

## Research Article

## Open Access

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

Cai X<sup>1</sup>, Perttula K<sup>1</sup>, Pajouh SK<sup>2</sup>, Hubbard A<sup>2</sup>, Nomura DK<sup>3</sup> and Rappaport SM<sup>1\*</sup><sup>1</sup>University of California, Berkeley; School of Public Health; Program in Environmental Health Sciences, USA<sup>2</sup>University of California, Berkeley; School of Public Health; Program in Biostatistics, USA<sup>3</sup>University 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- $\gamma$ -linolenoyl ethanolamide (DGLEA), an endogenous endocannabinoid receptor agonist and a glycerophosphocholine [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- $\gamma$ -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: Phosphoinositol; 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-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: [srappaport@berkeley.edu](mailto:srappaport@berkeley.edu)

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

**Citation:** Cai X, Perttula K, Pajouh SK, Hubbard A, Nomura DK, et al. (2014) Untargeted Lipidomic Profiling of Human Plasma Reveals Differences due to Race, Gender and Smoking Status. *Metabolomics* 4: 131. doi:[10.4172/2153-0769.1000131](https://doi.org/10.4172/2153-0769.1000131)

**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 [14,15]. 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 2000×g for 5 minutes. The bottom layer was collected, dried under N<sub>2</sub>, 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 (4.6×50 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, 5×10<sup>5</sup>. 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-minute 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 processed on the XCMS platform (<http://xcmsserver.nutr.berkeley.edu/>). All data-collection parameters were set to the "HPLC Orbitrap" default values (centwave 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 monoacylglycerols (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- $\gamma$ -linolenoyl ethanolamide (DGLEA), PC (16:0/18:1), and the sum of 6 plasmalogens [PE(P-16:0/20:4), PE(P-18:1/20:4), PE(P-18:0/20:4), 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 combinations of race, gender, smoking status, BMI and dietary fat (g) as predictor variables. (The sum of plasmalogen levels was used 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/m<sup>2</sup>) was significantly greater than that of white subjects (24.1 kg/m<sup>2</sup>) and black subjects had significantly higher consumption of all forms of fat. Also, smokers consumed significantly more dietary fats than nonsmokers.

### Profiling 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]<sup>+</sup>, while MGs, DGs, TGs and CEs were detected as [M+NH<sub>4</sub>]<sup>+</sup>.

### 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<sup>-8</sup>). As summarized in Table 2 and Figure 2, a total of 34 discriminating features was detected,

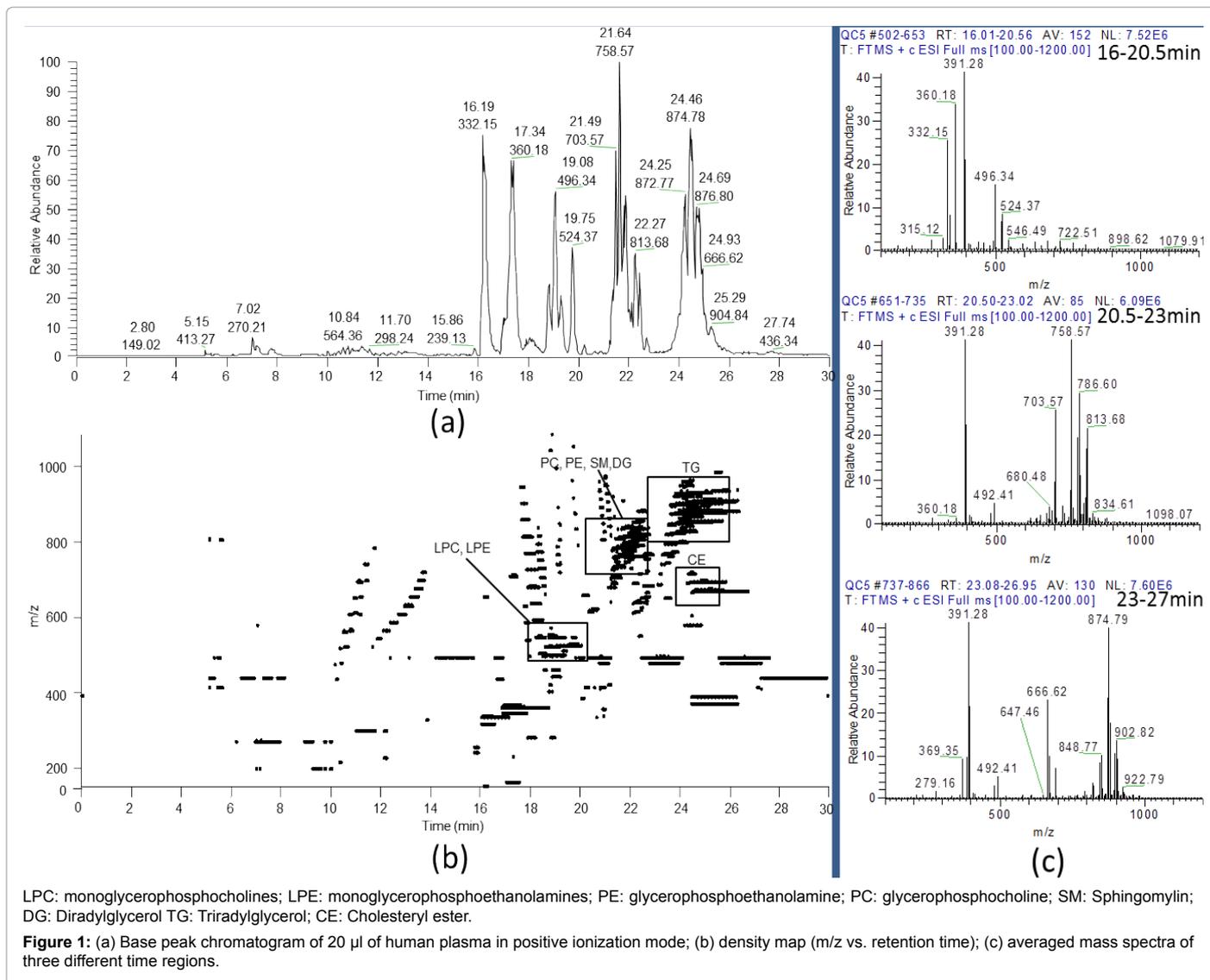
Characteristic <sup>a</sup>	Race			Gender			Smoking Status		
	Black	White	<i>P</i> -value <sup>c</sup>	Male	Female	<i>P</i> -value <sup>c</sup>	Smokers	Non-smokers	<i>P</i> -value <sup>c</sup>
Number of samples	18 (9M, 9F;	17 (9M, 8F;		18 (9B, 9W;	17 (9B, 8W;		19 (10M, 9F;	16 (8M, 8F;	
	9S, 9NS)	10S, 7NS)		10S, 8NS)	9S, 8NS)		9B, 10W)	9B, 7W)	
Age (y)	26 ± 2	25 ± 2	0.148	26 ± 3	25 ± 2	0.627	26 ± 3	25 ± 2	0.516
BMI (kg/m <sup>2</sup> )	28.0 ± 3.56	24.1 ± 1.64	2.20E-04	25.7 ± 2.21	26.6 ± 4.35	0.410	26.9 ± 3.94	25.2 ± 2.42	0.143
<b>Dietary Fat<sup>b</sup></b>									
Animal Fat (g)	59.7 ± 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

<sup>a</sup>Between 4 and 6 individual plasma specimens were pooled by race, gender and smoking status.

<sup>b</sup>Dietary fats and fatty acids were compiled from standardized food frequency questionnaires applied to individual subjects and averaged for pooled-plasma specimens in the current investigation.

<sup>c</sup>Based on Student's *t*-test.

**Table 1:** Characteristics and estimated intake of dietary fat across subjects represented by pooled plasma samples.



Factor	m/z	Ret. time (min)	Ionization-mode	MSMS fragments	Compound	Species	HMDB ID <sup>a</sup>	Appm	P-value <sup>b</sup>	Fold change <sup>c</sup>
Race	724.53126	21.861	ESI+	583.40, 364.22, 361.28	PE (P-16:0/20:4)	[M+H] <sup>+</sup>	08937	5.1	<1.00E-08	2.57
Race	750.54676	21.924	ESI+	609.42, 390.31, 361.23	PE (P-18:1/20:4)	[M+H] <sup>+</sup>	11419	4.7	<1.00E-08	1.77
Race	751.55602	21.945	ESI+	609.42, 390.27, 361.21	Isotope of PE (P-18:1/20:4)	[M+H] <sup>+</sup>	11419	6.4	<1.00E-08	1.98
Race	752.56303	22.071	ESI+	611.43, 392.32, 361.28	PE (P-18:0/20:4)*	[M+H] <sup>+</sup>	11386	5.6	<1.00E-08	1.81
Race	753.56640	22.062	ESI+	611.44, 392.28, 361.28	Isotope of PE (P-18:0/20:4)*	[M+H] <sup>+</sup>	11386	0.6	<1.00E-08	2.00
Race	754.58059	22.114	ESI+	612.46, 392.34, 362.30	Isotope of PE (P-18:0/20:4)*	[M+H] <sup>+</sup>	11386	7.6	<1.00E-08	2.31
Race	776.56270	22.044	ESI+	635.44, 392.33, 385.29	PE (P-18:0/22:6)	[M+H] <sup>+</sup>	11394	5.0	<1.00E-08	1.81
Race	778.57958	22.076	ESI+	637.45, 392.22, 387.39	PE (P-18:0/22:5)	[M+H] <sup>+</sup>	11394	6.5	<1.00E-08	1.49
Race	732.55708	21.525	ESI+	503.25, 311.18	Unknown				<1.00E-08	0.66
Race	733.56106	21.517	ESI+	504.33, 311.32	Unknown				<1.00E-08	0.68
Race	195.08792	10.185	ESI+	Undetected	Unknown				<1.00E-08	0.26
Race	410.36418	20.121	ESI+	392.29, 350.30, 326.28, 186.13, 168.11	8,11,14-Eicosatrienoylethanolamide (DGLEA)	[M+Isopro+H] <sup>+</sup>	13625	1.9	<1.00E-08	0.23
Race	766.57813	21.844	ESI+	Undetected	Unknown				0.026	1.41
Race	818.61083	21.944	ESI+	635.58, 550.35, 508.59	PC(P-18:0/22:6)*	[M+H] <sup>+</sup>	11262	6.2	0.026	1.45
Race	728.58275	21.400	ESI+	Undetected	Unknown				0.033	1.34
Race	917.84378	24.052	ESI+	900.62, 618.43, 604.57	Unknown				0.033	0.62

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, 279.17, 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

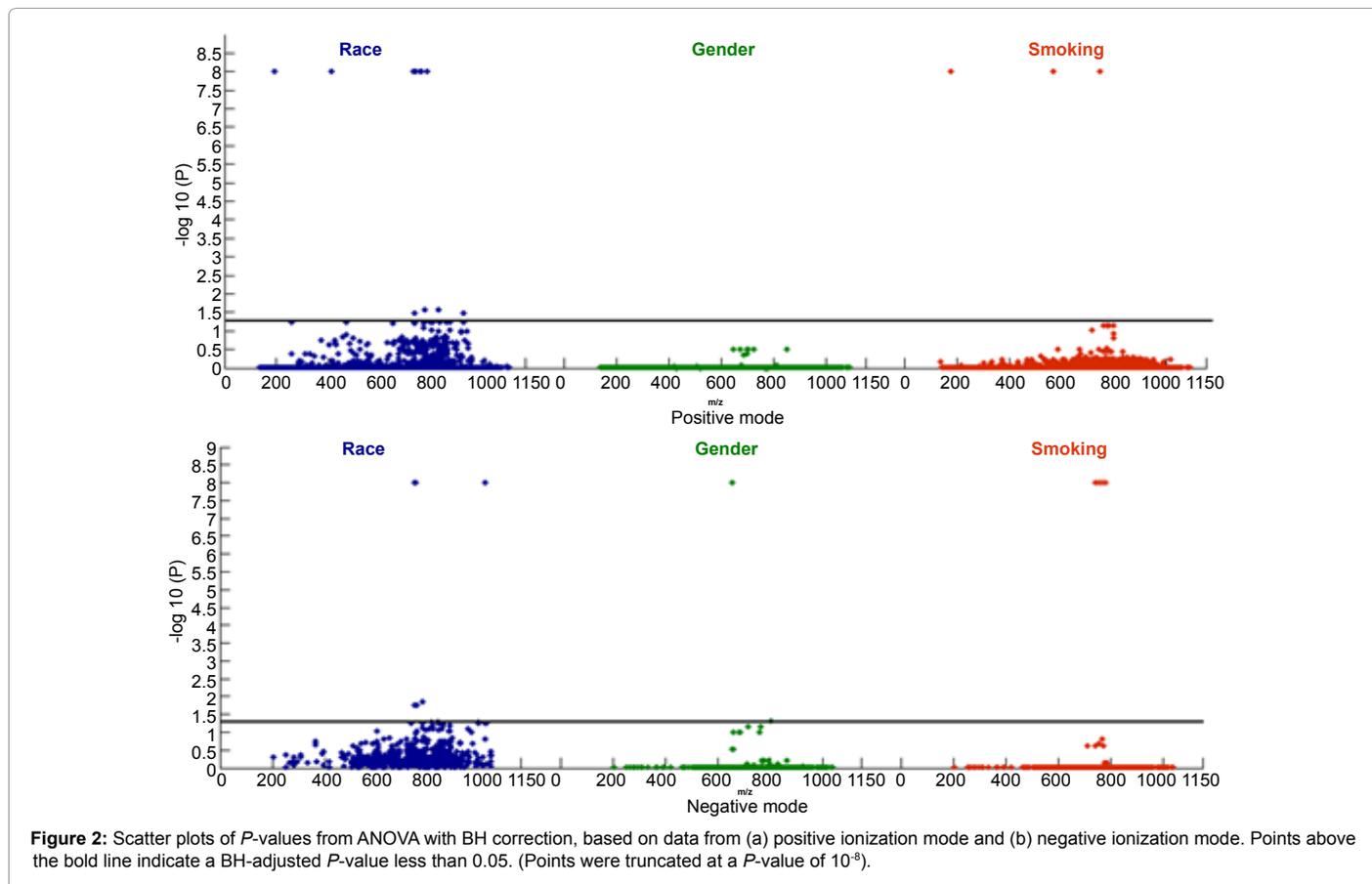
\* Confirmed by reference standard

<sup>a</sup> Human Metabolome Database

<sup>b</sup> P-values are based on ANOVA with BH correction

<sup>c</sup> Fold change based upon ratios of mean values for black/white, male/female or smoker/nonsmoker

**Table 2:** Significant lipidomic features associated with race, gender and smoking status.



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 plasmalogens [PE(P-16:0/20:4) [*m/z* 724.53126 [M<sup>+</sup>H]<sup>+</sup>]; PE(P-18:1/20:4) [*m/z* 750.54676 [M<sup>+</sup>H]<sup>+</sup>, 751.55602 [M<sup>+</sup>H]<sup>+</sup> (isotope)]; PE(P-18:0/20:4) [*m/z* 752.56303 [M<sup>+</sup>H]<sup>+</sup>, 750.54772 [M<sup>+</sup>H]<sup>+</sup>, 753.56640 [M<sup>+</sup>H]<sup>+</sup> (isotope), 754.58059 [M<sup>+</sup>H]<sup>+</sup> (isotope)]; PE(P-18:0/22:6) [*m/z* 776.56270 [M<sup>+</sup>H]<sup>+</sup>]; PE(P-18:0/22:5) [*m/z* 778.57958 [M<sup>+</sup>H]<sup>+</sup>, 776.56369 [M<sup>+</sup>H]<sup>+</sup>]; and PC(P-18:0/22:6) [*m/z* 818.61083 [M<sup>+</sup>H]<sup>+</sup>] that were present at approximately 2-fold greater concentrations in black subjects. Two other race-related features were tentatively identified from accurate masses, namely, DGLA, an agonist of the endocannabinoid receptors (CB1 and CB2) [23–25] that was present at 4-fold greater concentrations in white subjects and PC(16:0/18:1) [*m/z* 744.55806 [M<sup>+</sup>CH<sub>3</sub>]<sup>+</sup>, 745.56145 [M<sup>+</sup>CH<sub>3</sub>]<sup>+</sup> (isotope), *m/z* 746.56465 [M<sup>+</sup>CH<sub>3</sub>]<sup>+</sup> (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 consumption of fat as recorded by 6-month dietary recall. According to the R<sup>2</sup> values of the regression models (Table 3), race accounted for 68.0% of the summed plasmalogen levels, gender for 6.1% and consumption of animal fat for 2.2% (vegetable fat was not associated with plasmalogen levels). The model for putative DGLA showed that race accounted for 50% of the variation and that race, gender and their interaction jointly explained 62%, with white males having 7 times higher concentrations than black females. Race was the only significant predictor for putative PC (16:0/18:1) and explained 45% of the variance. With race in each model, BMI did not significantly contribute to explained variability.

Analyte	Parameter	Estimate	P-value	R <sup>2</sup>	ΔR <sup>2 a</sup>
Summed Plasmalogens	Intercept	3.786			
	Race black	3.336	<0.0001	0.6801	-
	Gender (male)	0.912	0.036	0.7414	0.0613
	Animal fat (per g)	0.022	0.093	0.7642	0.0228
DGLA <sup>b</sup>	Intercept	0.015			
	Race (black)	-0.063	0.016	0.504	-
	Gender (male)	0.074	0.0058	0.5343	0.0303
	Race×Gender	-0.091	0.014	0.6179	0.0836
PC(P-16:0/18:1)	Intercept	70.297			
	Race (black)	-13.07	<0.0001	0.4503	

<sup>a</sup> Change of R<sup>2</sup> value after adding the covariate to the model.

<sup>b</sup> Dihomo-γ-linolenoyl ethanolamide

**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 [3,31,32]. 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 plasmalogen 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 DGLA, which was found at 4-fold higher concentrations in white subjects, is an endocannabinoid that binds to receptors (CB1 or 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], a differentiating 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 heparinized 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

1. Rainville PD, Stumpf CL, Shockcor JP, Plumb RS, Nicholson JK (2007) Novel application of reversed-phase UPLC-*oa*TOF-MS for lipid analysis in complex biological mixtures: a new tool for lipidomics. *J Proteome Res* 6: 552-558.
2. Brown HA, Murphy RC (2009) Working towards an exegesis for lipids in biology. *Nat Chem Biol* 5: 602-606.
3. Smith RE, Lespi P, Di Luca M, Bustos C, Marra FA, et al. (2008) A reliable biomarker derived from plasmalogens to evaluate malignancy and metastatic capacity of human cancers. *Lipids* 43: 79-89.
4. Ma X, Yang J (2009) Lipidomics in Cancer Biomarker Discovery. Chapter B. Omi. Technol. Cancer Biomark. Discov. Cancer Gen.
5. Ritchie SA, Ahiahonu PWK, Jayasinghe D, et al. (2010) Reduced levels of hydroxylated, polyunsaturated ultra long-chain fatty acids in the serum of colorectal cancer patients: implications for early screening and detection.
6. Ritchie SA, Akita H, Takemasa I, Eguchi H, Pastural E, et al. (2013) Metabolic system alterations in pancreatic cancer patient serum: potential for early detection. *BMC Cancer* 13: 416.
7. Gross RW, Han X (2007) Lipidomics in diabetes and the metabolic syndrome. *Methods Enzymol* 433: 73-90.
8. Han X, Rozen S, Boyle SH, Hellegers C, Cheng H, et al. (2011) Metabolomics in early Alzheimer's disease: identification of altered plasma sphingolipidome using shotgun lipidomics. *PLoS One* 6: e21643.
9. Chan RB, Oliveira TG, Cortes EP, Honig LS, Duff KE, et al. (2012) Comparative lipidomic analysis of mouse and human brain with Alzheimer disease. *J Biol Chem* 287: 2678-2688.
10. Graessler J, Schwudke D, Schwarz PE, Herzog R, Shevchenko A, et al. (2009) Top-down lipidomics reveals ether lipid deficiency in blood plasma of hypertensive patients. *PLoS One* 4: e6261.
11. Cole LK, Dolinsky VW, Dyck JRB, Vance DE (2011) Impaired phosphatidylcholine biosynthesis reduces atherosclerosis and prevents lipotoxic cardiac dysfunction in ApoE<sup>-/-</sup> Mice. *Circ Res* 108:686-694.
12. Quehenberger O, Armando AM, Brown AH, Milne SB, Myers DS, et al. (2010) Lipidomics reveals a remarkable diversity of lipids in human plasma. *J Lipid Res* 51: 3299-3305.
13. Lin YS, McKelvey W, Waidyanatha S, Rappaport SM (2006) Variability of albumin adducts of 1,4-benzoquinone, a toxic metabolite of benzene, in human volunteers. *Biomarkers* 11: 14-27.
14. Rimm EB, Giovannucci EL, Stampfer MJ, Colditz GA, Litin LB, et al. (1992) Reproducibility and validity of an expanded self-administered semiquantitative food frequency questionnaire among male health professionals. *Am J Epidemiol* 135: 1114-1126.
15. Hu FB, Rimm E, Smith-Warner SA, Feskanich D, Stampfer MJ, et al. (1999) Reproducibility and validity of dietary patterns assessed with a food-frequency questionnaire. *Am J Clin Nutr* 69: 243-249.
16. Willett WC, Sampson L, Stampfer MJ, Rosner B, Bain C, et al. (1985) Reproducibility and validity of a semiquantitative food frequency questionnaire. *Am J Epidemiol* 122: 51-65.
17. Salvini S, Hunter DJ, Sampson L, Stampfer MJ, Colditz GA, et al. (1989) Food-based validation of a dietary questionnaire: the effects of week-to-week variation in food consumption. *Int J Epidemiol* 18: 858-867.
18. Nomura DK, Long JZ, Niessen S, Hoover HS, Ng SW, et al. (2010) Monoacylglycerol lipase regulates a fatty acid network that promotes cancer pathogenesis. *Cell* 140: 49-61.
19. Wang W, Zhou H, Lin H, Roy S, Shaler TA, et al. (2003) Quantification of proteins and metabolites by mass spectrometry without isotopic labeling or spiked standards. *Anal Chem* 75: 4818-4826.
20. Sysi-Aho M, Katajamaa M, Yetukuri L, Oresic M (2007) Normalization method for metabolomics data using optimal selection of multiple internal standards. *BMC Bioinformatics* 8: 93.
21. Anderson MJ, Ter Braak CJF (2003) Permutation tests for multi-factorial analysis of variance. *J Stat Comput Simul* 73: 85-113.
22. Benjamini Y, Hochberg Y (1995) Controlling the False Discovery Rate: a Practical and Powerful Approach to Multiple Testing. *J R Stat Soc* 57: 289-300.
23. Hanus L, Gopher A, Almog S, Mechoulam R (1993) Two new unsaturated fatty acid ethanolamides in brain that bind to the cannabinoid receptor. *J Med Chem* 36: 3032-3034.
24. Felder CC, Joyce KE, Briley EM, Mansouri J, Mackie K, et al. (1995) Comparison of the pharmacology and signal transduction of the human cannabinoid CB1 and CB2 receptors. *Mol Pharmacol* 48: 443-450.
25. Balvers MG, Verhoeckx KC, Bijlsma S, Rubingh CM, Meijerink J, et al. (2012) Fish oil and inflammatory status alter the n-3 to n-6 balance of the endocannabinoid and oxylipin metabolomes in mouse plasma and tissues. *Metabolomics* 8: 1130-1147.
26. Mokdad AH, Ford ES, Bowman BA, Nelson DE, Engelgau MM, et al. (2000) Diabetes trends in the U.S.: 1990-1998. *Diabetes Care* 23: 1278-1283.
27. Hayes DK, Greenland KJ, Denny CH, et al. (2005) Racial/ethnic and socioeconomic disparities in multiple risk factors for heart disease and stroke - United States, 2003. *Jama-Journal Am Med Assoc* 293: 1441-1443.
28. Rodriguez CJ, Diez-Roux AV, Moran A, Jin Z, Kronmal RA, et al. (2010) Left ventricular mass and ventricular remodeling among Hispanic subgroups compared with non-Hispanic blacks and whites: MESA (Multi-ethnic Study of Atherosclerosis). *J Am Coll Cardiol* 55: 234-242.
29. Wallner S, Schmitz G (2011) Plasmalogens the neglected regulatory and scavenging lipid species. *Chem Phys Lipids* 164: 573-589.
30. Braverman NE, Moser AB (2012) Functions of plasmalogen lipids in health and disease. *Biochim Biophys Acta* 1822: 1442-1452.
31. Colas R, Pruneta-Deloche V, Guichardant M, Luquain-Costaz C, Cugnet-Anceau C, et al. (2010) Increased lipid peroxidation in LDL from type-2 diabetic patients. *Lipids* 45: 723-731.
32. Igarashi M, Ma K, Gao F, Kim HW, Rapoport SI, et al. (2011) Disturbed choline plasmalogen and phospholipid fatty acid concentrations in Alzheimer's disease prefrontal cortex. *J Alzheimers Dis* 24: 507-517.
33. Felde R, Spittler G (1995) Plasmalogen oxidation in human serum lipoproteins. *Chem Phys Lipids* 76: 259-267.
34. Loidl-Stahlhofen A, Hannemann K, Felde R, Spittler G (1995) Epoxidation of

- plasmalogens: source for long-chain alpha-hydroxyaldehydes in subcellular fractions of bovine liver. *Biochem J* 309 : 807-812.
35. Hecht SS, Seow A, Wang M, Wang R, Meng L, et al. (2010) Elevated levels of volatile organic carcinogen and toxicant biomarkers in Chinese women who regularly cook at home. *Cancer Epidemiol Biomarkers Prev* 19: 1185-1192.
36. Wang G, Wang T (2010) The role of plasmalogen in the oxidative stability of neutral lipids and phospholipids. *J Agric Food Chem* 58: 2554-2561.
37. Nishimukai M, Wakisaka T, Hara H (2003) Ingestion of plasmalogen markedly increased plasmalogen levels of blood plasma in rats. *Lipids* 38: 1227-1235.
38. Nomura DK, Morrison BE, Blankman JL, Long JZ, Kinsey SG, et al. (2011) Endocannabinoid hydrolysis generates brain prostaglandins that promote neuroinflammation. *Science* 334: 809-813.
39. Kim J, Li Y, Watkins BA (2013) Fat to treat fat: emerging relationship between dietary PUFA, endocannabinoids, and obesity. *Prostaglandins Other Lipid Mediat* 104-105: 32-41.
40. Steffen BT, Steffen LM, Tracy R, Siscovick D, Jacobs D, et al. (2012) Ethnicity, plasma phospholipid fatty acid composition and inflammatory/endothelial activation biomarkers in the Multi-Ethnic Study of Atherosclerosis (MESA). *Eur J Clin Nutr* 66: 600-605.
41. Wang-Sattler R, Yu Y, Mittelstrass K, Lattka E, Altmair E, et al. (2008) Metabolic profiling reveals distinct variations linked to nicotine consumption in humans—first results from the KORA study. *PLoS One* 3: e3863.
42. Sadiq ST, Agranoff D (2008) Pooling serum samples may lead to loss of potential biomarkers in SELDI-ToF MS proteomic profiling. *Proteome Sci* 6: 16.
43. Gonzalez-Covarrubias V, Dane A, Hankemeier T, Vreeken RJ (2012) The influence of citrate, EDTA, and heparin anticoagulants to human plasma LC-MS lipidomic profiling. *Metabolomics* 9:337–348.

**Citation:** Cai X, Perttula K, Pajouh SK, Hubbard A, Nomura DK, et al. (2014) Untargeted Lipidomic Profiling of Human Plasma Reveals Differences due to Race, Gender and Smoking Status. *Metabolomics* 4: 131. doi:[10.4172/2153-0769.1000131](https://doi.org/10.4172/2153-0769.1000131)

### Submit your next manuscript and get advantages of OMICS Group submissions

#### Unique features:

- User friendly/feasible website-translation of your paper to 50 world's leading languages
- Audio Version of published paper
- Digital articles to share and explore

#### Special features:

- 350 Open Access Journals
- 30,000 editorial team
- 21 days rapid review process
- Quality and quick editorial, review and publication processing
- Indexing at PubMed (partial), Scopus, EBSCO, Index Copernicus and Google Scholar etc
- Sharing Option: Social Networking Enabled
- Authors, Reviewers and Editors rewarded with online Scientific Credits
- Better discount for your subsequent articles

Submit your manuscript at: <http://www.omicsonline.org/submission/>