Untargeted Lipidomic Profiling of Human Plasma Reveals Differences due to Race, Gender and Smoking Status

Cai X1, Perttula K1, Pajouh SK2, Hubbard A2, Nomura DK3 and Rappaport SM4*

1University of California, Berkeley; School of Public Health; Program in Environmental Health Sciences, USA
2University of California, Berkeley; School of Public Health; Program in Biostatistics, USA
3University of California, Berkeley; College of Natural Resources; Department of Nutritional Sciences and Toxicology, USA

Abstract

Lipidomic profiling can link genetic factors and exposures to risks of chronic diseases. Using untargeted liquid chromatography-Fourier Transform mass spectrometry (LC-FTMS), we explored differences in 3,579 lipidomic features in human plasma from 158 non-fasting subjects, pooled by race, gender and smoking status. Significant associations with race (23 features), smoking status (9 features) and gender (2 features) were detected with analysis of variance (ANOVA)-based permutation tests. Identities of several features were confirmed as plasmalogens (vinyl-ether phospholipids) that were present at 2-fold greater concentrations in black subjects. Other putative features, based on accurate masses, were more abundant in white subjects, namely, dihomo-γ-linolenoyl ethanolamide (DGLEA), an endogenous endocannabinoid receptor agonist and a glycero-phosphocholine [PC(16:0/18:1)]. After adjustment for race, multivariable linear regression models showed that gender was significantly associated with levels of plasmalogens and DGLEA and that consumption of animal fat was marginally associated with concentrations of plasmalogens. Interestingly, BMI did not explain additional variability in any race-adjusted model. Since plasmalogens are antioxidants that are generally regarded as health-promoting and DGLEA is an agonist of the cannabinoid receptor, our findings that these molecules differ substantially between black and white Americans and between men and women, could have health implications. The concentration of cotinine was greatly elevated in smoking subjects and 6 features with m/z values suggestive of phospholipids or sphingomyelins were present at significantly lower concentrations in smokers.

Keywords: Untargeted lipidomics; Race; Gender; Smoking status; Plasmalogens; Endocannabinoid receptor agonist

Abbreviations: BH: Benjamini-Hochberg; CE: Cholesteryl Esters; DAG: Diacylglycerol; DGLEA: Dihomo-γ-Linolenoyl Ethanolamide; FA: Fatty Acid; FTMS: Fourier Transform Mass Spectrometry; LPC: Monoglycerophosphocholine; LPE: Monoglycerophosphoethanolamine; PA: Glycerophosphate; MAG: Monoacylglycerol; PC: Glycerophosphocholine; PE: Glycerophosphoethanolamine; PG: Glycerophosphoglycerol; PI: Phosphoinositide; SM: Sphingomyelin; TAG: Triacylglycerol

Introduction

Since lipids are essential to functional membranes, energy storage and signaling [1,2], lipidomics provides an avenue for linking important biological processes with disease states. Indeed, differences in lipid profiles have been reported in investigations of cancer [3–6], diabetes [7], Alzheimer’s disease, [8,9] and cardiovascular disease [10,11]. Such studies increasingly rely on high-resolution mass spectrometry (MS) platforms that can detect thousands of lipidomic features in plasma while simultaneously providing accurate masses for annotation [12].

Given strong associations between blood-lipid levels and chronic diseases, it is surprising that baseline lipidomic profiles have not been reported across fundamental population characteristics such as race and gender as well as lifestyle factors such as smoking. Here, we used untargeted Fourier Transform (FT) MS to obtain lipidomic profiles containing over 3,000 features detected in plasma from healthy American subjects stratified by race (black and white), gender and smoking status. Race was the strongest classifying factor (23 significant features) followed by smoking status (9 features) and gender (2 features). Identities assigned to race-discriminating features included several plasmalogens (ether phospholipids containing fatty alcohols with vinyl-ether linkages in the sn-1 position and fatty acids with ester linkages in the sn-2 position) that were more abundant in black subjects. Tentative assignments, based on accurate masses, pointed to greater concentrations in white subjects of an endogenous endocannabinoid receptor agonist and a choline. Several unidentified features, with masses suggestive of phospholipids or sphingomyelins, were present at lower concentrations in smoking subjects. Since all of these lipids are physiologically important and some have been associated with chronic diseases, our results suggest that young American adults may be predisposed to diseases because of differing lipid concentrations associated with race, gender and smoking.

Materials and Methods

Reagents

Isopropanol, methanol, chloroform, formic acid, ammonium hydroxide and ammonium formate (10 M, pH 7.4) were from Fisher Scientific. Phosphate buffer saline (PBS) (pH 7.4) was from Invitrogen. Lipid standards of 1-octadecenyl-2-(5Z,8Z,11Z,14Z,17Z,20Z-eicosatetraenoyl)-sn-glycero-3-phosphoethanolamine PE(P-18:0/20:4) and 1-octadecenyl-2-(4Z,7Z,10Z,13Z,16Z,19Z-docosahexaenoyl)-sn-glycero-3-phosphocholine PC(18:0/22:6) were purchased from Avanti Polar Lipids (Alabaster, AL). Water was purified by a Milli-Q Gradient ultrapure water purification system (Millipore, Billerica, MA). All other chemicals were of analytical grade and used without purification.

*Corresponding author: Rappaport SM, School of Public Health, University of California, Berkeley, California, USA, Tel: 510-642-4355; Fax: 5106425815; E-mail: rrappaport@berkeley.edu

Received May 12, 2014; Accepted July 07, 2014; Published July 14, 2014


Copyright: © 2014 Cai X, et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.
Lipid nomenclature

Lipids were named according to Lipid Maps (http://www.lipidmaps.org); e.g. 1-hexadecanoyl-2-octadecenoyl-sn-glycero-3-phosphocholine is designated PC(16:0/18:1) and 1-hexadecenyl-2-eicosatetraenoyl-glycero-3-phosphoethanolamine is PE(P-16:0/20:4). When the fatty acid chain could not be determined, the total number of carbons and double bonds of all fatty acyl chains are given, e.g. PE (38:4).

Plasma samples

Non-fasting blood samples were obtained in heparin from 158 healthy subjects (78 males and 80 females), representing a subset from a previous study conducted by the corresponding author under an approved human-subjects protocol [13]. Within a few hours of collection, plasma was separated from red blood cells by centrifugation. Red cells were washed with an equal volume of PBS, which was added to the plasma and thus reduced plasma concentrations. Plasma samples were frozen and stored at -80°C for approximately 13 y prior to being aliquoted and pooled by combining aliquots from 4 to 6 subjects stratified by race, gender and smoking status. (Pooling was required by our institutional review board to ensure anonymity of subjects). A quality control sample was prepared by pooling 100 µl of each of these 35 pooled samples.

Demographics, smoking and dietary assessment

Demographic characteristics, including race, age, height and weight were obtained with a standardized questionnaire at the time of phlebotomy. Smoking status was based upon current smoking (yes/ no). A semi-quantitative food-frequency questionnaire containing 131 items was used to evaluate average daily consumption of fat (animal, vegetable and cholesterol) over the past six months for each individual. All dietary-intake values were compiled at the Channing Laboratory, Harvard Medical School [16,17].

Extraction of lipids

Lipids were extracted as described previously [18]. Briefly, 100 µl of plasma was thawed on ice and then mixed with 3 ml of chloroform:methanol (2:1:v/v) and 900 µl of phosphate buffered saline (PBS). After vortexing, the mixture was centrifuged at 2000xg for 5 minutes. The bottom layer was collected, dried under N₂, and dissolved in 100 µl chloroform. Extracts were stored at -80°C before LC-MS analysis.

LC-MS analysis

Liquid chromatography-MS analysis was performed with a Surveyor LC system coupled to an LTQ-FTMS, containing a heated electrospray ionization source (ESI) (Thermo Fisher Scientific, Waltham, MA). The MS was operated in both ESI+ and ESI- ionization modes with data collected from m/z 100 to 1200. For LC separation, a Luna C5 column (4.6x50 mm, 100 Å, 5 µm, Phenomenex, Los Angeles, CA) was selected with column and autosampler temperatures maintained at 25°C and 4°C, respectively. The C5 column was selected to elute all potential lipids in the samples, including hydrophobic triacylglycerides and cholesteryl esters. Injection volumes were 20 µl and 25 µl for ESI+ ionization and ESI- ionization, respectively. Mobile phases contained 0.1% formic acid for ESI+ ionization and 0.1% ammonium hydroxide for ESI- ionization. The column was eluted with a gradient of mobile phase A (methanol:50 mM ammonium formate 5:95) and mobile phase B (isopropanol:methanol:50 mM ammonium formate 60:35:5) as follows: 100% A for 5 minutes at 0.1 ml/min; 0-100% B over 15 minutes at 0.4 ml/min; 100% B for 5 minutes at 0.5 ml/min; 0-100% A for 5 minutes at 0.4 ml/min. Blank and QC samples were analyzed after 7 or 8 experimental samples to wash the column and monitor stability.

The MS was tuned with the following high-abundance lipids in several structural classes: tuning in positive mode employed LPC (16:0), LPE (18:1), PC (36:4), PE (34:1) and TAG (58:5) and tuning in negative mode employed FA (16:0), FA (20:4), PI (24:1) and PG (34:1). Several FTMS parameters, namely, mass resolution, maximum injection time, and maximum number of ions collected for each scan, were optimized for sensitivity while maintaining a mass resolution of 100,000. The following settings were used: vaporizer temperature, 280°C; sheath and auxiliary gases, 35 and 15 (arbitrary units); spray voltage, 3.5 kV; capillary temperature, 350°C; capillary voltage, 10 V; tube-lens voltage, 120 V; maximum injection time, 1000 ms; maximum number of ions collected for each scan, 5×10⁵. Mass calibration was carried out with a standard LTQ calibration mixture (Thermo Scientific, Waltham, MA). For untargeted analyses, a full scan was used for the FTMS with a mass resolution of 100,000, and data were recorded in centroid mode. To study structures of discriminating features, tandem MS/MS analyses were performed with the linear ion trap in low-resolution mode with a CID voltage of 30 V. Accurate masses were calculated using the lipid calculator (http://pharmacology.ucdenver.edu/lipidcalc/) and then extracted with a mass tolerance of 10 ppm in the total ion chromatogram (TIC).

Quantitation of analytes

Because PBS had been added to plasma (as erythrocyte washes) at the time of phlebotomy, volumes of diluted plasma varied across the pooled samples in our investigation. Thus, rather than quantifying peaks of unknown lipidomic features relative to internal standards, quantitation was based on dividing each peak intensity by the sum of all peak intensities detected in each pooled sample [19,20].

Data collection and processing

Data were collected continuously over the 30-minutes LC separation using Xcalibur software (Thermo Fisher Scientific). The raw data were converted to mzXML data format using proteoWizard software (Spielberg Family Center for Applied Proteomics, Los Angeles, CA). Peak detection, retention time collection and alignment were performed on the XCMS platform (http://xcmserver.nutr.berkeley.edu/). All data-collection parameters were set to the “HPLC Orbitrap” default values (centroid feature detection, loess non-linear retention time alignment, 0.5 minimum fraction of samples in one group to be a valid group, P-value thresholds=0.05, isotopic ppm error=5, m/z absolute error=0.015) except the following: maximal tolerated m/z deviation in consecutive scans=3.5 ppm; width of overlapping m/z slices (mzwid)=0.005; retention time window (bw)=15 s, minimum peak width=20, maximum peak width=80. Lists of retention times (RT), m/z values and peak intensities were exported to an Excel spreadsheet for processing. As noted previously, the intensity of each peak was normalized to the sum of total intensities in each sample and was then multiplied by 10,000 for statistical analysis.

Characteristics of lipid classes under our instrumental parameters were determined from LC-FTMS of a training set consisting of 193 common lipid species (Supplemental Information, Table S2) representing monooacylglycerols (MAG), diacylglycerols (DAG), triacylglycerols (TAG), glycerophosphocholines (PC), glycerophosphoethanolamines (PE), monoglycerophosphocholines (LPC), monoglycerophosphoethanolamines (LPE), sphingomyelins (SM), and cholesteryl esters (CE). Mass accuracy, precision, and
stability of the method were estimated from repeated analysis of 8 ion peaks representing lipids detected in the quality control sample that covered large ranges of masses, intensities and retention times. As shown in Supplemental Information, Table S1, mass accuracies were less than 6 ppm and coefficients of variation of retention times and peak intensities were 0.10%-0.56% and 4.08%-24.47%, respectively.

Statistical analysis

Because plasma samples were pooled for the current investigation, (4-6 subjects per pooled sample) mean values were used for statistical analyses. A combination of univariate and multivariate statistical models was used to investigate discriminating features. First, two-tailed Student’s t-tests were performed to screen for discriminating features by race, gender and smoking status. Then, analysis of variance (ANOVA) methods were applied using the R platform with significance determined using a non-parametric permutation test with 10,000 observations [21]. False discovery rates (FDR) were corrected using the Benjamini-Hochberg (BH) method to adjust P-values for false discovery involving multiple comparisons [22]. After application of the BH method, 34 significant features were detected.

After putative identification of discriminating features (described below), sources of variation of dihomo-γ-linolenoyl ethanolamide (DGLEA), PC (16:0/18:1), and the sum of 6 plasmalogens [PE(P-18:0/22:6), PE(P-18:0/22:5) and PC(P-18:0/22:6)] were evaluated with multivariable linear models that employed fat (g) as predictor variables. (The sum of plasmalogen levels was evaluated because concentrations of the 6 plasmalogens were highly correlated). Models were constructed using SAS software for Windows (v. 9.3, SAS Institute, Cary, NC).

Structural identification of discriminating features

Preliminary identification relied upon matching accurate masses from FTMS (with a mass tolerance of 10 ppm) with entries in the Human Metabolome Database (HMDB) (http://www.hmdb.ca/), the Structure Database of Lipid Maps (LMSD) (http://www.lipidmaps.org) and the Metabolite and Tandem MS Database (METLIN) (http://metlin.scripps.edu/). Since human plasma rarely contains lipids with odd-numbered fatty acyl chains, matches representing odd-numbered acyl chains were removed. Other filtering rules were constructed based on relative abundances of signals representing molecular ions and their common adducts, as determined from analyses of our training set of 193 lipid species (Supplemental Information, Table S2). Additional structural information was derived from MS/MS analysis and comparisons with reference standards (Supplemental Information, Section 2).

Results

Univariate analyses of demographic characteristics and dietary fat

Table 1 lists summary statistics for the subjects represented by the 35 pooled plasma samples in the current investigation. Subjects were young, with mean ages of 26 years and 25 years for black and white participants, respectively. The mean BMI for black subjects (28.9 kg/m²) was significantly greater than that of white subjects (24.1 kg/m²) and black subjects had significantly higher consumption of all forms of fat. Also, smokers consumed significantly more dietary fats than nonsmokers.

Proiling of plasma lipids

Untargeted lipidomic profiles of the 35 pooled plasma samples and QC samples were obtained by LC-FTMS in both ESI+ and ESI- modes. Many more features were detected in ESI+ mode (n=2,862) than in ESI- mode (n=717). Figure 1a represents a typical base-peak chromatogram of a lipid extract in ESI+ mode. Plasma lipids were mainly located in three time domains: 16-20 minutes, 20.5-23 min, 24-26 minutes. The averaged mass spectra of these three time domains, shown in Figure 1c, display prominent m/z ranges of 300-550, 700-820 and 800-910, respectively. The distribution of m/z is partially annotated in a density map (Figure 1b) which shows time domains of major lipid classes eluting between 19 and 26 minutes. Characteristics of these lipid classes were inferred from 193 lipid molecules in the training set (Supplemental Information, Table S2). Masses of PCs, PEs, SMs, LPCs, and LPEs were mainly detected as [M+H]^+. While MGs, DGs, TGs and CEs were detected as [M+H]^+. Discriminating lipidomic features

To screen for differences associated with race, gender and smoking status, ANOVA models were obtained for each m/z feature (2,862 in ESI+ mode and 717 in ESI- mode) and random permutation tests were performed to establish P-values. The significance of each feature for a given comparison was determined by its P-value after BH correction for false discovery (P-values were truncated at 10^-4). As summarized in Table 2 and Figure 2, a total of 34 discriminating features was detected,

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Black</th>
<th>White</th>
<th>P-value</th>
<th>Male</th>
<th>Gender</th>
<th>Female</th>
<th>P-value</th>
<th>Smokers</th>
<th>Non-smokers</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of samples</td>
<td>18 (9M, 9F); 17 (9M, 8F); 9S, 9NS</td>
<td>17 (9M, 8F); 10S, 7NS</td>
<td>0.148</td>
<td>18 (9B, 9W); 10S, 8NS; 9S, 8NS</td>
<td>26 ± 2; 25 ± 2; 27 ± 3</td>
<td>26 ± 3; 25 ± 2; 27 ± 2</td>
<td>0.143</td>
<td>26.9 ± 3.94; 25.2 ± 2.42; 2.20E-04</td>
<td>0.014</td>
<td></td>
</tr>
<tr>
<td>Age (y)</td>
<td>26 ± 2; 25 ± 2; 27 ± 3</td>
<td>26 ± 3; 25 ± 2; 27 ± 2</td>
<td>0.148</td>
<td>26.5 ± 2.21; 26.6 ± 4.35; 24.1 ± 1.64</td>
<td>0.410</td>
<td>26.9 ± 3.94; 25.2 ± 2.42; 2.20E-04</td>
<td>0.014</td>
<td>0.143</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BMI (kg/m²)</td>
<td>28.0 ± 3.56; 24.1 ± 1.64; 24.1 ± 1.64</td>
<td>25.7 ± 2.21; 26.6 ± 4.35; 24.1 ± 1.64</td>
<td>0.410</td>
<td>26.9 ± 3.94; 25.2 ± 2.42; 2.20E-04</td>
<td>0.014</td>
<td>0.143</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Between 4 and 6 individual plasma specimens were pooled by race, gender and smoking status.

| Dietary Fat | Animal Fat (g) | 50.7 ± 21.82; 39.2 ± 9.84; 1.38E-03 | 54.3 ± 22.47; 45.0 ± 15.78; 0.168 | 57.9 ± 22.03; 40.1 ± 11.01; 5.905E-03 |
| Vegetable Fat (g) | 39.2 ± 14.5; 27.3 ± 5.67; 3.77E-03 | 35.0 ± 13.0; 31.7 ± 12.1; 0.431 | 38.9 ± 14.0; 26.9 ± 6.17; 3.340E-03 |
| Cholesterol (mg) | 389 ± 142; 248 ± 64.7; 8.06E-04 | 356 ± 155; 283 ± 88.2; 0.097 | 371.2 ± 146.33; 261 ± 77.8; 0.010 |

*Based on Student’s t-test.

Table 1: Characteristics and estimated intake of dietary fat across subjects represented by pooled plasma samples.
Figure 1: (a) Base peak chromatogram of 20 μl of human plasma in positive ionization mode; (b) density map (m/z vs. retention time); (c) averaged mass spectra of three different time regions.
Table 2: Significant lipidomic features associated with race, gender and smoking status.

| Race     | 750.54772 | 21.853 | ESI- | 464.34, 303.26, 259.20 | PE (P-18:0/20:4)* | [M-H]- | 11386 | 4.6  | <1.00E-08  | 1.70  |
| Race     | 776.56369 | 21.839 | ESI- | 464.31, 329.24        | PE (P-18:0/22:5)  | [M-H]- | 11393 | 4.7  | 0.014     | 1.56  |
| Race     | 753.56665 | 21.715 | ESI- | Undetected            | Unknown            | 0.018  | 1.43  |
| Race     | 1013.78243| 22.614 | ESI- | 835.69, 726.53, 702.55, 700.54, 329.28, 303.29 | Unknown | <1.00E-08  | 1.52  |
| Race     | 744.55806 | 21.645 | ESI- | 480.36, 281.21, 255.21 | PC (16:0/18:1)    | [M-CH3]- | 5.1   | <1.00E-08  | 0.81  |
| Race     | 745.56145 | 21.648 | ESI- | 480.33, 281.20, 255.23 | Isotope of PC (16:0/18:1) | [M-CH3]- | 0.9   | 0.018     | 0.82  |
| Race     | 746.56465 | 21.669 | ESI- | 481.34, 281.22, 255.19 | Isotope of PC (16:0/18:1) | [M-CH3]- | 7.1   | 0.018     | 0.81  |
| Gender   | 657.49937 | 20.826 | ESI- | Undetected            | Unknown            | <1.00E-08  | 1.79  |
| Gender   | 806.61063 | 21.954 | ESI- | Undetected            | Unknown            | 0.048  | 1.35  |
| Smoking  | 177.10246 | 7.481  | ESI+ | 145.91, 120.01, 117.97, 97.85, 79.87 | Cotinine [M+H]+  | 1.3   | <1.00E-08  | 262  |
| Smoking  | 567.38180 | 7.475  | ESI+ | 391.04, 177.15        | Unknown            | <1.00E-08  | 190  |
| Smoking  | 568.38515 | 7.475  | ESI+ | 392.29, 177.09        | Unknown            | <1.00E-08  | 598  |
| Smoking  | 745.56131 | 21.475 | ESI+ | Undetected            | Unknown            | <1.00E-08  | 0.77  |
| Smoking  | 743.61051 | 21.868 | ESI- | 711.10, 462.31, 279.26 | Unknown            | <1.00E-08  | 0.84  |
| Smoking  | 757.62638 | 21.968 | ESI- | 739.04, 717.26, 697.40, 667.07, 278.17, 271.23 | Unknown            | <1.00E-08  | 0.74  |
| Smoking  | 769.62632 | 21.923 | ESI- | 751.03, 727.09, 709.24, 511.19, 279.24, 281.30 | Unknown            | <1.00E-08  | 0.85  |
| Smoking  | 783.64208 | 22.036 | ESI- | 767.05, 710.95, 521.06 | Unknown            | <1.00E-08  | 0.82  |
| Smoking  | 784.64551 | 22.036 | ESI- | 767.10, 711.98        | Unknown            | <1.00E-08  | 0.80  |

* Confirmed by reference standard
* Human Metabolome Database
* P-values are based on ANOVA with BH correction
* Fold change based upon ratios of mean values for black/white, male/female or smoker/nonsmoker

Figure 2: Scatter plots of P-values from ANOVA with BH correction, based on data from (a) positive ionization mode and (b) negative ionization mode. Points above the bold line indicate a BH-adjusted P-value less than 0.05. (Points were truncated at a P-value of 10⁻⁸).
namely, 23 for race, 9 for smoking status and 2 for gender. These features were concentrated in the m/z region between 650 Da and 850 Da, which is the domain of phospholipids and sphingomyelins (Figure 1b).

Sixteen of these features were putatively identified by combinations of accurate mass, retention time, MS/MS fragment ions and reference standards (details are given in Supplemental Information, Section 2). The sole non-lipid feature was identified as cotinine (m/z 177.10246), a metabolite of nicotine that was 262 times more abundant in smokers than in nonsmokers. The other tentatively identified features were all lipids that significantly discriminated for race. These race-discriminating lipids included 6 plasmalogenoids [PE(P-16:0/20:4) [m/z 724.53126 [M-H]-; PE(P-18:1/20:4) [m/z 750.54676 [M-H]-, 750.54772 [M-H], 753.56640 [M-H]+ (isotope), 754.58059 [M-H]+ (isotope)]; PE(P-18:0/22:6) [m/z 776.56270 [M-H]+]; PE(P-18:0/22:5) [m/z 778.57958 [M-H], 776.56369 [M-H]]; and PC(P-16:0/18:1) [m/z 744.55806 [M-CH3]-, 745.56145 [M-CH3]] that were present at 4-fold greater concentrations in black subjects than in white subjects and PC(16:0/18:1) [m/z 745.56415 [M-H], 745.56465 [M-H]] (isotope) that was present at moderately higher concentrations in white subjects (1.23-fold difference).

Multivariable linear regression models were used to investigate whether levels of the race-related lipids were affected by gender, BMI or difference). Using untargeted lipidomics with LC-FTMS, plasma-lipid changes that are present at 4-fold greater concentrations in white subjects and PC(16:0/18:1) [m/z 744.55806 [M-H], 745.56415 [M-H]] (isotope), m/z 746.56465 [M-H] (isotope)) that was present at 4-fold greater concentrations in white subjects (1.23-fold difference).

Table 3: Multivariable linear models of covariate effects on analytes representing significant lipidomic features. (Analyte levels were modeled as ratios of individual peak intensities divided by the sum of all peak intensities).

**Discussion**

Using untargeted lipidomics with LC-FTMS, plasma-lipid changes related to race, gender, and smoking status were detected in healthy young American adults. The fact that baseline concentrations of these lipids differ between black and white subjects could be relevant to interpretation of findings that chronic diseases are more prevalent in black Americans than white Americans [26–28].

Most of the race-discriminating lipids were plasmalogens that were present at 2-fold higher levels in black subjects. Plasmalogens are required for membrane integrity and messaging [29,30] and serve as free radical scavengers [31,32]. Thus, these lipids are generally regarded as health promoting and several plasmalogens were recently detected at significantly lower concentrations in subjects with pancreatic cancer than in control subjects [6]. On the other hand, some oxidation products of plasmalogens can be toxic [33–36]. Since animal fat is the major source of plasmalogens in Western diets [37], the observed differences could reflect higher dietary intake of animal fat in black and male subjects (Table 1). Indeed, dietary consumption of plasmalogens increased plasma levels of these lipids in rats [37]. Self-reported consumption of animal fat (but not vegetable fat) explained a small amount of the variability of plasmalogens concentrations in our subjects (Table 3) after adjusting for race and gender. The fact that race was a much stronger predictor of plasmalogen levels than animal fat in our study could point to imprecision in dietary assessment of fat consumption and from aggregation of subjects by race/gender pooling. Higher plasmalogen levels in black and male subjects could also point to differential plasmalogen biosynthesis, possibly related to peroxisome activity [29,30]. Although BMI was significantly greater in black subjects, it is noteworthy that BMI did not explain additional variability of identified features after adjustment for race and gender in multivariable models.

Putative DGLEA, which was found at 4-fold higher concentrations in white subjects, is an endocannabinoid that binds to receptors (CB1 and CB2) that are also the targets of tetrahydrocannabinol, the principal active component of marijuana [23]. Upon activation of at least one of these receptors, specific physiological short-range events can be triggered, including neurotransmitter release. Effects of these reactions include analgesia, increased appetite and neural tissue development [38]. Although the endocannabinoid system and its effects are not well understood, disruption of this system has been implicated in metabolic syndrome and accumulation of excess visceral fat [39]. Since the corresponding acid (DGLA) has been shown to have minimal differences across racial groups [40], the differentiation event may occur in the pathway between DGLA and the transformation to an ethanolamide.

Our untargeted lipid profiling discovered 6 features in the mass range between 740 and 790 Da that were present at lower concentrations in smoking subjects. Since this mass range is characteristic of phospholipids or sphingomyelins, our results lend support to the hypothesis that smoking interferes with metabolism of these lipid classes as suggested by targeted profiling of serum samples from smokers and non-smokers by Wang-Sattler et al. [41]. Interestingly, we also found that the level of the PC plasmalogen, PC(P-18:0/22:6), which was associated with race in our study, was approximately 33% lower in smokers, compared with non-smokers, consistent with the targeted study [41]. The two features with m/z 567.38180 and 568.38515 (Table 2) were highly correlated with cotinine (Spearman r=0.928, 0.948), suggesting that they are metabolites or reaction products of tobacco.
Because we used archived plasma from a previous investigation [13], our study has several limitations. First, it was necessary to pool the specimens - and thereby anonymize subjects' identities - while retaining testable factors (race, gender and smoking status). Although pooling is generally undesirable for small studies and could have reduced our ability to detect significant differences in population characteristics, those features that differed between races and genders (DGLEA and plasmalogens) are unlikely to be false positives [42]. Second, the blood samples had been obtained from non-fasting subjects, and this could have affected profiles of plasma lipids irrespective of race/gender/smoking status. Third, the blood sampling protocol employed hepaminized plasma, and differences in concentrations of numerous lipids have been observed across blood samples collected with different anticoagulants, including heparin [43]. Thus, the 6 plasmalogens and putative DGLEA and PC (16:0/18:1) should be interpreted as lipidomic features that differed significantly between races and genders in samples of heparinized plasma obtained from non-fasting subjects and stored for a prolonged period at -80°C. Finally, as noted previously, archived plasma samples from the 158 individual subjects in the original investigation had been diluted with varying volumes of erythrocyte washes. This effectively precluded quantitation based on internal standards and motivated us to normalize individual features by the sum of all detected peaks. While this method of quantitation could also have reduced precision – and the ability to detect discriminating features - it should not have generated false positives.

Acknowledgements

This research was supported by the National Institute for Environmental Health Sciences through grant P42ES04705. We thank Wendy McKelvey for assistance with collection of blood samples and food frequency data.

References


