Heritability and family-based GWAS analyses of the *N*-acyl ethanolamine and ceramide lipidome reveal genetic influence over circulating lipids

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

[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]Division of Population Health, Health Services Research & Primary Care, Faculty of Biology, Medicine and Health, Manchester Academic Health Science Centre, University of Manchester, UK
[4]Division of Cardiovascular Medicine, Radcliffe Department of Medicine, University of Oxford, Oxford, UK
[5]Wellcome Centre for Human Genetics, University of Oxford, Oxford, UK
[6]Institute of Genetic Medicine, Newcastle University, Newcastle upon Tyne, UK
[7]Manchester Heart Centre, Manchester University NHS Foundation Trust, UK
[8]These authors contributed equally to this work

Email addresses:

Kathryn A. McGurk: Kathryn.McGurk@manchester.ac.uk
Simon G. Williams: simon.williams2@manchester.ac.uk
Hui Guo: hui.guo@manchester.ac.uk
Hugh Watkins: hugh.watkins@rdm.ox.ac.uk
Martin Farrall: mfarrall@well.ox.ac.uk
Heather J. Cordell: heather.cordell@newcastle.ac.uk
Anna Nicolaou: Anna.Nicolaou@manchester.ac.uk
Bernard D. Keavney: Bernard.Keavney@manchester.ac.uk

Correspondence:

*kathryn.mcgurk@manchester.ac.uk @KathrynMcGurk
**bernard.keavney@manchester.ac.uk @KeavneyLab

Abstract

Certain signalling lipids of the *N*-acyl ethanolamine (NAE) and ceramide (CER) classes are emerging as novel biomarkers of cardiovascular disease. We sought to establish the heritability of plasma NAEs (including endocannabinoid anandamide) and CERs, and identify common DNA variants influencing the circulating concentrations of the heritable lipid species. Eleven NAEs and thirty CERs were analysed in plasma samples from 999 members of 196 British Caucasian families, using targeted mass spectrometry-based lipidomics (UPLC-MS/MS). Family-based heritability was estimated and GWAS analyses were undertaken. All lipids were significantly heritable over a wide range (h^2 = 18%-87%). A missense variant (rs324420) in the gene encoding the enzyme fatty acid amide hydrolase (*FAAH*), which degrades NAEs, associated at GWAS significance ($P<5x10^{-8}$) with four NAEs (DHEA, PEA, LEA, VEA). This SNP, previously reported to be associated with addictive behaviour, was associated with an approximately 10% per-allele difference in mean plasma NAE species. Additionally, we have extended the previously described association between rs680379 in the gene encoding the rate limiting step of CER biosynthesis (*SPTLC3)* and CERs to a wider range of species (e.g. CER[N(24)S(19)] and rs680379 ($P=4.82x10^{-27}$)). We have shown three novel associations (*CD83, SGPP1, FBXO28-DEGS1*) influencing plasma CER traits, two of which (*SGPP1* and *DEGS1*) implicate CER species in haematological phenotypes. This first genetic analysis of plasma NAE species, and a wide range of CER mediators, highlights these bioactive lipids as substantially heritable and influenced by SNPs in key metabolic enzymes.

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]. Although lipids are not DNA-encoded, their activities and metabolism are strongly controlled by DNA-encoded enzymes and other proteins. Recent advances in targeted bioanalytics 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 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 1A). 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 variation in NAEs has not yet been studied.

Ceramides are derivatives of sphingoid bases (e.g. sphingosine and dihydrosphingosine) and fatty acids (Figure 1B). The first and rate limiting step[10] of the *de novo* biosynthesis is catalysed by the enzyme serine palmitoyltransferase, a heterodimeric protein whose monomers are encoded by the *SPTLC1-3* genes. CER play important roles in apoptosis[11]. Recently, some circulating CER derivatives of 18-

carbon sphingosine and non-hydroxy fatty acids (e.g. CER[N(16)S(18)]) have been identified as novel biomarkers of cardiovascular death[12], type-2 diabetes, and insulin resistance[13]. However, the contribution of genetic factors to the circulating levels of CER has only been investigated for six species, namely: CER[N(16)S(18)], CER[N(20)S(18)], CER[N(22)S(18)], CER[N(23)S(18)], CER[N(24)S(18)], and CER[N(24:1)S(18)], found in plasma and serum via untargeted shotgun lipidomics[14,15].

In this study we investigate the role of genetics in determining plasma levels of 11 NAE and 30 CER species, in 196 British Caucasian families comprising 999 individuals. Using targeted, quantitative lipidomics, we show for the first time that these bioactive lipid mediators are substantially heritable, and demonstrate that plasma NAEs and a wide range of CERs are influenced by SNPs in key metabolic enzymes (*FAAH, SPTLC3, DEGS1, SGPP1).* Furthermore, we identify a novel inflammatory locus (*CD83*) associated with CER species, and implicate CERs in haematological phenotypes through *DEGS1* and *SGPP1*.

Subjects and Methods

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)[16]. 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 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. First, second and third degree relatives were then recruited to assemble a series of extended families. 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[17,18]. The participants were fully phenotyped for cardiovascular risk factors, blood biochemical measures, and anthropometric traits. Non-fasting blood samples were collected, plasma separated, and stored at -80°C. DNA was extracted from whole blood by standard methods.

UPLC/ESI-MS/MS mediator lipidomics

Plasma samples were extracted and analysed by mass spectrometry as previously described[19,20]. Briefly, lipids were extracted from plasma (1 mL) using chloroform-methanol in the presence of class specific internal standards: C17S, C17DS, and 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). CER 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 CER was based on class specific internals standards (Avanti Polar Lipids). Pooled plasma samples from healthy volunteers were used to create quality control samples that were extracted and analysed blindly alongside the familial samples.

Statistical analysis

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

158 covariates included mass spectrometry batch, sample abnormality, quality control
159 sample measures, age, age^2, sex, hypertension status, BMI, and total cholesterol. The
160 lipid measurements were assessed for effect of potential covariates using stepwise
161 multiple linear regression to identify the best set of predictors, using the 'caret'
162 package and 'leapSeq' method in R (version 3.5.2) (see Table S1 for predictors).
163 Multiple linear regression of the best predictors was undertaken using the 'lm'
164 function in R. Residuals from the covariate-adjusted regression models were
165 standardized to have a mean of 0 and a variance of 1. Outliers were assessed using the
166 R package 'car', assessing each observation by testing them as a mean-shift outlier
167 based on studentized residuals, to remove the most extreme observations (Bonferroni
168 P-value of $P<0.05$). Missing values were coded as such in the genetics analyses. As
169 lipid mediators can exert individual bioactivities, all lipid species were treated
170 uniquely for all analyses, intra-class correlation analyses are depicted in Figure S1.
171
172 *Genome-wide genotyping quality control*
173 Genotyping was performed using the Illumina 660W-Quad chip on 1,234 individuals
174 (580 males and 654 females) including 248 founders, at 557,124 SNPs. Quality
175 control of the genotyping data was undertaken using PLINK[21] (version 1.9). No
176 duplicate variants were found. SNPs that were identified as Mendelian inconsistencies
177 (--mendel-multigen) were marked as missing. Gender checks assessed by F-statistic
178 (--check-sex) showed that gender as inferred from 538,771 chromosomal SNPs
179 agreed with reported status. SNPs with low genotyping rates (--geno 0.05), low minor
180 allele frequency (--maf 0.01), and those that failed checks of Hardy-Weinberg
181 Equilibrium (--hwe 1e-8) were excluded. Individuals with low genotype rates (--mind
182 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[22], 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.

Heritability estimates

SNP-based heritability was estimated using GCTA software (version 1.26.0)[23,24]. 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)[25], 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 30 tests (CER) or 11 tests (NAE) via Bonferonni correction. The least significant adjusted P-value from the groups of lipid species described are depicted as $P_{adj}<X$.

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)[26], specifying pre-phasing with Eagle[27] (version 2.3) and imputation by Minimac3[26] using the European population of the Human Reference Consortium[28] (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.

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 S2. 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 portal browser (version 8), assessment of previously identified SNPs from GWAS was undertaken using the GWAS Catalog, and assessment of PheWAS with the UK Biobank[29] was undertaken using the Gene Atlas Browser (see Web Resources).

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/][30]. 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 CER 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), addiction (PEA), 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 S3. 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).

Results

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.

Lipidomics descriptive statistics

Of the 11 NAE species identified in plasma, palmitoyl ethanolamide (PEA) was at highest abundance (1.89 ± 1.36 ng/ml (mean ± SD)), and of the 30 plasma CER species, CER[N(24)S(18)] was most abundant (128.87 ± 61.00 pmol/ml). In some cases, the ratio of product/precursor metabolites involved in specific enzymatic reactions was measured; this allowed for examination of the genetic variants in corresponding enzymes (e.g. the ratio CER[NS] to CER[NDS] is indicative of *DEGS1* activity, Figure 1B). Structure-based summations of total abundance and ratios that have previously been associated in the literature with cardiovascular risk[12] were also assessed for genetic associations. A list of the species and measurements studied, and summary statistics are presented in Table S4.

Signalling lipid species are highly heritable

The NAE species had estimated heritabilities ranging from 45% to 82% ($P_{adj}<6.72x10^{-15}$), with pentadecanoyl ethanolamide (PDEA) having the highest estimated heritability. Ceramide species showed a wide range in estimated heritability. Of the CER classes examined, CER[NS] species had heritabilities between 18% - 62% ($P_{adj}<4.50x10^{-2}$), CER[NDS] species had estimated heritability of

32% - 52%, ($P_{adj}<3.00x10^{-11}$), while CER[AS], sphingosine-1-phosphate (C18S1P), sphingosine (C18S) and dihydrosphingosine (C18DS) were all significantly heritable. Heritability results are depicted in Figure 2 (a detailed list is provided in Table S5).

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; vaccinoyl ethanolamide, VEA), as well as the sum of all NAEs (sumEA), with SNPs in the gene encoding fatty acid amide hydrolase (*FAAH*; Figure 3, with details in Table S6), which catalyses the degradation of NAEs (Figure 1A). 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[31] by enhanced sensitivity of the enzyme to proteolytic degradation[32]. The magnitude of the genetic effect was considerable; for example, participants with the AA genotype of the lead SNP rs324420 had 22% increased mean plasma concentration of PEA (2.19 ± 1.57 ng/ml, n=51) compared to those carrying the CC genotype (1.79 ± 1.29 ng/ml, n=639). A LocusZoom plot of the association with the lipid PEA is depicted in Figure 4, with the 22-chromosome Manhattan plot depicted in Figure S3.

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 5, with example Manhattan plot depicted in Figure S4, and further details in Table S7). Assessing the SNPs using

GTEx confirmed them as liver eQTLs (Table S8) found 20,000 bases downstream of the gene encoding the third subunit of serine palmitoyltransferase (*SPTLC3*; Figure 6), which catalyses the rate-limiting step[10] of CER biosynthesis (Figure 1B). The SNPs are associated with differences in the expression of the *SPTLC3* gene in the liver, leading to changes in CER production that are reflected in altered plasma levels. Associated SNPs had considerable phenotypic effects, for example the AA genotype of the SNP rs680379 was associated with a 48% increase in the concentration of CER[N(24)S(19)] compared with the GG genotype (62.84 ± 24.86 pmol/ml [n=148], and 42.44 ± 18.48 pmol/ml [n=409], respectively). Furthermore, the summed total of all CER 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 S8). 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 S7).

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 signficiant SNPs identified here with the extensive number of phenotypes measured for the UK Biobank cohort. The ratio of CER[NS] to their precursor CER[NDS], is indicative of delta 4-desaturase, sphingolipid 1 (*DEGS1*) activity (Figure 1B). A set of SNPs in the upstream region of the *DEGS1* gene on chromosome 1 associated with the product/precursor ratio, CER[N(24)S(19)]/CER[N(24)DS(19)] ($P=4.34x10^{-8}$; depicted in Figure S6 with further details in Table S7). All significant SNPs were confirmed eQTLs of *DEGS1* in

whole blood (Table S8). This locus associated with numerous blood cell phenotypes in the UKBiobank data (e.g. rs4653568 and mean platelet (thrombocyte) volume; $P=4.77x10^{-12}$; Table S8). A further set of SNPs upstream of the gene encoding sphingosine-1 phosphate phosphatase (*SGPP1*) were associated with CER[N(24)S(16)] (e.g. rs7160525, $P=5.67x10^{-10}$; Figure S7). This enzyme is involved in the recycling of CER[NS] species from sphingosine-1-phosphate and ceramide-1-phosphate (C1P) (Figure 1B). All signficiant SNPs at this locus were also associated with blood cell phenotypes, identified in the UK Biobank data (e.g. rs7160525 and mean platelet (thrombocyte) volume, $P=3.28x10^{-29}$; Table S8). The significant SNPs identified at both *SGPP1* and *DEGS1* were assessed by 2SMR using published blood cell count GWAS as outcome variables. The SNPs at both loci were found significant using 2SMR ($P<0.05$) in influencing platelet, red blood cell, and white blood cell traits. Table S9 describes the results in detail.

No significant GWAS associations were found with the particular CER species that have been previously investigated as biomarkers of coronary artery disease and type-2 diabetes (e.g. CER[N(16)S(18)]). The *SPTLC3* locus has associated with these CERs previously. Therefore, an example CER (CER[N(22)S(19)]) which associated at GWAS significantly with the *SPTLC3* locus was used to investigate the causality via two-sample Mendelian randomisation between the identified SNPs in *SPTLC3* with coronary artery disease and type-2 diabetes. There was no signficiant association between the *SPTLC3* locus and the cardiovascular disease endpoints (Table S9).

Discussion

We show, for the first time, the substantial variation in heritability estimated for an array of signalling lipid mediators found in plasma, including low concentration members of the NAE and CER classes, and we identify GWAS significant associations between lipids and 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 plasma NAEs. Additionally, we have extended the previously described association between SNPs in the *SPTLC3* gene and plasma CERs 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, FBXO28-DEGS1)* influencing plasma CER species.

Genetic analyses of N-acyl ethanolamine species

We show for the first time associations at GWAS significance for NAE species (DHEA, LEA, PEA, and VEA) with 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[33], which reported the same direction of effect on plasma AEA, PEA, STEA and OEA species but with P-values insignificant at genome-wide levels ($0.003<P<0.04$). Here, in this study of 999 related participants, we identified additional NAE species significantly associated with the SNP (DHEA, LEA, and VEA), though we did not find significant association at the locus for AEA, STEA or OEA species. We observed a similar trend to the previous paper when comparing the plasma levels of NAE species between

participants with the missense AA genotype, and those with the wildtype CC genotype (Figure S8).

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[32,34,35] and to potentially contribute to pain insensitivity[36] [MIM: 618377], obesity[37], and anxiety[38,39]. 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) nor the further 2SMR analyses undertaken here using UK Biobank addiction treatment codes (18 cases, 462999 controls; Table S9). 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. 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 (CB1) antagonist drugs have caused severe adverse psychiatric effects[40], FAAH inhibitors are being evaluated as an alternative approach to modulating endocannabinoid signalling. However, in 2016, a FAAH inhibitor resulted in severe neurological side-effects in a Phase I trial, hypothesised due to off-target drug effects[41]. As the functional *FAAH* SNP rs324420 did not associate with

any adverse phenotypes in the UK Biobank, it is likely that on-target effects of FAAH inhibitor drugs do not have substantial risks of causing conditions that occurred with appreciable frequency in UK Biobank.

Genetic analyses of ceramides and other sphingolipid species

Narrow-sense heritability has been estimated for six CER (CER[N(16)S(18)], CER[N(18)S(18)], CER[N(20)S(18)], CER[N(22)S(18)], CER[N(24)S(18)], CER[N(24:1)S(18]) and the corresponding dihydroceramide species (CER[N(16)DS(18)], CER[N(18)DS(18)], CER[N(20)DS(18)], CER[N(22)DS(18)], CER[N(24)DS(18)], CER[N(24:1)DS(18)], showing estimated heritability of 0.37 - 0.51 ($P<0.01$) for CER[NS] and 0.09 - 0.34 ($P<0.01$) for CER[NDS] in 42 Mexican American families[42]. Here, we show assessing a larger array of species, that further CER[NS] and CER[NDS] were signficiantly heritable, expanding on these previous estimates.

The rs7157785 variant in sphingosine 1-phosphate phosphatase 1 (*SGPP1*), a CER metabolic enzyme, has been identified previously in GWAS of sphingomyelin[14,15,43], total cholesterol[44], glycerophospholipids[43], total cholesterol[44], and the ratio of an unknown blood lipid (X-08402) to cholesterol[45]. The novel association with CER[N(24)S(16)] we describe is consistent with the gene's role in influencing CER[NS] production, through the formation of sphingosine (C18S) for CER[NS] biosynthesis, and the production of CER[NS] from ceramide 1-phosphate (C1P) (Figure 1). The other significant SNPs identified at the same locus which associated with this CER species have been previously identified in further GWAS studies of blood phospholipids[46], red cell distribution width[47], sphingomyelin[46], and unknown

blood metabolite X-10510[45]. All SNPs identified at this locus associated in the UKBiobank PheWAS assessment, and 2SMR analyses, with multiple blood cell counts and other hematological phenotypes (Table S8). Association with hematological phenotypes was also identified for the locus on chromosome 1 at *DEGS1*, found via the GWAS results analysing the ratio of CER[N(24)S(19)] to CER[N(24)DS(19)]. CER have been previously shown to stimulate erythrocyte formation through platelet activating factor[48]. 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 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[49]. An interaction between CD83 and CER is currently unknown, but given the involvement of ceramide signalling in inflammation and immunity[50,51], it would be of interest to investigate further.

Association between some CER species and the *SPTLC3* SNP rs680379 has been identified previously through the use of shotgun lipidomics for five CER species (CER[N(16)S(18)], CER[N(22)S(18)], CER[N(23)S(18)], CER[N(24)S(18)], and CER[N(24:1)S(18)])[14,15]. 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 CER, this association may have wider implications.

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 CER biosynthesis. PheWAS analysis in UK Biobank, nor 2SMR analysis, did not identify significant disease associations with the *SPTLC3* locus. A number of CER[NS] species have been studied as potential biomarkers of cardiovascular disease and diabetes[12,52]. While GWAS significant associations were not found with these lipids, the extent to which specific species have a role in cardiovascular disease remains debated[53–60].

The sample size analysed here (999 participants) is the largest study analysing this number of plasma NAE and CER species to date[61]. However, it is still a modest study for GWAS analyses. The associations we have uncovered suggest that further investigation of heritable lipid species for which no GWAS association was found in this study would be of interest.

Supplemental Data

Supplemental Data include nine tables and eight figures.

Acknowledgements

We are very grateful to the families involved in this study. K.M. is supported by an MRC Doctoral Award (MR/K501311/1), the University of Manchester President's Doctoral Scholarship, and Doctoral Academy Society Membership Fund. AN is supported in part by the NIHR Manchester Biomedical Research Centre. BK is supported by a BHF Personal Chair.

Declaration of interests

The authors declare no competing interests.

Web Resources

OMIM: http://www.omim.org/

GWAS Catalog: https://www.ebi.ac.uk/gwas/

GTEx Portal: https://gtexportal.org/

UCSC Genome Browser: https://genome.ucsc.edu/cgi-bin/hgGateway/

Gene Atlas UKBiobank Browser: http://geneatlas.roslin.ed.ac.uk/

UKBiobank Data Show case [April 2019]: http://biobank.ndph.ox.ac.uk/showcase/

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

References

1. Teslovich, T.M., Musunuru, K., Smith, A. V, Edmondson, A.C., Stylianou, I.M., Koseki, M., Pirruccello, J.P., Ripatti, S., Chasman, D.I., Willer, C.J., et al. (2010). Biological, clinical and population relevance of 95 loci for blood lipids. Nature *466*, 707–713.

2. Willer, C.J., Schmidt, E.M., Sengupta, S., Peloso, G.M., Gustafsson, S., Kanoni, S., Ganna, A., Chen, J., Buchkovich, M.L., Mora, 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., Matsuzawa, Y., Izumiya, Y., Miyata, K., Nakayama, M., Kaikita, K., Obata, T., Takeya, M., 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., Pertwee, R.G., Stevenson, L.A., Griffin, G., Gibson, D., Mandelbaum, A., Etinger, A., and Mechoulam, R. (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., Neely, M.H., Fegley, D., Mangieri, R., Krey, J.F., Michael Walker, J., Holmes, P. V., Crystal, J.D., et al. (2005). An endocannabinoid mechanism for stress-induced analgesia. Nature *435*, 1108–1112.

9. Engeli, S., Böhnke, J., Feldpausch, M., Gorzelniak, K., Janke, J., Bátkai, S., Pacher, P., Harvey-White, J., Luft, F.C., Sharma, A.M., et al. (2005). Activation of the peripheral endocannabinoid system in human obesity. Diabetes *54*, 2838–2843.

10. Perry, D.K., Carton, J., Shah, A.K., Meredith, F., Uhlinger, D.J., and Hannun, Y.A. (2000). Serine palmitoyltransferase regulates de novo ceramide generation during etoposide-induced apoptosis. J. Biol. Chem. *275*, 9078–9084.

11. Perry, D.K., and Hannun, Y.A. (1998). The role of ceramide in cell signaling. Biochim. Biophys. Acta *1436*, 233–243.

12. Laaksonen, R., Ekroos, K., Sysi-Aho, M., Hilvo, M., Vihervaara, T., Kauhanen, D., Suoniemi, M., Hurme, R., März, W., Scharnagl, H., 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.

13. Haus, J.M., Kashyap, S.R., Kasumov, T., Zhang, R., Kelly, K.R., Defronzo, R.A., and Kirwan, J.P. (2009). Plasma ceramides are elevated in obese subjects with type 2 diabetes and correlate with the severity of insulin resistance. Diabetes *58*, 337–343.

14. Hicks, A.A., Pramstaller, P.P., Johansson, A., Vitart, V., Rudan, I., Ugocsai, P., Aulchenko, Y., Franklin, C.S., Liebisch, G., Erdmann, J., et al. (2009). Genetic determinants of circulating sphingolipid concentrations in European populations. PLoS Genet. *5*, e1000672.

15. Demirkan, A., van Duijn, C.M., Ugocsai, P., Isaacs, A., Pramstaller, P.P., Liebisch, G., Wilson, J.F., Johansson, Å., Rudan, I., Aulchenko, Y.S., et al. (2012). Genome-wide association study identifies novel loci associated with circulating phospho- and sphingolipid concentrations. PLoS Genet. *8*, e1002490.

16. Keavney, B., Mayosi, B., Gaukrodger, N., Imrie, H., Baker, M., Fraser, R., Ingram, M., Watkins, H., Farrall, M., Davies, E., et al. (2005). Genetic variation at the

529 locus encompassing 11-beta hydroxylase and aldosterone synthase accounts for
530 heritability in cortisol precursor (11-deoxycortisol) urinary metabolite excretion. J.
531 Clin. Endocrinol. Metab. *90*, 1072–1077.
532 17. Baker, M., Rahman, T., Hall, D., Avery, P.J., Mayosi, B.M., Connell, J.M.C.,
533 Farrall, M., Watkins, H., and Keavney, B. (2007). The C-532T polymorphism of the
534 angiotensinogen gene is associated with pulse pressure: A possible explanation for
535 heterogeneity in genetic association studies of AGT and hypertension. Int. J.
536 Epidemiol. *36*, 1356–1362.
537 18. Vickers, M.A., Green, F.R., Terry, C., Mayosi, B.M., Julier, C., Lathrop, M.,
538 Ratcliffe, P.J., Watkins, H.C., and Keavney, B. (2002). Genotype at a promoter
539 polymorphism of the interleukin-6 gene is associated with baseline levels of plasma
540 C-reactive protein. Cardiovasc. Res. *53*, 1029–1034.
541 19. Kendall, A.C., Pilkington, S.M., Massey, K. a, Sassano, G., Rhodes, L.E., and
542 Nicolaou, A. (2015). Distribution of Bioactive Lipid Mediators in Human Skin. J.
543 Invest. Dermatol. *135*, 1510–1520.
544 20. Cucchi, D., Camacho-Muñoz, D., Certo, M., Niven, J., Smith, J., Nicolaou, A.,
545 and Mauro, C. (2019). Omega-3 polyunsaturated fatty acids impinge on CD4+ T cell
546 motility and adipose tissue distribution via direct and lipid mediator-dependent
547 effects. Cardiovasc. Res. cvz208.
548 21. Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M.A.R., Bender, D.,
549 Maller, J., Sklar, P., de Bakker, P.I.W., Daly, M.J., et al. (2007). PLINK: A tool set
550 for whole-genome association and population-based linkage analyses. Am. J. Hum.
551 Genet. *81*, 559–575.
552 22. Auton, A., Abecasis, G.R., Altshuler, D.M., Durbin, R.M., Bentley, D.R.,
553 Chakravarti, A., Clark, A.G., Donnelly, P., Eichler, E.E., Flicek, P., et al. (2015). A

554 global reference for human genetic variation. Nature *526*, 68–74.

555 23. Yang, J., Benyamin, B., McEvoy, B.P., Gordon, S., Henders, A.K., Nyholt, D.R.,

556 Madden, P.A., Heath, A.C., Martin, N.G., Montgomery, G.W., et al. (2010). Common

557 SNPs explain a large proportion of the heritability for human height. Nat. Genet. *42*,

558 565–569.

559 24. Zaitlen, N., Kraft, P., Patterson, N., Pasaniuc, B., Bhatia, G., Pollack, S., and

560 Price, A.L. (2013). Using Extended Genealogy to Estimate Components of

561 Heritability for 23 Quantitative and Dichotomous Traits. PLoS Genet. *9*, e1003520.

562 25. Abecasis, G.R., Cardon, L.R., and Cookson, W.O. (2000). A general test of

563 association for quantitative traits in nuclear families. Am. J. Hum. Genet. *66*, 279–

564 292.

565 26. Das, S., Forer, L., Schönherr, S., Sidore, C., Locke, A.E., Kwong, A., Vrieze, S.I.,

566 Chew, E.Y., Levy, S., McGue, M., et al. (2016). Next-generation genotype imputation

567 service and methods. Nat. Genet. *48*, 1284–1287.

568 27. Loh, P.R., Danecek, P., Palamara, P.F., Fuchsberger, C., Reshef, Y.A., Finucane,

569 H.K., Schoenherr, S., Forer, L., McCarthy, S., Abecasis, G.R., et al. (2016).

570 Reference-based phasing using the Haplotype Reference Consortium panel. Nat.

571 Genet. *48*, 1443–1448.

572 28. McCarthy, S., Das, S., Kretzschmar, W., Delaneau, O., Wood, A.R., Teumer, A.,

573 Kang, H.M., Fuchsberger, C., Danecek, P., Sharp, K., et al. (2016). A reference panel

574 of 64,976 haplotypes for genotype imputation. Nat. Genet. *48*, 1279–1283.

575 29. Sudlow, C., Gallacher, J., Green, J., Sprosen, T., Pell, J., Burton, P., Matthews, P.,

576 Liu, B., Danesh, J., Downey, P., et al. (2015). UK Biobank: an open access resource

577 for identifying the causes of a wide range of complex diseases of middle and old Age.

578 PLOS Med. *12*, e1001779.

579 30. Smith, G.D., and Hemani, G. (2014). Mendelian randomization: Genetic anchors
580 for causal inference in epidemiological studies. Hum. Mol. Genet. *23*, R89–R98.
581 31. Chiang, K.P., Gerber, A.L., Sipe, J.C., and Cravatt, B.F. (2004). Reduced cellular
582 expression and activity of the P129T mutant of human fatty acid amide hydrolase:
583 Evidence for a link between defects in the endocannabinoid system and problem drug
584 use. Hum. Mol. Genet. *13*, 2113–2119.
585 32. Sipe, J.C., Chiang, K., Gerber, A.L., Beutler, E., and Cravatt, B.F. (2002). A
586 missense mutation in human fatty acid amide hydrolase associated with problem drug
587 use. Proc. Natl. Acad. Sci. *99*, 8394–8399.
588 33. Sipe, J.C., Scott, T.M., Murray, S., Harismendy, O., Simon, G.M., Cravatt, B.F.,
589 and Waalen, J. (2010). Biomarkers of endocannabinoid system activation in severe
590 obesity. PLoS One *5*, 1–6.
591 34. Flanagan, J.M., Gerber, A.L., Cadet, J.L., Beutler, E., and Sipe, J.C. (2006). The
592 fatty acid amide hydrolase 385 A/A (P129T) variant: Haplotype analysis of an ancient
593 missense mutation and validation of risk for drug addiction. Hum. Genet. *120*, 581–
594 588.
595 35. Sim, M.S., Hatim, A., Reynolds, G.P., and Mohamed, Z. (2013). Association of a
596 functional FAAH polymorphism with methamphetamine-induced symptoms and
597 dependence in a Malaysian population. Pharmacogenomics *14*, 505–514.
598 36. Cajanus, K., Holmström, E.J., Wessman, M., Anttila, V., Kaunisto, M.A., and
599 Kalso, E. (2016). Effect of endocannabinoid degradation on pain: role of FAAH
600 polymorphisms in experimental and postoperative pain in women treated for breast
601 cancer. Pain *157*, 361–369.
602 37. Sipe, J.C., Waalen, J., Gerber, A., and Beutler, E. (2005). Overweight and obesity
603 associated with a missense polymorphism in fatty acid amide hydrolase (FAAH). Int.

J. Obes. *29*, 755–759.

38. Dincheva, I., Drysdale, A.T., Hartley, C.A., Johnson, D.C., Jing, D., King, E.C., Ra, S., Gray, J.M., Yang, R., DeGruccio, A.M., et al. (2015). FAAH genetic variation enhances fronto-amygdala function in mouse and human. Nat. Commun. *6*, 1–9.

39. Lazary, J., Eszlari, N., Juhasz, G., and Bagdy, G. (2016). Genetically reduced FAAH activity may be a risk for the development of anxiety and depression in persons with repetitive childhood trauma. Eur. Neuropsychopharmacol. *26*, 1020–1028.

40. 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.

41. Mallet, C., Dubray, C., and Dualé, C. (2016). FAAH inhibitors in the limelight, but regrettably. Int. J. Clin. Pharmacol. Ther. *54*, 498–501.

42. Bellis, C., Kulkarni, H., Mamtani, M., Kent, J.W., Wong, G., Weir, J.M., Barlow, C.K., Diego, V., Almeida, M., Dyer, T.D., et al. (2014). Human plasma lipidome is pleiotropically associated with cardiovascular risk factors and death. Circ. Cardiovasc. Genet. *7*, 854–863.

43. Draisma, H.H.M., Pool, R., Kobl, M., Jansen, R., Petersen, A.-K., Vaarhorst, A.A.M., Yet, I., Haller, T., Demirkan, A., Esko, T., et al. (2015). Genome-wide association study identifies novel genetic variants contributing to variation in blood metabolite levels. Nat. Commun. *6*, 7208.

44. Klarin, D., Damrauer, S.M., Cho, K., Sun, Y. V., Teslovich, T.M., Honerlaw, J., Gagnon, D.R., DuVall, S.L., Li, J., Peloso, G.M., et al. (2018). Genetics of blood lipids among ~300,000 multi-ethnic participants of the Million Veteran Program. Nat. Genet. *50*, 1514–1523.

629 45. Shin, S.Y., Fauman, E.B., Petersen, A.K., Krumsiek, J., Santos, R., Huang, J.,
630 Arnold, M., Erte, I., Forgetta, V., Yang, T.P., et al. (2014). An atlas of genetic
631 influences on human blood metabolites. Nat. Genet. *46*, 543–550.
632 46. Li, Y., Sekula, P., Wuttke, M., Wahrheit, J., Hausknecht, B., Schultheiss, U.T.,
633 Gronwald, W., Schlosser, P., Tucci, S., Ekici, A.B., et al. (2018). Genome-wide
634 association studies of metabolites in patients with CKD identify multiple loci and
635 illuminate tubular transport mechanisms. J. Am. Soc. Nephrol. *29*, 1513–1524.
636 47. Astle, W.J., Elding, H., Jiang, T., Allen, D., Ruklisa, D., Mann, A.L., Mead, D.,
637 Bouman, H., Riveros-Mckay, F., Kostadima, M.A., et al. (2016). The Allelic
638 Landscape of Human Blood Cell Trait Variation and Links to Common Complex
639 Disease. Cell *167*, 1415–1429.
640 48. Lang, P.A., Kempe, D.S., Tanneur, V., Eisele, K., Klarl, B.A., Myssina, S.,
641 Jendrossek, V., Ishii, S., Shimizu, T., Waidmann, M., et al. (2005). Stimulation of
642 erythrocyte ceramide formation by platelet-activating factor. J. Cell Sci. *118*, 1233–
643 1243.
644 49. Ju, X., Silveira, P.A., Hsu, W.-H., Elgundi, Z., Alingcastre, R., Verma, N.D.,
645 Fromm, P.D., Hsu, J.L., Bryant, C., Li, Z., et al. (2016). The Analysis of CD83
646 Expression on Human Immune Cells Identifies a Unique CD83 + -Activated T Cell
647 Population. J. Immunol. *197*, 4613–4625.
648 50. Maceyka, M., and Spiegel, S. (2014). Sphingolipid metabolites in inflammatory
649 disease. Nature *510*, 58–67.
650 51. Hannun, Y.A., and Obeid, L.M. (2008). Principles of bioactive lipid signalling:
651 lessons from sphingolipids. Nat. Rev. Mol. Cell Biol. *9*, 139–150.
652 52. Hilvo, M., Salonurmi, T., Havulinna, A.S., Kauhanen, D., Pedersen, E.R., Tell,
653 G.S., Meyer, K., Teeriniemi, A.M., Laatikainen, T., Jousilahti, P., et al. (2018).

654 Ceramide stearic to palmitic acid ratio predicts incident diabetes. Diabetologia *61*,
655 1424–1434.
656 53. Wang, D.D., Toledo, E., Hruby, A., Rosner, B.A., Willett, W.C., Sun, Q.,
657 Razquin, C., Zheng, Y., Ruiz-Canela, M., Guasch-Ferré, M., et al. (2017). Plasma
658 ceramides, mediterranean diet, and incident cardiovascular disease in the PREDIMED
659 trial. Circulation *135*, 2028–2040.
660 54. Lopez, X., Goldfine, A.B., Holland, W.L., Gordillo, R., and Scherer, P.E. (2013).
661 Plasma ceramides are elevated in female children and adolescents with type 2
662 diabetes. J. Pediatr. Endocrinol. Metab. *26*, 995–998.
663 55. Jensen, P.N., Fretts, A.M., Yu, C., Hoofnagle, A.N., Umans, J.G., Howard, B. V.,
664 Sitlani, C.M., Siscovick, D.S., King, I.B., Sotoodehnia, N., et al. (2019). Circulating
665 sphingolipids, fasting glucose, and impaired fasting glucose: The Strong Heart Family
666 Study. EBioMedicine *41*, 44–49.
667 56. Lemaitre, R.N., Yu, C., Hoofnagle, A., Hari, N., Jensen, P., Fretts, A.M., Umans,
668 J.G., Howard, B. V, Sitlani, C.M., Siscovick, D.S., et al. (2018). Circulating
669 sphingolipids, insulin, HOMA-IR and HOMA-B: The Strong Heart Family Study
670 Running title: Sphingolipids and insulin resistance markers. Diabetes *67*, 1663–1672.
671 57. Yu, J., Pan, W., Shi, R., Yang, T., Li, Y., Yu, G., Bai, Y., Schuchman, E.H., He,
672 X., and Zhang, G. (2015). Ceramide Is Upregulated and Associated With Mortality in
673 Patients With Chronic Heart Failure. Can. J. Cardiol. *31*, 357–363.
674 58. de Carvalho, L.P., Tan, S.H., Ow, G.S., Tang, Z., Ching, J., Kovalik, J.P., Poh,
675 S.C., Chin, C.T., Richards, A.M., Martinez, E.C., et al. (2018). Plasma Ceramides as
676 Prognostic Biomarkers and Their Arterial and Myocardial Tissue Correlates in Acute
677 Myocardial Infarction. JACC Basic to Transl. Sci. *3*, 163–175.
678 59. Pan, W., Yu, J., Shi, R., Yan, L., Yang, T., Li, Y., Zhang, Z., Yu, G., Bai, Y.,

679 Schuchman, E.H., et al. (2014). Elevation of ceramide and activation of secretory acid
680 sphingomyelinase in patients with acute coronary syndromes. Coron. Artery Dis. *25*,
681 230–235.
682 60. Bergman, B.C., Brozinick, J.T., Strauss, A., Bacon, S., Kerege, A., Bui, H.H.,
683 Sanders, P., Siddall, P., Kuo, M.S., and Perreault, L. (2015). Serum sphingolipids:
684 relationships to insulin sensitivity and changes with exercise in humans. Am. J.
685 Physiol. Endocrinol. Metab. *309*, 398–408.
686 61. Hinterwirth, H., Stegemann, C., and Mayr, M. (2014). Lipidomics: Quest for
687 molecular lipid biomarkers in cardiovascular disease. Circ. Cardiovasc. Genet. *7*,
688 941–954.
689

Figure Legends

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

A) *N*-acyl ethanolamine species (NAE), 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. NAEs have signalling roles in pain, obesity, inflammation, and neurotransmission.

B) Ceramide (CER) species are biosynthesised via the enzyme serine palmitoyltransferase (*SPTLC1-3*) 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 (C18DS) 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 (C18S) and sphingosine 1-phosphate (C18S1P), with roles in cell survival, degrades ceramides through reversible reactions. CER[NS] are also reversibly converted to sphingomyelin (SM) or further metabolised to ceramide 1-phosphate [C1P]). In a similar way, addition of alpha-hydroxy fatty acids to C18DS, results in CER[ADS] species.

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 lipid species using SNP-based GCTA software (y-axis) and reported pedigree-based QTDT software (x-axis). This data is presented in detail in Table S5.

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

The radar plot depicts the P-value for association between the lead SNP and eQTL of *FAAH* (rs324420) and each NAE species. 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.

Figure 4: 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. 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 5: Family-based GWAS results for CER[NS] and precursor CER[NDS] with an exemplar SNP in serine palmitoyltransferase (*SPTLC3*).

The radar plot depicts the P-value for association between the lead SNP and liver eQTL of *SPTLC3* (rs680379) with CER species. 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.

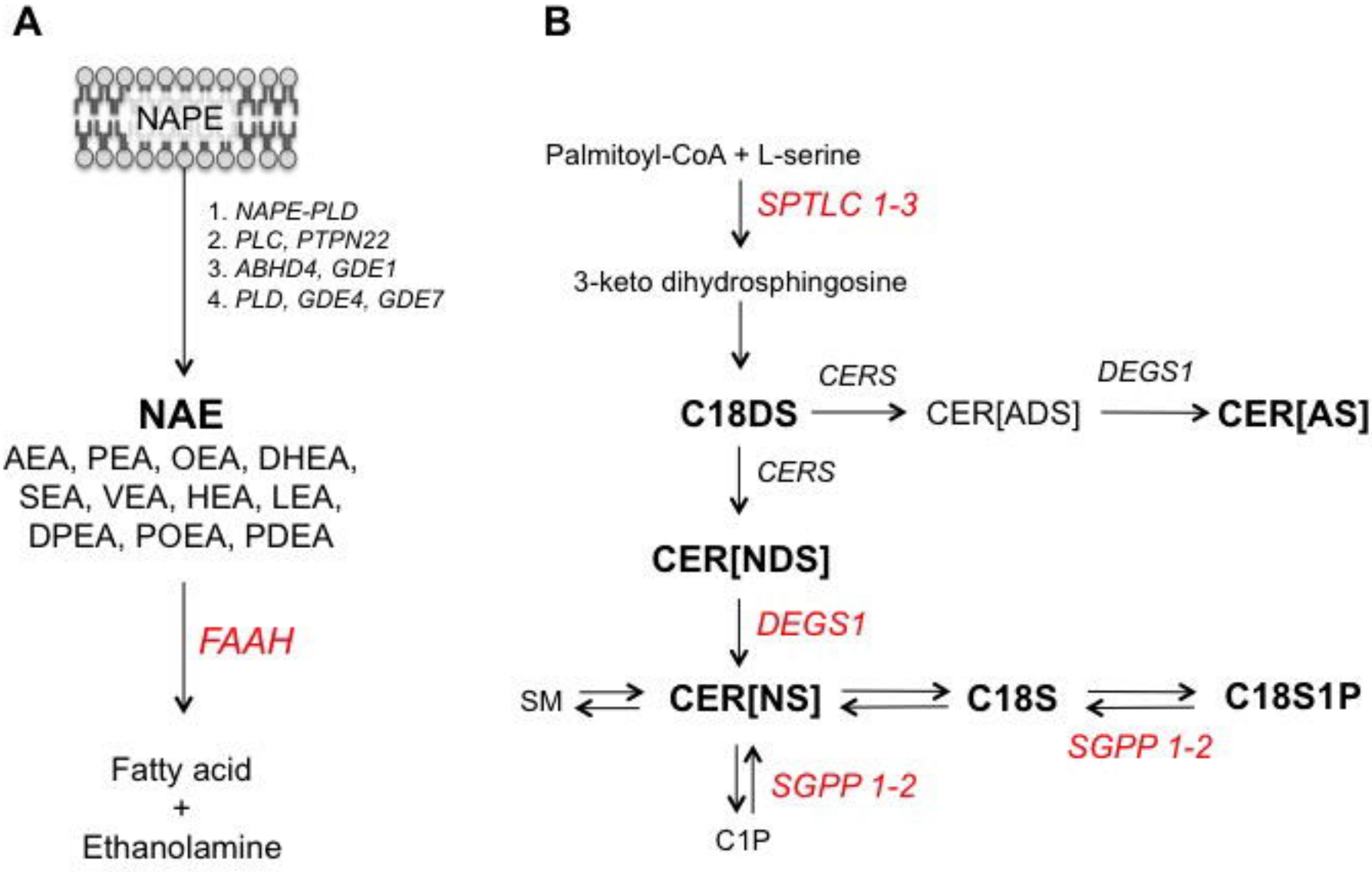
Figure 6: LocusZoom plot of the association of CER[N(24)S(19)] with *SPTLC3* SNP rs680379

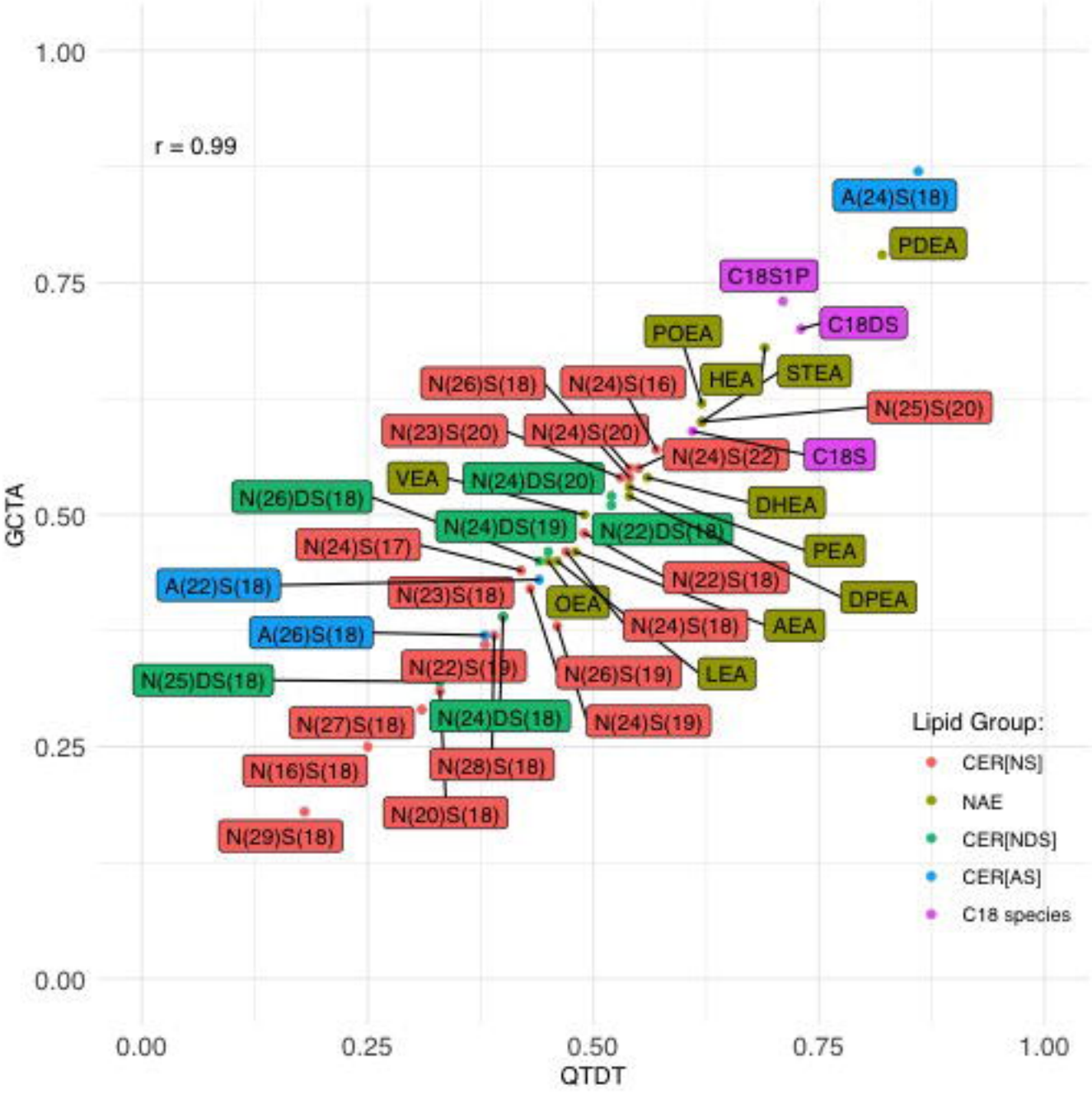
The LocusZoom plot depicts the association of CER[N(24)S(19)] with FAAH SNP rs680379 on chromosome 1. While there is a group of lead SNPs, this SNP was depicted as it has been identified previously to associate at GWAS with sphingolipid species. The r^2 for each SNP is depicted in colour. The plot was created using the LocusZoom plot tools at http://locuszoom.sph.umich.edu/.

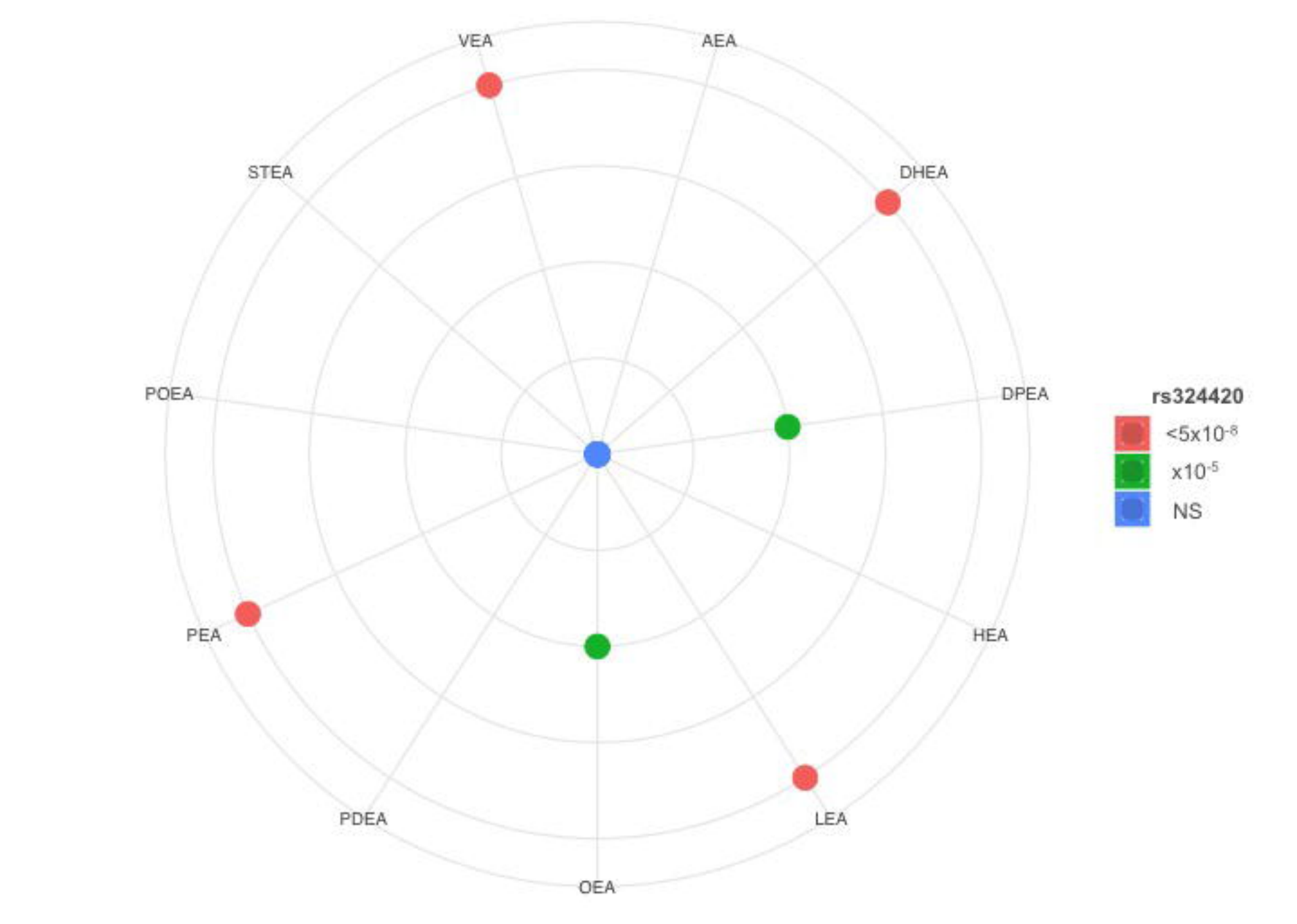
Tables

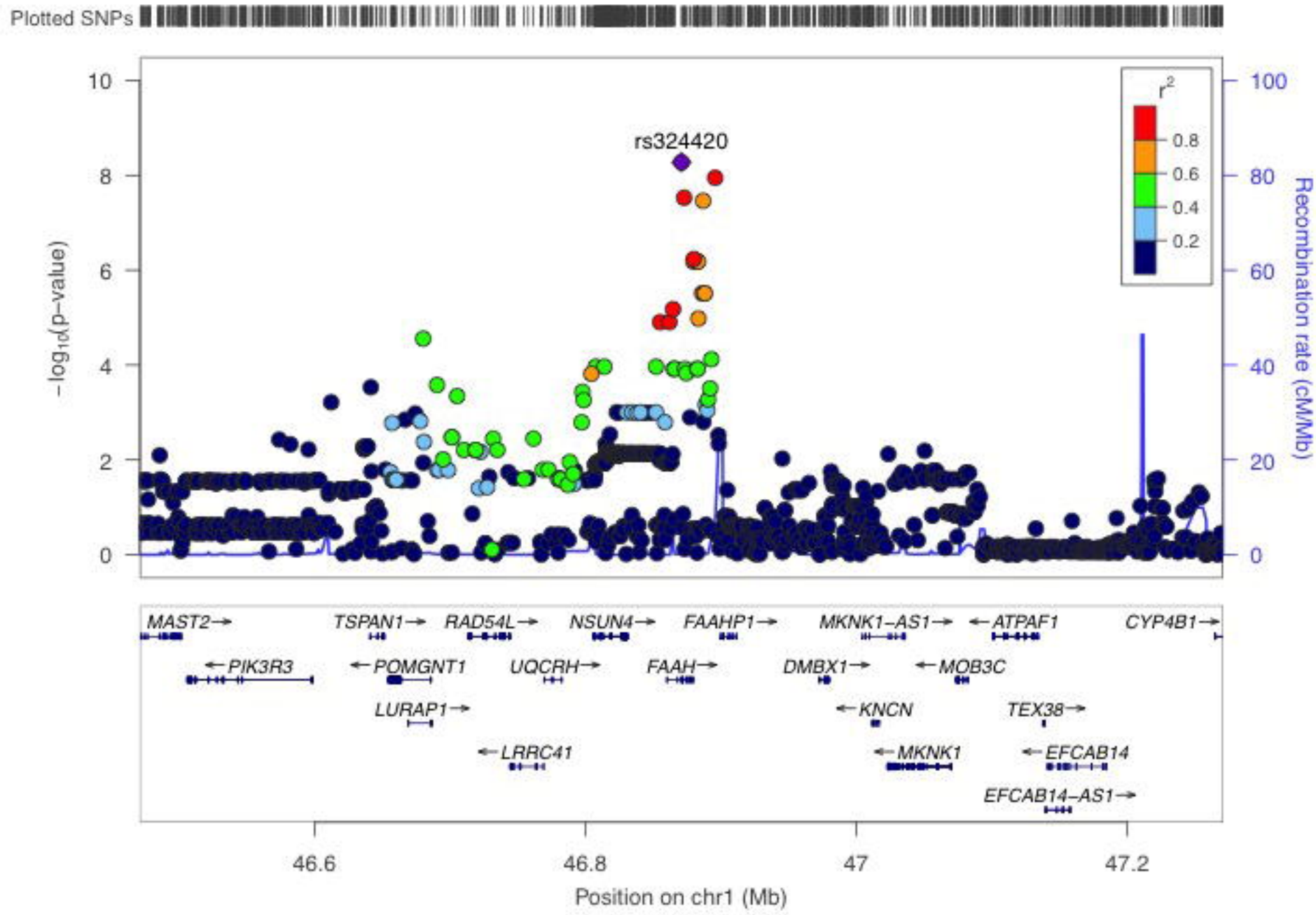
Table 1: Summary statistics for the study participants.
Data is shown as mean and standard deviation (SD) unless otherwise indicated; BMI, body mass index; WHR, waist-hip ratio.

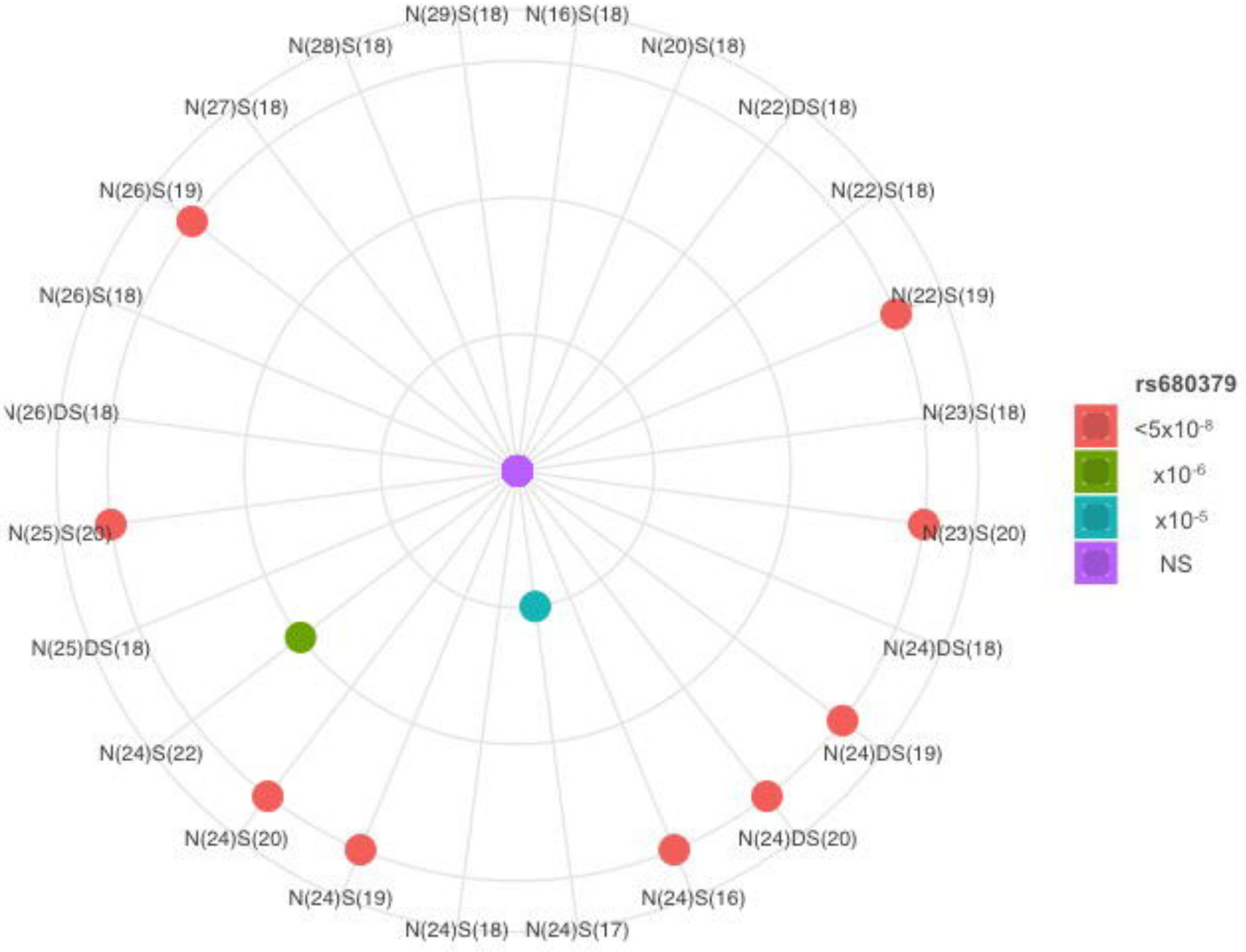
Trait	Mean (SD)
Gender	47% Male
Hypertensive	33%
Mean Blood Pressure	138/83 mmHg
Age (years)	49 (15)
BMI	26.04 (4.33)
WHR	0.86 (0.09)
Cholesterol (mmol/L)	5.61 (1.20)

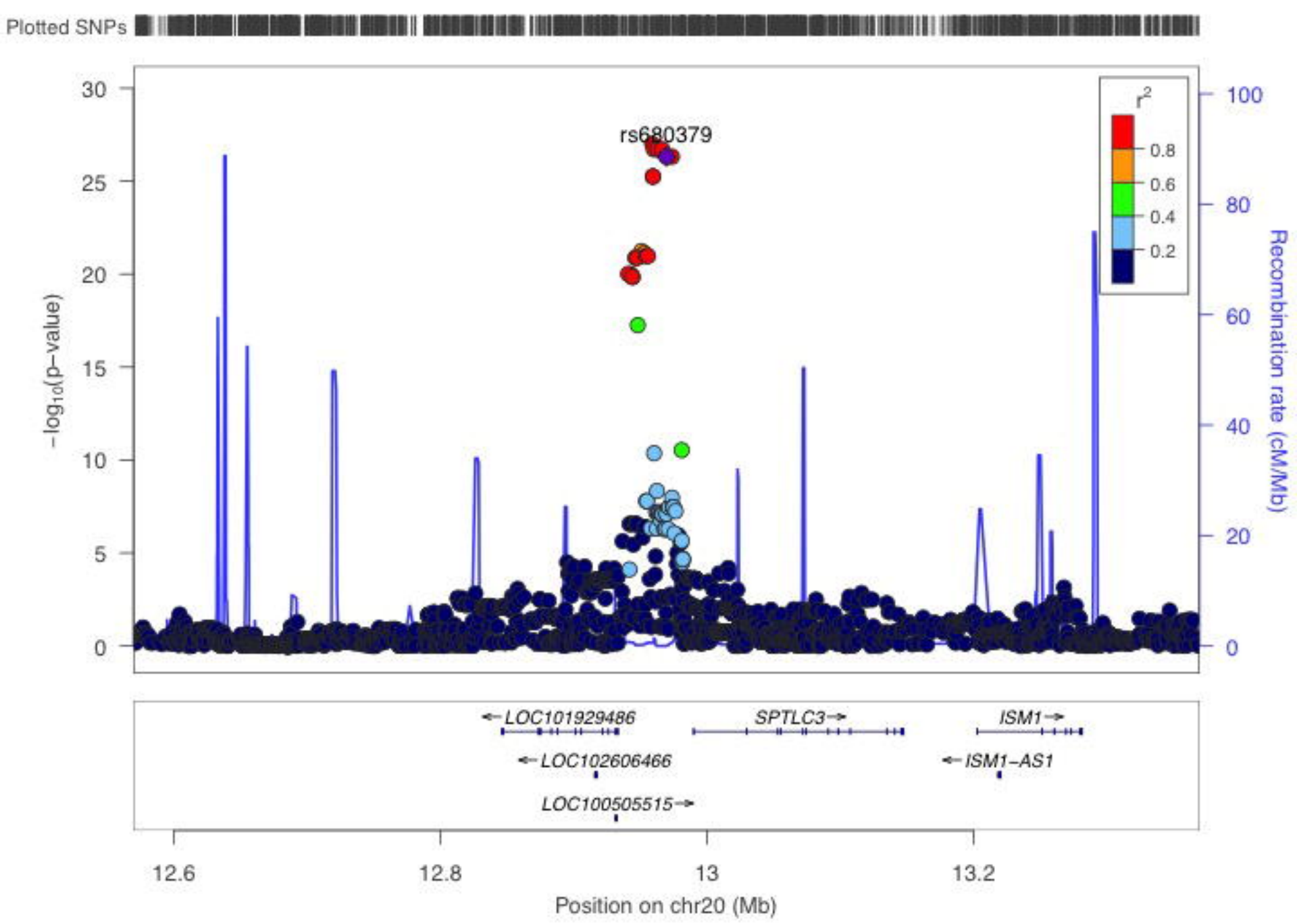

A
NAPE
1. NAPE-PLD
2. PLC, PTPN22
3. ABHD4, GDE1
4. PLD, GDE4, GDE7
NAE
AEA, PEA, OEA, DHEA,
SEA, VEA, HEA, LEA,
DPEA, POEA, PDEA
FAAH
Fatty acid
+
Ethanolamine
B
Palmitoyl-CoA + L-serine
SPTLC 1-3
3-keto dihydrosphingosine
CERS
DEGS1
C18DS
CER[ADS]
CER[AS]
CERS
CER[NDS]
DEGS1
SM
CER[NS]
C18S
C18S1P
SGPP 1-2
SGPP 1-2
C1P


r = 0.99
GCTA
QTDT
1.00
0.75
0.50
0.25
0.00
0.00
0.25
0.50
0.75
1.00
A(24)S(18)
PDEA
C18S1P
C18DS
POEA
HEA
STEA
N(26)S(18)
N(24)S(16)
N(25)S(20)
N(23)S(20)
N(24)S(20)
N(24)S(22)
C18S
VEA
N(24)DS(20)
N(26)DS(18)
DHEA
N(24)DS(19)
N(22)DS(18)
PEA
N(24)S(17)
N(22)S(18)
A(22)S(18)
N(23)S(18)
OEA
DPEA
N(24)S(18)
AEA
A(26)S(18)
N(22)S(19)
N(26)S(19)
LEA
N(25)DS(18)
N(27)S(18)
N(24)DS(18)
N(24)S(19)
N(16)S(18)
N(28)S(18)
N(20)S(18)
N(29)S(18)
Lipid Group:
CER[NS]
NAE
CER[NDS]
CER[AS]
C18 species


VEA
AEA
DHEA
DPEA
HEA
LEA
OEA
PDEA
PEA
POEA
STEA
rs324420
<5x10-8
x10-5
NS


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
Position on chr1 (Mb)


N(29)S(18)
N(16)S(18)
N(28)S(18)
N(20)S(18)
N(27)S(18)
N(22)DS(18)
N(26)S(19)
N(22)S(18)
N(26)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)DS(18)
N(26)DS(18)
rs680379
<5x10-8
x10-6
x10-5
NS


Plotted SNPs
rs680379
r²
0.8
0.6
0.4
0.2
−log10(p−value)
Recombination rate (cM/Mb)
30
25
20
15
10
5
0
100
80
60
40
20
0
LOC101929486
SPTLC3
ISM1
LOC102606466
ISM1−AS1
LOC100505515
12.6
12.8
13
13.2
Position on chr20 (Mb)