UNIVERSITÀ DEGLI STUDI DI SALERNO

UNIVERSITÀ DEGLI STUDI DI SALERNO

Dipartimento di Farmacia

PhD Program

in **Drug Discovery and Development**

XXXVIII Cycle

PhD Thesis in

Mass spectrometry-based multi-omics strategies to support the development of new therapeutic approaches

Candidate

Dott.ssa *Danila La Gioia*

Tutor

Chiar.mo Prof. *Pietro Campiglia*

Co-tutor

Chiar.mo Prof. *Eduardo Maria Sommella*

Coordinator: Chiar.ma Prof.ssa *Alessandra Tosco*

Index

Figure Captions

Chapter II

Figure 1 *Effect of mobile phase modifiers on metabolite coverage across different reversed-phase columns. Ammonium fluoride (1 mM) combined with the HSS T3 column resulted in the highest ionization efficiency and metabolite annotation counts. For Luna Omega C18 and Luna Omega Polar columns, acetic acid (0.1%) provided optimal performance, whereas formic acid (0.1%) consistently showed the poorest performance across all columns, yielding the lowest number of annotated metabolites.*

Figure 2 A-D *Stacked bars show the proportion of features classified as Insufficient (1.00–3.99), Acceptable (4.00–5.99), Good (6.00–7.99), and Excellent (8.00–10.00) according to the Peak Rating Factor. Panels **A–C** compare peak quality obtained with formic acid, acetic acid, and ammonium fluoride, respectively, across the three RP stationary phases. Ammonium fluoride (1 mM) yields the highest fraction of Good and Excellent peaks, particularly on HSS T3. Panel **D** reports peak quality distributions under the optimal additive condition for each column, confirming the superior chromatographic performance of HSS T3.*

Figure 3 *Heatmap illustrating the cumulative number of annotated features obtained by integrating all possible dataset combinations across chromatographic modes (RP and HILIC) and ionization polarities (ESI+ and ESI–). Color intensity reflects the total feature count for each combination, enabling visual comparison of metabolome coverage and identification of the most effective RP–HILIC pairing for comprehensive analysis. A: BEH Amide (+), B: ZIC HILIC (+), C: HSS T3 (+), D: Luna Omega C18 (+), E: Luna Omega Polar (+); 1: BEH Amide (-), 2: ZIC HILIC (-), 3: HSS T3 (-), 4: Luna Omega C18 (-), 5: Luna Omega Polar (-).*

Figure 4 *Pie charts illustrate the distribution of chromatographic peak quality categories (Poor, Acceptable, Good, Excellent) for HILIC columns under ESI+ conditions. The results indicate that the BEH Amide column achieved the highest proportion of Good and Excellent peaks compared to ZIC-HILIC, confirming its superior chromatographic performance in positive ionization mode.*

Figure 5 A-B *Distribution of annotated metabolites by chemical class and peak quality under the optimized RP–HILIC workflow (BEH Amide ESI^+ + HSS T3 ESI^-). The pie chart shows that 84.88% of features achieved a peak rating between "Good" and "Excellent." B. The stacked histogram illustrates the relative abundance of chemical classes across peak quality categories, highlighting amino acids and derivatives, lipids (glycerophosphocholines and glycerophosphoethanolamines), and carnitines/acylcarnitines as the most represented among high-quality peaks.*

Chapter III

Figure 1 A-C *Principal Component Analysis (PCA) of metabolomic and lipidomic profiles. PCA score and loading plots of (**A**) metabolomics and (**B**) lipidomics datasets after pre-processing and auto-scaling show a clear separation between mixed cryoglobulinemia (MC) and HCC-HCV samples, while HCC and HCV display overlapping trends. (**C**) SUM-PCA improves class discrimination, particularly between HCC and HCV, highlighting the complementary contribution of the two omics layers to disease-specific metabolic characterization.*

Figure 2 A-C *PLS-DA identified the top 20 discriminant features in (**A**) metabolomics and (**B**) lipidomics datasets, spanning amino acids and derivatives, dipeptides, purine derivatives, acylcarnitines, lysophosphatidylcholines (LPCs), and alkyl-LPCs. Univariate analysis identified significantly altered metabolites and lipids ($p < 0.05$; Table 3), which were mapped to HMDB identifiers and used for pathway enrichment analysis (**C**), highlighting disease-associated metabolic perturbations.*

Figure 3 A-G *Distribution in HCC, HCV and MC patients of short (**A**), medium (**B**) and long-chain (**C**) acylcarnitine (CAR) and (**D**) Ile-Pro, ADMA and MG, where *p-value < 0.05 and **p-value < 0.01. Distribution in HCC, HCV and MC patients of (**E**) lysophosphatydilcholines (LPCs), (**F**) ether-linked lysophosphatydilcholine (LPC-O) and (**G**) phosphatydilcholine (PC), where *p-value < 0.05 and **p-value < 0.001.*

Figure 4 A-D *Graphical representation of confusion matrices obtained from PLS-DA models (metabolomics **A**, lipidomics **B**) of both independent modalities for the test phase classifying AFP-negative HCC patients. The ROC curves (metabolomics **C**, lipidomics **D**) reported compares the models' performances reducing the number of variables to the VIPs and the first 20 VIP.*

Figure S1 *Graphical representation of distribution and class information for metabolites (HMDB) and lipids (LIPID MAPS) detected by UHPLC-TIMS-Q-TOF analysis in HCC, HCV and MC plasma samples.*

Figure S2 *Graphical representation of confusion matrices obtained from PLS-DA models of independent modality. The reported confusion matrices refer to PLS-DA performance for the test phase.*

Figure S3 *Graphical representation of confusion matrices obtained from SIMCA models of both independent modalities. The reported confusion matrices refer to SIMCA performance for the test phase.*

Figure S4 *Graphical representation of confusion matrices obtained for AFP classification.*

Figure S5 *PCA scores and loadings plots (PC1, PC2 and PC3) of metabolomics and lipidomics datasets including QC with confidence ellipses (95%) for each class. The datasets were pre-processed by 1) normalization using internal standard 2) the missing values and zeros were replaced with one-fifth of the minimum value recorded in the data set for that molecule 3) logarithm values of the base of 10 and 4) autoscaling.*

Figure S6 *Sphingomyelins rt/carbon number and m/z measured/CCS linearity.*

Figure S7 *Short-hand annotation for SMs.*

Figure S8 *Long-hand annotation for SMs.*

Figure S9 *Phosphatidylcholines rt/carbon number and m/z measured /CCS linearity.*

Figure S10 *Long-hand annotation for PCs.*

Figure S11 *Comparison between the acquired (Query) and the library spectrum (Spectral library) of statistically significant lipids and metabolites (A: LPC O-16:0; B: LPC 18:1; C: Asymmetric dimethylarginine; D: CAR 4:0).*

Chapter IV

Figure 1A-C. *Extracted ion chromatograms (EICs) of some target metabolites in human plasma obtained using three different sample extraction strategies: (A) protein precipitation with ACN 100%, (B) protein precipitation with MeOH 80 % (v/v %), and (C) solid-phase extraction (SPE) eluted with MeOH. The comparison highlights differences in signal intensity and chromatographic profiles among the extraction approaches. 1: CAR C16:0, 2: CAR C16:0-d_9, 3: CAR C5:0, 4: CAR C5:0-d_9, 5: CAR C8:0, 6: CAR C8:0-d_9, 7: LPC 17:0, 8: LPC 17.0-d_5.*

Figure 2A-C A. *Extracted ion chromatograms (EICs) intensities of CAR C8 under different mobile phase additives: 10 mM ammonium acetate (AmAc), formic acid (FA), AmAc + acetic acid (AcA), and ammonium formate + FA (AmF + FA). Signal intensity increased by 5.5-fold (450%), 3.5-fold (250%), and 2.67-fold (166.75%) for AmAc, FA, and AmAc + AcA, respectively, relative to AmF + FA.* ***B.*** *Extracted ion chromatograms (EICs) of CAR C16:0 under optimized solvent conditions, showing improved peak shape and signal intensity at ACN 80 % (v/v %) (green trace) compared to ACN 80 % (v/v %) (grey trace).* ***C.*** *FWHM values of some target metabolites detected in human plasma samples across different mobile phase modifiers. AmF: Ammonium Formate, FA: Formic Acid, AmAC: Ammonium Acetate, AcA: Ammonium Acetate*.

Figure 3A-B. A*. Extracted ion chromatogram (EIC) of LPC 17:0 under optimized HILIC conditions, showing the separation of LPC regioisomer (sn-1/ sn-2).* ***B.*** *Representative MS/MS spectra displaying diagnostic fragment ions and relative intensity used for regioisomer discrimination.*

Figure 4A-D *Comparative assessment of acquisition strategies for trace-level metabolite detection. S/N ratios and peak areas measured under different acquisition modes* ***A.*** *DDA* ***B.*** *SIM* ***C.*** *MPX-SIM. Representative chromatographic profiles demonstrating enhanced sensitivity and reproducibility with multiplexed SIM.* ***D.*** *Evaluation of Orbitrap MS1 acquisition parameters for optimized chromatographic sampling in SIM-MPX.*

Figure 5A-B A. *PCA scores plot (PC1 vs PC2) obtained after autoscaling of the concentration data. 95% confidence ellipses were calculated for each class based on Hotelling's T^2 statistic in the PCA score space.* ***B****. The loadings plot was derived from the PCA loading matrix and represents the contribution of each metabolite to the first two principal components. HCC: hepatocellular carcinoma; HCV: hepatitis C virus infection.*

Figure 6A-B *Supervised classification was performed using partial least-squares discriminant analysis (PLS-DA) coupled with linear discriminant analysis (LDA). The number of latent variables (nLVs = 2) was selected by five-fold cross-validation. Confusion matrices were generated from predicted versus true class memberships for both the* ***A.*** *training and* ***B.*** *independent test sets and are displayed with row- and column-normalized percentages.*

Figure 7A-B *Aligned line plot of CARs and LPCs quantified patients with HCC (A) and HCV (B) patients. Data are presented as mean ± SEM.*

Figure 8A-B *Receiver operating characteristic (ROC) curves. Each curve represents the univariate classification performance of a single LPC* ***A.*** *and CAR* ***B****. metabolite in discriminating HCC from HCV samples. The x-axis indicates the false positive rate, and the y-axis the true positive rate across varying thresholds. The corresponding area under the curve (AUC) values are reported in the legend, highlighting the strong discriminative capability of most LPC species.*

Figure S1 A-B. *Targeted MS/MS spectra of LPCs and CARs. A. LPCs were confirmed by the diagnostic phosphocholine fragment at m/z 184.0733 ($C_5H_{15}NO_4P^+$), while CARs were validated by the characteristic fragment at m/z 85.0284 ($C_4H_5O_2^+$).*

Figure S2 *Elution window for LPC 18:0, PC 18:1 and CAR C2:0 detected in human plasma.*

Figure S3 *Scatter plots for individual metabolites report log10 transformed concentration values. Statistical differences between HCC and HCV groups were assessed using the non-parametric Mann - Whitney U test followed by FDR correction. * pvalue < 0.05;*
*** pvalue < 0.001.*

Chapter V

Figure 1 A-C *Peak width at half maximum (FWHM) decreased progressively with increasing flow rates (50–100 µL/min), reaching a minimum of 0.14 min at 100 µL/min (**A**). Correspondingly, peak capacity increased, achieving a maximum value of Nc = 78 under these conditions (**B**). Injection-volume experiments (0.2, 0.5, and 1 µL at 200 and 20 ng/mL) (**C-D**) demonstrated that the optimal balance between chromatographic efficiency and sensitivity was obtained with a 200 nL injection at 20 ng/mL, yielding a peak capacity of Nc = 97.*

Figure 2 A-B *(**A**) Extracted ion chromatograms (EICs) of selected metabolite standards, L-Homocitrulline, Uric acid, 7-Methylguanine, N4-Acetylcytidine, and L-Tryptophan, acquired using 1.0-mm i.d. (red) and 2.1-mm i.d. (black) LC configurations under identical experimental conditions. (**B**) Comparative analysis of peak intensities for polar standard metabolites obtained with the 1.0-mm and 2.1-mm i.d. methods, highlighting the signal enhancement achieved with the microbore setup.*

Figure 3 A-C *Performance comparison between 1.0-mm and 2.1-mm i.d. LC configurations. (**A**) Extracted ion chromatograms (EICs) of N4-Acetylcytidine at 1 ng/mL, with corresponding MS/MS spectra, demonstrating successful precursor detection and fragmentation using the 1.0-mm i.d. method, while no precursor ion was detected with the 2.1-mm i.d. configuration. (**B**) Comparison of limits of detection (LOD) and limits of quantification (LOQ) for the investigated metabolites obtained with the 1.0-mm (red) and 2.1-mm (black) setups. (**C**) Repeatability assessment expressed as signal variability for both configurations, highlighting the performance of the 1.0-mm (red) relative to the 2.1-mm i.d. (black) method.*

Figure 4 A-B *Impact of the 1.0-mm i.d. configuration on feature detection and metabolite annotation in untargeted metabolomics across different biological matrices. (**A**) Percentage increase in MS^1 features, MS^2 events, and library-matched MS/MS spectra obtained with the 1.0-mm i.d. setup relative to the 2.1-mm configuration. The microbore method yielded, on average, +38.95% more MS^1 features and +39.26% more MS/MS events. (**B**) Proportion of metabolites exclusively annotated by each method across all matrices, demonstrating that over 34% of annotations were uniquely achieved using the 1.0-mm i.d. setup, underscoring its enhanced performance for untargeted metabolomics applications.*

Figure 5 A-C *Comparative evaluation of the 1.0-mm and 2.1-mm i.d. LC configurations in real biological matrices. (**A**) Three-dimensional principal component analysis (PCA) score plot showing clear clustering of the four investigated matrices and separation between the 1.0-mm and 2.1-mm i.d. approaches, highlighting method-dependent differences in metabolomic profiles. (**B**) Representative base peak chromatograms (BPCs) of a patient-derived organoid (PDO) sample acquired with the 1.0-mm (red) and 2.1-mm i.d. (black) setups, illustrating enhanced signal intensity and peak definition with the microbore configuration. (**C**) Quantitative comparison of peak intensities for metabolites commonly annotated in all four matrices, demonstrating the signal gain achieved with the 1.0-mm i.d. method in real-sample analyses.*

Figure S1 *Effect of flow rate and injection volume on FWHM and peak capacity in HILIC mode.*

Figure S2 *Effect of gradient length on peak capacity on both RP and HILIC.*

Figure S3 *Final comparison of FWHM and peak capacity between 1.0- and 2.1-mm.*

Figure S4 *Calibration curves for 1.0 mm and 2.1 mm i.d. setups.*

Figure S5 *EICs of kynurenine m/z=209.0921 and citrulline m/z=176.1030 in positive ionization mode employing vaporizer temperature at 100 °C in green, 150 °C in red and 200 °C in black.*

Figure S6 *Overlapped traces of base peak chromatogram and relative backpressure in HILIC-HRMS from #1 to 250 consecutive injections.*

Figure S7 *Dot plots of MS1 and MS/MS events on 2.1 mm i.d (top) and 1.0 mm i.d (bottom) setups.*

Figure S8A-C *Graphical representation of metabolome coverage and overlap between 1.0 and 2.1 mm i.d.*

Chapter VI

Figure 1 A-B *H&E morphological features associated with atrophic and involutional processes. (**A**) Tumor-associated atrophic areas defined by sharply demarcated regions with cell loss, ghost-cell morphology, and hyalinized stroma. (**B**) Involutional changes in adjacent thyroid parenchyma, marked by ectatic macrofollicles bordered by flattened epithelial cells.*

Figure 2 A-F *(**A**) PAX8 staining illustrating heterogeneous thyroid-lineage marker expression, with strong nuclear positivity in lower-grade areas and reduced, patchy expression in high-grade regions (**B**); atrophic foci show near-complete loss of signal (A). (**C**) PAX8 positive and negative cells. (**D**) CD56 expression across tumor compartments, with higher intensity in selected high-grade regions (**E**) and markedly reduced or absent signal in atrophic areas. (**F**) CD68-positive macrophages broadly distributed in both intratumoral and stromal compartments. CD163-positive macrophages enriched at the interface between viable tumor and atrophic zones and in stroma adjacent to high-grade foci.*

Figure 3A-B *(**A**) Hierarchical clustering segregating samples into distinct high-grade, involutional, and atrophic groups. (**B**) PCA representation of proteomic variability, with PC1 separating high-grade and atrophic lesions and PC2 reflecting intra-group diversity.*

Figure 3 C-D *(**C**) VIP score distribution from the PLS-DA model, indicating the proteins that most strongly contribute to class discrimination among atrophic, involutional, and high-grade tissue areas. (**D**) Network representation of 19 significantly enriched functional categories (adjusted $p < 0.05$), encompassing proteostasis-related processes (including proteasome and calpain pathways), RNA-processing mechanisms (such as spliceosome components and cap hypermethylation), metabolic pathways related to GDP-mannose and carbohydrate metabolism, and ERBB2-associated signaling modules.*

Figure 5 A-C *Boxplots of the top discriminant proteins identified through PLS-DA VIP scores, illustrating differential expression patterns across atrophic, involutional, and high-grade tumor areas, *p-value < 0.05, **p-value < 0.01 and ***p-value < 0.001.*

Table Captions

Chapter II

Table 1 *Overview of the reversed-phase chromatographic conditions evaluated in ESI positive and negative mode, including stationary phases (HSS T3, Luna Omega C18, and Luna Omega Polar) and mobile phase additive compositions (formic acid, acetic acid, and ammonium fluoride).*

Table 2 *Overview of the hydrophilic interaction liquid chromatographic conditions evaluated in ESI positive and negative mode.*

Table 3 *Total number of metabolite annotations detected for each RP column under varying mobile phase additive conditions, illustrating the impact of stationary phase and additive choice on metabolite coverage.*

Chapter III

Table 1 *Overview of the reversed-phase chromatographic conditions evaluated in ESI positive and negative mode, including stationary phases (HSS T3, Luna Omega C18, and Luna Omega Polar) and mobile phase additive compositions (formic acid, acetic acid, and ammonium fluoride).*

Table 2 *Overview of the hydrophilic interaction liquid chromatographic conditions evaluated in ESI positive and negative mode.*

Table 3 *Total number of metabolite annotations detected for each RP column under varying mobile phase additive conditions, illustrating the impact of stationary phase and additive choice on metabolite coverage.*

Table S1 *Significant metabolites and lipids annotated according to variable importance in projection (VIP) scores (from PLS-DA) and p-value (one-way ANOVA).*

Table S2 *List of significant metabolites based on PLS-DA classification and VIPS scores (>1). The presented table includes the metabolite names, corresponding m/z measurements, VIPs scores pertaining to the PLS-DA model, ANOVA-derived P-values, correlation coefficients with age, and the associated R-p-values indicating the significance of these correlations.*

Table S3 *List of significant lipids based on PLS-DA classification and VIPS scores (>1). The presented table includes the metabolite names, corresponding m/z measurements, VIPs scores pertaining to the PLS-DA model, ANOVA-derived P-values, correlation coefficients with age, and the associated R-p-values indicating the significance of these correlations.*

Table S4 *List of annotated metabolites by HILIC-UHPLC-TIMS-Q-TOF in HCC, HCV and MC plasma samples. Metabolites annotation was carried out with a spectral library approach following the metabolomics standard initiative guidelines3. Abbreviations: RT, retention time; m/z meas, mass to charge measured; Mass meas., measured mass; CCS, cross collisional section; Mob., mobility. Primary ions have been used for quantification and additional adducts, if detected, were reported in round brackets.*

Table S5. *List of annotated lipids by RP-UHPLC-TIMS-Q-TOF in HCC, HCV and MC plasma samples. Lipids are sorted by class. Lipid annotation was carried out with a combined rule based and spectral library approach following the lipidomic standard initiative guidelines[4]. Abbreviations: RT, retention time; m/z meas, mass to charge measured; Mass meas., measured mass; CCS, cross collisional section; Mob., mobility. Primary ion has been used for normalization and additional adducts, if detected, were reported in round brackets. Lipid class abbreviations: cholesteryl esters (CEs), ceramides (Cers), diacylglycerols (DGs), dihexosylceramide (Hex2cer), hexosylceramides (Hexcers), lysophosphatidylcholines (LPCs), ether-linked lysophosphatidylcholines (LPC-Os), lysophosphatidylethanolamines (LPEs), phosphatidylcholines (PCs), ether-linked phosphatidylcholine (PC-Os) phosphatidylethanolamines (PEs), ether-linked phosphatidylethanolamine (PE-Os), phosphatidylinositols (PIs), sphingomyelins (SMs), triacylglycerols (TGs).*

Table S6. *List of annotated lipids by RP-UHPLC-TIMS-Q-TOF in HCC, HCV and MC plasma samples. Lipids are sorted by class. Lipid annotation was carried out with a combined rule based and spectral library approach following the lipidomic standard initiative guidelines[4]. Abbreviations: RT, retention time; m/z meas, mass to charge measured; Mass meas., measured mass; CCS, cross collisional section; Mob., mobility. Primary ion has been used for normalization and additional adducts, if detected, were reported in round brackets. Lipid class abbreviations: cholesteryl esters (CEs), ceramides (Cers), diacylglycerols (DGs), dihexosylceramide (Hex2cer), hexosylceramides (Hexcers), lysophosphatidylcholines (LPCs), ether-linked lysophosphatidylcholines (LPC-Os), lysophosphatidylethanolamines (LPEs), phosphatidylcholines (PCs), ether-linked phosphatidylcholine (PC-Os) phosphatidylethanolamines (PEs), ether-linked phosphatidylethanolamine (PE-Os), phosphatidylinositols (PIs), sphingomyelins (SMs), triacylglycerols (TGs).*

Chapter IV

CHAPTER I

Introduction

The advent of omics sciences has profoundly transformed biomedical research, driving a paradigm shift toward personalized and precision medicine (1). By enabling the comprehensive characterization of biological systems at multiple molecular layers, omics approaches provide unprecedented opportunities to understand disease mechanisms, monitor progression, and tailor therapeutic interventions. Among these disciplines, metabolomics occupies a unique position at the bottom of the "omics cascade," offering a direct functional readout of biochemical processes influenced by genetic, epigenetic, and environmental factors (2). Unlike genomics or transcriptomics, which predict potential outcomes, metabolomics captures the real-time metabolic state of cells and organisms, reflecting dynamic responses to physiological or pathological stimuli such as disease, drug exposure, or dietary changes. Metabolites, as downstream products of gene expression and upstream mediators of environmental interactions, serve as sensitive indicators of phenotype and disease, making metabolomics a powerful tool for biomarker discovery and mechanistic insight.

The analytical complexity of metabolomics arises from the extraordinary chemical diversity of the metabolome, which spans several orders of magnitude in concentration and includes highly heterogeneous classes such as amino acids, organic acids, nucleotides, sugars, lipids, and xenobiotics (3). This diversity precludes the use of a single analytical platform for comprehensive coverage. While nuclear magnetic resonance (NMR) spectroscopy and gas chromatography–mass spectrometry (GC–MS) remain valuable, liquid chromatography coupled to mass spectrometry (LC–MS) has emerged as the gold standard for metabolomics and lipidomics due to its versatility, sensitivity, and compatibility with a wide range of biomolecules (4). LC–MS enables both untargeted and targeted strategies: untargeted workflows aim to detect and annotate all measurable features in a sample, including unknowns, whereas targeted approaches focus on

predefined metabolite panels for absolute quantification, often achieving attomole-level sensitivity (5).

Mass spectrometry, often in hyphenation with ultra-high performance liquid chromatography (UHPLC), represents nowadays the golden standard for both metabolomic and lipidomic analysis. Untargeted methods usually rely on the employment of high-resolution mass spectrometry, which is capable of delivering crucial information for the identification of known and unknown metabolites/lipids.

Specifically, hybrid instruments such as quadrupole/QTOF and quadrupole/orbital ion trap instruments are the most prevalent in untargeted metabolomics and lipidomics studies, due to their inherent advantages in terms of mass accuracy, resolution, dynamic range, acquisition rate, and sensitivity (6). In the context of MS/MS stages, a range of fragmentation strategies is at one's disposal. These include collision-induced dissociation (CID) (7), Higher-energy C-trap Dissociation (HCD) (8), and ultraviolet photo dissociation (UVDP) (9). Although classic data-dependent acquisition (DDA) methods still dominate the field, there has been a recent shift towards data-independent strategies (DIA). Examples of such strategies include Sequential Window Acquisition of All Theoretical Mass Spectra (SWATH) (10) and All Ion Fragmentation (AIF) (11). These DIA strategies have been shown to improve coverage for low-abundance precursor ions and retrospective analysis by dedicated software.

Despite the remarkable advances in LC-MS-based metabolomics, comprehensive coverage of the metabolome remains a significant analytical challenge. The broad polarity and physicochemical variability of metabolites limit the effectiveness of single chromatographic strategies. Highly polar and ionic metabolites-such as amino acids, organic acids, nucleotides, sugar phosphates, and small co-factors-are notoriously difficult to analyze using conventional reversed-phase liquid chromatography (RP-LC), due to insufficient retention and poor chromatographic resolution. To address this limitation, Hydrophilic Interaction Liquid Chromatography (HILIC) has become a cornerstone technique for metabolomics, offering improved retention and selectivity for polar species (12).

HILIC-based LC–MS workflows have demonstrated superior performance in clinical and biomedical applications, particularly for profiling central carbon metabolism and nucleotide

pathways. However, HILIC separations are inherently complex, requiring careful optimization of the mobile phase composition, buffer concentration and pH, as well as the appropriate selection of the stationary phase chemistry (13).

Complementary use of RP and HILIC chromatography is now recognized as essential for expanding metabolome coverage, as each technique interrogates distinct regions of chemical space. Nevertheless, chromatographic optimization remains a critical determinant of analytical performance, influencing retention stability, ionization efficiency, and ultimately metabolite detectability.

In parallel with chromatographic developments, technological advances in mass spectrometry have further strengthened metabolomics workflows. Improvements in Orbitrap and TOF analyzers, acquisition speed, and spectral resolution have enhanced the ability to resolve complex mixtures, reduce spectral interferences, and characterize low-abundance metabolites in challenging biological matrices (14). These developments have substantially improved confidence in metabolite annotation and structural elucidation.

More recently, increasing attention has been directed toward the miniaturization of LC–MS workflows as a strategy to enhance analytical sensitivity while reducing sample and solvent consumption. Micro- and nano-flow LC–MS approaches exploit reduced column internal diameters and lower flow rates to improve electrospray ionization efficiency, leading to enhanced signal-to-noise ratios and lower limits of detection compared to conventional analytical-scale LC–MS (15). These features make miniaturized platforms particularly appealing for metabolomics applications involving limited or precious samples, including three-dimensional organoid cultures, micro-dissected tissues, Dried Blood Spots (DBS), and rare clinical specimens (16).

Although micro- and nano-LC–MS systems are well established in proteomics, their adoption in metabolomics has been comparatively limited (17). This discrepancy is largely attributable to challenges related to system robustness, susceptibility to clogging, limited throughput, and the lack of standardized workflows compatible with large-scale metabolomics studies. In addition, the chemical diversity and wide concentration range of metabolites impose stringent requirements on chromatographic stability and reproducibility, which have historically favored analytical-flow LC–MS platforms.

Recent technological advances are progressively mitigating these limitations. The development of low-flow UHPLC systems with improved pumping accuracy, advances in column manufacturing technologies, and optimized electrospray interfaces have significantly improved the robustness and usability of micro-flow LC–MS. As a result, miniaturized LC–MS is emerging as a viable and powerful alternative for high-sensitivity metabolomics, enabling deeper metabolome coverage and improved detection of low-abundance metabolites while maintaining analytical reproducibility and throughput compatible with biological and clinical studies.

Collectively, these innovations in LC–MS technology-encompassing orthogonal chromatographic strategies for polar metabolites, high-resolution mass spectrometry with advanced acquisition methods, and miniaturized workflows-are driving the evolution of metabolomics within the broader multi-omics landscape. Continued optimization and standardization of these approaches are essential to expand metabolome coverage, improve quantitative performance, and enable metabolomics to fully realize its potential as a cornerstone of personalized medicine.

In this context, my PhD project focuses on the development and optimization of mass spectrometry-based multi-omics strategies to support the discovery of new therapeutic approaches. Specifically, the research aims to exploit state-of-the-art LC–MS technologies-including high-resolution mass spectrometry, HILIC separations for polar metabolites, and miniaturized micro- and nano-flow LC–MS workflows-to achieve comprehensive coverage of the metabolome, lipidome and proteome in limited biological samples. By integrating these analytical innovations with multi-omics data, the project aims to generate a more comprehensive understanding of disease-related molecular mechanisms, to identify potential biomarkers, and to support the development of strategies for improved therapeutic evaluation.

Liquid chromatography–high-resolution mass spectrometry (LC–HRMS) has become a cornerstone of untargeted metabolomics, enabling comprehensive profiling of small molecules in complex biological matrices. Despite continuous advances in instrumentation, the overall performance of untargeted LC-HRMS workflows remains strongly dependent on chromatographic selectivity, mobile phase composition, and ionization efficiency. The extreme chemical heterogeneity of the metabolome, spanning a wide polarity and

concentration range, precludes the use of a single separation strategy and necessitates the careful optimization of complementary chromatographic approaches. In this context, reversed phase (RP) and hydrophilic interaction liquid chromatography (HILIC) represents the most widely adopted and complementary platforms for untargeted metabolomics, each preferentially accessing distinct regions of chemical space.

Building on this premise, the first part of this thesis focuses on the systematic optimization of LC–HRMS workflows for untargeted metabolomics, with particular emphasis on chromatographic robustness and data quality. Using human plasma as a representative complex matrix, multiple RP and HILIC stationary phases were evaluated under different mobile phase additive conditions and ionization modes. Beyond conventional metrics based solely on feature counts, method performance was assessed through an explicit peak-quality scoring framework integrating peak shape, reproducibility, and robustness. This approach enabled a more nuanced and biologically meaningful comparison of chromatographic conditions. The integration of optimized RP and HILIC workflows maximized metabolome coverage while maintaining high chromatographic quality, demonstrating that peak quality–driven method selection represents a critical yet often overlooked dimension in untargeted metabolomics. These results establish a transferable analytical framework aligned with emerging community recommendations for rigor, reproducibility, and interpretability in large-scale metabolomics studies (**Chapter II**).

The optimized LC–HRMS platform was subsequently applied to a clinically relevant case study addressing one of the most pressing challenges in hepatology and oncology: the early detection of hepatocellular carcinoma (HCC). HCC frequently develops in the context of chronic liver diseases, such as Hepatitis C Virus (HCV) infection, and remains difficult to diagnose at early stages due to the limited sensitivity and specificity of currently available circulating biomarkers, including alpha-fetoprotein (AFP). This unmet clinical need highlights the urgency for novel, non-invasive biomarkers capable of improving diagnostic accuracy and patient stratification.

Within this framework, an integrated untargeted metabolomics and lipidomics strategy was employed to investigate metabolic alterations associated with HCC in HCV-positive patients. By combining complementary RP and HILIC separations with HRMS detection, distinct molecular signatures linked to mitochondrial dysfunction, altered fatty acid

oxidation, and membrane lipid remodeling were identified. Multivariate statistical analyses demonstrated that the integrated metabolic profiles outperformed AFP in discriminating against HCC from non-malignant HCV-related conditions, including AFP-negative cases. These findings underscore the value of metabolomics and lipidomics not only as discovery tools but also as powerful approaches for uncovering disease-associated biochemical phenotypes with direct translational potential (**Chapter III**).

To bridge the gap between untargeted discovery and clinical applicability, the following chapter introduces a two-tiered analytical strategy combining untargeted profiling with rapid targeted validation. Based on the most discriminant metabolite classes identified in the discovery phase-particularly acylcarnitines and lysophosphatidylcholines-a fast and robust targeted HILIC-HRMS assay was developed using a quadrupole–Orbitrap platform operating in multiplexed single ion monitoring mode. This targeted method enabled high-throughput quantification with excellent analytical performance in terms of sensitivity, accuracy, linearity, and robustness. Importantly, despite the reduced dimensionality inherent to targeted panels, the selected metabolites retained strong diagnostic power, confirming the biological relevance and clinical promise of the identified signatures. This integrated workflow illustrates a practical path from biomarker discovery to quantitative validation, paving the way for future large-scale clinical studies and precision diagnostics in HCC (**Chapter IV**).

While untargeted and targeted LC–HRMS workflows have traditionally relied on conventional narrow-bore columns (2.1 mm i.d.), the growing demand for higher sensitivity and the increasing prevalence of limited-volume samples pose significant analytical challenges. These limitations are particularly relevant for emerging biological models and clinical specimens, such as dried blood spots, organoids, and microdissected tissues. In this context, chromatographic miniaturization represents a promising strategy to enhance sensitivity while reducing sample and solvent consumption.

The subsequent chapter therefore explores the impact of column miniaturization on untargeted metabolomics performance by systematically comparing a microbore 1.0 mm i.d. UHPLC configuration with the conventional 2.1 mm i.d. format. Using both standard mixtures and diverse biological matrices, the microbore approach demonstrated substantial gains in sensitivity, lower limits of detection and quantification, improved

repeatability, and significantly enhanced metabolome coverage in both MS1 and MS/MS dimensions. Importantly, these improvements were achieved without compromising robustness or throughput, highlighting the feasibility of microbore LC–HRMS as a practical and scalable solution for high-sensitivity metabolomics applications (**Chapter V**).

Finally, the conceptual and technological advances explored throughout this thesis converge in the analysis of highly challenging, low-input clinical samples. The last chapter extends the principles of sensitivity-driven workflow optimization to microdissected tumor tissues, integrating mass spectrometry–based proteomics with advanced sample preparation strategies. Specifically, a microdissected-volume–preserving (mDVP) proteomics workflow was applied to thyroid tumor specimens, enabling deep molecular characterization from extremely limited material (**Chapter VI**).

This study exemplifies how analytical miniaturization, methodological rigor, and optimized MS workflows can collectively overcome sample-related constraints, ultimately broadening the scope of omics technologies in translational and precision medicine.

Bibliography Chapter I

[1]. Hasin Y, Seldin M, Lusis A. Multi-omics approaches to disease. Genome Biol. 2017 Dec;18(1):83.

[2]. Fiehn O. Metabolomics--the link between genotypes and phenotypes. Plant Mol Biol. 2002 Jan;48(1–2):155–71.

[3]. Patti GJ, Yanes O, Siuzdak G. Metabolomics: the apogee of the omics trilogy. Nat Rev Mol Cell Biol. 2012 Apr;13(4):263–9.

[4]. Zhou B, Xiao JF, Tuli L, Ressom HW. LC-MS-based metabolomics. Mol Biosyst. 2012 Feb;8(2):470–81.

[5]. Cajka T, Fiehn O. Toward Merging Untargeted and Targeted Methods in Mass Spectrometry-Based Metabolomics and Lipidomics. Anal Chem. 2016 Jan 5;88(1):524–45.

[6]. Dettmer K, Aronov PA, Hammock BD. Mass spectrometry-based metabolomics. Mass Spectrom Rev. 2007 Jan;26(1):51–78.

[7]. Li P, Hoffmann WD, Jackson GP. Multistage Mass Spectrometry of Phospholipids using Collision-Induced Dissociation (CID) and Metastable Atom-Activated Dissociation (MAD). Int J Mass Spectrom. 2016 June 1;403:1–7.

[8]. Hsu FF. Electrospray ionization with higher-energy collision dissociation tandem mass spectrometry toward characterization of ceramides as [M + Li]+ ions: Mechanisms of fragmentation and structural identification. Anal Chim Acta. 2021 Jan;1142:221–34.

[9]. Brodbelt JS, Morrison LJ, Santos I. Ultraviolet Photodissociation Mass Spectrometry for Analysis of Biological Molecules. Chem Rev. 2020 Apr 8;120(7):3328–80.

[10]. Shikshaky H, Ahmed EA, Anwar AM, Osama A, Ezzeldin S, Nasr A, et al. A Novel Approach of SWATH-Based Metabolomics Analysis Using the Human Metabolome Database Spectral Library. Int J Mol Sci. 2022 Sept 18;23(18):10908.

[11]. Graça G, Cai Y, Lau C, Vorkas PA, Lewis MR, Want EJ, et al. Automated Annotation of Untargeted All-Ion Fragmentation LC–MS Metabolomics Data with MetaboAnnotatoR. Anal Chem. 2022 Mar 1;94(8):3446–55.

[12]. Tang D, Zou L, Yin X, Ong CN. HILIC-MS for metabolomics: An attractive and complementary approach to RPLC-MS. Mass Spectrom Rev. 2016 Sept;35(5):574–600.

[13]. K Langová A, Manzi M, Brejchová J, Kuda O, Holčapek M, Jirásko R. Enhancing metabolite coverage using dedicated mobile phases for individual polarity modes in HILIC-MS. Anal Bioanal Chem [Internet].

[14]. Kind T, Fiehn O. Advances in structure elucidation of small molecules using mass spectrometry. Bioanal Rev. 2010 Dec;2(1–4):23–60.

[15]. Fitz V, El Abiead Y, Berger D, Koellensperger G. Systematic Investigation of LC Miniaturization to Increase Sensitivity in Wide-Target LC-MS-Based Trace Bioanalysis of Small Molecules. Front Mol Biosci. 2022 July 18;9:857505.

[16]. Harwood-Stamper AJ, Rowland CA, Dunn WB. Development of microflow ultra high performance liquid chromatography-mass spectrometry metabolomic assays for analysis of mammalian biofluids. Metabolomics. 2024 Oct 25;20(6):120. Nosé V, Gill A, Teijeiro JMC, Perren A, Erickson L. Overview of the 2022 WHO Classification of Familial Endocrine Tumor Syndromes. Endocr Pathol [Internet]. 2022;33. Available from: https://doi.org/10.1007/s12022-022-09705-5

[17]. Sanders KL, Edwards JL. Nano-liquid chromatography-mass spectrometry and recent applications in omics investigations. Anal Methods. 2020;12(36):4404–17.

CHAPTER II

Broadening Metabolome Coverage through RP-HILIC Integration: A Comparative Study of Chromatographic Conditions

2.1 Abstract

Liquid Chromatography–High-Resolution Mass Spectrometry (LC–HRMS) is a cornerstone of untargeted metabolomics; however, method performance remains highly dependent on chromatographic selectivity, mobile phase composition, and ionization efficiency. In particular, the broad polarity range of the metabolome necessitates careful optimization of both reversed phase (RP) and hydrophilic interaction liquid chromatography (HILIC) workflows to maximize coverage while maintaining robust chromatographic quality.

In this study, we systematically evaluated complementary RP and HILIC approaches for untargeted metabolomics using human plasma samples. Three RP stationary phases (HSS T3, Luna Omega C18, and Luna Omega Polar) were compared under multiple mobile phase additive conditions in negative electrospray ionization (ESI^-), including formic acid, acetic acid, and ammonium fluoride. Two HILIC phases (BEH Amide and ZIC-HILIC) were assessed under identical conditions in both positive and negative ionization modes. Method performance was evaluated not only by feature annotation counts but also through an explicit chromatographic peak-quality scoring framework integrating peak shape and robustness metrics.

Among RP conditions, the HSS T3 column consistently outperformed alternative phases in ESI^-, achieving the highest metabolite annotation counts and superior peak quality. The use of ammonium fluoride as a mobile phase additive markedly enhanced ionization efficiency and chromatographic fidelity compared to acidic modifiers, highlighting its utility for negative-mode metabolomics. For HILIC separations in ESI^+, BEH Amide and ZIC-HILIC

provided comparable metabolome coverage; however, BEH Amide yielded a higher proportion of well-defined peaks, indicating more robust chromatographic performance.

Integration of the optimized RP (HSS T3, ESI^-) and HILIC (BEH Amide, ESI^+) datasets maximized metabolome coverage while maintaining high chromatographic quality, with 84.9% of detected features classified as good or excellent peaks. High-quality features were predominantly associated with biologically relevant classes, including amino acids, glycerophospholipids, and carnitines.

Overall, this work establishes a peak-quality-driven framework for chromatographic method selection in untargeted LC-HRMS. The combined BEH Amide (ESI^+) and HSS T3 with ammonium fluoride (ESI^-) strategy provides a robust, transferable platform for large-scale metabolomics and biomarker discovery, aligning with emerging community recommendations for analytical rigor and reproducibility.

2.2 Introduction

Liquid chromatography coupled with mass spectrometry (LC-MS) represents one of the most powerful analytical platforms for metabolomics, enabling high sensitivity and broad coverage of metabolites in complex biological matrices (1). However, achieving robust and reproducible results requires addressing several critical challenges, including the selection of chromatographic stationary phases and mobile phases, which strongly influence retention, separation efficiency, and metabolite detectability. These factors become even more relevant when dealing with highly heterogeneous metabolomes, where compounds span a wide polarity range.

In this study, we focus on the comparative evaluation of two complementary chromatographic approaches-Reversed Phase (RP) and Hydrophilic Interaction Liquid Chromatography (HILIC) to maximize metabolomic coverage. RP chromatography is traditionally employed for semi-polar and hydrophobic metabolites, such as lipids and organic acids, whereas HILIC is particularly suited for highly polar compounds, including amino acids, nucleotides, and sugars. The intrinsic differences in stationary phase chemistry and mobile phase composition between these techniques provide orthogonal selectivity, which is essential for untargeted metabolomics workflows.

By systematically assessing different stationary phases and mobile phase modifiers, this work aims to identify optimal conditions for the separation of polar metabolites and lipid species, ensuring high peak quality and analytical throughput. This comparative approach not only enhances metabolome coverage but also supports the detection of clinically relevant biomarkers, paving the way for integrated multi-omics strategies in disease research**.**

2.3 Methods

2.3.1 Standard solutions

Mixtures of endogenous standards were prepared to assess instrument stability, analytical sensitivity, and chromatographic performance under controlled conditions. Standard solutions were prepared at two concentration levels (200 ng/mL and 20 ng/mL) and analyzed in both positive and negative heated electrospray ionization modes (H-ESI). The selected standards encompassed representative metabolites from diverse chemical classes and polarity ranges, including Nicotinamide (NAM), L-Phenylalanine (L-Phe), Kynurenine, L-Threonine (L-Thr), L-Carnitine (L-Car), L-Glutamate (L-Glu), Pyroglutamic acid, Succinic acid, and Palmitic acid.

These mixtures were employed as quality control references for routine monitoring of instrument performance and for systematic evaluation of chromatographic conditions. Specifically, they enabled comparative assessment of metabolite separation efficiency as a function of polarity and facilitated investigation of the impact of different mobile phase modifiers in reversed-phase chromatography on retention behaviour, peak shape, and overall analytical robustness.

2.3.2 Sample preparation

30 µL of human plasma sample were thawed on ice and extracted with 300 µL, of ice-cold MeOH/H_2O (80:20, *v/v*). Then, 5 µL of Deuterated Amino Acid Standard Mixture (50 µg/mL) were added to samples. The samples were vortexed for 12 min and incubated at −20 °C for 30 min, followed by centrifugation at 19,275 rcf for 10 min at 4 °C. Supernatants were collected and evaporated to dryness using a SpeedVac concentrator (Savant, Thermo Scientific, Milan, Italy). Dried extracts were reconstituted in 50 µL of ACN/H_2O (70:30, *v/v*) for HILIC analysis or MeOH/H_2O (10:90, *v/v*) for RP analysis. A Quality control (QC) sample

was prepared by pooling 5 µL of every study sample. This QC sample was used for monitoring stability, precision, random errors and correcting for instrument fluctuations during the analytical run.

2.3.3 Instrumentation and UHPLC-HRMS Acquisition

2.3.3.1 Reversed-Phase (RP) Chromatography Conditions

Reversed phase (RP) chromatography was performed using three different stationary phases: HSS T3 (150 × 2.1 mm; 1.8 µm), protected with a Vanguard precolumn (5 × 2.1 mm; 1.7 µm) (Waters, Milan, Italy), Luna Omega C18 (100 × 2.1 mm; 1.6 µm) Phenomenex, Torrance, CA, USA), protected with a Phenomenex SecurityGuard™ cartridge, and Luna Omega Polar (100 × 2.1 mm; 1.6 µm) protected with a Phenomenex SecurityGuard™ cartridge columns. All RP separations were conducted using water (H_2O) as mobile phase A and acetonitrile (ACN) as mobile phase B, with different mobile phase modifiers evaluated to optimize ionization efficiency and chromatographic performance.

For ESI positive ion mode (ESI^+), mobile phase A consisted of water containing 0.1% formic acid, while mobile phase B consisted of acetonitrile containing 0.1% formic acid for all three columns (HSS T3, Luna Omega C18, and Luna Omega Polar).

For ESI negative ion mode (ESI^-), three different mobile phase additive conditions were investigated for each RP column. In the first condition, mobile phase A was water with 0.1% formic acid and mobile phase B was acetonitrile with 0.1% formic acid. In the second condition, mobile phase A consisted of water with 0.1% acetic acid and mobile phase B of acetonitrile with 0.1% acetic acid. In the third condition, mobile phase A consisted of water containing 1 mM ammonium fluoride, while mobile phase B consisted of acetonitrile without additive.

The following gradient has been used: 0 min, 0% B; 6 min, 70% B; 8 min, 80% B; 9 min, 98% B, 10 min 98% B; 10.1 min, 0% B; and 3.9 min for column re-equilibration. The column temperature was set at 45 °C and a flow rate of 0.4 mL/min was used. The injection volume was 2µL.

These different additive strategies were systematically evaluated to assess their impact on chromatographic behavior, peak shape, and ionization efficiency in negative ion mode. The chromatographic conditions studied are shown in **Table 1**.

2.3.3.2 Hydrophilic Interaction Liquid Chromatography (HILIC) Conditions

Two HILIC columns were evaluated in this study: a SeQuant® ZIC®-HILIC (100 × 2.1 mm; 3.5 µm) (Merck, Darmstadt, Germany) column and an Acquity BEH Amide column (100 × 2.1 mm; 1.6 µm), protected by a VanGuard BEH Amide™ precolumn, both operated under identical mobile phase compositions and analyzed in both positive and negative electrospray ionization modes ($ESI^{+/-}$).

For both HILIC columns, mobile phase A consisted of water/acetonitrile (95:5, *v/v*) supplemented with 10 mM ammonium acetate, while mobile phase B consisted of water/acetonitrile (5:95, *v/v*) containing 10 mM ammonium acetate. The chromatographic conditions were optimized to ensure efficient retention and separation of polar metabolites across a broad polarity range.

The following gradient was used: 0-1 min, 99%B, 7-8 min 30% B, 8.1-11.5 min 99% B. The column temperature was set at 45 °C and a flow rate of 0.4 mL/min was used. The injection volume was 2µL.

Both the ZIC-HILIC and BEH Amide columns were evaluated under these conditions in combination with ESI operated in both positive and negative ion modes. The chromatographic conditions studied are shown in **Table 2**.

Reversed Phase Chromatographic Conditions		
H-ESI Positive		
HSS T3 ESI^+	**Luna Omega C18 ESI^+**	**Luna Omega Polar ESI^+**
Phase A: H_2O + 0,1% HCOOH	Phase A: H_2O + 0,1% HCOOH	Phase A: H_2O + 0,1% HCOOH
Phase B: ACN + 0,1% HCOOH	Phase B: ACN + 0,1% HCOOH	Phase B: ACN + 0,1% HCOOH
H-ESI Negative		
HSS T3 ESI^-	**Luna Omega C18 ESI^-**	**Luna Omega Polar ESI^-**
	1st additive	
Phase A: H_2O + 0,1% HCOOH	Phase A: H_2O + 0,1% HCOOH	Phase A: H_2O + 0,1% HCOOH
Phase B: ACN + 0,1% HCOOH	Phase B: ACN + 0,1% HCOOH	Phase B: ACN + 0,1% HCOOH
	2nd additive	
Phase A: H_2O + 0,1% CH_3COOH	Phase A: H_2O + 0,1% CH_3COOH	Phase A: H_2O + 0,1% CH_3COOH
Phase B: ACN + 0,1% CH_3COOH	Phase B: ACN + 0,1% CH_3COOH	Phase B: ACN + 0,1% CH_3COOH
	3rd additive	
Phase A: H_2O + NH_4F 1mM	Phase A: H_2O + NH_4F 1mM	Phase A: H_2O + NH_4F 1mM
Phase B: ACN	Phase B: ACN	Phase B: ACN

Table 1. Overview of the reversed-phase chromatographic conditions evaluated in ESI negative mode, including stationary phases (HSS T3, Luna Omega C18, and Luna Omega Polar) and mobile phase additive compositions (formic acid, acetic acid, and ammonium fluoride).

Hydrophilic Interaction Liquid Chromatography Conditions	
H-ESI Positive and Negative	
ZIC ESI$^{+/-}$	**BEH Amide ESI$^{+/-}$**
Phase A: H_2O/ACN (95:5 *v/v* %) + 10 mM CH_3COONH_4	Phase A: H_2O/ACN (95:5 *v/v* %) + 10 mM CH_3COONH_4
Phase B: H_2O/ACN (5:95 *v/v* %) + 10 mM CH_3COONH_4	Phase B: H_2O/ACN (5:95 *v/v* %) + 10 mM CH_3COONH_4

Table 2. Overview of the hydrophilic interaction liquid chromatographic conditions evaluated in ESI positive and negative mode.

2.3.3.3 High Resolution Mass Spectrometry (HRMS) parameters

MS data acquisition was performed in full scan-data dependent acquisition (FS-DDA) in the m/z range 70–800. MS1 scan OT resolution, 60,000; AGC, auto; maximum injection time, 100 ms. S-Lens RF level, 70; ddMS2 OT resolution, 15,000; isolation window, 1.5 Da; dynamic exclusion, 10 s; AGC, auto; maximum injection time, 22 ms. TopN, 4; HCD fragmentation normalized collision energies (NCE): 20, 40, 60. The HESI source parameters were as follows: sheath gas pressure, 40 a.u. and 50 a.u. for positive and negative ionization modes, respectively; aux gas flow, 15 a.u.; sweep gas flow, 0 a.u. Spray voltages were set to 3.3 kV and 3.0 kV for ESI^+ and ESI^-. Ion transfer tube (ITT) and vaporizer temperatures were set to 300 °C and 320 °C. The instrument was externally calibrated daily with FlexMix solution (ThermoFisher) while at the beginning of every LC run the internal calibrant was injected (IC run start mode).

2.3.4 Repeatability

Method repeatability was evaluated through intra- and inter-batch analyses conducted under the optimized chromatographic conditions for each column. Before initiating batch runs, columns were equilibrated with the initial mobile phase gradient and conditioned by ten consecutive injections of pooled QC plasma samples.

The intra-batch assessment comprised 20 sequential injections of QC plasma within a single batch, whereas the inter-batch evaluation involved three separate batches, each including 20 QC plasma injections, spaced by two-day intervals.

Repeatability was determined by calculating the relative standard deviation (RSD) for both retention time and peak area across all injections.

2.3.5 Data Processing and Statistical Analysis

FreeStyle (Thermo Fisher Scientific) was used to visualize RAW data, which were then imported to Compound Discoverer v.3.3 (Thermo Fisher Scientific) to normalize, align, detect, and identify compounds. Features were extracted from 0–10 min and 0–11 min of the HILIC and RP chromatography runs, respectively, in the m/z = 70–800 mass range. Data were aligned according to an adaptive curve alignment model. Compounds were detected using the following parameters settings: mass tolerance was set to 5 ppm, while retention

time tolerance was set to 0.2 min; minimum peak intensity was set to 100,000 AU and the signal to noise threshold for compound detection was set to 5. The peak rating filter was set to 3. To perform blank subtraction, we maintained max sample/max blank ratio > 5. To predict elemental compositions of the compounds, the relative intensity tolerance was set to 30% for isotope pattern matching. For the mzCloud database search, both the precursor and fragment mass tolerance were set to 5 ppm. The databases used for matching compounds in ChemSpider for structural search were BioCyc, the Human Metabolome Database, and KEGG, and the mass tolerance in ChemSpider Search was set to 5 ppm. The mass tolerance for matching compounds in Metabolika pathways was set to 5 ppm. Compounds were assigned by comparing annotations using the following nodes in order of priority: (1) mzCloud; (2) Predicted Compositions; (3) MassList search; (4) ChemSpider Search; (5) Metabolika search. Detailed information about spectral match parameters are reported in **Supplementary Material Chapter V.** Principal component analysis was performed by MetaboAnalyst 6.0; samples were log transformed and autoscaled prior to statistical analysis, all other graphs were built using GraphPad Prism 8.0 (GraphPad Software, Boston, MA, USA).

2.4 Results

2.4.1 Comparison of RP stationary phases and mobile phases additives

The comparative evaluation of reversed-phase (RP) columns under negative electrospray ionization (ESI^-) conditions revealed marked differences in annotation performance as a function of both stationary phase and mobile phase additive. Among the tested columns-HSS T3, Luna Omega C18, and Luna Omega Polar-the HSS T3 consistently achieved the highest total number of metabolite annotations, regardless of the additive employed, confirming its superior overall performance (**Table 3**).

Regarding mobile phase modifiers, ammonium fluoride (1 mM) provided the greatest improvement when combined with HSS T3, significantly enhancing ionization efficiency and metabolite coverage. Acetic acid (0.1%) delivered the best results for both Luna Omega C18 and Luna Omega Polar, whereas formic acid (0.1%) was the least effective across all columns, yielding the lowest annotation counts (**Figure 1**).

These findings underscore the importance of optimizing both stationary phase chemistry and mobile phase composition to maximize metabolome coverage. The histogram below illustrates the distribution of annotation counts for each column–additive combination, highlighting the synergistic effect of ammonium fluoride with HSS T3 and the variable performance of acetic acid and formic acid across different stationary phases

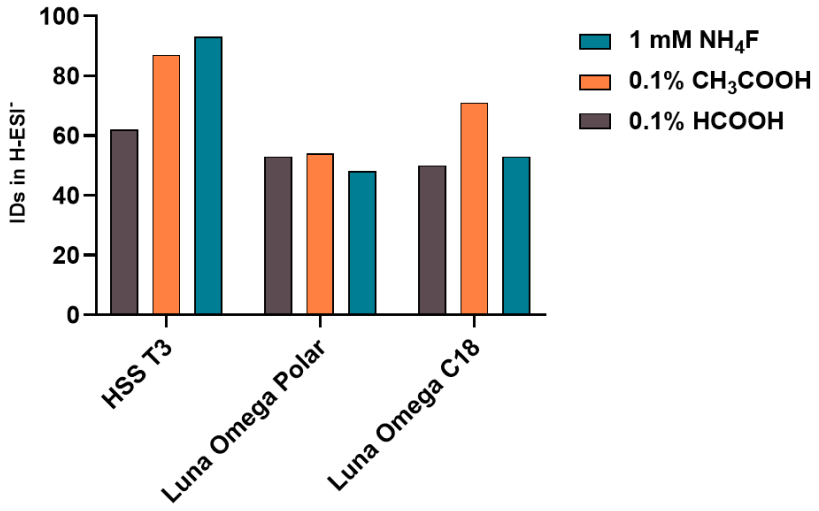


Figure 1. Effect of mobile phase modifiers on metabolite coverage across different reversed-phase columns. Ammonium fluoride (1 mM) combined with the HSS T3 column resulted in the highest ionization efficiency and metabolite annotation counts. For Luna Omega C18 and Luna Omega Polar columns, acetic acid (0.1%) provided optimal performance, whereas formic acid (0.1%) consistently showed the poorest performance across all columns, yielding the lowest number of annotated metabolites.

RP column	Additives			
	0.1% HCOOH	0.1% CH_3COOH	1 mM NH_4F	Total IDs
HSS T3	62	87	93	242
Luna Omega Polar	53	54	48	155
Luna Omega C18	50	71	53	174

Table 3. Total number of metabolite annotations detected for each RP column under varying mobile phase additive conditions, illustrating the impact of stationary phase and additive choice on metabolite coverage.

2.4.2 Chromatographic peak quality assessment on RP columns

To assess the quality of chromatographic peaks obtained under different reversed-phase (RP) conditions, we employed the Peak Rating Factor, which is calculated based on multiple parameters including FWHM² base, jaggedness, modality, and zigzag quality factor. This composite metric provides a robust evaluation of peak shape and overall chromatographic performance.

For interpretability, peak quality was categorized into four classes according to the peak rating score: Insufficient (1.00–3.99), Acceptable (4.00–5.99), Good (6.00–7.99), and Excellent (8.00–10.00), as illustrated in **Figure 2A-C**.

Our results indicate that the choice of mobile phase additives strongly influences peak quality. Ammonium fluoride (1 mM) consistently delivered the highest peak ratings when combined with the HSS T3 column, confirming its ability to enhance both ionization efficiency and chromatographic resolution. Acetic acid (0.1%) provided intermediate performance, particularly with Luna Omega C18 and Luna Omega Polar, whereas formic acid (0.1%) yielded the lowest peak ratings across all columns, indicating suboptimal conditions for negative ion mode.

Among the stationary phases tested, HSS T3 demonstrated the best overall performance, not only in terms of metabolite coverage but also for peak shape quality, as shown in **Figure 2D**. The histograms below report the distribution of peak rating values for each column–additive combination, highlighting the superior performance of ammonium fluoride with HSS T3 and the variable behavior of acetic and formic acid across different stationary phases.

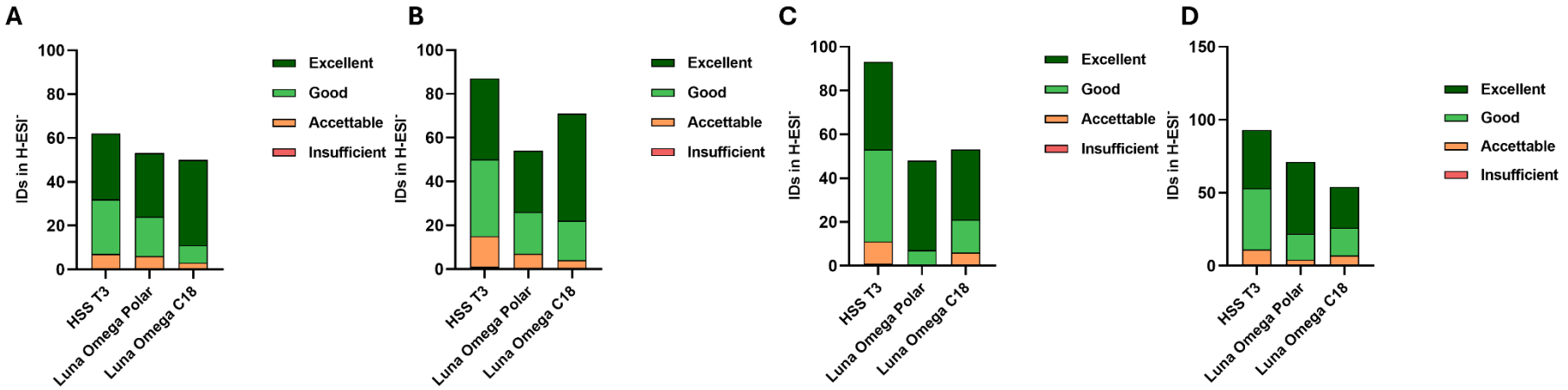


Figure 2A-D. Stacked bars show the proportion of features classified as Insufficient (1.00–3.99), Acceptable (4.00–5.99), Good (6.00–7.99), and Excellent (8.00–10.00) according to the Peak Rating Factor. Panels A–C compare peak quality obtained with formic acid, acetic acid, and ammonium fluoride, respectively, across the three RP stationary phases. Ammonium fluoride (1 mM) yields the highest fraction of Good and Excellent peaks, particularly on HSS T3. Panel D reports peak quality distributions under the optimal additive condition for each column, confirming the superior chromatographic performance of HSS T3.

2.4.3 Integration of RP and HILIC Data under Positive and Negative Ionization Modes

Following the identification of the most suitable mobile phase additive for RP analyses in ESI $^-$ mode, the evaluation was expanded to encompass data acquired in ESI^+ mode. This was then integrated with HILIC-based separations to ascertain the most efficacious chromatographic combination. Feature annotation datasets obtained from RP and HILIC (ESI^+) analyses were merged and systematically compared across the tested configurations.

The preliminary comparison yielded two promising configurations: ZIC-HILIC (ESI^+) + HSS T3 (ESI^-) or BEH Amide (ESI^+) + HSS T3 (ESI^-) (**Figure 3**). Despite the total number of annotated features being comparable between ZIC-HILIC and BEH Amide, a more in-depth investigation was conducted to differentiate between these two conditions by evaluating chromatographic peak quality. The evaluation of peak quality was conducted utilizing the Peak Rating Factor, a metric that integrates $FWHM^2$, jaggedness, modality, and zigzag quality factor.

The distribution of peak rating categories (Insufficient, Acceptable, Good, and Excellent) revealed that BEH Amide provided a higher proportion of Good and Excellent peaks compared to ZIC-HILIC, indicating superior chromatographic performance (**Figure 4**). This qualitative advantage, when combined with comparable feature coverage, resulted in the selection of the BEH Amide (ESI^+) and HSS T3 (ESI^-) combination for subsequent analyses, as it offered the optimal balance between metabolome coverage and peak annotation confidence.

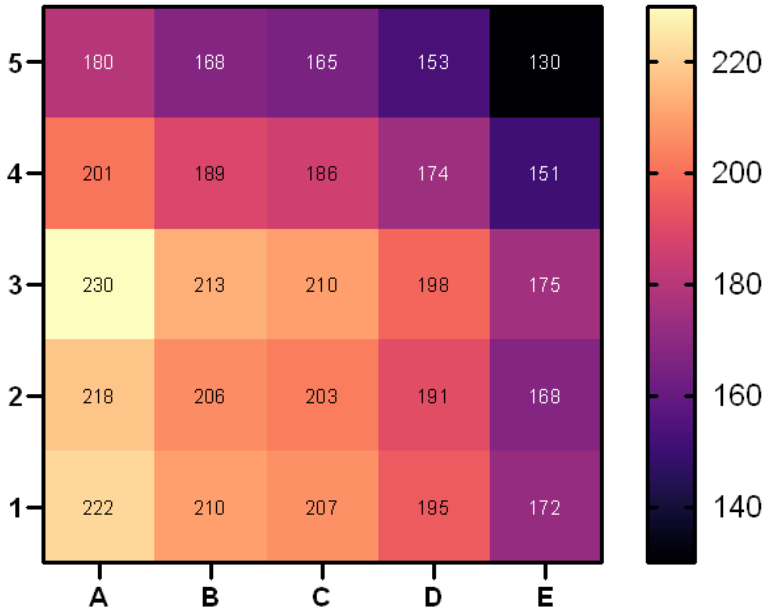


Figure 3. Heatmap illustrating the cumulative number of annotated features obtained by integrating all possible dataset combinations across chromatographic modes (RP and HILIC) and ionization polarities (ESI^+ and ESI^-). Color intensity reflects the total feature count for each combination, enabling visual comparison of metabolome coverage and identification of the most effective RP–HILIC pairing for comprehensive analysis. A: BEH Amide (+), B: ZIC HILIC (+), C: HSS T3 (+), D: Luna Omega C18 (+), E: Luna Omega Polar (+); 1: BEH Amide (-), 2: ZIC HILIC (-), 3: HSS T3 (-), 4: Luna Omega C18 (-), 5: Luna Omega Polar (-).

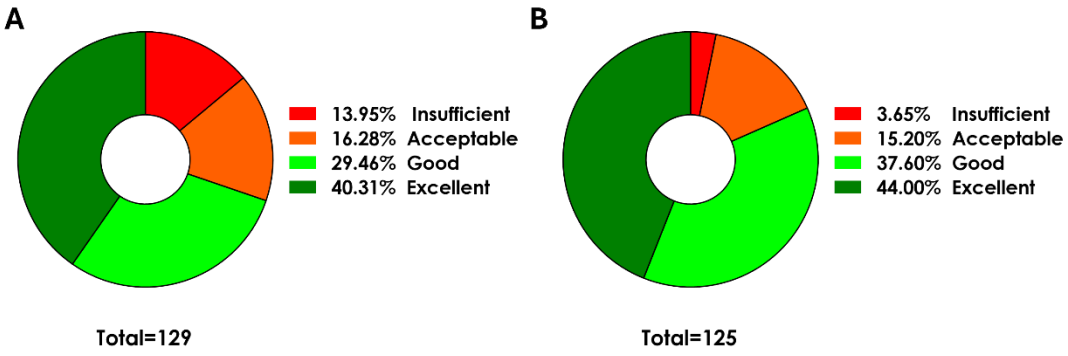


Figure 4. Pie charts illustrate the distribution of chromatographic peak quality categories (Poor, Acceptable, Good, Excellent) for HILIC columns under ESI^+ conditions. The results indicate that the BEH Amide column achieved the highest proportion of Good and Excellent peaks compared to ZIC-HILIC, confirming its superior chromatographic performance in positive ionization mode.

2.4.4 Integration of Chemical Class Distribution and Peak Quality

Following the selection of the optimal chromatographic combination (BEH Amide in ESI^+ and HSS T3 in ESI^-), the two datasets were merged to evaluate the distribution of annotated metabolites according to their chemical classification (**Figure 5A**)**.** This integrative analysis aimed to determine whether specific chemical classes exhibited distinct chromatographic behaviors under optimized conditions.

To visualize these trends, a stacked histogram was constructed, plotting the relative abundance of each chemical class against the corresponding peak rating categories (Insufficient, Acceptable, Good, Excellent). This approach enabled the identification of metabolite classes most frequently associated with high-quality peaks, providing insights into which compound families benefited most from the selected RP–HILIC configuration (**Figure 5B**).

Preliminary results indicate that 84.88% of the dataset consists of features with a peak rating between “Good” and “Excellent”, confirming the robustness of the optimized workflow. The most represented chemical classes among high-quality peaks were amino acids and derivatives, lipids (particularly glycerophosphocholines and glycerophosphoethanolamines), and carnitines and acylcarnitines. These findings suggest that the optimized RP–HILIC workflow not only maximizes overall metabolome coverage but also improves chromatographic fidelity for specific chemical subclasses, which is critical for downstream annotation confidence and quantitative reliability.

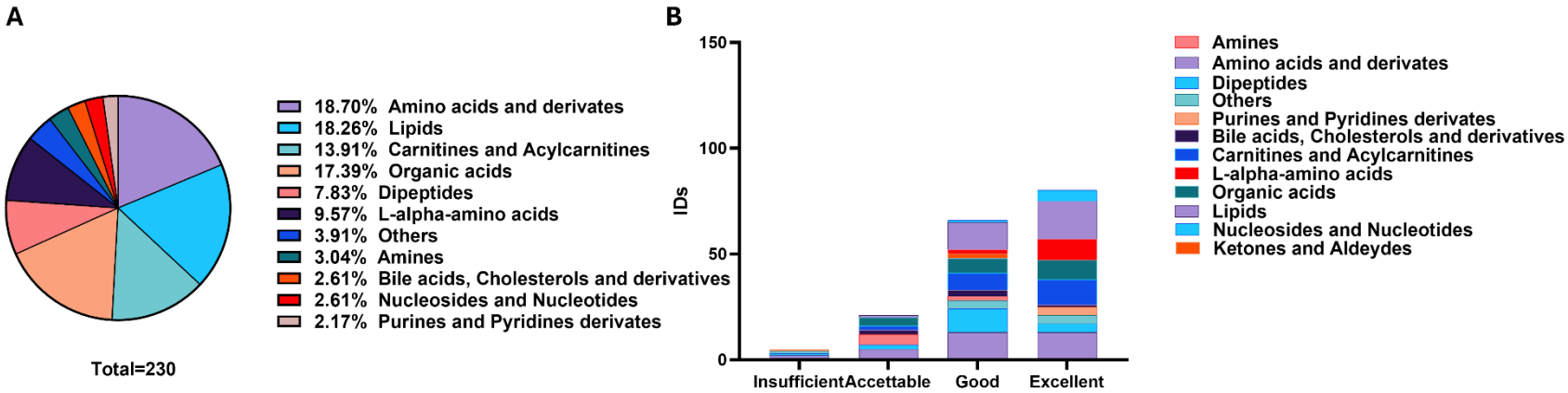


Figure 5 A-B. A. Distribution of annotated metabolites by chemical class and peak quality under the optimized RP–HILIC workflow (BEH Amide ESI^+ + HSS T3 ESI^-). The pie chart shows that 84.88% of features achieved a peak rating between "Good" and "Excellent." **B.** The stacked histogram illustrates the relative abundance of chemical classes across peak quality categories, highlighting amino acids and derivatives, lipids (glycerophosphocholines and glycerophosphoethanolamines), and carnitines/acylcarnitines as the most represented among high-quality peaks.

2.5 Discussion

A long-standing bottleneck in untargeted metabolomics is the simultaneous capture of highly polar small molecules and hydrophobic lipid species within a single analytical workflow. Because no single chromatographic mode can efficiently retain and resolve this chemical breadth, combining orthogonal separations has become best practice to expand coverage while maintaining data quality. In agreement with prior reviews and benchmarking studies, our side-by-side evaluation confirms that HILIC effectively retains very polar analytes (for example, amino acids, nucleotides, organic acids, sugars) that elute near the void in RP, whereas RP remains the workhorse for semi-polar and hydrophobic compounds, including numerous lipid classes. Integrating these modes therefore increases both the number and the chemical diversity of annotated features compared with single-mode LC–MS approaches (2).

Consistent with this rationale, merging datasets acquired under optimized RP and HILIC conditions in our study expanded metabolite coverage without compromising chromatographic quality, underscoring the practical orthogonality of the two modes when deployed in untargeted LC–MS.

A central finding of this work is the superiority of the HSS T3 stationary phase over Luna Omega C18 and Luna Omega Polar for RP–ESI$^-$, both in total annotations and in peak quality metrics. The T3 chemistry is designed to enhance retention of moderate polar species under high-aqueous conditions, a known limitation of conventional C18 phases, and this property likely contributes to its performance with metabolite classes commonly encountered in untargeted workflows (3).

Equally important, mobile phase additive chemistry exerted a pronounced effect on detectability in ESI $^-$. Ammonium fluoride (1 mM) delivered the largest gains in annotation counts and peak quality with HSS T3, outperforming acetic and formic acid; the latter was consistently suboptimal across columns. Mechanistically, NH_4F is known to boost negative-mode ionization-often by multiple-fold-through more efficient deprotonation and reduced ion-suppression, with gains reproduced across diverse small molecules and matrices. These trends align with recent systematic evaluations reporting 2–22-fold sensitivity improvements in ESI $^-$ with ammonium fluoride compared with conventional

additives, and with HILIC studies showing both higher S/N and improved robustness when switching from ammonium acetate to ammonium fluoride (4).

Together, these results emphasize that optimal RP performance in ESI $^-$ is not achieved by column selection alone; rather, column–additive pairing is decisive for maximizing annotation yield and chromatographic quality in untargeted metabolomics.

Beyond raw feature counts, chromatographic peak quality-shape, width, symmetry, and modality-directly affects downstream peak picking, alignment, and identification confidence. In our experiments, the HSS T3 + NH_4F condition not only maximized feature numbers but also produced the highest fraction of “Good” and “Excellent” peak rating scores, reducing the propagation of poorly formed peaks into statistical and identification pipelines. This is consistent with current best-practice guidance in metabolomics, which stresses that curation based on peak quality improves annotation confidence and reduces false positives, and with empirical studies showing that enforcing explicit peak quality thresholds can substantially enhance the explanatory power of downstream analyses (5,6).

When comparing two HILIC phases under ESI^+ conditions, BEH Amide and ZIC-HILIC provided comparable coverage; however, BEH Amide exhibited a higher proportion of well-defined peaks and more favorable peak quality distributions. This advantage is consistent with the amide phase’s neutral character and documented robustness across a range of analytes, as the mixed-mode retention mechanisms of BEH Amide have been shown to yield narrower peak widths and improved retention reproducibility compared to other HILIC stationary phases, leading to better overall chromatographic performance for a broad set of polar metabolites (7).

We note that the broader literature also documents cases in which zwitterionic HILIC phases (for example, ZIC-cHILIC/ZIC-pHILIC) excel for specific polar classes or isomer separations, highlighting that HILIC selectivity is strongly phase-dependent and study-specific. Such phase-dependent selectivity has been demonstrated in systematic comparisons of neutral and zwitterionic HILIC stationary phases, where zwitterionic chemistries showed superior separation and coverage for diverse polar compounds and isomeric features under certain pH and mobile phase conditions, underscoring the value of empirical screening during method development (8).

2.6 Conclusions

The present study proposes a practical and evidence-based framework for the selection of chromatographic methods in untargeted LC–MS metabolomics. The study emphasizes that optimal performance is achieved through the combined consideration of separation orthogonality, ionization efficiency, and chromatographic peak quality.

Primarily, the findings of this study serve to underscore the significance of leveraging chromatographic orthogonality in experimental design. The integration of reversed-phase and HILIC separations led to a substantial expansion in chemical coverage and annotation diversity. This confirmed that no single chromatographic mode can comprehensively capture the heterogeneous metabolome. Consequently, RP-HILIC pairing can be regarded as a rational and effective strategy for untargeted discovery workflows.

Furthermore, it is demonstrated that the judicious optimization of column–additive combinations is of particular significance in negative ionization mode. Among the conditions that were tested, the combination of an HSS T3 stationary phase with ammonium fluoride as a mobile phase additive provided superior sensitivity and chromatographic fidelity. This highlights the importance of aqueous-compatible RP chemistries and the selection of additives for enhancing metabolite detectability in ESI (−), where ionization efficiency often limits analytical depth.

Additionally, this study emphasizes the necessity of prioritizing chromatographic peak quality in conjunction with feature counts. The explicit use of peak-quality metrics enabled a more rigorous assessment of method performance and improved confidence in compound annotation and quantitative robustness. The integration of systematic peak-quality evaluation is consistent with the emerging community recommendations that are focused on enhancing transparency, reproducibility, and biological interpretability in untargeted metabolomics studies.

In conclusion, our comparison of HILIC stationary phases indicates that HILIC selectivity is inherently phase-dependent and study-specific. While BEH Amide delivered superior peak quality and overall robustness under the conditions tested here, zwitterionic HILIC phases remain advantageous for specific polar subclasses or isomeric separations. These

observations lend support to the value of limited empirical scouting during method development, with a view to tailoring HILIC selectivity to the analytical question at hand.

By jointly considering metabolome coverage, ionization efficiency, and chromatographic quality, the combined BEH Amide (ESI^+) and HSS T3 (ESI^-) workflow emerges as a robust and transferable platform for large-scale untargeted metabolomics, biomarker discovery, and integrative multi-omics applications. This strategy is fully aligned with evolving best-practice guidelines for rigorous metabolite identification, reporting, and analytical reproducibility.

Bibliography Chapter II

[1]. Hosseinkhani F, Huang L, Dubbelman AC, Guled F, Harms AC, Hankemeier T. Systematic Evaluation of HILIC Stationary Phases for Global Metabolomics of Human Plasma. Metabolites. 2022 Feb 9;12(2):165.

[2]. Öztepe T, Eylem CC, Reçber T, Nemutlu E. Advances in dual-column chromatography for metabolomics: Opportunities, challenges, and applications. J Chromatogr Open. 2025 Nov;8:100266.

[3]. Cífková E, Neuerová Z, Lísa M. Reversed-phase liquid chromatography/mass spectrometry approach for (un)targeted analysis of polar to mid-polar metabolites. Talanta. 2025 Aug;291:127853.

[4]. McFadden JR, Ames DM. Assessment of ammonium fluoride as a mobile phase additive for sensitivity gains in electrospray ionization. Anal Sci Adv. 2023 Dec;4(11–12):347–54.

[5]. Kumler W, Hazelton BJ, Ingalls AE. Picky with peakpicking: assessing chromatographic peak quality with simple metrics in metabolomics. BMC Bioinformatics. 2023 Oct 28;24(1):404.

[6]. Alseekh S, Aharoni A, Brotman Y, Contrepois K, D'Auria J, Ewald J, et al. Mass spectrometry-based metabolomics: a guide for annotation, quantification and best reporting practices. Nat Methods. 2021 July;18(7):747–56.

[7]. Langová A, Manzi M, Brejchová J, Kuda O, Holčapek M, Jirásko R. Enhancing metabolite coverage using dedicated mobile phases for individual polarity modes in HILIC-MS. Anal Bioanal Chem [Internet]. 2025 Nov 19

[8]. Hosseinkhani F, Huang L, Dubbelman AC, Guled F, Harms AC, Hankemeier T. Systematic Evaluation of HILIC Stationary Phases for Global Metabolomics of Human Plasma. Metabolites. 2022 Feb 9;12

CHAPTER III

A Metabolomics and Lipidomics Signature to Discriminate Hepatocellular Carcinoma (HCC) from Hepatitis C Virus (HCV) chronic infections and Lymphoproliferative disease

3.1 Abstract

Early detection of Hepatocellular Carcinoma (HCC) remains a major clinical challenge and is crucial for improving patient outcomes. Current circulating biomarkers, such as α-fetoprotein (AFP) and microRNAs, provide useful information but suffer from limited sensitivity and specificity, particularly in early-stage disease and AFP-negative cases. Therefore, novel non-invasive biomarkers are urgently needed to enhance diagnostic accuracy and patient stratification.

In this study, we applied an integrated Metabolomics and Lipidomics approach to investigate metabolic alterations associated with HCC in the context of chronic Hepatitis C Virus (HCV) infection. Untargeted plasma profiling was performed in a cohort of 102 HCV-positive subjects using Hydrophilic Interaction Liquid Chromatography (HILIC) and Reversed-Phase Ultra-High-Performance Liquid Chromatography (RP-UHPLC) coupled with High-Resolution Mass Spectrometry (HR-MS). Metabolic signatures were analyzed in relation to conventional liver function parameters and established tumor markers.

Our combined analysis revealed a distinct molecular phenotype associated with HCC, characterized by disruptions in mitochondrial fatty acid metabolism and phospholipid homeostasis, including alterations in acylcarnitines, amino acid-related metabolites, and lysophosphatidylcholines. Multivariate statistical modeling demonstrated that integrated

omics profiles outperformed AFP in distinguishing HCC from non-malignant HCV-related conditions, even in patients with low or negative AFP levels.

These findings highlight the potential of metabolomics and lipidomics as promising tools for identifying novel circulating biomarkers in HCC. This approach may contribute to improved early diagnosis and patient stratification, paving the way for the development of clinically applicable targeted assays.

3.2 Introduction

Hepatocellular carcinoma (HCC) is the most common histological subtype of primary liver cancer, accounting for nearly 90% of liver malignancies worldwide (1). Its development is a multistep process that typically occurs in the setting of chronic liver disease and cirrhosis. Major etiological factors include chronic hepatitis B and C virus (HBV and HCV) infections, alcohol abuse, non-alcoholic steatohepatitis (NASH), exposure to aflatoxins, and tobacco use. Chronic HCV infection also contributes to systemic metabolic and immunological alterations, sometimes manifesting as extrahepatic conditions such as mixed cryoglobulinemia (2,3).

Therapeutic options for HCC depend on tumour stage, liver function, and lesion burden. Curative strategies-such as surgical resection, local ablation, liver transplantation, and trans-arterial chemoembolization-are most effective in early or intermediate stages (4,5). Unfortunately, most cases are diagnosed at advanced stages, where treatment is limited to systemic therapies, including tyrosine kinase inhibitors and, more recently, immunotherapy-based regimens (6,7). Despite these advances, prognosis remains poor, with five-year survival rates below 20%, largely due to delayed diagnosis (8).

Early-stage HCC is often asymptomatic, and detection remains challenging even with advanced imaging, particularly for small lesions or in patients with confounding factors such as obesity or severe cirrhosis (1). Current surveillance relies on abdominal ultrasonography combined with serum biomarkers, including α-fetoprotein (AFP) and glypican-3 (GPC3). However, these markers lack adequate sensitivity and specificity, especially in early disease and AFP-negative patients, underscoring the need for more reliable diagnostic tools (9,10).

Recent efforts have focused on identifying novel biomarkers through genomic, transcriptomic, and proteomic approaches, such as gene expression profiling, microRNA

signatures, and protein-based markers (11,12). While these strategies have provided valuable insights into HCC biology, their clinical translation remains limited. In this context, metabolomics offers a powerful complementary approach, providing a direct functional readout of biochemical alterations associated with disease. By capturing the downstream effects of genetic, epigenetic, and environmental factors, metabolomics delivers a dynamic snapshot of cellular and systemic metabolism.

Previous metabolomic studies of HCC have revealed profound changes in energy metabolism, amino acid turnover, and lipid pathways (13,14). Lipid metabolism reprogramming is now recognized as a hallmark of cancer, with phospholipid remodelling playing a critical role in tumour growth, membrane dynamics, and signalling (15). Therefore, integrating metabolomic and lipidomic analyses holds promise for uncovering molecular signatures linked to HCC development and progression.

Building on this rationale, the present research aims to characterize metabolic and lipidomic alterations associated with HCC in patients with chronic HCV infection using high-resolution mass spectrometry. By combining untargeted metabolomics and lipidomics with multivariate statistical modelling, this study seeks to identify robust circulating signatures capable of distinguishing HCC from non-malignant HCV-related conditions, including AFP-negative cases. Ultimately, this approach may support the development of novel, non-invasive biomarkers for early HCC detection and improved clinical decision-making.

3.3 Methods

3.3.1 Participants' characteristics and collection of clinical samples

This retrospective study included 102 subjects with chronic hepatitis C virus (HCV) infection, stratified into three groups: patients diagnosed with hepatocellular carcinoma (HCC, *n* = 69), individuals with chronic HCV infection without evidence of malignancy (*n* = 23), and patients affected by mixed cryoglobulinemia (MC, *n* = 10). HCC and chronic HCV patients were enrolled at the Istituto Nazionale Tumori "Fondazione G. Pascale" (Naples, Italy), while MC patients were recruited at the Azienda Ospedaliera San Pio (Benevento, Italy).

HCC diagnosis was confirmed by imaging and histology, and all patients were classified as Barcelona Clinic Liver Cancer (BCLC) stage A or B. Surgical liver resection was performed

according to the Milan criteria. Chronic HCV infection was defined as persistent viremia for more than six months, confirmed by anti-HCV antibody detection (third-generation enzyme immunoassay) and HCV RNA quantification using the Cobas Amplicor assay (Roche).

Demographic and clinical data-including age, sex, viral load, and diagnosis-were retrospectively collected. Liver function tests and tumor biomarkers were measured using standardized assays: α-fetoprotein (AFP), carbohydrate antigen 19-9 (CA19-9), carcinoembryonic antigen (CEA), alanine aminotransferase (ALT), aspartate aminotransferase (AST), and gamma-glutamyl transferase (GGT). Reference limits were AFP ≤ 20 ng/L, CEA ≤ 3 ng/L, and CA19-9 ≤ 37 U/mL.

Peripheral blood (5 mL) was collected in EDTA tubes prior to any therapeutic intervention, including direct-acting antivirals. Samples were processed within two hours: plasma was separated by centrifugation at 1,200 × g for 15 minutes and stored at −80 °C until analysis.

Tumor size and number of nodules were assessed by computed tomography or magnetic resonance imaging. Histological grading followed Edmondson–Steiner criteria: most tumors were moderately differentiated (G2, *n* = 46) or poorly differentiated (G3, *n* = 2). Liver function was evaluated using the Child–Pugh classification (Child–Pugh A: *n* = 40; Child–Pugh B: *n* = 8) (16).

The study was approved by the Institutional Scientific Board and Ethics Committee of the Istituto Nazionale Tumori “Fondazione G. Pascale” (protocol number 51-OSS/21) and conducted in accordance with the Declaration of Helsinki.

3.3.2 Metabolome and Lipidome extraction

Plasma samples (20 µL) were thawed on ice and extracted using a biphasic protocol optimized for simultaneous metabolite and lipid recovery. Briefly, 225 µL of ice-cold methanol (CH_3OH) containing a mixture of deuterated internal standards (15:0-18:1(d7) PC (160 µg/mL), 15:0-18:1(d7) PE (5 µg/mL), 15:0-18:1(d7) PS (5 µg/mL), 15:0-18:1(d7) PG (30 µg/mL), 15:0-18:1(d7) PI (10 µg/mL), 15:0-18:1(d7) PA (7 µg/mL), 18:1(d7) LPC (25 µg/mL), 18:1(d7) LPE (5 µg/mL), 18:1(d7) CE (350 µg/mL), 18:1(d7) MG (2 µg/mL), 15:0-18:1(d7) DG (10 µg/mL), 15:0-18:1(d7)-15:0 TG (55 µg/mL), 18:1(d9) SM (30 µg/mL), Cholesterol (100 µg/mL), C16 Ceramide-d7 (d18:1-d7/16:0) (20 µg/mL), was added to each sample and

vortexed for 10 s. Subsequently, 750 µL of cold methyl tert-butyl ether (MTBE) were added, and the mixture was agitated in a thermomixer (Eppendorf, Milan, Italy) for 10 min at 300 rpm and 4 °C. Phase separation was induced by adding 188 µL of water (H_2O), followed by shaking for 20 s and centrifugation at 14,680 rpm for 10 min at 4 °C.

After centrifugation, the upper organic phase (MTBE, containing lipids) and the lower aqueous phase (MeOH/H_2O, containing metabolites) were carefully collected and evaporated to dryness using a SpeedVac concentrator (Savant, Thermo Scientific, Milan, Italy). To ensure analytical reproducibility and instrument stability, a quality control (QC) strategy was implemented. QC samples were prepared by pooling equal aliquots (10 µL) from each study sample. All injections were randomized, and blank samples were regularly analyzed to monitor carryover and background signals.

Dried extracts were reconstituted in 70 µL of acetonitrile/water (ACN/H_2O, 3:1 v/v) for metabolomics analysis and in 100 µL of butanol/isopropanol/water (BuOH/IPA/H_2O, 8:23:69 v/v) for lipidomics analysis.

3.3.3 Untargeted metabolomics and lipidomics profiling

Omics analyses were performed on a Thermo Ultimate RS 3000, which comprised a RS 3000 autosampler, a column oven and a binary pump with a 35 µL mixer. All connections were Viper (Thermo Fisher Scientific) stainless steel capillaries (0.100 mm I.D.). The column outlet was connected to the MS source with a peek tubing (0.100 mm I.D.) of the shortest length possible. The UHPLC system was coupled online to a TimsTOF Pro Quadrupole Time of Flight (Q-TOF) (Bruker Daltonics, Bremen, Germany) equipped with an Apollo II electrospray ionization (ESI) probe. The instrument was calibrated for both mass and mobility using the ESI-L Low Concentration Tuning Mix with the following composition: [m/z, $1/K_0$: (322.0481, 0.7318 Vs cm $^{-2}$), (622.0290, 0.9848 Vs cm^{-2}), (922.0098, 1.1895 Vs cm^{-2}), (1221.9906, 1.3820 Vs cm $^{-2}$)] in positive mode and [m/z, $1/K_0$: (301.99814, 0.6678 Vs cm^{-2}), (601.97897, 0.8781 Vs cm^{-2}), (1033.98811, 1.2525 Vs cm^{-2}), (1333.96894, 1.4015 Vs cm^{-2})] in negative mode. Before each LC-MS run a mixture (1:1 *v/v %*) of 10 mM sodium formate calibrant solution and ESI-L Low Concentration Tuning Mix was injected to recalibrate, respectively, the mass and mobility data.

3.3.3.1 Metabolome analyses

In detail, metabolome analyses were performed in HILIC mode, with an Acquity BEH Amide (150 × 2.1 mm; 1.7 µm) protected with a VanGuard amide precolumn (5 × 2.1 mm; 1.7 µm) (Waters, Milford, MA, U.S.A). The column temperature was set at 45 °C, and the flow rate was 0.350 mL/min. The mobile phase was (A): H_2O/ACN 95/5 (*v/v %*) and (B): in H_2O/ACN 5/ 95 (*v/v %*) both buffered with 10 mM CH_3COONH_4 plus 0.1% HCOOH (*v/v %*) for positive ionization while 10 mM CH_3COONH_4 plus 0.1% NH_4OH (*v/v %*, pH=8) additives were used for negative mode. The following gradient was employed: 0–0.1 min, 99 % B; 0.1–8 min, 99–50 % B; 8.0–8.5 min, 50-30 % B; 8.5-9.5 min isocratic at 30 % B; returning to 99% in 0.1 min, and then 4 min to recondition the column. The TIMS-MS analyses were performed in Data-Dependent Parallel Accumulation Serial Fragmentation (DDA-PASEF) with both positive and negative ionization in separate runs. 3 µL and 5 µL were injected for ESI^+ and ESI^- analysis, respectively. Source parameters: nebulizer gas (N_2) pressure: 3.0 bar, dry gas (N_2): 10 L/min, dry temperature: 220°C. Mass spectra were recorded in the range m/z 50–1000, with an accumulation and ramp time of 100 ms each. The ion mobility was scanned from 0.45 to 1.45 Vs/cm^2. Precursors for data-dependent acquisition were isolated within ± 2 m/z and fragmented with a TIMS-STEPPING ion mobility-dependent collision energy mode: CE [eV] #1: 50 and CE [eV] #2: 20. The total acquisition cycle was of 0.53 s and comprised one full TIMS-MS scan and two PASEF ramps. Exclusion time was set to 0.1 min, and Ion Charge Control (ICC) was set to 7.5.

3.3.3.2 Lipidome analyses

Lipidome analyses were performed with an Acquity UPLC CSH™ C18 column (50 × 2.1 mm; 1.7 µm, 130 Å) protected with a VanGuard CSH™ precolumn (5.0 × 2.1 mm; 1.7 µm, 130 Å) (Waters, Milford, MA, U.S.A). The column temperature was set at 65 °C, a flow rate of 0.55 mL/min was used, mobile phase consisted of (A): ACN/H_2O 60:40 (v/v %) and (B): IPA/ACN 90:10 (v/v %) both buffered with 10 mM $HCOONH_4$ and 0.1% HCOOH. The following gradient has been used: 0 min, 40% B; 0.4 min, 43% B; 0.425 min, 50% B; 0.9 min, 57% B; 2.0 min, 70% B; 2.950 min, 99% B; 3.3 min, 99% B; 3.301 min, 40% B and then 0.7 min for column re-equilibration. The TIMS-MS analyses were performed in DDA-PASEF with both positive and negative ionization, in separate runs. The injection volume was set at 2 µL and 4 µL for ESI^+ and ESI^-, respectively. Source parameters: Nebulizer gas (N_2) pressure:

3.0 Bar, Dry gas (N_2): 10 L/min, Dry temperature: 280°C. Mass spectra were recorded in the range m/z 100–1500, with an accumulation and ramp time of 100 ms each. The ion mobility was scanned from 0.55 to 1.70 Vs cm^{-2}. Precursors for data-dependent acquisition were isolated within ± 2 m/z and fragmented with a TIMS-STEPPING ion mobility-dependent collision energy mode: CE [eV] #1: 20-40 and CE [eV] #2: 35-50. The total acquisition cycle was 0.53 s and comprised one full TIMS-MS scan and two PASEF.

3.3.4 Metabolomics and Lipidomics pre-processing

Data alignment, feature detection, and annotation were performed using MetaboScape 2021 (Bruker) with the T-Rex 4D algorithm, which automatically extracts feature buckets from raw LC–MS files. For both metabolomics and lipidomics datasets, feature detection thresholds were set at 250 counts for positive and negative ionization modes, with a minimum of 100 data points in the 4D-TIMS space. A recursive feature extraction tool was applied with a minimum of 75 points to improve coverage.

Molecular formula assignment was carried out using SmartFormula™ (SF). Compound annotation employed stringent criteria: Mass accuracy: narrow 2 ppm, wide 10 ppm; mSigma: narrow 30, wide 250; MS/MS score: narrow 800, wide 150; Collision Cross Section (CCS): narrow ±2%, wide ±3.5%.

CCS values were validated against predictions from the CCSbase platform (17). Polar metabolite annotation utilized MS-DIAL libraries (18) (MSMS-Public-Pos and MSMS-Public-Neg), while lipid annotation combined rule-based approaches with the LipidBlast spectral library. Detailed information about lipid and metabolite annotations are reported in **Additional Section S1, Chapter III.**

Ion adduct handling: Lipidomics (positive mode): $[M+H]^+$, $[M+Na]^+$, $[M+K]^+$, $[M+H–H_2O]^+$, $[M+NH_4]^+$; Lipidomics (negative mode): $[M–H]^-$, $[M+Cl]^-$, $[M+HCOO]^-$, $[M–H_2O]^-$; Metabolomics (positive mode): primary $[M+H]^+$, seed ions $[M+Na]^+$, $[M+K]^+$, $[M+H–H_2O]^+$; Metabolomics (negative mode): primary $[M–H]^-$, seed ions $[M+Cl]^-$, $[M–H_2O]^-$

All spectra were manually curated to confirm annotation quality. Features missing in more than 75% of real samples or 50% of QC samples were excluded. Additionally, polar and non-polar molecules with a Coefficient of Variation (CV) >30% among QC replicates were discarded to ensure analytical robustness.

3.3.5 Chemometrics and Multivariate Data Analysis

Filtered data were processed using MATLAB R2022b (MathWorks Inc., Natick, MA, USA) through a combination of custom-developed routines and standard MATLAB functions for multivariate analysis. Each data set (metabolomics and lipidomics) was analyzed independently, and low-level data fusion was applied to integrate information across omics layers (44).

Metabolomics data were normalized by total ion current, while lipidomics data were normalized against class-specific internal standards. Missing values and zeros were imputed using one-fifth of the minimum value for the corresponding molecule. Data were then log-transformed (base 10) and autoscaled (mean-centered and divided by standard deviation) prior to chemometric modelling.

Principal Component Analysis (PCA) was performed on each data block after autoscaling to explore variance structure and detect patterns. A SUM-PCA approach was applied to fused datasets for integrated interpretation (19,20). PCA, an unsupervised dimensionality reduction technique, identifies principal components (PCs) that capture the most variance. Hotelling's T^2 confidence ellipses (95% confidence) were added to score plots for visual assessment of class separation.

To ensure comparability across omics modalities, the Kennard–Stone (KS) algorithm was used to partition data into training (70%) and test (30%) sets based on sample distances (21).

Two supervised models were applied to each modality: Partial Least Squares Discriminant Analysis (PLS-DA) (22) and Soft Independent Modelling of Class Analogy (SIMCA) (23,24).

The optimal number of latent variables (LVs) and PCs was determined via leave-one-out cross-validation to minimize misclassification and maximize accuracy. Model performance was evaluated using confusion matrices, calculating True Positive (TP), True Negative (TN), False Positive (FP), and False Negative (FN) rates. From these, sensitivity, specificity, and accuracy were derived for each class

Loadings and Variable Importance in Projection (VIP) scores were analyzed to identify discriminant molecules. These were compared with p-values from N-way ANOVA. Correlations with age were computed and tested for significance.

3.4 Results

3.4.1 Clinical characteristics of the study population

The demographic and clinical characteristics of the study population are summarized in **Table 1**. Most patients in both the HCC and HCV groups were male (71% and 61%, respectively; $p = 0.3354$). The mean age was 69 years for HCC patients and 57 years for HCV patients, with HCC cases predominantly older than 65 years ($p < 0.0001$). Chronic HCV infection represented the primary etiological factor for HCC.

Among HCC patients, 48 out of 69 presented with a single tumour nodule, and all tumours were classified as grade G2 according to the Edmondson–Steiner criteria. Biochemical analysis revealed no statistically significant differences in AST and AFP levels between groups, whereas ALT and GGT were significantly elevated ($p < 0.0196$ and $p < 0.001$, respectively). Furthermore, no significant correlation was observed between serum AFP levels and tumour characteristics (number of nodules or lesion size), consistent with previous findings by Carr et al. (25).

The modest association of low AFP concentrations (<20 ng/mL) with small HCC lesions underscores the limited sensitivity of AFP as a diagnostic marker and highlights the need for alternative biomarkers to improve early detection (26,27). AFP was not measured in patients with mixed cryoglobulinemia.

Patient characteristics	**HCC** (n=69)	**HCV** (n=23)	**MC** (n=10)	**p- value***
Age (years), n (%)				
≤ 65 years	17 (25%)	20 (87%)	4 (40%)	<0.0001
> 65 years	52 (75%)	3 (13%)	6 (60%)	
Gender, n (%)				
Male	49 (71%)	14 (61%)	5 (50%)	0.3354
Female	20 (29%)	9 (39%)	5 (50%)	
AFP (ng/mL), n (%)				
≤ 20 ng/mL	34 (49%)	22 (95%)	10 (100%)	0.4973
> 20 ng/mL	35 (51%)	1 (5%)	0	
CA19-9 (U/ml), n (%)				
≤37 U/mL	41 (59%)	nd	nd	
>37 U/mL	28 (41%)	nd	nd	
CEA (ng/mL), n (%)				
≤3 ng/mL	24 (35%)	nd	nd	
>3 ng/mL	45 (65%)	nd	nd	
ALT (U/l), n (%)				
≤33 U/l	6 (9%)	nd	4 (40%)	0.0196
>33 U/l	63 (91%)	nd	6 (60%)	
AST (U/l), n (%)				
≤32 U/l	6 (9%)	nd	2 (20%)	0.266
>32 U/l	63 (91%)	nd	8 (80%)	
GGT (U/l), n (%)				
≤40 U/l	17 (25%)	nd	7 (70%)	0.0071
>40 U/l	52 (75%)	nd	3 (30%)	
Tumor size, n (%)				

≤ 4 cm	46 (67%)	N/A	N/A
> 4 cm	23 (33%)	N/A	N/A
Tumor nodules, n (%)			
Single	48 (69.6%)	N/A	N/A
Multiple	21 (30.4%)	N/A	N/A
Tumor differentiation, n (%)			
G2	69 (100%)	N/A	N/A
G3	0	N/A	N/A
Child Pugh, n (%)			
A	60 (87%)	N/A	N/A
B	9 (13%)	N/A	N/A

Table 1. Laboratory results and clinical data of HCC, HCV and MC patients (*p-values were calculated using one-way ANOVA) nd: not detected, N/A: not applicable.

3.4.2 Untargeted metabo-lipidomics profiling

Dedicated high-resolution mass spectrometry (HRMS) strategies were employed for comprehensive metabolomic and lipidomic profiling. Metabolome analysis was performed using hydrophilic interaction liquid chromatography coupled with HRMS (HILIC-HRMS) (28), a suitable approach for polar metabolites that are poorly retained in reversed-phase systems. Lipidome analysis utilized a previously optimized reversed-phase UHPLC-HRMS (RP-UHPLC-HRMS) workflow (29). Compound annotation followed the guidelines of the Metabolomics Standards Initiative and Lipidomics Standards Initiative (30,31).

The initial untargeted workflow detected 15,757 metabolomic features and 3,488 lipidomic features. Each feature underwent stringent filtering based on multiple criteria: Mass accuracy: $\Delta m/z \leq 5$ ppm; Collision cross-section (CCS) error: $\Delta CCS \leq 3\%$; Peak shape and retention time consistency; Most probable adduct form; MS/MS spectral similarity score; Carryover assessment; Coefficient of Variation (CV) < 30% in QC samples.

Given the complexity of lipid annotation (32), each lipid species was manually curated to evaluate: Adduct formation under electrospray ionization; Retention behavior, applying the equivalent carbon number (ECN) model for RPLC; MS/MS spectrum quality: lipids were reported in shorthand form when MS^2 spectra contained fragments related only to the lipid class, and in longhand form when fragments corresponding to fatty acyl chains were present.

After manual curation, 280 compounds were annotated with high confidence, covering multiple classes and subclasses. The median annotation metrics were MS/MS score: 899.18; $\Delta m/z$: −0.21 ppm; ΔCCS: 1.35%.

Moreover, 90.57% of polar metabolites exhibited CV values < 20%, confirming analytical robustness. Principal Component Analysis (PCA) score plots demonstrated tight clustering of pooled QC samples, indicating excellent system stability throughout the batch (**Fig S5**).

3.4.3 Multi-omics Data Integration and Exploratory Analysis

Exploratory analysis was performed using Principal Component Analysis (PCA) on pre-processed and auto-scaled metabolomics and lipidomics datasets. PCA score and loading plots for each modality are shown in **Figure 1A–B.** The PCA results revealed a clear separation between mixed cryoglobulinemia (MC) and HCC-HCV samples in both

metabolomic and lipidomic datasets. Conversely, HCC and HCV groups exhibited overlapping trends, suggesting similar metabolic and lipidomic profiles. When integrating both omics layers, class separation improved slightly, particularly between HCC and HCV and relative to the MC group (**Figure 1C**). This enhanced discrimination likely reflects the complementary nature of metabolome and lipidome information, which together provide a more comprehensive representation of disease-specific biochemical alterations.

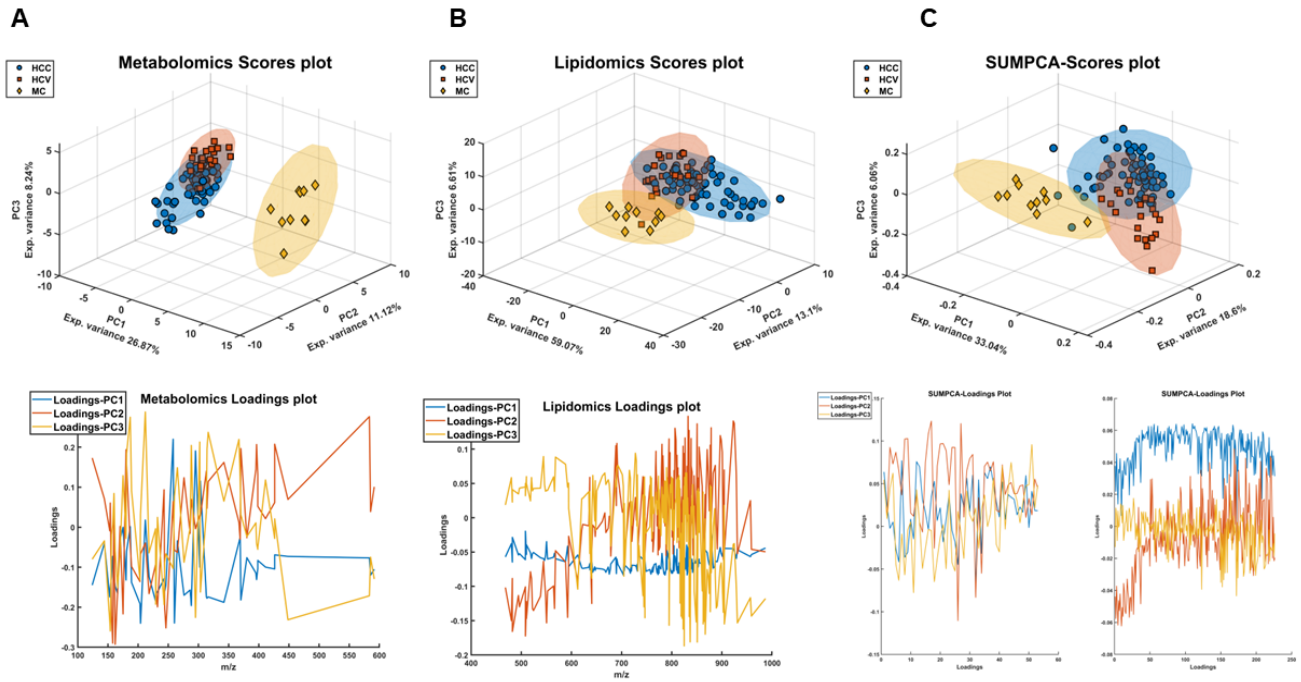


Figure 1 A–**C** Principal Component Analysis (PCA) of metabolomic and lipidomic profiles. PCA score and loading plots of (**A**) metabolomics and (**B**) lipidomics datasets after pre-processing and auto-scaling show a clear separation between mixed cryoglobulinemia (MC) and HCC-HCV samples, while HCC and HCV display overlapping trends. (**C**) SUM-PCA improves class discrimination, particularly between HCC and HCV, highlighting the complementary contribution of the two omics layers to disease-specific metabolic characterization.

3.4.4 Comparison of SIMCA and PLS-DA Performance in Classifying HCV-Related Conditions

To evaluate classification performance, two supervised multivariate models-Partial Least Squares Discriminant Analysis (PLS-DA) and Soft Independent Modeling of Class Analogy (SIMCA)-were independently applied to each dataset. The comparison aimed to highlight the distinct strengths of these approaches.

PLS-DA is primarily designed for discrimination, enabling characterization and differentiation among the three study groups representing different stages of HCV-related disease. In contrast, SIMCA focuses on modeling each class independently, addressing the challenge of distinguishing a specific category from all others. **Table 2** summarizes the comparative results of both methods, including sensitivity, specificity, and overall accuracy obtained through cross-validation and external testing. The table also reports the number of latent variables (LVs) used in PLS-DA and principal components (PCs) employed in SIMCA. Confusion matrices for the test set are provided in **Supplementary Figures S2** for PLS-DA and SIMCA (interpreted as a discriminant approach), respectively. PLS-DA generally achieved higher specificity and overall accuracy, consistent with its algorithmic design for discrimination. SIMCA exhibited superior sensitivity in certain scenarios, particularly when modeling individual classes. Performance varied across datasets and classes, underscoring the importance of selecting the appropriate model based on the clinical question.

		SIMCA					PLS-DA				
Method	Classes	PCs	Sensitivity (CV%)	Specificity CV (%)	Sensitivity Test (%)	Specificity Test (%)	LVs	Sensitivity CV (%)	Specificity CV (%)	Sensitivity Test (%)	Specificity Test (%)
Lipidomics	HCC	4	91.84	70.92	100	66.67	3	93.88	100	85	88.89
	MC	1	57.14	100	100	100		100	100	100	100
	HCV	3	94.12	82.25	100	69.57		100	94.64	83.33	86.96
Metabolomics	HCC	5	83.67	50	95	33.33	2	93.88	95.83	100	100
	MC	1	28.57	100	100	100		100	100	100	100
	HCV	4	47.06	97.79	100	91.3		94.12	94.64	100	100

Table 2. Values of SIMCA and PLS-DA performances. It reports on the values of sensitivity and specificity for both cross-validation (leave one out) and test sets. The abbreviations indicate respectively: PCs-Principal Components; CV-Cross Validation: LVs-Latent Variables.

3.4.5 Differential metabolites and lipids

PLS-DA applied to both metabolomics and lipidomics identified the top 20 discriminant compounds per modality (**Figure 2A–B**). These features span multiple classes-amino acids and derivatives, dipeptides, purine derivatives, acylcarnitines, lysophosphatidylcholines (LPCs), and alkyl-LPCs-consistent with metabolic rewiring in HCV-related disease and HCC. To support model-based rankings, we performed univariate testing and listed significant metabolites and lipids ($p < 0.05$) in Table 3. HMDB identifiers were then used for pathway-level interpretation via enrichment analysis (**Figure 2C; Supplementary Table S3**).

Pathway enrichment highlighted two significant processes ($p < 0.05$): (i) Mitochondrial β-oxidation of short-chain saturated fatty acids and (ii) Phospholipid biosynthesis. Both are biologically coherent with HCC pathophysiology: hepatic mitochondrial fatty-acid oxidation (mtFAO) underpins energy homeostasis via the carnitine shuttle, while phospholipid remodeling sustains membrane dynamics and signaling. Together, these axes capture complementary facets of tumor bioenergetics and membrane biogenesis.

We observed a chain-length–specific modulation of acylcarnitines across diagnostic groups: Short- and long-chain acylcarnitines (e.g., CAR 2:0, CAR 3:0, CAR 5:1; CAR 14:1, CAR 16:2, CAR 18:1) were higher in HCC vs. HCV and MC (**Figure 3A–C**). Medium-chain acylcarnitines (e.g., CAR 9:0, CAR 10:0, CAR 10:1) were higher in HCV vs. HCC and MC (**Figure 3B**). These patterns suggest segment-specific constraints along the β-oxidation spiral and altered substrate handling through the carnitine shuttle, consistent with lipid-centric reprogramming in HCC.

Beyond FAO, several polar metabolites were differentially abundant: the dipeptide isoleucyl-proline (Ile-Pro), the methylation by-product asymmetric dimethylarginine (ADMA), and methylguanine (MG) a modified purine. All reached their highest levels in HCC, intermediate in HCV, and lowest in MC (**Figure 3D**), pointing to enhanced amino-acid turnover, methyl-donor utilization, and purine pathway remodeling.

Lipidomics revealed a marked reduction of LPCs in HCC-both saturated (LPC 17:0, LPC 18:0) and unsaturated (LPC 18:1/18:2/18:3; LPC 20:3/20:4)-as well as alkyl-LPCs (LPC O-16:0, LPC O-16:1), relative to HCV and MC (**Figure 3E–F**). In parallel, several phosphatidylcholines (PCs) (PC 18:1_22:6, PC 16:1_18:2_A, PC 18:1_18:2, PC 40:8) tend to

be higher in HCV than HCC, with both HCC and HCV > MC (**Figure 3C**). The LPC↓/PC↑ divergence is consistent with increased LPC→PC reacylation (e.g., LPCAT activity) and altered Kennedy pathway flux, supporting membrane remodeling and proliferative demand in the tumor.

Correlation analyses (**Supplementary Tables S2–S3**) showed no meaningful age association for VIP>1 features (all $|r| \leq 0.63$ in metabolomics; ≤ 0.47 in lipidomics), supporting the disease-specific nature of the signatures employed for modeling.

The chain-length-specific acylcarnitine signature (mtFAO) and coordinated phospholipid remodeling (LPC↓/PC↑) provide a biologically coherent, multi-omics fingerprint of HCC within the HCV context. Such circulating markers-validated by convergent multivariate and univariate evidence-could enhance non-invasive stratification, particularly in AFP-negative cases, and motivate targeted assays for early detection.

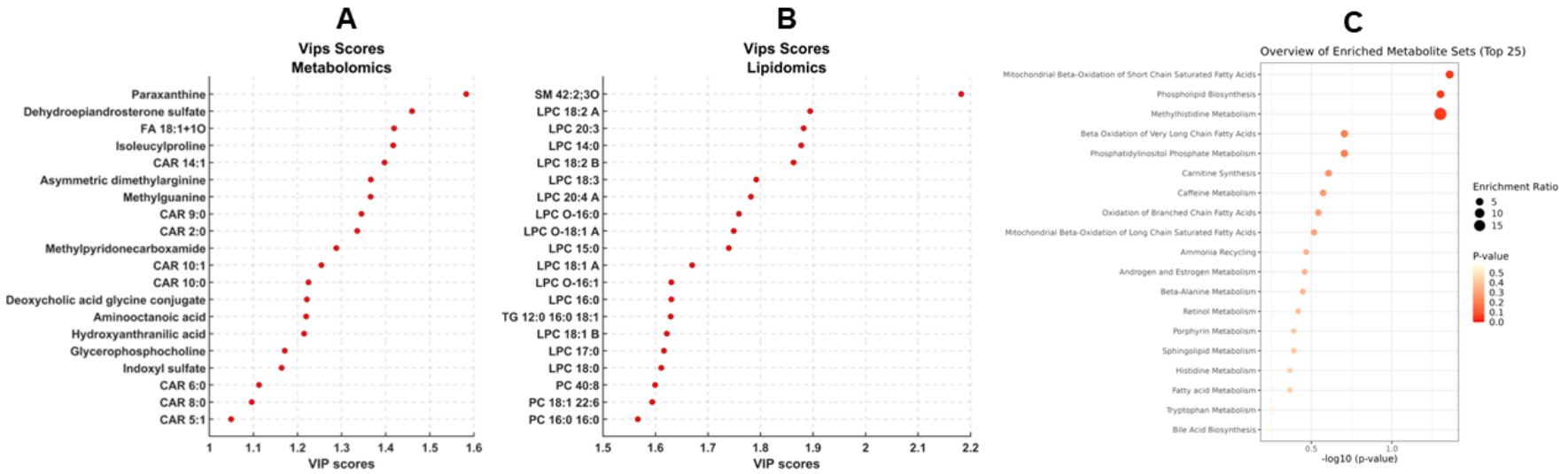


Figure 2A–C PLS-DA identified the top 20 discriminant features in (**A**) metabolomics and (**B**) lipidomics datasets, spanning amino acids and derivatives, dipeptides, purine derivatives, acylcarnitines, lysophosphatidylcholines (LPCs), and alkyl-LPCs. Univariate analysis identified significantly altered metabolites and lipids ($p < 0.05$; Table 3), which were mapped to HMDB identifiers and used for pathway enrichment analysis (**C**), highlighting disease-associated metabolic perturbations.

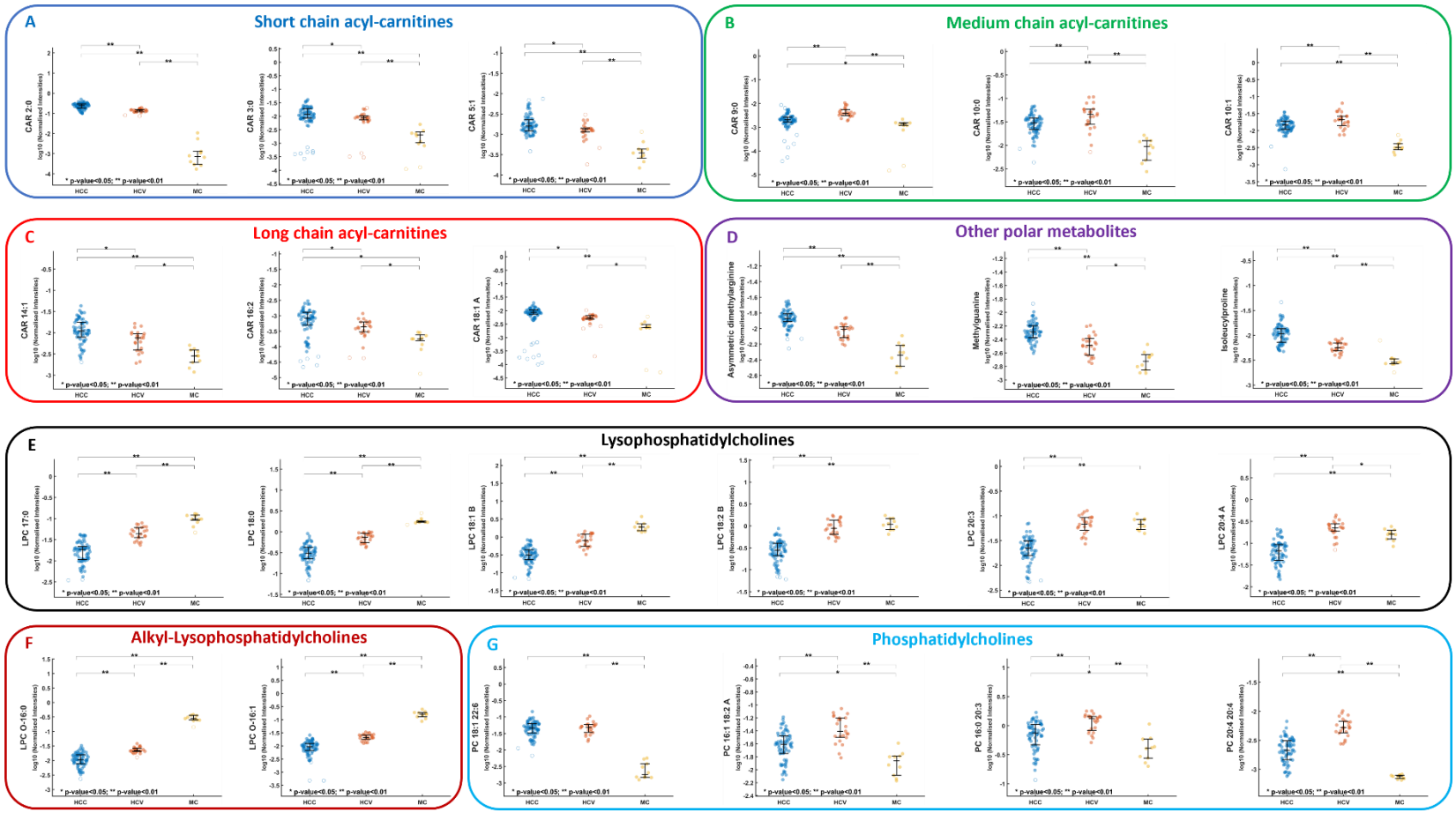


Figure 3 A-G Distribution in HCC, HCV and MC patients of short (**A**), medium (**B**) and long-chain (**C**) acylcarnitine (CAR) and **D** Ile-Pro, ADMA and MG, where *p-value < 0.05 and **p-value < 0.01. Distribution in HCC, HCV and MC patients of **E** lysophosphatydilcholines (LPCs), **F** ether-linked lysophosphatydilcholine (LPC-O) and **G** phosphatydilcholine (PC), where *p-value < 0.05 and **p-value < 0.001.

3.4.6 AFP Classification vs Multi-Omics PLS-DA Models

To benchmark the diagnostic performance of omics-based models against the conventional AFP approach, we compared classification accuracy across three strategies: AFP alone, Metabolomics + PLS-DA, and Lipidomics + PLS-DA.

The AFP dataset was partitioned into training and test sets using the same Kennard–Stone algorithm applied to metabolomics and lipidomics data. HCC samples with AFP values < 20 ng/mL were considered misclassified, reflecting the clinical challenge of AFP-negative tumors. Confusion matrices for AFP classification are shown in Supplementary **Figure S4**.

PLS-DA models were trained to discriminate against AFP-negative HCC patients from HCV and MC groups. Model performance was assessed via cross-validation and external testing, reporting sensitivity, specificity, and overall accuracy (**Table 3**). AFP alone achieved poor discrimination, with accuracy ≈ 50% (48.98% training; 55.00% test), confirming its limited utility in AFP-negative cases. Metabolomics + PLS-DA delivered near-perfect performance: 97.90% accuracy in training and 100.00% in test, underscoring the strong discriminative power of metabolic signatures. Lipidomics + PLS-DA also outperformed AFP, achieving > 94% accuracy in both training and test sets. VIP analysis identified the most influential variables driving PLS-DA separation, including acylcarnitines, LPC species, and amino-acid derivatives-consistent with the biological axes highlighted in pathway enrichment.

These results demonstrate that integrated omics approaches dramatically outperform AFP, particularly in AFP-negative HCC patients-a subgroup often missed by current surveillance protocols. The findings support the translational potential of metabolomic and lipidomic markers for non-invasive early detection and risk stratification, paving the way for targeted assays in clinical practice.

	PLS-DA							
	Lipidomics			**Metabolomics**			**AFP**	
LVs	Accuracy Training (%)	Accuracy CV (%)	Accuracy Test (%)	LVs	Accuracy Training (%)	Accuracy Test (%)	Accuracy Training (%)	Accuracy Test (%)
3	100		94.4	4	97.9	100	48.98	55

Table 3. PLS-DA and AFP accuracy for HCC sample classification. The abbreviations indicate respectively: CV-Cross Validation: LVs-Latent Variables.

3.4.7 Harnessing Omics PLS-DA Models for AFP-Negative HCC Classification

Confusion matrices (**Figure 4A–B**) demonstrate perfect classification accuracy (100.00%) for metabolomics, while lipidomics achieved 88.89% accuracy in AFP-negative HCC patients. These results underscore the superior discriminative power of omics-based models compared to AFP alone. To assess model robustness under dimensionality reduction, we compared diagnostic performance using: all VIP features (VIP > 1) and top 20 VIP-ranked metabolites and lipids only

Remarkably, metabolomics maintained an AUROC of 0.94 in both scenarios, while lipidomics achieved AUROC values of 0.89 (all VIPs) and 0.83 (top 20 VIPs) (**Figure 4C-D**). This indicates that even a minimal panel of highly discriminant molecules can deliver strong classification performance, paving the way for targeted assays in clinical practice.

These findings highlight a critical advantage: omics-driven models could accurately classify AFP-negative HCC cases, a subgroup frequently missed by current surveillance protocols. The ability to reduce the feature set without compromising performance suggests feasibility for cost-effective, high-throughput screening tools.

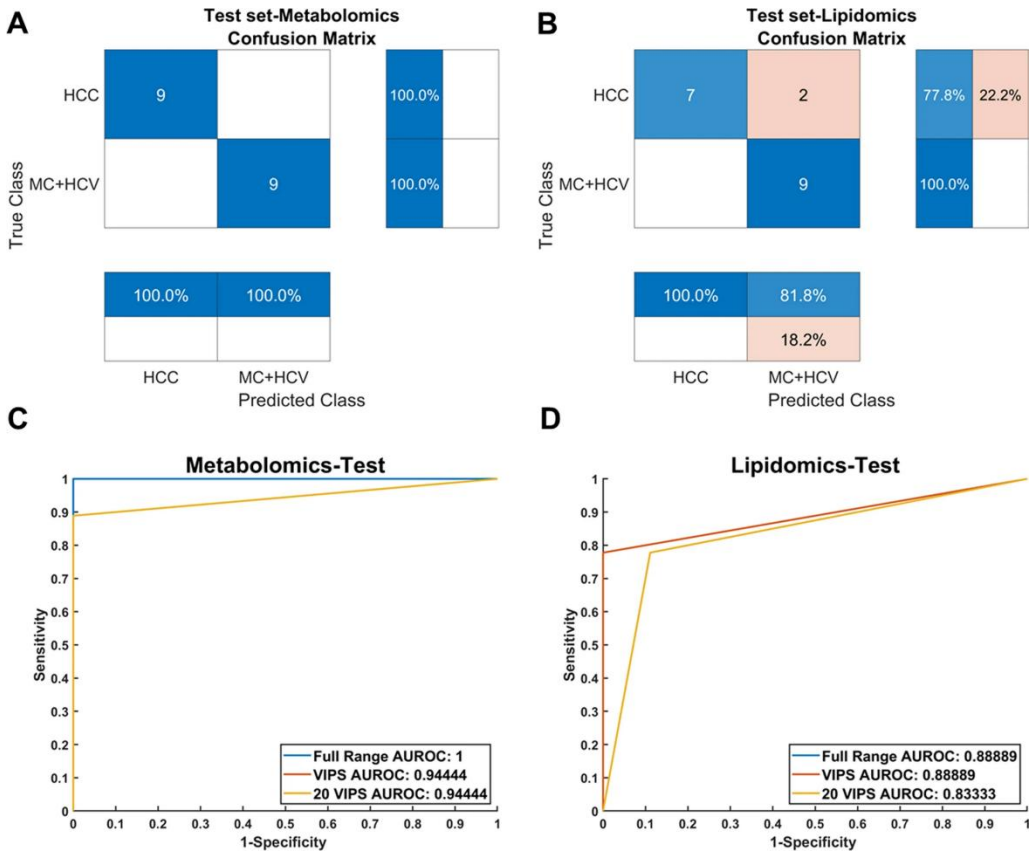


Figure 4 A-D. Graphical representation of confusion matrices obtained from PLS-DA models (metabolomics **A**, lipidomics **B**) of both independent modalities for the test phase classifying AFP-negative HCC patients. The ROC curves (metabolomics **C**, lipidomics **D**) reported compares the models' performances reducing the number of variables to the VIPs and the first 20 VIP.

3.5 Discussion

The identification of systematically altered metabolic targets is an imperative step toward exploiting metabolism in basic, translational, and clinical cancer studies. While genomic and epigenomic alterations have been associated with liver cancer (31) several shreds of evidence highlight that the tumour onset and progression are strongly characterized by metabolic reprogramming, such as central carbon metabolism, glycolysis, de novo lipogenesis, phosphatidylcholine synthesis (33). Targeting these mechanisms could represent both a therapeutic and diagnostic opportunity, and the combination of metabolomics and lipidomics holds great potential for the development of non-invasive diagnostics and tailored therapy for HCC patients (34). In this study, 102 HCV-positive patients were included, comprising 23 HCV, 69 HCC and 10 MC subjects. The combined metabolomic and lipidomic approach helped clarify the metabolic features clustering within the three different groups. It also allowed the recording of significant differences in the metabolic profile of HCC compared to the HCV and MC classes. Acylcarnitines play an important role in the transport of fatty acids into mitochondria during β-oxidation, and in cases of high energy demand, that cancer occur. The metabolic reprogramming observed in cancer is recognized for its role in regulating acylcarnitine levels across different chain lengths. It serves as a crucial mediator in cancer metabolic plasticity, intricately connecting essential pathways, factors, and metabolites to fulfil the energetic requirements of cancer cells (35). As the main organ responsible for endogenous carnitine synthesis and metabolism, the liver can experience notable changes in its acylcarnitine levels. Consequently, these variations may be closely linked to different stages of liver disease (36)**.** Accordingly, the alteration of acylcarnitines level has been associated with HCC. Short-chain (CAR 2:0, CAR 3:0, CAR 5:1) and long-chain acylcarnitines (CAR 14:1, 16:2, 18:1) levels were found increased in HCC with respect to both HCV and MC classes. Increased levels of long-chain acyl carnitines C14:1 and C18:1 have been reported in patients with NAFLD driven liver fibrosis and, additionally, even higher levels in the progression to HCC were found (59). On the contrary, a higher abundance of medium chain acylcarnitines (CAR 9:0, CAR 10:0, CAR 10:1) was observed in HCV class with respect to HCC and MC, in good accordance with previous findings (37). A similar trend has been observed in Chinese cohorts, also being able to discriminate across severe liver disease (CIR) and HCC (38,39). These results confirm the potential of acylcarnitines as potential biomarkers also for AFP

false-negative HCC patients (37), clearly a correlation of plasma and tissue levels would be further necessary to extend these findings. Within polar metabolites, three additional metabolites emerged as significant across the three classes, and globally elevated in HCC: ADMA, MG and the dipeptide Ile-Pro. ADMA is a byproduct of proteolysis of post-translationally methylated proteins, which has been found to increase in the plasma of cancer patients (40). Indeed, high plasma ADMA levels have been traced in patients with liver cirrhosis, alcoholic hepatitis and acute liver failure (41). MG is a modified purine derivative and has been associated with rapid turnover of nucleic acids which increases in the plasma and urine of cancer patients (42,43). Dipeptides have been recently investigated as potential markers of different cancers, including hepatocellular carcinoma, showing a different profile in tumour and non-tumour tissues (42,44). Alteration in lipid homeostasis represents an important hallmark for different diseases, especially cancer. Different lipid subclasses have been found altered in HCC. Notably, investigations into the lipidomic profiles of HCC have frequently noted a reduction in glycerophospholipids (GPLs) that incorporate polyunsaturated fatty acids (PUFAs), including arachidonic acid (C20:4), within human HCC specimens (45). Noteworthy, in our study LPCs containing PUFA were among the main lipid subclasses found reduced in the plasma of HCC patients. Recently the enzyme Lysophosphatidylcholine acyltransferase 1 (LPCAT1) has emerged as a novel diagnostic marker in HCC, being overexpressed in different cancers, including HCC (46,47). LPCAT1 catalyses the conversion of lysophosphatidylcholines into phosphatidylcholines. Notably, the lipidomic analysis of this study highlights a marked reduction in the overall LPCs profile, especially in the HCC group. This reduction is followed to a lesser extent by HCV and differs from MC. Interestingly, this reduction is independent of fatty acyl composition, indicating an up-regulation of the Land's cycle, resulting in further incorporation of fatty acids into PC. Our results are consistent with previous work carried out with an international cohort, illustrating a significant reduction in LPC in the serum of patients (48). On the contrary, the levels of different PCs have increased in both HCV and HCC groups with respect to MC. The latter confirms the potential alteration of the LPCAT activity, that induces the conversion of LPCs to PCs which is essential for tumorigenesis and promotes cancer cell growth and metastases (48,49). Certainly, further exploration is warranted given the extensive diversity of lipid species. While the prospect of identifying non-invasive circulating lipid biomarkers for HCC is exciting, the remodelling of tissue lipids in tumour and non-tumour liver tissues

could reveal pathways in HCC pathogenesis, whether related to viral factors or other mechanisms. Establishing correlations between plasma and tissue lipid changes, where possible, is crucial for gaining mechanistic insights into the regulation of lipid metabolism and homeostasis during the development and progression of liver cancer. Such insights may pave the way for the potential development of therapies targeting lipid pathways (50). Despite the growing interest in exploring lipidomic changes in liver cancer, attempts to target lipid metabolism in a therapeutic context have not yielded success. The dynamic nature of the lipidome, coupled with a lack of mechanistic insights into the specific role(s) of individual lipids in liver cancer development, poses a significant obstacle to the development of novel therapies. While our existing knowledge enables the utilization of the human lipidome as a non-invasive diagnostic and prognostic tool, it is imperative to delve. With respect to other studies that have been mainly focused on comparing tumor-free healthy controls vs HCC patients (13,51), the present study demonstrates how the developed model, based on the combined metabolome and lipidome signatures, shows high classification accuracy for distinguishing HCC vs other HCV related diseases. Furthermore, it was able to outperform AFP accuracy, especially for patients with values below 20 ng/mL (52), thus underlying its exploitability in the frame of diagnosis and prognosis of HCV-related landscape. Future studies with larger cohorts including different racial, ethnic, and geographical cohorts will be necessary for extending our current findings.

3.6 Conclusion

Our findings shed light on metabolomic and lipidomic alterations in the plasma of patients with HCC and chronic HCV infection, providing valuable insights for the identification of novel biomarkers. Currently, AFP measurement remains the most widely used screening test for HCC, yet its sensitivity and specificity are limited: early-stage tumors may not secrete AFP, and cirrhotic livers can produce elevated AFP levels, leading to false negatives and positives.

Understanding the molecular mechanisms underlying HCC offers opportunities not only for early diagnosis but also for predicting disease trajectories and selecting optimal therapies to maximize clinical benefit. The availability of new biomarkers addresses urgent clinical need and is pivotal for improving treatment effectiveness and patient survival.

This pilot study demonstrates the potential of a combined metabolomic–lipidomic signature to serve as the basis for targeted assays. Such panels could be implemented on widely available triple quadrupole MS platforms in clinical laboratories, enabling absolute quantitation and integration into routine workflows. These biomarkers may act as innovative risk factors for early HCC detection, supporting timely and effective interventions.

Bibliography Chapter III

[1]. Llovet JM, Kelley RK, Villanueva A, Singal AG, Pikarsky E, Roayaie S, et al. Hepatocellular carcinoma. Nat Rev Dis Primer. 2021 Jan 21;7(1):6.

[2]. Lauletta G, Russi S, Conteduca V, Sansonno L. Hepatitis C virus infection and mixed cryoglobulinemia. Clin Dev Immunol [Internet].

[3]. Westbrook RH, Dusheiko G. Natural history of hepatitis C. J Hepatol. 2014 Nov;61(1):S58–68.

[4]. Kulik L, El-Serag HB. Epidemiology and Management of Hepatocellular Carcinoma. Gastroenterology. 2019 Jan;156(2):477-491.e1.

[5]. Alvaro D, Gores GJ, Walicki J, Hassan C, Sapisochin G, Komuta M, et al. EASL-ILCA Clinical Practice Guidelines on the management of intrahepatic cholangiocarcinoma. J Hepatol. 2023 July;79(1):181–208.

[6]. Llovet JM, Pinyol R, Kelley RK, El-Khoueiry A, Reeves HL, Wang XW, et al. Molecular pathogenesis and systemic therapies for hepatocellular carcinoma. Nat Cancer. 2022 Apr 28;3(4):386–401.

[7]. Yang C, Zhang H, Zhang L, Zhu AX, Bernards R, Qin W, et al. Evolving therapeutic landscape of advanced hepatocellular carcinoma. Nat Rev Gastroenterol Hepatol. 2023 Apr;20(4):203–22.

[8]. Kim E, Viatour P. Hepatocellular carcinoma: old friends and new tricks. Exp Mol Med. 2020 Dec;52(12):1898–907.

[9]. Qu C, Wang Y, Wang P, Chen K, Wang M, Zeng H, et al. Detection of early-stage hepatocellular carcinoma in asymptomatic HBsAg-seropositive individuals by liquid biopsy. Proc Natl Acad Sci. 2019 Mar 26;116(13):6308–12.

[10]. Johnson P, Zhou Q, Dao DY, Lo YMD. Circulating biomarkers in the diagnosis and management of hepatocellular carcinoma. Nat Rev Gastroenterol Hepatol. 2022 Oct;19(10):670–81.

[11]. Okabe H, Satoh S, Kato T, Kitahara O, Yanagawa R, Yamaoka Y, et al. Genome-wide analysis of gene expression in human hepatocellular carcinomas using cDNA microarray: identification of genes involved in viral carcinogenesis and tumor progression. Cancer Res. 2001 Mar 1;61(5):2129–37.

[12]. Borel F, Konstantinova P, Jansen PLM. Diagnostic and therapeutic potential of miRNA signatures in patients with hepatocellular carcinoma. J Hepatol. 2012 June;56(6):1371–83.

[13]. Fages A, Duarte-Salles T, Stepien M, Ferrari P, Fedirko V, Pontoizeau C, et al. Metabolomic profiles of hepatocellular carcinoma in a European prospective cohort. BMC Med. 2015 Dec;13(1):242.

[14]. Di Poto C, Ferrarini A, Zhao Y, Varghese RS, Tu C, Zuo Y, et al. Metabolomic Characterization of Hepatocellular Carcinoma in Patients with Liver Cirrhosis for Biomarker Discovery. Cancer Epidemiol Biomarkers Prev. 2017 May 1;26(5):675–83.

[15]. Luo X, Cheng C, Tan Z, Li N, Tang M, Yang L, et al. Emerging roles of lipid metabolism in cancer metastasis. Mol Cancer. 2017 Dec;16(1):76.

[16]. Edmondson HA, Steiner PE. Primary carcinoma of the liver.A study of 100 cases among 48,900 necropsies. Cancer. 1954 May;7(3):462–503.

[17]. Collision cross section database and prediction, n.d. https://ccsbase.net/. Accessed 17 July 2023. Available from: https://ccsbase.net/

[18]. CompMS | MS-DIAL, n.d. http://prime.psc.riken.jp/compms/msdial/main.html. Accessed 17 July 2023. Available from: http://prime.psc.riken.jp/compms/msdial/main.html

[19]. Smilde AK, Westerhuis JA, De Jong S. A framework for sequential multiblock component methods. J Chemom. 2003 June;17(6):323–37.

[20]. Smilde AK, Næs T, Hovde LK. Multiblock data fusion in statistics and machine learning [Internet].

[21]. Kennard RW, Stone LA. Computer Aided Design of Experiments. Technometrics. 1969 Feb;11(1):137–48.

[22]. Geladi P, Kowalski BR. Partial least-squares regression: a tutorial. J Optoelectron Adv Mater. 1985;7.

[23]. Wold S, Sjöström M. SIMCA: a method for analyzing chemical data in terms of similarity and analogy, vol. 197. p. 243–82.

[24]. Vitale R, Cocchi M, Biancolillo A, Ruckebusch C, Marini F. Class modelling by Soft Independent Modelling of Class Analogy: why, when, how? A tutorial. Anal Chim Acta. 2023 Aug;1270:341304.

[25]. Carr BI, Akkiz H, Üsküdar O, Yalçın K, Guerra V, Kuran S. HCC with low- and normal-serum alpha-fetoprotein levels. Clin Pr [Internet].

[26]. Pezzuto F, Izzo F, Buonaguro L, Annunziata C, Tatangelo F, Botti G, et al. Tumor specific mutations in TERT promoter and CTNNB1 gene in hepatitis B and hepatitis C related hepatocellular carcinoma. Oncotarget. 2016 Aug 23;7(34):54253–62

[27]. Shang S, Plymoth A, Ge S, Feng Z, Rosen HR, Sangrajrang S, et al. Identification of Osteopontin as a Novel Marker for Early Hepatocellular Carcinoma. Hepatology. 2012 Feb;55(2):483–90.

[28]. Carbone D, Vestuto V, Ferraro MR, Ciaglia T, Pecoraro C, Sommella E, et al. Metabolomics-assisted discovery of a new anticancer GLS-1 inhibitor chemotype from a nortopsentin-inspired library: From phenotype screening to target identification. Eur J Med Chem. 2022 Apr;234:114233.

[29]. Merciai F, Musella S, Sommella E, Bertamino A, D'Ursi AM, Campiglia P. Development and application of a fast ultra-high performance liquid chromatography-trapped ion

mobility mass spectrometry method for untargeted lipidomics. J Chromatogr A. 2022 June;1673:463124.

[30]. lipidomicstandards.org, n.d. https://lipidomicstandards.org/. Accessed 17 July 2023. Available from: https://lipidomicstandards.org/

[31]. Sumner LW, Amberg A, Barrett D, Beale MH, Beger R, Daykin CA, et al. Proposed minimum reporting standards for chemical analysis: Chemical Analysis Working Group (CAWG) Metabolomics Standards Initiative (MSI). Metabolomics. 2007 Sept 19;3(3):211–21.

[32]. Köfeler HC, Eichmann TO, Ahrends R, Bowden JA, Danne-Rasche N, Dennis EA, et al. Quality control requirements for the correct annotation of lipidomics data. Nat Commun. 2021 Aug 6;12(1):4771.

[33]. Satriano L, Lewinska M, Rodrigues PM, Banales JM, Andersen JB. Metabolic rearrangements in primary liver cancers: cause and consequences. Nat Rev Gastroenterol Hepatol. 2019 Dec;16(12):748–66.

[34]. Wu X, Wang Z, Luo L, Shu D, Wang K. Metabolomics in hepatocellular carcinoma: From biomarker discovery to precision medicine. Front Med Technol. 2023 Jan 4;4:1065506.

[35]. Melone MAB, Valentino A, Margarucci S, Galderisi U, Giordano A, Peluso G. The carnitine system and cancer metabolic plasticity. Cell Death Dis. 2018 Feb 14;9(2):228.

[36]. Li S, Gao D, Jiang Y. Function, Detection and Alteration of Acylcarnitine Metabolism in Hepatocellular Carcinoma. Metabolites. 2019 Feb 21;9(2):36.

[37]. Enooku K, Nakagawa H, Fujiwara N, Kondo M, Minami T, Hoshida Y, et al. Altered serum acylcarnitine profile is associated with the status of nonalcoholic fatty liver disease (NAFLD) and NAFLD-related hepatocellular carcinoma. Sci Rep. 2019 July 23;9(1):10663.

[38]. Chen S, Kong H, Lu X, Li Y, Yin P, Zeng Z, et al. Pseudotargeted Metabolomics Method and Its Application in Serum Biomarker Discovery for Hepatocellular Carcinoma Based on Ultra High-Performance Liquid Chromatography/Triple Quadrupole Mass Spectrometry. Anal Chem. 2013 Sept 3;85(17):8326–33.

[39]. Lin X, Yang F, Zhou L, Yin P, Kong H, Xing W, et al. A support vector machine-recursive feature elimination feature selection method based on artificial contrast variables and mutual information. J Chromatogr B. 2012 Dec;910:149–55

[40]. Chen YL, Lowery AkT, Lin S, Walker AM, Chen KHE. Tumor cell-derived asymmetric dimethylarginine regulates macrophage functions and polarization. Cancer Cell Int. 2022 Nov 15;22(1):351.

[41]. Ferrigno A. Liver plays a central role in asymmetric dimethylarginine-mediated organ injury. World J Gastroenterol. 2015;21(17):5131.

[42]. Harahap Y, Tanujaya AT, Nurahman F, Vianney AM, Purwanto DJ. Determination of O6-Methylguanine in dried blood spot of breast cancer patients after cyclophosphamide administration. Heliyon. 2021 July;7(7):e07558.

[43]. Seow WJ, Shu XO, Nicholson JK, Holmes E, Walker DI, Hu W, et al. Association of Untargeted Urinary Metabolomics and Lung Cancer Risk Among Never-Smoking Women in China. JAMA Netw Open. 2019 Sept 20;2(9):e1911970.

[44]. Ozawa H, Hirayama A, Shoji F, Maruyama M, Suzuki K, Yamanaka-Okumura H, et al. Comprehensive Dipeptide Analysis Revealed Cancer-Specific Profile in the Liver of Patients with Hepatocellular Carcinoma and Hepatitis. Metabolites. 2020 Nov 1;10(11):442.

[45]. Tan SLW, Israeli E, Ericksen RE, Chow PKH, Han W. The altered lipidome of hepatocellular carcinoma. Semin Cancer Biol. 2022 Nov;86:445–56.

[46]. Morita Y, Sakaguchi T, Ikegami K, Goto-Inoue N, Hayasaka T, Hang VT, et al. Lysophosphatidylcholine acyltransferase 1 altered phospholipid composition and regulated hepatoma progression. J Hepatol. 2013 Aug;59(2):292–9.

[47]. Shen L, Gu P, Qiu C, Ding W tao, Zhang L, Cao W yue, et al. Lysophosphatidylcholine acyltransferase 1 promotes epithelial-mesenchymal transition of hepatocellular carcinoma via the Wnt/β-catenin signaling pathway. Ann Hepatol. 2022 May;27(3):100680.

[48]. Lu H, George J, Eslam M, Villanueva A, Bolondi L, Reeves HL, et al. Discriminatory Changes in Circulating Metabolites as a Predictor of Hepatocellular Cancer in Patients with Metabolic (Dysfunction) Associated Fatty Liver Disease. Liver Cancer. 2023;12(1):19–31.

[49]. Aboagye EO, Bhujwalla ZM. Malignant transformation alters membrane choline phospholipid metabolism of human mammary epithelial cells. Cancer Res. 1999;59.

[50]. Paul B, Lewinska M, Andersen JB. Lipid alterations in chronic liver disease and liver cancer. JHEP Rep. 2022 June;4(6):100479.

[51]. Jee SH, Kim M, Kim M, Yoo HJ, Kim H, Jung KJ. Metabolomics profiles of hepatocellular carcinoma in a Korean prospective cohort: the Korean cancer prevention study-II.

[52]. Attallah AM, Omran MM, Attallah AA, Abdallah SO, Farid K, Darwish H. HCC-ART score, a simple, highly sensitive and specific test for early diagnosis of hepatocellular carcinoma: a large-scale, multicentre study. Br J Cancer [Internet]. 2013;109.

CHAPTER IV

Beyond α-Fetoprotein: Development of a Targeted Metabolomics Assay for Accurate Discrimination of HCC from other Liver Diseases

4.1 Abstract

Early and accurate diagnosis of hepatocellular carcinoma (HCC) remains a major clinical challenge, particularly in patients with chronic hepatitis C virus (HCV) infection, where standard biomarkers such as alpha-fetoprotein (AFP) often lack sensitivity and specificity. Numerous biomarkers have emerged from untargeted metabolomics and lipidomics approaches, nevertheless, most of these studies provide only relative amounts and do not perform quantitative analysis as further validation. To address this challenge, we developed and validated a targeted method to quantitatively measure panels of biomarkers previously identified as predictive of HCC in untargeted analyses. A high-throughput (6 min) HILIC-HRMS method, based on a multiplexing-HRMS strategy was developed and validated to measure a panel of lysophosphatidylcholines (LPCs) and carnitines (CARs) in plasma samples from 92 HCC patients. The method showed satisfactory performances in terms of sensitivity (LOD_{avg}: 0.0558 ng/mL, LOQ_{avg}: 0.667 ng/mL), accuracy (98.75%), linearity (R^2 = 0.999), and stability (intra-day CV: 0.39%, inter-day CV: 1.46%), low matrix effects (10%), and recovery (95.82%). Despite analyzing a substantial reduced number of variables with respect to the untargeted approach (24 vs 280), the targeted panel confirmed the diagnostic value of the previously dentified signature and remains superior to the AFP classification. Among the measured compounds, LPC 14:0 and LPC 18:2 showed the highest discriminative power (AUC= 0.978, AUC= 0.967). The developed targeted approach demonstrated a promising potential tool for HCC discrimination from other liver diseases, and, by obtaining quantitative values of circulating analytes, could serve as a liquid biopsy approach for early diagnostics and therapy evaluation.

4.1 Introduction

Hepatocellular carcinoma (HCC), comprising approximately 75-85% of all primary liver cancers, represents a global health problem with a progressively increasing incidence (1,2). The major risk factors for HCC development are hepatitis B (HBV) and C (HCV) virus infection, excessive alcohol consumption, and metabolic dysfunction-associated steatotic liver disease (MASLD) (3). In addition, HCC is one of the leading causes of cancer-related mortality worldwide, mainly due to late diagnosis and limited treatment options in advanced stages (4). Therefore, the development of clinically useful biomarkers for hepatocellular carcinoma (HCC) diagnosis remains a major unmet clinical need. Although α-fetoprotein (AFP) remains the most widely used serum biomarker in clinical practice, its limited sensitivity and specificity, particularly in early-stage disease, severely restrict its diagnostic utility (5). Consequently, there is an urgent demand for alternative or complementary biomarkers capable of improving HCC early detection and patient stratification. In recent years, plasma cell-free DNA (cfDNA) and circulating tumor DNA (ctDNA) have emerged as promising non-invasive liquid biopsy approaches for early diagnosis, minimal residual disease assessment, and recurrence monitoring in HCC (6,7). However, while these strategies offer valuable genomic information, they may not fully capture the downstream functional consequences of tumor-associated metabolic reprogramming. In this context, metabolomics has gained increasing attention as a powerful tool for the discovery of circulating biomarkers across multiple diseases, owing to its ability to reflect dynamic biochemical alterations associated with pathophysiological states (8). Several liquid chromatography–mass spectrometry (LC–MS) based lipidomics and metabolomics studies, including untargeted, targeted, or pseudotargeted approaches, have reported circulating molecular signatures associated with HCC (9–11). These investigations consistently highlighted a complex metabolic dysregulation involving bile acids, glycerophospholipids, and the acylcarnitine (CAR) pathways, with multivariate and machine learning–based models often outperforming AFP in terms of predictive accuracy. However, most of these studies remain semi-quantitative, relying on normalized peak areas rather than absolute metabolite concentrations. While highly valuable for discovery, relative measurements are not sufficient for clinical implementation, where robust quantification is needed to define precision thresholds.

Enable inter-study comparability and support longitudinal monitoring (12,13). Translation of metabolomics-derived biomarkers into clinically actionable assays, therefore, requires analytically validated targeted methods capable of providing accurate and reproducible quantitative data in complex biological matrices. Targeted methods are typically performed in selected/multiple reaction monitoring mode (SRM/MRM) on triple quadrupole MS-analyzers that provide high sensitivity even for analytes in femtomolar concentrations. Nevertheless, this approach requires optimization of precursor and fragments transitions as well as collision energies; furthermore, several metabolites and lipids share the same fragment ions, leading to potential interferences and longer method development, such as for lysophosphatidylcholines (LPCs) and CARs. In contrast, high resolution mass spectrometry approaches (HRMS) benefit from increased confidence at the precursor level due to high resolving power and mass accuracy, increasing selectivity and robustness. Additionally, the multiplexing (MPX) capabilities of quadrupole-orbitrap analyzers allow multiple fills of the instrument's C-trap with quadrupole filtered precursors, prior to an Orbitrap mass analyzer scan (and optionally fragmented by HCD), which boosts sensitivity and increases duty cycle without sacrificing resolution, a strategy that has been shown to be superior to SRM, especially in proteomics (14).

In the context of HCC, through a previous integrated untargeted metabolomics and lipidomics study, we identified a circulating molecular signature that accurately discriminated HCC patients from chronically HCV-infected subjects, with LPCs and CARs emerging among the most significant discriminatory metabolites (15). Building on these discovery-phase findings, the present study aimed to develop and validate a high-throughput targeted quantitative metabo-lipidomics method for the accurate and sensitive measurement of selected LPCs and CARs in human plasma (15). A tailored analytical workflow for this analyte panel was optimized and validated, comprising a hydrophilic interaction liquid chromatography coupled to high-resolution mass spectrometry (HILIC–HRMS), operated in MPX-targeted-selected ion monitoring mode (MPX-T-SIM), and the ability of the resulting quantitative panel to discriminate HCC from chronic HCV infection was evaluated. By moving from discovery-based relative profiling to analytically validated absolute quantification, this work aims to advance a clinically translatable liquid biopsy strategy for HCC detection.

4.3 Methods

4.3.1 Sample preparation

4.3.1.1 Solid Phase Extraction (SPE)

Solid-phase extraction was performed using Strata-X polymeric reversed-phase cartridges (33 µm, Phenomenex, Torrance, CA, USA), selected for its broad retention characteristics and reported suitability for mixed-polarity analytes. Plasma samples were pre-treated according to the following workflow: 30 µL of plasma were thawed on ice and 300 µL of ice-cold ACN containing a mix of deuterated standards (C16-Carnitine, C8-Carnitine, C5-Carnitine, LPC 17:0, LPC 18:1, C16 Carnitine-d_9, C8 Carnitine-d_9, C5 Carnitine-d_9, LPC 15:0-d_5, LPC 17:0-d_5, LPC 19:0-d_5), were added and vortexed for 10 minutes. Samples were then incubated at −30 °C for 30 minutes and centrifuged at 14,680 rpm for 10 minutes at 4 °C to induce protein precipitation. Supernatants were loaded onto SPE cartridges after conditioning with 500 µL MeOH and equilibration with 500 µL H_2O. To enable simultaneous recovery of carnitines and LPCs-two classes with distinct physicochemical properties-different washing and elution conditions were tested: MeOH 5% (v/v), ACN 5% (v/v), and ACN 2% (v/v) for washing, and MeOH + 0.1% HCOOH or ACN + 0.1% HCOOH for elution.

4.3.1.2 Protein Precipitation (PP)

Plasma samples from 92 HCV-positive subjects diagnosed with HCC (n = 69) and chronic HCV infection (n = 23) were extracted as follows: 30 µL of plasma were thawed on ice and 300 µL of ice-cold ACN, containing a mix of deuterated standards, were added and vortexed for 10 minutes. Subsequently, samples were incubated at -30 °C for 30 minutes and centrifuged at 14,680 rpm, for 10 min at 4 °C to induce protein precipitation. Samples were injected in randomized order and blank samples were injected regularly and used to assess carryover and exclude background signals. Dried samples were dissolved in 50 µL of ACN/H_2O (80/20 *v/v* %).

4.3.2 Instrumentation and LC setup

Metabolome analyses were performed on a Thermo Vanquish Flex UHPLC (2.1-mm i.d. setup) coupled online to hybrid quadrupole Orbitrap Exploris 120 mass spectrometer (Thermo Fisher Scientific, Bremen, Germany) equipped with a heated electrospray

ionization probe (HESI II). For targeted metabolomics analyses the separation was performed with an Acquity UPLC BEH Amide™ (100 × 2.1 mm × 1.7 µm, 130 Å) protected with a VanGuard BEH Amide™ precolumn (5.0 × 2.1 mm; 1.7 µm, 130 Å) (Waters, Milford, MA, USA). The mobile phases were composed of (A): H_2O/ACN (95:5 v/v%) and (B): H_2O/ACN (5:95 v/v%), both with 10 mM CH_3COONH_4. The following gradient was used: 0-1 min, 99%B, 7-8 min 30% B, 8.1-11.5 min 99% B. The column temperature was set at 45 °C and a flow rate of 0.4 mL/min was used. The injection volume was 2µL.

4.3.3 High Resolution Mass Spectrometry (HRMS) parameters

MS data acquisition for both setups was performed in Scheduled-Single-Ion-Monitoring with Multiplexing (SIM-MPX). MS1 scan OT resolution, 30,000; Normalized AGC target (100%); maximum injection time, auto. S-Lens RF level, 70; isolation window, 2 Da; SIM Window Mode: Center Mass. The HESI source parameters were as follows: sheath gas, 40 a.u.; auxiliary gas, 15 a.u.; sweep gas, 0 a.u. Spray voltages were set to 3.3 kV and 3.0 kV for ESI (+) and ESI (−) respectively. Ion transfer tube and vaporizer temperatures were set to 280 °C and 300 °C respectively. The instrument was externally calibrated daily with FlexMix solution (ThermoFisher) while at the beginning of every LC run the internal calibrant was injected (IC run start mode). The SIM-MPX conditions for the quantification of CARs and LPCs are described in detail in **Table 1.**

Compound	Formula	Center Mass (m/z)	z	MSX ID
CAR C16:0	$C_{23}H_{45}NO_4$	400.3421	1	1
CAR C16:0-d_9	$C_{23}H_{36}D_9NO_4$	409.3986	1	1
CAR C8:0	$C_{15}H_{29}NO_4$	288.2169	1	1
CAR C8:0-d_9	$C_{15}H_{20}D_9NO_4$	297.2734	1	1
CAR C5:0	$C_{12}H_{23}NO_4$	246.17	1	1
CAR C5:0-d_9	$C_{12}H_{14}D_9NO_4$	255.2265	1	1
LPC 17:0	$C_{25}H_{52}NO_7P$	510.3554	1	2
LPC 17:0-d_5	$C_{25}H_{47}D_5NO_7P$	515.3868	1	2
LPC 18:1	$C_{26}H_{52}NO_7P$	522.3554	1	2
LPC 19:0-d_5	$C_{27}H_{51}D_5NO_7P$	543.4181	1	1
CAR C9:0	$C_{16}H_{31}NO_4$	302.2326	1	1
CAR C10:1	$C_{17}H_{31}NO_4$	314.2326	1	1
CAR C2:0	$C_9H_{17}NO_4$	204.123	1	1
CAR C10:0	$C_{17}H_{33}NO_4$	316.2482	1	1
CAR C6:0	$C_{13}H_{26}NO_4$	261.1935	1	1
CAR C5:1	$C_{12}H_{21}NO_4$	244.1543	1	1
LPC 18:2	$C_{26}H_{50}NO_7P$	520.3398	1	2
LPC 20:3	$C_{28}H_{52}NO_7P$	546.3554	1	2
LPC 14:0	$C_{22}H_{46}NO_7P$	468.3085	1	2
LPC 18:3	$C_{26}H_{48}NO_7P$	518.3241	1	2
LPC 20:4	$C_{28}H_{50}NO_7P$	544.3398	1	2
LPC 16:0	$C_{24}H_{50}NO_7P$	496.3398	1	2
LPC 18:0	$C_{26}H_{54}NO_7P$	524.3711	1	2

Table 1. Overview of the SIM-MPX acquisition parameters, including monitored ions, m/z values, charge, and MPX ID, for the targeted quantification of carnitines (CARs) and lysophosphatidylcholines (LPCs).

4.3.4 Method validation

The method validation was performed by the assessment of system suitability and performance, linearity, sensitivity, precision, accuracy and matrix effect, according to ICH Guidelines (7).

4.3.4.1 System Suitability and Performance

The system suitability was tested by 5 consecutive injections using a standard mixture of the analytes and internal standards. System performance was assessed by injecting 1 blank (without the analyte and IS) and one lower limit of quantification (LLOQ). System carry-over was investigated by injecting 2 blank samples immediately after a calibration standard at the upper limit of quantification (ULOQ). Carry-over was considered acceptable when the analyte signal in the blanks did not exceed 20% of the mean analyte response at the LLOQ level, Eq. (1). The re-injection reproducibility of extracted samples was verified by reanalyzing an entire batch after storage in the autosampler at 4 °C. The limit of detection (LOD) and limit of quantification (LOQ) were determined using the calibration curve approach, based on the ratio between the standard deviation of the intercept (σ) and the slope of the regression line (S). Specifically, LOD Eq. (2) and LOQ Eq. (3) were calculated according to the following equations:

$$Carry\ Over\ (\%) = \frac{Peak\ area\ blank}{Peak\ area\ ULOQ} \times 100$$

(1)

$$LOD = 3.3\ \times \frac{\sigma}{S}$$

(2)

$$LOQ = 10\ \times \frac{\sigma}{S}$$

(3)

where σ represents the standard deviation of the y-intercept of the calibration curve, and S is the slope.

4.3.4.2 Calibration samples

Stock solutions of each analyte were prepared by accurately weighing the appropriate amount of compound and dissolving it in methanol (MeOH) to obtain a final concentration of 1 mg/mL. Working standard solutions were freshly prepared on the day of analysis by serial dilution of the stock solutions with acetonitrile/water (ACN/H_2O, 80/20, v/v). From the stock solutions, a set of calibration standards covering the concentration range from 10 µg/mL to 1 ng/mL was prepared. All stock solutions were stored at −20 °C and protected from light.

4.3.4.3 Linearity

The linearity was assessed using a line of twelve concentration levels. The calibration curves were generated by plotting the peak area ratio of the analytes at each level to the peak area of IS vs. its concentration. The calculated concentration was used to determine the content of each analyte in each individual calibration standard and considered acceptable if the mean accuracy (%RE) were within ±20% of the nominal concentrations for all calibration points.

4.3.4.4 Sensitivity

The LOD and LOQ values were calculated through serial dilutions according to ICH Guidelines. The calculations were performed using the previous formulas. The determination of both LOD and LOQ was carried out in triplicate. The standard curve was prepared on each day of the analyses, and the calculated equations were used to determine the concentrations in all samples within the analytical runs.

4.3.4.5 Precision and Accuracy

The precision was evaluated based on the calibration curve by calculating the relative standard deviation (%RSD) of the back-calculated concentrations of the calibration standards across replicate injections. Intra-day (repeatability) precision was assessed within the same analytical run, while inter-day (reproducibility) precision was determined by analyzing calibration curves prepared and analyzed on three different

days. Precision values were considered acceptable when %RSD did not exceed 15% for all concentrations, except for the LLOQ, where a deviation of up to 20% was tolerated, in accordance with international bioanalytical method validation guidelines (e.g., EMA, FDA).

4.3.4.6 Recovery

Recovery was determined by comparing the analyte peak areas obtained from extracted samples spiked before extraction with those from samples spiked post-extraction at equivalent nominal concentrations. The recovery (%) was calculated as the ratio between the mean peak area of the pre-extraction spiked samples and that of the post-extraction spiked samples, multiplied by 100 Eq. 4.

$$Recovery\ (\%) = \frac{Peak\ area\ pre - extraction}{Peak\ area\ post - extraction} \times 100$$

(4)

This evaluation was performed at three concentration levels (low, medium, and high) of spiked standards within the calibration range, in triplicate. Consistent recovery across concentrations and low variability (expressed as %RSD) indicated adequate extraction efficiency and method reproducibility. The internal standard recovery was also monitored to ensure reliable normalization and correction of potential matrix effects or sample loss during preparation.

4.3.4.7 Matrix effect

The matrix effect was evaluated by comparing the calibration curve slopes obtained from standard solutions prepared in pure solvent with those prepared by spiking blank matrix extracts (post-extraction) at equivalent concentrations.

The matrix effect (%) was calculated as reported in Eq. (5):

$$Matrix\ effect\ (\%) = \frac{Slope\ matrix}{Slope\ solvent} \times 100$$

(5)

Values lower than 100% indicate ion suppression, while values higher than 100% indicate ion enhancement.

4.3.5 Data Analysis

All statistical analyses and graphical representations were performed using MATLAB R2025b (MathWorks, Natick, MA, USA) with in-house scripts and custom chemometric functions. Concentration data were $\log_{10}$-transformed when appropriate for univariate visualization.

Receiver operating characteristic (ROC) curve analysis was subsequently performed to evaluate the univariate discriminative ability of individual metabolites. For each quantified compound, ROC curves were computed using the corresponding concentration values as continuous predictors and the class label (HCC vs HCV) as the binary response variable, considering HCC as the positive class. The ROC curve represents the relationship between the true positive rate (sensitivity) and the false positive rate (1 − specificity) obtained by varying the decision threshold across the entire range of observed concentrations.

The discriminative performance of each metabolite was quantified by the area under the ROC curve (AUC), which provides a threshold-independent measure of classification capability. An AUC value of 0.5 indicates no discriminative power, whereas values approaching 1.0 indicate increasing separation between classes. Before multivariate analysis, the data matrix was autoscaled (mean-centred and scaled to unit variance) (18). Unsupervised exploration was performed using principal component analysis (PCA) (19).

Score plots were used to evaluate clustering behavior and detect potential outliers. 95% confidence ellipses were calculated in the score space based on Hotelling's T^2 statistic. Loading plots were examined to identify metabolites contributing most strongly to the PCs, and the percentage of explained variance was calculated for each component. For supervised modelling, samples were divided into

training (80% of the data) and test subsets using the duplex algorithm (20) applied in PCA score space, accounting for 98% of the explained variance, while preserving class distribution.

Supervised classification was performed by partial least-squares discriminant analysis (PLS-DA) (21). Models were trained on the autoscaled training set, and the optimal number of latent variables (LVs) was selected by minimizing classification error using 5-fold cross-validation. Model performance was evaluated on the training, cross-validation, and independent test sets in terms of accuracy, sensitivity, and specificity derived from confusion matrices.

4.4 Results

4.4.1 Characteristics of the study population

The study population was predominantly male in both HCC and HCV groups, with HCC patients being significantly older. HCV infection emerged as the main cause of HCC, and most patients presented with single tumor nodules and moderate (G2) differentiation. Biochemical analysis showed increased ALT and GGT levels, while AST and AFP were not significantly elevated in HCC group compared with the group of patients with chronic HCV (15). Additionally, AFP levels did not correlate with tumor characteristics, highlighting its limited sensitivity and the need for more reliable biomarkers for early HCC detection.

4.4.2 Optimization of Sample Preparation for Targeted Quantification of Carnitines and LPCs

To maximize recovery of the analytes of interest, different plasma sample-preparation strategies were initially compared, including solid-phase extraction (SPE) and protein precipitation using water-miscible organic solvents. Polymeric reversed phase sorbent was tested, with different wash and elution conditions, SPE workflows exhibited class-dependent selectivity, with variable recoveries across carnitines and LPCs. SPE protocol and recovery values are reported in supplementary material (**S2-TabS2**). Based on these observations, two different protein precipitation solutions, such

as ice-cold ACN (100% *v/v*) and MeOH (80% *v/v*), were compared. The results showed that ACN-based precipitation outperformed the corresponding MeOH-based approach, achieving a mean recovery of 95.82% and 60.21%, respectively, confirming the suitability of ACN-based precipitation for the quantitative recovery of carnitines and LPCs. The improved analytical performance obtained with ACN precipitation was further supported by representative extracted ion chromatograms, reported in **Fig. 1**.

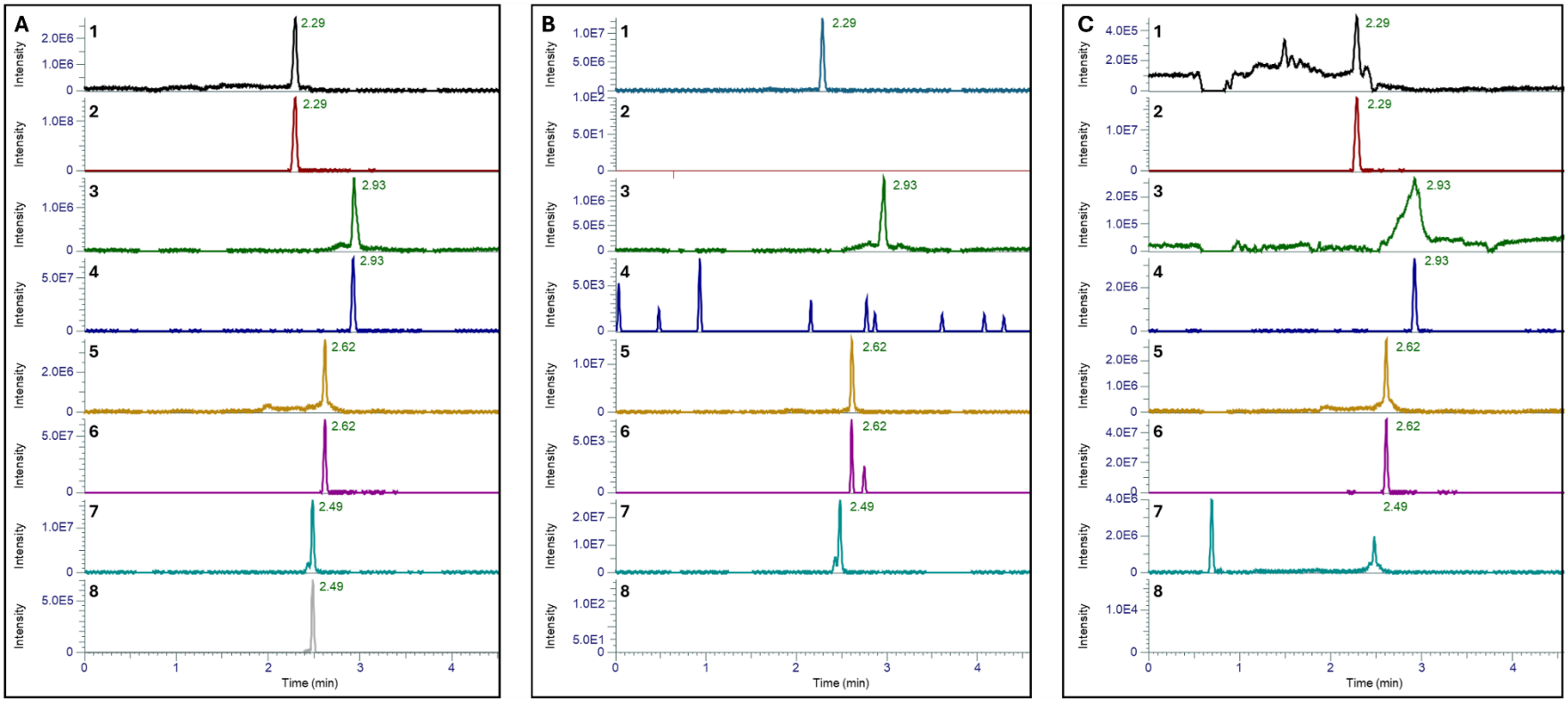


Fig. 1A-C. Extracted ion chromatograms (EICs) of some target metabolites in human plasma obtained using three different sample extraction strategies: (A) protein precipitation with ACN 100%, (B) protein precipitation with MeOH 80 % (v/v %), and (C) solid-phase extraction (SPE) eluted with MeOH. The comparison highlights differences in signal intensity and chromatographic profiles among the extraction approaches. 1: CAR C16:0, 2: CAR C16:0-d9, 3: CAR C5:0, 4: CAR C5:0-d9, 5: CAR C8:0, 6: CAR C8:0-d9, 7: LPC 17:0, 8: LPC 17.0- d5.

4.4.3 Optimization of Chromatographic Conditions for Targeted Quantification of Carnitines and LPCs

Since the previous untargeted approach was carried out using Hydrophilic Interaction Chromatography (HILIC) separation mode, the stationary phase chemistry was maintained while its geometry was modified using a shorter chromatographic column (150 vs 100 mm), to increase sample throughput. The gradient was accordingly scaled from 14 to 10 min. To maximize the ionization of the CARs and LPCs panel, multiple mobile phase additives based on volatile salts, namely 10 mM of ammonium formate (AmF) or 10 mM ammonium acetate (AmAc), with and without additional acidic modifiers (formic acid (HCOOH) or acetic acid (CH_3COOH), were evaluated. Each condition was assessed for its impact on signal-to-noise ratio (*S/N*) and peak symmetry. Among the tested conditions, the mobile phase composed of H_2O/ACN (95:5, *v/v %*) with 10 mM ammonium acetate as mobile phase A and H_2O/ACN (5:95, *v/v %*) with 10 mM ammonium acetate as mobile phase B, without additional acidic modifiers, provided the best overall performance, in terms of intensity and peak symmetry among all target analytes compared to other conditions. Signal intensity increased by 5.5-fold (450%), 3.5-fold (250%) and 2.67-fold (166.75%) when using AmAc, FA and AmAc + AcA, respectively, compared to AmF + FA **Fig. 2A**

Further optimization focused on sample resuspension solvent composition, which was adjusted using a higher proportion of organic solvent, specifically from ACN 70 % (*v/v%*) to ACN 80 % (*v/v%*), **Fig. 2B**. This optimization prevented peak splitting of some carnitines and enhanced analyte focusing on the column head.

Significantly, the optimized HILIC method allowed the chromatographic separation of LPCs regioisomers **Fig. 3A**. Different retention times were recorded for LPC positional isomers that varied in the acyl chain position (*sn*-1/ *sn*-2), and this was further supported by their MS/MS spectra **Fig. 3B.** Extracted ion chromatograms

showed near-baseline resolution, and the relative abundance of diagnostic fragment ions supported the structural assignment of each isomer.

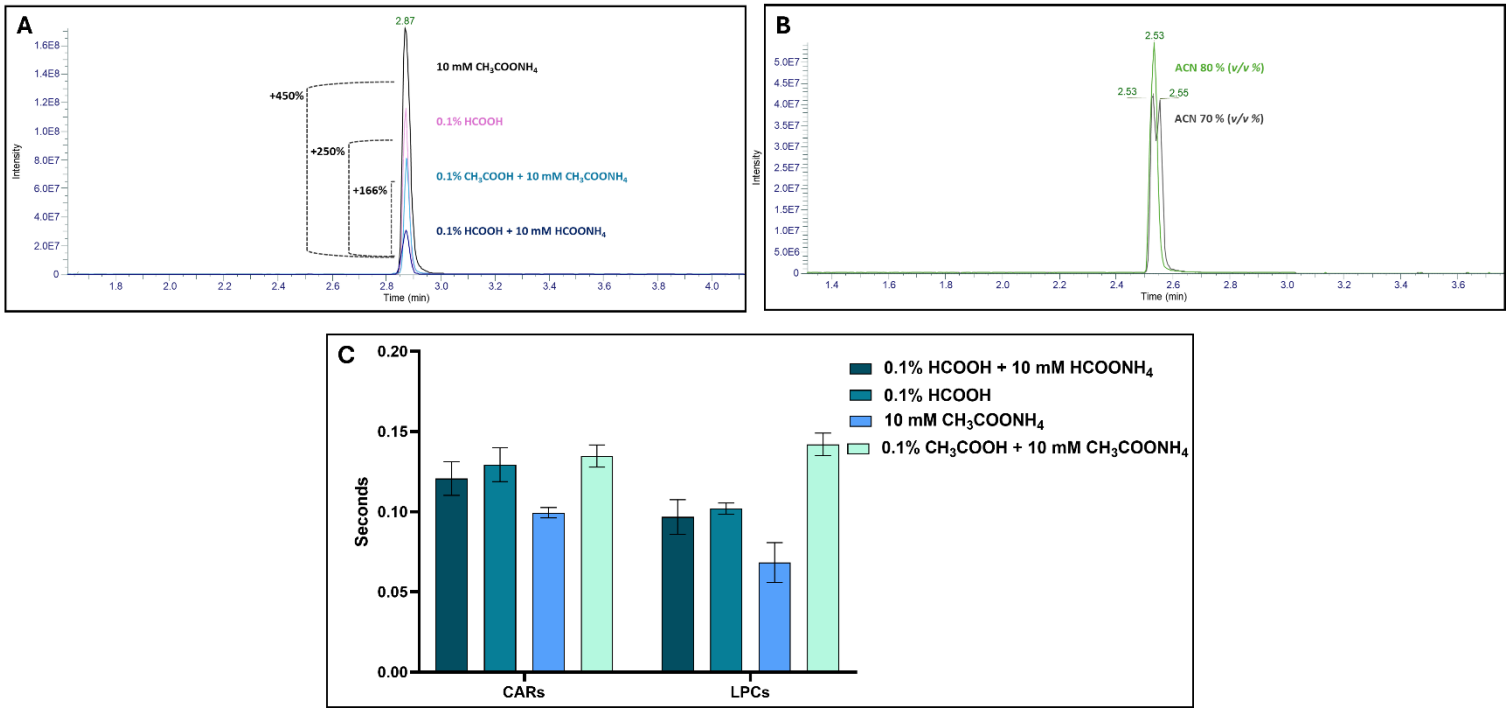


Fig. 2A-C A. Extracted ion chromatograms (EICs) intensities of CAR C8 under different mobile phase additives: 10 mM ammonium acetate (AmAc), formic acid (FA), AmAc + acetic acid (AcA), and ammonium formate + FA (AmF + FA). Signal intensity increased by 5.5-fold (450%), 3.5-fold (250%), and 2.67-fold (166.75%) for AmAc, FA, and AmAc + AcA, respectively, relative to AmF + FA. **B.** Extracted ion chromatograms (EICs) of CAR C16:0 under optimized solvent conditions, showing improved peak shape and signal intensity at ACN 80 % (v/v %) (green trace) compared to ACN 80 % (v/v %) (grey trace). **C.** FWHM values of some target metabolites detected in human plasma samples across different mobile phase modifiers. AmF: Ammonium Formate, FA: Formic Acid, AmAC: Ammonium Acetate, AcA: Ammonium Acetate.

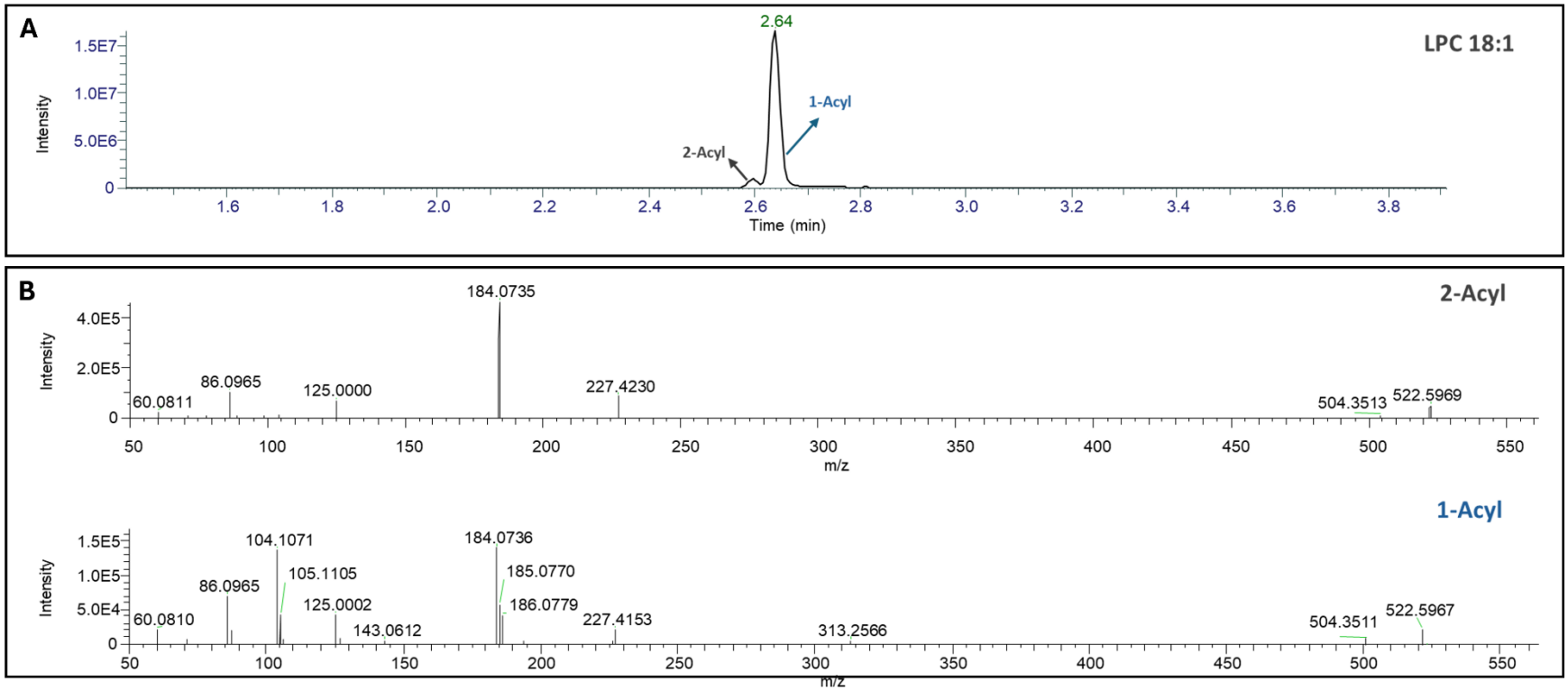


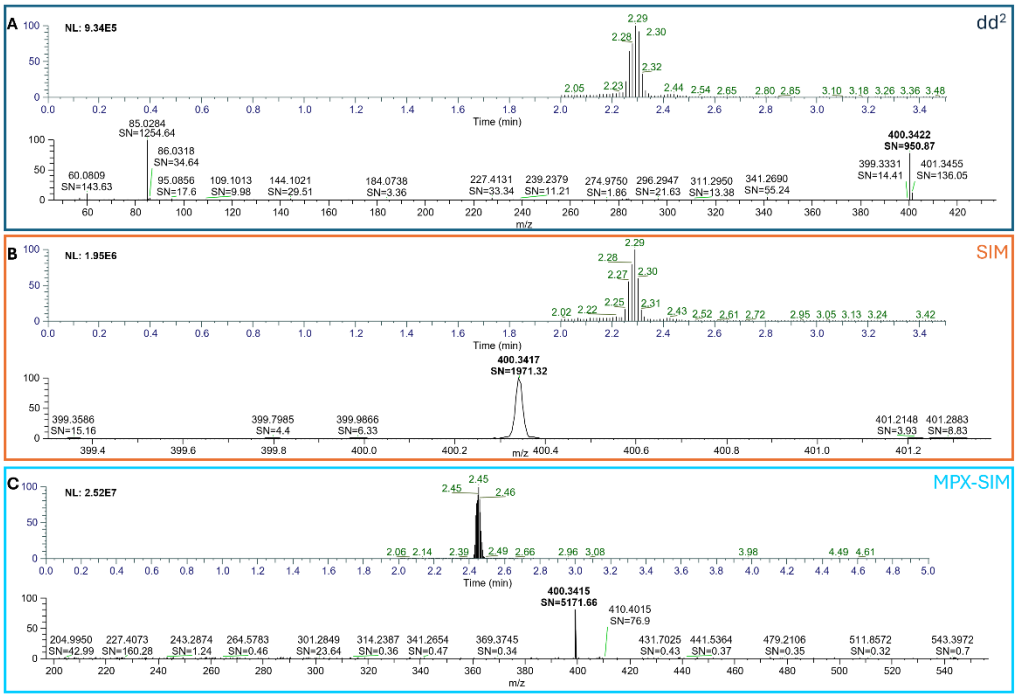
Fig 3A-B. A. Extracted ion chromatogram (EIC) of LPC 17:0 under optimized HILIC conditions, showing the separation of LPC regioisomer (*sn*-1/ *sn*-2). **B.** Representative MS/MS spectra displaying diagnostic fragment ions and relative intensity used for regioisomer discrimination.

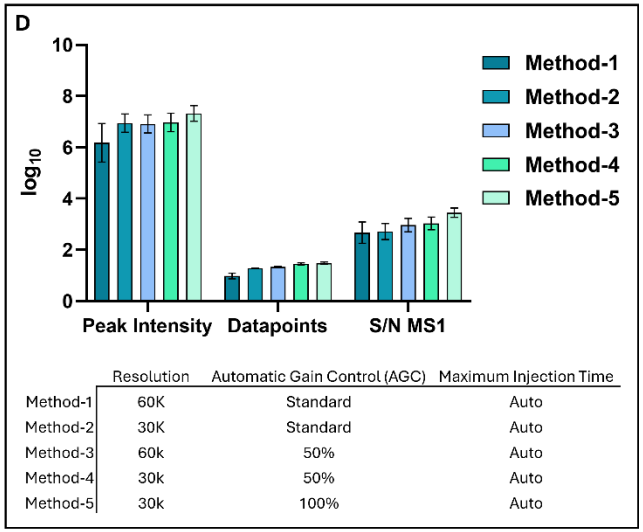
4.4.4 Selection of HRMS acquisition mode for targeted quantification of carnitines and LPCs

Different High-Resolution Mass Spectrometry (HRMS) acquisition strategies, including Full Scan-Data-Dependent Acquisition (FS-DDA), Selected Ion Monitoring (SIM), and Multiplexed Selected Ion Monitoring (MPX-SIM), were systematically assessed for the targeted assay of carnitines and LPCs in human plasma. Among these, the MPX-SIM method demonstrated superior analytical performance in terms of chromatographic sampling, sensitivity, and quantitative robustness. Notably, multiplexed SIM-MPX provided at least 10 data points per chromatographic peak for all monitored analytes, thereby meeting established criteria for robust targeted quantification and outperforming both FS-DDA and SIM acquisition modes, which resulted in reduced sampling frequency and compromised quantification for low-abundance analytes **Fig. 4A-B**. Comparative evaluation of *S/N* and peak areas across acquisition modes demonstrated that multiplexed SIM yielded superior analytical signals for all analytes, with marked improvements in both sensitivity and signal reproducibility **Fig. 4C**.

Additionally, targeted MS/MS experiments were performed to confirm their structural identities by monitoring class-specific diagnostic fragments. LPCs species were confirmed by the characteristic phosphocholine ion m/z 184.0733 ($C_5H_{15}NO_4P^+$), while carnitines were validated by the fragment at 85 m/z ($C_4H_5O_2^+$), which is the most significant feature peak of CAR spectra, **Fig.S1 A-B**.

In parallel, different Orbitrap MS1 resolution settings (30,000 and 60,000) were evaluated to balance mass accuracy with scan speed, ultimately selecting the resolution that provided optimal chromatographic sampling. Furthermore, various Automatic Gain Control (AGC) values were assessed (automatic, 50%, and 100%) to ensure ion accumulation of low-abundance species without extending the scan cycle time **Fig. 4D** Tab. S2. Consequently, multiplexed SIM was selected as the acquisition strategy for subsequent targeted analyses.





	Resolution	Automatic Gain Control (AGC)	Maximum Injection Time
Method-1	60K	Standard	Auto
Method-2	30K	Standard	Auto
Method-3	60k	50%	Auto
Method-4	30k	50%	Auto
Method-5	30k	100%	Auto

Fig 4A-D Comparative assessment of acquisition strategies for trace-level metabolite detection. S/N ratios and peak areas measured under different acquisition modes (A) DDA (B) SIM (C) MPX-SIM. Representative chromatographic profiles demonstrating enhanced sensitivity and reproducibility with multiplexed SIM. Evaluation of Orbitrap MS1 acquisition parameters for optimized chromatographic sampling in SIM-MPX.

4.4.5 Performance Evaluation of the Optimized SIM-MPX-Based LC–HRMS Method

The final targeted HILIC–HRMS method was validated according to ICH guidelines, assessing key performance parameters including sensitivity, accuracy, intra- and inter-day precision, matrix effects (ME), carry-over, and extraction recovery on the mixture of deuterated internal standards **Table 2**. Validation results confirmed that the optimized workflow met all acceptance criteria, with low variability and minimal matrix interference. In this context, the CARs and LPCs elution window was separated from phosphatidylcholines (PCs), thus avoiding ion suppression **Fig. S2**. The combination of tSIM acquisition, high mass resolution, and optimized MS parameters provided excellent sensitivity and reproducibility, establishing this approach as a robust solution for high-throughput targeted quantification of carnitines and LPCs in complex biological matrices.

		LOD	LOQ					Interday		Intraday	
Metabolite	R^2	ng/mL		Accuracy	Recovery	ME (%)	Carry Over	CV% Areas	CV% RT	CV% Areas	CV% RT
CAR C16:0-d_9	0.9965	3.873	11.738	98.551%	91.437%	5%	0.00%	5.892%	0.410%	0.137%	0.000%
CAR C8:0-d_9	0.9945	0.498	1.509	98.778%	106.621%	16%	0.00%	0.251%	0.369%	0.135%	7.098%
CAR C5:0-d_9	0.9993	0.089	0.270	98.108%	85.340%	13%	0.00%	15.492%	0.664%	0.126%	0.234%
LPC 17:0-d_5	0.9998	0.122	0.369	98.559%	96.393%	20%	0.00%	15.534%	0.445%	3.441%	0.000%
LPC 18:1-d_5	0.9987	0.185	0.560	98.658%	94.786%	20%	0.00%	1.013%	0.591%	0.067%	0.000%

Table 2. Analytical validation of the developed HILIC-SIM-MPX–HRMS method, reporting linearity, sensitivity (LOD and LOQ), intra- and inter-day precision, accuracy, repeatability, recovery, matrix effect and carry-over. LOD: Limit of Detection, LOQ: Limit of Quantification, ME: Matrix Effect, CV: Coefficient of Variation, RT: Retention Times.

4.4.7 Differential Metabolite and Lipid Profiles in HCC

PCA was applied to the autoscaled concentration data (**Fig 4 A-B**). PC1 explained 51.64% of the total variance, and PC2 accounted for 23.85%, retaining a cumulative variance of 75.49% for the first two components. The scores plot **(Fig 5A)** shows a clear separation between HCC and HCV samples, with minimal overlap. HCC samples were predominantly distributed along the negative side of PC1, whereas HCV samples were mainly located on the positive side, indicating that PC1 represents the main axis of discrimination. Intra-group variability was limited, particularly within the HCC cohort. The corresponding loadings plot **(Fig 5B)** highlights the metabolites contributing to this separation. LPCs are predominantly associated with the positive direction of PC1 and are characterized by lower p-values, indicating stronger statistical significance. In contrast, several carnitine species contribute to the opposite direction along PC1. Metabolites positioned at the extremes of the loading vectors had the smallest p-values, confirming their major influence on class discrimination and supporting the differential lipidomic signature observed between HCC and HCV.

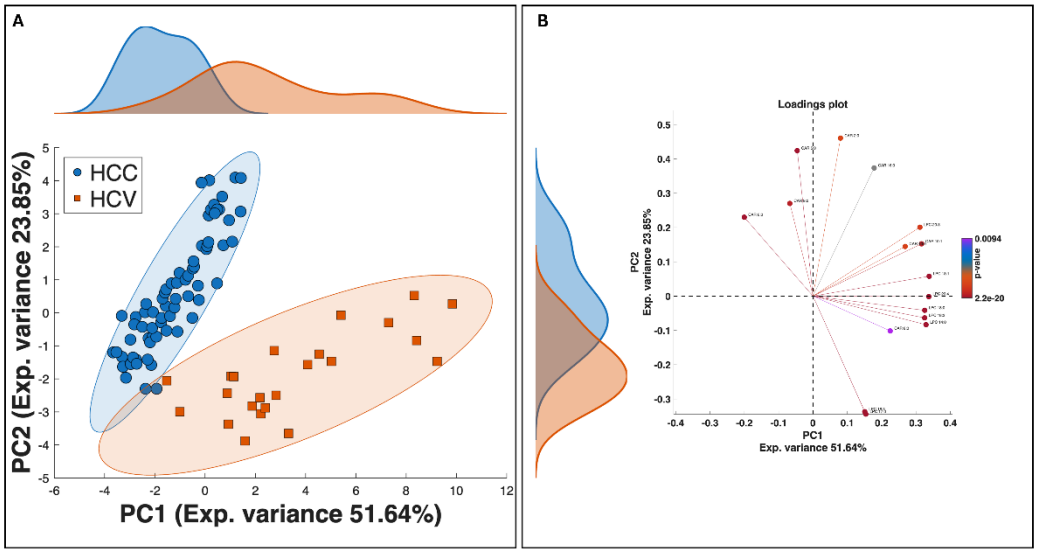


Fig 5 A-B A-PCA scores plot (PC1 vs PC2) obtained after autoscaling of the concentration data. 95% confidence ellipses were calculated for each class based on Hotelling's T^2 statistic in the PCA score space. **B-**The loadings plot was derived from the PCA loading matrix and represents the contribution of each metabolite to the first two principal components. HCC: hepatocellular carcinoma; HCV: hepatitis C virus infection.

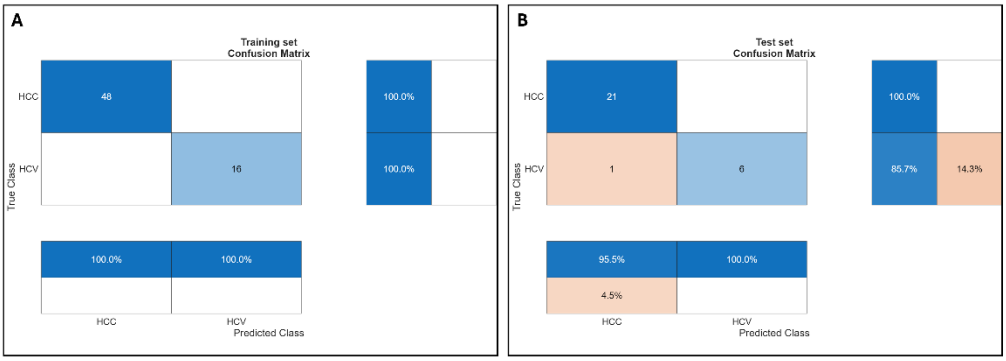


Fig 6 A-B Supervised classification was performed using partial least-squares discriminant analysis (PLS-DA) coupled with linear discriminant analysis (LDA). The number of latent variables (nLVs = 2) was selected by five-fold cross-validation. Confusion matrices were generated from predicted versus true class memberships for both the A - training and B - independent test sets and are displayed with row- and column-normalized percentages.

Supervised classification was subsequently performed using PLS-DA (**Fig. 6A-B**). The optimal model achieved 100.00% accuracy, sensitivity, and specificity in both the training and cross-validation phases. When applied to the independent test set, the model maintained high predictive performance, with an overall accuracy of 96.43%. Sensitivity for HCC was 100%, whereas sensitivity for HCV was 85.71%, due to a single misclassified HCV sample. Specificity values were 85.71% for HCC and 100.00% for HCV. Confusion matrices were generated from predicted versus true class assignments and are displayed with row- and column-normalized percentages (**Fig. 6**). The performance of the presented model is consistent with the previously reported omics-based PLS-DA results and remains superior to the AFP classification (15). In the earlier analysis, AFP classification showed limited discriminative power, with an accuracy of 48.98% in the training set and 55.00% in the test set (AFP < 20 ng/mL considered misclassified), confirming its modest effectiveness as a standalone marker for HCC identification, indicating limited discriminative capability with respect to the developed targeted assay.

Quantitative analysis of the analyte panel revealed clear stratification between clinical groups **(Tab S2, Fig 7**). Scatter plot representations of $\log_{10}$-transformed concentrations **(Fig. S3)** showed statistically significant differences for multiple LPCs and CARs species between HCC and HCV groups. Statistical comparisons were performed using the Mann-Whitney U test followed by FDR correction (*$p < 0.05$; **$p < 0.001$). LPCs levels were globally reduced in HCC vs HCV, while CARs showed acyl-chain length differences, and these results agree with previous untargeted analyses.

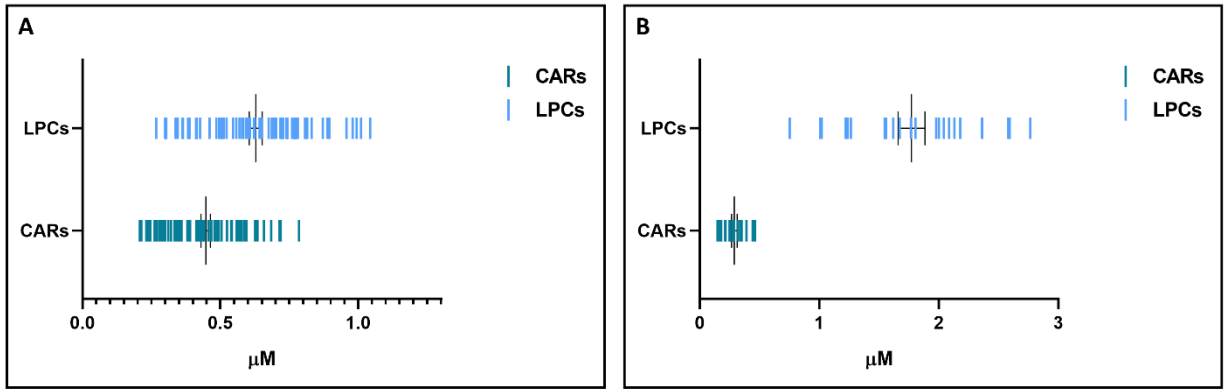


Fig. 7 A-B Aligned line plot of CARs and LPCs quantified in patients with HCC (A) and HCV (B) patients. Data are presented as mean ± SEM.

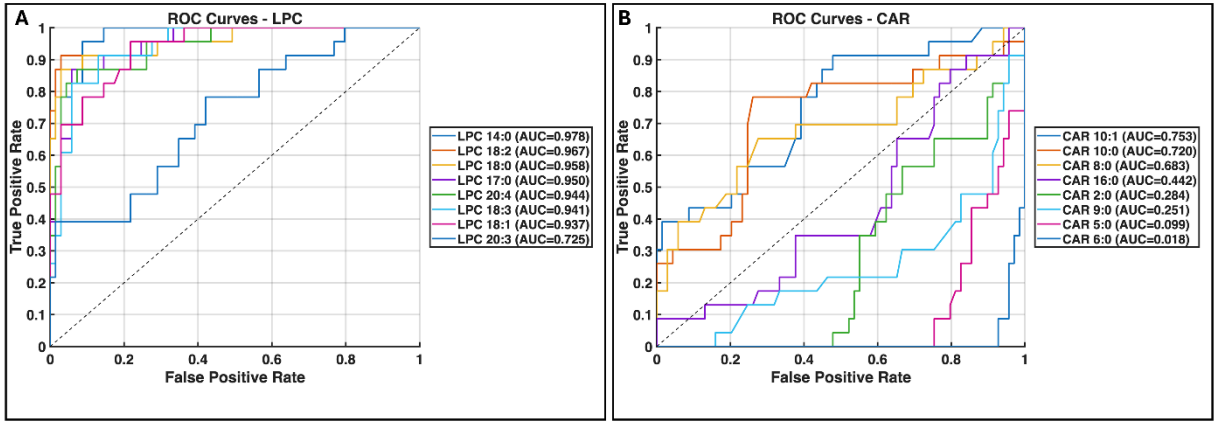


Fig 8A-B. Receiver operating characteristic (ROC) curves. Each curve represents the univariate classification performance of a single LPC (**A**) and CAR (**B**) metabolite in discriminating HCC from HCV samples. The x-axis indicates the false positive rate, and the y-axis the true positive rate across varying thresholds. The corresponding area under the curve (AUC) values are reported in the legend, highlighting the strong discriminative capability of most LPC species.

The discriminatory contribution of the monitored panel was further evaluated by ROC curve analysis **(Fig. 8 A-B).** The ROC curves obtained for LPC species (Fig. 8A) demonstrate a consistently strong classification performance between HCC and HCV samples. Most LPC variables exhibited high area under the curve (AUC) values, confirming their strong diagnostic potential. LPC 14:0 showed the highest discriminative power (AUC= 0.978), followed by LPC 18:2 (AUC= 0.967), LPC 18:0 (AUC= 0.958), and LPC 17:0 (AUC= 0.950). Additional species such as LPC 20:4 (AUC= 0.944), LPC 18:3 (AUC = 0.941), and LPC 18:1 (AUC= 0.937) also demonstrated excellent classification capability. Only LPC 20:3 displayed a comparatively lower performance (AUC= 0.725), although still indicating moderate discriminative ability. The consistently high AUC values across most LPCs corroborate and reinforce the key role of this lipid subclass in differentiating HCC from HCV. In contrast, the ROC analysis of CARs (**Fig. 8B**) revealed generally lower and more heterogeneous classification performance. Among them, the medium-chain carnitines displayed highest discriminative capability, such as CAR 10:1 (AUC = 0.753), CAR 10:0 (AUC = 0.720), and CAR 8:0 (AUC = 0.683), while other short-chain carnitines showed low performances CAR 6:0 (AUC = 0.018), CAR 5:0 (AUC = 0.099), CAR 2:0 (AUC = 0.284), indicating an inverse relationship with the class labels and suggesting that these metabolites are enriched in the opposite group relative to the defined positive class.

Overall, the ROC analysis confirms that LPC species provide the most robust univariate discrimination between HCC and HCV, whereas carnitines contribute more modestly and, in some cases, show inverse class associations.

Metabolite	HCC	HCV
	Avg ± Dev.Std	
CAR 10:0	0.113±0.082	0.225±0.223
CAR 10:1	0.188±0.124	0.373±0.274
CAR 16:0	0.248±0.053	0.234±0.054
CAR 2:0	0.855±0.448	0.496±0.219
CAR 5:0	0.9±0.432	0.356±0.154
CAR 6:0	1.195±0.402	0.337±0.17
CAR 8:0	0.06±0.048	0.138±0.148
CAR 9:0	0.062±0.021	0.04±0.025
LPC 14:0	1.004±0.636	4.287±2.108
LPC 17:0	0.915±0.494	2.462±0.941
LPC 18:0	0.121±0.076	0.379±0.152
LPC 18:1	1.162±0.551	2.507±0.92
LPC 18:2	0.115±0.073	0.386±0.145
LPC 18:3	0.888±0.508	2.175±0.887
LPC 20:3	0.799±0.454	1.297±0.727
LPC 20:4	0.111±0.076	0.348±0.15

Table 3. Absolute concentration values for CAR (nM) and LPCs (µM) expressed median ± standard deviation (SD).

4.5 Discussion

This study aimed to translate, confirm, and validate the findings obtained by an untargeted metabo-lipidomic approach for the stratification of HCC patients and their discrimination from other chronic liver diseases. Quantitative values in metabolomic and lipidomic studies are pivotal for defining diagnostic thresholds and obtaining values that can be compared across different labs and studies, which is of utmost importance for clinical employment. In this context, targeted assays are usually fit for purpose; consequently, this study tailored both sample preparation and LC-MS parameters to the specific analyte panel to maximize analytical sensitivity and robustness.

In this context, the comparative evaluation of SPE and solvent-based precipitation revealed an inherent trade-off between matrix clean-up and the extent of chemical coverage, a phenomenon that is frequently observed in SPE-based methods. While SPE offers selective enrichment and reduced matrix complexity, its performance was class-dependent, which limited the simultaneous recovery of carnitines and LPCs. The findings reported here are consistent with the literature, which suggests that polymeric reversed-phase SPE may preferentially retain or lose analytes depending on polarity and functional group composition, particularly when a single protocol is applied to heterogeneous metabolite classes (22).

Conversely, protein precipitation with ice-cold ACN provided a pragmatic solution for the targeted quantification of carnitines and LPCs, enhancing also reproducibility, as previously outlined in the context of metabolomics workflows (23). Furthermore, coupled with HILIC downstream separation, ACN precipitation resulted in a substantial enhancement of chromatographic performance, as evidenced by the observation of more defined and symmetrical peaks in comparison to those achieved through MeOH-based precipitation. This observation is in accordance with the extant literature, which suggests that ACN precipitation leads to a reduction in residual protein content and matrix complexity. Polson et al. (24) demonstrated that ACN forms larger protein pellets and drastically reduces residual protein in the supernatant, resulting in cleaner extracts and improved peak quality for LC–MS/MS analysis (25) resulting in higher average recovery and low matrix effect. These results are in line with best-practice guidelines for

bioanalytical method validation (ICH M10), which emphasize accuracy, precision, and matrix effect evaluation for clinical translation.

On the chromatographic side, the selection of a BEH Amide HILIC stationary phase was driven by our previous untargeted analyses as well as different analogue studies (26,27). Mobile phase composition plays a critical role in HILIC separations, especially with respect to buffering capacity and ionic strength. While acidic mobile phases containing formic acid are frequently employed due to their compatibility with ESI-MS, several studies have demonstrated that volatile ammonium salts, such as ammonium acetate or ammonium formate, can significantly improve peak shape and retention reproducibility by stabilizing the water-rich layer on the stationary phase and modulating electrostatic interactions (28). In this study, buffered mobile phases containing ammonium acetate outperformed acidic systems, likely due to improved control of ion–stationary phase interactions and reduced peak asymmetry. Finally, the observed chromatographic peak splitting when using an ACN 70% resuspension solvent is consistent with known HILIC injection effects, where mismatches between sample solvent strength and initial mobile phase composition can lead to poor analyte focusing. Increasing the organic content to ACN 80% enhanced on-column focusing and eliminated peak distortion, in agreement with established HILIC best practices (29).

Targeted LC-HRMS assays for CARs and LPCs must cope with broad dynamic ranges, low abundance features, isobaric/isotopic complexity and ion suppression. In quadrupole-Orbitrap MS analyzers, selected ion monitoring (SIM) offers a practical sensitivity gain versus full scan (FS) by enriching specific *m/z* windows in the Orbitrap analyzer while retaining high mass accuracy-an approach well suited to complex matrices and large analyte panels. Multiplexing SIM windows (MPX-SIM) further increases throughput without the need for extensive transition selection typical of triple-quadrupole SRM workflows (14). Additionally, this approach offers a high number of data points per chromatographic peak, improving peak integration stability and lowering CV% (30). While SRM/MRM remains a gold standard on triple quadrupoles, the multiplexed-SIM approach offers SRM comparable sensitivity with the enhanced selectivity at high resolution, simplified method development, and strong quantitative performance in complex matrices. Compared to traditional SIM, multiplexed-SIM monitors multiple targets within a single acquisition

event, reducing cycle time while maintaining sensitivity and quantitative accuracy when properly tuned. In fact, a known challenge of trapping analyzers is represented by space-charge effects, which were mitigated by optimizing AGC targets and maximum injection time (31).

Importantly, for lipid classes susceptible to positional isomerism, such as LPCs, chromatographic separation serves as an essential complement to high-resolution/accurate mass (HR/AM) detection. Consistent with established behavior of lysophospholipids (32,33), MS/MS spectra demonstrated a higher relative abundance of fragment ions corresponding to cleavage of the fatty acyl chain from the *sn-2* position, thereby providing an additional structural descriptor for regioisomer assignment.

The circulating signature showed elevation of short- and long-chain CARs in HCC, which supports the hypothesis of mitochondrial β-oxidation dysregulation, a well-recognized hallmark of cancer metabolism. Previous studies have linked altered CAR profiles to impaired fatty acid oxidation and increased energy demand in rapidly proliferating tumor cells, reinforcing their potential as diagnostic biomarkers for HCC (34). Conversely, the enrichment of medium-chain CARs in HCV samples suggests a distinct metabolic phenotype associated with chronic viral infection rather than malignant transformation, highlighting the discriminatory power of CAR profiling in differentiating disease states.

The most significant finding is represented by the pronounced reduction in LPCs, which could reflect accelerated membrane remodeling and altered phospholipid turnover, processes essential for sustaining tumor growth and metastasis. This lipidomic signature aligns with previous reports implicating LPC depletion and LPCAT-mediated reacylation in HCC progression (34,35). Notably, LPCs represent the strongest discriminative analytes in the measured panel.

While a limitation of this study is represented by both a small sample size and the absence of healthy controls, the obtained results confirm and reinforce previous untargeted findings, and strongly evidence that this quantitative non-invasive signature outperforms conventional markers such as AFP, particularly in AFP-negative cases offering potential results to be compared with prospective studies and future validation cohorts.

4.6 Conclusions

In this study, we established and validated an integrated two-tiered metabo-lipidomics workflow that bridges untargeted discovery with targeted quantitative validation, enabling the identification of clinically actionable biomarkers for hepatocellular carcinoma (HCC). By combining comprehensive untargeted profiling with a rapid, analytically robust multiplexed SIM-based HILIC–HRMS assay, we successfully translated disease-associated molecular signatures into absolute quantitative measurements suitable for clinical interpretation.

The optimized targeted method achieved simultaneous quantification of acylcarnitines and lysophosphatidylcholines within a 6-minute run, delivering outstanding analytical performance in sensitivity, linearity, precision, accuracy, recovery, and minimal matrix effects. Leveraging multiplexed SIM on a quadrupole–Orbitrap platform ensured high sampling density, enhanced detection of low-abundance metabolites, and reproducible quantitation, while maintaining high throughput and simplified method development compared to conventional SRM workflows.

Critically, quantitative evaluation confirmed the diagnostic relevance of the signature identified in untargeted analyses. Distinct alterations in acylcarnitine and LPC profiles enabled robust discrimination between HCC and chronic HCV patients, reflecting biologically coherent metabolic reprogramming linked to mitochondrial dysfunction and membrane lipid remodeling during tumor progression. ROC analysis and confusion matrices demonstrated excellent sensitivity and specificity, underscoring the potential of this targeted panel as a non-invasive liquid biopsy tool for HCC stratification.

Overall, this work highlights the power of high-resolution targeted metabolomics combined with rigorous validation and biologically informed marker selection to deliver clinically translatable solutions. The proposed workflow is scalable to larger cohorts and longitudinal studies, paving the way for precision diagnostics and improved clinical monitoring of HCC, particularly in AFP-negative cases where conventional biomarkers fail.

Bibliography Chapter IV

1. De Martel C, Georges D, Bray F, Ferlay J, Clifford GM. Global burden of cancer attributable to infections in 2018: a worldwide incidence analysis. Lancet Glob Health. 2020 Feb;8(2):e180–90. doi:10.1016/S2214-109X(19)30488-7

2. Rumgay H, Ferlay J, De Martel C, Georges D, Ibrahim AS, Zheng R, et al. Global, regional and national burden of primary liver cancer by subtype. Eur J Cancer. 2022 Jan;161:108–18. doi:10.1016/j.ejca.2021.11.023

3. Singal AG, Kanwal F, Llovet JM. Global trends in hepatocellular carcinoma epidemiology: implications for screening, prevention and therapy. Nat Rev Clin Oncol. 2023 Dec;20(12):864–84. doi:10.1038/s41571-023-00825-3

4. Bray F, Laversanne M, Sung H, Ferlay J, Siegel RL, Soerjomataram I, et al. Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin. 2024 May;74(3):229–63. doi:10.3322/caac.21834

5. Johnson P, Zhou Q, Dao DY, Lo YMD. Circulating biomarkers in the diagnosis and management of hepatocellular carcinoma. Nat Rev Gastroenterol Hepatol. 2022 Oct;19(10):670–81. doi:10.1038/s41575-022-00620-y

6. Buonopane IR, Saldanha EF, De Menezes JSA, Da Conceição LD, Reis CMDP, Leite LF, et al. Circulating tumour DNA for a minimal residual disease assessment and recurrence risk in hepatocellular carcinoma: a systematic review and meta-analysis. Br J Cancer. 2025 Dec 16. doi:10.1038/s41416-025-03296-8

7. Diaz LA, Bardelli A. Liquid Biopsies: Genotyping Circulating Tumor DNA. J Clin Oncol. 2014 Feb 20;32(6):579–86. doi:10.1200/JCO.2012.45.2011

8. Cui L, Lu H, Lee YH. Challenges and emergent solutions for LC-MS/MS based untargeted metabolomics in diseases. Mass Spectrom Rev. 2018 Nov;37(6):772–92. doi:10.1002/mas.21562

9. Chen S, Kong H, Lu X, Li Y, Yin P, Zeng Z, et al. Pseudotargeted Metabolomics Method and Its Application in Serum Biomarker Discovery for Hepatocellular Carcinoma Based on Ultra High-Performance Liquid Chromatography/Triple Quadrupole Mass Spectrometry. Anal Chem. 2013 Sep 3;85(17):8326–33. doi:10.1021/ac4016787

10. Luo P, Yin P, Hua R, Tan Y, Li Z, Qiu G, et al. A Large-scale, multicenter serum metabolite biomarker identification study for the early detection of hepatocellular carcinoma. Hepatology. 2018 Feb;67(2):662–75. doi:10.1002/hep.29561

11. Zhou L, Liao Y, Yin P, Zeng Z, Li J, Lu X, et al. Metabolic profiling study of early and late recurrence of hepatocellular carcinoma based on liquid chromatography-mass spectrometry. J Chromatogr B. 2014 Sep;966:163–70. doi:10.1016/j.jchromb.2014.01.057

12. Roberts I, Wright Muelas M, Taylor JM, Davison AS, Winder CL, Goodacre R, et al. Quantitative LC–MS study of compounds found predictive of COVID-19 severity and outcome. Metabolomics. 2023 Oct 18;19(11):87. doi:10.1007/s11306-023-02048-0

13. Torta F, Hoffmann N, Burla B, Alecu I, Arita M, Bamba T, et al. Concordant inter-laboratory derived concentrations of ceramides in human plasma reference materials via authentic standards. Nat Commun. 2024 Oct 3;15(1):8562. doi:10.1038/s41467-024-52087-x

14. Gallien S, Duriez E, Crone C, Kellmann M, Moehring T, Domon B. Targeted Proteomic Quantification on Quadrupole-Orbitrap Mass Spectrometer. Mol Cell Proteomics. 2012 Dec;11(12):1709–23. doi:10.1074/mcp.O112.019802

15. Caponigro V, Tornesello AL, Merciai F, La Gioia D, Salviati E, Basilicata MG, et al. Integrated plasma metabolomics and lipidomics profiling highlights distinctive signature of hepatocellular carcinoma in HCV patients. J Transl Med. 2023 Dec 18;21(1):918. doi:10.1186/s12967-023-04801-4

16. Kirwan JA, Gika H, Beger RD, Bearden D, Dunn WB, Goodacre R, et al. Quality assurance and quality control reporting in untargeted metabolic phenotyping: mQACC recommendations for analytical quality management. Metabolomics. 2022 Aug 27;18(9):70. doi:10.1007/s11306-022-01926-3

17. Mosley JD, Schock TB, Beecher CW, Dunn WB, Kuligowski J, Lewis MR, et al. Establishing a framework for best practices for quality assurance and quality control in untargeted metabolomics. Metabolomics. 2024 Feb 12;20(2):20. doi:10.1007/s11306-023-02080-0

18. Van Den Berg RA, Hoefsloot HC, Westerhuis JA, Smilde AK, Van Der Werf MJ. Centering, scaling, and transformations: improving the biological information content of metabolomics data. BMC Genomics. 2006 Dec;7(1):142. doi:10.1186/1471-2164-7-142

19. Wold S, Esbensen K, Geladi P. Principal component analysis. Chemom Intell Lab Syst. 1987 Aug;2(1–3):37–52. doi:10.1016/0169-7439(87)80084-9

20. Snee RD. Validation of Regression Models: Methods and Examples. Technometrics. 1977 Nov;19(4):415–28. doi:10.1080/00401706.1977.10489581

21. Ståhle L, Wold S. Partial least squares analysis with cross-validation for the two-class problem: A Monte Carlo study. J Chemom. 1987 Jul;1(3):185–96. doi:10.1002/cem.1180010306

22. Want EJ, Masson P, Michopoulos F, Wilson ID, Theodoridis G, Plumb RS, et al. Global metabolic profiling of animal and human tissues via UPLC-MS. Nat Protoc. 2013 Jan;8(1):17–32. doi:10.1038/nprot.2012.135

23. The Human Serum Metabolome (HUSERMET) Consortium, Dunn WB, Broadhurst D, Begley P, Zelena E, Francis-McIntyre S, et al. Procedures for large-scale metabolic profiling of serum and plasma using gas chromatography and liquid chromatography coupled to mass spectrometry. Nat Protoc. 2011 Jul;6(7):1060–83. doi:10.1038/nprot.2011.335

24. Polson C, Sarkar P, Incledon B, Raguvaran V, Grant R. Optimization of protein precipitation based upon effectiveness of protein removal and ionization effect in liquid chromatography–tandem mass spectrometry. J Chromatogr B. 2003 Mar 5;785(2):263–75. doi:10.1016/S1570-0232(02)00914-5

25. Bruce SJ, Jonsson P, Antti H, Cloarec O, Trygg J, Marklund SL, et al. Evaluation of a protocol for metabolic profiling studies on human blood plasma by combined ultra-performance liquid chromatography/mass spectrometry: From extraction to data analysis. Anal Biochem. 2008 Jan;372(2):237–49. doi:10.1016/j.ab.2007.09.037

26. Buszewski B, Noga S. Hydrophilic interaction liquid chromatography (HILIC)—a powerful separation technique. Anal Bioanal Chem. 2012 Jan;402(1):231–47. doi:10.1007/s00216-011-5308-5

27. Hemström P, Irgum K. Hydrophilic interaction chromatography. J Sep Sci. 2006 Aug;29(12):1784–821. doi:10.1002/jssc.200600199

28. Alpert AJ. Hydrophilic-interaction chromatography for the separation of peptides, nucleic acids and other polar compounds. J Chromatogr A. 1990 Jan;499:177–96. doi:10.1016/S0021-9673(00)96972-3

29. Taylor MR, Kawakami J, McCalley DV. Managing sample introduction problems in hydrophilic interaction liquid chromatography. J Chromatogr A. 2023 Jul;1700:464006. doi:10.1016/j.chroma.2023.464006

30. Zeng W, Bateman KP. Quantitative LC–MS/MS. 1. Impact of Points across a Peak on the Accuracy and Precision of Peak Area Measurements. J Am Soc Mass Spectrom. 2023 Jun 7;34(6):1136–44. doi:10.1021/jasms.3c00077

31. Lu W, McBride MJ, Lee WD, Xing X, Xu X, Li X, et al. Selected Ion Monitoring for Orbitrap-Based Metabolomics. Metabolites. 2024 Mar 25;14(4):184. doi:10.3390/metabo14040184

32. Lísa M, Cífková E, Holčapek M. Lipidomic profiling of biological tissues using off-line two-dimensional high-performance liquid chromatography–mass spectrometry. J Chromatogr A. 2011 Aug;1218(31):5146–56. doi:10.1016/j.chroma.2011.05.081

33. Pulfer M, Murphy RC. Electrospray mass spectrometry of phospholipids. Mass Spectrom Rev. 2003 Sep;22(5):332–64. doi:10.1002/mas.10061

34. Li S, Gao D, Jiang Y. Function, Detection and Alteration of Acylcarnitine Metabolism in Hepatocellular Carcinoma. Metabolites. 2019 Feb 21;9(2):36. doi:10.3390/metabo9020036

35. Guo W, Tan HY, Wang N, Wang X, Feng Y. Deciphering hepatocellular carcinoma through metabolomics: from biomarker discovery to therapy evaluation. Cancer Manag Res. 2018 Apr;Volume 10:715–34. doi:10.2147/CMAR.S15683

CHAPTER V

Harnessing the Potential of 1.0 mm ID columns in UHPLC-HRMS based Untargeted Metabolomics

5.1 Abstract

Untargeted metabolomics workflows based on UHPLC-HRMS commonly employ narrow-bore columns with 2.1 mm inner diameter (i.d.). However, the broad dynamic range of metabolite concentrations and the frequent need to analyze low-volume samples pose significant challenges to these approaches. Reducing column diameter represents a promising strategy to address these limitations. In this study, we systematically evaluated the performance of a microbore 1.0 mm i.d. setup compared to the conventional 2.1 mm i.d. format for untargeted metabolomics. The 1.0 mm i.d. configuration was implemented on a micro-UHPLC system, while the 2.1 mm i.d. setup was operated on a standard UHPLC, both coupled to quadrupole-orbitrap HRMS.

Analyses on a panel of polar standard metabolites revealed a marked sensitivity improvement with the 1.0 mm i.d. column, achieving an average 3.8-fold increase compared to 2.1 mm i.d., along with lower limits of detection (LOD_{avg}: 1.48 ng/mL vs. 6.18 ng/mL) and quantification (LOQ_{avg}: 4.94 ng/mL vs. 20.60 ng/mL). The microbore method enabled detection and quantification of all metabolites at LLOQ, while demonstrating superior repeatability for retention times (CV%: 0.29% vs. 0.63%) and peak areas (CV%: 4.65% vs. 7.27%).

Applications to diverse biological matrices, including plasma (various volumes), dried blood spots (DBS), and colorectal cancer patient-derived organoids (PDOs), in both RP and

HILIC modes under full-scan data-dependent acquisition (FS-DDA) confirmed the benefits of the microbore approach. Specifically, the 1.0 mm i.d. setup yielded significant increases in MS1 and MS2 features and MS/MS spectral matches by 38.95%, 39.26%, and 18.23%, respectively.

Overall, these findings demonstrate that 1.0 mm i.d. columns in UHPLC-HRMS workflows can substantially enhance metabolome coverage for low-amount samples while preserving analytical throughput and robustness typical of 2.1 mm i.d. formats, with the added advantage of reduced solvent consumption. This strategy holds promises for improving sensitivity and coverage in untargeted metabolomics, particularly in contexts where sample availability is limited.

5.2 Introduction

The growing interest in metabolomics has fueled the rise of metabolic phenotyping (1), which involves the comprehensive analysis of metabolites in biological fluids. The integration with other omics disciplines, such as proteomics and genomics, is driving the transition toward a patient phenotype-centric model, commonly referred to as personalized or precision medicine. Unlike proteomics and genomics, metabolomics analysis cannot be performed using a single analytical platform due to the extreme chemical complexity and wide dynamic range of metabolites in biological fluids, cells, and tissues.

Traditionally, nuclear magnetic resonance (NMR) spectroscopy, gas chromatography-mass spectrometry (GC-MS), and liquid chromatography-mass spectrometry (LC-MS) are employed-often in combination-to identify and quantify metabolites, each technique offering specific strengths and limitations (2). Nevertheless, LC-MS has become the technology of choice for metabolomics analysis because of its flexibility, sensitivity, and the availability of multiple chromatographic modes such as reversed phase (RP), hydrophilic interaction chromatography (HILIC), normal phase (NP), and supercritical fluid chromatography (SFC), which are extremely useful for handling the chemical and structural diversity of metabolite classes (3). Furthermore, the use of columns packed with sub-2 µm particles under ultra-high-pressure conditions (UHPLC), combined with the accuracy, sensitivity, and fast acquisition times of modern mass analyzers, has significantly enhanced LC-MS capabilities (4).

Despite these advances, one of the key challenges in metabolomics remains managing the vast dynamic range of the metabolome, which can vary significantly depending on sample type and amount, often available in very limited volumes. Conventionally, narrow-bore columns (2.1 mm i.d.) operated under analytical flow conditions (300–600 µL/min) are preferred in LC-MS setups (5). A potential strategy to improve LC sensitivity involves reducing the internal diameter of the column, which minimizes chromatographic dilution and thereby increases sensitivity (6). In metabolomics, this reduction, combined with mass spectrometry, could theoretically enable the annotation and quantification of a higher number of metabolites. This aspect is particularly relevant when restricted sample amounts are available, such as in 3D cell models (e.g., spheroids and organoids, <10^4 cells) and dried blood spots (DBS, 5-10 µL), as well as in pharmacometabolomic and toxicological studies involving mouse models, where sampling is critical and can often cause pain and stress to the animals (7).

Additionally, reducing column diameter lowers solvent consumption and waste production, which is a key point for sustainability, especially in large-scale studies. Nano-flow LC-MS (nLC-MS) is the cornerstone in proteomics; recently, the renaissance of microflow has gained popularity, particularly using 1.0-mm i.d. columns, which offer very high throughput, robustness, and excellent identification rates (8,9). Conversely, 1.0-mm i.d. columns have been scarcely used in metabolomics analysis, even though several authors have reported benefits such as reduced sample requirements, lower solvent consumption, and sensitivity gain, since a concentration factor of 4.4 can be achieved compared to standard 2.1-mm formats when the same amount is injected on column (10,11). The limited use of 1.0-mm i.d. columns has been attributed to several factors: reduced loading capacity compared to 2.1-mm i.d. columns, lower packing efficiency, and the significant impact of extra-column band broadening in standard UHPLC systems.

In this study, we aimed to assess the performance of microbore columns in untargeted metabolomics and compare them with the conventional 2.1-mm i.d. format. The objective was to evaluate the use of 1.0-mm i.d. columns in UHPLC-HRMS-based metabolomics workflows across various sample types, including plasma, cells, organoids, and dried blood spots. Both microflow and analytical flow systems were employed, exploring RP-UHPLC and HILIC conditions in conjunction with HRMS. The results demonstrate the benefits of using

1.0-mm i.d. columns in untargeted analyses, offering a practical alternative to the standard 2.1-mm i.d. format.

5.3 Methods

5.3.1 Metabolome Extraction Protocols

5.3.1.1 Plasma Metabolome Extraction

Two aliquots of human plasma (2 and 20 µL; *n* = 3) were obtained from a pooled quality control (QC) sample prepared using different aliquots of plasma from healthy individuals. Samples were thawed on ice and extracted with 20 µL or 200 µL, respectively, of ice-cold MeOH/H_2O (80:20, *v/v*) containing a mixture of deuterated internal standards (**Table S1**). The samples were vortexed for 12 min and incubated at −20 °C for 30 min, followed by centrifugation at 19,275 rcf for 10 min at 4 °C. Supernatants were collected and evaporated to dryness using a SpeedVac concentrator (Savant, Thermo Scientific, Milan, Italy). Dried extracts were reconstituted in 50 µL of ACN/H_2O (70:30, *v/v*) for HILIC analysis or MeOH/H_2O (10:90, *v/v*) for RP analysis. The study protocol was approved by the local Ethics Committee (prot./SCCE no. 71262, May 2020), and all procedures complied with the Declaration of Helsinki.

5.3.1.2 Dried Blood Spot (DBS) Metabolome Extraction

Whole blood was collected from a healthy volunteer among the authors (*n* = 5 spots; 10.0 ± 0.5 µL) using HemaXis DB10 (DBS Systems SA, Gland, Switzerland) according to the manufacturer's instructions and dried at room temperature overnight. DBS samples were extracted with 300 µL of ice-cold MeOH/H_2O (80:20, *v/v*) containing a mixture of deuterated standards. Samples were vortexed for 12 min, incubated at −20 °C for 30 min, and centrifuged at 19,275 rcf for 10 min at 4 °C. Supernatants were collected and evaporated using a SpeedVac. Dried extracts were reconstituted as described for plasma samples.

5.3.1.3 Patient Derived Organoids (PDOs) Metabolome Extraction

Non-treated (n = 3) and chemotherapy-treated (n = 3) PDO samples were provided by the Functional Genomics of Cancer Unit of San Raffaele. Extraction followed an MTBE-based biphasic protocol. Samples were transferred from a 96-well plate to tubes and mixed with 225 µL of ice-cold MeOH containing deuterated standards, followed by addition of 750

µL of cold MTBE. The mixture was shaken in a thermomixer (Eppendorf, Milan, Italy) for 1 h at 4 °C. Subsequently, 188 µL of H_2O was added, vortexed for 20 s, and centrifuged at 19,275 rcf for 10 min at 4 °C to induce phase separation. The lower aqueous phase containing polar metabolites was collected and evaporated using a SpeedVac. Dried extracts were reconstituted as described for the previous matrices.

5.3.2 Instrumentation and LC setups

Metabolome analyses were performed using two configurations: a Thermo Vanquish Neo nano/micro UHPLC system (1.0-mm i.d. setup) and a Vanquish Flex UHPLC system (2.1-mm i.d. setup). Each LC system was coupled online to a hybrid quadrupole Orbitrap Exploris 120 mass spectrometer (Thermo Fisher Scientific, Bremen, Germany), equipped with a heated electrospray ionization (HESI II) source.

For RP analyses, separations were carried out using an Acquity UPLC HSS T3™ column (150 × 2.1 mm or 1.0 mm for the micro-setup; 1.8 µm, 100 Å), protected by a VanGuard HSS T3™ precolumn (5 × 2.1 mm; 1.8 µm, 100 Å) (Waters, Milford, MA, USA). For HILIC analyses, an Acquity UPLC BEH Amide™ column (100 × 2.1 mm or 1.0 mm for the micro-setup; 1.7 µm, 130 Å) was used, protected by a VanGuard BEH Amide™ precolumn (5 × 2.1 mm; 1.7 µm, 130 Å) (Waters, Milford, MA, USA).

5.3.2.1 Microbore setup (1.0 mm ID)

For RP analyses, the column temperature was set at 55 °C, a flow rate of 100 µL/min was used, and mobile phases consisted of (A) H_2O + 0.1% HCOOH and (B) ACN + 0.1% HCOOH were used for positive ionization, while 0.1% CH_3COOH or 1 mM NH_4F was used for negative ionization mode. The following gradient has been used: 0 min, 0% B; 6 min, 70% B; 8 min, 80% B; 9 min, 98% B, 10 min 98% B; 10.1 min, 0% B; and 3.9 min for column re-equilibration.

For HILIC analyses, the column temperature was set at 55 °C, a flow rate of 90 µL/min was used, and mobile phases consisted of (A) H_2O + 0.1% HCOOH or 95/5 H_2O/ACN (v/v) + 10 mM CH_3COONH_4 and (B) ACN + 0.1% HCOOH or 95/5 ACN/H_2O (v/v) + 10 mM CH_3COONH_4 were used for positive and negative ionization. The following gradient has been used: 0–0.1 min, 99% B; 7–7.7 min, 30% B; 7.8 min, 99% B; and 3.4 min for column re-equilibration. All the connections were nanoViper of 50-µm i.d. as standard Vanquish neo

MicroLC configuration. An external oven was used and manually controlled (Phenomenex, Bologna, Italy).

5.3.2.2 Narrowbore setup (2.1 mm ID)

For RP mode, separation was carried out with a HSS T3 column (150 × 2.1 mm; 1.8 µm), protected with a Vanguard precolumn (5 × 2.1 mm; 1.7 µm) (Waters, Milan, Italy). The column temperature was set at 55 °C, and the flow rate was 500 µL/min. RP mobile phase was the same as reported for the microbore setup. The following gradient has been used: 0 min, 0% B; 1 min, 0% B; 1.5 min, 25% B; 4 min, 75% B; 6 min, 80% B; 6.1 min, 98% B; 7.1 min, 98%B; and 3.9 min for column re-equilibration.

For HILIC mode, separation was carried out with a BEH Amide column (100 × 2.1 mm; 1.7 µm) protected with a Vanguard precolumn (5 × 2.1 mm; 1.7 µm) (Waters, Milan, Italy). The column temperature was set at 45 °C, and the flow rate was 0.400 mL/min. HILIC mobile phase was the same as reported for the microbore setup. The following gradient was employed: 0–0.1 min, 99% B; 0.1–8 min, 99–50% B; 8.0–8.5 min, 50–30% B; 8.5–9.5 min isocratic at 30% B; returning to 99% in 0.1 min, and then 4 min to recondition the column.

5.3.2.3 High Resolution Mass Spectrometry (HRMS) parameters

MS data acquisition for both setups was performed in full scan-data dependent acquisition (FS-DDA) in the m/z range 70–800. MS1 scan OT resolution, 60,000; AGC, auto; maximum injection time, 100 ms. S-Lens RF level, 70; ddMS2 OT resolution, 15,000; isolation window, 1.5 Da; dynamic exclusion, 10 s; AGC, auto; maximum injection time, 22 ms. TopN, 4; HCD fragmentation normalized collision energies (NCE): 20, 40, 60. The HESI source parameters for 1.0-mm i.d. setup were as follows: sheath gas, 20 a.u.; auxiliary gas, 7 a.u.; sweep gas, 0 a.u. Spray voltages were set to 3.3 kV and 3.0 kV for ESI (+) and ESI (−) respectively. Ion transfer tube and vaporizer temperatures were set to 280 °C and 150 °C respectively. The instrument was externally calibrated daily with FlexMix solution (ThermoFisher) while at the beginning of every LC run the internal calibrant was injected (IC run start mode). For 2.1-mm i.d. setup, source parameters were as follows: sheath gas pressure, 40 a.u. and 50 a.u. for positive and negative ionization modes, respectively; aux gas flow, 15 a.u.; sweep gas flow, 0 a.u. Spray voltages were set to 3.3 kV and 3.0 kV for ESI (+) and ESI (−). Ion transfer tube (ITT) and vaporizer temperatures were set to 300 °C and

320 °C. The same MS and MS/MS acquisition parameters and calibration were used as reported above.

5.3.2.4 High Resolution Mass Spectrometry Analysis (HRMS)

MS data acquisition for both setups was performed in full scan-data dependent acquisition (FS-DDA) in the m/z range 70–800. MS1 scan OT resolution, 60,000; AGC, auto; maximum injection time, 100 ms. S-Lens RF level, 70; ddMS2 OT resolution, 15,000; isolation window, 1.5 Da; dynamic exclusion, 10 s; AGC, auto; maximum injection time, 22 ms. TopN, 4; HCD fragmentation normalized collision energies (NCE): 20, 40, 60. The HESI source parameters for 1.0-mm i.d. setup were as follows: sheath gas, 20 a.u.; auxiliary gas, 7 a.u.; sweep gas, 0 a.u. Spray voltages were set to 3.3 kV and 3.0 kV for ESI (+) and ESI (−) respectively. Ion transfer tube and vaporizer temperatures were set to 280 °C and 150 °C respectively. The instrument was externally calibrated daily with FlexMix solution (ThermoFisher) while at the beginning of every LC run the internal calibrant was injected (IC run start mode). For 2.1-mm i.d. setup, source parameters were as follows: sheath gas pressure, 40 a.u. and 50 a.u. for positive and negative ionization modes, respectively; aux gas flow, 15 a.u.; sweep gas flow, 0 a.u. Spray voltages were set to 3.3 kV and 3.0 kV for ESI (+) and ESI (−). Ion transfer tube (ITT) and vaporizer temperatures were set to 300 °C and 320 °C. The same MS and MS/MS acquisition parameters and calibration were used as reported above.

5.3.3 Data processing and statistical analysis

FreeStyle (Thermo Fisher Scientific) was used to visualize RAW data, which were then imported to Compound Discoverer v.3.3 (Thermo Fisher Scientific) to normalize, align, detect, and identify compounds. Features were extracted from 0–10 min and 0–11 min of the HILIC and RP chromatography runs, respectively, in the m/z = 70–800 mass range. Data were aligned according to an adaptive curve alignment model. Compounds were detected using the following parameters settings: mass tolerance was set to 5 ppm, while retention time tolerance was set to 0.2 min; minimum peak intensity was set to 100,000 AU and the signal to noise threshold for compound detection was set to 5. The peak rating filter was set to 3. To perform blank subtraction, we maintained max sample/max blank ratio > 5. For statistical comparison between the two setups, signal intensity was normalized by using the algorithm “Constant Sum.” To predict elemental compositions of the compounds, the

relative intensity tolerance was set to 30% for isotope pattern matching. For the mzCloud database search, both the precursor and fragment mass tolerance were set to 5 ppm. The databases used for matching compounds in ChemSpider for structural search were BioCyc, the Human Metabolome Database, and KEGG, and the mass tolerance in ChemSpider Search was set to 5 ppm. The mass tolerance for matching compounds in Metabolika pathways was set to 5 ppm. Compounds were assigned by comparing annotations using the following nodes in order of priority: (1) mzCloud; (2) Predicted Compositions; (3) MassList search; (4) ChemSpider Search; (5) Metabolika search. Principal component analysis was performed by MetaboAnalyst 6.0; samples were log transformed and autoscaled prior to statistical analysis, all other graphs were built using GraphPad Prism 8.0 (GraphPad Software, Boston, MA, USA).

5.4 Results and Discussion

5.4.1 Optimization of 1.0-mm i.d.-based approach

To compare the performance of microbore and narrowbore configurations, columns of identical length, stationary phase chemistry, and particle size were employed on a nano/micro-LC system and an equivalent analytical-flow UHPLC system, respectively. No hardware modifications were applied to the microbore system, which was equipped with 50-µm i.d. fluidics, including inlet and outlet column tubing. In the standard-flow UHPLC system, all fluidic connections remained at 0.1-mm i.d.

Benchmarking of the 1.0-mm i.d. column against the conventional 2.1-mm i.d. format was performed using a mixture of endogenous metabolite standards (**Table S1**). The LC method for the 2.1-mm column was based on a previously optimized in-house protocol (12), with minor modifications. The influence of flow rate, injection volume, analyte concentration, and gradient length was systematically assessed. Endogenous standards were injected at two concentration levels on the 1.0-mm i.d. column, and peak width at half maximum (FWHM) was measured to calculate peak capacity.

Initially, different flow rates (50, 80, 90, 95, and 100 µL/min) were tested. As expected, peak width decreased with increasing flow rate, reaching the minimum value of 0.14 min at 100 µL/min in RP mode, as illustrated in **Figure 1A-D**.

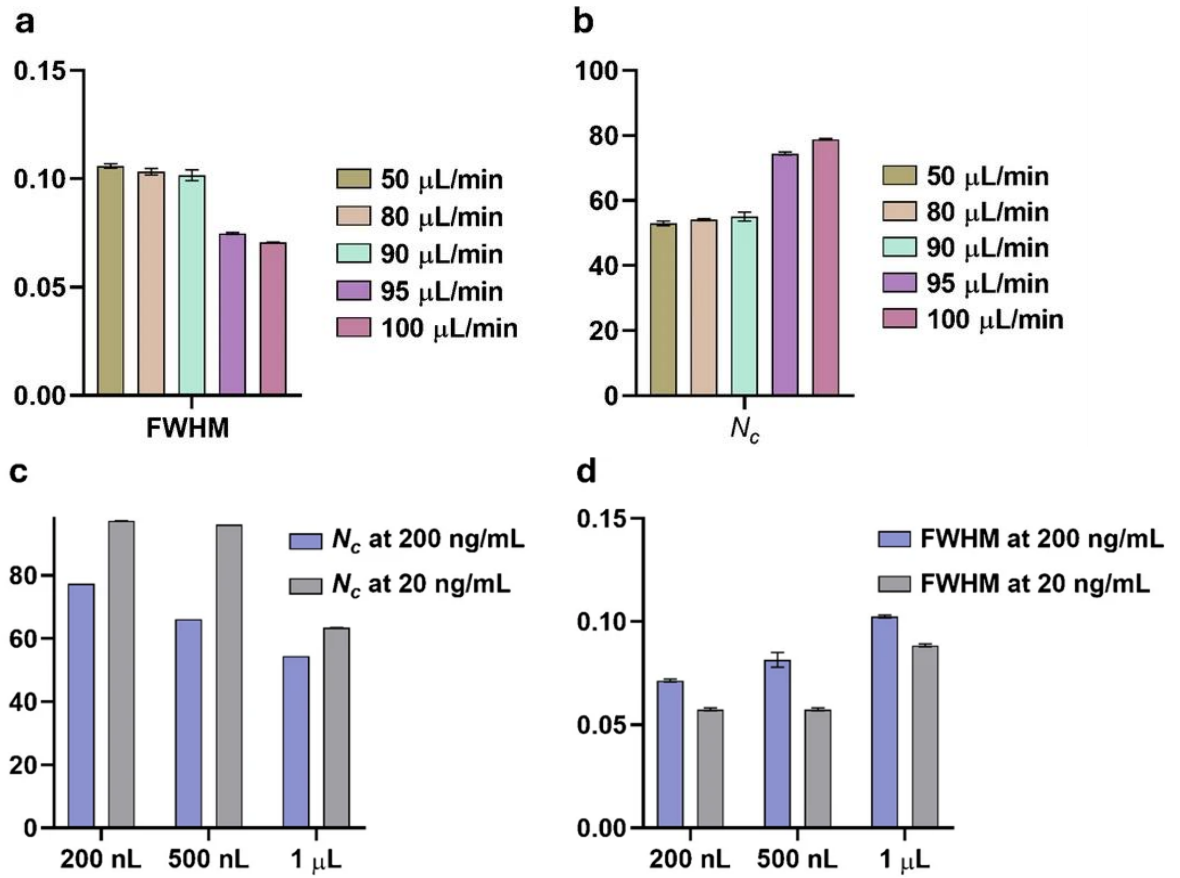


Figure 1A-D. Peak width at half maximum (FWHM) decreased progressively with increasing flow rates (50–100 µL/min), reaching a minimum of 0.14 min at 100 µL/min (**A**). Correspondingly, peak capacity increased, achieving a maximum value of *Nc* = 78 under these conditions (**B**). Injection-volume experiments (0.2, 0.5, and 1 µL at 200 and 20 ng/mL) (**C-D**) demonstrated that the optimal balance between chromatographic efficiency and sensitivity was obtained with a 200 nL injection at 20 ng/mL, yielding a peak capacity of *Nc* = 97.

Conversely, for the metabolites investigated, higher values were observed in HILIC mode (0.19 min, **Figure S1A-D**). Although HILIC is orthogonal to RP and capable of higher retention of the polar metabolome, it typically yields larger peak widths (13); additionally, it must be noted that the employed RP column was longer than the HILIC (150 vs. 100 mm).

Peak capacity values were calculated as follows (14):

$$nc = 1 + \frac{Tg}{w}$$

Consistent with the overall reduction in peak width, the highest peak capacity (*Nc*) values for the 1.0-mm i.d. column were observed at 100 µL/min in RP mode, reaching nc = 78 (Fig. 1b). According to the gradient transfer method calculator (15), the method transfers from the 2.1-mm i.d. column would correspond to a flow rate of 113 µL/min (F_1 to F_2). However, the upper flow rate limit of the Vanquish Neo system (100 µL/min) imposed a constraint, which was further reduced to 90 µL/min in HILIC mode due to system pressure limitations, resulting in a lower *Nc* value of 47 (**Figure S1 B**).

To overcome this limitation, we tested higher flow rates (up to 120 µL/min) by installing the 1.0-mm i.d. column on the analytical UHPLC system. Nevertheless, the results were significantly lower (−48.38%) compared to those obtained on the microflow system, confirming the critical impact of extra-column band broadening when using 1.0-mm i.d. columns on standard UHPLC platforms.

Next, the influence of injection volume and sample concentration was evaluated, as column overload represents a major challenge in microbore setups. Injection volumes of 0.2, 0.5, and 1 µL were tested at two concentration levels (200 and 20 ng/mL). The best performance was achieved with an injection volume of 200 nL at 20 ng/mL, yielding *Nc* = 97 (Fig. 1c–d, Fig. S1d). While reducing injection volume improves chromatographic efficiency, a balance with sensitivity is essential for untargeted metabolomics, where detection of low-abundance metabolites is critical (16).

Finally, gradient length was assessed. As shown in Fig. S2, extending the gradient to 30 min nearly doubled peak capacity (*Nc* = 141 for RP; *Nc* = 90 for HILIC), consistent with peak width compression and the general trend that longer gradients increase peak capacity,

although for small molecules this effect tends to plateau (17). However, analysis time remains a key consideration in untargeted workflows, especially with large sample sets. To maintain equivalent throughput between the 2.1-mm and 1.0-mm setups, we selected the shorter gradient for direct comparison under identical injection volume and concentration conditions.

Overall, comparison with the 2.1-mm i.d. column confirmed that the 1.0-mm i.d. configuration delivered slightly lower performance (FWHM: 0.14 min vs. 0.10 min; *Nc*: 78 vs. 103; **Figure S3**). This reduction can be attributed to operating below the optimal flow rate due to system constraints and potential additional band broadening at the MS source. Notably, we opted to use the high-flow HESI capillary to ensure robustness and minimize clogging risk during real sample analysis.

5.4.2 Sensitivity comparison between 1.0-mm and 2.1-mm i.d. setups on polar standard metabolites

Peak intensities investigated for the standard metabolites were compared between the 2.1-mm i.d. and 1.0-mm i.d. setups under the previously described conditions. As shown in **Figure 2A**, the 1.0-mm configuration provided a substantial sensitivity improvement, with an average fold-change (FC_{avg}) intensity gain of 3.79 relative to the 2.1-mm setup. The highest fold-change was observed for N4-Acetylcytidine (FC = 6.86), whereas the lowest was recorded for L-Carnosine (FC = 1.07) in RP (±) mode.

In contrast, HILIC-ESI (±) generally exhibited lower fold-change values, with an average FC of 1.48. Within this mode, L-Homocitrulline showed the highest gain (FC=3.04), while bilirubin displayed the lowest (FC=0.22). These results confirm that the microbore configuration significantly enhances sensitivity, particularly under RP conditions, although improvements are less pronounced in HILIC mode.

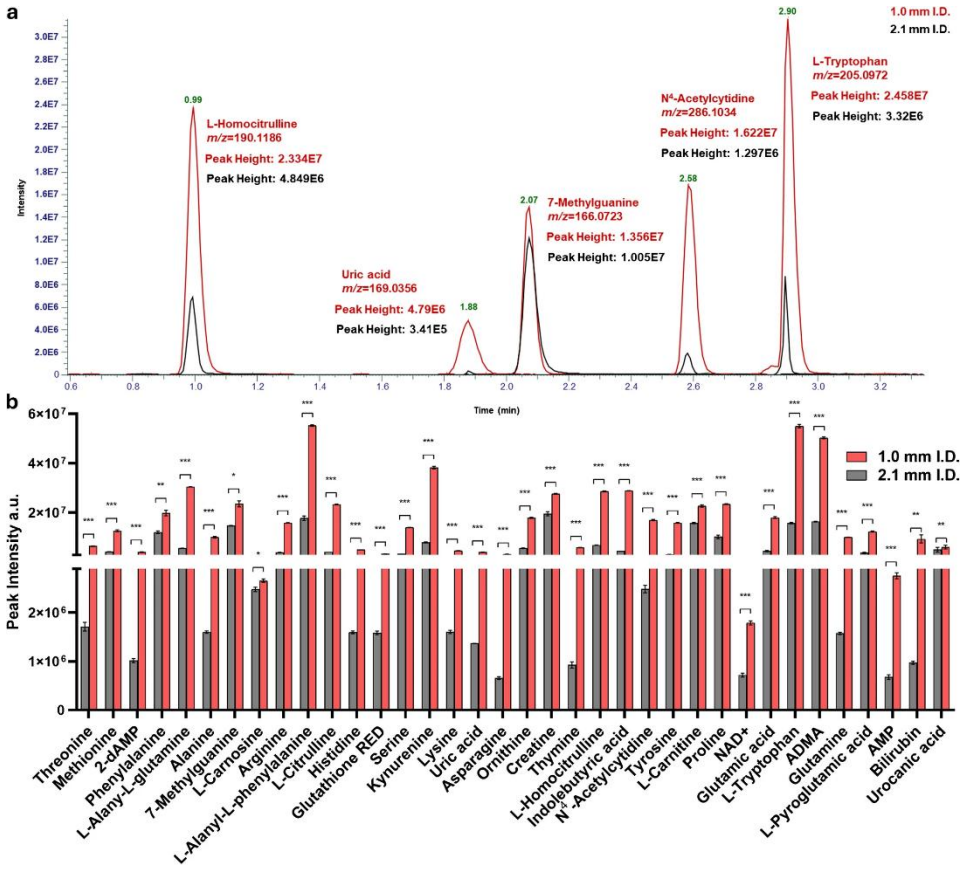


Fig. 2A-B. (**A**) Extracted ion chromatograms (EICs) of selected metabolite standards—L-Homocitrulline, Uric acid, 7-Methylguanine, N4-Acetylcytidine, and L-Tryptophan—acquired using 1.0-mm i.d. (red) and 2.1-mm i.d. (black) LC configurations under identical experimental conditions. (**B**) Comparative analysis of peak intensities for polar standard metabolites obtained with the 1.0-mm and 2.1-mm i.d. methods, highlighting the signal enhancement achieved with the microbore setup.

The differences in response were compound-dependent; however, as illustrated by the extracted ion chromatograms (EICs) in **Figure 2B**, they were also correlated with metabolite retention in both chromatographic modalities. Metabolites with longer retention times exhibited a more pronounced increase in peak intensity compared to early-eluting compounds. Only minor differences in retention times between the two LC systems were observed (±16%), indicating no significant additional delay volume in the micro-LC configuration.

Given that untargeted metabolomics often deals with a wide dynamic range of analyte concentrations in biospecimens, we evaluated the performance of the 1.0-mm i.d. method in terms of dynamic range by preparing calibration curves for standard metabolites spanning three orders of magnitude (**Figure S4**). Linearity was comparable and satisfactory for both methods; however, the 1.0-mm i.d. setup consistently detected and quantified all investigated analytes at the lowest concentration levels. In contrast, the 2.1-mm i.d. method failed to detect 10 out of 15 metabolites within the lower concentration range (1–5 ng/mL). This difference is clearly illustrated in Figure 3a, showing the EICs of N4-Acetylcytidine at the LLOQ: the metabolite was successfully detected and quantified using the 1.0-mm i.d. method, enabling precursor selection in DDA and acquisition of the corresponding MS/MS spectrum. Conversely, the 2.1-mm i.d. method did not detect the metabolite at this concentration.

Based on these observations, we proceeded to compare the limit of detection (LOD) for both approaches as follows:

$$LOD = 3 \times \frac{\sigma}{b}$$

where σ is the residual standard deviation of the calibration line in the LOD region and b is calibration graph slope. The results showed that the 1.0-mm i.d. on average possesses twofold lower limit of detection (LOD_{avg} 1.48 ng/mL vs. 6.18 ng/mL).

Similar results were obtained for the limit of quantification, which was calculated as follows:

$$LOQ = 10 \times \frac{\sigma}{b}$$

In this regard, almost three-fold lower values (LOQ_{avg} 4.94 ng/mL vs. 20.60 ng/mL) can be appreciated from **Figure 3B**. These data are compound dependent and can be influenced by the reproducibility of peak areas, but it is evident how the microbore setup possesses lower values. Complete data for the remaining metabolite standards employed in the optimization phase are reported in **Table 1.**

Clearly, signal to noise ratio (*S*/*N*) values are improved on the 1.0-mm i.d. setup; in this regard, the lower flow rate used for the 1.0-mm i.d. column increases desolvation and ionization efficiency in ESI (18). Complete data are reported in **Table S3**. Additionally, assessment of probe position as well as vaporizer temperature was also investigated to identify the most suitable conditions. The H-ESI source can be moved in different positions: *X* (side to side), *Y* (front to back), and *Z* (vertical). The different working positions in these three directions determine the proximity of the H-ESI probe to the ITT. We evaluated the effect of the *Z* positions by keeping the *X* direction at the center position and *Y* at 1. Our results revealed that a small increase in peak height can be obtained by moving the H-ESI source at low (L) *Z* position. Furthermore, we explored the potential impact of vaporization temperature on ionization by testing different vaporizer temperatures: 0 °C, 100 °C, 150 °C, and 200 °C. Our findings suggest that working at 150 °C vaporizer temperature may offer more favorable results; hence, the higher intensity was obtained with the probe position “L” and vaporizer temperature of 150 C° (**Figure S5**

Microbore							
Endogenous standard	**Linear range (1-1000 ng/mL)**	**Coefficient of correlation (R^2)**	**Limit of Detection (LOD) ng/mL**	**Limit of Quantification (LOQ) ng/mL**	**Intra-day stability**		**Accuracy**
					CV% RT	**CV% Areas**	
2'-dAMP	5-10 ng/mL	R^2 = 0.99691	1.742	5.805	0.46%	5.23%	**96.66%**
2'-Deoxyuridine	1-5 ng/mL	R^2 = 0.99957	3.025	10.085	0.00%	0.71%	**95.88%**
2-Oxoadipic acid	10-25 ng/mL	R^2 = 0.99992	4.219	14.064	0.01%	0.01%	**88.96%**
7-Methylguanine	1-5 ng/mL	R^2 = 0.99982	0.213	0.710	0.00%	6.73%	**90.27%**
Creatine	1-5 ng/mL	R^2 = 0.99998	0.256	0.854	0.00%	5.90%	**87.13%**
Guanine	25-50 ng/mL	R^2 = 0.99983	1.027	3.424	0.00%	0.06%	**91.68%**
Kynurenine	5-10 ng/mL	R^2 = 0.99985	0.875	2.918	0.23%	9.80%	**97.45%**
L-Carnitine	5-10 ng/mL	R^2 = 0.99995	0.452	1.506	0.71%	7.90%	**88.68%**
L-Carnosine	5-10 ng/mL	R^2 = 0.99856	0.282	0.941	0.82%	9.26%	**84.09%**
L-Homocitrulline	5-10 ng/mL	R^2 = 0.99786	0.891	2.970	0.00%	9.42%	**88.18%**
L-Tryptophan	1-5 ng/mL	R^2 = 0.99998	1.989	6.631	0.20%	5.77%	**99.18%**
N^4-Acetylcytidine	5-10 ng/mL	R^2 = 0.99995	0.703	2.344	0.23%	8.24%	**88.18%**
Thymine	5-25 ng/mL	R^2 = 0.99976	2.005	6.682	0.00%	6.44%	**98.07%**
Urocanic Acid	1-5 ng/mL	R^2 = 0.99990	3.083	10.277	0.00%	0.01%	**93.19%**
Narrowbore							
2'-dAMP	10-25 ng/mL	R^2 = 0.99990	2.905	9.683	0.31%	3.79%	**91.73%**
2'-Deoxyuridine	25-50 ng/mL	R^2 = 0.99995	4.24	14.14	1.02%	2.05%	**91.75%**
2-Oxoadipic acid	10-25 ng/mL	R^2 = 0.99996	4.76	15.86	0.82%	1.62%	**98.77%**
7-Methylguanine	1-5 ng/mL	R^2 = 0.99989	0.265	0.884	0.34%	2.04%	**94.72%**
Creatine	5-10 ng/mL	R^2 = 0.99953	0.250	0.835	0.72%	2.71%	**94.97%**
Guanine	25-50 ng/mL	R^2 = 0.99981	7.315	24.384	0.37%	5.17%	**95.78%**
Kynurenine	10-25 ng/mL	R^2 = 0.99995	7.679	25.595	0.00%	1.80%	**89.80%**

L-Carnitine	10-25 ng/mL	R^2 = 0.99993	1.432	4.773	0.94%	9.19%	**86.60%**
L-Carnosine	10-25 ng/mL	R^2 = 0.99788	0.800	2.666	0.94%	9.19%	**75.56%**
L-Homocitrulline	5-10 ng/mL	R^2 = 0.99989	0.16	0.52	0.66%	5.87%	**89.97%**
L-Tryptophan	1-5 ng/mL	R^2 = 0.99812	3.270	10.901	0.22%	2.43%	**90.01%**
N^4-Acetylcytidine	50-100 ng/mL	R^2 = 0.98942	8.265	27.550	1.02%	5.13%	**87.59%**
Thymine	50-100 ng/mL	R^2 = 0.99707	44.455	148.184	0.42%	16.50%	**96.09%**
Urocanic Acid	5-10 ng/mL	R^2 = 0.99995	0.739	2.465	0.57%	1.62%	**96.41%**

Table 1. Analytical performance parameters, including LOD and LOQ values, for all metabolite standards evaluated during method optimization.

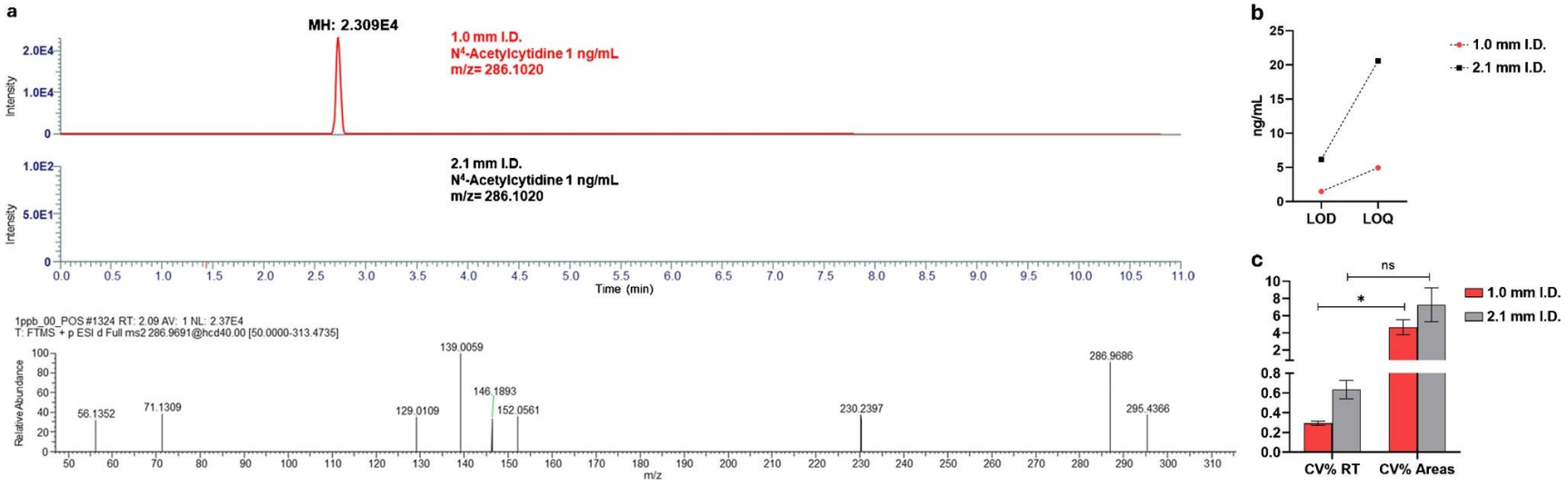


Figure 3A-C. (**A**) Extracted ion chromatograms (EICs) of N4-Acetylcytidine at 1 ng/mL, with corresponding MS/MS spectra, demonstrating successful precursor detection and fragmentation using the 1.0-mm i.d. method, while no precursor ion was detected with the 2.1-mm i.d. configuration. (**B**) Comparison of limits of detection (LOD) and limits of quantification (LOQ) for the investigated metabolites obtained with the 1.0-mm (red) and 2.1-mm (black) setups. (**C**) Repeatability assessment expressed as signal variability for both configurations, highlighting the performance of the 1.0-mm (red) relative to the 2.1-mm i.d. (black) method.

5.4.3 Robustness and Repeatability of the 1.0 mm i.d. setup

In untargeted metabolomics workflows, repeatability of retention times and peak areas is a critical parameter, particularly during pre-processing steps such as peak alignment. The 1.0-mm i.d. configuration demonstrated superior repeatability compared to the 2.1-mm i.d. setup, as evidenced by lower intraday CV% values for both retention time and peak area (CV_{RT}: 0.29% vs. 0.63%; CV_{areas}: 4.65% vs. 7.27%) (**Figure 3C**), with similar trends observed for inter-day measurements (data not shown).

Furthermore, robustness was confirmed using real samples: after 250 consecutive injections over 48 hours, system backpressure remained stable, as illustrated by the backpressure traces throughout the entire gradient (**Figure S6**). Complete data are reported in **Table 1**. Notably, solvent consumption in the microbore setup was significantly lower, totaling only 300 mL over 250 runs compared to 1.5 L in the narrowbore configuration, highlighting an additional advantage in terms of sustainability.

5.3.4 Metabolites annotation and filtering

In untargeted metabolomics using LC-MS/MS, the integration of comprehensive annotation and rigorous filtration steps is critical for obtaining reliable biological insights. Metabolite annotation benefits greatly from advanced methodologies, including the use of tandem mass spectral libraries, in silico fragmentation tools, and multi-layer metabolic networking to improve coverage and accuracy of both known and unknown metabolites (19, 20). These approaches address challenges posed by the structural diversity and complexity of metabolite features captured, enabling expansion of putative identifications beyond traditional spectral library matching.

Filtration workflows serve as an equally vital component, where the removal of noise, redundant features, and false positives reduces data complexity and enhances downstream statistical validity. A key refinement in our workflow involves the meticulous exclusion of duplicate features, both within individual polarity modes and between the two polarities after data merging, addressing the common issue of feature redundancy due to ionization variability. Moreover, exogenous features, often introduced from contaminants or background signals, are systematically filtered out to prevent confounding biological interpretations. This polarity bridging and duplicate removal step is supported by recent advances in cross-ionization mode similarity predictions such as MS2DeepScore, which

enable robust comparisons and deduplication of features across positive and negative ion datasets (21).

By combining these filtering steps with normalization, retention time consistency checks, and intensity thresholds, together with network-based metabolite annotation strategies for enhanced metabolite identification, the workflow ensures high confidence in the final list of annotated metabolites. This systematic approach not only improves the specificity of metabolite detection but also aids in revealing biologically meaningful metabolic pathways by reducing redundancy and noise, ultimately strengthening interpretation and reproducibility in untargeted metabolomics studies (22, 23). Such refined data processing pipelines represent a crucial advancement in tackling the inherent challenges of untargeted LC-MS/MS metabolomics data, fostering improved discovery and characterization of metabolomes across diverse biological systems.

5.3.5. Metabolome coverage and comparison with 2.1 mm i.d. setup over different samples

To further investigate the potential increase in metabolome coverage, we analyzed multiple sample types: human plasma at two different volumes, dried blood spots (DBS), and patient-derived colorectal cancer organoids. Metabolite annotation was performed using MS/MS spectral libraries mzCloud and mzVault (MSI level 2), and when available, through direct comparison with authentic standards (MSI level 1). Each sample was analyzed using a combination of reversed-phase (RP) and hydrophilic interaction liquid chromatography (HILIC) methods, as their complementary use significantly enhances metabolome coverage (24).

Different mobile phase additives were tested, and the optimal combination was identified as formic acid and ammonium fluoride for RP in ESI (+) and (−) modes, respectively, while ammonium acetate was selected for HILIC in both polarities (data not shown). RP and HILIC annotations were subsequently merged at each MS level and polarity, removing duplicates and reporting only the adduct with the highest intensity to express the effective gain of features in the 1.0-mm i.d. setup.

Figure 4A illustrates the percentage increase in MS^1 features, MS^2 features, and library-matched MS/MS spectra for each matrix. On average, the 1.0-mm i.d. setup annotated

more MS^1 features (+38.95%) and triggered more MS/MS events (+39.26%), as visualized in the dot-plot maps (Fig. S7) showing a higher number of precursors selected for HCD. Although the increase in library-matched MS/MS spectra was lower (+18.23%), this improvement is particularly relevant for detecting low-abundance features critical for identifying metabolites involved in key molecular pathways.

As highlighted in recent studies (25), the discrepancy between unannotated and annotated MS/MS spectrum is largely due to in-source fragmentation occurring prior to MS/MS events, resulting in unidentified metabolites. **Figure 4B** shows that, across all matrices, over 34% of metabolites were annotated exclusively with the 1.0-mm i.d. method, reinforcing its potential for untargeted approaches. PCA score plots (**Figure 5A**) demonstrate distinct clustering of samples analyzed with the two setups, reflecting the broader metabolic coverage achieved by the microbore approach. This is further supported by base peak chromatograms (**Figure 5B**), where the same organoid extract analyzed by RP-ESI (−) on the 1.0-mm setup exhibits a higher number of peaks, indicating increased biological information.

Overall, merging RP and HILIC results across all matrices and metabolite classes (e.g., nucleotides, amino acids and derivatives, carnitines, organic acids) yielded an average fold increase of 1.98 (**Figure 5C**), consistent with previous observations comparing 1.0-mm and 2.1-mm setups, although prior studies considered only RP mode (26). The combined annotations resulted in a global coverage of 507 metabolites. Among these, the 1.0-mm setup showed higher intensity for 130 of 184 shared annotations, mainly amino acids and derivatives, dipeptides, short-chain acylcarnitines, and fatty acid conjugates (**Figure S 8A**). Additionally, 219 metabolites were detected exclusively with the 1.0-mm setup, representing a broader range of subclasses (**Figure S 8B**). Conversely, 104 annotations were unique to the 2.1-mm setup (**Figure S 8C**), primarily glycerophospholipids, fatty acids, and long-chain acylcarnitines.

This difference likely reflects the semi-polar and non-polar nature of these compounds and the fact that HESI conditions in the 1.0-mm setup were optimized for polar metabolites, impacting ionization efficiency. In contrast, the 2.1-mm setup employed well-established analytical-flow parameters for Orbitrap systems. Furthermore, these compounds exhibit strong retention in RP and narrow peak widths, benefiting from the slightly higher peak

capacity of the narrowbore approach. This limitation underscores the need for further optimization of LC–MS conditions for semi-polar and non-polar metabolites in microbore workflows. Recent applications have demonstrated the potential of microbore setups for targeted lipidomics (27), using LC conditions tailored for lipid analysis.

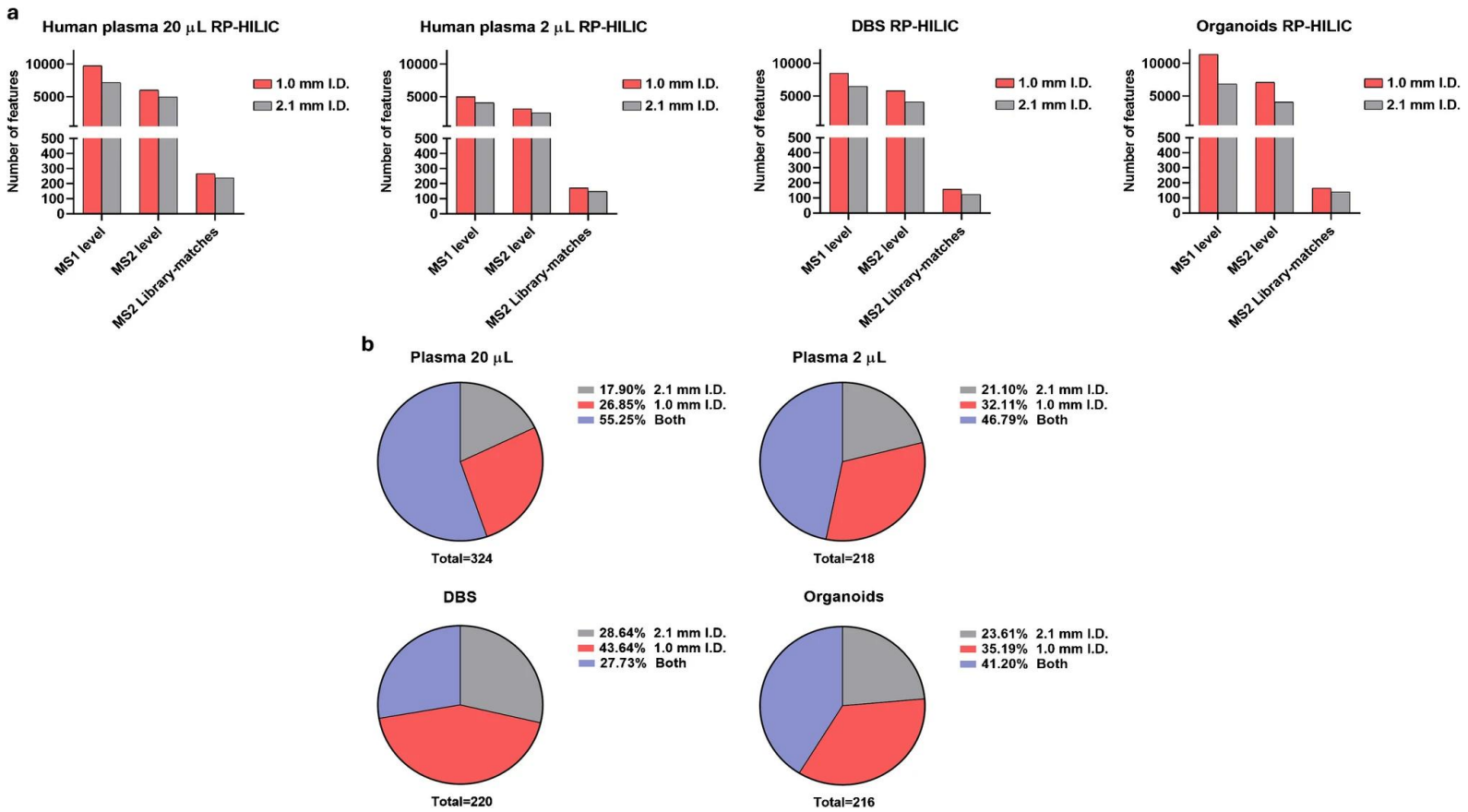


Figure 4A-B. Impact of the 1.0-mm i.d. configuration on feature detection and metabolite annotation in untargeted metabolomics across different biological matrices. (**A**) Percentage increase in MS^1 features, MS^2 events, and library-matched MS/MS spectra obtained with the 1.0-mm i.d. setup relative to the 2.1-mm configuration. The microbore method yielded, on average, +38.95% more MS^1 features and +39.26% more MS/MS events. (**B**) Proportion of metabolites exclusively annotated by each method across all matrices, demonstrating that over 34% of annotations were uniquely achieved using the 1.0-mm i.d. setup, underscoring its enhanced performance for untargeted metabolomics applications.

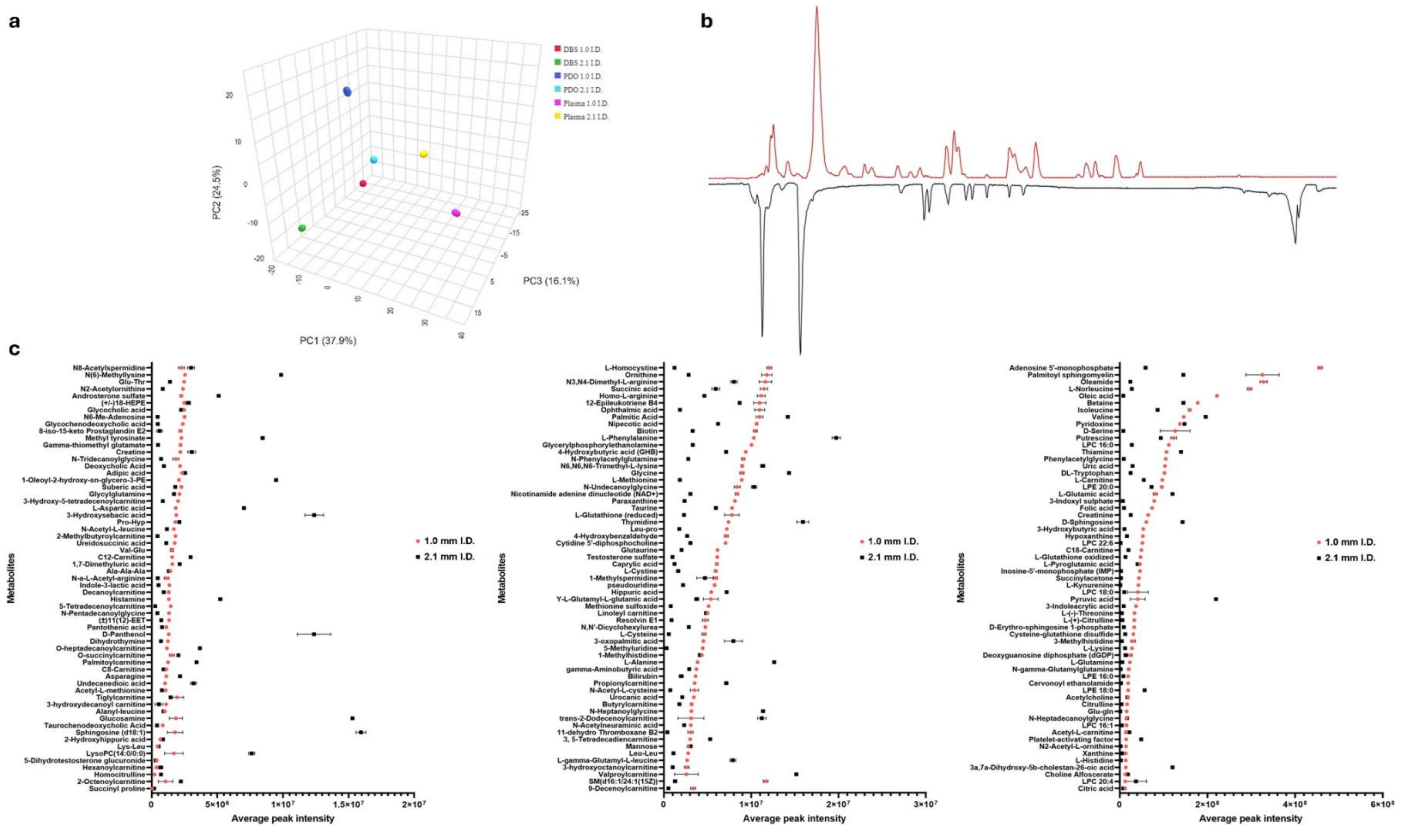


Figure 5a-c. **(A**) Three-dimensional principal component analysis (PCA) score plot showing clear clustering of the four investigated matrices and separation between the 1.0-mm and 2.1-mm i.d. approaches, highlighting method-dependent differences in metabolomic profiles. (**B**) Representative base peak chromatograms (BPCs) of a patient-derived organoid (PDO) sample acquired with the 1.0-mm (red) and 2.1-mm i.d. (black) setups, illustrating enhanced signal intensity and peak definition with the microbore configuration. (**C**) Quantitative comparison of peak intensities for metabolites commonly annotated in all four matrices, demonstrating the signal gain achieved with the 1.0-mm i.d. method in real-sample analyses.

5.5 Conclusions

In this work, we have explored and demonstrated the utility of 1.0-mm i.d. microbore column-based separation for untargeted metabolomics. Higher signal intensity, with lower values of LOD and LOQ, is obtained, resulting in appreciable gain in the overall coverage for polar metabolome with respect to 2.1-mm i.d. setup, together with better repeatability and robustness in the analysis of real samples. The developed method shows an average 1.3-fold increase in response compared to conventional narrowbore setup for several biospecimens, while maintaining similar throughput of the 2.1-mm i.d. approach. The method can have more metabolome information across different matrices, from conventional to low amounts. Finally, and noteworthy, a drastic reduction of solvent consumption is obtained. These results underline the potential employment of 1.0-mm i.d. microbore columns in untargeted metabolomics as a valuable alternative to narrowbore-based separations to achieve higher metabolome coverage and same analytical throughput. Further extension to non-polar metabolites and lipids could expand the utilization and coverage of 1.0-mm i.d. setup. Lastly, the drastic reduction of solvent consumption makes this approach environmentally friendly, especially when dealing with the screening of large cohorts of samples.

Bibliography Chapter V

[1]. Nicholson JK, Holmes E, Kinross JM, Darzi AW, Takats Z, Lindon JC. Metabolic phenotyping in clinical and surgical environments. Nature [Internet]. 2012;491. Available from: https://doi.org/10.1038/nature11708

[2]. Wishart DS. Emerging applications of metabolomics in drug discovery and precision medicine. Nat Rev Drug Discov [Internet]. 2016;15. Available from: https://doi.org/10.1038/nrd.2016.32

[3]. Theodoridis GA, Gika HG, Want EJ, Wilson ID. Liquid chromatography-mass spectrometry based global metabolite profiling: A review. Anal Chim Acta [Internet]. 2012;711. Available from: https://doi.org/10.1016/j.aca.2011.09.042

[4]. Zhou B, Xiao JF, Tuli L, Ressom HW. LC-MS-based metabolomics. Mol Biosyst [Internet]. 2012;8. Available from: https://doi.org/10.1039/C1MB05350G

[5]. Cajka T, Fiehn O. Toward merging untargeted and targeted methods in mass spectrometry-based metabolomics and lipidomics. Anal Chem [Internet]. 2016;88. Available from: https://doi.org/10.1021/acs.analchem.5b04491

[6]. Desmet G, Eeltink S. Fundamentals for LC miniaturization. Anal Chem [Internet]. 2013;85. Available from: https://doi.org/10.1021/ac303317c

[7]. Poitout-Belissent F, Aulbach A, Tripathi N, Ramaiah L. Reducing blood volume requirements for clinical pathology testing in toxicologic studies-points to consider. Vet Clin Pathol [Internet]. 2016;45. Available from: https://doi.org/10.1111/vcp.12429

[8]. Bian Y, Zheng R, Bayer FP, Wong C, Chang YC, Meng C. Robust, reproducible and quantitative analysis of thousands of proteomes by micro-flow LC–MS/MS. Nat Commun [Internet]. 2020;11. Available from: https://doi.org/10.1038/s41467-019-13973-x

[9]. Kuster B, Bian Y, Bayer FP, Chang YC, Meng C, Hoefer S. Robust microflow LC-MS/MS for proteome analysis: 38 000 runs and counting. Anal Chem [Internet]. 2021;93. Available from: https://doi.org/10.1021/acs.analchem.1c00257

[10]. Lenz EM, Williams RE, Sidaway J, Smith BW, Plumb RS, Johnson KA. The application of microbore UPLC/oa-TOF-MS and 1H NMR spectroscopy to the metabonomic analysis of rat urine following the intravenous administration of pravastatin. J Pharm Biomed Anal [Internet]. 2007;44. Available from: https://doi.org/10.1016/j.jpba.2007.04.035

[11]. Fitz V, Abiead Y, Berger D, Koellensperger G. Systematic investigation of LC miniaturization to increase sensitivity in wide-target LC-MS-based trace bioanalysis of small molecules. Front Mol Biosci [Internet]. 2022;9. Available from: https://doi.org/10.3389/fmolb.2022.857505

[12]. Mires S, Sommella E, Merciai F, Salviati E, Caponigro V, Giovanna M, et al. Plasma metabolomic and lipidomic profiles accurately classify mothers of children with congenital heart disease: an observational study. 2024. https://doi.org/10.1007/s11306-024-02129-8.

[13]. Song H, Adams E, Desmet G, Cabooter D. Evaluation and comparison of the kinetic performance of ultra-high performance liquid chromatography and high-performance liquid chromatography columns in hydrophilic interaction and reversed-phase liquid chromatography conditions. J Chromatogr A [Internet]. 2014;1369. Available from: https://doi.org/10.1016/j.chroma.2014.10.002

[14]. Neue UD. Theory of peak capacity in gradient elution. J Chromatogr A [Internet]. 2005;1079. Available from: https://doi.org/10.1016/j.chroma.2005.03.008

[15]. Guillarme D, Nguyen DTT, Rudaz S, Veuthey JL. Method transfer for fast liquid chromatography in pharmaceutical analysis: Application to short columns packed with small particle. Part II: Gradient experiments. Eur J Pharm Biopharm [Internet]. 2008;68. Available from: https://doi.org/10.1016/j.ejpb.2007.06.018

[16]. Werres T, Schmidt TC, Teutenberg T. The influence of injection volume on efficiency of microbore liquid chromatography columns for gradient and isocratic elution. J Chromatogr A [Internet]. 2021;1641. Available from: https://doi.org/10.1016/j.chroma.2021.461965

[17]. Soliven A, Haidar Ahmad IA, Filgueira MR, Carr PW. Optimization of gradient reversed phase chromatographic peak capacity for low molecular weight solutes. J Chromatogr A [Internet]. 2013;1273. Available from: https://doi.org/10.1016/j.chroma.2012.11.068

[18]. Uclés Moreno A, Herrera López S, Reichert B, Lozano Fernández A, Hernando Guil MD, Fernández-Alba AR. Microflow liquid chromatography coupled to mass spectrometry-an approach to significantly increase sensitivity, decrease matrix effects, and reduce organic solvent usage in pesticide residue analysis. Anal Chem [Internet]. 2015;87. Available from: https://doi.org/10.1021/ac5035852

[19]. Wang, X., Liu, Y., Jiang, C., Huang, Z., Yan, H., Wong, S. H., Johnson, C. H., Zhang, J., Ge, Y., Zhang, F., Zhang, J., Lai, R., Gao, P., Zhang, X., & Shen, X. (2026). TidyMass2: advancing LC-MS untargeted metabolomics through metabolite origin inference and metabolic feature-based functional module analysis. Nature communications, 10.1038/s41467-026-68464-7. https://doi.org/10.1038/s41467-026-68464-7

[20]. Zhou, Z., Luo, M., Zhang, H., Yin, Y., Cai, Y., & Zhu, Z. J. (2022). Metabolite annotation from knowns to unknowns through knowledge-guided multi-layer metabolic networking. Nature communications, 13(1), 6656. https://doi.org/10.1038/s41467-022-34537-6

[21]. de Jonge, N. F., Chekmeneva, E., Schmid, R., Joas, D., Truong, L. J., van der Hooft, J. J. J., & Huber, F. (2026). Cross ionization mode chemical similarity prediction between tandem mass spectra in metabolomics. Nature communications, 10.1038/s41467-026-69083-y. Advance online publication. https://doi.org/10.1038/s41467-026-69083-y

[22]. Shen, X., Wang, R., Xiong, X., Yin, Y., Cai, Y., Ma, Z., Liu, N., & Zhu, Z. J. (2019). Metabolic reaction network-based recursive metabolite annotation for untargeted metabolomics. Nature communications, 10(1), 1516. https://doi.org/10.1038/s41467-019-09550-x

[23]. Schiffman, C., Petrick, L., Perttula, K., Yano, Y., Carlsson, H., Whitehead, T., Metayer, C., Hayes, J., Rappaport, S., & Dudoit, S. (2019). Filtering procedures for untargeted LC-MS metabolomics data. BMC bioinformatics, 20(1), 334. https://doi.org/10.1186/s12859-019-2871-9

[24]. García-Cañaveras JC, López S, Castell JV, Donato MT, Lahoz A. Extending metabolome coverage for untargeted metabolite profiling of adherent cultured hepatic cells. Anal Bioanal Chem [Internet]. 2016;408. Available from: https://doi.org/10.1007/s00216-015-9227-8

[25]. Giera M, Aisporna A, Uritboonthai W, Siuzdak G. The hidden impact of in-source fragmentation in metabolic and chemical mass spectrometry data interpretation. Nat Metab. 2024. https://doi.org/10.1038/s42255-024-01076-x.

[26]. Gray N, Lewis MR, Plumb RS, Wilson ID, Nicholson JK. High-Throughput Microbore UPLC−MS Metabolic Phenotyping of Urine for Large-Scale Epidemiology Studies. J Prot Res [Internet]. 2015;14. Available from: https://doi.org/10.1021/acs.jproteome.5b00203

[27]. Gadara D, Berka V, Spacil Z. High-throughput microbore LC-MS lipidomics to investigate APOE phenotypes. Anal Chem [Internet]. 2024;96. Available from: https://doi.org/10.1021/acs.analchem.3c02652

CHAPTER VI

Spatial Deep Visual Proteomics Analysis of Low-Input Microdissected Thyroid Tumors from DICER1-Mutant Patients: A Pilot Study

6.1 Abstract

DICER1 syndrome is an autosomal dominant disorder with variable penetrance, characterized by a broad spectrum of developmental abnormalities, hamartomatous lesions, and benign, teratoid, or malignant tumors affecting multiple organs. DICER1 encodes a highly conserved RNase III endoribonuclease that is essential for microRNA (miRNA) biogenesis, mediating the maturation of precursor miRNAs into functional 5p and 3p strands and thereby regulating post-transcriptional gene expression. The canonical pathogenic mechanism involves a germline loss-of-function variant followed by a somatic hotspot missense mutation affecting metal-binding residues within the RNase IIIb domain. This “two-hit” model results in a selective defect in 5p miRNA processing-with preservation of partial DICER1 activity-ultimately producing a characteristic imbalance between 5p and 3p miRNA species.

In the thyroid gland, germline DICER1 pathogenic variants are strongly associated with multinodular follicular disease, differentiated thyroid carcinoma, and poorly differentiated thyroid carcinoma (PDTC), with particularly high penetrance in children and young adults. Mounting evidence indicates that DICER1-associated pediatric PDTC constitutes a distinct biological and molecular entity compared with adult PDTC, showing unique evolutionary trajectories and heterogeneous clinical behavior. Despite these advances, the molecular mechanisms driving tumor progression, intratumoral heterogeneity, and high-grade transformation in DICER1-mutant thyroid lesions remain poorly understood-largely due to the scarcity of analytical platforms suitable for spatially resolved, ultra-low-input clinical specimens.

In this study, we applied a spatial deep visual proteomics workflow (DVP) based on a laser-microdissected tissue to a cohort of DICER1-mutant thyroid tumors, including cases with germline and somatic hotspot variants. Laser microdissection enabled the precise isolation of high-grade tumor foci and adjacent lower-grade regions, allowing investigation of proteomic heterogeneity while preserving morphological context. The mDVP pipeline, optimized for archival tissue with minimal material, provided deep proteome coverage without compromising spatial resolution.

Using this approach, we achieved robust and reproducible proteomic profiling of microdissected thyroid tumor regions and integration with histopathological and immunohistochemical data, further elucidated differentiation states and active biological programs within tumor subregions.

Overall, this work demonstrates the feasibility and analytical power of applying spatial, sensitivity-driven proteomics to DICER1-mutant thyroid tumors, establishing mDVP as a robust and scalable workflow for molecular characterization of ultra-low-input clinical material. These findings underscore the value of spatial proteomics for dissecting tumor heterogeneity and disease evolution in rare hereditary cancer syndromes, with broader implications for biomarker discovery and translational research.

6.2 Introduction

DICER1 syndrome is a rare autosomal dominant cancer predisposition disorder with variable penetrance that presents with a broad spectrum of developmental abnormalities, hamartomatous lesions, and benign or malignant tumors affecting multiple organs. The disorder is caused by germline pathogenic variants in the DICER1 gene, which encodes a highly conserved RNase III endoribonuclease essential for microRNA (miRNA) biogenesis (1). DICER1 mediates the maturation of precursor miRNAs into functional 5p and 3p strands, thereby regulating post-transcriptional gene expression and maintaining cellular homeostasis (1).

The canonical pathogenic mechanism underlying DICER1 syndrome follows a non-classical "two-hit" model that distinguishes it from traditional tumor suppressors (2). Rather than complete loss of function, this model involves an initial germline loss-of-

function variant followed by a somatic hotspot missense mutation affecting metal-binding residues within the RNase IIIb domain. This dual-hit pattern results in a selective defect in 5p miRNA processing while preserving partial DICER1 activity, ultimately producing a characteristic imbalance between 5p and 3p miRNA species (1). This unique mechanism has important implications for tumorigenesis, as the dysregulated miRNA landscape drives oncogenic transformation through both loss of tumor-suppressive 5p miRNAs and unexpected gain of function from normally silent 3p miRNAs (3).

In the thyroid gland, germline DICER1 pathogenic variants are strongly associated with multinodular follicular disease, differentiated thyroid carcinoma, and poorly differentiated thyroid carcinoma (PDTC), with particularly high penetrance in children and young adults (4). DICER1-associated pediatric PDTC constitutes a distinct biological and molecular entity compared with adult PDTC, showing unique evolutionary trajectories and heterogeneous clinical behavior (3). The association between DICER1 variants and early-onset thyroid disease is particularly compelling, with pediatric follicular thyroid carcinomas demonstrating a prevalence of DICER1 alterations approaching 50% in some cohorts (4). Despite these advances in characterizing DICER1-associated thyroid pathology, the molecular mechanisms driving tumor progression, intratumoral heterogeneity, and high-grade transformation in DICER1 mutant thyroid lesions remain poorly understood, largely due to the scarcity of analytical platforms suitable for spatially resolved profiling of ultra-low-input clinical specimens.

Traditional proteomic approaches have been limited in their ability to preserve spatial context while achieving deep protein coverage from archival tissue samples with minimal material. Recent advances in spatial proteomics, particularly deep visual proteomics (DVP) combined with laser microdissection, offer unprecedented opportunities to investigate proteomic heterogeneity while maintaining morphological context (5). These approaches enable the precise isolation of distinct tumor regions with single cell to sub-millimeter resolution, facilitating the investigation of spatial proteomic variation within heterogeneous tissues. In this study, we applied a spatial deep visual proteomics workflow based on laser-microdissected tissue to a cohort of DICER1 mutant thyroid tumors, including cases with germline and somatic hotspot variants. Laser microdissection enabled the precise isolation of high-grade tumor foci and adjacent lower-grade regions, allowing

investigation of proteomic heterogeneity while preserving morphological context. Our optimized mDVP pipeline, designed for archival tissue with minimal material, provided deep proteome coverage without compromising spatial resolution.

6.3 Methods

6.3.1 Sample collection

Five pediatric patients (≤18 years) with thyroid tumors displaying high-grade morphological features were identified and included in this study. All patients were of Italian origin and underwent surgical treatment at the Endocrine Surgery Unit of IRCCS Ospedale San Raffaele (Milan, Italy) after evaluation by a multidisciplinary tumor board.

All surgical specimens were entirely submitted for histopathological examination in accordance with the standard protocol of the Pathology Unit for pediatric patients. Histological classification and diagnostic assessment were performed following the criteria established in the latest World Health Organization (WHO) classification of thyroid tumors (6,7).

During histological evaluation, particular attention was paid to morphological features previously associated with DICER1-mutated tumors and DICER1 syndrome. Specifically, two characteristic patterns were systematically assessed:

Atrophic changes, defined as well-demarcated ischemic-like areas within the tumor, characterized by vanishing or “ghost” cells embedded in dense hyaline stroma, in the absence of inflammation, apoptosis, or artifacts related to fine-needle aspiration procedures. These changes were predominantly localized in subcapsular regions and, in some cases, observed in multiple intranodular foci, including small nodules distinct from the main lesion (8,9).

Involutional changes, identified as clusters of ectatic macrofollicles lined by flattened epithelium in the non-neoplastic thyroid parenchyma, consistent with features previously described in association with DICER1 syndrome (10,11).

Tumor nuclear features were additionally evaluated and documented in descriptive terms.

6.3.2 Membrane slide preparation

Prior to tissue mounting, slides were washed with 70% EtOH for 5min, rinsed in ddH2O and incubated with poly-L-lysine at 0.01% (w/v) (Sigma-Aldrich, cat.no. P1524) for 10min to optimize tissue adherence. Slides were then dried overnight at room temperature. All samples were sectioned at a thickness of 5 µm and mounted on metal frame PPS membrane slides (Leica, 619 cat.no. 11600294), followed by overnight drying at 37 °C. Immediately prior to deparaffinization, slides were incubated at 60 °C for 30 min.

6.3.2 Tissue sectioning and Cyclic Immunofluorence (CyIF) staining

The FFPE blocks were deparaffinated and sequentially rehydrated as follows: two 5-min immersions in Neo-Clear buffer, two 2-min washes in 99% ethanol, and sequential 2-min washes in 80%, 70% ethanol, and 1x PBS (twice for 1min each). Heat-mediated antigen retrieval was performed in Tris-EDTA (pH 9) using a steamer for 25min, followed by cooling to room temperature in the retrieval solution. Slides were then washed three times in 1x PBS. To reduce tissue autofluorescence, sections were pre-bleached for 30min under direct white light in 4.5% H_2O_2 and 24 mM NaOH diluted in 1x PBS. Following three additional PBS washes, tissue sections were outlined with a PAP-pen (Science Services, N71312-N) to minimize the reaction volume and blocked with 3% BSA (Serva, 11948.01) in 1x PBS for 30 min at room temperature. Antibody incubation was performed overnight at 4 °C in a blocking buffer in a humidified staining chamber. Immunofluorescence staining was conducted over several cycles. All antibodies were directly conjugated (**Table 1**). Following imaging, coverslips were removed by soaking the slides in 1x PBS within a vertical Coplin jar on a platform shaker. Detached slides were washed in 1x PBS, bleached for 30 min under direct white light in 4.5% H2O2 and 24mM NaOH diluted in 1x PBS, and washed again in 1x PBS before proceeding to the next round of primary antibody staining. This process was repeated twice to achieve staining for all markers. After the final cycle, coverslips were removed by soaking the slides in 1x PBS within a vertical Coplin jar on a platform shaker. Once detached, the slides were rinsed with Milli-Q water, air-dried, and stored at 4 °C. Imaging was conducted using a Zeiss Axioscan 7 slide scanner equipped with the Colibri 7 LED light source and an EC Plan-Neofluar 20x/0.50 M27 objective at 2 × 2 binning. Stitching of the raw images was performed using ZEN software (version 3.5, Blue Edition), with the DAPI channel set as the reference for all channels and the following parameters: minimal

overlap of 5%, maximal shift of 15%, Comparer set to Optimized, and Global Optimizer set to Best.

Antibody	Company	Species	Dilution
CD56	Roche/Ventana	Mouse	1:100
CD163	Roche/Ventana	Mouse	1:100
CD68	Roche/Ventana	Mouse	1:100
PAX8	Roche/Ventana	Mouse	1:100

Table 1. Antibodies employed for Spatial Proteomics

6.3.3 Hematoxylin and Eosin (H&E) staining

Tissue sections were dehydrated by washing 2x5 minutes in Neo Clear (Sigma Aldrich, 1.09483.5000), followed by a series of 99%, 80% and 70% ethanol for 2 minutes, respectively, and rehydrated by immersing in milliQ water three times. Then, slides were stained in Mayer's hematoxylin for three minutes and immersed in tap water for another ten minutes, rinsed in milliQ water and stained with eosin for 30 seconds. Subsequently, the slides were dehydrated by submerging in 70%, 80%, and 99% ethanol serially. Samples were finally air-dried and stored at RT until imaging. Before imaging, a cover glass (Corning, CLS2980223, #1.5) was mounted with Aqua Poly Mount medium (Polysciences Europe GmbH, 18606-20).

6.3.4 Image analysis and contour export for laser microdissection (LMD)

QuPath (version 0.4.3) was utilized for conducting image analysis. Regions of interest (ROIs) were manually annotated in QuPath following the image analysis process. To ensure accurate contour transfer between the screening and laser microdissection microscopes, three tissue reference points (x-y coordinates) were selected. The contours and reference points were then exported in GeoJSON format and converted into.XML format, which is compatible with Leica LMD7 software.

6.3.5 Laser Microdissection

For laser microdissection–based tissue collection, the Leica LMD7 system equipped with Leica Laser Microdissection Software V 8.3.0.08259 was used. Tissue sections were cut using a 20× objective in either brightfield or fluorescence mode. For the 20× objective (HC

PL FL L 20×/0.40 CORR), the following laser settings were applied: power 60, aperture 15, cutting speed 20, middle pulse count 1, final pulse 0, head current 45-54% (adjusted according to tissue type and section thickness), pulse frequency 2328, and offset 107.

The microdissected contours were collected into a low-binding 384-well plate (Eppendorf 0030129547), configured using the "universal holder" function, with one empty well placed between samples to avoid cross-contamination.

6.3.6 Sample preparation for nanoLC-HRMS analysis

Automated cutting was employed to gather tissue specimens following contour import into 384-well plates (Eppendorf, 0030129547) with low-binding properties. To ensure tissue settled at the well bottoms post-LMD collection, each well received 15 µL of acetonitrile underwent brief vortexing and was vacuum dried (15 min at 60 °C). A subsequent well inspection was conducted prior to proteomics sample preparation to verify successful collection. The DDM-based protocol utilized a lysis buffer comprising 0.025% DDM, 5mM TCEP, 20mM CAA, and 0.1M TEAB in water. A MANTIS Liquid Dispenser with high-volume diaphragm chips was used to add 2 µl of lysis buffer to each sample well. The plate was sealed with PCR ComfortLid and heated at 95 °C for 60 min. After brief cooling, 1 µL of LysC (10 ng/ µL in 0.1M TEAB [pH 8.5] and 30% ACN in milli-Q water) were introduced, followed by digestion for at least 2 h at 37 °C in a thermal cycler (50 °C lid temperature). Next, 1 µL of trypsin (10 ng/µl containing 10% ACN and 0.1M TEAB [pH 8.5] in milli-Q water) was added, with overnight incubation at 37 °C in the thermal cycler. The following day samples underwent vacuum drying before peptide clean-up.

6.3.7 Peptide clean-up with C18 Evotips

Peptide purification was conducted using Evotips (Evosep, Odense, Denmark) following the manufacturer's instructions. The tips were prepared by adding 20 µl of buffer B (99.9% ACN, 0.1% FA) to each C18 tip (EV2013, Evotip Pure, Evosep) followed by centrifugation at 700 rpm for 1 min. Subsequently, 20 µl of buffer A (99.9% water, 0.1% FA) was added to the top of each C18 tip, which was then activated in isopropanol for 20 s and centrifuged again at 700 rpm for 1 min. Peptides were then applied to the Evotips, washed once with 20 µL of buffer A, and finally loaded with 200 µL of buffer A. The tips were also submerged at the bottom in buffer A before initiating the LC-MS analysis.

6.3.8 nanoLC-HRMS Analysis

Liquid chromatography was performed using the Evosep One LC system (Evosep, Odense, Denmark) connected to a trapped ion mobility spectrometer with quadrupole time-of-flight mass spectrometer (timsTOF Ultra 2, Bruker Daltonik, Bremen, Germany) with a nano-electrospray ion source (CaptiveSpray, Bruker Daltonik, Bremen, Germany). Digested peptides were loaded on the Evosep Performance column (EV1137, 150 µm inner diameter, packed with 1.5 µm C18 beads) at 40 °C. Chromatographic separation was performed using an Evosep 30SPD gradient. The solvents utilized were LC-MS-grade water containing 0.1% formic acid (buffer A) and acetonitrile with 0.1% formic acid (buffer B). For the DIA-PASEF analysis, we employed a method comprising eight DIA- PASEF scans divided into three ion mobility windows per scan. This covered a mass-to-charge ratio range of 400–1000 *m/z* using 25 Th windows and an ion mobility range from 0.64 to 1.37 Vs cm^{-2}. The mass spectrometer was run in high-sensitivity mode, with accumulation and ramp time set to 100ms, and capillary voltage at 1750 V. The collision energy was configured as a linear ramp, starting at 20 eV at $1/k_0$ = 0.6Vs cm^{-2} and increasing to 59 eV at $1/k_0$ = 1.6Vs cm^{-2}. This collision energy ramp was applied linearly as a function of ion mobility, decreasing from 59 eV at $1/k_0$ = 1.6Vs cm−2 to 20 eV at $1/k_0$ = 0.6Vs cm^{-2}.

6.3.9 Raw File processing

For the analysis of DIA-PASEF raw files and the generation of spectral libraries, DIA-NN was employed. The human reference proteome (FASTA format) was retrieved from the UniProt database. For the creation of in-silico–predicted spectral libraries, the human FASTA file was supplemented with a set of commonly observed contaminants. Deep-learning–based prediction modules for MS/MS spectra, retention times, and ion mobility values were activated across the 300–1200 m/z range.

N-terminal methionine excision and carbamidomethylation of cysteine residues were defined as fixed modifications. The digestion parameters allowed up to two missed cleavages, and precursor ion charges were restricted to the 2–4 range. DIA-NN was executed in default mode, with minor adjustments optimized for DIA-PASEF data. Specifically, MS1 and MS2 mass accuracies were set to 15.0 ppm, scan windows were left at 0 to enable automatic window assignment, isotopologue detection was enabled, and match-between-runs functionality was activated. Protein inference was conducted using

the heuristic approach, with shared spectra disabled, and grouping based on gene-level identifiers.

The neural-network classifier operated in single-pass mode. Quantification was performed using the "Robust LC (high precision)" strategy. Cross-run normalization was configured using the RT-dependent mode, whereas spectral library construction followed the "smart profiling" procedure. Computational settings for speed and memory usage were set to "optimal results," ensuring high-quality identification and quantitative robustness.

6.4 Results

6.4.1 Clinicopathological characteristics and histological features of DICER1-mutant thyroid tumors

Five pediatric patients (≤18 years) with thyroid tumors harboring pathogenic *DICER1* alterations and displaying high-grade morphological features were included in this study. All cases were classified according to the most recent WHO criteria for thyroid tumors.

Histological examination revealed marked architectural and cytological heterogeneity within individual tumors. Distinct areas with lower-grade follicular architecture coexisted with regions displaying high-grade features, including solid and trabecular growth patterns, increased cellular density, and variable nuclear atypia.

Two morphological patterns previously associated with DICER1-mutated thyroid lesions were systematically assessed.

Atrophic changes were identified within the tumor in three cases. These areas appeared as sharply demarcated regions characterized by loss of viable tumor cells, presence of ghost-like cellular remnants, and deposition of dense hyaline stroma. No accompanying inflammatory infiltrate, apoptotic bodies, or procedure-related artifacts were observed. Atrophic foci were most frequently localized in subcapsular regions but, in some cases, were also detected as multiple intranodular areas, including small nodules distinct from the main tumor mass (**Figure 1A**).

Involutional changes were observed in the adjacent non-neoplastic thyroid parenchyma in two cases. These consisted of clusters of ectatic macrofollicles lined by flattened epithelium, often spatially separated from the invasive tumor front (**Figure 1B**).

Nuclear features varied both between and within cases, with heterogeneous degrees of nuclear enlargement, chromatin clearing, and nucleolar prominence. This morphological heterogeneity provided the basis for spatially resolved downstream analyses.

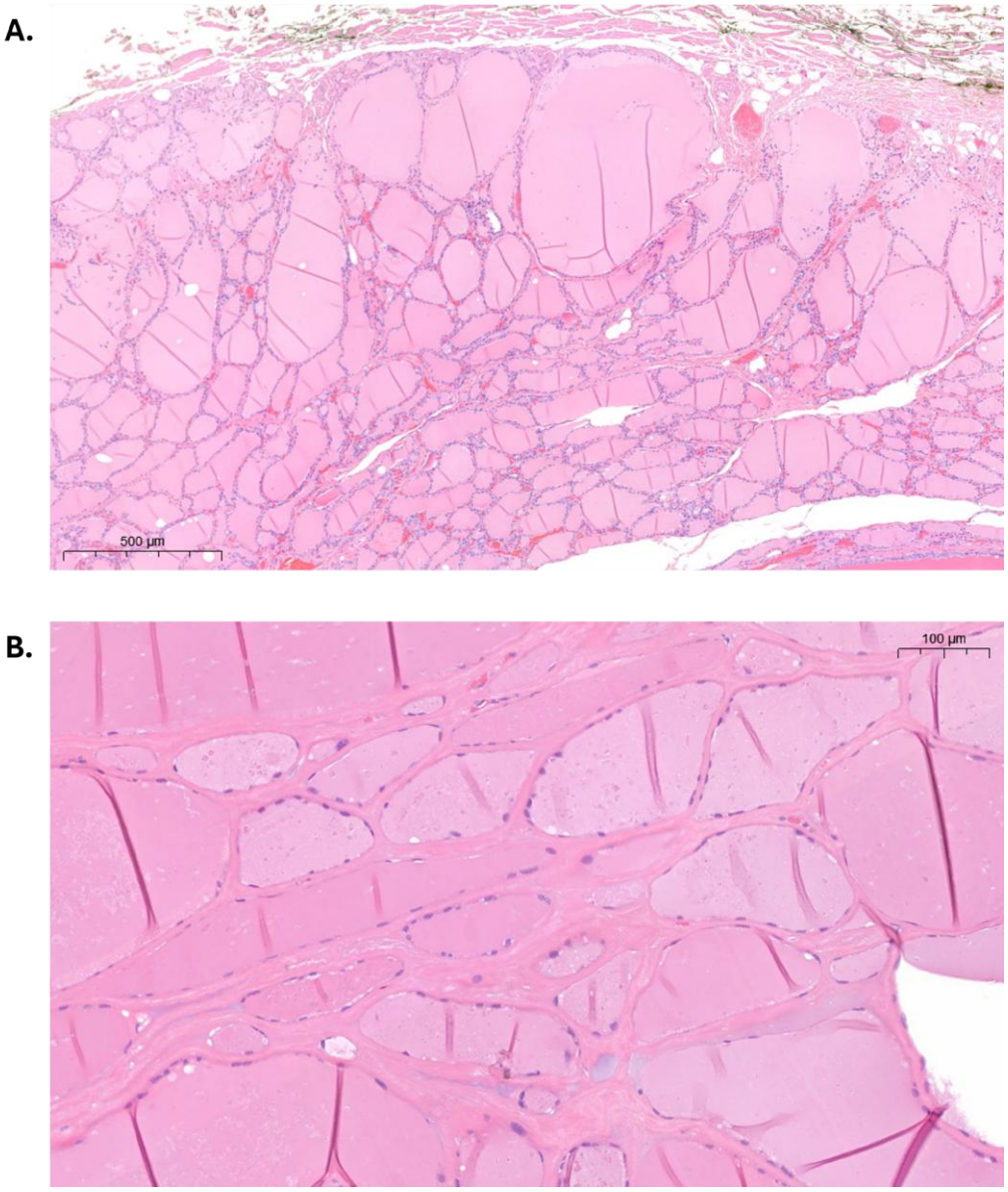


Figure 1 A-B. H&E morphological features associated with atrophic and involutional processes. **a.** Tumor-associated atrophic areas defined by sharply demarcated regions with cell loss, ghost-cell morphology, and hyalinized stroma. **b.** Involutional changes in adjacent thyroid parenchyma, marked by ectatic macrofollicles bordered by flattened epithelial cells.

6.4.2 Cyclic immunofluorescence reveals spatially heterogeneous tumor and microenvironmental features

To preserve tissue architecture and enable spatial annotation, cyclic immunofluorescence (CyIF) was performed on FFPE sections prior to laser microdissection. Markers were selected to characterize tumor cell populations and immune components of the tumor microenvironment.

PAX8 expression, confirming thyroid lineage, was detected in tumor cells across all cases, though its intensity and spatial distribution varied markedly between tumor subregions. Lower-grade areas exhibited strong, homogeneous nuclear PAX8 staining in most tumor cells. Conversely, high-grade regions showed reduced and heterogeneous PAX8 expression, characterized by patchy nuclear staining and focal areas of diminished signal intensity. Atrophic regions largely lacked PAX8 expression, consisting of scarce viable tumor cells and the presence of ghost-like remnants within hyaline stroma (**Figure 2A**).

CD56 expression varied across tumor compartments, with higher intensity observed in several high-grade areas compared to adjacent lower-grade regions, delineating distinct tumor microanatomical niches. In contrast, atrophic regions showed absent or markedly reduced CD56 staining, with no specific signal detected in tumor cells and only occasional weak background fluorescence associated with stromal components or cellular debris (**Figure 2B**).

Macrophage populations were characterized by CD68 and CD163 expressions. CD68-positive macrophages were broadly distributed within both intratumoral and stromal compartments. CD163-positive macrophages displayed a more restricted pattern, often enriched at the interface between viable tumor tissue and atrophic zones, as well as within stroma adjacent to high-grade tumor foci (**Figure 2F**). The density and localization of these populations varied between tumor subregions and cases.

The combined analysis of PAX8, CD56, CD68, and CD163 enabled precise spatial delineation of tumor compartments with distinct differentiation states, and immune microenvironment characteristics. Correlating CyIF data with corresponding H&E sections facilitated accurate annotation of proliferative, poorly differentiated high-grade foci; less proliferative, more differentiated tumor regions; and non-viable atrophic areas. These

spatially resolved annotations guided laser microdissection for region-specific proteomic profiling.

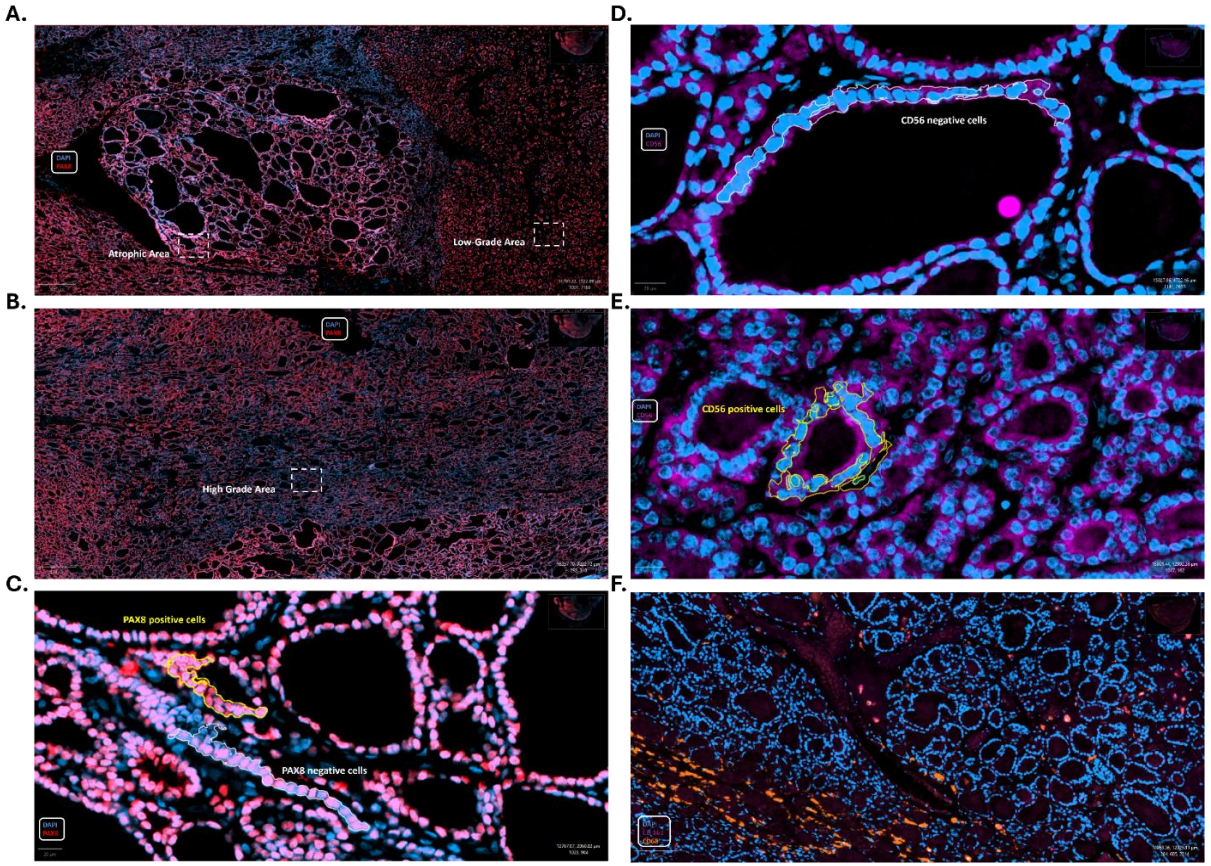


Figure 2 A-F. (**A**) PAX8 staining illustrating heterogeneous thyroid-lineage marker expression, with strong nuclear positivity in lower-grade areas and reduced, patchy expression in high-grade regions (**B**); atrophic foci show near-complete loss of signal (**A**). (**C**) PAX8 positive and negative cells. (**D**) CD56 expression across tumor compartments, with higher intensity in selected high-grade regions (**E**) and markedly reduced or absent signal in atrophic areas. (**F**) CD68-positive macrophages broadly distributed in both intratumoral and stromal compartments. CD163-positive macrophages enriched at the interface between viable tumor and atrophic zones and in stroma adjacent to high-grade foci.

6.4.4 Analytical performance and data quality of the DIA-PASEF workflow

Proteomic analyses were performed using an Evosep One LC system with a 30 samples-per-day (30SPD) method coupled to a timsTOF Ultra 2 mass spectrometer operating in DIA-PASEF mode (8 scans × 3 ion mobility windows).

Chromatographic performance was highly stable across all runs, with a median full width at half maximum (FWHM) of 5.1 seconds and a retention time coefficient of variation (RT CV) of 1.1%. Mass spectrometric acquisition showed high sampling efficiency, reflected by consistent identification rates across injections.

Across all regions of interest (ROIs), between 2300 and 4800 protein groups were identified per ROI (median 3300), corresponding to 12,000–26,000 peptides. Protein-level coefficients of variation for technical replicates ranged between 10–15%, indicating robust quantitative reproducibility. Missed cleavage rates were between 12–16%, and peptide identification rates per injection ranged from 88–93%.

These metrics demonstrate high analytical performance and reproducibility of the mDVP workflow when applied to ultra-low-input FFPE thyroid tissue.

Proteome coverage differed across morphologically defined tissue compartments. High-grade tumor regions yielded significantly deeper proteomic coverage compared with lower-grade areas, with a median increase of 350 protein groups per ROI (Wilcoxon test, $p = 0.006$).

Atrophic and involutional regions yielded reproducible but comparatively reduced proteome depth, consistent with differences in tissue composition and cellularity.

6.4.5 Global proteomic stratification of atrophic, involutional, and high-grade regions

Label-free quantification based on DIA data identified 3256 proteins that were significantly differentially abundant across the three tissue conditions, atrophic, involutional, and high-grade at a false discovery rate (FDR) < 0.05.

Unsupervised hierarchical clustering revealed a clear stratification of samples according to tissue lesion type. High-grade, involutional, and atrophic samples clustered into distinct groups, indicating the presence of lesion-specific proteomic signatures (**Figure 3A**).

To further assess global proteomic variability, principal component analysis (PCA) was performed. The first principal component (PC1), accounting for 41.69% of the total variance, separated high-grade from atrophic samples. The second principal component (PC2), explaining 12.37% of the variance, captured intra-class variability within each lesion type (**Figure 3B**).

To ascertain the molecular characteristics most influential in class discrimination, a partial least squares discriminant analysis (PLS-DA) model was employed. Variable Importance in Projection (VIP) scores were computed for all quantified proteins. Proteins with the highest VIP scores were identified as the most significant contributors to the differentiation among atrophic, involutional, and high-grade tissue classes, thereby delineating a subset of molecular features associated with lesion-specific proteomic profiles (**Figure 3C, Figure 4**).

Functional enrichment analysis of spatial proteomic signatures revealed 19 significantly dysregulated functional categories across atrophic, involutional, and high-grade regions (adjusted $p < 0.05$). These categories were organized into four principal functional networks: proteostasis and protein quality control, RNA processing, cellular metabolism, and drug resistance–associated signaling pathways (**Figure 3D**). Within the proteostasis network, several proteasome-related categories were significantly enriched, including proteasome-activating activity, cytosolic proteasome complex, proteasome, and proteasome binding.

Additionally, components of the calpain complex were significantly enriched. Proteins associated with localization to cytoplasmic stress granules were also enriched across lesion types. Within the RNA processing network, spliceosome-related pathways were significantly represented, including U4 snRNP and general spliceosome components. The 7-methylguanosine cap hypermethylation pathway was also enriched. Metabolic pathway enrichment encompassed processes related to carbohydrate metabolism, such as GDP-mannose synthesis, GDP-mannose metabolic process, and fructose and mannose metabolism (adjusted $p < 0.05$). Furthermore, pathways related to ERBB2 tyrosine kinase signaling were significantly enriched, including categories associated with resistance mechanisms to ERBB2-targeted therapies.

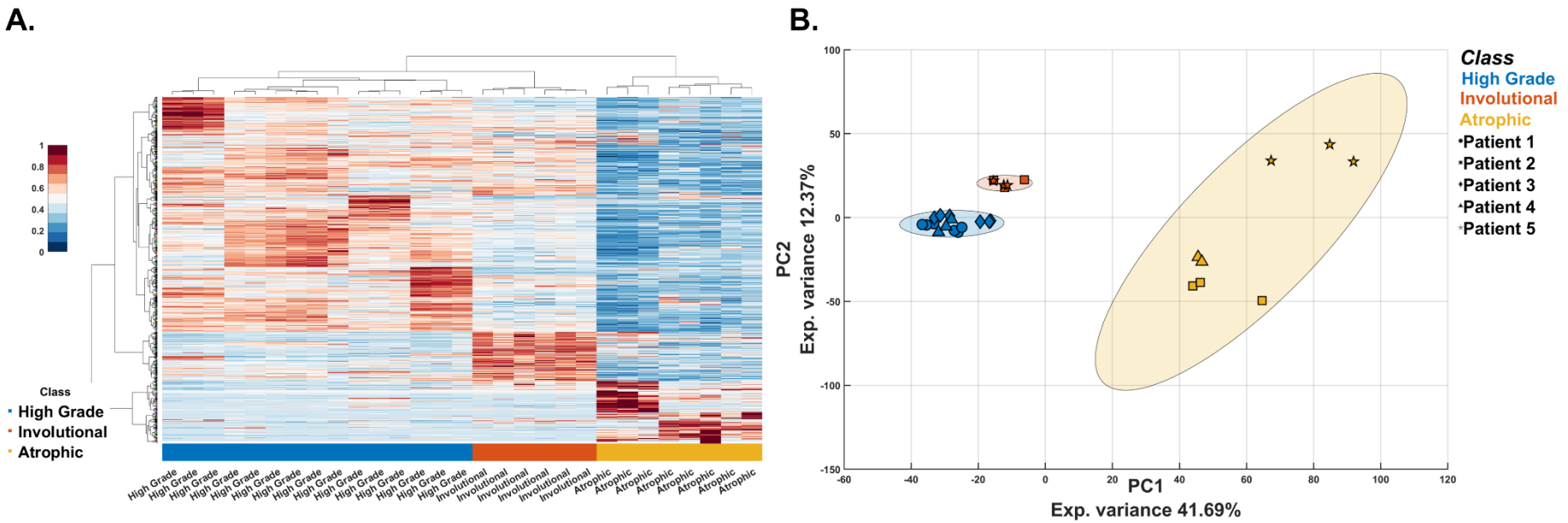


Figure 3A-B. (**A**) Hierarchical clustering segregating samples into distinct high-grade, involutional, and atrophic groups. (**B**) PCA representation of proteomic variability, with PC1 separating high-grade and atrophic lesions and PC2 reflecting intra-group diversity.

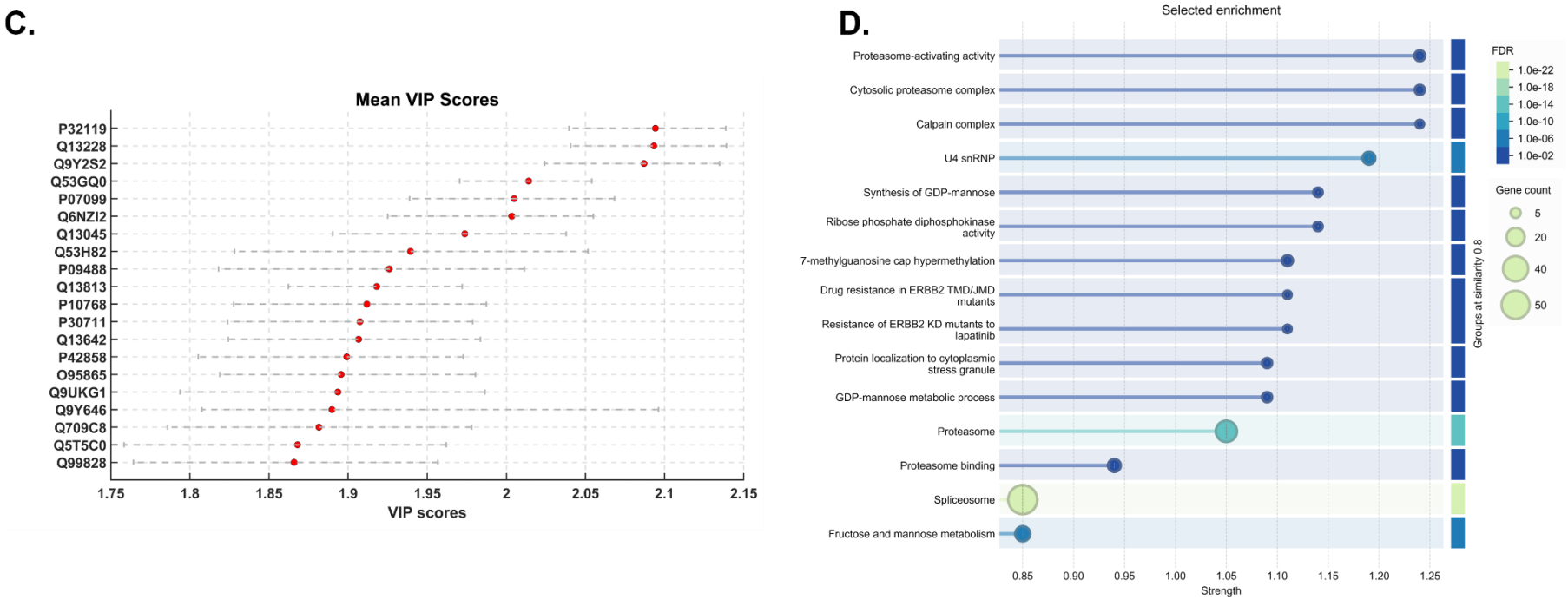


Figure 3C-D. (**C**) VIP score distribution from the PLS-DA model, indicating the proteins that most strongly contribute to class discrimination among atrophic, involutional, and high-grade tissue areas. (**D**) Network representation of 19 significantly enriched functional categories (adjusted $p < 0.05$), encompassing proteostasis-related processes (including proteasome and calpain pathways), RNA-processing mechanisms (such as spliceosome components and cap hypermethylation), metabolic pathways related to GDP-mannose and carbohydrate metabolism, and ERBB2-associated signaling modules.

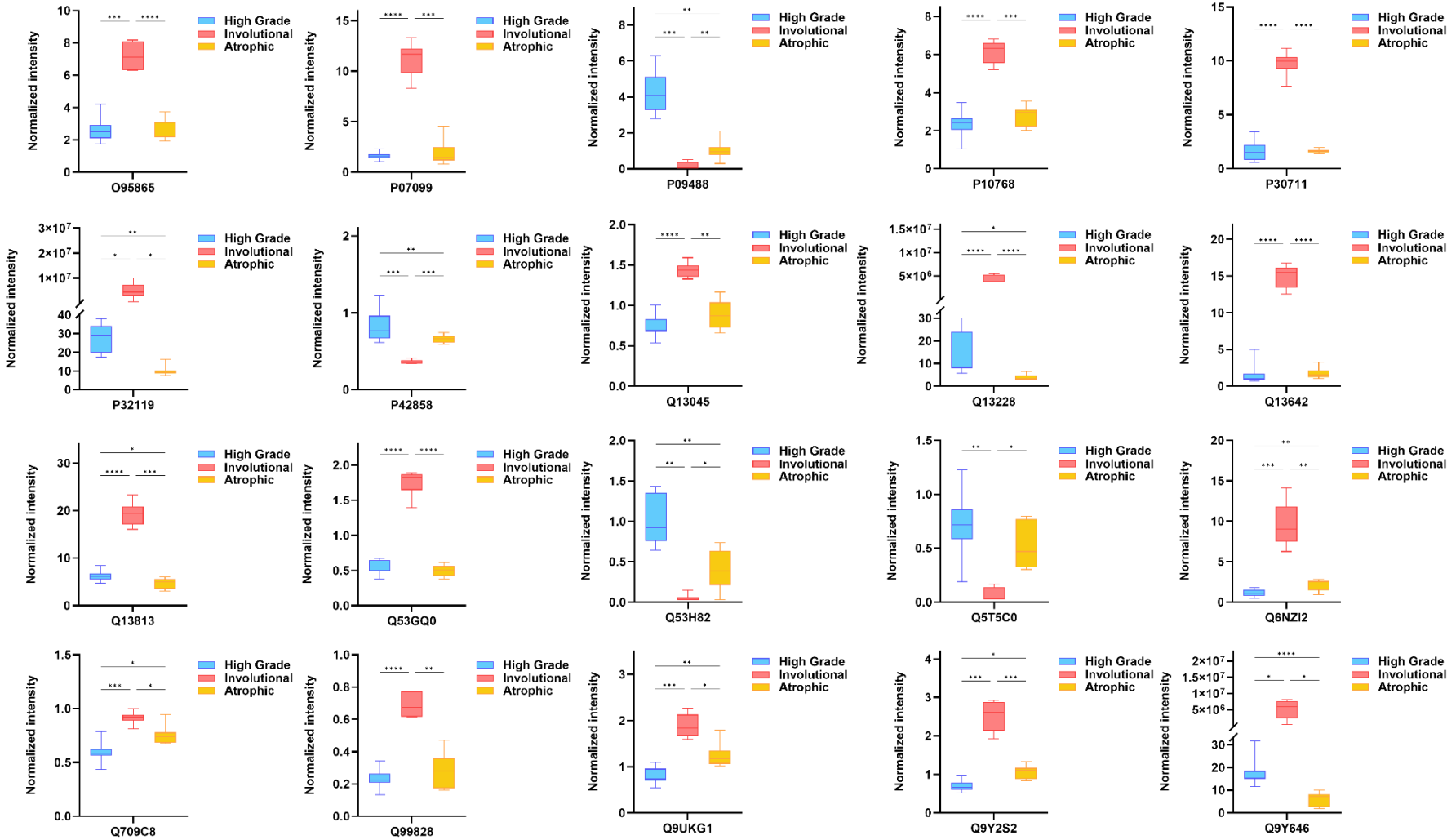


Figure 4. Boxplots of the top discriminant proteins identified through PLS-DA VIP scores, illustrating differential expression patterns across atrophic, involutional, and high-grade tumor areas, *p-value < 0.05, **p-value < 0.01 and ***p-value < 0.001.

6.5 Discussions

This investigation utilized a spatially resolved deep visual proteomics (mDVP) (5) method on DICER1-mutant pediatric thyroid tumors to dissect distinct tumor compartments under ultra-low input conditions. By integrating laser microdissection with DIA-PASEF-based proteomics, the study achieved robust profiling of the proteome in high-grade, involutional, and atrophic regions while maintaining morphological context and spatial resolution. This novel approach addresses a significant technical challenge in studying rare pediatric thyroid tumors, where limited tissue availability and intratumoral heterogeneity have hindered molecular investigations historically.

The spatial characterization of tumor heterogeneity revealed distinct microanatomical niches within thyroid tumors that reflect varying degrees of differentiation and cellular viability. PAX8 expression patterns served as a reliable marker of thyroid lineage and tumor cell identity, demonstrating heterogeneous distribution across tumor compartments with strong, homogeneous nuclear staining in lower-grade regions but reduced and patchy expression in high-grade areas (12).

The variable intensity of CD56 expression further delineated these compartments, with higher levels in select high-grade regions compared to adjacent lower-grade areas, while atrophic zones showed marked reduction or absence of signal (13). Macrophage populations displayed distinct localization patterns, with CD68-positive cells broadly distributed throughout intratumoral and stromal regions, whereas CD163-positive macrophages were preferentially enriched at the interface between viable tumor tissue and atrophic zones (14).

The combined multimarker analysis effectively stratified tumor regions into functionally distinct compartments characterized by different differentiation states and immune microenvironment signatures, which subsequently enabled precise region-specific proteomic profiling through laser microdissection (15). These findings demonstrate that spatial immunofluorescence profiling can effectively map tumor heterogeneity at the microanatomical level, providing a foundation for understanding how differentiation states and immune infiltration vary within the tumor microenvironment.

Unsupervised clustering and multivariate analyses revealed a clear stratification of samples based on lesion type, indicating that the morphological diversity in DICER1-mutant tumors is rooted in distinct proteomic states rather than purely histological differences. The segregation of high-grade, involutional, and atrophic regions along principal components suggests that these compartments represent discrete biological programs, likely corresponding to various stages of tumor evolution, differentiation status, and interactions within the microenvironment.

Proteomic analyses demonstrated that high-grade regions exhibit enhanced proteome coverage and unique molecular profiles compared to lower-grade and involutional areas. The heightened proteome depth in these regions likely mirrors increased cellular density, metabolic activity, and transcriptional output consistent with their aggressive histological characteristics (16).

Functional enrichment analysis indicated that high-grade regions are characterized by the upregulation of pathways involved in proteostasis, RNA processing, and cellular metabolism. Notably, the enrichment of proteasome-related categories and calpain complex components suggests an augmented reliance on protein quality control mechanisms and regulated proteolysis in high-grade tumor biology. These findings highlight the adoption of robust protein homeostasis strategies by high-grade DICER1-mutant thyroid tumors to sustain malignant progression.

DICER1 is a central component of the microRNA (miRNA) biogenesis pathway, functioning as a key RNase III endoribonuclease required for the processing of precursor miRNAs into mature, functional species (1). Pathogenic alterations in DICER1, particularly hotspot mutations affecting the RNase IIIb domain, represent a distinctive oncogenic mechanism in thyroid neoplasia and confer a molecular identity that is clearly separable from that driven by more common alterations such as BRAF or RET. These mutations result in a characteristic distortion of the miRNA landscape, marked by a global reduction in 5p-derived miRNAs accompanied by paradoxical accumulation of 3p miRNAs. The resulting imbalance profoundly reshapes post-transcriptional gene regulation and creates a permissive environment for tumor initiation and progression (17).

Despite sharing RAS-like transcriptomic features, DICER1-mutated thyroid lesions exhibit a strikingly broad clinicopathological spectrum, ranging from benign multinodular

goiter to poorly differentiated and high-grade carcinomas. This heterogeneity strongly suggests that DICER1 mutations alone are insufficient to dictate tumor aggressiveness, and that additional molecular and cellular adaptations are required to drive progression. The present enrichment analysis provides critical insight into these downstream processes, revealing a convergence of proteostatic, metabolic, and RNA-processing dysfunctions that collectively define high-grade disease (9,18,19).

One of the most prominent features emerging from the enrichment analysis is the strong overrepresentation of proteasome-related pathways, including proteasome-activating activity, cytosolic proteasome complexes, and proteasome binding functions. This finding points to a fundamental disruption of the cellular proteostasis network in DICER1-mutated tumors. The ubiquitin–proteasome system (UPS) is the primary mechanism for maintaining protein quality control, and its proper function is essential for preventing the accumulation of misfolded or damaged proteins (20-23).

Cancer cells are intrinsically dependent on robust proteostasis mechanisms due to their elevated rates of protein synthesis, rapid proliferation, and heightened metabolic demands. In this context, the pronounced enrichment of proteasome-associated pathways suggests that DICER1-mutated lesions experience an exceptional degree of proteotoxic stress. Rather than reflecting simple proteasome hyperactivity, these signatures are more consistent with a compensatory cellular response aimed at counteracting proteostasis collapse. Such a response is particularly relevant during progression from low-grade or involutional lesions to high-grade tumors, where increasing cellular density and biosynthetic activity amplify proteotoxic burden (24,25).

Mechanistically, proteasome dysfunction in this setting may be directly linked to defective miRNA processing. DICER1-dependent miRNAs regulate numerous transcripts encoding proteasome subunits, ubiquitin ligases, and regulatory cofactors. Disruption of this regulatory layer is therefore likely to impair coordinated proteasome assembly and function. The resulting accumulation of misfolded proteins further exacerbates cellular stress, creating a self-reinforcing cycle that drives selection for cells capable of activating alternative survival pathways (26–28).

In parallel with proteasome-related pathways, enrichment of calpain complex components highlights an additional axis of proteolytic dysregulation. Calpains are calcium-

dependent cysteine proteases that mediate limited, non-ubiquitin-dependent proteolysis and play critical roles in cytoskeletal remodeling, apoptosis regulation, and stress responses. Their activation is tightly controlled under physiological conditions but becomes dysregulated in states of sustained cellular stress (29,30).

The prominence of calpain-related pathways in DICER1-mutated high-grade lesions likely reflect elevated intracellular calcium levels associated with oncogenic transformation, mitochondrial dysfunction, and endoplasmic reticulum (ER) stress. Calpain overactivation has been documented in thyroid cancer and has been linked to altered cell survival, migration, and therapeutic resistance. Importantly, calpains target a wide range of substrates involved in DNA repair, mitochondrial integrity, and cytoskeletal dynamics, all of which contribute to the aggressive phenotype of high-grade tumors (31,32).

Taken together, the concurrent enrichment of proteasome and calpain pathways indicates that proteolytic imbalance in DICER1-mutated lesions is not confined to a single degradation system but instead reflects a global disruption of protein turnover mechanisms.

Another key feature of the enrichment analysis is the overrepresentation of spliceosome-related components, including U4 snRNP and core splicing machinery. Although DICER1 is classically defined by its role in miRNA biogenesis, accumulating evidence suggests that it participates more broadly in RNA quality control networks. The observed spliceosomal enrichment likely reflects indirect consequences of miRNA dysregulation, particularly through the loss of miRNAs that normally regulate splicing factor expression (33).

Defective splicing has emerged as a hallmark of cancer, capable of generating aberrant transcripts that encode dysfunctional or unstable proteins. In DICER1-mutated thyroid lesions, such splicing defects would further burden an already stressed proteostasis network by increasing the load of misfolded or truncated proteins. The enrichment of pathways related to 7-methylguanosine cap hypermethylation reinforces this interpretation, as cap modifications are essential for proper splicing, nuclear export, and mRNA stability (34,35).

The enrichment of pathways associated with protein localization to cytoplasmic stress granules provides further evidence that DICER1-mutated thyroid cells exist in a state of chronic cellular stress. Stress granules are dynamic ribonucleoprotein assemblies that form in response to translational arrest and serve to temporarily sequester mRNAs, thereby reducing protein synthesis and limiting proteotoxic damage (36).

Beyond their protective role, stress granules are increasingly recognized as active contributors to malignant phenotypes. By selectively regulating the translation of survival- and stress-related transcripts, they enable cancer cells to adapt to hostile microenvironments. In the context of DICER1-mutated lesions, stress granule dynamics may represent a critical adaptive mechanism that allows certain clones to survive proteostatic crisis and progress toward higher-grade disease (37).

Metabolic pathway enrichment, particularly involving GDP-mannose synthesis and fructose and mannose metabolism, reveals an additional layer of adaptation (38). These pathways are central to N-glycan biosynthesis, which is essential for protein folding, trafficking, and receptor signaling. Alterations in glycosylation are a well-established feature of cancer and can profoundly influence cell-cell interactions, immune evasion, and growth factor signaling (39).

Disruption of mannose metabolism is also tightly linked to ER function and proteostasis. Impaired N-glycan assembly can trigger the unfolded protein response, further amplifying ER stress and proteotoxic signaling. Thus, the metabolic rewiring observed in DICER1-mutated lesions likely intersects directly with proteostasis dysfunction, creating a coupled metabolic-proteostatic stress state that shapes tumor evolution (40).

The enrichment of ERBB2-related drug resistance pathways introduces a signaling dimension to this model. Although ERBB2 alterations are relatively uncommon in thyroid cancer, activation of ERBB2-dependent signaling may represent a secondary adaptive mechanism in DICER1-mutated lesions. ERBB2 signaling engages PI3K/AKT and MAPK pathways, promotes protein synthesis through mTOR activation, and supports metabolic reprogramming toward anabolic growth (41).

The association with lapatinib resistance pathways suggests that these tumors may harbor intrinsic or acquired mechanisms that blunt therapeutic targeting of receptor

tyrosine kinases. Importantly, ERBB2 signaling intersects with both proteostasis and metabolic networks, reinforcing the interconnected nature of the enriched pathways identified in this study (42).

Collectively, these findings support a unifying model in which DICER1 mutations initiate a cascade of molecular disruptions extending far beyond miRNA biogenesis. Loss of tumor-suppressive miRNAs destabilizes proteostasis, alters RNA processing, and rewires cellular metabolism. This creates a state of chronic cellular crisis characterized by proteotoxic stress, ER dysfunction, and metabolic imbalance.

Progression from benign or involutional lesions to high-grade carcinoma appears to depend on the ability of tumor cells to adapt to this crisis. Cells that successfully engage compensatory mechanisms, such as enhanced proteasome activity, calpain-mediated remodeling, stress granule formation, and activation of alternative survival signaling pathways-gain a selective advantage and progress toward more aggressive phenotypes. Conversely, cells that fail to adapt may undergo cell death or remain arrested at lower grades.

In this context, enrichment analysis captures not merely static molecular alterations but dynamic cellular processes that reflect an ongoing struggle between proteostatic collapse and adaptive survival. These insights highlight proteostasis as a central vulnerability in DICER1-mutated thyroid tumors and suggest that therapeutic strategies targeting protein quality control, stress response pathways, or metabolic–proteostatic coupling may hold promise for this aggressive disease subtype.

6.6 Conclusions

In conclusion, this study demonstrates that DICER1-mutated pediatric thyroid tumors are shaped not merely by aberrant miRNA biogenesis, but by a complex and spatially organized network of adaptive cellular responses that govern tumor evolution. Through the application of spatially resolved deep visual proteomics under ultra-low-input conditions, this work reveals that morphologically distinct tumor compartments correspond to discrete proteomic states, each defined by differential engagement of proteostasis mechanisms, RNA processing pathways, metabolic rewiring, and survival signaling networks.

High-grade regions emerge as focal points of intensified cellular stress, characterized by activation of proteasome and calpain pathways, perturbations in splicing and stress-granule dynamics, and metabolic adaptations linked to glycosylation and endoplasmic reticulum function. These alterations are accompanied by compensatory activation of pro-growth and resistance-related signaling modules, including ERBB2-associated pathways, underscoring the multifaceted nature of stress adaptation in aggressive DICER1-mutant disease.

Collectively, these findings support a model in which progression from low-grade or involutional lesions to high-grade carcinoma depends on the ability of tumor cells to withstand and adapt to chronic proteostatic and metabolic crises. Clones that successfully mobilize compensatory pathways gain a selective advantage and drive malignant progression, whereas cells unable to resolve these stresses remain arrested or eliminated. This conceptual framework positions proteostasis disruption as a central vulnerability in DICER1-mutated thyroid tumors.

Beyond advancing the biological understanding of this rare entity, the thesis underscores the unique value of spatial proteomics in disentangling tumor heterogeneity, resolving lineage and differentiation states, and reconstructing evolutionary trajectories that are otherwise concealed in bulk analyses. Altogether, these insights provide a mechanistic basis for how miRNA dysregulation translates into aggressive tumor behavior and highlight protein-quality-control and stress-adaptive pathways as promising therapeutic entry points for this clinically challenging disease.

Bibliography Chapter VI

[1]. Ricarte-Filho JC, Casado-Medrano V, Reichenberger E, Spangler Z, Scheerer M, Isaza A, et al. DICER1 RNase IIIb domain mutations trigger widespread miRNA dysregulation and MAPK activation in pediatric thyroid cancer. Front Endocrinol. 2023 Feb 21;14:1083382.

[2]. Lee YA, Im SW, Jung KC, Chung EJ, Shin CH, Kim JI, et al. Predominant DICER1 Pathogenic Variants in Pediatric Follicular Thyroid Carcinomas. Thyroid. 2020 Aug 1;30(8):1120–31.

[3]. Malagobadan S, Shi C, Yang A, Mondal I, Baldeh H, Spain K, et al. DICER1 hotspot mutation induces 3p microRNA gain of function via Argonaute strand switch. Nat Struct Mol Biol. 2025 Dec;32(12):2542–52.

[4]. Khan NE, Bauer AJ, Schultz KAP, Doros L, Decastro RM, Ling A, et al. Quantification of Thyroid Cancer and Multinodular Goiter Risk in the DICER1 Syndrome: A Family-Based Cohort Study. J Clin Endocrinol Metab. 2017 May 1;102(5):1614–22.

[5]. Mund A, Coscia F, Kriston A, Hollandi R, Kovács F, Brunner AD, et al. Deep Visual Proteomics defines single-cell identity and heterogeneity. Nat Biotechnol. 2022 Aug;40(8):1231–40.

[6]. WHO Classification of Tumours Editorial Board. Endocrine and neuroendocrine tumours. Lyon (France): International Agency for Research on Cancer; 2022. (WHO classification of tumours series, 5th ed.; vol. 10.

[7]. Khan NE, Bauer AJ, Schultz KAP, Doros L, Decastro RM, Ling A, Lodish MB, Harney LA, Kase RG, Carr AG, Rossi CT, Field A, Harris AK, Williams GM, Dehner LP, Messinger YH, Hill DA, Stewart DR. Quantification of Thyroid Cancer and Multinodular Goiter Risk in the DICER1 Syndrome: A Family-Based Cohort Study. J Clin Endocrinol Metab. 2017 May 1;102(5):1614-1622.

[8]. Juhlin CC, Stenman A, Zedenius J. Macrofollicular variant follicular thyroid tumors are DICER1 mutated and exhibit distinct histological features. Histopathology. 2021 Oct;79(4):661-666.

[9]. Condello V, Poma AM, Macerola E, Vignali P, Paulsson JO, Zedenius J, Basolo F, Juhlin CC. Prevalence, Molecular Landscape, and Clinical Impact of DICER1 and DGCR8 Mutated Follicular-Patterned Thyroid Nodules. J Clin Endocrinol Metab. 2024 Jun 17;109(7):1733-1744. https://doi.org/10.1210/clinem/dgae034.

[10]. Nosé V, Gill A, Teijeiro JMC, Perren A, Erickson L. Overview of the 2022 WHO Classification of Familial Endocrine Tumor Syndromes. Endocr Pathol [Internet]. 2022;33.

[11]. Wasserman JD, Sabbaghian N, Fahiminiya S, Chami R, Mete O, Acker M, Wu MK, Shlien A, de Kock L, Foulkes WD. DICER1 Mutations Are Frequent in Adolescent-Onset Papillary Thyroid Carcinoma. J Clin Endocrinol Metab. 2018 May 1;103(5):2009–2015.

https://doi.org/10.1210/jc.2017-02698. Erratum in: J Clin Endocrinol Metab. 2018 Aug 1;103(8):3114. https://doi.org/10.1210/jc.2018-01190.

[12]. Kang N, Zhao Z, Wang Z, Ning J, Wang H, Zhang W, et al. METTL3 regulates thyroid cancer differentiation and chemosensitivity by modulating PAX8. Int J Biol Sci. 2024;20(9):3426–41.

[13]. Cho U, Kim Y, Jeon S, Jung CK. CD56 Expression in Papillary Thyroid Carcinoma Is Highly Dependent on the Histologic Subtype: A Potential Diagnostic Pitfall. Appl Immunohistochem Mol Morphol. 2022 May;30(5):389–96.

[14]. Kim BH. The Expression of Tumor-Associated Macrophages in Papillary Thyroid Carcinoma. Endocrinol Metab. 2013;28(3):178.

[15]. Liotta LA, Conrads TP, Howard MA, Araujo RP, Petricoin EF. Laser capture microdissection. Nat Rev Methods Primer. 2025 Dec 24;5(1):80.

[16]. Cochrane DR, Negri GL, Huvila J, Kalantari F, Farnell DA, Mohammad N, et al. Proteomic analysis uncovers biological diversity in molecularly defined endometrial carcinomas. Neoplasia. 2025 Nov;69:101229.

[17]. Hata A, Kashima R. Dysregulation of microRNA biogenesis machinery in cancer. Crit Rev Biochem Mol Biol. 2016;51(3):121–34.

[18]. Shi C, Mu Z, Guo W, Zhang X, Sun D, Sun Y, et al. Distinctive role of DICER1 mutations in distant metastatic thyroid cancer. Chin J Cancer Res. 2024;36(6):700–12.

[19]. Minna E, Devecchi A, Pistore F, Paolini B, Mauro G, Penso DA, et al. Genomic and transcriptomic analyses of thyroid cancers identify DICER1 somatic mutations in adult follicular-patterned RAS-like tumors. Front Endocrinol. 2023 Oct 5;14:1267499.

[20]. Kandel R, Jung J, Neal S. Proteotoxic stress and the ubiquitin proteasome system. Semin Cell Dev Biol. 2024 Mar;156:107–20.

[21]. Zhang C, Li J, Tang Q, Li L, Cao D. Targeting proteostasis for cancer therapy: current advances, challenges, and future perspectives. Mol Cancer. 2025 Oct 21;24(1):265.

[22]. Lei L, Wu Z, Winklhofer KF. Protein quality control by the proteasome and autophagy: A regulatory role of ubiquitin and liquid-liquid phase separation. Matrix Biol. 2021 Jun;100–101:9–22.

[23]. Xu S, Gierisch ME, Schellhaus AK, Poser I, Alberti S, Salomons FA, et al. Cytosolic stress granules relieve the ubiquitin-proteasome system in the nuclear compartment. EMBO J. 2023 Feb;42(3):e111802.

[24]. Wu Y, Luo H, Pan Z, Chen W, Bi L. Bidirectional crosstalk between ER stress and lipid metabolism: From proteostasis to tumor adaptation. Cell Death Discov. 2025 Dec 5;12(1):37.

[25]. Ho Zhi Guang M, Kavanagh E, Dunne L, Dowling P, Zhang L, Lindsay S, et al. Targeting Proteotoxic Stress in Cancer: A Review of the Role that Protein Quality Control Pathways Play in Oncogenesis. Cancers. 2019 Jan 9;11(1):66.

[26]. Qu J, Lin Z. Autophagy Regulation by Crosstalk between miRNAs and Ubiquitination System. Int J Mol Sci. 2021 Nov 3;22(21):11912.

[27]. Carbonell T, Gomes AV. MicroRNAs in the regulation of cellular redox status and its implications in myocardial ischemia-reperfusion injury. Redox Biol. 2020 Sep;36:101607.

[28]. Akhlaghipour I, Moghbeli M. MicroRNAs as the critical regulators of ubiquitin-proteasome system through F-box protein targeting during tumor progression. Cancer Cell Int. 2025 Dec 24;25(1):440.

[29]. Shapovalov I, Harper D, Greer PA. Calpain as a therapeutic target in cancer. Expert Opin Ther Targets. 2022 Mar 4;26(3):217–31.

[30]. Nian H, Ma B. Calpain–calpastatin system and cancer progression. Biol Rev. 2021 Jun;96(3):961–75.

[31]. Samanta K, Ahel I, Kar P. Deciphering of the reactive oxygen species (ROS) induced calpain activation in cancer progression and its therapeutic potential. Adv Redox Res. 2025 Jun;15:100124.

[32]. Baek KH, Yu HV, Kim E, Na Y, Kwon Y. Calcium influx-mediated translocation of m -calpain induces Ku80 cleavage and enhances the Ku80-related DNA repair pathway. Oncotarget. 2016 May 24;7(21):30831–44.

[33]. Liu Y, Liu X, Lin C, Jia X, Zhu H, Song J, et al. Noncoding RNAs regulate alternative splicing in Cancer. J Exp Clin Cancer Res. 2021 Jan 6;40(1):11.

[34]. Niño C, Scotto Di Perrotolo R, Polo S. Recurrent Spliceosome Mutations in Cancer: Mechanisms and Consequences of Aberrant Splice Site Selection. Cancers. 2022 Jan 7;14(2):281.

[35]. Bradley RK, Anczuków O. RNA splicing dysregulation and the hallmarks of cancer. Nat Rev Cancer. 2023 Mar;23(3):135–55.

[36]. Li T, Zeng Z, Fan C, Xiong W. Role of stress granules in tumorigenesis and cancer therapy. Biochim Biophys Acta BBA - Rev Cancer. 2023 Nov;1878(6):189006.

[37]. Jia Y, Jia R, Dai Z, Zhou J, Ruan J, Chng W, et al. Stress granules in cancer: Adaptive dynamics and therapeutic implications. iScience. 2024 Aug;27(8):110359.

[38]. Lieu EL, Kelekar N, Bhalla P, Kim J. Fructose and Mannose in Inborn Errors of Metabolism and Cancer. Metabolites. 2021 Jul 25;11(8):479.

[39]. Xu X, Peng Q, Jiang X, Tan S, Yang W, Han Y, et al. Altered glycosylation in cancer: molecular functions and therapeutic potential. Cancer Commun. 2024 Nov;44(11):1316–36.

[40]. Ghemrawi R, Kremesh S, Mousa WK, Khair M. The Role of ER Stress and the Unfolded Protein Response in Cancer. Cancer Genomics Proteomics. 2025;22(3):363–81.

[41]. Pan L, Li J, Xu Q, Gao Z, Yang M, Wu X, et al. HER2/PI3K/AKT pathway in HER2-positive breast cancer: A review. Medicine (Baltimore). 2024 Jun 14;103(24):e38508.

[42]. Stuhlmiller TJ, Miller SM, Zawistowski JS, Nakamura K, Beltran AS, Duncan JS, et al. Inhibition of Lapatinib-Induced Kinome Reprogramming in ERBB2-Positive Breast Cancer by Targeting BET Family Bromodomains. Cell Rep. 2015 Apr;11(3):390–404.

Chapter VII

Conclusions

The PhD project aimed to explore and implement advanced mass spectrometry-based multi-omics strategies to investigate disease-related molecular alterations, with a particular focus on metabolomics. The overarching goal of this research was to develop robust analytical workflows capable of supporting precision medicine approaches by identifying novel biomarkers and improving the understanding of disease pathophysiology, ultimately contributing to more effective diagnostic and therapeutic strategies.

The identification of reliable biomarkers remains a critical challenge in biomedical research, especially for complex and heterogeneous diseases such as cancer and rare genetic disorders. In this context, a major part of this PhD project was devoted to the development and optimization of high-resolution LC-MS methodologies to expand metabolome coverage while ensuring analytical robustness and reproducibility.

The first part of the project focused on improving chromatographic performance in untargeted metabolomics through the integration of orthogonal separation techniques. A systematic comparison of reversed-phase and HILIC chromatographic conditions demonstrated that combining complementary separation modes significantly enhances metabolome coverage and peak quality. The results highlighted the importance of carefully optimizing stationary phases and mobile phase additives, showing that the BEH Amide HILIC column coupled with HSS T3 reversed-phase chromatography provided a robust and transferable analytical platform. This methodological framework contributes to improving data quality and reliability in large-scale metabolomics studies and represents a valuable tool for biomarker discovery (**Chapter II**).

Building upon this optimized analytical platform, the project was subsequently applied to clinically relevant case studies aimed at identifying metabolic and lipidomic signatures associated with hepatocellular carcinoma (HCC). Through an integrated untargeted metabolomics and lipidomics approach, distinct molecular alterations were identified in HCC patients compared with individuals affected by chronic hepatitis C virus infection and

mixed cryoglobulinemia. The results revealed significant dysregulation of mitochondrial fatty acid metabolism, amino acid turnover, and phospholipid remodeling. Importantly, multivariate models demonstrated that the identified multi-omics signatures outperformed conventional biomarkers such as alpha-fetoprotein, particularly in AFP-negative patients. These findings underline the potential of metabolomics and lipidomics as powerful tools for improving early disease detection and patient stratification (**Chapter III**).

To facilitate the translation of discovery-based findings into clinically applicable assays, a targeted metabolomics workflow was developed to quantify selected metabolite classes identified during the untargeted phase, including acylcarnitines and lysophosphatidylcholines. The optimized targeted HILIC-HRMS method demonstrated excellent analytical performance in terms of sensitivity, repeatability, and accuracy, while maintaining high throughput. This integrated discovery-to-validation strategy represents a practical pipeline for the implementation of metabolomics-based biomarkers in clinical research (**Chapter IV**).

Recognizing the growing need for enhanced analytical sensitivity and the increasing use of limited or precious biological samples, a further objective of this PhD project was to investigate the impact of chromatographic miniaturization on metabolomics workflows. The comparison between conventional narrow-bore columns and microbore UHPLC configurations demonstrated that reduced column internal diameter significantly improves detection sensitivity and metabolome coverage without compromising robustness or throughput. These findings support the adoption of micro-flow LC-MS platforms for high-sensitivity metabolomics applications, particularly when dealing with low-input samples such as organoids and dried blood spots (**Chapter V**).

Finally, the technological advancements developed throughout this PhD project were applied to the molecular characterization of highly challenging clinical samples, including microdissected thyroid tumors from patients with DICER1 syndrome. By integrating optimized low-input sample preparation workflows with advanced mass spectrometry-based proteomics, deep molecular characterization was achieved from extremely limited biological material. This study demonstrates how methodological innovations in sensitivity-driven omics workflows can expand the applicability of multi-omics approaches to rare

diseases and small clinical specimens, offering new insights into tumor biology and disease progression (**Chapter VI**).

Overall, the results obtained during this PhD project highlight the critical role of analytical innovation in advancing multi-omics research and biomarker discovery. The developed workflows demonstrate strong potential for translational applications, bridging the gap between fundamental metabolomics research and clinical implementation.

Nevertheless, the translation of these approaches into routine clinical practice requires further validation in larger and independent patient cohorts, as well as the development of standardized protocols to ensure method transferability across laboratories. Future research efforts will focus on further reducing analysis time while maintaining comprehensive molecular coverage, exploiting the increasing acquisition speed of modern mass spectrometry platforms. These advancements will facilitate the analysis of larger patient populations and accelerate the clinical adoption of omics-based diagnostic tools.

Collectively, this PhD work contributes to the advancement of mass spectrometry-based multi-omics strategies, supporting the development of innovative diagnostic and therapeutic approaches and reinforcing the role of omics technologies in precision medicine.

Supplementary Material

Supplementary Material Chapter III

Compound	m/z	Primary Adduct	Molecular Formula	VIPvalue	p-value*	HMDBCODE	KEGG
Metabolites							
Paraxanthine	181.07201	$[M+H]^+$	$C_7H_8N_4O_2$	1.5920	2.19^{E-05}	HMDB0001860	C13747
FA 18:1;1O	295.22836	$[M-H]^-$	$C_{18}H_{34}O_3$	1.5692	4.18^{E-28}	HMDB0030981	
Dehydroepiandrosterone sulfate	367.15889	$[M-H]^-$	$C_{19}H_{28}O_5S$	1.5499	4.18^{E-40}	HMDB0001032	C04555
Isoleucylproline	229.15457	$[M+H]^+$	$C_{11}H_{20}N_2O_3$	1.5264	1.98^{E-07}	HMDB0011174	
Asymmetric dimethylarginine	203.15007	$[M+H]^+$	$C_8H_{18}N_4O_2$	1.4982	7.12^{E-15}	HMDB0001539	C03626
Methylguanine	166.07248	$[M+H]^+$	$C_6H_7N_5O$	1.4922	1.09^{E-09}	HMDB0003282	C04152
CAR 9:0	302.23251	$[M+H]^+$	$C_{16}H_{31}NO_4$	1.4194	1.88^{E-08}	HMDB0013288	
Methylpyridonecarboxamide	153.06585	$[M+H]^+$	$C_7H_8N_2O$	1.3698	1.88^{E-08}	HMDB0004194	
CAR 10:1	314.23275	$[M+H]^+$	$C_{17}H_{31}NO_4$	1.3474	5.90^{E-08}	HMDB0250918	
CAR 2:0	204.12322	$[M+H]^+$	$C_9H_{17}NO_4$	1.3381	6.71^{E-19}	HMDB0240773	C02571
CAR 10:0	316.24831	$[M+H]^+$	$C_{17}H_{33}NO_4$	1.3152	4.28^{E-07}	HMDB0000651	C03299
Deoxycholic acid glycine conjugate	448.30692	$[M-H]^-$	$C_{26}H_{43}NO_5$	1.3128	8.75^{E-04}	HMDB0000631	C05464
Hydroxyanthranilic acid	154.05006	$[M+H]^+$	$C_7H_7NO_3$	1.3118	2.28^{E-07}	HMDB0001476	C00632
Indoxyl sulfate	212.00253	$[M-H]^-$	$C_8H_7NO_4S$	1.2672	2.14^{E-04}	HMDB0000682	

CAR 6:0	260.18546	$[M+H]^+$	$C_{13}H_{26}NO_4$	1.2195	5.61^{E-10}	HMDB0000705	
CAR 8:0	288.21724	$[M+H]^+$	$C_{15}H_{29}NO_4$	1.2038	1.87^{E-06}	HMDB0000791	C02838
Glycerophosphocholine	258.11014	$[M+H]^+$	$C_8H_{20}NO_6P$	1.1740	3.90^{E-61}	HMDB0000086	C00670
CAR 5:1	244.15388	$[M+H]^+$	$C_{12}H_{21}NO_4$	1.1487	1.23^{E-04}	HMDB0241656	
Aminooctanoic acid	160.13322	$[M+H]^+$	$C_8H_{17}NO_2$	1.1487	1.18^{E-04}	HMDB0247418	
Histidine	156.07681	$[M+H]^+$	$C_6H_9N_3O_2$	1.1124	7.08^{E-05}	HMDB0000177	C00135
Lipids							
SM 42:2;3O	829.67813	$[M+H]^+$	$C_{47}H_{93}N_2O_7P$	2.1819	1.71^{E-02}	HMDB0013469	C00550
LPC 18:2_A	520.34014	$[M+H]^+$	$C_{26}H_{50}NO_7P$	1.8945	1.68^{E-22}	HMDB0010386	C04100
LPC 20:3	546.355	$[M+H]^+$	$C_{28}H_{52}NO_7P$	1.8821	2.85^{E-21}	HMDB0010393	C03916
LPC 14:0	468.30896	$[M+H]^+$	$C_{22}H_{46}NO_7P$	1.8775	3.96^{E-20}	HMDB0010379	C03916
LPC 18:2_B	564.33085	$[M+HCOO]^-$	$C_{26}H_{50}NO_7P$	1.8629	6.72^{E-09}	HMDB0010386	C03916
LPC 18:3	518.32392	$[M+H]^+$	$C_{26}H_{48}NO_7P$	1.7920	5.64^{E-17}	HMDB0010387	C03916
LPC 20:4_A	544.33999	$[M+H]^+$	$C_{28}H_{50}NO_7P$	1.7817	4.99^{E-18}	HMDB0010396	C03916
LPC O-16:0	482.36072	$[M+H]^+$	$C_{24}H_{52}NO_6P$	1.7589	2.85^{E-21}	HMDB0243890	C13903
LPC O-18:1_A	508.37644	$[M+H]^+$	$C_{26}H_{54}NO_6P$	1.7491	5.65^{E-50}	HMDB0013122	C04317
LPC 15:0	482.32461	$[M+H]^+$	$C_{23}H_{48}NO_7P$	1.7397	2.28^{E-20}	HMDB0010381	C04230
LPC 18:1_A	522.3558	$[M+H]^+$	$C_{26}H_{52}NO_7P$	1.6699	2.26^{E-32}	HMDB0010385	C04230

LPC O-16:1	480.34516	$[M+H]^+$	$C_{24}H_{50}NO_6P$	1.6303	3.35^{E-50}	HMDB0010407	C04317
LPC 16:0	496.33988	$[M+H]^+$	$C_{24}H_{50}NO_7P$	1.6302	1.14^{E-35}	HMDB0010382	C04233
TG 12:0_16:0_18:1	794.72192	$[M+NH_4]^+$	$C_{49}H_{92}O_6$	1.6290	1.20^{E-03}		C00422
LPC 18:1_B	566.34631	$[M+HCOO]^-$	$C_{26}H_{52}NO_7P$	1.6217	4.87^{E-30}	HMDB0010385	C04230
LPC 17:0	510.35577	$[M+H]^+$	$C_{25}H_{52}NO_7P$	1.6162	3.80^{E-31}	HMDB0012108	C04230
LPC 18:0	524.37124	$[M+H]^+$	$C_{26}H_{54}NO7P$	1.6111	1.58^{E-40}	HMDB0010384	C04230
PC 40:8	830.56883	$[M+H]^+$	$C_{48}H_{80}NO_8P$	1.5994	1.98^{E-18}	HMDB0008443	C00157
PC 18:1_22:6	832.58371	$[M+H]^+$	$C_{48}H_{82}NO_8P$	1.5939	1.04^{E-07}	HMDB0008123	C1387
PC 16:0_16:0	734.56896	$[M+H]^+$	$C_{40}H_{80}NO_8P$	1.5665	6.52^{E-01}	HMDB0000564	C00157

Table S1. Significant metabolites and lipids are annotated according to variable importance in projection (VIP) scores (from PLS-DA) and p-value (one-way ANOVA).

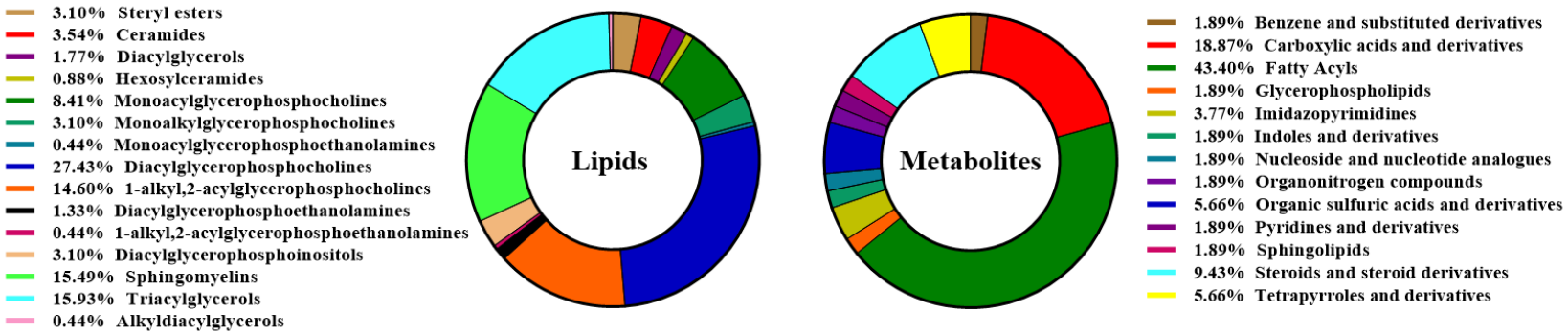


Figure S1. Graphical representation of distribution and class information for metabolites (HMDB) and lipids (LIPID MAPS) detected by UHPLC-TIMS-Q-TOF analysis in HCC, HCV and MC plasma samples.

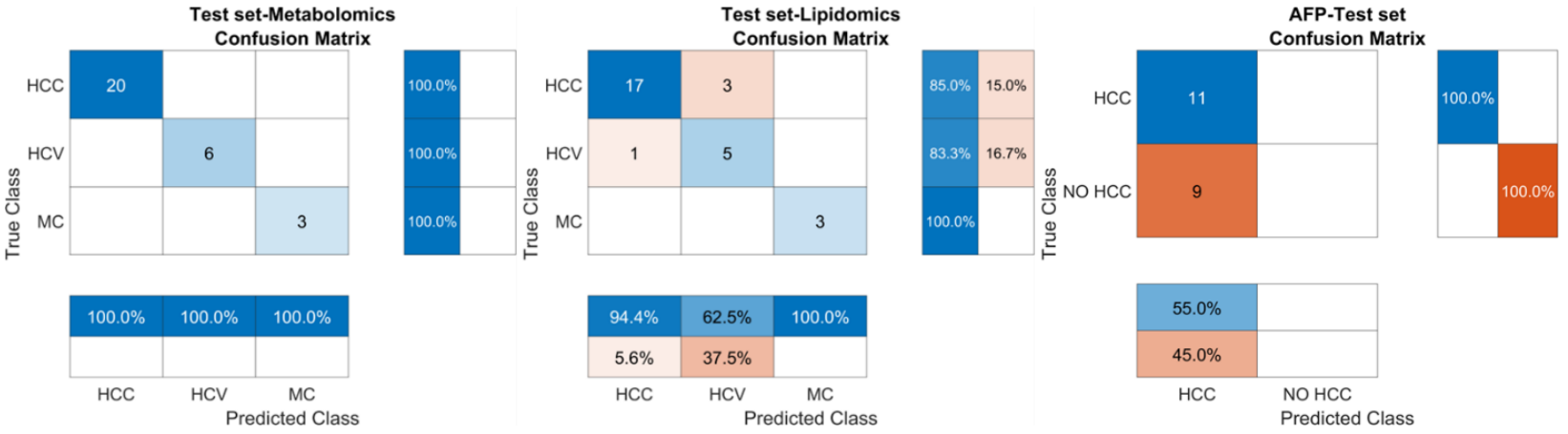


Figure S2. Graphical representation of confusion matrices obtained from PLS-DA models of independent modality. The reported confusion matrices refer to PLS-DA performance for the test phase

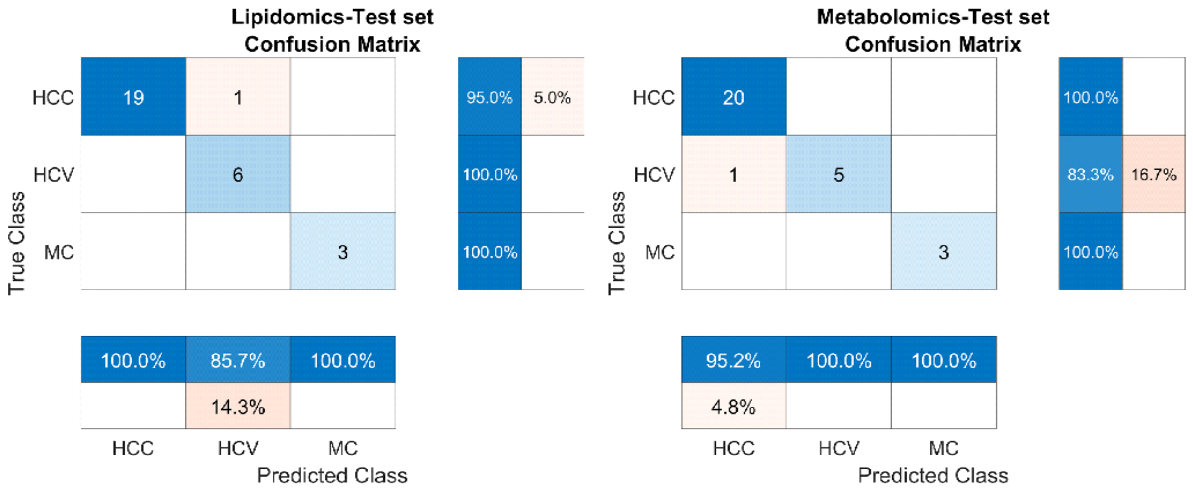


Figure S3. Graphical representation of confusion matrices obtained from SIMCA models of both independent modalities. The reported confusion matrices refer to SIMCA performance for the test phase.

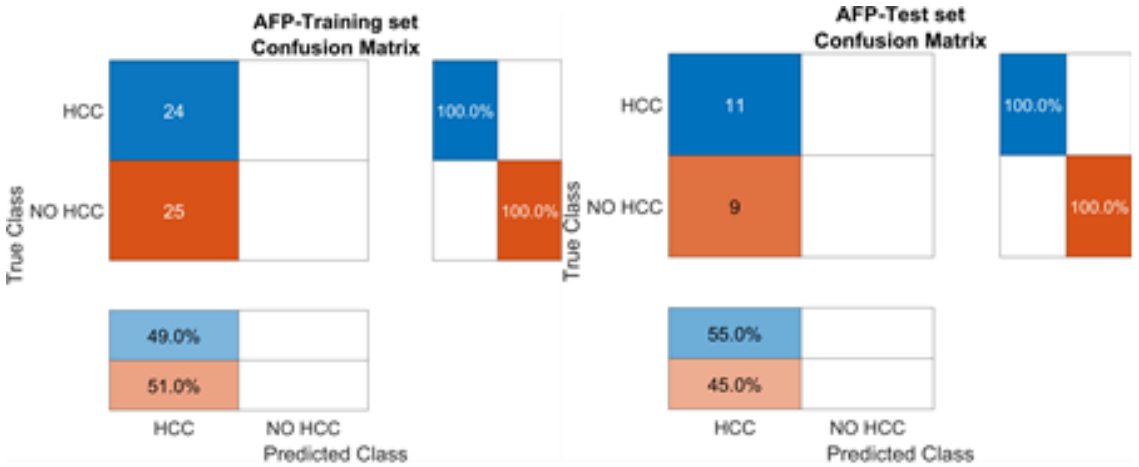


Figure S4. Graphical representation of confusion matrices obtained for AFP classification.

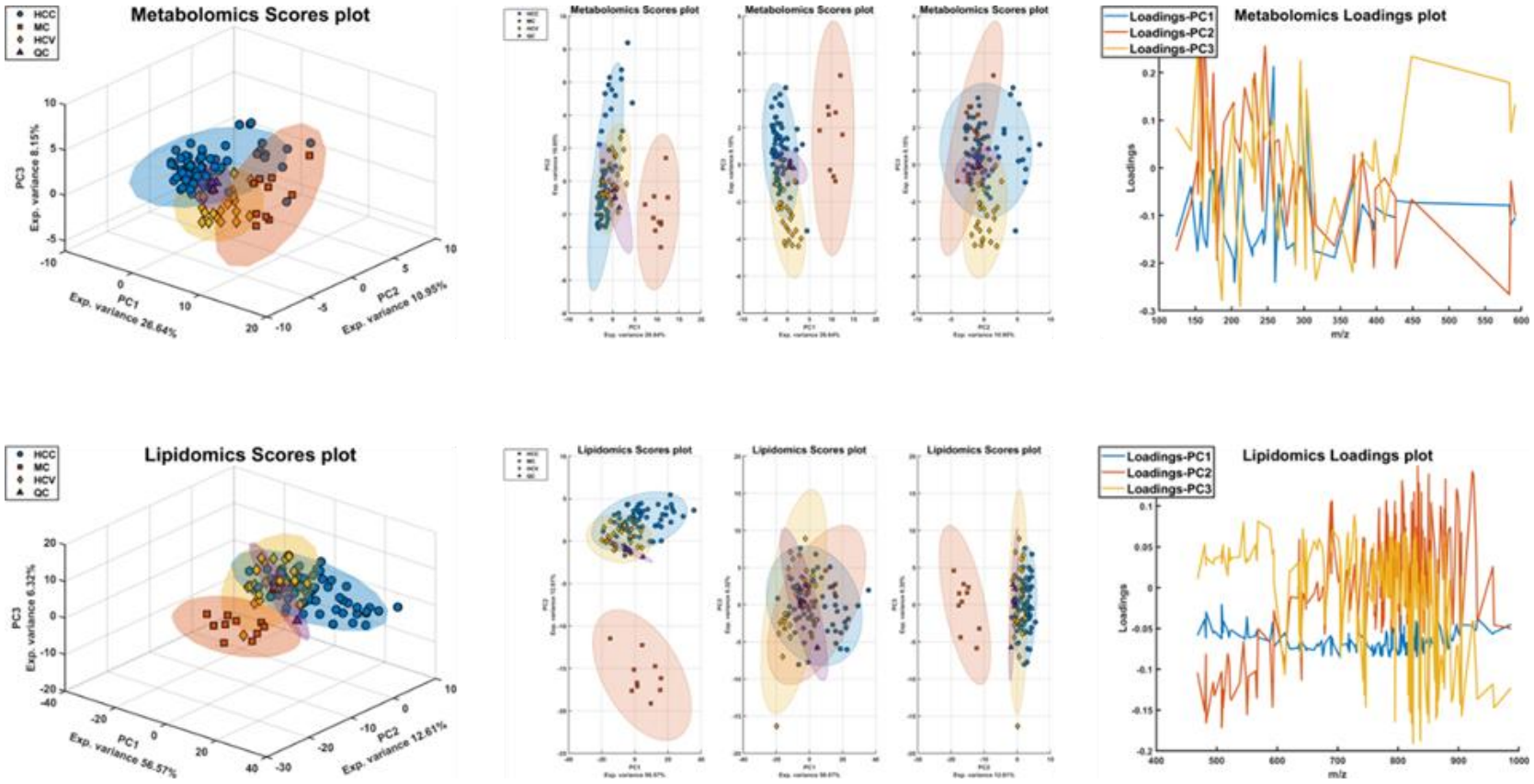


Figure S4. PCA scores and loadings plots (PC1, PC2 and PC3) of metabolomics and lipidomics datasets including QC with confidence ellipses (95%) for each class. The datasets were pre-processed by 1) normalization using internal standard 2) the missing values and zeros were replaced with one-fifth of the minimum value recorded in the data set for that molecule 3) logarithm values of the base of 10 and 4) autoscaling.

Name	m/z	VIPs	p-value	R	R p-value
Paraxanthine	181.1	1.592	2.19E-05	-0.159	1.11E-01
FA 18:1+1O	295.2	1.569	4.18E-40	0.036	7.20E-01
Dehydroepiandrosterone sulfate	367.2	1.55	1.14E-09	-0.626	2.00E-12
Isoleucylproline	229.2	1.527	1.98E-07	0.1155	2.48E-01
Asymmetric dimethylarginine	203.2	1.498	7.12E-15	0.2774	4.76E-03
Methylguanine	166.1	1.492	1.09E-09	0.2861	3.56E-03
CAR 9:0	302.2	1.42	1.88E-08	-0.143	1.52E-01
Methylpyridonecarboxamide	153.1	1.37	2.18E-10	0.1902	5.55E-02
CAR 10:1	314.2	1.348	5.90E-08	-0.117	2.40E-01
CAR 2:0	204.1	1.338	6.71E-19	0.2369	1.65E-02
CAR 10:0	316.3	1.315	4.28E-07	0.0182	8.56E-01
Deoxycholic acid glycine conjugate	448.3	1.313	8.75E-04	0.2815	4.16E-03
Hydroxyanthranilic acid	154.1	1.312	2.28E-07	0.2397	1.53E-02
Indoxyl sulfate	212	1.267	2.14E-04	-0.069	4.93E-01
CAR 6:0	260.2	1.22	5.61E-10	0.1596	1.09E-01
CAR 8:0	288.2	1.204	1.87E-06	-0.017	8.66E-01
Glycerophosphocholine	258.1	1.174	3.90E-61	0.0433	6.66E-01
CAR 5:1	244.2	1.149	1.23E-04	0.0254	8.00E-01
Aminooctanoic acid	160.1	1.142	1.18E-04	-0.04	6.92E-01
Histidine	156.1	1.113	7.08E-05	0.0783	4.34E-01
Carnitine	162.1	1.048	4.03E-07	0.0528	5.98E-01
Urobilinogen	591.3	1.013	1.35E-02	0.0931	3.52E-01

Table S2. List of significant metabolites based on PLS-DA classification and VIPS scores (>1). The presented table includes the metabolite names, corresponding m/z measurements, VIPs scores pertaining to the PLS-DA model, ANOVA-derived P-values, correlation coefficients with age, and the associated R-p-values indicating the significance of these correlations.

Name	m/z	VIPs	p-value	R	R p-value
SM 42:2;3O	829.68	2.182	1.71E-02	0.2939	2.72E-03
LPC 18:2_A	520.34	1.8945	1.68E-22	-0.407	2.18E-05
LPC 20:3	546.36	1.8821	2.85E-21	-0.2964	2.49E-03
LPC 14:0	468.31	1.8775	3.96E-20	-0.2266	2.20E-02
LPC 18:2_B	564.33	1.8629	6.72E-23	-0.3796	8.33E-05
LPC 18:3	518.32	1.792	5.64E-17	-0.2733	5.45E-03
LPC 20:4_A	544.34	1.7817	4.99E-18	-0.4623	1.00E-06
LPC O-16:0	482.36	1.7589	2.32E-57	0.0325	7.46E-01
LPC O-18:1_A	508.38	1.7491	5.65E-50	0.0577	5.65E-01
LPC 15:0	482.32	1.7397	2.28E-20	-0.189	5.72E-02
LPC 18:1_A	522.36	1.6699	2.26E-32	-0.1663	9.48E-02
LPC O-16:1	480.35	1.6303	3.35E-50	-0.0143	8.87E-01
LPC 16:0	496.34	1.6302	1.14E-35	-0.1578	1.13E-01
TG 12:0_16:0_18:1	794.72	1.629	1.24E-03	-0.1925	5.25E-02
LPC 18:1_B	566.35	1.6217	4.87E-30	-0.1794	7.12E-02
LPC 17:0	510.36	1.6162	3.80E-31	-0.202	4.17E-02
LPC 18:0	524.37	1.6111	1.58E-40	-0.1426	1.53E-01
PC 40:8	830.57	1.5994	1.98E-18	-0.3994	3.21E-05
PC 18:1_22:6	832.58	1.5939	1.04E-07	0.0556	5.79E-01
PC 16:0_16:0	734.57	1.5665	6.52E-01	0.0717	4.74E-01
LPC O-18:0	510.39	1.5381	1.03E-52	0.0386	7.00E-01
PC O-20:1	550.39	1.5349	9.15E-47	-0.0086	9.32E-01

LPC O-18:1_B	508.38	1.4866	2.96E-51	0.0134	8.94E-01
DG 18:0_18:1	640.59	1.4812	5.94E-17	-0.1279	2.00E-01
LPC O-24:1	592.47	1.4765	2.71E-41	0.0162	8.71E-01
LPC 16:1	494.32	1.4569	2.55E-20	-0.1271	2.03E-01
PC 16:0_22:6_B	850.56	1.4499	6.38E-08	-0.0268	7.89E-01
TG 15:0_16:0_18:2	834.75	1.3995	2.05E-02	-0.1609	1.06E-01
PC O-34:1	746.6	1.3778	7.35E-03	0.1925	5.26E-02
TG 18:1_18:2_20:4	922.78	1.3776	3.85E-05	-0.0902	3.67E-01
PC 18:0_22:6	878.59	1.3767	2.51E-07	-0.0151	8.81E-01
PC 16:0_22:6_A	806.57	1.3469	1.53E-07	-0.0485	6.28E-01
PC O-18:1_20:4	838.6	1.3376	1.23E-06	-0.0165	8.69E-01
TG 18:1_18:1_20:4	924.8	1.3216	1.08E-06	-0.1488	1.35E-01
PC O-38:5	794.6	1.3203	3.66E-06	-0.0204	8.38E-01
PC O-18:1_16:0	790.6	1.2987	1.21E-01	0.1728	8.24E-02
PC 40:6	834.6	1.2871	3.17E-06	-0.0232	8.17E-01
PC 18:2_18:2_A	782.57	1.2807	3.61E-11	-0.3662	1.53E-04
TG 14:0_16:0_18:3	818.72	1.2791	6.11E-04	-0.2173	2.82E-02
LPC 20:0	552.4	1.2706	1.15E-32	-0.1322	1.85E-01
LPC 20:2	548.37	1.265	7.05E-35	-0.1816	6.77E-02
Cer 18:1;2O/22:0	622.61	1.2524	3.43E-10	-0.2111	3.32E-02
LPC 19:0	538.39	1.2287	8.42E-35	-0.038	7.04E-01
SM 18:1;2O/22:0	787.67	1.2189	5.29E-12	-0.2709	5.90E-03
Cer 18:1;2O/24:0	650.64	1.2073	1.10E-11	-0.254	1.00E-02
TG 18:2_18:2_18:2	896.77	1.2037	4.71E-03	-0.0523	6.02E-01
TG 12:0_18:2_18:2	816.71	1.1877	1.08E-03	-0.2195	2.67E-02

TG 14:0_16:0_18:2	820.74	1.178	3.69E-04	-0.1861	6.10E-02
CE 20:5	688.6	1.1727	3.32E-07	-0.1148	2.51E-01
LPC 17:1	508.34	1.1722	3.75E-22	-0.1814	6.80E-02
CE 20:4	690.62	1.155	1.80E-09	-0.2331	1.84E-02
LPC 20:4_B	588.33	1.1548	3.26E-16	-0.4659	8.02E-07
PC 16:0_22:5	808.58	1.1527	8.81E-11	-0.2233	2.41E-02
LPC 22:6	568.34	1.1513	8.82E-07	-0.2434	1.37E-02
PC O-16:0_16:0	764.58	1.137	4.05E-01	0.1333	1.82E-01
PC 16:0_20:4_B	826.56	1.1328	2.75E-10	-0.2393	1.54E-02
SM 16:1;2O/22:0	759.64	1.1179	6.32E-12	-0.2676	6.55E-03
PC 16:0_20:5_A	780.55	1.1174	1.80E-05	-0.0213	8.32E-01
PC 13:0_22:4	768.55	1.1159	1.68E-10	-0.3181	1.12E-03
LPE 18:0	482.32	1.1044	3.82E-17	-0.2024	4.14E-02
TG 14:0_18:2_18:2	844.74	1.1022	1.25E-04	-0.2389	1.56E-02
PC O-40:5	822.64	1.0959	5.73E-08	-0.1505	1.31E-01
PC O-32:0	720.59	1.0891	1.44E-01	0.0789	4.31E-01
PC 16:0_16:1	732.55	1.0752	5.43E-02	0.2058	3.80E-02
DG 16:0_18:1	612.56	1.0731	1.19E-08	0.013	8.97E-01
SM 43 :1;2O	829.71	1.0614	2.20E-11	-0.2895	3.16E-03
PI 18:0_18:2	861.55	1.0516	3.29E-04	0.055	5.83E-01
TG 18:1_18:2_18:2	898.78	1.0484	2.08E-04	-0.178	7.35E-02
PC 18:2_18:2_B	826.56	1.0478	8.83E-11	-0.3113	1.45E-03
PC O-16:1_18:1	788.58	1.0464	2.35E-01	0.1096	2.73E-01
PC 14:0_18:2_B	730.54	1.0458	1.43E-11	-0.2634	7.49E-03
PC 36:0	790.63	1.0434	2.62E-08	-0.2483	1.18E-02

PC O-44:6	876.68	1.0422	1.93E-06	-0.2675	6.57E-03
TG 16:0_18:2_18:3	870.75	1.0403	1.61E-04	-0.2373	1.63E-02
PC 18:0_20:4_A	854.59	1.038	1.93E-10	-0.2809	4.24E-03
SM 34:0;2O	705.59	1.0335	5.63E-01	0.0824	4.10E-01
TG 18:2_18:2_22:0	956.86	1.0238	2.21E-01	-0.0137	8.91E-01
SM 18:1;2O/24:0	815.7	1.0134	5.98E-11	-0.2674	6.58E-03
PC 36:6	778.54	1.0088	1.59E-10	-0.2016	4.22E-02

Table S3. List of significant lipids based on PLS-DA classification and VIPS scores (>1). The presented table includes the metabolite names, corresponding m/z measurements, VIPs scores pertaining to the PLS-DA model, ANOVA-derived P-values, correlation coefficients with age, and the associated R-p-values indicating the significance of these correlations.

RT [min]	CCS (Å²)	m/z meas.	M meas.	Mob. 1/K0	Ions	Name	Molecula Formula	MS/MS score	Δm/z [ppm]	ΔCCS [%]
5.7	133.7	160.13322	159.12594	0.616	$[M+H]^+$	Aminooctanoic acid	$C_8H_{17}NO_2$	560.3	0.057	3
1.18	195.8	369.17463	370.1819	0.942	$[M-H]^-$	Androsterone sulfate	$C_{19}H_{30}O_5S$	899.1	-0.04	2.9
1.27	183.5	303.23318	304.24046	0.877	$[M-H]^-$	Arachidonic acid	$C_{20}H_{32}O_2$	830.1	1.096	3
7.66	135.8	175.11864	174.11137	0.63	$[M+H]^+$	Arginine	$C_6H_{14}N_4O_2$	986.7	-1.44	1.7
7.07	143.6	203.15007	202.14279	0.672	$[M+H]^+$	Asymmetric dimethylarginine	$C_8H_{18}N_4O_2$	872.1	-0.91	2.7
1.13	240.4	585.27057	584.26329	1.173	$[M+H]^+$	Bilirubin	$C_{33}H_{36}N_4O_6$	901.1	-0.35	0.6
2.51	242.4	583.25559	582.24831	1.182	$[M+H]^+$	Biliverdin	$C_{33}H_{34}N_4O_6$	598.2	0.795	1
3.76	190.5	316.24831	315.24103	0.912	$[M+H]^+$	CAR 10:0	$C_{17}H_{33}NO_4$	800.3	0.102	3
3.8	184.7	314.23275	313.22547	0.884	$[M+H]^+$	CAR 10:1	$C_{17}H_{31}NO_4$	960.3	0.439	1.5
3.63	204.2	370.29552	369.28964	0.983	$[M+H]^+$, ($[M+Na]^+$)	CAR 14:1	$C_{21}H_{39}NO_4$	900.3	0.9	1.1
3.59	211.5	398.32699	397.31971	1.021	$[M+H]^+$	CAR 16:1_A	$C_{23}H_{43}NO_4$	970.5	0.846	0.1
3.47	211.2	398.32648	397.3192	1.02	$[M+H]^+$	CAR 16:1_B	$C_{23}H_{43}NO_4$	990.1	-0.12	0.2
3.55	206	396.31189	395.30462	0.994	$[M+H]^+$	CAR 16:2	$C_{23}H_{41}NO_4$	877.7	3.347	2.4
3.39	219	426.35783	425.35055	1.059	$[M+H]^+$	CAR 18:1_A	$C_{25}H_{47}NO_4$	899.1	0.208	1.1
3.54	219.1	426.35883	425.35156	1.06	$[M+H]^+$	CAR 18:1_B	$C_{25}H_{47}NO_4$	943.5	3.205	1
5.5	143.4	204.12322	203.11594	0.671	$[M+H]^+$	CAR 2:0	$C_9H_{17}NO_4$	871.8	1.064	0.7
5.07	148.3	218.1381	217.13082	0.697	$[M+H]^+$	CAR 3:0	$C_{10}H_{19}NO_4$	967.3	-1.87	0.5
4.59	154.2	232.15408	231.1468	0.728	$[M+H]^+$	CAR 4:0	$C_{11}H_{21}NO_4$	921.6	-0.83	1.6
4.31	159.6	246.16982	245.16255	0.755	$[M+H]^+$	CAR 5:0	$C_{12}H_{23}NO_4$	996.5	-0.58	1.6
5.84	161	276.14418	275.1369	0.766	$[M+H]^+$	CAR 5:1;O2	$C_{12}H_{21}NO_6$	820.7	0.046	0.1
4.48	159	244.15388	243.1466	0.752	$[M+H]^+$	CAR 5:1	$C_{12}H_{21}NO_4$	910.5	-0.09	2.3
4.1	166.9	260.18546	259.17819	0.792	$[M+H]^+$	CAR 6:0	$C_{13}H_{25}NO_4$	655.9	-0.43	2.2

3.89	179.1	288.21724	287.20996	0.854	$[M+H]^+$	CAR 8:0	$C_{15}H_{29}NO_4$	765.3	1.015	3
3.97	175	286.20138	285.1941	0.834	$[M+H]^+$	CAR 8:1	$C_{15}H_{27}NO_4$	987.6	0.608	1.9
3.8	181.3	302.23251	301.22524	0.866	$[M+H]^+$	CAR 9:0	$C_{16}H_{31}NO_4$	897.8	-0.44	2.6
5.9	132.4	162.11254	161.10526	0.61	$[M+H]^+$	Carnitine	$C_7H_{15}NO_3$	588.1	0.485	0
1	135.1	187.00719	188.01446	0.629	$[M-H]^-$	Cresol sulfate	$C_7H_8O_4S$	943.9	0.997	1.7
1.15	196.2	367.15889	368.16616	0.944	$[M-H]^-$	Dehydroepiandrosterone sulfate	$C_{19}H_{28}O_5S$	999.8	0.944	4.7
4.25	200.5	448.30692	449.3142	0.972	$[M-H]^-$	Deoxycholic acid glycine conjugate	$C_{26}H_{43}NO_5$	998.8	-0.02	2.8
3.7	194.9	342.26504	341.25777	0.936	$[M+H]^+$	Dodecenoylcarnitine	$C_{19}H_{35}NO_4$	822.1	2.424	1.7
1.49	179.6	295.22836	296.23564	0.857	$[M-H]^-$	FA 18:1+1O	$C_{18}H_{32}O_3$	869.6	2.213	0.2
6.5	156.1	258.11014	257.10286	0.74	$[M+H]^+$	Glycerophosphocholine	$C_8H_{20}NO_6P$	992.1	0.151	0.1
7.61	130.7	156.07681	155.06953	0.601	$[M+H]^+$	Histidine	$C_6H_9N_3O_2$	899.2	0.359	0.8
7.53	140.1	189.13533	188.12805	0.653	$[M+H]^+$	Homoarginine	$C_7H_{16}N_4O_2$	920.6	2.934	1.5
1.3	180.9	313.23727	312.22999	0.866	$[M+H]^+$	Hydroperoxylinoleic acid	$C_{18}H_{32}O_4$	606.7	-0.11	1.5
2.46	124.4	154.05006	123.04278	0.571	$[M+H]^+$	Hydroxyanthranilic acid	$C_7H_7NO_3$	976.2	1.85	3
1.27	207.1	411.18539	412.19266	1.001	$[M-H]^-$	Hydroxypregnenolone sulfate	$C_{21}H_{32}O_6S$	619.9	0.84	3
1.14	128.7	158.06121	159.06849	0.592	$[M-H]^-$	Indole-acetaldehyde	$C_{10}H_9NO$	609.8	0.524	4.2
1.18	142.2	212.00253	213.00981	0.668	$[M-H]^-$	Indoxyl sulfate	$C_8H_7NO_4S$	752	1.094	0.2
5.93	149.3	229.15457	228.14729	0.704	$[M+H]^+$	Isoleucylproline	$C_{11}H_{20}N_2O_3$	500.2	-0.48	2.5
4.19	130.2	166.07248	165.06521	0.601	$[M+H]^+$	Methylguanine	$C_6H_7N_5O$	622.1	0.912	1.8
7.59	133.2	170.09257	169.08529	0.616	$[M+H]^+$	Methylhistidine	$C_7H_{11}N_3O_2$	849	0.81	1.7
2.29	130.3	153.06585	152.05858	0.598	$[M+H]^+$	Methylpyridonecarboxamide	$C_7H_8N_2O_2$	899.1	-0.11	1.2
1.77	133	181.07201	180.06473	0.618	$[M+H]^+$	Paraxanthine	$C_7H_8N_4O_2$	918.5	-0.37	1.1
4.42	159.2	265.11828	264.111	0.756	$[M+H]^+$	Phenylacetylglutamine	$C_{13}H_{16}N_2O_4$	677.9	-0.3	0.1
1.11	205.1	395.18855	396.19582	0.99	$[M-H]^-$	Pregnenolone sulfate	$C_{21}H_{32}O_5S$	711.7	-1.38	3

5.31	126.2	144.10192	143.09485	0.577	$[M+H]^+$, ($[M+Na]^+$, $[M+K]^+$)	Proline betaine	$C_7H_{13}NO_2$	959.9	0.779	0.2
6.59	151.2	229.11857	228.11129	0.713	$[M+H]^+$	Proline-hydroxyproline	$C_{10}H_{16}N_2O_4$	987.1	0.864	1.3
5.02	146.9	243.06252	244.0698	0.695	$[M-H]^-$	Pseudouridine	$C_9H_{12}N_2O_6$	936.5	0.93	3.2
4.54	203.4	380.25458	379.2473	0.98	$[M+H]^+$	Sphingosine 1-phosphate	$C_{18}H_{38}NO_5P$	980.3	-2.4	0.9
5.67	115.9	124.00733	125.0146	0.523	$[M-H]^-$	Taurine	$C_2H_7NO_3S$	998.1	-0.48	3
5.68	142.3	180.06673	181.07401	0.661	$[M-H]^-$	Tyrosine	$C_9H_{11}NO_3$	908.2	0.621	2.5
4.22	230.9	591.31872	590.31144	1.127	$[M+H]^+$	Urobilinogen	$C_{33}H_{42}N_4O_6$	721.2	1.084	3

Table S4. List of annotated metabolites by HILIC-UHPLC-TIMS-Q-TOF in HCC, HCV and MC plasma samples. Metabolites annotation was carried out with a spectral library approach following the metabolomics standard initiative guidelines. Abbreviations: RT, retention time; m/z meas, mass to charge measured; Mass meas., measured mass; CCS, cross collisional section; Mob., mobility. Primary ion has been used for quantification and additional adducts, if detected, were reported in round brackets.

RT [min]	CCS (Å²)	m/z meas.	M meas.	Mob. $1/K_0$	Ions	Name	Molecular Formula	MS/MS score	Δm/z [ppm]	ΔCCS [%]
3.32	284.2	642.6174	624.58385	1.389	$[M+NH4]^+$, ($[M+Na]^+$, $[M+K]^+$)	CE 16:0	$C_{43}H_{76}O_2$	872.1	-1.436	1.8
3.24	283.5	640.60243	622.56861	1.386	$[M+NH4]^+$	CE 16:1	$C_{43}H_{74}O_2$	803.1	-0.43	2.5
3.31	289.2	668.63442	650.60014	1.415	$[M+NH4]^+$, ($[M+H]^+$)	CE 18:1	$C_{45}H_{78}O_2$	820.7	0.616	2.1
3.25	289.5	666.61862	648.58479	1.416	$[M+NH4]^+$	CE 18:2	$C_{45}H_{76}O_2$	760.7	0.392	2.6
3.26	294.4	692.63127	674.59744	1.441	$[M+NH4]^+$	CE 20:3	$C_{47}H_{78}O_2$	931	-3.881	3.1
3.21	292.9	690.61797	672.58319	1.434	$[M+NH4]^+$, ($[M+H]^+$, $[M+K]^+$, $[M+Na]^+$)	CE 20:4	$C_{47}H_{76}O_2$	735.3	-0.215	2.5
3.15	290.7	688.60212	670.56818	1.423	$[M+NH4]^+$, ($[M+K]^+$, $[M+Na]^+$)	CE 20:5	$C_{47}H_{74}O_2$	784.1	-0.786	3.6
2.48	271.1	594.58198	593.5747	1.323	$[M+H]^+$	Cer 16:1;2O/22:0	$C_{38}H_{75}NO_3$	928.6	0.011	2.1
2.7	277	622.61327	621.60617	1.353	$[M+H]^+$, ($[M+K]^+$, $[M+Na]^+$, $[M+H-H2O]^+$)	Cer 18:1;2O/22:0	$C_{40}H_{79}NO_3$	932.7	-0.222	1.4
2.77	281.1	636.62848	635.62096	1.374	$[M+H]^+$, ($[M+H-H2O]^+$, $[M+Na]^+$)	Cer 18:1;2O/23:0	$C_{41}H_{81}NO_3$	967.8	-0.667	2.7
2.85	283	650.64423	649.63698	1.384	$[M+H]^+$, ($[M+H-H2O]^+$)	Cer 18:1;2O/24:0	$C_{42}H_{83}NO_3$	969.6	-0.527	2.3
2.69	279.8	648.62832	647.62105	1.368	$[M+H]^+$	Cer 18:1;2O/24:1	$C_{42}H_{81}NO_3$	966.8	-0.787	1.7
2.5	272.9	620.59698	619.5897	1.333	$[M+H]^+$	Cer 18:2;2O/22:0	$C_{40}H_{77}NO_3$	974.1	-1.036	0.7
2.62	276.2	634.61303	633.60575	1.35	$[M+H]^+$	Cer 18:2;2O/23:0	$C_{41}H_{79}NO_3$	968.9	-0.151	1
2.51	276.4	646.61191	645.60463	1.351	$[M+H]^+$	Cer 18:2;2O/24:1	$C_{42}H_{79}NO_3$	918.1	-2.111	1.4
2.45	265.8	612.55579	594.52223	1.298	$[M+NH4]^+$, ($[M+H-H2O]^+$, $[M+H]^+$, $[M+Na]^+$, $[M+K]^+$)	DG 16:0_18:1	$C_{37}H_{70}O_5$	939.4	-0.612	1.3
2.67	271.4	640.58702	622.5532	1.327	$[M+NH4]^+$, ($[M+Na]^+$, $[M+H-H2O]^+$, $[M+K]^+$)	DG 18:0_18:1	$C_{39}H_{74}O_5$	941.3	-0.651	2.4
2.47	268.3	638.57128	620.53739	1.311	$[M+NH4]^+$, ($[M+H]^+$, $[M+Na]^+$, $[M+H-H2O]^+$, $[M+K]^+$)	DG 18:1_18:1	$C_{39}H_{72}O_5$	968.5	-0.824	2.4
2.29	264.4	636.55552	618.52167	1.292	$[M+NH4]^+$, ($[M+K]^+$, $[M+Na]^+$, $[M+H-H2O]^+$)	DG 18:1_18:2	$C_{39}H_{70}O_5$	934.8	-0.997	1.8
1.68	306.3	862.62433	861.61705	1.505	$[M+H]^+$	Hex2Cer 18:1;2O/16:0	$C_{46}H_{87}NO_{13}$	499.6	-0.837	2.1
2.68	305.2	812.69395	811.68787	1.499	$[M+H]^+$, ($[M+Na]^+$)	HexCer 18:1;2O/24:0	$C_{48}H_{93}NO_8$	408.2	-4.13	1.3
0.4	225.4	468.30896	467.30169	1.093	$[M+H]^+$	LPC 14:0	$C_{22}H_{46}NO_7P$	584.5	1.053	0.4
0.45	228.6	482.32461	481.31733	1.11	$[M+H]^+$	LPC 15:0	$C_{23}H_{48}NO_7P$	715.8	1.016	0.5

0.51	232.7	496.33988	495.33261	1.131	$[M+H]^+$	LPC 16:0	$C_{24}H_{50}NO_7P$	717.5	0.237	1.9
0.42	227.7	494.32466	493.31721	1.106	$[M+H]^+$, ($[M+Na]^+$)	LPC 16:1	$C_{24}H_{48}NO_7P$	632.8	1.101	0.5
0.57	236.9	510.35577	509.34849	1.152	$[M+H]^+$	LPC 17:0	$C_{25}H_{52}NO_7P$	644.4	0.685	0.8
0.47	231.5	508.33989	507.33262	1.126	$[M+H]^+$	LPC 17:1	$C_{25}H_{50}NO_7P$	446.1	0.249	1.8
0.67	240.5	524.37124	523.36397	1.17	$[M+H]^+$	LPC 18:0	$C_{26}H_{54}NO_7P$	713.5	0.34	0.8
0.53	235.2	522.3558	521.34854	1.144	$[M+H]^+$, ($[M+K]^+$, $[M+Na]^+$)	LPC 18:1_A	$C_{26}H_{52}NO_7P$	643.2	0.747	0.1
0.53	241.8	566.34631	521.34811	1.179	$[M+HCOO]^-$	LPC 18:1_B	$C_{26}H_{52}NO_7P$	944.5	-0.047	0.6
0.45	229.8	520.34014	519.33287	1.118	$[M+H]^+$	LPC 18:2_A	$C_{26}H_{50}NO_7P$	703.7	0.747	1.1
0.45	238.9	564.33085	519.33264	1.164	$[M+HCOO]^-$	LPC 18:2_B	$C_{26}H_{50}NO_7P$	946.2	0.262	2.1
0.41	226.9	518.32392	517.31664	1.104	$[M+H]^+$	LPC 18:3	$C_{26}H_{48}NO_7P$	623.9	-0.13	0.5
0.78	244.3	538.38687	537.37959	1.189	$[M+H]^+$	LPC 19:0	$C_{27}H_{56}NO_7P$	783.3	0.223	0.8
0.93	247.4	552.40238	551.3951	1.205	$[M+H]^+$	LPC 20:0	$C_{28}H_{58}NO_7P$	665.5	0.113	0.8
0.56	237.5	548.37036	547.36308	1.157	$[M+H]^+$	LPC 20:2	$C_{28}H_{54}NO_7P$	666.3	-1.285	0.5
0.47	234.2	546.355	545.34773	1.14	$[M+H]^+$	LPC 20:3	$C_{28}H_{52}NO_7P$	669.3	-0.783	0.1
0.44	232.9	544.33999	543.33258	1.134	$[M+H]^+$, ($[M+Na]^+$)	LPC 20:4_A	$C_{28}H_{50}NO_7P$	617.9	0.404	0.5
0.44	242.9	588.33088	543.33268	1.185	$[M+HCOO]^-$	LPC 20:4_B	$C_{28}H_{50}NO_7P$	933.8	0.387	1.5
0.42	235.5	568.33965	567.33237	1.148	$[M+H]^+$	LPC 22:6	$C_{30}H_{50}NO_7P$	690.6	-0.184	0.5
0.58	234.1	482.36072	481.3535	1.137	$[M+H]^+$, ($[M+Na]^+$)	LPC O-16:0	$C_{24}H_{52}NO_6P$	880.7	0.468	1
0.56	228.7	480.34516	479.33789	1.11	$[M+H]^+$	LPC O-16:1	$C_{24}H_{50}NO_6P$	940.5	0.897	0.6
0.8	241.4	510.39204	509.38476	1.174	$[M+H]^+$	LPC O-18:0	$C_{26}H_{56}NO_6P$	884.6	0.5	0.4
0.6	236.7	508.37644	507.36811	1.151	$[M+H]^+$, ($[M+Na]^+$)	LPC O-18:1_A	$C_{26}H_{54}NO_6P$	874.1	0.564	1
0.75	236.5	508.37625	507.36897	1.15	$[M+H]^+$	LPC O-18:1_B	$C_{26}H_{54}NO_6P$	911.6	0.302	0.9
1.58	262.5	594.48541	593.47813	1.281	$[M+H]^+$	LPC O-24:0	$C_{32}H_{68}NO_6P$	871.7	-0.141	1.7
1.32	257.4	592.47003	591.46275	1.256	$[M+H]^+$	LPC O-24:1	$C_{32}H_{66}NO_6P$	885.1	-0.155	1.2
0.69	224.2	482.32452	481.31725	1.089	$[M+H]^+$	LPE 18:0	$C_{23}H_{48}NO_7P$	777.2	0.841	1.2
1.5	289.9	804.55269	803.54541	1.423	$[M+H]^+$	PC 38:7	$C_{46}H_{78}NO_8P$	958.0	-1.385	0.5
1.72	287.4	770.56811	769.56083	1.41	$[M+H]^+$	PC 35:3	$C_{43}H_{80}NO_8P$	900.0	-1.75	0.1
1.63	286.7	768.55289	767.54696	1.406	$[M+H]^+$, ($[M+Na]^+$)	PC 35:4	$C_{43}H_{78}NO_8P$	931.1	-1.155	0.6
1.47	276	678.50635	677.49908	1.351	$[M+H]^+$	PC 28:0	$C_{36}H_{72}NO_8P$	924.0	-0.712	1.1
1.51	278.4	704.52183	703.51456	1.363	$[M+H]^+$	PC 14:0_16:1	$C_{38}H_{74}NO_8P$	992.9	-0.825	0.9
1.56	283.8	774.52874	729.53054	1.393	$[M+HCOO]^-$	PC 14:0_18:2_A	$C_{40}H_{76}NO_8P$	784.3	-0.388	0.4

2.02	289.7	748.58411	747.57684	1.42	$[M+H]^+$	PC 33:0	$C_{41}H_{82}NO_8P$	515.0	-1.3	0.4
1.85	286.9	746.56892	745.56165	1.406	$[M+H]^+$	PC 33:1	$C_{41}H_{80}NO_8P$	854.0	-0.713	0.3
1.89	290.2	772.58468	771.5774	1.424	$[M+H]^+$	PC 35:2	$C_{43}H_{82}NO_8P$	964.0	-0.525	0.5
1.63	286.8	778.53522	777.52794	1.407	$[M+H]^+$	PC 36:6	$C_{44}H_{76}NO_8P$	310.0	-3.74	0.8
1.71	281.3	706.53764	705.53039	1.378	$[M+H]^+$, ($[M+Na]^+$)	PC 14 :0_16 :0	$C_{38}H_{76}NO_8P$	991.9	-0.696	1.1
1.67	288.5	788.54454	743.54634	1.416	$[M+HCOO]^-$	PC 15:0_18:2	$C_{41}H_{78}NO_8P$	747.4	-0.218	0.8
1.63	281	718.53708	717.5298	1.377	$[M+H]^+$	PC 16:0_15:1	$C_{39}H_{76}NO_8P$	992.1	-1.322	0.7
1.94	287.1	734.56896	733.56155	1.407	$[M+H]^+$, ($[M+Na]^+$)	PC 16:0_16:0	$C_{40}H_{80}NO_8P$	996.1	-0.645	1.4
1.74	283.7	732.55327	731.54599	1.39	$[M+H]^+$	PC 16:0_16:1	$C_{40}H_{78}NO_8P$	994	-0.698	0.7
1.85	287.8	790.56171	745.56351	1.412	$[M+HCOO]^-$	PC 16:0_17:1	$C_{41}H_{80}NO_8P$	863.1	1.673	0
1.97	290.2	760.58546	759.57818	1.423	$[M+H]^+$	PC 16:0_18:1_A	$C_{42}H_{82}NO_8P$	993.8	0.507	0.5
1.96	291.8	804.57683	759.57863	1.433	$[M+HCOO]^-$	PC 16:0_18:1_B	$C_{42}H_{82}NO_8P$	876.7	1.029	0.2
2.25	295.8	776.61471	775.60743	1.451	$[M+H]^+$	PC 35:0	$C_{43}H_{86}NO_8P$	720.0	-2.367	0.7
2.08	293.4	818.59205	773.59384	1.441	$[M+HCOO]^-$	PC 17:0_18:1_A	$C_{43}H_{84}NO_8P$	537.0	0.44	0.03
1.84	291.6	784.58443	783.57638	1.431	$[M+H]^+$, ($[M+Na]^+$, $[M+K]^+$, $[M+H-H2O]^+$)	PC 18:1_18:2	$C_{44}H_{82}NO_8P$	993.0	-0.761	0.8
1.74	290.3	782.56973	781.56245	1.424	$[M+H]^+$	PC 16:0_20:4_A	$C_{44}H_{80}NO_8P$	993.9	0.379	0.7
1.74	293.2	826.56065	781.56245	1.44	$[M+HCOO]^-$	PC 16:0_20:4_B	$C_{44}H_{80}NO_8P$	816.4	0.424	0.4
1.6	287.7	780.55348	779.54579	1.412	$[M+H]^+$, ($[M+K]^+$, $[M+Na]^+$)	PC 16:0_20:5_A	$C_{44}H_{78}NO_8P$	993.5	-0.218	1
1.59	291.2	824.54412	779.54592	1.43	$[M+HCOO]^-$	PC 16:0_20:5_B	$C_{44}H_{78}NO_8P$	844.8	-0.684	0.6
2.13	296.6	800.61637	799.60909	1.456	$[M+H]^+$	PC 37:2	$C_{45}H_{86}NO_8P$	856.0	-0.067	0.5
2.41	300.9	816.64667	815.63944	1.478	$[M+H]^+$, ($[M+Na]^+$)	PC 38:1	$C_{46}H_{90}NO_8P$	854.0	-1.253	0.6
1.69	292.8	806.5692	805.56156	1.437	$[M+H]^+$, ($[M+Na]^+$, $[M+K]^+$)	PC 16:0_22:6_A	$C_{46}H_{80}NO_8P$	994.1	-0.254	0.7
1.69	296.3	850.56039	805.56218	1.456	$[M+HCOO]^-$	PC 16:0_22:6_B	$C_{46}H_{80}NO_8P$	870.3	0.108	0.6
1.56	280.9	730.53751	729.53037	1.377	$[M+H]^+$, ($[M+Na]^+$)	PC 14:0_18:2_B	$C_{40}H_{76}NO_8P$	993.0	-0.606	0.8
1.61	284.3	756.55299	755.54572	1.394	$[M+H]^+$	PC 16:1_18:2_A	$C_{42}H_{78}NO_8P$	993.7	-1.017	0.4
1.61	287.8	800.54422	755.54602	1.413	$[M+HCOO]^-$	PC 16:1_18:2_B	$C_{42}H_{78}NO_8P$	797.1	-0.574	0.2
1.67	283.1	744.55333	743.54731	1.388	$[M+H]^+$, ($[M+Na]^+$)	PC 15:0_18:1	$C_{41}H_{78}NO_8P$	992.0	-0.622	0.3
2.18	293.6	762.59761	761.59033	1.44	$[M+H]^+$	PC 34:0	$C_{42}H_{84}NO_8P$	949.0	-4.099	1.1
2.08	293.5	774.60001	773.5936	1.44	$[M+H]^+$, ($[M+Na]^+$)	PC 17:0_18:1_B	$C_{43}H_{84}NO_8P$	992.2	-0.981	1
1.89	292.3	816.57603	771.57782	1.435	$[M+HCOO]^-$	PC 17:0_18:2	$C_{43}H_{82}NO_8P$	839.8	0.069	0.3

1.78	287	758.57004	757.56276	1.408	$[M+H]^+$	PC 16:0_18:2	$C_{42}H_{80}NO_8P$	994.0	0.917	0.58
2.4	298.8	790.62806	789.62079	1.466	$[M+H]^+$	PC 36:0	$C_{44}H_{88}NO_8P$	589.0	-4.989	1.0
2.2	296.1	788.61642	787.60915	1.453	$[M+H]^+$	PC 18:0_18:1	$C_{44}H_{86}NO_8P$	994.2	0.053	0.6
2.01	295.7	830.59237	785.59416	1.453	$[M+HCOO]^-$	PC 18:0_18:2_A	$C_{44}H_{84}NO_8P$	876.9	0.851	0.5
2.08	297.8	812.61603	811.60858	1.462	$[M+H]^+$, ($[M+Na]^+$, $[M+H-H2O]^+$, $[M+K]^+$)	PC 18:0_20:3	$C_{46}H_{86}NO_8P$	994.1	-0.433	0.8
1.98	299.3	854.59164	809.59343	1.471	$[M+HCOO]^-$	PC 18:0_20:4_A	$C_{46}H_{84}NO_8P$	827.5	0.072	0.8
2.3	303.9	840.64675	839.63947	1.493	$[M+H]^+$	PC 18:0_22:3	$C_{48}H_{90}NO_8P$	989.8	-1.064	0.9
2.15	303.4	882.62002	837.62182	1.492	$[M+HCOO]^-$	PC 18:0_22:4_A	$C_{48}H_{88}NO_8P$	869.6	-3.437	0.3
1.92	301.7	878.59111	833.59291	1.483	$[M+HCOO]^-$	PC 18:0_22:6	$C_{48}H_{84}NO_8P$	898.4	-0.597	0.8
2.01	293.3	786.60119	785.59391	1.439	$[M+H]^+$	PC 18: 0_18:2_B	$C_{44}H_{84}NO_8P$	994.0	0.59	0.7
1.88	298.2	854.59017	809.59197	1.465	$[M+HCOO]^-$	PC 18:1_20:3_A	$C_{46}H_{84}NO_8P$	730	-1.547	0.5
1.71	296.7	832.58371	831.57728	1.457	$[M+H]^+$, $[M+Na]^+$	PC 18:1_22:6	$C_{48}H_{82}NO_8P$	993.9	-1.614	1
1.64	288	782.56824	781.56095	1.413	$[M+H]^+$, ($[M+K]^+$, $[M+Na]^+$)	PC 18:2_18:2_A	$C_{44}H_{80}NO_8P$	996.2	-1.716	1.9
1.63	291.5	826.55949	781.56129	1.432	$[M+HCOO]^-$	PC 18:2_18:2_B	$C_{44}H_{80}NO_8P$	782.2	-1.143	0.3
2.13	297.8	844.60712	799.60892	1.463	$[M+HCOO]^-$	PC 19:0_18:2	$C_{45}H_{86}NO_8P$	840.8	-0.128	0.3
1.85	293.6	796.58381	795.57653	1.441	$[M+H]^+$	PC 37:4	$C_{45}H_{82}NO_8P$	909.0	-1.55	0.9
2.22	299.3	814.63052	813.6232	1.47	$[M+H]^+$, ($[M+Na]^+$)	PC 38:2	$C_{46}H_{88}NO_8P$	950.0	-1.829	0.6
1.87	295.1	810.59867	809.59139	1.449	$[M+H]^+$	PC 18:1_20:3_B	$C_{46}H_{84}NO_8P$	993.0	-2.495	0.5
1.98	296.4	810.59979	809.59229	1.455	$[M+H]^+$, ($[M+Na]^+$, $[M+K]^+$)	PC 18:0_20:4_B	$C_{46}H_{84}NO_8P$	994.0	-1.104	0.9
1.76	294	808.58314	807.57587	1.443	$[M+H]^+$	PC 16:0_22:5	$C_{46}H_{82}NO_8P$	994.0	-2.326	0.7
2.45	304.9	842.66354	841.65626	1.498	$[M+H]^+$	PC 40:2	$C_{48}H_{92}NO_8P$	741.0	0.231	0.6
2.15	302.1	838.63053	837.62264	1.484	$[M+H]^+$, ([M+Na] +)	PC 18:0_22:4_B	$C_{48}H_{88}NO_8P$	990.0	-1.792	0.87
1.91	299.2	834.5998	833.59238	1.469	$[M+H]^+$, ($[M+H-H2O]^+$, $[M+K]^+$, $[M+Na]^+$)	PC 40:6	$C_{48}H_{84}NO_8P$	981.0	-1.017	1.3
1.55	293.8	830.56883	829.56156	1.443	$[M+H]^+$	PC 40:8	$C_{48}H_{80}NO_8P$	936.0	-0.743	0.8
2.37	307.9	866.662	865.65472	1.513	$[M+H]^+$	PC 42:4	$C_{50}H_{92}NO_8P$	905.0	-1.317	1.7
2.31	299.1	802.63132	801.62404	1.468	$[M+H]^+$	PC 37:1	$C_{45}H_{88}NO_8P$	845.0	-0.661	0.5
2.09	286.4	764.58024	719.58203	1.405	$[M+HCOO]^-$	PC O-16:0_16:0	$C_{40}H_{82}NO_7P$	929.3	-1.109	0.6
1.89	292	812.58019	767.58199	1.433	$[M+HCOO]^-$	PC O-16:0_20:4	$C_{44}H_{82}NO_7P$	898.5	-1.129	0.5
2.05	286.6	762.56517	717.56697	1.406	$[M+HCOO]^-$	PC O-16:1_16:0	$C_{40}H_{80}NO_7P$	903.6	0.013	1.5

2.08	290.2	788.58063	743.58243	1.424	$[M+HCOO]^-$	PC O-16:1_18:1	$C_{42}H_{82}NO_7P$	899.1	-0.587	0.6
1.89	288.5	786.56519	741.56699	1.416	$[M+HCOO]^-$	PC O-16:1_18:2	$C_{42}H_{80}NO_7P$	921.6	-0.355	0.6
2.13	297.1	840.61164	795.61344	1.459	$[M+HCOO]^-$	PC O-18:0_20:4	$C_{46}H_{86}NO_7P$	911.1	-0.883	0.7
2.11	290	790.5969	745.59869	1.423	$[M+HCOO]^-$	PC O-18:1_16:0	$C_{42}H_{84}NO_7P$	920.9	-0.028	0.2
2.15	294.1	816.61309	771.61489	1.444	$[M+HCOO]^-$	PC O-18:1_18:1	$C_{44}H_{86}NO_7P$	644.5	0.214	0.3
2.12	294.3	814.59648	769.59828	1.445	$[M+HCOO]^-$	PC O-18:1_18:2	$C_{44}H_{84}NO_7P$	830.4	-0.324	0.8
1.9	296	838.59605	793.59785	1.454	$[M+HCOO]^-$	PC O-18:1_20:4	$C_{46}H_{84}NO_7P$	937.6	-0.873	0.6
0.69	242.7	550.3865	549.37922	1.182	$[M+H]^+$	PC O-20:1	$C_{28}H_{56}NO_7P$	975.5	-0.453	0.3
2.09	289.3	720.58895	719.58167	1.417	$[M+H]^+$	PC O-32:0	$C_{40}H_{82}NO_7P$	994.9	-1.707	1.8
1.89	285.8	718.57381	717.56653	1.4	$[M+H]^+$	PC O-32:1_A	$C_{40}H_{80}NO_7P$	999.8	-0.951	1.3
2.05	285.9	718.57358	717.5663	1.401	$[M+H]^+$	PC O-32:1_B	$C_{40}H_{80}NO_7P$	997.2	-1.228	1.4
2.11	293	746.60479	745.59752	1.436	$[M+H]^+$	PC O-34:1	$C_{42}H_{84}NO_7P$	996.2	-1.379	1.7
1.93	288.4	744.58827	743.58099	1.414	$[M+H]^+$	PC O-34:2_A	$C_{42}H_{82}NO_7P$	993.3	-2.515	0.8
2.08	289.3	744.58933	743.58205	1.418	$[M+H]^+$	PC O-34:2_B	$C_{42}H_{82}NO_7P$	988.5	-1.104	1.1
1.89	285.4	742.57371	741.56644	1.399	$[M+H]^+$, ($[M+Na]^+$)	PC O-34:3	$C_{42}H_{80}NO_7P$	999.5	-1.04	0.3
2.35	298.2	774.63669	773.62941	1.463	$[M+H]^+$	PC O-36:1	$C_{44}H_{88}NO_7P$	990.8	-0.545	1.8
2.15	295.4	772.6206	771.61332	1.449	$[M+H]^+$	PC O-36:2_A	$C_{44}H_{86}NO_7P$	996.7	-1.193	1.3
2.31	296	772.62094	771.61367	1.452	$[M+H]^+$	PC O-36:2_B	$C_{44}H_{86}NO_7P$	995.7	-0.613	1.5
1.96	293	770.6038	769.59653	1.437	$[M+H]^+$	PC O-36:3_A	$C_{44}H_{84}NO_7P$	997.2	-2.637	1.1
2.12	293.6	770.6047	769.59742	1.44	$[M+H]^+$	PC O-36:3_B	$C_{44}H_{84}NO_7P$	996.4	-1.408	1.4
1.89	291.5	768.5885	767.58111	1.43	$[M+H]^+$, ($[M+Na]^+$)	PC O-36:4	$C_{44}H_{82}NO_7P$	998	-2.242	1.3
2.13	297.6	796.62038	795.61305	1.461	$[M+H]^+$, ($[M+Na]^+$)	PC O-38:4	$C_{46}H_{86}NO_7P$	994.3	-1.473	1.7
1.91	295.2	794.60455	793.59727	1.449	$[M+H]^+$	PC O-38:5	$C_{46}H_{84}NO_7P$	995.6	-1.59	1.2
2.36	303.4	824.65167	823.6444	1.49	$[M+H]^+$	PC O-40:4	$C_{48}H_{90}NO_7P$	999.2	-1.326	2
2.13	300.9	822.63561	821.62833	1.478	$[M+H]^+$	PC O-40:5	$C_{48}H_{88}NO_7P$	991.2	-1.829	1.4
2.06	300.5	820.61974	819.61246	1.476	$[M+H]^+$	PC O-40:6	$C_{48}H_{86}NO_7P$	999.3	-2.15	1.5
2.58	309.7	852.68442	851.67714	1.522	$[M+H]^+$	PC O-42:4	$C_{50}H_{94}NO_7P$	994.6	0.364	1.5
2.35	306.9	850.66767	849.66039	1.508	$[M+H]^+$	PC O-42:5	$C_{50}H_{92}NO_7P$	998.4	-0.786	0.9
2.57	312.5	878.69857	877.69129	1.536	$[M+H]^+$	PC O-44:5	$C_{52}H_{96}NO_7P$	997.1	-1.266	0.9
2.38	310.8	876.68247	875.67519	1.528	$[M+H]^+$	PC O-44:6	$C_{52}H_{94}NO_7P$	998.1	-1.964	0.9
2.26	284.3	746.56838	745.56111	1.394	$[M+H]^+$	PE 18:0_18:1	$C_{41}H_{80}NO_8P$	922.5	-1.312	1.9

2.08	281.2	744.55307	743.54579	1.379	$[M+H]^+$	PE 18:0_18:2	$C_{41}H_{78}NO_8P$	942.3	-1	1.7
2.08	272.3	742.53884	743.54611	1.335	$[M-H]^-$	PE 18:2_18:0	$C_{41}H_{78}NO_8P$	724.4	-0.536	0.7
2.37	283	730.57309	729.56581	1.387	$[M+H]^+$	PE P -18:0 _18:1	$C_{41}H_{80}NO_7P$	895.3	-1.389	2.1
1.69	289.8	835.53412	836.5414	1.424	$[M-H]^-$	PI 16:0_18:1	$C_{43}H_{81}O_{13}P$	726.4	-0.151	2.6
1.52	286.8	833.51837	834.52564	1.409	$[M-H]^-$	PI 16:0_18:2	$C_{43}H_{79}O_{13}P$	601.6	-0.23	2.1
1.9	294.7	863.56516	864.57243	1.449	$[M-H]^-$	PI 18:0_18:1	$C_{45}H_{85}O_{13}P$	538.7	-0.389	2.5
1.74	294	861.54933	862.5566	1.445	$[M-H]^-$	PI 18:0_18:2	$C_{45}H_{83}O_{13}P$	585.1	-0.628	2.9
1.71	297.1	885.54924	886.55651	1.461	$[M-H]^-$	PI 18:0_20:4	$C_{47}H_{83}O_{13}P$	696	-0.647	2.8
1.5	300.1	876.55881	858.52498	1.475	$[M+NH4]^+$	PI 36:4	$C_{45}H_{79}O_{13}P$	806.4	-0.964	3.1
1.72	305.4	904.5903	886.55593	1.502	$[M+NH4]^+$, ($[M+Na]^+$, $[M+H]^+$)	PI 38:4	$C_{47}H_{83}O_{13}P$	888.9	-0.66	3.5
2.62	312.7	841.71316	840.70588	1.536	$[M+H]^+$	SM 44:2;2O	$C_{49}H_{97}N_2O_6P$	711.0	-3.266	1
1.54	287.1	727.57403	726.56675	1.407	$[M+H]^+$	SM 36:3;2O	$C_{41}H_{79}N_2O_6P$	873.0	-1.127	0.4
1.33	278.5	661.52752	660.52024	1.362	$[M+H]^+$	SM 31:1;2O	$C_{36}H_{73}N_2O_6P$	862.0	-0.571	2.4
2.09	301.3	797.65165	796.64437	1.479	$[M+H]^+$	SM 41:3;2O	$C_{46}H_{89}N_2O_6P$	828.0	-1.839	0.3
1.22	275.7	647.51194	646.50466	1.348	$[M+H]^+$	SM 30:1;2O	$C_{35}H_{71}N_2O_6P$	939.0	-0.538	2.8
1.45	281.7	675.54363	674.53594	1.378	$[M+H]^+$, ($[M+Na]^+$)	SM 18:1;2O/14:0	$C_{37}H_{75}N_2O_6P$	996.0	0.729	1.5
2.18	298.9	759.63664	758.62923	1.466	$[M+H]^+$, ($[M+Na]^+$, $[M+K]^+$)	SM 16:1;2O/22:0	$C_{43}H_{87}N_2O_6P$	994.1	-0.949	1.5
1.57	284.5	689.55891	688.55145	1.393	$[M+H]^+$, ($[M+Na]^+$)	SM 33:1;2O	$C_{38}H_{77}N_2O_6P$	840.0	-0.422	1.4
2.31	303.8	799.6679	798.66017	1.491	$[M+H]^+$, ($[M+Na]^+$)	SM 18:2;2O/23:0	$C_{46}H_{91}N_2O_6P$	996.0	-0.996	1.1
1.69	286.6	703.57539	702.56811	1.403	$[M+H]^+$	SM 18:1;2O/16:0	$C_{39}H_{79}N_2O_6P$	995.3	0.761	0.5
1.93	293.4	731.60547	730.59842	1.438	$[M+H]^+$, ($[M+Na]^+$)	SM 18:1;2O/18:0	$C_{41}H_{83}N_2O_6P$	993.5	-0.943	0.9
2.41	302.9	787.66808	786.66071	1.486	$[M+H]^+$, ($[M+Na]^+$, $[M+K]^+$)	SM 18:1;2O/22:0	$C_{45}H_{91}N_2O_6P$	996.7	-0.997	0.7
2.52	306.4	801.68384	800.67633	1.504	$[M+H]^+$, ($[M+Na]^+$)	SM 18:1;2O/23:0	$C_{46}H_{93}N_2O_6P$	996.7	-0.632	0.9
2.63	309.2	815.69944	814.69164	1.518	$[M+H]^+$, ($[M+Na]^+$)	SM 18:1;2O/24:0	$C_{47}H_{95}N_2O_6P$	996.9	-0.56	0.9
2.4	306.6	813.68454	812.67661	1.505	$[M+H]^+$, ($[M+Na]^+$, $[M+K]^+$)	SM 18:1;2O/24:1	$C_{47}H_{93}N_2O_6P$	997.3	0.17	1.1
1.49	283.8	701.55948	700.55263	1.39	$[M+H]^+$, ($[M+Na]^+$, $[M+K]^+$)	SM 18:2;2O/16:0	$C_{39}H_{77}N_2O_6P$	996.5	0.608	0.9
2.19	301.2	785.6522	784.64519	1.478	$[M+H]^+$, ($[M+Na]^+$, $[M+K]^+$)	SM 18:2;2O/22:0	$C_{45}H_{89}N_2O_6P$	996.2	-1.103	1.6
1.8	290.3	717.58978	716.5825	1.422	$[M+H]^+$	SM 35:1;2O	$C_{40}H_{81}N_2O_6P$	744.0	-1.011	0.4
2.48	309.4	827.69882	826.69155	1.519	$[M+H]^+$	SM 43:2;2O	$C_{48}H_{95}N_2O_6P$	886.0	-1.484	0.4
1.25	277.2	673.52752	672.52024	1.357	$[M+H]^+$	SM 32:2;2O	$C_{37}H_{73}N_2O_6P$	732.0	-0.561	0.8
1.61	286.1	715.57424	714.56696	1.402	$[M+H]^+$	SM 18 :2;2O/17 :0	$C_{40}H_{79}N_2O_6P$	998.0	-0.531	0.17

1.72	288.7	729.58967	728.58231	1.415	$[M+H]^+$, ($[M+Na]^+$)	SM 36:2;2O	$C_{41}H_{81}N_2O_6P$	816.0	-1.077	0.5
1.53	285.2	677.55861	676.55133	1.396	$[M+H]^+$	SM 32:0;2O	$C_{37}H_{77}N_2O_6P$	869.0	-0.875	2.3
2.05	295.7	745.62077	744.61349	1.45	$[M+H]^+$	SM 37:1;2O	$C_{42}H_{85}N_2O_6P$	855.0	-1.452	0.8
2	301.1	809.65192	808.64464	1.478	$[M+H]^+$	SM 42:4;2O	$C_{47}H_{89}N_2O_6P$	801.0	-1.51	0.2
1.77	290.5	705.59008	704.58281	1.423	$[M+H]^+$	SM 34:0;2O	$C_{39}H_{81}N_2O_6P$	824.0	-0.557	1.2
1.97	295.3	757.62082	756.61354	1.448	$[M+H]^+$	SM 38:2;2O	$C_{43}H_{85}N_2O_6P$	867.0	-1.28	0.3
2.3	301.1	773.65203	772.64302	1.477	$[M+H]^+$, ($[M+Na]^+$)	SM 39:1;2O	$C_{44}H_{89}N_2O_6P$	862.0	-1.314	0.1
2.2	303.9	811.66802	810.66074	1.492	$[M+H]^+$	SM 42:3;2O	$C_{47}H_{91}N_2O_6P$	885.0	-0.876	0.8
2.49	305.5	789.68137	788.67409	1.499	$[M+H]^+$	SM 40:0;2O	$C_{45}H_{93}N_2O_6P$	803.0	-3.717	0.7
2.31	306.9	825.68297	824.67569	1.507	$[M+H]^+$	SM 43:3;2O	$C_{48}H_{93}N_2O_6P$	745.0	-1.749	0.3
2.71	311.2	829.71423	828.70695	1.529	$[M+H]^+$	SM 43:1;2O	$C_{48}H_{97}N_2O_6P$	885.0	-1.874	0.8
1.61	294	719.569	718.56172	1.441	$[M+H]^+$	SM 34:1;3O	$C_{39}H_{79}N_2O_7P$	838.9	-0.954	3.2
2.29	309.8	831.69295	830.68567	1.522	$[M+H]^+$	SM 42:1;3O	$C_{47}H_{95}N_2O_7P$	894.1	-2.497	1
2.32	310.5	829.67813	828.67085	1.525	$[M+H]^+$	SM 42:2;3O	$C_{47}H_{93}N_2O_7P$	887.1	-1.454	1.8
3.14	306.1	794.72192	776.68764	1.503	$[M+NH4]^+$, ($[M+K]^+$, $[M+Na]^+$)	TG 12:0_16:0_18:1	$C_{49}H_{92}O_6$	814.1	-1.536	3
3.21	312.5	822.75334	804.71952	1.535	$[M+NH4]^+$	TG 14:0_16:0_18:1	$C_{51}H_{96}O_6$	899.4	-1.378	2.4
3.15	309.8	820.73802	802.70333	1.521	$[M+NH4]^+$, ($[M+K]^+$, $[M+Na]^+$)	TG 14:0_16:0_18:2	$C_{51}H_{94}O_6$	700.2	-1.015	1.8
3.09	307.8	818.72207	800.68823	1.511	$[M+NH4]^+$, ($[M+Na]^+$, $[M+K]^+$)	TG 14:0_16:0_18:3	$C_{51}H_{92}O_6$	234.7	-1.415	1.9
3.03	305.3	816.70584	798.67301	1.499	$[M+NH4]^+$, ($[M+K]^+$, $[M+Na]^+$)	TG 12:0_18:2_18:2	$C_{51}H_{90}O_6$	798.6	-2.104	2
3.1	311.5	844.73744	826.70361	1.53	$[M+NH4]^+$	TG 14:0_18:2_18:2	$C_{53}H_{94}O_6$	881.8	-1.697	2.6
3.23	315.2	836.7682	818.73436	1.548	$[M+NH4]^+$, ($[M+Na]^+$)	TG 15:0_16:0_18:1	$C_{52}H_{98}O_6$	870.2	-2.254	2.4
3.18	311	834.75329	816.71856	1.528	$[M+NH4]^+$, ($[M+Na]^+$, $[M+K]^+$)	TG 15:0_16:0_18:2	$C_{52}H_{96}O_6$	822.8	-1.462	1.7
3.18	316.9	860.7691	842.73504	1.557	$[M+NH4]^+$, ($[M+Na]^+$)	TG 15:0_18:1_18:2	$C_{54}H_{98}O_6$	875.9	-1.261	2.1
3.26	314.2	824.76757	806.73346	1.543	$[M+NH4]^+$, ($[M+Na]^+$)	TG 16:0_16:0_16:0	$C_{51}H_{98}O_6$	945.8	-3.022	0.8
3.27	318.2	850.78402	832.75074	1.563	$[M+NH4]^+$, ($[M+K]^+$, $[M+Na]^+$)	TG 16:0_16:0_18:1	$C_{53}H_{100}O_6$	951.1	-2.093	2
3.21	316.1	848.76959	830.73453	1.553	$[M+NH4]^+$, ($[M+Na]^+$, $[M+K]^+$)	TG 16:0_16:1_18:1	$C_{53}H_{98}O_6$	890.4	-0.683	2.3
3.15	313.6	846.75418	828.7194	1.541	$[M+NH4]^+$, ($[M+Na]^+$, $[M+K]^+$)	TG 16:0_16:1_18:2	$C_{53}H_{96}O_6$	834.1	-0.342	2
3.33	322.6	866.81221	848.77838	1.586	$[M+NH4]^+$	TG 16:0_17:0_18:0	$C_{54}H_{104}O_6$	911.2	-5.186	1
3.29	321.3	864.79978	846.76641	1.579	$[M+NH4]^+$, ($[M+Na]^+$)	TG 16:0_17:0_18:1	$C_{54}H_{102}O_6$	914.7	-1.931	2
3.24	318.4	862.78413	844.75026	1.565	$[M+NH4]^+$, ($[M+Na]^+$)	TG 16:0_17:1_18:1	$C_{54}H_{100}O_6$	803.4	-1.957	2.2
3.31	324.5	878.81702	860.78288	1.595	$[M+NH4]^+$, ($[M+Na]^+$, $[M+K]^+$)	TG 16:0_18:0_18:1	$C_{55}H_{104}O_6$	939.3	0.216	2.2

3.27	321.8	876.8005	858.76614	1.582	$[M+NH4]^+$, ($[M+Na]^+$, $[M+K]^+$)	TG 16:0_18:1_18:1	$C_{55}H_{102}O_6$	950.2	-1.105	2.5
3.22	319.8	874.78542	856.75073	1.572	$[M+NH4]^+$, ($[M+Na]^+$, $[M+K]^+$)	TG 16:0_18:1_18:2	$C_{55}H_{100}O_6$	954.5	-0.385	2.5
3.17	317.6	872.76989	854.7354	1.561	$[M+NH4]^+$, ($[M+Na]^+$, $[M+K]^+$)	TG 16:0_18:2_18:2	$C_{55}H_{98}O_6$	925.8	-0.298	2.9
3.11	315.7	870.75263	852.71858	1.552	$[M+NH4]^+$, ($[M+Na]^+$, $[M+K]^+$)	TG 16:0_18:2_18:3	$C_{55}H_{96}O_6$	882.5	-1.996	2.6
3.29	325.2	890.81448	872.78158	1.599	$[M+NH4]^+$, ($[M+Na]^+$)	TG 17:0_18:1_18:1	$C_{56}H_{104}O_6$	863.4	-2.932	1.8
3.24	322.5	888.79993	870.76611	1.586	$[M+NH4]^+$	TG 17:0_18:1_18:2	$C_{56}H_{102}O_6$	823.7	-1.719	2.6
3.19	321.5	886.78449	868.75066	1.581	$[M+NH4]^+$	TG 17:1_18:1_18:2	$C_{56}H_{100}O_6$	844.7	-1.241	3
3.32	327.9	904.8322	886.79764	1.613	$[M+NH4]^+$, ($[M+Na]^+$, $[M+K]^+$)	TG 18:0_18:0_18:2	$C_{57}H_{106}O_6$	431.7	-0.522	3.1
3.27	325.7	902.8153	884.78161	1.602	$[M+NH4]^+$, ($[M+K]^+$, $[M+Na]^+$)	TG 18:1_18:1_18:1	$C_{57}H_{104}O_6$	947.6	-1.882	2.4
3.22	323.8	900.79965	882.76303	1.592	$[M+NH4]^+$, ($[M+Na]^+$, $[M+K]^+$)	TG 18:1_18:1_18:2	$C_{57}H_{102}O_6$	914	-2.03	2.8
3.31	331	930.84521	912.81204	1.628	$[M+NH4]^+$, ($[M+Na]^+$)	TG 18:1_18:1_20:1	$C_{59}H_{108}O_6$	909	-3.324	2.6
3.19	327.9	924.79945	906.76419	1.613	$[M+NH4]^+$, ($[M+Na]^+$)	TG 18:1_18:1_20:4	$C_{59}H_{102}O_6$	630.3	-2.273	2.5
3.17	321.7	898.7841	880.74922	1.582	$[M+NH4]^+$, ($[M+Na]^+$, $[M+K]^+$)	TG 18:1_18:2_18:2	$C_{57}H_{100}O_6$	901.1	-1.818	2.3
3.15	326	922.78493	904.75111	1.604	$[M+NH4]^+$	TG 18:1_18:2_20:4	$C_{59}H_{100}O_6$	704.8	-1.002	2.5
3.35	335.4	958.87774	940.84391	1.651	$[M+NH4]^+$	TG 18:1_18:2_22:0	$C_{61}H_{112}O_6$	917.9	-2.076	2.2
3.39	338	986.90901	968.87518	1.664	$[M+NH4]^+$	TG 18:1_18:2_24:0	$C_{63}H_{116}O_6$	840.9	-2.013	1.9
3.12	319.6	896.7692	878.73433	1.572	$[M+NH4]^+$, ($[M+Na]^+$, $[M+K]^+$)	TG 18:2_18:2_18:2	$C_{57}H_{98}O_6$	936.1	-1.124	2.6
3.28	329.1	928.82878	910.79495	1.619	$[M+NH4]^+$	TG 18:2_18:2_20:0	$C_{59}H_{106}O_6$	323.7	-4.33	2.4
3.31	333.7	956.86203	938.82821	1.642	$[M+NH4]^+$	TG 18:2_18:2_22:0	$C_{61}H_{110}O_6$	777.5	-1.947	2.2
3.35	320.4	862.82002	844.7877	1.575	$[M+NH4]^+$, ($[M+Na]^+$)	TG O-18:1_16:0_18:1	$C_{55}H_{104}O_5$	856.6	-2.247	2.5

Table S6. List of annotated lipids by RP-UHPLC-TIMS-Q-TOF in HCC, HCV and MC plasma samples. Lipids are sorted by class. Lipid annotation was carried out with a combined rule based and spectral library approach following the lipidomic standard initiative guidelines[4]. Abbreviations: RT, retention time; m/z meas, mass to charge measured; Mass meas., measured mass; CCS, cross collisional section; Mob., mobility. Primary ion has been used for normalization and additional adducts, if detected, were reported in round brackets. Lipid class abbreviations: cholesteryl esters (CEs), ceramides (Cers), diacylglycerols (DGs), dihexosylceramide (Hex2cer), hexosylceramides (Hexcers), lysophosphatidylcholines (LPCs), ether-linked lysophosphatidylcholines (LPC-Os), lysophosphatidylethanolamines (LPEs), phosphatidylcholines (PCs), ether-linked phosphatidylcholine (PC-Os) phosphatidylethanolamines (PEs), ether-linked phosphatidylethanolamine (PE-Os), phosphatidylinositols (PIs), sphingomyelins (SMs), triacylglycerols (TGs).

Section S1: Lipid and Metabolite annotation

In this section some examples of lipid annotation workflow are proposed:

Sphingomyelins: 35 SMs were annotated, all of them were detected in positive mode and in the protonated form ([M+H]+), as indicated by by https://lipidomicstandards.org/lipid-class-specific-fragments/.

Few SMs (such as: SM 18:1;2O/24:1, SM 18:1;2O/22:0 etc) were detected also in [M+Na]+, [M+K]+ forms, but these were not the primary adducts or the adducts used to normalize them.

We annotated, at the end, 13 SMs (**Figure S6**) with 1 Double Bond (DB) and we checked that all of them followed the equivalent carbon number theory and that CCS values increased linearly with measured m/z. This strategy is useful to easily identify outliers, indeed at the beginning our dataset included 14 SMs (upper panel) with 1 DB and we discarded the SM with a total of 42 carbons eluting at 2.48 min (red circle). In this way our R2 improved from 0.9925 (upper panel) to 0.999 (lower panel), indicating a strong linearity and the respect of ECN model.

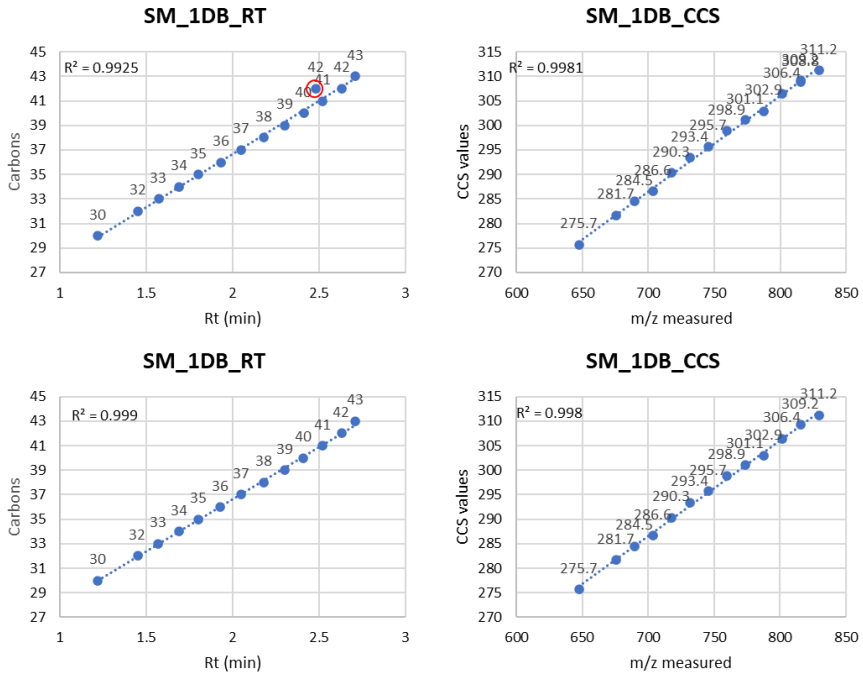


Figure S6. Sphingomyelins rt/carbon number and m/z measured/CCS linearity.

We here report some examples that we followed for short/long hand SMs' annotations.

If the MS2 spectrum contained only the phosphocoline head group (m/z 184) we annotated the lipid in its short-hand form, as for the SM 36:2;2O (**Figure S7).**

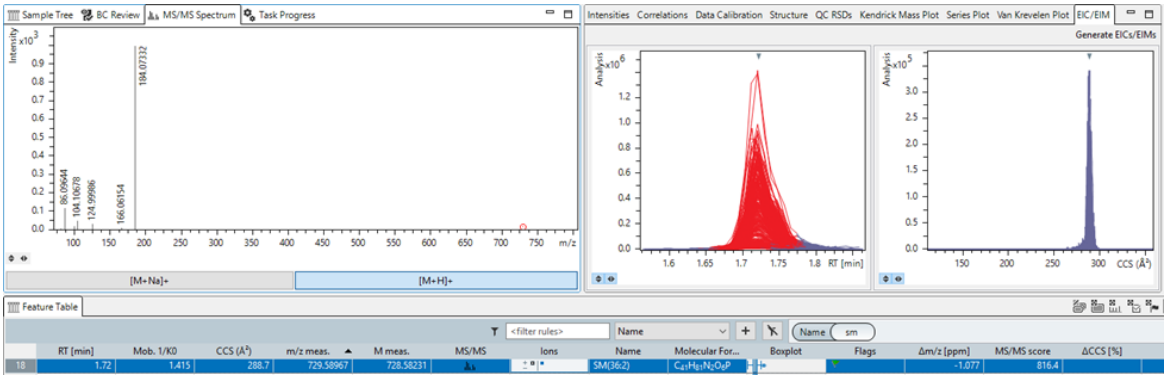


Figure S7. Short-hand annotation for SMs.

On the other hand, if our spectra contained additional fragments related to the lipid backbone we proposed it in the long-hand form, as for the SM 18:1;2O/22:0 where we observed a fragment at m/z 264.26932 due to LCB 18:1;2(-H3O2) (**Figure S8).**

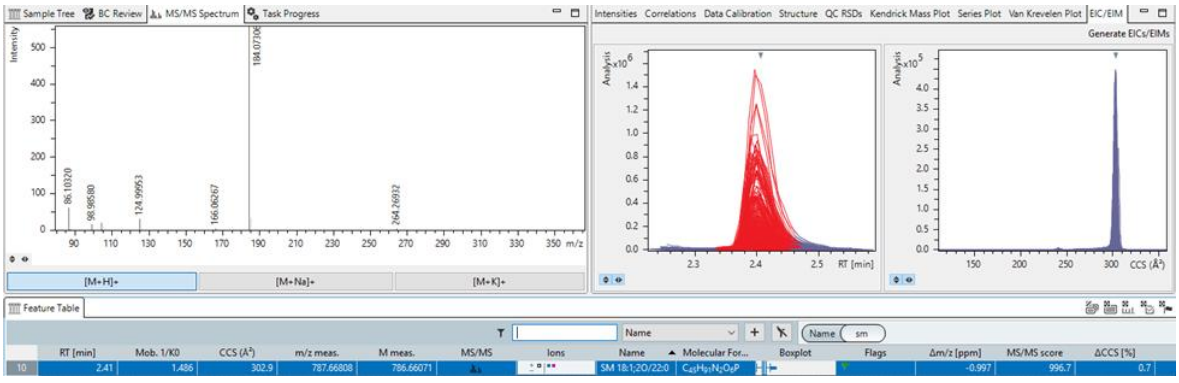


Figure S8. Long-hand annotation for SMs.

Below we report the order of elution and the CCS values of phosphocolines with 0, 1, 2 and 3 DBs in ESI+, plotting in this way 29 out of the 64 PCs (considering both polarities) reported at the end **(figure S9).**

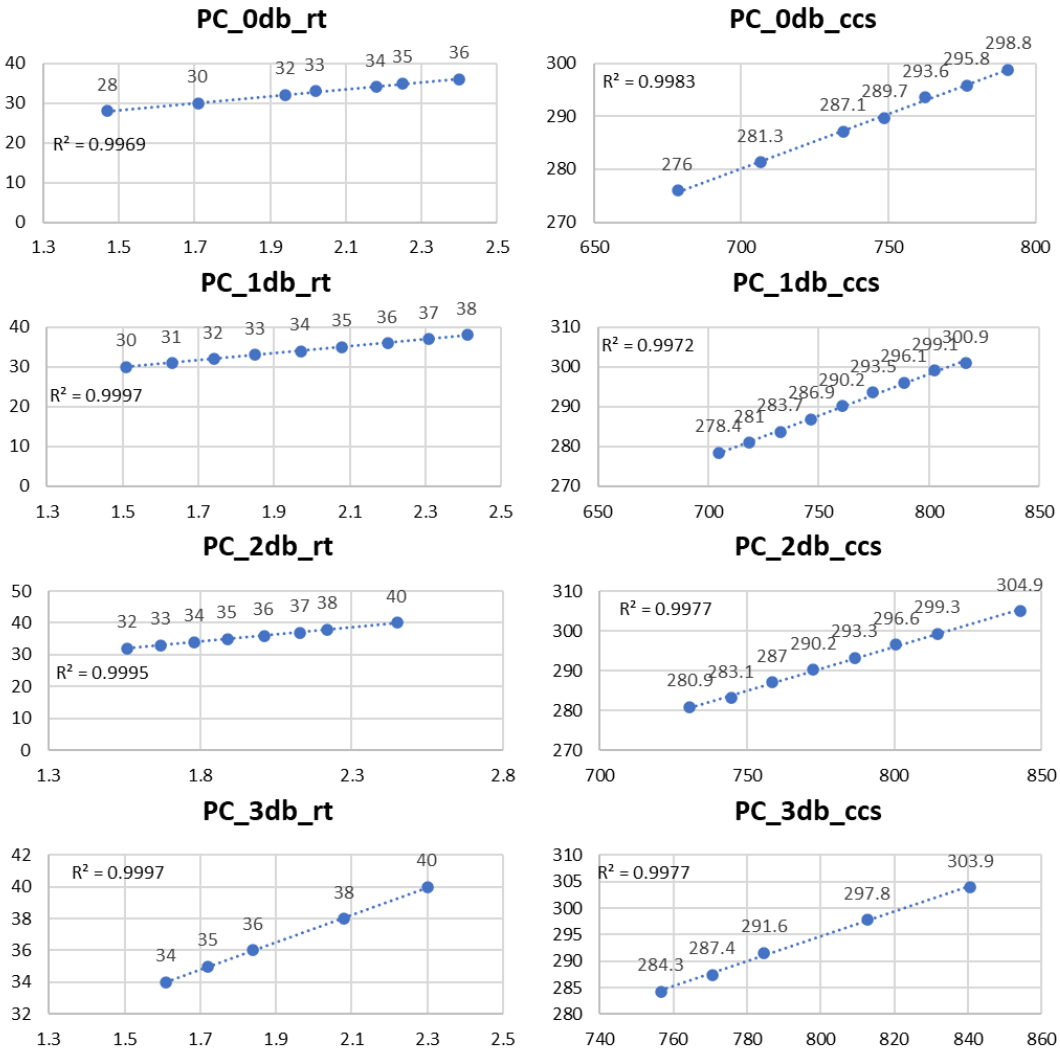


Figure S9. Phosphatidylcholines rt/carbon number and m/z measured /CCS linearity.

Also, for PCs we inspected the presence of fragments related to the acyl chains composing lipids. In the proposed case, PC 18:0_20:4, in addition to the m/z 184.07332 we have additional fragments at m/z 506.35904 and 524.36600 due to the NL of 20:4 chain and at m/z 544.33252 and 526.32088 conducible to the NL of 18:0 chain (**Figure S10**)

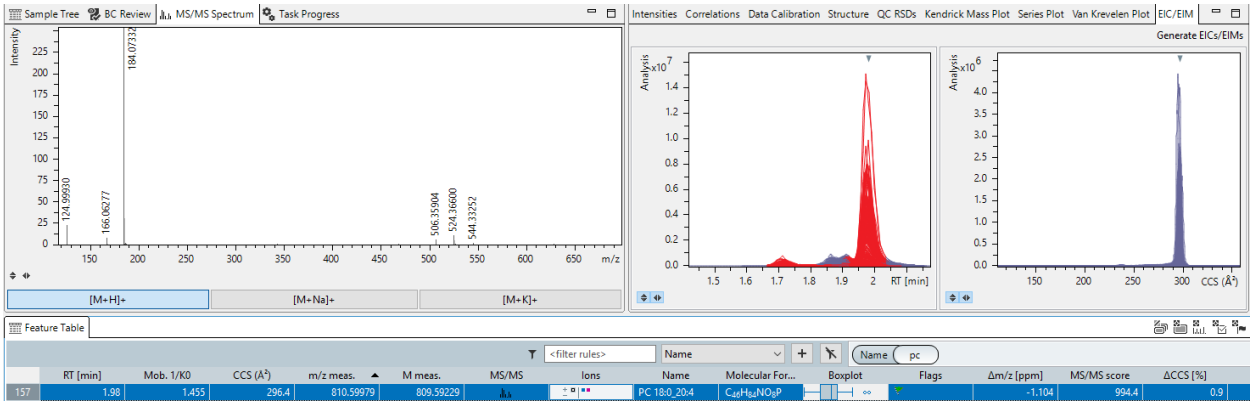


Figure S10. Long-hand annotation for PCs.

PCs have been listed in the short-hand form if the MS2 reported only the phosphocholine head group (m/z 184) in positive ionization mode.

Finally, we report some MS/MS library matches for LPC O-16:0, LPC 18:1, CAR 4:0, and ADMA.

In the MS2 spectrum of the LPC 18:1 and LPC O-16:0 are well present fragments at m/z 86.099, 184.074 and 104.108 due respectively to: deydrocholine, phosphocoline and choline.

CAR is characterized by the fragment at 85 m/z (C4H5O2+) which is the most significant feature peak of acylcarnitine spectrum. For ADMA it can be observed the characteristic fragmentation pattern of Arg, with ions at 70, 116 m/z, plus the ion at 158 m/z, which fits with reported fragmentation pattern and CCS (https://pubchem.ncbi.nlm.nih.gov/compound/N_N-dimethylarginine#section=Other-MS) even if the fragment at m/z 46 was not detected for the MS range employed (50-1500) (**Figure S11**).

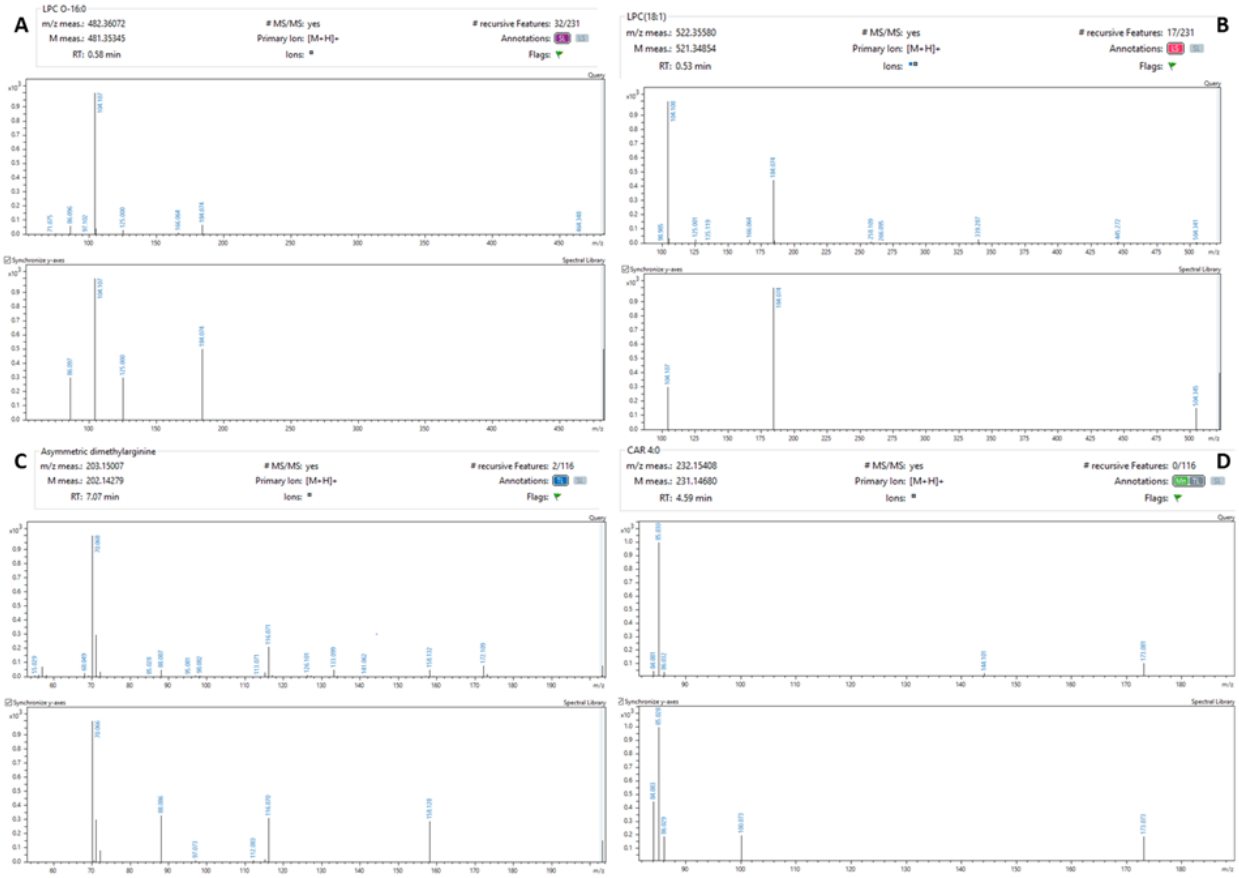


Figure S11. Comparison between the acquired (Query) and the library spectrum (Spectral library) of statistically significant lipids and metabolites (A: LPC O-16:0; B: LPC 18:1; C: Asymmetric dimethylarginine; D: CAR 4:0).

Supplementary Material Chapter IV

	$FWHM_{avg}$							
	CAR C16:0	**CAR C16:0-D9**	**CAR C5:0**	**CAR C5:0-D9**	**CAR C8:0**	**CAR C8:0-D9**	**LPC 17:0**	**LPC 17:0-D5**
AmF + FA	0.107	0.107	0.171	0.109	0.1055	0.1245	0.1075	0.086
FA	0.1615	0.1605	0.1085	0.1015	0.1235	0.1205	0.107	0.056
AmAc + AcA	0.1295	0.106	0.156	0.1325	0.144	0.1405	0.149	0.0485
AmAc	0.0905	0.0935	0.105	0.1115	0.1005	0.095	0.1055	0.0385

Tab S1 Comparative assessment of Full Width at Half Maximum (FWHM) under tested mobile phase conditions.

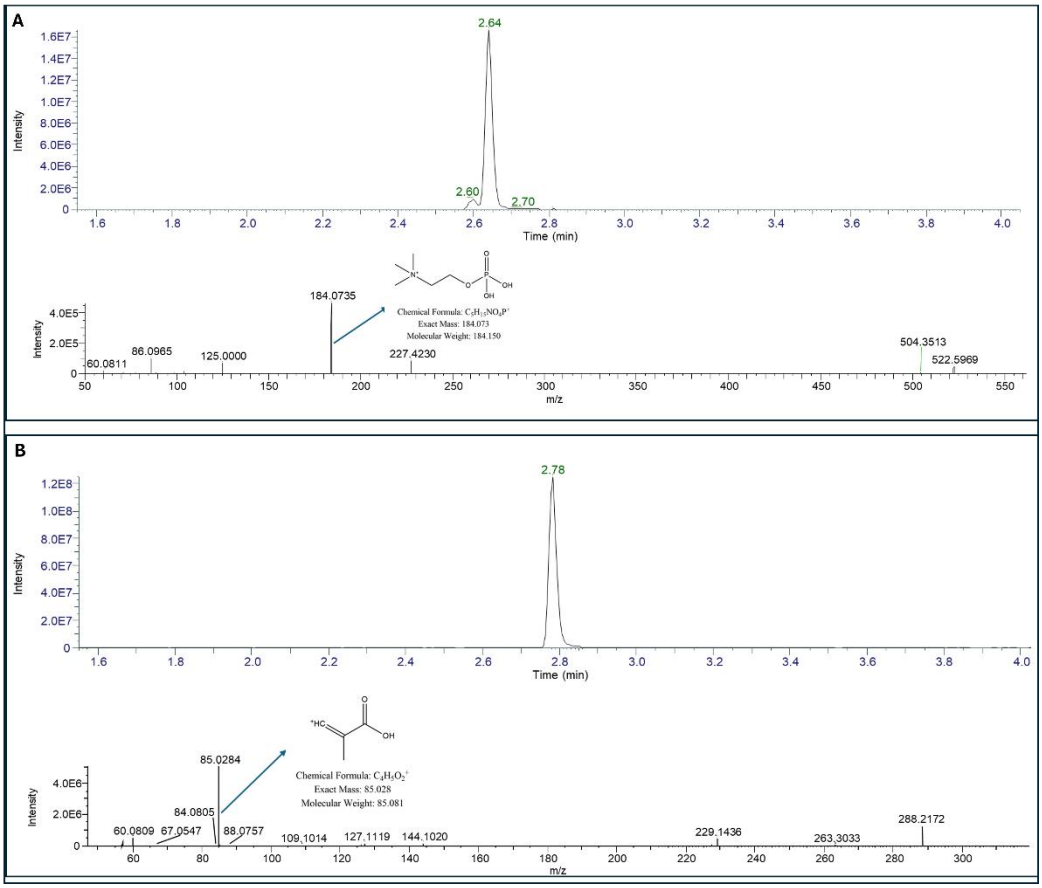


Fig. **S1 A-B** Targeted MS/MS spectra of LPCs and CARs. A. LPCs were confirmed by the diagnostic phosphocholine fragment at m/z 184.0733 ($C_5H_{15}NO_4P^+$), while CARs were validated by the characteristic fragment at m/z 85.0284 ($C_4H_5O_2^+$).

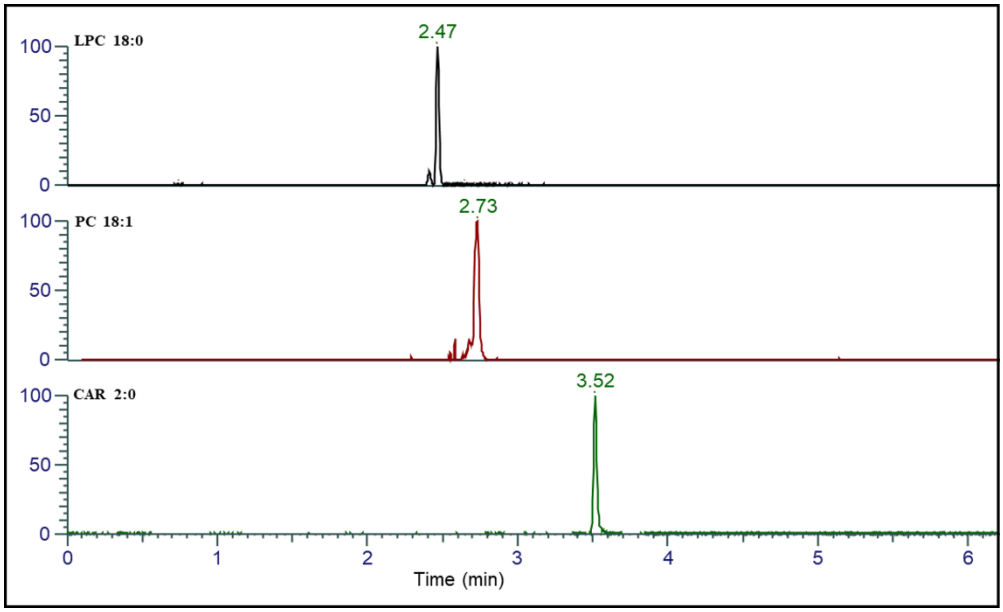


Fig. **S2** Elution window for LPC 18:0, PC 18:1 and CAR C2:0 detected in human plasma.

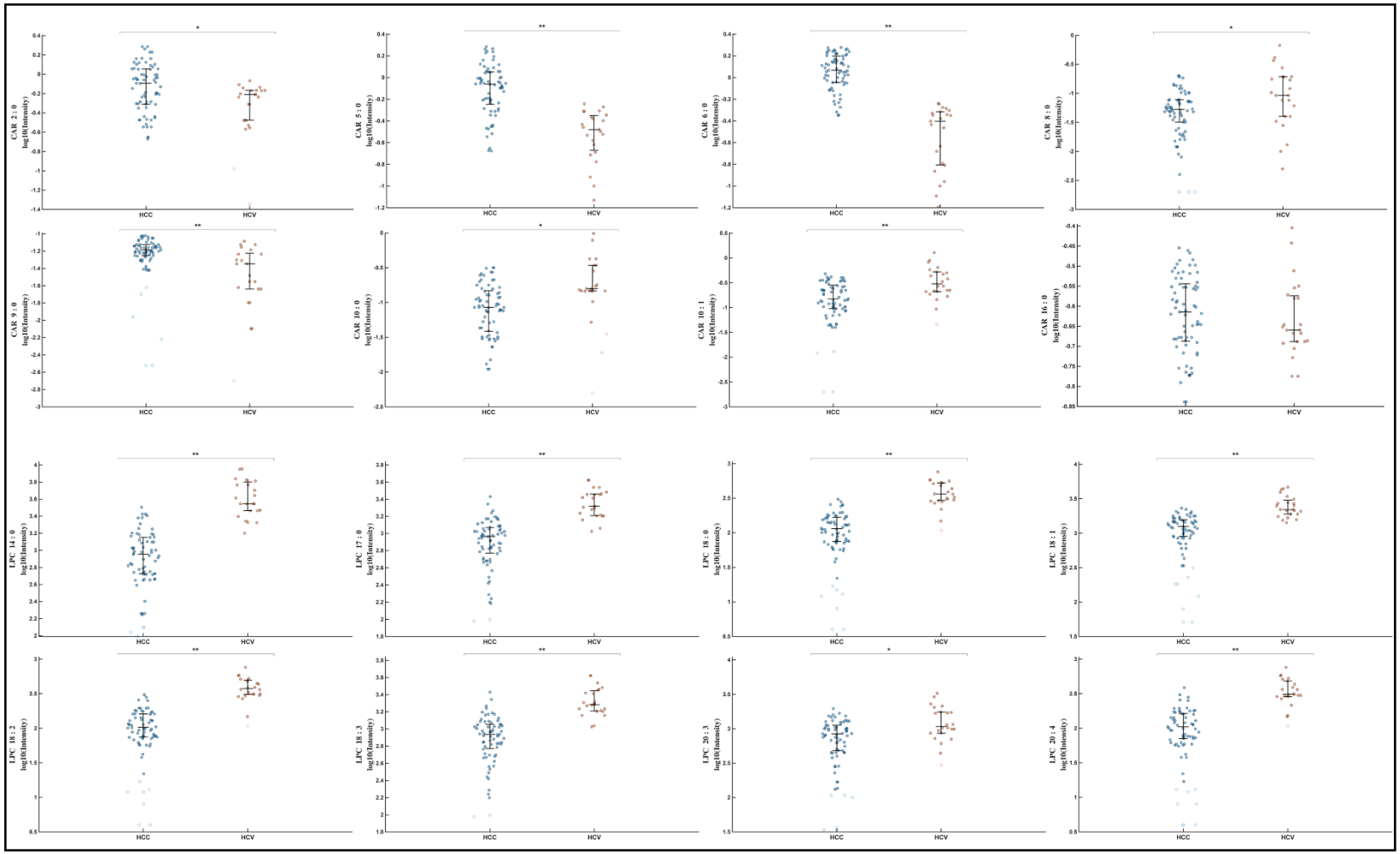


Fig. S3 Scatter plots for individual metabolites report log10 transformed concentration values. Statistical differences between HCC and HCV groups were assessed using the non-parametric Mann - Whitney U test followed by FDR correction. * pvalue < 0.05; ** pvalue < 0.001.

Supplementary Material Chapter V

Deuterated standards	**Molecular formula**
Butyric Acid-d7	$C_4HD_7O_2$
L-Carnitine-d9	$C_7H_7D_9NO_3.Cl$
L-Glutamic Acid-d5	$C_5H_4D_5NO_4$
L-Lysine-d3	$C_6H_{11}D_3N_2O_2.HCl$
L-Tryptophan-d5	$C_{11}H_7D_5N_2O_2$
Succinic Acid-d4	$C_4H_2D_4O_4$
Taurine-d4	$C_2H_3D_4NO_3S$
Taurohycholic Acid-d4	$C_{26}H_{40}D_4NO_7S.Na$
Endogenous standards	**Molecular formula**
2'-Deoxyadenosine-5'-monophosphate	$C_{10}H_{14}N_5O_6P$
7-Methylguanine	$C_6H_7N_5O$
Adenosine monophosphate (AMP)	$C_{10}H_{14}N_5O_7P$
Alanine	$C_3H_7NO_2$
Arginine	$C_6H_{14}N_4O_2$
Asparagine	$C_4H_8N_2O_3$
Asymmetric Dimethyl Arginine (ADMA)	$C_8H_{18}N_4O_2$
Bilirubin	$C_{33}H_{36}N_4O_6$
Creatine	$C_4H_9N_3O_2$
Glutamic acid	$C_5H_9NO_4$
Glutamine	$C_5H_{10}N_2O_3$
Glutathione RED	$C_{10}H_{17}N_3O_6S$
Histidine	$C_6H_9N_3O_2$
Indolebutyric acid	$C_{12}H_{13}NO_2$
Kynurenine	$C_{10}H_{12}N_2O_3$
L-Alany-L-glutamine	$C_8H_{15}N_3O_4$
L-Alanyl-L-phenylalanine	$C_{12}H_{16}N_2O_3$
L-Carnitine	$C_7H_{15}NO_3$
L-Carnosine	$C_9H_{14}N_4O_3$
L-Citrulline	$C_6H_{13}N_3O_3$
L-Homocitrulline	$C_7H_{15}N_3O_3$
L-Pyroglutamic acid	$C_5H_7NO_3$
L-Tryptophan	$C_{11}H_{12}N_2O_2$
Lysine	$C_6H_{14}N_2O_2$
Methionine	$C_5H_{11}NO_2S$
N^4-Acetylcytidine	$C_{11}H_{15}N_3O_6$
Nicotinamide Adenine Dinucleotide (NAD)	$C_{21}H_{27}N_7O_{14}P_2$
Ornithine	$C_5H_{13}N_2O_2.Cl$
Phenylalanine	$C_9H_{11}NO_2$
Proline	$C_5H_9NO_2$
Serine	$C_3H_7NO_3$

Threonine	$C_4H_9NO_3$
Thymine	$C_5H_6N_2O_2$
Tyrosine	$C_9H_{11}NO_3$
Uric acid	$C_5H_4N_4O_3$
Urocanic Acid	$C_6H_6N_2O_2$

Table S1. Deuterated internal standards and endogenous external standards employed for method optimization.

90 µL/min			
1.0 mm i.d. setup on microUHPLC		**1.0 mm i.d. setup on analytical UHPLC**	
FWHM	N_c	FWHM	N_c
0.105	54.425	0.143	39.658
100 µL/min			
0.07	78.800	0.136	40.672

Table S2. FWHM and peak capacity values obtained by installing the 1.0 mm i.d. column on the analytical UHPLC system.

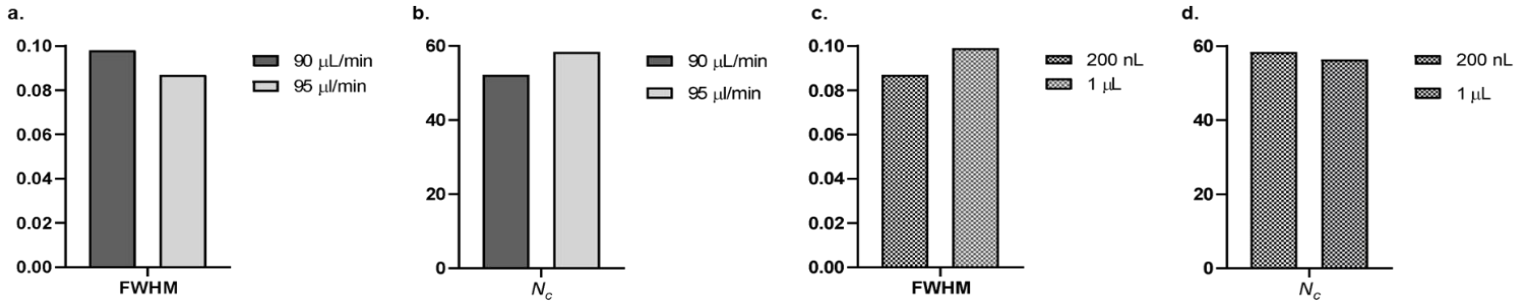


Figure S1A-D Effect of flow rate and injection volume on FWHM and peak capacity in HILIC mode.

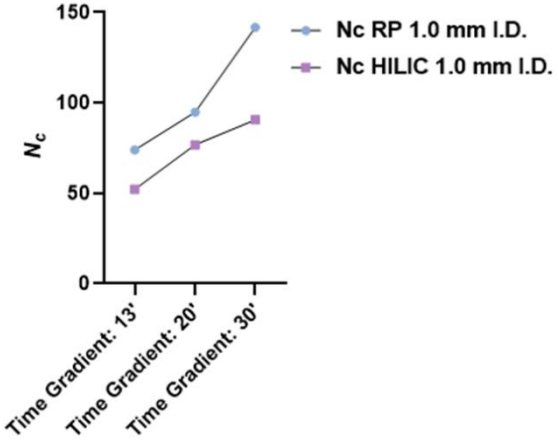


Figure S2: Effect of gradient length on peak capacity on both RP and HILIC.

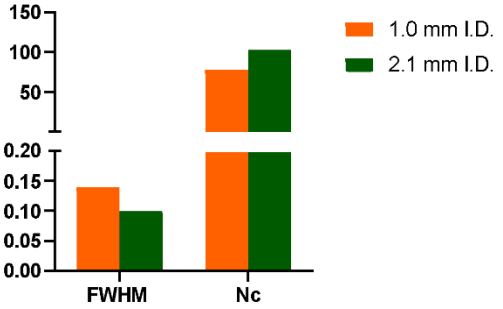


Figure S3: Final comparison of FWHM and peak capacity between 1.0- and 2.1-mm I.D.

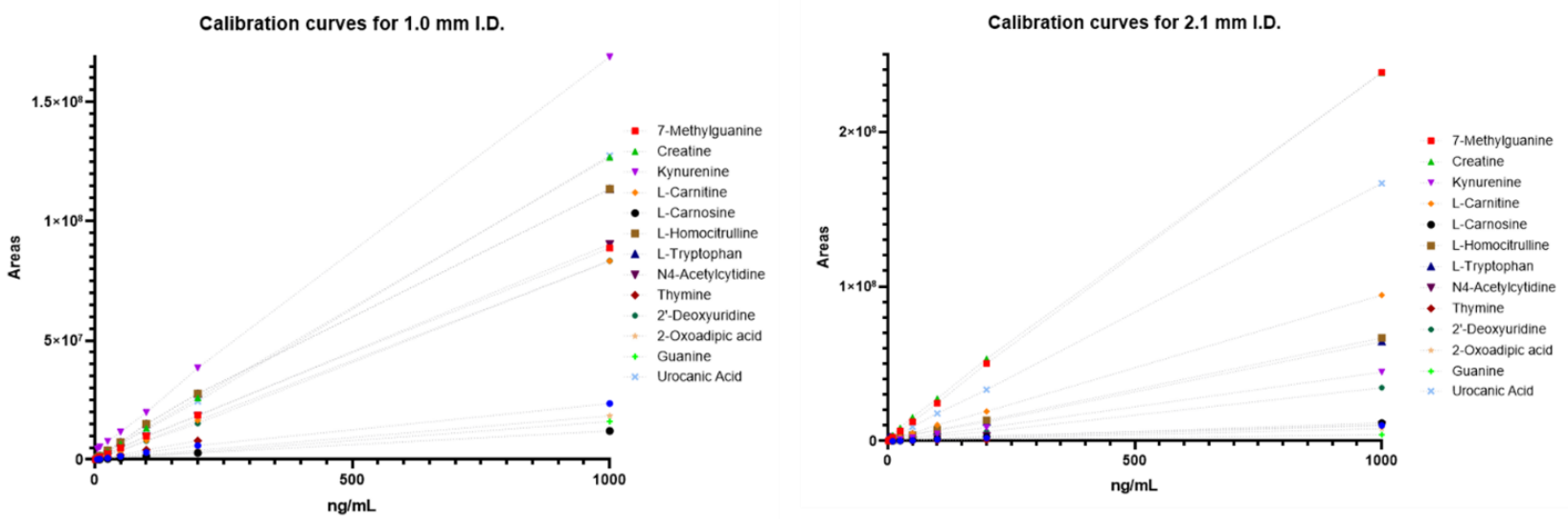


Figure S4. Calibration curves for 1.0 mm and 2.1 mm i.d. setups.

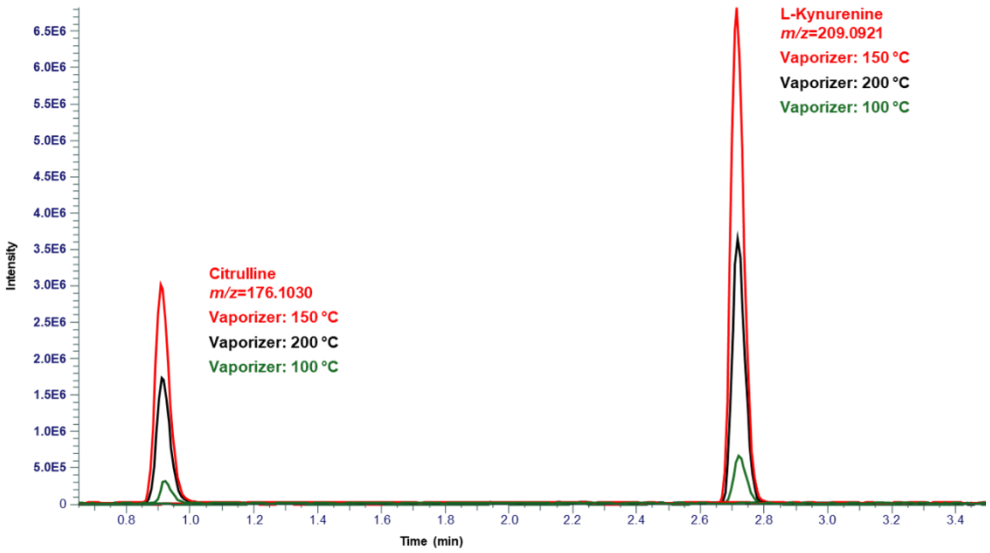


Figure S5. EICs of kynurenine m/z=209.0921 and citrulline m/z=176.1030 in positive ionization mode employing vaporizer temperature at 100 °C in green, 150 °C in red and 200 °C in black.

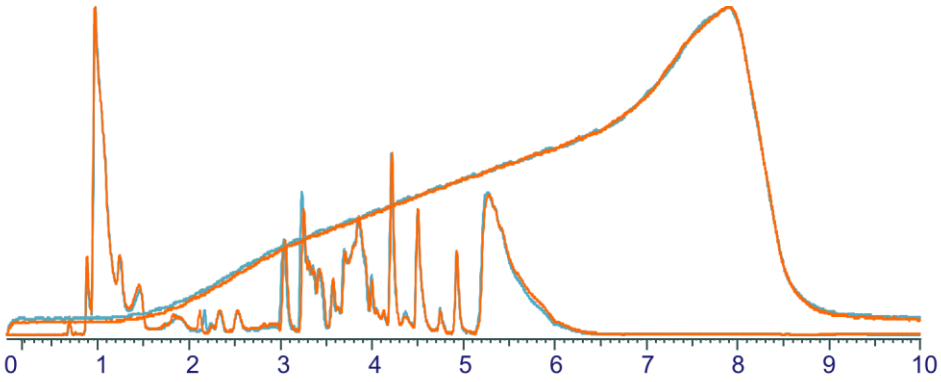


Figure S6. Overlapped traces of base peak chromatogram and relative backpressure in HILIC-HRMS from #1 to 250 consecutive injection.

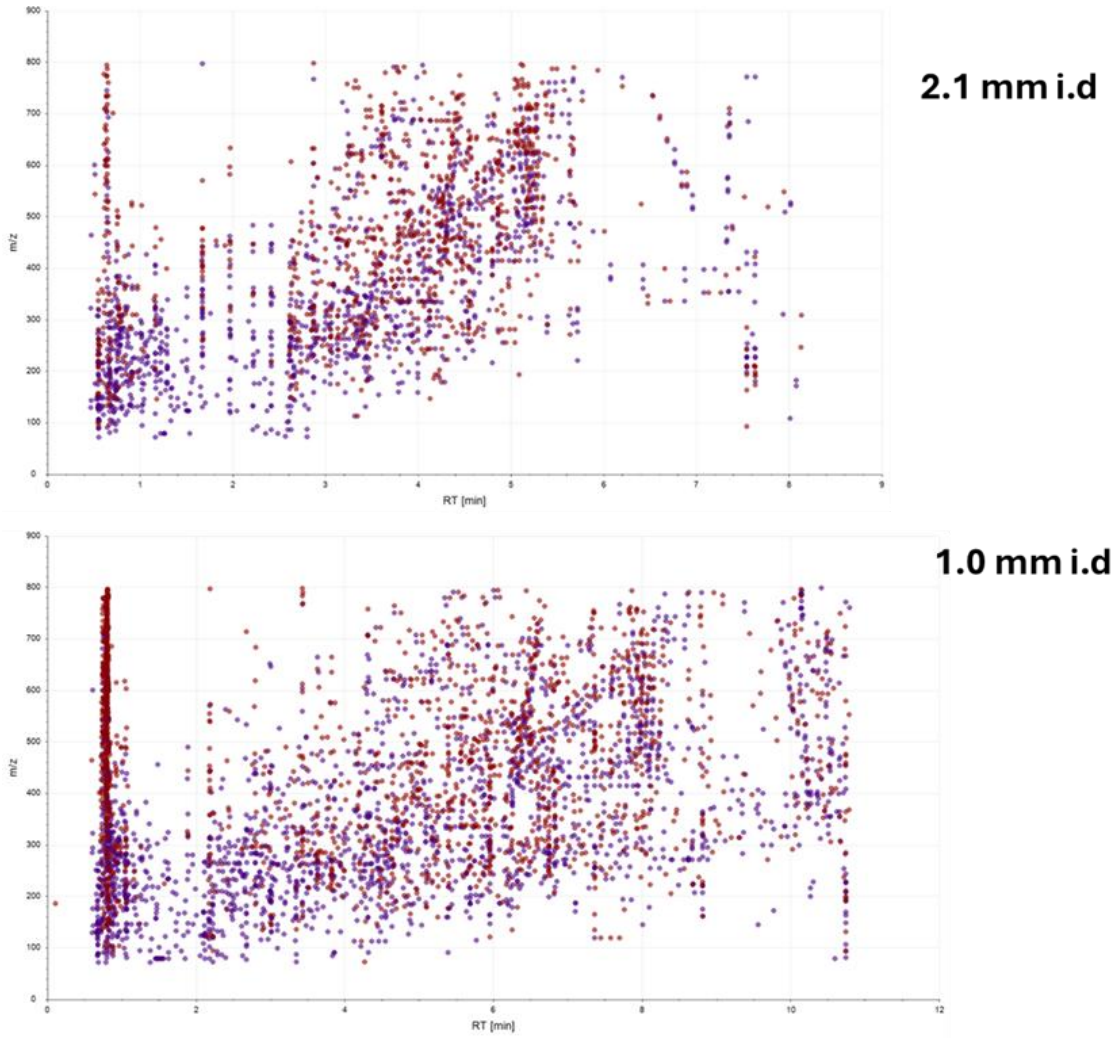


Figure S7. Dot plots of MS1 and MS/MS events on 2.1 mm i.d (top) and 1.0 mm i.d (bottom) setups.

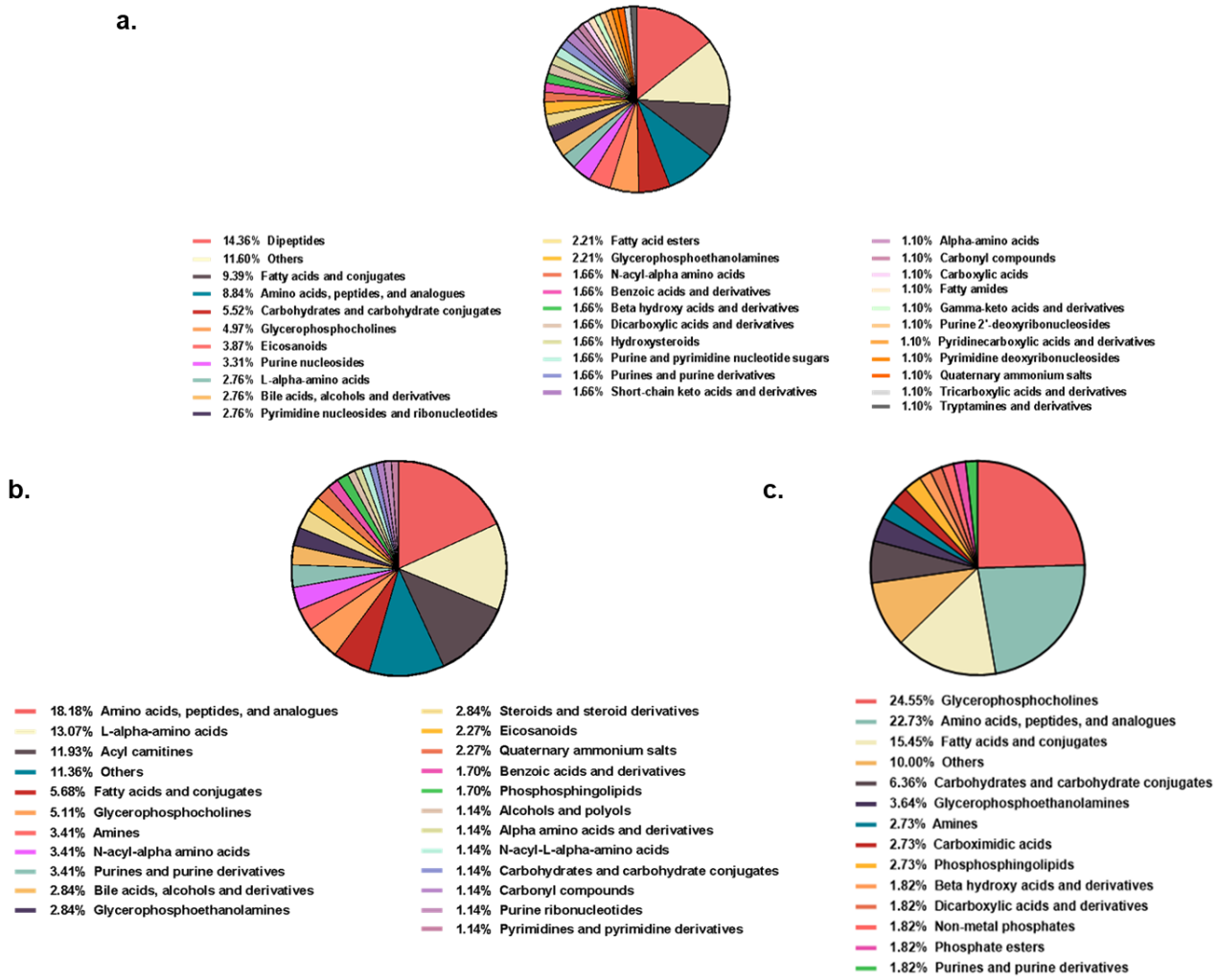


Figure S8 A-C. Graphical representation of metabolome coverage and overlap between 1.0 and 2.1 mm i.d.

Supplementary Section Chapter V

Spectral Matching Parameters

mzCloud Node

General Settings

Compound Classes: all; Precursor Mass Tolerance: 5 ppm; FT Fragment Mass Tolerance: 5 ppm: IT Fragment Mass Tolerance: 0.4 Da; Library: Autoprocessed, Reference; Post Processing: Recalibrated; Max # Results: 10; Annotate Matching Fragments: True; Search MSn Tree: True.

DDA Search

Identity Search: Cosine; Match Activation Type: True; Match Activation Energy: Match with Tolerance; Activation Energy Tolerance: 20; Apply Intensity Threshold: True; Similarity Search: Confidence Forward; Match Factor Threshold: 50.

DIA Search

Use DIA Scans for Search: False; Max. Isolation Width (Da): 500; Match Activation Type: False; Match Activation Energy: Any; Activation Energy Tolerance: 100; Apply Intensity Threshold: False; Match Factor Threshold: 20.

mzVault Node

Compound Classes: all; Match Ion Activation Type: True; Match Ion Activation Energy: Match with Tolerance; Ion Activation Energy Tolerance: 20; Match Ionization Method: True; Apply Intensity Threshold: True; Remove Precursor Ion: True; Precursor Mass Tolerance: 10 ppm; FT Fragment Mass Tolerance: 10 ppm; IT Fragment Mass Tolerance: 0.4 Da; Match Analyzer Type: True; Search Algorithm: HighChem HighRes; Match Factor Threshold: 50; Max # Results: 10; RT Tolerance (min): 2; Use Retention Time: False.

Name	Molecular formula
11(12)-EET	$C_{20}H_{32}O_3$
11(-HETE	$C_{20}H_{32}O_3$
18-HEPE	$C_{20}H_{30}O_3$
Ac-Ala-OH	$C_5H_9NO_3$
Acetylamino-amino-methyluracil	$C_7H_{10}N_4O_3$
Acetyl-arginine	$C_8H_{16}N_4O_3$
Acetylcholine	$C_7H_{15}NO_2$
Acetylcytidine	$C_{11}H_{15}N_3O_6$
Acetyl-D-lactosamine	$C_{14}H_{25}NO_{11}$
Acetyl-DL-serine	$C_5H_9NO_4$
Acetyl-L-carnitine	$C_9H_{17}NO_4$
Acetyl-L-cysteine	$C_5H_9NO_3S$
Acetyl-L-histidine	$C_8H_{11}N_3O_3$
Acetyl-L-leucine	$C_8H_{15}NO_3$
Acetyl-L-methionine	$C_7H_{13}NO_3S$
Acetylneuraminic acid	$C_{11}H_{19}NO_9$
Acetylornithine	$C_7H_{14}N_2O_3$
Acetylputrescine	$C_6H_{14}N_2O$
Acetylserine	$C_5H_9NO_4$
Acetylspermidine	$C_9H_{21}N_3O$
Acetylsphinganine	$C_{20}H_{41}NO_3$
Acetylvaline	$C_7H_{13}NO_3$
Acetyl-β-methylcholine	$C_8H_{17}NO_2$
Adenine	$C_5H_5N_5$
Adenosine	$C_{10}H_{13}N_5O_4$
Adenosine diphosphate	$C_{10}H_{15}N_5O_{10}P_2$
Adenosine-monophosphate	$C_{10}H_{14}N_5O_7P$
Adenosylhomocysteine	$C_{14}H_{20}N_6O_5S$
Adipic acid	$C_6H_{10}O_4$
Adrenic acid	$C_{22}H_{36}O_2$
Ala-Ala-Ala	$C_9H_{17}N_3O_4$
ALA-GLU	$C_8H_{14}N_2O_5$
Alanine	$C_3H_7NO_2$
Alanine betaine	$C_6H_{13}NO_2$
Alanyl-leucine	$C_9H_{18}N_2O_3$
Ala-Val	$C_8H_{16}N_2O_3$
Allantoin	$C_4H_6N_4O_3$
alpha-ketoadipic acid	$C_6H_8O_5$
alpha-Ketoglutaric acid	$C_5H_6O_5$
Aminoadenosine	$C_{10}H_{14}N_6O_4$

Aminoadipic acid	$C_6H_{11}NO_4$
Aminobutyric acid	$C_4H_9NO_2$
Amino-D-isoleucine	$C_6H_{14}N_2O_2$
Aminohippuric acid	$C_9H_{10}N_2O_3$
Aminolevulinic acid	$C_5H_9NO_3$
Aminonicotinic acid	$C_6H_6N_2O_2$
Aminoundecanoic acid	$C_{11}H_{23}NO_2$
Aminovaleric acid	$C_5H_{11}NO_2$
AMPA	$C_7H_{10}N_2O_4$
Androsterone sulfate	$C_5H_6O_5$
Arabinosylhypoxanthine	$C_{10}H_{12}N_4O_5$
Arachidic acid	$C_{20}H_{40}O_2$
Arachidonic acid	$C_{20}H_{32}O_2$
Arachidoyl Ethanolamide	$C_{22}H_{45}NO_2$
Arachidyl carnitine	$C_{27}H_{53}NO_4$
Arg-Val	$C_{11}H_{23}N_5O_3$
Asn-Pro	$C_9H_{15}N_3O_4$
Asn-Val	$C_9H_{17}N_3O_4$
Asparagine	$C_4H_8N_2O_3$
Aspartylphenylalanine	$C_{13}H_{16}N_2O_5$
Asp-Gln	$C_9H_{15}N_3O_6$
b-Ala-Lys	$C_9H_{19}N_3O_3$
beta-Aspartylaspartic acid	$C_8H_{12}N_2O_7$
Betaine	$C_5H_{11}NO_2$
Bilirubin	$C_{33}H_{36}N_4O_6$
Biotin	$C_{10}H_{16}N_2O_3S$
Bis(ethylhexyl)adipate	$C_{22}H_{42}O_4$
Butenyl-Methyl-threonine	$C_7H_8N_4O_3$
Butyrylcarnitine	$C_{11}H_{21}NO_4$
C12Carnitine	$C_{19}H_{37}NO_4$
C14Carnitine	$C_{21}H_{41}NO_4$
C18Carnitine	$C_{25}H_{49}NO_4$
Calcitriol	$C_{27}H_{44}O_3$
Caprolactam	$C_6H_{11}NO$
Caproylglycine	$C_{12}H_{23}NO_3$
Caprylic acid	$C_8H_{16}O_2$
Capryloylglycine	$C_{10}H_{19}NO_3$
CCarnitine	$C_{15}H_{29}NO_4$
Cervonoyl ethanolamide	$C_{24}H_{36}O_3$
Cholesterol sulfate	$C_{27}H_{46}O_4S$
Cholic acid	$C_{24}H_{40}O_6$
Choline	$C_5H_{13}NO$
Choline Alfoscerate	$C_8H_{20}NO_6P$

Citric acid	$C_6H_8O_8$
CMPF	$C_{12}H_{16}O_6$
Corticosterone	$C_{21}H_{30}O_4$
Cortisol	$C_{21}H_{30}O_6$
Creatine	$C_4H_9N_3O_3$
Creatinine	$C_4H_7N_3O$
Cysteine-glutathione disulfide	$C_{13}H_{22}N_4O_8S_3$
Cytidine-diphosphocholine	$C_{14}H_{26}N_4O_{11}P_3$
Cytidine-monophosphate (hydrate)	$C_9H_{14}N_3O_8P$
Decadienoylcarnitine	$C_{17}H_{29}NO_4$
Decanamide	$C_{10}H_{21}NO$
Decanoic acid	$C_{10}H_{20}O_2$
Decanoylcarnitine	$C_{17}H_{33}NO_4$
Decenoylcarnitine	$C_{17}H_{31}NO_4$
Dehydro Thromboxane B2	$C_{20}H_{32}O_6$
Dehydroadenosine	$C_{10}H_{11}N_5O_4$
Deoxycholic Acid	$C_{24}H_{40}O_4$
Deoxycholic acid glycine conjugate	$C_{26}H_{43}NO_5$
Deoxycytidine	$C_9H_{13}N_3O_4$
Deoxyguanosine diphosphate (dGDP)	$C_{10}H_{15}N_5O_{10}P_2$
Deoxyguanosine-monophosphate (dGMP)	$C_{10}H_{14}N_5O_7P$
Deoxyinosine	$C_{10}H_{12}N_4O_4$
Deoxypentose	$C_5H_{10}O_4$
Deoxy-phosphonopent-ulose	$C_5H_{11}O_7P$
Deoxy-prostaglandin A1	$C_{20}H_{30}O_3$
Deoxyuridine	$C_9H_{12}N_2O_5$
Deoxyuridine-monophosphate	$C_9H_{13}N_2O_8P$
D-Erythrose	$C_4H_8O_4$
D-Erythro-sphingosine 1-phosphate	$C_{18}H_{38}NO_5P$
D-Fructose	$C_6H_{12}O_6$
D-Galactose	$C_6H_{12}O_6$
D-Glucose	$C_6H_{12}O_6$
D-Glucose phosphate	$C_6H_{13}O_9P$
D-Glutamine	$C_5H_{10}N_2O_3$
D-Homoproline	$C_6H_{11}NO_2$
Dicyclohexylurea	$C_{13}H_{24}N_2O$
Diethylethanolamine	$C_6H_{15}NO$
Dihydro-15-keto Prostaglandin F2α	$C_{20}H_{34}O_5$
Dihydro-keto Prostaglandin A2	$C_{20}H_{30}O_4$
Dihydrotestosterone	$C_{19}H_{30}O_2$
Dihydrotestosterone glucuronide	$C_{25}H_{38}O_8$
Dihydrothymine	$C_5H_8N_2O_2$
Dihydrouridine	$C_9H_{14}N_2O_6$

Dihydroxybutyric acid	$C_4H_8O_4$
Dihydroxycholanoic acid	$C_{24}H_{40}O_4$
Dihydroxy-cholestanoic acid	$C_{27}H_{46}O_4$
Dihydroxy-cholestenoate	$C_{27}H_{44}O_4$
Dihydroxyphenylglycine	$C_8H_9NO_4$
Dimethyl adipate	$C_8H_{14}O_4$
Dimethylglycine	$C_4H_9NO_2$
Dimethylguanosine	$C_{12}H_{17}N_5O_5$
Dimethyl-L-arginine	$C_8H_{18}N_4O_2$
Dimethylsphingosine	$C_{20}H_{37}NO_4$
Dimethyluric acid	$C_7H_8N_4O_3$
Diphosphoglyceric acid	$C_3H_8O_{10}P_2$
DL-Arginine	$C_6H_{14}N_4O_2$
DL-Homocysteine	$C_4H_9NO_2S$
DL-Homoserine	$C_4H_9NO_3$
DL-Leucine	$C_6H_{13}NO_3$
DL-Mevalonic acid	$C_6H_{12}O_4$
DL-Stachydrine	$C_7H_{13}NO_3$
DL-α-Aminocaprylic acid	$C_8H_{17}NO_2$
D-Mannose	$C_6H_{12}O_6$
D-Mannose phosphate	$C_6H_{13}O_9P$
D-myo-Inositol-bisphosphate	$C_6H_{14}O_{12}P_2$
Docosahexaenoic Acid	$C_{22}H_{32}O_2$
Docosanoylglycine	$C_{24}H_{47}NO_3$
Dodecanedioic acid	$C_{12}H_{22}O_4$
Dodecenoic acid	$C_{12}H_{22}O_2$
D-Panthenol	$C_9H_{19}NO_4$
D-Pipecolinic acid	$C_6H_{11}NO_2$
D-Ribose-1-phosphate	$C_5H_{11}O_8P$
D-Tryptophan	$C_{11}H_{12}N_2O_2$
D-Xylose	$C_5H_{10}O_5$
Eicosadienoic Acid	$C_{20}H_{36}O_2$
Eicosapentaenoic acid	$C_{20}H_{30}O_2$
Eicosatrienoic acid	$C_{20}H_{34}O_2$
Epileukotriene B4	$C_{20}H_{32}O_4$
Ethyl argininate	$C_8H_{18}N_4O_2$
Ethyl lactate	$C_5H_{10}O_3$
Ethyl lysine	$C_8H_{18}N_2O_2$
Ethyl malate	$C_8H_{14}O_5$
Ethylglycine	$C_4H_9NO_2$
Folic acid	$C_{19}H_{19}N_7O_6$
Formyl-hydroxykynurenamine	$C_{10}H_{12}N_2O_3$
Formylkynurenine	$C_{11}H_{12}N_2O_4$

gamma-Aminobutyric acid	$C_4H_9NO_2$
gamma-Aminobutyryl-lysine	$C_{10}H_{21}N_3O_3$
gamma-Butyrobetaine	$C_7H_{15}NO_2$
gamma-Glu-gln	$C_{10}H_{17}N_3O_6$
Gamma-Hydroxybutyric acid	$C_4H_8O_3$
gamma-L-glutamyl-L-tyrosine	$C_{14}H_{18}N_2O_6$
gamma-thiomethyl glutamate	$C_6H_{11}NO_4S$
Gln-Gln	$C_{10}H_{18}N_4O_5$
Gluconic acid	$C_6H_{12}O_7$
Glucosamine	$C_6H_{13}NO_5$
Glucose 1-phosphate	$C_6H_{13}O_9P$
Glucuronic acid-lactone	$C_6H_8O_6$
Glu-Gly	$C_7H_{12}N_2O_5$
Glu-Ser	$C_8H_{14}N_2O_6$
Glutamyl phosphate	$C_5H_{10}NO_7P$
Glutaric acid	$C_5H_8O_4$
Glutaric anhydride	$C_5H_6O_3$
Glutarylcarnitine	$C_{12}H_{21}NO_6$
Glutaurine	$C_7H_{14}N_2O_6S$
Glu-Thr	$C_9H_{16}N_2O_6$
Glycerylphosphorylethanolamine	$C_5H_{14}NO_6P$
Glycine	$C_2H_5NO_2$
Glycochenodeoxycholic acid	$C_{26}H_{43}NO_5$
Glycocholic acid	$C_{26}H_{43}NO_6$
Glycodeoxycholic acid	$C_{26}H_{43}NO_5$
Glycylglutamic acid	$C_7H_{12}N_2O_5$
Glycylglutamine	$C_7H_{13}N_3O_4$
Glycylproline	$C_7H_{12}N_2O_3$
Gly-DL-Phe	$C_{11}H_{14}N_2O_3$
Gly-Lys	$C_8H_{17}N_3O_3$
Glyoxylic acid	$C_2H_2O_3$
guanidinobutanal	$C_5H_{11}N_3O$
Guanidinobutyric acid	$C_5H_{11}N_3O_2$
guanidino-oxopentanoic acid	$C_6H_{11}N_3O_3$
Guanidinosuccinic acid	$C_5H_9N_3O_4$
Guanidoacetic acid	$C_3H_7N_3O_2$
Guanosine	$C_{10}H_{13}N_5O_5$
Guanosine monophosphate	$C_{10}H_{14}N_5O_8P$
Guanosine-diphosphate (GDP)	$C_{10}H_{15}N_5O_{11}P_2$
Guanosine-diphospho-D-mannose	$C_{16}H_{25}N_5O_{16}P_3$
Guanosine-diphospho-L-fucose	$C_{16}H_{25}N_5O_{15}P_3$
Heptadecanoylcarnitine	$C_{24}H_{47}NO_4$
Heptadecanoylglycine	$C_{19}H_{37}NO_3$

Heptanoylglycine	$C_9H_{17}NO_3$
Hexadecanamide	$C_{16}H_{33}NO$
Hexadecenoylcarnitine	$C_{23}H_{43}NO_4$
Hexanoylcarnitine	$C_{13}H_{25}NO_4$
Hexanoylglycine	$C_8H_{15}NO_3$
Hexanoylsphingosine	$C_{24}H_{47}NO_3$
Hexenoylcarnitine	$C_{13}H_{23}NO_4$
Hippuric acid	$C_9H_9NO_4$
His-Lys	$C_{12}H_{21}N_5O_3$
Histamine	$C_5H_9N_3$
Homocitrulline	$C_7H_{15}N_3O_3$
Homo-L-arginine	$C_7H_{16}N_4O_2$
Homovanillic acid	$C_9H_{10}O_4$
Hydroxy dehydroepiandrosterone-sulfate	$C_{19}H_{28}O_6S$
Hydroxy myristic acid	$C_{14}H_{28}O_4$
Hydroxybenzaldehyde	$C_7H_6O_2$
Hydroxybenzoic acid	$C_7H_6O_3$
Hydroxybutyric acid	$C_4H_8O_3$
Hydroxybutyric acid (GHB)	$C_4H_8O_3$
Hydroxybutyrylcarnitine	$C_{11}H_{21}NO_5$
Hydroxycaproic acid	$C_6H_{12}O_3$
Hydroxy-cholenoic acid	$C_{24}H_{38}O_3$
hydroxydecanoyl carnitine	$C_{17}H_{33}NO_5$
hydroxy-deoxyguanosine	$C_{10}H_{13}N_5O_5$
hydroxydodecanoyl carnitine	$C_{19}H_{37}NO_5$
Hydroxyhexadecanoylcarnitine	$C_{23}H_{45}NO_5$
Hydroxyhexanoylcarnitine	$C_{13}H_{25}NO_6$
Hydroxyhippuric acid	$C_9H_9NO_4$
Hydroxyindole-acetic acid	$C_{10}H_9NO_3$
hydroxyisouric acid	$C_5H_4N_4O_4$
Hydroxyisovalerylcarnitine	$C_{12}H_{23}NO_6$
Hydroxy-leukotriene E4	$C_{23}H_{37}NO_6S$
Hydroxy-L-leucine	$C_6H_{13}NO_3$
Hydroxymyristoylcarnitine	$C_{21}H_{41}NO_5$
hydroxyoctanoylcarnitine	$C_{15}H_{29}NO_5$
Hydroxyphenylacetylglycine	$C_{10}H_{11}NO_4$
Hydroxyproline	$C_5H_9NO_3$
Hydroxypurine	$C_5H_4N_4O$
Hydroxypyridine	C_5H_5NO
Hydroxysebacic acid	$C_{10}H_{18}O_5$
Hydroxy-tetradecenoylcarnitine	$C_{21}H_{39}NO_5$
Hydroxytryptophan	$C_{11}H_{12}N_2O_3$
Hydroxyvaleric acid	$C_5H_{10}O_3$

Hypoxanthine	$C_5H_4N_4O$
Indole-acetic acid	$C_{10}H_9NO_2$
Indole-lactic acid	$C_{11}H_{11}NO_3$
Indoxyl sulphate	$C_8H_7NO_4S$
Inosine	$C_{10}H_{12}N_4O_5$
Inosine-monophosphate (IMP)	$C_{10}H_{13}N_4O_8P$
Inosinic acid	$C_{10}H_{13}N_4O_8P$
Isobutyric acid	$C_4H_8O_2$
Isocitric acid	$C_6H_8O_7$
Isoleucine	$C_6H_{13}NO_2$
Isoprostaglandin E2	$C_{20}H_{32}O_5$
Isovalerylglycine	$C_7H_{13}NO_3$
Itaconic acid	$C_5H_6O_4$
Keto prostaglandin A1	$C_{20}H_{30}O_4$
Ketobutyric acid	$C_4H_6O_3$
ketodeoxycholic acid	$C_{24}H_{38}O_5$
Keto-glutaramic acid	$C_5H_7NO_4$
Ketoleucine	$C_6H_{10}O_3$
keto-Prostaglandin E2	$C_{20}H_{30}O_5$
Lactamide	$C_3H_7NO_2$
L-Alanine	$C_3H_7NO_2$
L-Alanine methyl ester	$C_4H_9NO_2$
L-Alanyl-L-proline	$C_8H_{14}N_2O_3$
L-Aspartic acid	$C_4H_7NO_4$
Lauroylglycine	$C_{14}H_{27}NO_3$
LCarnitine	$C_7H_{15}NO_3$
L-Citrulline	$C_6H_{13}N_3O_3$
L-Cysteine	$C_3H_7NO_2S$
L-Cystine	$C_6H_{12}N_2O_4S_2$
Leu-Asp	$C_{10}H_{18}N_2O_5$
Leu-Leu	$C_{12}H_{24}N_2O_3$
Leu-Phe	$C_{15}H_{22}N_2O_3$
Leu-Pro	$C_{11}H_{20}N_2O_3$
L-gamma-Glutamyl-L-leucine	$C_{11}H_{20}N_2O_5$
L-gamma-Glutamyl-L-valine	$C_{10}H_{18}N_2O_5$
L-Glutamic acid	$C_5H_9NO_4$
L-Glutamine	$C_5H_{10}N_2O_3$
L-Glutamyl-L-glutamic acid	$C_{10}H_{16}N_2O_7$
L-Glutathione (reduced)	$C_{10}H_{17}N_3O_6S$
L-Glutathione oxidized	$C_{20}H_{32}N_6O_{12}S_2$
L-Histidine	$C_6H_9N_3O_2$
L-Homocystine	$C_8H_{16}N_2O_4S_2$
L-Iditol	$C_6H_{14}O_6$

Linoleamide	$C_{18}H_{33}NO$
Linolenic acid	$C_{18}H_{30}O_2$
Linoleyl carnitine	$C_{25}H_{45}NO_4$
Lipoxin B4	$C_{20}H_{32}O_5$
L-Isoleucine	$C_6H_{13}NO_2$
L-Kynurenine	$C_{10}H_{12}N_2O_3$
L-Lactic acid	$C_3H_6O_3$
L-Leucine	$C_6H_{13}NO_2$
L-Lysine	$C_6H_{14}N_2O_2$
L-Methionine	$C_5H_{11}NO_2S$
L-Norleucine	$C_6H_{13}NO_2$
LPA 16:0	$C_{19}H_{39}O_7P$
LPA 18:0	$C_{21}H_{43}O_7P$
LPC 14:0	$C_{22}H_{46}NO_7P$
LPC 16:0	$C_{24}H_{50}NO_7P$
LPC 16:1	$C_{24}H_{48}NO_7P$
LPC 18:0	$C_{26}H_{54}NO_7P$
LPC 18:1	$C_{26}H_{52}NO_7P$
LPC 20:1	$C_{28}H_{54}NO_7P$
LPC 20:2	$C_{28}H_{54}NO_7P$
LPC 20:3	$C_{28}H_{52}NO_7P$
LPC 20:4	$C_{28}H_{50}NO_7P$
LPC 20:5	$C_{28}H_{48}NO_7P$
LPC 22:4	$C_{30}H_{54}NO_7P$
LPC 22:5	$C_{30}H_{52}NO_7P$
LPC 22:6	$C_{30}H_{50}NO_7P$
LPC O-16:0	$C_{24}H_{52}NO_6P$
LPC O-16:1	$C_{24}H_{50}NO_6P$
LPE (22:6)	$C_{27}H_{44}NO_7P$
LPE 16:0	$C_{21}H_{44}NO_7P$
LPE 18:0	$C_{23}H_{48}NO_7P$
LPE 18:1	$C_{23}H_{46}NO_7P$
LPE 18:2	$C_{23}H_{44}NO_7P$
LPE 20:0	$C_{25}H_{52}NO_7P$
LPE 20:4	$C_{25}H_{44}NO_7P$
L-Phenylalanine	$C_9H_{11}NO_2$
L-Phenyllactic acid	$C_9H_{10}O_3$
L-Pyroglutamic acid	$C_5H_7NO_3$
L-Serine	$C_3H_7NO_3$
L-Threonine	$C_4H_9NO_3$
Lys-Glu	$C_{11}H_{21}N_3O_5$
Lys-Leu	$C_{12}H_{25}N_3O_3$
Maleic acid	$C_4H_4O_4$

Malic acid	$C_4H_6O_5$
Malonic acid	$C_3H_4O_4$
Me-Adenosine	$C_{11}H_{15}N_5O_4$
Met-Glu	$C_{10}H_{18}N_2O_5S$
Methionine sulfoxide	$C_5H_{11}NO_3S$
Met-His	$C_{11}H_{18}N_4O_3S$
Methoxy-L-tryptophan	$C_{12}H_{14}N_2O_3$
Methoxyphenylacetic acid	$C_9H_{10}O_3$
Methyl isonicotinate	$C_7H_7NO_2$
Methyl tyrosinate	$C_{10}H_{13}NO_3$
Methylbutyroylcarnitine	$C_{12}H_{23}NO_4$
Methyldioctylamine	$C_{17}H_{37}N$
Methylene-oxoglutarate	$C_6H_6O_5$
Methylglutarate	$C_6H_{10}O_5$
Methylguanine	$C_6H_7N_5O$
Methylguanosine	$C_6H_7N_5O$
Methylhippuric acid	$C_{10}H_{11}NO_3$
Methylhistidine	$C_7H_{11}N_3O_2$
Methylhistidine	$C_7H_{11}N_3O_2$
Methylinosine	$C_{11}H_{14}N_4O_5$
Methylmalonic acid	$C_4H_6O_4$
Methylnicotinamide	$C_7H_8N_2O$
Methylspermidine	$C_8H_{21}N_3$
methylsphingosine	$C_{19}H_{37}NO_3$
Methyluric acid	$C_6H_6N_4O_3$
Methyluridine	$C_{10}H_{14}N_2O_6$
Methylxanthine	$C_6H_6N_4O_3$
Mevalonic acid	$C_6H_{12}O_4$
Myristic Acid	$C_{14}H_{28}O_2$
Myristoleic acid	$C_{14}H_{26}O_2$
Myristyl sulfate	$C_{14}H_{30}O_4S$
Niacinamide	$C_6H_6N_2O$
Nicotinamide	$C_6H_6N_2O$
Nicotinamide 1-oxide	$C_6H_6N_2O_2$
Nicotinamide adenine dinucleotide (NAD+)	$C_{21}H_{27}N_7O_{14}P_2$
Nicotinamide mononucleotide	$C_{11}H_{15}N_2O_8P$
Nipecotic acid	$C_6H_{11}NO_2$
N-Methyllysine	$C_7H_{16}N_2O_2$
Nonanoylcarnitine	$C_{16}H_{31}NO_4$
Nonanoylglycine	$C_{11}H_{21}NO_3$
O-beta-D-galactosyl-sn-glycerol	$C_9H_{18}O_9$
Octenoylcarnitine	$C_{15}H_{27}NO_4$
Oleamide	$C_{18}H_{35}NO$

Oleic acid	$C_{18}H_{34}O_2$
Oleoylcarnitine	$C_{25}H_{47}NO_4$
Ophthalmic acid	$C_{11}H_{19}N_3O_6$
Ornithine	$C_5H_{12}N_2O_2$
Oxalomalic acid	$C_6H_6O_8$
Oxodecanoyl cysteamine	$C_{12}H_{23}NO_2S$
Oxononanoic acid	$C_9H_{16}O_3$
Oxopalmitic acid	$C_{16}H_{30}O_4$
Oxovaleric acid	$C_5H_8O_4$
Palmitamide	$C_{16}H_{33}NO$
Palmitelaidic acid	$C_{16}H_{30}O_2$
Palmitic Acid	$C_{16}H_{32}O_2$
Palmitoleic acid	$C_{16}H_{30}O_2$
Palmitoleoyl ethanolamide	$C_{18}H_{37}NO_2$
Palmitoyl sphingomyelin	$C_{39}H_{79}N_2O_6P$
Palmitoylcarnitine	$C_{23}H_{45}NO_4$
Palmitoylglycine	$C_{18}H_{35}NO_3$
Pantothenic acid	$C_9H_{17}NO_5$
Paraxanthine	$C_7H_8N_4O_2$
PC 21:1; O2	$C_{29}H_{56}NO_{10}P$
PC 22:0	$C_{28}H_{56}NO_8P$
PC 24:0	$C_{32}H_{64}NO_8P$
PC 30:0	$C_{38}H_{76}NO_8P$
PC 32:1	$C_{40}H_{78}NO_8P$
PC 34:2	$C_{42}H_{80}NO_8P$
PC 34:3	$C_{42}H_{78}NO_8P$
PC 36:3	$C_{44}H_{82}NO_8P$
PC 36:4	$C_{44}H_{80}NO_8P$
PC 36:5	$C_{44}H_{78}NO_8P$
PE 12:0	$C_{17}H_{34}NO_8P$
PE 38:4	$C_{43}H_{78}NO_8P$
PE 38:5	$C_{43}H_{76}NO_8P$
Pentadecanoic acid	$C_{15}H_{30}O_2$
Pentadecanoylcarnitine	$C_{22}H_{43}NO_4$
Pentadecanoylglycine	$C_{17}H_{33}NO_3$
Phe-Ile	$C_{15}H_{22}N_2O_3$
Phenylacetylglutamine	$C_{13}H_{16}N_2O_4$
Phenylacetylglycine	$C_{10}H_{11}NO_3$
Phenylalanylphenylalanine	$C_{18}H_{20}N_2O_3$
Phenylbutyric acid	$C_{10}H_{12}O_2$
Phenylbutyrylglutamine	$C_{15}H_{20}N_2O_4$
Phenylvaleric acid	$C_{11}H_{14}O_2$
Phosphoenolpyruvic acid	$C_3H_5O_6P$

phosphoglyceric acid	$C_3H_7O_7P$
Phosphoric acid	H_3O_4P
Pimelylcarnitine	$C_{14}H_{25}NO_6$
Platelet-activating factor	$C_{26}H_{54}NO_7P$
Pro-Hyp	$C_{10}H_{16}N_2O_4$
Prolinamide	$C_5H_{10}N_2O$
Proline	$C_5H_9NO_2$
Prolylleucine	$C_{11}H_{20}N_2O_3$
Propionylcarnitine	$C_{10}H_{19}NO_5$
Propylurea	$C_4H_{10}N_2O$
Propynyluracil	$C_7H_6N_2O_2$
Prostaglandin E2	$C_{20}H_{32}O_5$
Prostaglandin E3	$C_{20}H_{30}O_5$
Pseudouridine	$C_9H_{12}N_2O_6$
Putrescine	$C_4H_{12}N_2$
Pyridoxic acid	$C_8H_9NO_4$
Pyridoxine	$C_8H_{11}NO_3$
Pyruvic acid	$C_3H_4O_3$
Resolvin D1	$C_3H_4O_3$
Resolvin E1	$C_{20}H_{30}O_5$
Ribothymidine	$C_{10}H_{14}N_2O_6$
ser-thr	$C_7H_{14}N_2O_5$
SM 32:1; O2	$C_{37}H_{75}N_2O_6P$
SM 36:2; O2	$C_{41}H_{81}N_2O_6P$
SM 40:2; O2	$C_{45}H_{89}N_2O_6P$
S-Methyl-thioadenosine	$C_{11}H_{15}N_5O_3S$
Spermidine	$C_7H_{19}N_3$
Sphingosine (d18:1)	$C_{18}H_{37}NO_2$
Stearamide	$C_{18}H_{37}NO$
Suberic acid	$C_8H_{14}O_5$
Succinic acid	$C_4H_6O_4$
Succinic semialdehyde	$C_4H_6O_3$
Succinyl proline	$C_9H_{13}NO_5$
Succinylacetone	$C_7H_{10}O_4$
Succinyladenosine	$C_{14}H_{17}N_5O_8$
Succinylcarnitine	$C_{11}H_{19}NO_6$
Sulfosuccinic acid	$C_4H_6O_7S$
Taurine	$C_2H_7NO_3S$
Taurochenodeoxycholic Acid	$C_{26}H_{45}NO_6S$
Testosterone sulfate	$C_{19}H_{28}O_5S$
Tetracosanoylglycine	$C_{26}H_{51}NO_3$
Tetradecadiencarnitine	$C_{21}H_{37}NO_4$
Tetradecenoylcarnitine	$C_{21}H_{39}NO_4$

Tetrahydrobiopterin	$C_9H_{15}N_5O_3$
Tetrahydrocortisone	$C_{21}H_{32}O_5$
Thiamine	$C_{12}H_{16}N_4OS$
Thymidine	$C_{10}H_{14}N_2O_5$
Tiglylcarnitine	$C_{12}H_{21}NO_4$
Tilarginine	$C_7H_{16}N_4O_2$
trans-Aconitic acid	$C_6H_6O_6$
trans-Dodecenoylcarnitine	$C_{19}H_{35}NO_4$
trans-Indoleacrylic acid	$C_{11}H_9NO_2$
Tricosanoylglycine	$C_{25}H_{49}NO_3$
Tridecanoylglycine	$C_{15}H_{29}NO_3$
trimethylglycine	$C_5H_{11}NO_2$
Trimethyl-L-lysine	$C_9H_{20}N_2O_2$
Trp-Glu	$C_{16}H_{19}N_3O_5$
Trp-His	$C_{17}H_{19}N_5O_3$
Trp-Phe	$C_{20}H_{21}N_3O_3$
Tyramine	$C_8H_{11}NO$
Tyr-Lys	$C_{15}H_{23}N_3O_4$
Undecanedioic acid	$C_{11}H_{20}O_4$
Undecanoic acid	$C_{11}H_{22}O_2$
Undecanoylglycine	$C_{13}H_{25}NO_3$
Uracil	$C_4H_4N_2O_2$
Ureidosuccinic acid	$C_5H_8N_2O_5$
Uric acid	$C_5H_4N_4O_3$
Uridine	$C_9H_{12}N_2O_6$
Uridine 5'-diphosphogalactose	$C_{15}H_{24}N_2O_{17}P_2$
Uridine monophosphate (UMP)	$C_9H_{13}N_2O_9P$
Urocanic acid	$C_6H_6N_2O_2$
Val-Asn	$C_9H_{17}N_3O_4$
Valeric acid	$C_5H_{10}O_2$
Val-Glu	$C_{10}H_{18}N_2O_5$
Valine	$C_5H_{11}NO_2$
Val-lys	$C_{11}H_{23}N_3O_3$
Valyl-hydroxyproline	$C_{10}H_{18}N_2O_4$
Valylvaline	$C_{10}H_{20}N_2O_3$
Xanthine	$C_5H_4N_4O_2$

Table S3. List of annotated metabolites using 1.0 mm i.d. setup.

Supplementary Material Chapter VI

UniProt Ids	VIP Mean	Mean High Grade	Mean Involutional	Mean Atrophic
P32119	2.0942	Medium	High	Low
Q13228	2.0932	Medium	High	Low
Q9Y2S2	2.087	Low	High	Medium
Q53GQ0	2.0141	Medium	High	Low
P07099	2.0049	Low	High	Medium
Q6NZI2	2.0034	Low	High	Medium
Q13045	1.9738	Low	High	Medium
Q53H82	1.9394	High	Low	Medium
P09488	1.926	High	Low	Medium
Q13813	1.918	Medium	High	Low
P10768	1.9119	Low	High	Medium
P30711	1.9075	Medium	High	Low
Q13642	1.9066	Medium	High	Low
P42858	1.8991	High	Low	Medium
O95865	1.8956	Medium	High	Low
Q9UKG1	1.8935	Low	High	Medium
Q9Y646	1.8898	Medium	High	Low
Q709C8	1.8816	Low	High	Medium
Q5T5C0	1.868	High	Low	Medium
Q99828	1.8659	Low	High	Medium
Q6VY07	1.8514	Low	High	Medium
Q07812	1.8472	High	Low	Medium
Q15404	1.8472	Low	High	Medium
Q9BTV4	1.8454	Low	High	Medium
P42574	1.8367	High	Low	Medium
Q01082	1.8351	Medium	High	Low
Q9NVD7	1.8308	Low	High	Medium
P49327	1.8274	High	Low	Medium
Q9NZN4	1.823	Low	High	Medium
P38606	1.8195	High	Low	Medium
O15037	1.8166	Medium	Low	High
Q9Y3E0	1.8166	High	Low	Medium
Q05655	1.8162	High	Low	Medium
P09525	1.8074	Low	High	Medium
Q9Y6M1	1.803	High	Low	Medium
P14618	1.7937	Low	Medium	High
P63167;Q96FJ2	1.7889	High	Low	Medium
P23352	1.7888	Medium	Low	High
Q9Y4X5	1.7851	High	Low	Medium
P07355	1.7783	Low	Medium	High
Q9NYL9	1.7781	Low	High	Medium
Q8WWX9	1.7744	Low	High	Medium
P30740	1.7682	Low	High	Medium

P36871	1.7608	Low	High	Medium
Q9HBH0	1.7605	High	Low	Medium
P50453	1.7597	Low	High	Medium
Q9Y5K8	1.7567	High	Low	Medium
P27361	1.7523	Low	High	Medium
O43182	1.7517	Low	High	Medium
Q5R372	1.7482	High	Low	Medium
P09972	1.7418	Low	High	Medium
O60701	1.7354	Low	High	Medium
P35611	1.7353	Medium	High	Low
P07205	1.734	Low	High	Medium
P60842	1.7316	High	Low	Medium
P28065	1.7281	Low	High	Medium
Q9BVC6	1.7184	Low	High	Medium
P19525	1.715	Medium	High	Low
O14773	1.713	High	Low	Medium
Q9BSA4	1.7117	High	Low	Medium
Q8NC96	1.7107	High	Low	Medium
P01008; cRAP-P01008	1.7071	High	Low	Medium
Q9UEY8	1.703	Medium	High	Low
Q15181	1.6997	Medium	High	Low
P30046	1.6991	Medium	High	Low
O15460	1.6988	High	Low	Medium
Q15631	1.6966	Low	High	Medium
O43488	1.6922	Medium	High	Low
Q63ZY3	1.6906	Low	High	Medium
O94985	1.6905	Medium	Low	High
Q86W92	1.6864	Low	High	Medium
P17655	1.6849	Low	High	Medium
O14950; P19105	1.6842	Low	High	Medium
P60903	1.6832	Low	Medium	High
Q15121	1.6814	Low	High	Medium
Q9ULC3	1.679	Low	High	Medium
P04181	1.678	Low	High	Medium
Q9H4G0	1.6752	Low	High	Medium
O75369	1.6746	Low	High	Medium
P11117	1.6697	High	Low	Medium
O60229	1.6689	High	Low	Medium
Q15417	1.668	Low	High	Medium
Q8TF05	1.6669	High	Low	Medium
P01703	1.6665	Medium	Low	High
P06280	1.6651	Low	High	Medium
Q9NPQ8	1.6615	Low	High	Medium
O75348	1.6601	High	Low	Medium
Q04760	1.6593	Medium	High	Low
Q09666	1.6592	Low	High	Medium
Q15847	1.6572	Low	High	Medium

Q8TE77	1.6572	Low	High	Medium
Q92536	1.6535	Medium	Low	High
Q96T76	1.6526	High	Low	Medium
P02461	1.6502	Low	Medium	High
Q8N5I2	1.645	Medium	Low	High
O95202	1.6424	High	Medium	Low
P16278	1.6393	High	Low	Medium
Q16864	1.6373	High	Low	Medium
Q96P48	1.6355	Low	High	Medium
P61981	1.6336	High	Low	Medium
Q9BY32	1.6336	Medium	High	Low
P12277	1.633	High	Medium	Low
Q9BVJ7	1.6307	Low	High	Medium
P08123	1.6298	Low	High	Medium
Q9UH99	1.6292	Low	High	Medium
Q9BY67	1.6291	Medium	High	Low
O60256	1.6259	Medium	High	Low
P15121	1.6238	Low	High	Medium
P48735	1.6237	High	Low	Medium
P29218	1.6215	Medium	High	Low
P54577	1.6182	High	Low	Medium
P36551	1.614	High	Low	Medium
P21281	1.6103	High	Low	Medium
Q05682	1.6081	Low	High	Medium
P11908	1.6067	Medium	High	Low
P51511	1.6064	Medium	Low	High
Q9NQ88	1.6009	High	Low	Medium
O00622	1.5996	Medium	Low	High
P22105	1.5992	Low	Medium	High
Q9UI12	1.5989	High	Low	Medium
O60504	1.596	Low	High	Medium
P55263	1.5925	High	Low	Medium
P02452	1.5909	Low	Medium	High
O00339	1.5904	Medium	High	Low
Q93052	1.5898	Low	High	Medium
O95025	1.5896	Medium	Low	High
Q9BVC4	1.5876	High	Low	Medium
Q9Y305	1.5859	Low	High	Medium
P37837	1.5807	Medium	High	Low
Q9NQR4	1.5801	Low	High	Medium
P08294	1.5792	Low	High	Medium
P41250	1.578	High	Low	Medium
Q1HG44	1.5728	High	Low	Medium
O43739	1.5724	Medium	High	Low
Q9H1E5	1.5702	Low	High	Medium
Q9BPX5	1.5655	High	Low	Medium
P51884	1.5652	Low	Medium	High

Q9UQB8	1.5648	High	Low	Medium
O00159	1.5645	Low	Medium	High
P07384	1.5628	Medium	High	Low
P07686	1.5625	Low	Medium	High
P36269	1.5605	Low	High	Medium
Q8NCE2	1.5603	High	Low	Medium
Q16555	1.5602	Low	High	Medium
Q14185	1.5594	Medium	High	Low
Q99536	1.5586	Low	High	Medium
Q9HB40	1.5586	Low	High	Medium
Q00796	1.5533	Medium	High	Low
Q05DH4	1.552	Medium	Low	High
Q13418	1.5495	Low	High	Medium
Q9UH65	1.5493	Low	High	Medium
P29373	1.5486	High	Low	Medium
P61086	1.548	High	Medium	Low
Q8N271	1.5475	Medium	Low	High
P28074	1.5452	High	Medium	Low
O95810	1.5448	Medium	High	Low
Q9UM22	1.5441	Low	High	Medium
P36543	1.5418	High	Low	Medium
Q9H0Q0	1.5409	High	Low	Medium
Q9UBV8	1.5397	High	Low	Medium
Q14195	1.5394	Low	High	Medium
O95372	1.5388	High	Low	Medium
Q2PPJ7	1.5384	High	Low	Medium
Q96C23	1.5372	Low	High	Medium
Q969X5	1.5364	Low	High	Medium
P22234	1.5357	Medium	High	Low
P30101	1.5342	Medium	High	Low
P21397	1.5339	Medium	High	Low
P21796	1.5319	High	Low	Medium
Q96PD2	1.5294	Medium	Low	High
P42658	1.5269	High	Low	Medium
Q02952	1.5252	Low	High	Medium
P78417	1.5246	Medium	High	Low
Q08623	1.5245	High	Low	Medium
Q07960	1.5241	Low	High	Medium
Q9NS86	1.5236	High	Low	Medium
P23610	1.5233	High	Low	Medium
P26440	1.5214	Medium	High	Low
Q9HBL0	1.5184	Low	High	Medium
P36915	1.5177	Medium	High	Low
Q6P1M3	1.5172	High	Low	Medium
Q16619	1.5158	Low	High	Medium
O95758	1.509	High	Low	Medium
Q9Y490	1.5088	Low	High	Medium

O00192	1.5074	High	Low	Medium
P15144	1.5066	Low	High	Medium
Q9BUH6	1.5062	High	Medium	Low
Q16881	1.5057	Medium	High	Low
P31150	1.5052	Medium	High	Low
O75608	1.5043	High	Low	Medium
Q9UBX5	1.5042	Low	Medium	High
Q9UHX1	1.5008	High	Medium	Low
A0A0B4J1Y8	1.5007	Medium	Low	High
P22413	1.4991	Low	High	Medium
P46782	1.494	High	Low	Medium
O00151	1.4925	Medium	High	Low
P12955	1.4922	Low	High	Medium
Q93063	1.488	Medium	Low	High
Q13641	1.4868	Low	Medium	High
Q9H773	1.4859	Low	High	Medium
Q9NUD5	1.4853	High	Low	Medium
Q9UJ70	1.4828	High	Low	Medium
P10619	1.4823	High	Low	Medium
O76021	1.482	High	Medium	Low
P02794	1.4814	Low	Medium	High
A0A0A0MS14	1.4809	Medium	Low	High
Q9NQ48	1.4776	Medium	High	Low
Q9GZU7	1.4771	Medium	Low	High
Q99439	1.4733	Low	High	Medium
Q14554	1.4729	Low	High	Medium
Q9Y5P6	1.4721	Medium	High	Low
Q9UQN3	1.4714	High	Low	Medium
Q9P258	1.4691	High	Low	Medium
P31751	1.4687	High	Medium	Low
Q8TEX9	1.4647	High	Medium	Low
Q9H0W9	1.464	Medium	High	Low
O94832	1.4626	Low	Medium	High
Q05707	1.4626	Low	High	Medium
Q9NS15	1.4602	Medium	Low	High
Q9H2C0	1.4598	Medium	Low	High
Q13188	1.4582	High	Low	Medium
O76070	1.4575	Low	Medium	High
A0A075B6I9; P04211	1.4571	Medium	Low	High
P62942	1.4568	High	Medium	Low
P11279	1.4562	Low	High	Medium
P30043	1.456	Medium	High	Low
O94919	1.4555	Low	Medium	High
P17844	1.4535	High	Medium	Low
Q9UKE5	1.4531	Medium	Low	High
Q9H6V9	1.4511	High	Medium	Low
Q9Y6W3	1.4509	Medium	Low	High

P02545	1.4497	Low	High	Medium
O75787	1.4466	Medium	Low	High
Q8NC56	1.4464	Low	High	Medium
O95470	1.446	High	Low	Medium
Q96L46	1.4452	Low	High	Medium
O00264	1.4432	Medium	High	Low
Q9H223	1.4432	Medium	High	Low
Q9UBW8	1.443	Medium	High	Low
P15927	1.4425	Medium	High	Low
Q99873	1.4421	High	Medium	Low
P26885	1.4416	Medium	High	Low
P35237	1.4411	Medium	High	Low
Q12792	1.4396	High	Medium	Low
P33121	1.4394	High	Low	Medium
P35555	1.439	Low	Medium	High
A0AVT1	1.4364	Medium	High	Low
P01743	1.4341	Medium	Low	High
P39656	1.4338	Medium	High	Low
O60271	1.4334	Medium	High	Low
P01593; P01594	1.4318	Medium	Low	High
Q5VYK3	1.4316	High	Medium	Low
Q07866	1.4312	High	Low	Medium
P29590	1.4307	Medium	High	Low
P54619	1.4293	High	Low	Medium
P61960	1.4277	Medium	High	Low
Q96CX2	1.4268	Low	High	Medium
O14618	1.425	Medium	High	Low
Q9UQ13	1.4237	High	Low	Medium
O43617	1.4235	High	Medium	Low
P08651	1.423	Medium	High	Low
Q15628	1.4212	Low	High	Medium
Q9ULH0	1.4209	Medium	Low	High
Q15435	1.4202	Medium	High	Low
Q3YEC7	1.4194	High	Low	Medium
O75368	1.4187	Low	High	Medium
Q9UGT4	1.418	Low	Medium	High
Q6DT37	1.4174	Medium	Low	High
O95571	1.4172	Low	High	Medium
Q8N4C8	1.416	Medium	Low	High
P35573	1.4159	High	Medium	Low
P04424	1.4139	High	Medium	Low
Q9NR99	1.4134	Medium	Low	High
Q16774	1.4119	Medium	High	Low
O75351	1.4101	High	Low	Medium
P00492	1.4092	Medium	High	Low
P02787; cRAP-P02787	1.4088	High	Low	Medium
Q86UX2	1.408	Low	Medium	High

Q9NUQ8	1.4079	High	Medium	Low
P09417	1.4058	Medium	High	Low
P05106	1.4054	Low	Medium	High
Q99653	1.4035	Medium	High	Low
Q9UHG3	1.4022	Medium	High	Low
Q14558	1.4018	Medium	High	Low
P02786	1.4014	Medium	Low	High
P09455	1.4012	Low	Medium	High
P28072	1.4009	High	Medium	Low
O15127	1.4004	High	Low	Medium
O75352	1.3992	High	Medium	Low
Q27J81	1.3992	Low	Medium	High
Q16658	1.399	Low	High	Medium
O43399	1.3985	Medium	High	Low
P53007	1.3965	High	Low	Medium
O00469	1.3956	High	Low	Medium
Q9UNM6	1.3939	High	Medium	Low
Q02790	1.3934	High	Medium	Low
Q16531	1.3922	High	Medium	Low
Q7LBR1	1.3903	High	Low	Medium
P05198	1.3896	High	Medium	Low
Q9H2U1	1.3893	High	Low	Medium
Q16394	1.3882	Medium	Low	High
O43301	1.3874	Medium	High	Low
P45877	1.3874	Medium	Low	High
Q9Y624	1.3846	Medium	Low	High
Q9Y259	1.384	Low	High	Medium
Q9UNZ2	1.3831	Medium	High	Low
Q8NCG7	1.3823	High	Low	Medium
P13987	1.3789	Low	Medium	High
Q9Y5Z4	1.3789	Medium	High	Low
P40763	1.3788	Low	High	Medium
Q13547	1.3782	High	Medium	Low
Q96CD0	1.378	Low	High	Medium
O94886	1.376	Low	Medium	High
P09958	1.3758	Medium	Low	High
Q5T013	1.3756	Medium	High	Low
Q9BZM4	1.3719	Medium	Low	High
Q6PHR2	1.3705	High	Low	Medium
Q9UNF0	1.3689	Medium	High	Low
O95841	1.3668	Medium	Low	High
Q14050	1.3666	High	Low	Medium
Q13310	1.3652	High	Medium	Low
P35052	1.3647	Medium	Low	High
Q14974	1.3647	High	Medium	Low
Q14126	1.3642	High	Low	Medium
P19021	1.3639	Medium	Low	High

P62140	1.3635	High	Low	Medium
Q14C86	1.3624	High	Low	Medium
P23786	1.3622	High	Medium	Low
Q96IJ6	1.362	Medium	High	Low
P36957	1.3616	High	Medium	Low
Q8N1N4	1.3609	High	Low	Medium
P54709	1.36	High	Low	Medium
P62917	1.3599	High	Low	Medium
P20810	1.3584	Low	High	Medium
P49006	1.3583	High	Medium	Low
Q9H477	1.358	Medium	High	Low
P16152	1.3554	Medium	High	Low
P20700	1.3551	High	Medium	Low
P20962	1.3544	High	Medium	Low
Q9H3G5	1.3542	Low	High	Medium
P18433	1.3532	Low	High	Medium
P0DMV8; P0DMV9	1.3519	Medium	High	Low
Q08379	1.3517	High	Medium	Low
Q9Y3Q3	1.3512	High	Low	Medium
Q14160	1.3494	High	Medium	Low
Q01105	1.3463	High	Medium	Low
Q15008	1.3447	High	Medium	Low
P43686	1.3446	High	Medium	Low
Q8TCD5	1.3445	Medium	High	Low
P54578	1.3412	High	Medium	Low
P27824	1.3405	Medium	High	Low
P55957; cRAP-P55957	1.3381	High	Low	Medium
Q9P265	1.3378	Medium	Low	High
P20591	1.337	Low	High	Medium
Q14515	1.3368	Medium	Low	High
P19367	1.3358	Medium	High	Low
P29762	1.3354	Medium	High	Low
O43681	1.3348	Medium	High	Low
Q8NFF5	1.3347	High	Medium	Low
Q32MZ4	1.3345	Low	High	Medium
Q9Y4L1	1.3343	Medium	High	Low
Q9Y4W6	1.3332	High	Medium	Low
P08572	1.3328	Low	Medium	High
P04179	1.3319	Low	High	Medium
Q14847	1.3313	Low	High	Medium
Q9NVA2	1.331	Low	High	Medium
Q9UBR2	1.3293	Low	Medium	High
O43143	1.329	High	Medium	Low
P22695	1.329	High	Medium	Low
Q9BWS9	1.3289	Medium	High	Low
Q16853	1.3281	Low	High	Medium
P34949	1.3267	Medium	High	Low

P08727	1.3264	Medium	Low	High
Q04637	1.3252	High	Medium	Low
Q8WUY1	1.3247	Medium	High	Low
Q8TEM1	1.3239	High	Medium	Low
Q09161	1.3232	High	Medium	Low
O75695	1.3229	Low	High	Medium
Q92973	1.3227	High	Medium	Low
O95295	1.3226	Medium	High	Low
Q8N2K0	1.3224	Medium	High	Low
Q9HA82	1.3216	High	Low	Medium
P31749	1.3209	High	Medium	Low
P61088	1.3207	High	Medium	Low
Q86X29	1.3205	Medium	Low	High
O00487	1.3193	High	Medium	Low
Q9H008	1.3176	Medium	High	Low
P17096	1.3104	Medium	High	Low
Q13617	1.3094	High	Medium	Low
O14786	1.3093	Low	Medium	High
Q12797	1.3082	Medium	High	Low
P62854	1.3078	High	Low	Medium
P06756	1.307	Low	Medium	High
P54920	1.3067	Medium	High	Low
P54652	1.3066	High	Medium	Low
P16473	1.3065	Low	High	Medium
Q9Y570	1.3065	High	Medium	Low
O60513	1.3062	Medium	Low	High
P42566	1.3062	Low	High	Medium
Q9H788	1.3052	High	Low	Medium
O43491	1.3048	Low	High	Medium
P62195	1.3036	High	Medium	Low
Q53EL6	1.3033	High	Medium	Low
Q92817	1.3033	Medium	Low	High
P62820	1.3032	Low	High	Medium
Q9H4M9	1.3026	Low	High	Medium
P52907	1.3022	High	Medium	Low
Q8TER0	1.3007	Low	Medium	High
P50225	1.3001	Medium	High	Low
Q13057	1.2994	High	Medium	Low
P62191	1.2986	High	Medium	Low
Q9Y5X3	1.2968	Medium	High	Low
Q92616	1.2966	High	Medium	Low
Q9HCE1	1.2964	High	Low	Medium
O00764	1.2963	Low	Medium	High
O95425	1.2959	Low	High	Medium
P13489	1.2953	Medium	High	Low
P17301	1.2951	Medium	Low	High
P49441	1.2945	Medium	High	Low

Q13576	1.2944	Medium	High	Low
Q16769	1.2944	Medium	Low	High
P06454	1.2934	High	Medium	Low
P28066	1.2932	High	Medium	Low
P22392	1.2917	Medium	High	Low
P17980	1.2901	High	Medium	Low
Q9NZU5	1.2901	Low	Medium	High
Q13724	1.29	High	Low	Medium
P51688	1.289	Low	High	Medium
P58107	1.2888	Medium	Low	High
O95433	1.2881	High	Medium	Low
Q9NR28	1.2873	High	Medium	Low
A0A0J9YX35	1.2864	Medium	Low	High
O00232	1.2858	High	Medium	Low
Q9HC38	1.285	Medium	High	Low
P24534	1.2849	Medium	High	Low
Q14145	1.2842	High	Low	Medium
O75390	1.2833	High	Medium	Low
O75843	1.2833	High	Low	Medium
O60762	1.2815	Medium	High	Low
Q9Y383	1.2806	High	Medium	Low
P20073	1.2801	Medium	High	Low
P20774	1.2801	Low	Medium	High
P00558	1.2794	Medium	High	Low
O75400	1.2792	High	Medium	Low
Q08945	1.279	High	Medium	Low
Q8TD55	1.2788	Low	High	Medium
Q9H330	1.278	High	Medium	Low
Q9H9B4	1.2778	High	Medium	Low
P62136	1.2777	High	Medium	Low
A5YKK6	1.2771	High	Medium	Low
Q7LG56	1.2771	High	Low	Medium
O43242	1.2756	High	Medium	Low
Q6P1X6	1.2755	Medium	High	Low
P55809	1.2709	Medium	High	Low
P21359	1.2705	High	Medium	Low
Q13126	1.2699	High	Medium	Low
Q9BS40	1.2698	Medium	High	Low
Q96NY7	1.2686	Low	Medium	High
Q99598	1.2683	Medium	High	Low
A0FGR8	1.2678	Medium	High	Low
Q15819	1.2677	High	Medium	Low
P50851	1.2673	High	Medium	Low
Q8TCJ2	1.2668	Medium	High	Low
Q5VW36	1.266	High	Medium	Low
P60891	1.2645	Medium	High	Low
Q9H4A4	1.2645	Medium	High	Low

O43175	1.2644	High	Medium	Low
O60684	1.2641	High	Low	Medium
P34931	1.2637	Medium	High	Low
O00231	1.2633	High	Medium	Low
Q08257	1.2615	Medium	High	Low
P62826	1.2614	High	Medium	Low
Q13595	1.2613	High	Medium	Low
Q16186	1.2604	High	Medium	Low
P30085	1.2598	High	Medium	Low
Q14114	1.2596	Medium	Low	High
O15162	1.2593	Low	Medium	High
O95479	1.2591	Low	High	Medium
Q8WVC6	1.259	Medium	High	Low
Q9H9Q2	1.2584	High	Low	Medium
Q9UHF1	1.2581	Low	Medium	High
P26640	1.2575	High	Medium	Low
Q9BXW6	1.2566	Low	Medium	High
P01782; P0DP04	1.2558	Medium	Low	High
P36873	1.2556	High	Medium	Low
Q8IXB1	1.2553	Medium	High	Low
Q15369	1.2541	High	Medium	Low
Q13200	1.2538	High	Medium	Low
P52943	1.2536	Low	High	Medium
Q8NDA8	1.2534	High	Medium	Low
P14384	1.253	Medium	Low	High
O14818	1.2529	High	Medium	Low
P04626	1.2527	Medium	Low	High
P09651	1.2523	High	Medium	Low
Q02218	1.2522	High	Medium	Low
Q7Z460	1.252	High	Medium	Low
P67809	1.2518	High	Medium	Low
P11766	1.2514	Medium	High	Low
Q9Y2X3	1.2513	High	Medium	Low
O95747	1.251	Medium	High	Low
Q92692	1.2509	Medium	Low	High
Q9UK41	1.2499	Medium	Low	High
Q969G5	1.2497	Low	High	Medium
Q03135	1.2491	Low	High	Medium
Q9UQ35	1.2484	High	Medium	Low
Q5EB52	1.2481	Medium	Low	High
P60174	1.2479	Low	High	Medium
Q96TA1	1.2476	High	Low	Medium
P61619	1.2473	High	Medium	Low
Q15172	1.2472	High	Medium	Low
P62333	1.2468	High	Medium	Low
P06865	1.2456	Low	Medium	High
Q9HAV4	1.2453	High	Medium	Low

O60503	1.2448	High	Low	Medium
Q4KMQ2	1.2435	Low	Medium	High
P01137	1.2427	Medium	Low	High
P57735	1.2419	Medium	Low	High
Q15746	1.2419	Low	Medium	High
Q8NG11	1.2412	Medium	Low	High
Q01813	1.2409	High	Medium	Low
Q13630	1.2408	High	Medium	Low
Q99460	1.2393	High	Medium	Low
Q9NVS9	1.2392	Medium	High	Low
Q9NR12	1.2382	Low	Medium	High
Q96SM3	1.2381	Low	Medium	High
Q7Z2K6	1.2378	High	Low	Medium
Q96I24	1.2377	High	Medium	Low
O75165	1.2372	Low	High	Medium
P54727	1.2368	High	Medium	Low
Q92688	1.2368	High	Medium	Low
P00387	1.2364	Medium	High	Low
Q96EZ8	1.2364	High	Low	Medium
O00571	1.2357	High	Medium	Low
P05543	1.2356	Medium	Low	High
Q9UL18	1.2356	Low	High	Medium
P24844	1.2355	Low	High	Medium
O14908	1.2347	Medium	Low	High
Q8NDZ4	1.2346	Medium	Low	High
Q13492	1.2325	Low	Medium	High
Q9Y5L0	1.232	High	Medium	Low
Q8NCW5	1.231	Medium	High	Low
P62495	1.2307	High	Medium	Low
Q8NFX7	1.2304	High	Low	Medium
Q9Y4G6	1.2301	Medium	High	Low
P43378	1.2288	Medium	High	Low
P50402	1.2277	Medium	High	Low
Q9H7B2	1.2277	High	Medium	Low
A0A075B6I0	1.2275	Medium	Low	High
A5D8V6	1.2269	Medium	Low	High
A8MV23	1.2266	Low	Medium	High
P15289	1.2258	Medium	High	Low
Q7KZ85	1.2258	High	Medium	Low
Q86WV6	1.2257	Low	Medium	High
P05388	1.2248	High	Medium	Low
O15305	1.2244	High	Low	Medium
P38646	1.2243	High	Medium	Low
Q3LXA3	1.2237	Medium	High	Low
Q9NX46	1.2231	High	Medium	Low
Q02880	1.223	High	Medium	Low
Q9NYU2	1.2229	Medium	High	Low

O75439	1.2227	High	Medium	Low
Q99569	1.2225	Medium	High	Low
Q9H0U3	1.2214	Medium	High	Low
Q16698	1.22	High	Medium	Low
P16930	1.2197	Medium	High	Low
P22061	1.2195	Medium	High	Low
P45974	1.2195	High	Medium	Low
O43847	1.2194	High	Low	Medium
P53602	1.2194	High	Medium	Low
P17900	1.2193	Medium	Low	High
Q8WWY3	1.219	High	Medium	Low
O00567	1.2189	High	Medium	Low
Q9Y282	1.2184	High	Low	Medium
Q6ZT21	1.218	Medium	High	Low
P06744	1.2172	Medium	High	Low
O96008	1.2164	High	Medium	Low
Q9Y6E2	1.2158	High	Medium	Low
P15924	1.2154	High	Low	Medium
A0A0C4DH35	1.2147	Medium	Low	High
Q9NRG7	1.2145	Medium	High	Low
Q9H9P8	1.2127	High	Medium	Low
P04439	1.2125	Low	High	Medium
O95716	1.2105	Medium	High	Low
Q9BZQ8	1.2097	Low	Medium	High
P49746	1.2096	Medium	Low	High
Q66K74	1.2091	Medium	High	Low
P35998	1.209	High	Medium	Low
Q9NVJ2	1.209	Medium	High	Low
P07900	1.2079	High	Medium	Low
O15212	1.2077	Medium	High	Low
Q96GK7	1.2069	Medium	High	Low
P55786	1.2068	High	Medium	Low
P12110	1.2063	Low	Medium	High
Q8NCA5	1.2056	High	Low	Medium
Q96FW1	1.2046	High	Medium	Low
Q9HB19	1.2044	Low	High	Medium
Q9H329	1.2038	Medium	Low	High
Q6NUK1	1.2033	High	Low	Medium
P62318	1.2029	High	Medium	Low
Q9H444	1.2028	Medium	Low	High
Q9NV96	1.2025	High	Low	Medium
P26373	1.2022	High	Medium	Low
P31689	1.2007	High	Low	Medium
P47985	1.2006	High	Medium	Low
P14927	1.2003	High	Medium	Low
Q10713	1.2003	High	Medium	Low
Q92685	1.1999	High	Medium	Low

Q12906	1.1997	High	Medium	Low
Q8IUZ5	1.1993	Medium	High	Low
P38919	1.1991	High	Medium	Low
P55769	1.1987	High	Medium	Low
P48059	1.1983	Low	High	Medium
P84157	1.1979	Medium	High	Low
Q15029	1.1973	High	Medium	Low
P02730	1.1971	Medium	High	Low
Q14980	1.1964	High	Medium	Low
O15042	1.1959	High	Medium	Low
Q13435	1.195	High	Medium	Low
P53041	1.1946	High	Low	Medium
P12081;cRAP-P12081	1.1944	High	Medium	Low
cRAP-P22629	1.1944	High	Medium	Low
Q5T0Z8	1.1942	High	Low	Medium
P20042	1.1938	High	Medium	Low
Q66K14	1.193	Medium	High	Low
Q08211	1.1929	High	Medium	Low
P00167;cRAP-P00167	1.1927	High	Medium	Low
P18583	1.1922	High	Medium	Low
P09874	1.1916	High	Medium	Low
P23246	1.1891	High	Medium	Low
Q9BXP5	1.189	High	Medium	Low
Q86UU1	1.1889	Low	High	Medium
Q07021	1.1888	High	Medium	Low
Q8NBF2	1.1887	Medium	High	Low
P49589	1.1878	High	Medium	Low
Q9Y3F4	1.1867	High	Medium	Low
P40926	1.1866	High	Medium	Low
Q7L014	1.1865	High	Medium	Low
P22087	1.186	High	Medium	Low
P02753; cRAP-P02753	1.1858	Medium	Low	High
P55072	1.1849	High	Medium	Low
P80303	1.1849	Medium	High	Low
P01701	1.1847	Medium	Low	High
P07311	1.1838	Medium	High	Low
Q15366	1.1837	High	Medium	Low
Q86YB7	1.1836	Medium	High	Low
Q9C005	1.1835	High	Medium	Low
P01780	1.1833	Medium	Low	High
P02790	1.1829	Medium	Low	High
P09622	1.1828	High	Medium	Low
P61970	1.1827	Medium	High	Low
A0A0C4DH68	1.1818	Medium	Low	High
Q15942	1.1817	Low	Medium	High
P68036	1.1816	High	Medium	Low
Q9H9H4	1.1816	Medium	Low	High

P01042	1.181	Medium	Low	High
P68371	1.181	High	Medium	Low
P31483	1.1806	High	Medium	Low
Q02809	1.1806	Medium	Low	High
Q86VP6	1.1802	High	Medium	Low
P0C0L5	1.18	Medium	Low	High
Q9BRX8	1.18	Medium	High	Low
A0A0C4DH25	1.1798	Medium	Low	High
P10606	1.1795	High	Medium	Low
P02765	1.179	Medium	Low	High
Q9P0V9	1.179	Medium	High	Low
P07948	1.1788	Low	Medium	High
Q9NRR5	1.1788	High	Medium	Low
Q8IX12	1.1785	High	Medium	Low
Q9UMX5	1.1782	Medium	High	Low
P02774	1.178	Medium	Low	High
Q02252	1.178	Medium	High	Low
P61081	1.1779	High	Medium	Low
P30041	1.1775	Low	High	Medium
P42704	1.1774	High	Medium	Low
P05114	1.177	High	Medium	Low
Q99424	1.177	High	Medium	Low
P38159	1.1769	High	Medium	Low
P02649	1.1764	Medium	Low	High
O43866	1.1763	Low	Medium	High
P01024	1.1762	Medium	Low	High
P11021	1.176	Medium	High	Low
P01876	1.1757	Medium	Low	High
A0A0B4J1X5	1.175	Medium	Low	High
P52566	1.1748	Low	Medium	High
P42126	1.1741	High	Medium	Low
P61009	1.1741	Medium	High	Low
Q00341	1.1739	Medium	High	Low
O95490	1.1731	Low	Medium	High
P08603	1.173	Low	Medium	High
Q9UHA4	1.173	Medium	High	Low
O00468	1.1729	Low	Medium	High
Q6UXI9	1.1723	Low	Medium	High
P02656	1.1721	Low	Medium	High
Q969P0	1.1717	Medium	Low	High
P84090	1.1709	High	Medium	Low
Q15382	1.1705	Low	Medium	High
P34741	1.1703	Medium	High	Low
Q9HB07	1.1701	High	Medium	Low
P53384	1.1694	High	Medium	Low
Q9Y3C6	1.1689	High	Medium	Low
Q9Y4E8	1.1689	High	Medium	Low

A6NHQ2	1.1684	High	Medium	Low
Q9BZE9	1.1678	High	Medium	Low
Q14166	1.1677	High	Medium	Low
Q9BXB4	1.1672	Medium	Low	High
Q07666	1.1669	High	Medium	Low
O75347	1.1668	High	Medium	Low
P0C0S5; Q71UI9	1.1663	Medium	High	Low
P61978	1.1663	High	Medium	Low
Q9UJS0	1.1657	High	Medium	Low
P17931	1.1655	Low	Medium	High
P01742	1.1654	Medium	Low	High
Q01469	1.1654	Medium	High	Low
P05455	1.1651	Medium	High	Low
Q15582	1.1648	Medium	Low	High
Q9UQ80	1.1648	High	Medium	Low
Q8WXF1	1.1646	High	Medium	Low
Q02878	1.1633	High	Medium	Low
P10599; cRAP-P10599	1.1632	High	Medium	Low
P22352	1.1629	Low	Medium	High
A8MWD9; P62308	1.1627	High	Medium	Low
Q14498	1.1624	High	Medium	Low
Q8TC07	1.1619	High	Medium	Low
P49755	1.1618	High	Low	Medium
Q07955	1.1618	High	Medium	Low
P25787	1.1617	High	Medium	Low
Q86U42	1.1614	High	Medium	Low
O94979	1.1608	High	Medium	Low
Q13557	1.1606	High	Medium	Low
P04844	1.1601	Medium	High	Low
Q16610	1.1601	Medium	Low	High
Q969U7	1.1598	High	Medium	Low
P42285	1.1596	High	Medium	Low
P16401	1.1595	High	Medium	Low
O14497	1.1594	High	Medium	Low
Q96AX1	1.1587	High	Medium	Low
Q7Z4H8	1.1578	Medium	High	Low
P38432	1.1576	High	Low	Medium
P55268	1.1568	Low	Medium	High
P62263	1.1567	High	Medium	Low
P68400	1.1567	High	Medium	Low
Q9Y3I0	1.1566	High	Medium	Low
Q8IY17	1.1561	Medium	High	Low
Q9Y6D6	1.1558	High	Medium	Low
P26368	1.1556	High	Medium	Low
P55001	1.1556	Low	Medium	High
P08237	1.1555	High	Medium	Low
Q9NTJ5	1.1551	Medium	High	Low

O00422	1.1547	High	Medium	Low
P14866	1.1543	High	Medium	Low
P13861	1.1542	High	Medium	Low
Q7Z4Q2	1.1537	High	Medium	Low
Q5SZK8	1.1536	Low	Medium	High
P49189	1.1528	Medium	High	Low
Q9UHQ9	1.1526	Medium	High	Low
Q14011	1.1525	High	Medium	Low
Q07507	1.1524	Low	Medium	High
P39687	1.152	High	Medium	Low
Q16181	1.1519	Medium	High	Low
P62424	1.1518	High	Medium	Low
P23470	1.1517	Low	Medium	High
Q15084	1.1517	Medium	High	Low
P42766	1.1514	High	Medium	Low
P15559; cRAP-P15559	1.1511	High	Medium	Low
P61353	1.151	High	Medium	Low
Q7Z4W1	1.15	High	Medium	Low
Q96I15	1.1497	Medium	High	Low
P51608	1.1494	Medium	High	Low
O75874	1.1493	High	Low	Medium
P48681	1.1487	Medium	High	Low
Q9UJZ1	1.1485	High	Medium	Low
P22314	1.1483	Medium	High	Low
P31930	1.1483	High	Medium	Low
Q96AE4	1.148	High	Medium	Low
Q8IXI2	1.1479	High	Medium	Low
Q93034	1.1475	Medium	High	Low
O95336	1.1473	Medium	High	Low
Q96CM8	1.147	High	Medium	Low
P30153	1.1468	High	Medium	Low
Q86V81	1.1464	High	Medium	Low
P12111	1.1462	Low	Medium	High
Q9BVG4	1.1455	High	Medium	Low
P01768; P0DP03	1.1449	Medium	Low	High
P51665	1.1445	High	Medium	Low
P25788	1.1444	High	Medium	Low
O76041	1.1436	Medium	High	Low
Q6IAA8	1.1432	Medium	High	Low
P29144	1.1427	High	Medium	Low
P62306	1.1425	High	Medium	Low
Q6UX71	1.1424	Low	Medium	High
Q15437	1.1421	High	Medium	Low
P62753	1.1418	High	Medium	Low
Q16630	1.1418	High	Medium	Low
Q8N1G4	1.1418	Medium	High	Low
P31948	1.1413	High	Medium	Low

Q8ND30	1.1413	High	Medium	Low
P25789	1.14	High	Medium	Low
Q16539	1.1397	High	Medium	Low
Q4G0J3	1.1394	High	Medium	Low
P60900	1.1385	High	Medium	Low
Q16629	1.1383	High	Medium	Low
Q9UNS2	1.1379	Medium	High	Low
Q13263	1.1376	High	Medium	Low
P35232	1.1369	High	Medium	Low
Q9UDY2	1.1366	High	Medium	Low
P35637	1.1365	High	Medium	Low
Q7L5N7	1.1363	Medium	High	Low
Q9Y3Y2	1.1363	High	Medium	Low
Q96GA7	1.136	Medium	High	Low
Q969Z3	1.1353	Medium	High	Low
Q9H6T0	1.1351	High	Low	Medium
Q6BCY4	1.1349	Low	Medium	High
O75340	1.1347	High	Low	Medium
P47897	1.1338	High	Low	Medium
P50897	1.1337	Medium	Low	High
Q9BQG0	1.1337	High	Medium	Low
P43304	1.1336	High	Medium	Low
Q9Y277	1.1332	High	Medium	Low
Q3I5F7	1.1329	High	Medium	Low
Q9H0N5	1.1327	Medium	High	Low
Q9Y587	1.1327	High	Medium	Low
P11940	1.1318	High	Medium	Low
P51532	1.1317	High	Medium	Low
Q14103	1.1311	High	Medium	Low
Q6IAN0	1.131	High	Medium	Low
P08865	1.1308	High	Medium	Low
Q6ZMZ3	1.1305	Medium	High	Low
P49748	1.1302	Medium	High	Low
P61326; Q96A72	1.1297	High	Medium	Low
O75431	1.1296	High	Medium	Low
Q13363	1.1296	High	Medium	Low
O75506	1.1291	Medium	High	Low
P36776	1.1284	High	Medium	Low
Q13586	1.1282	Medium	High	Low
P67936	1.128	Low	Medium	High
P49368	1.1279	High	Medium	Low
P12109	1.1278	Low	Medium	High
Q9NP58	1.1278	Medium	High	Low
P46977	1.1276	Medium	High	Low
Q6PHW0	1.1275	High	Medium	Low
Q9H4A6	1.127	Medium	High	Low
O15264	1.1268	High	Medium	Low

P62316	1.1268	High	Medium	Low
Q9Y237	1.1267	Medium	High	Low
Q9NX61	1.1264	High	Medium	Low
Q13247	1.1262	High	Medium	Low
O43149	1.1261	High	Medium	Low
Q8NFI3	1.1259	Medium	High	Low
Q9H3N1	1.1258	High	Medium	Low
P52597	1.1256	High	Medium	Low
P43243	1.1254	High	Medium	Low
P06733	1.1253	Low	High	Medium
P01859	1.125	Medium	Low	High
Q15056	1.1249	High	Medium	Low
P62304	1.1248	High	Medium	Low
P99999; cRAP-P99999	1.1245	High	Medium	Low
Q9UBQ6	1.1243	Medium	Low	High
O60341	1.1242	High	Medium	Low
P62899	1.1241	High	Medium	Low
Q5JU69	1.124	High	Medium	Low
Q9UM47	1.124	Low	Medium	High
P49916	1.1239	High	Medium	Low
Q99615	1.1238	High	Medium	Low
A0A0C4DH38	1.1231	Medium	Low	High
Q96PK6	1.1231	High	Medium	Low
P08574	1.1229	High	Medium	Low
P38435	1.1224	Medium	High	Low
Q15043	1.1219	High	Low	Medium
Q9BUT1	1.1219	Medium	High	Low
P26927	1.1218	Medium	Low	High
P62314	1.1216	High	Medium	Low
Q5T6V5	1.1216	High	Medium	Low
O94903	1.1212	High	Medium	Low
O94766	1.1207	Medium	Low	High
Q13131	1.1205	High	Medium	Low
P52272	1.1202	High	Medium	Low
Q14257	1.1199	High	Medium	Low
Q13098	1.1197	High	Medium	Low
A0A0A0MS15	1.1196	Medium	Low	High
P26447	1.1194	Low	Medium	High
O00267	1.1192	High	Medium	Low
P04843	1.119	Medium	High	Low
Q03426	1.1185	High	Medium	Low
Q96AQ6	1.1181	High	Medium	Low
P63162	1.118	High	Medium	Low
B5ME19; Q99613	1.1175	High	Medium	Low
P50502	1.1174	High	Medium	Low
O94822	1.1173	High	Medium	Low
Q9UL46	1.1173	High	Medium	Low

Q9Y5B9	1.1173	High	Medium	Low
Q5T5P2	1.1172	Medium	High	Low
Q01130	1.1166	High	Medium	Low
O76024	1.1165	Medium	High	Low
P38117	1.1159	High	Medium	Low
P34932	1.1155	High	Medium	Low
Q8NCH0	1.1153	Medium	Low	High
Q9NZL9	1.1153	High	Medium	Low
P00488	1.1152	Low	Medium	High
E9PAV3; Q13765	1.115	High	Medium	Low
P06681	1.1142	Medium	Low	High
P0DI81; P0DI82	1.1138	High	Medium	Low
Q96P70	1.1137	High	Medium	Low
Q9UKU7	1.1137	High	Medium	Low
P08559	1.1131	High	Medium	Low
Q9UKV3	1.1131	High	Medium	Low
P40189	1.1128	Low	Medium	High
Q9H078	1.1128	High	Medium	Low
Q9NZM1	1.1124	Low	Medium	High
Q9Y3B4	1.1121	High	Medium	Low
Q96E39	1.1119	Medium	High	Low
P30040	1.1118	High	Medium	Low
Q15293	1.1116	Medium	High	Low
Q8NC51	1.1112	High	Medium	Low
Q9NQX3	1.1111	High	Medium	Low
P25786	1.1099	High	Medium	Low
A0A0A0MRZ8; P04433	1.1097	Medium	Low	High
Q14152	1.1094	High	Medium	Low
O14979	1.1092	High	Medium	Low
P31942	1.1091	High	Medium	Low
O15382	1.109	High	Medium	Low
P62995	1.1089	High	Medium	Low
Q9C0C9	1.1088	High	Medium	Low
P82909	1.1084	High	Medium	Low
Q9H0S4	1.1084	High	Medium	Low
Q9P260	1.1081	High	Medium	Low
A0A075B6S2; A2NJV5	1.1079	Medium	Low	High
P02647	1.1075	Low	Medium	High
P32321	1.1075	Medium	High	Low
Q8N4Q0	1.1074	High	Medium	Low
O75533	1.1069	High	Medium	Low
P49756	1.1067	High	Medium	Low
P25686	1.1061	Medium	High	Low
P63279; cRAP-P63279	1.1053	High	Medium	Low
P63220	1.105	High	Medium	Low
Q3ZCM7	1.1049	High	Medium	Low
P13667	1.1045	High	Medium	Low

Q13895	1.1043	High	Medium	Low
P22626	1.1042	High	Medium	Low
Q2VPB7	1.1042	High	Medium	Low
P61313	1.1041	High	Medium	Low
Q5SSJ5	1.1038	High	Medium	Low
P31943	1.1037	High	Medium	Low
Q5K4L6	1.1032	Medium	High	Low
P04632	1.1031	Medium	High	Low
P80404	1.103	High	Medium	Low
P24666	1.1029	High	Medium	Low
P06727	1.1028	Medium	Low	High
P02766	1.1027	Medium	Low	High
P39019	1.1026	Medium	High	Low
Q8TF74	1.1025	High	Medium	Low
P21810	1.1014	Low	Medium	High
Q9UBS4	1.1014	Medium	High	Low
Q00839	1.1008	High	Medium	Low
Q6ZVM7	1.1004	High	Medium	Low
O43678	1.0995	Medium	High	Low
P62750	1.0995	High	Medium	Low
Q8IYB5	1.0995	Medium	High	Low
O60936	1.0993	Medium	High	Low
P25398	1.099	High	Medium	Low
P27695	1.0988	High	Medium	Low
Q9Y5L4	1.0987	Medium	High	Low
P05155	1.0985	Medium	Low	High
O75494	1.0984	High	Medium	Low
Q9H5X1	1.0983	High	Medium	Low
O43815	1.0982	Medium	High	Low
P27105	1.0982	Low	Medium	High
P26599	1.0981	High	Medium	Low
Q9UQE7	1.0981	High	Medium	Low
Q8IW45	1.0978	Medium	High	Low
O94760	1.0975	Medium	High	Low
P21912	1.0975	High	Medium	Low
Q6ZMI0	1.0975	Medium	High	Low
P62857	1.0974	High	Medium	Low
Q9H1E3	1.0973	High	Medium	Low
Q9NNW7	1.0972	High	Medium	Low
A0A0C4DH67	1.0971	Medium	Low	High
Q8WYA6	1.097	High	Medium	Low
P55010	1.0969	High	Medium	Low
P02760	1.0968	Low	Medium	High
P62851	1.0965	High	Medium	Low
Q14683	1.0964	High	Medium	Low
Q16401	1.0961	High	Medium	Low
P33240	1.0958	High	Medium	Low

Q15637	1.0958	High	Medium	Low
P01011	1.0955	Medium	Low	High
P35659	1.0955	High	Medium	Low
O75146	1.0954	High	Medium	Low
Q9Y2V2	1.0948	Medium	High	Low
A0A075B6K5; P80748	1.0945	Medium	Low	High
P78527	1.094	High	Medium	Low
Q15717	1.0933	High	Medium	Low
P42167	1.093	High	Medium	Low
Q9BRP4	1.0928	Medium	High	Low
P98160	1.0922	Medium	Low	High
P60866	1.0919	High	Medium	Low
Q8IYB3	1.0918	High	Medium	Low
Q14112	1.0916	Medium	Low	High
Q9Y265	1.0912	High	Medium	Low
Q8TDM6	1.0909	High	Medium	Low
O14647	1.0908	High	Medium	Low
P08238	1.0907	High	Medium	Low
P61758	1.09	Medium	High	Low
P55060	1.0899	High	Medium	Low
Q92841	1.0899	High	Medium	Low
Q9Y287	1.0897	Low	Medium	High
P51991	1.0896	High	Medium	Low
P21953	1.0893	Medium	High	Low
Q12805	1.0893	Medium	Low	High
Q9NRZ7	1.0887	High	Medium	Low
P12270	1.0886	High	Medium	Low
Q99714	1.0886	Medium	High	Low
Q9H7B4	1.0885	High	Medium	Low
O60739	1.0884	High	Medium	Low
O00519	1.0883	Medium	High	Low
A0M8Q6	1.0881	Medium	Low	High
P39060	1.088	Medium	Low	High
Q9UQ16	1.0879	Medium	High	Low
Q9NUS5	1.0877	High	Medium	Low
P20339	1.0876	Medium	High	Low
P54819	1.087	High	Medium	Low
Q9NWV4	1.0869	Medium	High	Low
P01009	1.0868	Medium	Low	High
Q53H96	1.0863	High	Medium	Low
A1L0T0	1.086	Medium	High	Low
P07954	1.0854	High	Medium	Low
Q9Y285	1.0853	High	Medium	Low
A0A0C4DH72; A0A0C4DH73; P01597; P01611; P04432	1.0852	Medium	Low	High
Q99623	1.085	High	Medium	Low
Q6ZT12	1.0847	High	Medium	Low

P05387	1.0845	High	Medium	Low
Q8NFH5	1.0845	High	Medium	Low
P27169	1.0844	Low	Medium	High
P46777	1.0843	High	Medium	Low
Q96C01	1.084	High	Medium	Low
Q9NU22	1.0839	High	Medium	Low
P42345	1.0837	Medium	High	Low
Q8TBX8	1.0837	High	Medium	Low
Q9Y230	1.0837	High	Medium	Low
P00746	1.0836	Medium	Low	High
Q9UBQ5	1.0834	High	Medium	Low
P36578	1.0833	High	Medium	Low
Q16543	1.0832	Medium	High	Low
P04004	1.083	Medium	Low	High
P49411	1.083	High	Medium	Low
P33993	1.0829	High	Medium	Low
Q14978	1.0827	High	Medium	Low
A0A087WW87;P01614	1.0826	Medium	Low	High
Q9BSL1	1.0826	Medium	High	Low
P78346	1.0822	High	Medium	Low
Q9UKG9	1.0822	Medium	High	Low
Q9UKM9	1.0819	High	Medium	Low
P51572	1.0804	Medium	High	Low
O43760	1.0803	Medium	Low	High
P12694	1.0803	Medium	High	Low
Q9H3U1	1.0803	High	Medium	Low
P10586	1.0802	Medium	Low	High
P51148	1.0798	Medium	High	Low
Q9GZZ9	1.0798	Medium	High	Low
Q7KZF4	1.0797	Medium	High	Low
Q9BT78	1.0795	High	Medium	Low
Q15526	1.0793	High	Medium	Low
Q92522	1.0791	High	Medium	Low
Q6UW02	1.0788	Medium	High	Low
P07910	1.0782	High	Medium	Low
O00358	1.078	High	Medium	Low
Q9NXG6	1.078	High	Medium	Low
Q86TN4	1.0774	High	Medium	Low
P29692	1.0771	High	Medium	Low
P0C0L4	1.0769	Low	Medium	High
P62906	1.0769	High	Medium	Low
P78347	1.0769	High	Medium	Low
Q9BQ52	1.0769	High	Medium	Low
Q9P2G3	1.0769	High	Medium	Low
Q9NQG5	1.0767	High	Medium	Low
P01834	1.0766	Medium	Low	High
Q9Y2U8	1.0762	Medium	High	Low

P30044	1.0761	Medium	High	Low
Q13838	1.0756	High	Medium	Low
Q9BRP8	1.0752	High	Medium	Low
Q03252	1.0751	Medium	High	Low
Q9UBQ7	1.075	High	Medium	Low
Q9Y5K5	1.0749	Medium	High	Low
P37108	1.0748	Medium	High	Low
P45984	1.0745	Medium	High	Low
Q9H3K2	1.0738	High	Medium	Low
Q4G176	1.0736	High	Medium	Low
O94874	1.0735	Medium	High	Low
O43290	1.0733	High	Medium	Low
O00244	1.0732	Medium	High	Low
Q9H1B7	1.0732	High	Medium	Low
P19338	1.0731	High	Medium	Low
P62081	1.0731	High	Medium	Low
P43246	1.0729	High	Medium	Low
Q7Z478	1.0726	High	Medium	Low
Q8IY81	1.0726	High	Medium	Low
P07305	1.0724	Medium	High	Low
P12532	1.0724	Medium	High	Low
Q9UBB6	1.0723	High	Medium	Low
Q9BT40	1.0721	Medium	High	Low
P49721	1.0717	High	Medium	Low
P82979	1.0717	High	Medium	Low
Q8IXK0	1.0717	High	Medium	Low
Q6P6C2	1.071	High	Medium	Low
P43699	1.0708	High	Medium	Low
Q9NUU7	1.0708	High	Medium	Low
P41223	1.0704	High	Medium	Low
Q8IZE3	1.0704	High	Medium	Low
O95292	1.0702	High	Medium	Low
Q8WX92	1.0702	High	Medium	Low
Q15643	1.0701	High	Medium	Low
Q9BS26	1.0697	High	Medium	Low
Q00765	1.0696	High	Medium	Low
Q06546	1.0696	Medium	High	Low
P48643	1.0695	Medium	High	Low
Q96EK7	1.069	High	Medium	Low
P30050	1.0689	High	Medium	Low
Q12789	1.068	High	Medium	Low
Q96P16	1.0678	High	Medium	Low
Q9H8H3	1.0674	Medium	High	Low
P00918; cRAP-P00918	1.0666	High	Medium	Low
P14314	1.0666	Medium	High	Low
Q9UMS0	1.0666	High	Medium	Low
Q8N2F6	1.0655	Medium	High	Low

Q92600	1.0654	High	Medium	Low
Q9HD26	1.0654	High	Medium	Low
Q9NRY6	1.065	Low	Medium	High
Q9H553	1.0647	Medium	High	Low
Q92945	1.0646	High	Medium	Low
O14787	1.0644	High	Medium	Low
Q93008	1.0641	Medium	High	Low
O15230	1.0637	Medium	Low	High
P40925	1.0635	Medium	High	Low
P62310	1.0634	High	Medium	Low
Q7L2E3	1.0624	High	Medium	Low
O43390	1.0623	High	Medium	Low
Q7L099	1.0617	High	Medium	Low
Q15459	1.0615	High	Medium	Low
P14625	1.0611	Medium	High	Low
P15374	1.0611	Medium	High	Low
Q9P2M7	1.0609	High	Medium	Low
P51659	1.0607	High	Medium	Low
Q14657	1.0602	High	Medium	Low
P13639	1.0598	High	Medium	Low
Q96I99	1.0593	Medium	High	Low
Q9UKN8	1.0588	High	Medium	Low
P08621	1.0587	High	Medium	Low
P49720	1.0586	High	Medium	Low
Q86YS7	1.0584	High	Medium	Low
Q9NZJ7	1.0581	High	Medium	Low
Q9NZJ9	1.0559	High	Medium	Low
P02763	1.0558	Medium	Low	High
Q9NYR9	1.0558	High	Medium	Low
P09661	1.0557	High	Medium	Low
Q15477	1.0551	High	Medium	Low
Q93074	1.0548	High	Medium	Low
P46379	1.0542	High	Medium	Low
P43897	1.0536	High	Medium	Low
Q96T60	1.0523	High	Medium	Low
Q12929	1.0522	High	Medium	Low
P61289	1.0511	High	Medium	Low
Q86YP4	1.0503	High	Medium	Low
Q6PJG6	1.0501	High	Medium	Low
P51888	1.05	Low	Medium	High
Q8N8N7	1.0496	Medium	High	Low
O75884	1.0482	Medium	High	Low
Q9H488	1.0474	Medium	High	Low

Table S1. Summary of Variable Importance in Projection (VIP) values calculated from the optimized PLS-DA model, highlighting the most influential features selected for class separation.

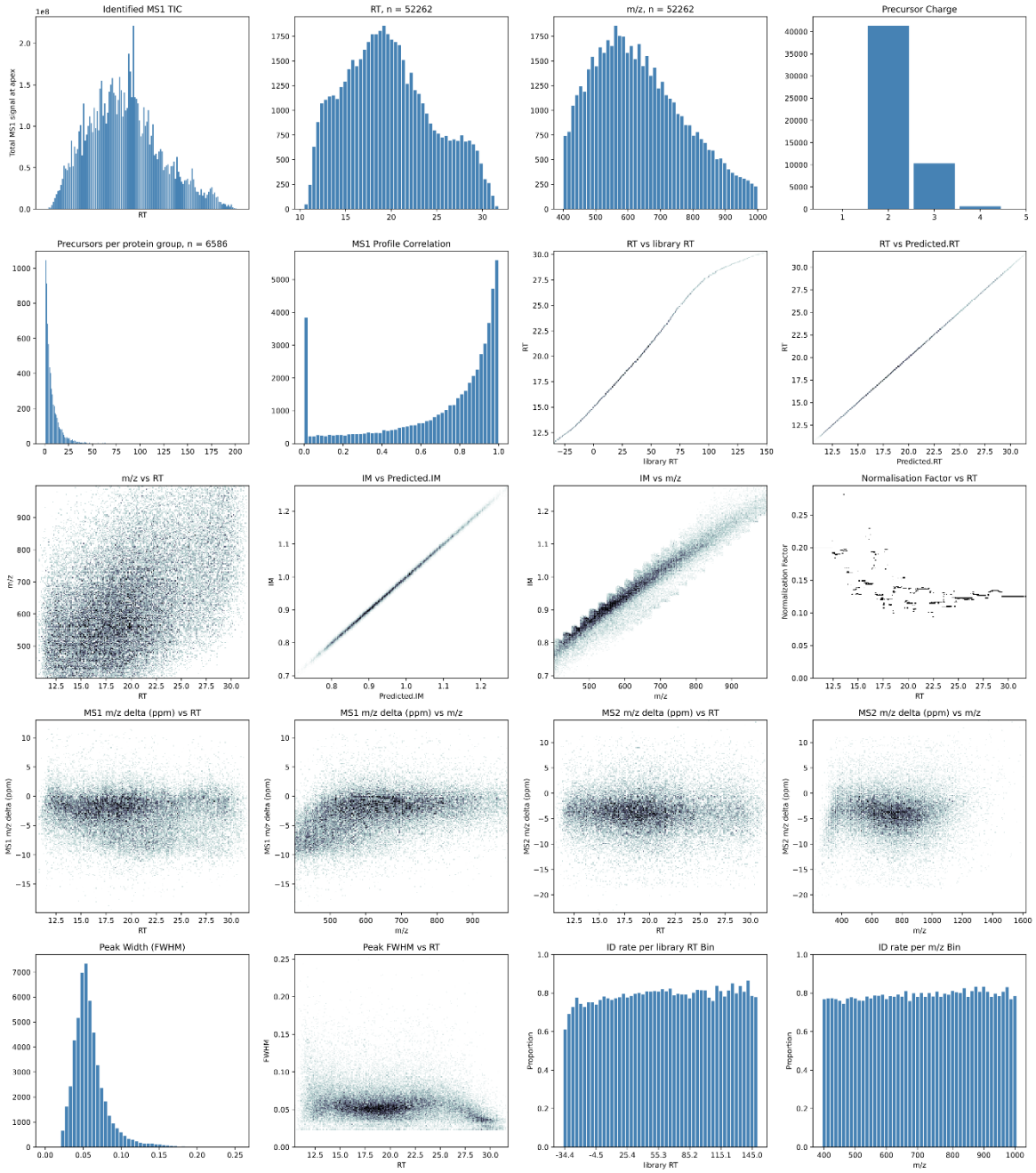


Figure S1. Quality control and identification metrics generated by DIA-NN. The panels display the distribution of identified MS1 intensities, retention times (RT), m/z values, precursor charge states, and the number of precursors per protein group. Additional metrics include mass accuracy, RT alignment, MS1 profile correlation, and ion mobility. Overall, the plots indicate high data quality, adequate chromatographic separation, and strong consistency between experimental data, the spectral library, and predictive models.