Heritability and family-based GWAS analyses of the *N*-acyl ethanolamine and ceramide plasma lipidome

Kathryn A. McGurk[1,2,3], Simon G. Williams[1], Hui Guo[4], Hugh Watkins[5,6], Martin Farrall[5,6], Heather J. Cordell[7], Anna Nicolaou[2,¶], Bernard D. Keavney[1,8,¶,*]

[1]Division of Cardiovascular Sciences, Faculty of Biology, Medicine and Health, Manchester Academic Health Science Centre, University of Manchester, Manchester, UK

[2]Laboratory for Lipidomics and Lipid Biology, Division of Pharmacy and Optometry, Faculty of Biology, Medicine and Health, Manchester Academic Health Science Centre, University of Manchester, Manchester, UK

[3]Current address: National Heart and Lung Institute, Faculty of Medicine, Imperial College London, UK

[4]Division of Population Health, Health Services Research & Primary Care, Faculty of Biology, Medicine and Health, Manchester Academic Health Science Centre, University of Manchester, UK

[5]Division of Cardiovascular Medicine, Radcliffe Department of Medicine, University of Oxford, Oxford, UK

[6]Wellcome Centre for Human Genetics, University of Oxford, Oxford, UK

[7]Population Health Sciences Institute, Faculty of Medical Sciences, Newcastle University, Newcastle upon Tyne, UK

[8]Manchester Heart Centre, Manchester University NHS Foundation Trust, UK

[¶]These authors contributed equally to this work

***To whom correspondence should be addressed**

Bernard D. Keavney, AV Hill Building, University of Manchester, Manchester M13 9NT

T: +44 161 275 1225, F: +44 161 275 1183, E: bernard.keavney@manchester.ac.uk

Abstract

Signalling lipids of the *N*-acyl ethanolamine (NAE) and ceramide (CER) classes are emerging as novel cardiovascular disease biomarkers. We sought to establish the heritability of plasma NAEs (including the endocannabinoid anandamide) and CERs, and identify common DNA variants influencing the circulating concentrations of the heritable lipid species. Nine NAE and sixteen CER species were analysed in plasma samples from 999 members of 196 British Caucasian families, using targeted mass spectrometry (UPLC-MS/MS). Heritability was estimated and GWAS analyses were undertaken; all target lipids were significantly heritable (h^2 = 36%-62%). A missense variant (rs324420) in the gene encoding the enzyme fatty acid amide hydrolase (*FAAH*), which degrades NAEs, associated at GWAS significance ($P<2.15x10^{-8}$) with four NAEs (DHEA, PEA, LEA, VEA). The A allele of this SNP was associated with a 0.23 SD per-allele increase in plasma NAE species. Additionally, we found association between rs680379 in the *SPTLC3* gene, which encodes a subunit of the rate limiting enzyme in CER biosynthesis, and a range of CER species (e.g. CER[N(24)S(19)]; $P = 4.82x10^{-27}$). We also observed three novel associations (*CD83, SGPP1, FBXO28-DEGS1*) influencing plasma CER traits, two of which (*SGPP1* and *DEGS1*) implicate CER species in haematological phenotypes. NAE and CER are substantially heritable bioactive lipids, influenced by SNPs in key metabolic enzymes.

1. Introduction

Genetic studies in large numbers of individuals have identified loci where common genetic variation influences the prevalence of the major plasma lipid species, such as HDL- and LDL-cholesterol, and triglycerides (1,2). Recent advances in targeted lipidomics have enabled quantitative analyses of a greater proportion of the mediator lipidome in blood, supporting attempts to identify genetic associations for low-concentration bioactive lipid species to potentially find cardiovascular disease biomarkers.

Bioactive lipids of the *N*-acyl ethanolamine (NAE) and ceramide (CER) classes have potent roles in inflammation and immunity (3–5). NAEs are fatty acid derivatives, derived from membrane phospholipids and degraded by the enzyme fatty acid amide hydrolase (*FAAH*; Figure 1). This class of bioactive lipids includes the endocannabinoid anandamide (AEA), the nuclear factor agonist palmitoyl ethanolamide (PEA), and a number of other species with roles in neuronal signalling, pain and obesity (6–9). The contribution of genetic factors to the blood levels of NAEs has only been studied for one NAE species, oleoyl ethanolamide (OEA), in a study using untargeted mass spectrometry (10).

Ceramides are members of the sphingolipid class, being derivatives of sphingoid bases (e.g. sphingosine and dihydrosphingosine) and fatty acids (Figure 1). The first and rate limiting step (11) of their *de novo* biosynthesis is catalysed by the enzyme serine palmitoyltransferase, a heterodimeric protein whose monomers are encoded by the *SPTLC1-3* genes. Ceramides play important roles in apoptosis (12). Recently, some circulating ceramide derivatives of 18-carbon sphingosine and non-hydroxy fatty acids (e.g. CER[N(16)S(18)]) have been suggested as novel biomarkers of cardiovascular death (13), type-2 diabetes, and insulin resistance (14). Furthermore, the contribution of genetic factors to these ceramides species has also been investigated (15–18).

In this study we analysed plasma NAEs and ceramides by mass spectrometry-based targeted quantitative lipidomics in 196 British Caucasian families comprising 999 individuals, and then investigated the heritability and common genetic variant associations. We show that these bioactive lipid mediators are substantially heritable, and demonstrate that plasma NAEs and a wide range of ceramides are influenced by SNPs in key metabolic enzymes (*FAAH, SPTLC3, DEGS1, SGPP1)*. Furthermore, we identify a novel inflammatory locus (*CD83*) associated with ceramide species, and implicate this group of bioactive lipid mediators in haematological phenotypes through Mendelian randomisation, using SNPs in *DEGS1* and *SGPP1* as instruments.

2. Results

2.1 Population characteristics

Plasma samples of 999 participants from 196 British Caucasian families were included in the genetic analyses. The families consisted of 1-24 members (mean of 5 members) with plasma available for lipidomics analyses (Figure S2). Participant descriptions are listed in Table 1.

2.2 Lipidomics descriptive statistics

Of the nine NAE species identified in plasma, palmitoyl ethanolamide (PEA) was at highest abundance (1.89 ± 1.36 ng/ml [mean ± SD]), and of the sixteen plasma ceramide species, CER[N(24)S(18)] was most abundant (2.72 ± 1.29 nmol/ml), similar to previous studies (19–23).

2.3 Signalling lipid species are highly heritable

The signalling lipid species studied in this project were estimated to be substantially heritable (Figure 2, Table S8). The NAE species had estimated heritabilities ranging from 45% to 56% ($P_{adj}<4.28x10^{-15}$), with *N*-docosahexaenoylethanolamine (DHEA) producing the highest estimated heritability. The ceramide species had estimated heritabilities ranging from 36% to 62% ($P_{adj}<4.40x10^{-13}$), with CER[N(25)S(20)] and CER[N(24)S(16)] producing the highest estimated heritability.

2.4 Genome-wide association study of N-acyl ethanolamines

There were conventionally GWAS significant ($P<5x10^{-8}$) associations between four NAEs (*N*-docosahexaenoyl ethanolamide, DHEA; *N*-linoleoyl ethanolamide, LEA; *N*-palmitoyl ethanolamide PEA; *N*-vaccinoyl ethanolamide, VEA), as well as the sum of all NAEs (sumNAE), with SNPs in the gene encoding fatty acid amide hydrolase (*FAAH*), which catalyses the degradation of NAEs (Figure 1, Table S9, Figure S3). The leading SNP is a missense variant (rs324420; C385A; P129T) and eQTL of *FAAH* in multiple tissues including whole blood. Presence of the missense variant causes the enzyme to display normal catalytic properties but decreased cellular stability (24) by enhanced sensitivity of the enzyme to proteolytic degradation (19). The magnitude of the genetic effect was

considerable (Figure 3); the A allele of the lead SNP rs324420 caused a 0.23 SD per-allele increase in the plasma NAE species.

2.5 Genome-wide association study of ceramides and related sphingolipids

Seven CER[NS] and two CER[NDS] species were significantly associated with SNPs in an intergenic region on chromosome 20 (Figure 4, with example Manhattan plot depicted in Figure S4, and further details in Table S11). Assessing the SNPs using GTEx confirmed them as liver eQTLs (Table S11) found 20,000 bases downstream of the gene encoding the third subunit of serine palmitoyltransferase (*SPTLC3*; Figure 6), which catalyses the rate-limiting step (11) of CER biosynthesis (Figure 1). The SNPs are associated with differences in the expression of the *SPTLC3* gene in the liver. Associated SNPs had considerable phenotypic effects, for example the A allele of the SNP rs680379 was associated with a 0.30 SD per-allele increase in plasma ceramides. Furthermore, the summed total of all ceramide species with 24-carbon non-hydroxy fatty acids, and, independently, those with 19- and 20-carbon sphingosine bases, were found associated with the same SNPs at the *SPTLC3* locus (Table S11). A novel association was identified for CER[N(26)S(19)] at a locus on chromosome 6, upstream to the gene for inflammatory protein CD83 (e.g. rs6940658, $P=2.07x10^{-8}$; depicted in Figure S5, with further details in Table S11).

2.6 Association of ceramides and related traits with hematological phenotypes

The Gene Atlas Browser of PheWAS in the UK Biobank study was used to assess the association of significant SNPs identified here with the extensive number of phenotypes measured for the UK Biobank cohort. The ratio of CER[N(24)S(19)] to its biochemical precursor CER[N(24)DS(19)], is indicative of delta 4-desaturase, sphingolipid 1 (*DEGS1*) activity (Figure 1). A set of SNPs in the upstream region of the *DEGS1* gene on chromosome 1 associated with this ratio ($P=4.34x10^{-8}$; Figure S6, Table S10). All significant SNPs were confirmed eQTLs of *DEGS1* in whole blood (Table S11). This locus associated with numerous blood cell phenotypes in the UKBiobank data (e.g. rs4653568 and mean platelet volume; $P=4.77x10^{-12}$; Table S11).

The GWAS results for species CER[N(24)S(16)] showed an association with the *SPTLC3* region and also with a further set of SNPs upstream of the gene encoding sphingosine-1 phosphate phosphatase (*SGPP1*) (e.g. rs7160525, $P=5.67x10^{-10}$; Figure S7). This enzyme is involved in the recycling of CER[NS] species from sphingosine and sphingosine-1-phosphate (Figure 1). All significant SNPs at this locus were also associated with blood cell phenotypes, identified in the UK Biobank data (e.g. rs7160525 and mean platelet volume, $P=3.28x10^{-29}$; Table S11).

The significant SNPs identified at *SGPP1*, *DEGS1*, and *SPTLC3* genomic loci were used as instruments for two-sample Mendelian randomisation analyses (2SMR) using published blood cell count GWAS as outcome variables (25). The SNPs in *SGPP1* that associated with CER[N(24)S(16)] were significant ($P_{adj}<0.05$) in influencing platelet, red blood cell, and white blood cell traits (Table S12). 2SMR analyses using the significant SNPs identified for *SPTLC3, FAAH, SGPP1, and DEGS1* as genetic instruments, did not suggest a causal role of the lipid species for which we detected GWAS significant association, in CVD (26). Neither did we find evidence for a causal association of the ceramide species measured here with type 2 diabetes using associated SNPs in *SPTLC3* (27).

3. Discussion

We show substantial heritability estimated for an array of signalling lipid mediators found in plasma and we identify GWAS significant associations between lipid species and DNA variants of the enzymes in their respective metabolic pathways. We have provided the first GWAS significant evidence of association between SNPs in the *FAAH* gene and four plasma NAEs (DHEA, LEA, PEA, and VEA). Additionally, we have extended the previously described association between SNPs in the *SPTLC3* gene and plasma ceramides to a wider range of species. Our results indicate that these two genes are the major loci influencing plasma levels of NAEs and CERs, respectively. In addition, we have shown novel SNP associations (*CD83, SGPP1, DEGS1)* influencing plasma ceramides species, that implicate ceramides in haematological phenotypes.

The NAE species DHEA, LEA, PEA, and VEA, associated with the SNP rs680379, a missense change in the NAE degradation enzyme FAAH. The association with PEA was identified previously in a single candidate gene study of mutations in *FAAH* in 114 subjects (28), which reported the same direction of effect on plasma AEA, PEA, and OEA species but with P-values insignificant at genome-wide levels ($0.003<P<0.04$). OEA is the only NAE species that has been previously associated with DNA variants at GWAS significance. An eQTL of *FAAH* (rs1571138, upstream to *FAAHP1*, $P=5.15\times10^{-23}$) that is in complete linkage disequilibrium with the lead SNP in our study, was identified in an untargeted study of blood lipids; OEA was the only NAE species measured in that study (10). Here, we found only a suggestive association with OEA ($P=5.80\times10^{-5}$), although we observed non-significant trends in the same direction for all NAE species with genotype at this SNP (Figure 3).

While the *FAAH* missense SNP rs324420 is not associated with any disease endpoints identified from GWAS to date, the A allele, associated with higher NAE levels, has been reported to increase the risk of polysubstance addiction and abuse [MIM: 606581] in three candidate gene studies totaling 863 cases and 2,170 controls (19,29,30). PheWAS analysis using the Gene Atlas UK Biobank online

browser however, did not identify significant association in a similar number of UK Biobank cases of substance abuse/dependency (OR for A allele = 1.10; P = 0.14; 746 cases and 451,518 controls). It is possible that misclassification bias has affected the UK Biobank PheWAS; among the 451,518 UK Biobank participants assigned as controls, some reported dependencies on other substances and behaviours, such as coffee, cigarettes, prescription drugs, and gambling [UKBiobank data show case; http://biobank.ndph.ox.ac.uk/showcase/, accessed April 2019]. The potential implication of NAE species in addiction through the association with the *FAAH* SNP, warrants further investigation in larger numbers of cases.

As direct cannabinoid receptor 1 antagonist drugs have caused severe adverse psychiatric effects (31), FAAH inhibitors are being evaluated as an alternative approach to modulating endocannabinoid signalling (32,33). However, In 2016, a FAAH inhibitor resulted in severe neurological side-effects in a Phase I trial; this was hypothesised due to off-target drug effects rather than adverse on-target effects (34). As the functional *FAAH* SNP rs324420, which substantially impacts FAAH activity, did not associate with any adverse phenotypes in the UK Biobank, it is likely that on-target effects of FAAH inhibitors do not have substantial risks of causing conditions that occurred with appreciable frequency in UK Biobank.

Narrow-sense heritability has been estimated for ceramides in previous studies, showing estimated heritability of 9% to 51% (17,18). Here, we assess a different array of species and expand on these previous estimates to show that further CER[NS] and CER[NDS] are significantly heritable. The rs7157785 variant in sphingosine 1-phosphate phosphatase 1 (*SGPP1*), a ceramide metabolic enzyme, has been identified previously in GWAS of sphingomyelin (15,16,35), total cholesterol (36), glycerophospholipids (35), and the ratio of an unknown blood lipid (X-08402) to cholesterol (37). The novel association with CER[N(24)S(16)] we describe is consistent with the enzyme's role in influencing CER[NS] production, through the formation of sphingosine (C18S) for CER[NS] biosynthesis (Figure 1). The other significant SNPs identified at the same locus and in linkage disequilibrium with the lead SNP, associated with this ceramide species and have been previously

identified in further GWAS studies of blood phospholipids (38), red cell distribution width (25), sphingomyelin (38), and unknown blood metabolite X-10510 (37). All SNPs identified at this locus associated in the UKBiobank PheWAS assessment with multiple blood cell counts and other hematological phenotypes. Ceramides have been previously shown to stimulate erythrocyte formation through platelet activating factor (39). However, further studies will be required to identify the mechanism of the association between genetically determined plasma ceramide levels and blood cell phenotypes.

CER[N(26)S(19)] associated at GWAS significance with SNPs at a novel locus on chromosome 6, upstream to the gene encoding the inflammatory protein CD83 ($P=2.07x10^{-8}$), a member of the immunoglobulin superfamily of membrane receptors expressed by antigen-presenting white blood cells, leukocytes, and dendritic cells (40). An interaction between CD83 and ceramides is currently unknown, but given the involvement of ceramide signalling in inflammation and immunity (41,42), it would be of interest to investigate further.

Association between some ceramide species and the *SPTLC3* SNP rs680379 has been identified previously through the use of shotgun lipidomics for four ceramide species (CER[N(22)S(18)], CER[N(23)S(18)], CER[N(24)S(18)], and CER[N(24:1)S(18)]) at GWAS significance (15,16). Here, we identify associations between an additional seven CER[NS] and two CER[NDS] plasma species and this SNP, and with other eQTLs of serine palmitoyltransferase at the same locus; as this enzyme is the rate limiting step for the *de novo* biosynthesis of ceramides, this association may have wider implications. While we did not find a signficiant association with all CER and *SPTLC3* at GWAS, we observed non-significant trends in the same direction for all ceramide species with genotype at the rs680379 SNP (Figure 4). The information gathered from the eQTL analysis highlights all of the *SPTLC3* confirmed eQTLs act in the liver, which is a major site for plasma ceramide biosynthesis. Neither PheWAS analysis in UK Biobank, nor 2SMR analysis, identified significant disease associations with the *SPTLC3* locus. A number of CER[NS] species have been studied as potential biomarkers of cardiovascular disease and diabetes (13,43), and data from others has suggested that the

SPTLC3 locus is associated with these ceramides (15,16). The extent to which specific species have a role in cardiovascular disease remains debated (44,45).

A limitation of this study is its size (999 participants), which limits the power to detect small effects; however, our study is the largest of which we are aware that has analysed this range of plasma NAE and ceramide species to date. The inclusions of non-fasting samples that are subjectively haemolysed or contain white blood cells may add noise to the analysis that would impair the detection of weaker genetic effects. However, this may not be a signficiant issue as a study on blood ceramide lipidomics analysis did not find differences between serum and plasma samples, and fasting and non-fasting samples (46).

The associations we have uncovered suggest that further investigation of heritable lipid species for which no GWAS association was found in this study, in larger cohorts and ethnically diverse populations, would be of interest.

4. Materials and Methods

4.1 Family recruitment

Families were recruited for a quantitative genetic study of hypertension and other cardiovascular risk factors, and selected via a proband with essential hypertension (secondary hypertension was excluded using standard clinical criteria) as previously described (47). Probands were recruited from outpatients attending the John Radcliffe Hospital, Oxford hypertension clinic, or via their family doctors. Included family members were U.K. residents of self-reported White European ancestry and were required to consist of 3 or more siblings quantitatively assessable for blood pressure if one parent of the sibship was available for blood sampling, or 4 or more siblings if no parent was available. The hypertensive proband could be either in the sibship or parental generation. First, second and third degree relatives were then recruited to assemble a series of extended families. The participants were fully phenotyped for blood pressure (using ambulatory monitoring), cardiovascular risk factors, blood biochemical measures, and anthropometric traits. Non-fasting blood samples were collected, plasma separated, and stored at -80°C until lipidomic extraction. DNA was extracted from whole blood by standard methods. The collection protocol obtained ethical clearance from the Central Oxford Research Ethics Committee (06/Q1605/113) and it corresponds with the principles of the Declaration of Helsinki. Written informed consent was obtained from all participants. This cohort of extended families has previously been shown to have adequate power to detect moderate-sized genetic influences on quantitative traits (48,49).

4.2 UPLC/ESI-MS/MS mediator lipidomics

Plasma samples were extracted and analysed by mass spectrometry as previously described (50–52). Briefly, lipids were extracted from plasma (1 mL) using chloroform-methanol in the presence of internal standards: CER[N(25)S(18)] (50 pmol/sample; Ceramide/Sphingoid Internal Standard Mixture I, Avanti Polar Lipids, USA) for CER, and AEA-*d*8 (20 ng/sample; Cayman Chemical Co., USA) for NAE. Targeted lipidomics was performed on a triple quadrupole mass spectrometer (Xevo TQS, Waters, UK) with an electrospray ionisation probe coupled to a UPLC pump (Acquity UPLC,

Waters, UK). Ceramide species were separated on a C8 column (2.1 x 100 mm) and NAE were separated on a C18 column (2.1 x 50 mm) (both Acquity UPLC BEH, 1.7µm, Waters, UK). NAE species were quantified using calibration lines of synthetic standards (Cayman Chemical); relative quantitation of ceramides was based on the internal standard (Avanti Polar Lipids) (51,53). Thus, the concentration of plasma NAE species are reported in pg/ml, while the relative abundance of plasma ceramide species is reported in pmol/ml. Pooled plasma samples from healthy volunteers were used to create quality control samples that were extracted and analysed blindly alongside the familial samples. Detailed quality control information can be found in the Supplemental Methods. The CER[NS] notation (54) denotes a ceramide that contains a non-hydroxy fatty acid attached to a sphingosine base, for example, a 16-carbon non-hydroxy fatty acid joined to a 18-carbon sphingoid base is denoted as CER[N(24)S(18)], where N(24) represents a 24-carbon non-hydroxy fatty acid, and S(18) represents a 18-carbon sphingosine base attached. This species is also denoted as Cer(d18:1/24:0) in literature.

4.3 Statistical analysis

4.3.1 Covariate adjustment

Systematic error was considered from a variety of sources and assessed for collinearity; mass spectrometry batch and a trait created to adjust for sample abnormality (haemolysis or presence of white blood cells; present in 14% of samples) were included as potential covariates. Ascertainment selection was modelled via binary hypertension status. The resulting concentrations for each lipid species from the pooled quality control samples were used for adjustment of systematic errors during extraction, quantitation, and data processing. The final set of potential covariates included mass spectrometry batch, sample abnormality, quality control sample measures, age, age^2, sex, hypertension status, BMI, and total cholesterol. The lipid measurements were assessed for effect of potential covariates using stepwise multiple linear regression to identify the best set of predictors, using the ‘caret’ package and ‘leapSeq’ method in R (version 3.5.2) (see Table S4 for predictors). Multiple linear regression of the best predictors was undertaken using the ‘lm’ function in R.

Residuals from the covariate-adjusted regression models were standardized to have a mean of 0 and a variance of 1. Outliers were assessed using the R package 'car', assessing each observation by testing them as a mean-shift outlier based on studentized residuals, to remove the most extreme observations (Bonferroni P-value of $P<0.05$). Missing values were coded as such in the genetics analyses. As lipid mediators can exert individual bioactivities, all lipid species were treated uniquely for all analyses, intra-class correlation analyses are depicted in Figure S1.

4.3.2 Genome-wide genotyping quality control

Genotyping was performed using the Illumina 660W-Quad chip on 1,234 individuals (580 males and 654 females) including 248 founders, at 557,124 SNPs. Quality control of the genotyping data was undertaken using PLINK (55) (version 1.9). No duplicate variants were found. SNPs that were identified as Mendelian inconsistencies (--mendel-multigen) were marked as missing. Gender checks assessed by F-statistic (--check-sex) showed that gender as inferred from 538,771 chromosomal SNPs agreed with reported status. SNPs with low genotyping rates (--geno 0.05), low minor allele frequency (--maf 0.01), and those that failed checks of Hardy-Weinberg Equilibrium (--hwe 1e-8) were excluded. Individuals with low genotype rates (--mind 0.05) and outlying heterozygosity were removed (0.31 - 0.33 included). Relatedness was assessed by high levels of IBD sharing (--genome and --rel-check) and by visualisation of pairs of individuals' degree of relatedness (through plotting the proportion of loci where the pair shares one allele IBD (Z1) by the proportion of loci where the pair shares zero alleles IBD (Z0)), and two outlier individuals were removed. Ethnicity was assessed via principal components analysis with genotype data from the 1,000 Genome Project (56), which confirmed all participants were of European/CEU origin. Following quality control, 503,221 autosomal SNPs from 1,219 individuals (216 families) were available for SNP-based heritability assessments, of which 999 individuals (196 families; 198 founders and 801 non-founders) had plasma available for lipidomics.

4.3.3 Heritability estimates

SNP-based heritability was estimated using GCTA software (version 1.26.0) (57). A genetic relationship matrix was created from the quality controlled genotyping data and the --reml command was used to estimate variance of the traits explained by the genotyped SNPs. A complementary estimation of pedigree-based heritability was undertaken using the QTDT software (version 2.6.1) (58), by specifying the -we and -veg options to compare an environmental only variance model with a polygenic and environmental variances model. The P-values presented are adjusted for multiple comparisons via Bonferroni correction. The least significant adjusted P-value from the groups of lipid species described are depicted as $P_{adj}<X$.

4.3.4 Genotyping imputation

Following genotyping quality control, 503,221 autosomal SNPs were available to inform imputation. Imputation was performed through the Michigan Imputation Server (version v1.0.4), specifying pre-phasing with Eagle (59) (version 2.3) and imputation by Minimac3 (60) using the European population of the Human Reference Consortium (61) (version hrc.r1.1.2016). Following imputation, duplicate SNPs and SNPs with $r^2<0.8$ were removed to generate a final set of 10,652,600 SNPs. Quality control was undertaken for the imputed data on the 999 individuals with lipidomics available, as follows: SNPs that were identified as Mendelian inconsistencies (--mendel-multigen) were marked as missing. SNPs with low call rates (--geno 0.05), low minor allele frequency (--maf 0.05), and those that failed checks of Hardy-Weinberg Equilibrium (--hwe 1e-8) were excluded, resulting in a final count of 5,280,459 SNPs available for genome-wide association analyses.

4.3.5 Family-based genome-wide association studies

Linear mixed modelling approaches were used to account for family structure. Family-based genome-wide association analyses were undertaken for each lipid trait using GCTA software (version 1.26.0), specifying mixed linear model association analyses (--mlma). Genomic control inflation factors from the GWAS analyses can be found in Table S5. The least significant P-values of the significantly associated SNPs ($P<5x10^{-8}$) are depicted as P<X in the manuscript. Significantly associated SNPs were analysed by Ensembl API Client (version 1.1.5 on GRCh37.p13) to identify neighbouring genes.

Further analyses were undertaken of the significantly associated SNPs; expression quantitative trait loci (eQTL) were identified using the GTEx eQTL Browser (version 8), assessment of previously identified SNPs from GWAS was undertaken using the GWAS Catalog, review of variants on OMIM, visualisation of variants using UCSC Genome Browser, and assessment of PheWAS with the UK Biobank (62) was undertaken using the Gene Atlas Browser.

4.3.6 Two-sample Mendelian randomisation analysis

Two sample Mendelian randomisation (2SMR) analysis was undertaken in R following the guidelines provided by Davey Smith et al [https://mrcieu.github.io/TwoSampleMR/] (63). Briefly, selected examples of the significant associations identified for each class of lipid were analysed by 2SMR for a number of previously published GWAS of interest. The GWAS significant associations ($P<5x10^{-8}$) identified for NAE species PEA, and ceramide traits CER[N(22)S(19)], CER[N(24)S(16)], and CER[N(24)S(19)]/CER[N(24)DS(19)] ratio, were assessed for coronary artery disease (all), Type-2 Diabetes (CER[N(22)S(19)]), and blood cell counts (CER[N(24)S(16)] and CER[N(24)S(19)]/CER[N(24)DS(19)] ratio). Details on the published GWAS used as outcomes are presented in Table S6. As many GWAS associated SNPs were in linkage disequilibrium, the following SNPs remained in the analysis after the data clumping step; rs324420 (*FAAH*; PEA), rs438568 (*SPTLC3*; CER[N(22)S(19)]), rs7160525 (*SGPP1*; CER[N(24)S(16)]), and rs4653568 (*DEGS1*; CER[N(24)S(19)]/CER[N(24)DS(19)] ratio).

Acknowledgements and funding support

We are grateful to the families involved in this study. This work was supported by MRC Doctoral Award (MR/K501311/1 to K.M.). K.M. is supported by a University of Manchester President's Doctoral Scholarship. A.N. is supported in part by the NIHR Manchester Biomedical Research Centre. B.K. is supported by a British Heart Foundation Personal Chair.

Conflict of interest statement

None declared.

Author contributions

KM: design, analysis and interpretation of data; drafting the article; final approval of the version to be submitted. SW: analysis of data; revising the article; final approval of the version to be submitted. HG: analysis of data; revising the article; final approval of the version to be submitted. HW: acquisition of data; revising the article; final approval of the version to be submitted. MF: analysis and interpretation of data; revising the article; final approval of the version to be submitted. HC: analysis and interpretation of data; revising the article; final approval of the version to be submitted. AN, BK: conception, design, acquisition, and interpretation of data; supervision; revising the article; final approval of the version to be submitted.

References

1. Teslovich, T. M., Musunuru, K., Smith, A. V, et al. (2010) Biological, clinical and population relevance of 95 loci for blood lipids. *Nature*, **466**, 707–13.
2. Willer, C. J., Schmidt, E. M., Sengupta, S., et al. (2013) Discovery and refinement of loci associated with lipid levels. *Nat. Genet.*, **45**, 1274–1285.
3. MacEyka, M. and Spiegel, S. (2014) Sphingolipid metabolites in inflammatory disease. *Nature*, **510**, 58–67.
4. Kendall, A. C. and Nicolaou, A. (2013) Bioactive lipid mediators in skin inflammation and immunity. *Prog. Lipid Res.*, **52**, 141–164.
5. Sugamura, K., Sugiyama, S., Nozaki, T., et al. (2009) Activated endocannabinoid system in coronary artery disease and anti inflammatory effects of cannabinoid 1 receptor blockade on macrophages. *Circulation*, **119**, 28–36.
6. Devane, W. A., Hanuš, L., Breuer, A., et al. (1992) Isolation and structure of a brain constituent that binds to the cannabinoid receptor. *Science (80-.).*, **258**, 1946–1949.
7. Wilson, R. I. and Nicoll, R. a (2001) Endogenous cannabinoids mediate retrograde signalling at hippocampal synapses. *Nature*, **410**, 588–592.
8. Hohmann, A. G., Suplita, R. L., Bolton, N. M., et al. (2005) An endocannabinoid mechanism for stress-induced analgesia. *Nature*, **435**, 1108–1112.
9. Engeli, S., Böhnke, J., Feldpausch, M., et al. (2005) Activation of the peripheral endocannabinoid system in human obesity. *Diabetes*, **54**, 2838–2843.
10. Long, T., Hicks, M., Yu, H. C., et al. (2017) Whole-genome sequencing identifies common-to-rare variants associated with human blood metabolites. *Nat. Genet.*, **49**, 568–578.
11. Perry, D. K., Carton, J., Shah, A. K., et al. (2000) Serine palmitoyltransferase regulates de novo ceramide generation during etoposide-induced apoptosis. *J. Biol. Chem.*, **275**, 9078–9084.
12. Perry, D. K. and Hannun, Y. A. (1998) The role of ceramide in cell signaling. *Biochim. Biophys. Acta*, **1436**, 233–243.

13. Laaksonen, R., Ekroos, K., Sysi-Aho, M., et al. (2016) Plasma ceramides predict cardiovascular death in patients with stable coronary artery disease and acute coronary syndromes beyond LDL-cholesterol. *Eur. Heart J.*, **37**, 1967–1976.

14. Haus, J. M., Kashyap, S. R., Kasumov, T., et al. (2009) Plasma ceramides are elevated in obese subjects with type 2 diabetes and correlate with the severity of insulin resistance. *Diabetes*, **58**, 337–343.

15. Demirkan, A., van Duijn, C. M., Ugocsai, P., et al. (2012) Genome-wide association study identifies novel loci associated with circulating phospho- and sphingolipid concentrations. *PLoS Genet.*, **8**, e1002490.

16. Hicks, A. A., Pramstaller, P. P., Johansson, A., et al. (2009) Genetic determinants of circulating sphingolipid concentrations in European populations. *PLoS Genet.*, **5**, e1000672.

17. Bellis, C., Kulkarni, H., Mamtani, M., et al. (2014) Human plasma lipidome is pleiotropically associated with cardiovascular risk factors and death. *Circ. Cardiovasc. Genet.*, **7**, 854–863.

18. Cadby, G., Melton, P. E., McCarthy, N. S., et al. (2020) Heritability of 596 lipid species and genetic correlation with cardiovascular traits in the Busselton Family Heart Study. *J. Lipid Res.*, jlr.RA119000594.

19. Sipe, J. C., Chiang, K., Gerber, A. L., et al. (2002) A missense mutation in human fatty acid amide hydrolase associated with problem drug use. *Proc. Natl. Acad. Sci.*, **99**, 8394–8399.

20. Fanelli, F., Di Lallo, V. D., Belluomo, I., et al. (2012) Estimation of reference intervals of five endocannabinoids and endocannabinoid related compounds in human plasma by two dimensional-LC/MS/MS. *J. Lipid Res.*, **53**, 481–493.

21. Bowden, J. A., Heckert, A., Ulmer, C. Z., et al. (2017) Harmonizing lipidomics: NIST interlaboratory comparison exercise for lipidomics using SRM 1950–Metabolites in Frozen Human Plasma. *J. Lipid Res.*, **58**, 2275–2288.

22. Quehenberger, O., Armando, A. M., Brown, A. H., et al. (2010) Lipidomics reveals a remarkable diversity of lipids in human plasma. *J. Lipid Res.*, **51**, 3299–3305.

23. Kauhanen, D., Sysi-Aho, M., Koistinen, K. M., et al. (2016) Development and validation of a high-throughput LC–MS/MS assay for routine measurement of molecular ceramides. *Anal.*

Bioanal. Chem., **408**, 3475–3483.

24. Chiang, K. P., Gerber, A. L., Sipe, J. C., et al. (2004) Reduced cellular expression and activity of the P129T mutant of human fatty acid amide hydrolase: Evidence for a link between defects in the endocannabinoid system and problem drug use. *Hum. Mol. Genet.*, **13**, 2113–2119.

25. Astle, W. J., Elding, H., Jiang, T., et al. (2016) The Allelic Landscape of Human Blood Cell Trait Variation and Links to Common Complex Disease. *Cell*, **167**, 1415–1429.

26. Nikpay, M., Goel, A., Won, H.-H., et al. (2015) A comprehensive 1,000 Genomes-based genome-wide association meta-analysis of coronary artery disease. *Nat. Genet.*, **47**, 1121–30.

27. Mahajan, A., Go, M. J., Zhang, W., et al. (2014) Genome-wide trans-ancestry meta-analysis provides insight into the genetic architecture of type 2 diabetes susceptibility. *Nat. Genet.*, **46**, 234–244.

28. Sipe, J. C., Scott, T. M., Murray, S., et al. (2010) Biomarkers of endocannabinoid system activation in severe obesity. *PLoS One*, **5**, 1–6.

29. Flanagan, J. M., Gerber, A. L., Cadet, J. L., et al. (2006) The fatty acid amide hydrolase 385 A/A (P129T) variant: Haplotype analysis of an ancient missense mutation and validation of risk for drug addiction. *Hum. Genet.*, **120**, 581–588.

30. Sim, M. S., Hatim, A., Reynolds, G. P., et al. (2013) Association of a functional FAAH polymorphism with methamphetamine-induced symptoms and dependence in a Malaysian population. *Pharmacogenomics*, **14**, 505–514.

31. Mach, F., Montecucco, F. and Steffens, S. (2009) Effect of blockage of the endocannabinoid system by CB1 antagonism on cardiovascular risk. *Pharmacol. Reports*, **61**, 13–21.

32. Deutsch, D. G. (2016) A personal retrospective: Elevating anandamide (AEA) by targeting fatty acid amide hydrolase (FAAH) and the fatty acid binding proteins (FABPs). *Front. Pharmacol.*, **7**, 1–7.

33. Griebel, G., Stemmelin, J., Lopez-Grancha, M., et al. (2018) The selective reversible FAAH inhibitor, SSR411298, restores the development of maladaptive behaviors to acute and chronic stress in rodents. *Sci. Rep.*, **8**, 1–25.

34. Mallet, C., Dubray, C. and Dualé, C. (2016) FAAH inhibitors in the limelight, but regrettably.

Int. J. Clin. Pharmacol. Ther., **54**, 498–501.

35. Draisma, H. H. M., Pool, R., Kobl, M., et al. (2015) Genome-wide association study identifies novel genetic variants contributing to variation in blood metabolite levels. *Nat. Commun.*, **6**, 7208.

36. Klarin, D., Damrauer, S. M., Cho, K., et al. (2018) Genetics of blood lipids among ~300,000 multi-ethnic participants of the Million Veteran Program. *Nat. Genet.*, **50**, 1514–1523.

37. Shin, S. Y., Fauman, E. B., Petersen, A. K., et al. (2014) An atlas of genetic influences on human blood metabolites. *Nat. Genet.*, **46**, 543–550.

38. Li, Y., Sekula, P., Wuttke, M., et al. (2018) Genome-wide association studies of metabolites in patients with CKD identify multiple loci and illuminate tubular transport mechanisms. *J. Am. Soc. Nephrol.*, **29**, 1513–1524.

39. Lang, P. A., Kempe, D. S., Tanneur, V., et al. (2005) Stimulation of erythrocyte ceramide formation by platelet-activating factor. *J. Cell Sci.*, **118**, 1233–1243.

40. Ju, X., Silveira, P. A., Hsu, W.-H., et al. (2016) The Analysis of CD83 Expression on Human Immune Cells Identifies a Unique CD83 + -Activated T Cell Population. *J. Immunol.*, **197**, 4613–4625.

41. Maceyka, M. and Spiegel, S. (2014) Sphingolipid metabolites in inflammatory disease. *Nature*, **510**, 58–67.

42. Hannun, Y. A. and Obeid, L. M. (2008) Principles of bioactive lipid signalling: lessons from sphingolipids. *Nat. Rev. Mol. Cell Biol.*, **9**, 139–50.

43. Hilvo, M., Salonurmi, T., Havulinna, A. S., et al. (2018) Ceramide stearic to palmitic acid ratio predicts incident diabetes. *Diabetologia*, **61**, 1424–1434.

44. Neeland, I. J., Singh, S., McGuire, D. K., et al. (2018) Relation of plasma ceramides to visceral adiposity, insulin resistance and the development of type 2 diabetes mellitus: the Dallas Heart Study. *Diabetologia*, **61**, 2570–2579.

45. Anroedh, S., Hilvo, M., Akkerhuis, K. M., et al. (2018) Plasma concentrations of molecular lipid species predict long-term clinical outcome in coronary artery disease patients. *J. Lipid Res.*, **59**, 1729–1737.

46. Hammad, S. M., Pierce, J. S., Soodavar, F., et al. (2010) Blood sphingolipidomics in healthy humans: impact of sample collection methodology. *J. Lipid Res.*, **51**, 3074–87.

47. Keavney, B., Mayosi, B., Gaukrodger, N., et al. (2005) Genetic variation at the locus encompassing 11-beta hydroxylase and aldosterone synthase accounts for heritability in cortisol precursor (11-deoxycortisol) urinary metabolite excretion. *J. Clin. Endocrinol. Metab.*, **90**, 1072–1077.

48. Baker, M., Rahman, T., Hall, D., et al. (2007) The C-532T polymorphism of the angiotensinogen gene is associated with pulse pressure: A possible explanation for heterogeneity in genetic association studies of AGT and hypertension. *Int. J. Epidemiol.*, **36**, 1356–1362.

49. Vickers, M. A., Green, F. R., Terry, C., et al. (2002) Genotype at a promoter polymorphism of the interleukin-6 gene is associated with baseline levels of plasma C-reactive protein. *Cardiovasc. Res.*, **53**, 1029–1034.

50. Kendall, A. C., Pilkington, S. M., Massey, K. a, et al. (2015) Distribution of Bioactive Lipid Mediators in Human Skin. *J. Invest. Dermatol.*, **135**, 1510–1520.

51. Cucchi, D., Camacho-Muñoz, D., Certo, M., et al. (2019) Omega-3 polyunsaturated fatty acids impinge on CD4+ T cell motility and adipose tissue distribution via direct and lipid mediator-dependent effects. *Cardiovasc. Res.*, cvz208.

52. Kendall, A. C., Pilkington, S. M., Murphy, S. A., et al. (2019) Dynamics of the human skin mediator lipidome in response to dietary ω-3 fatty acid supplementation. *FASEB*, **33**, 13014–13027.

53. Poolman, T. M., Gibbs, J., Walker, A. L., et al. (2019) Rheumatoid arthritis reprograms circadian output pathways. *Arthritis Res. Ther.*, **21**, 1–13.

54. Masukawa, Y., Narita, H., Shimizu, E., et al. (2008) Characterization of overall ceramide species in human *stratum corneum*. *J. Lipid Res.*, **49**, 1466–1476.

55. Purcell, S., Neale, B., Todd-Brown, K., et al. (2007) PLINK: A tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.*, **81**, 559–575.

56. Auton, A., Abecasis, G. R., Altshuler, D. M., et al. (2015) A global reference for human

genetic variation. *Nature*, **526**, 68–74.

57. Yang, J., Benyamin, B., McEvoy, B. P., et al. (2010) Common SNPs explain a large proportion of the heritability for human height. *Nat. Genet.*, **42**, 565–569.

58. Abecasis, G. R., Cardon, L. R. and Cookson, W. O. (2000) A general test of association for quantitative traits in nuclear families. *Am. J. Hum. Genet.*, **66**, 279–92.

59. Loh, P. R., Danecek, P., Palamara, P. F., et al. (2016) Reference-based phasing using the Haplotype Reference Consortium panel. *Nat. Genet.*, **48**, 1443–1448.

60. Das, S., Forer, L., Schönherr, S., et al. (2016) Next-generation genotype imputation service and methods. *Nat. Genet.*, **48**, 1284–1287.

61. McCarthy, S., Das, S., Kretzschmar, W., et al. (2016) A reference panel of 64,976 haplotypes for genotype imputation. *Nat. Genet.*, **48**, 1279–1283.

62. Sudlow, C., Gallacher, J., Green, J., et al. (2015) UK Biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old Age. *PLOS Med.*, **12**, e1001779.

63. Smith, G. D. and Hemani, G. (2014) Mendelian randomization: Genetic anchors for causal inference in epidemiological studies. *Hum. Mol. Genet.*, **23**, R89–R98.

Figure Legends

Figure 1: Schematic overview of the biosynthetic pathways for (A) *N*-acyl ethanolamines and (B) ceramides

A) *N*-acyl ethanolamine (NAE) species, including the endocannabinoid anandamide (AEA), are produced through four independent enzymatic pathways from the membrane phospholipid precursor (*N*-acyl phosphatidylethanolamine; NAPE). Fatty acid amide hydrolase (*FAAH*) degrades NAEs to free fatty acids (such as arachidonic acid for AEA) and ethanolamine.

B) Ceramide (CER) species are biosynthesised via the enzyme serine palmitoyltransferase (*SPTLC 1-3*) that converts palmitoyl-CoA and L-serine to 3-keto dihydrosphingosine, in the rate-limiting step of the sphingolipid *de novo* pathway. The resulting dihydrosphingosine is coupled to various fatty acids via ceramide synthases (*CERS*) to generate dihydroceramides CER[NDS] that are further converted to CER[NS] via the enzyme delta 4-desaturase (*DEGS1*). Conversion of these pro-apoptotic CER[NS] species to sphingosine and sphingosine 1-phosphate, with roles in cell survival, degrades ceramides through reversible reactions. CER[NS] are also reversibly converted to sphingomyelin or further metabolised to ceramide 1-phosphate). Measured lipid species are in bold; genes encoding enzymes are in italics; genes identified through SNPs that associated at GWAS with circulating lipid levels are in red.

Figure 2: Heritability estimates of *N*-acyl ethanolamines and ceramides found in human plasma.

This figure depicts the heritability estimated for each of 25 lipid species in 999 plasma samples using SNP-based GCTA software (y-axis) and reported pedigree-based QTDT software (x-axis). This data is presented in detail in Table S8.

Figure 3: Family-based GWAS results for *N*-acyl ethanolamines and the lead SNP in fatty acid amide hydrolase (*FAAH*).

A) The radar plot depicts the P-value for association between the lead SNP and eQTL of *FAAH* (rs324420) and each of 9 NAE species in 999 plasma samples. The P-values were grouped into "$<5x10^{-8}$" ($P<5x10^{-8}$, outermost ring), "$x10^{-6}$" ($P=5.0x10^{-8}$ - $9.9x10^{-6}$ [of which there are no NAE species]), "$x10^{-5}$" ($1.0x10^{-5}$ - $9.9x10^{-5}$), and "NS" (not significant) at the center of the radar.

B) Trend in concentrations of plasma NAE species separated by FAAH rs324420 genotype. The figure depicts the mean standardised residuals of the NAE species in participants with the three genotypes at rs324420. The mean is shown with standard error. 51 participants had the AA genotype, 310 had the AC genotype, and 638 had the CC genotype in the cohort.

C) LocusZoom plot of the association of PEA with *FAAH* SNP rs324420. The LocusZoom plot depicts the association of *N*-acyl ethanolamine lipid species PEA with *FAAH* SNP rs324420 on chromosome 1 in 993 plasma samples. The r^2 for each SNP is depicted in colour. The plot was created using the LocusZoom plot tools at http://locuszoom.sph.umich.edu/.

Figure 4: Family-based GWAS results for CER[NS] and precursor CER[NDS] with an exemplar SNP in serine palmitoyltransferase (*SPTLC3*).

A) The radar plot depicts the P-value for association between the lead SNP and liver eQTL of *SPTLC3* (rs680379) with the 13 CER[NS] and 3 CER[NDS] species in 999 plasma samples. The P-values were grouped into "$<5x10^{-8}$" ($P<5x10^{-8}$, outermost ring), "$x10^{-6}$" ($P=5.0x10^{-8}$ - $9.9x10^{-6}$, "$x10^{-5}$" ($1.0x10^{-5}$ - $9.9x10^{-5}$), and "NS" (not significant) at the center of the radar.

B) Trend in concentrations of plasma ceramide species separated by *SPTLC3* rs680379 genotype. The figure depicts the mean standardised residuals of the ceramide species in participants with the three genotypes at rs680379. The mean is shown with standard error. 409 participants had the GG genotype, 442 had the AG genotype, and 148 had the AA genotype in the cohort.

C) LocusZoom plot of the association of CER[N(24)S(19)] with *SPTLC3* SNP rs680379

The LocusZoom plot depicts the association of CER[N(24)S(19)] with *SPTLC3* SNP rs680379 on chromosome 20 in 991 plasma samples. While there is a group of lead SNPs, this SNP was depicted as it has been identified previously to associate at GWAS with sphingolipid species. The r^2 for each SNP is depicted in colour. The plot was created using the LocusZoom plot tools at

http://locuszoom.sph.umich.edu/.

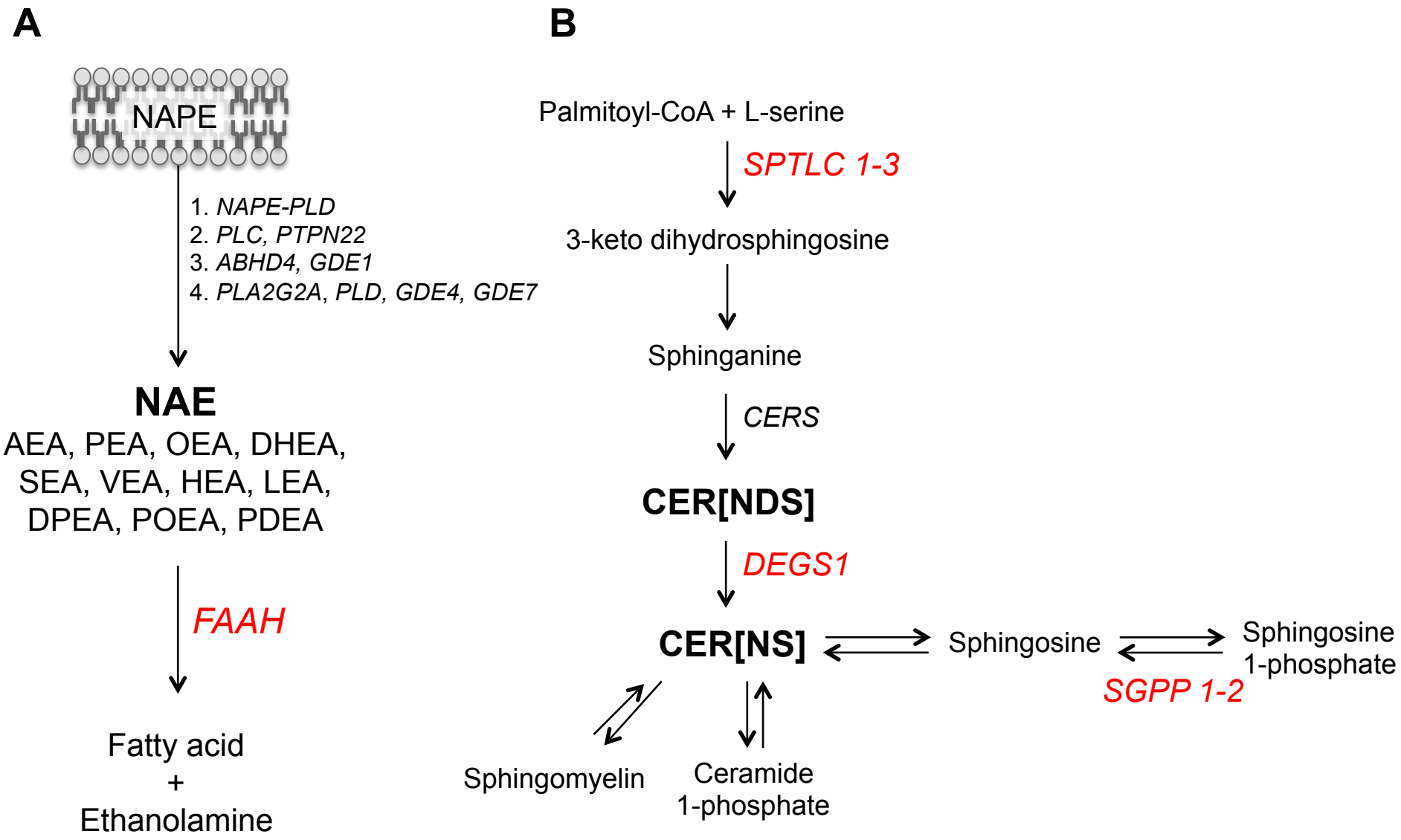
Tables

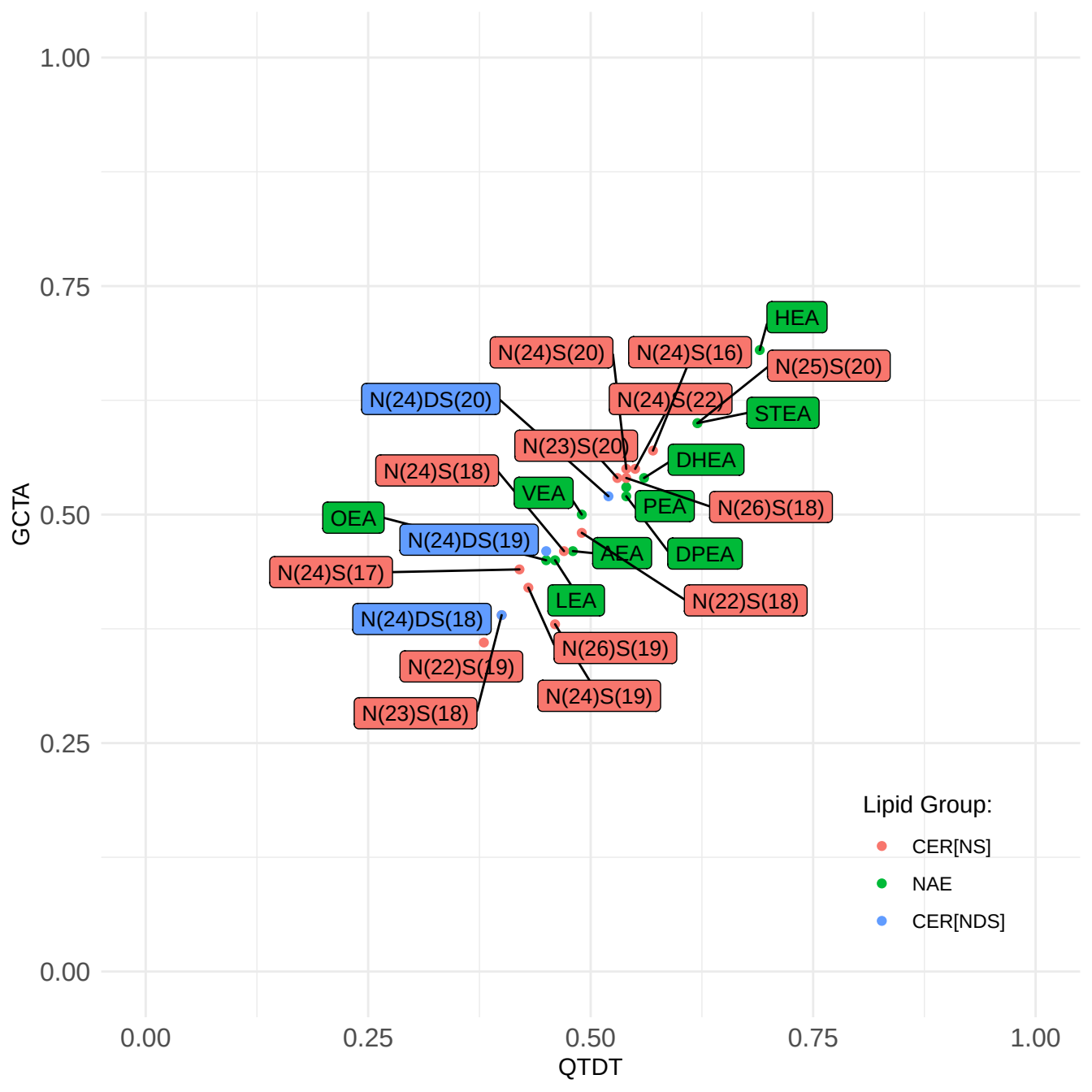
Table 1: Summary statistics for the study participants.
Data is shown as mean and standard deviation (SD) unless otherwise indicated; BMI, body mass index; WHR, waist-hip ratio.

Trait	Mean (SD)
Gender	47% Male
Hypertensive	33%
Mean Blood Pressure	138/83 mmHg
Age (years)	49 (15)
BMI	26.04 (4.33)
WHR	0.86 (0.09)
Cholesterol (mmol/L)	5.61 (1.20)

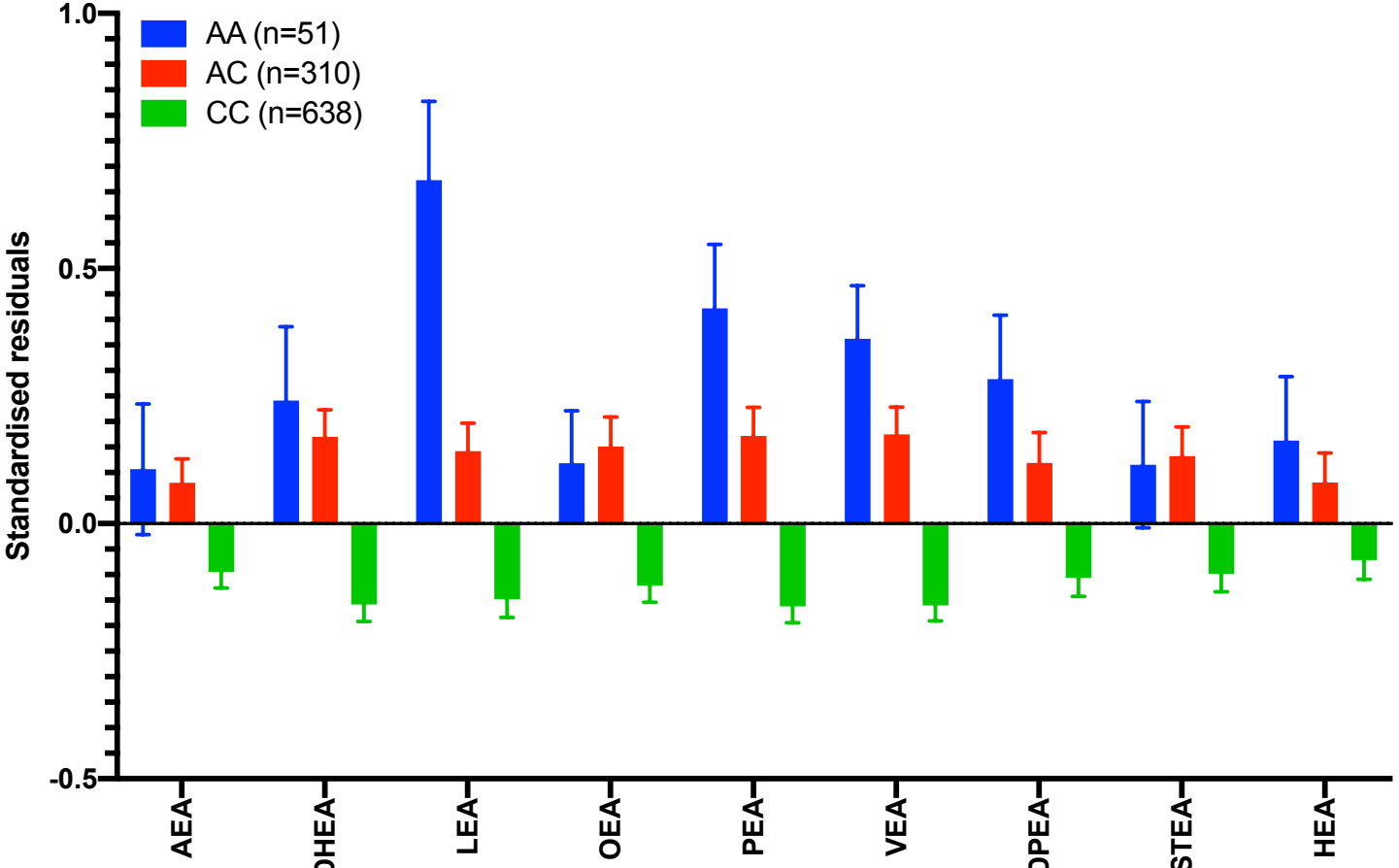
Abbreviations

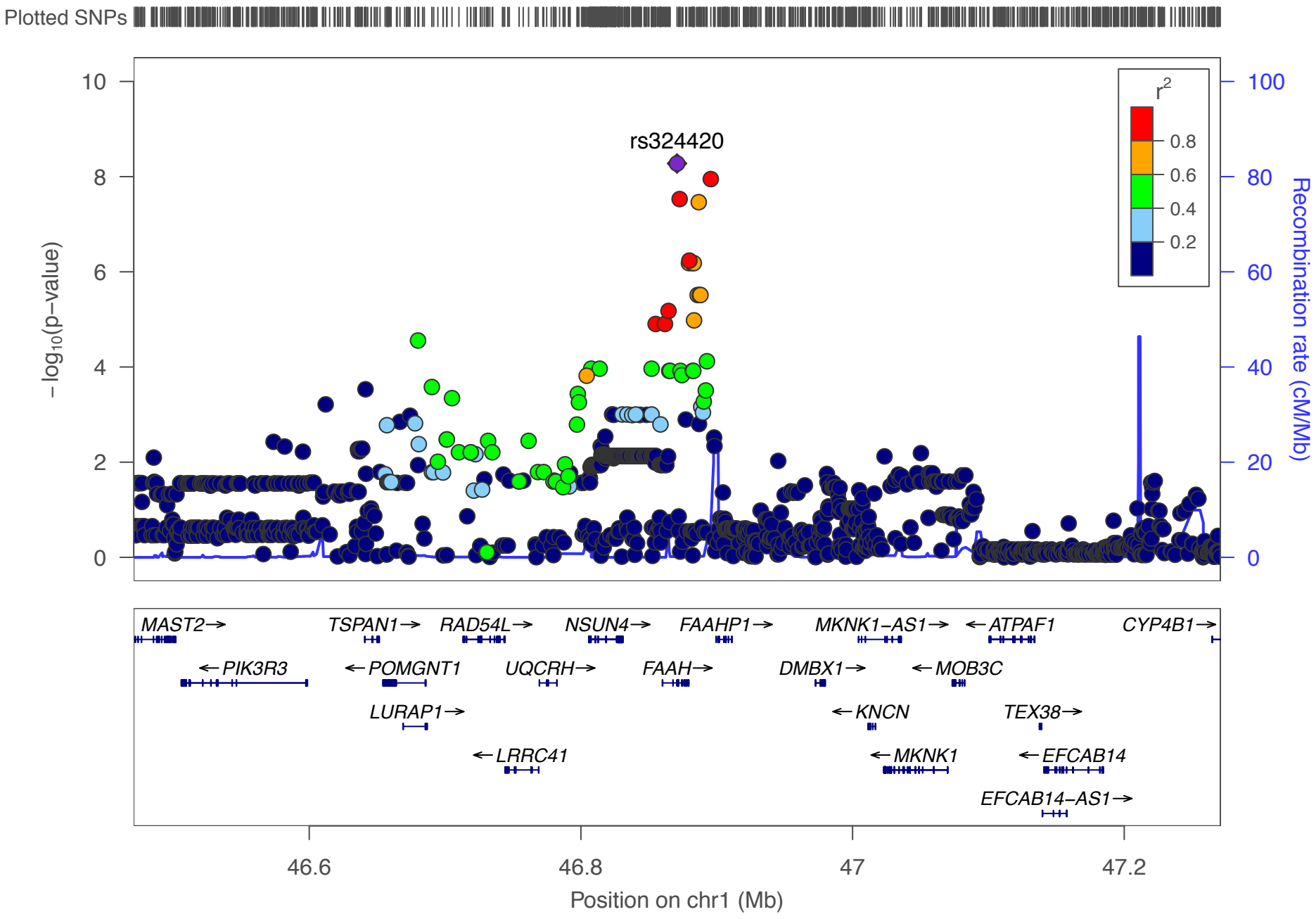
N-acyl ethanolamine (NAE); ceramide (CER); expression quantitative trait loci (eQTL); heritability (h^2); single nucleotide polymorphism (SNP); ultra performance liquid chromatography and tandem mass spectrometry (UPLC-MS/MS); Anandamide (*N*-arachidonoyl ethanolamide) (AEA); *N*-docosahexaenoyl ethanolamide (DHEA); *N*-docosapentaenoyl ethanolamine (DPEA); *N*-linoleoyl ethanolamide (LEA); *N*-oleoyl ethanolamide (OEA); *N*-palmitoyl ethanolamide (PEA); *N*-vaccinoyl ethanolamide (VEA); *N*-stearoyl ethanolamide (STEA); *N*-heptadecanoyl ethanolamide (HEA).

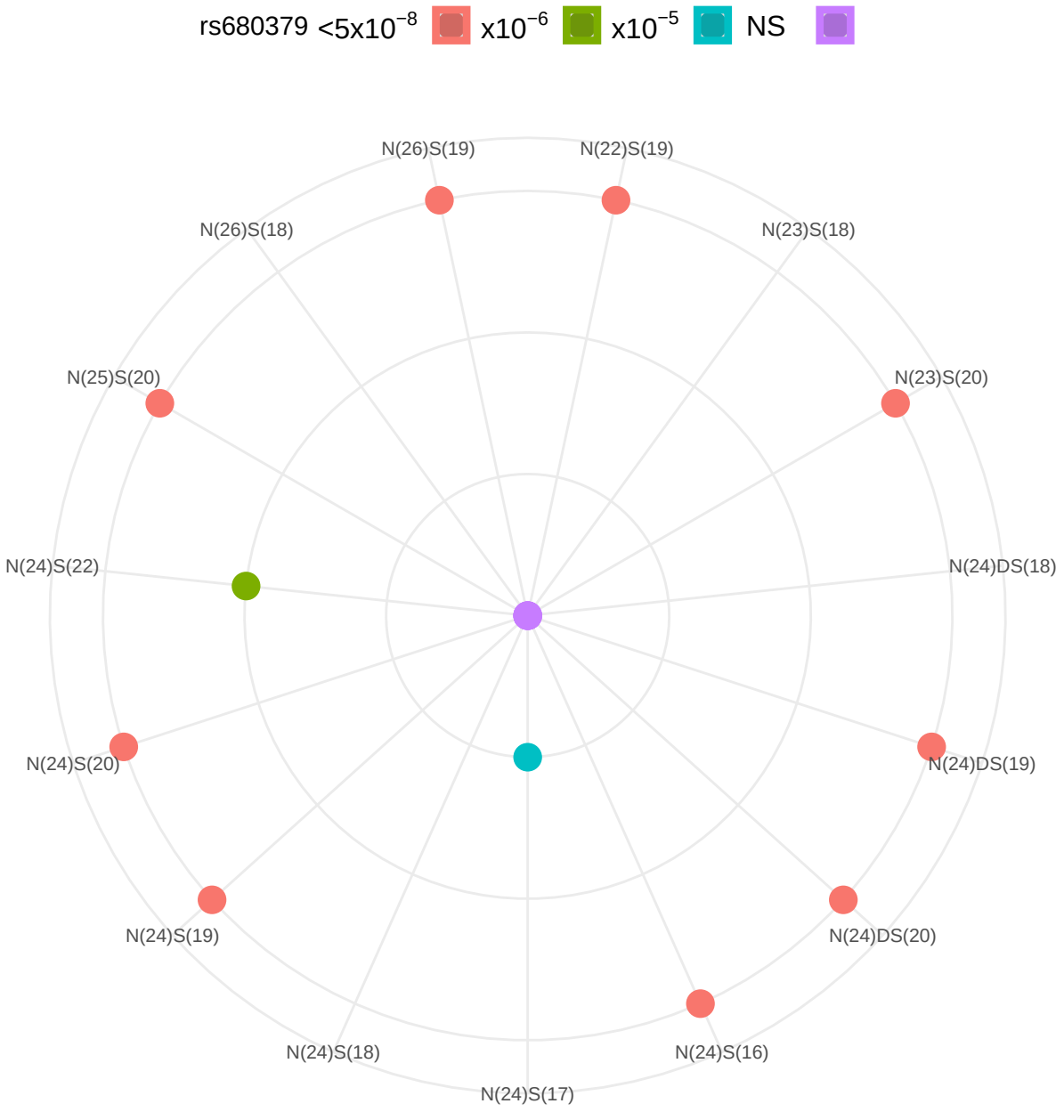

A
NAPE
1. NAPE-PLD
2. PLC, PTPN22
3. ABHD4, GDE1
4. PLA2G2A, PLD, GDE4, GDE7
NAE
AEA, PEA, OEA, DHEA,
SEA, VEA, HEA, LEA,
DPEA, POEA, PDEA
FAAH
Fatty acid
+
Ethanolamine
B
Palmitoyl-CoA + L-serine
SPTLC 1-3
3-keto dihydrosphingosine
Sphinganine
CERS
CER[NDS]
DEGS1
CER[NS]
Sphingosine
Sphingosine
1-phosphate
SGPP 1-2
Sphingomyelin
Ceramide
1-phosphate

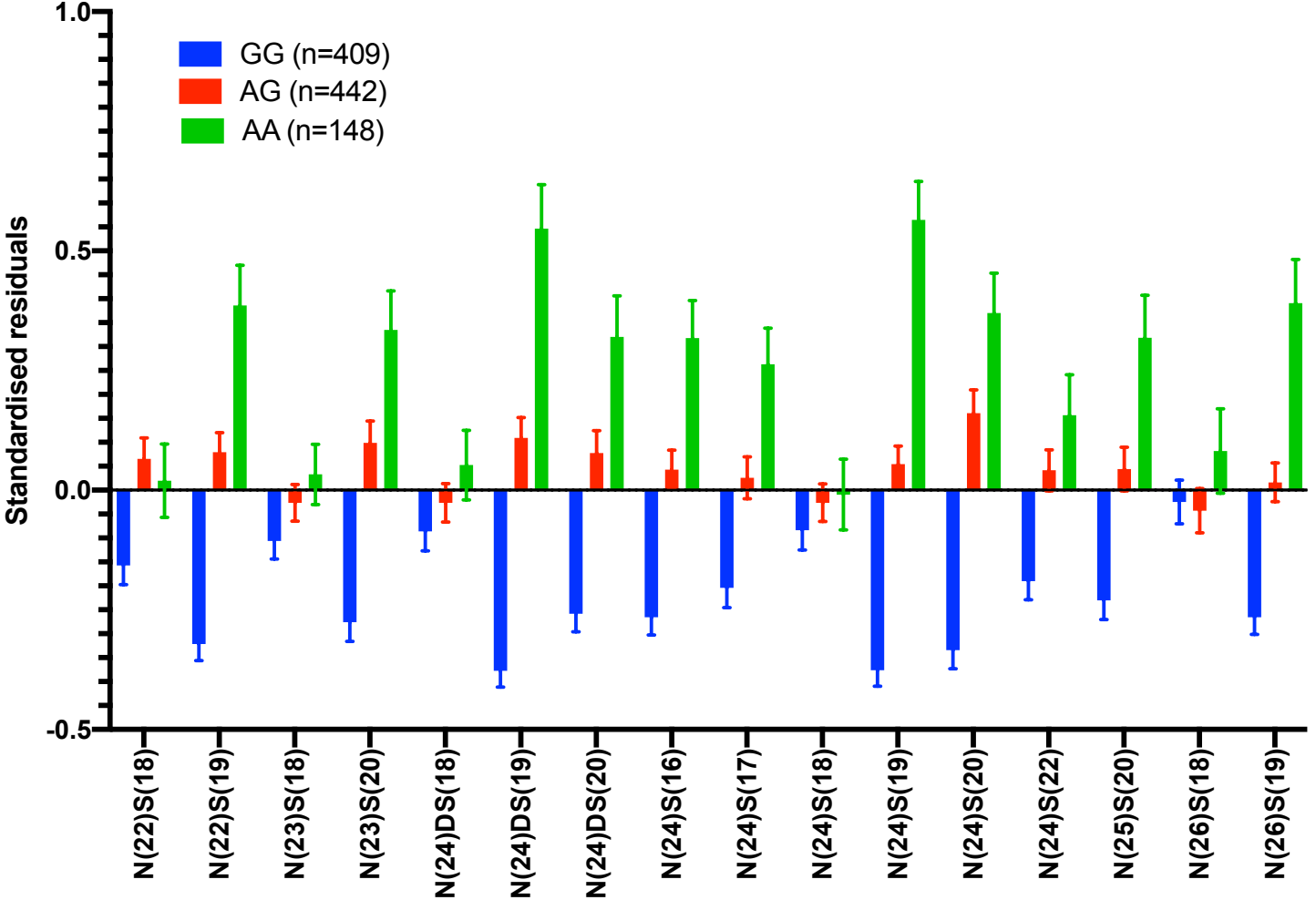

1.00
0.75
0.50
0.25
0.00
GCTA
0.00
0.25
0.50
0.75
1.00
QTDT
HEA
N(24)S(20)
N(24)S(16)
N(25)S(20)
N(24)DS(20)
N(24)S(22)
STEA
N(23)S(20)
DHEA
N(24)S(18)
VEA
PEA
N(26)S(18)
OEA
N(24)DS(19)
AEA
DPEA
N(24)S(17)
LEA
N(22)S(18)
N(24)DS(18)
N(26)S(19)
N(22)S(19)
N(24)S(19)
N(23)S(18)
Lipid Group:
CER[NS]
NAE
CER[NDS]

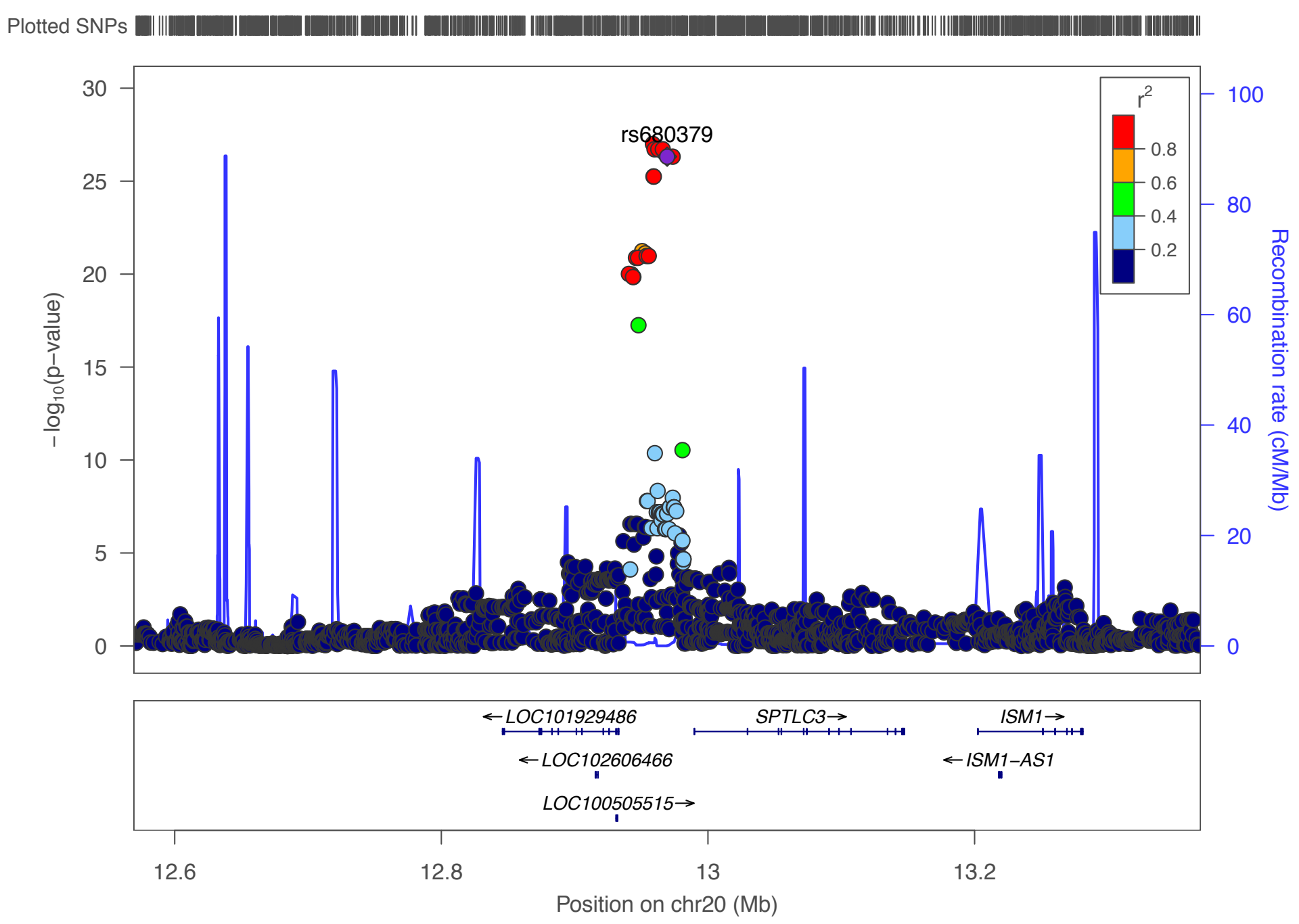
rs324420 <5x10^{-8}
x10^{-5}
NS
VEA
AEA
STEA
DHEA
PEA
DPEA
OEA
HEA
LEA


AA (n=51)
AC (n=310)
CC (n=638)
Standardised residuals
1.0
0.5
0.0
-0.5
AEA
DHEA
LEA
OEA
PEA
VEA
DPEA
STEA
HEA


Plotted SNPs
rs324420
r2
0.8
0.6
0.4
0.2
−log10(p−value)
Recombination rate (cM/Mb)
MAST2
TSPAN1
RAD54L
NSUN4
FAAHP1
MKNK1−AS1
ATPAF1
CYP4B1
PIK3R3
POMGNT1
UQCRH
FAAH
DMBX1
MOB3C
LURAP1
KNCN
TEX38
LRRC41
MKNK1
EFCAB14
EFCAB14−AS1
46.6
46.8
47
47.2
Position on chr1 (Mb)


rs680379
<5x10−8
x10−6
x10−5
NS
N(26)S(19)
N(22)S(19)
N(26)S(18)
N(23)S(18)
N(25)S(20)
N(23)S(20)
N(24)S(22)
N(24)DS(18)
N(24)S(20)
N(24)DS(19)
N(24)S(19)
N(24)DS(20)
N(24)S(18)
N(24)S(16)
N(24)S(17)


GG (n=409)
AG (n=442)
AA (n=148)
Standardised residuals
1.0
0.5
0.0
-0.5
N(22)S(18)
N(22)S(19)
N(23)S(18)
N(23)S(20)
N(24)DS(18)
N(24)DS(19)
N(24)DS(20)
N(24)S(16)
N(24)S(17)
N(24)S(18)
N(24)S(19)
N(24)S(20)
N(24)S(22)
N(25)S(20)
N(26)S(18)
N(26)S(19)


Plotted SNPs
rs680379
r2
0.8
0.6
0.4
0.2
−log10(p−value)
Recombination rate (cM/Mb)
LOC101929486
SPTLC3
ISM1
LOC102606466
ISM1−AS1
LOC100505515
12.6
12.8
13
13.2
Position on chr20 (Mb)