Validation and automation of a high-throughput multi-targeted method for semi-quantification of endogenous metabolites from different biological matrices using tandem mass spectrometry

Jatin Nandania [1+], Gopal Peddinti [2#], Alberto Pessia [1&], Meri Kokkonen [1§], Vidya Velagapudi [1*]

[1] Metabolomics unit, Institute for Molecular Medicine Finland (FIMM), University of Helsinki, HiLIFE, Tukholmankatu 8, Biomedicum 2U, 00290, Helsinki, Finland.
[2] Computational Systems Medicine group, University of Helsinki, 00290, Helsinki, Finland.

Present addresses:

+Roche Diagnostics GmbH, Nonnenwald 2, 82377 Penzberg, Germany.
#VTT Technical Research Center of Finland, P.O Box 1000, 02044 Espoo, Finland.
&Network Pharmacology for Precision Medicine group, University of Helsinki, 00290, Helsinki, Finland.
§Finnish Customs Laboratory, Tekniikantie 13, 02150 Espoo, Finland.

* Corresponding author: vidya.velagapudi@helsinki.fi; Tel.: +358 50 317 5087

Abstract: The use of metabolomics profiling to understand metabolism under different physiological states has increased in recent years, which created the need for robust analytical platforms. Here, we present a validated method for targeted and semi-quantitative analysis of 102 polar metabolites that covers major metabolic pathways from 24 classes in a single 17.5-min assay. The method has been optimized for a wide range of biological matrices from various organisms, and involves automated sample preparation, and data processing using in-house developed R package. To ensure reliability, the method was validated for accuracy, precision, selectivity, specificity, linearity, recovery, and stability according to European Medicines Agency guidelines. We demonstrated excellent repeatability of the retention times (CV<4%), calibration curves ($R^2 \geq 0.980$) in their respective wide dynamic concentration ranges (CV<3%), and concentrations (CV<25%) of quality control samples interspersed within 25 batches analyzed over a period of one-year. The robustness was demonstrated through high correlation between metabolite concentrations measured using our method and NIST reference values ($R^2=0.967$), including cross-platform comparability against the BIOCRATES AbsoluteIDQp180 kit ($R^2=0.975$) and NMR analyses ($R^2=0.884$). We have shown that our method can be successfully applied in many biomedical research fields and clinical trials, including epidemiological studies for biomarker discovery. In summary, a thorough validation demonstrated that our method is reproducible, robust, reliable, and suitable for metabolomics studies.

Keywords: High-throughput; targeted; semi-quantitation; metabolomics; LC-MS; multi-analyte method; validation; cross-platform comparability; automation; biomarkers.

1. Introduction

Metabolomics has a great influence on many disciplines, as metabolites are intermediates or end products of cellular functions. Hence, metabolomics can be used as a powerful tool to generate data for understanding, diagnosing, and managing different pathophysiological conditions. It is therefore essential to be able to identify and measure metabolites from different biological matrices [1]. Although global metabolomics has been widely used in discovery studies for understanding cellular responses to normal and abnormal biological conditions, targeted metabolomics has more advantages for addressing biological questions in a more hypothesis-driven manner than global untargeted metabolomics [2]. Furthermore, targeted metabolomics can quantify metabolites that are low in abundance, which are difficult to assess using an untargeted approach.

An appropriate sample pretreatment is required to obtain reproducible and high quality quantitative data in targeted metabolomics. However, metabolites are present in a wide dynamic range with great diversity in physicochemical properties in the biological matrices [3]. Recent advancements in extraction techniques and automated approaches for sample preparation have partially satisfied the demands of targeted metabolomics. However, there are still many outstanding challenges, such as the matrix effect and laboratory-to-laboratory variation associated with sample preparation. Hence, standardization of sample preparation is a fundamental requirement in metabolomics studies [4].

Secondly, robust analytical methodology is required for accurate quantification of metabolites with good reproducibility over an extended period of time [5]. Molecular diversity is a major problem that hinders the separation of all pre-selected metabolites in a single chromatographic run and the detection of all separated metabolites with minimum technical variation [6]. Tandem mass spectrometry (MS) is a technique used predominantly due to its high sensitivity and high throughput for the detection of metabolites. A combination of MS and separation techniques is used to increase the sensitivity and reliability of analytical methods for the analysis of metabolites from complex biological matrices [7]. In addition, latest developments in triple quadrupole instrumentation strengthened the possibilities to develop multi-analyte methods in a single injection that yield reliable and quantitative data [8].

Even though liquid chromatography-mass spectrometry (LC-MS) is a method of choice in targeted metabolomics, obtaining accurate quantification and long-term data reproducibility remains an analytical challenge. This is due to limitations such as matrix effect, MS performance drift, and LC column contamination and aging [9]. Finally, although instrument vendor software for data processing provides some crucial functions such as peak integration (rendering the data from high-throughput metabolomics experiments into numerical values that represent metabolite concentrations), the lack of automation in downstream data processing and quality control remains a major bottleneck in high-throughput analyses. Thus, an efficient pipeline is necessary to enable rapid, accurate, and standardized processing of these data. Such a pipeline should facilitate automated analyses, while at the same time allowing the user to fine-tune the parameters for accurate data processing. Taken together, there is a need for development and validation of standardized, robust, and quantitative methods for large-scale targeted metabolomics studies in a high-throughput manner to minimize bias associated with sample preparation and the analytical technique used.

We have previously developed a robust, reproducible, and high-throughput targeted method for measuring 102 polar metabolites from various biological classes semi-quantitatively in a single injection [10,11]. The metabolites were selected according to the following criteria: have important roles in many biological processes, are known biomarkers in several diseases and technical feasibility for developing an analytical method that covers all the metabolites in a single assay. The selected metabolites come from 24 different classes (Table 1), covering a wide range of metabolic pathways. We created an in-house metabolite database by manually curating all the available information (i.e., names, HMDB, PUBCHEM, KEGG Ids, chemical properties, reported normal and abnormal concentration ranges, links to their structures) from the Human Metabolome Database (HMDB).

The selected metabolites were separated by using Hydrophilic Interaction Liquid Chromatography (HILIC) and measured using triple quadrupole mass spectrometry (MS). Briefly,

after addition of an internal standards working solution, biofluid samples were extracted with 1% formic acid in 99% acetonitrile. Tissue samples were extracted using homogenization in two steps. Extraction was first performed with 1% formic acid in 99% acetonitrile; the second extraction used 80/20% ACN/H_2O+1% formic acid by protein precipitation followed by filtration through Ostro plates. After this step, 5 µL of filtered sample extract was injected into an Acquity UPLC system coupled to a Xevo® TQ-S triple quadrupole mass spectrometer (Waters Corporation, Milford, MA, USA), which was operated in both positive and negative polarities with a polarity switching time of 20 msec for metabolite separation and quantification. Multiple Reaction Monitoring (MRM) acquisition mode was selected for metabolite quantification. A detailed method description and instrument parameters are provided elsewhere [10,11]. Our method involves automated sample preparation for biofluid samples and minimal manual steps for non-biofluid samples. This reduces the analysis time and inter-batch variation and minimizes human error. Semi-quantification of the metabolites was performed using external calibration curves ($R^2 \geq 0.980$) in their respective wide dynamic concentration ranges and 12 labeled internal standards were used for most of the metabolites to minimize matrix effects.

Several analytical methods have also been developed for semi-quantitative measurement of large number of metabolites in a single run [12-16]. Li et al [12] measured 610 metabolites from 60 biochemical pathways. However, their method requires approximately 40 minutes to separate all the metabolites (with poor resolution for some metabolites). Furthermore, their method was not applicable to all biological matrices. Similarly, Wei et al [15] detected approximately 200 metabolites in 10 minutes from plasma. However, their method lacks information on validation and for other biological matrices. Yuan et al [16] also provided a protocol for detection of approximately 250 metabolites in a single method. However, they did not use an automated platform and a drying step in their method increases the analysis time of the method overall. Our method has clear advantages compared to other published methods, including shorter analysis time, high throughput, automation, and semi-quantitation using individual external 11-point calibration curves along with 12 labeled internal standards. Furthermore, the method has been used in many biomedical and clinical studies for biomarker discovery [10,11,17-37].

The primary objective of this work is to show the robustness of our method through a thorough validation of our previously developed analytical method according to European Medicines Agency (EMA) guidelines. We also demonstrate automation of tedious and manual data processing tasks in high-throughput metabolomics analyses using an in-house developed R package. The R package automates various corrections and normalization steps to convert the raw peak area data to molecular concentrations for each compound in each sample and also provides quality evaluation of the data and reduces the manual work load significantly.

2. Materials and Methods

2.1. Chemicals and reagents

All metabolite standards were purchased from Sigma-Aldrich (St. Louis, MO, USA). Internal standards were ordered from Cambridge Isotope Laboratory. Inc. (Tewksbury, MA, USA). LC-MS-grade solvents, 2-proponol, acetonitrile, and methanol (HiPerSolv) were obtained from VWR International (Helsinki, Finland). Analytical-grade chemicals (formic acid, ammonium formate, and ammonium hydroxide) were obtained from Sigma-Aldrich. Deionized water (18 MΩ.cm at 25°C) used for solution preparation was made using a Milli-Q water purification system (Bamstead EASYpure RoDi ultrapure water purification system, Thermo scientific, Ohio, USA). Mouse tissues, including heart, liver, brain, spleen, and muscles were obtained from Innovative Research Laboratory (Novi, MI, USA). Whole blood was obtained from the Finnish Red Cross blood service (Helsinki, Finland), from which serum was prepared during method optimization and validation. Cell samples

were provided by our research collaborators. NIST Standard reference material (SRM) 1950 plasma was purchased from Sigma-Aldrich (Gillingham, UK).

2.2. Metabolite extraction protocol and instrumentation

All metabolites were extracted, separated with HILIC chromatography (Acquity BEH amide, 2.1 X 100 mm, 1.7µ), and analyzed with a Waters Xevo TQ-S mass spectrometer using our previously published protocol [11]. The protocol for tissues and adherent cells was optimized for better recovery and chromatography and to cover a wide range of tissue and cell types with a single protocol. For tissue sample extractions, 90/10% ACN/H2O + 1% formic acid was used instead of 80/20% ACN/H2O + 1% formic acid during the second step of extraction. Additionally, during cell pellet sample extraction, 80/20% ACN/H2O+1% formic acid was replaced with 90/10% ACN/H2O + 1% formic acid. After optimization, we used the tissues protocol for analysis of various biological matrices, such as heart, liver, placenta, brain, muscles, spleen, *C. elegans*, *Drosophila* larvae, dental carries, dried blood spots, and fecal samples. The cell pellet protocol was used for all types of adherent cells and *E. coli* and *S. cerevisiae* samples. The biofluid protocol was used for all types of biofluids, such as blood, plasma, serum, cell culture supernatant, CSF, and urine.

2.3. Method Validation

Validation of the method was performed to verify various parameters and the reliability of the developed method for analysis of a large number of samples. The method was validated according to EMA guidelines for bioanalytical method validation in terms of selectivity, specificity, linearity, accuracy, precision, extraction recovery, matrix effect, and stability [38]. In addition, we used pooled healthy human serum samples as internal quality control (QC) samples in all studies to correct signal drift during sample runs and to improve confidence in the statistical data. QC samples at high, medium and low (for serum) or high and low concentration levels (for tissues) were prepared by spiking a mixed standard solution in their respective homogenized biological matrices to perform all the method validation experiments. We performed validation for commonly used biological samples in metabolomics analyses, such as biofluid (serum), tissue (liver, brain and spleen), and cell samples. An aqueous calibration curve was used to calculate the concentration values during the method validation. The instrument performance for response reproducibility and sensitivity was always verified by six consecutive injections of medium concentration solution at the start of any experiment.

2.3.1. Selectivity and specificity

The selectivity and specificity for each metabolite were investigated using serum-spiked samples (N=6) with known amount of standard. Chromatographic interferences from other endogenous compounds of the biological matrix at the retention time of the target analyte for a particular metabolite were verified. The chromatographic peaks from spiked samples were compared with the standards by the retention times and if required from their respective MRM spectra.

2.3.2. Linearity, accuracy, and precision

To assess the linearity, accuracy, and precision, six replicates of spiked QC samples at high, medium, and low concentrations along with calibration curve were injected on three separate days. Calibration curve standards of 11 points were prepared via serial dilution. The curve was plotted by using the peak area response ratios (standard/labeled standard) versus the concentrations of the individual metabolites. The calibration curve was constructed using the regression equations (linear and quadratic) by applying appropriate weighing factor and by transforming the axis (both instrument response and theoretical concentration) into logarithmic or square root function. The

accuracy was calculated as measured value divided by the nominal value at each concentration level of the calibration curve standards in all three batches. Inter- and intra-batch variability was calculated by measuring coefficient of variation (%CV) at each QC concentration level.

2.3.3. Recovery and matrix effect

The recovery efficiencies for each analyzed metabolite was determined by comparing analytical results from QC samples spiked with standards mixture before and after extraction using different concentrations. The spike concentrations covered the calibration range. The matrix effect (percentage of ion suppression or enhancement of the MS signal) was determined by comparing the analytical response of QC samples that were spiked after extraction with the analytical response of aqueous spiked samples (diluent spiked with respective concentrations of QCs). Since there were endogenous metabolites, we subtracted the endogenous concentrations from the samples that were spiked. This experiment was performed using six QC replicates.

2.3.4. Stability of the metabolites

Wet extract, freeze-thaw, and stock solution stability for all metabolites were determined to check the integrity of the analytes in solvents and in QC samples at different conditions. To determine the wet extract stability, six replicates of extracted QC samples were kept in the auto-sampler at 5°C. The same samples in the same sequence were reinjected with freshly extracted QC samples and the results were compared.

Freeze-thaw stability was evaluated up to three cycles by freezing and thawing the spiked QC samples stored at -80°C and comparing the concentrations against the freshly thawed and spiked QC samples.

Long-term stock solution stability for metabolite stock solutions and intermediate solutions were checked by comparing the mean peak area of freshly prepared solutions with stored solutions at 4°C. All stability experiments were performed with six replicates of QCs.

2.3.5. Sample carry-over

Sample carry-over was evaluated by injecting the highest standard concentration (ULOQ) of the metabolites in the calibration curve followed by a series of blank injections and lowest standard concentration (LLOQ). The blank samples were evaluated for any signal at the retention time of particular metabolites and signal intensities of the blank samples were compared with the LLOQ samples. The acceptance criteria for carry-over was set at 20% of the peak area corresponding to the LLOQ level as per the EMA guidelines for bioanalysis.

2.3.6. Quality control samples

Internal QC samples were prepared after separating serum from pooled healthy human blood samples. A volume of 350 µL of serum was aliquoted and stored at -80°C after providing a lot number and QC number. The concentration of QC samples that were incorporated in batches during the metabolomics studies were calculated for all the metabolites along with the experimental samples. Average concentrations (µmol/L) and %CV of the QC samples were calculated for each metabolite. The data were saved along with QC lot numbers, batch name, and run date. The QC data were collected from six different lots for a period of 5.5 years (N=539 replicates). An internal QC database has been maintained and used for quality checks.

2.3.7. Comparison with reference material

To evaluate the performance of our semi-quantitative method, commercially available standard reference plasma (NIST SRM 1950) [39] was analyzed using our method (N=8 replicates). The concentration values from the matched 17 metabolites were compared with the given standard reference values.

2.3.8. Cross-platform comparison

To further evaluate the robustness and performance of our method, we performed a cross-platform comparison using two completely different analytical platforms: (1) the commercially available AbsoluteIDQ p180 targeted metabolomics assay kit using LC-MS/MS and (2) a nuclear magnetic resonance (NMR) platform. We sent our internal QC samples to the BIOCRATES Life Sciences AG (Innsbruck, Austria) (N=3 replicates) and to the NMR Metabolomics Laboratory, School of Pharmacy, University of Eastern Finland (Kuopio, Finland) (N=3 replicates). Our QC samples were extracted and analyzed as described previously for the AbsoluteIDQ p180 kit [40] and for the NMR analysis of small molecules [41]. We compared these results with the results obtained from our method (N=4-5 replicates).

2.4. Statistical analyses

To estimate the median concentration values (μmol/L) of the metabolites from QC samples (N=539), we fitted a linear mixed model with the MCMCglmm R package [42] using an expanded parameter formulation and default settings. Observed data was assumed to be log-normally distributed and corrected for the six different QC lots. Credibility intervals (95%) of median concentration values were computed from 20 000 samples of the posterior distribution. Error bars shown in the scatter plots are 95% confidence intervals, while the coefficient of determination R^2 refers to the simple linear regression between the two plotted variables. Coefficient of variation (CV) percentages were calculated as a measure of variability. To automate the downstream processing of the data produced by the instrument vendor software (TargetLynx), we built a data processing package called "Unlynx" in R statistical programming language.

2.5. Automated data processing

The "Unlynx" package parses the output of TargetLynx software (i.e., raw data containing the concentration values in PPB units) and produces a processed dataset in an Excel spreadsheet after performing a series of preprocessing operations.
The preprocessing steps included the following:

(i) Molecular weight normalization, in which the ppb values are normalized by the molecular weight of each compound, thereby converting the data from ppb units to μmoles.
(ii) Process efficiency correction for the semi-quantification of metabolites without internal standards.
(iii) Normalization using dilution factor for specific sample type if dilution was needed.
(iv) Cell number normalization (for cell samples) to convert the concentration values per million cells.
(v) Calculation of mean, standard deviation, and relative standard deviation (RSD) of molecular concentrations (resulting from the previous steps) for each phenotypic group.
(vi) Outlier detection in each phenotypic group; if the concentration value of a compound in a sample is more than one or two standard deviation (SD) away from the mean of the phenotypic group, then it is marked as an outlier in the Excel data set in two different colors.
(vii) QC check by comparing the RSD of QC samples in the current dataset against the internal database of QC sample RSDs (based on inter-day RSDs recorded over one year).

3. Results and Discussions

3.1. Extraction method optimization

The primary objective of this work was to optimize and validate our previously published protocol for different types of biological matrices. For tissue samples (placenta, liver, heart, brain, spleen, and muscles), the sample volumes of the tissues and extraction solvent volumes were optimized to fit the concentrations of most of the metabolites within the linearity of calibration curve for reliable results. We observed that most of the metabolites can be semi-quantified within the calibration curve range with 20±5mg of sample weight.

Furthermore, we optimized the protocol with extraction solvent for tissues and adherent cells. Some of the metabolites (in particular inositol, GABA, asymmetric dimethylarginine, symmetric dimethylarginine, spermidine, ribose-5-phosphate, and orotic acid) had poor separation and irreproducible chromatography. Interference of isobaric compounds with other metabolites was also observed due to poor separation. Thus, different compositions of the extraction solvent were assessed to achieve the acceptable chromatography. We observed that modification of acetonitrile content from 80% to 90% and applying longer equilibration time for the HILIC column yielded acceptable chromatography and also good separation for most of the metabolites.

We also optimized the extraction protocol for different sample types from various organisms, such as tissue types (adipose, endometrium, testicles, lung), fecal samples, dried blood samples, dental carries, biofilm, extracellular vesicles, mitochondrial isolates, drosophila, *C. elegans*, *E. coli*, and *S. cerevesiae* samples (Figure 1).

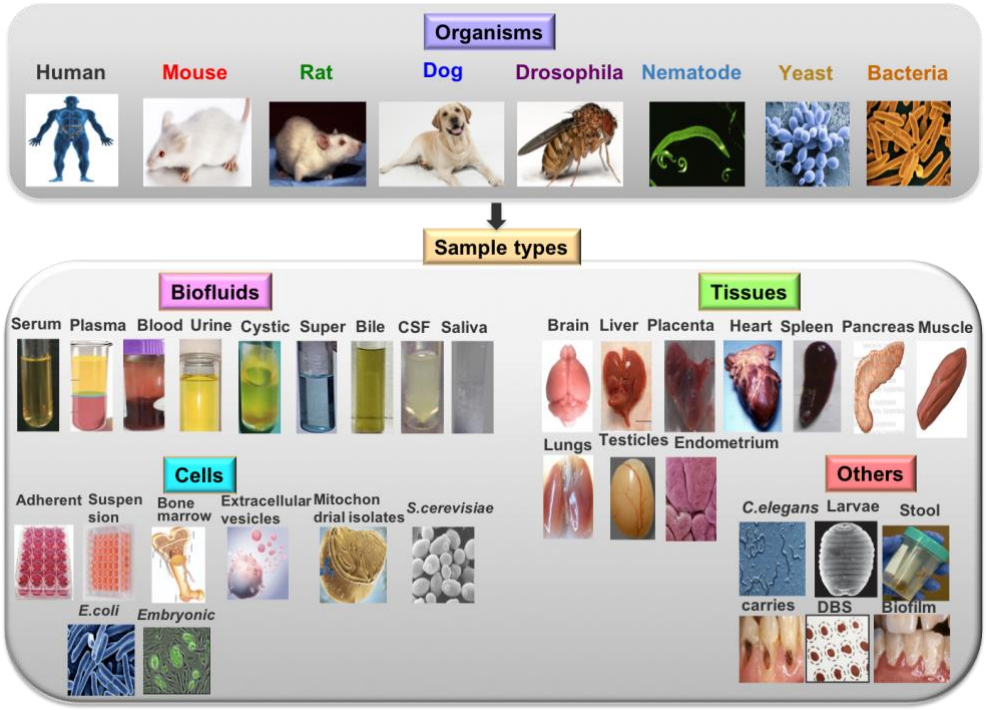

Figure 1. Illustration of different sample types from various organisms that have been used for extraction method optimization.

3.2. Method Validation

3.2.1. Selectivity and specificity

There were no significant interference peaks from the matrix components in their respective retention time windows, indicating the selectivity of the metabolites in our method. We repeated the injections for five times from all different serum samples and confirmed that the peak eluted was only from the target analyte, indicating that they are specific to their corresponding MRM transitions (Figure 2).

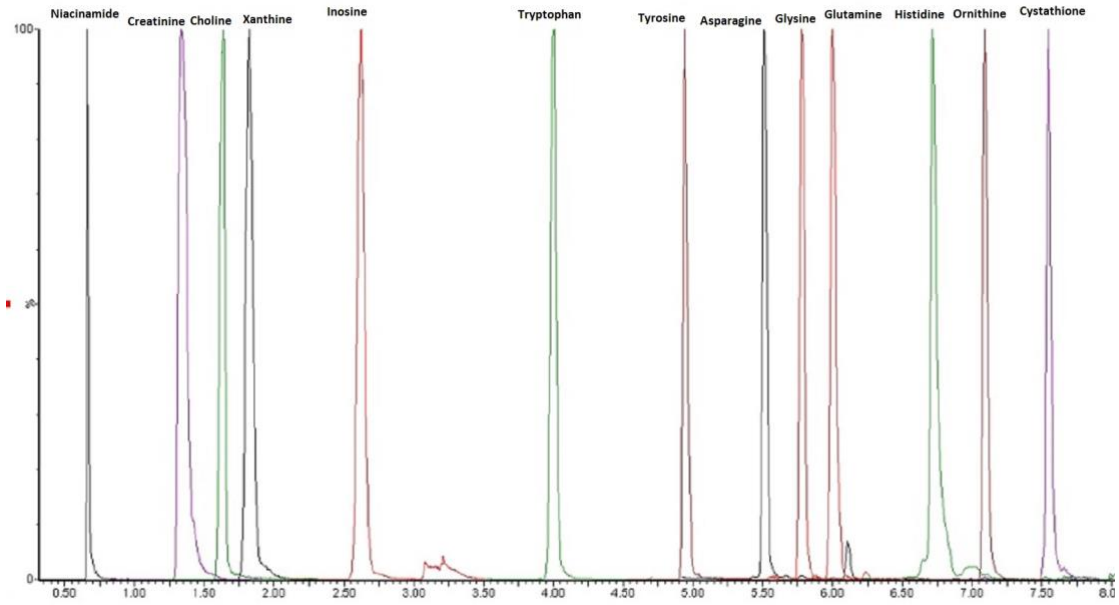

Figure 2. Representative chromatographic peaks of selected metabolites at their respective retention times in the QC serum sample.

3.2.2. Linearity, accuracy, and precision

To cover a broad concentration range, the linear or quadratic models were used and the variables from X and Y axes were logarithm or squareroot transformed to fit the calibration data [43]. The coefficient of determination (R^2) value for each metabolite was greater than 0.980 at their respective concentration range, except for some metabolites such as aspartate, uracil, 2-deoxyuridine sucrose, and chenodeoxycholic acid (likely due to their broad peak shapes and poor recovery at lower concentration, Table S1).

Concentration precision for QC samples was calculated by measuring %CV at high, medium, and low concentration level of QCs (N=6 replicates). In general, intra- and inter-day precision (CV) values were within 15% for most of the metabolites except acetoacetic acid, folic acid, sucrose, homoserine, 2-deoxyuridine, and cholic acid (Figure 3). However, at low concentrations, more than 20% CV were observed for NAD and myo-inositol. This might be due to the low recoveries of these compounds.

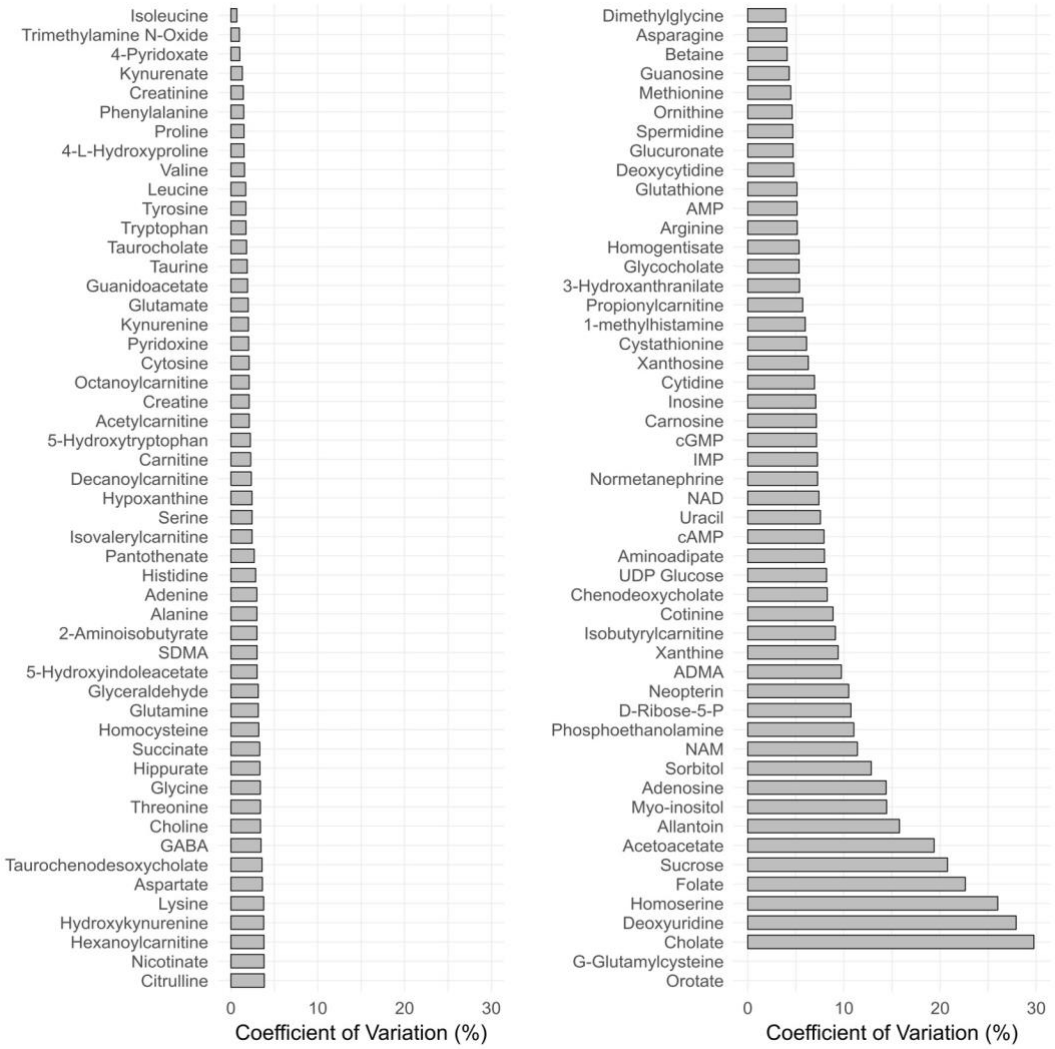

Figure 3. Inter-batch variation of high concentrations of spiked QC serum samples (N=6) analyzed on 3 different days. Data not shown for g-glutamylcysteine and orotic acid (CV>30%).

3.2.3. Recovery and matrix effect

For most of the metabolites, recoveries were found to be between 50% to 120% with good repeatability at all three concentrations levels (low, medium, and high) in both biofluid (serum) and tissues (brain, liver, and spleen). However, compounds such as UDP-glucose, IMP, cGMP, D-ribose 5-phosphate, NAD, AMP, homocysteine, carnosine, and glutathione had recoveries less than 30% in serum. However, CV of recoveries at low, medium, and high concentration levels were within 25% except for cGMP in serum. Metabolites such as 1-methylhistamine, aspartate, glutamine, adenosine, and glutathione had recoveries over 120%. Histidine, ornithine, cystathionine, 3-OH-DL-kynurenine, carnosine, AMP, NAD, cGMP, IMP, and UDP-glucose had less than 30% recovery in some tissues types, indicating matrix effect or degradation (Figure 4). However, CV of repeatability at all concentration levels were within 15% except for 1-methylhistamine, aspartate, glutamine, glutathione, NAD, and UDP-glucose in some tissues.

The matrix effect values were observed to be within the range of 0.6 to 1.8 (below 1 indicates ion suppression and above 1 indicates ion enhancement) for the metabolites in serum and tissues. The challenge of the matrix effect can be overcome by having individual isotope-labeled internal standards for each individual compound for true quantification. However, this is not practically possible for high-throughput metabolomics analyses, where usually the aim is relative comparison of cases *vs* controls. This is due to high costs and also because not all internal standards are

352 commercially available. In our method, we selected 12 labeled internal standards (Table S1), which
353 represent chemically similar classes for optimal correction. This is because the matrix effect was
354 expected to be the same for an analyte and its labeled isotope analogue. The process efficiency
355 percentages were calculated for the metabolites without internal standards. The analyte
356 concentrations determined through the external calibration were divided with the total process
357 efficiency values to correct the concentration values of the analyte in the given biological sample.
358 Also, the repeatability of matrix effect in terms of CV was less than 25% for most of the compounds.
359 Reliable measurements are accordingly possible.

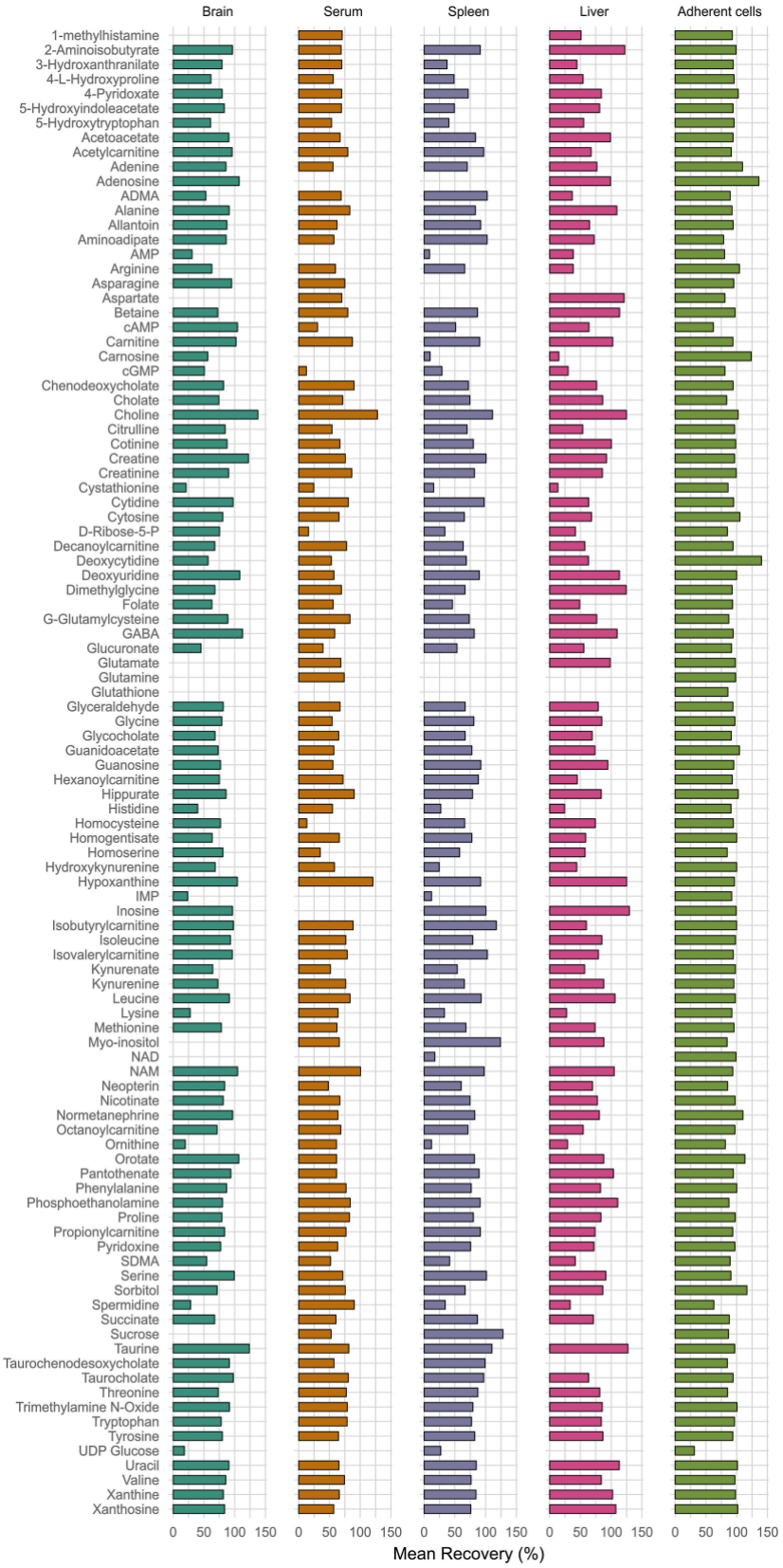

Figure 4. Percentage mean recoveries for all metabolites at low, medium, and high concentration levels of QCs (N=6 at each level) spiked in serum, brain, spleen, liver, and adherent cells. Data not shown for metabolites with poor chromatography and irreproducible results at different concentration levels.

365 3.2.4. Stability

366 Endogenous metabolites are not stable due to degradation or conversion reactions. Hence, the
367 stabilities of all metabolites were assessed under different conditions. For wet extract stability,
368 approximately 90% of the metabolites were stable (stabilities range between 85% and 115%) for 35
369 hours at 5°C in the auto-sampler (Figure 5A).
370 For freeze and thaw cycle stability, most of the metabolites were stable even after three freeze
371 and thaw cycles, with the exception of cGMP, succinate, glutathione, and homocysteine (stability
372 below 30%, Figure 5B). This information is particularly important for clinical studies, where samples
373 are often thawed once or twice.
374 To determine the stability of working solutions, we started evaluating the stability from
375 intermediate solutions for all the metabolites. Most of the metabolites were stable for 245 days at
376 intermediate concentration when stored at 4°C. However, 16 metabolites had low stability at
377 intermediate concentration; stabilities were thus determined at stock-level concentration for these
378 compounds. We observed that the stock solutions were stable for 56 days except taurocholic acid,
379 sucrose, UDP-glucose, and glutamine (Figure 6). Hence, these stock solutions were freshly prepared
380 during the analysis. The stability of internal standards solutions was also assessed and they were
381 stable for one year.

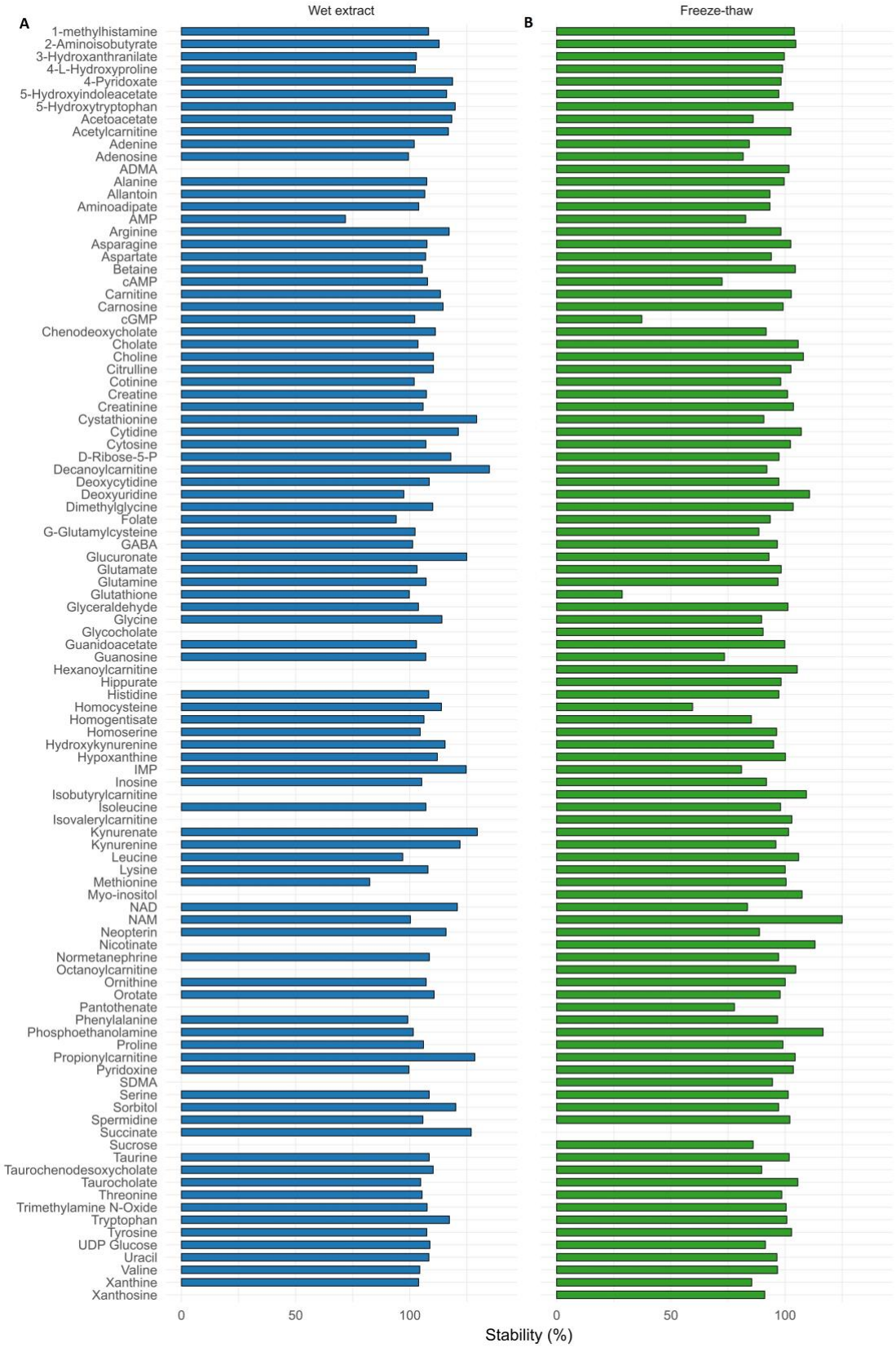

Figure 5. Percentage mean stability was calculated for three different concentrations (low, medium, and high) spiked in QC serum (N=6 at each level) during (A) wet extract and (B) freeze-thaw stability for all the metabolites.

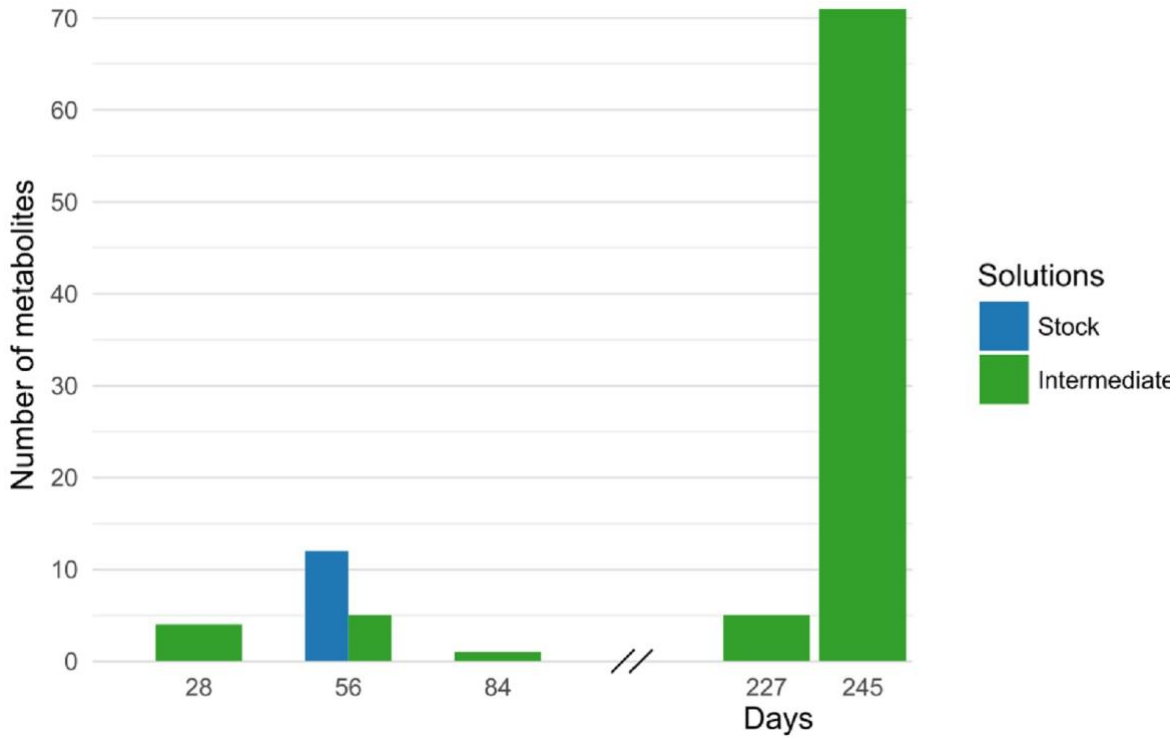

Figure 6. Bar graph representing stability of stock solutions and intermediate solutions of metabolites that were used to freshly prepare calibration curve standards during the analysis.

3.2.5. Sample carryover

In general, for majority of the analyte MRM channels neither peak nor any interference in the blank samples was detected after injection of metabolite standard with high concentration. For compounds such as spermidine, succinate, AMP, and IMP, carryover eluted constantly even after washing but was not significantly high. Other than these compounds, we can conclude that the column, needle, syringe, and seal washes were sufficient to avoid any inter-sample carryover.

3.2.6. Reproducibility

To ensure good quality of the data, internal QC samples were incorporated to a batch of samples and run after every tenth experimental sample. QC data were collected from 25 different batches that were performed during various metabolomics studies over a period of 1 year. Mean concentrations and %CV values of QC replicates within each batch were calculated for all the 25 batches. We observed that approximately 80% to 85% of the metabolites were always present within 25% of CV values (Figure 7). The higher %CV values for the remaining metabolites could be partially explained by low abundance in human serum, low recovery, or poor chromatography; these were consistently found to be below LLOQ within the 25 batches.

In addition, %CV values for retention times and R^2 values of calibration curves for each metabolite in all the 25 batches were calculated to verify the reproducibility. Based on these results, the repeatability was excellent except for a few compounds over a period of 1 year. No drifting effect for the retention times (CV<4%) was observed, and excellent reproducibility was observed for R^2 values of calibration curves (CV<3%) (Figure 8). On the basis of these results, our method can be considered accurate, reliable, and reproducible.

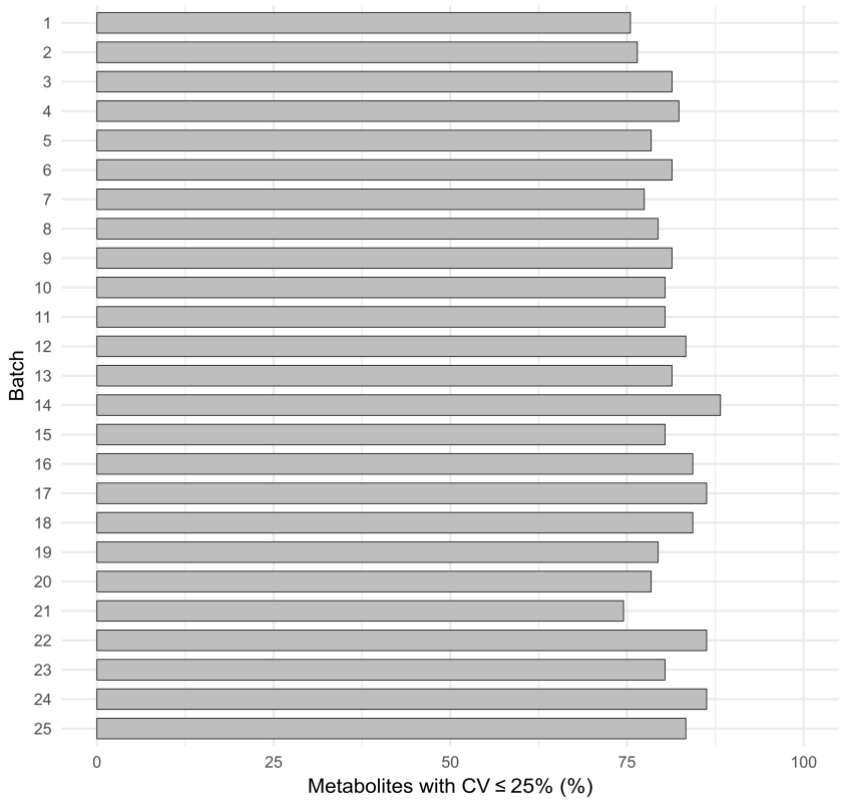

Figure 7. Percentage of metabolites with less than 25% CV values of QC concentrations in 25 different batches analyzed over a period of 1 year.

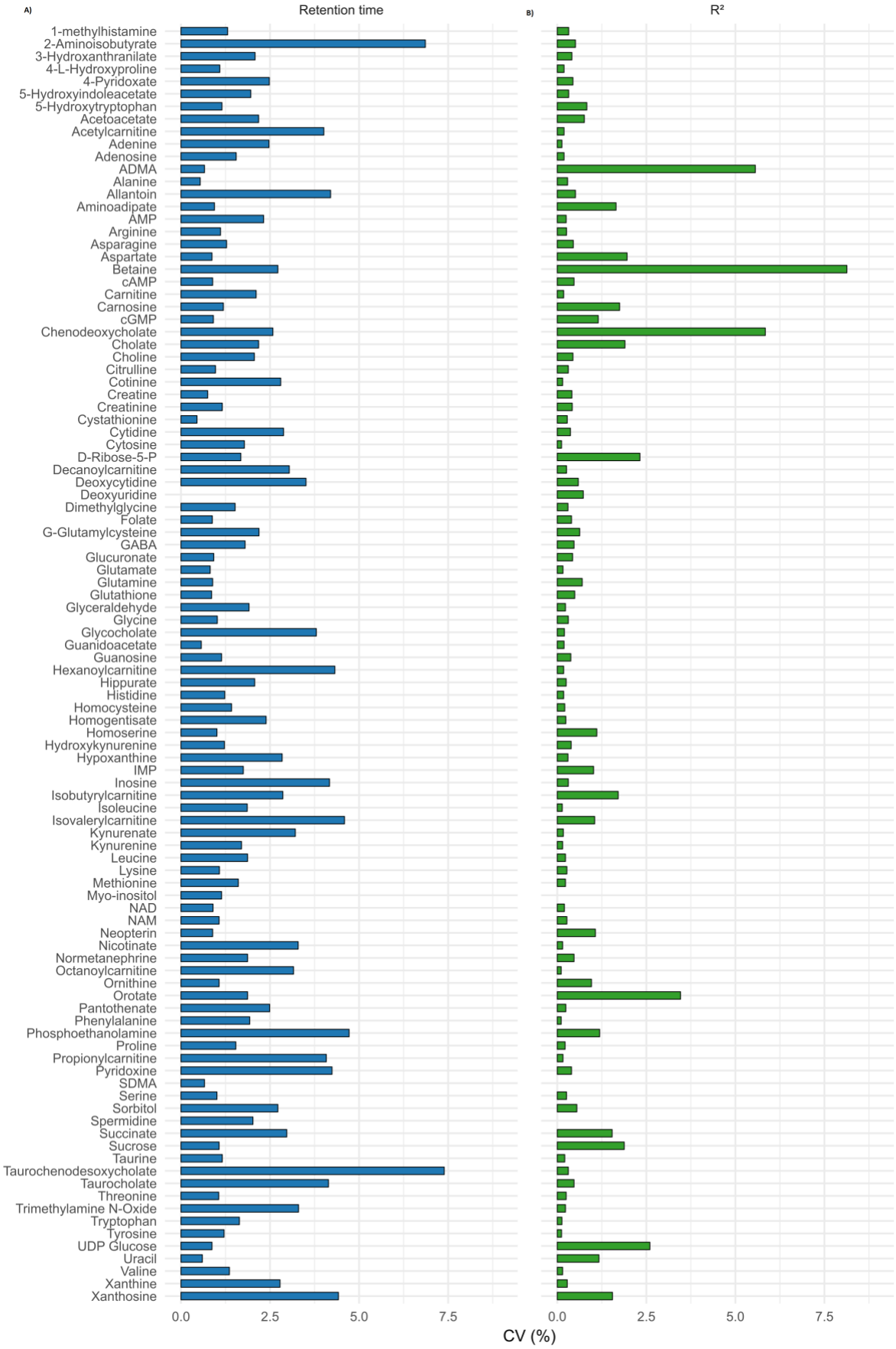

Figure 8. Reproducibility of (A) retention time of respective chromatographic peaks and (B) regression coefficient values (R^2) from external calibration curve standards for each metabolite analyzed in 25 different batches over a period of 1 year.

3.2.7. Quality management

To obtain reproducible and accurate data, we set up a strict quality management and electronic lab notebook system. To reduce the bias from sample analysis, we always double-randomized the samples (i.e., one before the sample extraction step and one before injecting into the LCMS system across different phenotypes of the samples). For stabilization of response and retention time, we always verified a few runs of highest calibration level 11 before injecting the experimental samples. During the stabilization process, we also verified the chromatography including peak shape, retention time, and response of all the metabolites. Any significant changes in the intensity, peak shape, retention time, and system pressure were thoroughly investigated and corrected by resolving the problems before injection of experimental samples.

To ensure the integrity of LCMS runs, QC samples were run at every tenth experimental sample and a blank sample at every fifth run during all the metabolomics studies within a batch. Furthermore, chromatography and response of QC samples (including chromatography of some metabolites and IS response variation) and blank runs were always verified after completion of the runs and before starting data processing. In case of any abnormality observed for particular samples, those samples were reinjected or reanalyzed. Only after passing these quality checks we proceed further to process the data. This included verifying the accuracy of calibration curve standards, chromatography peak integrations, IS response variation, and verifying LLOQ and ULOQ for each sample for all metabolites within a batch. The high-throughput targeted metabolomics workflow is shown in Figure 9.

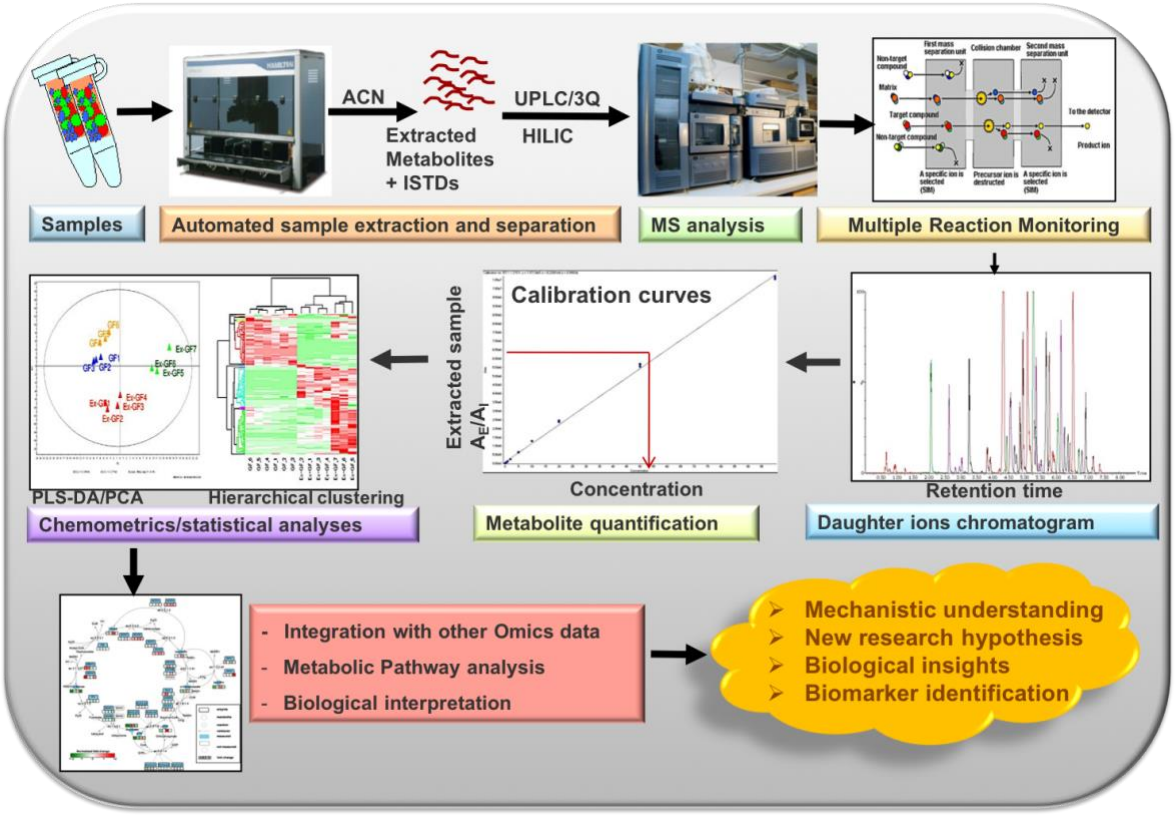

Figure 9. A typical workflow for high-throughput targeted metabolomics analysis.

We collected concentration values (μmol/L) for our QC samples within metabolomics studies conducted over a period of 5.5 years from six different lots (N=539 replicates). The median values of each metabolite together with a 95% credibility interval are presented in Table 1. These represent a reference level for a population of healthy adult individuals.

446 **Table 1.** Median concentration levels (μmol/L) measured in pooled healthy adult serum samples.

		Population median (μmol/L)		
			95% credibility interval	
Class and Metabolite Name	**HMDB Id**	**Estimate**	**Lower**	**Upper**
1. Alpha amino acids and derivatives				
2-Aminoisobutyrate	HMDB0001906	1.128	0.809	1.509
4-L-Hydroxyproline	HMDB0000725	15.895	11.300	21.184
5-Hydroxytryptophan	HMDB0000472	0.043	0.031	0.058
ADMA	HMDB0001539	0.963	0.248	2.230
Alanine	HMDB0000161	477.946	339.667	635.383
Aminoadipate	HMDB0000510	2.121	1.502	2.818
Arginine	HMDB0000517	84.902	60.795	113.172
Asparagine	HMDB0000168	47.204	33.635	62.727
Aspartate	HMDB0000191	26.932	18.786	35.424
Betaine	HMDB0000043	100.045	72.251	134.758
Citrulline	HMDB0000904	27.815	19.812	36.972
Creatine	HMDB0000064	61.180	44.481	82.161
Creatinine	HMDB0000562	66.085	47.095	88.066
Cystathionine	HMDB0000099	0.131	0.093	0.174
Dimethylglycine	HMDB0000092	3.541	2.521	4.733
GABA	HMDB0000112	0.195	0.138	0.259
G-Glutamylcysteine	HMDB0001049	2.966	2.116	3.946
Glutamate	HMDB0000148	53.446	37.691	70.613
Glutamine	HMDB0000641	791.142	560.999	1050.605
Glutathione	HMDB0000125	0.021	0.015	0.028
Glycine	HMDB0000123	243.717	172.929	325.417
Guanidoacetate	HMDB0000128	2.635	1.880	3.498
Histidine	HMDB0000177	88.616	62.718	117.848
Homocysteine	HMDB0000742	0.480	0.124	1.153
Homoserine	HMDB0000719	0.337	0.240	0.448
Hydroxykynurenine	HMDB0000732	0.096	0.068	0.128
Isoleucine	HMDB0000172	83.316	58.352	110.232
Kynurenine	HMDB0000684	1.146	0.813	1.520
Leucine	HMDB0000687	126.072	90.139	167.353
Lysine	HMDB0000182	176.519	127.474	236.393
Methionine	HMDB0000696	29.201	21.081	39.249
Ornithine	HMDB0000214	90.177	65.043	120.796
Phenylalanine	HMDB0000159	88.326	63.916	117.977
Proline	HMDB0000162	251.689	180.529	335.079
SDMA	HMDB0003334	2.862	2.060	3.834
Serine	HMDB0000187	148.482	107.012	198.507
Threonine	HMDB0000167	152.013	108.810	203.651
Tryptophan	HMDB0000929	33.961	24.104	45.261

Tyrosine	HMDB0000158	65.558	45.778	86.262
Valine	HMDB0000883	394.932	282.460	526.776
2. Benzoic acids and derivatives				
3-Hydroxanthranilate	HMDB0001476	0.188	0.134	0.251
Hippurate	HMDB0000714	6.101	4.362	8.124
3. Beta amino acids and derivatives				
Carnosine	HMDB0000033	0.014	0.010	0.019
Pantothenate	HMDB0000210	0.310	0.220	0.411
4. Bile acids, alcohols and derivatives				
Chenodeoxycholate	HMDB0000518	53.661	38.588	71.987
Cholate	HMDB0000619	0.676	0.483	0.900
Glycocholate	HMDB0000138	0.373	0.267	0.498
Taurochenodesoxycholate	HMDB0000951	0.507	0.359	0.673
5. Carbohydrates and carbohydrate conjugates				
D-Ribose-5-P	HMDB0001548	1.273	0.919	1.710
Glyceraldehyde	HMDB0001051	239.946	172.630	322.782
Sucrose	HMDB0000258	1.417	1.015	1.886
6. Dialkylamines				
Spermidine	HMDB0001257	33.601	23.958	44.759
7. Dicarboxylic acids and derivatives				
Succinate	HMDB0000254	7.912	5.597	10.491
8. Fatty Acyls				
Acetylcarnitine	HMDB0000201	9.709	2.323	22.264
Decanoylcarnitine	HMDB0000651	0.305	0.087	0.720
Hexanoylcarnitine	HMDB0000705	0.055	0.015	0.129
Isobutyrylcarnitine	HMDB0000736	0.248	0.061	0.578
Isovalerylcarnitine	HMDB0000688	0.102	0.027	0.240
Octanoylcarnitine	HMDB0000791	0.297	0.076	0.691
Propionylcarnitine	HMDB0000824	0.423	0.119	0.998
9. Folates				
Folate	HMDB0000121	0.011	0.003	0.027
10. Glucuronic acid and derivatives				
Glucuronate	HMDB0000127	1.960	1.404	2.627
11. Imidazoles				
1-Methylhistamine	HMDB0000898	0.006	0.004	0.008
Allantoin	HMDB0000462	2.447	1.741	3.254
12. Indoles and derivatives				
5-Hydroxyindoleacetate	HMDB0000763	0.074	0.053	0.099
13. Keto acids and derivatives				
Acetoacetate	HMDB0000060	6.713	4.798	8.988
14. Organic phosphoric acids and derivatives				

Phosphoethanolamine	HMDB0000224	3.316	2.356	4.420
15. Organosulfonic acids				
Taurine	HMDB0000251	221.359	156.971	294.936
Taurocholate	HMDB0000036	0.083	0.060	0.112
16. Oxides				
Trimethylamine N-oxide	HMDB0000925	1.477	1.055	1.973
17. Phenols				
Homogentisate	HMDB0000130	0.115	0.082	0.153
Normetanephrine	HMDB0000819	0.0010	0.0007	0.0014
18. Pteridines and derivatives				
Neopterin	HMDB0000845	0.005	0.004	0.007
19. Purines and derivatives				
Adenine	HMDB0000034	0.007	0.005	0.009
Adenosine	HMDB0000050	0.008	0.006	0.011
AMP	HMDB0000045	0.106	0.075	0.140
cAMP	HMDB0000058	0.005	0.003	0.006
cGMP	HMDB0001314	0.009	0.002	0.026
Guanosine	HMDB0000133	0.445	0.315	0.593
Hypoxanthine	HMDB0000157	58.838	41.908	78.087
IMP	HMDB0000175	0.212	0.151	0.282
Inosine	HMDB0000195	36.061	25.798	48.008
Xanthine	HMDB0000292	3.926	2.790	5.211
Xanthosine	HMDB0000299	0.352	0.249	0.466
20. Pyridines and derivatives				
4-Pyridoxate	HMDB0000017	0.057	0.042	0.077
Cotinine	HMDB0001046	0.531	0.382	0.710
NAD	HMDB0000902	0.015	0.011	0.020
Niacinamide	HMDB0001406	0.395	0.279	0.524
Nicotinate	HMDB0001488	0.012	0.009	0.017
Pyridoxine	HMDB0000239	0.0007	0.0005	0.0009
21. Pyrimidines and derivatives				
Cytidine	HMDB0000089	0.003	0.002	0.004
Cytosine	HMDB0000630	0.080	0.056	0.106
Deoxycytidine	HMDB0000014	0.871	0.440	1.461
Deoxyuridine	HMDB0000012	0.543	0.391	0.726
Orotate	HMDB0000226	0.036	0.006	0.096
UDP Glucose	HMDB0000286	0.232	0.165	0.308
Uracil	HMDB0000300	0.058	0.042	0.078
22. Quaternary ammonium salts				
Carnitine	HMDB0000062	85.669	61.210	114.060
Choline	HMDB0000097	95.170	67.565	126.279
23. Quinolines and derivatives				
Kynurenate	HMDB0000715	0.044	0.031	0.058

24. Sugar alcohols				
Myo-inositol	HMDB0000211	16.989	12.075	22.606
Sorbitol	HMDB0000247	3.692	2.617	4.904

3.2.8. Robustness and cross-platform comparison

To verify the performance of our method, we analyzed the NIST standard reference material SRM 1950 plasma. The correlation coefficient for 17 matched metabolites between the given reference values and from our semi-quantitative method was 0.967, indicating the high performance of our method (Figure 10A).

Furthermore, we verified cross-platform comparability. This was achieved by comparing metabolite concentrations analyzed using our method against two completely different analytical platforms (BIOCRATES AbsoluteIDQ p180 kit and NMR) in our QC samples. We obtained a high correlation coefficient for matched 38 metabolites measured using BIOCRATES AbsoluteIDQ p180 kit and our method (R^2=0.975) (Figure 10B) and for matched 22 metabolites measured using NMR and our method (R^2=0.884) (Figure 10C). These results demonstrate the robustness of our method.

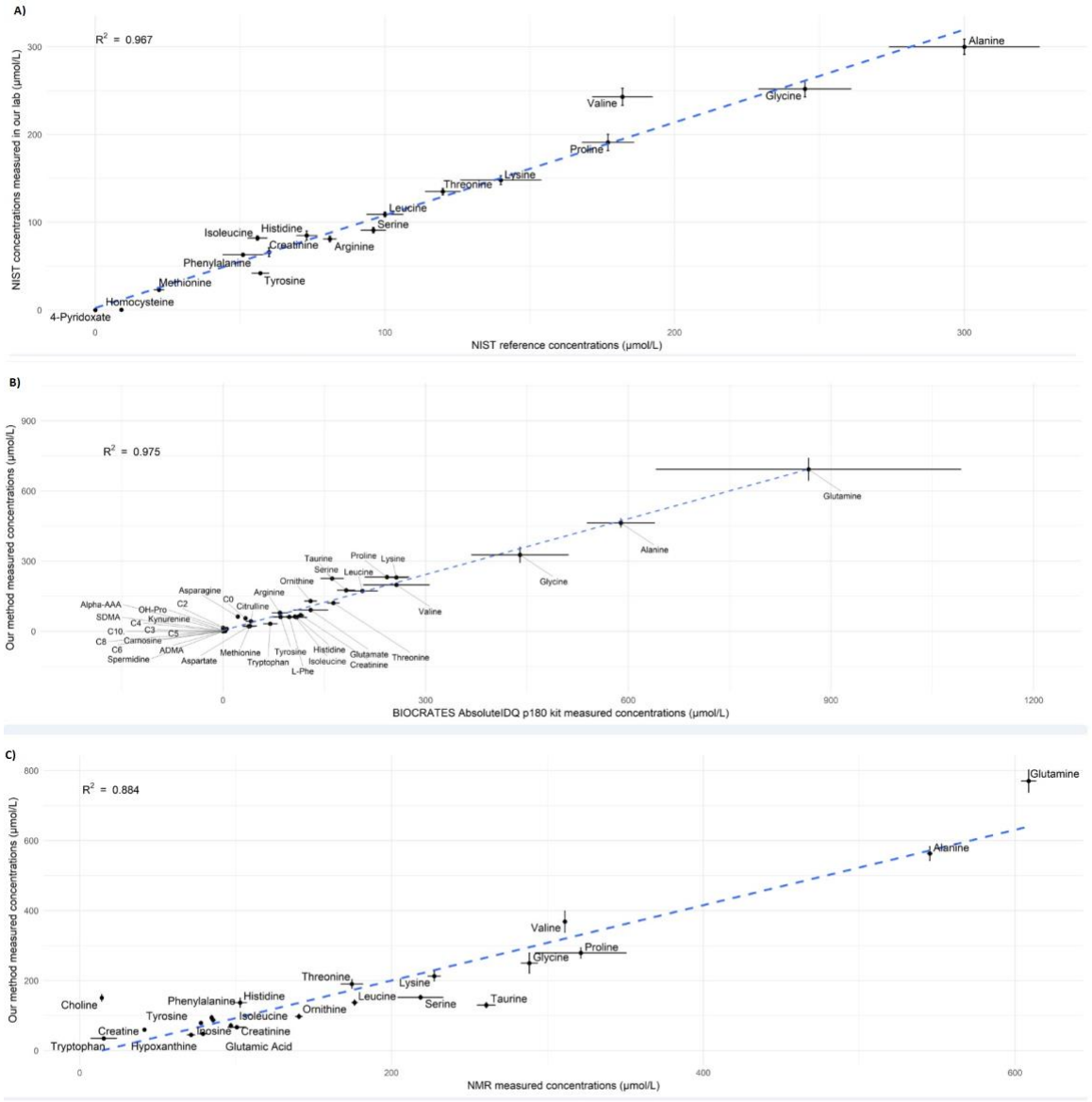

Figure 10. (A) Comparison of metabolite concentration reference values given in the NIST SRM 1950 plasma against our method. Comparison of metabolite concentrations in our QC samples measured between our method and (B) BIOCRATES AbsoluteIDQ p180 kit, and (C) NMR analyses.
Error bars represent 95% confidence intervals.

3.3 Automated data processing

After the raw data processing using the instrument-coupled software (TargetLynx), a single-sheet excel file containing information such as sample ids, file name, and concentrations (PPB) is generated in a complex format for all metabolites. For example, in a single-batch run of 85 samples, after data processing the concentration values for all samples are obtained under each metabolite separately. This means that if a data matrix of 85 samples x 100 metabolites is desired, each individual concentration value must be copied and pasted in another sheet. PPB values are then converted into µM and corrected for process efficiencies and dilution factors (if any) and normalized with tissue weight or cell number depending on the sample type. Apart from these steps, as the samples are analyzed in a randomized manner, the next step is to rearrange the experimental samples according to phenotypic group and also to separate the QC samples. Manual data formatting to create a ready-

to-use data matrix for visualization and for downstream statistical analyses is a tedious and time-consuming task and (most importantly) is prone to errors.

We have implemented a software package "Unlynx" in R statistical language. This package takes the raw data produced by TargetLynx software as input and produces processed data into ready-to-use spreadsheets. Using the R package, all the PPB values are converted into µmol/L, µmol/g, or µmol/million cells by diving with molecular weight of the respective metabolites and correction factor (weight of tissues, number of cells, and dilution factor) with respect to sample types. In addition, mean concentrations and %CV of QC samples for all the metabolites are calculated and used to evaluate quality by comparing with the *in-house* QC database constructed based on inter-day %CV. Furthermore, retrieving LLOQ, ULOQ, and outlier values in each phenotypic group according to one and two standard deviations, retention time values, and R^2 values for each metabolite is implemented in the R software. After implementing the automated data processing, we had reduced the 2-day manual workload to a few minutes.

Thus, after processing the data using TargetLynx, we routinely export all results to an excel file for automated data processing using "Unlynx" software. Typically, the data resulting from such automated data processing is suitable for more specialized data analyses (such as statistical hypotheses testing, classification, regression, and clustering) aimed at answering specific scientific questions related to the study design.

3.4 Applicability of the method

We have applied our fully validated analytical methodology in various international and national biomedical research projects, epidemiological studies, clinical studies including dietary interventions, and clinical trials. We have successfully implemented our technology in the following research fields, including but not limited to: mitochondrial metabolism/disorders [11,17-23], cancer [24,25], bone metabolism [12], endocrinology [26,27], psychiatric disorders [28], inflammatory bowel disease [29], viral infections [30-33], allergies [34,35], circadian rhythms [36], and pain research [37].

4. Conclusions

The developed high-throughput targeted and semi-quantitative method was optimized for various biological matrices (biofluids, tissues, cells) from different organisms. We validated the analytical method according to EMA guidelines for bioanalytical methods and showed good accuracy, reproducibility, selectivity, specificity, recoveries, and stability. We have also implemented a strict quality management and electronic notebook system. Reproducibility was demonstrated by consistent results for retention time and correlation coefficient of calibration curves, and concentrations of QC samples over a period of 1 year. Reliability was shown by the excellent correlation between metabolite concentrations measured using our method and the NIST reference values. Moreover, robustness was shown via good cross-platform comparability between two completely different analytical platforms. Furthermore, we have automated the downstream data processing steps to handle sample analyses in a high-throughput manner, which is particularly valuable for analyzing population cohorts and large clinical samples for metabolomics studies. We have successfully applied this method in many biomedical research projects and clinical trials, including epidemiological studies for biomarker discovery.

Supplementary Materials: The following are available online at www.mdpi.com/link, Table S1: List of metabolites showing molecular weight, retention time, linearity of calibration, and compound dependent MS parameters.

Acknowledgments: The authors would like to thank all the collaborators, especially Prof. Anu Wartiovaara for providing the biological samples; Jean-Christophe Yorke, Vasudev Kantae, Kenneth Nazir, and Bharat Gajera for technical assistance, and Dr. Janne Wallenius for reviewing the manuscript. The Biocenter Finland and HiLIFE supported this work.

Author Contributions: J.N. performed all the validation experiments, processed the data, and wrote the manuscript. M.K. and J.N. were involved in extraction protocol optimization. G.P. developed the R package (Unlynx) for automated data processing. A.P. performed the QC data analysis, prepared the related table, and wrote the corresponding section. V.V. conceived, designed, and supervised the study, and wrote and edited the manuscript. All authors edited the manuscript.

Conflicts of Interest: The authors declare no conflict of interest. The sponsors had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; and in the decision to publish the results.

5. References

1. Gowda, G.A.N.; Djukovic, D. Overview of mass spectrometry-based metabolomics: Opportunities and challenges. *Methods Mol Biol.* **2014**, *1198*, 3-12, DOI:10.1007/978-1-4939-1258-2_1.
2. Dudley, E.; Yousef, M.; Wang, Y.; Griffiths, W.J. Targeted metabolomics and mass spectrometry. *Adv Protein Chem Struct Biol.* **2010**, *80*, 45-83, DOI:10.1016/B978-0-12-381264-3.00002-3.
3. Raterink, R.-J.; Lindenburg, P.W.; Vreeken, R.J.; Ramautar, R.; Hankemeier, T. Recent developments in sample-pretreatment techniques for mass spectrometry-based metabolomics. *TrAC Trends in Analytical Chemistry* **2014**, *61*, 157-167, DOI:10.1016/j.trac.2014.06.003.
4. Gong, Z.-G.; Hu, J.; Wu, X.; Xu, Y.-J. The recent developments in sample preparation for mass spectrometry-based metabolomics. *Crit Rev Anal Chem.* **2017**, *47*, 325-331, DOI:10.1080/10408347.2017.1289836.
5. Dettmer, K.; Aronov, P.A.; Hammock, B.D. Mass spectrometry-based metabolomics. *Mass Spectrom Rev.* **2007**, *26*, 51-78, DOI:10.1002/mas.20108.
6. Abdel Rahman, A.M.; Pawling, J.; Ryczko, M.; Caudy, A.A.; Dennis, J.W. Targeted metabolomics in cultured cells and tissues by mass spectrometry: Method development and validation. *Anal Chim Acta.* **2014**, *845*, 53-61, DOI:10.1016/j.aca.2014.06.012.
7. Villas-Bôas, S.G.; Mas, S.; Åkesson, M.; Smedsgaard, J.; Nielsen, J. Mass spectrometry in metabolome analysis. *Mass Spectrom Rev.* **2005**, *24*, 613-646, DOI:10.1002/mas.20032.
8. Virgiliou, C.; Sampsonidis, I.; Gika, H.G.; Raikos, N.; Theodoridis, G.A. Development and validation of a hilic-ms/ms multitargeted method for metabolomics applications. *Electrophoresis* **2015**, *36*, 2215-2225, DOI:10.1002/elps.201500208.
9. Peng, J.; Chen, Y.-T.; Chen, C.-L.; Li, L. Development of a universal metabolome-standard method for long-term lc–ms metabolome profiling and its application for bladder cancer urine-metabolite-biomarker discovery. *Anal Chem.* **2014**, *86*, 6540-6547, DOI:10.1021/ac5011684.
10. Khan, N.A.; Auranen, M.; Paetau, I.; Pirinen, E.; Euro, L.; Forsström, S.; Pasila, L.; Velagapudi, V.; Carroll, C.J.; Auwerx, J.; Suomalainen, A. Effective treatment of mitochondrial myopathy by nicotinamide riboside, a vitamin B3. *EMBO Mol Med.* **2014,** *6*, 721, DOI:10.1002/emmm.201403943.
11. Roman-Garcia, P.; Quiros-Gonzalez, I.; Mottram, L.; Lieben, L.; Sharan, K.; Wangwiwatsin, A.; Tubio, J.; Lewis, K.; Wilkinson, D.; Santhanam, B.; Sarper, N.; Clare, S.; Vassiliou, G.S.; Velagapudi, V.R.; Dougan, G.; Yadav, V.K. Vitamin b12–dependent taurine synthesis regulates growth and bone mass. *J Clin Invest.* **2014**, *124*, 2988-3002, DOI:10.1172/JCI72606.

12. Li, K.; Naviaux, J.C.; Bright, A.T.; Wang, L.; Naviaux, R.K. A robust, single-injection method for targeted, broad-spectrum plasma metabolomics. *Metabolomics* **2017**, *13*, 122, DOI:10.1007/s11306-017-1264-1.
13. Liu, P.; Qi, C.-B.; Zhu, Q.-F.; Yuan, B.-F.; Feng, Y.-Q. Determination of thiol metabolites in human urine by stable isotope labeling in combination with pseudo-targeted mass spectrometry analysis. *Scientific Reports* **2016,** *6*, 21433, DOI: 10.1038/srep21433.
14. Rosebrock, A.P. Targeted full-scan lc-ms metabolomics: Simultaneous quantitation of knowns and feature discovery provide the best of both worlds. *Bioanalysis* **2017**, *9*, 5-8, DOI:10.4155/bio-2016-0256.
15. Wei, R.; Li, G.; Seymour, A.B. High-throughput and multiplexed lc/ms/mrm method for targeted metabolomics. *Anal Chem.* **2010**, *82*, 5527-5533, DOI:10.1021/ac100331b.
16. Yuan, M.; Breitkopf, S.B.; Yang, X.; Asara, J.M. A positive/negative ion–switching, targeted mass spectrometry–based metabolomics platform for bodily fluids, cells, and fresh and fixed tissue. *Nat Protoc.* **2012**, *7*, 872, DOI:10.1038/nprot.2012.024.
17. Nikkanen, J.; Forsström, S.; Euro, L.; Paetau, I.; Kohnz, R.A.; Wang, L.; Chilov, D.; Viinamäki, J.; Roivainen, A.; Marjamäki, P.; Liljenbäck, H.; Ahola, S.; Buzkova, J.; Terzioglu, M.; Khan, N.A.; Pirnes-Karhu, S.; Paetau, A.; Lönnqvist, T.; Sajantila, A.; Isohanni P.; Tyynismaa, H.; Nomura, D.K.; Battersby, B.; Velagapudi, V.; Carroll, C.J.; Suomalainen, A. Mitochondrial DNA Replication Defects Disturb Cellular dNTP Pools And Remodel One-Carbon Metabolism. *Cell Metab.* **2016**, *23*, 635-648, DOI:10.1016/j.cmet.2016.01.019.
18. Ahola, S.; Auranen, M.; Isohanni, P.; Niemisalo, S.; Urho, N.; Buzkova, J.; Velagapudi, V.; Lundbom, N.; Hakkarainen, A.; Muurinen, T.; Piirilä, P.; Pietiläinen, K.; Suomalainen, A. Modified Atkins diet induces subacute selective ragged red fiber lysis in mitochondrial myopathy patients. *EMBO Mol Med.* **2016**, *11*, 1234-1247, DOI:10.15252/emmm.201606592.
19. Rajendran, J.; Tomašić, N.; Kotarsky, H.; Hansson, E.; Velagapudi, V.; Kallijärvi, J.; Fellman, V. Effect of high-carbohydrate diet on plasma metabolomics in mice with mitochondrial respiratory chain complex III deficiency due to a Bcs1l mutation. *Int J Mol Sci.* **2016**, *17*, 1824, DOI:10.3390/ijms17111824.
20. Szibor, M.; Dhandapani, P.K.; Dufour, E.; Holmström, K.M.; Zhuang, Y.; Salwig, I.; Wittig, I.; Heidler, J.; Gizatullina, Z.; Gainutdinov, T.; Consortium, G.M.; Fuchs, H.; Gailus-Durner, V.; de Angelis, M.H.; Nandania, J.; Velagapudi, V.; Wietelmann, A.; Rustin, P.; Gellerich, F.N.; Jacobs, H.T.; Braun, T. Broad AOX expression in a genetically tractable mouse model does not disturb normal physiology. *Dis Model Mech*. **2017**, *2*, 163-171, DOI:10.1242/dmm.027839.
21. Schatton, D.; Pla-Martin, D.; Marx, M.C.; Hansen, H.; Mourier, A.; Nemazanyy, I.; Pessia, A.; Zentis, P.; Corona, T.; Kondylis, V.; Barth, E.; Schauss, A.C.; Velagapudi, V.; Rugarli, E. CLUH regulates mitochondrial metabolism by controlling translation and decay of target mRNAs. *J Cell Biol.* **2017**, *216*, 675-693, DOI:10.1083/jcb.201607019.
22. Purhonen, J.; Rajendran, J.; Mörgelin, M.; Uusi-Rauva, K.; Katayama, S.; Krjutskov, K.; Einarsdottir, E.; Velagapudi, V.; Kere, J.; Jauhiainen, M.; Fellman, V.; Kallijärvi, J. Ketogenic diet attenuates hepatopathy in mouse model of respiratory chain complex III deficiency caused by a Bcs1l mutation. *Sci Rep.* **2017**, *7*, 957, DOI:10.1038/s41598-017-01109-4.
23. Khan, N.A.; Nikkanen, J.; Yatsuga, S.; Wang, L.; Jackson, C.; Pessia, A.; Riikka K.; Velagapudi, V.; Suomalainen, A. mTORC1 Regulates Mitochondrial Folate Cycle and the Integrated Stress Response in Mitochondrial Disease. *Cell Metab.* **2017**, *26*, 419-428, DOI:10.1016/j.cmet.2017.07.007.

24. Scott, T.A.; Quintaneiro, L.M.; Norvaisas, P.; Lui, P.P.; Wilson, M.P.; Leung, K.Y.; Herrera-Dominguez, L.; Sudiwala, S.; Pessia, A.; Clayton, P.T.; Bryson, K.; Velagapudi, V.; Mills, P.B.; Typas, A.; Greene, N.D.E.; Cabreiro, F. Host-Microbe Co-metabolism Dictates Cancer Drug Efficacy in C.elegans. *Cell* **2017**, *169*, 442-456, DOI:10.1016/j.cell.2017.03.040.

25. Puhka, M.; Takatalo, M.; Nordberg, M.E.; Valkonen, S.; Nandania, J.; Aatonen, M.; Yliperttula, M.; Laitinen, S.; Velgapudi, V.; Mirtti, T.; Kallioniemi, O.; Rannikko, A.; Siljander, P.R.-M.; Hallstrom, T.M. Metabolomic profiling of extracellular vesicles and alternative normalisation methodsreveal enriched metabolites and strategies to study prostate cancer related changes. *Theranostics* **2017**, *7*, 3824-3841, DOI:10.7150/thno.19890.

26. Schrade, A.; Kyrönlahti, A.; Akinrinade, O.; Pihlajoki, M.; Häkkinen, M.; Fischer, S.; Alastalo, T.-P.; Velagapudi, V.; Toppari, J.; Wilson, D.B.; Heikinheimo, M. GATA4 is a key regulator of steroidogenesis and glycolysis in mouse Leydig cells. *Endocrinology* **2015**, *156*, 1860-1872, DOI:10.1210/en.2014-1931.

27. Schrade, A.; Kyrönlahti, A.; Akinrinade, O.; Pihlajoki, M.; Fischer, S.; Rodriguez, V.M.; Velagapudi, V.; Toppari, J.; Wilson, D.B.; Heikinheimo, M. GATA4 regulates blood-testis barrier function and lactate metabolism in mouse Sertoli cells. *Endocrinology* **2016**, *157*, 2416-2431, DOI:10.1210/en.2015-1927.

28. Ali-Sisto, T.; Tolmunen, T.; Toffol, E.; Viinamäki, H.; Mäntyselkä, P.; Valkonen-Korhonen, M.; Honkalampi, K.; Ruusunen, A.; Velagapudi, V.; Lehto, S.M. Purine metabolism is dysregulated in patients with major depressive disorder. *Psychoneuroendocrinology* **2016**, *70*, 25-32, DOI:10.1016/j.psyneuen.2016.04.017.

29. Kolho, K.-L.; Pessia, A.; Jaakkola, T.; de Vos, W.M.; Velagapudi, V. Fecal and serum metabolomics in pediatric inflammatory bowel disease. *J Crohns Colitis.* **2017**, *11*, 321-334, DOI:10.1093/ecco-jcc/jjw158.

30. Fu, Y.; Gaelings, L.; Söderholm, S.; Belanov, S.; Nandania, J.; Nyman, T.A.; Matikainen, S.; Anders, S.; Velagapudi, V.; Kainov, D.E. NJ872 inhibits influenza A virus replication without altering cellular antiviral responses. *Antiviral Res.* **2016**, *133*, 23-31, DOI:10.1016/j.antiviral.2016.07.008.

31. Gaelings, L.; Söderholm, S.; Bugai, A.; Fu, Y.; Nandania, J.; Schepens, B.; Lorey, M.B.; Tynell, J.; Ginste, L.V.; Goffic, R.L.; Miller, M.S.; Kuisma, M.; Marjomäki, V.; Brabander, J.D.; Matikainen, S.; Nyman, T.A.; Bamford, D.; Saelens, X.; Julkunen, I.; Paavilainen, H.; Hukkanen, V.; Velagapudi, V.; Kainov, D.E. Regulation of Kynurenine Biosynthesis during Influenza Virus Infection. *FEBS J.* **2017**, *284*, 222-236, DOI:10.1111/febs.13966.

32. Kuivanen, S.; Bespalov, M.M.; Nandania, J.; Ianevski, A.; Velagapudi, V.; De Brabander, J.K.; Kainov, D.E.; Vapalahti, O. Obatoclax, saliphenylhalamide and gemcitabine inhibit Zika virus infection in vitro and differentially affect cellular signaling, transcription and metabolism. *Antiviral Res.* **2017**, *139*, 117-128, DOI:10.1016/j.antiviral.2016.12.022.

33. Bulanova, D.; Ianevski, A.; Bugai, A.; Akimov, Y.; Kuivanen, S.; Paavilainen, H.; Kakkola, L.; Nandania, J.; Turunen, L.; Ohman, T.; Ala-Hongisto, H.; Pesonen, H.M.; Kuisma, M.S.; Honkimaa, A.; Walton, E.L.; Oksenych, V.; Lorey, M.B.; Guschin, D.; Shim, J.; Kim, J.; Than, T.T.; Chang, S.Y.; Hukkanen, V.; Kulesskiy, E.; Marjomaki, V.S.; Julkunen, I.; Nyman, T.A.; Matikainen, S.; Saarela, J.S.; Sane, F.; Hober, D.; Gabriel, G.; De Brabander, J.K.; Martikainen, M.; Windisch, M.P.; Min, J.Y.; Bruzzone, R.; Aittokallio, T.; Vähä-Koskela, M.; Vapalahti, O.; Pulk, A.; Velagapudi, V.; Kainov, D.E. Antiviral Properties of Chemical Inhibitors of Cellular Anti-Apoptotic Bcl-2 Proteins. *Viruses* **2017**, *9*, 271, DOI:10.3390/v9100271.

655 34. Zinkevičienė, A.; Kainov, D.; Girkontaitė, I.; Lastauskienė, E.; Kvedarienė, V.; Fu, Y.; Anders,
656 S.; Velagapudi, V. Activation of Tryptophan and Phenylalanine Catabolism in the Remission Phase of
657 Allergic Contact Dermatitis: A Pilot Study. *Int Arch Allergy Immunol.* **2016**, *170*, 262-268,
658 DOI:10.1159/000450789.
659 35. Haapaniemi, E.M.; Fogarty, C.L.; Katayama, S.; Vihinen, H.; Keskitalo, S.; Ilander, M.; Krjutškov, K.;
660 Mustjoki, S.; Lehto, M.; Hautala, T.; Eriksson, O.; Jokitalo, E.; Velagapudi, V.; Varjosalo, M.; Seppänen,
661 M.; Kere, J. Combined immunodeficiency with hypoglycemia caused by mutations in hypoxia up-
662 regulated 1. *J Allergy Clin Immunol.* **2017**, *139*, 1391-1393, DOI:10.1016/j.jaci.2016.09.050.
663 36. Rey, G.; Valekunja, U.K.; Feeney, K.A.; Wulund, L.; Milev, N.B.; Stangherlin, A.; Bollepalli,
664 L.; Velagapudi, V.; O'Neill, J.S.; Reddy, A.B. The Pentose Phosphate Pathway Regulates the Circadian
665 Clock. *Cell Metab.* **2016**, *24*, 462-473, DOI:10.1016/j.cmet.2016.07.024.
666 37. Mäntyselkä, P.: Ali-Sisto, T.; Kautiainen, H.; Savolainen, J.; Niskanen, L.; Viinamäki, H.; Velagapudi,
667 V.; Lehto, S.M. The association between musculoskeletal pain and circulating ornithine: A population-
668 based study. *Pain Medicine* **2017**, *18*, 1145-1151, DOI:10.1093/pm/pnw285.
669 38. Guideline on bioanalytical method validation. European Medicines Agency: **2011**.
670 39. Yamil, S.-M.; Mark, S.L.; Lisa, E.K.; Maureen, L.S.; Kelly, H.T.; Paul, A.R.; William, G.M.; Daniel W.B.;
671 Tracey, B.S.; Dmitrii, V.T.; Niksa, B.; Xinjian, Y.; Yuxue, L.; Yufang, Z.; William, E.W.; Pedatsur, N.;
672 Karen, W.P.; Alan, T.R.; Stephen, E.S. Metabolite Profiling of a NIST Standard Reference Material for
673 Human Plasma (SRM 1950): GC/MS, LC/MS, NMR and Clinical Laboratory Analyses, Libraries and
674 Web-based resources. *Anal Chem.* **2013**, *24*. 11725-11731, DOI:10.1021/ac402503m.
675 40. Siskos, A.P.; Jain, P.; Römisch-Margl W.; Bennett, M.; Achaintre, D.; Asad, Y.; Marney, L.;, Richardson,
676 L.; Koulman, A.; Griffin, J.L.; Raynaud, F.; Scalbert, A.; Adamski, J.; Prehn, C.; Keun, H.C.
677 Interlaboratory Reproducibility of a Targeted Metabolomics Platform for Analysis of Human Serum
678 and Plasma. *Anal Chem.* **2017**, *89*, 656-665, DOI:10.1021/acs.analchem.6b02930.
679 41. Mihaleva, V.V.; Korhonen, S.P.; van Duynhoven, J.; Niemitz, M.; Vervoort, J.; Jacobs, D.M. Automated
680 quantum mechanical total line shape fitting model for quantitative NMR-based profiling of human
681 serum metabolites. *Anal Bioanal Chem.* **2014**, *406*, 3091-3102, DOI:10.1007/s00216-014-7752-5.
682 42. Hadfield, J.D. Mcmc methods for multi-response generalized linear mixed models: The mcmcglmm r
683 package. *Jouranl of statistical software* **2010**, *33*, 22, DOI:10.18637/jss.v033.i02.
684 43. Kimanani, E.K. Bioanalytical calibration curves: proposal for statistical criteria. *J. Pharm. Biomed. Anal.*
685 **1998**, *16*, 1117-1124, DOI:10.1016/S0731-7085(97)00064-2.

686