analytical chemistry

Article

pubs.acs.org/ac

Untargeted UPLC-MS Profiling Pipeline to Expand Tissue Metabolome Coverage: Application to Cardiovascular Disease

Panagiotis A. Vorkas,[†] Giorgis Isaac,[‡] Muzaffar A. Anwar,[§] Alun H. Davies,[§] Elizabeth J. Want,[†] Jeremy K. Nicholson,[†,∥] and Elaine Holmes*[,†,∥]

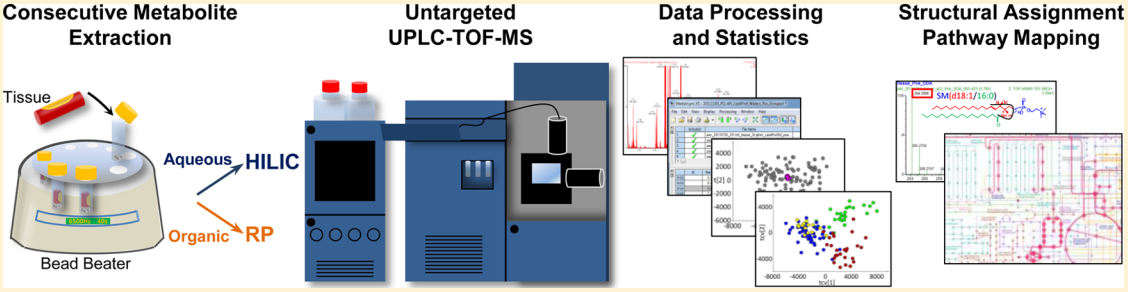
[†]Biomolecular Medicine, Division of Computational and Systems Medicine, Department of Surgery and Cancer, Faculty of Medicine, Imperial College London, London SW7 2AZ, U.K.

[‡]Pharmaceutical Discovery and Life Sciences, Waters Corporations, Milford, Massachusetts 01757, United States

[§]Academic Section of Vascular Surgery, Division of Surgery, Department of Surgery and Cancer, Faculty of Medicine, Imperial College London, London W6 8RF, U.K.

[∥]MRC-NIHR National Phenome Centre, IRDB Building, Imperial College London, Hammersmith Hospital, London W12 0NN, U.K.

S *Supporting Information*



ABSTRACT: Metabolic profiling studies aim to achieve broad metabolome coverage in specific biological samples. However, wide metabolome coverage has proven difficult to achieve, mostly because of the diverse physicochemical properties of small molecules, obligating analysts to seek multiplatform and multimethod approaches. Challenges are even greater when it comes to applications to tissue samples, where tissue lysis and metabolite extraction can induce significant systematic variation in composition. We have developed a pipeline for obtaining the aqueous and organic compounds from diseased arterial tissue using two consecutive extractions, followed by a different untargeted UPLC-MS analysis method for each extract. Methods were rationally chosen and optimized to address the different physicochemical properties of each extract: hydrophilic interaction liquid chromatography (HILIC) for the aqueous extract and reversed-phase chromatography for the organic. This pipeline can be generic for tissue analysis as demonstrated by applications to different tissue types. The experimental setup and fast turnaround time of the two methods contributed toward obtaining highly reproducible features with exceptional chromatographic performance (CV % < 0.5%), making this pipeline suitable for metabolic profiling applications. We structurally assigned 226 metabolites from a range of chemical classes (e.g., carnitines, α-amino acids, purines, pyrimidines, phospholipids, sphingolipids, free fatty acids, and glycerolipids) which were mapped to their corresponding pathways, biological functions and known disease mechanisms. The combination of the two untargeted UPLC-MS methods showed high metabolite complementarity. We demonstrate the application of this pipeline to cardiovascular disease, where we show that the analyzed diseased groups (n = 120) of arterial tissue could be distinguished based on their metabolic profiles.

Metabolic profiling relies on the application of a range of analytical technologies to measure simultaneously differential levels of multiple metabolites in biological matrices.[1] It is important to ensure wide metabolome coverage and consequently enhance biomarker detection probability. For this an untargeted format is the approach of choice warranting the ability to detect unmapped metabolites and pathways, as well as compounds originating from environmental interactions unrelated to the biology of the system studied.[2,3] To compensate for the different capabilities of each technique/method, their bias toward specific classes of metabolites,[4] and the wide physicochemical diversity of the metabolome in a biological matrix, multiple platforms[5−8] and methods[4] are required to expand metabolite coverage.

Hydrophilic interaction liquid chromatography coupled to mass spectrometry (HILIC-MS) is a relatively new chromatographic tool applied in the effort to expand metabolome coverage. The importance of HILIC is attributed to its ability to

Received: October 8, 2014
Accepted: February 9, 2015
Published: February 9, 2015

DOI: 10.1021/ac503775m
Anal. Chem. 2015, 87, 4184−4193

generate profiles of mainly polar metabolites and is therefore highly complementary to the traditionally used reversed-phase (RP) chromatography.[4] Several studies have demonstrated superior partitioning abilities of HILIC columns for polar compounds in biological matrices,[4,9−11] predominantly urine samples.[9,10,12,13] However, the application of HILIC in metabolic profiling studies continues to present unresolved analytical challenges particularly with respect to chromatographic performance.[4,11,14,15]

A previously described study aiming to expand metabolome coverage combining RP-LC-MS and HILIC-MS was applied using an HPLC system. This was translated into an extended 60 min run-time, and did not address tissue analysis.[4] Moreover, although a combination of RP-, HILIC-, and CE-MS has been described by Saric et al.,[5] it focused on evaluating the abilities of single-extraction methodologies for *Fasciola hepatica* worms rather than assessing biological pathway coverage of metabolites detected.

Herein, we describe a combination of RP- and HILIC-UPLC-MS methodologies functioning in an untargeted mode with the aim of expanding metabolome coverage. We focused on applying a UPLC-MS-based setup to tissue samples. Tissue samples can provide unique information on physiological or pathological mechanisms elucidation and biomarker discovery.[16,17] However, tissue lysis and metabolite extraction, as well as establishing the appropriate analysis combination can be challenging. We utilized a simple extraction format, based on consecutive extraction steps,[18] followed by a rational choice of UPLC-MS analysis, according to the lipophilicity of the extracts, wherein the aqueous extracts were analyzed using a HILIC and organic extracts by a RP method. We applied the method to profile tissue from patients with cardiovascular disease. The types of tissue used, such as atherosclerotic plaque, have complex structure, intense presence of lipid molecules and are biologically active. We evaluated the feasibility of analyzing these complex tissue extracts, with emphasis on robust column performance, stability, and longevity. The generic nature of the presented pipeline was demonstrated with applications in additional tissue types. A comprehensive structural assignment of detected ions was performed, resulting in the generation of a database of 226 unique structurally assigned metabolites. We further demonstrated the complementarity of the two chromatographic methods applied, and employed pathway mapping tools to display their contribution to expanding the coverage of metabolic pathways, biological functions, and known disease mechanisms.

EXPERIMENTAL SECTION

A schematic of the pipeline applied in the present study is illustrated in Figure 1.

Chemical Materials. Description of the solvents, chemicals, and authentic standards used can be found in Supporting Information.

Samples. Research ethics committee approval (RREC 2989 and RREC 3199) and patient informed consent was obtained for the collection of tissue specimens from abdominal aortic aneurysm repair surgery and carotid and femoral endarterectomy. Analyzed tissue samples consisted of four groups of patients: 26 abdominal aortic aneurysm (AAA), 52 carotid stenosing plaques (CAR), 26 femoral stenosing plaques (FEM) and 16 intimal thickening (INT). Samples obtained from endarterectomies have been previously described.[19] A table

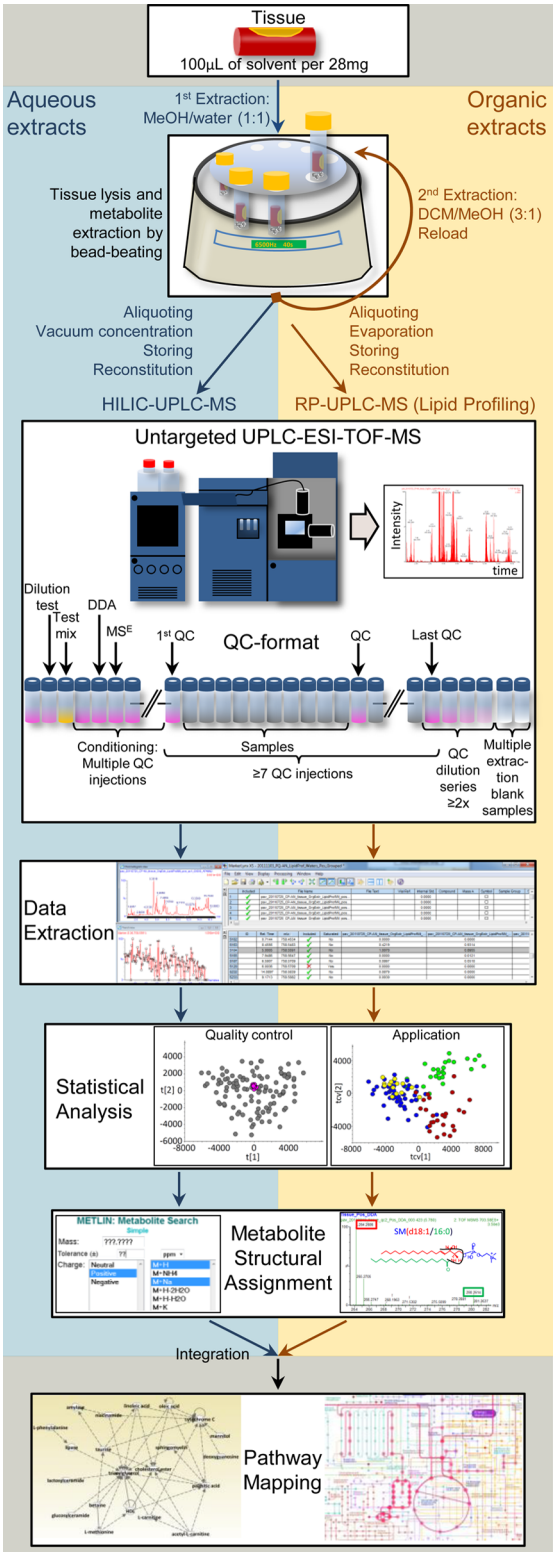


Figure 1. Schematic of the analytical pipeline applied in the present study. Following two consecutive tissue extraction procedures (aqueous followed by organic) each extract was handled separately up to the point of data integration, where a pathway mapping step was further applied to the combined data set. MeOH: Methanol. DCM: Dichloromethane. QC: Quality control. DDA: Data dependent acquisition (unbiased precursor ion selection for MS/MS), MS^E: Application of collision induced dissociation without precursor ion selection.

DOI: 10.1021/ac503775m
Anal. Chem. 2015, 87, 4184−4193

with patient demographics can be found in Supporting Information Table S-1.

Metabolite Extraction. *Aqueous Extraction.* Tissue samples (153−416 mg) were loaded into bead beating tubes (Percellys Steel-Kit) preloaded with three steel beads. A prechilled methanol(MeOH)/water solution (1:1), was added to the tissue samples. The volume of the solution was adjusted according to weight of the sample starting at a maximum weight with 1.5 mL, and reduced proportionally to sample weight (100 μL of solvent/28 mg of tissue). Tissue lysis and metabolite extraction was performed using a bead beater (Bertin Technologies) after freezing on dry ice. The bead beater was vibrating at 6500 Hz for 40 s and 2 plus 2 cycles were performed separated by freezing of the samples on dry ice. Bead beating was followed by centrifugation (Eppendorf, Centrifuge 5417R, Germany) at 13000*g* for 20 min, at 4 °C. Aliquots of 100 μL (∼25 mg of tissue/100 μL of aliquot) of the supernatant were obtained into Eppendorf tubes. Samples were spun in a vacuum concentrator for 3 h at 45 °C (Eppendorf Concentrator Plus, V-AQ mode) until dry and stored at −40 °C until analysis.

Organic Extraction. Following aqueous extraction a solution of prechilled dichloromethane(DCM)/MeOH (3:1) was added to the residual pellet. The volume of the solution was proportional to the sample weight (as described in previous paragraph; aqueous extraction). Samples were frozen on dry ice and reloaded into the bead beater (2 cycles, 6500 Hz, 40 s). Samples were then centrifuged at 13000*g* for 20 min, and 100 μL (∼25 mg of tissue/100 μL of aliquot) of organic phase supernatant was subsequently aliquoted into glass vials. Samples were allowed to evaporate at room temperature in an extractor hood overnight and stored at −40 °C until analysis.

HILIC-UPLC-MS Analysis of Aqueous Extracts. *UPLC-MS Analysis.* The aqueous extracts of the tissue samples were reconstituted in 200 μL of solvent mixture of H_2O/ACN (5:95) and transferred into Total Recovery vials (Waters Corp, USA), after centrifugation for 20 min at 13000*g*, 4 °C.

UPLC separation was conducted using an Acquity UPLC System (Waters Corp, USA). An Acquity UPLC BEH HILIC 2.1 × 100 mm, 1.8 μm, column (Waters Corp, USA) was used. Column temperature was set at 35 °C. Mobile phase A consisted of acetonitrile (ACN)/water (95:5) and mobile phase B ACN/water (50:50). In both solutions ammonium acetate was diluted to 10 mM and formic acid to 0.1%. The elution gradient was set as follows: 99% A (0.0−2.0 min; 0.4 mL/min), 99−45% A (2.0−8.0 min; 0.4 mL/min), 45−1% A (8.0−9.0 min; 0.4 mL/min), 1% A (9.0−9.1 min; 0.4−0.8 mL/min), 1% A (9.1−11.0 min; 0.8 mL/min), 1−99% A (11.0−11.1 min; 0.8 mL/min), 99% A (11.1−19.0 min; 0.8 mL/min), 99% A (19.0−19.1 min; 0.8−0.4 mL/min), 99% A (19.1−23.0 min; 0.4 mL/min). An injection volume of 10 μL was used for both positive and negative ionization polarity modes. The autosampler was set at 4 °C. Mass spectrometry was performed using a Premier Q-TOF (Waters MS Technologies Ltd., UK) with an electrospray ionization (ESI) source. MS conditions can be found in Supporting Information.

A standard QC strategy[20] was used for the UPLC-MS analysis. Briefly, a pooled sample (Quality Control Sample, QC) of the reconstituted extracts was prepared. This sample was injected at least 10 times before initiating the run, in order to condition the column. Then the sample was reinjected once at the beginning, every 10 sample injections, and at the end of the run (total of 13 injections) to assess instrument stability and analyte reproducibility. Following sample analysis a QC sample dilution series (1:2, 1:4, 1:8) in the reconstitution solvent mixture was performed and followed by extraction blank sample. This strategy is also illustrated in Figure 1.

Data Extraction. Collected data were subjected to peak-picking and grouping using MarkerLynx XS (Waters Inc., v4.1) software. Parameters used are presented in Supporting Information Table S-2. Values were reported as height of intensity peaks. Samples were normalized to total intensity. Values were multiplied by 10 000 prior to statistical analyses. The dilution series was used here to remove peaks that did not respond to dilution. This was achieved by applying multivariate statistics to the QC samples and dilutions and removing features (variables) that did not vary in intensity according to the dilutions applied. Additionally, as the local anesthetic lidocaine is typically locally administered to patients prior to performing carotid endarterectomy, features attributed to lidocaine and its metabolite hydroxylidocaine were removed from further statistical analyses.

RP-UPLC-MS Analysis (Lipid Profiling) of Organic Extracts. *UPLC-MS Analysis.* The organic extracts of the tissue samples were reconstituted in 500 μL of the solvent mixture of water/ACN/isopropanol (ISP) (1:1:2) and transferred into Total Recovery vials (Waters Corp, USA), after centrifugation for 10 min at 5000*g* and 4 °C.

UPLC separation was conducted using an Acquity UPLC System (Waters Corp, USA). An Acquity UPLC CSH C18 2.1 × 100 mm, 1.7 μm, column (Waters Corp, USA) was used. Column temperature was set at 55 °C and flow rate at 0.4 mL/min. Mobile phase A consisted of ACN/water (60:40) and mobile phase B ISP/ACN (90:10). In both solutions ammonium formate was diluted to 10 mM and formic acid to 0.1%. The elution gradient was set as follows: 60−57% A (0.0−2.0 min), 57−50% A (2.0−2.1 min; curve 1), 50−46% A (2.1−12.0 min), 46−30% A (12.0−12.1 min; curve 1), 30−1% A (12.1−18 min), 1−60% A (18.0−18.1 min), 60% A (18.1−20.0 min). Injection volumes of 3 and 7 μL were used for positive and negative ionization modes, respectively. The autosampler was set at 4 °C. Mass spectrometry was performed using a Xevo G2 QTof (Waters MS Technologies, U.K.) with an electrospray ionization (ESI) source. Both MS and MS^E data scans were acquired. MS conditions can be found in Supporting Information. The same QC strategy was followed as for the HILIC-UPLC-MS analysis.

Data Extraction. After acquisition, data were centroided (m/z spectra peaks are automatically detected and their centroid is calculated based on the average m/z value and weighted by the intensity). This was followed by peak-picking and grouping using MarkerLynx XS (Waters Inc., v4.1) software. Parameters used are presented in Supporting Information Table S-2. Values were reported as peak intensity (area). Saturated peaks (as identified by the MassLynx software) were removed prior to total area normalization. Values were multiplied by 10 000 prior to statistical analyses.

Safety Considerations. MeOH and ACN are known to be harmful while DCM is harmful, considered toxic and may be carcinogenic. These solvents should be handled in a fume hood and avoid skin contact and inhalation. Safety considerations include obtaining appropriate training prior to using these solvents and using appropriate waste disposal since they can impose an environmental hazard.

DOI: 10.1021/ac503775m
Anal. Chem. 2015, 87, 4184−4193

Statistical Analysis. Multivariate data analysis (MVDA) was conducted using the SIMCA package (version 13.0.2.; Umetrics, Sweden). Principal components analysis (PCA) and orthogonal projection to latent structures discriminant analysis (OPLS-DA) were employed to examine UPLC-MS data in a multivariate setting. Prior to model fitting, features were subjected to Pareto scaling. Two-tailed t tests (assuming unequal variance), and coefficients of variation % (CV%) were calculated in Microsoft Office Excel 2007.

Metabolite Structural Assignment. For metabolite structure assignment, accurate m/z measurements of detected chromatographic peaks were first matched to metabolites from online MS databases (Metlin,[21] HMDB,[22] and Lipidmaps[23]). After an assessment of retention time and isotopic pattern, tandem MS (UPLC-MS^E and UPLC-MS/MS) fragmentation pattern was employed for further structural elucidation. Assignment according to fragmentation pattern was dependent on the ability to obtain spectra with adequate signal. Further, an authentic standard of the metabolite was run using identical UPLC-MS/MS conditions and the detected m/z was matched to (1) the retention time and, where possible, (2) the MS/MS spectrum obtained from the sample under identical experimental conditions. Matching to an authentic standard was dependent on commercial availability and was pursued for small molecules. In Supporting Information Table S-3, the level of assignment is provided for every annotation according to the following criteria: (1) Accurate mass matched to database indicating tentative assignment, (2) accurate mass matched to database and tandem MS spectrum matched to in silico fragmentation pattern, (3) tandem MS spectrum matched to database or literature, (4) retention time (RT) matched to standard compound, and (5) MS/MS spectrum matched to standard compound.

Pathway Mapping. For pathway mapping the KEGG database[24] (July 10, 2014 Update) and Ingenuity Pathway Analysis software (IPA) (Build, 313398M; Version, 18841524) were used.

Venn Diagrams. Venn diagrams were constructed using online available software (http://bioinfogp.cnb.csic.es/tools/venny/index.html).

RESULTS AND DISCUSSION

In the present study a simple pipeline (Figure 1) was designed in order to expand metabolome coverage and analyte reproducibility in tissue samples. The well recognized limitations of a single chromatographic mode to deliver broad metabolite coverage motivated the utilization of both RP- and HILIC-UPLC-MS columns. Chromatographic conditions were optimized to enhance separation of the detected compounds. Further, we provide an experimental setup for untargeted metabonomic settings, in order to support efficient metabolite assignment and increase confidence of potential biomarkers. In addition, analytical and RT reproducibility were also assessed using the 13 QC samples, followed by an evaluation of metabolome coverage. Pathway mapping methodologies were used to assess the extent of coverage of biological pathways and functions. Lastly, an application to cardiovascular disease related samples (n = 120) illustrated the ability of the pipeline to deliver disease related metabolic profiles. The successful analysis of a biologically active tissue with high lipid content as a matrix ensures the robustness and wide applicability of the developed methodologies.

Tissue Lysis and Extraction. Previous work from our laboratory[18] has demonstrated that consecutive extractions, that is, polar extraction followed by organic (rather than a bilayer format) delivers higher reproducibility in tissue extraction. Thus, consecutive extraction was applied here (Figure 1). Moreover, the use of bead-beating for tissue lysis has also been demonstrated to be superior to other lysis methodologies, such as manual grinding and electric homogenizer.[6] For tissue lysis, bead-beating was used with preference for steel beads, as from preliminary studies, it was observed that steel beads were superior in homogenizing the calcified parts of the plaque tissues, as compared to zirconium beads.

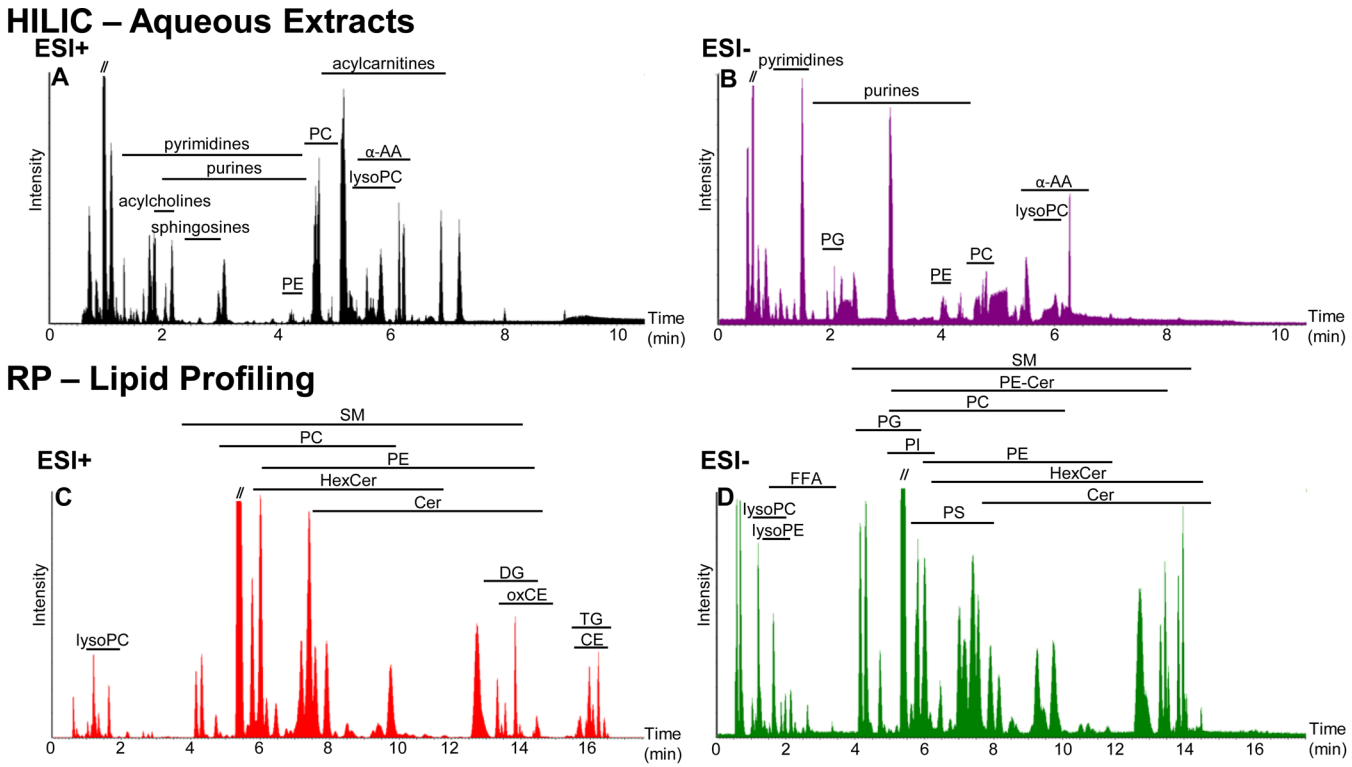
Optimization of HILIC-UPLC-MS Analysis of Aqueous Extracts. For HILIC chromatographic conditions, no pH mobile phase adjustment was chosen, opting to keep the method unbiased in terms of basic or acidic compounds. By using arterial tissue extracts, tests with longer isocratic initial conditions did not result in improved chromatographic separation.

The ability of the gradient program to adequately re-equilibrate the system was also assessed. The re-equilibration part of the gradient program was extended to accommodate adequate stability prior to each injection, while ensuring analysis did not exceed 23 min per sample making it attractive for application in large-scale studies.

The effect of the reconstitution solvent on chromatographic performance was also assessed using different percentages of ACN. Specifically, three different proportions of ACN were used (50%, 75% and 95% ACN) to reconstitute pooled tissue extracts. This test aimed to explore the hypothesis that low percentages of ACN in the reconstitution solvent may disturb the major partitioning mechanism between the aqueous coating formed in the HILIC column on the surface of the stationary phase of the HILIC column and the analytes.[25] It was found that when ACN was present at <95%, a number of chromatographic peaks were split, resulting in two peaks -one eluting at the beginning of the gradient program, and another at a higher retention time (Supporting Information Figure S-1). It was further observed that low percentages of ACN could also distort the peak shape (Supporting Information Figure S-1). These observations were further verified using standard compounds dissolved in different proportions of ACN (Supporting Information Figure S-2). Another advantage of using 95% ACN was that it made the column less exposed to precipitating proteins, as proteins would have been precipitated prior to injection. Preventing proteins from precipitating and eventually blocking the column can be a challenge in HILIC analysis, due to the highly organic initial solvent proportions of the mobile phase system, causing the proteins dissolved in the sample to precipitate as they enter the column. With the described methodology it was observed that no increase in system pressure was detected after 300 injections of tissue extracts (Supporting Information Figure S-3), indicating that there was likely no buildup of proteins in the column. However, using a highly organic solution for reconstitution could influence the desolvation of polar compounds. It was observed that the sensitivity for some of the polar compounds (detected toward the end of the chromatogram) was affected, although these compounds were still detectable (Supporting Information Figure S-4).

Using the HILIC method it was noted that low levels of lipid classes, such as triacylglycerol (TG), diacylglycerol (DG), and cholesteryl ester (CE) eluted with the solvent front

DOI: 10.1021/ac503775m
Anal. Chem. 2015, 87, 4184−4193



α-AA:alpha-amino acid, CE:cholesteryl ester, Cer:ceramide, DG:diacylglycerol, FFA:free fatty acid, HexCer:hexosylceramide, oxCE:oxidized cholesteryl ester, PC:phosphatidylcholine, PE:phosphatidylethanolamine, PE-Cer:phosphatidylethanolamine-ceramide, PG:phosphatidylglycerol, PI:phosphatidylinositol, PS:phosphatidylserine, SM:sphingomyelin, TG:triacylglycerol.

Figure 2. Base peak intensity chromatograms demonstrating the elution times of some of the major metabolite classes detected in (A) positive electrospray ionization mode (ESI+), (B) negative electrospray ionization mode (ESI−) from the HILIC-UPLC-MS analysis of the aqueous extracts; (C) ESI+ and (D) ESI− from the lipid profiling analysis of the organic extracts.

(Supporting Information Figure S-5), while several phospholipids eluted in a narrow range in the middle part of the chromatogram (Figure 2C and D). This characteristic demonstrates an advantage of HILIC over RP chromatography since lipid build-up in the column is a known weakness of RP. Additionally, in RP methods, lipid elution requires longer gradients and can span a wide range of the chromatographic run, which could affect the detection of smaller molecules. Lipids are highly abundant in biological matrices and are expected in higher concentrations in tissue relevant to cardiovascular disease, such as atherosclerotic plaques. Most importantly, it became clear from early optimization experiments that the RP column could adequately retain polar compounds (Supporting Information Figure S-6). The capabilities of the HILIC method were also assessed using adipose tissue aqueous extracts (Supporting Information Figure S-7).

Optimization of Lipid Profiling via RP-UPLC-MS. Optimization of the chromatographic separation stage of the lipid profiling methodology was initially conducted using lipid standard mixtures. A charged surface hybrid (CSH) column was used as it improved the separation of lipid moieties between classes as well as within each class, spanning a wide range of lipophilicity. Interclass separation is important as it can assist in structural assignment of the lipids. Moreover, intraclass separation can reduce ion suppression induced by coeluting lipids of the same class. In fact, the use of HILIC chromatography for lipid profiling was abandoned due to the insufficient intraclass lipid separation (Figure 2) and inadequate partitioning of neutral lipids (Supporting Information Figure S-5).

The CSH column also demonstrated short re-equilibration times between injections, reducing analytical run length and increasing throughput. Lastly, the use of ISP in the mobile phase and high column temperature provided the ability to elute highly lipophilic compounds such as TGs and CEs that would otherwise be strongly partitioned into the stationary phase, thus reducing chromatographic performance and causing carry over phenomena. We did not observe significant carry over effects of any molecular class (Supporting Information Figure S-8).

The separation ability of the CSH column was further assessed by using isomers and cis- and trans- stereoisomers of the phosphatidylglycerol (PG) lipid species. Firstly, separation the cis- compounds of the PG(18:1(9Z)/18:1(9Z)) against PG(18:0(9Z)/18:2(9Z)) was demonstrated. Additionally, separation of cis-PG(18:1(9Z)/18:1(9Z)) and trans-PG(18:1-(9E)/18:1(9E)) was achieved (Supporting Information Figure S-9). Lastly, the applicability of this column for analysis of real biological samples was assessed by testing chromatographic performance using rat, murine and human plasma organic extracts and bovine liver, human arterial and adipose tissue organic extracts (Figure 2 and Supporting Information Figure S-7).

UPLC-MS Experimental Setup. The use of QC samples to assess reproducibility and instrument performance and stability is preferred in metabolic profiling studies.[10,18,20] We herein employed the QC format and demonstrated the usefulness of this approach to assess reproducibility using both MVDA (Figure 3) and univariate statistics (Table 1 and Figure 4). We also included in the analysis several extraction blank samples (Figure 1). The sample preparation procedure and solvent

DOI: 10.1021/ac503775m
Anal. Chem. 2015, 87, 4184−4193

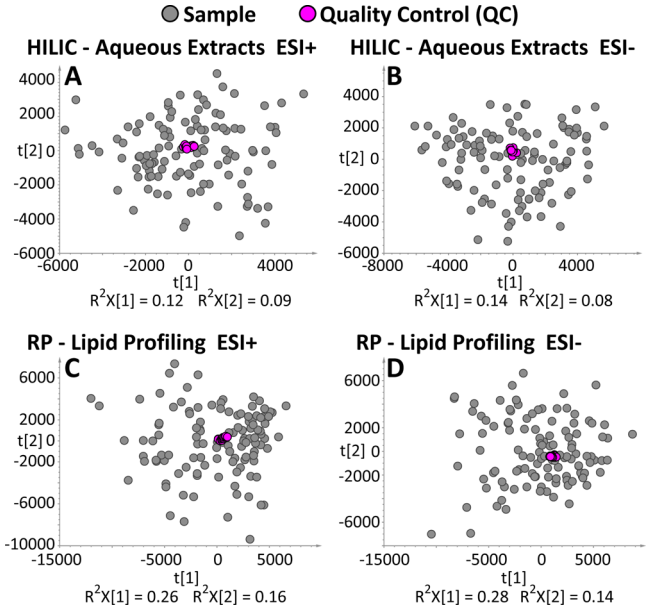


Figure 3. Scores plots of principal components analysis (PCA) models generated from the UPLC-MS analyses of tissue extracts, demonstrating the biological samples and the quality control samples (QC). The tight grouping of QCs as compared to the distribution of the biological samples highlights the reproducibility and instrument stability through the run. (A) Positive electrospray ionization mode (ESI+) and (B) negative electrospray ionization mode (ESI−) from the HILIC-UPLC-MS analysis of the aqueous extracts. (C) ESI+ and (D) ESI− from the lipid profiling analysis of the organic extracts.

impurities can introduce ions unrelated to the analyzed matrix. By including extraction blank samples we were able to easily identify and remove features present in the blank samples but not the matrix of interest, using multivariate statistics (Supporting Information Figure S-10). Verifying that the origin of the ion of interest is matrix-related can minimize the possibility of artifacts.

We also included dilutions of the pooled (QC) samples to assess the ability of each ion to respond to fluctuations in concentration (Figure 1). An initial assessment of the response to dilutions was also performed—using only the 2-fold dilution - prior to initiating analysis. The injection volume was adjusted to optimize intensity response to dilution. The dilution series—injected at the end of the run—assisted in identifying features that were not responding to the subsequent dilutions (Supporting Information Figure S-11). Features demonstrating an erratic behavior were observed in the HILIC analysis. This was predominantly the case with lipid moieties of the same class. We hypothesized that this may be due to the coelution of lipids and a resulting suppression of their signals as a result of simultaneous ionization in the ESI source. These features were removed from further analysis (Supporting Information Figure S-11). In literature, there are suggestions on how the dilution series can assist in increasing biomarker confidence.[26]

Collision-induced dissociation (CID) experiments were applied to the QC sample in order to obtain structural information for hundreds of detected ions. This was conducted during the conditioning step (Figure 1). Obtaining fragmentation patterns of molecules of interest during the original profiling run can save the analyst time and frustration, since variations in retention time and instrument performance can occur when returning to the instrument for tandem-MS experiments after long periods. Both unbiased UPLC-MS/MS acquisition (DDA, data dependent acquisition), as well as acquisition with no precursor ion selection (UPLC-MS^E) were employed. The DDA MS/MS experiments can provide fragments specifically attributed to the precursor ion. On the other hand, the MS^E approach,[27] in combination with chromatographic separation, proved very useful for structural assignment of metabolite classes, as it can provide the analyst with retention times of characteristic secondary ions upon fragmentation, known to be specific for metabolite classes.

Figure 5 demonstrates how a characteristic fragment of carnitines (Figure 5B) assisted in identifying several molecules from this group of compounds. Lastly, we use the same gradient program for analyzing in positive and negative ionization mode as it can provide valuable complementary information aiding toward an easier metabolite structural assignment.

Reproducibility. *Chromatographic Reproducibility.* It has been reported that chromatographic reproducibility can be a limiting factor especially with HILIC columns,[4,15] ultimately affecting data quality. Peak retention time (RT) reproducibility was evaluated using the 13 QC samples interspersed throughout the run to calculate the CV%. The most abundant ions per RT window in the QC samples were used for the evaluation (Supporting Information Figure S-12). The peaks of these ions demonstrated an RT CV % values of <0.5%.

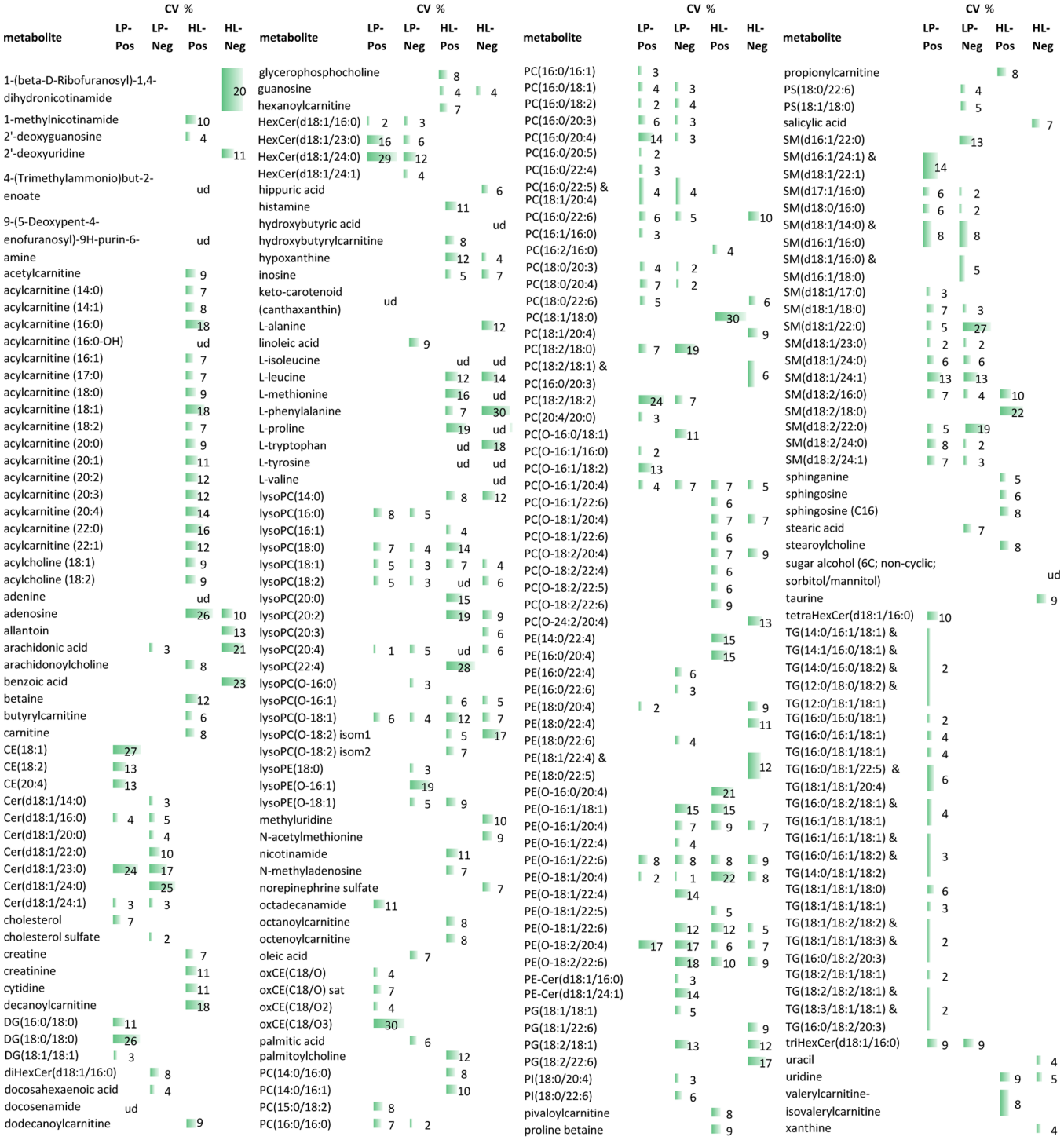
Analytical Reproducibility. Using the 13 injections of the QC samples, two statistical approaches were applied to evaluate the analytical reproducibility of the two methodologies in terms of detected intensities: (1) using PCA as a MVDA method to test reproducibility and (2) by calculating the CV% of detected peaks as a univariate way. The PCA demonstrated a tight grouping of the QC samples (Figure 3), indicative of their high reproducibility across the run and confirming that the variation observed between the samples (n = 120) was nonsystematic, but biologically related. Additionally, >80% of the abundant features demonstrated CV values of <30% (Table 1). Abundant features were considered to be the features for which chromatographic peaks could be detected in all of the 13 QC

Table 1. Reproducibility Assessment of Detected Features

	detected features	abundant features	abundant reproducible features (% to all abundant)	abundant features with CV %[a] < 10	abundant features with CV %[a] 10−20	abundant features with CV %[a] 20−30	abundant features with CV %[a] > 30
lipid profiling −positive mode	9884	2582	2095 (81)	1080	702	313	487
lipid profiling −negative mode	4804	1410	1164 (83)	619	338	207	246
HILIC−positive mode	5479	1378	1123 (81)	403	464	256	255
HILIC−negative mode	3483	1240	1096 (88)	344	525	227	144

[a]CV % = Coefficient of variation %

DOI: 10.1021/ac503775m
Anal. Chem. 2015, 87, 4184−4193



CV: Coefficient of variation; LP-Pos: RP-Lipid Profiling-Positive Mode; LP-Neg: RP-Lipid Profiling-Negative Mode; HILIC-Pos: HILIC-Aqueous Extracts-Positive Mode; HILIC-Neg: HILIC-Aqueous Extracts-Negative Mode; ud: unable to determine
CE: cholesteryl ester; Cer: ceramide; DG: diacylglycerol; HexCer: hexosylceramide; oxCE: oxidized cholesteryl ester; PC:phosphatidylcholine; PE:phosphatidylethanolamine; PE-Cer: phosphatidylethanolamine-ceramide; PG: phosphatidylglycerol; PI: phosphatidylinositol; PS: phosphatidylserine; SM: sphingomyelin; TG:triacylglycerol

Figure 4. Structurally assigned metabolites with coefficient of variation (CV%) values for each analysis where they were detected. HL: HILIC-UPLC-MS. LP: Lipid profiling. Pos: Positive electrospray ionization mode. Neg: Negative electrospray ionization mode.

samples. It is important to exclude features for which the peak intensity did not pass the peak-picking intensity S/N threshold (or otherwise not detected) as they could skew the value of the CV % calculation. Table 1 summarizes the CV% reproducibility analysis of all employed methodologies.

The high retention time and analytical reproducibility of the detected ions demonstrate that the presented methodologies can deliver the robustness required by a metabolic profiling study.

Improvement in Metabolome Coverage by Combining HILIC and RP Methodologies. The scientific community is constantly striving to expand metabolome coverage. This forces analysts to combine different analytical platforms and methods in order to increase the number of robustly detected

DOI: 10.1021/ac503775m
Anal. Chem. 2015, 87, 4184−4193

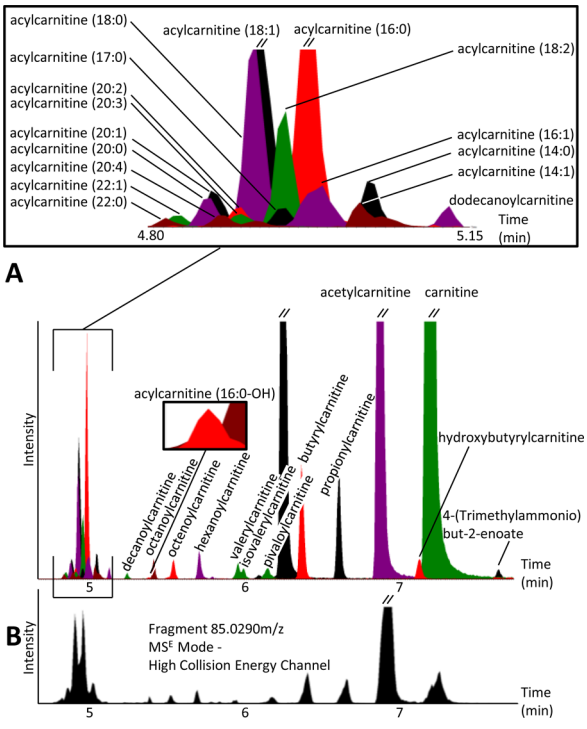


Figure 5. (A) Profile of extracted ion chromatograms (XIC) of carnitine, acylcarnitines, and other carnitine derivatives (inset shows expansion of the retention time window of 4.80−5.15 min) using the HILIC-UPLC-MS method on aqueous extracts in positive mode. Unlabeled peaks represent unrelated isobaric ions. (B) XIC of a characteristic acyl carnitine fragment obtained from MS^E analysis.

metabolites. Multiplatform analyses are particularly relevant when both polar and nonpolar metabolites require analysis. Here, two UPLC-MS chromatographic methods were combined to expand metabolome coverage resulting in a total of 226 unique metabolite structural assignments (Figure 4 and Supporting Information Table S-3), ranging from polar (e.g., amino acids, creatine, carnitines, purines, and pyrimidines) to highly lipophilic metabolites (e.g., TGs and CEs). The two methods demonstrated high complementarity, with only 20 intermethod metabolites detected (Figure 6). These metabolites corresponded to highly abundant lipid compounds with amphiphilic properties. With the lipid profiling method, 97 metabolites were uniquely assigned, with 29 commonly detected in both polarities (Figure 6). The HILIC method provided 109 unique metabolites with 18 detected by both of the HILIC polarity modes (Figure 6).

Several molecules were preferentially ionized in only one of the MS polarities. This was particularly obvious with compound classes such as TGs, CEs and acyl-carnitines preferentially ionizing in positive mode, while classes such as free fatty acids, phosphatidylinositols, phosphatidylserines and sulfated compounds preferentially ionized in negative mode. Additionally, metabolites detected by both polarities can frequently provide essential information to guide structural assignment. In fact, for several species tandem MS spectra from both polarities were required for achieving structural assignment, thus demonstrating the necessity of analyzing both intramethod polarities with the same chromatographic gradient. Overall, a substantial amount of extra information was obtained by using both polarity modes, resulting in expansion of metabolome coverage and assisting metabolite structure assignment.

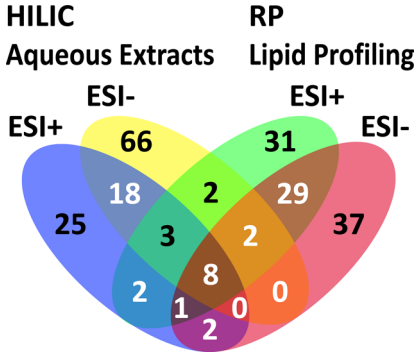


Figure 6. A Venn diagram demonstrating commonly and exclusively detected metabolites for the employed analyses and electrospray ionization modes. The vast number of metabolites robustly detected by only one of the methods demonstrates the complementarity of the HILIC-UPLC-MS and the lipid profiling methods when applied on the aqueous and organic extracts, respectively. The diagram is based only on structurally assigned metabolites. ESI+: positive electrospray ionization mode, ESI-: negative electrospray ionization mode.

It should be noted that the detected metabolite list (Figure 6 and Supporting Information Table S-3) does not represent all the metabolites that can be detected by the described experiments, nor exhausts the ability of these methods for detection of additional metabolites. Factors such as tissue specificity can affect the concentrations or presence of specific molecules. Moreover, several ions could not be structurally assigned due to ambiguity of their m/z, inadequate tandem MS information, or absence from online databases and are not further described here. At the same time, several administered drugs were also detected, such as the local anesthetics bupivacaine, lidocaine and their metabolites such as hydroxylidocaine. However, they are not further discussed as the study focuses predominantly on endogenous compounds.

Using the HILIC method, carnitine, carnitine derivatives, such as acylcarnitines incorporating fatty acyl chains (FAC) in the range of 22C to acetylcarnitine, were detected. The chromatographic profile of the structurally assigned carnitine derivatives and carnitine is presented in Figure 5. Additionally, nine α-amino acids (α-AA), and α-AA derivatives, such as *N*-acetylmethionine, purines and pyrimidines (Figure 2), as well as sphinganine and the 16C and 18C sphingosines, organic acids (benzoic and salicylic), nicotinamide and 1-methylnicotinamide, creatine and creatinine, betaine, glycerophosphate derivatives, sugars, acylcholines, could be indentified in the aqueous tissue extracts.

Using the lipid profiling method we detected phospholipids from the classes of: phosphatidylcholines, phosphatidylethanolamines, phosphatidylinositols, phosphatidylserines, phosphatidylglycerols, lysophosphatidylcholines, and lysophosphatidylethanolamines (Figure 2). Additionally, sphingolipids with different backbone lengths and degree of unsaturation were identified (d18:0, d18:1, d18:2, d16:1 and d17:1). Sphingolipids detected included ceramides, sphingomyelins, phosphatidylethanolamine-ceramides (PE-Cer), mono-, di-, tri-, and tetra-hexosylceramides. Free fatty acids were another lipid class detected using the negative mode of the lipid profiling method (Figure 2B). FFA in the range of 16−22C were detected, while cis- and trans- stereoisomers (elaidic and oleic acid; linoelaidic and linoleic acid) could be chromatographically resolved (Supporting Information Figure S-13). Glycerolipids were detected in the form of DGs and TGs and in a wide range of

DOI: 10.1021/ac503775m
Anal. Chem. 2015, 87, 4184−4193

incorporated FAC. Finally, sterol lipids such as cholesterol, cholesteryl esters, oxidized cholesteryl esters (Figure 2A) and cholesterol sulfate were identified. Lipid structural assignments would typically include elucidation of FAC length and degree of unsaturation. However, the double bond and sn- positions could not be elucidated using the MS approach employed here.

Finally, it should be highlighted that the untargeted format of the two methodologies is capable of identifying previously uncharacterized compounds, a critical advantage compared to targeted approaches. An example is the previously unreported forms of PE-Cers (PE-Cer(d18:1/16:0) and PE-Cer(d18:1/24:1)) (Figure 4 and Supporting Information Table S-3). Additionally, acylcholines, such as arachidonoylcholine and palmitoylcholine (Figure 4 and Table S-3), were not listed in the MS databases used.

Pathway Mapping. Out of the 229 assigned metabolites, 187 could be mapped using the KEGG[24] database and a KEGG ID could be assigned (Supporting Information Table S-3). A total of 100 KEGG IDs were considered unique. Using the KEGG mapper (http://www.genome.jp/kegg/mapper.html), a mapped overview of the primary human metabolic pathways demonstrated 49 mapped metabolites (Supporting Information Figure S-14). However, it was observed that some primary metabolic pathways were lacking representation. This was evident with the central carbohydrate metabolism, fatty acid metabolism (fatty acids <16C) and glycan metabolism. Coverage of these pathways will be subsequently pursued with appropriate platforms and methods, along with efforts for further unbiased expansion of the range of the detected and structurally assigned metabolites. Nonetheless, using the IPA software a wide range of canonical pathways, biological functions and disease pathways could be covered by the assigned metabolites. Up to 89 out of the 100 KEGG IDs imported in the IPA software were recognized. These metabolites could be mapped by IPA to 182 canonical pathways (Supporting Information Figure S-15 and Spreadsheet S-1). Most importantly, 60 unique metabolites could be associated and mapped to vast number of diseases and biological functions (Supporting Information Figure S-16 and Spreadsheet S-2). These included lipid accumulation and concentration, cancer, inflammation, cell death and survival, proliferation, apoptosis, necrosis, peroxisomal disorder and cell differentiation. Additionally, 36 assigned metabolites could be mapped being involved in 99 toxicity functions (Supporting Information Figure S-17 and Spreadsheet S-3), such as liver damage, cardiac damage, and renal failure.

Application to Cardiovascular Disease. According to the World Health Organization (WHO), CVD is the leading cause of mortality in the western world. Deaths associated with CVD are mainly related to atherosclerosis and hypertension. Here, we used diseased human tissue from the arterial tree to assess the ability of the RP lipid profiling and the HILIC polar phenotyping methods to deliver metabolic profiles relevant to tissue phenotype. Human tissues samples (n = 120) were used, obtained from 26 abdominal aortic aneurysm (AAA) tissue, 52 carotid stenosing plaque (CAR), 26 femoral stenosing plaques (FEM), and 16 from intimal thickening plaque-free tissue (INT). The tissue used is known to be biologically active with presence of several inflammatory factors contributing to the disease process. Additionally, plaque tissue incorporates high lipid content with macroscopically discrete substructure. These features (high lipid and protein content) of the chosen analyzed tissue demand robust pipeline and tools to avoid compromising optimal performance. The successful metabolic profiling of such tissue, as evidenced by the high reproducibility and number of metabolites detected, further demonstrates the robustness of our pipeline and developed methodologies.

Using OPLS-DA fitted models to explore differences in tissue metabolic phenotypes, separation of the four tissue groups was apparent as demonstrated by the cross-validated scores plots in Figure 7, along with models characteristics. The model characteristics presented well-fitted models with high predictive values (Q^2Y:0.44−0.52; Figure 7). Intriguing findings from this

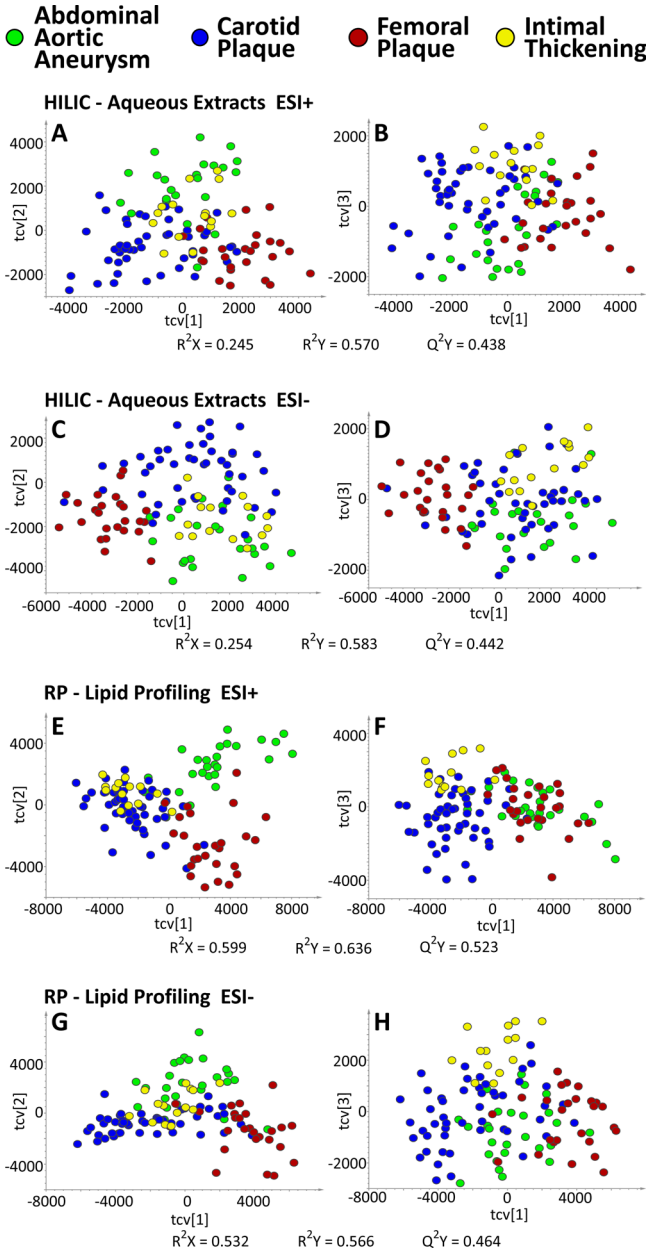


Figure 7. Cross-validated scores plots of orthogonal projection to latent structures−discriminant analysis (OPLS-DA), demonstrating separation between the UPLC-MS analyses of the tissue extracts from the four diseased groups. (A, B) Positive electrospray ionization mode (ESI+) and (C, D) negative electrospray ionization mode (ESI−) from the HILIC-UPLC-MS analysis of the aqueous extracts. (E, F) ESI + and (G, H) ESI− from the lipid profiling analysis of the organic extracts.

DOI: 10.1021/ac503775m
Anal. Chem. 2015, 87, 4184−4193

data set were (1) manifested differences between the two plaque groups CAR and FEM, since it demonstrates that plaques from different locations express distinct metabolic activity, and (2) the similarities observed by the FEM and AAA groups. Loadings plots showing the metabolites driving the separations between the sample groups can be found in Supporting Information Figure S-18. However, biological interpretation of these findings exceeds the scope of this paper and is not further discussed. Biological interpretation for the comparison between INT tissue and CAR and FEM plaques has been previously described.[19]

CONCLUSIONS

We demonstrate that a combination of RP- and HILIC-UPLC-MS methodologies delivers improved ability to achieve broad metabolome coverage of compounds with diverse physico-chemical properties. Additionally, these chromatographic methods are highly complementary, further illustrating the necessity for rationally combining RP and HILIC methods for analyzing organic and aqueous tissue extracts, respectively. Importantly, this pipeline can cover a wide range of metabolic pathways and biological functions, while their untargeted format allows for the detection of previously unknown metabolites. The implementation of DDA and MS^E acquisitions on pooled QC samples within the analytical run, permits simultaneous collection of metabolite structural information and reduces the time of the analytical pipeline. The described pipeline is generic and applicable to other tissue types. In future, the described methodologies will be complemented with additional methods and platforms to address gaps in metabolome coverage identified by the pathway mapping procedures.

ASSOCIATED CONTENT

Supporting Information

Further information on chemical materials, MS and data processing parameters, tables of patient demographics and assigned metabolites, and figures of: HILIC and RP optimization and performance; HILIC and RP comparison; additional applications on adipose tissue, liver tissue and plasma samples; further data processing; and pathways analysis. This material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Author

*Tel: +44 (0)20 7594 3220. Fax: +44 (0) 20 759 43226. E-mail: elaine.holmes@imperial.ac.uk.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

This study was funded by the Royal Society of Chemistry (RSC) (Grant number: P16557). PAV acknowledges the RSC for supporting his PhD studentship. Additional support was received by the National Institute for Health Research (NIHR) Biomedical Research Centre based at Imperial College Healthcare NHS Trust and Imperial College London. The views expressed are those of the authors and not necessarily those of the NHS, the NIHR or the Department of Health. EJW would like to acknowledge Waters Corporation for her funding.

REFERENCES

(1) Nicholson, J. K.; Lindon, J. C.; Holmes, E. *Xenobiotica* **1999**, *29*, 1181−1189.
(2) Waldram, A.; Holmes, E.; Wang, Y.; Rantalainen, M.; Wilson, I. D.; Tuohy, K. M.; McCartney, A. L.; Gibson, G. R.; Nicholson, J. K. *J. Proteome Res.* **2009**, *8*, 2361−2375.
(3) Ismail, N. A.; Posma, J. M.; Frost, G.; Holmes, E.; Garcia-Perez, I. *Electrophoresis* **2013**, *34*, 2776−2786.
(4) Ivanisevic, J.; Zhu, Z. J.; Plate, L.; Tautenhahn, R.; Chen, S.; O'Brien, P. J.; Johnson, C. H.; Marletta, M. A.; Patti, G. J.; Siuzdak, G. *Anal. Chem.* **2013**, *85*, 6876−6884.
(5) Saric, J.; Want, E. J.; Duthaler, U.; Lewis, M.; Keiser, J.; Shockcor, J. P.; Ross, G. A.; Nicholson, J. K.; Holmes, E.; Tavares, M. F. *Anal. Chem.* **2012**, *84*, 6963−6972.
(6) Geier, F. M.; Want, E. J.; Leroi, A. M.; Bundy, J. G. *Anal. Chem.* **2011**, *83*, 3730−3736.
(7) Naz, S.; Garcia, A.; Barbas, C. *Anal. Chem.* **2013**, *85*, 10941−10948.
(8) Beltran, A.; Suarez, M.; Rodriguez, M. A.; Vinaixa, M.; Samino, S.; Arola, L.; Correig, X.; Yanes, O. *Anal. Chem.* **2012**, *84*, 5838−5844.
(9) Gika, H. G.; Theodoridis, G. A.; Wilson, I. D. *J. Sep. Sci.* **2008**, *31*, 1598−1608.
(10) Spagou, K.; Wilson, I. D.; Masson, P.; Theodoridis, G.; Raikos, N.; Coen, M.; Holmes, E.; Lindon, J. C.; Plumb, R. S.; Nicholson, J. K.; Want, E. J. *Anal. Chem.* **2011**, *83*, 382−390.
(11) Spagou, K.; Tsoukali, H.; Raikos, N.; Gika, H.; Wilson, I. D.; Theodoridis, G. *J. Sep. Sci.* **2010**, *33*, 716−727.
(12) Want, E. J.; Wilson, I. D.; Gika, H.; Theodoridis, G.; Plumb, R. S.; Shockcor, J.; Holmes, E.; Nicholson, J. K. *Nat. Protoc.* **2010**, *5*, 1005−1018.
(13) Idborg, H.; Zamani, L.; Edlund, P. O.; Schuppe-Koistinen, I.; Jacobsson, S. P. *J. Chromatogr. B: Anal. Technol. Biomed. Life Sci.* **2005**, *828*, 14−20.
(14) Cubbon, S.; Antonio, C.; Wilson, J.; Thomas-Oates, J. *Mass Spectrom Rev.* **2010**, *29*, 671−684.
(15) Gray, N.; Heaton, J.; Musenga, A.; Cowan, D. A.; Plumb, R. S.; Smith, N. W. *J. Chromatogr. A* **2013**, *1289*, 37−46.
(16) Schmidt, C. *J. Natl. Cancer Inst.* **2004**, *96*, 732−734.
(17) Mirnezami, R.; Spagou, K.; Vorkas, P. A.; Lewis, M. R.; Kinross, J.; Want, E.; Shion, H.; Goldin, R. D.; Darzi, A.; Takats, Z.; Holmes, E.; Cloarec, O.; Nicholson, J. K. *Mol. Oncol.* **2014**, *8*, 39−49.
(18) Masson, P.; Alves, A. C.; Ebbels, T. M.; Nicholson, J. K.; Want, E. J. *Anal. Chem.* **2010**, *82*, 7779−7786.
(19) Vorkas, P. A.; Shalhoub, J.; Isaac, G.; Want, E. J.; Nicholson, J. K.; Holmes, E.; Davies, A. H. *J. Proteome Res.* **2015**, *14*, 1389−1399.
(20) Gika, H. G.; Theodoridis, G. A.; Wingate, J. E.; Wilson, I. D. *J. Proteome Res.* **2007**, *6*, 3291−3303.
(21) Smith, C. A.; O'Maille, G.; Want, E. J.; Qin, C.; Trauger, S. A.; Brandon, T. R.; Custodio, D. E.; Abagyan, R.; Siuzdak, G. *Ther. Drug Monit.* **2005**, *27*, 747−751.
(22) Wishart, D. S.; Knox, C.; Guo, A. C.; Eisner, R.; Young, N.; Gautam, B.; Hau, D. D.; Psychogios, N.; Dong, E.; Bouatra, S.; Mandal, R.; Sinelnikov, I.; Xia, J.; Jia, L.; Cruz, J. A.; Lim, E.; Sobsey, C. A.; Shrivastava, S.; Huang, P.; Liu, P.; Fang, L.; Peng, J.; Fradette, R.; Cheng, D.; Tzur, D.; Clements, M.; Lewis, A.; De Souza, A.; Zuniga, A.; Dawe, M.; Xiong, Y.; Clive, D.; Greiner, R.; Nazyrova, A.; Shaykhutdinov, R.; Li, L.; Vogel, H. J.; Forsythe, I. *Nucleic Acids Res.* **2009**, *37*, D603−610.
(23) Fahy, E.; Sud, M.; Cotter, D.; Subramaniam, S. *Nucleic Acids Res.* **2007**, *35*, W606−612.
(24) Kanehisa, M.; Goto, S. *Nucleic Acids Res.* **2000**, *28*, 27−30.
(25) Hao, Z.; Xiao, B.; Weng, N. *J. Sep. Sci.* **2008**, *31*, 1449−1464.
(26) Eliasson, M.; Rannar, S.; Madsen, R.; Donten, M. A.; Marsden-Edwards, E.; Moritz, T.; Shockcor, J. P.; Johansson, E.; Trygg, J. *Anal. Chem.* **2012**, *84*, 6869−6876.
(27) Plumb, R. S.; Rainville, P. D.; Potts, W. B., 3rd; Johnson, K. A.; Gika, E.; Wilson, I. D. *J. Proteome Res.* **2009**, *8*, 2495−2500.

DOI: 10.1021/ac503775m
Anal. Chem. 2015, 87, 4184−4193