Human Plasma Lipidome Is Pleiotropically Associated With Cardiovascular Risk Factors and Death

Claire Bellis, PhD; Hemant Kulkarni, MD; Manju Mamanti, MD; Jack W. Kent, Jr, PhD; Gerard Wong, PhD; Jacquelyn M. Weir, BAppSci (Chem); Christopher K. Barlow, PhD; Vincent Diego, PhD; Marcio Almeida, PhD; Thomas D. Dyer, PhD; Harald H.H. Göring, PhD; Laura Almasy, PhD; Michael C. Mahaney, PhD; Anthony G. Comuzzie, PhD; Sarah Williams-Blangero, PhD; Peter J. Meikle, PhD; John Blangero, PhD; Joanne E. Curran, PhD

Background—Cardiovascular disease (CVD) is the most common cause of death in the United States and is associated with a high economic burden. Prevention of CVD focuses on controlling or improving the lipid profile of patients at risk. The human lipidome is made up of thousands of ubiquitous lipid species. By studying biologically simple canonical lipid species, we investigated whether the lipidome is genetically redundant and whether its genetic influences can provide clinically relevant clues of CVD risk.

Methods and Results—We performed a genetic study of the human lipidome in 1212 individuals from 42 extended Mexican American families. High-throughput mass spectrometry enabled rapid capture of precise lipidomic profiles, providing 319 unique species. Using variance component–based heritability analyses and bivariate trait analyses, we detected significant genetic influences on each lipid assayed. Median heritability of the plasma lipid species was 0.37. Hierarchical clustering based on complex genetic correlation patterns identified 12 genetic clusters that characterized the plasma lipidome. These genetic clusters were differentially but consistently associated with risk factors of CVD, including central obesity, obesity, type 2 diabetes mellitus, raised serum triglycerides, and metabolic syndrome. Also, these clusters consistently predicted occurrence of cardiovascular deaths during follow-up.

Conclusions—The human plasma lipidome is heritable. Shared genetic influences reduce the dimensionality of the human lipidome into clusters that are associated with risk factors of CVD. (Circ Cardiovasc Genet. 2014;7:854-863.)

Key Words: cardiovascular diseases • genetics • lipids

Mexican Americans have a high prevalence of cardiovascular morbidity and mortality. The high risk for cardiovascular disease (CVD) in this ethnic group is partly explained by a high propensity to metabolic syndrome, which is a constellation of clinical states that also contribute to cardiovascular death. However, because the prediction models for cardiovascular deaths have limited accuracy, the search for important biomarkers of CVD continues. There is now a renewed interest in the potential contribution of lipids to CVD. Revolutionary advances in the methods for measuring the wide spectrum of lipid molecules in different tissues have helped to characterize the lipidome, the complete universe of fundamental lipid species. There are >1000 such lipid species comprising the lipidome. The emerging field of lipidomics allows the simultaneous assay of large numbers of these canonical lipids. The lipidome is a rich compilation of lipid species comprising the lipidome. By studying biologically simple canonical lipid species, we investigated whether the lipidome is genetically redundant and whether its genetic influences can provide clinically relevant clues of CVD risk. However, because research on lipidome is still in its infancy, the putative role of many lipids in health and disease is far from understood. Thus, a better appreciation of the lipidomic landscape, especially on the changing backdrop of cardiovascular health, is required.
aforementioned clinical states when examined in a multivariate context. Although such interlipid species correlations have been conceptually implied, their complete characterization is currently lacking. To understand the genetic influence on CVD, it is imperative to understand (1) whether the human plasma lipidome is itself genetically controlled and (2) whether the plasma lipidome and CVD share genetic influences. Because the SAFHS participants represent large and complex pedigrees, this study population permits assessment of the possible genetic basis of cardiovascular and other diseases with a focus on the plasma lipidome.

In this study, we set out to characterize the genetic correlations among the lipid species comprising the plasma lipidome. Our central question was whether or not these fundamental lipid components represent unique phenotypes that are closely related to polygenes which may be causally involved in cardiovascular risk. We investigated the potential genetic redundancy in the human plasma lipidome in Mexican Americans and the association of these genetically derived phenotypes with common risk factors of CVD. We also examined the association of human plasma lipidome with prospectively monitored cardiovascular deaths.

Methods

Study Participants
The SAFHS began in 1991 and has enrolled large, extended Mexican American families residing in San Antonio. The enrollment procedures, inclusion and exclusion criteria, and phenotypic assessments of the study participants have been described in detail previously.17,18 This is an ongoing longitudinal observational investigation, which has had 4 phases of data collection for a 23-year period. The data and samples used in this study were collected during the first phase of data collection that lasted from 1992 to 1996. Informed consent was obtained from all participants before collection of samples. The Institutional Review Board of the University of Texas Health Science Center at San Antonio approved the study. Plasma lipidomic data were available for 1212 participants (from 42 families). Other phenotypic data were available for 1198 participants. The biological relationships observed in the study sample were monozygotic twins (1 pair), parents−offspring (922 pairs), siblings (1111 pairs), grandparents−grandchildren (310 pairs), avuncular relatives (2064 pairs), half siblings (148 pairs), double first cousins (8 pairs), third-degree relatives (5321 pairs), fourth-degree relatives (2876 pairs), fifth-degree relatives (1204 pairs), and sixth-degree relatives (316 pairs). The clinical characteristics of the study participants are shown in Table 1.

Lipidomic Studies
We estimated the concentrations of a total of 319 lipid species (representing 23 lipid classes and subclasses shown in Table 2) in fasting plasma samples by combining high performance liquid chromatography and mass spectroscopy. These assays were conducted in the Metabolomics Laboratory, Baker IDI Heart and Diabetes Institute, Melbourne, Australia. The experimental protocols used have been described elsewhere.19

Analytic Approach

Heritability of the Plasma Lipidome
Using rich phenotypic and kinship data, we first examined the degree to which plasma levels of individual lipid species are heritable within a variance components framework. This analytic framework does not require large-scale genotyping data because kinships can inform about the extent of shared genetic information between individuals. We used polygenic regression models that predict the inverse-normalized plasma concentration of each lipid species after accounting for the kinship structure as follows:

\[ \Omega = 2\sigma_G^2 + \sigma_E^2, \]

where, \( \Omega \) is the total phenotypic covariance matrix of a trait, \( \sigma_G^2 \) is the matrix of kinship coefficients, \( i \) is the identity matrix, \( \sigma_E^2 \) is additive genetic variance, and \( \sigma_G^2 \) is residual environmental variance. All the models were adjusted for the following covariates: age, age2, sex, age × sex interaction, age2 × sex interaction, receipt of lipid-lowering and antihypertensive drugs. Heritability was estimated as \( \rho_G^2 = \sigma_G^2 / \sigma_\Omega^2 \) and represented the proportion of phenotypic variance explained by the genetic similarity (represented in the kinship matrix). Inverse-normalization of lipid species was achieved by ranking the observations, generating a cumulative density function, and then converting this probability function into a standardized deviation. Thus, all transformed lipid species concentrations had a mean of 0 and SD of unity.

Genetic Redundancy Among the Plasma Lipid Species
The variance components approach has been extended to permit bivariate trait analyses20–22 in which it is possible to further partition the phenotypic variance into genetic and environmental components. Specifically, in the context of bivariate trait analyses, the phenotypic covariance \( \rho_{ij}^2 = \sigma_{ij}^2 / \sigma_\Omega^2 \) is regarded as a function composed of the additive genetic \( \rho_{G_i}^2 \) and environmental \( \rho_{E_i}^2 \) covariances between 2 traits (denoted below as \( i \) and \( j \)).
Table 2. Summary of Lipid Classes Included in the Plasma Lipidome

<table>
<thead>
<tr>
<th>Lipid Class</th>
<th>Acronym</th>
<th>No. of Species</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dihydroceramide</td>
<td>dhCer</td>
<td>6</td>
</tr>
<tr>
<td>Ceramide</td>
<td>Cer</td>
<td>6</td>
</tr>
<tr>
<td>Monohexosylceramide</td>
<td>MHC</td>
<td>6</td>
</tr>
<tr>
<td>Dihexosylceramide</td>
<td>DHC</td>
<td>6</td>
</tr>
<tr>
<td>Trihexosylceramide</td>
<td>THC</td>
<td>6</td>
</tr>
<tr>
<td>G₂₃ ganglioside</td>
<td>GM</td>
<td>6</td>
</tr>
<tr>
<td>Sphingomyelin</td>
<td>SM</td>
<td>19</td>
</tr>
<tr>
<td>Phosphatidylcholine</td>
<td>PC</td>
<td>46</td>
</tr>
<tr>
<td>Alkenyphosphatidylcholine</td>
<td>PC(P)</td>
<td>18</td>
</tr>
<tr>
<td>Alkenyphosphatidylcholine</td>
<td>PC(O)</td>
<td>8</td>
</tr>
<tr>
<td>Lyso phosphatidylcholine</td>
<td>LPC</td>
<td>21</td>
</tr>
<tr>
<td>Lysoalkylphosphatidylcholine</td>
<td>LPC(O)</td>
<td>6</td>
</tr>
<tr>
<td>Phosphatidylethanolamine</td>
<td>PE</td>
<td>18</td>
</tr>
<tr>
<td>Alkenyphosphatidylethanolamine</td>
<td>PE(P)</td>
<td>12</td>
</tr>
<tr>
<td>Alkenyphosphatidylethanolamine</td>
<td>PE(O)</td>
<td>9</td>
</tr>
<tr>
<td>Lyso phosphatidylethanolamine</td>
<td>LPE</td>
<td>6</td>
</tr>
<tr>
<td>Phosphatidylinositol</td>
<td>PI</td>
<td>17</td>
</tr>
<tr>
<td>Phosphatidylserine</td>
<td>PS</td>
<td>7</td>
</tr>
<tr>
<td>Phosphatidylglycerol</td>
<td>PG</td>
<td>4</td>
</tr>
<tr>
<td>Cholesteryl ester</td>
<td>CE</td>
<td>26</td>
</tr>
<tr>
<td>Cholesterol</td>
<td>COH</td>
<td>1</td>
</tr>
<tr>
<td>Diacylglycerol</td>
<td>DG</td>
<td>22</td>
</tr>
<tr>
<td>Triacylglycerol</td>
<td>TG</td>
<td>43</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>319</td>
</tr>
</tbody>
</table>

Table 2. Summary of Lipid Classes Included in the Plasma Lipidome

These parameters are estimated using an estimation-maximization algorithm by jointly using all available pedigree information with a significant genetic correlation between 2 phenotypes indicates shared polygenic effect of causal genes. We used this approach to estimate the genetic correlation between each pair of lipid species. We generated a genetic correlation matrix of all the 319 lipid species with each other. We then used hierarchical clustering methods to reduce the dimensionality of this correlation matrix into meaningful clusters that are hereinafter referred to as genetic clusters of lipid species.

Clinical Associations of Genetic Clusters

We investigated the usefulness of the genetic clusters through their association with 7 prevalent clinical states, and observed cardiovascular deaths in the study participants using polygenic regression models. The clinical states that we studied were as follows: central obesity (waist circumference ≥102 cm in men and ≥88 cm in women), obesity (body mass index ≥30 kg/m²), raised triglycerides (serum triglycerides ≥150 mg/dL [1.7 mmol/L] or receipt of lipid-lowering drugs), low high-density lipoprotein cholesterol (HDL-C; serum HDL-C <40 mg/dL [1.03 mmol/L] in men, <50 mg/dL [1.29 mmol/L] in women or receipt of lipid-lowering drugs), hypertension (systolic blood pressure ≥140 or diastolic blood pressure ≥90 mm Hg or receipt of antihypertensive drugs), type 2 diabetes mellitus (American Diabetes Association criteria), and metabolic syndrome defined as follows: presence of central obesity combined with any of the following: raised triglycerides, low HDL-C, high blood pressure and raised fasting plasma glucose (≥5.6 mmol/L), previously diagnosed type 2 diabetes mellitus, or receiving antidiabetic medication (International Diabetes Federation definition). In addition, we had information on the deaths of study participants, which was derived from death certificates provided by the San Antonio Metro Health Department, and the causes of death in deceased SAFHS participants. Using the International Classification of Diseases coding scheme (ICD-10), we found that 73 deaths were reported to have CVDs (ICD code category I) as primary or contributory cause of death in participants followed up till October 31, 2009. Deaths with following primary or contributory causes (ICD-10 codes) were defined as cardiovascular deaths: Rheumatic fever with heart involvement (I01), essential hypertension (I10), rheumatic mitral valve disease (I05.9), hypertensive heart disease (I11.9), hypertensive chronic kidney disease (I12.9), ST-segment–elevation myocardial infarction (I21.3 and I21.9), atherosclerotic heart disease of native coronary artery (I25.0 and I25.1), ischemic cardiomyopathy (I25.5), chronic ischemic heart disease (I25.9), diseases of pericardium (I31.9), hypertrophic cardiomyopathy (I42.2), cardiomyopathy (I42.9), cardiac arrest (I46.9), paroxysmal tachycardia (I47.2), cardiac arrhythmia (I49.9), heart failure (I50.0), heart failure (I50.9), nontraumatic intracerebral hemorrhage (I61.9 and I62.0), sequelae of nontraumatic intracranial hemorrhage (I69.2 and I69.4), atheroembolic disease (I70.9), peripheral vascular disease (I73.9), other disorders of arteries and arterioles (I77.6), and hypertension (I85.9). In the subsample on which lipidomic studies were done, there were 52 cardiovascular deaths during follow-up.

Statistical Analysis

To compare distribution of continuous variables across 2 groups, we used the nonparametric Mann–Whitney U test. To test the assumptions of a normal distribution of a variable, we used the skewness/kurtosis test of D’Agostino et al. To test the variability of estimated parameters across subgroups, we used Cochrane’s Q statistic. We used the R packages hclust (for hierarchical clustering) and corrplot (for depicting the large genetic correlation matrix). Enrichment of lipid classes within genetic clusters was assessed using Fisher exact test and applying Benferroni correction for multiple comparisons.

Association analyses were conducted using the SOLAR software package. These analyses also used the variance components approach in a polygenic regression model as follows.

\[ CT = m + \sum a_i \alpha_i + \varepsilon \]

where, \( CT \) is the liability of a clinical trait; \( m \) is the mean; \( a \) is the covariate vector of dimension \( k \), with \( b \) as the vector of corresponding regression coefficients; \( g \) is the polygenic effect; and \( e \) is the residual error for an individual indexed by \( i \). Because all the 7 clinical traits were discrete in nature, we used the liability threshold approach to model the likelihood of these traits. We modeled the term \( g \) as a random variable on the basis of the coefficients of relationship in the kinship matrix. All models included adjustments for age, age2, sex, age × sex interaction, and age2 × sex interaction, and use of lipid-lowering and antihypertensive drugs and 12 cluster scores as covariates. We generated a cluster score by calculating the average of the inverse-normalized plasma concentrations of all lipid species belonging to that genetic cluster. Statistical significance of the association was tested by constraining the regression coefficient to 0 and comparing the log-likelihoods of the constrained and unconstrained regression models in a likelihood ratio test by \( \chi^2 \) test. Statistical significance was tested at a global type I error rate (\( \alpha \)) of 0.05; however, to correct for multiple tests, we used the Benjamini–Hochberg method of controlling false-discovery rates. We used Stata 12.0 (Stata Corp; College Station, TX) package for the Mann–Whitney U test and multiple test correction.

Results

The SAFHS participants were middle-aged with a majority of women (Table 1). The prevalence of obesity, central obesity,
type 2 diabetes mellitus, and metabolic syndrome was high in this sample. Approximately 10% of the study participants were already receiving antihypertensive treatment and another 54 subjects had clinically detectable hypertension at the time of enrollment. Only a small proportion of the study participants (<2%) were already receiving lipid-lowering medications. Follow-up of 9314 person-years revealed that there were 52 cardiovascular deaths in this sample with an estimated cardiovascular mortality incidence of 5.58 deaths per 1000 participants per year. The class membership of each lipid species and its relative plasma concentration are given in Table I in the Data Supplement. The observed relative concentrations indicate a substantial variability of the plasma lipidome across lipid species, classes, subclasses, and individuals.

**Human Plasma Lipidome Is Heritable**

We estimated the narrow-sense heritability of each lipid species in the study participants. We found that each of the 319 lipid species had a statistically significant heritability even after correction for multiple testing (Table I in the Data Supplement). The heritability estimates ranged from a minimum of 0.09 ($P=0.0226$ after multiple test correction) for dihydroceramide 16:0 to a maximum of 0.60 ($P=4.2\times10^{-34}$ after multiple test correction) for dihexosylceramide 24:1. The histogram of the estimated heritabilities (Figure 1A) indicated a potential asymmetry around the central value. When we tested the assumption for normal distribution of the heritability estimates using the skewness/kurtosis test, we found that the skewness significantly deviated from normality ($P=0.011$) but the kurtosis was normal ($P=0.816$). We therefore generated a box plot of this distribution (Figure 1B), which showed that the median heritability of the plasma lipidome was 0.3705 with an interquartile range of 0.1255.

We then explored whether or not the estimated heritabilities were similar across the lipid classes and subclasses. A box plot (Figure 1C) showed that there was considerable variation in the estimated heritability across lipid classes and subclasses, with the alkylphosphatidylethanolamines (PE[O]) showing the least median heritability (0.2220) and the monohexosylceramides (MHC) showing the highest median heritability (0.5002). There was a statistically significant heterogeneity in heritability across lipid classes and subclasses ($Q=69.64$; degree of freedom=22; $P=7.53\times10^{-7}$). In general, phospholipids had lower heritabilities than the sphingolipid or glycerolipid classes.

**Genetic Correlations Among Lipid Species**

We next conducted a series of analyses to characterize the complex genetic correlations among the plasma lipid species. First, we estimated pairwise genetic correlation coefficients for all pairs of the plasma lipidome (total 50721 pairs) using bivariate trait analyses. Because presence of obesity, type 2 diabetes mellitus, metabolic syndrome, and hypertension can influence the plasma levels of lipid species, we adjusted the genetic correlations between species by including these clinical states as covariates. We also adjusted for the receipt of lipid-lowering and antihypertensive drugs. The resulting genetic correlation matrix is provided fully in Table II in the Data Supplement and shown pictorially in Figure 2. We observed that a total of 23477 (46.3% of lipid species pairs) genetic correlations were statistically significant at a nominal $P$ value of 0.05, and 3492 (6.9%) genetic correlations were significant at a false-discovery rate–corrected $P$ value of 0.05. Figure 2 shows that there was generally a high positive genetic correlation between lipid species of the same class (concentration of blue squares along the diagonals).

Second, we used unsupervised hierarchical clustering method to reduce the complex genetic correlation matrix into genetically meaningful clusters. Using 8 criteria for clustering validation, we chose a solution that yielded 12
species can provide clinically meaningful information. We assessed whether or not the pleiotropically related lipid classes and subclasses completely corresponded to any single genetic cluster. Thus, we reasoned that the genetic clusters might contain information that is different from that contained in lipid species pairs that belonged to the same cluster with those that belonged to different clusters. We observed (Figure 3B) that the median absolute genetic correlation between lipid species of the same cluster was twice that of the lipid species belonging to different clusters (0.52 versus 0.26), a difference that was highly significant. The average genetic correlation within each specific cluster was also high (Figure 3C).

A cross-tabulation based on lipid classes and genetic clusters (Figure 3D) indicated that none of the 23 lipid classes and subclasses completely corresponded to any single genetic cluster. Thus, we reasoned that the genetic clusters might contain information that is different from that contained in lipid classes and subclasses. Statistical test for enrichment of lipid species pairs that belonged to the same cluster with those that belonged to different clusters was associated with a significantly increased likelihood of most of the clinical states. Fourth, cluster 5 scores were significantly associated with an increased likelihood of low HDL-C. Sixth, cluster 1 scores (composed primarily of lipids containing polyunsaturated fatty acids) demonstrated remarkable specificity of association—significantly increased likelihood of obesity but a significantly reduced likelihood of type 2 diabetes mellitus. It is noteworthy that certain dihydroceramides, ceramides, and cholesteryl esters were the most common lipid classes that defined cluster 7. Furthermore, clusters 6 and 7 were consistently associated with increased likelihood of most of the clinical states. Fourth, cluster 5 scores were significantly associated with an increased likelihood of central obesity, obesity, raised serum triglycerides, and type 2 diabetes mellitus. This cluster, as can be seen from Figure 3D, is mainly composed of lysophosphatidylcholine species. Third, cluster 7 scores were associated with an increased likelihood of central obesity, raised serum triglycerides, and type 2 diabetes mellitus.

There were strong associations found with raised serum triglycerides. Second, cluster 10 scores were consistently associated with a significantly reduced likelihood of central obesity, obesity, raised triglycerides, low HDL-C, and metabolic syndrome. This cluster, as can be seen from Figure 3D, is mainly composed of lysophosphatidylcholine species. Third, cluster 7 scores were associated with an increased likelihood of central obesity, raised serum triglycerides, and type 2 diabetes mellitus.

found interesting patterns of association (Figure 4). First, the strongest associations were found with raised serum triglycerides. Second, cluster 10 scores were consistently associated with a significantly reduced likelihood of central obesity, obesity, raise triglycerides, low HDL-C, and metabolic syndrome. This cluster, as can be seen from Figure 3D, is mainly composed of lysophosphatidylcholine species. Third, cluster 7 scores were associated with an increased likelihood of central obesity, raised serum triglycerides, and type 2 diabetes mellitus.

Finally, when we examined the association between genetic cluster scores and risk of cardiovascular deaths during follow-up, we found that cluster 1 scores were associated with a significantly reduced risk, whereas cluster 6 scores were associated with a significantly increased risk of cardiovascular deaths. This cluster is dominated by sphingomyelins and both the ether-linked and plasmalogen derivatives.
of phosphatidylethanolamines. This pattern of association was similar to that with type 2 diabetes mellitus. Of interest, type 2 diabetes mellitus (polygenic regression coefficient =0.56; P=0.0002) was a highly significant predictor of cardiovascular deaths in the SAFHS participants (data not shown).

Discussion

We report 3 novel and important findings. First, all the plasma lipid species were significantly heritable—a finding that strongly places the plasma lipidome in a genetic context. To our knowledge, such a finding has not been reported previously and implies that a significant proportion of the interindividual variability in plasma levels of bioactive lipids may be determined by polygenes. Considering that CVD and its risk factors are themselves genetically determined,34 our findings raise the possibility that there may be a genetic concordance between the plasma lipid species and clinical states predisposing to CVD in Mexican Americans.

Second, our results of hierarchical clustering suggest that there may be a substantial promiscuity in the genetic control of plasma lipid species. Moreover, the fact that the genetically derived clusters were conceptually different from the chemically defined lipid classes indicates a significant pleiotropy in the genetic control of the plasma lipidome that may not be fully explained by the chemical structure of the lipid class or subclass. For example, chemically similar phosphatidylcholine species (PC 38:6a and PC 38:6b) had distinct cluster memberships (to clusters 1 and 11, respectively; Table III in the Data Supplement). Also, the fatty acid composition of these 2 species is different (primary constituent of PC 38:6a is PC 16:0/22:6, whereas that for PC 38:6b is PC 18:2/20:4). Interestingly, several other members of cluster 1 contain the omega-3 fatty acid, docosahexaenoic acid (DHA, C22:6), although some also contain the omega-6 fatty acid, arachidonic acid (AA, C20:4). This is further supported by
the distribution of diacylglycerol and triacylglycerol species among clusters 11 and 12; the species in cluster 12 contain primarily saturated and monounsaturated fatty acids, whereas cluster 11 contains a high proportion of polyunsaturated species particularly reflecting the omega-6 fatty acid, linoleic acid (C18:2) in the triacylglycerols (Table V in the Data Supplement). These results raise the possibility that some of the genetic control of the plasma lipids may reside at the level of fatty acid metabolism. Future studies are needed to evaluate this possibility.

Figure 4. Association of genetic clusters with clinical conditions. For each indicated trait, the results are from a single polygenic regression model that was adjusted for age, age², sex, age × sex interaction, age² × sex interaction, and receipt of antihypertensive and lipid-lowering drugs. Average inverse-normalized scores for lipid species belonging to each genetic cluster were then included as covariates. The plots report the polygenic regression coefficients. The squares and error bars represent the point and 95% confidence intervals. The squares are color-coded as follows: red indicates significantly increased liability of the trait, gray, statistically not significant, and blue, significantly decreased liability of the trait. The significance values are indicated at the top of each plot.
Third, the genetically derived clusters of plasma lipid species were not only associated with the risk factors of CVD but also with the prospectively measured, hard outcome of cardiovascular deaths confirmed from death certificates. With this regard, it is conceivable that cluster 1 may reflect an association with DHA metabolism which has been implicated in CVD pathogenesis, whereas cluster 6 may relate to the known association of sphingomyelins with CVD. It is noteworthy that recent studies have demonstrated strong associations between the plasma lipidome and cardiovascular mortality, but our findings show the importance of genetics in characterization of the plasma lipidome and the relevance of this genetic information in potential risk-stratification of CVD.

When interpreting the association results, it should be remembered that the clusters were derived on the basis of shared genetic influences and therefore are inherently designed to detect associations with clinical states that have a high likelihood of a strong genetic basis. For example, in a large study from 2 cohorts, we recently demonstrated that type 2 diabetes mellitus is significantly associated with dihydriceramides, ceramides, and cholesterol ester species. The fact that, in this study, cluster 7 was significantly associated with type 2 diabetes mellitus points to a possible genetic basis for these associations. Concordantly, we had observed that 4 dihydroceramide species (18:0, 20:0, 22:0, and 24:1) were significantly associated with waist circumference in Mexican Americans, and here we find that clusters 5 and 7 (which contain these 4 lipid species) showed significant association with central obesity again, indicating a possible genetic basis to the previously reported associations. Finally, our observation about the consistent negative association of cluster 10 with several risk factors of CVD indicates a likely genetic explanation for the reported association of reduced levels of plasma lysophosphatidylcholine with obesity and type 2 diabetes mellitus.

Some limitations of the present study need to be considered. First, our findings should be seen as indicative and need confirmation by replications across cohorts. Demonstration of heritability and genetic correlation is a useful initial step in the quest to uncover genetic underpinnings of CVD. Deciphering the inheritance patterns and underlying molecular composition of risk traits influencing common, yet genetically complex, diseases remains largely unachieved. A logical future research direction is to identify specific sequence variants and other epigenetic control mechanisms regulating the plasma lipidome. Second, with the exception of the prospective component of cardiovascular deaths, all other inferences in this study are based on cross-sectional data. The associations therefore do not automatically imply a causal role of plasma lipid species in the pathogenesis of CVD. Rather, the main goal of the study was to query the existence of potential correlations in the, as yet latent, genetic regulators of CVD. That we obtained a consistent pattern of associations with the prospectively monitored cardiovascular deaths further the likelihood that the genetic clusters derived in the study are clinically meaningful. Third, because the study participants are all Mexican Americans, it is not possible to generalize these results to other ethnic groups. Future studies need to investigate the similarities and differences of the genetic clusters on the background of differing ethnicity. Fourth, lipid concentrations were estimated in stored samples and, in theory, storage can influence the estimated concentrations. However, it has been demonstrated that quantitative lipidomic techniques such as the ones used in this study are unlikely to be affected by storage even if the samples were exposed to multiple freeze/thaw cycles. Of note, the samples used in this study did not undergo any freeze/thaw cycles before analysis. Thus, the results presented here are unlikely to have been influenced by storage of samples. Fifth, it is conceivable that the variability of the plasma lipid species across individuals is only partly explained by genetics. Environmental factors such as shared households, dietary profiles, and lifestyle factors can all contribute to both lipidomic variability and CVD risk. Our study did not evaluate these aspects but future studies need to dissect out these additional contributors and likely confounders of CVD. Finally, the choice of number of clusters can influence the strength of association of the clusters with clinical states. Therefore, the results shown here should be considered indicative of patterns rather be considered conclusive about a genetic structure. Alternative definitions and strategies for cluster identification can be envisioned and need to be evaluated.

CVD management currently accounts for 17% of the national health expenditure in the United States, and the costs associated with CVD have been projected to triple by 2030. Therefore, prevention of CVD is a feasible and economically viable alternative to treatment programs. To that end, novel insights into biological mechanisms that predispose individuals to CVD hold the promise of potential new therapies and significant reduction of this considerable economic burden. Modern genomic technologies can be exploited to rapidly identify genes involved in disease susceptibility. However, the cost-effectiveness of such exploratory endeavors can be greatly augmented if genetic basis to a phenotype is strongly suspected. That was the motivation for the present study. Identifying novel lipid-related endophenotypes that are genetically correlated with CVD offers the potential to discover biomarkers which will quickly lead us to causal genes. Because the axis of management is now heavily tilted in favor of personalized medicine, identification of genetic predilection to CVD is important. Our study represents a preliminary step in that direction. Specifically, our results urge that the plasma lipidome needs to be carefully examined in future studies as a potential harbinger of CVD risk because of shared genetic influence and concurrence with the risk factors of CVD.

Acknowledgments
We would like to acknowledge the late Dr Jeremy Jowett. Dr Jowett was a very close collaborator and colleague for many years and was a key member of the research team at Baker IDI in Melbourne, Australia. He helped to foster the initial lipidomic collaboration between the Metabolomics group at Baker IDI and the Texas Biomed, and worked closely as part of this collaboration until his death in September of 2012. We are extremely thankful to Dr Jowett for his valuable scientific contributions and leadership over the years. We are also very grateful to the participants of the San Antonio...
Family Heart Study for their continued involvement in our research programs.

Sources of Funding
This work was supported in part by National Institutes of Health (NIH) grants R01 HL113323, R01 DK079169, and R01 DK088972; by National Health and Medical Research Council of Australia Grants 1029754 and 1042095; and by the OIS Program of the Victorian Government, Australia. Data collection for the San Antonio Family Heart Study was supported by NIH grant R01 HL045522. The development of the analytical methods and software used in this study was supported by NIH grant R37 MH059490. The AT&T Genomics Computing Center supercomputing facilities used for this work were supported in part by a gift from the AT&T Foundation and with support from the National Center for Research Resources Grant Number S10 RR029392. This investigation was conducted in facilities constructed with support from Research Facilities Improvement Program grants C06 RR13556 and C06 RR10751 from the National Center for Research Resources of NIH.

Disclosures
None.

References
Circulating lipid levels are generally considered as an important indicator of cardiovascular disease risk. However, the relative contribution of genetic and environmental influences on these lipid levels is unknown. It is possible to discern these relative contributions through the use of family studies, a powerful genetic epidemiological tool. We used data from the ongoing San Antonio Family Heart Study of high-risk Mexican American families and surveyed their plasma lipidome for a total of 319 lipid species. We observed that median heritability of the plasma lipidome was 0.37, indicating that ≈37% of the variation in plasma lipid species can be attributed to genetic similarity among the study participants. Furthermore, we found that on the basis of genetic correlations between pairs of lipid species, the entire lipidome could be faithfully represented by 11 hierarchical clusters. These clusters were conceptually distinct from the generally used classification of lipid species on the basis of their chemical structure. The clusters were differentially but consistently associated with several cardiovascular disease risk factors like central obesity, obesity, type 2 diabetes mellitus, raised serum triglycerides, metabolic syndrome, and cardiovascular death. In the continued quest for personalized solutions to disease diagnosis and treatment, our results place plasma lipid species in a genetic context. Because the clusters were based on genetic correlations, our results highlight the pleiotropic nature of the association of plasma lipid species with indicators of cardiovascular disease and death.