Analyst

Accepted Manuscript

This is an *Accepted Manuscript*, which has been through the Royal Society of Chemistry peer review process and has been accepted for publication.

Accepted Manuscripts are published online shortly after acceptance, before technical editing, formatting and proof reading. Using this free service, authors can make their results available to the community, in citable form, before we publish the edited article. We will replace this *Accepted Manuscript* with the edited and formatted *Advance Article* as soon as it is available.

You can find more information about *Accepted Manuscripts* in the Information for Authors.

Please note that technical editing may introduce minor changes to the text and/or graphics, which may alter content. The journal's standard Terms & Conditions and the Ethical guidelines still apply. In no event shall the Royal Society of Chemistry be held responsible for any errors or omissions in this *Accepted Manuscript* or any consequences arising from the use of any information it contains.

ROYAL SOCIETY OF CHEMISTRY

www.rsc.org/analyst

Optimization of Metabolite Extraction of Human Vein Tissue for Ultra Performance Liquid Chromatography-Mass Spectrometry and Nuclear Magnetic Resonance-Based Untargeted Metabolic Profiling

Muzaffar A Anwar, Panagiotis A Vorkas[†], Jia V. Li[δ†], Joseph Shalhoub*, Elizabeth J Want[†], Alun H Davies*, Elaine Holmes[δ†]*

Academic Section of Vascular Surgery*

Section of Computational and Systems Medicine[†]

Department of Surgery and Cancer

Faculty of Medicine

Centre for Digestive and Gut Health, Institute of Global Health Innovation[δ]

Imperial College London

Corresponding Author:

Professor Elaine Holmes

Address: Room 661, 6th floor, Sir Alexander Fleming Building, Imperial College London, South Kensington Campus, London. SW7 2AZ, UK

Phone: +44 (0) 20 7594 3220

E-mail: Elaine.holmes@imperial.ac.uk

Key Words: Human Vein, Varicose Veins, Metabolomics, Metabonomics, Metabolic Profiling, Metabolic phenotyping, Lipidomics, Metabolite Extraction, NMR, UPLC-MS

Analyst Accepted Manuscript

ABSTRACT

Human vein tissue is an important matrix to examine when investigating vascular diseases with respect to understanding underlying disease mechanisms. Here, we report the development of an extraction protocol for multi-platform metabolic profiling of human vein tissue. For the first stage of the optimization, two different ratios of methanol/water and 5 organic solvents – namely dichloromethane, chloroform, isopropanol, hexane, methyl *tert*-butyl ether (MTBE) solutions with methanol were tested for polar and organic compound extraction, respectively. The extraction output was assessed using ^{1}H Nuclear Magnetic Resonance (NMR) spectroscopy and a panel of Ultra Performance Liquid Chromatography-Mass Spectrometry (UPLC-MS) methodologies. On the basis of the reproducibility of extraction replicates and metabolic coverage, the optimal aqueous (methanol/water) and organic (MTBE/methanol) solvents identified from the first stage were used in a sequential approach for metabolite extraction, altering the order of solvent-mixture addition. The combination of organic metabolite extraction with MTBE/methanol (3:1) followed by extraction of polar compounds with methanol/water (1:1) was shown to be the best method in terms of reproducibility and number of signals detected for extracting metabolites from human vein tissue and could be used as a single extraction procedure to serve both NMR and UPLC-MS analyses. Molecular classes such as triacylglycerols, phosphatidylcholines, phosphatidylethanolamines, sphingolipids, purines, and pyrimidines were reproducibly extracted. This study enabled an optimal extraction protocol for robust and more comprehensive metabolome coverage for human vein tissue. Metabolic profiles and pathophysiological processes affecting human vein tissue can resemble those affecting several tissue types and hence the extraction method developed in this study can be generically applied.

Analyst Accepted Manuscript

Introduction

Metabolic profiling approaches to characterizing biological fluids and tissues involve application of methods such as ultra-performance liquid chromatography-mass spectrometry (UPLC-MS) and nuclear magnetic resonance spectroscopy (NMR) to provide comprehensive metabolic phenotypes of individuals[1]. With the aid of chemometric tools, key information relating to the influential metabolites in the context of disease pathogenesis and diagnosis can be recovered from complex spectral datasets[2]. Comprehensive coverage of the metabolic phenotype by NMR and UPLC-MS relies on efficient recovery of polar and organic metabolites from biological tissues, for which a number of extraction protocols have been developed[3, 4]. Parallel use of these two analytical platforms on extracted metabolites can increase metabolome coverage and hence provide a better tool for detecting pathway dysregulation and identifying disease diagnostic markers.

Each tissue or biofluid has different properties and therefore extraction protocols can be optimized depending on the structural and chemical properties of the tissue. Veins distributed extensively in the human body serve as a blood reservoir. Furthermore, veins are used as homograft conduits for cardiac and limb arterial bypass surgery[5]. Morphologically, vein walls are divided into three layers; intima, media and adventitia. Within these three layers there are three main cells types, which are endothelial cells in intima, smooth muscle cells in media and fibroblasts in adventitia. In addition, the adventitial layer contains mainly type III collagen and elastin, which provide elasticity to the vein wall. Venous tissue can be affected by pathological conditions such as varicose veins and venous thrombosis, involving processes such as inflammation [6, 7] along with recruited factors and alteration of the metabolic context. Under these pathological conditions processes in vein walls such as inflammation and muscular hypertrophy or hypotrophy and intimal hyperplasia are affected[6]. Under arterial pressure, vein conduits used in bypass surgery show intimal hyperplasia and changes akin to

Analyst Accepted Manuscript

atherosclerosis observed in carotid, coronary and other peripheral arteries[8, 9]. Therefore, the use of vein tissue and diseased vein tissue for optimization of the extraction procedure can be considered generic for a variety of tissue types and diseases. Applicability of this extraction methodology to additional tissue types is also demonstrated. Recently, we studied the metabolic signature of human varicose vein disease using magic angle spinning NMR and identified differential metabolites of potential significance in characterizing the pathology including lactate, creatine and *myo*-inositol[10]. This has prompted further metabolic profiling studies to more comprehensively characterize the metabolic signature of human blood vessel tissue extracts by both NMR and UPLC-MS. However, an extraction method for blood vessels has not yet been assessed and optimized. This study aimed to develop and optimize a tissue extraction method, which will be largely valid to other tissues that are affected by similar pathological processes (inflammation, intimal hyperplasia, hypoxia, cell death) as observed in veins. Various studies report different solvents for the extraction of polar metabolites from tissues. For example, a mixture of ethanol and phosphate buffer has also been shown to demonstrate adequate reproducibility for LC-MS-based profiling of brain tissue[11], while methanol/water (v:v, 4:1) has been used for LC-MS-based metabolite recovery in wide range of human tissue including muscle, adrenal gland, colon, lung, pancreas, small intestine, spleen, stomach, prostate and kidney[12]. Lin *et al* found that for NMR analysis of liver extracts, the mixture of methanol/chloroform and water was considered the best combination in terms of metabolic yield and reproducibility[13]. In contrast, Masson *et al.* showed that the optimal protocol for metabolic profiling of liver extracts by UPLC-MS was methanol/water (v:v, 1:1) followed by an organic extraction with dichloromethane/water (v:v, 3:1)[14]. Bligh and Dyer described a rapid and simple method for lipid extraction from biological material using chloroform/methanol in 1959[15]. Since then, metabolite extraction using chloroform/methanol has been considered as the gold standard protocol[16, 17], although

Analyst Accepted Manuscript

chloroform carries health and environmental hazards. In literature, a range of solvents including chloroform[16], dichloromethane (DCM), methyl *tert*-butyl ether (MTBE)[18], hexane and isopropanol (ISP) have been used for organic metabolite extraction[15, 19, 20]. DCM/methanol has been found to be comparable to chloroform/methanol in terms of efficiency of extraction, whilst being less toxic[21]. Various studies have focused on recovery of lipid metabolites from different biological fluids and tissues, including human blood[22, 23], feces[24], colonic tissue[25], mouse brain and different bacterial strains[19]. Le Belle *et al* reported that methanol/chloroform/water extraction was superior to perchloric acid as a solvent for NMR-based analysis of aqueous extracts from rat brain tissue[26]. Moreover, Want *et al* demonstrated that methanol-based extraction methodologies precipitate proteins from serum and hence improve the chromatographic performance when differentiating signals from metabolites eluting at a similar retention time[27, 28]. Likewise, Masson *et al* and Geier *et al* found methanol/water (v:v, 1:1) and methanol/water (v:v, 4:1) to be the most efficient for aqueous metabolite extraction for analysis of rat liver and nematodes, respectively[14, 29].

For venous disease, both lipids and small polar metabolites have been shown to be either mechanistically important or to have potential as biomarkers of disease presence or stage[10]. Thus, the optimal extraction procedure should be appropriate for both the polar and the hydrophobic components of tissue. There are two broad approaches employed for metabolite extraction when wishing to capture both the polar and organic components, namely the bilayer and consecutive approaches. The bilayer approach involves simultaneous extraction of polar and organic metabolites using a combination of water/methanol/chloroform; this result in two layers separated by a protein pellet[13, 14]. The consecutive approach comprises of an aqueous extraction followed by the organic extraction or *vice versa*[30]. It has been shown that the consecutive approach has the advantage over the bilayer approach in terms of

Analyst Accepted Manuscript

reproducibility and metabolite yields for both liver extracts and *Caenorhabditis elegans* tissue extracts[14, 29].

The aim of the current study was to identify the best solvent system (aqueous and organic) and subsequently to assess the influence of the sequence of solvent use on the robustness, reproducibility, recovery and metabolite coverage for both aqueous and organic phases with respect to phenotyping venous pathologies across two metabolic profiling platforms (NMR and UPLC-MS).

Experimental

Vein tissue collection and preparation

Vein tissue was collected from patients who underwent varicose vein surgery (research ethics committee approval RREC 3092). A total of approximately 10.5 grams of great saphenous vein tissue was collected from 12 patients, with the purpose of preparing a homogenate mixture. Human vein tissue was snap frozen in liquid nitrogen and stored at -80 °C. All the frozen tissue was combined in a 15 cm mortar (VWR, UK), immersed in liquid nitrogen and mixed using a pestle and mortar in a class II biological cabinet. The frozen homogenate was then further ground into powder using a cryogenic impact mill (freezer mill 6870, SPEX, Stanmore, UK)[31] with a cooling step (3 min) and a grinding cycle (2 min, at 10 Hz). A total of 70 tissue aliquots, each weighing 145 +/- 5 mg, were obtained. Each group consisted of 10 tissue aliquots and was treated with a solvent system. Each tissue aliquot was used to produce 5 aliquots of extracts each corresponding to ~25mg of tissue. A total of 7 solvents systems were used: 2 for aqueous extraction and 5 for organic extraction. The study comprises of two stages which are outlined Figure 1 a and b.

Analyst Accepted Manuscript

Methodology of aqueous phase extraction

Samples were kept on dry ice throughout the procedure. Details of chemicals used in the study are given in supporting information section. Extraction of aqueous metabolites was performed by adding 1.5 mL of methanol/water (v:v, 1:3) or methanol/water (v:v, 1:1) in each 2 mL microtube (VWR, UK) containing tissue sample (145 ±5 mg) and zirconium beads with a diameter of 1 mm (BioSpec, USA). A blank control sample, containing only solvent and beads, was also prepared for each group and run in parallel with the tissue samples. Samples were loaded onto a bead beater (Precellys 24, Bertin Technologies) and a homogenization cycle, consisting of 40 s shaking at 6500 Hz followed by 5 min cooling on dry ice, was repeated 4 times to maximize dissolution of the powder. Samples were centrifuged (Eppendorf 5417R) at 17949 x *g* for 20 min at 4 °C and the supernatant was taken into an Eppendorf tube. A total of 1.25 mL of supernatant was obtained from each sample and further divided into 5 x 250 µL aliquots (corresponding to ~25mg of tissue per aliquot). For each of these aliquots the methanol concentration was increased to 75%, in order to improve protein precipitation and prevent column degradation. This was followed by 1 min of vortex and centrifugation at 17949 x *g* for 20 min at 4 °C. The supernatant from each aliquot was then transferred into a new Eppendorf tube. Samples were dried in a speed vacuum for 10 hours at 30 °C and stored in a -40 °C freezer pending NMR and UPLC-MS analysis.

Methodology of organic phase extraction

Extraction of organic metabolites was performed by adding 1.5 mL of mixed organic solvent (chloroform, DCM, hexane, ISP or MTBE)/methanol (v:v, 3:1) to each sample in a microtube with zirconium beads for extraction on a bead beater. ISP was also added in the hexane/methanol solvent mixture to ensure homogeneity; the final proportions were

Analyst Accepted Manuscript

hexane/methanol/ISP (v:v:v, 13:5:2). A blank sample was also prepared and run in parallel to detect contaminants introduced during extraction. The remaining steps involving bead-beating samples and centrifugation were as described in the aforementioned aqueous metabolite extraction. From each bead beating tube, 200 µL aliquots (corresponding to ~25mg of tissue per aliquot) were transferred into 5 glass vials (Fisher, UK). Samples were left overnight in a fume hood at room temperature to allow solvent evaporation, and then stored in a -40 °C freezer until analysis.

Determination of the optimal order of consecutive extraction

For the second stage of the experiment, the optimal aqueous and organic solvents identified from stage 1 were used. A total of 3 g of the aforementioned human vein tissue powder was weighed and divided into 20 samples. Ten aliquots underwent consecutive aqueous extraction followed by organic extraction (C-A-O group), whereas the other ten samples were extracted consecutively by organic solvent followed by the aqueous extraction (C-O-A group). See Figure 1 b.

The extraction procedure was performed by addition of 1.5 mL of the first solvent system to each 2 mL microtube (VWR, UK) containing the tissue sample and 1 mm zirconium beads (BioSpec, USA). This was followed by bead beating and centrifugation (using the same protocol detailed in stage 1). For the aqueous extraction a total of 5 aliquots were obtained from each sample, each containing 250 µL of aqueous supernatant. Aqueous extracts were dried in a speed vacuum for 10 hours at 30 °C and frozen at -40 °C.

Following decanting of aqueous extracts, 1.5 mL of the chosen organic solvent mixture was added to the microtube and loaded onto bead beater. The same bead beating and centrifugation protocol as described in the preceding paragraphs was performed except for bead beating cycles, which were reduced to 2. A total of 1.2 mL supernatant from each

Analyst Accepted Manuscript

sample was transferred into 5 glass vials, each containing 200 µL. Samples were dried in vacuum hood overnight at room temperature and then frozen at -40 °C. For the C-O-A group of samples, the organic extraction was performed first, and with the same protocols as detailed above.

^{1}H-NMR spectroscopic analysis of aqueous and organic extracts

A detailed protocol for preparation of aqueous and organic extracts for NMR analysis is given in the supporting information section. Aqueous and organic vein extracts were analyzed using a 600 MHz NMR spectrometer (Bruker BioSpin, Rheinstetten, Germany) at the operating ^{1}H frequency of 600.13 MHz at a temperature of 300 K. To acquire one-dimensional (1D) ^{1}H NMR spectra of aqueous extracts a Carr Purcell Meiboom Gill (CPMG) pulse sequence [RD–90°–(τ–180°–τ) n–FID, 2nτ=64 ms] to suppress broad signals from macromolecules was used. A 90 degree pulse was adjusted to 10 µs. A total of 512 scans were accumulated into 64 k data points with a spectral width of 20 ppm.

For organic extracts a 1D (Zg30pr) experiment was applied wherein 256 scans were attained into 32 k data points. The spectral width was 20.00 ppm with a RD of 2 s, acquisition time 1.36 s, spin-echo delay τ=400 µs and total echo time of 64 ms for all organic and aqueous experiments.

UPLC-MS analysis of aqueous and organic extracts

A detailed protocol for reconstitution of aqueous and organic extracts for UPLC-MS analysis is given in supporting information section. A total of 50 µl from each sample was added together to make a quality control (QC) sample[32] (see detailed protocol in supporting information section). Analysis of aqueous extracts with hydrophilic interaction liquid chromatography (HILIC) was performed as previously described[33], using an Acquity UPLC

Analyst Accepted Manuscript

System (Waters, Ltd. Elstree, UK), coupled with LCT Premier mass spectrometer (Waters MS Technologies, Ltd., Manchester, U.K.). An Acquity UPLC BEH HILIC column (1.7 µm, 2.1 x 100 mm, Waters, USA) was used and maintained at 35 ^{o}C. For RP chromatography of aqueous extracts, an Acquity UPLC HSS T3 column (1.8 µm, 2.1 x 100 mm, Waters, USA) was used on acquity UPLC System (Waters, Ltd. Elstree, UK), coupled with LCT Premier mass spectrometer (Waters MS Technologies, Ltd., Manchester, UK). One replicate from the methanol/water (1:3) group in RP-UPLC-MS analysis ESI+ mode experienced an injection failure and was removed from further analysis. Organic extracts (lipid profiling) were analyzed as previously described[33] using an Acquity UPLC system (Waters Ltd Elstree, UK) coupled to a Q-TOF Premier mass spectrometer (Waters Technologies, Ltd. Manchester, UK). For chromatography of organic extracts, an Acquity UPLC column CSH (1.7 µm, 2.1 x 100 mm, Waters, USA) was used. Detailed parameters of instruments are mentioned in the supporting information section.

The gradient programs for all UPLC-MS analyses are given in the supporting information (Table S 1 a,b,c in supporting information). The order of injection of samples was randomized. QC samples were used to monitor the performance of UPLC-MS system, and were run at the beginning of the run (to condition the chromatographic column) and periodically after every 3 aqueous samples and 10 organic samples in stage 1, and after every 3 samples in stage 2 during the experiment. Analyses were conducted separately for positive (ESI+) and negative (ESI-) ESI modes. Two extraction and two solvent blanks were injected at the end of each run to identify any features introduced from the extraction process and solvent systems. The injection from one sample from the DCM/methanol group failed and this sample was removed from further analysis.

Analyst Accepted Manuscript

Data processing and statistical analysis of NMR and UPLC-MS data

Aqueous and organic NMR spectra were phased, corrected for baseline distortions and calibrated to chemical shift of TSP (δ 0.00) in the aqueous phase or TMS (δ 0.00) in the organic phase respectively using TOPSPIN 3.0 software (Bruker BioSpin, Rheinstetten, Germany). Spectra were imported into MATLAB R2009b (Mathworks™, 2009) using in-house developed scripts. Regions containing water resonance (from δ 4.68 to δ 5.24) and TSP or TMS (from δ -1 to δ 0.2) were removed from all spectra. The resulting spectra were aligned using recursive segment-wise peak alignment[34] followed by probabilistic quotient normalization[35] of the spectral data.

For UPLC-MS, data were processed using the MarkerLynx package (MassLynx V4.1 software, SCN 857 Build 26, Waters). Parameters used in Markerlynx for data analysis are given in Table S 2 in supporting information. After peak-picking of the chromatographic peaks a three-dimensional table with features being characterized by their m/z, retention time and signal intensity, was produced. The dataset was subjected to total area normalization.

NMR and UPLC-MS processed data for each stage were transferred to SIMCA-P+ 12.5 statistical software (UMETRICS™, Sweden). Models were constructed using principal component analysis (PCA) with Pareto scaling. PCA scores plots display the overall variation in the dataset based on the minimum number of components; samples are mapped based on their feature similarities or differences to other samples. PCA scores plots were used to assess the effects of the different extraction solvent systems on the metabolic profiles (using the extraction replicates) and to identify any runtime or machine variance and analytical performance (using the QC samples). Outliers reflecting experimental anomalies were identified and excluded. Reproducibility was assessed by using the coefficient of variation (CV) expressed as a percentage of the extraction replicates. CVs were calculated for each individual metabolic feature among all replicates of each extraction group (n=10). The

Analyst Accepted Manuscript

distribution of these CVs were then compared between the groups to overview the differences in reproducibility[14, 36]. For data acquired from UPLC-MS, reproducibility was assessed by calculating the percentage of metabolites with CV within a 30% CV cut-off ($CV_{30\%}$) range (n=10 per group). This cut-off value has been used as a standard to assess reproducibility[32, 37, 38]. For data acquired from NMR, the number of features with their CVs within a 5% cut-off limit (n=10 per group) was calculated for each extraction solvent and method. Extraction solvents or methods with higher number of features with CVs within that limit were considered to have a better reproducibility. The low cut-off limit of 5% for NMR analysis relative to the $CV_{30\%}$ value employed for UPLC-MS was chosen because instrument related disparities are small in NMR. Additionally, in the case of UPLC-MS data, the features present in the extraction blanks of each group were removed from the data. In NMR analysis, only those spectral peaks that were not present in the respective blank samples were included.

For UPLC-MS, metabolite structural assignments were conducted by: 1) matching accurate mass measurements to theoretical values from on-line databases including METLIN (*http://metlin.scripps.edu/metabolites*), HMDB *http://www.hmdb.ca*) and LIPID MAPS, (http://www.lipidmaps.org), 2) isotopic patterns, 3) in-house developed libraries of standards and 4) MS^E and/or MS/MS spectra, by matching to tandem MS experiments from online databases.

Results and discussion

We developed a robust workflow to address the needs of metabolic phenotyping studies in the context of metabolic characterization of human vascular tissue but this analytical pipeline would also be applicable to study other types of tissue. The first stage involved comparison of two different solvents systems for extraction of aqueous metabolites and five different

Analyst Accepted Manuscript

solvents for organic metabolite extraction from human vein tissue homogenate. The optimal aqueous and organic solvent mixtures were chosen based on reproducibility regarding the number of metabolites recovered from the tissue and their signal intensities detected by both NMR and UPLC-MS analyses. Methanol/water (1:1) was found to be the best solvent for aqueous extraction compared with Methanol/water (1:3), whereas MTBE/methanol (3:1) was the most robust solvent in organic extraction compared with chloroform/methanol (3:1), DCM/methanol (3:1), hexane/methanol/ISP (13:5:2), ISP/methanol (3:1).

Optimization of aqueous phase metabolite extraction

Based on reproducibility and feature recovery from NMR spectra

The global PCA model of the NMR aqueous extracts of human veins (Figures 2 a and 1b) showed a skewed profile, influenced by two outliers, which belonged to the group extracted using methanol/water (1:3). Inspection of individual spectra showed that one of the outliers possessed a lower signal to noise level, and the other had generally reduced peak intensities with a markedly lower intensity for lactate as compared to the other samples.

Reproducibility was evaluated using the CV, calculated for each group (see Table 1 for reproducibility). The number of NMR peak intensities with their CVs within the cut-off of 5% was higher for replicates of methanol/water (1:1) group as compared to replicates of methanol/water (1:3) group (264 data points *vs.* 11). Visual examination of the NMR spectra from each group showed that most features for aqueous metabolites were common between the two extraction solvents.

Based on reproducibility and feature recovery from UPLS-MS data

The PCA score plots of UPLC-MS (HILIC and RP) in ESI+ and ESI- modes performed on aqueous extracts (Figures 2 c, d, e and f) showed tight clustering of the QC samples in all

Analyst Accepted Manuscript

models, suggesting a high stability of the instrument during the analytical runs. For features acquired from HILIC- and RP-UPLC-MS data (both ESI+ and ESI-), features from replicates of methanol/water (1:1) group had significantly higher percentage of their CVs within 30% cut-off compared to the other solvents: for example, 53% of metabolites from methanol/water (1:1) as compared to 35% from replicates of methanol/water (1:3) group. Features driving the variation – thus, the differences between groups - in the PCA scores plots were identified using the PCA loadings plots. This way, we were able to highlight which features were unique to one group and to ascertain which metabolites or classes of metabolites may have diminished recovery under certain solvents (Figure S 1 and 2 and Table S 3 a,b, c in supporting information). Features including phosphatidylcholine (PC) and hypoxanthine (in HILIC ESI+), phosphorylethanolamine (in ESI-) and PC and monoacylglycerol (in RP ESI+) were detected differentially in higher intensities by the methanol/water (1:3) solvent system.

For both ESI modes of the HILIC-UPLC-MS analysis, 95% and 91% of their features in the methanol/water (1:3) and methanol/water (1:1) solvents were common to both groups. Similarly, >90% of features were shared by the two groups in RP-UPLC-MS ESI+ and ESI- modes. The total numbers of features detected by each group analyzed by UPLC-MS (ESI +/-) are listed in Table 1. More features were detected in HILIC ESI+ mode by methanol/water (1:1) (n=1143) as compared to methanol/water (1:3) (n=1082), whilst the reverse was true in HILIC ESI- mode (1578 features for methanol/water (1:1) as compared to 1700 for methanol/water (1:3)). There was no difference in the number of features detected between the two groups in ESI+ mode measured by RP-UPLC-MS analysis (n=560 features detected in both solvent systems), while methanol/water (1:1) had a greater number of features in ESI- mode (n=884 versus n=722).

Optimization of organic phase metabolite extraction

Analyst Accepted Manuscript

Based on reproducibility and features recovery in NMR

PCA models of the NMR spectra from organic extracts (Figure 3a) showed an overlap between the replicates extracted by chloroform/methanol (3:1) and DCM/methanol (3:1), which is in keeping with both these solvents sharing similar chemical structure, and thus physicochemical properties. The other 3 solvent groups clustered independently in the PCA scores plot. The number of features with their CVs within the set limit of 5% was higher for the MTBE/methanol (3:1) extracted group (439) followed by DCM/methanol (3:1) method with only 8 features within the cut-off limit of CV ≤ 5% for DCM.

Based on reproducibility and features recovery on UPLC-MS

The PCA scores plots of RP-UPLC-MS ESI+ and ESI- modes of the organic extracts (lipid profiling; Figures 3b and 3c) showed good clustering of QCs, demonstrating satisfactory instrument analytical stability. During the UPLC-MS experiment run, one sample from the ISP/methanol group generated considerably lower intensity chromatograms (both ESI+ and ESI-) as compared to the remaining replicates from the same group. The pattern of distribution of replicates from each group in the PCA scores plots was similar to that produced by NMR analysis, notably, an overlap between chloroform/methanol (3:1) and DCM/methanol (3:1) groups was observed. Replicates extracted by MTBE/methanol (3:1) showed the tightest clustering, whereas the rest of the groups were widely spread. This indicates a better reproducibility for the MTBE/methanol (3:1) group. Likewise the percentage of CVs of features within $CV_{30\%}$ cut-off was markedly higher for replicates of MTBE/methanol (3:1) group as compared to the rest of the organic solvents on UPLC-MS ESI+ and ESI- (lipid profiling), further supporting the findings demonstrating that MTBE/methanol (3:1) solvent extraction has superior reproducibility for vein tissue.

Analyst Accepted Manuscript

Greater than 80% of features identified were common to all five solvent groups used for the extraction of organic metabolites in ESI+, and approximately 90% of features were in common to all solvents in the ESI- mode. MTBE/methanol (3:1) and chloroform/methanol (3:1) shared the highest proportion of features (91%) while chloroform/methanol (3:1) and DCM/methanol (3:1) shared 85% features. Chloroform/methanol (3:1) provided the largest number of features both in ESI+ and ESI– modes, n=734 and n=105 features, respectively, followed by MTBE: methanol (3:1), 600 and 98 features, respectively.

Several studies have suggested that solvent extraction has a greater effect on metabolite profiling quality than other methodological considerations such as tissue disruption method or temperature of the solvent mixture[29, 39]. ISP, ether, DCM, MTBE and chloroform have all been used to extract metabolites from different tissues types[22, 23 13, 24], and the consensus is that methanol/water/chloroform and methanol/water based extraction solvents provide good recovery for wide range of animal or human tissues including liver, brain and colonic tissues in terms of yield and reproducibility. Here we evaluated the potential of various solvents with a view to combined UPLC-MS and NMR coverage of the metabolome. In terms of extraction of polar metabolites, methanol/water (1:1) provided higher reproducibility for UPLC-MS and NMR based analysis of human vein tissue aqueous extracts. Since the performance of the two solvent systems was similar with respect to NMR analysis, the methanol/water (1:1) was selected as the optimum solvent for aqueous extraction.

Samples extracted by MTBE/methanol (3:1) have better reproducibility both in NMR and UPLC-MS analysis (comparison of their CVs is listed in the Table 1). The chloroform/methanol (3:1) solvent mixture – which is widely considered as the gold standard solvent system for extraction of organic components - performed better only in terms of the number of features detected (734 in ESI+ and 105 in ESI-). Both chloroform/methanol (3:1)

Analyst Accepted Manuscript

and DCM/methanol (3:1) have preference towards picking specific classes of metabolites such as phosphatidylcholine whereas replicates extracted by MTBE/methanol (3:1) appear to be clustered in the center of the PCA scores plot (see figures S3 and tables S3 e and f in supporting information). Based on the superior reproducibility MTBE/methanol (3:1) was selected as the most appropriate method for human vein tissue profiling.

Establishing the order of consecutive extraction of organic and aqueous phases

It has previously been shown that the consecutive approach to extraction of metabolites from tissue has the advantage over the bilayer approach[4]. Therefore, we compared the order of solvent extraction for the optimal solvent systems selected in stage 1 (aqueous: methanol/water (1:1); organic: MTBE/methanol (3:1)) as follows: (i) aqueous extraction followed by organic (C-A-O), and (ii) organic extraction followed by aqueous (C-O-A).

Reproducibility and features recovery as measured by NMR spectroscopy

The PCA scores plots of the NMR analyzed aqueous (Figure 4a) and organic (Figure 4b) extracts demonstrated that the spread of replicates had a similar pattern for both extraction sequences. Reproducibility was assessed by measuring the number of metabolites with CVs within the 5% cut-off limit (n=10 per group). With NMR analysis of aqueous extracts, the C-A-O method yielded better reproducibility with CVs of 629 features within 5% as compared to 258 features for the C-O-A method. For the NMR-based analysis of organic extracts, the number of features with their CVs within 5% cut-off limit was higher for the C-O-A method (347 features) as opposed to 69 features recovered for the C-A-O sequence.

Recovery of metabolites based on reproducibility and features in the UPLC-MS data

Analyst Accepted Manuscript

For the UPLC-MS-based analysis, QC samples were well clustered in PCA scores plots derived from HILIC and RP data, regardless of polarity (Figures 4c to 4f). For the aqueous extracts, the two groups demonstrated different profiles in the PCA scores plots with better clustering among replicates of the C-O-A group. Multivariate analysis of the organic metabolite extraction from the two protocols (Figures 4 g and h) showed similar distribution and clustering between the groups indicating comparable reproducibility for organic metabolites in the PCA scores plots.

Analysis of the UPLC-MS data showed that the C-O-A group had a higher reproducibility for polar metabolites. This high reproducibility was demonstrated by the greater percentage of features within $CV_{30\%}$ limit for the C-O-A group (34%) against the C-A-O group (16%) in HILIC ESI+ mode (see Table 1 for reproducibility and features detected for each group). For the UPLC-MS-based analysis of organic extracts (lipid profiling), there was no difference between the two extraction methods in terms of reproducibility, as demonstrated by the number of metabolites within the $CV_{30\%}$ cut-off limit.

For polar metabolite extraction, the C-A-O method gave superior results in terms of the number of features detected in UPLC-MS analysis (HILIC and RP), although perceived differences were small. On the contrary, the number of features detected from organic phase extraction was higher for the C-O-A method as compared to the C-A-O method (1114 features versus 878 features). For UPLC-MS (HILIC and RP) analysis of aqueous extracts, 85% of features were common between the two solvent groups used in the stage 2 analysis. For organic analysis of samples ran using the RP-UPLC-MS ESI+ and ESI- modes (lipid profiling), both extraction sequences retrieved exactly the same features in 85% of cases in ESI+ and 97% in ESI– modes.

To summarize the second stage of study, where we compared the order of solvent extraction, the polar metabolites UPLC-MS-based analyses supported the C-O-A extraction method over

Analyst Accepted Manuscript

the C-A-O method. The reverse was true for the NMR-based analysis. With respect to organic metabolites extractions, NMR analysis produced more favorable results for extraction of organic metabolites first using MTBE/methanol (3:1), whereas there was no difference between the two groups for UPLC-MS-based analysis. However, more organic features were detected by the C-O-A method. Aqueous and organic NMR spectra and UPLC-MS chromatograms are provided in the supplementary data (Supplementary figure S 4 and S 5).

The classical use of metabonomics usually relies on untargeted analyses of a large number of analytes. Therefore, when using these approaches to measure a vast range of metabolites in a complex mixture, for example tissue or biofluid, optimized protocols are required to achieve maximum and reproducible coverage of the metabolome. Recent work on intact human vein tissue biopsies using ^{1}H magic angle spinning-NMR has been used to characterize a range of metabolites including alanine, lactate, *myo*-inositol, glutamate, glucose, small amino acids and different species of triglycerides in human vein tissue samples[10]. More importantly, these metabolites may have value as potential biomarkers, which could influence the treatment of varicose vein disease[10]. Here, we focused on optimizing the first and likely the most important step in terms of the induction of systematic variation and metabolite extraction yield for two spectroscopic platforms to broaden the scope of metabolites detected. We showed recovery of a wide range of molecular species in addition to those reported from NMR analysis, adding numerous lipid moieties such as sphingomyelins, triacylglycerol species and phosphatidylcholines to the list. It must be noted that analysis required ~25mg of tissue for each UPLC-MS method and ~50mg (2 x 25mg aliquots) for NMR spectroscopy. It was essential that the tissue had been already homogenized. This approach adds value to metabolic phenotyping of tissue by enabling direct comparison between analytical platforms and resolves any challenges occurring as a result of tissue heterogeneity. In situations where

Analyst Accepted Manuscript

sample weight is limited we would suggest using a combined organic phase RP- and aqueous phase HILIC- UPLC-MS analysis only, since they were the most informative and sensitive methods. Additionally, the extraction method developed in this study for veins can be used for other types of tissue. We have recently applied the described optimized extraction protocol for metabolite extraction and analysis of adipose tissue[33].

Conclusions

This study comprehensively evaluated and optimized sample extraction protocols for human vein tissue, with a view to providing a single method suited to both NMR and UPLC-MS analysis. For extraction of human vein tissue samples for multi-platform metabonomic analysis, a consecutive approach with extraction of organic metabolites using MTBE/ methanol (v:v, 3:1), followed by extraction of polar metabolites using methanol/water (v:v, 1:1) was found to be the optimal solution. This was evident predominantly in terms of reproducibility, whilst remaining comparable to the consecutive aqueous extraction followed by organic in terms of metabolic features acquired. The optimized protocol will ultimately provide a robust platform for studying not only venous disease but also diseases affecting several tissue types, by enhancing our understanding of the underlining pathological mechanisms which are currently poorly understood.

Acknowledgments

MA acknowledges Imperial College Healthcare Trust Charity, The Graham Dixon Charitable Trust, and The European Venous Forum for research grants. PAV acknowledges the Royal Society of Chemistry for funding his PhD studentship. We acknowledge National institute for Health Research (NIHR) and Imperial Biomedical Research Centre (BRC) funding for Clinical Phenotyping Centre at Imperial College London.

Analyst Accepted Manuscript

References

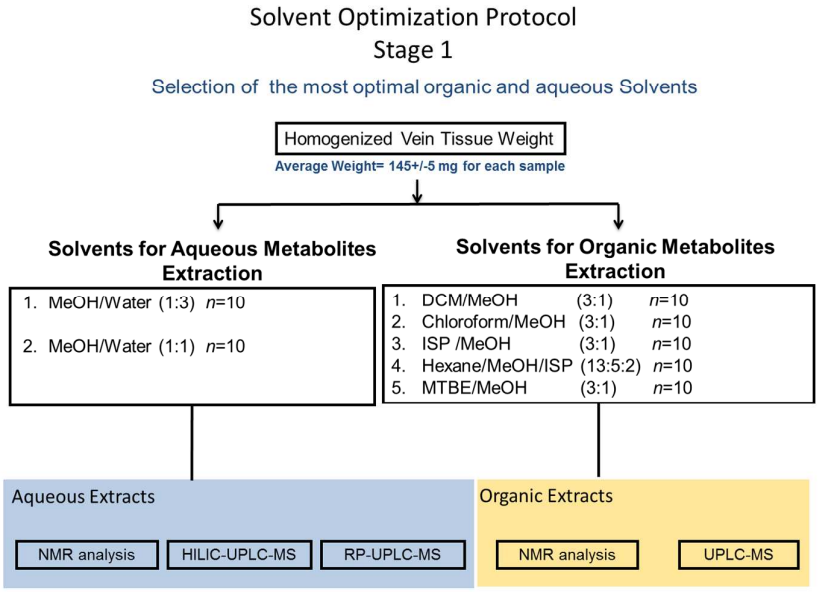
1. J. K. Nicholson and I. D. Wilson, *Nat Rev Drug Discov*, 2003, **2**, 668-676.
2. G. S. Parab, R. Rao, S. Lakshminarayanan, Y. V. Bing, S. M. Moochhala and S. Swarup, *Anal Chem*, 2009, **81**, 1315-1323.
3. O. Beckonert, H. C. Keun, T. M. Ebbels, J. Bundy, E. Holmes, J. C. Lindon and J. K. Nicholson, *Nat Protoc*, 2007, **2**, 2692-2703.
4. P. Masson, K. Spagou, J. K. Nicholson and E. J. Want, *Anal Chem*, 2011, **83**, 1116-1123.
5. J. L. Ballard and J. L. Mills, Sr., *Tech Vasc Interv Radiol*, 2005, **8**, 169-174.
6. C. S. Lim and A. H. Davies, *Br J Surg*, 2009, **96(11)**, 1231-1242.
7. P. Saha, J. Humphries, B. Modarai, K. Mattock, M. Waltham, C. E. Evans, A. Ahmad, A. S. Patel, S. Premaratne, O. T. Lyons and A. Smith, *Arterioscler Thromb Vasc Biol*, 2011, **31**, 506-512.
8. E. M. Boyle, Jr., S. T. Lille, E. Allaire, A. W. Clowes and E. D. Verrier, *Ann Thorac Surg*, 1997, **63**, 885-894.
9. E. Allaire and A. W. Clowes, *Ann Thorac Surg*, 1997, **63**, 582-591.
10. M. A. Anwar, J. Shalhoub, P. A. Vorkas, C. S. Lim, E. J. Want, J. K. Nicholson, E. Holmes and A. H. Davies, *Eur J Vasc Endovasc Surg*, 2012, **44**, 442-450.
11. M. Urban, D. P. Enot, G. Dallmann, L. Korner, V. Forcher, P. Enoh, T. Koal, M. Keller and H. P. Deigner, *Analytical biochemistry*, 2010, **406**, 124-131.
12. M. V. Brown, J. E. McDunn, P. R. Gunst, E. M. Smith, M. V. Milburn, D. A. Troyer and K. A. Lawton, *Genome medicine*, 2012, **4**, 33.
13. W. H. Lin C.Y., Tjeerdema RS, Viant MR, *Metabolomics*, 2007, **3**, 55-67.
14. P. Masson, A. C. Alves, T. M. Ebbels, J. K. Nicholson and E. J. Want, *Anal Chem*, 2010, **82**, 7779-7786.
15. E. G. Bligh and W. J. Dyer, *Canadian journal of biochemistry and physiology*, 1959, **37**, 911-917.
16. H. Nygren, T. Seppanen-Laakso, S. Castillo, T. Hyotylainen and M. Oresic, *Methods Mol Biol*, 2011, **708**, 247-257.
17. R. Salek, K. K. Cheng and J. Griffin, *Methods in enzymology*, 2011, **500**, 337-351.
18. L. Whiley, J. Godzien, F. J. Ruperez, C. Legido-Quigley and C. Barbas, *Anal Chem*, 2012, **84**, 5992-5999.
19. V. Matyash, G. Liebisch, T. V. Kurzchalia, A. Shevchenko and D. Schwudke, *Journal of lipid research*, 2008, **49**, 1137-1146.
20. J. Folch, M. Lees and G. H. Sloane Stanley, *J Biol Chem*, 1957, **226**, 497-509.
21. L. A. Carlson, *Clinica chimica acta; international journal of clinical chemistry*, 1985, **149**, 89-93.
22. S. J. Bruce, P. Jonsson, H. Antti, O. Cloarec, J. Trygg, S. L. Marklund and T. Moritz, *Analytical biochemistry*, 2008, **372**, 237-249.
23. Y. Sato, T. Nakamura, K. Aoshima and Y. Oda, *Anal Chem*, 2010, **82**, 9858-9864.
24. D. M. Jacobs, N. Deltimple, E. van Velzen, F. A. van Dorsten, M. Bingham, E. E. Vaughan and J. van Duynhoven, *NMR Biomed*, 2008, **21**, 615-626.
25. M. Mal, P. K. Koh, P. Y. Cheah and E. C. Chan, *Rapid Commun Mass Spectrom*, 2011, **25**, 755-764.
26. J. E. Le Belle, N. G. Harris, S. R. Williams and K. K. Bhakoo, *NMR Biomed*, 2002, **15**, 37-44.
27. E. J. Want, M. Coen, P. Masson, H. C. Keun, J. T. Pearce, M. D. Reily, D. G. Robertson, C. M. Rohde, E. Holmes, J. C. Lindon, R. S. Plumb and J. K. Nicholson, *Anal Chem*, 2010, **82**, 5282-5289.
28. E. J. Want, I. D. Wilson, H. Gika, G. Theodoridis, R. S. Plumb, J. Shockcor, E. Holmes and J. K. Nicholson, *Nat Protoc*, 2010, **5**, 1005-1018.
29. F. M. Geier, E. J. Want, A. M. Leroi and J. G. Bundy, *Anal Chem*, 2011, **83**, 3730-3736.
30. M. Coen, E. M. Lenz, J. K. Nicholson, I. D. Wilson, F. Pognan and J. C. Lindon, *Chem Res Toxicol*, 2003, **16**, 295-303.
31. M. L. J. G. Bundy, *Metabolomics*, 2011, DOI: DOI 10.1007/s11306-011-0377-1.

Analyst Accepted Manuscript

32. K. Spagou, I. D. Wilson, P. Masson, G. Theodoridis, N. Raikos, M. Coen, E. Holmes, J. C. Lindon, R. S. Plumb, J. K. Nicholson and E. J. Want, *Anal Chem*, 2011, **83**, 382-390.
33. P. A. Vorkas, G. Isaac, M. A. Anwar, A. H. Davies, E. J. Want, J. K. Nicholson and E. Holmes, *Anal Chem*, 2015, **87**, 4184-4193.
34. K. A. Veselkov, J. C. Lindon, T. M. Ebbels, D. Crockford, V. V. Volynkin, E. Holmes, D. B. Davies and J. K. Nicholson, *Anal Chem*, 2009, **81**, 56-66.
35. F. Dieterle, A. Ross, G. Schlotterbeck and H. Senn, *Anal Chem*, 2006, **78**, 4281-4290.
36. W. B. Dunn, D. Broadhurst, M. Brown, P. N. Baker, C. W. Redman, L. C. Kenny and D. B. Kell, *J Chromatogr B Analyt Technol Biomed Life Sci*, 2008, **871**, 288-298.
37. T. Sangster, H. Major, R. Plumb, A. J. Wilson and I. D. Wilson, *The Analyst*, 2006, **131**, 1075-1078.
38. FDA, MAY 2001.
39. A. Beltran, M. Suarez, M. A. Rodriguez, M. Vinaixa, S. Samino, L. Arola, X. Correig and O. Yanes, *Anal Chem*, 2012, **84**, 5838-5844.

Analyst Accepted Manuscript

a



b.

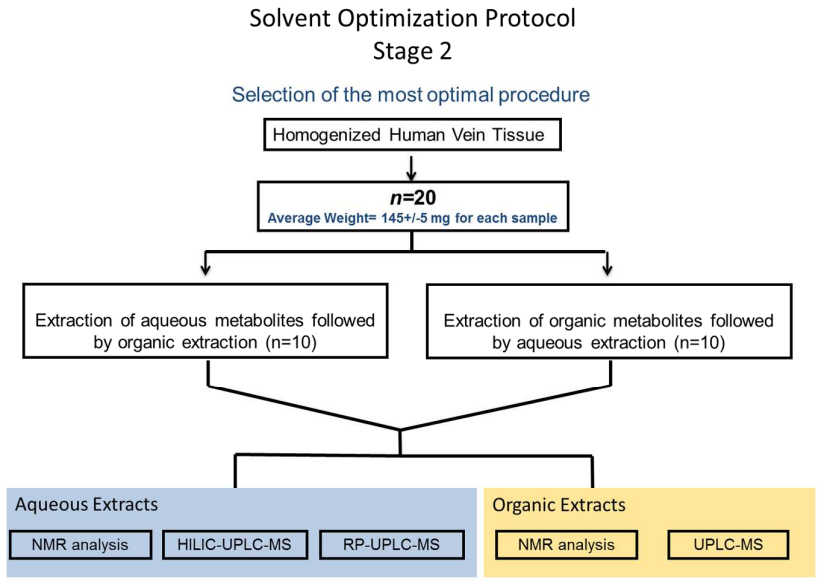


Figure 1. Schematic diagrams of the workflows followed in order to evaluate a) the aqueous and organic solvent mixtures and b) the optimal order of aqueous and organic phase extractions in consecutive mode. DCM; Dichloromethane, ISP; Isopropanol, MeOH; methanol, MTBE; Methyl *tert*-butyl ether, NMR; Nuclear magnetic resonance spectroscopy, UPLC-MS; Ultra performance liquid chromatography – mass spectrometry, HILIC; hydrophilic interaction liquid chromatography, RP; Reversed phase.

Analyst Accepted Manuscript

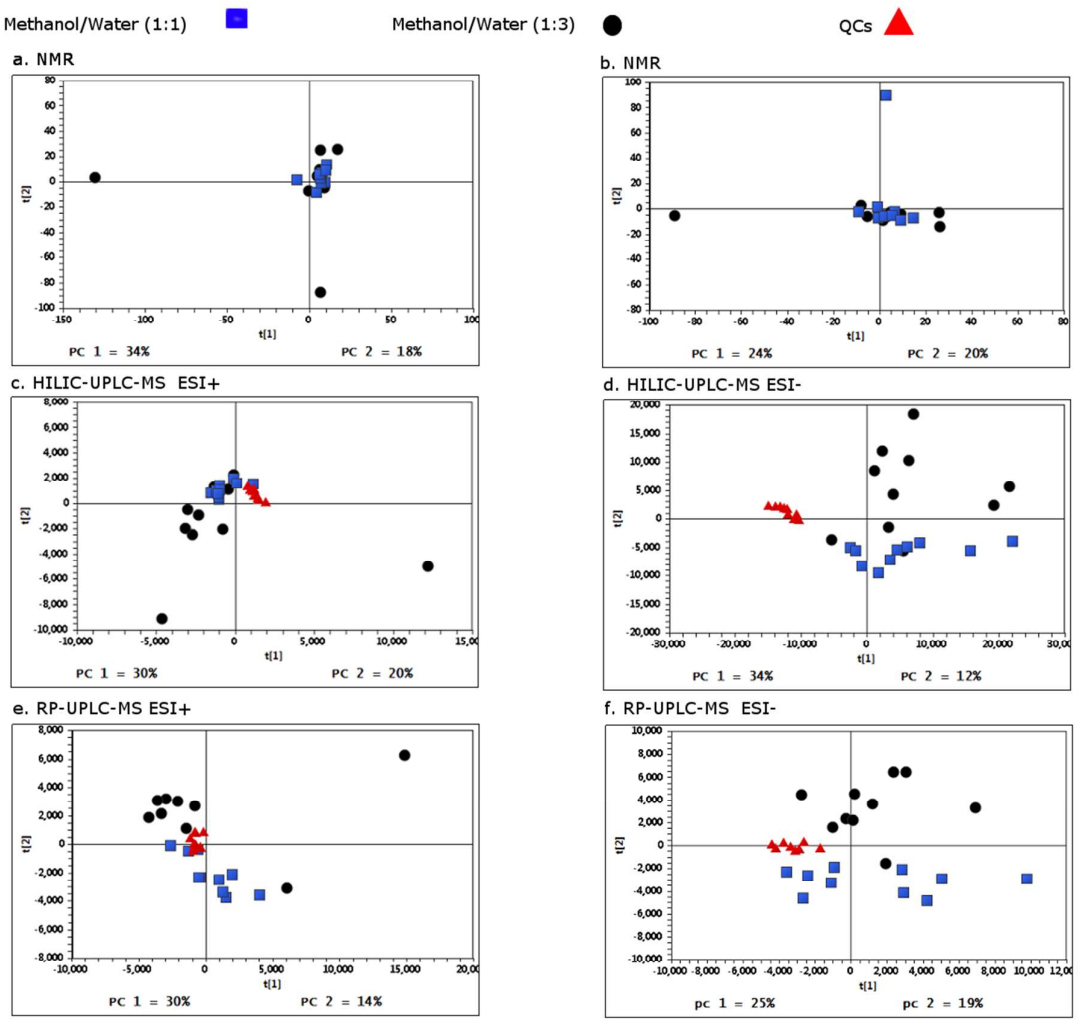


Figure 2. Principal component analysis (PCA) scores plots showing the variations and trends of the data acquired from the analysis of aqueous extracts, for the optimization of the aqueous extraction solvent mixture ***(stage 1).*** (a) NMR data model with outliers (b) NMR data model after outlier removal and model re-fitting, (c) and (d) HILIC-UPLC-MS and (e) and (f) RP-UPLC-MS. For all UPLC-MS experiments samples were analyzed in both ESI polarity modes: positive (ESI+) and negative (ESI-). Reproducibility can be assessed by observing the grouping of replicas for each group. NMR; Nuclear magnetic resonance spectroscopy, HILIC; Hydrophilic interaction liquid chromatography, UPLC-MS; Ultra performance liquid chromatography – mass spectrometry, RP; Reversed phase, ESI; Electrospray ionization, QCs; Quality controls.

Analyst Accepted Manuscript

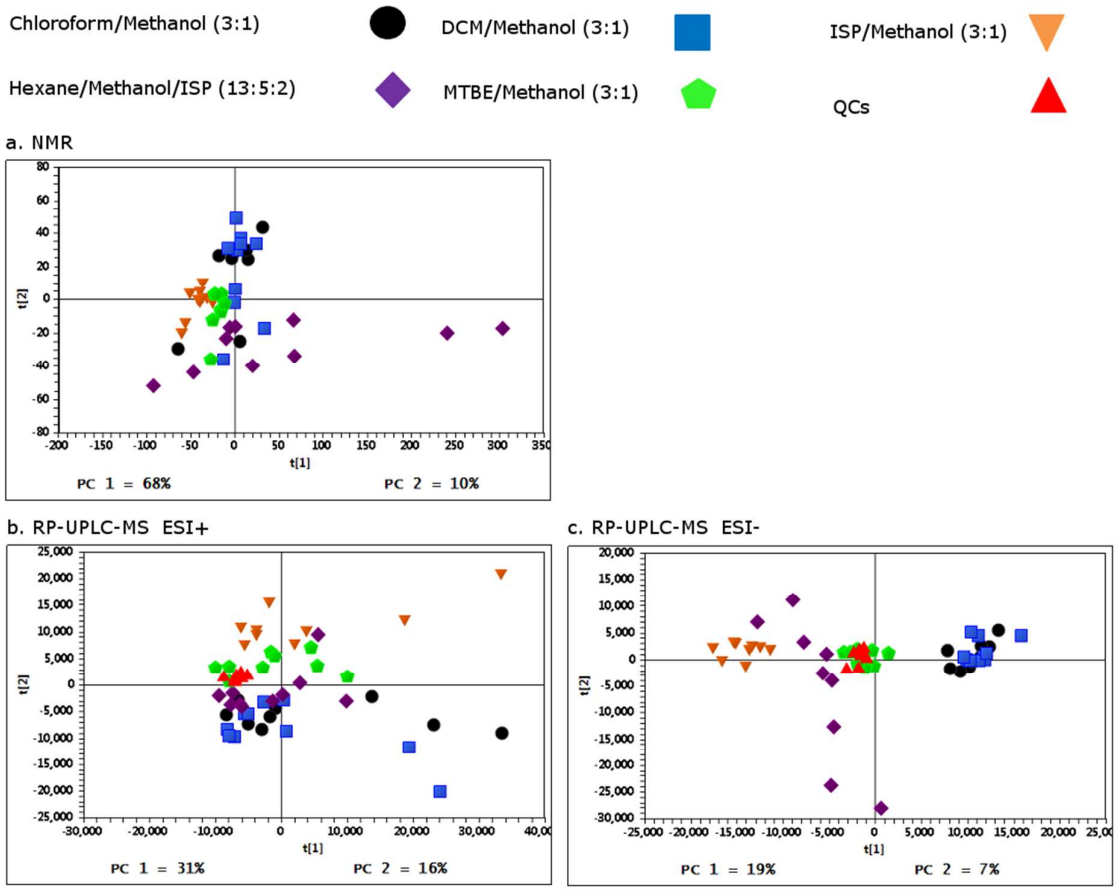


Figure 3. Principal component analysis scores plots of the data obtained from the optimization of the organic phase extraction after analyses using (a) NMR, (b) reversed phase-UPLC-MS ESI+ and (c) ESI-. Five different organic solvent systems were evaluated (***stage 1)***. Reproducibility can be assessed by observing the grouping of replicates for a given extraction method. DCM; Dichloromethane, ISP; Isopropanol, MTBE; Methyl *tert*-butyl ether, NMR; Nuclear magnetic resonance spectroscopy, RP; Reversed phase, UPLC-MS; Ultra performance liquid chromatography – mass spectrometry, ESI; Electrospray ionization, QCs; Quality controls.

Analyst Accepted Manuscript

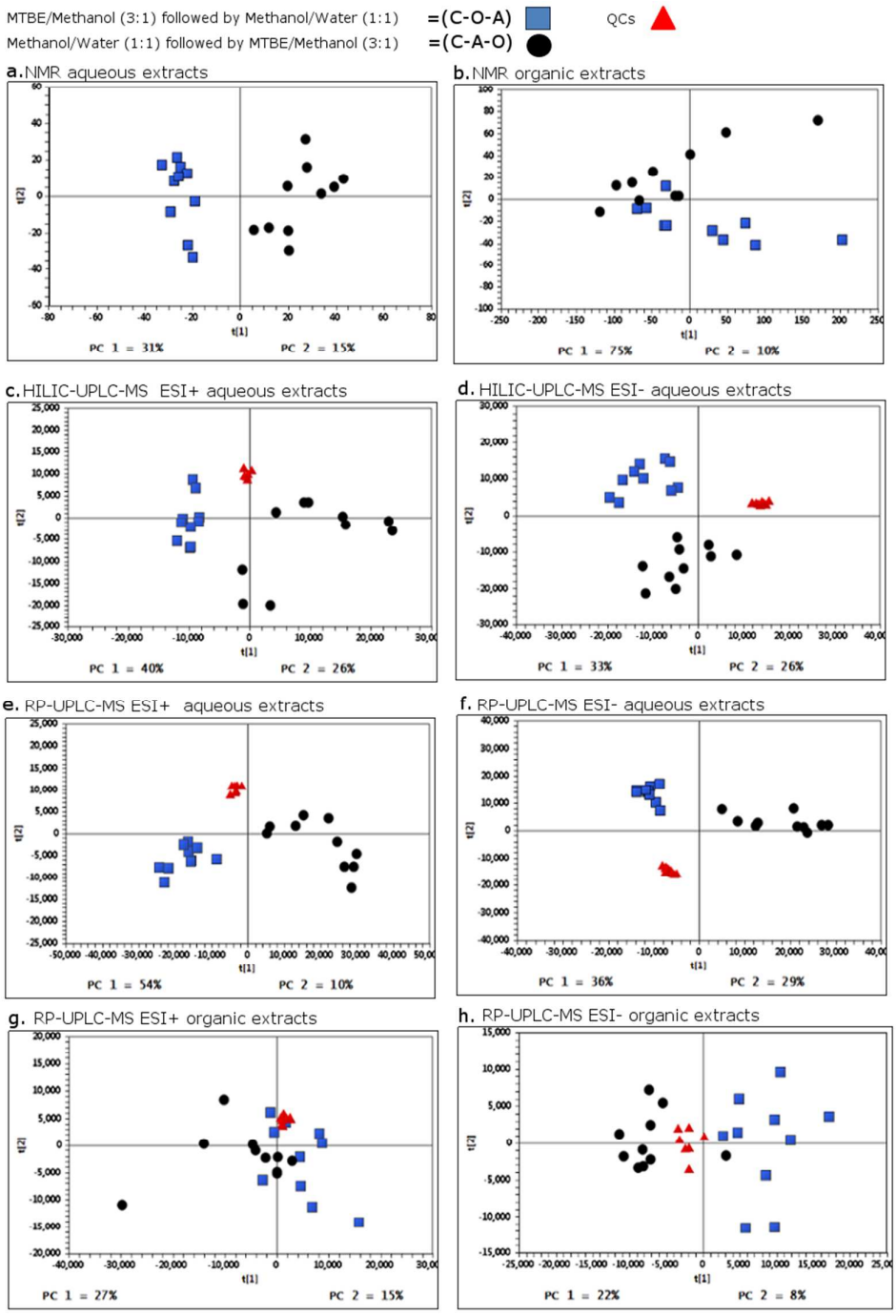


Figure 4. Principal component analysis scores plots of the data acquired from aqueous and organic extracts, in the stage of optimal extraction order assessment, using consecutive mode (***Stage 2)***. (a) NMR of aqueous extracts, (b) NMR of organic extracts, (c) HILIC-UPLC-MS of aqueous extracts in ESI+ and (d) ESI-, (e) RP-UPLC-MS of aqueous extracts in ESI+ and (f) ESI-, and (g) RP-UPLC-MS of organic extracts in ESI+ and (h) ESI-. The grouping of replicates for each extraction protocol gives an indication of the reproducibility. MTBE; methyl *tert*-butyl ether, C-A-O; consecutive aqueous extraction followed by organic, C-O-A;

Analyst Accepted Manuscript

consecutive organic extraction followed by aqueous, NMR; Nuclear magnetic resonance spectroscopy, HILIC; Hydrophilic interaction liquid chromatography, UPLC-MS; Ultra performance liquid chromatography – mass spectrometry, ESI; Electrospray ionization, RP; Reversed phase, QCs; Quality controls.

Analyst Accepted Manuscript

Table 1 Number of features detected and percentage of features with coefficient of variation (CV) within the predetermined cut-offs (reproducible), calculated using the replicates (n=10) of each extraction group. For UPLC-MS-based profiling, the CV cut-off was set at 30% ($CV_{30\%}$). For NMR-based profiling, reproducible features were considered if their CVs were within the 5% cut-off limit ($CV_{5\%}$)[a]

	NMR Reproducible features	HILIC-UPLC-MS ESI+ Features detected – reproducible	HILIC-UPLC-MS ESI– Features detected – reproducible	RP-UPLC-MS ESI+ Features detected – reproducible	RP-UPLC-MS ESI- Features detected – reproducible	OrgEx RP-UPLC-MS ESI+ Features detected – reproducible	OrgEx RP-UPLC-MS ESI- Features detected – reproducible
Methanol/Water (1:1)	$CV_{5\%}$= 264	1143 $CV_{30\%}$= 53%	1578 $CV_{30\%}$= 25%	560 $CV_{30\%}$= 39%	884 $CV_{30\%}$= 30%	–	–
Methanol/Water (1:3)	$CV_{5\%}$= 11	1082 $CV_{30\%}$= 34%	1700 $CV_{30\%}$= 15%	560 $CV_{30\%}$= 6%	722 $CV_{30\%}$= 18%	–	–
DCM/Methanol (3:1)	$CV_{5\%}$= 8	–	–	–	–	516 $CV_{30\%}$= 8%	90 $CV_{30\%}$= 27%
Chloroform/Methanol (3:1)	$CV_{5\%}$= 0	–	–	–	–	734 $CV_{30\%}$= 13%	105 $CV_{30\%}$= 34%
ISP/Methanol (3:1)	$CV_{5\%}$= 2	–	–	–	–	499 $CV_{30\%}$= 2.5%	76 $CV_{30\%}$= 25%
Hexane/Methanol/ISP (13:5:2)	$CV_{5\%}$= 0	–	–	–	–	522 $CV_{30\%}$= 10%	98 $CV_{30\%}$= 15%
MTBE/Methanol (3:1)	$CV_{5\%}$= 439	–	–	–	–	600 $CV_{30\%}$= 22%	98 $CV_{30\%}$= 40%
MTBE/Methanol (3:1) followed by Methanol/Water (1:1) (C-O-A)	Aqueous $CV_{5\%}$= 258 Organic $CV_{5\%}$= 347	1013 $CV_{30\%}$= 34%	580 $CV_{30\%}$= 35%	1283 $CV_{30\%}$= 34%	695 $CV_{30\%}$= 24%	1114 $CV_{30\%}$= 30%	261 $CV_{30\%}$= 55%
Methanol/Water (1:1) followed by MTBE/Methanol (3:1) (C-A-O)	Aqueous $CV_{5\%}$= 629 Organic $CV_{5\%}$= 69	934 $CV_{30\%}$= 16%	629 $CV_{30\%}$= 27%	1391 $CV_{30\%}$= 24%	784 $CV_{30\%}$= 21%	878 $CV_{30\%}$= 32%	192 $CV_{30\%}$= 50%

[a]Dichloromethane, ISP; Isopropanol, MTBE; methyl *tert*-butyl ether, NMR; Nuclear magnetic resonance spectroscopy, HILIC; Hydrophilic interaction liquid chromatography, UPLC-MS; Ultra performance liquid chromatography – mass spectrometry, ESI; Electrospray ionization, RP; Reversed phase, C-A-O; consecutive aqueous extraction followed by organic, C-O-A; consecutive organic extraction followed by aqueous.

Analyst Accepted Manuscript