analytical chemistry

Article

pubs.acs.org/ac

# Comparative Analysis of Sample Preparation Methods To Handle the Complexity of the Blood Fluid Metabolome: When Less Is More

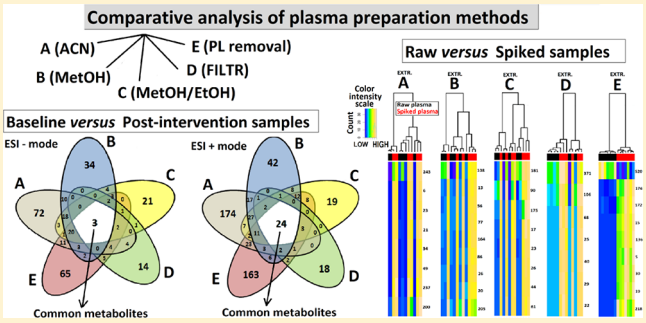
Sara Tulipani, Rafael Llorach, Mireia Urpi-Sarda, and Cristina Andres-Lacueva*

Biomarkers and Nutritional & Food Metabolomics Research Group, Department of Nutrition and Food Science, XaRTA, INSA, Faculty of Pharmacy, University of Barcelona, Av. Joan XXIII s/n, 08028, Barcelona, Spain

INGENIO−CONSOLIDER Program, Fun-C-Food CSD2007-063, Ministry of Science and Innovation, Barcelona, Spain

Supporting Information

**ABSTRACT:** Blood sample preparation before LC-MS metabolomic fingerprinting is one of the most challenging and error-prone parts of the analytical procedure. Besides proteins, phospholipids contained in blood fluids are known to cause matrix effects and ion suppression phenomena, thus masking biological variation. Nevertheless, the commonly used sample preparation techniques do not consider their removal prior to analysis. Pooled plasma and serum samples were used as biological material, partly as raw samples and partly spiked with distinct concentrations of a metabolite mix (1−5 μg/mL). Prior to LC-ESI-qToF-MS-driven metabolomic analysis, samples were subjected to different preparation methods consisting of three extractions with organic solvents (acetonitrile, methanol, and methanol/ethanol), a membrane-based solvent-free technique, and a hybrid method combining solvent extraction and SPE-mediated removal of phospholipids. The comparative analysis among sample preparation procedures was based on the capacity to detect endogenous compounds in raw samples, differentiate raw versus spiked samples, and reveal real-life metabolomic changes, following a dietary intervention. Method speed, minimum sample handling, compatibility to automation, and applicability to large-scale metabolomic studies were also considered. The combination of solvent deproteinization and the selective removal of phospholipids was revealed to be the most suitable method, in terms of improvement of nonlipid metabolite coverage, extraction reproducibility, quickness, and compatibility with automation, the minimization of matrix effects being among the most probable causes for the good extraction performance associated with the removal of phospholipid species. The main advantage of conventional solvent extraction procedures was the metabolite information coverage for lipid low-molecular-weight species, and extraction with acetonitrile was generally the second choice for sample preparation. Ultrafiltration was the least effective method for plasma and serum preparation; thus, its use without a previous solvent extraction step of the samples should be discarded. According to the presented data, there is no apparent reason to believe that sacrificing information on lipid compounds is too high of a price to pay in order to gain more information on nonlipid LMW metabolites.




To successfully face the challenge of biomarker discovery and hypothesis generation, technology-driven approaches such as metabolomics are needed to provide a comprehensive overview (fingerprinting) of all the low-molecular-weight (LMW) metabolites present in a cellular or biological system at a given moment.[1−3] Liquid chromatography/mass spectrometry (LC-MS) techniques have proven to be the most suitable for examining the largest number of metabolites in biosamples;[4−6] however, globality is a massive challenge when sample treatments are needed prior to analysis. Sample preparation before chromatographic separation is still a time-consuming and error-prone analytical step, particularly when managing complex biological matrices such as blood fluids.[7,8] Consequently, at present, no universal technique suitable for blood fluid sample preparation applied to metabolomic fingerprinting exists.[9] Deproteinization is the minimum sample treatment required prior to analysis, to preserve the LC-MS system integrity and drastically reduce matrix effects, and is commonly achieved by organic solvent precipitation,[10−13] by protein denaturation (using heat or acids), or through membrane-based techniques such as ultrafiltration.[7,14] Nevertheless, the conventional sample preparation techniques do not take into account that, besides proteins, other components endogenous to the biological samples are a primary challenge in LC-ESI-MS bioanalytical methods. Above all, phospholipid species (mainly glycerophosphocholines, the major phospholipids in plasma) might prejudice LMW metabolite coverage and reproducibility, particularly for polar compounds, thus masking biological variation. In the absence of proper mobile phase gradients, they accumulate in the analytical column (accel-

Received: October 3, 2012
Accepted: November 28, 2012
Published: November 28, 2012

 © 2012 American Chemical Society  dx.doi.org/10.1021/ac302919t | Anal. Chem. 2013, 85, 341−348

eration of column degradation) and are slowly released from the column in subsequent analytical runs, together with other metabolites (gradual increase of baseline noise, matrix effects induction, alteration of separation characteristics).[15] Phospholipids are also known to cause ion suppression or enhancement phenomena, to adversely influence the mass accuracy of the coeluting species (particularly in ES+), to interact with other LMW metabolites,[16] and to often represent noisy variables, given their significant gender-, individual- or even time-related variability in blood. Despite these considerations, their removal from blood fluids prior to untargeted metabolomic analysis has not been considered yet.

The central objective of the present study was to investigate the effects of common and newly proposed methods for plasma and serum extraction, through the application of an LC-ESI-qToF-MS-driven metabolomic workflow. The effects of three conventional solvent-based protein precipitation procedures, a membrane-based solvent-free technique, and the combination of solvent deproteinization with the removal of phospholipids were evaluated in terms of extraction efficiency, minimization of matrix effects, and extraction repeatability, in the framework of two nutrimetamolomic studies.[17−19] Other practical aspects were also considered, such as method quickness, complexity, and compatibility with automation. Chemometric data analysis included independent intramethod comparisons, intermethod second-order (meta-)analysis, and the computationally assisted structural identification of diet-related biomarker candidates.

## EXPERIMENTAL SECTION

An overview of the sample preparation and the metabolomic workflow applied for comparative analysis is shown in the Supporting Information (Figure S-1).

**Solvents and Reagents.** All aqueous solutions were prepared using ultrapure water obtained from a Millipore Milli-Q Gradient A10 system (Millipore, Bedford, MA, USA). HPLC-grade methanol, *o*-phosphoric acid (85%), and formic acid were purchased from Scharlau Chemie S.A. (Barcelona, Spain), and LC-MS-grade acetonitrile 0.1% formic acid was purchased from Fluka. The following chemicals were obtained commercially: D-L-carnitine, L-phenylalanine, L-tryptophan, glycochenodeoxycholic acid, gallic acid, syringic acid, (−)-epicatechin, leucine, isoleucine, palmitic acid, acetylcholine chloride, acetyl-L-carnitine hydrochloride, 1-*O*-palmitoyl-sn-glycero-3-phosphocholine, 1-*O*-stearoyl-sn-glycero-3-phosphocholine, α-hydroxybutyric acid (Sigma−Aldrich, St Louis, MO), 4-hydroxyhippuric acid (PhytoLab GmbH & Co. KG), and naringenin (Extrasynthèse, Genay, France).

**Aqueous 16-Component Metabolite Mix.** An aqueous 16-component metabolite standard mix of biologically relevant compounds was prepared, including three amino acids, two carnitines, three organic acids, two acyl glycine conjugates, an ester of acetic acid and choline, a fatty acid, two lysophosphatidylcholines and two flavonoid compounds (details are given in the Supporting Information). The standard mix was used for data quality control, to monitor mass accuracies, retention times, signal intensities, and eventual drifts in instrumental sensitivity, and to evaluate the ability of each sample preparation method to highlight differences among unspiked, spiked, and differently spiked biological samples.

**Biological Samples.** Plasma and serum samples of healthy donors were obtained from two dietary intervention trials, namely at the beginning (T0) and at the end (T1) of an acute cocoa consumption study and a 4-week red wine consumption study, respectively. Comprehensive details on study design and sample collection have been previously described.[20,21] In both cases, pooled biological samples were used to avoid biological variability in the comparative analysis among sample preparation procedures. Aliquots were stored at −80 °C until analysis and processed as described in the next section.

**Sample Preparation.** Preparation of plasma and serum samples was carried out separately. In both cases, the order of sample preparation was randomized, to avoid bias. On the day of the analysis, aliquots of pooled T0 and T1 samples were thawed on ice and alternatively spiked with the 16-component metabolite mix, at two final standard concentrations (1 and 5 μg/mL), or with Milli-Q water (raw), to maintain the same dilution factor across the samples. All aliquots (1 mL) were then added, along with 20 μL of concentrated *o*-phosphoric acid, to break down noncovalent intermolecular interactions and maximize the recovery of strongly protein-bound metabolites. Samples were vortexed for 1 min, left to stand for 10 min at 4 °C, and then subjected to the different high-throughput metabolite extraction procedures, in triplicate, to provide technical replicates.

*Solvent Extraction.* Samples were vortexed for 2 min with three times the volume of, respectively, acetonitrile (method A), methanol (method B), or a combination of methanol and ethanol at a ratio of 50:50 (method C), and kept at 4 °C for 10 min to promote further protein precipitation. The mixtures were then centrifuged at 10 000*g* at 4 °C, for 10 min, and the supernatants were collected.

*Ultrafiltration.* In method D, protein removal was obtained by sample dilution with three times the volume of Milli-Q water, followed by ultrafiltration at 9100 *g* at 4 °C for 1 h, through HMT filter units holding a microporous regenerated cellulose membrane with a 5000 Da nominal molecular weight limit, according to the manufacturer's suggestions (Ultrafree-MC-PLHCC for Metabolome Analysis, HMT Co., Ltd.).

*Solvent Extraction and SPE-Mediated Phospholipid Removal.* Sample preparation was carried out by acidic solvent extraction, followed by glycerophospholipid solid-phase extraction (SPE) (method E), using Ostro 96-well plates with pressure valves (Ostro plates, Waters). An Ostro plate was fixed on top of a 96-well collection plate in a manifold with a pressure gauge, and was connected to vacuum. Samples were pipetted into the wells, followed by the forceful addition of 1% formic acid in acetonitrile (3:1 solvent/sample) for in-well protein precipitation. The mixture was then quickly mixed to promote solubilization of plasma components from the plasma protein precipitate, and kept at 4 °C for 10 min to promote further protein precipitation. Vacuum (15 in. (∼381 mm) Hg) was then applied to the Ostro plate for 10 min, through a vacuum manifold, causing the valve mat tips in the plate to open; the precipitation solvent containing the nonphospholipid plasma components was filtered out and collected in the 96-well collection plate. No external standards were added to the sample matrices after extraction.

All solvents were kept at 4 °C prior to their use, and all procedures were carried out in a cold room, assuming that a 4 °C extraction temperature and the relatively short extraction time (10 min) may be favorable for avoiding biological sample degradation and reducing the risk of metabolite precipitation. 50 μL sample aliquots were used, except with the Ostro 96-well plate extraction, which required a minimum sample volume of 100 μL. On the assumption that plasma and serum are 100% aqueous matrices, the maximum percentage of organic solvents


dx.doi.org/10.1021/ac302919t | *Anal. Chem.* 2013, 85, 341−348

in the extracted samples was considered, overall, to be 75%, except for ultrafiltration (100% aqueous). In addition, sample lyophilization or evaporation to dryness was bypassed, since it was considered to be a time-consuming and critical step, because of the risk of incomplete solubilization of the dry residues during sample reconstitution. Therefore, aliquots of all sample extracts (120 $\mu$L) were diluted 1:1 with Milli-Q water within the HPLC injection plate, prior to analysis, in order to reduce the organic percentage and, thus, the risk of column flooding and poor chromatography.[22]

The following quality control (QC) samples were also prepared prior to analysis: QC1, Milli-Q water samples; QC2, aqueous 16-component metabolite mix (5 $\mu$g/mL final standard concentration); and QC3, randomly selected biological samples reinjected in opposite positions within each batch, along the sequence of analysis.

**LC-ESI-q-ToF-MS Data Acquisition.** Liquid chromatography was performed on an HPLC Agilent series 1200RR system using a Phenomenex RP 18 Luna column (50 mm × 2.0 mm, 5 $\mu$m) at 40 °C. MS experiments were carried out on a hybrid quadrupole time-of-flight (q-ToF) QSTAR Elite (Applied Biosystems/MDS Sciex) equipped with a TurboIonspray source operating in positive and negative ionization modes with a scan range from 70 $m/z$ to 700 $m/z$. The LC-MS operating conditions previously established for urine metabolomic fingerprinting[20,23,24] were optimized for the different biological matrices (details given in the Supporting Information). Because of the organic percentage in the extracted samples, the injection volume was reduced (5 $\mu$L) in order to preserve the peak shape of early eluting compounds from distortion. The time of flight (ToF) was calibrated with taurocholic acid (ions at $m/z$ 79.9568 and $m/z$ 514.2844) (1 pmol/$\mu$L) and reserpine (ions at $m/z$ 195.1651 and $m/z$ 609.2812) (1 pmol/$\mu$L) for negative and positive mode calibration, respectively. Prior to analysis, at least two QC2 samples were analyzed to check for system suitability and stability of standard RT, signal intensity, and mass accuracy values. Solvent mixtures were also injected alone, to produce an ion exclusion list of solvent-dependent mass features not to be considered during comparative data analysis. After that, 5−10 biological samples were injected for system conditioning with the sample matrix (plasma or serum). To avoid possible bias, a doubly rerandomized batch design was used for run sequence order. Samples were subdivided into homogeneous sub-batches (~10 injections each), separated by the regular analysis of QC samples consisting of 30% of the total runs. Per each biological matrix, 340 consecutive injections were acquired (170 × 2 ionization modes). The system was controlled by the software Analyst 2.0 that was supplied by Applied Biosystems (Foster City, CA, USA).

**Data Conversion and Preprocessing.** As shown in the Supporting Information (Figure S-1), raw chromatographic data files were converted from the original Analyst format (*.wiff) into peaks (*.peaks) and mzXML (*.mzXML) formats, respectively, to ensure format compatibility with MarkerView software version 1.2.1 (Applied Biosystems, MDS Sciex, Toronto, Ontario, Canada) and with R-based packages. MarkerView software and R-based packages were then used in parallel, in order to apply alternative/complementary tools for data preprocessing and further chemometric analysis. MarkerView peak finding options were set as follows: subtraction offset, 5 scans; subtraction multiplication factor, 1.5; minimum spectral peak width, 1 ppm, minimum retention time peak width, 3 scans; and noise threshold, 5. Peak alignment options were set as follows: retention time tolerance, 0.07 min; mass tolerance, 0.02 Da; intensity threshold, 5. The dataset was transformed logarithmically and Pareto-scaled (each variable weighted according to $1/\sqrt{\mathrm{SD}}$) before unsupervised and supervised statistical models (PCA, PCA-DA) were applied (MarkerView software version 1.2.1). The in-house R-based package known as MAIT (Metabolite Automatic Identification Toolkit, European patent pending) was used to apply a modularly designed and almost fully automatic metabolomic workflow including peak detection and annotation, dimensionality reduction, and mass spectra building, until significant spectra finding and identification. The settings used in the workflow were established during MAIT development as optimal for the preprocessing of metabolomic data coming from biological matrices, specifically acquired under our experimental conditions,[24] and their adequacy was verified by checking that the expected isotope peaks, adducts, and fragment ions generated from the standard metabolites spiked in the samples were correctly clustered in spectra. For peak detection, the R XCMS function[25] was applied by the MAIT workflow by using the default settings, except for the S/N threshold (snThres =3), bandwidth grouping (bwGroup =3), and mass width grouping (mzWidGrou =0.05). Then, the R CAMERA function[26] was applied to cluster the different mass features coming from the same metabolite (including the molecular ion, $^{13}$C isotopes, adducts, and in-source fragments) into spectra. Among the dimensionality reduction methods proposed by the MAIT workflow to treat ions colinearity and improve the statistical power of the gathered data, non-negative matrix factorization[27] was applied, to switch from a dataset where the variables are individual mass features to a dimensionally reduced data matrix composed of spectra, then subjected to univariate and multivariate analysis for significant spectra identification.

**Chemometric Analysis.** The use of commercial software and R-based packages was combined for method comparison. MarkerView software was used to evaluate data quality assurance, overall extraction-dependent effects, and in a first-pass comparative analysis of sample preparation techniques. Unsupervised (PCA) and supervised (PCA-DA) multivariate analysis were carried out for evaluation of the overall extraction-dependent effects. Univariate Student's $t$ tests between pairwise sample classes for each extraction method were used to evaluate the capacity of each procedure in distinguishing between raw (unspiked), spiked and differently spiked biological samples ($p < 0.05$). The detection of endogenous metabolites commonly present in blood fluids was investigated to compare the method extraction efficiencies on raw samples.[13] R-based packages were used for further pairwise and second-order sample class comparisons[28] (XCMS, metaXCMS), cross-validation of sample classification, and computationally assisted identification of significant metabolites up- or down regulated following the dietary intervention, within the two nutrimetabolomic case studies (raw T0 versus T1 samples). PCA scoreplot, heatmap, boxplot and Venn diagram functionalities were helpful in results visualization.

## RESULTS AND DISCUSSION

**Data Quality Assurance.** The stability of the chromatographic system throughout the data acquisition phase was first examined through PCA of the global datasets (biological and QC samples) and confirmed by batches overlapping; the


dx.doi.org/10.1021/ac302919t | Anal. Chem. 2013, 85, 341−348

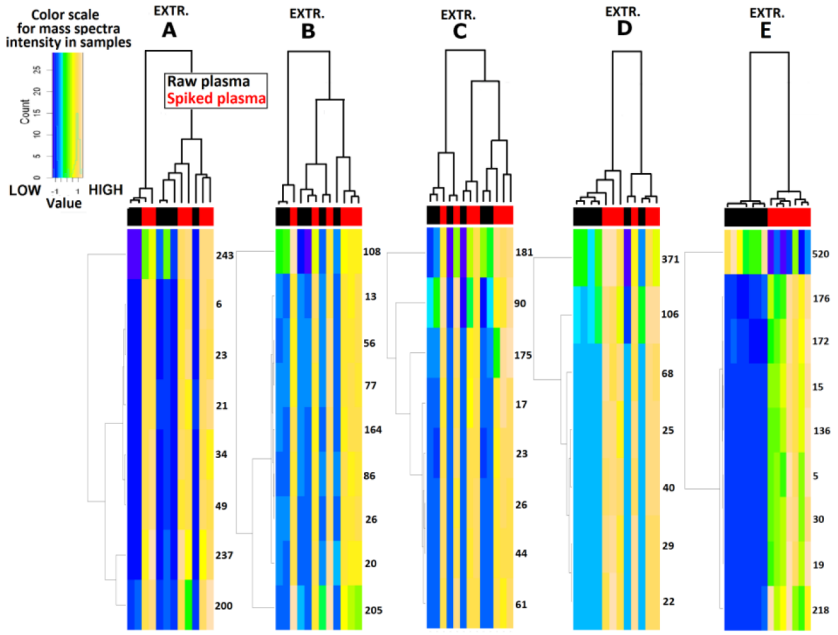



**Figure 1.** Heatmaps reflecting the capacity to differentiate raw versus spiked (1 $\mu$g/mL) plasma samples according to the initial sample preparation (threshold $p$-value = 0.001; Pearson clustering distance, negative ion mode). Each colored cell represents the intensity of a mass spectrum in one sample, according to the color scale on the left. Rows are mass spectra, and columns are raw (black) or spiked (red) plasma samples. The heatmap representation was obtained by the signHeatMap function of the MAIT algorithm.

absence of apparent carryover (QC1, QC2) or sample artifactual clustering; the clear spatial separation among the different sample classes, regardless of the batch injection order and randomization (PC1 versus PC2 scores plots in Figure S-2 in the Supporting Information); and the positioning of the QC3 samples (reinjections within batches) close to the corresponding first injected samples. The analytical variability across the runs was then evaluated by monitoring the standard metabolite components of QC2 samples injected along the entire datasets, covering the RT range from 0.27 min (earlier-eluting standard: D-L-carnitine) to 7.30 min (later-eluting standard: palmitic acid). The run-to-run repeatability of metabolites RT and mass accuracies successfully exceeded the quality criteria proposed for metabolomic analysis protocol (intrabatch and interbatch retention time shifts of <0.02 min, mass accuracy deviations of <5 mDa), and the signal intensity variation was satisfactory (peak areas of CV < 20%),[29,30] except for metabolites with very low response to ionization (palmitic acid) (see Table S-1 in the Supporting Information). No instrument failure indicating a decline in sensitivity, RT shifts, or changes in mass accuracy was observed. Therefore, the data gave confidence both about the robustness of the HPLC-q-ToF-MS system operating conditions and the reliability of the data for further statistical analysis.

**Overall Extraction-Based Effects.** The clustering of the biological samples within the PCA scores plot space (PC1 versus PC2) seems to be more dependent on the preparation method used than on the type of samples (unspiked, spiked, baseline, post-intervention), confirming the influence of sample extraction on the final metabolomic profile detected (see Figure S-2 in the Supporting Information).

The effects of the extraction conditions were first evaluated on the RT and peak areas of the 16 standard metabolites spiked in biological samples ("procedure" *plus* "biological matrix" effects), compared to aqueous samples (only "procedure effects") (see Table S-2 in the Supporting Information). Generally, a similar behavior was observed for plasma and serum samples, compared to aqueous samples, although plasma was associated with greater matrix effects for all of the tested metabolites. The use of EDTA anticoagulant during plasma separation from whole blood probably played a role in the differences observed between the two matrices, since EDTA is known to impact the peak response of metabolites via LC–ESI-MS analysis, by inducing matrix effects and affecting extraction.[31]

For both biological matrices, sample preparation by Ostro 96-well plate technology (E) determined the removal of the lysophospholipid species (palmitoyl-sn-glycero-3-phosphocholine and steroyl-sn-glycero-3-phosphocholine), as expected, but also a drastic increase in signal intensity (up to 30-fold higher peak areas) for most of the remaining spiked compounds, compared to the other methods (CV < 20%). Among the conventional solvent extraction procedures tested, protein precipitation with acetonitrile (A) was generally the second choice for sample preparation, while solvent-free ultrafiltration (D) resulted the least-effective method for plasma and serum preparation, particularly for the lower extraction of metabolites containing a nonpolar hydrocarbon alkane chain (lysophosphatidylcholines, fatty acids) or relatively lipophilic metabolites (flavonoids, cholic acids). These findings suggested that a previous solvent extraction step would always be required prior to filtration, in order to avoid those less polar compounds getting glued to the container walls and to maximize their solubilities and the filtering potentialities.

**Raw versus Spiked Samples.** The efficiency of the sample extraction procedures was then evaluated for the capacity to reveal significant differences among raw biological samples and samples spiked with different concentrations of the standard component mix (1 and 5 $\mu$g/mL), and to give the highest fold change in the mass signal intensities, through paired $t$-tests (see Table S-3 in the Supporting Information). No method enabled 1-ppm-scale changes to be detected for all targeted metabolites,


dx.doi.org/10.1021/ac302919t | *Anal. Chem.* 2013, 85, 341–348

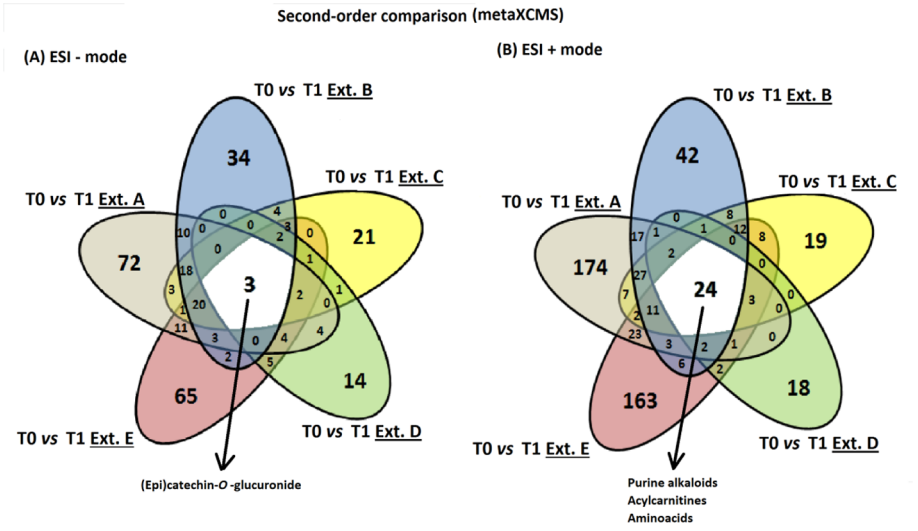



**Figure 2.** Venn diagrams showing the overlap among the five pairwise comparisons between baseline and post-dietary intervention samples (plasma T0 versus T1) obtained by each sample preparation method, in order to distinguish the diet-associated significant mass features only detected after unique sample preparation procedures from those shared among different extraction methods, in negative (A) and positive (B) ion mode. Defined thresholds: $p$ = 0.05, fold change = 1.5, mass tolerance = 0.02 amu, RT tolerance = 9 s.

and no optimal sample preparation method was definable for palmitic acid and the two lysophosphatidylcholines, at either of the spiked concentrations tested. These findings suggested that, although these most unpolar classes of compounds were more abundant in samples after solvent extraction procedures (methods A, B, and C), they are not quantitatively extracted by the organic solvents commonly used for sample preparation in metabolomic studies (methanol, ethanol, acetonitrile, or a combination of them), confirming that caution should be warranted when evaluating these metabolites among potential candidate metabolomic biomarkers.[32,33] In contrast, the combination of solvent extraction and SPE-mediated removal of phospholipids (method E) was the most suitable sample preparation for detecting subtle quantitative changes for the majority of the remaining metabolites monitored, particularly for highly polar (acetylcholine, acetyl-L-carnitine, (iso)leucine) and polar (L-phenylalanine) compounds (see Table S-3 in the Supporting Information). Heatmap visualizations confirmed that only samples subjected to this hybrid extraction were correctly clustered in separate classes (raw versus spiked samples) (see the example given in Figure 1), at a highly significant threshold ($p$ < 0.001). (−)-Epicatechin, naringenin, 4-hydroxyhippuric acid, and acetyl-L-carnitine were among the more highly significant standard metabolites responsible for class separation.

**Raw Samples among Different Extraction Methods.** In raw samples subjected to phospholipids removal, significantly higher signal intensities were observed for several compounds, including amino acids, carnitines, organic acids, and acyl glycine conjugates. Common (lyso)phospholipids were only detected after classical solvent extraction (methods A, B, and C), although with high variation among technical replicates (see Figures S-3 and S-4 in the Supporting Information).

**Baseline versus Post-Dietary Intervention Samples.** We finally compared the sample preparation methods according to the ability to reveal real-life metabolomic changes following a dietary intervention. Five independent intramethod comparisons between baseline and post-dietary intervention samples (paired $t$-test between T0 and T1) were first obtained. The results of the independently generated data matrices from intramethod comparisons were then subjected to second-order (meta-)analysis, in order to distinguish the diet-associated significant mass features only detected after unique sample preparation procedures from those shared among different extraction methods. The Venn diagrams in Figure 2 (metaXCMS graphical outputs) show the intra- and inter-method comparative analysis (see Figure S-5 in the Supporting Information shows outputs for serum sample analysis). Significant changes in the human plasma and serum metabolome following the interventions were revealed by all the sample extraction methods ($p$ < 0.05), although few of the significant metabolites were shared among them. A list of markers tentatively identified in plasma following the acute intake of cocoa is provided as Supporting Information (Table S-4). For each compound, the level of significance, fold, and type of change (increase/decrease) obtained through the different sample preparation procedures are presented, and a first-choice preparation method is proposed. The most typical markers of cocoa consumption were shared through all the sample preparation procedures, namely, purine alkaloids (caffeine and theobromine) and (epi)catechin-$O$-glucuronide, a phase II metabolite of cocoa polyphenols (flavan-3-ols).[23] A significant reduction of plasma acylcarnitines was also observed, regardless of the sample preparation procedure. Our data were in accordance with recent results from a large-scale quantitative and targeted metabolomic study,[34] where a strong negative association was observed between coffee consumption and the concentration of acylcarnitines in plasma, suggesting that components shared by both coffee and cocoa (i.e., purine alkaloids, niacin) may be responsible for these plasma effects. The meta-analysis of the results highlighted that acetonitrile-based extraction methods (A, E) gave the most information-rich samples, confirming them as the two best choice methods in our study. In particular, the combination of acidic acetonitrile extraction and SPE-mediated removal of phospholipids was the first-choice method to detect changes in compounds commonly targeted in nutrimetabolomic studies, such as phase II and microbial-derived metabolites of cocoa phenolics, and bile acids and their glycine conjugates, covering a wide range of RT and $m/z$. Figure 3 shows an example of the most highly significant


dx.doi.org/10.1021/ac302919t | Anal. Chem. 2013, 85, 341−348

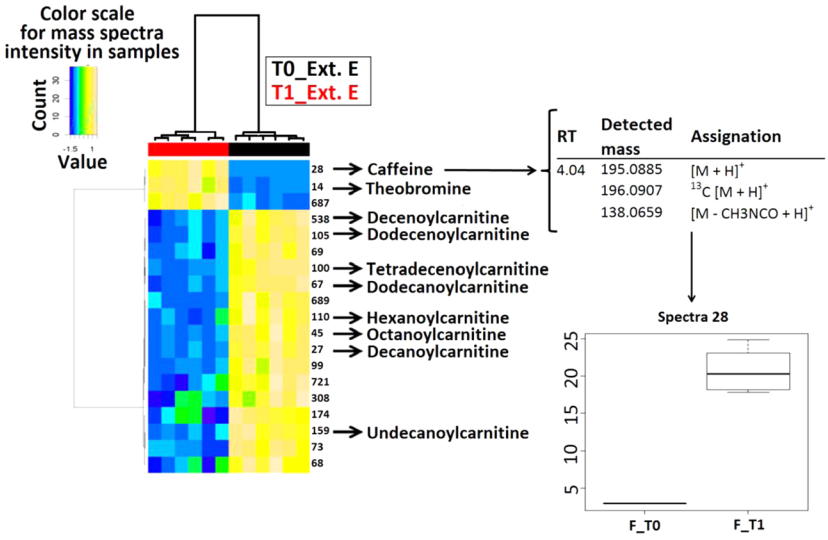



**Figure 3.** Heatmap showing examples of metabolites significantly up- and down-regulated in plasma following acute cocoa consumption ($p < 0.0001$; Pearson clustering distance; positive ion mode). Rows are mass spectra, and columns are plasma samples before cocoa consumption (T0) and 2 h following cocoa consumption (T1). An example of computationally assisted structural identification of significant mass spectra (caffeine). The heatmap and boxplot visualization were respectively obtained using the signHeatMap and drawBoxplot functions of the MAIT algorithm.

**Table 1. Overall Evaluation of the Sample Preparation Techniques**

| criteria | Method A (acetonitrile) | Method B (methanol) | Method C (methanol/ethanol) | Method D (solvent-free filtration) | Method E (acidic acetonitrile + SPE) |
|---|---|---|---|---|---|
| extraction repeatibility (CV %)[a] | 15 (11) | 10 (10) | 9 (8) | 13 (12) | 12 (12) |
| extraction efficiency[b] | 3-fold higher (2) | 3-fold higher (2) | 3-fold higher (2) | lowest | 4-fold higher (5) |
| detection of 1 ppm-scale changes[c] | 39% | 44% | 33% | 44% | 55% |
| detection of real-life diet-related metabolic changes[d]: | | | | | |
| amino acids | moderate | poor | poor | moderate | high |
| purine alkaloids | moderate | moderate | moderate | moderate | high |
| flavan-3-ols and microbial-derived metabolites | poor | poor | poor | poor | high |
| acylcarnitines | high | moderate | moderate | poor | moderate |
| fatty acids (C8−C18) | high | moderate | poor | very poor | poor |
| bile acids/bile acids glycine conjugates | poor | moderate | poor | very poor | high |
| simplicity | moderate | moderate | moderate | high | moderate |
| speed | moderate | moderate | moderate | low | high |
| environmentally friendly | moderate | moderate | moderate | high | moderate |
| instrumentation needed | lab centrifuge | lab centrifuge | lab centrifuge | lab centrifuge | vacuum system and manifold |
| compatibility with automation | low | low | low | high | high |
| key factors to be optimized/ major drawbacks | extraction repeatibility; poor automation | extraction efficiency; poor automation | extraction efficiency; poor automation | extraction efficiency; solvent extraction required | loss of PL[e] |

[a]Methods comparison based on peak areas CV (%) of the spiked metabolites. [b]Methods comparison based on mean peak areas of the spiked metabolites. For fold-change definition, pairwise comparisons with the least efficient method was carried out. Data about serum matrix are presented between parentheses. [c]Percentage of spiked metabolites detected as significantly different ($p < 0.01$) between raw and spiked samples. [d]Results from a nutrimetabolomic case-study (detection of plasma metabolomic changes following the acute consumption of cocoa powder (40 g) dissolved in water (250 mL). Terms are defined as follows: "high", when more than 50% of the total metabolites identified in a given chemical class are detected as significantly different between T0 and T1 samples, and with high statistical significance (at least $p < 0.001$); "moderate", when more than 50% of the total metabolites identified in a given chemical class are detected as significantly different between T0 and T1 samples, but without high statistical significance; "poor", when a number equal to or less than 50% of the total metabolites identified in a given chemical class are detected as significantly different between T0 and T1 samples, regardless of the significance threshold; and "very poor", when no metabolites identified in a given chemical class are detected as significantly different between T0 and T1 samples. More details are presented in Table S-3 in the Supporting Information. [e]PL = phospholipids.

metabolites associated with the acute intake of cocoa (ESI+ mode), when samples were prepared by method E ($p < 0.0001$; Pearson clustering distance). Besides (epi)catechin-*O*-glucuronide, three sulfate derivatives were identified in ESI- mode, namely, (epi)catechin-*O*-sulfate and two *O*-methyl-(epi)-catechin-*O*-sulfates, already recognized markers of cocoa or


dx.doi.org/10.1021/ac302919t | *Anal. Chem.* 2013, 85, 341−348

chocolate polyphenol intake and early bioavailability.[20,35] In addition, a strong negative association between cocoa consumption and a series of circulating medium- and long-chain acylcarnitines was particularly highlighted through this sample preparation procedure, deserving further attention, because of the influence of carnitine/acylcarnitine turnover on innumerable aspects of carbohydrate and lipid metabolism, including the regulation of insulin secretion by pancreatic $\beta$-cells and the determination of tissue insulin sensitivity.[36]

The minimization of matrix effects was among the most probable causes for the good extraction performance associated with the SPE removal of phospholipid species. A lower intensity of the EDTA-related peaks ($m/z$ 293.0978 and $m/z$ 315.0802, corresponding to EDTA and its sodium adduct, in positive ionization mode, respectively) was also observed in all plasma samples subjected to SPE, compared to the other methods, possibly due to interactions of the anticoagulant with the stationary phase of the Ostro 96-well plate (data not shown). In addition, given their significant gender-, individual-, or even time-related variability in plasma/serum, the removal of phospholipid species may have contributed to the discarding of noisy variables and allowed us to observe more clearly the subtle metabolomic changes that are occurring in blood following real-life situations, such as food consumption.

The removal of phospholipid species has never been included so far among the sample preparation steps prior to blood fluid LC-MS metabolic fingerprinting, probably because these LMW molecules are actually part of the plasma and serum metabolome and their selective removal would imply a conscious loss of information. Nevertheless, it is also known that it is practically impossible to measure the levels of all metabolites present in a biological sample simultaneously, with a single analytical platform, so that multiple analytical strategies are generally applied in parallel to handle the complexity of the metabolome and increase metabolite coverage in biosamples.[37] Moreover, a single extraction and detection method for all metabolites from biomatrices is already known to be unfeasible, and one-step extractions separating a lipophilic and hydrophilic layers intended for distinct analytical approaches have also been proposed, to optimize sample exploitation.[33] In this scenario, when considering that an emerging and self-standing -omics field (lipidomics) now "*comes to grips with lipid diversity*"[38] and is entirely dedicated to the systems-level scale analysis of all lipids and interacting moieties present in a biosample,[39,40] there is no apparent reason to believe that sacrificing information on lipid compounds is too high a price to pay in order to gain more information on nonlipid LMW metabolites.

## CONCLUSIONS

In the present study, we have applied a metabolomic workflow to compare the effects of conventional and newly proposed procedures for blood fluid extraction. In summary (Table 1), although no universal technique suitable for the optimal detection of all types of LMW metabolites exists, the combination of solvent extraction and SPE-mediated removal of phospholipids was revealed to (a) improve the signal response of most nonlipid LMW metabolites of blood fluids, compared to the other sample preparation techniques tested, probably by reducing the background of interfering species and minimizing ion suppression phenomena; (b) remove a class of metabolites (phospholipids) that are not quantitatively extracted by the organic solvents commonly used in metabolomics, may represent noisy variables given their significant gender-, individual- or even time-related variability in blood, and are among the target compounds of a self-standing "omics" (lipidomics); (c) be a quick and moderately simple method allowing multisample simultaneous preparation, thus compatible with robotic liquid handlers and automation. Certainly, the loss of information about the variation of phospholipids and other lipid LMW species in blood fluids needs consideration, particularly in terms of sample exploitation, and a future implementation of Ostro technology for phospholipid recovery following SPE should be planned.

## ASSOCIATED CONTENT

### Supporting Information

This material is available free of charge via the Internet at http://pubs.acs.org.

## AUTHOR INFORMATION

### Corresponding Author

*Tel.: +34 934034840. Fax: +34934035931. E-mail: candres@ub.edu. Website: www.nutrimetabolomics.com.

### Notes

The authors declare no competing financial interest.

## ACKNOWLEDGMENTS

This work was supported by national grants from the Spanish MICINN (CICYT AGL2009-13906-C02-01, AGL2010-10084-E, Ingenio-CONSOLIDER program FUN-C-Food CSD2007-063). S.T. would like to thank the "Juan de la Cierva" program; R.L. and M.U.S would like to thank the "Ramon y Cajal"program, from the MEC and Fondo Social Europeo (FSE).

## REFERENCES

(1) German, J. B.; Hammock, B. D.; Watkins, S. M. *Metabolomics* **2005**, *1* (1), 3−9.
(2) Metabolomics Society Metabolomics. http://www.metabolomicssociety.org/metabolomics.html (accessed Sept. 15, 2012).
(3) Mamas, M.; Dunn, W. B.; Neyses, L.; Goodacre, R. *Arch. Toxicol.* **2011**, *85* (1), 5−17.
(4) Dettmer, K.; Aronov, P. A.; Hammock, B. D. *Mass Spectrom. Rev.* **2007**, *26* (1), 51−78.
(5) Dunn, W. B.; Broadhurst, D.; Brown, M.; Baker, P. N.; Redman, C. W.; Kenny, L. C.; Kell, D. B. *J. Chromatogr. B* **2008**, *871* (2), 288−298.
(6) Dunn, W. B.; Broadhurst, D.; Begley, P.; Zelena, E.; Francis-McIntyre, S.; Anderson, N.; Brown, M.; Knowles, J. D.; Halsall, A.; Haselden, J. N.; Nicholls, A. W.; Wilson, I. D.; Kell, D. B.; Goodacre, R. *Nat. Protoc.* **2011**, *6* (7), 1060−1083.
(7) Hyotylainen, T. *Anal. Bioanal. Chem.* **2009**, *394* (3), 743−58.
(8) Psychogios, N.; Hau, D. D.; Peng, J.; Guo, A. C.; Mandal, R.; Bouatra, S.; Sinelnikov, I.; Krishnamurthy, R.; Eisner, R.; Gautam, B.; Young, N.; Xia, J.; Knox, C.; Dong, E.; Huang, P.; Hollander, Z.; Pedersen, T. L.; Smith, S. R.; Bamforth, F.; Greiner, R.; McManus, B.; Newman, J. W.; Goodfriend, T.; Wishart, D. S. *PLoS One* **2011**, *6* (2), No. e16957.
(9) Want, E. J.; Wilson, I. D.; Gika, H.; Theodoridis, G.; Plumb, R. S.; Shockcor, J.; Holmes, E.; Nicholson, J. K. *Nat. Protoc.* **2010**, *5* (6), 1005−1018.
(10) Bruce, S. J.; Jonsson, P.; Antti, H.; Cloarec, O.; Trygg, J.; Marklund, S. L.; Moritz, T. *Anal. Biochem.* **2008**, *372* (2), 237−249.
(11) Michopoulos, F.; Lai, L.; Gika, H.; Theodoridis, G.; Wilson, I. *J. Proteome Res.* **2009**, *8* (4), 2114−2121.

(12) Zelena, E.; Dunn, W. B.; Broadhurst, D.; Francis-McIntyre, S.; Carroll, K. M.; Begley, P.; O'Hagan, S.; Knowles, J. D.; Halsall, A.; Wilson, I. D.; Kell, D. B. *Anal. Chem.* **2009**, *81* (4), 1357−1364.
(13) Pereira, H.; Martin, J.-F.; Joly, C.; Sébédio, J.-L.; Pujos-Guillot, E. *Metabolomics* **2010**, *6* (2), 207−218.
(14) Ooga, T.; Sato, H.; Nagashima, A.; Sasaki, K.; Tomita, M.; Soga, T.; Ohashi, Y. *Mol. Biosyst.* **2011**, *7* (4), 1217−1223.
(15) Ismaiel, O. A.; Zhang, T.; Jenkins, R. G.; Karnes, H. T. *J. Chromatogr. B* **2010**, *878* (31), 3303−3316.
(16) Calbiani, F.; Careri, M.; Elviri, L.; Mangia, A.; Zagnoni, I. *J. Mass Spectrom.* **2006**, *41* (3), 289−294.
(17) Wishart, D. S. *Trends Food Sci. Technol.* **2008**, *19* (9), 482−493.
(18) Scalbert, A.; Brennan, L.; Fiehn, O.; Hankemeier, T.; Kristal, B. S.; van Ommen, B.; Pujos-Guillot, E.; Verheij, E.; Wishart, D.; Wopereis, S. *Metabolomics* **2009**, *5* (4), 435−458.
(19) Llorach, R.; Garcia-Aloy, M.; Tulipani, S.; Vazquez-Fresno, R.Andres-LacuevaC. Nutrimetabolomic Strategies To Develop New Biomarkers of Intake and Health Effects*J. Agric. Food Chem.* **2012**, *60* (36), 8797−8808.
(20) Llorach, R.; Urpi-Sarda, M.; Jauregui, O.; Monagas, M.; Andres-Lacueva, C. *J. Proteome Res.* **2009**, *8* (11), 5060−5068.
(21) Rotches-Ribalta, M.; Urpi-Sarda, M.; Llorach, R.; Boto-Ordóñez, M.; Jauregui, O.; Chiva-Blanch, G.; Perez-Garcia, M.; Jaeger, W.; Guillen, M.; Corella, D.; Tinahones, F.; Estruch, R.; Andres-Lacueva, C. Gut and microbial resveratrol metabolite profiling after moderate long-term consumption of red wine versus dealcoholized red wine in humans by an optimized ultra-high-pressure liquid chromatography tandem mass spectrometry method. *J. Chromatogr. A* **2012**, *1265*, 105−113.
(22) Keunchkarian, S.; Reta, M.; Romero, L.; Castells, C. *J. Chromatogr. A* **2006**, *1119* (1−2), 20−28.
(23) Llorach, R.; Garrido, I.; Monagas, M.; Urpi-Sarda, M.; Tulipani, S.; Bartolome, B.; Andres-Lacueva, C. *J. Proteome Res.* **2010**, *9* (11), 5859−5867.
(24) Tulipani, S.; Llorach, R.; Jauregui, O.; Lopez-Uriarte, P.; Garcia-Aloy, M.; Bullo, M.; Salas-Salvado, J.; Andres-Lacueva, C. *J. Proteome Res.* **2011**, *10* (11), 5047−5508.
(25) Smith, C. A.; Want, E. J.; O'Maille, G.; Abagyan, R.; Siuzdak, G. *Anal. Chem.* **2006**, *78* (3), 779−787.
(26) Kuhl, C.; Tautenhahn, R.; Bottcher, C.; Larson, T. R.; Neumann, S. *Anal. Chem.* **2012**, *84* (1), 283−289.
(27) Brunet, J. P.; Tamayo, P.; Golub, T. R.; Mesirov, J. P. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101* (12), 4164−4169.
(28) Patti, G. J.; Tautenhahn, R.; Siuzdak, G. *Nat. Protoc.* **2012**, *7* (3), 508−516.
(29) Guy, P. A.; Tavazzi, I.; Bruce, S. J.; Rramadan, Z.; Kochhar, S. *J. Chromatogr. B* **2008**, *871* (2), 253−260.
(30) Bruce, S. J.; Tavazzi, I.; Parisod, V.; Rezzi, S.; Kochhar, S.; Guy, P. A. *Anal. Chem.* **2009**, *81* (9), 3285−3296.
(31) Gonzalez-Covarrubias, V.; Dane, A.; Hankemeier, T.; Vreeken, R. The influence of citrate, EDTA, and heparin anticoagulants to human plasma LC−MS lipidomic profiling. *Metabolomics* **2012**, DOI: 10.1007/s11306-012-0450-4.
(32) Barri, T.; Holmer-Jensen, J.; Hermansen, K.; Dragsted, L. O. *Anal. Chim. Acta* **2012**, *718*, 47−57.
(33) Whiley, L.; Godzien, J.; Ruperez, F. J.; Legido-Quigley, C.; Barbas, C. *Anal. Chem.* **2012**, *84* (14), 5992−5999.
(34) Altmaier, E.; Kastenmuller, G.; Romisch-Margl, W.; Thorand, B.; Weinberger, K. M.; Adamski, J.; Illig, T.; Doring, A.; Suhre, K. *Mol. Nutr. Food Res.* **2009**, *53* (11), 1357−1365.
(35) Urpi-Sarda, M.; Monagas, M.; Khan, N.; Llorach, R.; Lamuela-Raventos, R. M.; Jauregui, O.; Estruch, R.; Izquierdo-Pulido, M.; Andres-Lacueva, C. *J. Chromatogr. A* **2009**, *1216* (43), 7258−7267.
(36) Huffman, K. M.; Shah, S. H.; Stevens, R. D.; Bain, J. R.; Muehlbauer, M.; Slentz, C. A.; Tanner, C. J.; Kuchibhatla, M.; Houmard, J. A.; Newgard, C. B.; Kraus, W. E. *Diabetes Care* **2009**, *32* (9), 1678−1683.
(37) Oresic, M.; Vidal-Puig, A.; Hanninen, V. *Expert Rev. Mol. Diagn.* **2006**, *6* (4), 575−585.
(38) Shevchenko, A.; Simons, K. *Nat. Rev. Mol. Cell Biol.* **2010**, *11* (8), 593−8.
(39) Wenk, M. R. *Nat. Rev. Drug Discovery* **2005**, *4* (7), 594−610.
(40) Quehenberger, O.; Armando, A. M.; Brown, A. H.; Milne, S. B.; Myers, D. S.; Merrill, A. H.; Bandyopadhyay, S.; Jones, K. N.; Kelly, S.; Shaner, R. L.; Sullards, C. M.; Wang, E.; Murphy, R. C.; Barkley, R. M.; Leiker, T. J.; Raetz, C. R.; Guan, Z.; Laird, G. M.; Six, D. A.; Russell, D. W.; McDonald, J. G.; Subramaniam, S.; Fahy, E.; Dennis, E. A. *J. Lipid Res.* **2010**, *51* (11), 3299−3305.