

# SCIENTIFIC REPORTS



OPEN

## Obesity, adipokines, and C-peptide are associated with distinct plasma phospholipid profiles in adult males, an untargeted lipidomic approach

C. Austin Pickens<sup>1</sup>, Ana I. Vazquez<sup>2</sup>, A. Daniel Jones <sup>3,4</sup> & Jenifer I. Fenton<sup>1</sup>

Obesity is associated with dysregulated lipid metabolism and adipokine secretion. Our group has previously reported obesity and adipokines are associated with % total fatty acid (FA) differences in plasma phospholipids. The objective of our current study was to identify in which complex lipid species (i.e., phosphatidylcholine, sphingolipids, etc) these FA differences occur. Plasma lipidomic profiling ( $n = 126$ , >95% Caucasian, 48–65 years) was performed using chromatographic separation and high resolution tandem mass spectrometry. The responses used in the statistical analyses were body mass index (BMI), waist circumference (WC), serum adipokines, cytokines, and a glycemic marker. High-dimensional statistical analyses were performed, all models were adjusted for age and smoking, and p-values were adjusted for false discovery. In Bayesian models, the lipidomic profiles (over 1,700 lipids) accounted for >60% of the inter-individual variation of BMI, WC, and leptin in our population. Across statistical analyses, we report 51 individual plasma lipids were significantly associated with obesity. Obesity was inversely associated lysophospholipids and ether linked phosphatidylcholines. In addition, we identify several unreported lipids associated with obesity that are not present in lipid databases. Taken together, these results provide new insights into the underlying biology associated with obesity and reveal new potential pathways for therapeutic targeting.

Obesity is associated with insulin resistance, oxidative stress, chronic low-grade inflammation, and dysregulated secretion of adipose-derived cytokines (i.e., adipokines)<sup>1</sup>. Obesity is also associated with altered lipid metabolism (reviewed in detail ref. 2) and many lipids function as signaling molecules involved in inflammation pathways. The chronic low-grade inflammation in obesity can alter insulin receptor downstream signaling leading to insulin resistance (reviewed in detail ref. 3), and insulin resistance can affect lipid storage and lipid metabolism<sup>4</sup>.

The human plasma lipidome is comprised of lipid species such as glycerolipids, fatty acids (FAs), cholesterol esters, and phospholipids (PLs) including the glycerol-based diacylphospholipid and lysophospholipid (LPL), and the sphingosine-based sphingolipids<sup>5</sup>. PL contain esterified fatty acids (FAs), for instance, there are two esterified FAs in diacylphospholipids such as: phosphatidylcholine (PC), phosphatidylethanolamine (PE), phosphatidylserine (PS), phosphatidylglycerol (PG), and phosphatidylinositol (PI), and there is one esterified FA in their respective LPLs: LPC, LPE, LPS, LPG, and LPI. FAs are endogenously synthesized except for the essential polyunsaturated FA (PUFAs) linoleic acid (LA), a C18:2<sup>9Z,12Z</sup> omega-6 ( $\omega$ -6), and alpha-linolenic acid (ALA), a C18:3<sup>9Z,12Z,15Z</sup> omega-3 ( $\omega$ -3). Obesity is associated with increased elongation, desaturation, and oxygenation of FAs, which can lead to higher levels of LA-derived long chain PUFAs such as arachidonic acid (ARA), a C20:4<sup>5Z,8Z,11Z,14Z</sup>  $\omega$ -6<sup>2,6</sup>, and ARA-derived proinflammatory signaling lipids<sup>7</sup>. Despite FA metabolites being

<sup>1</sup>Department of Food Science and Human Nutrition, Michigan State University, 469 Wilson Road, East Lansing, MI 48824, USA. <sup>2</sup>Department of Epidemiology and Biostatistics, Michigan State University, 909 Fee Road, East Lansing, MI 48824, USA. <sup>3</sup>Department of Biochemistry and Molecular Biology, Michigan State University, 603 Wilson Road, East Lansing, MI 48824, USA. <sup>4</sup>Department of Chemistry, Michigan State University, 578 S Shaw Lane, East Lansing, MI 48824, USA. Correspondence and requests for materials should be addressed to J.I.F. (email: [imigjeni@msu.edu](mailto:imigjeni@msu.edu))

Parameter	Overall n = 126 <sup>a</sup>	Lean n = 28 <sup>a</sup>	Overweight n = 46 <sup>a</sup>	Obese n = 52 <sup>a</sup>	BH FDR p-value <sup>b</sup>
Age (years)	58 [48–65]	58 [50–65] <sup>A</sup>	54 [48–65] <sup>B</sup>	59 [50–65] <sup>A</sup>	p < 0.005
Ever Smoked (% total)	31	25	24	27	—
BMI (kg/m <sup>2</sup> )	29.1 [19.2–45.6]	23.6 [19.2–25.0] <sup>A</sup>	28.1 [25.4–29.9] <sup>B</sup>	33.7 [30.0–45.6] <sup>C</sup>	p < 0.0001
WC (cm)	104.8 [75.6–146.1]	85.7 [76.2–101.6] <sup>A</sup>	101.6 [82.5–113.7] <sup>B</sup>	119.4 [75.6–146.1] <sup>C</sup>	p < 0.0001
Leptin (ng/mL)	6.1 [0.3–49.5]	2.1 [0.3–6.0] <sup>A</sup>	5.2 [0.7–10.2] <sup>B</sup>	13.7 [3.8–49.5] <sup>C</sup>	p < 0.0001
Total Adiponectin (µg/mL)	4.2 [1.0–13.2]	5.4 [1.9–13.2] <sup>A</sup>	4.8 [2.2–8.5] <sup>A</sup>	3.7 [1.0–8.5] <sup>B</sup>	p < 0.005
C-peptide (ng/mL)	2.4 [0.7–9.9]	1.4 [0.7–5.1] <sup>A</sup>	2.1 [0.9–8.1] <sup>B</sup>	3.3 [1.0–9.9] <sup>C</sup>	p < 0.0001
IL-6 (pg/mL)	1.6 [0.0–137]	0.2 [0.0–40.9] <sup>A</sup>	1.0 [0.0–96.6] <sup>A</sup>	3.4 [0.0–137] <sup>A</sup>	—
TNF-α (pg/mL)	7.4 [1.6–67.1]	5.7 [2.8–26.2] <sup>A</sup>	7.3 [1.6–24.1] <sup>A</sup>	8.6 [1.6–67.1] <sup>B</sup>	p < 0.0005
IP-10 (pg/mL)	306 [120–1192]	257.5 [120–557] <sup>A</sup>	271 [145–1029] <sup>A</sup>	378 [147–1192] <sup>B</sup>	p < 0.0005
MCP-1 (pg/mL)	483 [85–1018]	428 [218–875] <sup>A</sup>	488 [285–1018] <sup>AB</sup>	511 [85–922] <sup>B</sup>	p < 0.05

**Table 1.** Median [range] of the participant characteristics, in the sample, and divided by BMI category. Body mass index (BMI), waist circumference (WC), tumor necrosis factor-alpha (TNF-α), interferon gamma-induced protein-10 (IP-10), interleukin-6 (IL-6), monocyte chemoattractant protein-1 (MCP-1). <sup>a</sup>Values expressed as median [range]. Cytokines quantified from serum. <sup>b</sup>Kruskal Wallis one-way ANOVA was conducted across BMI categories along with Dunn's test for multiple comparison. The superscripted letters A, AB, B, and C represent the multiple comparison test for each parameter for the lean, overweight, and obese categories. P-values were adjusted using Benjamini-Hochberg false discovery rate (BH FDR) correction.

directly involved in inflammatory signaling, the metabolism of PLs are also involved in activating signal transducing pathways associated with inflammation regulation [reviewed in detail ref. 8].

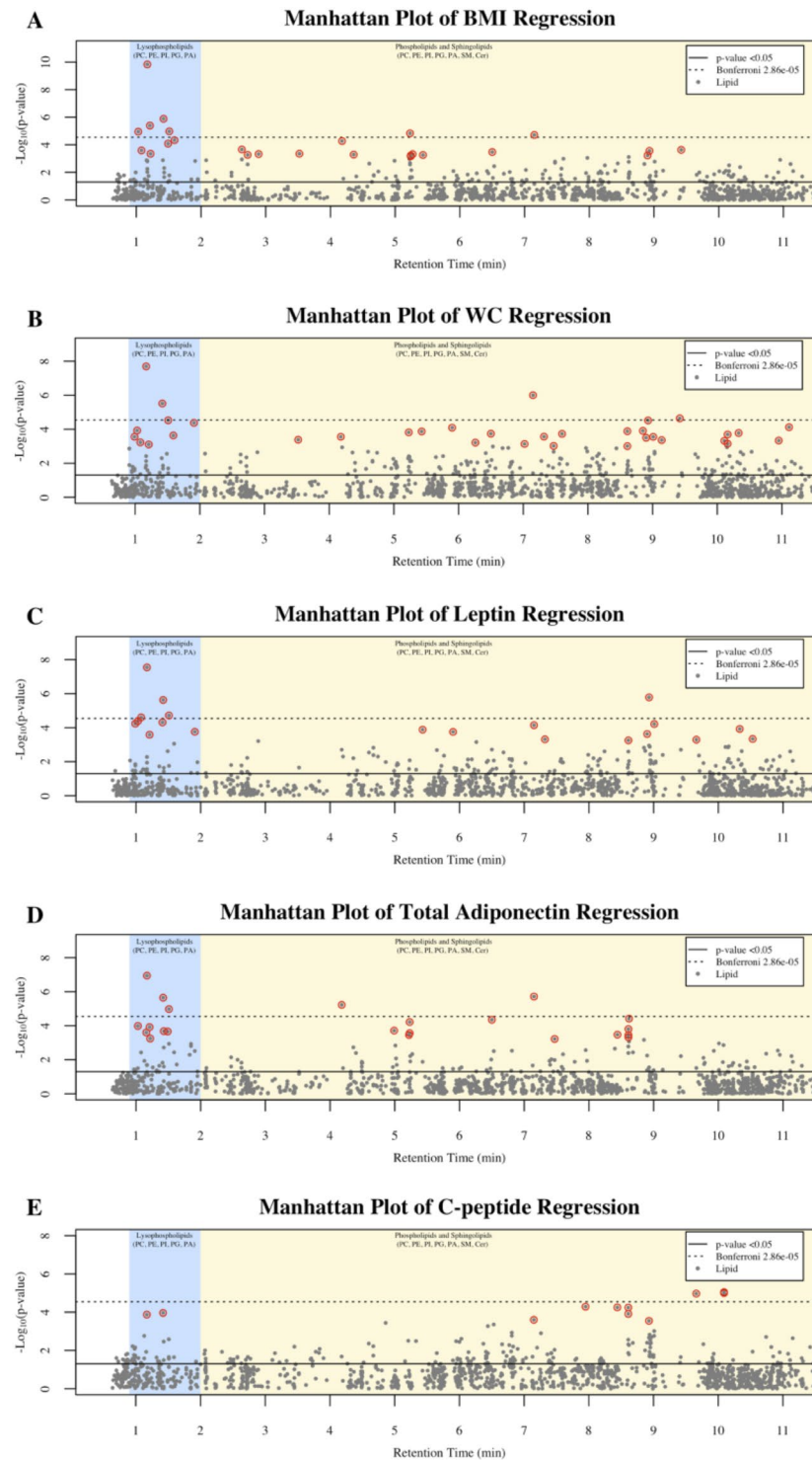
The metabolism of diacylphospholipids results in the production of LPLs, which are generated by lecithin:cholesterol acyltransferase (LCAT)<sup>9</sup>, and phospholipase A<sub>1</sub> (PLA<sub>1</sub>) and A<sub>2</sub> (PLA<sub>2</sub>)<sup>10</sup>. LPLs such as LPCs are ligands for G-protein coupled receptors (reviewed in detail ref. 11) and can exert either pro-inflammatory<sup>12,13</sup> or anti-inflammatory responses depending on the FA chain esterified<sup>14</sup>. Several studies report obesity is inversely associated with LPLs. For instance, overweight and obese individuals have lower LPC(18:1) and (18:2)<sup>15</sup>, and Rauschert *et al.* reported WC is inversely associated with LPC(18:1) and (18:2)<sup>16</sup>. Higher levels of LPC(18:2) were independently associated with a decreased risk of type-2 diabetes in a European Prospective Investigation into Cancer and Nutrition (EPIC)-Potsdam study<sup>17</sup>. In a recent randomized control trial, supplementation with ω-3 PUFAs increased ω-3s in LPLs and lowered ω-6s in LPLs, but overall LPLs were still lower in obese compared with normal and overweight individuals<sup>18</sup>. Some studies suggest that increases in some LPLs such as LPC(18:0) are indicators of high fat diet induced obesity<sup>19</sup>, and in obesity resistant mice, HFD increases plasma LPLs in particular LPC(18:0), (18:1), (18:2), (22:6)<sup>20</sup>. The decrease in plasma LPLs may result from altered lipid metabolism in the liver<sup>21</sup>, since LPLs changes are associated with fatty liver<sup>22,23</sup>. Although there are conflicting reports on whether specific plasma LPLs increase<sup>15,21,24</sup> or decrease<sup>8,15,25</sup> in obesity, there is a clear link between LPLs and obesity-associated comorbidities.

Our group has previously reported that plasma PL FA levels were associated with obesity<sup>26</sup>, adipokines, and the glycemic marker C-peptide<sup>27</sup>. In these studies, the plasma PL FA measurements were obtained using a FA methyl ester (FAME) analysis and by analyzing PL FAs as % of total FAs. PL FAME analyses involve hydrolyzing FAs from their parent PL (i.e., glycerol-based PLs, sphingolipids, etc), meaning the PLs class (i.e., LPLs) associated with the FA differences cannot be determined. We hypothesized that our previously reported plasma PL FA differences associated with obesity<sup>26</sup> and cytokines<sup>27</sup> occur in specific PLs. In our current study, we analyzed plasma collected from the same patients at the same time point as the plasma analyzed in our previous FAME studies<sup>26,27</sup>, to determine these specific PLs with differing FAs. We investigated associations between the traits obesity, adipokines, proinflammatory cytokines, and C-peptide, and intact plasma PLs by using negative-mode high-resolution tandem mass spectrometry (MS/MS) to 1) identify specific plasma PL species with FA differences associated with the traits, and 2) assess the contribution of the entire plasma lipidome on the variation of traits using multidimensional statistical methods.

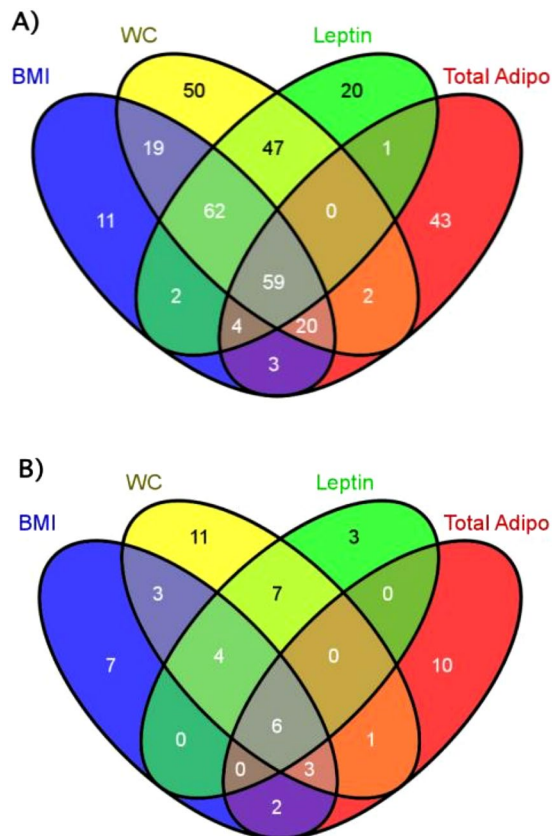
## Results

**Participant characteristics.** Median and range values of age, smoking, anthropometric, and serum adipokines, glycemic, and inflammatory markers of the overall population (n = 126) and separated by BMI categories are presented in Table 1. In brief, lean and obese participants were older than overweight participants. In the obese group, 19 patients had BMI > 35 and only 6 patients had BMI > 40. Both BMI and WC increased with increasing BMI category. Serum leptin and C-peptide significantly increased across BMI category, and serum adiponectin was decreased in obese participants compared to lean and overweight participants. Serum TNF-α and IP-10 were significantly elevated in obese participants compared to overweight and lean participants, while MCP-1 was only elevated in obese participants compared to lean.

**Lipidomic profiles explain a large proportion of inter-individual variation of obese traits.** The lipidomic profiles accounted for an important proportion of the inter-individual variation of all responses (Table 2). For several traits (i.e., BMI, WC, and leptin) lipidomic profiles accounted for over 60% of the variation in the samples (Table 2). Lipidomic profiles accounted for roughly 40% of the variation in total-adiponectin



**Figure 1.** Manhattan plots of the  $-\log_{10}(p\text{-value})$  for the traits: (A) BMI, (B) WC, (C) Leptin, (D) Total Adiponectin, and (E) C-peptide. Single lipid regression models defined as:  $\text{Trait}_i = \text{Age}_i + \text{Smoking}_i + \text{Lipid}_j$ , where  $i = (1, \dots, 126)$  and  $j = (1, \dots, 1745)$ . Each y-axis represents  $-\log_{10}(p\text{-value})$  for each respective model, and each x-axis represent the retention time of the metabolite in minutes. Plasma lipid Benjamini-Hochberg false discovery corrected p-values  $< 0.05$  are circled in red. Body mass index (BMI), waist circumference (WC), tumor necrosis factor-alpha (TNF- $\alpha$ ), interferon gamma-induced protein-10 (IP-10), interleukin-6 (IL-6), monocyte chemoattractant protein-1 (MCP-1).

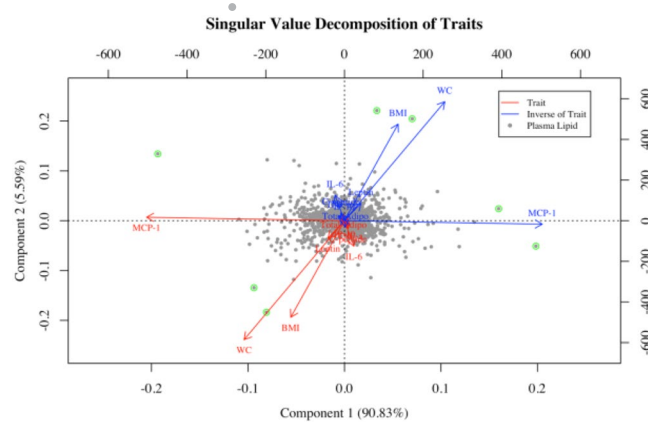


**Figure 2.** Venn diagrams displaying the relationship of lipids significantly associated between the traits body mass index (BMI), waist circumference (WC), leptin, and total adiponectin (Total Adipo). (A) Relationship of lipids with p-values < 0.05 for the traits. (B) Relationship of lipids with Benjamini-Hochberg p-values < 0.05.

and C-peptide concentrations. For the inflammatory cytokines IL-6, TNF- $\alpha$ , IP-10, and MCP-1, lipidomic profiles accounted for about 25% of variation in the samples. The 95% confidence intervals for most responses was roughly 10–20% for each respective trait (Table 2).

**Phospholipid classes and individual phospholipids are associated with obese traits.** Results of age and smoking adjusted single lipid regressions are outlined in Fig. 1 for the traits BMI, WC, leptin, total adiponectin, and C-peptide, and for each trait the primary id, estimated effects, and p-values of lipids with BH FDR p-values < 0.05 are displayed in Supplementary Tables 2, 3, 4, 5 and 6, respectively. Since no lipids were significantly associated with any response after 11.5 mins, the Manhattan plots were created to span from 0.5 (i.e., the start of MS/MS data collection) to 11.5 mins for clarity. In BMI regressions, 25 plasma lipids had BH FDR p-values < 0.05 and LPLs accounted for over 70% of lipids with Bonferroni p-values < 0.05 (Fig. 1A). WC was significantly associated with 35 lipids below the BH FDR (Fig. 1B). Leptin was significantly associated with 20 plasma lipids below the BH FDR (Fig. 1C). In total adiponectin regressions there were 22 lipids below the BH FDR (Fig. 1D). For WC, leptin, and total adiponectin, over 50% of lipids with Bonferroni p-values < 0.05 were LPLs (Fig. 1B,C and D, and Supplementary Tables 3, 4 and 5). C-peptide was significantly associated with 11 plasma lipids below the BH FDR (Fig. 1E). Neither IL-6, IP-10, or MCP-1 had any lipid significant below the BH FDR (data not shown).

Next, lipids significant across several traits from single lipid regression were inputted into the software Venny<sup>28</sup>. This was performed to illustrate the number of significant lipids with un-adjusted p-values and the number of significant lipids with BH FDR adjusted p-values, from each trait using Venn diagrams. Since Venny creates a Venn diagram for up to 4 traits, the anthropometrics BMI and WC, and the adipokines leptin and total adiponectin were selected for the illustration. Relationships between the traits and lipids with p-values < 0.05 are presented in Fig. 2A. There were numerous lipids with p-values < 0.05 associated with traits (Figs 1 and 2A), for instance: 180 lipids with BMI, 259 lipids with WC, 195 lipids with leptin, and 132 lipids with total adiponectin. In total, there were 59 lipids associated between the four traits (Fig. 2A). A Venn diagram of only lipids with BH FDR p-values < 0.05 for the traits are presented in Fig. 2B. There were 6 lipids with BH FDR p-values < 0.05 associated between the four traits (Fig. 2B), and the primary ids were: X1.17\_564.3289, X1.42\_566.3497, X1.21\_476.2768, X1.51\_592.3513, X1.03\_562.3132, and X7.15\_786.5626. A complete list of lipids significant in at least 2 or more single marker regression models are outlined in Table 3.



**Figure 3.** Singular value decomposition modeled on the 1,745 regression beta coefficients from plasma lipids in age and smoking regression for each respective parameter. (•) represent the plasma lipid loading score into vectors 1 and 2. Plasma lipids are highlighted if radius >0.15. Parameters are listed in red text with red arrows and the inverse of the parameters are listed in blue text with blue arrows. The score of MCP-1 in vector 1 is scaled by a factor 25 (i.e., original score >12,000). Body mass index (BMI), waist circumference (WC), tumor necrosis factor-alpha (TNF- $\alpha$ ), interferon gamma-induced protein-10 (IP-10), interleukin-6 (IL-6), monocyte chemoattractant protein-1 (MCP-1).

Parameter	$\sigma_{\mu}^2/\sigma_{\mu}^2 + \sigma_{\epsilon}^2$ [95% CI]
BMI	64% $\pm$ [38%, 84%]
WC	69% $\pm$ [46%, 85%]
Leptin	76% $\pm$ [56%, 90%]
Total Adiponectin	42% $\pm$ [21%, 66%]
C-peptide	48% $\pm$ [23%, 73%]
IL-6	23% $\pm$ [10%, 43%]
TNF- $\alpha$	28% $\pm$ [12%, 54%]
IP-10	23% $\pm$ [10%, 43%]
MCP-1	24% $\pm$ [11%, 45%]

**Table 2.** Percent of the inter-individual differences in response variables that can be attributed to lipidome profiles, posterior mean [95% Confidence region]. Body mass index (BMI), waist circumference (WC), tumor necrosis factor-alpha (TNF- $\alpha$ ), interferon gamma-induced protein-10 (IP-10), interleukin-6 (IL-6), monocyte chemoattractant protein-1 (MCP-1).

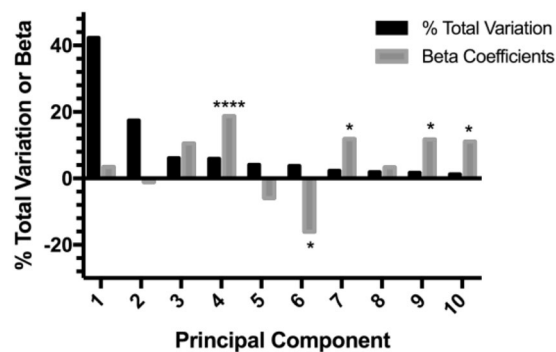
**Biplot.** A biplot analysis (Fig. 3) was derived from the 1,745 beta coefficients from each respective lipid in single marker regressions and the 9 traits (i.e., BMI, WC, leptin, total adiponectin, C-peptide, IL-6, IP-10, and MCP-1). Data were multidimensional, thus a biplot was used to extend the bivariate scatter plot. Biplots allow displaying information on samples and variables simultaneously. Here, a matrix with the lipid effects was built with samples in the rows (beta coefficients of the 1,745 samples) and 9 different lipid traits in the columns (i.e., BMI, WC, leptin, total adiponectin, C-peptide, IL-6, IP-10, and MCP-1). That matrix was decomposed in three using SVD, and in the biplot presented in Fig. 3, samples are displayed as points while traits are displayed as vectors. The vectors 1 and 2 from the SVD derived from the plasma lipid estimated effects, are represented with dots, and the loadings of the traits are represented with red arrows (in general, leftmost and bottom axes), and the inverse loadings of the traits are represented with blue arrows (in general, uppermost and right axes). Vectors 1 and 2 accounted for 96% of the variation in the lipidomics profiles. The orthogonal projection of the dots in the vectors are an approximation of the effect of each lipid for that trait. For instance, X12.43\_934.7918 contributes to higher MCP-1 concentrations, while X8.17\_830.6155 and 12.14\_850.7682 both contribute to lower MCP-1 concentrations. The lipids X8.24\_880.6145 and 3.52\_421.3076 contributed to higher BMI, WC, and MCP-1 concentrations, and the lipids X1.51\_592.3513 and X1.35\_669.3989 contributed to lower BMI, WC, and MCP-1 concentrations. In addition, the score of MCP-1 in the first vector was scaled by a factor 25 (i.e., original score >12,000) to display it within the range of the other traits in the figure. The traits IL-6, leptin, total adiponectin, C-peptide, and IP-10 did not significantly load into vectors 1 and 2.

**Different BMIs display different lipidomic profiles.** Next, BMI was regressed on Pc scores derived from the 1,745 lipid abundances (Fig. 4). The first 10 Pc scores were analyzed individually and models were adjusted for age and smoking. The 4<sup>th</sup> principal component accounted for 5.9% of Eigen value percent total variation and was positively associated (estimated effect = 18.7) with BMI (Bonferroni  $p < 0.05$ , data not shown). The 6<sup>th</sup> principal

Primary ID	Common name	Models where significant
X0.99_586.3136	LPC(20:5)	WC, leptin, Pc4
X1.03_562.3132	LPC(18:3)	BMI, WC, leptin, total adipo
X1.08_612.3280	LPC(22:6)	BMI, WC, leptin
X1.17_564.3289	LPC(18:2)	BMI, WC, leptin, total adipo, C-peptide
X1.21_476.2768	LPE(18:2)	BMI, WC, leptin, total adipo
X1.22_544.2648	LPS(20:4)	BMI, total adipo
X1.35_669.3989	Unknown	SVD
X1.42_566.3497	LPC(18:1)	BMI, WC, leptin, total adipo, C-peptide
X1.49_478.2927	LPE(18:1)	BMI, total adipo
X1.51_592.3513	LPC(20:2)	BMI, WC, leptin, total adipo, SVD
X1.59_554.3446	LPC(17:0)	BMI, WC
X1.91_594.3757	LPC(20:1)	WC, leptin
X2.65_255.2331	Palmitic acid	Pc4
X3.52_421.3076	TLTI	BMI, WC, SVD
X4.18_794.5050	Unknown	BMI, WC, total adipo
X4.37_881.5140	PI(16:0_22:6)	Pc4
X5.23_826.5592	PC(18:2_18:2)	BMI, WC, total adipo
X5.43_1068.6601	Unknown	BMI, WC, leptin
X5.90_810.5600	PC(O-16:0_20:5)	WC, leptin
X6.50_828.5729	PC(18:1_18:2)	BMI, WC, total adipo
X7.15_786.5626	PC(O-16:1_18:2)	BMI, WC, leptin, total adipo, C-peptide
X7.32_864.6056	PC(O-18:1_22:5)	WC, leptin
X7.47_788.5783	PC(O-16:0_18:2)	WC, total adipo
X8.17_830.6155	PC(18:0_18:2)	SVD
X8.23_950.6408	TLTI	Pc4
X8.24_880.6145	PC(20:1_20:4)	SVD
X8.44_716.5227	PE(16:0_18:1)	total adipo, C-peptide
X8.61_1038.473	TLTI	WC, C-peptide
X8.61_766.5363	PE(18:0_20:4)	WC, leptin, total adipo, C-peptide
X8.67_864.6048	PC(O-18:0_22:6)	Pc4
X8.90_880.6027	PC(18:0_22:5)	BMI, WC, leptin
X8.93_856.6036	PC(18:0_20:3)	BMI, WC, leptin, C-peptide
X9.01_1128.5388	TLTI	WC, leptin
X9.03_842.5446	TLTI	Pc4
X9.42_814.5927	PC(O-18:1_18:2)	BMI, WC
X9.45_1261.8069	TLTI	Pc4
X9.66_768.5494	PE(18:0_20:3)	leptin, C-peptide
X9.96_1034.5762	TLTI	Pc4
X10.11_1381.848	Unknown	WC, Pc4
X10.19_802.5695	TLTI	Pc4
X10.33_871.6904	SM(d19:1/24:1)	WC, leptin, Pc4
X11.20_1012.7670	TLTI	Pc4
X12.14_850.7682	TLTI	SVD
X12.43_934.7918	TLTI	SVD
X13.07_988.794	TLTI	Pc4
X13.10_920.7410	TLTI	Pc4
X13.11_919.7343	TLTI	Pc4
X14.31_881.7973	TLTI	Pc4
X14.33_869.3325	TLTI	Pc4
X14.35_1062.754	TLTI	Pc4
X14.38_880.7349	TLTI	Pc4

**Table 3.** Identification of plasma lipid biomarkers. The structural of identification of plasma lipids was performed based if  $m/z$ : was significant in two or more models from single marker regression analyses of responses, or had a radii  $\geq 0.15$  in singular value decomposition analysis, or was one of the top ten lipids positively and inversely associated with principal component 4. Plasma lipids meeting this criterion that were too low to identify (TLTI) are listed. Plasma lipid common names determined by lipid structure identification using Lipid Maps and Human Metabolome databases, and manual confirmation of mass spectra. The “\_”

between fatty acids in glycerophospholipid structures is used, since the position at the sn1 and sn2 position cannot be determined. Plasma lipids listed as unknown did not match a lipid library database search or exhibit a fragmentation pattern indicative of phospholipids in ms/ms analyses. Spectra of unknowns are listed in Supplementary Table 9. Body mass index (BMI), waist circumference (WC), total adiponection (total adipo), principal component 4 (Pc4) and singular value decomposition (SVD).



**Figure 4.** X-axis represent the principal component, and Y-axis represents the percent total variation of principal components or beta coefficient, respectively. Only the first 10 principal components (Pc) are displayed. Black bars represent the percentage of the sum of all possible eigenvalues; the Eigen value % of total variation. Gray bars represent the beta coefficients from regressions; model:  $(BMI)_i = Age_i + Smoking_i + Pc_k$  where  $i = (1, \dots, 126)$  and  $k = (1, \dots, 10)$ . \*p-value < 0.05 and \*\*\*\*p-value < 0.0001. Body mass index (BMI).

component was inversely associated with BMI, and principal components 7, 8, 9, and 10 were positively associated with BMI; estimated effects:  $-16.1, 11.9, 11.8,$  and  $11.0,$  respectively. In order to determine which lipids were driving the 4<sup>th</sup> principal component scores, the 4<sup>th</sup> principal component scores were regressed on the 1,745 lipids individually, and the top 10 beta coefficients positively and inversely associated with the 4<sup>th</sup> principal component scores, and the lipids respective Bonferroni corrected p-values are presented in Supplementary Table 8.

**Structural identification of primary ids.** The primary ids significantly associated with traits across several models were selected for structural identification. The primary ids, lipid common name, and list of models where each respective primary id was significant, are outlined in Table 3. PL head groups were determined as follows: PCs were confirmed by loss 60.02 Da (methyl formate), PEs were confirmed by presence of 140.01  $m/z$  (ethanolamine phosphate), PIs were confirmed by presence of 241.01  $m/z$  (inositol phosphate- $H_2O$ ), and PSs were confirmed by neutral loss of 87.03 Da (loss of serine). FA chains length and degree of saturation were determined as follows: C16:0 was confirmed by presence of 255.233  $m/z$ , C17:0 was confirmed by presence of 269.248  $m/z$ , C18:0 was confirmed by presence of 283.264  $m/z$ , C18:1 was confirmed by presence of 281.248  $m/z$ , C18:2 was confirmed by presence of 279.233  $m/z$ , C18:3 was confirmed by presence of 277.217  $m/z$ , C20:1 was confirmed by presence of 309.279  $m/z$ , C20:2 was confirmed by presence of 307.264  $m/z$ , C20:3 was confirmed by presence of 305.248  $m/z$ , C20:4 was confirmed by presence of 303.233  $m/z$ , C20:5 was confirmed by presence of 301.217  $m/z$ , C22:5 was confirmed by presence of 329.248  $m/z$ , and C22:6 was confirmed by presence of 327.233  $m/z$ .

Eleven LPLs were associated with traits from several lipid classes including LPC, LPE, and LPS. These LPLs identified included: LPC(17:0), (18:1), (18:2), (18:3), (20:1), (20:2), (20:5), and (22:6); LPE(18:1) and (18:2); and LPS(20:4). The primary id X1.35\_669.3989 fell within the retention time range of LPLs but did not exhibit a fragmentation pattern indicative of PLs (Supplemental Table 9) and was not available in lipid databases. The free FA palmitic acid was confirmed by 255.2331  $m/z$  and matching the retention time (2.65 min) with a pure standard. The primary ids X4.18\_794.5050, X5.43\_1068.6601, and X10.11\_1381.8480 did not have a fragment pattern indicative of PLs, although X5.43\_1068.6601 did produce the  $m/z$  1023.6949 indicative of formate (45.99 Da) loss.

Six PCs with ether linked chains were associated with several traits. These PCs with ether linked lipids included: PC(O-16:0\_20:5), PC(O-16:1\_18:2), PC(O-18:1\_22:5), PC(O-16:0\_18:2), PC(O-18:0\_22:6), and PC(O-18:1\_18:2). Ether linked chains in PCs were confirmed as follows: 466.3303  $m/z$  for O-16:0, 464.3147  $m/z$  for O-16:1, 494.3616  $m/z$  for O-18:0, and 492.3460  $m/z$  for O-18:1. In addition, there were several diacyl PLs associated with traits that contained FA chains  $\geq$  C18:0 and FA chains indicative of  $\omega$ -6s, including: PC(18:2\_18:2), PC(18:0\_18:2), PC(20:1\_20:4), PE(18:0\_20:4), PC(18:0\_22:5), PC(18:0\_20:3), and PE(18:0\_20:3). Since sphingolipids do not fragment well in negative mode, primary id X10.33\_871.6904 was determined to be SM(d19:1/24:1) as follows: by confirmation of methyl formate loss of parent ion and presence of daughter ion 811.6074  $m/z$  in negative mode, and the presence of 827.70  $m/z$  (the SM(d19:1/24:1)  $[M + H]^+$ ) at 10.33 min in positive mode, and the fragments 278.28  $m/z$  (indicative of SM d19:1 [long chain base- $H_2O$ ] $^+$ ) and 336.37  $m/z$  (indicative of SM with 24:1 FA in positive mode). A complete table of the primary ids, common name, molecular adducts, theoretical  $m/z$ , and mass error (<0.05 for all identified lipids) are presented in Supplementary Table 10.

Common name <sup>a</sup>	Palmitic <sup>b</sup> C16:0	Stearic <sup>b</sup> C18:0	Oleic <sup>b</sup> C18:1	Elaidic <sup>b</sup> C18:1	Eicosenoic <sup>b</sup> C20:1	Linoleic <sup>b</sup> C18:2	Linoelaidic <sup>b</sup> C18:2	ALA <sup>b</sup> C18:3	Eicosadienoic <sup>b</sup> C20:2	DGLA <sup>b</sup> C20:3	ARA <sup>b</sup> C20:4	EPA <sup>b</sup> C20:5	DPA <sup>b</sup> ω-3 <sup>b</sup> C22:5	DPAω-6 <sup>b</sup> C22:5	DHA <sup>b</sup> C22:6	Nervonic <sup>b</sup> C24:1
LPC(20:5)	—	—	—	—	—	—	—	—	—	—	—	<b>1.02</b>	—	—	—	—
LPC(18:3)	—	—	—	—	—	—	—	<b>0.24</b>	—	—	—	—	—	—	—	—
LPC(22:6)	—	—	—	—	—	—	—	—	—	—	—	—	—	—	<b>0.96</b>	—
LPC(18:2)	—	—	—	—	—	<b>0.90</b>	0.01	—	—	—	—	—	—	—	—	—
LPE(18:2)	—	—	—	—	—	<b>1.33</b>	0.03	—	—	—	—	—	—	—	—	—
LPS(20:4)	—	—	—	—	—	—	—	—	—	—	<b>-0.85</b>	—	—	—	—	—
LPC(18:1)	—	—	<b>0.36</b>	0.02	—	—	—	—	—	—	—	—	—	—	—	—
LPE(18:1)	—	—	<b>0.48</b>	0.09	—	—	—	—	—	—	—	—	—	—	—	—
LPC(20:2)	—	—	—	—	—	—	—	—	<b>0.28</b>	—	—	—	—	—	—	—
LPC(17:0)	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
LPC(20:1)	—	—	—	—	0.11	—	—	—	—	—	—	—	—	—	—	—
Palmitic acid	1.24	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
PI(16:0_22:6)	<b>0.09</b>	—	—	—	—	—	—	—	—	—	—	—	—	—	<b>0.97</b>	—
PC(18:2_18:2)	—	—	—	—	—	<b>2.61</b>	<b>-0.29</b>	—	—	—	—	—	—	—	—	—
PC(O-16:0_20:5)	—	—	—	—	—	—	—	—	—	—	—	<b>0.99</b>	—	—	—	—
PC(18:1_18:2)	—	—	0.15	-0.17	—	<b>1.05</b>	<b>-0.14</b>	—	—	—	—	—	—	—	—	—
PC(O-16:1_18:2)	—	—	—	—	—	<b>0.93</b>	-0.12	—	—	—	—	—	—	—	—	—
PC(O-18:1_22:5)	—	—	—	—	—	—	—	—	—	—	—	<b>0.70</b>	0.21	—	—	—
PC(O-16:0_18:2)	—	—	—	—	—	<b>1.14</b>	-0.13	—	—	—	—	—	—	—	—	—
PC(18:0_18:2)	—	0.54	—	—	—	<b>1.04</b>	<b>-0.33</b>	—	—	—	—	—	—	—	—	—
PC(20:1_20:4)	—	—	—	—	0.31	—	—	—	—	—	<b>-0.11</b>	—	—	—	—	—
PE(16:0_18:1)	<b>2.53</b>	—	0.33	-0.21	—	—	—	—	—	—	—	—	—	—	—	—
PE(18:0_20:4)	—	<b>0.77</b>	—	—	—	—	—	—	—	—	0.12	—	—	—	—	—
PC(O-18:0_22:6)	—	—	—	—	—	—	—	—	—	—	—	—	—	—	<b>0.74</b>	—
PC(18:0_22:5)	—	<b>1.86</b>	—	—	—	—	—	—	—	—	—	-0.28	<b>0.63</b>	—	—	—
PC(18:0_20:3)	—	<b>0.76</b>	—	—	—	—	—	—	—	<b>0.68</b>	—	—	—	—	—	—
PC(O-18:1_18:2)	—	—	—	—	—	0.65	-0.11	—	—	—	—	—	—	—	—	—
PE(18:0_20:3)	—	<b>1.62</b>	—	—	—	—	—	—	—	<b>0.82</b>	—	—	—	—	—	—
SM(d19:1/24:1)	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	<b>0.81</b>

**Table 4.** Relationship between untargeted PL FA chains and the % of geometric and positional FA isomers in plasma PL. Each plasma PL was regressed on the % of each FA isomer in the plasma PL, and models were adjusted for BMI. This was performed to determine the relationship between the abundance of each plasma PL significantly associated with the responses in our study and the % of FA isomers in plasma PL from our previous study. For instance, LPC(18:2) and LPE(18:2) were significantly associated with the % of linoleic acid (LA, ω-6, C18:2<sup>Δ9Z,12Z</sup>) in plasma PL, whereas LPC(18:2) and LPE(18:2) were not correlated with the % of trans isomer linoelaidic acid (ω-6, C18:2<sup>Δ9E,12E</sup>) in plasma PL. Numeric values represent the estimated effects (beta coefficients) and are bolded if  $p < 0.05$ . The plasma analyzed from both data sets were collected at the same time point from the same patients. Models were defined as:  $\log(\text{plasma PL}) = \log(\% \text{ FA isomer in plasma PL}) + \text{BMI}$ . <sup>a</sup>The significant PL associated with the responses that were structurally characterized in our study. The experimental methodology employed was the UPLC-ESI-MS<sup>E</sup> analysis of crude lipid extracts outlined in this manuscript. <sup>b</sup>The geometric and positional FA isomers that were targeted in our previous study (Pickens *et al.* PLEFA. 2015). The experimental methodology employed was the FAME analysis of isolated plasma PL by GC-FID. The specific isomers investigated were: oleic acid (ω-9C18:1<sup>Δ9Z</sup>), elaidic acid (ω-9C18:1<sup>Δ9E</sup>), eicosanoic acid (ω-9, C20:1<sup>Δ11Z</sup>), linoleic acid (LA, ω-6, C18:2<sup>Δ9Z,12Z</sup>), linoelaidic acid (ω-6, C18:2<sup>Δ9E,12E</sup>), dihomogamma-linolenic acid (DGLA, ω-6, C20:3<sup>Δ8Z,11Z,14Z</sup>), arachidonic acid (ARA, ω-6, C20:4<sup>Δ5Z,8Z,11Z,14Z</sup>), docosapentaenoic acid ω-6 (DPA ω-6, C22:5<sup>Δ4Z,7Z,10Z,13Z,16Z</sup>), alpha-linolenic acid (ALA, ω-3, C18:3<sup>Δ9Z,12Z,15Z</sup>), eicosapentaenoic acid (EPA, ω-3, C20:5<sup>Δ5Z,8Z,11Z,14Z,17Z</sup>), docosapentaenoic acid ω-3 (DPA ω-3, C22:5<sup>Δ7Z,10Z,13Z,16Z,19Z</sup>), docosahexaenoic acid (DHA, ω-3, C20:5<sup>Δ4Z,7Z,10Z,13Z,16Z,19Z</sup>), and nervonic acid (ω-9, C24:1<sup>Δ15Z</sup>).

**Plasma PL FA chains are associated to FAs with specific double bond position and geometry in plasma PL.** The untargeted analysis employed in our study does not determine the double bond geometry (i.e., cis/trans) or position (i.e., between carbons 6–7, 10–11, etc) on the unsaturated FA chains. The significant plasma PL structurally characterized in our current study (Table 3) were regressed on the % total plasma PL FA levels from our previous study<sup>26</sup>, since the plasma analyzed in our current and previous study were collected from the same patients at the same time point. Thus, researchers can infer the ω-family, and double bond geometry and position of the FA chains from our untargeted results. For instance, each one unit change in the % of LA



( $\omega$ -6, C18:2 $^{\Delta 9Z,12Z}$ ) in plasma PL was associated with an average one unit change in LPC(18:2) and LPE(18:2) levels (Table 4). These 18:2 containing LPLs were not associated with the % of the trans-isomer linoleic acid ( $\omega$ -6, C18:2 $^{\Delta 9E,12E}$ ) in plasma PL as outlined in Table 4. For each one unit change in the % of eicosapentaenoic acid (EPA,  $\omega$ -3, C20:5 $^{\Delta 5Z,8Z,11Z,14Z,17Z}$ ) in plasma PL, the average change in mean LPC(20:5) and PC(O-16:0\_20:5) levels were roughly one unit. Similarly, LPC(22:6), PI(16:0\_22:6), and PC(O-18:0\_22:6) were highly associated with the % of docosahexaenoic acid (DHA,  $\omega$ -3, C22:6 $^{\Delta 4Z,7Z,10Z,13Z,16Z,19Z}$ ). PC(O-18:1\_22:5) was highly associated with the % of docosapentaenoic acid  $\omega$ -3 (DPA  $\omega$ -3, C22:5 $^{\Delta 7Z,10Z,13Z,16Z,19Z}$ ) in plasma PL, and PC(O-18:1\_22:5) and was not associated with % of DPA  $\omega$ -6 (C22:5 $^{\Delta 4Z,7Z,10Z,13Z,16Z}$ ) in plasma PL (Table 4). On the other hand, PC(18:0\_22:5) was highly associated with the % of DPA  $\omega$ -6 (C22:5 $^{\Delta 4Z,7Z,10Z,13Z,16Z}$ ) in plasma PL and was not associated with % of DPA  $\omega$ -3 (C22:5 $^{\Delta 7Z,10Z,13Z,16Z,19Z}$ ) in plasma PL. Each model in Table 4 was adjusted for BMI to account for the underlying association between BMI, and the plasma PL and % of FA isomer. The results presented in Table 4 were similar to the spearman correlations between each PL and the % of each FA isomer (Supplementary Table 11).

## Discussion

This study characterized plasma PL profiles related to obesity and obesity-associated serum adipokines, cytokines, and the glycemic marker C-peptide. Specifically, we report that several classes of LPLs (i.e., LPC, LPE, and LPS) are inversely associated with BMI, WC, leptin, and C-peptide. Plasma ether linked PCs were also significantly decreased in obesity, including 16 and 18 carbon saturated and unsaturated ethers. On the other hand, total adiponectin, an anti-inflammatory adipokine inversely associated with obesity, was positively associated with LPLs and C18:2 containing PLs. Overall we report that C18:2 is lower in several lipid classes (i.e., LPLs, diacyl PCs, and ether linked PCs) in obese individuals. The significant PL containing C18:2 chains were highly associated with the % of LA ( $\omega$ -6, C18:2 $^{\Delta 9Z,12Z}$ ) in plasma PL. We previously reported in this study population, that median plasma PL LA ( $\omega$ -6, C18:2 $^{\Delta 9Z,12Z}$ ) levels are roughly 3% lower in our obese patients compared to lean<sup>26</sup>. Results of our current study indicate that plasma PL FA differences occur in specific PL species, and since obesity is associated with altered lipid metabolism, our results provide insight into metabolic pathways that may be of interest for future research.

Associations between LPLs and overweight/obesity are well described, in particular the decrease in LPC(18:2)<sup>15,16</sup>. After weight loss in obese individuals there is an increase in LPCs, in particular C18:2, suggesting the inverse association between LPLs and obesity could be related to diet or weight loss<sup>29,30</sup>. Interestingly, the seminal paper by Pietilainen *et al.* reported in monozygotic twins that obesity was associated with increases serum LPCs, including LPC(18:2), and decreases in serum ether lipids, which suggests that changes in LPL profiles are likely associated with genetic factors related with obesity rather than environmental and lifestyle factors<sup>31</sup>. In experimental models with high fat diets (HFD), there are rapid decreases in plasma LPC(18:1), (18:2), (20:1), and (20:5)<sup>25</sup>. Del Bas *et al.* recently reported, in a randomized control trial, that  $\omega$ -3 PUFA supplementation only increases  $\omega$ -3s in LPLs and lowers  $\omega$ -6s in LPLs, but has no effect on overall LPL concentrations in overweight individuals<sup>18</sup>. Therefore, the observed decrease in plasma LPLs may suggest that our obese patients were consuming high-fat diets prior to study recruitment, and we did not assess dietary status in our study. C-peptide, a marker of insulin production, was also inversely associated with LPLs including LPC(18:2). Higher C-peptide indicates a higher production of insulin<sup>32</sup>. Our patient exclusion criteria excluded individuals with type-1 or type-2 diabetes, and those being treated for insulin resistance. Lower levels of plasma LPCs are associated with obese individuals<sup>25</sup>, obese individuals with type-2 diabetes<sup>25</sup>, and have been associated with insulin resistance independent of obesity<sup>33</sup>. Interestingly, LPCs activate PPAR $\delta$  to reduce skeletal muscle inflammation and ER stress<sup>34</sup>, and accumulations of fat in skeletal muscle<sup>35</sup> and liver<sup>36</sup> are associated with decreased insulin sensitivity.

Studies have shown that fat accumulation in the liver occurs independent of obesity and is more associated with insulin resistance<sup>37</sup>. Insulin resistance is associated with increased hepatic fat accumulation, which alters hepatic lipoprotein secretion such as increasing secretion of low-density and very low-density lipoproteins, and decreasing secretion of high-density lipoproteins [reviewed in detail ref. 38]. In fact, high-density lipoprotein particles are rich in LPLs and obesity is associated with decreases in high-density lipoproteins<sup>39</sup>. We did not assess lipoprotein levels or screen our patients for fatty liver disorders, however, LPC(18:2) and ether lipid concentrations are inversely associated with liver fat percentages and are predictive of NAFLD<sup>40</sup>. Furthermore, the enzyme LCAT is largely responsible for generating plasma LPLs, and LCAT is present in low-density lipoproteins, which are increased in individuals with obesity, insulin resistance, and fatty liver disorders [reviewed in detail refs 8 and 38]. Sansone *et al.* reported that intra-patient concentrations of LA (C18:2 $^{\Delta 9Z,12Z}$ ) were higher in plasma cholesterol esters and lower in RBC PLs, suggesting that LA is being shunted from the PL to neutral lipids such as cholesterol esters<sup>41</sup>. It is likely our observed decrease in plasma LPLs could be associated with a combination of factors including dietary intake, altered lipoprotein metabolism, shunting of specific FAs from PL to cholesterol esters, insulin resistance, and fatty liver.

Overall most of the significant plasma PL FAs differences in our results were PUFAs, in particular C18:2 which decreased in LPLs, ether linked PCs, and diacyl PCs. The specific double bond location and geometry of unsaturated FA chains are associated with metabolic diseases<sup>42</sup>. Correlations between the significant PL containing 18:2 chains and FA data from our plasma PL FAME analysis indicated that LA (C18:2 $^{\Delta 9Z,12Z}$ ) is likely decreasing in these specific LPLs, ether lipids, and diacyl PCs (Table 4). A recent meta-analysis reported that long chain PUFA profiles, in particular  $\omega$ -6s, are dysregulated in overweight and obese individuals<sup>6</sup>. In obesity, decreases in LA are associated with increases in LA-derived DGLA by the enzyme delta-6-desaturase (D6D). Taking into account the results from Sansone *et al.*<sup>41</sup>, LA decreases in obesity are likely due to a combination of LA shunting to neutral lipid species and increased metabolism of LA to long chain  $\omega$ -6 PUFAs through D6D. We have previously reported in our study population that PL DGLA levels and D6D enzyme activity estimates (EAE), D6D EAE = ratio of DGLA/LA, are increased in obese individuals<sup>26</sup>. In our current study we report C20:3 (i.e.,

DGLA) is increased specifically in PC(18:0\_20:3) and PE(18:0\_20:3), and these PL contains 20:3 chains are highly associated with the % of DGLA ( $\omega$ -6, C20:3 $\Delta^{8Z,11Z,14Z}$ ) in plasma PL. In addition, we have also reported in our study population that C-peptide concentrations are inversely associated with plasma PL delta-5-desaturase (D5D) EAAs (i.e., the ratio of ARA/DGLA) even after adjusting for obesity (i.e., BMI and WC)<sup>27</sup>. In our current study, for each one unit change in PE(18:0\_20:3), the average change in mean C-peptide concentrations was roughly 2.13 units (Supplementary Table 6). It is likely that PLs such as PE(18:0\_20:3) and PC(18:0\_20:3) are important lipids contributing to the increases in PL DGLA levels that is observed in numerous obesity related studies (reviewed in detail refs 2 and 6).

The purpose of our study was to identify specific plasma PL species with FA differences associated with traits (i.e., BMI, WC, serum C-peptide), and assess the contribution of the entire plasma lipidome on the variation of these traits, in a cross-sectional study ( $n = 126$ , >96% Caucasian males, ages 48–65). We recognize the generalizability of our study and these results should be verified in larger, more diverse populations. We acknowledge not assessing diet, % energy from fat, and supplement usage are limitations to our study. Although our patient exclusion criteria excluded those with diabetes and those being treated for insulin resistance, it is possible some of our obese patients were insulin resistant. Our untargeted analysis identified FA chains esterified in plasma PL, and the double bond position and geometry could not be determined using our method. Therefore, we correlated patient's PL FAME levels with their levels of significant untargeted lipids, so the double bond geometry and position can be inferred. In conclusion, the whole plasma lipidomic profiles accounted for over 60% of the inter-individual variation in obesity. Our results are in agreement that obesity is associated with decreases in LPLs<sup>15–18, 25</sup>, ether linked lipids<sup>31, 43–45</sup>, and an overall decrease in plasma PL LA<sup>2, 6, 41</sup>. We report plasma PL LA likely decreases in the following lipid species: the LPLs LPC(18:2) and LPE(18:2), the diacyl PCs PC(18:0\_18:2), PC(18:1\_18:2), and PC(18:2\_18:2), and the ether linked PCs PC(O-16:0\_18:2), PC(O-16:1\_18:2), and PC(O-18:1\_18:2). Future research should quantify the levels of these specific lipids using synthetic standards and determine their feasibility for use in clinical diagnostics screenings for obesity-associated co-morbidities.

## Materials and Methods

**Study population.** Data are described elsewhere<sup>26, 46–50</sup>. In brief, healthy and asymptomatic male subjects ( $n = 126$ , >96% Caucasian) 48 to 65 years of age were enrolled between 2009 and 2011 in a cross-sectional study. Data are comprised of clinical metadata on individuals' co-morbidities, family history, and use of medications. Individuals with the following medical conditions were excluded: (1) cancer in the previous two years, (2) surgery in the previous two years, (3) type-1 and type-2 diabetes, and (4) autoimmune diseases. Immediately after enrollment, trained staff collected anthropometric measurements and venous blood of study participants. In brief, body mass index (BMI) was assessed by recording patient's height using a stadiometer and weight using digital platform scale. The plasma fraction was separated from whole blood by centrifugation and stored at  $-80^{\circ}\text{C}$  until time of analysis. Smoking status was assessed as "ever smoked" or "never smoked". A previous complete description of this study can be found elsewhere<sup>46</sup>. All participants provided written informed consent, and all experiments were performed in accordance with relevant guidelines and regulations. The study was approved by the Biomedical and Health Institutional Review Board of Michigan State University (IRB# 08-786).

**Plasma preparation.** A detailed list of chemicals, solvents, and internal standards used in plasma lipid extractions are presented in Supplementary Table 1. Plasma lipid extraction was performed following a Waters application note<sup>51</sup>, but modified as specified. A quality control (QC) sample was made by pooling 10  $\mu\text{L}$  plasma from all 126 samples, and this QC sample was extracted and processed alongside individual samples. Plasma samples were thawed on ice. A 50  $\mu\text{L}$  aliquot of each plasma was transferred to a labeled microcentrifuge tube containing 200  $\mu\text{L}$  of the extraction solution (2:1 v/v, 100  $\mu\text{g}/\mu\text{L}$  BHT, 5  $\text{ng}/\mu\text{L}$  PLs and 0.2  $\text{ng}/\mu\text{L}$  Cer). These mixtures were then incubated on ice for 15 minutes, then vortexed for 30 s. After vortexing, protein crash was performed by centrifugation in an accuSpin<sup>TM</sup> Micro R centrifuge at  $13,000 \times g$  for 5 min at  $4^{\circ}\text{C}$  (Fisher Scientific, Waltham, MA). Following centrifugation, each lower organic phase was collected and transferred to new labeled microcentrifuge tube, evaporated to dryness in a Savant SpeedVac (ThermoQuest, Holbrook, NY) for 3 h, with no applied heating, then stored under high-purity nitrogen at  $4^{\circ}\text{C}$  for no longer than 1 day. Dried residues were re-dissolved in 500  $\mu\text{L}$  HPLC-grade isopropanol:acetonitrile:water (2:1:1 v/v/v), then vortexed gently. Re-dissolved extracts were again centrifuged  $13,000 \times g$  for 5 min at  $4^{\circ}\text{C}$  (Fisher Scientific). After centrifugation, 100  $\mu\text{L}$  drawn from the top of the supernatant was transferred to an amber autosampler vial with glass insert, and the vials were purged with high-purity nitrogen and sealed.

**Lipidomic profiling by UPLC-MS/MS analyses.** Mass spectrometric analysis was accomplished following Waters application note<sup>51</sup>, but modified as specified. In brief, reverse phase ultra-performance liquid chromatography (UPLC)-ESI<sup>neg</sup>-MS/MS was performed on a Waters ACQUITY UPLC (Waters, Milford, MA) coupled to a Waters Xevo G2-XS quadrupole time-of-flight mass spectrometer (Waters). Chromatographic separation was performed with an ACQUITY UPLC CSH C18 1.7  $\mu\text{m}$  2.1  $\times$  100 mm column (Waters) held at  $55^{\circ}\text{C}$ . The autosampler (Waters) temperature was held at  $10^{\circ}\text{C}$ . The UPLC method was shortened to 15 min and the mobile phases and gradients, along with mass spectrometer settings are outlined in Supplementary Table 1. Data acquisition was performed using MS<sup>E</sup> in continuum mode with leucine enkephalin as lock mass for mass correction. The QC sample was injected after every 10th sample throughout the entire analysis.

**Data alignment and processing of plasma lipids.** Results from UPLC-MS<sup>E</sup> analyses were imported into Progenesis QI v2.0 (Nonlinear Dynamics, Durham, NC). Peaks were aligned using one of the QC samples as references. Peak picking was performed under normal conditions, collecting signals from 4802 ions. The Progenesis

QI generated data set was subjected to relative mass defect filtering to narrow ions to those of potential lipids<sup>52</sup>, and ions with relative mass defects <350 and >950 parts-per-million (ppm) were excluded from the data set. The mass defect filtered data set was then imported into R for multivariate statistical analyses. Since large -omic data sets often have 10–20% missing values<sup>53,54</sup>, 1126 ions with >20% missing values were excluded from the data set, the 10 IS were split from the data set, and peak areas for the remaining 1,745 ions were mean imputed. Next, due to a time of injection affect (details are presented in Supplementary Table 1), the data matrix peak areas were normalized to the IS PC(8:0/8:0) signal. After IS normalization of data matrix, the coefficient of variation between QC injections was <5% for the most significant lipids outlined in this manuscript.

**Plasma PL FAME data.** Plasma was extracted and PL were isolated as previously described<sup>55, 56</sup>. Experimental methods and PL FAME acquisition are outlined in detail as previously described<sup>26</sup>.

**Statistical analyses.** Our study was powered (0.8) to detect differences in plasma FAs, as previously described<sup>49</sup>. Statistical analysis of the data was performed using R v3.2.2<sup>57</sup>, and all R-code used in these analyses are included in the Supplementary Table 12 and publicly available at <https://github.com/AustinPickens/Untargeted-Lipidomics>. The responses used in the statistical analyses were the traits: BMI, waist circumference (WC), log transformed serum leptin, log transformed serum total adiponectin, log transformed serum c-peptide, log transformed serum interleukin-6 (IL-6), log transformed serum tumor necrosis factor-alpha (TNF- $\alpha$ ), log transformed serum interferon gamma-induced protein-10 (IP-10), and serum monocyte chemoattractant protein-1 (MCP-1). The analysis consisted of two parts: (1) Single lipid regressions on each trait associating the abundance of the lipid in plasma with the responses, to identify individual plasma lipids significantly associated with each response; (2) High dimensional analyses including principal components (Pc), singular value decomposition (SVD), and a model regressing the responses in all lipids in plasma, to identify whether the inter-individual differences can be attributed to the lipidomic profile. P-value correction was performed according to Benjamini-Hochberg false discovery rate (BH FDR) and Bonferroni<sup>58, 59</sup>.

**Single Lipid Regressions.** Response variables include: BMI, WC, log transformed leptin, log transformed total adiponectin, log transformed c-peptide, log transformed IL-6, log transformed TNF- $\alpha$ , log transformed IP-10, and MCP-1. We herein refer to these responses as traits, and these traits consist of phenotype  $y_i (i = 1, \dots, 126)$  indexed by individuals ( $i$ ) and the set of predictors coming from the lipidomic data as well as clinical covariates including age of the participant ( $a_i$ )  $56.9 \pm 4.7$  (mean  $\pm$  s.d.), the status of smoking behavior ( $s_i$ ) 25.40% ever smoked, 57.14% never smoked, 17.46% smoking unknown. In each regression only one of the lipids was included from the pareto standardized lipidomic data ( $l_{ij}$ ). The statistical model was adjusted for each lipid,  $j = 1, \dots, 1745$  as follows in Eq. (1):

$$y_i = \beta_{0j} + a_i\beta_{1j} + s_i\beta_{2j} + l_{ij}\beta_{3j} + \varepsilon_{ij}, \quad (1)$$

let  $\beta_{0j}$  be a general intercept, and  $\beta_{1j}$ ,  $\beta_{2j}$  and  $\beta_{3j}$  the corresponding regression coefficients and  $\varepsilon_{ij}$  a random residual following a IID normal distribution centered at zero. A total of 1,745 regressions, one per lipid, were adjusted in Eq. (1).

**High dimensional analyses.** The lipidomic data is a high-dimension dataset with a higher number of parameters to estimate ( $p = 1745$ ) than observations ( $n = 126$ ); for example  $p \gg n$ . Thus, the following analysis were performed: A) Pc derived from lipidomic data abundances, B) SVD derived from the lipids estimated effects obtained in the single marker regressions, and C) whole lipidome regressions, as described in refs 60–62.

- A) The transposed cross product (i.e.,  $X'X$ ) of the  $n \times m$  ( $n = 1, \dots, 126$ ;  $m = 1, \dots, 1745$ ) matrix of lipidomic abundances was used to produce an  $n \times n$  matrix of distances. Next, an Eigen value decomposition of the  $n \times n$  matrix was employed to generate Eigen values and Pc scores for each of the Pcs-derived from lipid abundances. There were 126 Pcs generated and we evaluated only the first ten Pcs. The effects of each Pc were evaluated by regressing BMI against Pc scores individually. The statistical model was adjusted for each Pc,  $k = (1, \dots, 10)$  as follows in Eq. (2):

$$BMI_i = \beta_{0k} + a_i\beta_{1k} + s_i\beta_{2k} + Pc_{ik}\beta_{3k} + \varepsilon_{ik} \quad (2)$$

let BMI be the phenotype ( $i = 1, \dots, 126$ ) indexed by individuals ( $i$ ) and the set of predictors coming from the Pc scores derived from lipid abundances and let  $\beta_{0j}$  be a general intercept, and  $\beta_{1k}$ ,  $\beta_{2k}$ , and  $\beta_{3k}$  the corresponding regression coefficients and  $\varepsilon_{ik}$  a random residual following a IID normal distribution centered at zero. A total of 10 regressions, one per Pc, were adjusted in Eq. (2). Finally, since Pc4 was highly associated with BMI, Pc4 scores were regressed on the individual lipids, one at a time to identify the lipids driving Pc4. The statistical model adjusted for each lipid,  $j$  as follows in Eq. (3):

$$Pc4_i = \beta_{0j} + l_{ij}\beta_{1j} + \varepsilon_{ij} \quad (3)$$

in this analysis, Pc4 loadings are derived from lipid abundances and the set of predictors coming from lipidomic data and let  $\beta_{0j}$  be a general intercept and  $\beta_{1j}$  the corresponding regression coefficients and  $\varepsilon_{ij}$  a random residual following a IID normal distribution centered at zero. A total of 1,745 regressions, one per lipid, were adjusted in Eq. (3).

- B) Data were multidimensional, thus, a biplot was used to extend the bivariate scatter plot. Biplots allow displaying information on samples and variables simultaneously. Here, a matrix with the lipid effects

was constructed with lipids in the rows (beta coefficients of the 1,745 lipids) and 9 different traits in the columns (i.e., BMI, WC, leptin, total adiponectin, C-peptide, IL-6, IP-10, and MCP-1). That matrix was decomposed in three using SVD, and the biplot results, displayed in Fig. 3, are the first two vectors that spanned the rows based on the column space of the SVD<sup>63</sup>.

- C) Finally, the lipidome was fully accounted for in a whole genome regression analysis<sup>61,62</sup>. The probabilistic model assumed used in each of the responses was as follows in Eq. (4),

$$y_i = \beta_0 + a\beta_1 + s\beta_2 + \sum_{j=1}^{j=1745} x_{ij}\alpha_j + \varepsilon_i, \quad (4)$$

where  $u_i = \sum_{j=1}^{j=1745} x_{ij}\alpha_j$  represents the total effect of the lipids on trait  $y_i$ ,  $x_{ij}$  is the lipidomic abundance for participant  $i$  and lipid  $j$ , and  $\alpha_j$  is the estimated effect. In Eq. (4), assume that the responses consist of the traits indexed by individuals ( $i$ ) and the set of predictors coming from the lipidomic data.  $u_i$  follows a normal distribution centered zero with variance covariance  $u_i \sim N(0, G\sigma_u^2)$ , and  $G$  is an  $n \times n$  matrix of distances to measure similarities between participants with respect to their lipid profiles, as previously described<sup>64</sup>. These analyses were conducted using the R-package BGLR<sup>65</sup>, using a long Markov Chain of 200,000 iterations and 50,000 samples were discarded for burn in. Inference was done based on one of every 5 samples of the last 150,000. High dimensional regressions were adjusted to the lipidomic data as describe before for other omics<sup>60,66</sup>.

**Structural Identification of Plasma Lipids.** Lipids in this study are presented in the form of Xretention time<sub>mass-to-charge ratio</sub> (i.e., X1.17\_564.3289 [1.17 mins, 564.3298 m/z]). The structural of identification of plasma lipids was performed if the integrated peak areas for the ion: were significant in two or more models from single marker regression analyses of responses, or had radii  $\geq 0.15$  in SVD analysis (Fig. 3, and Supplementary Table 7), or was one of the top ten lipids positively and inversely associated with Pc4 (Supplementary Table 8), and the  $m/z$  exhibited a fragmentation pattern indicative of a free FA, LPL, or PL species. Since the objective of this study was to identify free FA and PL differences related to obesity and obesity-associated adipokines, significant lipid ions that did not have fragment patterns matching free FA and PLs are listed as unknown (Table 3), and their spectra are listed in Supplementary Table 9. Plasma lipids meeting our criteria for identification that are not listed in Table 3 were too low in abundance to obtain a useful MS/MS spectrum and are listed as too low to identify (TLTI). Lipid structure identification was confirmed using Lipid Maps (<http://lipidmaps.org>) and Human Metabolome database (<http://www.hmdb.ca>), and manual evaluation of MS/MS spectra. The “\_” denotation is used between FAs in glycerolPL structures, since the FA position (i.e., either sn1 and sn2) were not determined.

## References

- Emanuela, F. *et al.* Inflammation as a Link between Obesity and Metabolic Syndrome. *J Nutr Metab* **2012**, 476380, doi:10.1155/2012/476380 (2012).
- Kroger, J. & Schulze, M. B. Recent insights into the relation of Delta5 desaturase and Delta6 desaturase activity to the development of type 2 diabetes. *Curr Opin Lipidol* **23**, 4–10, doi:10.1097/MOL.0b013e32834d2dc5 (2012).
- Shoelson, S. E., Lee, J. & Goldfine, A. B. Inflammation and insulin resistance. *J Clin Invest* **116**, 1793–1801, doi:10.1172/JCI29069 (2006).
- van Diepen, J. A., Berbee, J. F., Havekes, L. M. & Rensen, P. C. Interactions between inflammation and lipid metabolism: relevance for efficacy of anti-inflammatory drugs in the treatment of atherosclerosis. *Atherosclerosis* **228**, 306–315, doi:10.1016/j.atherosclerosis.2013.02.028 (2013).
- Quehenberger, O. *et al.* Lipidomics reveals a remarkable diversity of lipids in human plasma. *J Lipid Res* **51**, 3299–3305, doi:10.1194/jlr.M009449 (2010).
- Fekete, K. *et al.* Long-chain polyunsaturated fatty acid status in obesity: a systematic review and meta-analysis. *Obes Rev* **16**, 488–497, doi:10.1111/obr.12280 (2015).
- Pickens, C. A., Sordillo, L. M., Zhang, C. & Fenton, J. I. Obesity is positively associated with arachidonic acid-derived 5- and 11-hydroxyeicosatetraenoic acid (HETE). *Metabolism* **70**, 177–191, doi:10.1016/j.metabol.2017.01.034 (2017).
- Grzelczyk, A. & Gendaszewska-Darmach, E. Novel bioactive glycerol-based lysophospholipids: new data—new insight into their function. *Biochimie* **95**, 667–679, doi:10.1016/j.biochi.2012.10.009 (2013).
- Sutherland, W. H., Temple, W. A., Nye, E. R. & Herbison, P. G. Lecithin:cholesterol acyltransferase activity, plasma and lipoprotein lipids and obesity in men and women. *Atherosclerosis* **34**, 319–327 (1979).
- Garces, F. *et al.* High plasma phospholipase A2 activity, inflammation markers, and LDL alterations in obesity with or without type 2 diabetes. *Obesity (Silver Spring)* **18**, 2023–2029, doi:10.1038/oby.2010.9 (2010).
- Xu, Y. Sphingosylphosphorylcholine and lysophosphatidylcholine: G protein-coupled receptors and receptor-mediated signal transduction. *Biochim Biophys Acta* **1582**, 81–88 (2002).
- Stock, C., Schilling, T., Schwab, A. & Eder, C. Lysophosphatidylcholine stimulates IL-1beta release from microglia via a P2X7 receptor-independent mechanism. *J Immunol* **177**, 8560–8568 (2006).
- Nishi, E. *et al.* Lysophosphatidylcholine enhances cytokine-induced interferon gamma expression in human T lymphocytes. *Circ Res* **83**, 508–515 (1998).
- Riederer, M. *et al.* Acyl chain-dependent effect of lysophosphatidylcholine on endothelial prostacyclin production. *J Lipid Res* **51**, 2957–2966, doi:10.1194/jlr.M006536 (2010).
- Kim, J. Y. *et al.* Metabolic profiling of plasma in overweight/obese and lean men using ultra performance liquid chromatography and Q-TOF mass spectrometry (UPLC-Q-TOF MS). *J Proteome Res* **9**, 4368–4375, doi:10.1021/pr100101p (2010).
- Rauschert, S. *et al.* Lipidomics Reveals Associations of Phospholipids With Obesity and Insulin Resistance in Young Adults. *J Clin Endocrinol Metab* **101**, 871–879, doi:10.1210/jc.2015-3525 (2016).
- Floegel, A. *et al.* Identification of serum metabolites associated with risk of type 2 diabetes using a targeted metabolomic approach. *Diabetes* **62**, 639–648, doi:10.2337/db12-0495 (2013).
- Del Bas, J. M. *et al.* Impairment of lysophospholipid metabolism in obesity: altered plasma profile and desensitization to the modulatory properties of n-3 polyunsaturated fatty acids in a randomized controlled trial. *Am J Clin Nutr* **104**, 266–279, doi:10.3945/ajcn.116.130872 (2016).

19. Li, F. *et al.* Lipidomics reveals a link between CYP1B1 and SCD1 in promoting obesity. *J Proteome Res* **13**, 2679–2687, doi:10.1021/pr500145n (2014).
20. Kim, H. Y. *et al.* Lysophospholipid profile in serum and liver by high-fat diet and tumor induction in obesity-resistant BALB/c mice. *Nutrition* **30**, 1433–1441, doi:10.1016/j.nut.2014.04.013 (2014).
21. Kim, H. J. *et al.* Metabolomic analysis of livers and serum from high-fat diet induced obese mice. *J Proteome Res* **10**, 722–731, doi:10.1021/pr100892r (2011).
22. Lehmann, R. *et al.* Circulating lysophosphatidylcholines are markers of a metabolically benign nonalcoholic fatty liver. *Diabetes Care* **36**, 2331–2338, doi:10.2337/dc12-1760 (2013).
23. Tanaka, N., Matsubara, T., Krausz, K. W., Patterson, A. D. & Gonzalez, F. J. Disruption of phospholipid and bile acid homeostasis in mice with nonalcoholic steatohepatitis. *Hepatology* **56**, 118–129, doi:10.1002/hep.25630 (2012).
24. Aleksandrova, K. *et al.* Metabolic syndrome and risks of colon and rectal cancer: the European prospective investigation into cancer and nutrition study. *Cancer Prev Res (Phila)* **4**, 1873–1883, doi:10.1158/1940-6207.CAPR-11-0218 (2011).
25. Barber, M. N. *et al.* Plasma lysophosphatidylcholine levels are reduced in obesity and type 2 diabetes. *PLoS One* **7**, e41456, doi:10.1371/journal.pone.0041456 (2012).
26. Pickens, C. A. *et al.* Plasma phospholipids, non-esterified plasma polyunsaturated fatty acids and oxylipids are associated with BMI. *Prostaglandins Leukot Essent Fatty Acids* **95**, 31–40, doi:10.1016/j.plefa.2014.12.001 (2015).
27. Pickens, C. A., Matsuo, K. H. & Fenton, J. I. Relationship between Body Mass Index, C-Peptide, and Delta-5-Desaturase Enzyme Activity Estimates in Adult Males. *PLoS One* **11**, e0149305, doi:10.1371/journal.pone.0149305 (2016).
28. C., O. J. *An interactive tool for comparing lists with Venn Diagrams*, <http://bioinfogp.cnb.csic.es/tools/venny/index.html> (2007).
29. Reinehr, T. *et al.* Changes in the serum metabolite profile in obese children with weight loss. *Eur J Nutr* **54**, 173–181, doi:10.1007/s00394-014-0698-8 (2015).
30. Graessler, J. *et al.* Lipidomic profiling before and after Roux-en-Y gastric bypass in obese patients with diabetes. *Pharmacogenomics J* **14**, 201–207, doi:10.1038/tpj.2013.42 (2014).
31. Pietilainen, K. H. *et al.* Acquired obesity is associated with changes in the serum lipidomic profile independent of genetic effects—a monozygotic twin study. *PLoS One* **2**, e218, doi:10.1371/journal.pone.0000218 (2007).
32. Jones, A. G. & Hattersley, A. T. The clinical utility of C-peptide measurement in the care of patients with diabetes. *Diabet Med* **30**, 803–817, doi:10.1111/dme.12159 (2013).
33. Tonks, K. T. *et al.* Skeletal muscle and plasma lipidomic signatures of insulin resistance and overweight/obesity in humans. *Obesity (Silver Spring)* **24**, 908–916, doi:10.1002/oby.21448 (2016).
34. Klingler, C. *et al.* Lysophosphatidylcholines activate PPARdelta and protect human skeletal muscle cells from lipotoxicity. *Biochim Biophys Acta* **1861**, 1980–1992, doi:10.1016/j.bbali.2016.09.020 (2016).
35. Virkamaki, A. *et al.* Intramyocellular lipid is associated with resistance to *in vivo* insulin actions on glucose uptake, antilipolysis, and early insulin signaling pathways in human skeletal muscle. *Diabetes* **50**, 2337–2343 (2001).
36. Falck-Ytter, Y., Younossi, Z. M., Marchesini, G. & McCullough, A. J. Clinical features and natural history of nonalcoholic steatosis syndromes. *Semin Liver Dis* **21**, 17–26 (2001).
37. Seppala-Lindroos, A. *et al.* Fat accumulation in the liver is associated with defects in insulin suppression of glucose production and serum free fatty acids independent of obesity in normal men. *J Clin Endocrinol Metab* **87**, 3023–3028, doi:10.1210/jcem.87.7.8638 (2002).
38. Fon Tacer, K. & Rozman, D. Nonalcoholic Fatty liver disease: focus on lipoprotein and lipid deregulation. *J Lipids* **2011**, 783976, doi:10.1155/2011/783976 (2011).
39. Rashid, S. & Genest, J. Effect of obesity on high-density lipoprotein metabolism. *Obesity (Silver Spring)* **15**, 2875–2888, doi:10.1038/oby.2007.342 (2007).
40. Oresic, M. *et al.* Prediction of non-alcoholic fatty-liver disease and liver fat content by serum molecular lipids. *Diabetologia* **56**, 2266–2274, doi:10.1007/s00125-013-2981-2 (2013).
41. Sansone, A. *et al.* Hexadecenoic Fatty Acid Isomers in Human Blood Lipids and Their Relevance for the Interpretation of Lipidomic Profiles. *PLoS One* **11**, e0152378, doi:10.1371/journal.pone.0152378 (2016).
42. Ferreri, C. *et al.* Fatty Acids in Membranes as Homeostatic, Metabolic and Nutritional Biomarkers: Recent Advancements in Analytics and Diagnostics. *Diagnostics (Basel)* **7**, doi:10.3390/diagnostics7010001 (2016).
43. Zhang, L., Keung, W., Samokhvalov, V., Wang, W. & Lopaschuk, G. D. Role of fatty acid uptake and fatty acid beta-oxidation in mediating insulin resistance in heart and skeletal muscle. *Biochim Biophys Acta* **1801**, 1–22, doi:10.1016/j.bbali.2009.09.014 (2010).
44. Donovan, E. L., Pettine, S. M., Hickey, M. S., Hamilton, K. L. & Miller, B. F. Lipidomic analysis of human plasma reveals ether-linked lipids that are elevated in morbidly obese humans compared to lean. *Diabetol Metab Syndr* **5**, 24, doi:10.1186/1758-5996-5-24 (2013).
45. Yamazaki, Y. *et al.* The proportion of nervonic acid in serum lipids is associated with serum plasmalogen levels and metabolic syndrome. *J Oleo Sci* **63**, 527–537 (2014).
46. Comstock, S. S. *et al.* Adipokines and obesity are associated with colorectal polyps in adult males: a cross-sectional study. *PLoS One* **9**, e85939, doi:10.1371/journal.pone.0085939 (2014).
47. Comstock, S. S. *et al.* Cross-sectional analysis of obesity and serum analytes in males identifies sRAGE as a novel biomarker inversely associated with diverticulosis. *PLoS One* **9**, e95232, doi:10.1371/journal.pone.0095232 (2014).
48. Comstock, S. S. *et al.* Association of insulin-related serum factors with colorectal polyp number and type in adult males. *Cancer Epidemiol Biomarkers Prev* **23**, 1843–1851, doi:10.1158/1055-9965.EPI-14-0249-T (2014).
49. Pickens, C. A., Lane-Elliott, A., Comstock, S. S. & Fenton, J. I. Altered Saturated and Monounsaturated Plasma Phospholipid Fatty Acid Profiles in Adult Males with Colon Adenomas. *Cancer Epidemiol Biomarkers Prev* **25**, 498–506, doi:10.1158/1055-9965.EPI-15-0696 (2016).
50. Pickens, C. A., Albuquerque Pereira, M. F. & Fenton, J. I. Long-chain omega-6 plasma phospholipid polyunsaturated fatty acids and association with colon adenomas in adult men: a cross-sectional study. *Eur J Cancer Prev*. doi:10.1097/CEJ.0000000000000312 (2016).
51. Issac, G., McDonald, S. & Astarita, G. (Waters Corporation, Milford, MA, USA., 2011).
52. Ekanayaka, E. A., Celiz, M. D. & Jones, A. D. Relative mass defect filtering of mass spectra: a path to discovery of plant specialized metabolites. *Plant Physiol* **167**, 1221–1232, doi:10.1104/pp.114.251165 (2015).
53. Gromski, P. S. *et al.* Influence of missing values substitutes on multivariate analysis of metabolomics data. *Metabolites* **4**, 433–452, doi:10.3390/metabo4020433 (2014).
54. Armitage, E. G., Godzien, J., Alonso-Herranz, V., Lopez-Gonzalez, A. & Barbas, C. Missing value imputation strategies for metabolomics data. *Electrophoresis* **36**, 3050–3060, doi:10.1002/elps.201500352 (2015).
55. Rose, H. G. & Oklander, M. Improved Procedure for the Extraction of Lipids from Human Erythrocytes. *J Lipid Res* **6**, 428–431 (1965).
56. Agren, J. J., Julkunen, A. & Penttila, I. Rapid separation of serum lipids for fatty acid analysis by a single aminopropyl column. *J Lipid Res* **33**, 1871–1876 (1992).
57. Team, R. C. R. A language and environment for statistical computing, <http://www.r-project.org/> (2015).
58. Benjamini, Y. & Hochberg, Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society B* **57**, 289–300 (1995).

59. Bland, J. M. & Altman, D. G. Multiple significance tests: the Bonferroni method. *BMJ* **310**, 170 (1995).
60. Vazquez, A. I. *et al.* Multi-Omic Prediction of Disease Risk and Progression Using Bayesian Generalized Additive Models. *Genetics* (2016).
61. de los Campos, G., Gianola, D. & Allison, D. B. Predicting genetic predisposition in humans: the promise of whole-genome markers. *Nat Rev Genet* **11**, 880–886, doi:10.1038/nrg2898 (2010).
62. Vazquez, A. I. *et al.* Increased Proportion of Variance Explained and Prediction Accuracy of Survival of Breast Cancer Patients with Use of Whole-Genome Multiomic Profiles. *Genetics* **203**, 1425–1438, doi:10.1534/genetics.115.185181 (2016).
63. Greenacre, M. J. Biplots: the joy of singular value decomposition. *Wiley Interdisciplinary Reviews: Computational Statistics* **4**, 399–406, doi:10.1002/wics.1200 (2012).
64. VanRaden, P. M. Efficient methods to compute genomic predictions. *J Dairy Sci* **91**, 4414–4423, doi:10.3168/jds.2007-0980 (2008).
65. Perez, P. & de los Campos, G. Genome-wide regression and prediction with the BGLR statistical package. *Genetics* **198**, 483–495, doi:10.1534/genetics.114.164442 (2014).
66. Vazquez, A. I., Wiener, H., Shrestha, S., Tiwari, G. & de los Campos, G. Integration of Multi-Layer Omic Data for Prediction of Disease Risk in Humans. *WCGALP* 1–6 (2014).

## Acknowledgements

The authors would like to thank Anthony Schillmiller and Michigan State University's Mass Spectrometry and Metabolomics Core Facility. A.I.V. acknowledges financial support from NIH grant 7-R01-DK-062148-10-S1, NIG grants R01GM09992 and R01GM101219, and NSF grant IOS-1444543 sub-award UFDSP00010707.

## Author Contributions

Concept and design: C.A.P. and J.I.F. Develop of methodology: A.D.J. and C.A.P. Statistical analyses: A.I.V. and C.A.P. Writing, review, and manuscript revision: C.A.P., A.I.V., A.D.J., J.I.F. Study supervision: C.A.P. and J.I.F.

## Additional Information

**Supplementary information** accompanies this paper at doi:10.1038/s41598-017-05785-0

**Competing Interests:** The authors declare that they have no competing interests.

**Publisher's note:** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2017