Published in final edited form as:

Analyst. 2018 October 08; 143(20): 5014–5022. doi:10.1039/c8an01017j.

Multiple Reaction Monitoring (MRM)-Profiling with Biomarker Identification by LC-QTOF to Characterize Coronary Artery Disease

Karen E. Yannell¹, Christina R. Ferreira¹, Shane Tichy², and R. Graham Cooks¹

¹Chemistry Department, Purdue University, West Lafayette, IN 47907, US

²Agilent Technologies, Santa Clara, CA 95051, US

Abstract

Metabolite profiling by mass spectrometry (MS) is an area of interest for disease diagnostics, biomarker discovery, and therapeutic evaluation. A recently developed approach, multiple reaction monitoring (MRM)-profiling, searches for metabolites with precursor (Prec) and neutral loss (NL) scans in a representative sample and creates a list of ion transitions. These are then used in an MRM method for fast screening of individual samples and discrimination between healthy and diseased. A large variety of functional groups are considered and all signals discovered are recorded in the individual samples, making this a largely unsupervised method.

MRM-profiling is described here and then demonstrated with data for over 900 human plasma coronary artery disease (CAD) samples. Representative pooled samples for each condition were interrogated using a library of over a hundred Prec and NL scans on a triple quadrupole MS. The data from the Prec and NL experiments were converted into ion transitions, initially some 1266 transitions. Each ion transition was examined in the individual samples on a time scale of milliseconds per transition, which allows for rapid screening of large sample sets (<5 days for 1000 samples). Use of univariate and multivariate statistics allowed classification of the sample set with high accuracy.

The metabolite profiles classified the CAD female, CAD male, and peripheral artery disease (PAD) samples relative to controls with an accuracy of 90%, 78%, and 85%, respectively. The compounds responsible for informative ion transitions were identified by chromatography and high resolution MS; some have been previously reported and found to be associated with coronary artery disease metabolism, indicating that the methodology generates a meaningful metabolite profile while being faster than traditional methodologies.

Graphical Abstract

⁷·CONFLICTS OF INTEREST

There are no conflicts to declare.

2. INTRODUCTION

Metabolites are characteristic of an organism's phenotype and include information on diet, environment, genetics, and health.^{1, 2} When performed appropriately, a metabolomics study can uncover biomarkers that are specific to a particular disease state. These can then be used to increase the understanding of the disease mechanism, to help design new drugs, or to assist in disease diagnosis.^{1, 3} MS is often employed for metabolomic studies due to its sensitivity and specificity. It may be used in either an untargeted and in a targeted fashion.

Targeted methodologies use tandem MS (MS/MS) experiments or MRMs, typically in combination with chromatography, to measure specific analytes that are known to be present in a sample.^{2, 4, 5} These can be measured quantitatively but this requires standard solutions and stable isotopically labeled internal standards for signal normalization. Targeted methods require *a priori* knowledge or assumptions of analytes or pathways affected by a disease.² It can be challenging for these methods to take into account heterogeneity of diseases where multiple pathways can be affected.^{5, 6} Untargeted exploratory methods are not based on biological understanding of the disease, and do not require prior knowledge of the disease or metabolic pathways. They are typically performed by high resolution mass spectrometry (HRMS) to provide accurate masses for molecular formula and/or by using data dependent product ion scans.^{2, 7, 8} Data dependent product ion scan methods utilize algorithms which fragment peaks over a certain threshold (counts). Therefore, these methods are often biased toward peaks with the highest intensity, which causes data to be lost. Moreover, when product ions covering a large mass range are fragmented, a very large amount of data is generated, much of which is not biologically relevant.^{6, 9} Both targeted and untargeted metabolomics methods typically employ chromatographic separation (*e.g.* LC). Although chromatography provides additional information about an analyte (*i.e.* retention time, RT) the methods are time consuming, expensive, and may have errors (*e.g.* RT shift).^{1, 10}

MRM-profiling is a simplified and informative exploratory metabolomics method that initially interrogates a sample by Prec and NL scans in order to discover metabolites based on known and/or common functional groups. These experiments which are not completely unsupervised, facilitate the discovery of the biologically significant signals. After the metabolite responses are recorded, the masses involved are converted into ion transitions (Prec/product ion pair) and incorporated into a MRM method to scan for each transition on the millisecond time scale for rapid acquisition of metabolite signals. Interrogating a thousand samples take only a few days for the data acquisition and sample classification by multivariate statistics.

Omics data analysis workflows are readily adapted to processing MRM-profiling data. Univariate statistics, *e.g.* t-test and fold change (FC), are applied so that only significant

transitions remain. Subsequently, multivariate statistical data treatments, *e.g.* partial least squared discriminant analysis (PLSDA), are used to build classification models and determine the specificity, sensitivity, and accuracy of the MRM-profile.^{7, 11} The net interactions of all affected metabolites create the metabolite profile, which is conveniently summarized by the metabolites that are either up or down regulated in a diseased relative to a healthy population.

With MRM-profiling, structural information is accessed through ionic fragmentation. Because neither high resolution nor chromatography is used, isobaric analytes are not resolved. Importantly, no attempt is made to remove matrix effects which are an intrinsic feature of the sample and as such add information. The ion transitions are measured as relative signal intensities, they are not associated with single discrete analytes so modulation by the matrix does not interfere with the analysis. Direct information on how particular analytes vary in different samples is not provided. Finally, data analysis workflows are simpler than LC and HRMS methods since no retention time (RT) or mass correction is necessary.

MRM-profiling methods require very small amounts of sample, often <20 μ L of a biofluid, making them ideal for analyzing a drop of blood from a finger prick, for example. Previously, MRM-profiling has been demonstrated with cerebrospinal fluid analysis for Parkinson's disease, follicular fluid analysis for polycystic ovarian syndrome, and diet alterations in humans amongst other studies.^{12–14} Each of these studies successfully discovered ion transitions that separated different populations. In some cases, analyte identification (not necessarily part of the MRM-profiling workflow) was performed subsequently.

As interest grows in this methodology, it is desirable to confirm that this workflow improves biomarker discovery and finds metabolite signals related to particular populations. This is achieved here by studying a well-characterized disease, coronary artery disease (CAD), relating the significant ion transitions to biomarkers that are biologically relevant, and confirming the identity of the responsible compounds by LC-MS. CAD is a well-studied and highly prevalent disease. In the US alone, over 90 million people suffer from some form of cardiovascular disease and one in three deaths is attributed to CAD.¹⁵ Worldwide, escalations and complications from CAD caused 7.4 million deaths in 2015.¹⁶ Currently, generic blood panels and physical examinations are performed for initial screening followed by invasive stress testing and imaging to confirm the diagnosis.¹⁷ The discovery of new molecular biomarkers for diagnosing CAD and evaluating prescription or lifestyle therapies could benefit this patient population and reduce healthcare costs.

Here we apply MRM-profiling to a set (N>900) of CAD and control plasma samples, quickly interrogate and classify the samples, and confirmed biomarker identities by LC-HRMS. High levels of disease/healthy sample discrimination was achieved and the observed biomarkers were previously reported as related to CAD pathophysiology.

3. EXPERIMENTAL

3.1 Instrumentation, Software, and System Configuration

An Agilent 1290 Infinity series pump and a 6470 QQQ (Agilent Technologies, Santa Clara, CA) were used for MRM-profiling. For sample preparation, an Agilent Bravo Automatic Liquid Handling Platform was used, a Thermo Fisher Scientific SpeedVac was used for drying samples, and an Agilent PlateLoc was used for sealing 96-well plates. An Agilent 6545 QTOF was used for HRMS data acquisition. MassHunter Acquisition (B.08.00), Qualitative (B.08.00), and Quantitative (B.08.00) software programs were used for MRM-profiling and LC-MS experiments. For the univariate and multivariate statistics, Metaboanalyst 3.0 (metaboanalyst.ca, accessed 08/2017–09/2017) and Mass Profiler Professional (B.14.8) were used. Microsoft Excel was used to organize and sort discovered transitions.

Electrospray ionization was employed. Sample introduction used a model 1290 HPLC pump for flow injection (FI).^{13, 18} Additional discussion of methods for introducing samples is found in the ESI (electronic supplementary information). For FI, the pump was connected using a short length of small diameter tubing. A 20 μ L sample loop was used but it was switched to a 40 μ L sample loop as needed. Since a column was not used, the sample was sent directly from the auto sampler to the Agilent Jet Stream (AJS) source, by-passing the MS switching valve and sample filter. A low flow rate (<0.1 mL/min) was desired to maximize the number of scans collected during one injection. To create enough back pressure on the pump, restrictive tubing was added prior to the auto sampler switching valve. Even multiple lines of restrictive tubing did not achieve the required pressure (>30 bar) so a short C18 column was added prior to the injection port. This column was intended only to build up pressure, not for separation.

3.2 Chemicals and Materials

Agilent HPLC vials with 200 μ L glass sample inserts were used for single sample analysis. Deep and shallow poly propylene 96 well plates were used for high throughput sample preparation. HPLC grade acetonitrile, methanol, and chloroform were used for sample preparation and analysis. Millipore ultra-pure water (18.2 MOhm) was procured from a house system. Analytical grade formic acid, ammonium formate, and reserpine were from Sigma Aldrich. Alprazolam-D₅ (0.1 mg/mL) was purchased from Cerilliant. Plates were sealed with an Agilent PlateLoc using the foil seal. An Agilent ZORBAX Eclipse Plus RRHD C18, 2.1 $\times$ 150 mm, 1.8 μ m column was used for separation.

3.3 Samples

De-identified human plasma samples (N=1081) were provided by the Fairbanks Institute for Healthy Communities (Indianapolis, IN). The Institute obtained patient consent for broad, unspecified future research use of their samples issued an RFA for such use. Purdue University and Fairbanks Institute entered into a use agreement on these samples on 22nd July, 2014. Some clinical data was also provided for each sample (*e.g.* case/control, gender, age, medical history, lab results). This data is extensive but incomplete for many samples. Samples were stored at -80 C.

Dilute and shoot sample preparation is encouraged for MRM-profiling so that no metabolites or potential biomarkers are lost to complex sample preparation steps. However, dilute and shoot of human plasma does not give acceptable reproducibility. Therefore, a modified Bligh Dyer extraction procedure was used to extract lipids and polar metabolites.^{13, 19} This automated protocol is detailed in Table S1.

The MRM-profiling methodology does not intrinsically have an analytical quality control check (*e.g.* correct sample preparation, no bubbles in the pump lines, sensitivity of MS). For this reason an inexpensive deuterated drug (alprazolam-D₅) was spiked at 100 ng/mL into the solvent used for sample dilution (90% methanol, 10% chloroform, 10 ppm ammonium formate, 0.1% formic acid). This quality control (QC) analyte was used to check for proper sample preparation and performance of the instrument. This analyte was measured in every time segment and if it was statistically high or low in a sample (outliers calculated by plate, $Q \pm 1.5ICR$) then the sample was re-prepared and reanalyzed. If the sample was rejected a second time then the sample was eliminated from the study (16% rejection rate).

3.4 Method Development

Considerations for implementing MRM-profiling and parameters for optimization are discussed in the ESI. Specifically, a FI-MRM-profiling method and its optimization for studying CAD is described. The sample preparation, solvent systems, and MS parameters were optimized and reproducibility was evaluated.

3.5 MRM-profiling

Pooled samples were used in the discovery phase of MRM-profiling to reduce effects of biological variability while averaging the signals relevant to the disease state. Pooling also provided a sufficient volume of sample without using all of any one individual sample. Pooled samples were prepared individually as described in Table S1. Data were acquired for each pool using the Prec and NL scans (Table S5). After data acquisition a list of transitions was created. Details on these points are found in the ESI along with an example of the Prec and NL parameters used during MassHunter MS data acquisition (Figure S8).

The software does not allow examination of more than 500 transitions per time segment. Thus, multiple injections must be made since 1266 transitions were discovered. This limits throughput and justifies use of a discrimination study to reduce the number of transitions. Discrimination has been described previously with MRM-profiling.¹² Data from a small representative number of samples was chosen and univariate statistics with broad parameters (*e.g.* p value = 0.1) was used to reduce the number of transitions in the study. This produced a final method that required <5 minutes per sample. In detail, for the discrimination study, a subset of 383 samples was randomly selected from the control, CAD and PAD samples and arranged in four, 96-well plates. These samples were prepared as described in Table S1 and analyzed using a method with four segments (3 positive ion mode and 1 negative ion mode). Each time segment (including the negative mode injection) included the alprazolam-D₅ transition (m/z 314.1→210.1). MassHunter Quantitative software Spectrum Summation feature was used to quickly integrate the response of each transition. Then each transitions' signal area was exported to Excel. Analytical outliers were eliminated so as not to skew the

data. This was done by analyzing the data by principal component analysis (PCA) to check if any of the plates appeared to be an outlier as a result of being prepared incorrectly. Next, batched by 96-well plate, alprazolam-D₅ signal was analyzed in individual samples. If the alprazolam-D₅ signal was an outlier then the sample was removed.

To determine which transitions were significant, the samples were analyzed by univariate statistics. In this step, broad, non-strict univariate parameters were used to avoid false negatives. The strategy here was to prefer to keep unimportant transitions (false positive) rather than risking the elimination of important ones.^{20, 21} The univariate statistics were performed with Metaboanalyst 3.0 and included t-test ($p=0.1$), fold change ($FC = 1.5$), and receiving operating characteristic (ROC, $AUC > 0.7$). The samples were grouped in several different ways to avoid loss of metabolites important to different populations. The groups compared to controls were (i) CAD, (ii) PAD, (iii) high low density lipoproteins (LDL), (iv) high triacylglycerides (TAG), (v) low high density lipoproteins (HDL), and (v) 'severe' CAD. Additionally, each of these groups were then separated by gender and filtered again. The transitions found to be significant for each of these comparisons were combined. Duplicate transitions were eliminated.

The final MRM method for the screening phase contained 485 transitions. These were measured in three consecutive injections. MassHunter Acquisition parameters are reported in Figure S9. Individual samples ($N=956$) were organized randomly into 96 well plates, prepared no more than twelve hours prior to injection, and kept at 4°C before and during data acquisition. The final method was 4.7 minutes injection to injection. In total, 956 samples were tested and data were collected over five days. Each plate was analyzed for analytical outliers as described above. If any were found, the sample was re-prepared and re-acquired in a later plate.

3.6 Data Analysis

The clinical data associated with the samples, as provided by the Fairbanks Institute, described the samples as 'case' or 'control'. The gender and age of the subject was always reported. Additional medical history, laboratory test results, and clinical information was sometimes reported. For analysis, biological outliers were removed by using this clinical data so as not to skew the data on the basis of other diseases or habits.²² Smokers, those with a history of cancer or stroke, and those with PAD were removed. The PAD samples were matched with random controls. No biological outliers were removed and no further grouping was performed with PAD samples because of the small sample set ($N=103$).

Data sets were analyzed using Mass Profiler Professional (MPP). The significant transitions between case and controls were found using univariate statistics ($p=0.05$ and $FC= 1.25$). The following tests and parameters were used for all univariate analysis: unpaired t-test, asymptotic p-value computation, Benjamini-Hochberg correction, and 1000 absolute count cutoff.

After univariate statistics, PLS-DA modeling was performed using the significant transitions. Using MPP, a cross validation study was done for each sample group by leaving one third of the samples out, modeling with two thirds, and classifying the one third left out based on a

model. This cross validation was done 100 times for each model and the accuracy of the classification was used to construct an averaged cross confusion matrix. This was used to calculate the sensitivity, specificity, and total accuracy of the method for the CAD and PAD samples.

3.7 LC-MS Methods and Identification Strategy

LC was used to identify compounds responsible for strong up or down regulation. A method previously described for lipid separation was adjusted for separation of lyso-lipids, phospholipids, and TAGs in the plasma extract.²³ An Agilent ZORBAX Eclipse Plus RRHD C18, 2.1 × 150 mm, 1.8 μm column was used. Isopropyl alcohol/methanol/water (5:1:4) with 5 mM ammonium acetate and 0.1 % acetic acid (Pump A) and isopropyl alcohol/water (99:1) with 5 mM ammonium acetate and 0.1 % acetic acid (Pump B) were used as mobile phases. The gradient and source conditions are reported in Figure S12.

Using this LC method, data were collected using the triple quadrupole over the whole concentration gradient for the transitions of interest. Using MassHunter Qualitative software an expected RT associated with for every transition was found. Once a RT had been determined, the LC system and pump method were implemented using a QTOF and a full scan HRMS method was performed. Using MassHunter Qualitative software the chromatographic peaks at the known RT were integrated, spectra extracted, and an exact mass for each precursor was determined. Next, data was collected using a targeted MS/MS method with a narrow isolation width (~1.7 *m/z*) which produced high resolution product ion data for each precursor ion. Agilent purine and 922 mass calibration sample were infused in the dual AJS source at 10 μL/min and used for mass calibration during data acquisition. This experiment provided the exact mass, RT, and high resolution product ion spectra for many precursors. With this information, PCDL, Metlin, HMDB, and LipidMaps databases were used to identify the transitions. Some of the transitions were not identified using this chromatographic method. Alternative mobile phases or columns may be needed to identify these further.

4. RESULTS & DISCUSSION

The discovery phase experiments found 1266 ion transitions in the plasma extracts, 1122 in positive mode and 144 in negative mode. A discrimination study was needed because of an instrumental limitation on transitions per scan. This was performed on a random subset of samples and it reduced the number of transitions to 485. Data for all these transitions were collected for the whole sample study (N=956). After analytical outliers were removed, 800 samples remained. After biological outliers were removed (*i.e.* smokers, stroke, and PAD), 535 CAD samples remained. A final grouping by age and gender was performed (Table 1). The 46–65 year old age range represented had the largest sample set. This age range, therefore, was analyzed for this study.

4.1 Modeling, Accuracy, Specificity, and Sensitivity

After the univariate statistics, the CAD samples from females between 46 and 65 years old had 62 significant transitions, the CAD samples from males between 46 and 65 years old

had 44 significant transitions, and the PAD samples had 55 significant transitions. Many transitions overlapped between the three groups and there was a total of 104 unique transitions.

The PLSDA models for female and male case verses control samples are shown in Figure 1. Females between 46–65 years old had a sensitivity of 95%, specificity of 87% and an accuracy of 90%. Males in the same age group had a sensitivity of 70%, specificity of 96% and an accuracy of 78% (Figure 2). The PAD model (Figure 3) had a sensitivity of 83%, a specificity of 86%, and an accuracy of 85% (Figure 2). Because of the small sample population biological outliers were not removed from the PAD population. However, the separation for this group is good even without removing the outliers.

The MRM-profiling results showed good separation with the PLSDA model. The PLSDA plots are not expected to show perfect separation of the groups because the disease is a spectrum. The sensitivity, specificity, and accuracy of the models are highly encouraging. Similarly, 100% accuracy is not expected because of the nature of the disease. The sample population analyzed consists of patients with different concurrent diseases, varying medical history, and differing therapies for CAD (and/or other diseases). All of these would cause differences in plasma metabolites. This is one explanation for the high number of false positives in the male samples (Figure 2). Reducing this variability was the motivation for blocking the population by age and removing biological outliers. Working with a clinician who is knowledgeable about the disease and patients would likely guide this blocking and improve the results.

4.2 Analyte Identification

The signal from the MRM-profile is sufficient for classification of a sample. However, exact chemical identification was desired to show that the metabolite profile consisted of signal relevant to a plasma sample and important to the disease. Many of the lipid transitions can be identified easily without chromatography because the Prec and NL scans are very specific. But many others cannot be identified by the QQQ transitions alone.

Retention time and HRMS data revealed that the 104 discriminating transitions consisted of 39 unique compounds (Table S4). Many compounds had a number of transitions (same precursor and different product ion) in the method and several of the lipids were represented by both the protonated and sodiated forms. The analytes were separated over a one hour gradient (Figure S13). Since the chromatographic method used was intended for lipids and triglycerides, the more polar analytes were not retained but many exact mass and product ion spectra were still collected. The lyso-phospholipids, PCs, sphingomyelin (SM) lipids, and TAGs separated clearly.

One of the most important findings of this study is that the separation of the samples is occurring on the basis of signals relevant to the plasma sample and to the disease. Choline and carnitine are upregulated in the CAD case samples. These have previously been reported to be altered in CAD populations and are taken as supplements by those with heart disease.^{24, 25} Another small metabolite, p-cresol-sulfate was also upregulated. This metabolite is found at higher concentrations in CAD patients with renal disease.²⁶ TAGs are well known

to be higher in populations with heart disease and many were upregulated in these case samples.²⁷ Lyso-chain PCs, SMs, and PC lipids are all down regulated in the case samples. This is indicative of cellular stress and inflammation which can be expected in CAD and PAD populations.^{10, 28–31} Trimethylamine N-oxide (TMAO) was not confirmed by HRMS but its transition (+ 76.0→58.0) matches those found in literature and, just as in other studies, it is up regulated in CAD.^{32, 33} Interestingly, cholesterol, an upregulated molecular biomarker for heart disease, was found to be down regulated in CAD and PAD populations. This is likely due to medications that the case population takes and the opposite would be expected for a population that is not taking cholesterol lowering medications.

Salicylic acid was found in the negative mode and was up regulated in PAD and CAD samples. It is the active ingredient in aspirin which is commonly taken by people with heart disease.³⁴ This is appropriate to include in the transition list and in modeling if the purpose of the MRM-profile is to study the signal profile of a population known to be on drugs (*e.g.* epidemiological studies). When using MRM-profiling for diagnostic purposes, transitions that may be drugs should be removed from the method prior to modeling. When this exogenous analyte was removed, the female group had a sensitivity of 93%, a specificity of 82%, and an accuracy of 86% and the male group had a sensitivity of 70%, a specificity of 91%, and an accuracy of 86%. Removing the drug did not alter the classification ability of MRM-profiling significantly in this study. Thus, the method continues to separate populations on endogenous metabolites and can be used for diagnostics or therapy monitoring.

The MRM-profiling methodology is easy to develop and implement. In a three month period, the workflow for MRM-profiling was integrated on the Agilent MS system, a method was optimized, and tested on 900+ samples. Much of the work reported in this article and ESI may be implemented directly in other labs so shortening this development time. Furthermore, the workflow was carried out in a high throughput fashion: eleven, 96 sample plates took less than five days to acquire data for 485 transitions.

Because of the all-inclusive measurement aim that this methodology embraces, tradeoffs are significant and parameters for every project should be evaluated to ensure that the most diverse signals (in terms of number of analytes) and those of the highest quality are being measured. This is especially true for the solvent system. Fortunately, the development experiments proved to be simple, straight forward, and fast because of the absence of chromatography in the first phases.

One major consideration is the selection of the Prec and NL scans to test (Table S5). All the scans should be performed to achieve an unsupervised methodology that promotes discovery with minimal bias. Many of these scans are highly specific (*e.g.* Prec 184 for PC lipids), however, others are less specific so exogenous analytes, like salicylic acid, might be found. If these less specific scans are not performed then it is possible to miss small metabolites (*e.g.* TMAO, choline, carnitine, and p-cresol-sulfate).

The discrimination study removed transitions thought to be unimportant before the full sample set was screened. To prevent eliminating important transitions at this step,

discrimination should be avoided in MRM-profiling studies. This could be done by choosing another instrument configuration, like direct infusion or nano-infusion, which continuously spray sample. Many samples in the final screening were eliminated as analytical outliers. This number could have been reduced by performing all the sample preparation pipetting steps with the Bravo liquid handler. Furthermore, protein contamination could be minimized by compacting the protein layer better. To do this, less plasma sample could be used and/or the plates could be centrifuged at a higher speed (20,000 RPM not 4,000 RPM). If less plasma is used, the extract dilution (200X) could be reduced.

Some analytes found by MRM-profiling have no RT and exact mass listed (Table S4) since they were not amenable to the chromatographic method. This demonstrates that MRM-profiling can measure more analytes in a single injection than traditional LC-MS methods which is advantageous to metabolomics discovery methods.^{2, 8} Furthermore, a few of the transitions which do have RT and exact mass were not found in the databases. Identification is not needed for classifying with MRM-profiling.

5. CONCLUSIONS

MRM-profiling of CAD successfully demonstrated this technique as a high throughput metabolomic method for biomarker discovery and sample classification. The pooled CAD and control plasma samples were analyzed for biologically relevant signal with Prec and NL scans. Then, over 900 individual samples were screened in under five days for the metabolite signals. The MRM-profile separated the CAD female, CAD male, and PAD samples from controls with high accuracy: 90%, 78%, and 85%, respectively. Furthermore, LC-MS methods showed that the characteristic signals discovered are related to the disease.

- MRM-profiling quickly identifies a metabolite profile with which to classify a population. It may be used for diagnosis or therapy evaluation.
- Biologically relevant signal is discovered in an unbiased manner and only that signal is targeted in an MRM method. Therefore, unimportant signal or background noise is not collected. By eliminating noise collection the method can be very fast. This also gives biological specificity to the data collection and results.
- This methodology is fast and simple to develop. Experiments described in the ESI Method Development section are straightforward but are essential for measuring meaningful signal in a high throughput and robust manner.
- Large sample sets can be screened very quickly.
- No libraries or standards are needed which reduces the cost of material and aids in the discovery of analytes based on functional groups and not a known exact structure.
- Some bias may be present in this method because of the Bligh Dyer sample preparation. Dilute and shoot methodologies are preferred. Here, we extracted lipids (chloroform phase) and polar metabolites (methanol and water phase) and combine them in the same injection. However, in future studies should evaluate

the sample preparation procedure to ensure optimal extraction and reconstitution of polar metabolites.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGEMENTS

The authors acknowledge the Fairbanks Institute for Healthy Communities (Indianapolis, IN) for the samples, Dr. Tommy Sors, Discovery Park, Purdue University, for arranging our use of these samples in this study and Agilent Technologies for their support and access to instrumentation. We acknowledge the support from the Indiana Clinical and Translational Sciences Institute (CTSI) who supplied the samples donated by the Fairbanks Institute for Healthy Communities. The CTSI is in part funded by Award Number UL1TR002529 from the National Institutes of Health, National Center for Advancing Translational Sciences, Clinical and Translational Sciences Award.

ABBREVIATIONS

6.

MS	Mass spectrometry
ESI	Electronic supplementary information
LC	Liquid chromatography
QQQ	Triple quadrupole
MS/MS	Tandem mass spectrometry
MRM	Multiple reaction monitoring
CE	Collision energy
V	Voltage
RSD	Relative standard deviation
TIC	Total ion current
Prec	Precursor
NL	Neutral Loss
HRMS	High resolution mass spectrometry
QTOF	Quadrupole time of flight
DI	Direct infusion
FI	Flow injection
AJS	Agilent jet stream
PLSDA	Partial least squares discriminant analysis
CAD	Coronary artery disease

PAD	Peripheral artery disease
MPP	Mass Profiler Professional
FC	Fold change
RT	Retention time
PC	Phosphatidylcholine
SM	Sphingomyelin
TAG	Triacylglyceride

9. REFERENCES

- González-Domínguez R, Sayago A and Fernández-Recamales A, *Bioanalysis*, 2017, 9, 131–148. [PubMed: 27921460]
- Aretz I and Meierhofer D, *International journal of molecular sciences*, 2016, 17.
- Wood PL, *Neuropsychopharmacology : official publication of the American College of Neuropsychopharmacology*, 2014, 39, 24–33. [PubMed: 23842599]
- Han X and Gross RW, *Mass Spectrometry Reviews*, 2005, 24, 367–412. [PubMed: 15389848]
- Sethi S and Brietzke E, *Prostaglandins & other lipid mediators*, 2017, 128–129, 8–16.
- Rath CM, Yang JY, Alexandrov T and Dorrestein PC, *Journal of the American Society for Mass Spectrometry*, 2013, 24, 1167–1176. [PubMed: 23568029]
- Anand S, Young S, Esplin MS, Peadar B, Tolley HD, Porter TF, Varner MW, D'Alton ME, Jackson BJ and Graves SW, *Journal of lipid research*, 2016, 57, 687–696. [PubMed: 26891737]
- Garcia-Sevillano MA, Garcia-Barrera T, Navarro F, Montero-Lobato Z and Gomez-Ariza JL, *Biomaterials : an international journal on the role of metal ions in biology, biochemistry, and medicine*, 2015, 28, 341–351.
- Prasain JK, Wilson L, Hoang HD, Moore R and Miller MA, *Metabolites*, 2015, 5, 677–696. [PubMed: 26569325]
- Basak T, Varshney S, Hamid Z, Ghosh S, Seth S and Sengupta S, *Journal of proteomics*, 2015, 127, 169–177. [PubMed: 25790721]
- Gromski PS, Muhamadali H, Ellis DI, Xu Y, Correa E, Turner ML and Goodacre R, *Analytica chimica acta*, 2015, 879, 10–23. [PubMed: 26002472]
- Ferreira CR, Yannell KE, Mollenhauer B, Espy RD, Cordeiro FB, Ouyang Z and Cooks RG, *Analyst*, 2016, 141, 5252–5255. [PubMed: 27517482]
- Cordeiro FB, Ferreira CR, Sobreira TJP, Yannell KE, Jarmusch AK, Cedenho AP, Lo Turco EG and Cooks RG, *Rapid Communications in Mass Spectrometry*, 2017, 31, 1462–1470. [PubMed: 28656689]
- Dhillon J, Ferreira CR, Sobreira TJP and Mattes RD, *Current Developments in Nutrition*, 2017, 1.
- Benjamin EJ, Blaha MJ, Chiuve SE, Cushman M, Das SR, Deo R, de Ferranti SD, Floyd J, Fornage M, Gillespie C, Isasi CR, Jimenez MC, Jordan LC, Judd SE, Lackland D, Lichtman JH, Lisabeth L, Liu S, Longenecker CT, Mackey RH, Matsushita K, Mozaffarian D, Mussolino ME, Nasir K, Neumar RW, Palaniappan L, Pandey DK, Thiagarajan RR, Reeves MJ, Ritchey M, Rodriguez CJ, Roth GA, Rosamond WD, Sasson C, Towfighi A, Tsao CW, Turner MB, Virani SS, Voeks JH, Willey JZ, Wilkins JT, Wu JHY, Alger HM, Wong SS and Muntner P, *Heart Disease and Stroke Statistics 2017 At-a-Glance*, American Heart Association Statistics Committee and Stroke Statistics Subcommittee, 2017.
- Cardiovascular diseases (CVDs) Fact sheet, <http://www.who.int/mediacentre/factsheets/fs317/en/>.
- Coronary artery disease, <https://www.mayoclinic.org/diseases-conditions/coronary-artery-disease/diagnosis-treatment/drc-20350619>.
- Morand KL, *High Throughput Flow Injection Analysis- Mass Spectrometry*, Elsevier, 2004.

19. Patterson RE, Ducrocq AJ, McDougall DJ, Garrett TJ and Yost RA, *Journal of chromatography. B, Analytical technologies in the biomedical and life sciences*, 2015, 1002, 260–266. [PubMed: 26343017]
20. Banerjee A, Chitnis UB, Jadhav SL, Bhawalkar JS and Chaudhury S, *Indian Journal of Psychiatry*, 2009, 18, 127–131.
21. Rothman KJ, *European Journal of Epidemiology*, 2010, 25, 223–224. [PubMed: 20232112]
22. Mueller DC, Piller M, Niessner R, Scherer M and Scherer G, *Journal of proteome research*, 2014, 13, 1602–1613. [PubMed: 24354774]
23. Sartain M and Sana T, *Journal*, 2015.
24. Wang ZY, Liu YY, Liu GH, Lu HB and Mao CY, *Life Science*, 2018, 194, 88–97.
25. Shah SH, Kraus WE and Newgard CB, *Circulation*, 2012, 126, 1110–1120. [PubMed: 22927473]
26. Meijers BJI, Claes K, Bammens B, de Loo H, Viaene L, Verbeke K, Kuypers D, Vanrenterghem Y and Evenepoelcorresponding P, *Clinical Journal of the American Society of Nephrology*, 2010, 5, 1182–1189. [PubMed: 20430946]
27. Austin MA, *Arteriosclerosis, Thrombosis, and Vascular Biology*, 1991, 11, 2–14.
28. Byeon SK, Lee JY, Lim S, Choi D and Moon MH, *Journal of Chromatography A*, 2012, 1270, 246–253. [PubMed: 23195705]
29. Cui S, Li K, Ang L, Liu J, Cui L, Song X, Lv S and Mahmud E, *JACC Cardiovasc Interv*, 2017, 10, 1307–1316. [PubMed: 28624380]
30. Sutter I, Klingenberg R, Othman A, Rohrer L, Landmesser U, Heg D, Rodondi N, Mach F, Windecker S, Matter CM, Luscher TF, von Eckardstein A and Hornemann T, *Atherosclerosis*, 2016, 246, 130–140. [PubMed: 26773473]
31. Ganna A, Salihovic S, Sundstrom J, Broeckling CD, Hedman AK, Magnusson PK, Pedersen NL, Larsson A, Siegbahn A, Zilmer M, Prenti J, Arnlov J, Lind L, Fall T and Ingelsson E, *PLOS Genetics*, 2014, 10, e1004801. [PubMed: 25502724]
32. Wang Z, Levison BS, Hazen JE, Donahue L, Li XM and Hazen SL, *Analytical Biochemistry*, 2014, 455, 35–40. [PubMed: 24704102]
33. Guasch-Ferre M, Hu FB, Ruiz-Canela M, Bullo M, Toledo E, Wang DD, Corella D, Gomez-Gracia E, Fiol M, Estruch R, Lapetra J, Fito M, Aros F, Serra-Majem L, Ros E, Dennis C, Liang L, Clish CB, Martinez-Gonzalez MA and Salas-Salvado J, *Journal of the American Heart Association*, 2017, 6.
34. Würtz M, *Danish Medical Journal*, 2015, 2015, 4.
35. Beckmann M, Parker D, Enot DP, Duval E and Draper J, *Nature Protocols*, 2008, 3, 486–504. [PubMed: 18323818]
36. Cajka T and Fiehn O, *Metabolomics : Official journal of the Metabolomic Society*, 2016, DOI: 10.1007/s11306-015-0929-x, 12–34.
37. Murphy RC, Fiedler J and Hevko J, *Chemical Reviews*, 2001, 101, 479–526. [PubMed: 11712255]
38. McAnoy AM, Wu CC and Murphy RC, *Journal of the American Society for Mass Spectrometry*, 2005, 16, 1498–1509. [PubMed: 16019221]
39. Li M, Butka E and Wang X, *Scientific Reports*, 2014, 4, 6581. [PubMed: 25301200]
40. Milne S, Ivanova P, Forrester J and Alex Brown H, *Methods*, 2006, 39, 92–103. [PubMed: 16846739]
41. Lieser B, Liebisch G, Drobnik W and Schmitz G, *Journal of lipid research*, 2003, 44, 2209–2216. [PubMed: 12897185]
42. Colsch B, Afonso C, Popa I, Portoukalian J, Fournier F, Tabet JC and Baumann N, *Journal of lipid research*, 2004, 45, 281–286. [PubMed: 14595000]
43. Gross RW and Han X, in *Lipidomics and Bioactive Lipids: Specialized Analytical Methods and Lipids in Disease*, 2007, DOI: 10.1016/s0076-6879(07)33004-8, pp. 73–90.
44. BRUGGER B, ERBEN G, SANDHOFF R, WIELAND FT and LEHMANN WD, *Proceedings of the National Academy of Sciences of the United States of America*, 1997, 94, 2339–2344. [PubMed: 9122196]
45. Han X, Yang K, Yang J, Cheng H and Gross RW, *Journal of lipid research*, 2006, 47, 864–879. [PubMed: 16449763]

46. Merrill AH, Jr., Sullards MC, Allegood JC, Kelly S and Wang E, *Methods*, 2005, 36, 207–224. [PubMed: 15894491]
47. Ma Y, Kind T, Yang D, Leon C and Fiehn O, *Analytical chemistry*, 2014, 86, 10724–10731. [PubMed: 25263576]
48. Delgado de la Torre MP, Ferreiro-Vera C, Priego-Capote F and Luque de Castro MD, *Journal of Agricultural and Food Chemistry*, 2013, 61, 12539–12548. [PubMed: 24295391]
49. Taguchi R, Houjou T, Nakanishi H, Yamazaki T, Ishida M, Imagawa M and Shimizu T, *Journal of chromatography. B, Analytical technologies in the biomedical and life sciences*, 2005, 823, 26–36. [PubMed: 15990370]
50. Isaac G, Jeannotte R, Esch SW and Welti R, *Genetic Engineering*, 2007, 28, 129–157. [PubMed: 17153937]
51. Liebisch G, Binder M, Schifferer R, Langmann T, Schulz B and Schmitz G, *Biochimica et Biophysica Acta*, 2006, 1761, 121–128. [PubMed: 16458590]
52. Christie WW, *Journal Lipid Research*, 1985, 26, 507–512.

Figure 1.

CAD PLSDA models of case (red) and control (yellow) samples made with filtered transition sets in MPP. The female samples (top) between 46–65 years old separated on the basis of 62 transitions and the male samples (bottom) in the same age group separated on the basis of 44 transitions.

		Predicted			
		Case	Control	Accuracy:	
Actual	Case	61	3	Sensitivity:	95%
	Control	13	93	Specificity:	88%

A

		Predicted			
		Case	Control	Accuracy:	
Actual	Case	106	44	Sensitivity:	71%
	Control	2	57	Specificity:	97%

B

		Predicted			
		Case	Control	Accuracy:	
Actual	Case	86	17	Sensitivity:	83%
	Control	14	93	Specificity:	87%

C

Figure 2.

Cross confusion matrices for female (A) and male (B) case verses control (age 46 to 65) and PAD verses control (C). Sensitivity, selectivity, and accuracy for the models calculated from the PLSDA cross validation are reported.

Figure 3. PLSDA model showing the separation of PAD samples (red) and control samples (yellow) using 55 transitions.

Table 1CAD cases and controls sorted by age and gender (N=535)^a

	F		M	
	Case	Control	Case	Control
Total Samples:	107	143	223	81
Age Range:	37–84	39–85	22–82	33–78
22–45 year old:	8	12	17	9
46–65 year old ^b :	64	106	150	59
66–85 year old:	34	18	46	12

^aSmokers, those with a history of stroke or cancer, and PAD samples were removed.^bThis age group was used for further analysis

Table 2

Summary of identified analytes, their exact mass, and if they are up-regulated (↑) or down-regulated (↓) in female or male CAD or PAD samples

ID	Exact mass (<i>m/z</i>)	Mass error (ppm)	F CAD	M CAD	PAD
TMAO	N/A	N/A	↑	↑	↑
Choline	(+) 104.1071	−3.8			↑
Octylamine	(+) 130.1586	−7.7	↑	↑	↑
Carnitine	(+) 162.1123	−4.3	↑		↑
γ-hydroxy-L-homoarginine	(+) 205.1271	−14.6	↑		↑
Cholesterol	(+) 369.3511	6.8	↓		↓
MonoChain-PC	(+) 520.3371	−6.1	↓	↓	
MonoChain-PC	(+) 521.3427	13.6	↓	↓	↓
MonoChain-PC	(+) 542.3215	−5.9	↓	↓	
20:5 Cholesteryl ester	(+) 671.5742	−3.7	↓	↓	↓
SM(34:1)	(+) 725.5563	−1.4	↓	↓	
PC(34:2)	(+) 759.6366	−1.2	↓	↓	
SM(38:1)	(+) 759.5731	−1.8	↓	↓	
SM(40:2)	(+) 785.6530	−0.9	↓		
PC(36:2)	(+) 786.6005	−1.0	↓		
SM(40:1)	(+) 787.6692	−0.1	↓		
PC(37:4)	(+) 796.5239	43.4	↓	↓	
PC(38:6)	(+) 806.5662	−4.7	↓	↓	
TAG	(+) 848.7687	−2.4		↑	
TAG	(+) 850.7868	0.5	↑	↑	
TAG	(+) 851.7101	−3.3		↑	
TAG	(+) 874.7828	−4.1		↑	
TAG	(+) 876.7995	−2.9	↑	↑	
TAG	(+) 877.803	−22.1	↑	↑	
TAG	(+) 878.7289	5.8		↑	
Salicylic acid	(−) 137.0252	4.4	↑	↑	↑
p-cresol sulfate	(−) 187.0064	−0.5		↑	↑