Longitudinal Lipidomic Signature of All-cause and CVD Mortality in American Indians: Findings from the Strong Heart Family Study

Guanhong Miao,[1,2] Mingjing Chen,[1,2] Ying Zhang,[3] Jason G. Umans,[4,5] Shelley A. Cole,[6] Elisa T. Lee,[3] Barbara V. Howard,[4] Oliver Fiehn,[7] Jinying Zhao[1,2]

[1]Department of Epidemiology, College of Public Health & Health Professions and College of Medicine, University of Florida, Gainesville, FL

[2]Center for Genetic Epidemiology and Bioinformatics, University of Florida, Gainesville, FL

[3]Department of Biostatistics and Epidemiology, University of Oklahoma Health Sciences Center, Oklahoma City, OK

[4]MedStar Health Research Institute, Hyattsville, MD

[5]Georgetown-Howard Universities Center for Clinical and Translational Science, Washington, DC

[6]Texas Biomedical Research Institute, San Antonio, TX

[7]West Coast Metabolomics Center, University of California-Davis, CA

Correspondence to: Jinying Zhao, MD, PhD
Center for Genetic Epidemiology and Bioinformatics, University of Florida
Department of Epidemiology, College of Public Health and Health Professions and College of Medicine, University of Florida
2004 Mowry Rd, Gainesville, FL 32610
Tel: 352-273-5933; Fax: 352-273-5365
Email: jzhao66@ufl.edu

Early and accurate identification of individuals with high risk of mortality could benefit preventive therapies and intervention before the death. Traditional risk factors such as high bad cholesterol and low good cholesterol have limitations in predicting mortality. To improve quality of life and increase life expectancy, it is important to identify novel biomarkers that can predict people who have high risk of death many years before the mortality. The main purpose of this study is to identify novel lipids that can predict risk of all-cause and CVD mortality using a new technology called lipidomics.

We measured the blood levels of 1,542 lipids in around 2,000 American Indians who attended both the Strong Heart Study (SHS) Phase IV and Phase V. Statistical analyses were performed to identify lipids that can predict the risk of all-cause and CVD mortality during the follow-up (mean 17.8 years). We first did the analyses in the SHS and then verified results in European Caucasians from the Malmö Diet and Cancer-Cardiovascular Cohort (MDC-CC). We found that 20 baseline lipids were associated with risk of all-cause mortality in both cohorts, after taking into consideration of traditional risk factors including age, sex, BMI, smoking, hypertension, diabetes, CVD, low good cholesterol (i.e., LDL-c) and kidney function (i.e., eGFR) at baseline. In addition, three lipids were associated with CVD mortality in both SHS and MDC-CC. We also found the longitudinal change in lipids significantly associated with risk of all-cause and CVD mortality. These results suggest that the newly detected blood lipids are more accurate than traditional risk factors in predicting the risk of all-cause and CVD mortality. Thus, they are likely to serve as new biomarkers for early prevention of all-cause and CVD mortality many years before the death.

Longitudinal Lipidomic Signature of All-cause and CVD Mortality in American Indians: Findings from the Strong Heart Family Study

Guanhong Miao,[1,2] Mingjing Chen,[1,2] Ying Zhang,[3] Jason G. Umans,[4,5] Shelley A. Cole,[6] Elisa T. Lee,[3] Barbara V. Howard,[4] Oliver Fiehn,[7] Jinying Zhao[1,2]

[1]Department of Epidemiology, College of Public Health & Health Professions and College of Medicine, University of Florida, Gainesville, FL

[2]Center for Genetic Epidemiology and Bioinformatics, University of Florida, Gainesville, FL

[3]Department of Biostatistics and Epidemiology, University of Oklahoma Health Sciences Center, Oklahoma City, OK

[4]MedStar Health Research Institute, Hyattsville, MD

[5]Georgetown-Howard Universities Center for Clinical and Translational Science, Washington, DC

[6]Texas Biomedical Research Institute, San Antonio, TX

[7]West Coast Metabolomics Center, University of California-Davis, CA

Correspondence to: Jinying Zhao, MD, PhD
Center for Genetic Epidemiology and Bioinformatics, University of Florida
Department of Epidemiology, College of Public Health and Health Professions and College of Medicine, University of Florida
2004 Mowry Rd, Gainesville, FL 32610
Tel: 352-273-5933; Fax: 352-273-5365
Email: jzhao66@ufl.edu

Abstract

Background: Dyslipidemia contributes to a greater all-cause and CVD mortality, but a comprehensive assessment of molecular lipid species predictive of risk for all-cause and CVD mortality is lacking in large-scale community-dwelling individuals.

Methods: We sought to identify fasting plasma lipids associated with risk of all-cause and CVD mortality among American Indians in the Strong Heart Family Study (SHFS), a large-scale community-dwelling of individuals, followed by replication in European Caucasians from the Malmö Diet and Cancer-Cardiovascular Cohort (MDC-CC). Moreover, we performed repeated measurement analysis to examine the longitudinal change in lipid species (mean ~5.5 years apart) associated with risk of all-cause and CVD mortality. Network analysis was conducted to identify differential lipid coregulations associated with risk of all-cause and CVD mortality.

Results: We found that multiple baseline lipid species, including unsaturated cholesterol esters and lyso-glycerophospholipids, very long-chain phosphatidylcholines, ether-phosphatidylcholines, unsaturated phosphatidylinositols and sphingomyelins, was significantly associated with risk of all-cause mortality in both SHFS and MDC-CC, independent of clinical factors. Phosphatidylethanolamines and lysophosphatidylethanolamine were positively associated with CVD mortality in both cohorts. Longitudinal change in the plasma lipidome was significantly associated with risk of all-cause or CVD mortality after adjusting for covariates and baseline lipids. Network analysis identified highly correlated lipid module differentiating high from low-risk groups.

Conclusions: Altered lipid metabolism precedes all-cause and CVD mortality in American Indians. Our results shed light on the mechanisms through which dyslipidemia contributes to

mortality and provide evidence for targeting lipid metabolism to guide novel preventive interventions.

Keywords: all-cause mortality, CVD mortality, longitudinal lipidomic profiling, American Indians, Strong Heart Study, Malmö Diet and Cancer Study

Introduction

The American Indian or Alaska Native population experienced substantially higher death rates than other populations in the United States.[1-4] Cardiovascular disease (CVD) is the leading causes of deaths in American Indians.[5, 6] Epidemiological studies have shown that individuals with metabolic diseases suffer from higher mortality rate than those without.[7-11] In addition, high levels of low-density lipoprotein cholesterol (LDL-c) and total cholesterol, and a low level of high-density lipoprotein cholesterol (HDL-c) have previously been associated with increased cardiovascular disease (CVD) mortality or all-cause mortality in different populations,[12-16] although results are mixed.[17-20] These findings suggest that dyslipidemia may contribute to all-cause or CVD mortality. Deciphering the metabolic pathways underlying the association between lipid metabolism and mortality is likely to lead to novel biomarkers for risk prediction and stratification. However, routine lipid panel cannot capture all molecular lipids in blood (i.e., blood lipidome), and thus has limited ability to detect perturbed lipid metabolism associated with mortality. Lipidomics is an emerging high-throughput biochemical technique that can identify and quantify hundreds to thousands of lipid species in the biospecimens, and thus is well suited to characterize disturbed lipid metabolism that precedes mortality.

There has been a growing interest in using lipidomics to better understand the associations between altered lipid metabolism and mortality. Several studies have reported associations of altered blood lipid species, such as phosphatidylcholines, ceramides, ratio of polyunsaturated fatty acids to total fatty acids, phosphatidylethanolamines, with all-cause or cause-specific mortality in various populations.[21-26] However, existing studies were either limited to smaller sample size or only examined a smaller number of lipids at a single time point. To date, no study

has investigated the relationship between fasting plasma lipid species and mortality in American Indians. Moreover, no large-scale longitudinal epidemiological studies have examined the association between the longitudinal change in lipid species and mortality in any racial/ethnic group.

Here we report findings from the first longitudinal lipidomic profile in 3,821 fasting blood samples from 1,930 American Indians attending two clinical examinations (mean follow-up: 5.5 years) in the Strong Heart Family Study (SHFS). Our primary goals here are to (1) identify individual lipid species associated with risk of all-cause and CVD mortality beyond traditional risk factors, and (2) examine the association between longitudinal changes in plasma lipidome and risk of all-cause and CVD mortality.

Research design and methods

Study populations

Discovery cohort. The Strong Heart Family Study (SHFS, 2001-ongoing), a family-based prospective study designed to identify genetic and lifestyle factors for CVD and risk factors in American Indians, as previously described.[5, 27, 28] Briefly, 2,780 tribal members (aged 14 and older) in three geographic regions (Arizona, North/South Dakota, Oklahoma) were initially examined in 2001-2003 and re-examined in 2006-2009 (mean 5.5 years apart) using the same protocols. Information for demography, family history, medical records, and lifestyle was collected at each visit. A total of 1,930 individuals (1,203 women, mean age at baseline: 40.4) with complete information for clinical and lipidomic data were included in the current analysis. More information for covariate assessments was described elsewhere.[29] All the participants provided informed consents. The SHFS protocols were approved by the Institutional Review Boards of participating institutions and the American Indian tribes.

Replication cohort. The Malmö Diet and Cancer-Cardiovascular Cohort (MDC-CC), a sub-cohort of the Malmö Diet and Cancer (MDC) Study,[30] is a prospective population-based cohort designed to study the epidemiology of carotid artery disease (CAD) and its risk factors in European Caucasians in Sweden, as previously described.[31, 32] A subset of 3,943 participants (2,319 women, mean age at baseline: 57.7) who had complete information for clinical phenotypes and lipidomic data were included in the analysis. All MDC-CC participants provided written informed consent, and the study was approved by the Ethics Committee at Lund University.

Follow-up and ascertainment of mortality

In the SHFS, baseline information was collected in 2001-2003 and all living participants were followed through 2020. Detailed methods for ascertainment of deaths and causes of deaths in the SHFS have been described previously.[5, 6, 27] Briefly, death of a participant was identified from the Indian Health Service hospital records and direct contact with the family members. The cause of death was determined by physicians on the SHS Mortality Review Committee using medical records, autopsy reports, and informant interview. Information on mortality was retrieved from the mortality decision form used for cardiovascular disease surveillance in the SHS. The cause of death was determined independently by two members of the SHS Mortality Committee after reviewing information including death certificates and medical records including pathology reports as well as informant interviews, when needed. Death certificate codes were recorded according to the International Classification of Diseases Injuries, and Causes of Death, 9th Revision (ICD-9).[5] CVD mortality was defined as death caused by myocardial infarction, stroke, sudden death from coronary heart disease, or congestive heart

failure (ICD-9 codes 390-448). Cancer mortality was defined as death with a primary cause of death listed as malignant neoplasm (ICD-9 codes 140-239).

In the MDC-CC, all subjects were followed from the baseline examination until death, emigration from Sweden, or December 31 2020, whichever came first. Causes of death and the vital status of participants were retrieved through record linkage of the personal identification number and the Swedish Cause of Death Register and the National Tax Board.[33, 34] CVD mortality was defined as death caused by CVD on the basis of ICD-9 codes 390-459 or ICD-10 codes I00-I99. Cancer mortality was defined based on ICD-9 codes 140-239 or ICD-10 codes D00-D48.

Assessments of clinical covariates

In the SHFS, demographic information (age and sex), lifestyle habits (smoking/drinking status, physical activity), medical history, family history of illnesses, and use of prescription medications were collected using structured questionnaires.[5, 28] Smoking status will be categorized as current versus non-current smokers (former and never smokers combined). Drinking status will be categorized as current versus non-current drinkers (former and never drinkers combined). Anthropometric measures including height, weight and waist circumferences were obtained through physical examinations at each visit. Body mass index (BMI) was calculated as body weight in kilograms divided by the square of height in meters. Fasting glucose and blood lipids, including total cholesterol, triglycerides, LDL-c and HDL-c were measured by standard laboratory methods.[5] Hypertension will be defined as blood pressure levels ≥140/90 mmHg or use of antihypertension medications. Type 2 diabetes will be defined as fasting plasma glucose ≥126 mg/dL or use of hypoglycemic medications. Estimated glomerular filtration rate (eGFR) was calculated using the chronic kidney disease (CKD)

Epidemiology Collaboration (CKD-EPI).[35] Physical activity was assessed by the mean number of steps per day calculated by averaging the total number of steps recorded each day during the 7-day period. Information on use of lipid-lowering drugs was also collected at each visit.

In the MDC-CC, information on lifestyle and clinical factors, such as BMI, blood pressure, use of antihypertensive treatment, current smoking, clinical lipids, and blood glucose, were obtained as previously described.[36]

Lipidomic data acquisition, pre-processing and quality control

In the SHFS, relative abundance of fasting plasma lipid species at two time points (mean follow-up: 5.5 years) was quantified by untargeted liquid chromatography-mass spectrometry (LC-MS) as described previously.[29] After pre-processing and quality control, we obtained 1,542 lipids (518 known, 1,024 unknown) in 3,977 plasma samples (1,983 at baseline, 1,994 at follow-up). After further excluding outlier samples or those with missing covariates, 1,930 participants (1,930 at baseline, 1,891 at follow-up) with complete clinical and lipidomic data were included in the analysis (**Figure S1**).

Lipidomic profiling in the MDC-CC was performed using fasting plasma samples collected at enrollment (i.e., baseline) by mass spectrometry as previously described.[37, 38] Spectra were analyzed with in-house developed lipid identification software based on LipidXplorer.[39] Data processing, normalization and batch correction was performed as described previously.[31, 38, 40] Lipids that were present in at least 70% of the participants were included in the subsequent statistical analysis, resulting in a total of 184 lipids.[31, 38] Missing values of the remaining lipids were imputed using the NIPALS algorithm.[41] A total of 3,943 participants (58.8% females) with available lipidomic data were included in the current study to replicate findings

from the SHFS. The mean age of MDC-CC participants was 57.7 years at baseline and 81.3 years at follow-up (mean 23.7 years of follow-up).

Statistical analyses

Figure S1 illustrates the procedures for participant selection and statistical analyses. All the continuous variables including lipids were standardized to zero mean and unit variance. Multiple testing was controlled by false discovery rate (FDR) using the Storey's q-value method.[42, 43]

Prospective association analysis. To identify baseline plasma lipids predictive of risk for all-cause or CVD mortality, we constructed Cox proportional hazards model (SHFS, n=1,930; MDC-CC, n=3,943), in which baseline lipid was the predictor and time to all-cause or CVD mortality was the outcome, adjusting for age, sex, BMI, smoking, hypertension, diabetes, cardiovascular disease and LDL-c at baseline. The analysis in the SHFS additionally adjusted for eGFR. Family relatedness in the SHFS was accounted for by including a frailty term in the model. The analysis was first conducted in the SHFS, followed by replication of top hits ($P<0.05$) in the MDC-CC. Replication was defined as lipids with $P<0.05$ and consistent directions of association in both cohorts. Results from two cohorts were combined by random-effects meta-analysis.

Repeated measurement analysis. To examine whether longitudinal change in plasma lipidome was associated with risk of all-cause or CVD mortality, we constructed Cox proportional hazards model with a frailty term accounting for the family relatedness. In the model, time to all-cause or CVD mortality was the outcome and change in the relative abundance of lipid was the predictor, adjusting for age, sex, smoking, cardiovascular disease at baseline and longitudinal change in continuous traits (i.e., BMI, fasting glucose, systolic blood pressure, LDL-c, eGFR) plus baseline

lipid. This analysis was performed in the SHFS only as MDC-CC only measured blood lipid species at baseline.

Differential lipid network analysis. To identify lipid networks (i.e., sets of lipids that are highly correlated) associated with all-cause or CVD mortality, we constructed lipid modules (subnetworks) using the Weighted Correlation Network Analysis (WGCNA).[44] Briefly, signed weighted lipid co-regulation networks were constructed using all 1,542 baseline lipids among individuals died of all causes, those died of CVD and survivors, separately. Lipid species were hierarchically clustered, and those with a high topological overlap similarity were grouped into a same module. Differential modular analysis was performed to dissect intra-module difference (i.e., difference of connectivity among lipids within a module) between individuals died of all causes and survivors, and those died of CVD and survivors. To quantify the intra-module difference, we calculated modular differential connectivity (MDC),[45, 46] i.e., the difference in the total connectivity of all lipid pairs for a specific lipid module between individuals died of all causes and survivors, and those died of CVD and survivors. Gain of connectivity (GOC) was defined if MDC>0 and loss of connectivity (LOC) if MDC<0. Statistical significance of MDC was assessed by 1,000 permutation tests.[47, 48]

Sensitivity analysis. To examine whether the level of physical activity (steps/day), use of lipid-lowering drugs (yes/no) or diet quality (assessed by the Alternate Healthy Eating Index[49]) affect our results, we conducted sensitivity analysis by additionally adjusting for these variables in the above described prospective analysis. This analysis focused on lipids that are significant in both SHFS and MDC-CC.

Results

In the SHFS, 295 out of 1,930 participants died during the follow-up (mean 17.8 years). Among these, 66 participants died of CVD. In the MDC-CC, 1,845 out of 3,943 participants died during the follow-up (mean 23.7 years) with 566 of them died of CVD. The age-standardized all-cause and CVD mortality rates in the SHFS are 18.2% and 4.4%, respectively. The MDC-CC has the age-standardized rate of all-cause mortality 22.9% and CVD mortality 12.5%. **Table 1** presents baseline characteristics of participants in these two populations. In both cohorts, those who died of all causes or CVD only were significantly older, and more likely to be male, and had higher hypertension and diabetes rates, and higher baseline levels of systolic blood pressure, total cholesterol, triglycerides and fasting glucose compared to survivors.

Baseline plasma lipid species predict risk of all-cause mortality. In the SHFS (discovery stage), 552 baseline lipids (192 known) were associated with all-cause mortality at $P<0.05$ (**Figure 1**, **Table 2**). Of these, higher baseline levels of 9 acylcarnitines, 3 cholesterol esters (i.e., CE(16:1), CE(18:1), CE(20:4)), 6 fatty acids, 45 phosphatidylcholines (PCs), 17 phosphatidylethanolamines (PEs), 1 phosphatidylglycerol (PG), 9 phosphatidylinositols (PIs), 1 phosphatidylserine (PS), 3 diacylglycerols (DAGs), 16 sphingomyelins (SMs), 6 ceramides (CERs), 3 glucosylceramides (GlcCers), and 1 lactosylceramide (LacCer) were associated with an increased risk of all-cause mortality. In contrast, CE(18:0), LPE(20:6), 19 phosphatidylcholines, 1 diacylglycerol, 30 triacylglycerols, 2 ceramides, and 18 sphingomyelins were inversely associated with risk of all-cause mortality. 183 known lipids remained significant after multiple testing ($q<0.05$) and only one lipid class, phosphatidylserine, became non-significant.

Of the 192 known lipid species identified in the SHFS ($P<0.05$), 58 lipids were also available in the MDC-CC. Of these, 15 baseline lipids, including CE(16:1), CE(18:1), LPE(18:2), 9

glycerophospholipids (e.g., PC(32:0), PC(34:1), PC(38:2)) and 3 phosphatidylinositols, were positively, whereas 5 lipids, including PC(38:6) B and 4 sphingomyelins (i.e., SM(d38:1) B, SM(d38:2) A, SM(d40:1) B, SM(d40:2) B), were inversely associated with all-cause mortality ($q<0.05$) in the MDC-CC with same direction of association.

Transethnic meta-analysis identified 25 lipids significantly associated with risk of all-cause mortality ($P<0.05$). After multiple testing, 2 cholesterol esters (i.e., CE(16:1), CE(18:1)), 5 phosphatidylcholines (i.e., LPC(20:1), PC(32:1), PC(34:1), PC(34:3) A, PC(36:1)), LPE(18:2) and 3 phosphatidylinositols (i.e., PI(18:0/18:1), PI(18:0/18:2), PI(18:0/20:3) A) were positively, whereas PC(38:6) B and 3 sphingomyelins (i.e., SM(d38:2) A, SM(d38:1) B, SM(d40:2) B) were inversely associated with risk of all-cause mortality at $q<0.05$. **Figure S2** and **Table 2** schematically illustrate the associations between baseline lipid species and risk of all-cause mortality in both cohorts.

Baseline plasma lipid species predict risk of CVD mortality. In the SHFS, 328 baseline lipids (106 known) were associated with CVD mortality at $P<0.05$ (**Figure 1**, **Table 3**). Of these, higher baseline levels of 3 acylcarnitines (i.e., AC(18:0), AC(24:0), AC(26:0)), 2 cholesterol esters (i.e., CE(22:5) A, CE(22:5) B), 68 glycerophospholipids (e.g., 43 phosphatidylcholines (PCs), 17 phosphatidylethanolamines (PEs), 8 phosphatidylinositols (PIs)), and 31 sphingolipids (e.g., 15 sphingomyelins (SMs), 10 ceramides (CERs), 5 glucosylceramides (GlcCers), 1 lactosylceramide (LacCer)) were associated with an increased risk of all-cause mortality. In contrast, LPC(22:5) and LPG(17:0) were inversely associated with risk of CVD mortality. After multiple testing, 10 lipids (i.e., AC(26:0), CER(d42:2) B, GlcCer(d40:1), LPC(24:0), PC(p-14:0/22:1)/PC(o-14:0/22:1), PC(p-36:1)/PC(o-36:2) A, PE(16:0/20:5), PE(p-36:2)/PE(o-36:3) B, SM(d34:0) A, SM(d34:1) A) were positively associated with CVD mortality at $q<0.05$.

Of the 106 known lipid species identified in the SHFS, 34 lipids were also available in the MDC-CC. Of these, 3 baseline lipids (i.e., LPE(18:2), PE(18:1/18:1), PE(38:4)) were positively associated with CVD mortality ($P<0.05$) in the MDC-CC with same direction of association. One lipid (i.e., PE(18:1/18:1)) remained significant with $q<0.05$.

Transethnic meta-analysis identified baseline levels of 5 lipids (i.e., LPE(18:2), PE(18:1/18:1), PE(38:4), PI(18:0/18:2), PI(18:0/18:1)) associated with an increased risk of CVD mortality ($P<0.05$). Of these, 3 lipids (LPE(18:2), PE(18:1/18:1), PE(38:4)) reached $q<0.05$. **Figure S3** and **Table 3** present the associations between baseline lipids and risk of CVD mortality in both cohorts.

Figure S4 shows 57 baseline lipids associated with risk of both all-cause and CVD mortality ($P<0.05$). Specifically, higher baseline levels of 2 acylcarnitines, 18 sphingolipids (e.g., sphingomyelins, ceramides, glucosylceramides, lactosylceramide) and 35 glycerophospholipids (e.g., phosphatidylcholines, phosphatidylethanolamines, phosphatidylinositols) were associated with an increased risk of both all-cause and CVD mortality. In contrast, two lipids, CER(d40:2) and SM(d36:3) B, were inversely associated with risk of all-cause mortality but positively associated with CVD mortality.

Longitudinal change in plasma lipid species predict risk of all-cause mortality. After adjusting for covariates and baseline lipids, longitudinal changes in 624 lipids (224 known) were associated with risk of all-cause mortality at $P<0.05$. Of these, changes in 7 acylcarnitines, 10 fatty acids, 50 glycerophospholipids (e.g., 29 phosphatidylcholines, 14 phosphatidylethanolamines, 7 phosphatidylinositols) and 16 sphingolipids (e.g., 11 sphingomyelins, 3 ceramides, 1 glucosylceramide, 1 lactosylceramide) were positively associated with risk of all-cause mortality. Changes in 2 cholesterol esters, 46

glycerophospholipids (e.g., 42 phosphatidylcholines, 1 phosphatidylethanolamine, 3 phosphatidylinositols), 36 sphingolipids (e.g., 31 sphingomyelins, 5 ceramides), 57 glycerolipids (5 diacylglycerols, 52 triacylglycerols) were inversely associated with risk of all-cause mortality. All the 224 known lipids reached q<0.05 after multiple testing (**Figure 2**, **Table 4**).

Longitudinal change in plasma lipid species predict risk of CVD mortality. Longitudinal changes in 201 lipids (60 known) were associated with risk of CVD mortality at P<0.05 (**Figure 2**, **Table 4**). Specifically, changes in 3 acylcarnitines, 8 fatty acids, 1 lactosylceramide, 3 sphingomyelins, 14 phosphatidylcholines, 2 phosphatidylethanolamines were positively, whereas changes in 2 cholesterol esters, 1 glucosylceramide, 12 phosphatidylcholines, 4 phosphatidylethanolamines, 1 phosphatidylinositol and 9 sphingomyelins were inversely associated with risk of CVD mortality. After multiple testing, only CE(18:0) and SM(d42:1) B remained significant at q<0.05.

Of these identified lipids, changes in 43 known lipids were associated with both all-cause mortality and CVD mortality with P<0.05 (**Figure S5)**. Changes in 2 acylcarnitines, 6 fatty acids, 12 phosphatidylcholines, 2 phosphatidylethanolamines, 1 sphingomyelin and 1 lactosylceramide were positively, whereas 9 phosphatidylcholines, 9 sphingomyelins and 1 phosphatidylethanolamine were inversely associated with risk of both all-cause mortality and CVD mortality.

Differential lipid networks associated with risk of all-cause and CVD mortality. Network analysis in the SHFS identified 9, 11 and 12 lipid modules among participants died of all causes, CVD and survivors, respectively (**Figure S6**). The connectivity (i.e., corrections) between lipids in one of the modules (i.e., module turquoise) exhibited significant difference between all-cause mortality and survivors (**Figure S7**). Lipids in this module include ceramides, cholesterol ester,

phosphatidylcholines, phosphatidylinositols, and sphingomyelins. Compared to survivors, all-cause mortality exhibited gain of connectivity (GOC) for lipids in the module turquoise (modular differential connectivity (MDC)=202.2, P=0.046). Hub lipids in this module included SM(d17:0/18:2) A, SM(d39:1) A, SM(d41:1) A and SM(d38:1) A. The connectivity between lipids in one of the modules (i.e., module green) exhibited significant difference between CVD mortality and survivors (**Figure S8**). Lipids in this module include cholesterol esters, fatty acids, glucosylceramides, phosphatidylcholines, and sphingomyelins. Compared to survivors, CVD mortality exhibited gain of connectivity (GOC) for lipids in the module green (modular differential connectivity (MDC)=92.9, P=0.026). Hub lipids in this module included PC(p-14:0/26:2)/PC(o-14:1/26:2), GlcCer(d40:1), PC(p-16:1/26:2)/PC(o-16:2/26:2) and PC(p-16:0/26:2)/PC(o-16:1/26:2).

Sensitivity analysis. Additional adjustments for the level of physical activity, use of lipid-lowering drugs and diet quality slightly attenuated the associations between identified lipids and risk of all-cause or CVD mortality, but most lipids remained to be significant (**Table 5**).

Discussion

In this large-scale longitudinal lipidomic profiling from diverse cohorts, we had several key findings. First, multiple lipid species were associated with risk of all-cause or CVD mortality beyond clinical factors. Specifically, prospective association analysis demonstrated that higher baseline levels of 9 acylcarnitines, 3 cholesterol esters, 3 diacylglycerols, 6 fatty acids, 73 glycerophospholipids, and 26 sphingolipids, and lower baseline levels of CE(18:0), 20 glycerophospholipids, 31 glycerolipids, and 20 sphingolipids were associated with increased risk of all-cause mortality in American Indians, independent of clinical factors. Of these, 2 unsaturated cholesterol esters (i.e., CE(16:1), CE(18:1)), 2 unsaturated lyso-

glycerophospholipids (i.e., LPC(20:1), LPE(18:2)), 7 very long-chain phosphatidylcholines (e.g., PC(32:0), PC(34:1), PC(38:2)), 2 very long-chain ether-phosphatidylcholines (i.e., PC(o-32:0), PC(p-34:0)/PC(o-34:1)), 3 very long-chain unsaturated phosphatidylinositols (i.e., PI(18:0/18:1), PI(18:0/18:2), PI(18:0/20:3) A), and 4 very long-chain unsaturated sphingomyelins (i.e., SM(d38:1) B, SM(d38:2) A, SM(d40:1) B, SM(d40:2) B) were replicated (with same direction of association) in the MDC-CC. Moreover, increased risk of CVD mortality was associated with higher baseline levels of 3 acylcarnitines, 2 cholesterol esters, 43 phosphatidylcholines, 17 phosphatidylethanolamines, 8 phosphatidylinositols, 15 sphingomyelins, 10 ceramides, 5 glucosylceramides, 1 lactosylceramide and lower levels of 2 lyso-glycerophospholipids (i.e., LPC(22:5), LPG(17:0)). Of these, 3 lipids (i.e., LPE(18:2), PE(18:1/18:1), PE(38:4)) were replicated in the MDC-CC with consistent direction. Second, our repeated measurement analysis identified, for the first time, the longitudinal change in fasting plasma lipidome predictive of risk for all-cause and CVD mortality, independent of clinical factors and baseline lipids. Longitudinal changes in 224 known lipids largely included acylcarnitines, fatty acids, phosphatidylcholines, phosphatidylethanolamines, sphingomyelins, and triacylglycerols predicted risk of all-cause mortality. Changes in 60 known lipids, largely in fatty acids, phosphatidylcholines, phosphatidylethanolamines, and sphingomyelins, were significantly associated with risk of CVD mortality. Third, our network analysis identified differential lipid clusters (i.e., modules) associated with risk of all-cause and CVD mortality. Together, our results revealed a distinct lipidomic signature associated with all-cause or CVD mortality and provide insight into the mechanisms through which dyslipidemia contributes to the mortality. Moreover, because aberrant expression of plasma lipidome associated with mortality is clearly present years prior to

the death, the identified lipids may provide potential biomarkers for identifying individuals with high risk of death at earlier stages.

It is challenging to replicate individual lipid species across populations.[50] First, different mass spectrometry platforms employed in different studies usually result in different resolution and/or coverage of the lipidome. Second, blood lipidome is regulated by genetic, lifestyle and environmental factors which may vary across populations. Moreover, American Indians in the SHFS were younger than participants in the MDC-CC, the lipidomic findings in the SHFS may reflect early perturbation of lipid metabolism compared to MDC-CC. Although these factors all contribute to the difficulties of individual lipid species replication, we found that around one third of the significant lipids predictive of all-cause mortality in the SHFS were replicated in the MDC-CC, suggesting the robustness of our results.

Several limitations of our study should be noted. First, despite the large number of lipids detected in our discovery cohort, many lipids are unknown and we were unable to distinguish isomeric lipids either. Additional experiments are needed to characterize these unknowns if considered of interest. Also, the lack of absolute quantification does not allow clinical utility. Moreover, the static lipidomic platforms may not disclose the contribution of upstream lipid regulators with differential flux and yet steady state concentrations. Second, our findings were obtained from a cohort of American Indians with a high rate of diabetes. However, all our analyses adjusted for diabetes status and excluding diabetic participants did not affect our results. Moreover, diabetes prevalence is rising in many other populations, suggesting that our findings may be replicated in other settings. Third, although our statistical models controlled many clinical factors known to be associated with mortality, we cannot exclude the possibility of residual confounding by unknown or unmeasured factors. Finally, as all other observational

studies, the prospective associations of perturbed lipids with risk of mortality identified in our study may not necessarily be causal because baseline factors increasing risk of mortality, either unmeasured or imperfectly measured, may influence this apparently “causal” relationship.

However, our study has several strengths. First, the current study included nearly 6,000 community-based individuals of diverse cohorts. Second, the longitudinal profiling of plasma lipidome in a large community-based cohort represents the major strength of this study. Third, our statistical analyses in the SHFS and MDC-CC adjusted for chronic conditions associated with mortality (e.g., obesity, diabetes and hypertension). The effect of confounders on comparison of two ethnic groups was minimized by using the same model and adjusting for the same covariates. Moreover, we performed sensitivity analysis to additionally adjust for the level of physical activity, the use of lipid-lowering medications, and diet quality in the SHFS. Thus, lipids identified in our study should be independent of these conventional risk factors. Finally, our high-resolution lipidomic platforms in both discovery and replication cohorts identified a larger number of molecular lipid species, allowing us to identify novel lipid species associated with risk of mortality and offering unprecedented opportunities for future investigations.

In summary, we identified a range of novel molecular lipids associated with risk of all-cause and CVD mortality beyond clinical factors. Our results demonstrated that dysregulated lipid metabolism occurs years before mortality. Thus, the newly identified lipids may help identify individuals with risk of all-cause and CVD mortality years before the death. Our findings offer potential opportunities for new intervention strategies (e.g., lifestyle/drug) to prevent mortality by modifying lipid metabolism. Our results further highlight the need for mechanistic studies to characterize the role of lipid species in mortality.

REFERENCES

[1] (2019) Disparities fact sheet. Available from https://www.ihs.gov/newsroom/factsheets/disparities/. Accessed October 7 2022

[2] Arias E, Xu J, Curtin S, Bastian B, Tejada-Vera B (2021) Mortality Profile of the Non-Hispanic American Indian or Alaska Native Population, 2019. Natl Vital Stat Rep 70(12): 1-27

[3] Ahmad FB, Cisewski JA, Miniño A, Anderson RN (2021) Provisional Mortality Data - United States, 2020. MMWR Morb Mortal Wkly Rep 70(14): 519-522. 10.15585/mmwr.mm7014e1

[4] Espey DK, Jim MA, Cobb N, et al. (2014) Leading causes of death and all-cause mortality in American Indians and Alaska Natives. Am J Public Health 104 Suppl 3(Suppl 3): S303-311. 10.2105/AJPH.2013.301798

[5] Lee ET, Welty TK, Fabsitz R, et al. (1990) The Strong Heart Study. A study of cardiovascular disease in American Indians: design and methods. Am J Epidemiol 132(6): 1141-1155. 10.1093/oxfordjournals.aje.a115757

[6] Lee ET, Cowan LD, Welty TK, et al. (1998) All-cause mortality and cardiovascular disease mortality in three American Indian populations, aged 45-74 years, 1984-1988. The Strong Heart Study. Am J Epidemiol 147(11): 995-1008. 10.1093/oxfordjournals.aje.a009406

[7] Guize L, Thomas F, Pannier B, Bean K, Jego B, Benetos A (2007) All-cause mortality associated with specific combinations of the metabolic syndrome according to recent definitions. Diabetes Care 30(9): 2381-2387. 10.2337/dc07-0186

[8] Gu K, Cowie CC, Harris MI (1998) Mortality in adults with and without diabetes in a national cohort of the U.S. population, 1971-1993. Diabetes Care 21(7): 1138-1145. 10.2337/diacare.21.7.1138

[9] Roglic G, Unwin N, Bennett PH, et al. (2005) The burden of mortality attributable to diabetes: realistic estimates for the year 2000. Diabetes Care 28(9): 2130-2135. 10.2337/diacare.28.9.2130

[10] Lakka HM, Laaksonen DE, Lakka TA, et al. (2002) The metabolic syndrome and total and cardiovascular disease mortality in middle-aged men. JAMA 288(21): 2709-2716. 10.1001/jama.288.21.2709

[11] Dunn W, Xu R, Wingard DL, et al. (2008) Suspected nonalcoholic fatty liver disease and mortality risk in a population-based cohort study. Am J Gastroenterol 103(9): 2263-2271. 10.1111/j.1572-0241.2008.02034.x

[12] Wadhera RK, Steen DL, Khan I, Giugliano RP, Foody JM (2016) A review of low-density lipoprotein cholesterol, treatment strategies, and its impact on cardiovascular disease morbidity and mortality. J Clin Lipidol 10(3): 472-489. 10.1016/j.jacl.2015.11.010

[13] Lewington S, Whitlock G, Clarke R, et al. (2007) Blood cholesterol and vascular mortality by age, sex, and blood pressure: a meta-analysis of individual data from 61 prospective studies with 55,000 vascular deaths. Lancet 370(9602): 1829-1839. 10.1016/S0140-6736(07)61778-4

[14] Silbernagel G, Schöttker B, Appelbaum S, et al. (2013) High-density lipoprotein cholesterol, coronary artery disease, and cardiovascular mortality. Eur Heart J 34(46): 3563-3571. 10.1093/eurheartj/eht343

[15] Hunt KJ, Resendez RG, Williams K, Haffner SM, Stern MP, Study SAH (2004) National Cholesterol Education Program versus World Health Organization metabolic syndrome in relation to all-cause and cardiovascular mortality in the San Antonio Heart Study. Circulation 110(10): 1251-1257. 10.1161/01.CIR.0000140762.04598.F9

[16] Sierra-Johnson J, Fisher RM, Romero-Corral A, et al. (2009) Concentration of apolipoprotein B is comparable with the apolipoprotein B/apolipoprotein A-I ratio and better than routine clinical lipid measurements in predicting coronary heart disease mortality: findings from a multi-ethnic US population. Eur Heart J 30(6): 710-717. 10.1093/eurheartj/ehn347

[17] Ravnskov U, Diamond DM, Hama R, et al. (2016) Lack of an association or an inverse association between low-density-lipoprotein cholesterol and mortality in the elderly: a systematic review. BMJ Open 6(6): e010401. 10.1136/bmjopen-2015-010401

[18] Briel M, Ferreira-Gonzalez I, You JJ, et al. (2009) Association between change in high density lipoprotein cholesterol and cardiovascular disease morbidity and mortality: systematic review and meta-regression analysis. BMJ 338: b92. 10.1136/bmj.b92

[19] Horwich TB, Hamilton MA, Maclellan WR, Fonarow GC (2002) Low serum total cholesterol is associated with marked increase in mortality in advanced heart failure. J Card Fail 8(4): 216-224. 10.1054/jcaf.2002.0804216

[20] Kaysen GA, Ye X, Raimann JG, et al. (2018) Lipid levels are inversely associated with infectious and all-cause mortality: international MONDO study results. J Lipid Res 59(8): 1519-1528. 10.1194/jlr.P084277

[21] Sigruener A, Kleber ME, Heimerl S, Liebisch G, Schmitz G, Maerz W (2014) Glycerophospholipid and sphingolipid species and mortality: the Ludwigshafen Risk and Cardiovascular Health (LURIC) study. PLoS One 9(1): e85724. 10.1371/journal.pone.0085724

[22] Deelen J, Kettunen J, Fischer K, et al. (2019) A metabolic profile of all-cause mortality risk identified in an observational study of 44,168 individuals. Nat Commun 10(1): 3346. 10.1038/s41467-019-11311-9

[23] Qin M, Zhu Q, Lai W, et al. (2020) Insights into the prognosis of lipidomic dysregulation for death risk in patients with coronary artery disease. Clin Transl Med 10(5): e189. 10.1002/ctm2.189

[24] 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(6): 854-863. 10.1161/CIRCGENETICS.114.000600

[25] Walker ME, Xanthakis V, Peterson LR, et al. (2020) Dietary Patterns, Ceramide Ratios, and Risk of All-Cause and Cause-Specific Mortality: The Framingham Offspring Study. J Nutr 150(11): 2994-3004. 10.1093/jn/nxaa269

[26] Peterson LR, Xanthakis V, Duncan MS, et al. (2018) Ceramide Remodeling and Risk of Cardiovascular Events and Mortality. J Am Heart Assoc 7(10). 10.1161/JAHA.117.007931

[27] Howard BV, Lee ET, Cowan LD, et al. (1999) Rising tide of cardiovascular disease in American Indians. The Strong Heart Study. Circulation 99(18): 2389-2395. 10.1161/01.cir.99.18.2389

[28] North KE, Howard BV, Welty TK, et al. (2003) Genetic and environmental contributions to cardiovascular disease risk in American Indians: the strong heart family study. Am J Epidemiol 157(4): 303-314. 10.1093/aje/kwf208

[29] Miao G, Zhang Y, Huo Z, et al. (2021) Longitudinal Plasma Lipidome and Risk of Type 2 Diabetes in a Large Sample of American Indians With Normal Fasting Glucose: The Strong Heart Family Study. Diabetes Care 44(12): 2664-2672. 10.2337/dc21-0451

[30] Berglund G, Elmstähl S, Janzon L, Larsson SA (1993) The Malmo Diet and Cancer Study. Design and feasibility. J Intern Med 233(1): 45-51. 10.1111/j.1365-2796.1993.tb00647.x

[31] Ottosson F, Emami Khoonsari P, Gerl MJ, Simons K, Melander O, Fernandez C (2021) A plasma lipid signature predicts incident coronary artery disease. Int J Cardiol 331: 249-254. 10.1016/j.ijcard.2021.01.059

[32] Fernandez C, Sandin M, Sampaio JL, et al. (2013) Plasma lipid composition and risk of developing cardiovascular disease. PLoS One 8(8): e71846. 10.1371/journal.pone.0071846

[33] Hallengren E, Almgren P, Engström G, et al. (2014) Fasting levels of high-sensitivity growth hormone predict cardiovascular morbidity and mortality: the Malmö Diet and Cancer study. J Am Coll Cardiol 64(14): 1452-1460. 10.1016/j.jacc.2014.03.063

[34] Drake I, Gullberg B, Sonestedt E, et al. (2013) Scoring models of a diet quality index and the predictive capability of mortality in a population-based cohort of Swedish men and women. Public Health Nutr 16(3): 468-478. 10.1017/S1368980012002789

[35] Lee JS, Chang PY, Zhang Y, Kizer JR, Best LG, Howard BV (2017) Triglyceride and HDL-C Dyslipidemia and Risks of Coronary Heart Disease and Ischemic Stroke by Glycemic Dysregulation Status: The Strong Heart Study. Diabetes Care 40(4): 529-537. 10.2337/dc16-1958

[36] Rosvall M, Ostergren PO, Hedblad B, Isacsson SO, Janzon L, Berglund G (2000) Occupational status, educational level, and the prevalence of carotid atherosclerosis in a general population sample of middle-aged Swedish men and women: results from the Malmö Diet and Cancer Study. Am J Epidemiol 152(4): 334-346. 10.1093/aje/152.4.334

[37] Surma MA, Herzog R, Vasilj A, et al. (2015) An automated shotgun lipidomics platform for high throughput, comprehensive, and quantitative analysis of blood plasma intact lipids. Eur J Lipid Sci Technol 117(10): 1540-1549. 10.1002/ejlt.201500145

[38] Lauber C, Gerl MJ, Klose C, Ottosson F, Melander O, Simons K (2022) Lipidomic risk scores are independent of polygenic risk scores and can predict incidence of diabetes and cardiovascular disease in a large population cohort. PLoS Biol 20(3): e3001561. 10.1371/journal.pbio.3001561

[39] Herzog R, Schuhmann K, Schwudke D, et al. (2012) LipidXplorer: a software for consensual cross-platform lipidomics. PLoS One 7(1): e29851. 10.1371/journal.pone.0029851

[40] Fernandez C, Surma MA, Klose C, et al. (2020) Plasma Lipidome and Prediction of Type 2 Diabetes in the Population-Based Malmö Diet and Cancer Cohort. Diabetes Care 43(2): 366-373. 10.2337/dc19-1199

[41] Tenenhaus M (1998) La régression PLS: théorie et pratique. Editions Technip

[42] Storey JD (2002) A Direct Approach to False Discovery Rates. Journal of the Royal Statistical Society Series B (Statistical Methodology) 64(3): 479-498

[43] Storey John D, Tibshirani R (2003) Statistical significance for genomewide studies. Proceedings of the National Academy of Sciences 100(16): 9440-9445. 10.1073/pnas.1530509100

[44] Langfelder P, Horvath S (2008) WGCNA: an R package for weighted correlation network analysis. BMC Bioinformatics 9: 559. 10.1186/1471-2105-9-559

[45] Horvath S, Dong J (2008) Geometric interpretation of gene coexpression network analysis. PLoS Comput Biol 4(8): e1000117. 10.1371/journal.pcbi.1000117

[46] Langfelder P, Luo R, Oldham MC, Horvath S (2011) Is my network module preserved and reproducible? PLoS Comput Biol 7(1): e1001057. 10.1371/journal.pcbi.1001057

[47] Zhang B, Gaiteri C, Bodea LG, et al. (2013) Integrated systems approach identifies genetic nodes and networks in late-onset Alzheimer's disease. Cell 153(3): 707-720. 10.1016/j.cell.2013.03.030

[48] Peschel S, Müller CL, von Mutius E, Boulesteix AL, Depner M (2021) NetCoMi: network construction and comparison for microbiome data in R. Brief Bioinform 22(4). 10.1093/bib/bbaa290

[49] Chiuve SE, Fung TT, Rimm EB, et al. (2012) Alternative dietary indices both strongly predict risk of chronic disease. J Nutr 142(6): 1009-1018. 10.3945/jn.111.157222

[50] Zeng W, Beyene HB, Kuokkanen M, et al. (2022) Lipidomic profiling in the Strong Heart Study identified American Indians at risk of chronic kidney disease. Kidney Int. 10.1016/j.kint.2022.06.023

Figure S1. Flowchart describing the procedures for participants' selection and statistical analyses. We used participants in the SHFS as discovery sample and participants in the MDC-CC as replication sample. To identify baseline plasma lipids significantly associated with all-cause or CVD mortality, we first conducted Cox proportional hazard frailty model in the SHFS, adjusting for age, sex, BMI, smoking, hypertension, diabetes, cardiovascular disease, LDL-c and eGFR at baseline. The frailty term accounted for the family relatedness. The putative lipids (raw $P<0.05$) in the SHFS were validated in the MDC-CC using Cox proportional hazard model, adjusting for age, sex, BMI, smoking, hypertension, diabetes, cardiovascular disease and LDL-c at baseline. Random-effects meta-analysis was performed to combine results across cohorts. In addition, we conducted repeated measurement analysis to examine longitudinal change in lipid species associated with all-cause or CVD mortality in the SHFS. Finally, we performed lipid network analysis to identify co-expressed lipid modules and modular differential connectivity associated with all-cause or CVD mortality in the SHFS and then validated in the MDC-CC.

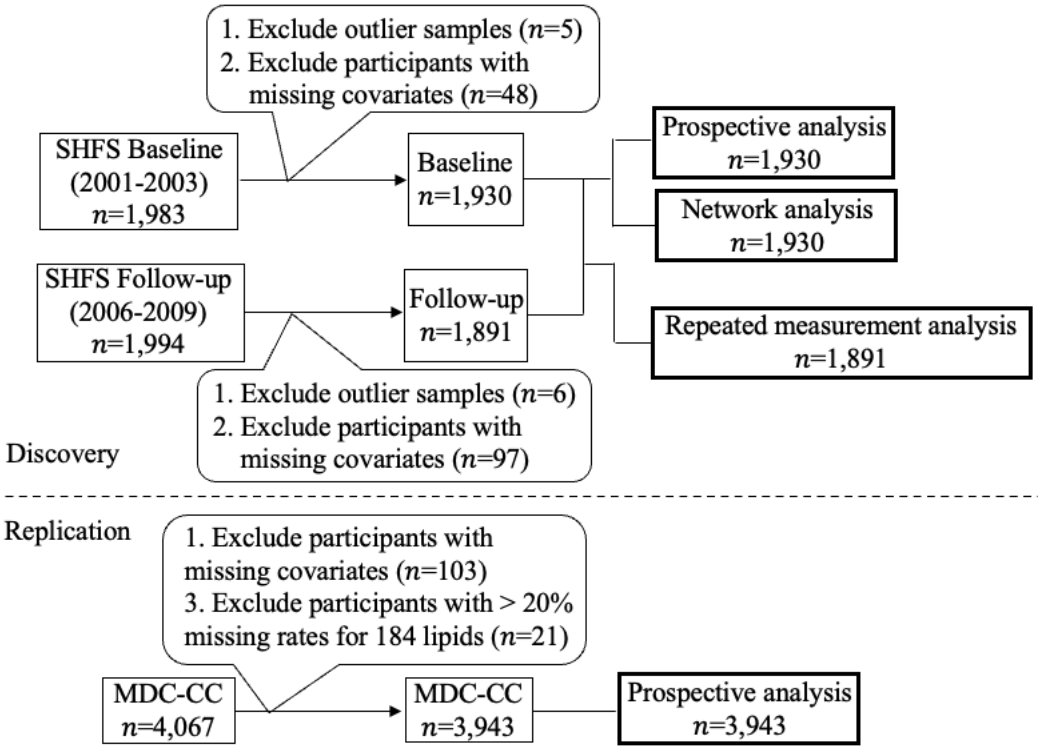


Figure S2. Baseline plasma lipid species predictive of all-cause mortality. Hazard ratios (HRs) and 95% confidence intervals (CIs) were obtained by Cox proportional hazard frailty model (SHFS) or Cox proportional hazard model (MDC-CC), adjusting for age, sex, BMI, smoking, hypertension, diabetes, cardiovascular disease and LDL-c at baseline. The analysis in the SHFS additionally adjusted for eGFR. Family relatedness in the SHFS was accounted for by including a frailty term in the model. Only known lipids with P<0.05 in the SHFS and those measured in the MDC-CC are shown. Lipids confirmed in the MDC-CC (q<0.05) are highlighted in blue. The letter A, B, or C in the name of lipids indicates isomers.

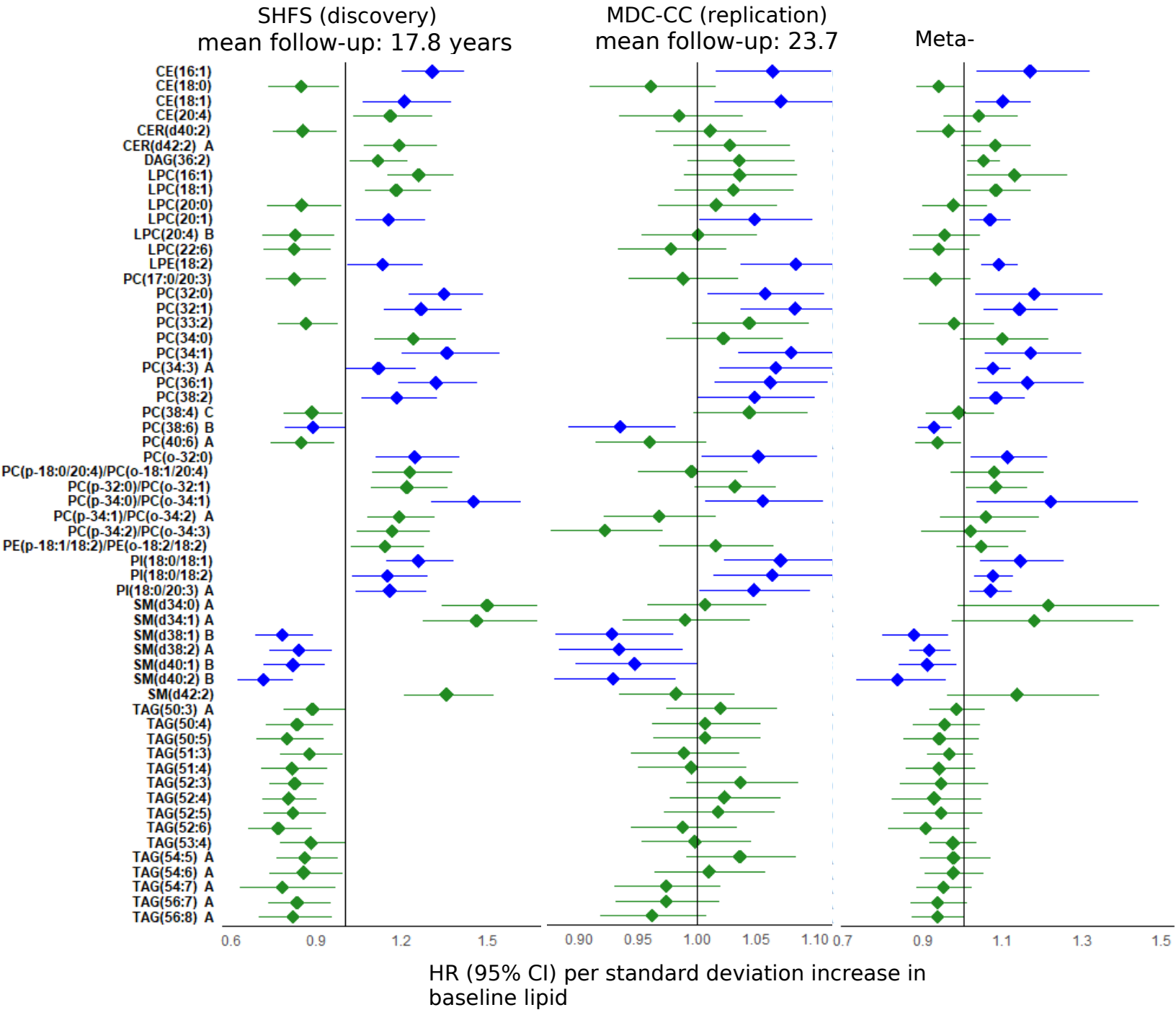


Figure S3. Baseline plasma lipid species predictive of CVD mortality. Hazard ratios (HRs) and 95% confidence intervals (CIs) were obtained by Cox proportional hazard frailty model (SHFS) or Cox proportional hazard model (MDC-CC), adjusting for age, sex, BMI, smoking, hypertension, diabetes, cardiovascular disease and LDL-c at baseline. The analysis in the SHFS additionally adjusted for eGFR. Family relatedness in the SHFS was accounted for by including a frailty term in the model. Only known lipids with P<0.05 in the SHFS and those measured in the MDC-CC are shown. Lipids confirmed in the MDC-CC (P<0.05) are highlighted in blue. The letter A or B in the name of lipids indicates isomers.

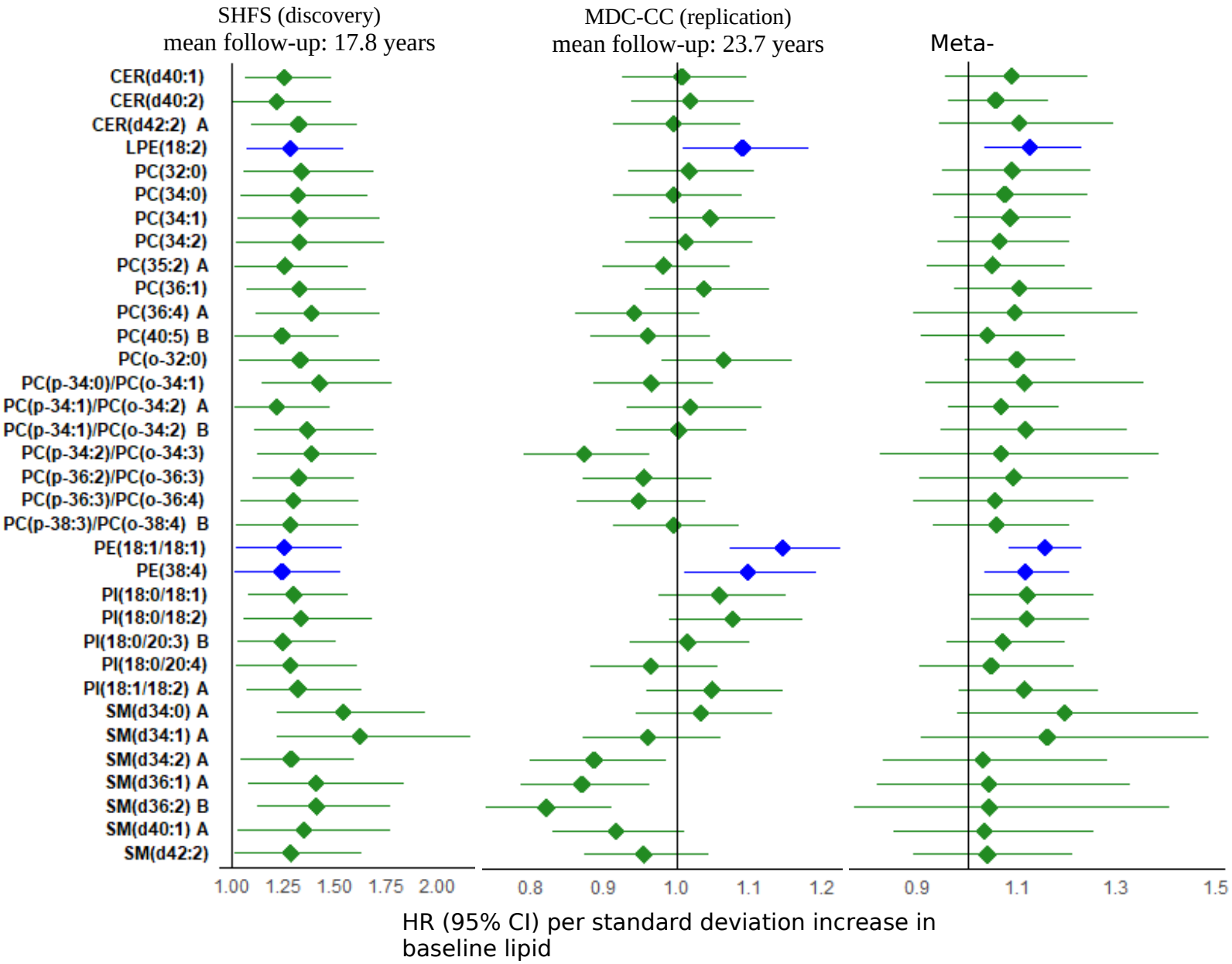


Figure S4. Baseline plasma lipid species predictive of both all-cause and CVD mortality in the SHFS (P<0.05). Hazard ratios (HRs) and 95% confidence intervals (CIs) were obtained by Cox proportional hazard frailty model, adjusting for age, sex, BMI, smoking, hypertension, diabetes, cardiovascular disease, LDL-c and eGFR at baseline. Family relatedness was accounted for by including a frailty term in the model. Only known lipids with P<0.05 are shown. The letter A or B in the name of lipids indicates isomers.

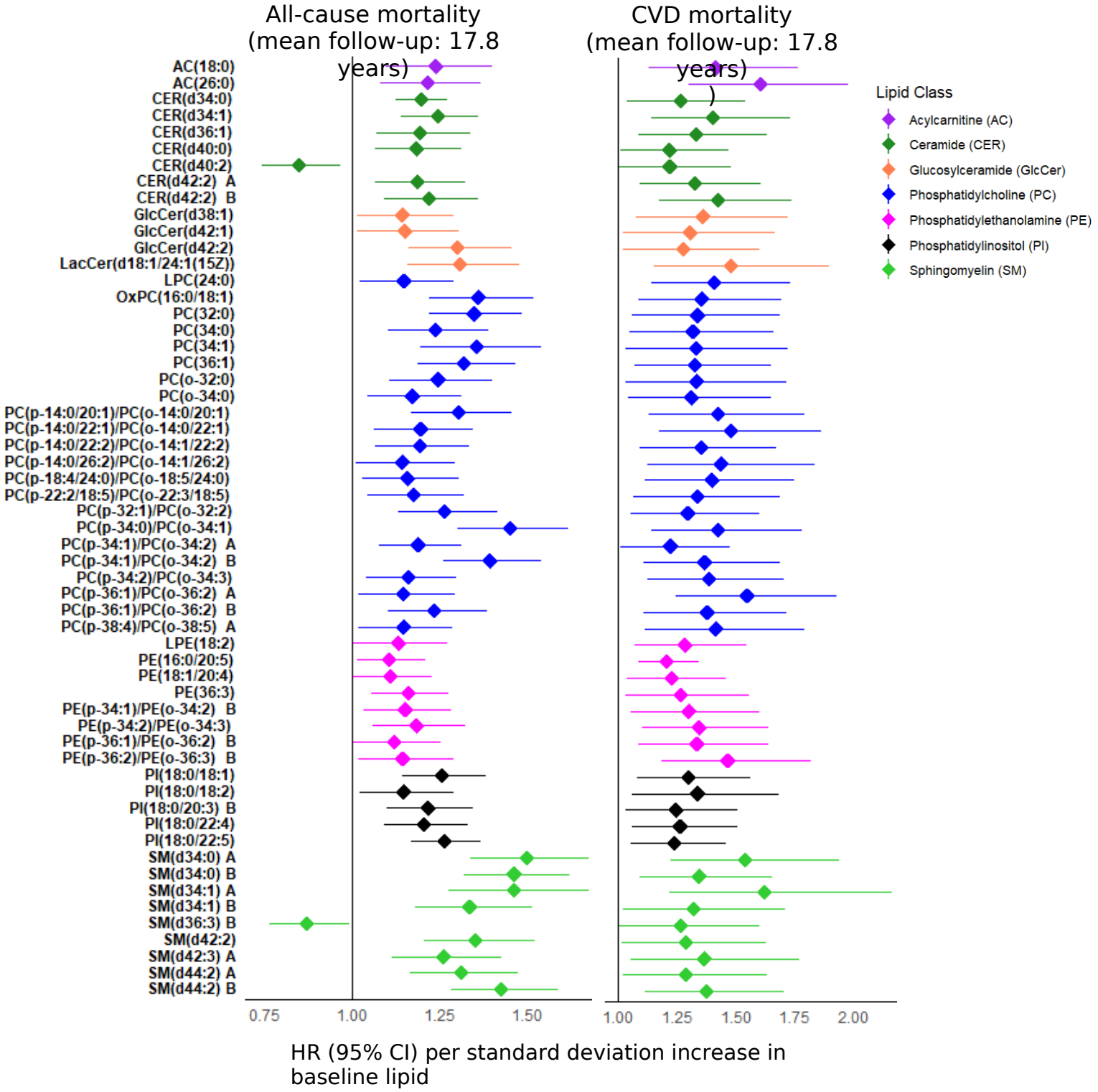


Figure S5. Longitudinal change in plasma lipid species predictive of both all-cause mortality and CVD mortality in the SHFS (P<0.05). Hazard ratios (HRs) and 95% confidence intervals (CIs) were obtained by Cox proportional hazard frailty model, adjusting for age, sex, smoking, cardiovascular disease at baseline and change in continuous traits (i.e., BMI, fasting glucose, systolic blood pressure, LDL-c, eGFR). Family relatedness was accounted for by including a frailty term in the model. Only known lipids with P<0.05 are shown. The letter A, B, or C in the name of lipids indicates isomers.

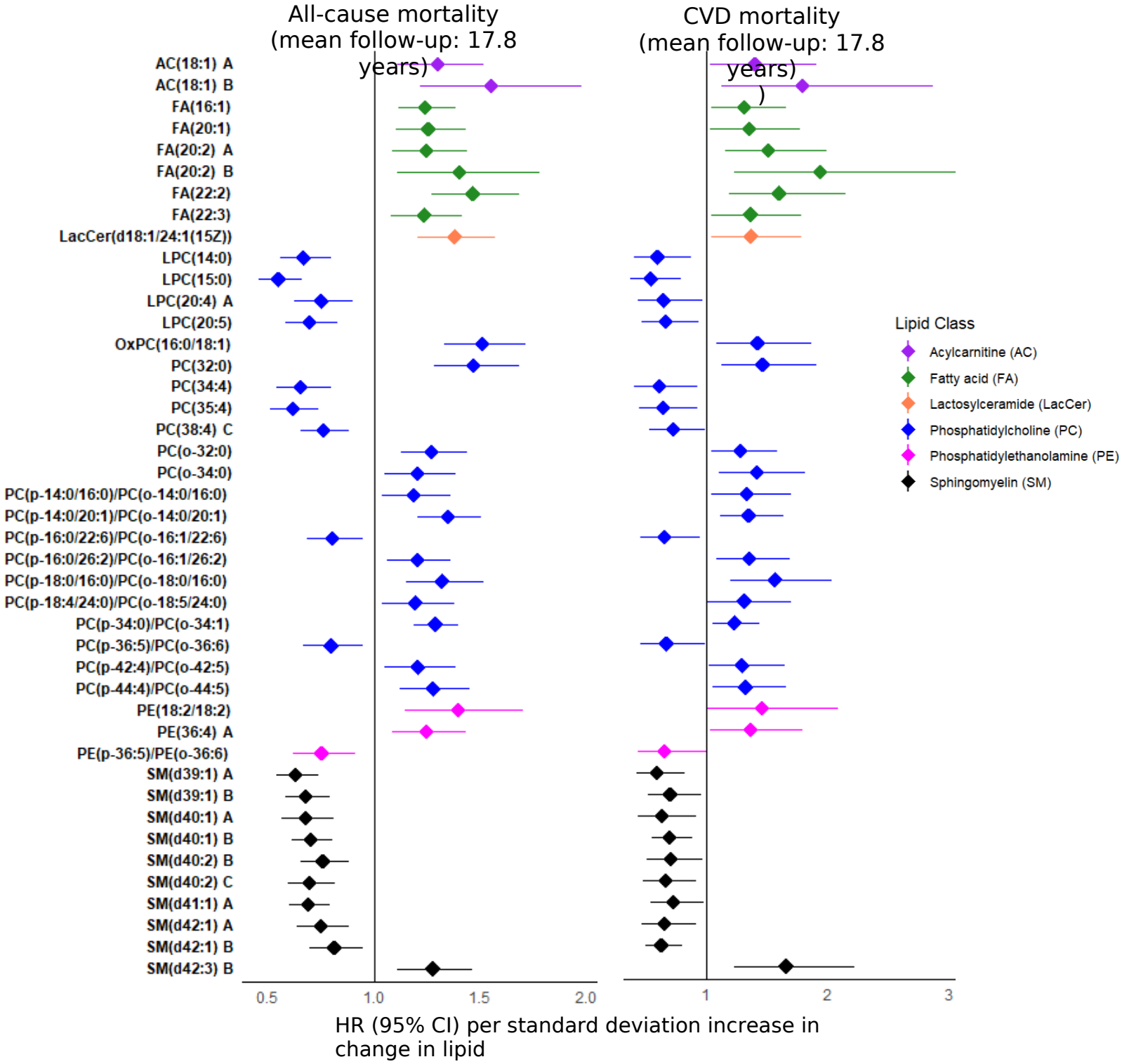
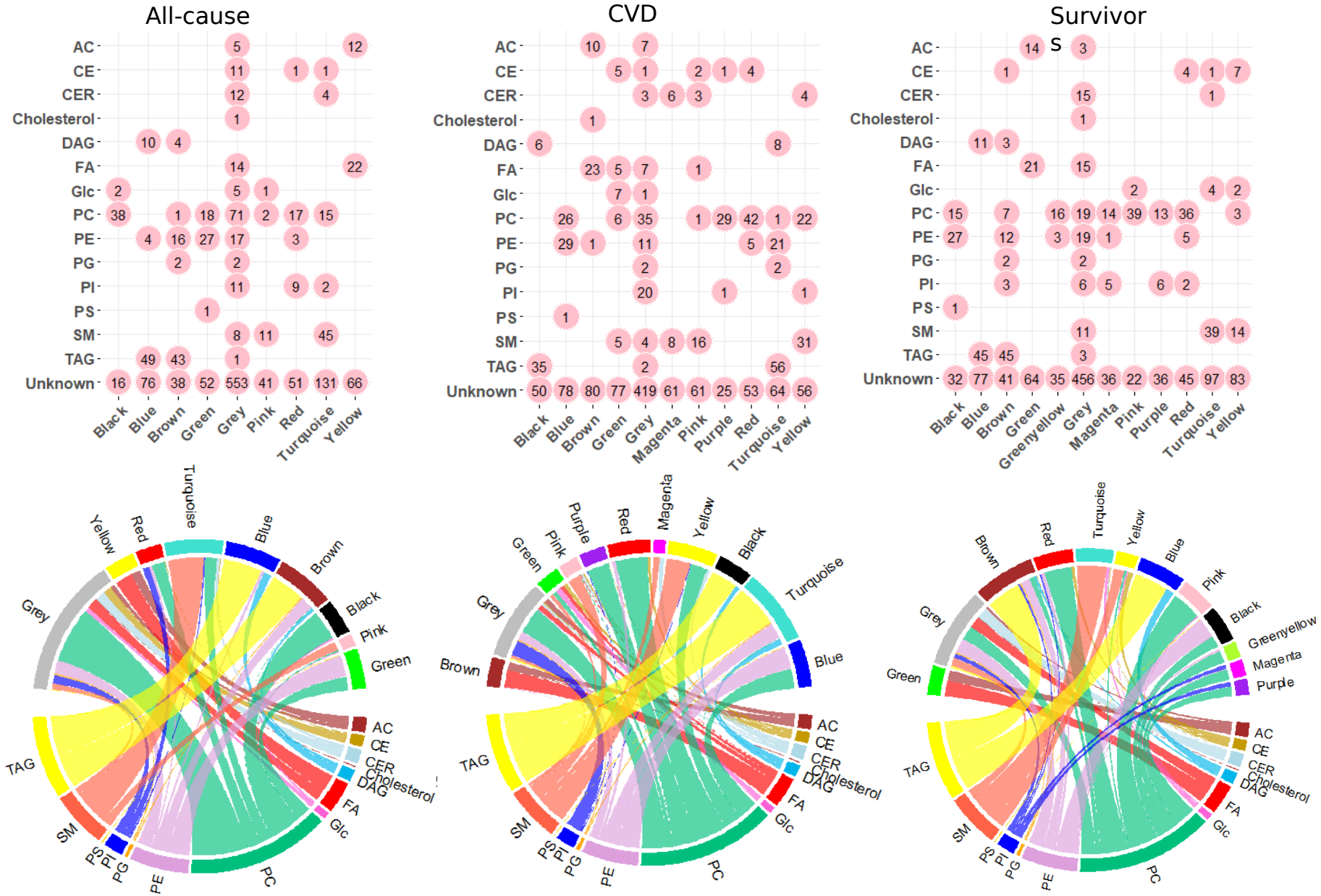


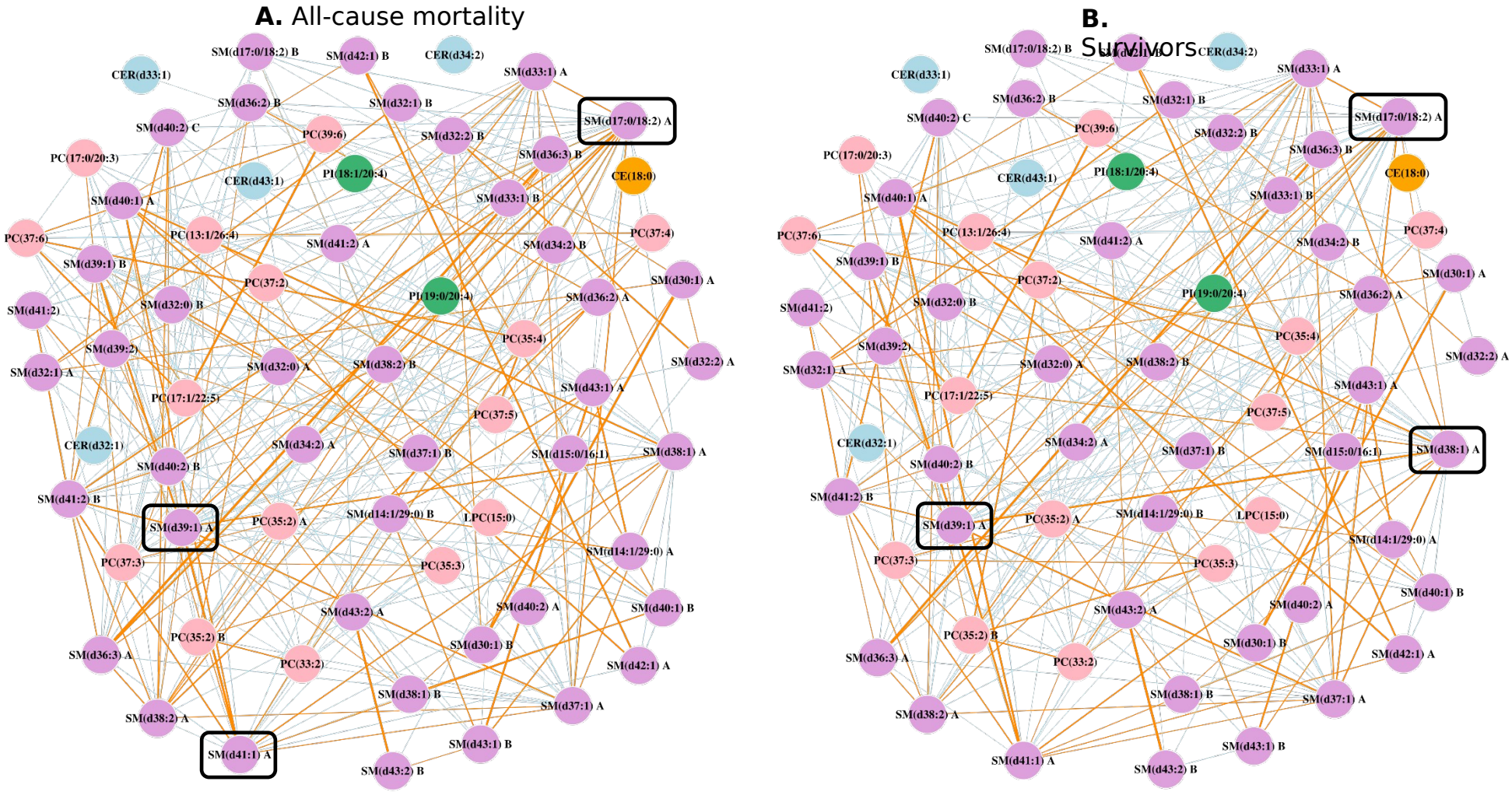
Figure S6. Co-regulated lipid modules among all-cause mortality, CVD mortality and survivors in the SHFS. Lipid network analysis identified 9 co-regulated lipid modules in all-cause mortality (n=295), 11 modules in CVD mortality (n=66) and 12 modules in survivors (n=1,635). **Upper panel**: Number of lipids (both known and unknown) included in each module; **Lower panel**: Lipid classes of 518 known lipids included in each module.



Each node represents a lipid species, and nodes in different colors represent different lipid classes. The edge colors reflect the strength of correlation between lipids (orange shows the strongest correlation followed by light blue). Hub lipids were highlighted by squares.

Turquoise module: Gain of connectivity (GOC) among participants who died (**A**) compared to survivors (**B**). Modular differential connectivity (MDC) [all-cause mortality – survivors]=202.2, P=0.046.

The letter A or B in the name of lipids indicates isomers.



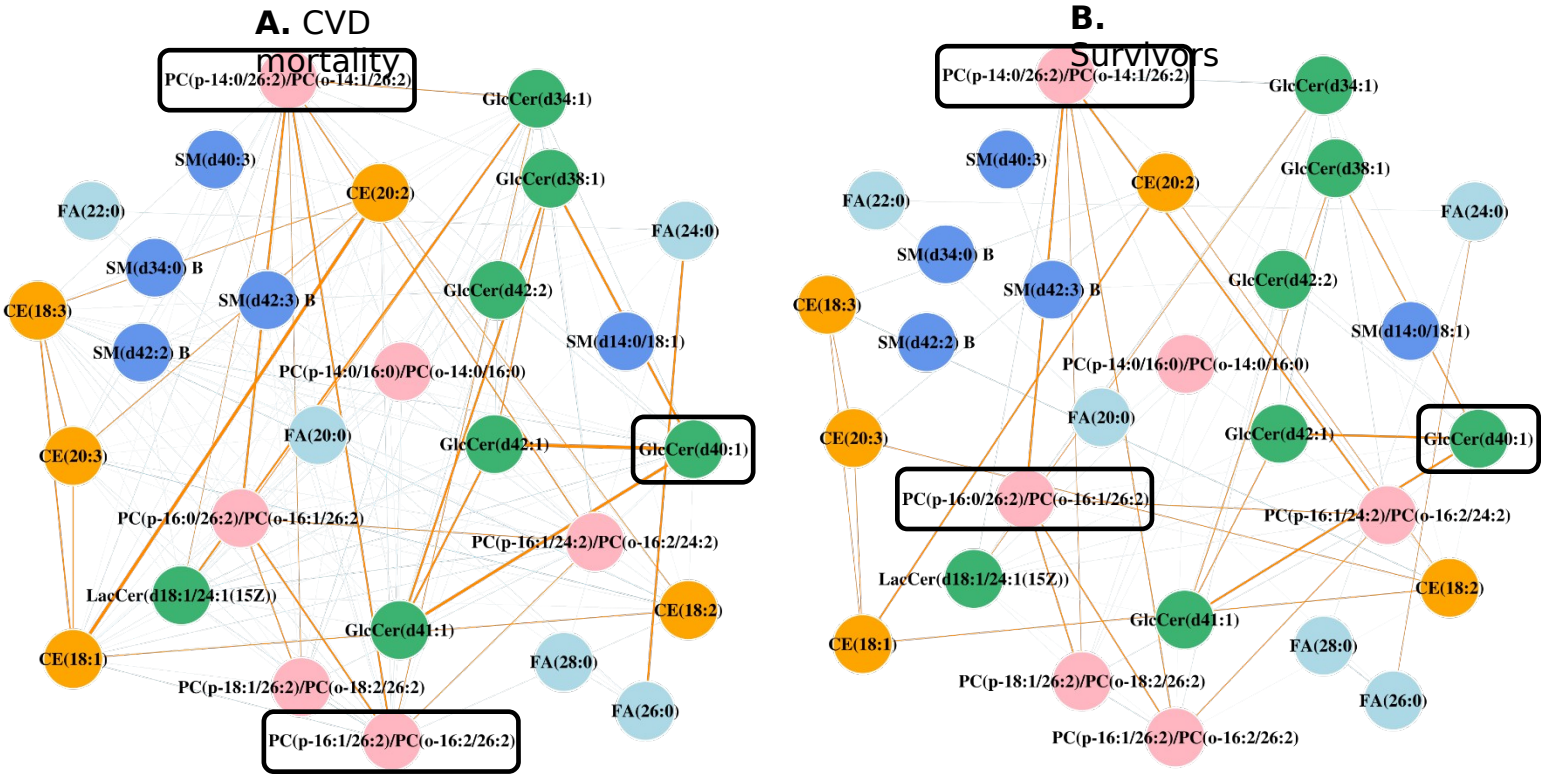
Abbreviations: CER: ceramide (blue color), CE: cholesterol ester (orange color), PC: phosphatidylcholine (pink color), PI: phosphatidylinositol (green color), SM: sphingomyelin (purple color).

Figure S8. Differential lipid networks associated with CVD mortality in the SHFS.

Each node represents a lipid species, and nodes in different colors represent different lipid classes. The edge colors reflect the strength of correlation between lipids (orange shows the strongest correlation followed by light blue). Hub lipids were highlighted by squares.

Green module: Gain of connectivity (GOC) among participants who died of CVD (**A**) compared to survivors (**B**). Modular differential connectivity (MDC) [CVD mortality – survivors]=92.9, P=0.026.

The letter A, B or C in the name of lipids indicates isomers.



Abbreviations: CE: cholesterol ester (orange color), FA: fatty acid (lightblue color), GlcCer: glucosylceramide (green color), LacCer: lactosylceramide (green color), PC: phosphatidylcholine (pink color), SM: sphingomyelin (blue color).

Figure 1. Baseline plasma lipid species predictive of all-cause or CVD mortality in the SHFS (P<0.05). Hazard ratios (HRs) and 95% confidence intervals (CIs) were obtained by Cox proportional hazard frailty model, adjusting for age, sex, BMI, smoking, hypertension, diabetes, cardiovascular disease, LDL-c and eGFR at baseline. Family relatedness was accounted for by including a frailty term in the model. Only known lipids with P<0.05 are shown. Top significant lipids with q<0.05 were labelled in black. The letter A or B in the name of lipids indicates isomers.

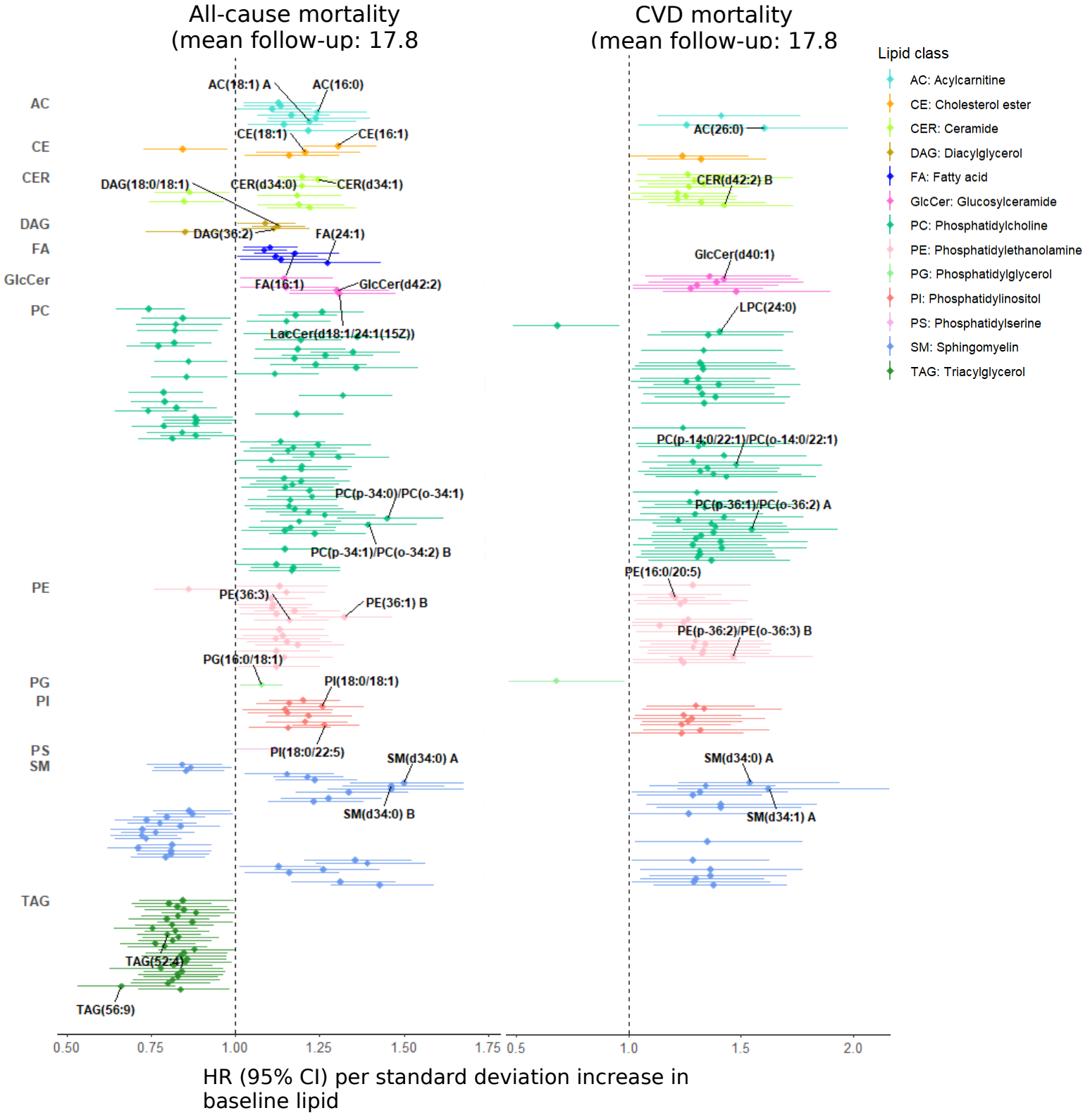
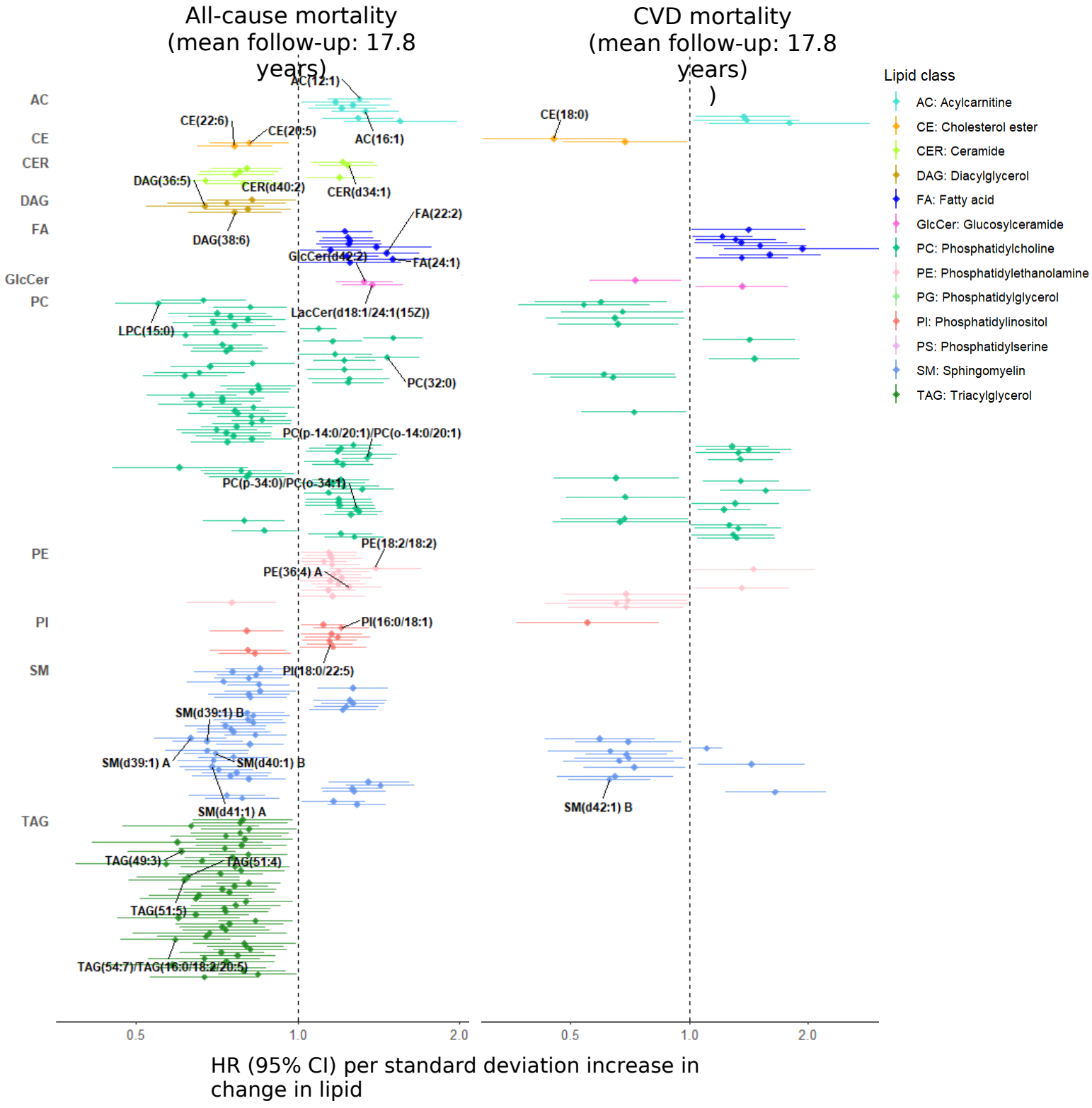


Figure 2. Longitudinal change in plasma lipid species predictive of all-cause mortality or CVD mortality in the SHFS (P<0.05). Hazard ratios (HRs) and 95% confidence intervals (CIs) were obtained by Cox proportional hazard frailty model, adjusting for age, sex, smoking, cardiovascular disease at baseline and change in continuous traits (i.e., BMI, fasting glucose, systolic blood pressure, LDL-c, eGFR). Family relatedness was accounted for by including a frailty term in the model. Only known lipids with P<0.05 are shown. Top significant lipids with q<0.05 were labelled in black. The letter A or B in the name of lipids indicates isomers.



Supplementary Tables

Table 1 - Baseline characteristics of participants in the SHFS (n=1,930) and MDC-CC (n=3,943)

Table 2 - Baseline plasma lipid species associated with all-cause mortality (P<0.05). Only known lipids in the SHFS with P<0.05 are listed below.

Table 3 - Baseline plasma lipid species associated with CVD mortality (P<0.05). Only known lipids in the SHFS with P<0.05 are listed below.

Table 4 - Longitudinal changes in fasting plasma lipid species associated with all-cause and CVD mortality in the SHFS (P<0.05).

Table 5 - Results for sensitivity analyses in the SHFS (baseline lipid species predict all-cause and CVD mortality)