, we used a modified Bonferroni correction based on the effective number of independent tests.³⁹ Briefly, we computed the number of principal components necessary to explain 99.5% of metabolite variability ($N_{\text{effective}} = 72$), yielding a multiple testing-corrected p-value threshold of $5E-09/72 = 6.94E-11$. We used EasyStrata software⁴⁰ to identify independent loci and corresponding top variants, using 500 KB as the physical position threshold for clumping. We used the Open Targets Genetics tool (<https://genetics.opentargets.org>) to annotate the nearest genetic locus for each of our top variants.

Main effects conditional analyses

We performed stepwise exact conditional analyses to identify additional independent secondary signals for each genome-wide significant locus identified in the main effects GWAS. Briefly, we included the top genome-wide significant variant at each significant locus as a covariate in our SUGEN analysis and iterated until no novel significant secondary signals ($p < 6.94E-10$) remained. We then estimated the proportion of variance in the phenotype explained (PVE) by a given variant for each independent index and secondary variants using the following equation, as previously described^{41,42}:

$$PVE = \frac{2\hat{\beta}^2 MAF(1-MAF)}{2\hat{\beta}^2 MAF(1-MAF) + se(\hat{\beta})^2 2NMAF(1-MAF)}$$

Joint effects GWAS

To increase the power to identify additional loci in which genetic variants exhibit both a main effect and an interaction effect,^{28,29} we also conducted joint two-degree-of-freedom tests accounting for waist circumference as a measure of central adiposity. Because sex-specific waist circumference thresholds have been recommended for defining central obesity,^{43,44} we conducted sex-stratified joint effect analyses, including inverse rank normalized residuals of waist circumference adjusted for body mass index (BMI) as an interaction term in our analyses. Waist circumference, weight, and height were measured using standardized protocols at the baseline study visit. Waist circumference adjusted for BMI was calculated in sex-stratified analyses, regressing waist circumference on age, age², and BMI, and then inverse rank normalizing the residuals.⁴⁵ We then meta-analyzed the p-values for the joint variant and variant x waist circumference effect (“joint p-value”) using METAL,⁴⁶ weighting by sample size.

All figures were generated using R version 4.1.0, using the colorblind-friendly color palette from the following resource: <https://zenodo.org/record/3381072>.

Results

There were $n=11,610$ SOL participants with measured genetic and metabolite data at the baseline visit. After excluding $n=26$ participants with missing self-reported background group, the final main effects analysis sample was $n=11,584$. This analytic sample was 58.6% female, and the mean age of participants was 46.1 years (standard deviation [SD]=13.8). After excluding an additional $n=106$ with missing BMI-adjusted-waist circumference, the final joint effect analysis sample was $n_{\text{Females}}=6,725$ and $n_{\text{Males}}=4,753$ (total $n=11,478$). The mean waist circumference for females was 97.6 centimeters (SD=14.2 cm), and the mean waist circumference for males was 99.0 centimeters (SD=13.3 cm).

Spearman correlation of fatty acids and oxylipins

Spearman correlations of the \log_2 -transformed values of the 81 oxylipins and fatty acids ranged between -0.27 to 0.98 (mean=0.21), with the strongest correlations observed between fatty acids (range: -0.22 to 0.90, mean=0.34) (**Figure 1, Supplementary Figure 2**). Among the oxylipins, correlations varied both within and between precursor fatty acid groupings and were highest for oxylipins derived from linoleic acid (range: -0.12 to 0.70, mean=0.28) and lowest for oxylipins putatively derived from arachidonic acid or eicosapentaenoic acid (range: -0.06 to 0.18, mean=0.07) (**Figure 1, Supplementary Figure 2**).

Heritability of fatty acids and oxylipins

Forty-eight of the 81 (59.3%) fatty acids and oxylipins exhibited significant ($p < 0.05$) heritability (significant h_{SNP}^2 range: 4.6% to 47.9%), the magnitude of which varied across the 34 fatty acids and 47 oxylipins (**Supplementary Table 1**). Among metabolites with significant heritability, fatty acids generally exhibited higher heritability (mean $h_{\text{SNP}}^2 = 12.6\%$, SD = 10.2%, range: 5.0% to 47.9%) than oxylipins (mean $h_{\text{SNP}}^2 = 8.2\%$, SD=3.4%, range: 4.6% to 16.9%)

(**Figure 2, Supplementary Table 1**). Among oxylipins, heritability estimates were generally consistent across fatty acid precursor groups.

Main effects GWAS

Of the 81 fatty acids and oxylipins, 40 (49.4%) had at least one genome-wide significant ($p < 6.94E-11$) variant. These metabolites included 33 of the 48 (68.8%) significantly heritable fatty acids and oxylipins (**Supplementary Table 1**). After accounting for variants that mapped to the same locus, there were 19 independent loci, and 55 unique metabolite-locus pairs (**Figure 3a, Supplementary Table 2, Supplementary Figure 3**). The top variants were all common (minor allele frequency (MAF) > 0.01), and all of these variants or their high linkage disequilibrium (LD) proxies have been previously associated with metabolites, though not necessarily oxylipins (**Supplementary Table 2**). As previously reported, allele frequencies for several of the loci we identified varied across reference continental ancestry populations, including *FADS*¹⁸, the *CYP* supergene family,⁴⁷ *SLCO1B1*,^{48,49} *FUT2*,⁵⁰ and *MCM6*.⁵¹ Moreover, *ADH1A* top variant rs28600890 was monomorphic in 1000 Genomes European and South and East Asian reference populations (**Supplementary Table 2, Supplementary Figure 4**).

Shared Genetic Architecture

The genetic architecture of the $n = 40$ fatty acids and oxylipins with at least one genome-wide significant variant was characterized by both shared and distinct genetic effects. Of the 20 identified loci, four pleiotropic loci (*SLCO1B1*, *FADS*, *PKD2L1*, and *PPP1R3B*) were associated with four or more metabolites (**Figure 3b**). These loci accounted for 38 (69.1%) of the 55 unique metabolite-locus associations, included common variants (MAF range: 0.08 – 0.50, mean: 0.33), and generally belonged to similar fatty acid or precursor fatty acid groupings, though *PKD2L1* was associated with saturated and monounsaturated fatty acids as well as oxylipins (see **Figure**

3b, Supplemental Table 5). Moreover, most of the oxylipins and fatty acids (71.4%) associated with these pleiotropic loci did not have significant associations with any other additional loci. Furthermore, there were 12 highly correlated metabolite pairs (Spearman correlation coefficient > 0.80) for which one metabolite had a genome-wide significant association with one of these pleiotropic loci, but the other did not have a genome-wide or near genome-wide significant association for the same locus (see **Supplementary Table 3**). For example, a fatty acid annotated as dihomogamma-linolenic acid/eicosatrienoic acid ($m/z = 305.25$, $RT=0.386$) was associated with both the *FADS* and *NTAN1* loci ($p_{FADS} = 1.20E-22$, $p_{NTAN} = 5.03E-27$), but correlated eicosadienoic acid (Spearman correlation = 0.81, $m/z=307.26$, $RT=0.401$) was associated with neither locus ($p_{FADS} = 0.71$, $p_{NTAN} = 0.57$). This may reflect non-genetic drivers, such as diet, of the correlations between these fatty acids and oxylipins.

These four pleiotropic loci also showed little evidence of allelic heterogeneity (**Supplementary Tables 2, 4**). For example, lead missense variant rs4149056 was shared by all 16 *SLCO1B1*-associated metabolites, and variant rs11045856 was identified as the top secondary variant in all conditional analyses. Similarly, lead variant rs174570 was shared by five of the 11 *FADS*-associated metabolites, including three of the four oxylipins associated with *FADS*. Variants underlying a shared genetic architecture also explained the highest observed estimated percentage of trait variance. For example, *FADS* variant rs174564 explained 9.4% of arachidonic acid ($m/z = 303.23$, $RT=0.374$) trait variance and *SLCO1B1* variant rs4149056 explained 6.78% of HpODE ($m/z = 311.22$, $RT=0.343$) trait variance (**Supplementary Table 2**).

Distinct Genetic Architecture

The remaining 15 significant loci were each associated with three or fewer metabolites, including 13 only associated with a single metabolite (*MCM6*; *SCARB1*; *SEC14L6*; *THEM4*;

CYP3A5, ZSCAN25; ADH1A; FUT2; PECR; CYP2C9; HACL1; NTAN1, PDXDC2, NPIPA8; UGT2B7; CYP4A11, CYP4Z2P) (**Figure 3b, Supplementary Table 2**). Some of these loci belonged to shared gene families, such as the four CYP loci associated with only two (*CYP4F2*) or one (*CYP2C9; CYP4A11, CYP4Z2P; CYP3A5, ZSCAN25*) metabolites. Lead variants at these loci were common (MAF range: 0.18-0.25, mean = 0.23) and demonstrated evidence of distinct genetic effects within a broad biosynthetic pathway among modestly correlated metabolites (Spearman correlation range: 0.11 - 0.74) (**Supplementary Table 5**).

We also identified unique genetic loci for metabolites that were highly correlated. For example, pentadecanoic acid (m/z = 241.22, RT=0.376) was highly correlated with myristic acid (m/z = 227.20, RT=0.368) (Spearman correlation = 0.83), but only pentadecanoic acid was significantly associated with the *MCM6* locus (rs4988235, MAF = 0.21, $p_{\text{pentadecanoic acid}} = 6.57\text{E-}16$, $p_{\text{myristic acid}} = 3.91\text{E-}04$) (**Supplementary Table 3**). For 13 of the 15 loci associated with only one or two metabolites, the lead variant (MAF range: 0.02 – 0.45, mean: 0.22) explained less than 1% of the trait variance (**Supplementary Table 2**).

Unique putative novel oxylipin was associated with 5 genetic loci

Thirty-seven of the 40 (92.5%) fatty acids and oxylipins were only significantly associated with one or two genetic loci. However, one oxylipin annotated as a “putative novel oxylipin” (m/z = 365.27, RT=0.351), exhibited the highest heritability of all oxylipins ($h_{\text{SNP}}^2 = 16.9\%$ [95% CI: 10.9% - 23.5%]) and was associated with five loci, including shared genetic locus *SLCO1B1*. Three of these loci were specific only to this metabolite: *PECR* (missense variant rs9288513, MAF = 0.13), *CYP2C9* (rs9332172, MAF = 0.17), and *FUT2* (stop gained variant rs601338, MAF = 0.36). Together, these five lead variants explained 5.0% of trait variance. (**Supplementary Table 2**).

Joint effects GWAS

Joint two-degree-of-freedom tests, which were conducted to increase the power to identify additional loci in which genetic variants exhibit both a main effect and interaction effect with waist circumference,^{28,29} identified 35 fatty acids and oxylipins with at least one genome-wide significant locus ($p_{joint} < 6.94E-11$) (**Supplementary Table 6**). The chromosome 13 *CARS2* locus, which was associated with a leukotriene (m/z=480.25, RT=0.351), was the only locus not previously identified in our main effect GWAS ($p_{main\ effects} = 0.16$). The *CARS2* top variant rs183091953 was significant among females (joint effects $p_{joint, females} = 2.00E-15$), but it did not pass the effective N > 30 filter in males (**Supplementary Table 6**). Top variant rs183091953 was infrequent in HCHS/SOL participants (MAF = 0.002, effective N = 31), and is monomorphic in 1000 Genomes African, European, and South and East Asian reference continental ancestry groups.

Discussion

In this study, we performed one of the largest GWAS of fatty acids and oxylipins, which contribute to the initiation and resolution of inflammation. Our GWAS identified 19 genetic loci association with 40 oxylipins and fatty acids, including several associations driven by missense or stop gained variants. This work builds on prior GWAS of polyunsaturated fatty acids and a small number of oxylipins and illustrates the shared and distinct genetic architecture affecting fatty acid and oxylipin levels. These results provide insights into the genetic factors influencing circulating fatty acids and oxylipins, and motivate future multi-omics work to better characterize these metabolites and their roles in disease pathways.

Previous studies have reported a wide range in heritability estimates (0% to ~80%)^{12,52,53} across a variety of metabolites, with significant differences often reported for different

metabolite classes.⁵² One previous study of serum oxylipins and polyunsaturated fatty acids in fasting samples from 138 participants from the University College Dublin twin study⁵⁴ reported heritability estimates up to 74% for oxylipins, which is much larger than our oxylipin heritability estimates. However, given that all heritability estimates are population specific, we know that distinct estimates may be driven by genetic and environmental factors, such as diet. For example, the highest heritability estimates in our sample were reported for polyunsaturated fatty acids that are far along the enzymatically regulated biosynthesis pathway from essential dietary fatty acids linoleic acid and alpha linolenic-acid (e.g., arachidonic acid, adrenic acid), whereas the heritability estimates of essential dietary fatty acids linoleic acid and alpha-linolenic acid were much lower.

All of the variants we identified in our main effects GWAS, or high LD proxies, have been previously associated with fatty acid metabolites, phospholipids or other unannotated metabolites, as reported in the GWAS Catalog.⁵⁵ Furthermore, two recent studies in the Framingham and Atherosclerosis in Communities (ARIC) cohorts also identified many of the same loci associated with a variety of oxylipins and fatty acids as we report here.^{56,57} Many of the loci underlying the genetic architecture shared between multiple oxylipins and fatty acids are involved in downstream effects, such as small molecule transportation or signaling, that have important consequences for many endogenous compounds. For example, the *SLCO1B1* locus, which was associated with multiple oxylipins from both n-6 (e.g., arachidonic acid, linoleic acid) and n-3 (e.g., alpha-linolenic acid) fatty acid precursor groups, encodes the hepatocyte-expressed OATP1B1 protein, a membrane influx transporter.⁴⁹ OATP1B1 has both exogenous substrates, including statin drugs, and endogenous substrates like oxylipins, bilirubin, and conjugated steroids.^{49,58} Several variants, including our top variant rs4149056, a missense variant, have been

associated with statin-induced myopathy and other related drug interactions^{48,49} due to their impact on OATP1B1 transport activity.^{48,59} Previous studies on the effect of *SLCO1B1* variants on endogenous substrate transportation have been limited to bilirubin and were largely inconclusive.^{49,60-62} However, the associations we identified between *SLCO1B1* and many oxylipins motivate the need for future experimental studies to determine whether rs4149056 and other variants also impact oxylipin transportation.

In addition to shared genetic architecture, we also identified 15 loci associated with three or fewer metabolites, indicating that there were also distinct genetic effects influencing circulating fatty acid and oxylipin levels. Among these distinct effect loci, the average percent of trait variance explained by our top variants was 0.94%, suggesting that many of the metabolites associated with these loci may be more polygenic, with additional unidentified small effect variants. These loci are involved in a range of biological functions, including oxylipin synthesis and metabolism (*CYP* genes),⁶³⁻⁶⁵ and elimination and detoxification of drugs and other endogenous compounds (*UGT2B7*).^{66,67} Indeed, several of these distinct loci may offer insights into the potential identity of putative novel oxylipins. For example, in addition to associations with OATP1B1 transporter protein locus *SLCO1B1* and *CYP* locus *CYP2C9*, one putatively novel oxylipin was also associated with loci implicated in fatty acid elongation (*PECR*),^{68,69} hydrolysis of activated fatty acids into their non-esterified forms (*ACOT6/ACOT2*, *HEATR4*),⁷⁰ and catalyzation of the transfer of fucose residues to cell surface acceptor molecules such as glycoproteins, glycolipids, and oligosaccharides (*FUT2*).^{71,72} Collectively, the activities of these associated loci suggest that this oxylipin could be a long chain fatty acid compound or related metabolite, though more formal analyses will be necessary to characterize putative novel oxylipin metabolites. Emerging methods that integrate genomics and metabolomics are showing

promise for metabolite annotation,⁷³ offering novel means to improve annotation given the ubiquity of large-scale genomics.

Notably, contrary to previous studies,^{12,19} we did not identify associations between genes encoding lipoxygenase (LOX) enzymes (*ALOX5*, *ALOX12*, *ALOX12B*, *ALOXE3*, *ALOX15*, *ALOX15B*)⁷⁴ and oxylipins derived from these enzymatic reactions. Our inability to replicate these associations may be due to the composition of oxylipin species quantified in our samples; for example, there were relatively few HETE oxylipins. This finding likely reflects both heterogeneity in the composition of fatty acid and oxylipin profiles in our data due to diet and other factors, as well as limitations in the non-targeted LC-MS methods used to specifically capture and annotate bioactive metabolites. Similarly, while other studies have reported associations between the *ELOVL2* locus and eicosapentaenoic acid, docosapentaenoic acid, docosahexaenoic acid, and adrenic acid,^{16,57} we did not identify those associations here.

However, we did identify two variants (rs28600890-*ADH1A* and rs183091953-*CARS2*) that are monomorphic in at least one other 1000 Genomes continental ancestry reference population, but not in the admixed American HCHS/SOL sample. *ADH1A* top variant rs28600890, which is monomorphic in European and South and East Asian 1000 Genomes reference continental ancestry populations, is in high LD ($D' = 1$) with variant rs28864441, which was previously associated with unannotated metabolite levels in an African ancestry population.⁷⁵ Furthermore, by extending our GWAS models to perform two-degree-of-freedom tests to detect both variant and variant-waist circumference-interaction effects on circulating oxylipins, we identified a rare variant specific to admixed American continental ancestry populations near the *CARS2* locus associated with a leukotriene ($m/z = 480.25$, $RT=0.351$). Although this finding requires external replication, particularly considering the low frequency of

the top variant, one recent study demonstrated an anti-inflammatory role for *CARS2* in macrophages and smooth muscle cells,⁷⁶ suggesting a biologically plausible link between *CARS2* and the leukotriene oxylin.

Strengths and Limitations

Utilizing a high-throughput and sensitive LC-MS method for comprehensively profiling fatty acids and oxylin and conducted in a sample of over 11,000 participants, this study is one of the largest GWAS of oxylin and related fatty acids to-date. Moreover, it was conducted in the ancestrally diverse HCHS/SOL cohort, which is important given that most genetic association and metabolomics studies have been historically conducted in European ancestry populations. Indeed, we identified two variants (rs28600890-*ADHA1* and rs183091953-*CARS2*) that are monomorphic in the 1000 Genomes European continental ancestry reference population, but not in admixed American reference populations, thus enabling identification of these signals in the HCHS/SOL cohort.

However, there are also several limitations to this study. First, formal replication of metabolite-genetic locus associations in independent populations is challenging, due to differences in measurement platforms and annotation of metabolites. In particular, we identified multiple significant associations with putative novel oxylin, and therefore formal replication will not be possible until these can be more definitively annotated. However, the following loci that we identified here were also recently associated with many annotated and novel oxylin compounds in the Framingham and ARIC cohorts^{56,57}: *FADS*, *SLCO1B1*, *UGT2B7*, *PKD2L1*, *ACOT6*, *PECR*, *THEM4*, *ADH1A*, and multiple *CYP* loci. This supports our results and suggests that the associations between these loci and oxylin and fatty acids are real signals, though future work will be needed to better characterize the specific oxylin subgroups that are

influenced by each locus. Second, although circulating oxylipin concentrations can become elevated due to secretion or spillover from inflamed tissues,⁷⁷ the actions of oxylipins are likely highly localized and specific to tissue and cell types, which can be challenging to measure in non-experimental settings. Thus, our results may not capture genetic associations that underly tissue-specific oxylipin activity and profiles. However, they still provide insights into systemic oxylipin and fatty acid profiles and the pathways in which they operate.

Conclusions

Our study revealed shared and distinct genetic architecture underlying correlated and modest-to-moderately heritable fatty acids and bioactive oxylipins. The genetic architecture included loci involved in a variety of biological functions, including fatty acid and oxylipin synthesis, exogenous and endogenous substrate transport and metabolism, and inflammation regulation. These results provide insights into the genetic factors influencing circulating bioactive lipids, and motivate future multi-omics work to better characterize these compounds and elucidate their roles in disease pathways.

Acknowledgements: The authors thank the staff and participants of HCHS/SOL for their important contributions. A complete list of staff and investigators has been provided by Sorlie P., et al. in *Ann Epidemiol.* 2010 Aug; 20: 642-649 and is also available on the study website <http://www.csc.unc.edu/hchs/>.

Funding: The Hispanic Community Health Study/Study of Latinos was carried out as a collaborative study supported by contracts from the National Heart, Lung, and Blood Institute (NHLBI) to the University of North Carolina (N01-HC65233), University of Miami (N01-HC65234), Albert Einstein College of Medicine (N01-HC65235), Northwestern University

(N01-HC65236), and San Diego State University (N01-HC65237). The following Institutes/Centers/Offices contribute to the HCHS/SOL through a transfer of funds to the NHLBI: National Center on Minority Health and Health Disparities, the National Institute of Deafness and Other Communications Disorders, the National Institute of Dental and Craniofacial Research, the National Institute of Diabetes and Digestive and Kidney Diseases, the National Institute of Neurological Disorders and Stroke, and the Office of Dietary Supplements. CGD was funded by T32HL129982. CLA was funded by 5R01HL147853-04. KEN was funded by R01HL142302, R01HL151152, R01 DK122503, R01HD057194, R01HG010297, R01HL143885. AGH was funded by R01HL143885. AGL was funded by T32TD091058. RCK was funded by 5R01MD011389-02.

Author Contributions: CGD and CLA designed the study. CGD performed the statistical analysis. MJ, TL, and SC generated the oxylipin data. CGD and CLA wrote the manuscript with input from all authors.

Disclosures: CGD and CLA had full access to the study data and take responsibility for the integrity of the data and accuracy of analyses. All authors have reviewed and approved the final manuscript. Mohit Jain, Tao Long, and Nick Miller are employees and equity holders of Sapient Bioanalytics. None of the other authors had any financial or other conflicts of interest.

Data Availability Statement: All summary statistics will be available upon publication through the NHGRI-EBI GWAS catalog (<https://www.ebi.ac.uk/gwas/>)

Figure Legends:

Figure 1. Spearman correlation matrix of log₂-transformed values of n=81 putative fatty acids and oxylipin metabolites. Positive correlations are displayed in blue, negative correlations are displayed in orange. Metabolites are annotated with the putative general class of compound to which they belong, their measured mass-to-charge ratio (m/z), and their retention time (RT) in the form: metabolite name_m/z_RT. When annotations indicated multiple potential compound mappings due to compounds with similar chemical formulas and mass-to-charge ratios co-eluting at the same time, we have listed all of them, separated by backslashes (e.g., HETE/EET). Metabolites are ordered in the matrix based on these general groupings, with putative free fatty acids first (polyunsaturated fatty acids, monounsaturated fatty acids, saturated fatty acids), followed by oxylipin metabolites.

Figure 2. Estimated proportion of variance due to kinship (heritability) and 95% confidence intervals of n=81 putative fatty acids and oxylipin metabolites. Metabolites are annotated with the putative general class of compound to which they belong, their measured mass-to-charge ratio (m/z), and their retention time (RT) in the form: metabolite name_m/z_RT. Estimates are colored by the putative grouping to which the metabolite belongs: polyunsaturated fatty acid, monounsaturated fatty acid, saturated fatty acid, or the putative fatty precursor from which the putative oxylipin is derived. Putative novel and known oxylipins are grouped in “Unknown precursor fatty acid” grouping.

Figure 3. Summary of unique loci identified in main effects GWAS. (a) Manhattan plot of unique loci identified in main effects GWAS, color coded by the putative grouping to which the metabolite belongs: polyunsaturated fatty acid, monounsaturated fatty acid, saturated fatty acid, or the putative fatty precursor from which the putative oxylipin is derived. Oxylipins annotated as ‘putative novel’ or ‘putative known’ are colored in gray, and precursor fatty acid group is marked as “Unknown precursor fatty acid.” -log₁₀ p-values were capped at 50 to aid visualization of results (maximum -log₁₀pvalue observed was 262, for arachidonic acid associated with *FADS2*). Some annotations included multiple putative compounds which are derived from different fatty acid precursors, so both are listed (e.g., dihomo-gamma-linolenic acid or eicosadienoic acid; arachidonic acid or eicosapentaenoic acid). Black line indicates genome-wide significance threshold ($5E-09/72 = 6.94E-11$); all variants with p-value $>1E-7$ are colored in gray to improve visualization of results. **(b)** Network plot of the 56 unique metabolite-locus associations to illustrate pleiotropic and distinct genetic effects. Genetic loci are in boxed labels, metabolites are colored based on the putative fatty acid or precursor fatty acid grouping from which the oxylipin is derived. Metabolites are numbered (1-40) and these numbers correspond to the “metabolite id number” column in Supplementary Table 2.

References

1. Saltiel, A.R. & Olefsky, J.M. Inflammatory mechanisms linking obesity and metabolic disease. *J Clin Invest* **127**, 1-4 (2017).
2. Hotamisligil, G.S. Inflammation, metaflammation and immunometabolic disorders. *Nature* **542**, 177-185 (2017).
3. Lontchi-Yimagou, E., Sobngwi, E., Matsha, T.E. & Kengne, A.P. Diabetes mellitus and inflammation. *Curr Diab Rep* **13**, 435-44 (2013).
4. Minihane, A.M. *et al.* Low-grade inflammation, diet composition and health: current research evidence and its translation. *Br J Nutr* **114**, 999-1012 (2015).
5. Pickup, J.C. & Crook, M.A. Is type II diabetes mellitus a disease of the innate immune system? *Diabetologia* **41**, 1241-8 (1998).
6. Stoughton, R.B. Mediators of inflammation. *Arch Dermatol* **93**, 601-7 (1966).
7. Quaranta, A., Revol-Cavalier, J. & Wheelock, C.E. The octadecanoids: an emerging class of lipid mediators. *Biochem Soc Trans* **50**, 1569-1582 (2022).
8. Dennis, E.A. & Norris, P.C. Eicosanoid storm in infection and inflammation. *Nat Rev Immunol* **15**, 511-23 (2015).
9. Buczynski, M.W., Dumlao, D.S. & Dennis, E.A. Thematic Review Series: Proteomics. An integrated omics analysis of eicosanoid biology. *J Lipid Res* **50**, 1015-38 (2009).
10. Guimaraes, R.C., Goncalves, T.T. & Leiria, L.O. Exploiting oxidized lipids and the lipid-binding GPCRs against cardiometabolic diseases. *Br J Pharmacol* **178**, 531-549 (2021).
11. Newgard, C.B. Metabolomics and Metabolic Diseases: Where Do We Stand? *Cell Metab* **25**, 43-56 (2017).
12. Feofanova, E.V. *et al.* A Genome-wide Association Study Discovers 46 Loci of the Human Metabolome in the Hispanic Community Health Study/Study of Latinos. *Am J Hum Genet* **107**, 849-863 (2020).
13. Menni, C. *et al.* Targeted metabolomics profiles are strongly correlated with nutritional patterns in women. *Metabolomics* **9**, 506-514 (2013).
14. Lee, J.M., Lee, H., Kang, S. & Park, W.J. Fatty Acid Desaturases, Polyunsaturated Fatty Acid Regulation, and Biotechnological Advances. *Nutrients* **8**(2016).
15. Guan, W. *et al.* 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 (2014).
16. Lemaitre, R.N. *et al.* 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 (2011).
17. Tanaka, T. *et al.* Genome-wide association study of plasma polyunsaturated fatty acids in the InCHIANTI Study. *PLoS Genet* **5**, e1000338 (2009).
18. Ameur, A. *et al.* Genetic adaptation of fatty-acid metabolism: a human-specific haplotype increasing the biosynthesis of long-chain omega-3 and omega-6 fatty acids. *Am J Hum Genet* **90**, 809-20 (2012).
19. Shin, S.Y. *et al.* An atlas of genetic influences on human blood metabolites. *Nat Genet* **46**, 543-550 (2014).

20. Alseekh, S. *et al.* Mass spectrometry-based metabolomics: a guide for annotation, quantification and best reporting practices. *Nat Methods* **18**, 747-756 (2021).
21. Watrous, J.D. *et al.* Directed Non-targeted Mass Spectrometry and Chemical Networking for Discovery of Eicosanoids and Related Oxylipins. *Cell Chem Biol* **26**, 433-442 e4 (2019).
22. Villar, J. *et al.* Association between fetal abdominal growth trajectories, maternal metabolite signatures early in pregnancy, and childhood growth and adiposity: prospective observational multinational INTERBIO-21st fetal study. *Lancet Diabetes Endocrinol* **10**, 710-719 (2022).
23. de Heredia, F.P., Gomez-Martinez, S. & Marcos, A. Obesity, inflammation and the immune system. *Proc Nutr Soc* **71**, 332-8 (2012).
24. Festa, A. *et al.* The relation of body fat mass and distribution to markers of chronic inflammation. *Int J Obes Relat Metab Disord* **25**, 1407-15 (2001).
25. Hermsdorff, H.H., Zulet, M.A., Puchau, B. & Martinez, J.A. Central adiposity rather than total adiposity measurements are specifically involved in the inflammatory status from healthy young adults. *Inflammation* **34**, 161-70 (2011).
26. Nishida, M., Moriyama, T., Sugita, Y. & Yamauchi-Takahara, K. Abdominal obesity exhibits distinct effect on inflammatory and anti-inflammatory proteins in apparently healthy Japanese men. *Cardiovasc Diabetol* **6**, 27 (2007).
27. Hardwick, J.P. *et al.* Eicosanoids in metabolic syndrome. *Adv Pharmacol* **66**, 157-266 (2013).
28. Graff, M. *et al.* Genome-wide physical activity interactions in adiposity - A meta-analysis of 200,452 adults. *PLoS Genet* **13**, e1006528 (2017).
29. Justice, A.E. *et al.* Genome-wide meta-analysis of 241,258 adults accounting for smoking behaviour identifies novel loci for obesity traits. *Nat Commun* **8**, 14977 (2017).
30. Sorlie, P.D. *et al.* Design and implementation of the Hispanic Community Health Study/Study of Latinos. *Ann Epidemiol* **20**, 629-41 (2010).
31. Lavange, L.M. *et al.* Sample design and cohort selection in the Hispanic Community Health Study/Study of Latinos. *Ann Epidemiol* **20**, 642-9 (2010).
32. van den Berg, R.A., Hoefsloot, H.C., Westerhuis, J.A., Smilde, A.K. & van der Werf, M.J. Centering, scaling, and transformations: improving the biological information content of metabolomics data. *BMC Genomics* **7**, 142 (2006).
33. Bien, S.A. *et al.* Strategies for Enriching Variant Coverage in Candidate Disease Loci on a Multiethnic Genotyping Array. *PLoS One* **11**, e0167758 (2016).
34. Caruso, J.C. & Cliff, N. Empirical Size, Coverage, and Power of Confidence Intervals for Spearman's Rho. *Educational and Psychological Measurement* **57**, 637-654 (1997).
35. Sofer, T. Confidence intervals for heritability via Haseman-Elston regression. *Stat Appl Genet Mol Biol* **16**, 259-273 (2017).
36. Conomos, M.P. *et al.* Genetic Diversity and Association Studies in US Hispanic/Latino Populations: Applications in the Hispanic Community Health Study/Study of Latinos. *Am J Hum Genet* **98**, 165-84 (2016).
37. McCaw, Z.R., Lane, J.M., Saxena, R., Redline, S. & Lin, X. Operating characteristics of the rank-based inverse normal transformation for quantitative trait analysis in genome-wide association studies. *Biometrics* **76**, 1262-1272 (2020).

38. Lin, D.Y. *et al.* Genetic association analysis under complex survey sampling: the Hispanic Community Health Study/Study of Latinos. *Am J Hum Genet* **95**, 675-88 (2014).
39. Gao, X., Starmer, J. & Martin, E.R. A multiple testing correction method for genetic association studies using correlated single nucleotide polymorphisms. *Genet Epidemiol* **32**, 361-9 (2008).
40. Winkler, T.W. *et al.* EasyStrata: evaluation and visualization of stratified genome-wide association meta-analysis data. *Bioinformatics* **31**, 259-61 (2015).
41. Shim, H. *et al.* A multivariate genome-wide association analysis of 10 LDL subfractions, and their response to statin treatment, in 1868 Caucasians. *PLoS One* **10**, e0120758 (2015).
42. Teslovich, T.M. *et al.* Biological, clinical and population relevance of 95 loci for blood lipids. *Nature* **466**, 707-13 (2010).
43. Ross, R. *et al.* Waist circumference as a vital sign in clinical practice: a Consensus Statement from the IAS and ICCR Working Group on Visceral Obesity. *Nat Rev Endocrinol* **16**, 177-189 (2020).
44. Zimmet, P., Magliano, D., Matsuzawa, Y., Alberti, G. & Shaw, J. The metabolic syndrome: a global public health problem and a new definition. *J Atheroscler Thromb* **12**, 295-300 (2005).
45. Justice, A.E. *et al.* Genome-wide association study of body fat distribution traits in Hispanics/Latinos from the HCHS/SOL. *Hum Mol Genet* **30**, 2190-2204 (2021).
46. Willer, C.J., Li, Y. & Abecasis, G.R. METAL: fast and efficient meta-analysis of genomewide association scans. *Bioinformatics* **26**, 2190-1 (2010).
47. Zhou, S.F., Liu, J.P. & Chowbay, B. Polymorphism of human cytochrome P450 enzymes and its clinical impact. *Drug Metab Rev* **41**, 89-295 (2009).
48. Cooper-DeHoff, R.M. *et al.* The Clinical Pharmacogenetics Implementation Consortium Guideline for SLCO1B1, ABCG2, and CYP2C9 genotypes and Statin-Associated Musculoskeletal Symptoms. *Clin Pharmacol Ther* **111**, 1007-1021 (2022).
49. Kalliokoski, A. & Niemi, M. Impact of OATP transporters on pharmacokinetics. *Br J Pharmacol* **158**, 693-705 (2009).
50. Koda, Y., Soejima, M. & Kimura, H. The polymorphisms of fucosyltransferases. *Leg Med (Tokyo)* **3**, 2-14 (2001).
51. Anguita-Ruiz, A., Aguilera, C.M. & Gil, A. Genetics of Lactose Intolerance: An Updated Review and Online Interactive World Maps of Phenotype and Genotype Frequencies. *Nutrients* **12**(2020).
52. Rhee, E.P. *et al.* A genome-wide association study of the human metabolome in a community-based cohort. *Cell Metab* **18**, 130-43 (2013).
53. Hagenbeek, F.A. *et al.* Heritability estimates for 361 blood metabolites across 40 genome-wide association studies. *Nat Commun* **11**, 39 (2020).
54. Bermingham, K.M. *et al.* Genetic and environmental influences on serum oxylipins, endocannabinoids, bile acids and steroids. *Prostaglandins Leukot Essent Fatty Acids* **173**, 102338 (2021).
55. Sollis, E. *et al.* The NHGRI-EBI GWAS Catalog: knowledgebase and deposition resource. *Nucleic Acids Res* **51**, D977-D985 (2023).

56. Granados, J.C. *et al.* Regulation of Human Endogenous Metabolites by Drug Transporters and Drug Metabolizing Enzymes: An Analysis of Targeted SNP-Metabolite Associations. *Metabolites* **13**(2023).
57. Rhee, E.P. *et al.* A genome-wide association study identifies 41 loci associated with eicosanoid levels. *Commun Biol* **6**, 792 (2023).
58. Tamai, I. *et al.* Molecular identification and characterization of novel members of the human organic anion transporter (OATP) family. *Biochem Biophys Res Commun* **273**, 251-60 (2000).
59. Kameyama, Y., Yamashita, K., Kobayashi, K., Hosokawa, M. & Chiba, K. Functional characterization of SLCO1B1 (OATP-C) variants, SLCO1B1*5, SLCO1B1*15 and SLCO1B1*15+C1007G, by using transient expression systems of HeLa and HEK293 cells. *Pharmacogenet Genomics* **15**, 513-22 (2005).
60. Huang, M.J. *et al.* Risk factors for severe hyperbilirubinemia in neonates. *Pediatr Res* **56**, 682-9 (2004).
61. Ho, R.H. *et al.* Effect of drug transporter genotypes on pravastatin disposition in European- and African-American participants. *Pharmacogenet Genomics* **17**, 647-56 (2007).
62. Ieiri, I. *et al.* Influence of common variants in the pharmacokinetic genes (OATP-C, UGT1A1, and MRP2) on serum bilirubin levels in healthy subjects. *Hepatol Res* **30**, 91-95 (2004).
63. Kim, W.Y. *et al.* Identification of novel CYP4F2 genetic variants exhibiting decreased catalytic activity in the conversion of arachidonic acid to 20-hydroxyeicosatetraenoic acid (20-HETE). *Prostaglandins Leukot Essent Fatty Acids* **131**, 6-13 (2018).
64. Nebert, D.W. & Russell, D.W. Clinical importance of the cytochromes P450. *Lancet* **360**, 1155-62 (2002).
65. Nebert, D.W., Wikvall, K. & Miller, W.L. Human cytochromes P450 in health and disease. *Philos Trans R Soc Lond B Biol Sci* **368**, 20120431 (2013).
66. GeneCards. UGT2B7. (2023).
67. Picard, N., Ratanasavanh, D., Premaud, A., Le Meur, Y. & Marquet, P. Identification of the UDP-glucuronosyltransferase isoforms involved in mycophenolic acid phase II metabolism. *Drug Metab Dispos* **33**, 139-46 (2005).
68. GeneCards. PECCR.
69. Das, A.K., Uhler, M.D. & Hajra, A.K. Molecular cloning and expression of mammalian peroxisomal trans-2-enoyl-coenzyme A reductase cDNAs. *J Biol Chem* **275**, 24333-40 (2000).
70. Brocker, C., Carpenter, C., Nebert, D.W. & Vasiliou, V. Evolutionary divergence and functions of the human acyl-CoA thioesterase gene (ACOT) family. *Hum Genomics* **4**, 411-20 (2010).
71. GeneCards. FUT2. (2023).
72. Ma, B., Simala-Grant, J.L. & Taylor, D.E. Fucosylation in prokaryotes and eukaryotes. *Glycobiology* **16**, 158R-184R (2006).
73. Rueedi, R. *et al.* Metabomatching: Using genetic association to identify metabolites in proton NMR spectroscopy. *PLoS Comput Biol* **13**, e1005839 (2017).

74. Horn, T. *et al.* Evolutionary aspects of lipoxygenases and genetic diversity of human leukotriene signaling. *Prog Lipid Res* **57**, 13-39 (2015).
75. Tahir, U.A. *et al.* Whole Genome Association Study of the Plasma Metabolome Identifies Metabolites Linked to Cardiometabolic Disease in Black Individuals. *Nat Commun* **13**, 4923 (2022).
76. Dang, A.T. *et al.* A novel anti-inflammatory role links the CARS2 locus to protection from coronary artery disease. *Atherosclerosis* **348**, 8-15 (2022).
77. Quehenberger, O. & Dennis, E.A. The human plasma lipidome. *N Engl J Med* **365**, 1812-23 (2011).

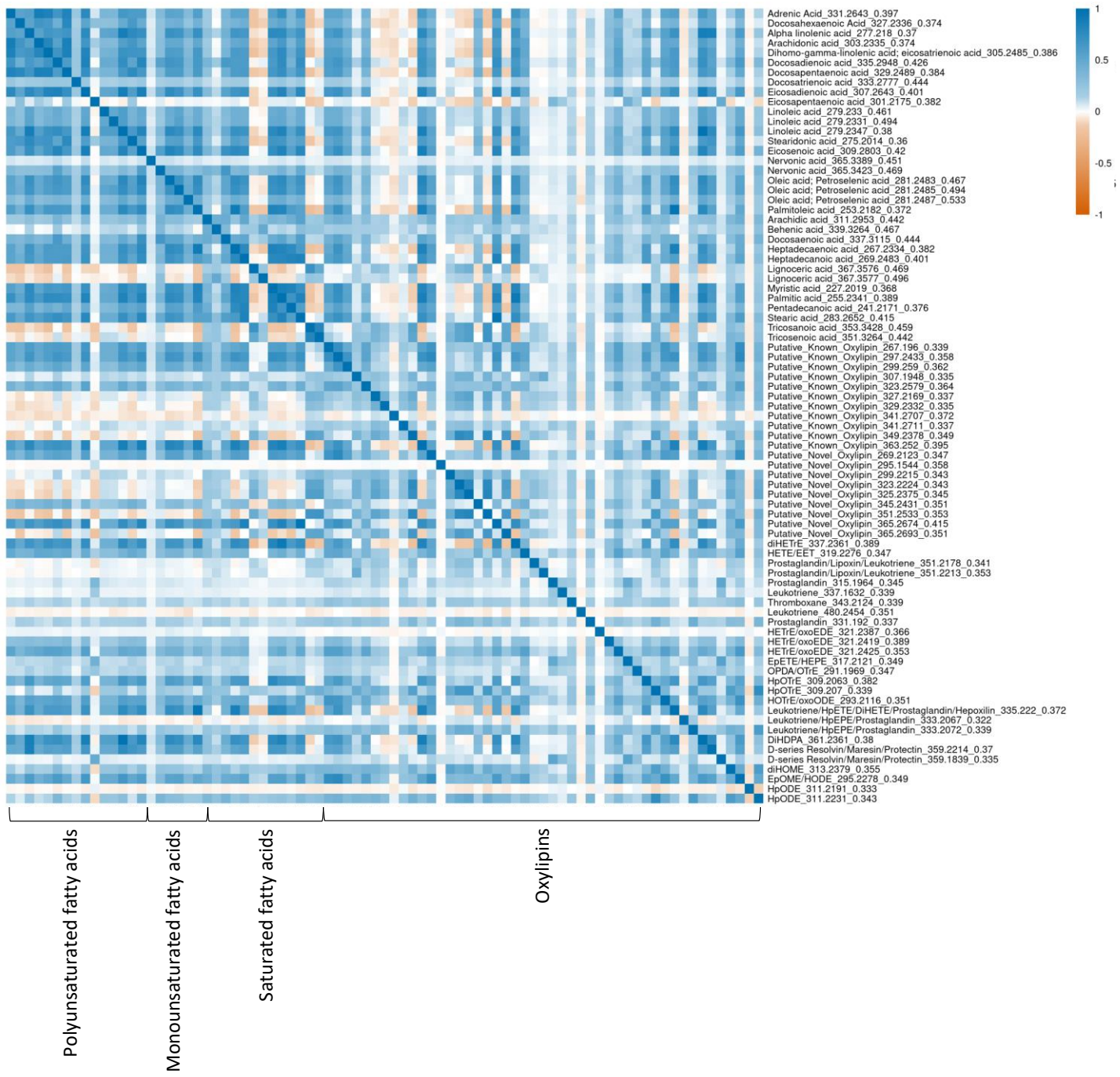


Figure 1. Spearman correlation matrix of log₂-transformed values of n=81 putative fatty acids and oxylipin metabolites.

Positive correlations are displayed in blue, negative correlations are displayed in orange. Metabolites are annotated with the putative general class of compound to which they belong, their measured mass-to-charge ratio (m/z), and their retention time (RT) in the form: metabolite name_m/z_RT. When annotations indicated multiple potential compound mappings due to compounds with similar chemical formulas and mass-to-charge ratios co-eluting at the same time, we have listed all of them, separated by backslashes (e.g., HETE/EET). Metabolites are ordered in the matrix based on these general groupings, with putative free fatty acids first (polyunsaturated fatty acids, monounsaturated fatty acids, saturated fatty acids), followed by oxylipin metabolites.

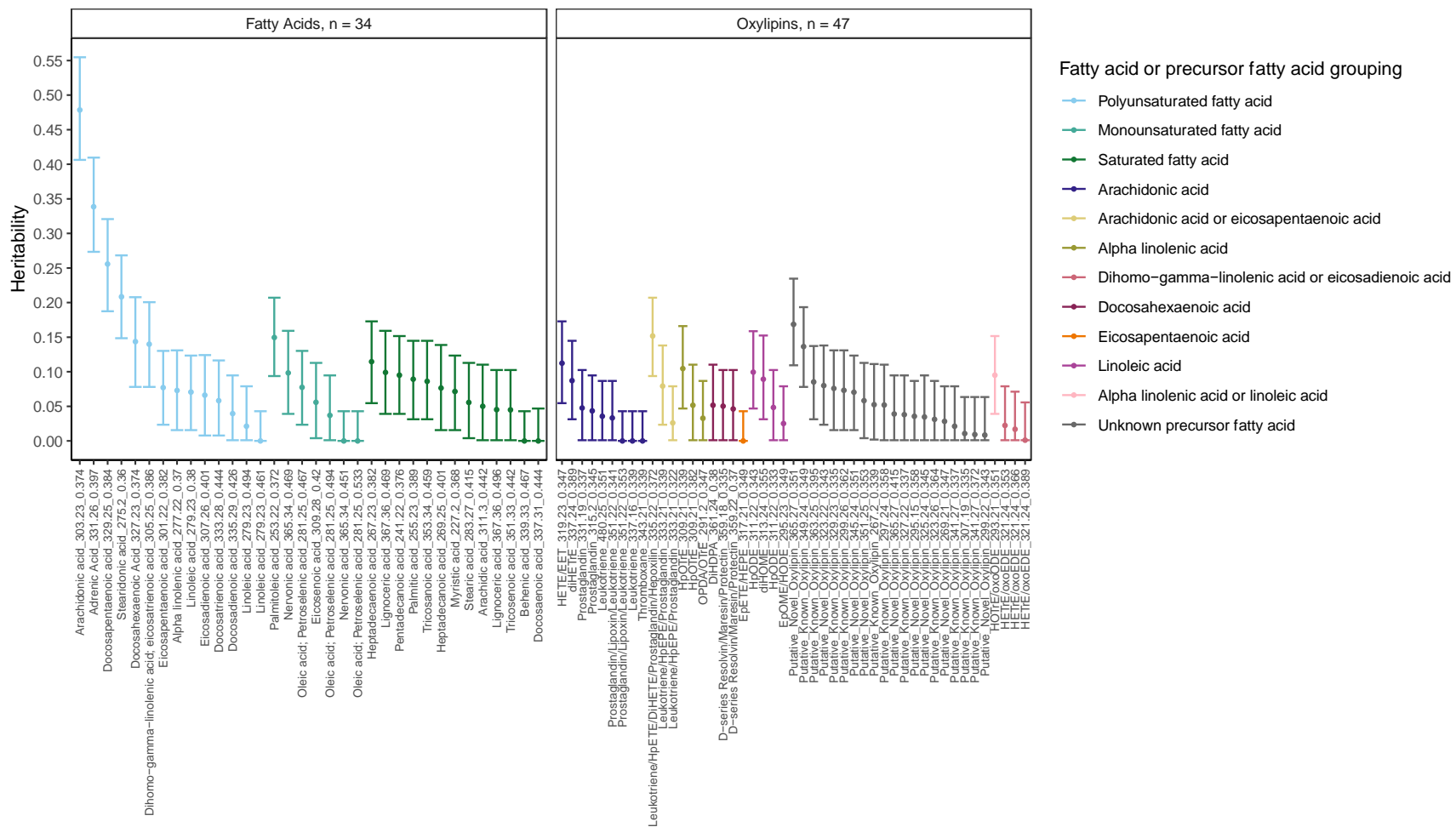
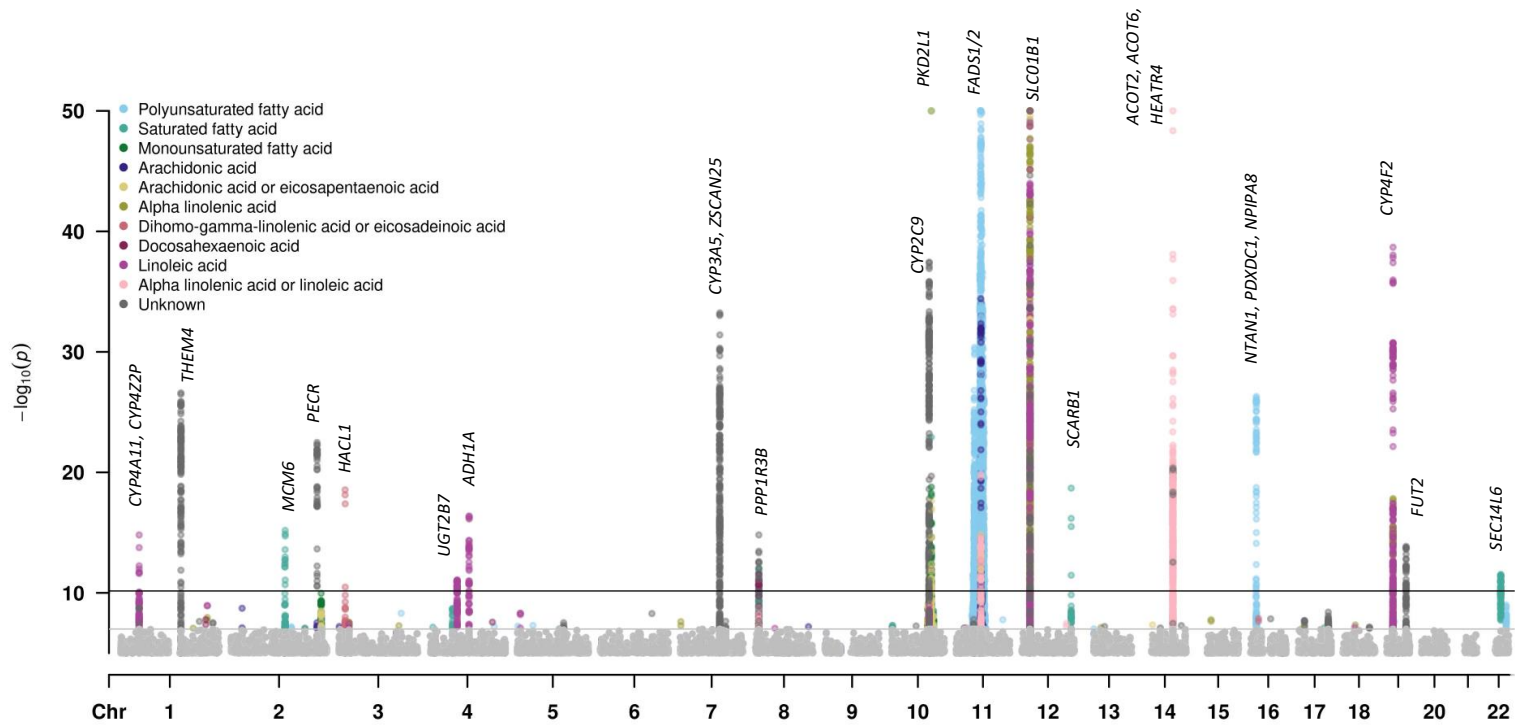


Figure 2. Estimated proportion of variance due to kinship (heritability) and 95% confidence intervals of n=81 putative fatty acids and oxylipin metabolites. Metabolites are annotated with the putative general class of compound to which they belong, their measured mass-to-charge ratio (m/z), and their retention time (RT) in the form: metabolite name_m/z_RT. Estimates are colored by the putative grouping to which the metabolite belongs: polyunsaturated fatty acid, monounsaturated fatty acid, saturated fatty acid, or the putative fatty precursor from which the putative oxylipin is derived. Putative novel and known oxylipins are grouped in “Unknown precursor fatty acid” grouping.

a)



b)

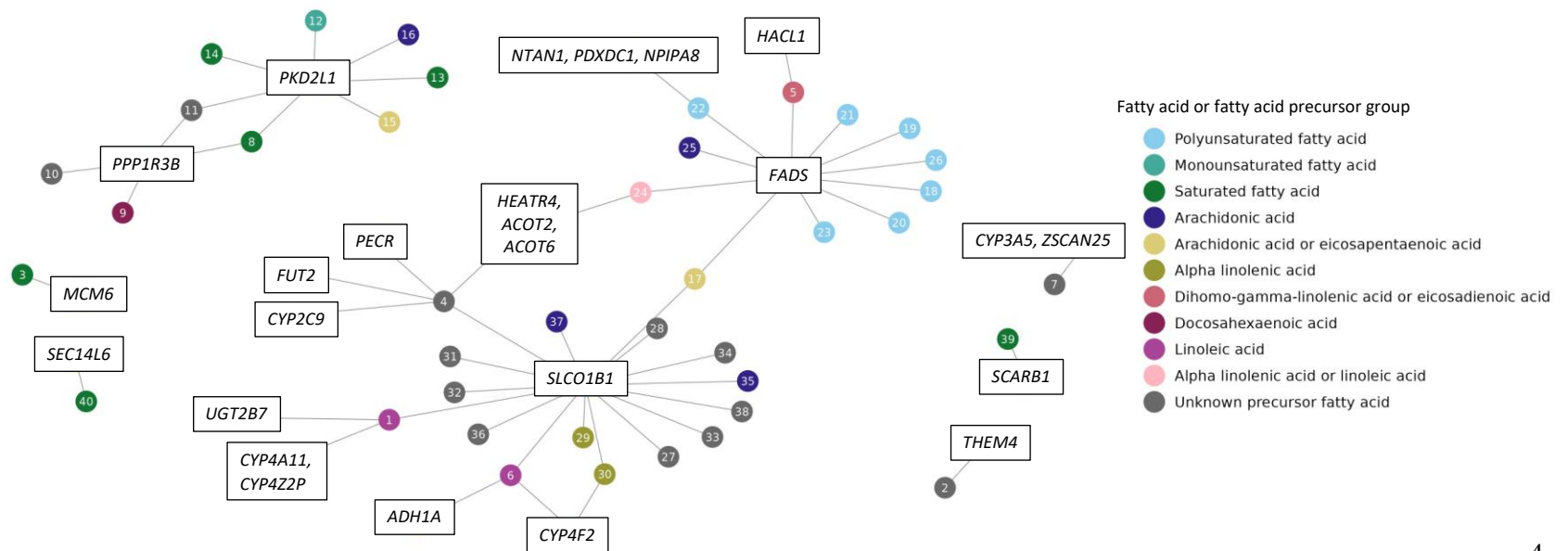


Figure 3. Summary of unique loci identified in main effects GWAS. (a) Manhattan plot of unique loci identified in main effects GWAS, color coded by the putative grouping to which the metabolite belongs: polyunsaturated fatty acid, monounsaturated fatty acid, saturated fatty acid, or the putative fatty precursor from which the putative oxylipin is derived. Oxylipins annotated as ‘putative novel’ or ‘putative known’ are colored in gray, and precursor fatty acid group is marked as “Unknown precursor fatty acid.” $-\log_{10}$ p-values were capped at 50 to aid visualization of results (maximum $-\log_{10}$ pvalue observed was 262, for arachidonic acid associated with *FADS2*). Some annotations included multiple putative compounds which are derived from different fatty acid precursors, so both are listed (e.g., dihomo-gamma-linolenic acid or eicosadienoic acid; arachidonic acid or eicosapentaenoic acid). Black line indicates genome-wide significance threshold ($5E-09/72 = 6.94E-11$); all variants with p-value $>1E-7$ are colored in gray to improve visualization of results. (b) Network plot of the 56 unique metabolite-locus associations to illustrate pleiotropic and distinct genetic effects. Genetic loci are in boxed labels, metabolites are colored based on the putative fatty acid or precursor fatty acid grouping from which the oxylipin is derived. Metabolites are numbered (1-40) and these numbers correspond to the “metabolite id number” column in Supplementary Table 2.