UCSF
UC San Francisco Previously Published Works

Title

Trans-ethnic genome-wide association study of blood metabolites in the Chronic Renal Insufficiency Cohort (CRIC) study

Permalink

https://escholarship.org/uc/item/9cs8d43x

Journal

Kidney International, 101(4)

ISSN

0085-2538

Authors

Rhee, Eugene P
Surapaneni, Aditya
Zheng, Zihe
et al.

Publication Date

2022-04-01

DOI

10.1016/j.kint.2022.01.014

Peer reviewed

eScholarship.org Powered by the California Digital Library
University of California

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

HHS Public Access

Author manuscript

Kidney Int. Author manuscript; available in PMC 2023 April 01.

Published in final edited form as:

Kidney Int. 2022 April ; 101(4): 814–823. doi:10.1016/j.kint.2022.01.014.

Trans-ethnic genome-wide association study of blood metabolites in the chronic renal insufficiency cohort (CRIC) study

Eugene P Rhee, MD[1], **Aditya Surapaneni, PhD**[2],[3], **Zihe Zheng, MBBS**[4], **Linda Zhou, MS**[2],[3], **Diptavo Dutta, PhD**[5], **Dan E Arking, PhD**[6], **Jingning Zhang, MS**[5], **ThuyVy Duong, BS**[6], **Nilanjan Chatterjee, PhD**[5], **Shengyuan Luo, MD**[7], **Pascal Schlosser, PhD**[2],[8], **Rupal Mehta, MD**[9], **Sushrut S Waikar, MD**[10], **Santosh L Saraf, MD**[11], **Tanika N Kelly, PhD**[12], **Lee L Hamm, MD**[13], **Panduranga S Rao, MBBS**[14], **Anna V Mathew, MD**[14], **Chi-yuan Hsu, MD**[15], **Afshin Parsa, MD**[16], **Ramachandran S Vasan, MD**[17],[18], **Paul L Kimmel, MD**[19], **Clary B Clish, PhD**[20], **Josef Coresh, MD PhD**[2],[3], **Harold I Feldman, MD**[4],[21],[22], **Morgan E Grams, MD PhD**[2],[3],[23], **CKD Biomarkers Consortium and the Chronic Renal Insufficiency (CRIC) Study Investigators**

[1]Nephrology Division and Endocrine Unit, Massachusetts General Hospital, Boston, MA

[2]Department of Epidemiology, Johns Hopkins University Bloomberg School of Public Health, Baltimore, MD

[3]Welch Center for Prevention, Epidemiology, and Clinical Research, Johns Hopkins University, Baltimore, MD

[4]Department of Biostatistics, Epidemiology, and Informatics, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA

[5]Department of Biostatistics, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD

[6]McKusick-Nathans Institute, Department of Genetic Medicine, Johns Hopkins University School of Medicine, Baltimore, MD, USA

[7]Department of Medicine, Rush University Medical Center, Chicago, IL

[8]Institute of Genetic Epidemiology, Faculty of Medicine and Medical Center - University of Freiburg, Freiburg, Germany

[9]Division of Nephrology and Hypertension, Northwestern University Feinberg School of Medicine, Chicago, IL

Corresponding author(s): Eugene P Rhee, Thier Research Building, Room 1051, 50 Blossom Street, Boston, MA 02214, (P) 617-643-2888, (F) 617-724-1122, eprhee@partners.org; Morgan E Grams, 1830 E. Monument St, 4th Floor, Suite 416, Baltimore, MD 21287, (P) 443-287-1827, (F) 410-955-0485, mgrams2@jhmi.edu.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Data sharing statement

The CRIC GWAS data are uploaded to dbGAP (Study Accession: phs000524.v1.p1).

DISCLOSURE STATEMENT

None

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

[10]Section of Nephrology, Boston University School of Medicine, Boston Medical Center, Boston, MA

[11]Division of Hematology and Oncology, University of Illinois at Chicago, Chicago, IL

[12]Department of Epidemiology, Tulane University School of Public Health and Tropical Medicine, New Orleans, LA

[13]Department of Medicine, Tulane University School of Medicine, New Orleans, LA

[14]Division of Nephrology, University of Michigan Medical School, Ann Arbor, MI

[15]Division of Nephrology, University of California, San Francisco School of Medicine, San Francisco, CA

[16]Division of Kidney, Urologic, and Hematologic Diseases, National Institute of Diabetes and Digestive and Kidney Diseases, Bethesda, MD

[17]Sections of Preventive Medicine and Epidemiology and Cardiology, Department of Medicine, Boston University School of Medicine, Boston, MA

[18]Department of Epidemiology, Boston University School of Public Health, Boston, MA

[19]Division of Kidney Urologic and Hematologic Diseases, National Institutes of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, MD

[20]Broad Institute of MIT and Harvard, Cambridge, MA

[21]Department of Medicine, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA

[22]Center for Clinical Epidemiology and Biostatistics, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA

[23]Division of Nephrology, Department of Medicine, Johns Hopkins University, Baltimore MD

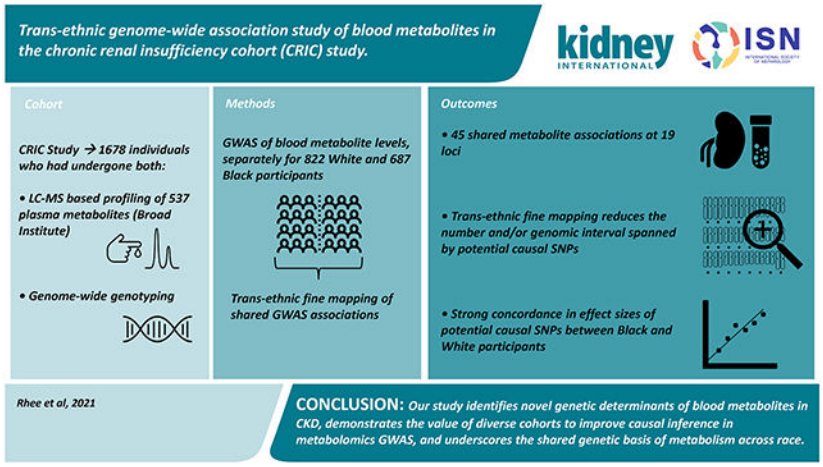
Abstract

Metabolomics genome wide association study (GWAS) help outline the genetic contribution to human metabolism. However, studies to date have focused on relatively healthy, population-based samples of White individuals. Here, we conducted a GWAS of 537 blood metabolites measured in the Chronic Renal Insufficiency Cohort (CRIC) Study, with separate analyses in 822 White and 687 Black study participants. Trans-ethnic meta-analysis was then applied to improve fine-mapping of potential causal variants. Mean estimated glomerular filtration rate was 44.4 and 41.5 mL/min/1.73m^2 in the White and Black participants, respectively. There were 45 significant metabolite associations at 19 loci, including novel associations at *PYROXD2, PHYHD1, FADS1-3, ACOT2, MYRF, FAAH, and LIPC*. The strength of associations were unchanged in models additionally adjusted for estimated glomerular filtration rate and proteinuria, consistent with a direct biochemical effect of gene products on associated metabolites. At several loci, trans-ethnic meta-analysis, which leverages differences in linkage disequilibrium across populations, reduced the number and/or genomic interval spanned by potentially causal single nucleotide polymorphisms compared to fine-mapping in the White participant cohort alone. Across all validated associations, we found strong concordance in effect sizes of the potentially

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

causal single nucleotide polymorphisms between White and Black study participants. Thus, our study identifies novel genetic determinants of blood metabolites in chronic kidney disease, demonstrates the value of diverse cohorts to improve causal inference in metabolomics GWAS, and underscores the shared genetic basis of metabolism across race.

Graphical Abstract



Keywords

metabolomics; GWAS; trans-ethnic meta-analysis

INTRODUCTION

With the application of metabolomic profiling in large cohorts, several independent GWAS have identified more than 100 variants associated with plasma, serum, and urine metabolite levels.[1-11] Many of these loci encode enzymes or transporters directly involved with the given metabolite's synthesis, reabsorption, or secretion, including within the kidney. Because of the direct relationship between gene product and metabolite, many loci have large effect sizes on metabolite levels and the proportion of metabolite variance explained, as compared to GWAS for common diseases.[12] Taken together, these studies have enhanced our understanding of the genetic basis of human metabolism and identified both genes and metabolites relevant to human disease.[13]

One major limitation of the studies conducted to date is that most of the genetic and metabolomics data have been generated in White individuals. Although rare studies have been conducted among Black individuals,[14, 15] no study has conducted GWAS in a racially heterogeneous cohort. This represents a missed opportunity to illustrate concordance in the genetic architecture of human metabolism across diverse groups. Further, the identification of shared signals across different racial groups has the potential to narrow the list of potential causal variants underlying select gene-metabolite associations,[16, 17] an approach that has been applied to racially diverse GWAS of glycemic traits, erythropoietic indices, and estimated glomerular filtration rate (eGFR), among others.[18-20] Finally, the majority of

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

cohorts examined in metabolomics GWAS to date have been relatively healthy, population-based samples. Recent GWAS that focused on urine metabolites in a European chronic kidney disease (CKD) cohort have illuminated loci that are highly informative for renal tubular physiology and for the absorption, distribution, metabolism, and excretion of numerous metabolites, underscoring the value of examining disease-specific cohorts.[9, 10]

The Chronic Renal Insufficiency Cohort (CRIC) Study is a prospective cohort study designed to include a racially diverse group of adults with mild to moderate CKD.[21,22] Approximately half of those recruited had diabetes, and in the first phase of enrollment, were less than 75 years of age, limiting the number of individuals with age-related diminutions of glomerular filtration rate but otherwise non-progressive disease. As a result, the cohort was enriched for individuals at relatively greater risk of subsequent CKD progression and adverse cardiovascular outcomes.[23] Here, we report the results of GWAS of blood metabolites assayed by liquid chromatography-mass spectrometry (LC-MS) in the CRIC Study, with separate analyses in self-identified Black and White study participants. Unlike prior metabolomics GWAS, this analysis synthesizes results from a racially diverse study cohort, enriched for CKD and its metabolic correlates.

METHODS

CRIC Study.

Between 2003-2008, 3939 individuals with mild to moderate kidney disease were recruited at 13 CRIC sites across the US.[21, 22] Study participants were all between the age of 21 and 74 years, with an eGFR of 20-70 ml/min/1.73m^2. Patients with polycystic kidney disease or on active immunosuppressive agents for glomerulonephritis were excluded from the study. A total of 1800 randomly selected participants who attended the year 1 visit underwent blood metabolomic profiling. Of these individuals, 1678 had also undergone genome-wide genotyping, including 687 self-identified Black study participants and 822 self-identified White study participants. All participants provided written informed consent, and the study adhered to the Declaration of Helsinki and was approved by the institutional review board of the Perelman School of Medicine at the University of Pennsylvania.

Clinical characteristics.

Estimated GFR was calculated from serum creatinine and cystatin C using the CRIC Study equation.[24] Urine protein-creatinine ratio was determined from 24-hr urine collection or random spot measures. Diabetes was defined as self-reported use of insulin or oral hypoglycemic medications, fasting blood glucose ≥126 mg/dL or a non-fasting level ≥200 mg/dL, or an HbA1c ≥6.5%. Hypertension was defined as self-reported antihypertensive medication use, systolic blood pressure ≥140 mmHg, or diastolic blood pressure ≥90 mmHg. Cardiovascular disease was self-reported.

Metabolomics.

Detailed methods, including QC characteristics, of the Broad Institute metabolomics methods have been published.[25] In brief, a combination of three LC-MS injections were used to profile metabolites in year 1 visit fasting plasma that had been stored at −80°C.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Metabolomics analyses were conducted 2018-2019. All data were acquired using LC-MS systems comprised of Nexera X2 U-HPLC systems (Shimadzu Scientific Instruments) coupled to Q Exactive/Exactive Plus orbitrap mass spectrometers (Thermo Fisher Scientific). Positively charged polar analytes were measured in 10 μL of plasma via protein precipitation with the addition of nine volumes of 74.9:24.9:0.2 v/v/v acetonitrile/methanol/formic acid, with supernatants injected onto a 150 x 2 mm Atlantis HILIC column (Waters) and MS analyses carried out using electrospray ionization in the positive ion mode using full scan analysis over m/z 70-800. Negatively charged polar analytes were measured in 30 μL of plasma via protein precipitation with the addition of four volumes of 80% methanol, with supernatants injected onto a 150 x 2.0 mm Luna NH2 column (Phenomenex) and MS analyses carried out using electrospray ionization in the negative ion mode using full scan analysis over m/z 60-750. Positively charged lipids were measured in 10 μL of plasma via protein precipitation with the addition of nineteen volumes of isopropanol, with supernatants injected onto a 100 x 2.1 mm ACQUITY BEH C8 column (1.7 μm; Waters) and MS analyses carried out using electrospray ionization in the positive ion mode using full scan analysis over m/z 200-1100. Pairs of pooled plasma quality control samples were analyzed at intervals of approximately 20 study samples, with one sample from each pair used for MS drift correction and the other for evaluation of analytical reproducibility. Raw metabolomics data were processed using TraceFinder (Thermo Fisher Scientific) and Progenesis QI (Nonlinear Dynamics). Identification of metabolite peaks was conducted by matching measured retention times and masses to mixtures of reference metabolites analyzed in each batch and to an internal database of >600 compounds that have been characterized.

Metabolomic data processing.

Drug metabolites and metabolites missing in >50% of samples were excluded from the present analysis, leaving 537 metabolites for evaluation (median % missing, 0%; range 0% to 47%). Missing values were imputed with the lowest observed level, as has been done previously.[26, 27] Metabolites were log(base2)-transformed to normalize their skewed distributions. The residuals of the linear regression of the log-transformed metabolites on age, sex, the first 10 principal components (PCs) of the log-transformed metabolites (to account for any factors related to batch of processing), and the first 10 race-specific genetic PCs (to account for population sub-structure) were used in subsequent GWAS. In supplemental analyses, metabolites were additionally adjusted for eGFR and urine protein-creatinine ratio.

Genotyping and imputation.

Genotyping was conducted in 3641 CRIC participants using the HumanOmni1-Quad arrays (Illumina).[28] Data cleaning was performed following a protocol that included evaluations of per-SNP and per-individual call rates, phenotypic and genotypic sex mismatches, relatedness, and genetic ancestry. Altogether, 99,457 (SNPs with duplicate positions or across-individual call rate <95%) out of 1,050,925 SNPs and 27 samples (across-SNP call rate <95%, sex mismatch, or related) were removed during quality control and data cleaning. We then filtered for minor allele frequency (>5%) and Hardy-Weinberg Equilibrium p-value (>0.00001). Genotype data were then imputed to a common set of SNPs using Trans-Omics

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

for Precision Medicine (TOPMed) (Freeze 5 on GRCh38) as the reference panel,[29, 30] leading to a final set of 9,161,457 markers in White study participants and 15,508,234 markers in Black study participants for GWAS. Mean African ancestry was 0.81 for the Black cohort and 0.01 for the White cohort.

Genome-wide association studies.

Associations were estimated under an additive genetic assumption separately in the White and Black cohorts using Fast Association Tests (FAST) software.[31] Starting with the larger sample (White cohort), the association between metabolite and SNP was considered statistically significant at the Bonferroni adjusted genome-wide threshold of $P < 9.3 \times 10^{-11}$ (5×10^{-8} / 537 metabolites). For each metabolite, we identified the index SNP with the lowest statistically significant P-value within a 1-Mb genomic locus. Associations identified in the White cohort were considered shared across cohorts if there was a corresponding association within 500 kb on either side of the index SNP in the Black cohort at $P < 1.0 \times 10^{-3}$ (0.05 / 48 significant metabolites in the White cohort). Gene mapping and annotation was performed by querying Ensembl Regulatory Build, with significant SNPs assigned to overlapping, upstream, or downstream genes.[32] We estimated the Pearson correlations (r) of metabolites and eGFR and metabolites and log urine protein-creatinine ratio, and for both analyses evaluated whether there was a difference in r between metabolites with validated genetic associations versus all other metabolites by Fisher transformed t-test.

Trans-ethnic fine mapping.

For metabolites with shared genetic associations across White and Black cohorts, we performed fine-mapping separately in White and Black cohorts using SuSiE, assuming that each genetic locus contained one causal variant.[33, 34] For loci with more than one metabolite association, e.g. *NAT8, AGXT2*, and *FADS1-3*, we focused on the most significantly associated metabolite among White study participants. Fine-mapping establishes credible sets of SNPs that have 99% probability of containing a variant with a non-zero causal effect regulating plasma metabolite levels. When the 99% credible set contained more than one SNP in the White cohort, we tried to narrow the results by performing trans-ethnic fine-mapping across the White and Black cohorts using MANTRA,[16] which allowed us to calculate the posterior probability of a variant being included in the credible set across groups. To compare effect sizes across race, we plotted GWAS beta coefficients and coded allele frequencies in the White and Black cohorts for the 99% credible set of SNPs from fine-mapping of the White cohort.

RESULTS

Study sample.

Characteristics of the study sample are shown in Table 1, with 687 Black and 822 White study participants. Mean age in both groups was 58.9 years, with a lower percentage of women in the White subcohort of the CRIC Study. The prevalence of hypertension, diabetes, and cardiovascular disease were higher in the Black subcohort of the CRIC Study, and the mean eGFR was 44.4 and 41.5 mL/min/1.73m^2 in the White and Black groups, respectively.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

GWAS.

In the White cohort, 48 associations achieved Bonferroni adjusted genome-wide significance. Of these, 45 associations—spanning 45 metabolites at 19 different genetic loci—were also significant in the Black cohort (Figure 1, Table 2). Our results replicate several well-established gene-metabolite associations. Select examples include *NAT8*, which encodes N-acetyltransferase 8, and the acetylated amino acids N-acetylornithine, N-alpha-acetylarginine, N6-acetyllysine, and N-acetylphenylalanine; *AGXT2*, which encodes Alanine-glyoxylate aminotransferase 2, and its substrates 2-aminoisobutyric acid, dimethylguanidino valeric acid (DMGV), and symmetric dimethylarginine (SDMA); *ASPG*, which encodes asparaginase and its substrate asparagine; and *FADS1-3*, which encodes Fatty acid desaturases 1-3, and its highly unsaturated lipid products spanning various lipid classes, including PC (phosphatidylcholines), CE (cholesterol esters), and TAG (triacylglycerols).

Previously unreported associations were identified based on review of the published literature and by individual look-up of each association in the NHGRI-EBI Catalog of human GWAS (https://www.ebi.ac.uk/gwas/).[1-11] Based on this review, we highlight 11 previously unreported metabolite associations across seven genetic loci (Table 2): N6-methyllysine and N6,N6-dimethyllysine at *PYROXD2*; guanine and 5-methylcytidine at *PHYHD1*; two diacylglycerols (DAG 38:5, DAG 38:4) at *FADS1-3*; PS plasmalogen 36:3 at *ACOT2*; CE 22:4 at *MYRF;* oleoyl glycine at *FAAH*; and two phosphatidylethanolamines (PE 36:4, PE 38:5) at *LIPC*. Notably, other metabolites (i.e. distinct from the metabolites above) have previously been reported at *PYROXD2, PHYHD1, FADS1-3, MYRF*, and *LIPC*, including several other PEs at the latter locus. No previous metabolite associations have been reported at *FAAH* or *ACOT2*.

Figure 2A shows the 45 metabolites with significant associations validated across the Black and White study participants distributed on the basis of their cross-sectional correlations with eGFR. Figure 2B shows the distribution of correlations for these 45 metabolites with urine protein-creatinine ratio. Compared to all other metabolites, the 45 validated GWAS hits demonstrated a similar distribution of correlations with eGFR and proteinuria. In addition, the P-value of associations for the 45 significant associations were not attenuated in GWAS models additionally adjusted for eGFR and proteinuria (Supplementary Figure S1).

Trans-ethnic meta-analysis.

Next, we conducted trans-ethnic meta-analysis, with the goal of improving the resolution of fine-mapping of potential causal variants. To illustrate the value of this approach in metabolomics GWAS, we began with the White cohort, where the large majority of studies have been conducted to date. For 8 of the 19 loci examined, the White specific analysis yielded a single likely causal SNP, such that there was no opportunity for further improvement through trans-ethnic analysis. For the remaining 11 loci, the White-specific analysis yielded 2 or more potentially causal SNPs (spanning a mean genomic distance of 30,050 bp, range 1472 to 139,855 bp), whereas Black-specific analysis yielded 1 to 128 potentially causal SNPs (spanning a mean genomic distance of 466,777 bp, range 0 to

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

993,254 bp). Trans-ethnic meta-analysis reduced the number of potentially causal SNPs and the genomic interval spanned by these SNPs at seven loci compared to the White specific analysis. As shown, trans-ethnic meta-analysis highlighted between 1 to 70 potential causal SNPs spanning a mean genomic distance of 27,360 bp, with range 0 to 152,544 bp (Table 3, Supplementary Table S1).

Illustrative graphs for glycine at *CPS1* and PE 38:6 at *LIPC* are shown in Figure 3. At *CPS1*, trans-ethnic meta-analysis identified a single SNP rs1047891 as potentially causal in this region for the association with glycine. Notably, this is a coding variant that results in a change in protein sequence, pThr1405Asn, that has been implicated as a cause of neonatal pulmonary hypertension and hyperhomocysteinemia.[35, 36] At *LIPC*, trans-ethnic meta-analysis highlighted the SNPs rs2043085 and rs261291, both non-coding SNPs upstream of the *LIPC* gene.

Effect size comparison.

In addition to providing an opportunity to enhance fine-mapping, our trans-ethnic study design permits the direct comparison of biologic effect sizes at shared SNPs across Black and White cohorts. Focusing on the validated metabolite associations highlighted in Table 2, we identified the plausible causal SNPs across all 45 associations among White study participants using SuSiE and then compared the beta estimate at each SNP between White and Black cohorts. As shown in Figure 4A, we observed strong concordance in the beta estimates of each SNP on corresponding metabolite level across groups. Figure 4B plots the coding allele frequency at each of these same SNPs, highlighting the known heterogeneity in allele frequencies across White and Black cohorts.

DISCUSSION

Since the first published metabolomics GWAS in 2008,[12] the literature has grown in terms of sample size, depth of genome coverage, and number of metabolites measured, significantly expanding our understanding of the genetic determinants of human metabolism. Here, we sought to extend current knowledge by examining a racially heterogeneous cohort enriched for CKD.

Consistent with studies to date, the validated loci in our study encode proteins that are directly involved in biochemical processes, providing a high degree of plausibility for the locus-metabolite associations. More specifically, of the 19 loci underlying the 45 validated associations in our study, 17 encode enzymes, many that act directly on the associated metabolite(s). Some of these associations highlight the high degree of specificity of metabolite measurements: for example, the association of *ACADS*, which encodes the short-chain acyl-CoA dehydrogenase, with the short-chain C4 carnitine and the association of *ACADM*, which encodes the medium-chain acyl-CoA dehydrogenase, with the medium-chain C6 and C8 carnitines. Two of our novel associations were at loci previously not associated with metabolite levels. *FAAH* encodes Fatty acid amide hydrolase, an enzyme that hydrolyses the carbon-nitrogen bond in fatty acid amides and ethanolamines.[37] Whereas prior work applied metabolomics to *FAAH* knock-out mice to identify FAAH substrates,[38] we show how leveraging common variation at the human locus

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

is complementary, nominating oleoyl glycine as a novel substrate. *ACOT2* encodes Acyl-CoA thioesterase 2, an enzyme involved in mitochondrial fatty acid metabolism; however, it is not known to act directly on the associated lipid metabolite PS plasmalogen 36:3.

Other novel associations identified in our study highlighted new metabolites at loci that had previously been linked to other metabolites. These results nevertheless provide new insight. For example, *PYROXD2* encodes Pyridine nucleotide-disulphide oxidoreductase domain 2, a mitochondrial enzyme of undetermined function.[39] Oxidoreductases, enzymes that catalyze the transfer of electrons from one molecule to another, fall into several different categories: those that act on CH-OH groups, CH-CH groups, sulfur groups, heme groups, etc as electron donors. Prior GWAS have identified associations at this locus with blood levels of asymmetric dimethylarginine[5] and urinary levels of trimethylamine,[10, 40] metabolites of interest in CKD research. Our study identified new associations at this locus for N6-methyllysine and N6,N6-dimethyllysine. The common feature across these disparate molecules is the presence of a methylated amine, suggesting a specific role for PYROXD2 oxidoreductase action on CH-NH groups that would lead to the removal of a methyl group from these molecules.

Our findings do not appear to be strongly influenced by CKD. In theory, metabolite GWAS among CRIC study participants could have highlighted metabolites that are normally low abundance in circulation but accumulate with GFR loss, or genes that are upregulated when kidney function declines.[41] However, we did not observe a skew in the density plots of GWAS results compared to all other metabolites in relation to correlation with eGFR or proteinuria. Further, the strength of association for validated metabolites was unaffected by additional adjustment for eGFR and proteinuria. These findings underscore the close proximity of gene and trait for our validated associations—in most cases the associated metabolite interacts directly with the gene products as its metabolic substrate or product. Another theoretical advantage of metabolomics GWAS in the CRIC study is the ability to link metabolites and genetic loci that are associated with CKD pathogenesis. Of 19 loci highlighted by our study, *CPS1* and *NAT8* have previously been associated with eGFR.[42] *CPS1* encodes carbamoyl-phosphate synthase 1, which is involved in the catabolism of its associated metabolite glycine.

Because glycine is required for creatine biosynthesis, the association between *CPS1* and eGFR has been attributed to its effect on creatinine production rather than kidney function, supported by the observation that this locus is associated with GFR estimated by serum creatinine but not cystatin C. By contrast, *NAT8* has been associated with eGFR calculated using creatinine and cystatin C, as well as with CKD status. We recently conducted a targeted analysis of *NAT8* genetic variants in the African American Study of Kidney Disease and Hypertension, Atherosclerosis Risk in Communities study, and BioMe and found significant associations between the *NAT8* locus and 14 acetylated amino acids, five of which were also associated with kidney failure.[43] Our current results reinforce the key biologic impact of *NAT8*, a gene highly expressed in the kidney proximal tubule, on amino acid acetylation and motivate further inquiry into its role in CKD pathogenesis.[15]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

The ability to conduct trans-ethnic analyses is a major strength of our study. Local linkage disequilibrium (LD) structure is a key obstacle to identifying causal variants that underlie select associations. Increasing sample size in the same population does not circumvent this challenge. By contrast, leveraging differences in local LD structure between populations enriched for European and African ancestry can enhance the signal at the causal variant because SNPs strongly linked to the causal variant in White individuals may not be strongly linked in Black individuals, or vice versa.[16, 17] Because of relatively low average LD, cohorts enriched for African ancestry may offer a particular advantage in these analyses. As noted, we began with the White cohort—where most studies have been conducted to date—and focused on the loci where the White specific analysis identified two or more potential causal SNPs. At seven of these eleven loci, we found that trans-ethnic meta-analysis improved the resolution of fine-mapping. Whereas one of the illustrative examples confirms a known coding variant in *CPS1*, the other SNPs highlighted by trans-ethnic meta-analysis are non-coding, and we acknowledge further functional analysis would be required to establish causality. Nevertheless, our results build on the existing literature that support the value of racially diverse cohorts to improve causal inference in GWAS.[18-20]

In addition to performing an analysis across the White and Black subcohorts of the CRIC Study, we compared the effect sizes of shared SNPs at all of the validated associations identified in our study. In our view, a metabolomics GWAS is particularly well suited for this analysis for two reasons. First and as already noted, the close proximity of gene and trait translates into larger effect sizes. This proximity reduces the possibility that differential environmental interactions could drive variability in allelic effects between groups. Second, the metabolites and loci of interest span a range of biochemical pathways, providing a broader base of comparison across race than would be possible with a single disease or phenotype GWAS. Across all 45 validated associations, our results demonstrate that each SNP has nearly the same biologic effect, i.e. impacts circulating blood metabolite levels to the same extent, regardless of race. This does not preclude the possibility that differences in allele frequencies, in addition to differences in environment, could contribute to variation in the blood metabolome across populations. Further, we note that our analysis across racial groups is unable to confirm ancestry specific signals.

In sum, expanding metabolomics GWAS to consider disease-specific and racially diverse populations yields new insights. Although modest in size relative to other recent metabolomics GWAS, our analyses highlight the particular advantage of diverse study cohorts for fine mapping. More work is required to verify causality at select SNPs highlighted by our trans-ethnic meta-analysis, including metabolomic analysis of tissue to understand how genetic variation impacts metabolism within the kidney.[44] Future studies should also further expand the racial diversity of our study cohorts and further integrate omics data in CRIC beyond the genome and metabolome, i.e. to include the epigenome and proteome.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

ACKNOWLEDGMENTS

Alan S. Go, MD; James P. Lash, MD; Robert G. Nelson, MD, PhD, MS; Mahboob Rahman, MD; Vallabh O Shah, PhD, MS; Raymond R. Townsend, MD; Mark L. Unruh, MD, MS

The opinions presented do not necessarily represent those of the National Institute of Diabetes Digestive and Kidney Diseases, the National Institutes of Health, the Department of Health and Human Services or the government of the United States.

Sources of support:

This study was supported by the CKD Biomarkers Consortium, including U01 DK106981 (PI: Dr. Rhee) and U01 DK085689 (PI: Dr. Coresh); additional funding was from R01DK108803 and R01DK124399 (PI: Dr. Grams). Funding for the CRIC Study was obtained under a cooperative agreement from National Institute of Diabetes and Digestive and Kidney Diseases (U01DK060990, U01DK060984, U01DK061022, U01DK061021, U01DK061028, U01DK060980, U01DK060963, U01DK060902 and U24DK060990). In addition, this work was supported in part by: the Perelman School of Medicine at the University of Pennsylvania Clinical and Translational Science Award NIH/NCATS UL1TR000003, Johns Hopkins University UL1 TR-000424, University of Maryland GCRC M01 RR-16500, Clinical and Translational Science Collaborative of Cleveland, UL1TR000439 from the National Center for Advancing Translational Sciences (NCATS) component of the National Institutes of Health and NIH roadmap for Medical Research, Michigan Institute for Clinical and Health Research (MICHR) UL1TR000433, University of Illinois at Chicago CTSA UL1RR029879, Tulane COBRE for Clinical and Translational Research in Cardiometabolic Diseases P20 GM109036, Kaiser Permanente NIH/NCRR UCSF-CTSI UL1 RR-024131, Department of Internal Medicine, University of New Mexico School of Medicine Albuquerque, NM R01DK119199.

REFERENCES

1. Illig T, Gieger C, Zhai G, et al. A genome-wide perspective of genetic variation in human metabolism. Nat Genet 2010; 42: 137–141. [PubMed: 20037589]
2. Suhre K, Shin S-Y, Petersen A-K, et al. Human metabolic individuality in biomedical and pharmaceutical research. Nature 2011; 477: 54–60. [PubMed: 21886157]
3. Suhre K, Wallaschofski H, Raffler J, et al. A genome-wide association study of metabolic traits in human urine. Nat Genet 2011; 43: 565–569. [PubMed: 21572414]
4. Kettunen J, Tukiainen T, Sarin A-P, et al. Genome-wide association study identifies multiple loci influencing human serum metabolite levels. Nat Genet 2012; 44: 269–76. [PubMed: 22286219]
5. Rhee EP, Ho JE, Chen MH, et al. A genome-wide association study of the human metabolome in a community-based cohort. Cell Metab 2013; 18: 130–143. [PubMed: 23823483]
6. Shin SY, Fauman EB, Petersen AK, et al. An atlas of genetic influences on human blood metabolites. Nat Genet 2014; 46: 543–550. [PubMed: 24816252]
7. Rhee EP, Yang Q, Yu B, et al. An exome array study of the plasma metabolome. Nat Commun 2016; 7: 12360. [PubMed: 27453504]
8. Long T, Hicks M, Yu HC, et al. Whole-genome sequencing identifies common-to-rare variants associated with human blood metabolites. Nat Genet 2017; 49: 568–578. [PubMed: 28263315]
9. Li Y, Sekula P, Wuttke M, et al. Genome-Wide Association Studies of Metabolites in Patients with CKD Identify Multiple Loci and Illuminate Tubular Transport Mechanisms. J Am Soc Nephrol 2018; 29: 1513–1524. [PubMed: 29545352]
10. Schlosser P, Li Y, Sekula P, et al. Genetic studies of urinary metabolites illuminate mechanisms of detoxification and excretion in humans. Nat Genet 2020; 52: 167–176. [PubMed: 31959995]
11. Lotta LA, Pietzner M, Stewart ID, et al. A cross-platform approach identifies genetic regulators of human metabolism and health. Nat Genet 2021; 53: 54–64. [PubMed: 33414548]
12. Gieger C, Geistlinger L, Altmaier E, et al. Genetics meets metabolomics: a genome-wide association study of metabolite profiles in human serum. PLoS Genet 2008; 4: e1000282. [PubMed: 19043545]
13. Suhre K, Gieger C. Genetic variation in metabolic phenotypes: study designs and applications. Nat Rev Genet 2012; 13: 759–769. [PubMed: 23032255]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

14. Yu B, Zheng Y, Alexander D, et al. Genetic determinants influencing human serum metabolome among African Americans. PLoS Genet 2014; 10: e1004212. [PubMed: 24625756]
15. Luo S, Feofanova EV, Tin A, et al. Genome-wide association study of serum metabolites in the African American Study of Kidney Disease and Hypertension. Kidney Int 2021; 100: 430–439. [PubMed: 33838163]
16. Morris AP. Transethnic meta-analysis of genomewide association studies. Genet Epidemiol 2011; 35: 809–822. [PubMed: 22125221]
17. Zaitlen N, Pasaniuc B, Gur T, et al. Leveraging genetic variability across populations for the identification of causal variants. Am J Hum Genet 2010; 86: 23–33. [PubMed: 20085711]
18. Liu CT, Raghavan S, Maruthur N, et al. Trans-ethnic Meta-analysis and Functional Annotation Illuminates the Genetic Architecture of Fasting Glucose and Insulin. Am J Hum Genet 2016; 99: 56–75. [PubMed: 27321945]
19. Morris AP, Le TH, Wu H, et al. Trans-ethnic kidney function association study reveals putative causal genes and effects on kidney-specific disease aetiologies. Nat Commun 2019; 10: 29. [PubMed: 30604766]
20. van Rooij FJA, Qayyum R, Smith AV, et al. Genome-wide Trans-ethnic Meta-analysis Identifies Seven Genetic Loci Influencing Erythrocyte Traits and a Role for RBPMS in Erythropoiesis. Am J Hum Genet 2017; 100: 51–63. [PubMed: 28017375]
21. Feldman HI, Appel LJ, Chertow GM, et al. The Chronic Renal Insufficiency Cohort (CRIC) Study: Design and Methods. J Am Soc Nephrol 2003; 14: S148–153. [PubMed: 12819321]
22. Lash JP, Go AS, Appel LJ, et al. Chronic Renal Insufficiency Cohort (CRIC) Study: baseline characteristics and associations with kidney function. Clin J Am Soc Nephrol 2009; 4: 1302–1311. [PubMed: 19541818]
23. Denker M, Boyle S, Anderson AH, et al. Chronic Renal Insufficiency Cohort Study (CRIC): Overview and Summary of Selected Findings. Clin J Am Soc Nephrol 2015; 10: 2073–2083. [PubMed: 26265715]
24. Anderson AH, Yang W, Hsu CY, et al. Estimating GFR among participants in the Chronic Renal Insufficiency Cohort (CRIC) Study. Am J Kidney Dis 2012; 60: 250–261. [PubMed: 22658574]
25. Rhee EP, Waikar SS, Rebholz CM, et al. Variability of Two Metabolomic Platforms in CKD. Clin J Am Soc Nephrol 2019; 14: 40–48. [PubMed: 30573658]
26. Yu B, Heiss G, Alexander D, et al. Associations Between the Serum Metabolome and All-Cause Mortality Among African Americans in the Atherosclerosis Risk in Communities (ARIC) Study. Am J Epidemiol 2016; 183: 650–656. [PubMed: 26956554]
27. Luo S, Coresh J, Tin A, et al. Serum Metabolomic Alterations Associated with Proteinuria in CKD. Clin J Am Soc Nephrol 2019; 14: 342–353. [PubMed: 30733224]
28. Parsa A, Kanetsky PA, Xiao R, et al. Genome-Wide Association of CKD Progression: The Chronic Renal Insufficiency Cohort Study. J Am Soc Nephrol 2017; 28: 923–934. [PubMed: 27729571]
29. Kowalski MH, Qian H, Hou Z, et al. Use of >100,000 NHLBI Trans-Omics for Precision Medicine (TOPMed) Consortium whole genome sequences improves imputation quality and detection of rare variant associations in admixed African and Hispanic/Latino populations. PLoS Genet 2019; 15: e1008500. [PubMed: 31869403]
30. Naj AC. Genotype Imputation in Genome-Wide Association Studies. Curr Protoc Hum Genet. 2019; 102: e84. [PubMed: 31216114]
31. Chanda P, Huang H, Arking DE, et al. Fast association tests for genes with FAST. PLoS One 2013; 8: e68585. [PubMed: 23935874]
32. Zerbino DR, Wilder SP, Johnson N, et al. The ensembl regulatory build. Genome Biol 2015; 16: 56. [PubMed: 25887522]
33. Schaid DJ, Chen W, Larson NB. From genome-wide associations to candidate causal variants by statistical fine-mapping. Nat Rev Genet 2018; 19: 491–504. [PubMed: 29844615]
34. Wang G, Sarkar A, Carbonetto P, Stephens M. A simple new approach to variable selection in regression, with application to genetic fine mapping. J R Stat Soc Series B Stat Methodol 2020; 82: 1273–1300

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

35. Pearson DL, Dawling S, Walsh WF, et al. Neonatal pulmonary hypertension--urea-cycle intermediates, nitric oxide production, and carbamoyl-phosphate synthetase function. N Engl J Med 2001; 344: 1832–1838. [PubMed: 11407344]

36. Lange LA, Croteau-Chonka DC, Marvelle AF, et al. Genome-wide association study of homocysteine levels in Filipinos provides evidence for CPS1 in women and a stronger MTHFR effect in young adults. Hum Mol Genet 2010; 19: 2050–2058. [PubMed: 20154341]

37. Giang DK, Cravatt BF. Molecular characterization of human and mouse fatty acid amide hydrolases. Proc Natl Acad Sci U S A 1997; 94: 2238–2242. [PubMed: 9122178]

38. Saghatelian A, Trauger SA, Want EJ, et al. Assignment of endogenous substrates to enzymes by global metabolite profiling. Biochemistry 2004; 43: 14332–14339. [PubMed: 15533037]

39. Wang T, Xie X, Liu H, et al. Pyridine nucleotide-disulphide oxidoreductase domain 2 (PYROXD2): Role in mitochondrial function. Mitochondrion 2019; 47: 114–124. [PubMed: 31170524]

40. Rueedi R, Ledda M, Nicholls AW, et al. Genome-wide association study of metabolic traits reveals novel gene-metabolite-disease links. PLoS Genet 2014; 10: e1004132. [PubMed: 24586186]

41. Kottgen A, Raffler J, Sekula P, et al. Genome-Wide Association Studies of Metabolite Concentrations (mGWAS): Relevance for Nephrology. Semin Nephrol 2018; 38: 151–174. [PubMed: 29602398]

42. Kottgen A, Pattaro C, Boger CA, et al. New loci associated with kidney function and chronic kidney disease. Nat Genet 2010; 42: 376–384. [PubMed: 20383146]

43. Luo S, Surapaneni A, Zheng Z, et al. NAT8 Variants, N-Acetylated Amino Acids, and Progression of CKD. Clin J Am Soc Nephrol 2020; 16: 37–47. [PubMed: 33380473]

44. Hasegawa S, Tanaka T, Saito T, et al. The oral hypoxia-inducible factor prolyl hydroxylase inhibitor enarodustat counteracts alterations in renal energy metabolism in the early stages of diabetic kidney disease. Kidney Int 2020; 97: 934–950. [PubMed: 32171449]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

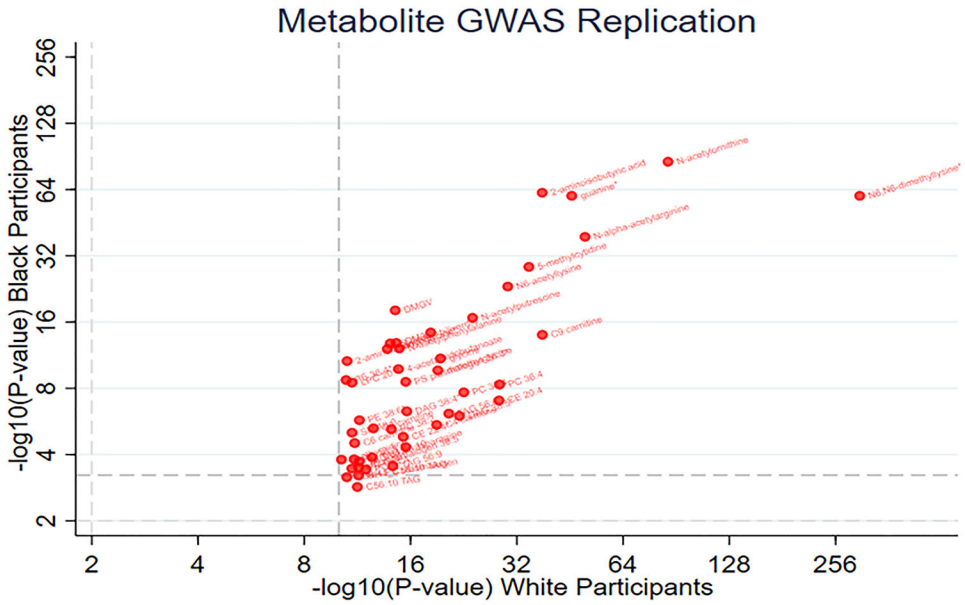


Figure 1.

Metabolite GWAS in Black and White study participants in the CRIC Study. In the plot, each circle represents a distinct metabolite plotted along the x-axis for its lowest GWAS P-value of association in the White cohort (index SNP), and then along the y-axis for the lowest P-value of association in the Black cohort within 500 kb on either side of the index SNP. For validated associations, the size of the circle correlates with the distance between the White and Black index SNPs.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

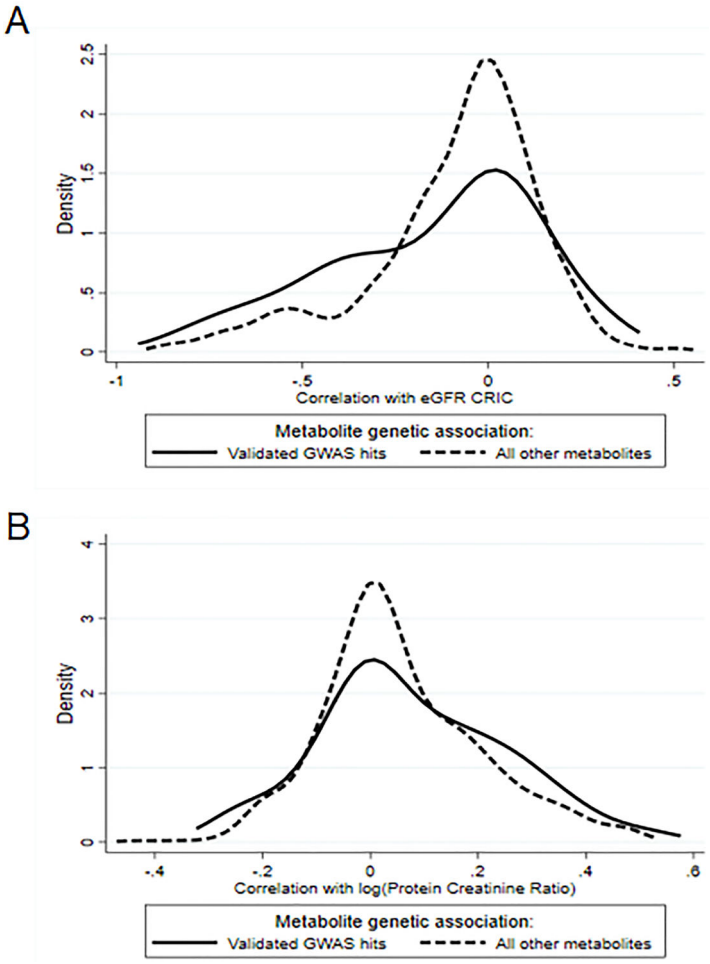


Figure 2.

Distribution of Pearson correlations of metabolites with CKD measures. Kernel density plot of metabolite associations with eGFR (A) and proteinuria (B), shown for metabolites with validated GWAS associations (solid line) and all other metabolites (dotted line).

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

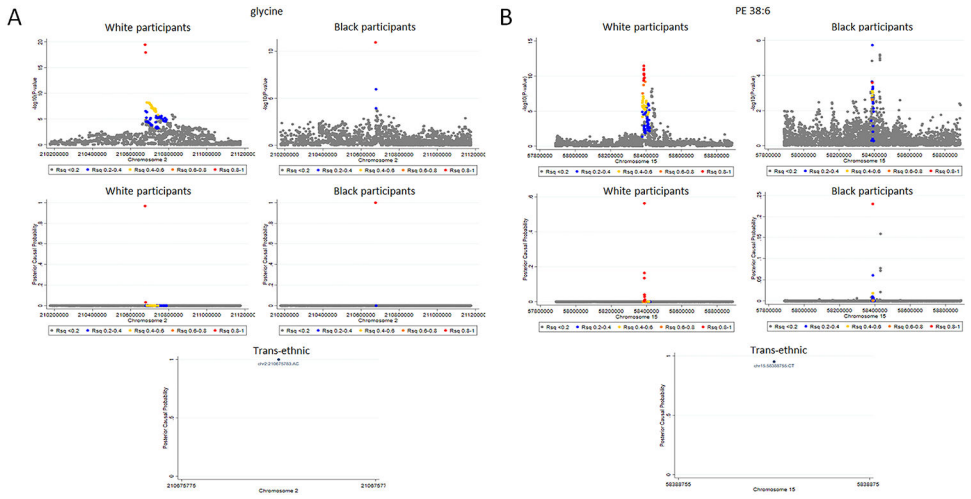


Figure 3.

Trans-ethnic fine mapping at *CPS1* and *LIPC*. For the glycine association at *CPS1* (A) and the PE 38:6 association at *LIPC* (B): regional plots for White and Black cohorts showing P-values and SNP correlations (top), credible sets of potential causal SNPs within White and Black cohorts (middle), and output from trans-ethnic meta-analysis (bottom).

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

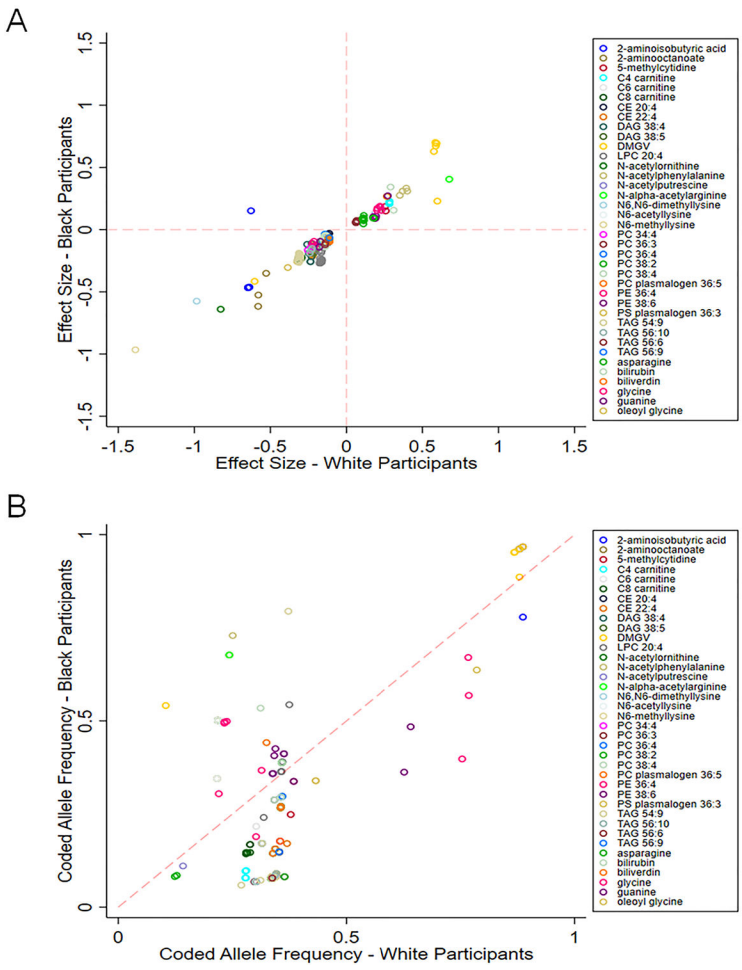


Figure 4.

Comparison of validated metabolite loci between White and Black cohorts. Scatter plots of effect size of coded allele on metabolite levels (A) and coded allele frequencies (B) for the 99% credible set of SNPs from the White cohort fine-mapping.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 1.

CRIC study sample for metabolomics GWAS

	Black	White
Sample size	687	822
Age, years	58.9 (10.2)	58.9 (11.0)
Female	348 (50.7%)	324 (39.4%)
eGFR, ml/min/1.73m^2	41.5 (17.1)	44.4 (16.8)
Urine Protein/Creatinine Ratio, g/g (IQR)	0.2 (0.1 - 0.9)	0.1 (0.1 - 0.5)
Total Cholesterol (mg/dL)	182.5 (43.2)	180.1 (42.8)
High-density Lipoprotein (mg/dL)	49.8 (16.6)	47.4 (15.1)
Low-density Lipoprotein (mg/dL)	102.5 (36.0)	96.3 (32.8)
Triglycerides (mg/dL)	138.5 (99.4)	167.1 (114.7)
Statin use	387 (56.6%)	499 (61.0%)
Hypertension	657 (95.6%)	683 (83.3%)
Diabetes	359 (52.3%)	377 (45.9%)
Cardiovascular disease	281 (40.9%)	278 (33.8%)

Data represents means (standard deviation) unless otherwise noted eGFR, estimated glomerular filtration rate

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Kidney Int. Author manuscript; available in PMC 2023 April 01.

Table 2.

Significant GWAS results validated across White and Black study participants

		White							Black						
Metabolite	**Locus**	**rsid**	**SNP**	**Pos**	**CAF**	**Beta**	**SE**	**P**	**rsid**	**SNP**	**Pos**	**CAF**	**Beta**	**SE**	**P**
N6-methyllysine *	*PYROXD2*	rs5787280	ch10:98384252:T/TGAC	intron	0.37	−1.41	0.04	0.0E+00	rs6584191	ch10:98392298:C/G	intron	0.79	−0.97	0.05	3.9E-81
N6,N6-dimethyllysine *	*PYROXD2*	rs5787280	ch10:98384252:T/TGAC	intron	0.37	−1.00	0.03	3.3E-300	rs7905265	ch10:98383436:C/G	down	0.79	−0.58	0.04	1.2E-60
N-acetylornithine	*NAT8*	rs10168931	ch2:73650092:G/A	up	0.21	−0.88	0.04	1.7E-86	rs13431529	ch2:73648914:G/C	up	0.51	−0.73	0.04	2.2E-86
N-alpha-acetylarginine	*NAT8*	rs13410232	ch2:73644562:A/T	intron	0.21	0.76	0.05	1.2E-50	rs7596773	ch2:73618873:C/T	down	0.52	0.59	0.04	9.3E-40
guanine *	*PHYHD1*	**rs55758160**	ch9:128922736:A/G	intron	0.39	0.28	0.02	1.7E-46	**rs55758160**	ch9:128922736:A/G	intron	0.49	0.38	0.02	1.1E-60
C9 carnitine	*AGXT2*	**rs3764913**	ch2:210210185:T/C	intron	0.33	0.49	0.04	1.6E-38	**rs3764913**	ch2:210210185:T/C	intron	0.15	0.40	0.05	9.9E-15
2-aminoisobutyric acid	*AGXT2*	rs40199	ch5:35037576:G/A	intron	0.88	−0.65	0.05	1.8E-38	rs180749	ch5:35033500:G/A	exon	0.83	0.90	0.05	1.1E-62
5-methylcytidine *	*PHYHD1*	**rs55758160**	ch9:128922736:A/G	intron	0.39	0.26	0.02	2.4E-35	**rs55758160**	ch9:128922736:A/G	intron	0.49	0.29	0.03	3.2E-29
N6-acetyllysine	*NAT8*	rs7594511	ch2:73454512:A/G	up	0.22	0.20	0.02	6.9E-31	rs10168931	ch2:73650092:G/A	up	0.52	0.17	0.02	6.4E-24
PC 36:4	*FADS1-3*	rs174583	ch11:61842278:C/T	inter	0.36	−0.14	0.01	2.5E-29	rs174564	ch11:61820833:A/G	inter	0.15	−0.12	0.02	4.7E-09
CE 20:4	*FADS1-3*	rs174564	ch11:61820833:A/G	inter	0.35	−0.12	0.01	3.2E-29	rs102274	ch11:61790354:T/C	down	0.09	−0.10	0.02	9.0E-08
N-acetylputrescine	*NAT2*	rs1495741	ch8:18415371:G/A	down	0.76	−0.26	0.03	1.1E-24	rs11784251	ch8:18402503:A/G	down	0.57	−0.20	0.02	1.9E-17
PC 38:4	*FADS1-3*	rs28456	ch11:61822009:A/G	inter	0.32	−0.15	0.02	2.3E-23	rs102274	ch11:61790354:T/C	down	0.09	−0.13	0.02	2.1E-08
DAG 38:5 *	*FADS1-3*	rs174583	ch11:61842278:C/T	inter	0.36	−0.21	0.02	9.8E-23	rs174564	ch11:61820833:A/G	inter	0.15	−0.17	0.04	1.0E-06
TAG 56:6	*FADS1-3*	rs174583	ch11:61842278:C/T	inter	0.36	−0.14	0.02	2.9E-21	rs174564	ch11:61820833:A/G	inter	0.15	−0.12	0.02	7.2E-07
glycine	*CPS1*	**rs1047891**	ch2:210675783:C/A	exon	0.31	0.20	0.02	3.6E-20	**rs1047891**	ch2:210675783:C/A	exon	0.37	0.17	0.03	1.2E-11
dimethylglycine	*DMGDH*	rs2947610	ch5:79064382:T/G	intron	0.80	−0.26	0.03	7.7E-20	rs74603482	ch5:79040460:A/G	intron	0.07	−0.35	0.06	2.2E-10
C4 carnitine	*ACADS*	rs1799958	ch12:120738280:G/A	exon	0.28	0.28	0.03	1.1E-19	rs12829722	ch12:120717819:C/T	up	0.08	0.27	0.06	3.5E-06

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

		White							Black						
Metabolite	**Locus**	**rsid**	**SNP**	**Pos**	**CAF**	**Beta**	**SE**	**P**	**rsid**	**SNP**	**Pos**	**CAF**	**Beta**	**SE**	**P**
biliverdin	*UGT1A1*	rs4148325	ch2:233764663:C/T	intron	0.33	0.27	0.03	5.7E-19	rs35754645	ch2:233755940:ATC/A	intron	0.46	0.29	0.04	4.5E-15
DAG 38:4*	*FADS1-3*	rs174583	ch11:61842278:C/T	inter	0.36	−0.24	0.03	2.4E-16	rs108499	ch11:61779765:C/T	down	0.07	−0.27	0.05	5.2E-07
asparagine	*ASPG*	rs12587599	ch14:104108793:C/T	intron	0.13	0.18	0.02	3.0E-16	rs10873540	ch14:103896106:T/C	up	0.57	−0.08	0.02	4.8E-05
PS plasmalogen 36:3*	*ACOT2*	**rs11626972**	ch14:73566890:A/G	up	0.43	−0.32	0.04	3.2E-16	**rs11626972**	ch14:73566890:A/G	up	0.34	−0.26	0.04	2.8E-09
CE 22:4*	*MYRF*	rs174535	ch11:61783884:T/C	exon	0.35	−0.22	0.03	5.5E-16	rs102274	ch11:61790354:T/C	down	0.09	−0.20	0.05	1.5E-05
N-acetylphenylalanine	*NAT8*	rs10185080	ch2:73570380:C/A	up	0.22	0.40	0.05	1.4E-15	rs10168931	ch2:73650092:G/A	up	0.52	0.39	0.05	7.3E-13
4-acetamidobutanoate	*NAT2*	rs1495741	ch8:18415371:G/A	down	0.76	−0.17	0.02	1.6E-15	rs721399	ch8:18401856:C/T	down	0.56	−0.14	0.02	1.6E-10
bilirubin	*UGT1A1*	rs7567229	ch2:233703893:C/A	up	0.31	0.31	0.04	2.7E-15	rs887829	ch2:233759924:C/T	up	0.46	0.38	0.05	1.4E-13
DMGV	*AGXT2*	rs40199	ch5:35037576:G/A	intron	0.88	0.60	0.08	3.3E-15	rs37369	ch5:35037010:C/T	exon	0.54	−0.42	0.05	8.2E-19
TAG 56:9	*FADS1-3*	rs174578	ch11:61838027:T/A	inter	0.36	−0.11	0.01	5.5E-15	rs2238003	ch11:61756263:C/T	down	0.07	−0.11	0.03	2.8E-04
PC 38:2	*FADS1-3*	rs174564	ch11:61820833:A/G	inter	0.35	0.11	0.01	7.6E-15	rs1729387	ch11:62217305:T/C	up	0.50	−0.07	0.02	6.1E-06
SDMA	*AGXT2*	rs40199	ch5:35037576:G/A	intron	0.88	−0.17	0.02	9.9E-15	rs37369	ch5:35037010:C/T	exon	0.54	0.12	0.02	1.6E-13
oleoyl glycine*	*FAAH*	rs324420	ch1:46405089:C/A	exon	0.21	0.39	0.05	1.7E-14	rs35686500	ch1:46414344:C/CAG	up	0.58	−0.34	0.05	8.2E-13
C8 carnitine	*ACADM*	rs61799988	ch1:75773001:G/A	down	0.29	−0.32	0.04	2.8E-13	rs2881786	ch1:75868721:G/T	down	0.69	0.24	0.05	5.5E-06
TAG 56:10	*FADS1-3*	rs1535	ch11:61830500:A/G	inter	0.34	−0.24	0.03	3.5E-13	rs10897208	ch11:62024415:A/T	up	0.19	−0.16	0.04	1.3E-04
PC 34:4	*FADS1-3*	rs174564	ch11:61820833:A/G	inter	0.35	−0.25	0.04	1.1E-12	None	ch11:62278486:C/T		0.17	0.23	0.07	3.8E-04
PE 38:6*	*LIPC*	rs261291	ch15:58387979:T/C	up	0.34	0.19	0.03	3.5E-12	rs2043085	ch15:58388755:T/C	up	0.36	−0.14	0.03	1.9E-06
TAG 54:9	*FADS1-3*	**rs174564**	ch11:61820833:A/G	inter	0.35	−0.32	0.05	3.5E-12	**rs174564**	ch11:61820833:A/G	inter	0.15	−0.25	0.07	2.0E-04
PC plasmalogen 36:5	*FADS1-3*	rs174594	ch11:61852357:C/A	inter	0.60	0.12	0.02	4.0E-12	rs99780	ch11:61829161:C/T	inter	0.39	−0.07	0.02	3.2E-04
C6 carnitine	*ACADM*	rs61799988	ch1:75773001:G/A	down	0.29	−0.31	0.05	7.8E-12	rs12134854	ch1:75671921:T/C	up	0.22	−0.23	0.06	3.1E-05

Kidney Int. Author manuscript; available in PMC 2023 April 01.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

		White							Black						
Metabolite	**Locus**	**rsid**	**SNP**	**Pos**	**CAF**	**Beta**	**SE**	**P**	**rsid**	**SNP**	**Pos**	**CAF**	**Beta**	**SE**	**P**
PC 36:3	*FADS1-3*	rs174560	ch11:61814292:T/C	inter	0.32	0.07	0.01	8.5E-12	rs2429864	ch11:61602239:C/T	down	0.06	0.09	0.02	1.6E-04
SM 18:0	*CERS4*	rs2016145	ch19:8205552:T/C	up	0.34	0.10	0.01	1.2E-11	rs12972697	ch19:8247728:G/A	intron	0.19	0.09	0.02	9.4E-06
LPC 20:4	*FADS1-3*	rs28456	ch11:61822009:A/G	inter	0.32	−0.18	0.03	1.2E-11	rs102274	ch11:61790354:T/C	down	0.09	−0.25	0.04	3.3E-09
carnosine	*KLKB1*	rs1973612	ch4:186248013:C/T	intron	0.51	0.91	0.13	1.2E-11	rs114467684	ch4:186681328:T/G	down	0.06	1.04	0.29	3.5E-04
2-aminooctanoate	*NAT8*	rs111540621	ch2:73627267:C/G	down	0.21	−0.62	0.09	2.7E-11	rs6736866	ch2:73434858:C/G	down	0.35	−0.65	0.10	2.3E-11
PE 36:4 *	*LIPC*	rs35980001	ch15:58430391:G/GC	intron	0.22	0.23	0.03	3.1E-11	rs1077835	ch15:58431227:A/G	intron	0.50	0.19	0.03	1.9E-09
ribothymidine	*TYMP*	rs140521	ch22:50539595:C/A	up	0.71	−0.16	0.02	6.5E-11	rs73172241	ch22:50517186:C/T	down	0.08	−0.22	0.06	1.6E-04

Ch, chromosome; SNP, single nucleotide polymorphism; rsid, reference SNP ID (bolded if identical between Black and White cohorts); Pos, SNP position relative to locus (up, upstream; down, downstream; inter, intergenic); CAF, coded allele frequency; PC, phosphatidylcholine; CE, cholesterol ester; DAG, diacylglycerol; PS, phosphatidylserine; DMGV, dimethylguanidino valeric acid; TAG, triacylglycerol; SDMA, symmetric dimethylarginine; PE, phosphatidylethanolamine; SM, sphingomyelin; LPC, lysophosphatidylcholine

* indicated previously unreported associations; CAF refers to second allele listed for each SNP

Kidney Int. Author manuscript; available in PMC 2023 April 01.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Table 3.

Trans-ethnic fine-mapping of validated loci

			White		Black		Trans-ethnic	
Metabolite	**Locus**	**Ch**	**SNPs**	**bp range**	**SNPs**	**bp range**	**SNPs**	**bp range**
2-aminoisobutyric acid	*AGXT2*	5	5	8,064	3	893	3	2,808
PC 36:4	*FADS1-3*	11	13	61,725	30	80,329	2	30,480
glycine	*CPS1*	2	2	2,548	1	0	1	0
C4 carnitine	*ACADS*	12	4	4,495	67	753,456	6	24,400
asparagine	*ASPG*	14	6	11,332	128	817,791	4	3,512
PS plasmalogen 36:3	*ACOT2*	14	2	1,472	9	219,349	1	0
CE 22:4	*MYRF*	11	30	67,251	51	701,385	14	41,712
oleoyl glycine	*FAAH*	1	2	22,943	10	89,828	3	24,880
C8 carnitine	*ACADM*	1	14	139,855	62	631,161	70	152,544
PE 38:6	*LIPC*	15	8	3,698	40	993,254	2	776
SM 18:0	*CERS4*	19	6	7,180	73	847,099	11	19,847

Ch, chromosome; PC, phosphatidylcholine; PS, phosphatidylserine; CE, cholesterol ester; PE, phosphatidylethanolamine; SM: sphingomyelin

Kidney Int. Author manuscript; available in PMC 2023 April 01.