Evaluation of extraction solvents for untargeted metabolomics to decipher the DOM of Antarctic cryoconite holes

Swapnil Mundhe[a,b]; Saborni Maiti[c,h]; Aritri Sanyal[e]; Narendra Y Kadoo[c,d]; Dhiraj Dhotre[a]; Vitthal T Barvkar[f]; Shamim A. Shaikh[b]; Runa Antony[e,g]; Dhiraj Paul[a,h#]

[a]National Center for Cell Science; Pune, Maharashtra-411008, India

[b]Rajiv Gandhi Institute of IT and Biotechnology, Bharati Vidyapeeth (Deemed University), Katraj, Pune, Maharashtra-411046, India

[c]Biochemical Sciences Division, CSIR-National Chemical Laboratory, Pune, Maharashtra 411008, India

[d]Academy of Scientific and Innovative Research (AcSIR), Ghaziabad, Uttar Pradesh-201002, India

[e]ESSO–National Centre for Polar and Ocean Research, Ministry of Earth Sciences, Vasco-Da-Gama, Goa-403804, India

[f]Department of Botany, Savitribai Phule Pune University, Pune, Maharashtra, 411007, India

[g]GFZ German Research Center for Geosciences; 14473 Potsdam, Germany

[h]Department of Environmental Science and Biology, University of Eastern Finland, Kuopio-70210, Finland

#Corresponding author:
Dhiraj Paul; Department of Environmental Science and Biology, University of Eastern Finland, Kuopio, 70210, Finland, Email: dhiraj.paul@uef.fi

Highlights

- This is the first report on LCMS-based untargeted metabolomics to decipher the small molecules present in the DOM of Antarctic cryoconite holes
- Extraction solvents were critically assessed for the highest metabolic recovery in terms of quantity (number) and quality (chemical diversity) for RPLC-MS
- We discussed the reproducibility and efficiency of the solvents based on their polarity in extracting the metabolites and the use of dual LC (RP and HILIC) and ionization polarity (positive and negative) to increase further the breadth of the detection of the metabolome of cryoconite holes.
- We suggested the best single and two solvent combinations for the optimum extraction of the metabolites and future studies to improve the accuracy of the solvent assessment further.
- This data can be integrated with other 'omics' approaches to gain deeper insights into underlying carbon and nutrient cycling mechanisms in global cryoconite holes and similar ecosystems in response to climate change.

Abstract

Cryoconite holes (CHs) are biological hotspots with a high biogeochemical turnover rate, contributing significantly to the glacial ecosystem's overall carbon cycles and net fluxes. There is limited information regarding the composition of low molecular weight (LMW) molecules formed through the metabolic processes of cryoconite-dwelling microbes. These molecules constitute a substantial portion of the dissolved organic matter (DOM) within cryoconite holes. Here, we evaluated different solvents to extract low molecular weight (LMW) compounds for untargeted metabolomics using reverse phase liquid chromatography (RP-LC) coupled with high-resolution tandem mass spectrometry in positive- and negative-ionization modes. We prepared single, binary, and ternary mixtures of highly polar to relatively non-polarsolvents like water, methanol, and acetonitrile to extract intra- and extracellular metabolites from CHs sediment. The biological replicates (n=4) of each identical solvent, showed high reproducibility in metabolite diversity while substantial differences were observed among different solvent types. Among the single solvents, organic-rich 70:30 MeOH: water and in parallel 2-single solvent combination of 70:30 MeOH: Water and 40:40:20 Acetonitrile: Methanol: Water provided increased number and chemical diversity of extracted metabolites. Combining RP with the hydrophilic interaction liquid chromatography (HILIC) technique provided the highest number of unique metabolites. HILIC and RP detected polar and mid- to non-polar molecules at high intensity, respectively. This dual-LC and ionization polarity combination increased the detection of metabolic features by 46.96% and 24.52% in single- and two-solvent combinations compared to RP alone. This study developed a simple untargeted metabolomics workflow that is highly sensitive and robust, detecting and potentially identifying a large number of broad chemically diverse molecules present in the DOM (extracellular) and microbes (intracellular)

from the CH's environment. This method can better characterize DOM's chemical composition and, after integrating with other 'omics' approaches, can be used to examine the link between metabolic pathways and microbial communities in global CHs or other similar ecosystems, revealing how these earthy systems and their microbial flora control carbon or nutrient storage or release in response to global climate change.

Key Words
Cryoconite hole
Dissolved organic matter
Untargeted metabolomics
Reverse-phase liquid chromatography
Hydrophilic interaction liquid chromatography

Abbreviations
CHs: Cryoconite holes
DOM: Dissolved organic matter
DOC: Dissolved organic carbon
EIC: Extracted ion chromatogram
ESI: Electrospray ionization
FTICR-MS: Fourier transform ion cyclotron resonance-mass spectrometry
GC-MS: Gas chromatography-mass spectrometry
HESI: Heated electrospray ionization
HILIC-LCMS: Hydrophilic interaction liquid chromatography-mass spectrometry
HRMS: High-resolution mass spectrometry
LC: Liquid chromatography
LC-MS/MS: Liquid chromatography-tandem mass spectrometry
LMW: Low molecular weight
NMR: Nuclear Magnetic Resonance
PCA: Principal component analysis
QC: Quality control
RP-LCMS: Reverse phase liquid chromatography-mass spectrometry
UHPLC: Ultra high-pressure liquid chromatography
Water: 100% water
30MW: 30:70 Methanol: Water

30AW: 30:70 Acetonitrile: Water
MeOH: 100% Methanol
ACN: 100% Acetonitrile
70AW: 70:30 Acetonitrile: Water
70MW: 70:30 Methanol: Water
AMW: 40:40:20 Acetonitrile: Methanol: Water
60HILIC: 60: 40 acetonitrile: water
75HILIC: 75: 25 acetonitrile: water
90HILIC: 90: 10 acetonitrile: water
95HILIC: 95: 5 acetonitrile: water

Introduction

Low molecular weight (LMW, 50-1500 Da) molecules like small peptides, carbohydrates, nucleic acids, organic acids, lipids, and other metabolites present in the dissolved organic matter (DOM) fractions that are readily available for microbial processing (Ladd et al., 2019). Microorganisms are a critical component in the terrestrial biosphere as they are actively involved in the cycling of DOM (i.e., production of microbial products and metabolic decomposition) (Schmidt et al., 2022). Therefore, fingerprinting of LMW DOM provides a detailed snapshot of organic molecules that are consumed or produced by microbes. This is especially relevant for Antarctic ecosystems that experience rapid changes due to climate change. However, the heterogeneity of this pool, coupled with consistently low concentrations of these molecules in the Antarctic or such extreme ecosystems, poses significant challenges in detection and quantitation.

Cryoconite holes (CHs) are self-contained water-filled cylindrical holes with cryoconite (dark-colored sediment) deposited at their undermost, and are formed on ice sheets and glaciers worldwide (Poniecka et al., 2020; Rozwalak et al., 2022). They are formed when organic or inorganic particles are deposited on the glacial ice surface. Being darker, they absorb more energy from the incoming solar irradiation compared to the surface ice environments, melting downwards, and creating water-filled cylindrical depressions (Gribbon, 1979; Cook et al., 2016). Because of dark mineral and organic/inorganic matter (cryoconite) at the bottom covered by ice melt-water, CHs can have a high biogeochemical turnover rate (Bagshaw et al., 2013; Zawierucha et al., 2018a, 2018b, and 2022). Despite their small individual size, several microbial-mediated biogeochemical processes within them can significantly contribute to the

glacier ecosystem's overall nutrient cycles and net fluxes (Anesio et al., 2009,2017). Due to this unique nature containing high DOM and organic compounds, CHs allow diverse microbial communities to flourish in this unique and harsh environment (Sanyal et al., 2018; Sommers et al., 2019). While several studies have reported microbial diversity and its composition in the CHs ecosystem only a few have reported the metabolic potential of the indigenous microbial communities of this ecosystem (Gokul et al., 2023) and supraglacial snow (Lutz et al., 2015; Doting et al., 2024). Despite the importance of microbial metabolites in nutrient cycles and ecosystem dynamics, optimization of an untargeted metabolomics approach to investigate the LMW DOM (henceforth referred to as small molecules), and their role in the cryoconite hole systems of both northern and southern hemispheres is very limited.

Untargeted metabolomics is a robust and rapidly growing method to analyze complex extracts by providing comprehensive data-driven metabolism analyses (Di Minno A et al., 2021; Bowen et al., 2023; Bittremeiux W et al., 2023). Recent studies indicate that an untargeted metabolomics approach can reveal a wide range of compounds in the DOM of terrestrial and marine ecosystems, increasing our understanding of the significance of DOM in structuring microbial populations and its role in global climate change. Even though relatively underdeveloped, these studies are gaining much recent interest and rising as an emerging area of study (Swenson et al., 2015; Ladd et al., 2019,2021; Catalá et al., 2021; Yao et al., 2023). Glacial ecosystems like CHs, as mentioned before, can produce significant amounts of dissolved organic matter because of the biochemical activities, which can potentially seed this bio-available material to nearby aquatic environments (Foreman et al., 2007; Feng et al., 2016). Although the high DOC and nutrients are known to be present in CHs, the interactions between the various organic substrates and microbes in the area are poorly understood (Sanyal et al.,2018; 2020). While some studies have explored DOM characterization at the compound level in various terrestrial ecosystems (Mann P et al., 2012; Singer et al., 2012; Nebbioso et al., 2013; Pautler et al., 2013; Antony et al., 2014; Mann B et al., 2015; Swenson et al., 2015; Grewer et al., 2016), the number of such investigations still remain limited. In these studies, Nuclear Magnetic Resonance (NMR) spectroscopy, Fourier transform ion cyclotron resonance mass spectrometry (FTICR-MS), UV-visible or excitation-emission matrix fluorescence spectroscopy and gas chromatography-mass spectrometry (GC-MS) measurements were used to apply an untargeted metabolomics approach to reveal a range of DOM compounds. Even though the choice of analytical platform is study-specific, these techniques have inherent limitations. For example, FTICR-MS can offer the highest resolution and mass accuracy but at

the expense of longer runtimes due to low data acquisition rates and are less suitable for high-throughput data acquisition than orbitrap instruments (Hohenester et al., 2020). UV-visible or excitation-matrix spectroscope requires detectable chromophores in each analyte, and highly complex mixtures may contain molecules that don't have chromophores (Vitale et al., 2024). NMR is less sensitive, and the number of metabolites detected is also low; GC-MS-based experiments require volatile analytes or prior derivatization of samples (Martin et al., 2014; Ladd et al., 2019).

Consequently, in recent years, LC-ESI-MS has emerged as a powerful tool for analyzing mixtures of small molecules. Its broad detection range, high-throughput capabilities, and sensitivity make it an ideal alternative to the traditional methods of characterizing DOM (Patti, 2011; Dunn et al., 2013; Chetwynd et al., 2018). Although RP-liquid chromatography (RP-LC) in positive ionization mode is preferred, drawbacks of using only a single chromatography phase or polarity mode are reported (Gika et al., 2014). The HILIC (hydrophilic interaction liquid chromatography) technique is more suitable for retaining and separating highly polar molecules than RP-LC, which can retain and separate mid- and non-polar molecules (Boudah et al., 2014; Contrepois et al., 2015). Combining RP- and HILIC-LC-MS techniques is an attractive complementary approach to identifying compounds for metabolomics to cover a broad spectrum of metabolome (Tang et al., 2016). This untargeted global metabolite profiling would bridge the knowledge gap that limits our understanding of DOM composition and metabolic pathways in microbial communities(Swenson et al., 2015; Ladd et al., 2019).

Here, we perform a comparative evaluation of the different binary and ternary mixtures of the most commonly available MS grade solvents for RPLC-MS/MS-based untargeted metabolomics (positive and negative ESI mode) to extract a broad range of metabolites (intra- and extracellular) from CHs samples collected from Larsemann Hills, Antarctica. Further, we combine results with HILIC-LC-MS/MS in ESI(+)ve and ESI(-)ve ionization modes to increase the metabolite detection range. To the best of our knowledge, this is the first time a systematic optimization study has delineated an LC-MS/MS-based untargeted metabolomics approach to uncover DOM composition in cryoconite holes. This study lays the technical foundation for future experiments aiming to perform untargeted metabolomics studies to understand the low molecular weight metabolites composition of the cryoconite holes worldwide, aiming to integrate LMW DOM molecular data into process-based ecological models.

Materials and Methods

Chemicals

LC-MS grade solvents Methanol (CAS 67-56-1), Acetonitrile (CAS 75-05-8), Water (CAS 7732-18-5), and Isopropanol (CAS 67-63-0) were from JT Baker (USA). LC-MS grade chemicals Ammonium formate (CAS 540-69-2) and formic acid (CAS 64-18-6) were from Sigma (Germany). Authentic metabolic standards like L-alanine, L-glutamic acid, L-methionine, L-tryptophan, L-tyrosine, L-phenylalanine, L-proline, L-malic acid, L-isoleucine were from Sigma (Germany).

Cryoconite hole sample collection and processing

Sediment samples from cryoconite holes in the Larsemann Hills region of East Antarctica were obtained during the 38th Indian Scientific Expedition to Antarctica during the austral summer of 2019-20 **(69° 25' 13.09"S, 76° 13' 22.85 E)**. Dissolved oxygen, pH, temperature, oxidation-reduction potential, salinity, total dissolved solids, and conductivity were measured during sample collection (Table S1). The frozen samples were initially transported to the National Centre for Polar and Ocean Research (India) for sub-sampling and subsequently transferred under dry conditions to the National Center for Cell Science-Pune, Indiastored at -80 °C until analysis. Before metabolite extraction, samples were lyophilized, and coarse sand and large stones were removed. Samples were pulverized with a sterile, liquid N_2 chilled mortar and pestle, and the powdered samples were stored at -80 °C.

Metabolites extraction conditions for the Cryoconite sediment

An equal amount of homogenized powdered samples from 20 CHs were mixed and used to evaluate extraction solvents. Investigations were performed using the following extraction method to optimize experimental conditions of metabolite extraction for reverse phase RP-LC-MS/MS analysis. Before extraction, all the LC-MS grade solvents were filtered via a 0.22 µm syringe filter. Eight extraction solvents of single, binary, and ternary mixtures of water, methanol, and acetonitrile were prepared as 100% water (Water), 30:70 Methanol: Water (30MW), 30:70 Acetonitrile: Water (30AW), 100% Methanol (MeOH), 100% Acetonitrile (ACN), 70:30 Acetonitrile: Water (70AW) and 70:30 Methanol: Water (70MW) and 40:40:20 Acetonitrile: Methanol: Water (AMW). The samples were extracted in quadruplets. Sample pre-preparation, data acquisition, and analysis are shown in Workflow 1. All extraction was done using 2g of sample biomass (Swenson et al., 2015). A finely powdered sample was

transferred into a 2 mL centrifuge tube. One mL of ice-chilled extraction solvent was added, followed by vortexing for 1 min to mix the sample and solvent, with subsequent homogenization in TissueLyser II (QIAGEN, Germany) and incubation (at 4 °C) for 5 min, respectively. Samples were sonicated for 10 min on ice to keep the temperature low and incubated again at 4 °C for 5 min, followed by vortexing for 5 min and centrifugation at 13500 rpm for 15 min at 4 °C. The supernatant was collected and filtered through a 0.22 µm syringe filter. Samples were evaporated to near dryness and volume made upto 500 µL using the respective extraction solvent. For HILIC phase LC-MS/MS analysis, 2g of sample was extracted in water using the above-mentioned extraction method. Dried extracts were reconstituted in 500 µL of 60% (60HILIC), 75% (75HILIC), 90% (90HILIC), and 95 % (95HILIC) of the acetonitrile-water solvent mixture. Samples were stored at -80 °C till further use.

Evaluation of extraction yield

Absolute metabolites extraction yield was estimated for RP extracts by measuring the remaining dried residue after evaporating 500 µL of solvent supernatant (extract) to dryness (Savant DNA-120 SpeedVac concentrator; Thermo Scientific, USA). The extraction yield was determined as a percent of the initial dry weight of the sample utilized to produce 500 µL of extract. All the extracts were digitally photographed and visually inspected for color, clarity, and resemblance.

LC-MS/MS data acquisition

Thermo Q-Exactive-Orbitrap™ high-resolution mass spectrometer (HRMS) coupled to an Accela™ 1200 ultra high-performance liquid chromatography (UHPLC) system (Thermo Fisher) was used for data acquisition in ESI (+)ve and ESI (-)ve mode. A heated electrospray ionization (HESI) source settings were as follows: capillary and gas temperature was 320 ºC; auxiliary, sheath, and spare gas were at 45, 12, and 2 arbitrary units, respectively, in both the modes; spray capillary voltage was 4.2 kV and 3.8 kV in positive and negative mode, respectively. The tube lens was set to 45 V, and the mass scan range was set from 50-1000 m/z. The resolution power of the Orbitrap was set at 70,000. The instrument was calibrated before running samples as per the vendor's instruction. A Hypersil GOLD C18 column (1.9 µm, 2.1mm × 150, Thermo Fisher) was utilized for reverse phase chromatography (RP). The column oven temperature was maintained at 40 ºC, and the sample manager temperature was set at 4 ºC. Mobile phases A (0.1% formic acid) and B (ACN, 0.1% formic acid) were employed to separate metabolites. Linear gradient elution was set to 20 min with a 600 µL/min flow rate.

Eluent A was reduced from 98% to 60% between the start and 5 min, 40% at 7 min, 20% at 10 min, and 2% at 13 min. It was restored to its beginning condition after 16 min and held for another 4 min. For HILIC phase chromatography, a Syncronis C18 column (1.7 µm, 2.1mm × 100, Thermo Fisher) was used and maintained at 40 ºC. Mobile phase A (95% ACN, 5% 5 mM ammonium formate, 0.1% formic acid) and B (5 mM ammonium formate, 0.1% formic acid) were used to separate metabolites. The gradient was: at 0-2 min, eluent A was 95 %; at 14 min, eluent A was brought to 60%; maintained until 16 min, and to 95 % at 16.5 min and held until 20 min before injecting the following sample. The flow rate was set at 600 µL /min. The mass spectrometer parameters were the same as the reverse phase.

Thermo Xcalibur version 3.0.63 was used to record chromatograms and spectra in profile mode. The instrument was periodically checked for ambient operating conditions. Samples were injected randomly with an injection volume of 3 µL. Pooled quality control (QC) samples were prepared by mixing all the samples to ensure quality throughout the run and the quality of the metabolic profile. Before beginning actual sample data acquisition, a series of QC samples were injected to equilibrate and condition the system, followed by an interval of nine samples. QC samples were also used for the tandem MS data acquisition. In RP, MS/MS data acquisition was conducted at 40 NCE collision energy in PRM mode using an inclusion list containing the m/z of detected ions and their associated RTs acquired from prior data analysis, and in HILIC data was acquired in data-dependent MS/MS (ddMS^2) mode with top N=10.

LC-MS data processing and analysis

Raw data processing, peak picking, and conversion of raw data from profile (.raw) to centroid mode format, i.e., mzXML of both ESI (+) ve and ESI (-) ve ionization modes were done by ProteoWizard v 3.0.21082 (Chambers et al., 2012). Peak detection, retention time correction, peak grouping, and probabilistic putative metabolic annotation were performed using R-based (Version 3.3.1) ProbMetab tool (Silva et al., 2014) equipped with packages like xcms (Smith C et al., 2006), CAMERA (Kuhl C et al., 2012), mzMatch, Comb2+ algorithm, and metExplore. The different metabolic classes were assigned using R Package omu (Tiffany et al., 2019). Data for the ESI (+) ve and ESI (-) ve ionization modes were processed separately for RP and HILIC and merged for further analysis (Table S5, S6, S7, and S8).

For multivariate statistical analysis of the RP data (all solvents combined), duplicate and unknown features were removed first. Artificial artifacts, noises, and contaminations were removed manually using the blanks as a reference. The resulting data frame was uploaded to

the freely available web-based platform MetaboAnalyst 5.0 (http://www.metaboanalyst.ca). The data filtering step applied a 10% variance filter based on the interquartile range (IQR) to remove detected near-constant features across all solvents. Further, data was normalized by sum to adjust systematic differences among samples followed by log transformation (base 10) and Pareto scaling (mean-centered and divided by the square root of the standard deviation of each variable). Fig. S4 shows the data distribution pattern before and after normalization. Next, parametric one-way ANOVA was applied to select 499 significant features ($p<0.05$), followed by principal component analysis (PCA). For hierarchical cluster analysis (Heatmap), we selected the top 65 features with variable importance in projections, with a VIP score >1.5 (Fig S6) via partial least square discriminant analysis (PLS-DA, Fig S5)(Eriksson et al., 2006). Of 65 features, 8 were confirmed (level 1), and 19 were putatively confirmed (level 2). Confirmed identification of the m/z was performed using MS/MS spectral match with authentic standards (Table S3 and Fig. S8). Where authentic standards were not available, putative MS/MS spectral match was carried out by extracting MS2 spectra (fragment list) of desired parent ion using the Qual browser of Xcalibur (3.0.63, Thermo Fisher) and matching with theoretically generated fragments of the same parent ion from Mass Frontier (7.0, Thermo Fisher) software (Table S4). The workflow for data analysis for RP hand HILIC is shown in Workflow S1. Data were reported at levels 1 and 2 according to the metabolite standard initiative guidelines (Sumner et al., 2007).

The QCViewer function of the RawHummus tool, an R package, was used to evaluate data quality, instrument stability, and reproducibility (Dong et al., 2022). R packages like ggplot2, reshape, and pheatmap were used to generate scatter plots, stack bar plots, and heat maps. Venn diagrams were generated using InteractiVenn, a web-based tool (Heberle H et al., 2015). The metabolites were putatively assigned kegg IDs and compound names at the MS1 level by searching against the KEGG (Kyoto Encyclopedia of Genes and Genomes) (http://www.kegg.com) database.

Detection of low abundance metabolic features

A cumulative abundance plot was used to evaluate the solvent extraction and chromatographic technique in extracting and detecting the least abundant metabolic features. Total number of detected features in individual analysis of all eight solvents were used for evaluation. For RP

and HILIC techniques highest detected features (i. e 70MW and 75HILIC) were compared in both positive and negative modes.

The proportional abundance of each feature contributing to the total abundance of all metabolic features detected was calculated using the following formula:

$PA = M_a/M_t$

PA=Proportional abundance
M_a=abundance of each metabolic feature
M_t= total abundance of all metabolic features

Cumulative abundance was calculated using the following formula:

$CA=PA_x+PA_h$

CA= cumulative abundance
PA_x= proportional abundance of the xth rank metabolic feature
PA_h=proportional abundance of metabolic feature $(x+1)^{th}$ higher in rank

Results

The present study aims to optimize and demonstrate a high-throughput, sensitive, untargeted approach to detect, quantify (relatively), and putatively annotate variations in LMW DOM availability across space in cryoconite holes. Therefore, to enable characterization and to expand the maximum coverage of metabolites, we evaluate several extraction solvent mixtures for RP-LCMS-based untargeted metabolomics of LMW compounds in CHs in combination with HILIC-LCMS data.

Extraction Yield

The extraction efficiencies of the eight solvent systems used to extract metabolites from the DOM of the CHs were evaluated (Fig. 1). According to percentage extraction yield, solvents were grouped as high (70AW, 70MW, ACN), moderate (MeOH, AMW, Water), and low (30MW, 30AW) yield. A visual inspection showed that all extracts were clear and transparent but varied in color (Fig. S1 A and S2 B). Aqueous and aqueous-rich solvent extracts such as Water, 30AW, and 30MW were colorless, whereas organic and organic-rich solvent extracts like MeOH, ACN, 70AW, and AMW were colored (light yellow and green). Despite 70MW, being an organic-rich solvent, the extract was colorless. The percentage extract yield ranged

between 0.66 to 1.56 %. Solvent 70 AW showed the highest extraction yield (1.56%). Overall, across all the solvent systems, extraction efficiency increased in the order: 30AW<30 MW<Water<AMW<MeOH<ACN<70MW<70AW.

Stability, accuracy, and reproducibility of the LC-MS system

To evaluate instrument stability, reproducibility, and accuracy, a detailed analysis of quality control (QC) samples was performed using the RawHummus tool. An overlaid TIC (Total ion chromatogram) analysis revealed the least retention time and intensity fluctuations throughout the run (Fig. S2 A). The summed TIC analysis (i.e., the sum of the intensities of all the scan points across the TIC) revealed that global ion intensity variations among QC samples were within an acceptable range (Fig. S2 B). The TIC correlation analysis revealed that the QC samples had good metabolic profile similarity (Pearson correlation coefficient R >0.90; acceptable range R >0.85) in RT (retention time) and chromatogram peak shape (Table S2), showing the least variance from run to run.

Throughout the chromatogram, RawHummus selected seven putatively annotated ion characteristics (two manually and five automatically) to monitor and assess the LC-MS system for RT, mass, and ion intensity variations. Maximum RT, mass difference, ion intensity ratio, and intensity coefficient of variation (CV) of the extracted ion chromatogram (EIC) of selected seven ion features are presented in Fig. S3 A and B. The maximum RT variation (which should not exceed 1 min) was less than 0.12 min, showing high retention time reproducibility. The mass difference between 0.25 and 1.34 ppm indicates good mass accuracy (maximum acceptable mass variation is up to 5 ppm). The ion intensity ratio (which should be between 1 and 2) was less than 1.6, indicating good stability. The ion intensity coefficient of variation for all m/z was less than 30 %, indicating better instrument stability. One of the m/z features among the seven was 148.0602, which was identified and confirmed as L-Glutamic acid after matching MS/MS spectra with authentic standards (Fig. S3 C and D).

Detection of features and evaluation of a solvent-wise number of metabolites coverage

A broad range of metabolic features were detected using different combinations of solvents. Table 1 depicts the metabolic features discovered and annotated for each extract in both positive and negative modes during early data processing. In RP, metabolic features ranged from 4898 to 6479 in the ESI (+)ve mode and 2381 to 2759 in the ESI (-)ve mode. Extracts containing organic or organic-rich solvents, such as methanol and acetonitrile mixtures, were

shown to have more detectable features. The 70MW and 70AW extracts showed the highest detected metabolic features (8868 and 8860, respectively). The order of extracts in terms of total detected metabolic features was 70MW > 70AW > AMW > MeOH > ACN > Water > 30MW > 30AW. The putative metabolites annotations using the KEGG database were between 23.45 and 23.87% in the ESI (+)ve mode and 18.92 to 19.25 % in the ESI (-)ve mode. In the HILIC phase, metabolic features ranged from 793-1033 in the ESI (+)ve mode and 2339-3275 in the ESI (-)ve mode. 75HILIC extract showed the highest number of metabolic features detected, annotated, and was used to compare and combine with RP in subsequent data analysis. The percentage of putative metabolites annotations using the KEGG database was between 28.84-29.82 % in the ESI (+)ve and 33.37-33.91 % in the ESI (-)ve mode. A large proportion (>75% in RP- and >65% in HILIC-LC-MS) of the detected features remained unknown after the database match.

One of the aims of this study is to find out the best solvent system, i.e., a single or combined solvent system, that would approach the maximum chemical diversity of CHs. In a single solvent system (Fig. 3A), 70MW solvent extracted the highest number of metabolites. Examining the features from any separately prepared parallel single solvent extractions enables comparing shared and unique features between two or more solvent systems. Shared and unique features in the two- and three- solvent combination are shown in figure 3B and Fig. S7, respectively.

Efficiency for extraction and detection of low-abundance metabolic features

To inspect the sensitivity and dynamic range of our untargeted approach, we examined the proportion for which each metabolic feature contributed to the total singal of metabolic features extracted by the solvents and detected by LC-MS techniques (Ladd et al., 2019). Fig. 2A showed that of the eight distinct solvents, Water had highest number of low abundant features (267), and MeOH had lowest (64) accounting for 50% of total signal, respectively. In Fig. 2B, of the four LC-MS techniques, RP (+) detected highest (177), and HILIC (+) detected lowest (8) low abundant features accounting for 50% of total signal. RP detected more low abundant features than HILIC, consistent with it having more favourable retention and ionisation conditions resulting in enhanced MS detection sensitivity.

Effect of extraction solvents on chemical diversity of metabolites in DOC of CHs

A principal component analysis (PCA, Fig. 4) was performed to understand how solvent extracts are globally related. The study revealed a tight primary clustering among four

biological replicates of each identical extraction solvent, indicating that LC/MS fingerprinting is reproducible. PC1 and PC2 accounted for more than 80% variation for all extracts. Notably, the PC1 component alone accounted for 70.4% variation, forming three groups: group 1 (aqueous and aqueous-rich solvents such as Water, 30MW, and 30AW), group 2 (organic-rich solvents such as 70MeOH, 70W, and AMW), group 3 (organic solvents such as MeOH and ACN). Therefore, secondary grouping, i.e., three distinct clusters, was generated using various extraction solvents. Interestingly, extract clusters are formed away from the PCA plot origin, revealing a particular set of metabolite-enriched clusters.

Heatmap (Fig. 5) analysis shows abundance of the metabolites of different chemical classes based on the polarity of the extraction solvent, from pure aqueous (polar) to pure organic (relatively non-polar) from right to left. The figure shows that adding organic solvents to water up to 30 percent (water, 30MW, and 30AW) extracted highly abundant polar metabolites, for example, polar (S)-malate, pyridoxal, and L-glutamate are more intense in aqueous and aqueous-rich solvents than in the rest of the solvent systems. As the proportion of the organic solvent increases to 70% (70MW and 70AW), a broad spectrum of polar and relatively nonpolar metabolites (l-methionine, phytosphinganine, and hexadecasphingosine etc.) were extracted in high abundnace. AMW showed more intense non-polar molecules and a few polar molecules. Pure organic solvents like MeOH and ACN showed highly abundant non-polar molecules like Ginkgetin, octadecanamide, and dodecanamide etc. For special questions, like targeted metabolomics studies, one can use pure organic solvents like acetonitrile or methanol to extract lipophilic fractions and water for highly hydrophilic fractions. The line plots on the left side of the heatmap showed the pattern of solvent-based metabolite extraction. The pink lines (upper left of the figure) represent a group of metabolites with high intensities from the top left portion of the heatmap, gradually decreasing from organic and organic-rich to aqueous-rich and aqueous extraction solvents. In contrast, the plot of the green lines (lower left of the figure) showed the opposite trend.

Combination of HILIC and Reverse phase chromatography for global metabolomics

The HILIC technique was combined with RP to increase the spectrum of chemically diverse metabolites from CHs. In addition to quantity, qualitative aspects of metabolites are an essential comparison measure. Therefore, to illustrate the distribution of extracted metabolites using both RP and HILIC chromatography, a scatter plot analysis was carried out about retention

time (rt) and mass (m/z) dimensions (Fig. 6). In RP, the majority of m/z eluted in the middle phase (from 5 to 15 min) of the chromatographic gradient. In HILIC, most m/z eluted at the initial phase (from 0.5 to 10 min) of the gradient.

75% HILIC solvent in HILIC extracted the highest number of metabolites (Table 1). The Venn analysis of the best solvent extract from HILIC (75HILIC) with RP best single-solvent (70MW) and two-solvent combination (70MW and AMW), respectively, revealed that both chromatographic conditions shared the least number of metabolites (Fig. 7A and B). The HILIC best vs. RP best single-solvent comparison showed that the RP (+ve) mode detected the most unique features (5229), followed by the HILIC (-ve) mode (2457). Both modes accounted for 73.68% of the metabolic features across the entire dataset. A similar pattern was observed when HILIC's best solvent extract was compared to the RP's best two-solvent combination system. It was noted that most features were detected in a positive mode in the RP and a negative mode in the HILIC technique.

The current investigation found amino acids, nucleic acids, sugars, and secondary metabolites in RP-LC and HILIC. Polar compounds like cytidine, ectoine, and L-glutamic acid and relatively non-polar molecules like gingerol, L-carnitine, L-tryptophan, L-tyrosine, and L-proline were detected at higher intensity in RP-LC (Fig. 8 A). Highly polar molecules like sugars such as D-ribose, D-erythrose, and rhamnose, and amino acids like L-serine L-glutamine, and L-phenylalanine, nucleic acids like adenosine and uracil were detected in higher intensity in HILIC (Fig. 8 B). This result is consistent with the fact that the HILIC approach is better suited for detecting highly polar molecules, while RP-LC is better at detecting relatively non-polar molecules.

Combining HILIC and RP chromatography significantly boosted the total number of metabolic features and their diversity (Fig. 7C). The number of features in HILIC best with the single best RP solvent was the least (10432). Overlapping these two techniques showed that few metabolites (1323-5363 features) were shared, showing that the combination of both this chromatographic phase covers the maximum feature diversity of the metabolome. As a result, to approach maximum coverage and chemical diversity of the LMW metabolites present in the DOM of CHs, the option of using both (RP and HILIC) chromatographic techniques should always be considered whenever possible.

Discussion

Extraction solvent compositions

In this study, we evaluated a defined set of extraction solvents. We aimed to detect a broad spectrum of metabolites in terms of their number and chemical diversity present in an extractable fraction of the DOM (extra- and intracellular) of CHs. Selecting the best extraction solvent is a pivotal aspect of metabolomics experiments, as efficient extraction relies on the polarity and physicochemical properties of both the metabolites and solvent (Moosmang et al., 2017). Typical solvents for metabolomics sample extraction include but are not limited to perchloric acid, ethyl acetate, chloroform, 2-propanol, and n-hexane. However, the perchlorate salts in perchloric acids can be problematic if not completely removed during the extraction procedure, as they are incompatible with analytical platforms like LC and GC (Mushtaq et al., 2014). Ethyl acetate is a more suitable solvent for GC-MS. Chloroform, 2-propanol, and n-hexane are preferred to extract non-polar metabolites in lipidomics studies (Reis et al., 2013). The ideal solvent for an LC-MS-based analytical platform should be less toxic, have high solubility power, have good pH tolerance, and be compatible with the instrument. Many of these properties can be controlled or modified using appropriate options, e.g., toxicity can be controlled by practicing routine safety measures, and pH can be modified using approved buffers (Mushtaq et al., 2014). However, some properties, like polarity and solubility power, are intrinsic to the solvent and cannot be modified.

For our study, we carefully selected the most commonly and widely used, less toxic, and highly RPLC-MS-compatible aqueous and organic solvents, such as water, methanol, and acetonitrile. Both methanol and acetonitrile have good miscibility with water and produce homogenous monophasic solutions. Another advantage of using these solvents is their good wetting behavior, which allows rapid mixing with the sample matrix. A study by Snyder 1978 (Table 1) classified various solvents according to their polarity index and selectivity group. Water, for instance, belongs to the selectivity group of VII and is the most polar solvent. Methanol, an aliphatic alcohol, belongs to group II and has a lower polarity index than water; it is a proton donor (i.e., protic polar solvent) and mid-polar solvent. Acetonitrile, a nitrile, belongs to group VIa with the highest dipole moment and slightly lower polarity index than methanol, is not a proton donor (i.e., aprotic polar solvent), and is considered a relatively non-polar solvent. Also, water has the highest solubility power, followed by methanol and acetonitrile. We prepared a set of eight solvent systems containing single, binary, and ternary mixtures of these solvents,

which can produce solvents of decreasing polarity strength from most polar to relatively non-polar such as Water > 30MW > 30AW >70MW > 70AW > AMW > MeOH > ACN. Each solvent was evaluated for extraction yield, the efficiency of extracting low abundant features, maximum number, and chemical diversity of metabolites.

Increased number of metabolites using appropriate extraction solvent and their combination

Methanol, acetonitrile, and aqueous mixtures have been the choices for several reported studies. For example, for the fecal metabolite extraction, Huang et al. (2013) used MeOH, Deda et al. (2017) used aqueous ACN, and Jiménez-Girón et al. (2015) used saline water. The choice of the extraction solvent is study-specific to meet the desired objectives for our present study. Methanol, particularly aqueous methanol, is the most popular choice as an extraction solvent in many studies due to its ability to extract a broad and diverse set of metabolites (like amino acids, organic and nucleic acids, secondary metabolites, and sugars) than water and ACN (Cheng et al., 2020). Fang et al., (2024) found ACN a suitable extraction solvent for saliva samples.

In the current investigation, binary mixtures of organic-rich solvents extracted the highest number of metabolites (70MW and 70AW, Fig. 3A). We have found that single organic solvents (MeOH and ACN) extracted more metabolites than aqueous solvents (Water). Adding a low fraction of organic solvents in water (30AW and 30MW) resulted in slightly decreased metabolic features. We also noted that organic-rich binary mixtures of MeOH and ACN (70MW and 70AW) are not complementary but mostly similar in terms of the number of metabolic features extracted, contrary to Doppler et al. (2024), where methanol and acetonitrile were found to be complementary solvents. Differences in organic and aqueous extract can be due to more penetrability of organic solvents by disrupting microbial cells, resulting in leakage of intracellular contents.

To maximize the number of metabolic features, we considered combining separately prepared single-solvent extracts with pairs of two-solvent combinations. The total number of features and shared features differed significantly, depending upon which solvent combination was used. The 70MW: AMW was the most effective extraction solvent pair in two-solvent combinations, as evidenced by its tallest bar height (Fig. 3B) and heatmap analysis (Fig. 5). In heatmap 70MW extracted high-intensity metabolites in the region where AMW showed less

intense metabolite extraction. This could be due to differences in the polarity. Aqueous-rich solvents like 30AW and 30MW shared the least metabolites with organic solvents like MeOH in two solvent combinations. These solvent combinations, as they share the least number of metabolites, extracted more unique features by individual solvents displaying high chemical diversity. The heatmap showed different patterns; MeOH showed the high intensity of non-polar, whereas 30AW and 30MW showed high intensity for polar metabolites complementing each other.

This study shows that the polarity of the extraction solvent can impact the number of the extracted metabolites. One can select complementary two-solvent combinations of parallel or repeated extractions to increase this number further. One can look into least shared metabolites solvent combination systems to approach maximum chemical diversity. But, combinations of solvent extraction, either parallel or repeated, are well suited to small-scale studies. Implementing large-scale experiments or routine analysis makes work more time-consuming and laborious and may result in diluted metabolic extract (Moosmang et al., 2017). Also, multistep protocols involving repeated extractions may introduce relatively higher deviations and variations in metabolite recovery (Karu et al., 2018).

Influence of the extraction solvent on the chemical diversity of the metabolites

LC-MS-based untargeted metabolomics, a high throughput technique, can detect a large number of metabolic features of different chemical classes. One of the objectives of this study is to evaluate the ability of extraction solvents to extract chemically diverse metabolic features. The polarity of the extraction solvent, whether polar, mid-polar, or non-polar, is a critical factor in extracting these diverse metabolites. The type of sample also influences the chemical nature of the extracted metabolic feature. For instance, the metabolite diversity in human blood (or plasma) differs from that in plant tissue or environmental samples (CHs). Therefore, we propose that the extracted metabolites' overall chemical diversity is a combined effect of the extraction solvent's polarity and the sample type.

In the DOM of CHs, water provides the natural environment containing water-soluble metabolites, particularly extracellular, which microbes can access. This applies to various environmental ecosystems like terrestrial soils etc. So, water, a highly polar solvent, can be a good starting point for determining these environmental ecosystems' chemical diversity of metabolomes (Swenson et al., 2015). Moosmang et al. (2017) have discussed the scope of using

aqueous (Water) and organic (methanol and acetonitrile) as an extraction solvent for different chemical classes of metabolites. According to this study, water can extract highly polar (hydrophilic) molecules, whereas methanol and acetonitrile mainly cover lipophilic constituents of the sample metabolome. The main advantages of water as an extraction solvent are that it is cost-effective and reproducible; the extract can be directly used for LC-MS with minimal further treatment and is non-toxic. In our study, to increase chemical diversity from hydrophilic to lipophilic, we prepared extraction solvents ranging from water (aqueous, highly polar) to methanol (organic, mid-polar) and acetonitrile (organic, relatively non-polar) and mixtures of them (aqueous, and organic). Aqueous solvents extracted more low-abundant features than organic solvents like ACN and MeOH (Fig. 2A).

Combining HILIC and RP is an excellent approach to increase metabolome coverage.

To this point, we have discussed how the meticulous selection of an extraction solvent can enhance the overall number and chemical diversity of the metabolites. Careful selection of an appropriate analytical platform is simultaneously crucial to identify this large number of metabolites. With high-throughput technology and software advancements in the last couple of decades, LC-MS has emerged as the top choice as an analytical platform and has been extensively used for many studies involving targeted or untargeted metabolomics of the plant and human tissues, urine, and bacterial cultures. RP-LCMS stands out among all for its robust, high throughput, and reproducibility for scientific findings. It is especially adept at detecting mid-polar to non-polar molecules. However, highly polar and ionic molecules (e.g., amino, organic acids, and sugars) are often poorly retained and resolved on RP-LC columns, hindering their detection and accurate quantification (Bieber et al., 2017). For a complex mixture such as soil, sediment, and similar environmental samples where polar molecules (water-soluble fraction of the DOM of these ecosystems) constitute a significant portion of the metabolome, it is advised to combine RP-LCMS with HILIC technique, which, on the other hand, is an excellent analytical platform to retain and detect highly polar molecules, thus complementing RP-LCMS (Ladd et al., 2021).

In our study, we used dual LC-MS (RP and HILIC) and dual ionization mode (positive and negative) approaches to increase the coverage of the metabolome in Antarctic CHs. Qualitative characteristics of the metabolic features regarding their distribution in chromatographic and mass spectral dimensions (Fig. 6) varied in RPLC and HILIC. In RPLC, the feature distribution

of extracts is heavily weighted between RT 5 to 15 min, indicating retention of relatively polar and non-polar molecules. In HILIC, the feature distribution of extracts was denser between 0.5 to 10 min, indicating well-retained highly polar molecules. To verify the differential detection of highly polar and relatively non-polar molecules in these techniques, we selected some putatively annotated highly polar and relatively non-polar metabolites for comparison (Fig. 8 A and B). Mid-polar molecules like gingerol, L-carnitine, L-tryptophan, etc., were detected at high intensity in RP-LCMS. While highly polar molecules like D-ribose, adenosine, L-serine, etc., were high in HILIC. Analyzing the overlap between the two conditions (Fig. 7 A, B, C, and Fig. 9) revealed that the least metabolic features were shared between these two techniques. Only around 0.68% of metabolic features were shared in more than one technique in a single solvent and two-solvent, respectively. A study by Contrepois et al., (2015) showed that combining the HILIC and RPLC-MS approaches greatly expanded metabolome coverage, with 44% and 108% new metabolic features detected compared to RPLC-MS alone for urine and plasma samples, respectively. Similarly, in the present study, combining HILIC with RP-LC for a single, two-solvent combination led to the detection of 46.96 % and 24.52% of unique metabolic features, respectively, compared to RP-LC alone. The decrease in percentage by combining more than one solvent was due to the contribution of the HILIC remaining the same while increasing the number of metabolites in RP in a two-solvent combination.

The use of LC-MS platforms to decipher the metabolome of the DOM of environmental samples is in an early stage of development. Furthermore, the novel approach of employing dual-LC and dual ionization polarity, which has not been reported to characterize the range of the LMW of DOM from CHs, adds a unique perspective to our study. The choice of using this innovative approach of combining HILIC and RPLC in both positive and negative modes will depend on the study's merit and the availability of resources.

Extraction yield, reproducibility and detection of low abundance metabolites

Yield measurements indicate that 70AW and 70MW are the most efficient solvents in extraction yield compared to pure methanol and acetonitrile for extracting metabolites from cryoconite hole samples. 70MW, despite being colorless, produced a high extraction yield. This result shows a similar observation in a previous study where aqueous organic extracts yielded more than absolute organic extracts (Martin et al., 2014). Although the percentage yield reflects crude extraction efficiency in quantity, it may not provide exact chemical diversity. In

conclusion, a high-yielding solvent may produce extracts of low overall quality due to limited diversity.

The reproducibility of the extracted metabolites is another parameter used to assess the extraction solvent. Our study showed good reproducibility for all eight solvents (Fig. 4). The tight primary clustering formed between replicates of each solvent showed good reproducibility of the solvent, extraction method, and analytical conditions. The secondary grouping (groups 1,2 and 3) formed due to the varying metabolic features extracted by the eight extracts, viz., Aqueous (polar; Water) and aqueous rich (polar; 30MW and 30AW) solvents, organic-rich (mid-polar; 70MW,70AW, and AMW) and organic (relatively non-polar; MeOH and ACN).
In addition, an untargeted approach should reliably identify a diverse group of compounds, including both high- and low-abundant metabolites. The climate in Antarctica is severe, and metabolic traits with low abundance could also be biologically significant.

Roles of detected metabolites in CHs microbial community and their adaptation to extreme environment

In the present study, CHs sediment LC-MS-based untargeted metabolomics workflow detected a wide range of small molecules that have crucial roles in the biological activity of cryoconite holes. Table S5 lists putatively annotated metabolites detected in this study and previously reported, which play an important role in adaptation in cold environments. Production of small osmoprotectants and cryoprotectants like glycine, betaine aldehyde, glycerol, trehalose, sucrose, and D-sorbitol can allow cryoconite microbial community to reduce the freezing point of cytoplasm thereby preventing macromolecule aggregation and can also stabilize cell membrane under cold conditions (Tribelli and Lopez, 2018). Therefore, in cold and challenging environments like CHs, microbes have found a way to survive and grow by developing adaptation strategies to adapt to their surroundings. It is known that microbial growth is mediated more by nitrogen rather than soil carbon content (Eisenhauer et al., 2010), indicating that amino acids play a crucial role in regulating these microbial communities. Many amino acids and carbohydrates in the present study were also supported. The present study also

detected branched and unsaturated fatty acids, which play a crucial role in microbial cell membrane modifications in response to cold stress (Siliakus et al., 2017).

Conclusion

Here, we comprehensively assess the defined set of solvents for an RP-LCMS-based high-throughput, sensitive, and robust untargeted metabolomics workflow. This method can be feasibly applied to extract a broad spectrum of the LMW molecules in the DOM of global CHs. Aqueous solvents like water extracted highly polar metabolites, while pure organic solvents like methanol and acetonitrile extracted more non-polar metabolites. The organic-rich solvent mixtures (70MW and 70AW) led to the extraction of a large number and broad range of chemically diverse metabolites. All solvents showed good reproducibility. No solvent could extract a complete feature set, so combining parallel single or repeated extractions with multiple solvent systems is required to enhance the number and capture maximum chemical diversity. A combination of HILIC and RP phase chromatography expanded coverage of a diverse and large-scale chemical feature, providing qualitative and relative quantitative data, yielding an information-rich chemical snapshot of biogeochemical activity in cryoconite holes. We conclude that the choice of the solvent and technique is exclusively study-specific, and solvents can be preferred as per the requirement of the experiment.

Our research opens up exciting possibilities for future studies. We suggest the application of a 13C isotope-assisted workflow that would allow the direct consideration of detectable metabolites of biological origin to assess different extraction mixtures. This could significantly enhance the accuracy and reliability of future studies. The Future work should include applying the technique across multiple northern and southern hemisphere cryoconite hole samples, correlating shifts in LMW DOM chemistry with microbial community composition and environmental variables to map compounds to metabolic pathways. Furthermore, the larger-scale studies integrating metabolomics with other ‘omics’ approaches will provide deep insight into critical biogeochemical transformations, e.g., C and nutrient cycling over time or space. Detailed stoichiometric information for mechanistic models will help reduce uncertainty in predictions of how cryoconite hole ecosystems of the northern and southern hemispheres will respond to climate change.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

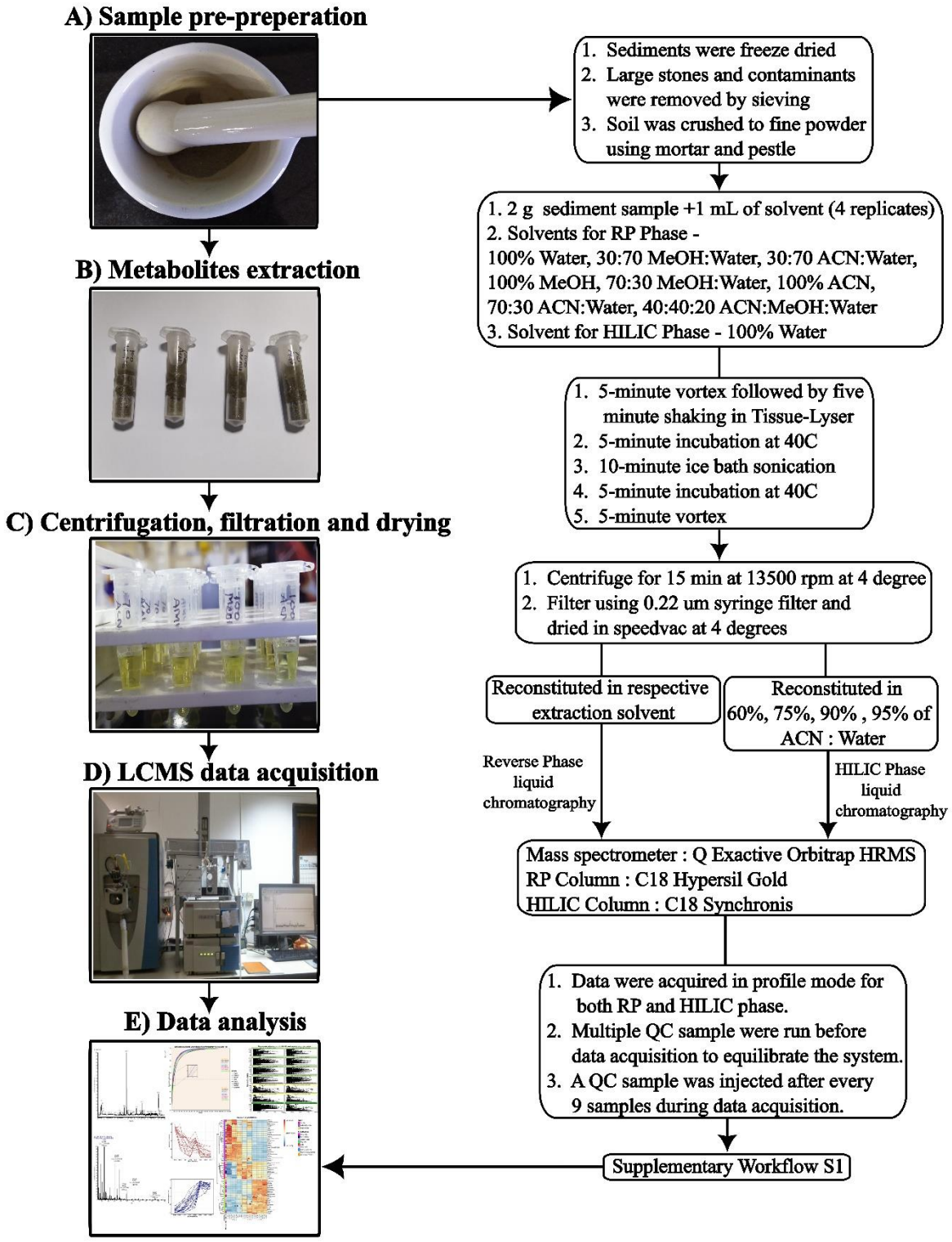
Data availability

The high-resolution LC-MS data from the present study were deposited on the massIVE repository, the link will be provided on demand.

Acknowledgments

We thank the Director, National Centre for Polar and Ocean Research (NCPOR), Goa, India, for providing samples, Dr. B Santhakumari for granting access to the Mass spectrometry facility at CSIR-National Chemical Laboratory, and Mr. Anirban Majumdar for helping with the R scripts. This work was funded by the Ministry of Earth Sciences (MoES), Government of India, under a central sector umbrella scheme named "Polar Science and Cryosphere Research (PACER)." We also thank the Director-National Center for Cell Science, an autonomous institute funded by the Department of Biotechnology, Government of India, for allowing access to the resources for wet and dry lab work and the Director of National Centre for Polar and Ocean Research (NCPOR), Goa, India, for support.

Figures:


A) Sample pre-preperation
1. Sediments were freeze dried
2. Large stones and contaminants were removed by sieving
3. Soil was crushed to fine powder using mortar and pestle
B) Metabolites extraction
1. 2 g sediment sample +1 mL of solvent (4 replicates)
2. Solvents for RP Phase - 100% Water, 30:70 MeOH:Water, 30:70 ACN:Water, 100% MeOH, 70:30 MeOH:Water, 100% ACN, 70:30 ACN:Water, 40:40:20 ACN:MeOH:Water
3. Solvent for HILIC Phase - 100% Water
1. 5-minute vortex followed by five minute shaking in Tissue-Lyser
2. 5-minute incubation at 40C
3. 10-minute ice bath sonication
4. 5-minute incubation at 40C
5. 5-minute vortex
C) Centrifugation, filtration and drying
1. Centrifuge for 15 min at 13500 rpm at 4 degree
2. Filter using 0.22 um syringe filter and dried in speedvac at 4 degrees
Reconstituted in respective extraction solvent
Reconstituted in 60%, 75%, 90% , 95% of ACN : Water
Reverse Phase liquid chromatography
HILIC Phase liquid chromatography
D) LCMS data acquisition
Mass spectrometer : Q Exactive Orbitrap HRMS
RP Column : C18 Hypersil Gold
HILIC Column : C18 Synchronis
1. Data were acquired in profile mode for both RP and HILIC phase.
2. Multiple QC sample were run before data acquisition to equilibrate the system.
3. A QC sample was injected after every 9 samples during data acquisition.
E) Data analysis
Supplementary Workflow S1

Workflow 1: Workflow of the untargeted metabolomics approach established in the present study to analyze LMW DOM from Cryoconite holes samples. Several solvent systems were used for extraction for reverse phase and HILIC phase chromatography-MS analysis. After filtration, quadruple extracts for each solvent system were done separately. The resulting concentrated aliquots (90 injections or samples) were run on two LC phases and in two MS modes, resulting in four analytical conditions per sample. RT: retention time. ESI (+)ve and ESI (-)ve electrospray ionization positive and negative mode.

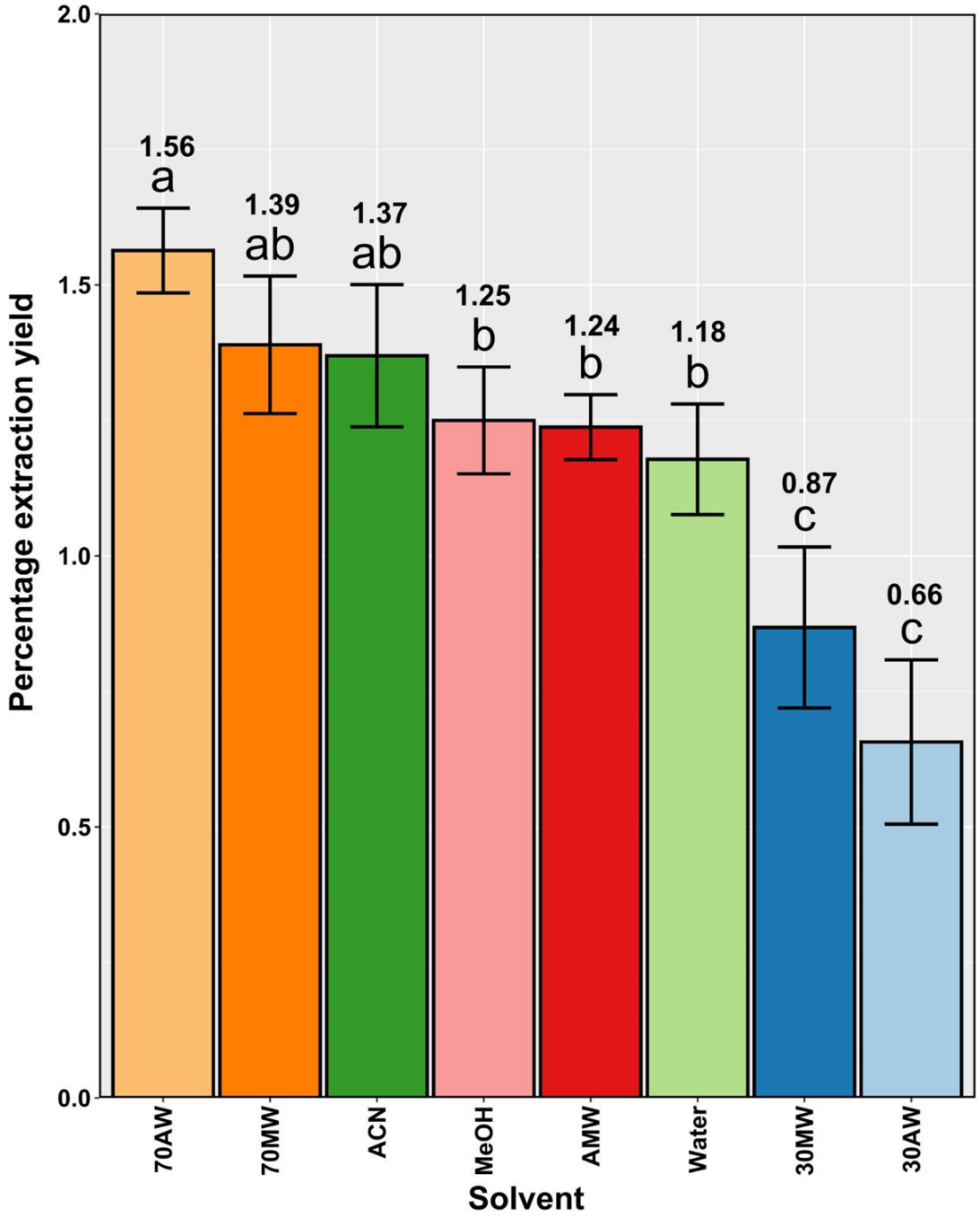
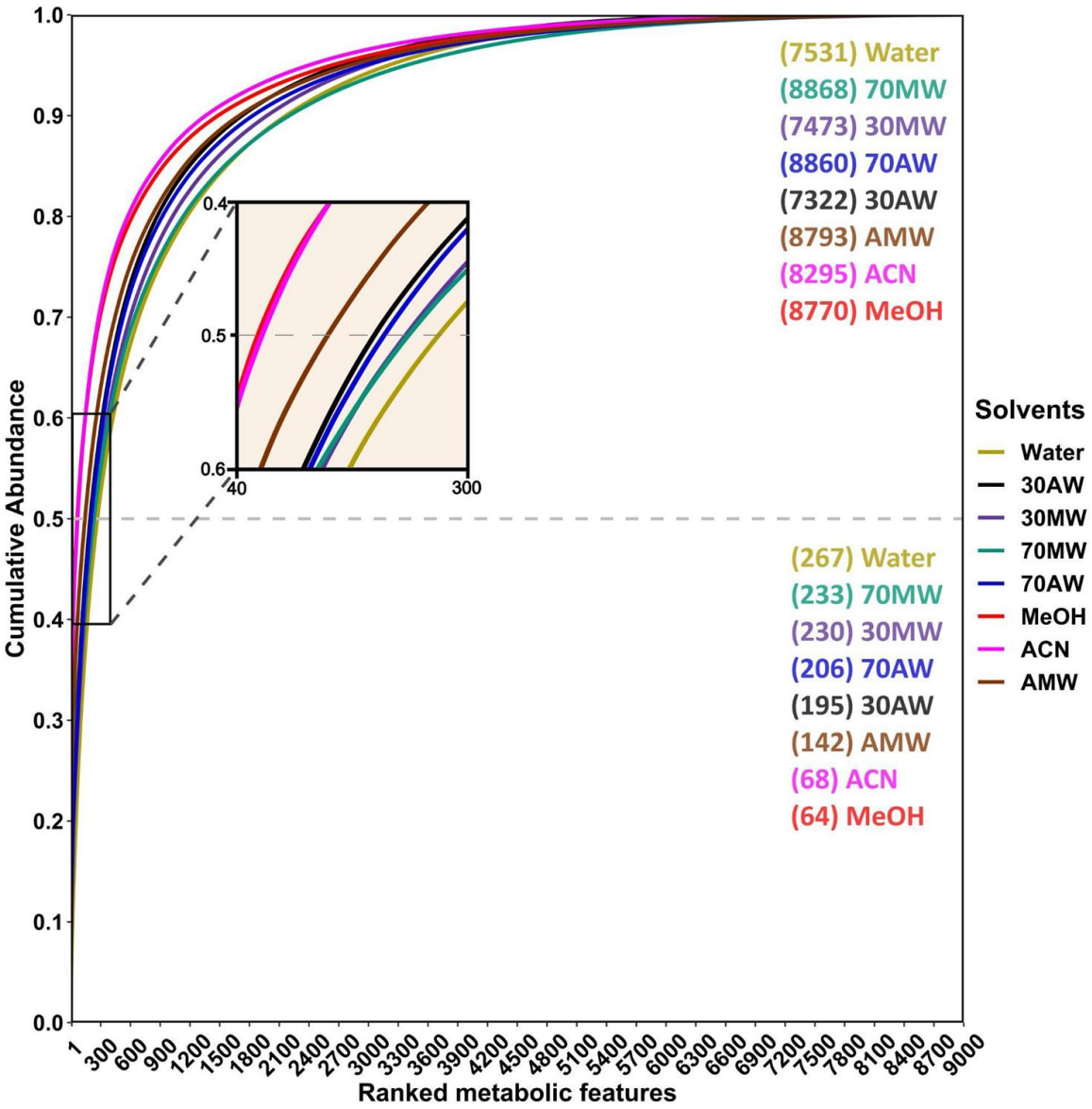
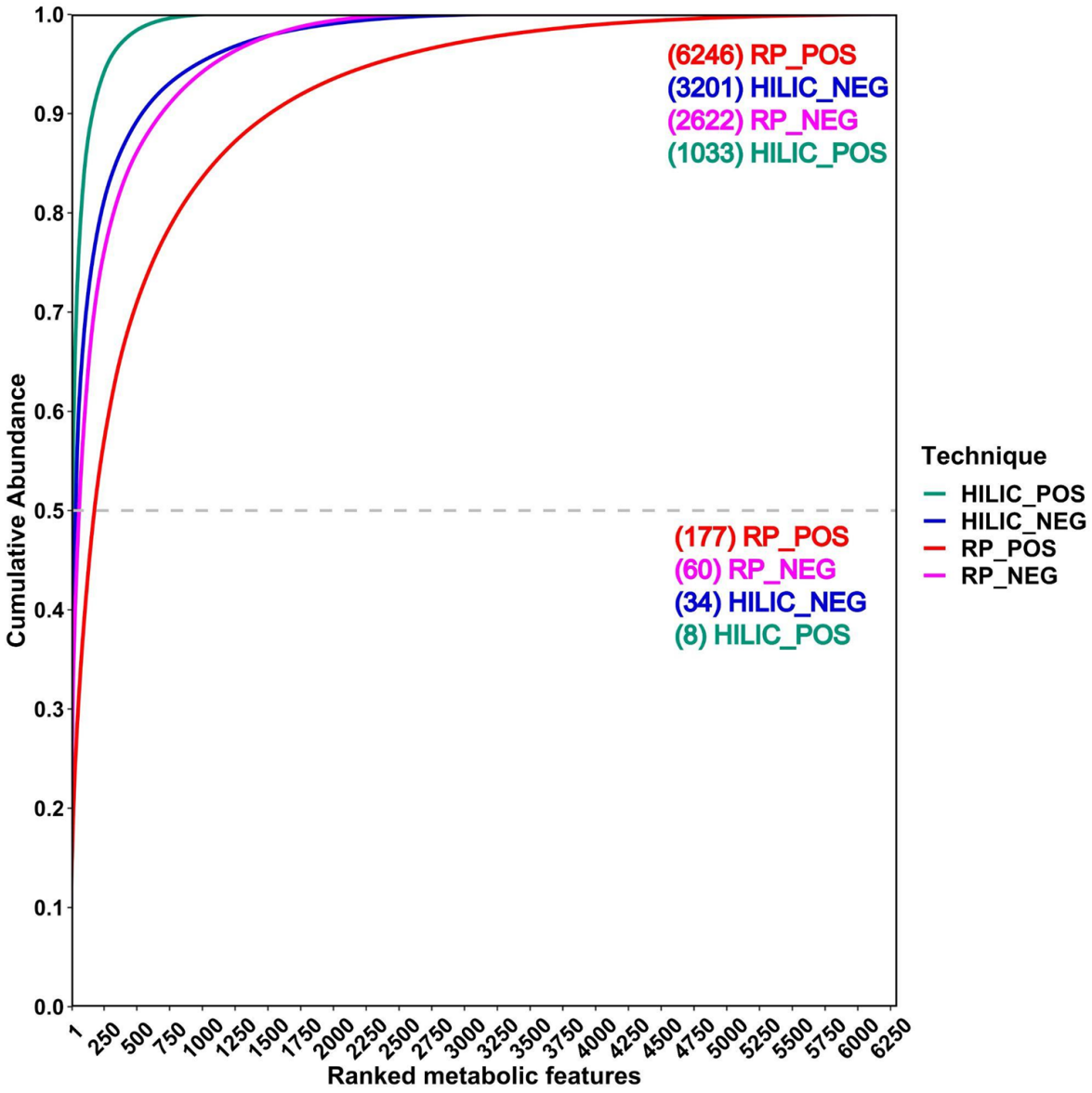


Figure 1: Percentage extraction yield. Comparison of the extraction yield for extracts prepared from 8 extraction solvents using CHs sediments. Bars show the average percent yield of 4 replicate extractions related to the initial dry weight of the sediments, and the error bars represent the standard error. Letter superscripts indicate statistically different groups at a p-value < 0.01 according to Tukey's HSD test. The average percent extraction yield of the different solvent extracts is displayed above the letter superscripts.


(7531) Water
(8868) 70MW
(7473) 30MW
(8860) 70AW
(7322) 30AW
(8793) AMW
(8295) ACN
(8770) MeOH
(267) Water
(233) 70MW
(230) 30MW
(206) 70AW
(195) 30AW
(142) AMW
(68) ACN
(64) MeOH
Solvents
Water
30AW
30MW
70MW
70AW
MeOH
ACN
AMW
Cumulative Abundance
Ranked metabolic features

(Figure 2A)



(Figure 2B)

Figure 2: Cumulative abundance plot for (A) individual solvents and (B) chromatographic phase.

High-quality features are ranked by abundance: 1=most abundant and 2300=least abundant in the case of individual solvents (A) and 1=most abundant and 1400=least abundant in the case of the chromatographic phase (B). The number of LMW DOM features detected by each solvent system and each LC-MS condition accounts for half, and the total cumulative abundance is reported.

Tabel 1A: Solvents and their strength and polarity indices (adapted from Snyder 1978)

Solvent	Selectivity group	Solvent strength (solubility power)	P′	Xe	Xd	Xn
Acetonitrile	VIa (nitriles)	0.65	6.2	0.3	0.26	0.41
Methanol	II (aliphatic alcohols)	0.95	6.6	0.51	0.19	0.3
Water	VIII (water)	>>1	9	0.4	0.34	0.26

P′ is the polarity index of the solvent, and Xe, Xd, and Xn represent the fraction of P′ contributed by interactions associated with proton-donor, proton-acceptor, and dipole moment characteristics of the solvent, respectively.

Table 1B: Extraction solvent compositions and polarity index

Chromatography phase	Extraction solvents and compositions					Number of replicates
	Extraction solvent	Reconstitution solvent	Combination	Category	Polarity index	
RP	Water	Water	Single	Aqueous	9	4
	30MW	30MW	Binary	Aqueous rich	8.28	4
	30AW	30AW	Binary	Aqueous rich	8.16	4
	70MW	70MW	Binary	Organic rich	7.32	4
	70AW	70AW	Binary	Organic rich	7.04	4
	AMW	AMW	Ternary	Organic rich	6.92	4
	MeOH	MeOH	Single	Organic	6.6	4
	ACN	ACN	Single	Organic	6.2	4
HILIC	Water	60HILIC	Single	Aqueous	9	4
	Water	75HILIC	Single	Aqueous	9	4
	Water	90HILIC	Single	Aqueous	9	4
	Water	95HILIC	Single	Aqueous	9	4

The polarity index of the mixture was calculated using the following formula (Adapted from Snyder 1978):

$P' = Ø_a P'_a + Ø_b P'_b$

Where,

P′ = polarity index of the resultant mixture

Øa = fraction volume of solvent a

P′a = polarity index of solvent a

Øb = fraction volume of solvent b

P′b = polarity index of solvent b

Table 1C: Detected and putatively annotated metabolic features

Chromatography phase	Solvent	Detected and putatively annotated metabolic features					
		Positive ESI mode		Negative ESI mode		Total features annotated (POS ESI + NEG ESI)	Total features detected
		Features detected	Features annotated	Features detected	Features annotated		
RP	Water	5089	1506	2442	446	1952	7531
	30MW	4922	1084	2551	465	1549	7473
	30AW	4898	1169	2424	437	1606	7322
	70MW	6246	1704	2622	561	2265	8868
	70AW	6479	1506	2381	460	1966	8860
	AMW	6267	1437	2526	498	1935	8793
	MeOH	6011	1457	2759	531	1988	8770
	ACN	5782	1320	2513	455	1775	8295
HILIC	60HILIC	939	272	3275	1110	1382	4214
	75HILIC	1033	298	3201	1149	1447	4234
	90HILIC	936	287	2730	985	1272	3666
	95HILIC	793	233	2339	780	1013	3132

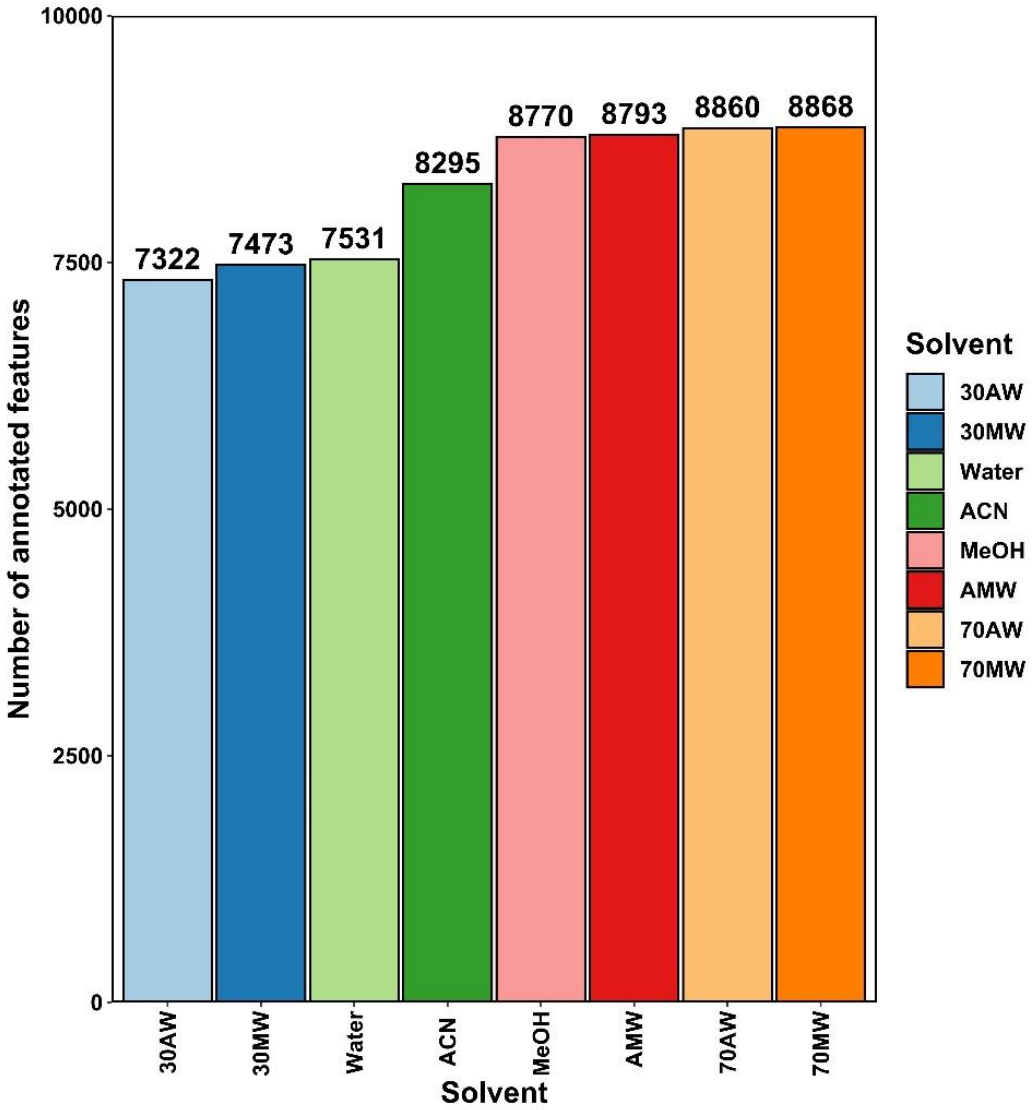


Figure 3 (A) Individual solvent extract bar plot

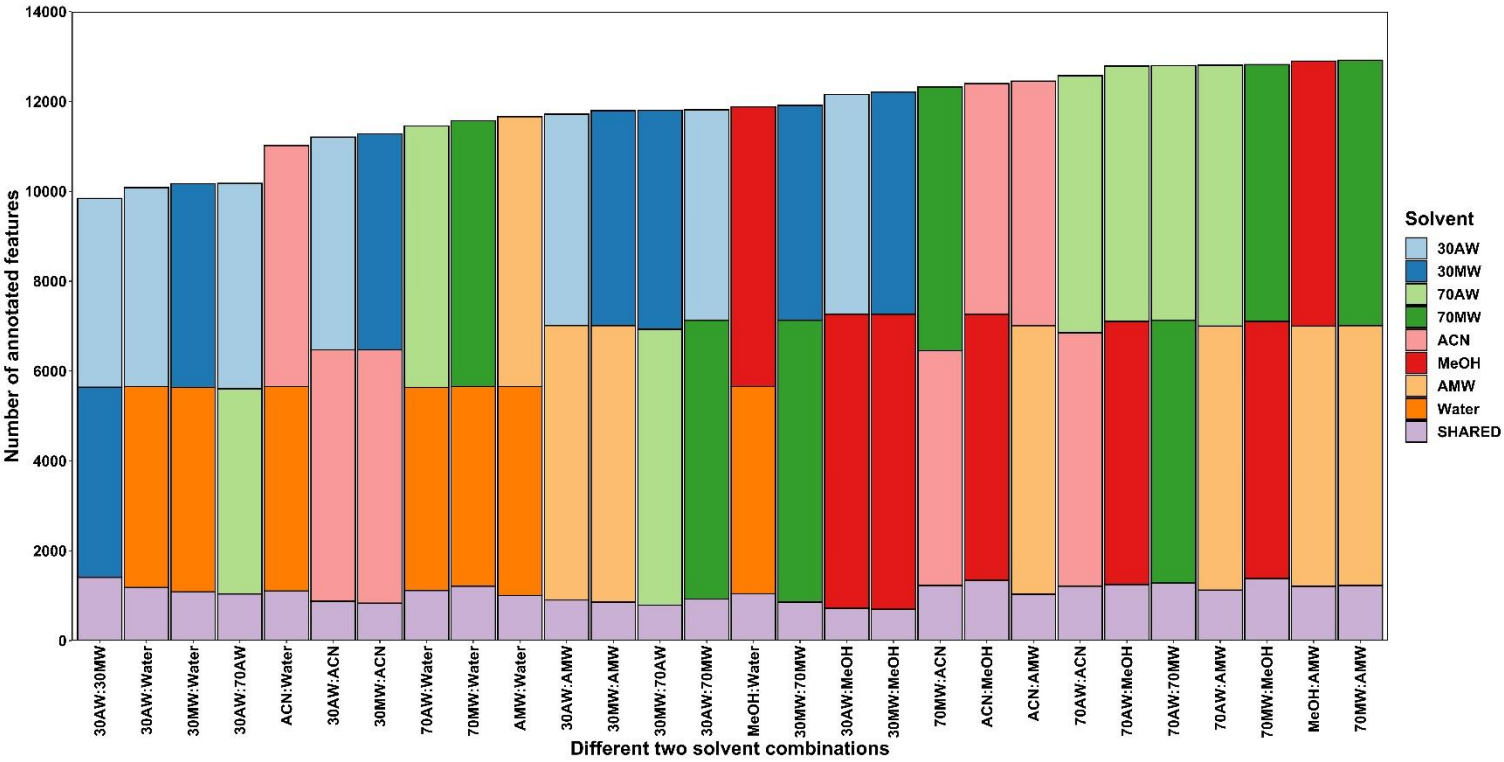


Figure 3 (B) two parallel single-solvent extraction combinations

Figure 3: (A) Individual solvent extract bar plot: Plot showing a number of metabolic features extracted by eight solvents. Each bar represents a solvent and is displayed with a unique color. The height of each bar indicates the total number of features detected; the exact number is shown at the top of each bar. Comparison of solvent extraction efficiency via chemical diversity for combinations of **(B) two parallel single-solvent extractions** The total number of unique features is shown for combinations of two and three parallel single-solvent extracts. Unique features detected by a single solvent from each combination are displayed by each color bar. The common features detected in two solvent systems are indicated in pink and named shared features.

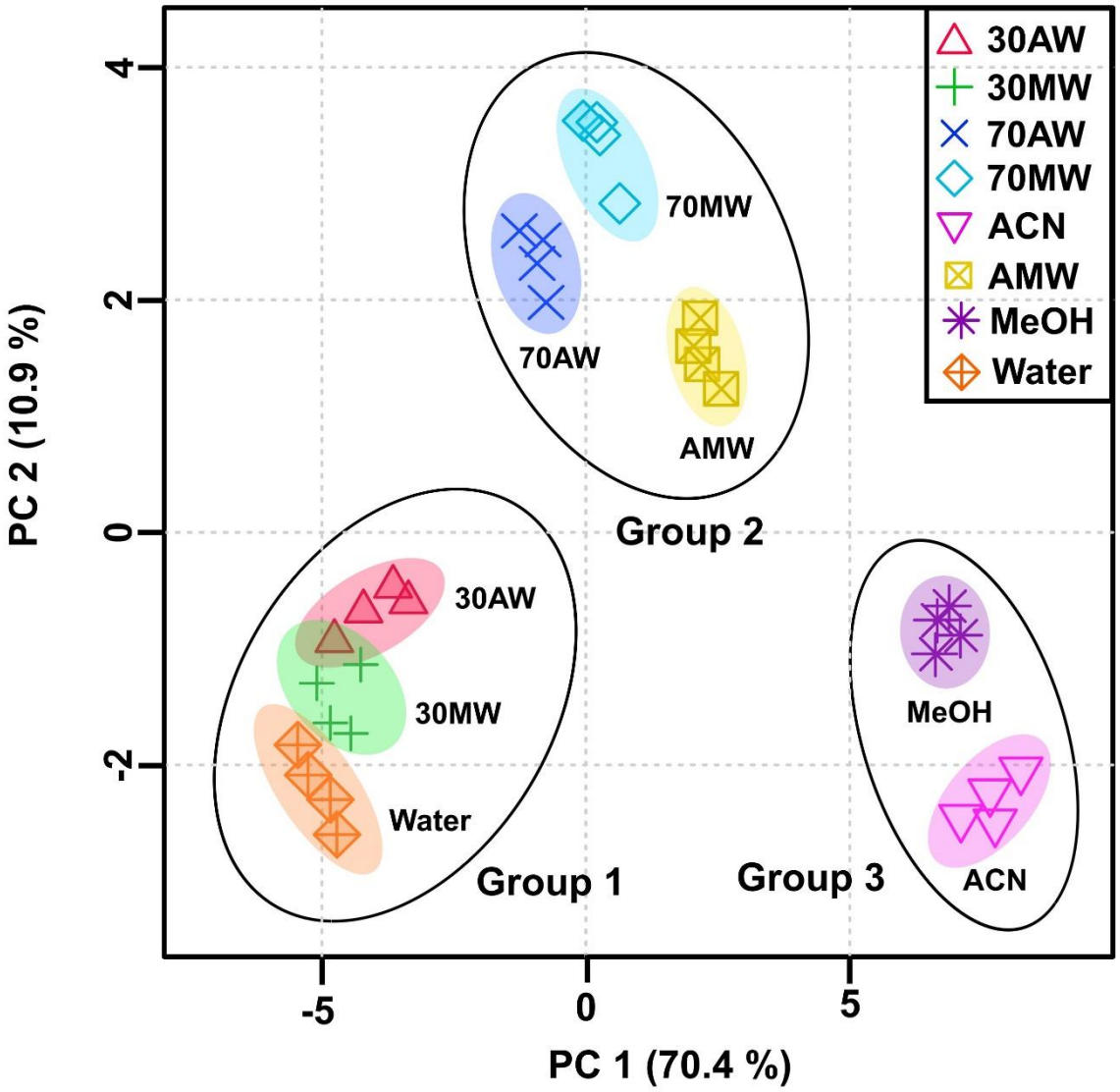


Figure 4: Principal component analysis of eight solvent extracts. PC1 and PC2 accounted for around 80% of the variation in this sample set. Tight cluster groups are produced by replicating a single solvent system, indicating that each has many highly reproducible unique parameters. Secondary clustering, i.e., groups 1, 2, and 3, are formed to replicate similar and clustered extracts.

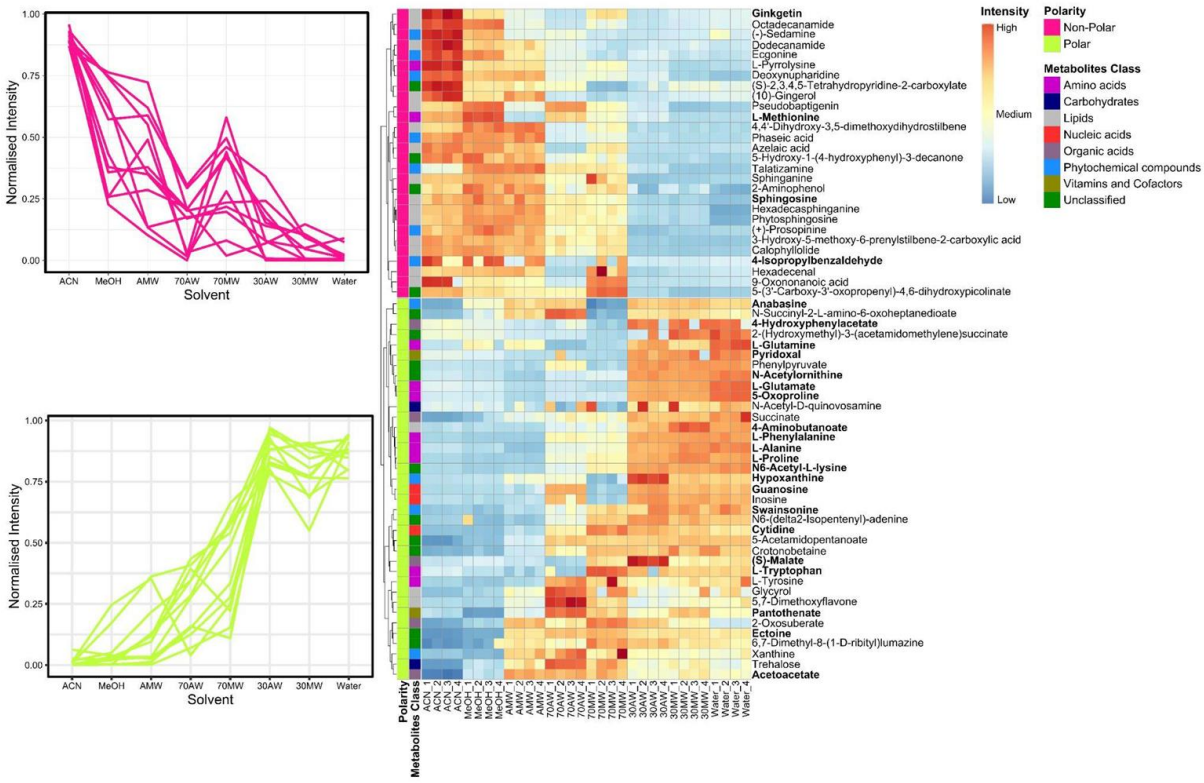


Figure 5: Comparison of extraction conditions and chemical diversity of detected metabolic features. The dendrogram on the left clusters similarly extracted metabolic features based on hierarchical clustering. The heat map displays the intensity of metabolites normalized to the most intense peak within each row (metabolic feature). Differences in metabolic feature peak intensity across extraction conditions are visible as top to bottom – based on polarity and left to right – based on solvents. The line plot at the far left of the figure shows decreasing (brown lines) and increasing (green lines) peak intensities in extraction conditions from left to right. Metabolic features were roughly divided into polar and nonpolar and categorized into different metabolic classes. Each class was given a different color, as the figure legend shows. Putatively confirmed metabolic features after MS/MS spectra match with the standards/METLIN database are shown in bold. N = 4 for each extraction condition.

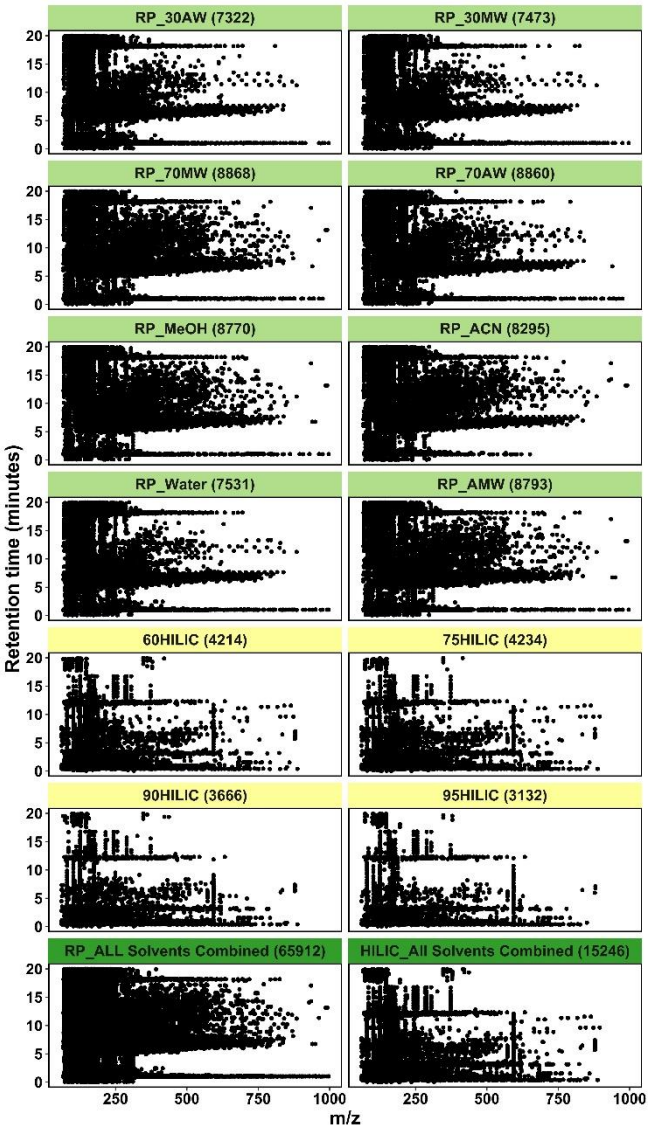


Figure 6: RP and HILIC solvents scatter plot. Plot showing the distribution of metabolic features with retention time. Distribution is observed for different solvent extracts and RP and HILIC chromatography—individual features detected by black dots. A total number of ion features detected from individual solvents is given parenthetically.

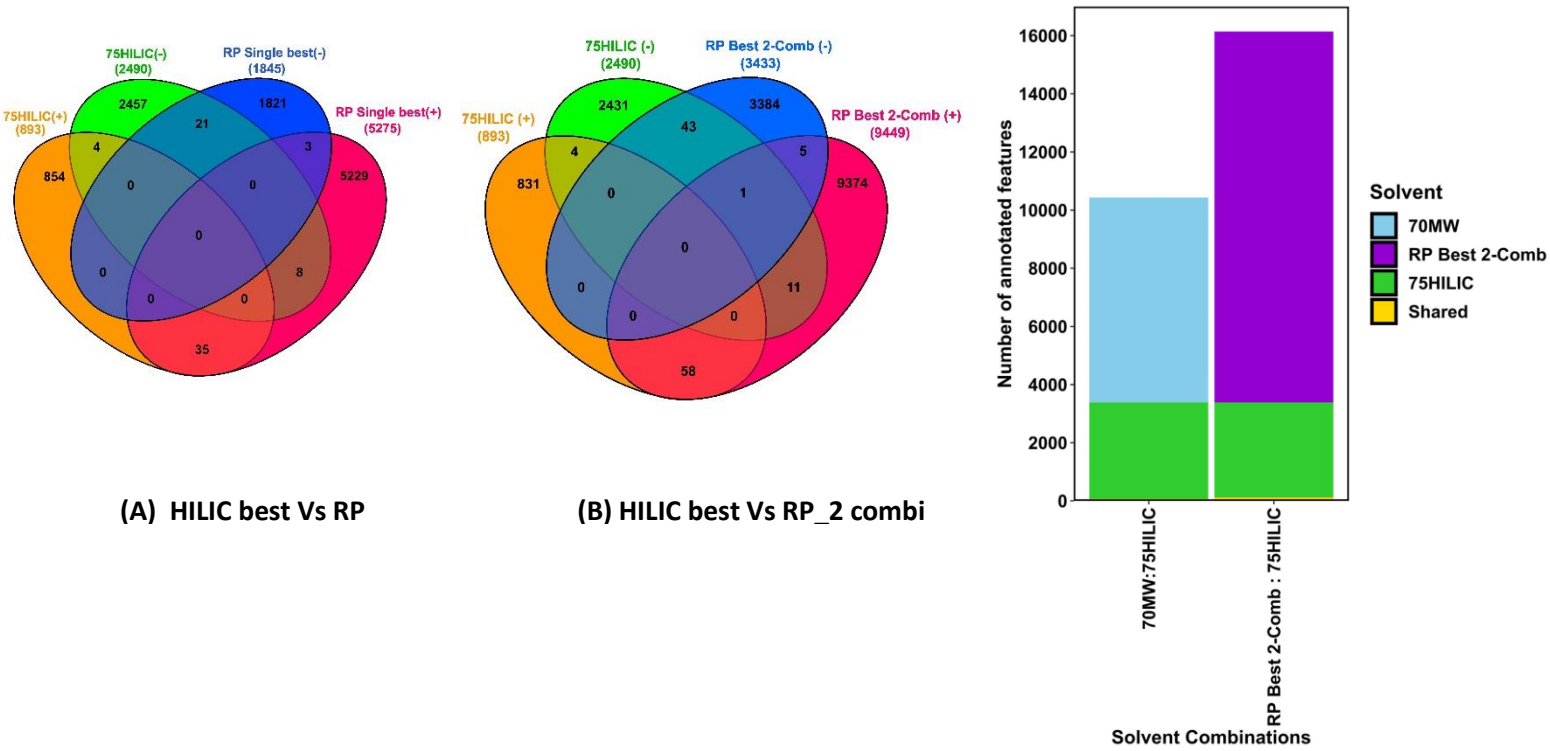


Figure 7. Overlap of High-quality features detected by HILIC and RP in positive- and negative-ion MS polarities (based on the MS1 neutral mass for the corresponding $[M + H]^+$ or $[M - H]^-$ ion, +/−0.001 Da). **(A)** feature detected between HILIC best (75HILIC) and RP best (70MW); **(B)** feature detected between HILIC best (75HILIC) and RP two best solvent combination (70MW and AMW); and **(C)** The bar plot shows unique and shared features between the best HILIC (75HILIC) and RP single solvent (70MW), two solvent combinations, and three solvent combinations. The total number of metabolites detected is displayed on top of each bar. Note that the number of shared metabolites (red) between both techniques is the least in all three comparisons.

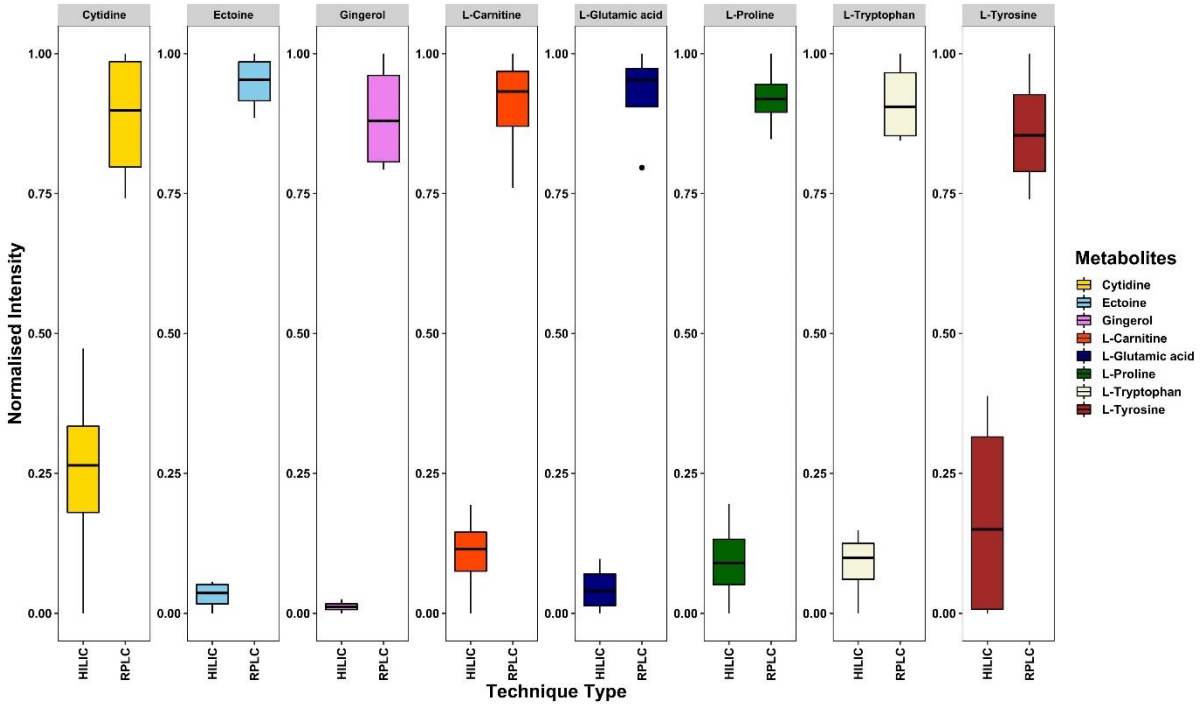


Figure 8 (A) Metabolic features with high intensity in RP phase LCMS

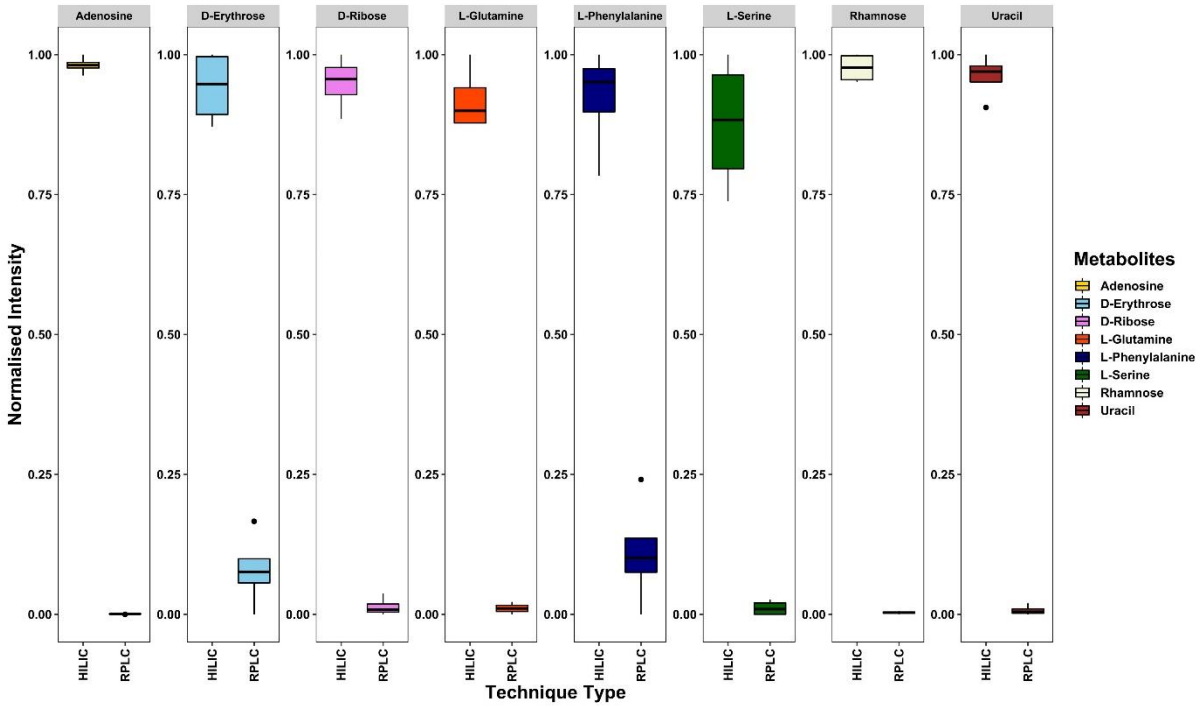


Figure 8 (B) Metabolic features with high intensity in HILIC phase LCMS

Fig. 8 (A) shows metabolic features detected in RP-LC-MS with higher intensity than HILIC-LC-MS and vice-versa in **Fig. 8 (B)**. The HILIC phase detected highly

polar metabolic features compared to the RP phase, whereas the RP phase detected relatively polar to relatively nonpolar metabolites.

References

Anesio, A.M., Hodson, A.J., Fritz, A., Psenner, R., Sattler, B., 2009. High microbial activity on glaciers: Importance to the global carbon cycle. Global Change Biology 15, 955–960.

Anesio, A.M., Lutz, S., Chrismas, N.A.M., Benning, L.G., 2017. The microbiome of glaciers and ice sheets. npj Biofilms and Microbiomes 3, 10.

Bagshaw, E.A., Tranter, M., Fountain, A.G., Welch, K., Basagic, H.J., Lyons, W.B., 2013. Do cryoconite holes have the potential to be significant sources of C, N, and P to downstream depauperate ecosystems of Taylor Valley, Antarctica? Arctic, Antarctic, and Alpine Research 45, 440–454.

Bieber, S., Greco, G., Grosse, S., Letzel, T., 2017. RPLC-HILIC and SFC with Mass Spectrometry: Polarity-Extended Organic Molecule Screening in Environmental (Water) Samples. Analytical Chemistry 89, 7907–7914.

Bittremieux, W., Avalon, N.E., Thomas, S.P., Kakhkhorov, S.A., Aksenov, A.A., Gomes, P.W.P., Aceves, C.M., Caraballo-Rodríguez, A.M., Gauglitz, J.M., Gerwick, W.H., Huan, T., Jarmusch, A.K., Kaddurah-Daouk, R.F., Kang, K. Bin, Kim, H.W., Kondić, T., Mannochio-Russo, H., Meehan, M.J., Melnik, A. V., Nothias, L.F., O'Donovan, C., Panitchpakdi, M., Petras, D., Schmid, R., Schymanski, E.L., van der Hooft, J.J.J., Weldon, K.C., Yang, H., Xing, S., Zemlin, J., Wang, M., Dorrestein, P.C., 2023. Open access repository-scale propagated nearest neighbor suspect spectral library for untargeted metabolomics. Nature Communications 14, 8488.

Boudah, S., Olivier, M.F., Aros-Calt, S., Oliveira, L., Fenaille, F., Tabet, J.C., Junot, C., 2014. Annotation of the human serum metabolome by coupling three liquid chromatography methods to high-resolution mass spectrometry. Journal of Chromatography B: Analytical Technologies in the Biomedical and Life Sciences 966, 34–47.

Bowen, T.J., Southam, A.D., Hall, A.R., Weber, R.J.M., Lloyd, G.R., Macdonald, R., Wilson, A., Pointon, A., Viant, M.R., 2023. Simultaneously discovering the fate and biochemical effects of pharmaceuticals through untargeted metabolomics. Nature Communications 14, 4653.

Catalá, T.S., Shorte, S., Dittmar, T., 2021. Marine dissolved organic matter: a vast and unexplored molecular space. Applied Microbiology and Biotechnology 105, 7225–7239.

Chambers, M.C., MacLean, B., Burke, R., Amodei, D., Ruderman, D.L., Neumann, S., Gatto, L., Fischer, B., Pratt, B., Egertson, J., Hoff, K., Kessner, D., Tasman, N., Shulman, N., Frewen, B., Baker, T.A., Brusniak, M.Y., Paulse, C., Creasy, D., Flashner, L., Kani, K., Moulding, C., Seymour, S.L., Nuwaysir, L.M., Lefebvre, B., Kuhlmann, F., Roark, J., Rainer, P., Detlev, S., Hemenway, T., Huhmer, A., Langridge, J., Connolly, B., Chadick, T., Holly, K., Eckels, J., Deutsch, E.W., Moritz, R.L., Katz, J.E., Agus, D.B., MacCoss, M., Tabb, D.L., Mallick, P., 2012. A cross-platform toolkit for mass spectrometry and proteomics. Nature Biotechnology 10, 918–920.

Cheng, K., Brunius, C., Fristedt, R., Landberg, R., 2020. An LC-QToF MS based method for untargeted metabolomics of human fecal samples. Metabolomics 16, 46.

Chetwynd, A.J., David, A., 2018. A review of nanoscale LC-ESI for metabolomics and its potential to enhance the metabolome coverage. Talanta 182, 380–390.

Contrepois, K., Jiang, L., Snyder, M., 2015. Optimized analytical procedures for the untargeted metabolomic profiling of human urine and plasma by combining hydrophilic interaction (HILIC) and reverse-phase liquid chromatography (RPLC)-mass spectrometry. Molecular and Cellular Proteomics 14, 1684–1695.

Cook, J., Edwards, A., Takeuchi, N., Irvine-Fynn, T., 2016. Cryoconite: The dark biological secret of the cryosphere. Progress in Physical Geography 40, 66–111.

Deda, O., Chatziioannou, A.C., Fasoula, S., Palachanis, D., Raikos, N., Theodoridis, G.A., Gika, H.G., 2017. Sample preparation optimization in fecal metabolic profiling. Journal of Chromatography B: Analytical Technologies in the Biomedical and Life Sciences 1047, 115–123.

Di Minno, A., Gelzo, M., Stornaiuolo, M., Ruoppolo, M., Castaldo, G., 2021. The evolving landscape of untargeted metabolomics. Nutrition, Metabolism and Cardiovascular Diseases 31, 1645–1652.

Dong, Y., Kazachkova, Y., Gou, M., Morgan, L., Wachsman, T., Gazit, E., Birkler, R.I.D., 2022. RawHummus: An R Shiny app for automated raw data quality control in metabolomics. Bioinformatics 38, 2072–2074.

Doppler, M., Kluger, B., Bueschl, C., Schneider, C., Krska, R., Delcambre, S., Hiller, K., Lemmens, M., Schuhmacher, R., 2016. Stable isotope-assisted evaluation of different extraction solvents for untargeted metabolomics of plants. International Journal of Molecular Sciences 17, 1017.

Doting, E.L., Jensen, M.B., Peter, E.K., Ellegaard-Jensen, L., Tranter, M., Benning, L.G., Hansen, M., Anesio, A.M., 2024. The exometabolome of microbial communities inhabiting bare ice surfaces on the southern Greenland Ice Sheet. Environmental Microbiology 26, e16574.

Dunn, W.B., Erban, A., Weber, R.J.M., Creek, D.J., Brown, M., Breitling, R., Hankemeier, T., Goodacre, R., Neumann, S., Kopka, J., Viant, M.R., 2013. Mass appeal: Metabolite identification in mass spectrometry-focused untargeted metabolomics. Metabolomics 9, 44–66.

Eisenhauer, N., Beßler, H., Engels, C., Gleixner, G., Habekost, M., Milcu, A., Partsch, S., Sabais, A.C.W., Scherber, C., Steinbeiss, S., Weigelt, A., Weisser, W.W., Scheu, A.S., 2010. Plant diversity effects on soil microorganisms support the singular hypothesis, Ecology 91, 485–496.

Eriksson, L., Johansson, E., Kettaneh-Wold, N., Trygg, J., Wikstrom, C., Wold, S., 2006. Multivariate and Megavariate Data Analysis Volume 1: Basic Principles and Applications. Umetrics Academy, Sweden.

Fang, L., Zhai, Q., Zhang, Hua, Ji, P., Chen, C., Zhang, Hongmei, 2024. Comparisons of different extraction methods and solvents for saliva samples. Metabolomics 20, 38.

Feng, L., Xu, J., Kang, S., Li, X., Li, Y., Jiang, B., Shi, Q., 2016. Chemical composition of microbe-derived dissolved organic matter in cryoconite in tibetan plateau glaciers: Insights from fourier transform ion cyclotron resonance mass spectrometry analysis. Environmental Science and Technology 50, 13215–13223.

Foreman, C.M., Sattler, B., Mikucki, J.A., Porazinska, D.L., Priscu, J.C., 2007. Metabolic activity and diversity of cryoconites in the Taylor Valley, Antarctica. Journal of Geophysical Research: Biogeosciences 112, G04S32.

Gika, H.G., Theodoridis, G.A., Plumb, R.S., Wilson, I.D., 2014. Current practice of liquid chromatography-mass spectrometry in metabolomics and metabonomics. Journal of Pharmaceutical and Biomedical Analysis 87, 12–25.

Gokul, J.K., Mur, L.A.J., Hodson, A.J., Irvine-Fynn, T.D.L., Debbonaire, A.R., Takeuchi, N., Edwards, A., 2023. Icescape-scale metabolomics reveals cyanobacterial and topographic control of the core metabolism of the cryoconite ecosystem of an Arctic ice cap. Environmental Microbiology 25, 2549–2563.

Grewer, D.M., Lafrenière, M.J., Lamoureux, S.F., Simpson, M.J., 2016. Redistribution of soil organic matter by permafrost disturbance in the Canadian High Arctic. Biogeochemistry 128, 397–415.

Gribbon, P.W.F., 1979. Cryoconite Holes on Sermikavsak, West Greenland. Journal of Glaciology 22, 177–181.

Heberle, H., Meirelles, V.G., da Silva, F.R., Telles, G.P., Minghim, R., 2015. InteractiVenn: A web-based tool for the analysis of sets through Venn diagrams. BMC Bioinformatics 16,169.

Huang, H. jun, Zhang, A. ye, Cao, H. cui, Lu, H. feng, Wang, B. hong, Xie, Q., Xu, W., Li, L.J., 2013. Metabolomic analyses of faeces reveals malabsorption in cirrhotic patients. Digestive and Liver Disease 45, 677–682.

Jiménez-Girón, A., Ibáñez, C., Cifuentes, A., Simó, C., Muñoz-González, I., Martín-Álvarez, P.J., Bartolomé, B., Victoria Moreno-Arribas, V., 2015. Faecal metabolomic fingerprint after moderate consumption of red wine by healthy subjects. Journal of Proteome Research 14, 897–905.

Karu, N., Deng, L., Slae, M., Guo, A.C., Sajed, T., Huynh, H., Wine, E., Wishart, D.S., 2018. A review on human fecal metabolomics: Methods, applications and the human fecal metabolome database. Analytica Chimica Acta 1030, 1–24.

Kuhl, C., Tautenhahn, R., Böttcher, C., Larson, T.R., Neumann, S., 2012. CAMERA: An integrated strategy for compound spectra extraction and annotation of liquid chromatography/mass spectrometry data sets. Analytical Chemistry 84, 283–289.

Ladd, M.P., Giannone, R.J., Abraham, P.E., Wullschleger, S.D., Hettich, R.L., 2019. Evaluation of an untargeted nano-liquid chromatography-mass spectrometry approach to expand coverage of low molecular weight dissolved organic matter in Arctic soil. Scientific Reports 9, 5810.

Ladd, M.P., Reeves, D.T., Poudel, S., Iversen, C.M., Wullschleger, S.D., Hettich, R.L., 2021. Untargeted exometabolomics provides a powerful approach to investigate biogeochemical hotspots with vegetation and polygon type in arctic tundra soils. Soil Systems 5, 1–19.

Lutz, S., Anesio, A.M., Field, K., Benning, L.G., 2015. Integrated "Omics", targeted metabolite and single-cell analyses of arctic snow algae functionality and adaptability. Frontiers in Microbiology 6, 1323.

Mann, B.F., Chen, H., Herndon, E.M., Chu, R.K., Tolic, N., Portier, E.F., Chowdhury, T.R., Robinson, E.W., Callister, S.J., Wullschleger, S.D., Graham, D.E., Liang, L., Gu, B., 2015. Indexing permafrost soil organic matter degradation using high-resolution mass spectrometry. PLoS ONE 10, e0130557.

Mann, P.J., Davydova, A., Zimov, N., Spencer, R.G.M., Davydov, S., Bulygina, E., Zimov, S., Holmes, R.M., 2012. Controls on the composition and lability of dissolved organic matter in Siberia's Kolyma River basin. Journal of Geophysical Research: Biogeosciences 117, G1.

Martin, A.C., Pawlus, A.D., Jewett, E.M., Wyse, D.L., Angerhofer, C.K., Hegeman, A.D., 2014. Evaluating solvent extraction systems using metabolomics approaches. RSC Advances 4, 26325–26334.

Moosmang, S., Pitscheider, M., Sturm, S., Seger, C., Tilg, H., Halabalaki, M., Stuppner, H., 2019. Metabolomic analysis—Addressing NMR and LC-MS related problems in human feces sample preparation. Clinica Chimica Acta 489, 169–176.

Mushtaq, M.Y., Choi, Y.H., Verpoorte, R., Wilson, E.G., 2014. Extraction for metabolomics: Access to the metabolome. Phytochemical Analysis 25, 291–306.

Nebbioso, A., Piccolo, A., 2013. Molecular characterization of dissolved organic matter (DOM): A critical review. Analytical and Bioanalytical Chemistry 405, 109–124.

Patti, G.J., 2011. Separation strategies for untargeted metabolomics. Journal of Separation Science 34, 3460–3469.

Pautler, B.G., Dubnick, A., Sharp, M.J., Simpson, A.J., Simpson, M.J., 2013. Comparison of cryoconite organic matter composition from Arctic and Antarctic glaciers at the molecular-level. Geochimica et Cosmochimica Acta 104, 1–18.

Poniecka, E.A., Bagshaw, E.A., Sass, H., Segar, A., Webster, G., Williamson, C., Anesio, A.M., Tranter, M., 2020. Physiological Capabilities of Cryoconite Hole Microorganisms. Frontiers in Microbiology 11, 1783.

Rozwalak, P., Podkowa, P., Buda, J., Niedzielski, P., Kawecki, S., Ambrosini, R., Azzoni, R.S., Baccolo, G., Ceballos, J.L., Cook, J., Di Mauro, B., Ficetola, G.F., Franzetti, A., Ignatiuk, D., Klimaszyk, P., Łokas, E., Ono, M., Parnikoza, I., Pietryka, M., Pittino, F., Poniecka, E., Porazinska, D.L., Richter, D., Schmidt, S.K., Sommers, P., Souza-Kasprzyk, J., Stibal, M., Szczuciński, W., Uetake, J., Wejnerowski, Ł., Yde, J.C., Takeuchi, N., Zawierucha, K., 2022. Cryoconite – From minerals and organic matter to bioengineered sediments on glacier's surfaces. Science of the Total Environment 807, 150874.

Sanyal, A., Antony, R., Samui, G., Thamban, M., 2018. Microbial communities and their potential for degradation of dissolved organic carbon in cryoconite hole environments of Himalaya and Antarctica. Microbiological Research 208, 32–42.

Schmidt, S.K., Johnson, B.W., Solon, A.J., Sommers, P., Darcy, J.L., Vincent, K., Vimercati, L., Fountain, A.G., Porazinska, D.L., 2022. Microbial biogeochemistry and phosphorus

limitation in cryoconite holes on glaciers across the Taylor Valley, McMurdo Dry Valleys, Antarctica. Biogeochemistry 158, 313–326.

Siliakus, M.F., van der Oost, J., Kengen, S.W.M., 2017. Adaptations of archaeal and bacterial membranes to variations in temperature, pH and pressure. Extremophiles 21, 651–670.

Silva, R.R., Jourdan, F., Salvanha, D.M., Letisse, F., Jamin, E.L., Guidetti-Gonzalez, S., Labate, C.A., Vêncio, R.Z.N., 2014. ProbMetab: An R package for Bayesian probabilistic annotation of LC-MS-based metabolomics. Bioinformatics 30, 1336–1337.

Singer, G.A., Fasching, C., Wilhelm, L., Niggemann, J., Steier, P., Dittmar, T., Battin, T.J., 2012. Biogeochemically diverse organic matter in Alpine glaciers and its downstream fate. Nature Geoscience 5, 710–714.

Smith, C.A., Want, E.J., O'Maille, G., Abagyan, R., Siuzdak, G., 2006. XCMS: Processing mass spectrometry data for metabolite profiling using nonlinear peak alignment, matching, and identification. Analytical Chemistry 78, 779–787.

Snyder, L.R., 1978. Classification off the Solvent Properties of Common Liquids. Journal of Chromatographic Science 16, 223–234

Sommers, P., Darcy, J.L., Porazinska, D.L., Gendron, E.M.S., Fountain, A.G., Zamora, F., Vincent, K., Cawley, K.M., Solon, A.J., Vimercati, L., Ryder, J., Schmidt, S.K., 2019. Comparison of microbial communities in the sediments and water columns of frozen cryoconite holes in the McMurdo Dry Valleys, Antarctica. Frontiers in Microbiology 10, 65.

Sumner, L.W., Amberg, A., Barrett, D., Beale, M.H., Beger, R., Daykin, C.A., Fan, T.W.M., Fiehn, O., Goodacre, R., Griffin, J.L., Hankemeier, T., Hardy, N., Harnly, J., Higashi, R., Kopka, J., Lane, A.N., Lindon, J.C., Marriott, P., Nicholls, A.W., Reily, M.D., Thaden, J.J., Viant, M.R., 2007. Proposed minimum reporting standards for chemical analysis: Chemical Analysis Working Group (CAWG) Metabolomics Standards Initiative (MSI). Metabolomics 3, 211–221.

Swenson, T.L., Jenkins, S., Bowen, B.P., Northen, T.R., 2015. Untargeted soil metabolomics methods for analysis of extractable organic matter. Soil Biology and Biochemistry 80, 189–198.

Tang, D.Q., Zou, L., Yin, X.X., Ong, C.N., 2016. HILIC-MS for metabolomics: An attractive and complementary approach to RPLC-MS. Mass Spectrometry Reviews 35, 574-600.

Tiffany, C.R. and Bäumler, A.J., 2019. omu, a Metabolomics Count Data Analysis Tool for Intuitive Figures and Convenient Metadata Collection DATABASES AND SOFTWARE crossm 8.

Tribelli, P.M. and López, N.I., 2018. Reporting key features in cold-adapted bacteria. Life 8.

Yao, S., Bian, Y., Jiang, X., Song, Y., 2023. Characterization of dissolved organic matter distribution in forestland and farmland of mollisol based on untargeted metabolomics. Soil Ecology Letters 5, 230179.

Zawierucha, K., Buda, J., Pietryka, M., Richter, D., Łokas, E., Lehmann-Konera, S., Makowska, N., Bogdziewicz, M., 2018a. Snapshot of micro-animals and associated biotic and abiotic environmental variables on the edge of the south-west Greenland ice sheet. Limnology 19, 141–150.

Zawierucha, K., Stec, D., Lachowska-Cierlik, D., Takeuchi, N., Li, Z., Michalczyk, Ł., 2018b. High Mitochondrial Diversity in a New Water Bear Species (Tardigrada: Eutardigrada) from Mountain Glaciers in Central Asia, with the Erection of a New Genus Cryoconicus. Annales Zoologici, 68, 179–201.

Zawierucha, K., Trzebny, A., Buda, J., Bagshaw, E., Franzetti, A., Dabert, M., Ambrosini, R., 2022. Trophic and symbiotic links between obligate-glacier water bears (Tardigrada) and cryoconite microorganisms. PLoS ONE 17, e0262039.